



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE CIÊNCIAS**  
**DEPARTAMENTO DE COMPUTAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO**  
**DOUTORADO EM CIÊNCIA DA COMPUTAÇÃO**

**HENRIQUE SÉRGIO LIMA PEQUENO**

**DIGITAL CLASS: UM MODELO PARA A CRIAÇÃO DE CURSOS NO PARADIGMA**  
**DE ENSINO HÍBRIDO**

**FORTALEZA**

**2017**

HENRIQUE SÉRGIO LIMA PEQUENO

DIGITAL CLASS: UM MODELO PARA A CRIAÇÃO DE CURSOS NO PARADIGMA DE  
ENSINO HÍBRIDO

Tese apresentada ao Curso de Doutorado em Ciência da Computação do Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Ciência da Computação. Área de Concentração: Sistemas de Informação

Orientador: Prof<sup>o</sup>. Dr. Miguel Franklin de Castro

Co-Orientador: José Gilvan Rodrigues Maia

FORTALEZA

2017

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

- P1d Pequeno, Henrique Sérgio Lima.  
Digital Class: Um Modelo Para a Criação de Cursos no Paradigma de Ensino Híbrido / Henrique Sérgio Lima Pequeno. – 2017.  
202 f. : il. color.
- Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Ciência da Computação, Fortaleza, 2017.  
Orientação: Prof. Dr. Miguel Franklin de Castro.  
Coorientação: Prof. Dr. José Gilvan Rodrigues Maia.
1. Ensino Híbrido. 2. Sistemas de Recomendação. 3. Ambiente Virtual de Aprendizagem. I. Título.  
CDD 005
-

HENRIQUE SÉRGIO LIMA PEQUENO

DIGITAL CLASS: UM MODELO PARA A CRIAÇÃO DE CURSOS NO PARADIGMA DE  
ENSINO HÍBRIDO

Tese apresentada ao Curso de Doutorado em Ciência da Computação do Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Ciência da Computação. Área de Concentração: Sistemas de Informação

Aprovada em: 31 de Agosto de 2017

BANCA EXAMINADORA

---

Prof<sup>o</sup>. Dr. Miguel Franklin de Castro (Orientador)  
Universidade Federal do Ceará (UFC)

---

José Gilvan Rodrigues Maia (Co-Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Francisco Milton Mendes Neto  
Universidade Federal Rural do Semi-Árido (UFERSA)

---

Prof. Dr. Cidcley Teixeira de Souza  
Instituto Federal de Educação, Ciência e Tecnologia do  
Ceará (IFCE)

---

Prof. Dr. Ernesto Trajano de Lima Neto  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Gabriel Antoine Louis Paillard  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. José Neuman de Souza  
Universidade Federal do Ceará (UFC)

Dedico este trabalho a Luís Inácio Lula da Silva, o maior responsável pelo crescimento da universidade pública do país e defensor incansável dos interesses do povo brasileiro.

## AGRADECIMENTOS

Agradeço a Deus e à Nossa Senhora de Fátima por ter permitido que ao longo da vida tivesse a oportunidade de acesso a estudos de qualidade.

Agradeço a meus pais por desde sempre estimular e orientar incansavelmente na busca de um futuro digno, sempre enaltecendo a importância dos estudos em minha vida. Meus irmãos, fonte de amor e força, agradeço pela paciência e respeito em razão de tempos prolongados de reclusão aos quais me submeti.

Agradeço a minha esposa e filhos. Eles compartilharam, diariamente, as consequências inerentes à realização deste objetivo de vida, com momentos de renúncia à vida social. Minha esposa, minha companheira incansável e revisora de primeira hora, meu agradecimento especial.

Agradeço a meu Tio Prof. Mauro Pequeno, que sempre enxergou potencial em meu trabalho. Seu exemplo de simplicidade e denodo para com a Universidade sempre foram fonte de inspiração. A meu Tio Prof. Marcelino pela paciência e atenção nas aulas de Lógica.

Agradeço a meus avós, Prof. Haroldo Pequeno e Dona Magela, pela acolhida como filho ao vir estudar em Fortaleza. Os ensinamentos de matemática, física e ciências nos tempos de escola foram encantadores e repousam em minhas mais felizes memórias. Obrigado a ambos pela formação e valores que recebi. Por fim meu avô, agradeço pelo exemplo de Professor que foi e, com isso, pude me associar hoje ao seu amor, admiração e respeito pela Universidade Federal do Ceará.

A meu orientador e amigo, Prof. Miguel Franklin, pela confiança sempre dada e orientações qualificadas para a obtenção deste trabalho. A meu co-orientador, também amigo, Prof. José Gilvan Rodrigues Maia, pelo auxílio incondicional em todo o doutorado.

A meus colegas e amigos do Instituto UFC Virtual e Laboratório de Mídias Educacionais que direta ou indiretamente me apoiaram nessa jornada. Em especial, agradeço a Izabella Sombra, Ricardo Palácio, Elifrânio Cruz, Lisboa Jr., Solange Macedo e Gilson Jr pelas inestimáveis e incontáveis ajudas técnicas oferecidas ao longo do desenvolvimento deste trabalho.

Agradeço aos Professores Júnior Vieira (Prof. Titular da Disciplina) e Chistiane Cruz (Diretora da Escola) pelo apoio pleno para que esse trabalho pudesse ser desenvolvido em sua escola. Estes professores tem em comum o respeito para com a educação pública de qualidade, de mente e corações abertos para o melhoramento constante de sua escola.

Agradeço aos amigos Wellington Wagner Sarmiento e George Allan pelo apoio e discussões de ideias, sempre oportunas.

Agradeço imensamente a todos os professores do Programa MDCC da UFC, pela valorosa formação à qual tive o privilégio de me submeter. Agradeço aos funcionários pela paciência e gentileza de sempre. Aos alunos monitores das disciplinas, pela generosidade em em nos ajudar com tanto desprendimento durante a oferta de disciplinas.

Agradeço ao Presidente LULA por ter ampliado o ensino superior público deste país, dando oportunidades a todas as classes sociais, sobretudo as mais desprovidas economicamente. A este verdadeiro brasileiro, meu muito obrigado.



## RESUMO

A popularização das tecnologias da informação e comunicação (TIC) tem suscitado sua utilização em ambientes educacionais como escolas e universidades. No entanto, o emprego isolado de tecnologias, sem metodologia adequada, não tem garantido um efetivo ganho do aluno em termos de aprendizagem. Na última década, um campo de investigação tem ganho destaque na reflexão da questão do uso de tecnologias no contexto educacional: o Ensino Híbrido (*Blended Learning*). O Ensino Híbrido implica na associação de tecnologias às experiências de ensino presencial e virtual, promovendo um novo cenário que seja disruptivo ao modelo clássico de educação. A mudança paradigmática possibilitada por uma experiência real de ensino híbrido deve modificar o formato de sala de aula, transcendendo o escopo de atuação espacial e temporal. Diversos modelos de ensino híbrido têm sido estabelecidos pela comunidade acadêmica mundial, sem ainda o estabelecimento de um padrão. Neste sentido, a presente tese apresenta o novo modelo conceitual, denominado *Digital Class*, voltado à criação de cursos no paradigma de ensino híbrido. O Modelo define seis dimensões que objetivam promover, verdadeiramente, um comportamento reformista frente ao ensino tradicional. Dentre as dimensões estabelecidas, destacam-se a personalização de atendimento e a predição de desempenho do aluno. Foram utilizadas técnicas de sistemas de recomendação para atuar tanto na oferta de objetos educacionais, como na predição de desempenho. Este modelo foi validado em duas etapas. A primeira culminou com o desenvolvimento de uma arquitetura computacional que implementa os conceitos definidos pelo Modelo *Digital Class*. A segunda consistiu na utilização da arquitetura por uma disciplina de Física, pertencente à dois cursos de uma Escola Pública Profissionalizante do Estado do Ceará. Os resultados obtidos dos experimentos de validação do Modelo *Digital Class*, para todas as suas dimensões, foram positivos e norteadores para o aprimoramento dos conceitos do Modelo e melhoramentos na Arquitetura.

**Palavras-chave:** Ensino Híbrido. Sistemas de Recomendação. Ambiente Virtual de Aprendizagem.

## ABSTRACT

The popularization of information and communication technologies (ICT) has promoted its use in educational environments such as schools and universities. However, the isolated use of technologies, without an adequate methodology, has not guaranteed an effective gain of the student in terms of learning. In the last decade, a field of research has gained prominence in the subject of the use of technologies in the educational context: Blended Learning. Blended Learning implies in the association of technologies, the experience of face-to-face and virtual teaching, promoting a new scenario that is disruptive to the classic model of education. A paradigmatic change, that has been made possible by a real experience of hybrid teaching, should modify the classroom format, transcending the scope of spatial and temporal performance. Several models of hybrid education have been established by the global academic community, yet without establishing a standard. In this sense, the current thesis presents the new conceptual model, called Digital Class, aimed at the creation of courses in the paradigm of hybrid teaching. The Model defines six dimensions that aim to truly promote a reformist behavior towards traditional teaching. Among the dimensions, we highlight the personalization of student assistance and the prediction of student performance. This model was validated in two steps: the first finished in the development of a computational architecture that implements the concepts defined by the Digital Class Model. The second one is the use of the architecture by a discipline of Physics, which belongs to the two courses of a Professional Public School of the State of Ceará. The results obtained from the validity tests of the Digital Class Model, were positive and helped to improve both the Model concepts and the Architecture.

**Keywords:** Blended Learning. Recommender Systems. Learning Management System.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Modelos de Ensino Híbrido . . . . .	29
Figura 2 – Exemplo do Modelo Rotação de Estação na KIPP LA Empower Academy. . .	30
Figura 3 – Exemplo de Laboratório <i>Rotacional Rocketship Educational</i> . . . . .	31
Figura 4 – Exemplo de Sala de Aula Invertida <i>Stillwater Area Public Schools</i> . . . . .	31
Figura 5 – Exemplo de Rotação Individual <i>Carpe Diem Collegiate High School and MiddleSchool</i> . . . . .	32
Figura 6 – Exemplo de Modelo <i>Flex, San Francisco Flex Academy</i> . . . . .	33
Figura 7 – Exemplo de Modelo à <i>La Carte, Quakertown Community School District</i> . . .	34
Figura 8 – Exemplo de Modelo Virtual Enriquecido, Albuquerque <i>e-Cademy</i> . . . . .	35
Figura 9 – Exemplo de um Modelo Classificador. . . . .	38
Figura 10 – Exemplo de uma Árvore de Decisão. . . . .	39
Figura 11 – Exemplo de um Classificador Bayesiano. . . . .	39
Figura 12 – Exemplo de Naive Bayes. . . . .	40
Figura 13 – Exemplo de uma arquitetura de RNA. . . . .	41
Figura 14 – Funcionamento do SVM e de seus vetores de Suporte. . . . .	43
Figura 15 – Representação de uma Regressão Linear. . . . .	44
Figura 16 – Diferentes formas de representar <i>clusters</i> . . . . .	45
Figura 17 – Fluxo de processo em quatro etapas AEH. . . . .	52
Figura 18 – Arquitetura Genérica de AEH. . . . .	53
Figura 19 – Fluxo do Protocolo de Revisão . . . . .	56
Figura 20 – Arquitetura ALAS-KA . . . . .	64
Figura 21 – Processo KDD . . . . .	66
Figura 22 – Número de publicações usando cada tipo de análise de dados. . . . .	68
Figura 23 – Fluxo de atividades. . . . .	68
Figura 24 – Relação de Atividades e Objetos segundo a condição temporal. . . . .	69
Figura 25 – Um <i>framework</i> para análise de abordagens instrucionais em BL. . . . .	70
Figura 26 – Componentes do Modelo de Múltiplas Perspectivas. . . . .	71
Figura 27 – Modelo de <i>Flipped Classroom</i> . . . . .	72
Figura 28 – Procedimento utilizado para caracterização do sistema de aprendizado web personalizado. . . . .	76
Figura 29 – Arquitetura <i>Digital Class</i> . . . . .	89

Figura 30 – Criação de Atividades de um Curso . . . . .	90
Figura 31 – Visão do aluno no módulo Gryphon. . . . .	93
Figura 32 – Plano de Estudos . . . . .	97
Figura 33 – Tela de exibição de mídias educacionais. . . . .	97
Figura 34 – Filtro de Itens na Seção Biblioteca . . . . .	99
Figura 35 – Gráficos de situação do desenvolvimento do estudante nos temas do curso. . .	100
Figura 36 – Ferramenta de Autoria de Slides (HTML 5) . . . . .	116
Figura 37 – Gravação de uma aula no Chimera. . . . .	116
Figura 38 – Aplicação de um teste <i>online</i> . . . . .	117
Figura 39 – Relatório pormenorizado após a aplicação de um teste online. . . . .	118
Figura 40 – Gráfico de Evolução do Desempenho Discente no tempo. . . . .	119
Figura 41 – Algoritmo SVD <sup>++</sup> utilizado para modelar a sistemática de predição. . . . .	126
Figura 42 – Fluxo de atividades sugerido para aplicação do Modelo Digital Class. . . . .	127
Figura 43 – Nota dos alunos obtida pelos grupos. . . . .	135
Figura 44 – Relação entre o número médio de temas trabalhados pelos grupos ao longo das aplicações. . . . .	138
Figura 45 – Gráficos de Dispersão das Aplicações (Relação de Notas e Número de Temas Trabalhados) . . . . .	139
Figura 46 – Gráficos de Evolução dos sete alunos com maiores médias. . . . .	141
Figura 47 – Gráficos de Evolução das notas médias. . . . .	141
Figura 48 – Fonte de Acerto × Resultado Observado na Questão-Prova . . . . .	143
Figura 49 – Chance de Acerto x Proporção de Acerto na Bimestral (Alunos que erraram a questão-prova) . . . . .	145
Figura 50 – Chance de Acerto x Proporção de Acerto na AV Bimestral (alunos que acertaram a questão-prova) . . . . .	145
Figura 51 – Chance de Acerto vs Resultado da Questão-prova (Tema 86). . . . .	146
Figura 52 – Chance de Acerto x Proporção de Acerto na Bimestral (alunos que erraram)	147
Figura 53 – Chance de Acerto x Proporção de Acerto na AV Bimestral (alunos que acertaram a questão-prova) . . . . .	147
Figura 54 – RMSE x Número de interações . . . . .	148
Figura 55 – RMSE x Número de interações . . . . .	148
Figura 56 – RMSE x Número de variáveis latentes . . . . .	149

Figura 57 – RMSE x Número de variáveis latentes . . . . .	149
Figura 58 – RMSE x Termo de Regularização ( $\lambda$ ) . . . . .	150
Figura 59 – Boxplot dos Totais de Questões Resolvidas, Tentativas e Acertos obtidos. . . . .	152
Figura 60 – Dispersão do viés (erro) ao longo das iterações. . . . .	153
Figura 61 – Dispersão do viés (erro) ao longo das iterações. . . . .	153
Figura 62 – Dispersão do viés (erro) ao por usuário (Temas 85 e 86). . . . .	155
Figura 63 – SVD <sup>++</sup> (Treinamento $\times$ Validação) . . . . .	158
Figura 64 – RMSE Regressor Logístico x SVD <sup>++</sup> (Temas 85 e 86) . . . . .	159
Figura 65 – Evolução de notas de cada grupo de alunos ao longo das aplicações. . . . .	161
Figura 66 – Planos de estudo (Aluno 67 e 94). . . . .	162
Figura 67 – Visão dos alunos sobre a utilidade de ferramentas tecnológicas na escola. . . . .	163
Figura 68 – Auxílio da <i>Digital Class</i> na organização das atividades. . . . .	164
Figura 69 – Auxílio ao Aprendizado . . . . .	165
Figura 70 – Auxílio ao Aprendizado . . . . .	165
Figura 71 – Opinião do aluno com relação à quantidade de questões recomendadas pela <i>Digital Class</i> . . . . .	166
Figura 72 – Utilidade dos vídeos para o aprendizado. . . . .	167
Figura 73 – Visão do aluno em relação à quantidade de vídeos disponibilizados. . . . .	168
Figura 74 – Utilização da <i>Digital Class</i> fora do ambiente escolar. . . . .	169
Figura 75 – Utilização da <i>Digital Class</i> para revisão de conteúdos. . . . .	170
Figura 76 – Dificuldade de utilização da plataforma. . . . .	170
Figura 77 – Recomendação de adoção da <i>Digital Class</i> a outras disciplinas . . . . .	172

## LISTA DE TABELAS

Tabela 1 – Exemplos de aplicação de gamificação em processos educativos . . . . .	65
Tabela 2 – Conteúdos abordados. . . . .	121
Tabela 3 – Relação entre estudante $\times$ tentativa $\times$ acerto de questões . . . . .	122
Tabela 4 – Relação entre estudante $\times$ tentativa $\times$ acerto de questões $\times$ nota Geral $\times$ Nota Padronizada . . . . .	123
Tabela 5 – Matriz esparsa com as 10 primeiras observações das performances observadas.	123
Tabela 6 – Número de performances observadas por cada conteúdo . . . . .	124
Tabela 7 – Amostra de tuplas dos dados que do conjunto de treinamento $\mathbb{D}^{treino}$ . . . . .	124
Tabela 8 – Amostra de tuplas dos dados do conjunto de validação $\mathbb{D}^{validação}$ . . . . .	124
Tabela 9 – Avaliação de Dimensões do Modelo . . . . .	132
Tabela 10 – Avaliação média dos alunos aos objetos didáticos recomendados ( <i>rating</i> ). . . . .	134
Tabela 11 – Relação de dados obtidos das aplicações (temas trabalhados $\times$ nota). . . . .	135
Tabela 12 – Teste não paramétrico para a comparação das notas de desempenho dos grupos 1 e 2 em cada uma das aplicações. . . . .	136
Tabela 13 – Estatísticas básicas das notas médias de desempenho de cada aluno na aplicação 6. . . . .	137
Tabela 14 – Estatísticas básicas das notas médias de desempenho de cada aluno na aplicação 5. . . . .	137
Tabela 15 – Estatísticas básicas das notas médias de desempenho de cada aluno na aplicação 4. . . . .	137
Tabela 16 – Média de Temas Trabalhados por alunos durante as aplicações. . . . .	138
Tabela 17 – Análise de Estimativas . . . . .	143
Tabela 18 – Comparação entre a chance de acerto dimensionada com a nota obtida na bimestral. . . . .	144
Tabela 19 – Análise das estimativas . . . . .	146
Tabela 20 – Dados da interação dos alunos e previsão de acerto do modelo. . . . .	147
Tabela 21 – Tabela de valores de parâmetros utilizados para geração do modelo . . . . .	151
Tabela 22 – Tabela de valores de parâmetros utilizados para geração do modelo. . . . .	152
Tabela 23 – Dados de Treinamento . . . . .	154
Tabela 24 – Dados de Validação . . . . .	154
Tabela 25 – Notas observadas dos exercícios (conjunto treinamento). . . . .	154

Tabela 26 – Estimativa do modelo e resultado efetivo na bimestral (Temas 85). . . . .	155
Tabela 27 – Estimativa do modelo e resultado efetivo na bimestral (Temas 86). . . . .	155
Tabela 28 – Teste de Normalidade. . . . .	156
Tabela 29 – Teste Não Paramétrico Wilcoxon. . . . .	157
Tabela 30 – do RMSE (Temas 85 e 86). . . . .	157
Tabela 31 – Valores de RMSE obtidos através das comparação das notas observadas na avaliação bimestral com as previsões fornecidas pelos métodos SVD <sup>++</sup> e Regressão Logística. . . . .	159
Tabela 32 – Mapeamento de componentes entre Modelo Conceitual e Arquitetura Digital Class . . . . .	193
Tabela 33 – Mapeamento de componentes entre Modelo Conceitual e Arquitetura Digital Class (continuação) . . . . .	194
Tabela 34 – Principais técnicas de filtragem contribuem com sistemas AEH. . . . .	196

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>18</b>
<b>1.1</b>	<b>Motivação e caracterização do problema</b>	<b>18</b>
<b>1.2</b>	<b>Hipótese e Questões de Partida</b>	<b>22</b>
<b>1.3</b>	<b>Objetivos</b>	<b>23</b>
<b>1.4</b>	<b>Metodologia</b>	<b>24</b>
<b>1.5</b>	<b>Escopo</b>	<b>24</b>
<b>1.6</b>	<b>Contribuição da tese</b>	<b>25</b>
<b>1.7</b>	<b>Organização da tese</b>	<b>26</b>
<b>2</b>	<b>ANÁLISE FATORIAL CLÁSSICA E ALGUNS RESULTADOS</b>	<b>28</b>
<b>2.1</b>	<b>Ensino híbrido (<i>blended learning</i>)</b>	<b>28</b>
<b>2.2</b>	<b>Mineração de dados educacionais</b>	<b>35</b>
<b>2.2.1</b>	<i>Predição</i>	<b>37</b>
<b>2.2.2</b>	<i>Agrupamento</i>	<b>44</b>
<b>2.3</b>	<b>Sistemas de recomendação</b>	<b>45</b>
<b>2.3.1</b>	<i>Sistemas de Recomendação Sensíveis ao Contexto (Context-aware Recommender Systems – CARS)</i>	<b>47</b>
<b>2.3.2</b>	<i>Sistemas de recomendação educacionais – SRE</i>	<b>50</b>
<b>2.3.2.1</b>	<i>Sistemas Adaptativos de Hiperídia Educacional (AEH)</i>	<b>51</b>
<b>2.3.2.2</b>	<i>Técnicas utilizadas em SRE'S</i>	<b>53</b>
<b>3</b>	<b>TRABALHOS RELACIONADOS</b>	<b>56</b>
<b>3.1</b>	<b>Processo de revisão sistemática</b>	<b>56</b>
<b>3.1.1</b>	<i>Planejamento e revisão</i>	<b>57</b>
<b>3.1.2</b>	<i>Condução da revisão</i>	<b>58</b>
<b>3.1.3</b>	<i>Resultado da revisão</i>	<b>59</b>
<b>3.2</b>	<b>Visão geral dos estudos</b>	<b>59</b>
<b>3.3</b>	<b>Resultado das questões da pesquisa</b>	<b>67</b>
<b>3.3.1</b>	<i>Questão principal: qual o estado da arte de modelos de blended learning?</i>	<b>67</b>
<b>3.3.2</b>	<i>QP01: De que forma a tecnologia tem propiciado a implantação de inovações metodológicas para a melhoria do processo de ensino e aprendizagem em cenários de Blended Learning?</i>	<b>73</b>



3.3.3	<i>QP02: Quais aspectos do campo educacional podem ser diretamente trabalhados e aprimorados pela utilização de TEL no contexto de Blended Learning?</i> . . . . .	77
3.3.4	<i>QP03: Como viabilizar um processo de ensino e aprendizagem personalizado?</i> . . . . .	78
3.3.5	<i>QP04: Como atender aos principais requisitos necessários à implantação de uma solução tecnológica de Blended Learning?</i> . . . . .	79
4	<b>MODELO CONCEITUAL PARA ELABORAÇÃO DE CURSOS EM BLENDED LEARNING – DIGITAL CLASS</b> . . . . .	81
4.1	<b>Apresentação do Modelo Conceitual <i>Digital Class</i></b> . . . . .	81
4.1.1	<i>Adequabilidade a objetivos do curso</i> . . . . .	82
4.1.2	<i>Avaliação contínua</i> . . . . .	83
4.1.3	<i>Paralelização de estudos</i> . . . . .	84
4.1.4	<i>Ritmo próprio</i> . . . . .	85
4.1.5	<i>Predição de desempenho</i> . . . . .	86
4.1.6	<i>Personalização</i> . . . . .	87
4.2	<b>Arquitetura <i>Digital Class</i></b> . . . . .	88
4.2.1	<i>Módulo de autenticação</i> . . . . .	89
4.2.2	<i>Gryphon</i> . . . . .	89
4.2.3	<i>Sistemas de recomendação</i> . . . . .	100
4.2.3.1	<i>Questões técnicas</i> . . . . .	101
4.2.3.2	<i>Modelo matemático</i> . . . . .	104
4.2.3.3	<i>Modelo Preditivo</i> . . . . .	106
4.2.4	<i>Chimera</i> . . . . .	115
4.2.5	<i>Sphinx</i> . . . . .	117
4.2.6	<i>Pegasus</i> . . . . .	118
5	<b>METODOLOGIA DA VALIDAÇÃO DO EXPERIMENTO</b> . . . . .	120
5.1	<b>Primeira Etapa da Validação</b> . . . . .	120
5.1.1	<i>Construção da arquitetura</i> . . . . .	120
5.1.2	<i>Métodos Utilizados para a Previsão de Desempenho dos Estudantes</i> . . . . .	121
5.2	<b>Segunda Etapa da Validação</b> . . . . .	126
5.2.1	<i>Processo de Aplicação do Modelo <i>Digital Class</i></i> . . . . .	126

<b>5.3</b>	<b>Caracterização Metodológica do Estudo</b> . . . . .	130
<b>5.3.1</b>	<i>Descrição da Avaliação do Modelo</i> . . . . .	131
<b>5.3.2</b>	<i>Análise de Satisfação dos Usuários com a Arquitetura</i> . . . . .	131
<b>6</b>	<b>RESULTADOS E DISCUSSÕES</b> . . . . .	133
<b>6.1</b>	<b>Análise da Aplicação da Arquitetura <i>Digital Class</i> e Validação do Modelo</b>	133
<b>6.1.1</b>	<i>Validação da dimensão adequabilidade a objetivos do curso</i> . . . . .	133
<b>6.1.2</b>	<i>Validação da dimensão avaliação contínua</i> . . . . .	134
<b>6.1.3</b>	<i>Validação da dimensão paralelização de estudos</i> . . . . .	137
<b>6.1.4</b>	<i>Validação da dimensão ritmo próprio</i> . . . . .	139
<b>6.1.5</b>	<i>Validação da dimensão predição de desempenho</i> . . . . .	142
<b>6.1.5.1</b>	<i>Modelo de regressão linear logística</i> . . . . .	142
<b>6.1.5.2</b>	<i>Modelo SVD<sup>++</sup></i> . . . . .	148
<b>6.1.5.3</b>	<i>Discussão sobre os modelos preditivos adotados</i> . . . . .	157
<b>6.1.6</b>	<i>Validação da Dimensão Personalização</i> . . . . .	160
<b>6.2</b>	<b>Discussão do instrumento de satisfação aplicado aos alunos</b> . . . . .	162
<b>7</b>	<b>CONSIDERAÇÕES FINAIS</b> . . . . .	174
<b>APÊNDICES</b> . . . . .		192
	<b>APÊNDICE A</b> – Mapeamento de componentes entre Modelo Conceitual e Arquitetura Digital Class. . . . .	192
	<b>APÊNDICE B</b> – Arquitetura Genérica de AEH . . . . .	195

## 1 INTRODUÇÃO

O presente trabalho apresenta uma solução metodológica e tecnológica, denominada *Digital Class*, que está inserida na área de pesquisa de Ensino Híbrido - EH (Blended Learning). Este capítulo de introdução versa sobre questões aderentes ao desenvolvimento desta investigação, como aspectos motivacionais, metodológicos e organizacionais. O capítulo está organizado da seguinte forma: na seção 1.1 a motivação e caracterização do problema é abordada; na seção 1.2 são elencadas as hipóteses que originaram o desenvolvimento desse estudo; na seção 1.3 são apresentados os objetivos e metas estabelecidos; na seção 1.4 é detalhada a metodologia que norteou o desenvolvimento deste trabalho; na seção 1.5 é definido o escopo de investigação; na seção 1.6 são descritas as principais contribuições deste estudo; o capítulo se encerra na seção 1.7, com a apresentação da organização dos demais capítulos que compõem esta tese.

### 1.1 Motivação e caracterização do problema

A popularização das tecnologias da informação e comunicação (TIC) tem ocasionado mudanças comportamentais por parte de pessoas e instituições. Esta constatação pode ser feita de forma trivial ao observarmos as modalidades de serviços permitidos pelos recentes avanços tecnológicos. São exemplos destes novos formatos de serviços, compras *online*, bancos eletrônicos, acesso a vídeos sob demanda, educação *online*, etc.

Tomando como exemplo o contexto educacional, tem-se observado que nas últimas décadas este setor tem feito uso de recursos tecnológicos, como microcomputadores, *tablets*, *smartphones*, *softwares* educativos, dentre outros. A modalidade de Educação a Distância (EaD), em razão de seu contexto distribuído, vem utilizando de maneira mais expressiva as novas tecnologias como forma de instrumentalizar sua realização. Já a modalidade presencial, via de regra, tem repetido modelos tradicionais, em que a informática surge como uma disciplina curricular e não como um instrumento elementar para que as diversas atividades didáticas possam ser desenvolvidas de maneira mais rica e atraente ao estudante.

As escolas atuais, comumente, ainda adotam um modelo em que a infraestrutura tecnológica é uma estrutura a parte, com forte controle de acesso, e que não se integra a um planejamento que vise qualificar o processo educacional, instrumentalizando uma nova postura por parte de alunos, professores e instituições. Muitas escolas têm atualizado seus laboratórios com *tablets*. Entretanto, é muito comum nesses casos a não definição prévia de um propósito

pedagógico que fundamente o porquê da adoção de determinados recursos tecnológicos. Assim, essas instituições mais utilizam estes dispositivos como peças figurativas em atendimento à uma proposta de propaganda do que efetivamente como parte de estratégia metodológica de ensino-aprendizagem (REAL et al.,2013).

O cenário descrito acima começa a ser revisto a partir de iniciativas mundo afora quanto à ressignificação do espaço escolar/acadêmico pela adoção de tecnologias. Tais experiências tratam de desenvolver um estruturado planejamento pedagógico em que o uso de instrumentos computacionais (hardware e software) seja feito de forma criteriosa, possibilitando a mudança de paradigmas presentes historicamente no modelo presencial de ensino (MORAN, 2015). Desta forma, tem ganho projeção uma área de pesquisa denominada de Ensino Híbrido (*Blended Learning* - BL). Hoic-Bozic et al. (2009) descrevem *Blended Learning* como o modelo de aprendizado que se baseia em combinações diversas de aulas clássicas presenciais aliadas a atividades didáticas virtuais, apoiada por tecnologias, objetivando atingir um cenário de maior eficiência da aprendizagem.

Para Bacich & Moran (2015), o híbrido significa misturado, mesclado e, historicamente, o processo educacional sempre se deu desta forma, aliando espaços, tempos, atividades formais e informais, planejamentos intencionais ou não. Com a tecnologia, evidencia-se essa mistura, permitindo um ambiente mais amplo e propenso à criatividade.

O ensino híbrido implica que a associação de tecnologias às experiências de ensino presencial deve se dar de uma forma reformista. Do contrário, a tecnologia a ser empregada nada mais será que uma peça figurativa ou, quando muito, subutilizada. A mudança paradigmática possibilitada por uma experiência real de ensino híbrido deve modificar o formato de sala de aula, transcendendo o escopo de atuação espacial e temporal. Uma sala de aula nesse contexto não contempla mais o enfileiramento linear de cadeiras. A organização dos espaços da escola (salas de aula, laboratórios, bibliotecas, pátios, etc.) passa a ser utilizada de forma associada e mais harmônica. Uma aula já não implica em começar e terminar em determinado horário. A escola já não se limita a seus muros (VALENTE,2014). E sobretudo o professor já não é mais a fonte, quase que exclusiva de informação.

O professor reduzirá seu esforço de explanação para ampliar sua condição de mediador. O aluno terá um papel principal na obtenção de seu conhecimento. Perderá sua passividade para adquirir postura de autonomia e protagonismo de sua formação (CASTRO et al., 2015). A realização de uma disciplina não se dará mais da forma uníssona como hoje é realizada. Ela

se dará de forma segmentada, respeitando individualidades e condições diferentes de acordo com o conhecimento do aluno. Portanto, essas e outras características ensejam uma mudança substancial frente aos atuais moldes da educação, dando significado à implantação desse novo paradigma educacional (DZIUBANet al., 2006).

Como apoio à perseguição das características de ensino híbrido, boa parte das experiências educacionais nesse campo investigativo tem feito uso de sistemas *online* comumente adotados em cursos a distância: os AVAs (Ambientes Virtuais de Aprendizagem) (SHEN et al., 2009). Os AVAs popularizaram-se no final da década de 1990 como importantes instrumentos acadêmico-administrativos. A diversidade de funcionalidades desses sistemas, aliada à facilidade de uso de seus recursos, fizeram com que os cursos a distância se popularizassem através da Web. Diversos AVAs foram desde então desenvolvidos no Brasil, destacando-se o Eproinfo (MEC) <sup>1</sup>, Teleduc (Unicamp) <sup>2</sup>, AulaNet (PUC-RIO) <sup>3</sup> e SOLAR (UFC) <sup>4</sup>. Destaque é dado para este último, que foi desenvolvido pela Universidade Federal do Ceará no ano de 2003, e está hoje em sua segunda versão, atendendo a cerca de 50.000 usuários. Esse sistema atende sobretudo aos cursos a distância da instituição. No entanto, já se inicia seu uso para os cursos presenciais, como uma ferramenta complementar às atividades face a face. Apesar dos casos anteriormente indicados de autoria própria de instituições nacionais, a plataforma *Moodle*, sistema de código aberto, é um dos AVAs mais adotados em todo o mundo (MAIA et al., 2015).

Os AVAs, em muitos casos, têm sido utilizados como repositórios de arquivos de uma disciplina. Muitas experiências relatam o uso de AVA como um elemento extra e não como parte importante de um planejamento didático mais eficiente. Nesses casos, não há ruptura paradigmática.

Com a popularização dos cursos abertos, grandes universidades, sobretudo as americanas, desenvolveram de forma consorciada suas plataformas para a oferta desses cursos. Os MOOCs (*Massive Open Online Course*) representaram um grande avanço na disponibilização de cursos *online* no mundo, sobretudo pela chancela trazida por grandes Universidades como Stanford, MIT, Harvard, dentre outras. Sua proposta é atender a um vasto número de alunos ao redor do mundo. Críticas são feitas ao aspecto auto-instrucional de sua formação, sem o acompanhamento dos cursistas. Há cursos que possuem mais de 20.000 alunos (FERGUSON; CLOW, 2015). Portanto, não há como dimensionar e dedicar esforço humano para atender

---

<sup>1</sup> <http://www.eproinfo.mec.gov.br>

<sup>2</sup> <http://gcte.unicamp.br/teleduc/>

<sup>3</sup> <http://web.ccead.puc-rio.br/aulanet2/>

<sup>4</sup> <http://solar.virtual.ufc.br>

a esse alto contingente de cursistas. No entanto, o lado positivo é ter acesso a cursos oferecidos por professores de referência no mundo, fato que, sem a tecnologia, não seria possível (MARGARYAN et al., 2015).

As experiências de BL encontradas na literatura têm sido feitas, em sua grande parte, sobre plataformas tecnológicas que não foram originalmente desenvolvidas para esse domínio. Se, por um lado, isso garante uma série de facilidades aos professores, por outra é responsável por causar descontentamentos e frustrações por não permitir que determinadas necessidades pertinentes ao ensino híbrido sejam atendidas. Um exemplo disto é a ferramenta *Khan Academy* (KHAN, 2012). A ferramenta é vastamente utilizada por experiências de BL e apresenta uma gama robusta de exercícios, aulas e relatórios. Sua abordagem também prevê a sistemática de “gamificação”, o que proporciona um engajamento motivacional bem maior dos estudantes. Mesmo essa solução sendo bem mais adequada ao contexto de ensino híbrido, ela possui uma natureza que não permite que professores ajustem de forma mais refinada a questões mais específicas, como, por exemplo, a personalização de atendimento. Ainda que a plataforma seja rica na oferta de exercícios e materiais didáticos em diversos temas, ela ainda deixa a desejar em muitas situações, pois não permite que o professor utilize ou elabore materiais próprios. A plataforma, por exemplo, é muito apropriada para um professor de matemática adotá-la, mas para um de física, já deixa a desejar. Bernatek et al. (2012) verificaram que a plataforma *Khan* é muito útil para se trabalhar o nivelamento de alguns assuntos com os estudantes e auxiliar o professor no diagnóstico constante sobre a evolução do aluno a partir da realização dos exercícios. No entanto, não se mostra tão eficiente para aprofundamentos de assuntos. É comum que alunos avancem rapidamente na conclusão de objetivos estabelecidos e fiquem frustrados ao não serem mais desafiados, quando ainda há muito o que se aprofundar em determinados assuntos.

Os ambientes virtuais de aprendizagem, quando adotados por cursos de ensino híbrido, atendem parcialmente aos requisitos que norteiam uma boa prática nesse domínio. Destaca-se nesses sistemas a ausência de instrumentos que permitam um atendimento sob medida às necessidades de cada cursista. Ainda que muitos teóricos da educação a distância ponderem que esta seja uma premissa desta modalidade, o fato concreto é que os cursos de EaD, reproduzem os mesmos aspectos generalistas vistos na modalidade presencial (HILL, 2012). Exemplo disto é um planejamento único da disciplina, com material didático igual a todos os cursistas, mesmas atividades avaliativas, etc.

As plataformas de cursos abertos, por exemplo, devido a seu atendimento a cursistas

em larga escala, necessitaria de uma condição de atendimento didático automático que pudesse auxiliar os cursistas e acompanhá-los ao longo de um curso (ALEVEN et al., 2016). Uma vez que é recomendável que o professor virtualize parte ou totalmente suas aulas expositivas em BL, os AVAs comumente não oferecem esse recurso ao professor. A elaboração do conteúdo didático tem que ser feita via um outro sistema e ao final publicado no AVA (MAIA et al., 2015).

As restrições listadas anteriormente demonstram a importância de que os ambientes virtuais de aprendizagem sejam mais aderentes aos requisitos presentes em ensino híbrido. Desta forma, o presente estudo contribui com a geração de um modelo que estabelece bases conceituais para a realização de experiências consistentes em ensino híbrido. No entanto, este modelo precisou ser validado e, dadas as restrições citadas anteriormente pelas plataformas tecnológicas existentes, foi necessário o desenvolvimento de uma arquitetura própria, modelada e desenvolvida especificamente para atender as dimensões do modelo. Dentre as funcionalidades não encontradas nos AVA's, destaca-se o uso de técnicas de aprendizado de máquina e sistemas de recomendação para auxiliar na avaliação dos alunos e na realização de previsões de desempenho estudantil. O sistema oferece também um atendimento personalizado, apresentando a cada cursista de uma turma um plano de estudos próprio que leva em consideração a condição cognitiva do estudante e o planejamento e objetivos da disciplina. O professor, além de ter um grande aliado no atendimento discente, pode acompanhar a condição dos estudantes em tempo real, ao longo de toda a disciplina. Um módulo de autoria de material didático, cadastro de conteúdos externos e repositório de itens também compõem a solução computacional proposta ao domínio de EH.

A solução foi aplicada em dois cursos (Informática e Produção de Áudio e Vídeo) de uma escola profissionalizante do Estado do Ceará, em uma disciplina de Física. A experiência consistiu na utilização da Arquitetura pelas turmas de alunos durante o primeiro semestre de 2017.

## **1.2 Hipótese e Questões de Partida**

Dado que a inovação trazida pelo ensino híbrido à escolas e universidades passa, necessariamente, pela mudança paradigmática trazida por um planejamento estratégico em que coloca a tecnologia como vetor desta mudança, esta pesquisa testa a Hipótese Inicial definida abaixo:

**Hipótese Inicial:** *É possível atingir um estado de satisfação de professores e alunos de um curso, em formato de ensino híbrido, através da oferta de uma solução metodológica, instrumentada por tecnologia, que respeita os objetivos do curso e oferece aos alunos uma condição de personalização de estudos, com apoio automático para a sua preparação, planos de estudos individualizados e a garantia de um estudo com ritmo próprio, além de auxiliar o professor na avaliação contínua e previsão de desempenho de seus alunos.*

A partir desta hipótese, quatro questões de partida foram definidas:

- QP01: Quais são as principais demandas existentes no contexto de ensino híbrido que devem ser atendidas para termos uma experiência mais rica nesse domínio?
- QP02: É possível utilizar técnicas de aprendizagem de máquina que permitam auxiliar o professor e aluno nas principais demandas de EH?
- QP03: É possível implementar uma sistemática para que o aluno tenha uma experiência personalizada?
- QP04: É possível desenvolver uma solução unificada em que todas as principais demandas de EH sejam reunidas em uma só solução computacional?

### 1.3 Objetivos

Esta tese tem como objetivo principal propor um modelo conceitual voltado ao desenvolvimento de ações educativas no contexto de ensino híbrido. Para a obtenção deste objetivo, foram definidos os seguintes objetivos específicos:

- a. Realizar revisão bibliográfica sobre experiências de ensino híbrido (QP01);
- b. Realizar revisão bibliográfica sobre técnicas de aprendizagem de máquina voltadas ao contexto educacional (QP02);
- c. Desenvolver sistema de recomendação para oferta de planos de estudos personalizados aos estudantes (QP03);
- d. Elaborar um modelo conceitual e suas dimensões para orientar a criação de cursos híbridos (QP04);
- e. Desenvolver módulos de arquitetura e integração entre eles de forma a validar o modelo e a permitir que os usuários (alunos e professores) tenham uma experiência completa e facilitada de uso da solução (QP04).



## 1.4 Metodologia

A elaboração deste trabalho seguiu a seguinte metodologia:

- a. Estudo da fundamentação teórica necessária para o desenvolvimento da tese;
- b. Realização de revisão bibliográfica, segundo uma revisão sistemática de trabalhos relacionados;
- c. Definição de componentes do Modelo *Digital Class* a partir das principais demandas exigidas ao domínio de ensino híbrido;
- d. Implementação dos módulos da arquitetura, garantindo a integração e facilidade de uso por parte dos usuários;
- e. Validação da solução em grupo de teste pertencente à uma escola pública profissionalizante (ensino médio);
- f. Análise de resultados da aplicação e ajustes da proposta.

## 1.5 Escopo

Este trabalho é voltado à discussão sobre a proposição de um modelo conceitual para o desenvolvimento de ações educacionais inseridas na abordagem metodológica de ensino híbrido. Neste sentido, serão apresentados a seguir os assuntos que fazem parte do escopo deste trabalho:

**Modelo de Ensino Híbrido:** este trabalho define um conjunto de dimensões (seis) que associadas compreendem um inovador cenário de aplicação de cursos na metodologia de EH.

**Arquitetura Computacional:** este trabalho valida o Modelo *Digital Class*, através da implementação e aplicação de uma arquitetura computacional voltada à realização de cursos segundo a abordagem de EH. No entanto, não é objetivo deste trabalho enveredar no campo de Engenharia de Software para estabelecer processo analítico sobre as diversas perspectivas de estruturas análogas.

**Aprendizado de Máquina:** este trabalho apresenta o estado da arte do emprego de diversas técnicas de aprendizado de máquina no contexto educacional. Essa área é denominada de Edumining (*educational data mining*). Não é objetivo deste trabalho o desenvolvimento de uma

nova técnica ou aprimoramento em termos de ganho de desempenho computacional de nenhuma das principais técnicas existentes.

**Sistema de Recomendação para Aprendizado Personalizado:** uma das contribuições deste trabalho foca-se no desenvolvimento de um sistema de recomendação para que o aluno receba um roteiro de estudos (plano de estudos), voltado à sua condição particular em termos de conhecimento sobre os conteúdos de um curso.

**Predição de desempenho:** este trabalho implementa uma funcionalidade de previsão de desempenho do estudante. Isto orienta o professor sobre a melhor forma de agir diante da detecção de problemas e sinalizar aos sistemas de recomendação quando e de que forma agir. Modelos preditivos a partir de técnicas sistemas de recomendação poderão ser aprimorados a partir do acúmulo histórico de dados.

## 1.6 Contribuição da tese

Esta tese contribui para o campo investigativo de EH ao oferecer uma solução metodológica e tecnológica que cobre as principais demandas de cursos a serem realizados sob a abordagem de EH. Ao contrário de outras iniciativas, originalmente desenvolvidas para outras finalidades, a *Digital Class* é específica para o EH.

Ela permite que o professor planeje seu curso dimensionando as atividades virtuais e presenciais, construa itens avaliativos, elabore conteúdos didáticos (vídeo-aulas, vídeos de resoluções de exercícios, slides e páginas HTML5), acompanhamento em tempo real de alunos com relatórios e gráficos, sistemas de recomendação para estudantes (recomendação de planos de aula, recomendação de materiais didáticos, recomendação de exercícios/atividades), agrupamento de alunos (orientação ao professor de como distribuir os alunos para a realização de atividades em sala de aula levando em consideração a condição cognitiva de cada cursista), predição de desempenho (permite ao professor visualizar com antecedência a provas qual a expectativa de desempenho de cada aluno), etc.

As oportunidades que poderão ser trazidas a partir da utilização da *Digital Class* são:

- a. Campo exploratório para educadores avaliarem a evolução de indicadores de desempenho discente a partir a realização de EH;
- b. Campo de trabalho para especialistas em mineração de dados educacionais, observando

a base de dados montada a cada temporada de oferta;

- c. A plataforma desenvolvida poderá ser utilizada em rede, compartilhando recursos didáticos entre educadores e instituições (repositório de conteúdos, repositório de questões/itens, etc.);
- d. Os dados gerados com as ofertas dos cursos poderão direcionar a definição de políticas públicas em promoção de inovação metodológica mediada por tecnologia nas instituições de ensino;
- e. Construção de cursos autoinstrucionais para as modalidades presencial e/ou a distância. A plataforma pode atender a um dos pontos mais criticados de cursos autoinstrucionais, a falta de acompanhamento e interação. Os processos automatizados da solução *Digital Class* permitem que o aluno tenha uma orientação guiada à suas necessidades e respeita os objetivos de um curso, oferecendo a dinâmica necessária para auxiliar o estudante no cumprimento dos requisitos exigidos por um curso.

## 1.7 Organização da tese

Este capítulo apresentou as questões motivacionais para a escrita desta tese de doutorado, bem como a hipótese, questões de pesquisa, objetivos, metodologia, escopo e contribuições. O restante deste trabalho é composto por sete capítulos, organizados da seguinte forma:

- a. Capítulo 2: apresenta a fundamentação teórica, com alguns conceitos necessários para o melhor entendimento do trabalho, abordando Ensino Híbrido, Mineração de dados educacionais e Sistemas de Recomendação;
- b. Capítulo 3: apresenta uma revisão sistemática e uma bibliográfica, com a identificação de trabalhos relacionados à Ensino Híbrido, Mineração de dados educacionais e Sistemas de Recomendação;
- c. Capítulo 4: são apresentados o Modelo *Digital Class* com suas dimensões e a arquitetura (homônima) gerada a partir do modelo com seus componentes;
- d. Capítulo 5: apresenta a metodologia utilizada para promover a validação da solução desenvolvida;
- e. Capítulo 6: este capítulo apresenta e discute os resultados obtidos quanto à aplicação da solução na escola, dificuldades encontradas e ações corretivas empreendidas.
- f. Considerações finais: esta seção é dedicada às considerações finais e conclusões desta

tese de doutorado. Nela, são descritos os resultados alcançados e uma discussão sobre as questões da pesquisa. Ao final, trabalhos futuros derivados desta pesquisa são apresentados.

## 2 ANÁLISE FATORIAL CLÁSSICA E ALGUNS RESULTADOS

Neste capítulo são descritos conceitos de base para o entendimento do objeto desta tese. Os assuntos Ensino Híbrido, Mineração de Dados Educacionais e Sistemas de Recomendação serão trabalhados nas seções a seguir.

### 2.1 Ensino híbrido (*blended learning*)

Graham (2006) definiu *Blended Learning* como o processo (sistema) em que são combinados os formatos de instrução presencial (*face-to-face*) com instrução mediada pelo computador. Desta forma, o autor estabelece em sua definição a ideia de que BL une dois modelos de ensino e aprendizagem historicamente separados, o tradicional sistema de ensino presencial com a educação a distância.

Laumakis et al. (2009) definem *Blended Learning* como uma zona de fronteira, em que componentes pertencentes à contextos presencial e a distância se unem, promovendo um novo objeto como novas características.

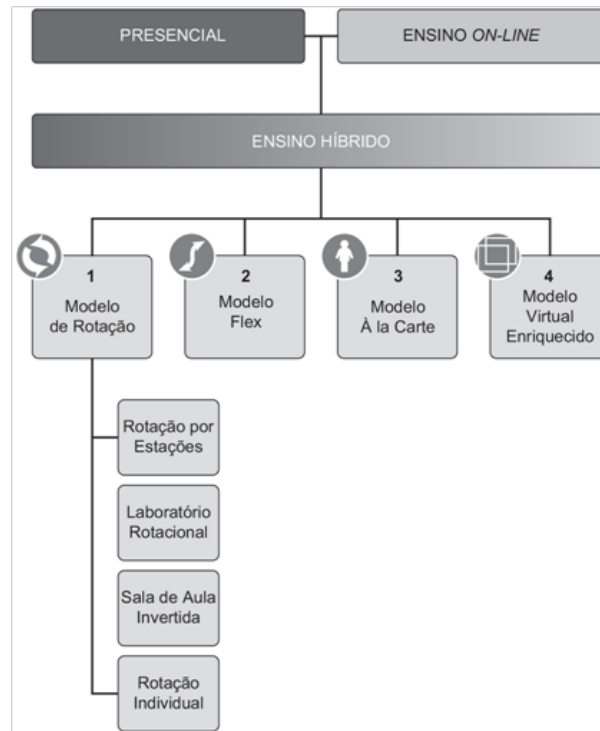
Segundo Bacich et al. (2015) o ensino híbrido está relacionado à ideia de educação híbrida, em que não existe uma forma única de aprender e na qual a aprendizagem é um processo contínuo, que ocorre de diferentes formas, em diferentes espaços.

Staker e Horn (2012) conceituam ensino híbrido como um programa de educação formal no qual um estudante aprende, ao menos em parte, através da fruição do ensino de forma online. Além disso, este programa prevê um acompanhamento do estudante pelo professor, em relação a seu tempo de estudos, ritmo de estudos, locais disponíveis para acompanhamento e trajetórias de estudos adotadas. A condição de formalismo defendida pelos autores, indica que esses materiais devem ser de cunho autoral, ou seja, desenvolvidos especificamente para um curso e não acervos da internet.

O ensino híbrido configura-se, portanto, como uma associação metodológica que impacta em professores e alunos, quanto a mudança na proposta de ensino e na aprendizagem.

Staker e Horn (2012) definem quatro modelos que são comumente utilizados pelas principais experiências de BL, são eles: modelo rotacional, modelo *flex*, modelo *à la carte* (*self-blend*), modelo virtual enriquecido (Figura 1).

Figura 1 – Modelos de Ensino Híbrido

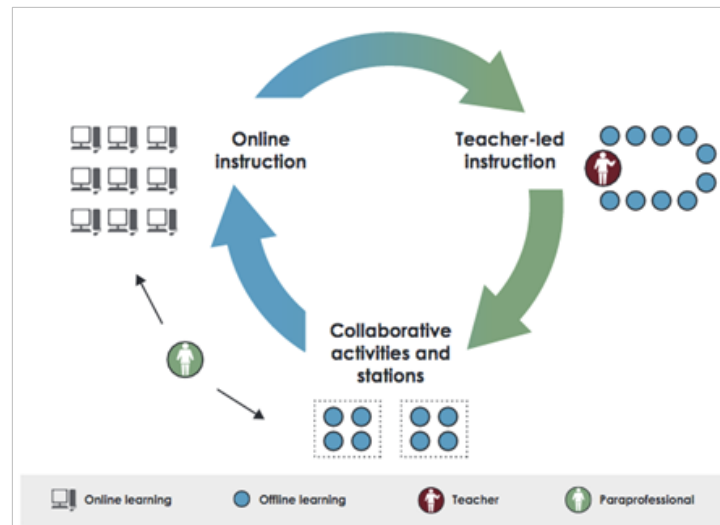


Fonte: Bacich et al.(2015), baseado em Staker e Horn (2012).

O Modelo Rotacional é um programa no qual os estudantes de um curso revezam as atividades a serem realizadas sob orientação ou não do professor, em períodos determinados de tempo. As tarefas podem envolver discussões em grupo, atividades escritas, leituras e, ao menos, uma atividade *online*. Faz parte deste modelo, as seguintes proposições:

- Rotação de Estação (*station rotation*):** neste cenário são definidos pelo professor a divisão da sala de aula em  $n$  estações, representando cada uma, uma experiência didática específica. Os alunos por sua vez, são distribuídos em grupos que são organizados prioritariamente em razão do nível de conhecimento sobre os assuntos a serem trabalhados. Esses grupos ficarão em tempos determinados em cada estação, promovendo um rodízio ao longo do tempo da aula. Dentre as estações, uma que não pode deixar de existir é a *online*. Ela poderá ser composta de conteúdos multimídia como vídeos, animações, jogos, exercícios etc. O professor deve participar de alguma(s) estações, além de utilizar, quando possível, monitores para o acompanhamento de alguns grupos. Há, portanto, a possibilidade de que os alunos trabalhem colaborativamente ou individualmente, com o professor e/ou com monitores, com tecnologia/sem tecnologia, enfim, o ambiente de sala de aula passa a ser múltiplo rompendo com a passividade e desenho único instrucional, fatores característicos do ensino tradicional (Figura 2).

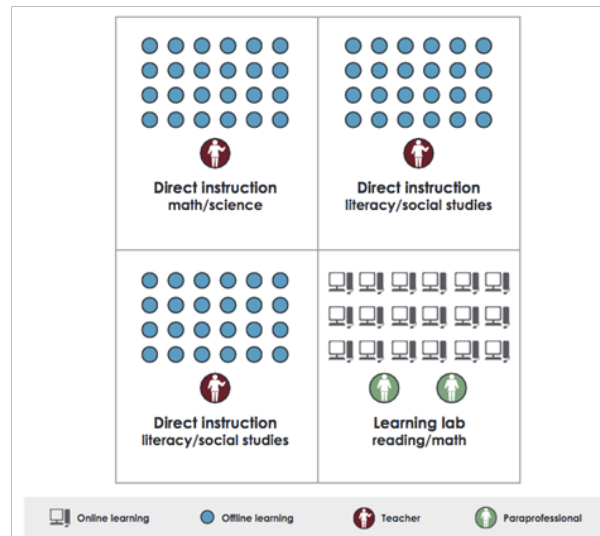
Figura 2 – Exemplo do Modelo Rotação de Estação na KIPP LA Empower Academy.



Fonte – Staker e Horn(2012).

- **Laboratório Rotacional (*rotational laboratory*):** neste modelo, o professor divide a turma, ficando com parte da turma em sala e a outra sendo destinada ao laboratório de informática. Este modelo se assemelha ao de rotação de estações, no entanto, ao invés de fazer subdivisões do espaço de uma mesma sala de aula, separa a turma em dois ambientes distintos (sala de aula e laboratório). Comumente, essa proposta não rompe com o tradicionalismo no que concerne às atividades em sala de aula. Já a prática *online*, contribui com a metodologia tradicional para melhorar o atendimento ao aluno. Um professor tutor deve auxiliar os alunos que se deslocaram para o laboratório de informática.
- **Sala de Aula Invertida (*flipped classroom*):** nesse modelo, o aluno que tradicionalmente recebe instrução através do professor em sala de aula, passa a receber essa instrução no ambiente virtual. Com a liberação do professor da responsabilidade de apresentar conteúdos, este pode desenvolver novas perspectivas didáticas com seus alunos em sala de aula. Desta forma, o professor passa realizar atividades práticas, elaborar projetos, auxiliar alunos de forma segmentada etc. Com isso, também, o aluno ganha maior controle sobre como, quando e onde irá acessar os conteúdos. Esse modelo é tido como a porta de entrada do ensino híbrido, podendo ser adaptado e aprimorado. Nesse sentido, possibilitar ao estudante um contexto de experimentação prévia, estimula os estudantes a pensar de forma crítica, para daí acessar sob uma condição mais adequada os formatos de transmissão de instrução tradicionais, ainda que por formato de vídeos, textos, etc (BACICH et al., 2015). A Figura 4 apresenta um exemplo de sala de aula invertida. A metodologia tradicional de

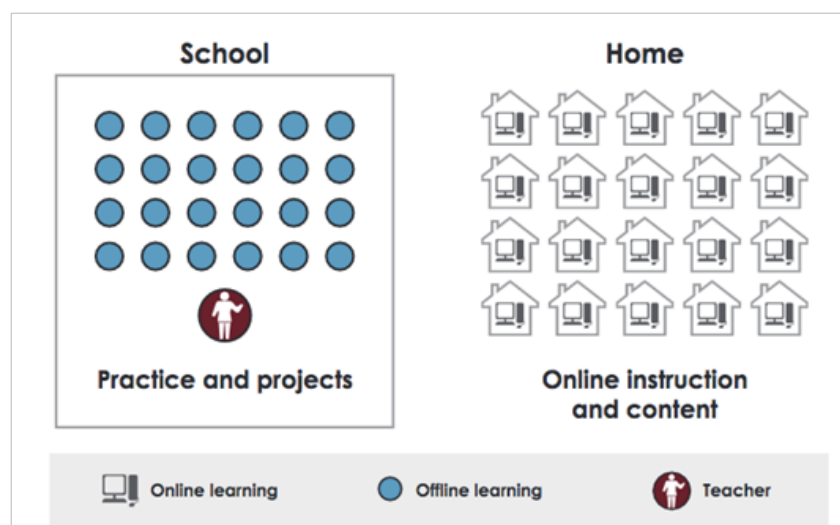
Figura 3 – Exemplo de Laboratório *Rotacional Rocketship Educational*.



Fonte – Staker e Horn, 2012.

ensino oferece um primeiro contato do estudante à um tema através do professor, ficando o aprofundamento no assunto uma tarefa para se fazer em casa pelo aluno, geralmente de forma individual, resolvendo exercícios, fazendo leituras complementares, etc. Já na sala de aula invertida, o primeiro contato com o conteúdo passa a ser *online* e o aprofundamento passa a ser coletivo em sala de aula com colegas e o professor.

Figura 4 – Exemplo de Sala de Aula Invertida *Stillwater Area Public Schools*.



Fonte – Staker e Horn (2012)

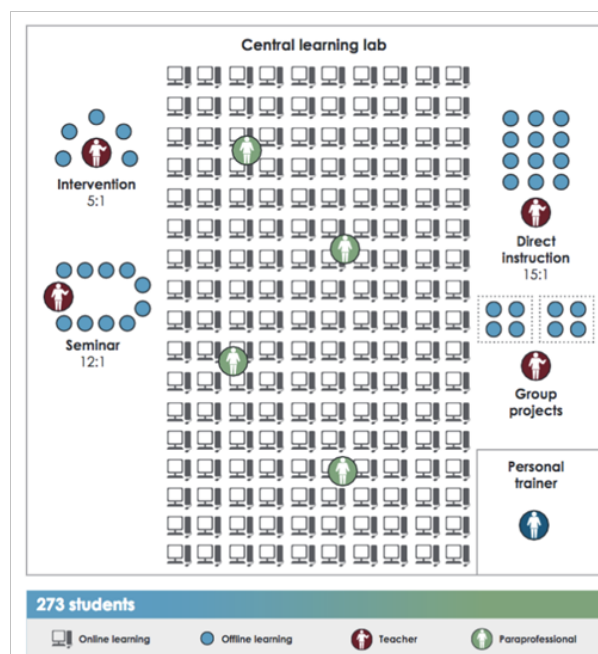
- **Rotação Individual:** neste modelo evidencia-se um dos principais aspectos que proporc-



onam uma ruptura de BL em relação à metodologia tradicional de ensino e aprendizagem, a personalização do processo de aprendizagem dos alunos. Assim, são desenvolvidos planos individuais de estudo a cada aluno. Desta forma, a personalização necessária para a geração desses planos, se dá a partir da avaliação de cada estudante, pois é preciso conhecer primeiro a condição de cada um, para em seguida poder elaborar o planos de estudos adequados aos cursistas. Esses planos podem ser elaborados computacionalmente ou por um ou mais professores. Este modelo difere dos demais, pois não requer que o estudante se desloque entre estações ou modalidade.

Bacich e Moran (2015) apontam que é possível especificar atividades diferentes para grupos de alunos distintos, com ritmos próprios sob o devido acompanhamento docente. Segundo os autores, nos dias atuais há uma grande capacidade de geração e análise de dados que permitem a criação de relatórios personalizados. Além disto, o desenvolvimento de plataformas e aplicativos adaptativos orientam os professores sobre como os alunos aprendem e em que estágio se encontra cada aluno e o que mais o motiva (GOMES, 2013).

Figura 5 – Exemplo de Rotação Individual *Carpe Diem Collegiate High School and MiddleSchool*

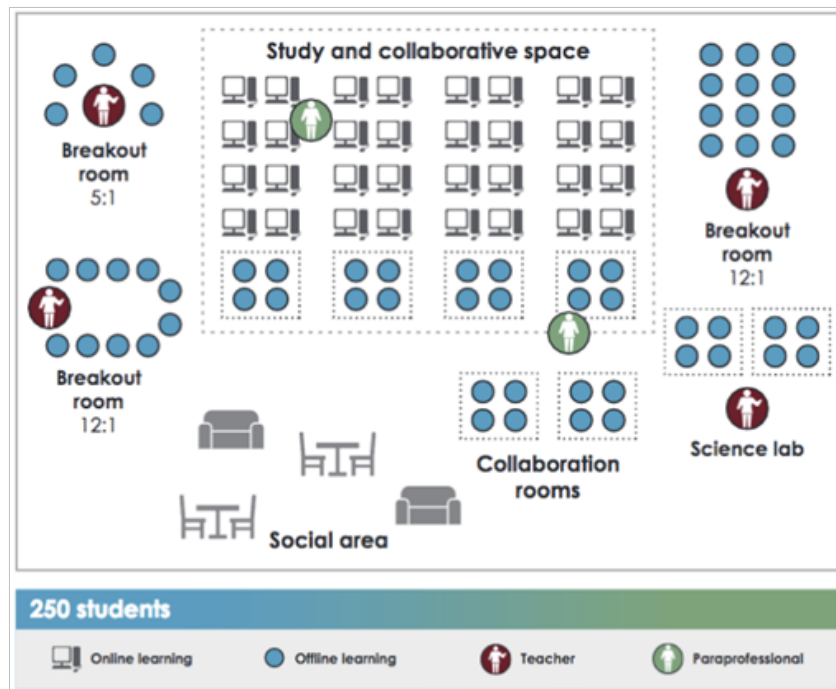


Fonte – Staker e Horn (2012).

O Modelo *Flex* estabelece ao aluno uma lista de atividades para cumprir, ritmo personalizado e apoiado por professores e/ou tutores. São criados grupos específicos para se

trabalhar projetos, auxílio a dúvidas etc. A intensidade de atividades de apoio presencial varia de instituição para instituição, havendo registros de instituições com alta presença do professor junto aos alunos e outras que possuem uma mínima atenção face a face.

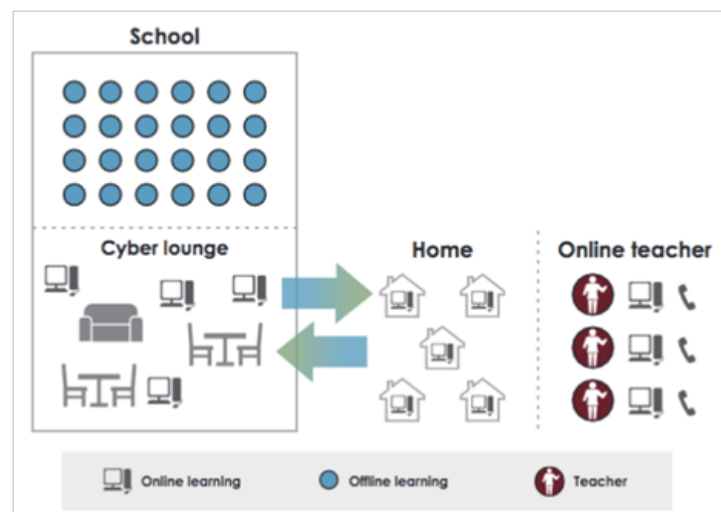
Figura 6 – Exemplo de Modelo *Flex*, *San Francisco Flex Academy*.



Fonte – Fonte: Staker e Horn (2012).

No Modelo à *La Carte* o aluno possui cursos inteiramente *online* que podem ser pré-requisitos ou complementares aos presenciais nos quais esteja matriculado. O aluno contará com suporte *online* de professores. Espera-se também que o aluno tenha franca autonomia para organizar seu tempo de estudos e cumprimento dos requisitos exigidos para os mesmos. O acesso ao curso *online* poderá se dar a partir de ambientes da própria instituição de ensino ou de qualquer outro lugar que o aluno preferir (ex: casa, *lan house*, entre outros).

Figura 7 – Exemplo de Modelo à *La Carte*, *Quaker-town Community School District*.



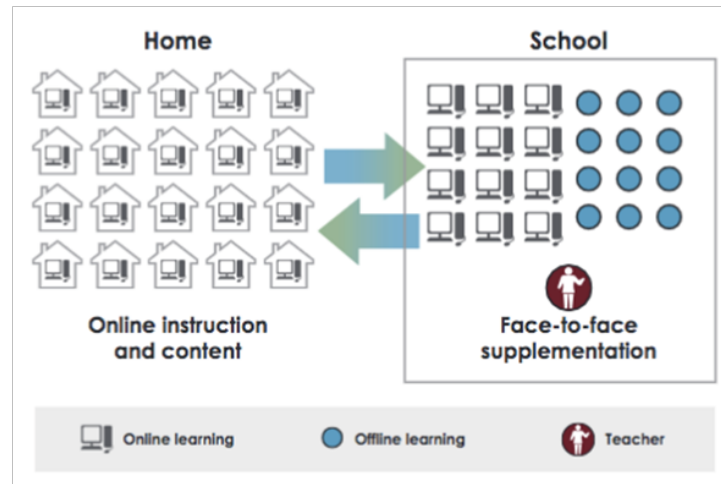
Fonte – Staker e Horn, 2012.

O Modelo Virtual Enriquecido especifica que o estudante divide seu tempo entre atividades *online* e presenciais, sendo o primeiro predominante. No entanto, a parcela presencial é pontual, como, por exemplo, a ida do aluno uma única vez por semana à escola/universidade, ficando o restante do tempo com acesso *online* à disciplina.

Bacich et al. (2015) destacam que não há uma ordem estabelecida para aplicação e desenvolvimento desses modelos em sala de aula, tampouco uma hierarquia entre eles. Experiências demonstram o uso conjugado desses modelos, podendo, por exemplo, um professor adotar uma atividade de sala de aula invertida em um dia e, na aula seguinte, utilizar rotação por estações.

Bacich et al. (2015) também ressaltam que grande parte das propostas apresentadas assemelham-se à trabalhos anteriores, como os centros de interesse de Decroly e os complexos de interesse de Freinet. Ambos já propunham uma organização da sala de aula em espaços que atuavam de forma diferenciada, de acordo com as necessidades dos estudantes, aproximando-se

Figura 8 – Exemplo de Modelo Virtual Enriquecido, Albuquerque *e-Cademy*.



Fonte – Staker e Horn(2012).

do que é proposto na rotação por estações e no laboratório rotacional. Muitas das técnicas propostas por Freinet baseiam-se no respeito ao ritmo de cada criança, bem como estimulam a livre expressão, cooperação e a comunicação. O plano de trabalho é outra técnica apresentada por Freinet e que se aproxima da rotação individual, ao permitir que cada aluno tenha um roteiro moldado à sua condição, abrindo espaço para que o desenvolva sob uma ordem que lhe aprouver (IMBERNÓN, 2010). Nesse cenário, o professor acompanha o progresso de cada aluno e, diariamente, propõe novas atividades à luz dos resultados obtidos.

## 2.2 Mineração de dados educacionais

Mineração de dados (MD) consiste na aplicação de métodos computacionais (algoritmos) voltados ao processamento de grandes quantidades de dados, com o intuito de dar evidência a informações contidas nos dados e de difícil detecção. Essas informações são geradas a partir da identificação de relações entre os dados que podem produzir novos conhecimentos e novas descobertas científicas (BAKER et al.,2011).

A utilização de técnicas estatísticas e computacionais para o processamento e a análise de grandes volumes de dados tem ganhado destaque, sobretudo, pelo uso comercial. São representativos os casos exitosos que fazem uso dessas técnicas para estímulo do desejo de compras das pessoas, ou mesmo para obtenção de insumos para a definição de novas estratégias de negócio (BRAMER, 2013).

Vale ressaltar, porém, que diversas outras áreas tem feito uso de análise de dados

inteligentes, como educação, saúde e segurança (ELMARI;NAVATHE, 2010). Notadamente na área educacional, diversos estudos têm sido realizados com o intuito de contribuir direta ou indiretamente com o aprendizado de alunos. O campo de pesquisa que aplica técnicas de mineração de dados ao contexto educacional é denominado de Educational Data Mining (EDM) (ROMERO; VENTURA, 2007), (BAKER, 2010).

A popularização nas últimas décadas do uso de ambientes virtuais de aprendizagem (AVA's) na educação à distância, oferece a pesquisadores em EDM e gestores acadêmicos uma oportunidade ímpar, uma vez que a quantidade de dados gerados nesses sistemas, permitem que novos indicativos acadêmicos sejam apresentados, como: análise comportamental de estudantes, tutores e professores; definição de quais elementos constituintes dos cursos são efetivamente mais importantes à formação do estudante; quais variáveis são mais significativas para indicar uma possível evasão; elaboração de modelos preditivos de desempenho do estudante; personalização de atendimento pedagógico etc.

São diversas as finalidades possíveis quanto à aplicação das técnicas de MD voltadas ao contexto educacional. Baker (2010) apresenta uma taxonomia das principais sub-áreas de pesquisa em EDM:

- Predição
  - Classificação
  - Regressão
  - Estimação de Densidade
- Agrupamento (clustering)
- Mineração de Relações
  - Regras de Associação
  - Mineração de Correlações
  - Mineração de padrões sequenciais
  - Mineração de causas
- Destilação de dados
- Descobertas com modelos

Como forma de discutir os principais pontos indicados na taxonomia de Baker, fazemos a descrição das técnicas que de forma direta ou indireta associam-se ao objeto deste trabalho.

### 2.2.1 Predição

A atividade de predição consiste na elaboração de modelos que, a partir da definição de variáveis preditoras, possa inferir sobre futuros valores a serem adotados por essas variáveis. A predição precisa de uma certa quantidade de dados para a geração do modelo. Na tarefa de classificação no contexto educacional, é preciso enquadrar o estudante em alguma classe. Por exemplo, qual o nível de motivação do aluno a partir do registro comportamental em um AVA? Quais alunos estão prestes a abandonar o curso dadas as notas de suas atividades e/ou frequência? (HÄMÄLÄINEN; VINI, 2010)

Diversos algoritmos são trabalhados por pesquisadores para aplicação aos dados do domínio educacional. Sua adoção deve considerar, dentre outras questões, qual a natureza da variável preditora, ou seja, se ela é binária ou categórica. Para casos binários são utilizados, por exemplo, algoritmos como árvores de decisão e regressão logística. Já se for numérica, são comumente utilizados redes neurais supervisionadas, máquinas de suporte vetorial (SVM), etc. Essa abordagem preditiva é importante, pois ela auxilia no desenvolvimento e uso de atividades instrucionais, através da geração de estimativas sobre os benefícios educacionais antes mesmo da atividade ser aplicada com os alunos (BAKER et al., 2011).

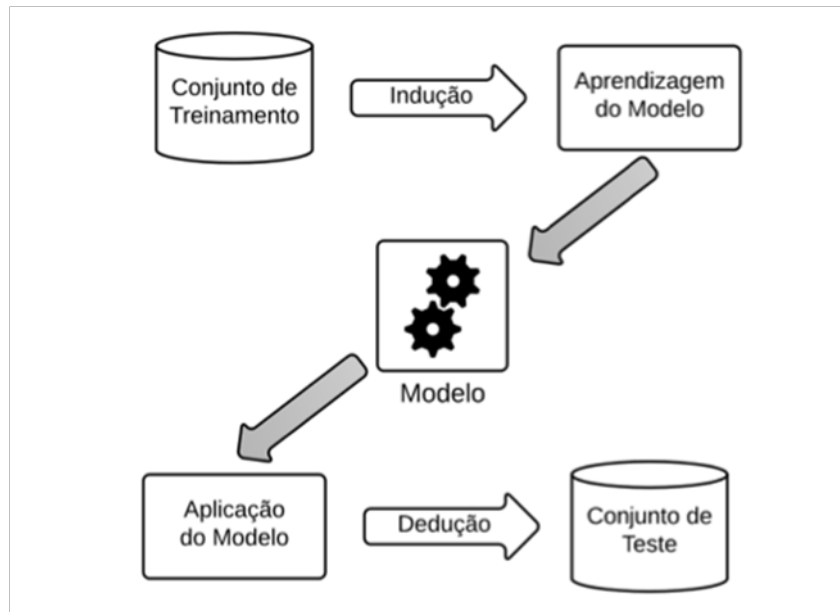
A construção de modelos computacionais de classificação, comumente, utiliza um dentre dois paradigmas a seguir:

- *Top-down*: modelo gerado a partir de informações colhidas com especialistas;
- *Bottom-up*: modelo gerado a partir da identificação de relacionamentos entre as variáveis dependentes e independentes nas bases de dados rotuladas. O classificador opera sob a perspectiva da generalização, a partir de casos específicos contidos no banco de dados (supervisionado). Há também a possibilidade de generalização a partir de dados não-rotulados (não-supervisionados).

A geração de modelos de classificação a partir de bases de dados é um processo que reúne duas etapas, a etapa de aprendizado e a etapa de teste. O processo de geração de um modelo classificador é ilustrado na Figura 9. Nela observa-se como entrada um conjunto de treinamento, composto por instâncias de dados rotulados (contendo o indicativo da classe para cada linha da tabela). A partir desse conjunto de dados, o processo de aprendizagem gera um modelo classificador que é validado via um conjunto de dados de teste.

Dentre as abordagens (técnicas) de classificação frequentemente utilizadas em EDM, poderemos destacar:

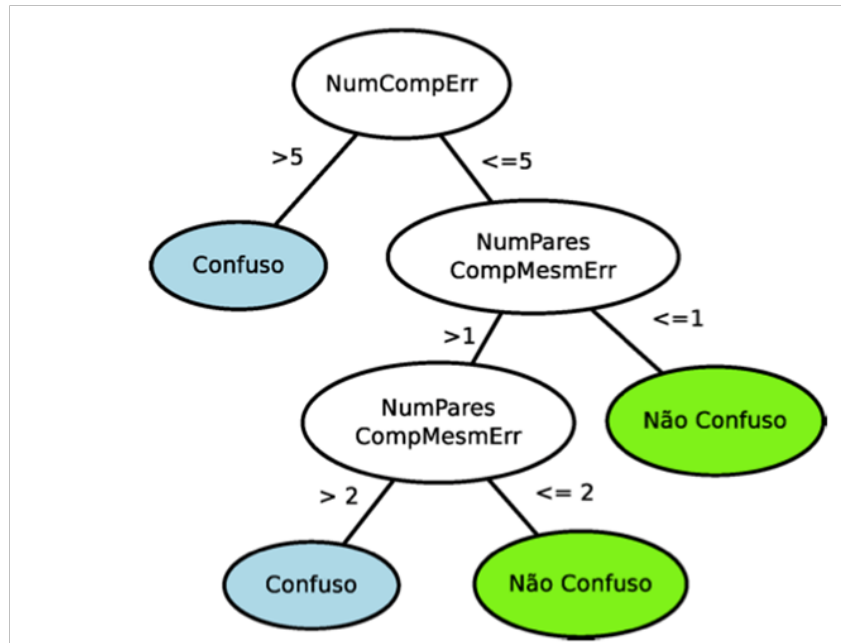
Figura 9 – Exemplo de um Modelo Classificador.



Fonte – Adaptado de Costa et al. (2012).

- **Árvores de Decisão:** representa um conjunto de regras de classificação no formato de uma árvore. Cada percurso de uma árvore (da raiz até a folha) corresponde à uma regra da forma  $T_{i1} \wedge \dots \wedge T_{in}(C = c)$ , onde  $c$  é o valor da classe na folha e cada  $T_{in}$  é um valor booleano testado sobre o atributo  $A_{ij}$ . Cada folha é uma classe. Dentre as vantagens de se fazer uso de árvores de decisão está sua simplicidade e facilidade de entendimento, a manipulação de diferentes tipos de variáveis (tanto variáveis numérica como categóricas), a rapidez para se classificar novos dados e sua flexibilidade. Elas possuem um grande poder de generalização. Como desvantagem, elas estão muito sujeitas ao problema de sobre-ajuste (*overfitting*). *Overfitting* consiste no fenômeno em que o modelo gerado ajusta-se muito aos dados utilizados para o seu treinamento. Desta feita, a ocorrência de *overfitting* prejudica o desempenho do modelo, pois sua capacidade de generalização fica comprometida para analisar futuras amostras não presentes no conjunto de dados utilizado para seu treino. Logo, no domínio educacional, que é caracterizado pelo uso não tão grande de dados em comparação à domínios como o comércio eletrônico, esse pode ser um fator decisivo para se considerar a adoção ou não de árvores de decisão. A Figura 10 apresenta graficamente a classificação característica de uma árvore de decisão.

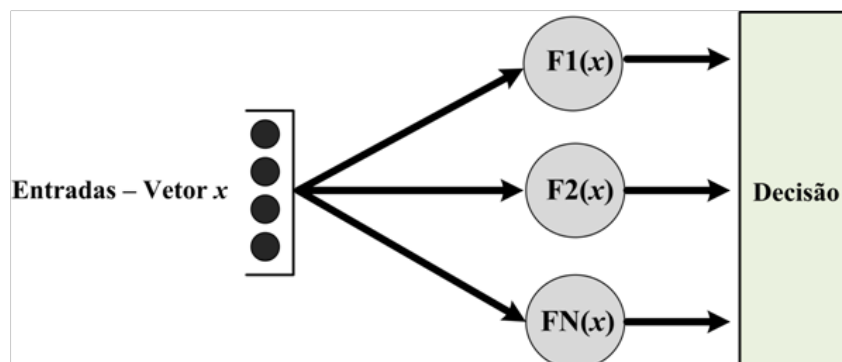
Figura 10 – Exemplo de uma Árvore de Decisão.



Fonte – Fonte: Adaptado de Costa et al., 2012.

- **Classificadores Bayesianos:** tipo de classificadores estatísticos que analisam o pertencimento de um determinado elemento à uma dada classe, utilizando para isso da estimativa de probabilidade do elemento pertencer à esta classe. É um tipo de aprendizado supervisionado que se baseia no teorema de Bayes (REVISTABW,2015). Dentro deste grupo de classificadores, encontra-se a técnica de redes *bayesianas*. Essa técnica caracteriza-se por modelar as dependências estatísticas através de uma estrutura de grafo, onde cada vértice de um grafo corresponde à um atributo, e um arco de entrada à ele corresponde ao atributo que ele depende. A força da dependência é definida por probabilidades condicionais. A Figura 11 apresenta o funcionamento de um Classificador Bayesiano.

Figura 11 – Exemplo de um Classificador Bayesiano.



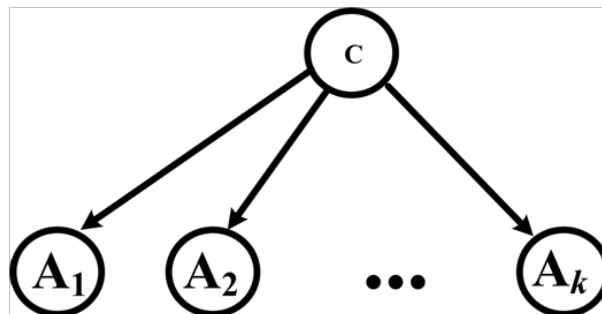
Fonte – Adaptado de Revistabw (2015).



A figura 11 retrata a atuação conjunta das funções  $F_i(x)$  e do módulo de decisão, o que faz com que o espaço dos dados seja dividido em regiões associadas a diferentes classes. Um problema desta abordagem é a necessidade de estimar um grande número de probabilidades, que requer, conseqüentemente, a necessidade de um grande conjunto de treinamento para estimar a probabilidade conjunta precisamente. Por exemplo, se todos os atributos  $A_1, A_2, \dots, A_k$  tem  $v$  diferentes valores e todos  $A_i$ 's são mutuamente dependentes, tem-se a necessidade de estimar  $O(v^k)$ . Outro problema é o decréscimo da precisão da classificação a partir do uso da função de tamanho de decréscimo mínimo. Esta função mede o erro no modelo sobre as variáveis, mas ela não necessariamente minimiza o erro na classe de variável.

Outra importante abordagem é Naive Bayes. Ela resolve os dois problemas citados anteriormente da abordagem de redes bayesianas. Nela, a complexidade para o cálculo de estimativas de probabilidades é linear em  $O(k^v)$  por classe. A Figura 12 mostra Naive Bayes como uma abordagem que pode ser representada por duas camadas de redes bayesianas. São vantagens de Naive Bayes: simplicidade, eficiência, robustez a ruído e facilidade de interpretação. É muito adequado para cenários de tamanho pequeno de dados. Desta forma, é bem adequado ao contexto educacional, onde, comumente, os conjuntos de dados são de pequena ordem.

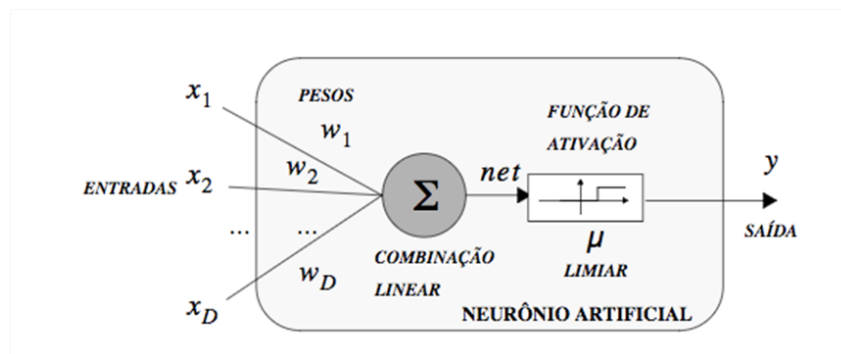
Figura 12 – Exemplo de Naive Bayes.



Fonte – Adaptado de Hämmäläinen e Vini (2010).

- **Redes Neurais Artificiais (RNA):** consiste em uma importante abordagem trabalhada no campo de reconhecimento de padrões e que foi inspirada pelo funcionamento cerebral, estabelecendo a lógica de sinapses e neurônios para a transmissão e processamento de informações. A maioria das RNA's utilizam arquiteturas padronizadas, projetadas especialmente para a resolução de um grupo de problemas. Sua estrutura é composta, basicamente, por três camadas: uma camada de entrada, uma camada intermediária e uma camada de saída (CASTRO;ZUBEN, 2001). O processamento básico ocorre em diversas unidades simples denominadas de neurônios artificiais ou neurônios (nós). Os neurônios estão interconectados gerando as redes neurais. A informação (sinal) é transmitida entre os neurônios através das sinapses. Às sinapses são atribuídos pesos, e a informação armazenada encontra-se nelas. Esta técnica tem sido utilizada com muito sucesso em diversas áreas atuando eficazmente para resolver problemas de predição, aproximação, classificação e reconhecimento de padrões (OSMANBEGOVIC; SULJIC, 2012). Porém, ao comparar-se com a adoção de outras técnicas de aprendizagem de máquina por trabalhos em EDM, não se verifica uma significativa adoção de RNA. Isso se dá em razão da dificuldade de se trabalhar com poucos dados, realidade comum ao cenário educacional e dificuldade em treinar corretamente o modelo, uma vez que há uma série de parâmetros a se trabalhar, como número de camadas e nós ocultos, pesos iniciais e critério de parada. A seleção da arquitetura adequada a um certo problema (topologia da rede) e o critério de parada são as questões mais críticas desta técnica, uma vez que redes neurais são muito suscetíveis à *overfitting*.

Figura 13 – Exemplo de uma arquitetura de RNA.



Fonte – RAUBER(2005).

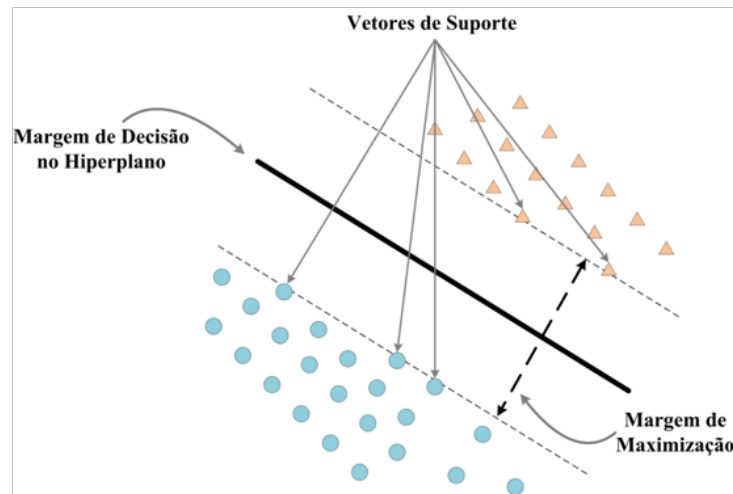
- **Classificadores  $K$  vizinhos mais próximos ( $k$ -nearest neighbor):** representa uma abordagem peculiar para se fazer classificação, pois ao invés de construir um modelo global para

generalização, atua localmente. É um modelo supervisionado, ou seja, durante seu treinamento possui a identificação das classes de cada instância. Sua ideia principal é, portanto, a classificação de um novo elemento pela análise dos valores das classes dos  $k$  elementos mais próximos. A classe que será escolhida será a mais comum entre os vizinhos ou uma distribuição de classe na vizinhança. Este método possui uma série de vantagens, como o fato de ser simplesmente configurável, uma vez que só possui dois parâmetros  $k$  e  $d$  que, respectivamente, correspondem ao número de vizinhos e a métrica de distância. É também considerada uma classificação de alta precisão, sendo robusta a ruídos e à esparsidade de dados. A técnica possui um alto poder de generalização. A principal desvantagem é a dificuldade de selecionar uma função de distância  $d$ . O contexto educacional trabalhar geralmente com dados numéricos ou categóricos, e atributos numéricos podem ser em diferentes escalas. Isto significa que, para calibrar precisamente a função de distância, será preciso contar com uma quantidade significativa de dados para treinamento. Deve ser feito antes, um processo de tratamento de dados para remover os atributos irrelevantes, uma vez que a técnica utiliza todos os atributos e isso pode gerar processamento computacional desnecessário.

- **Máquina de Vetores de Suporte (*support vector machines*):** Esta técnica é robusta para se fazer a separação de dados não separáveis linearmente. Ela consegue fazer a geração de um modelo não linear a partir de poucos dados de treinamento. A ideia principal é mapear os dados em uma dimensão superior, para tornar as classes linearmente separáveis. Esse mapeamento, portanto, é feito de forma implícita pelo uso de funções *kernel*. A principal vantagem do SVM é no fato dele sempre encontrar o ótimo global, uma vez que não há ótimo local na maximização das fronteiras das classes. Outro grande ponto positivo desta abordagem é a robustez quanto à *overfitting*. Sua precisão de classificação e generalização são altas mesmo diante de um reduzido conjunto de dados. Essa é uma característica importante que essa técnica oferece ao campo de EDM, uma vez que o contexto educacional, via de regra, é complexo em termos possuir muitos atributos e restrito quanto ao volume dados para geração do modelo. Como desvantagens, pode-se destacar a necessidade de trabalhar com dados numéricos contínuos, a dificuldade de entendimento de seu funcionamento, e a dificuldade para fazer a seleção de parâmetros apropriados às funções de *kernel* (Hämäläinen; Vini, 2010).

alta

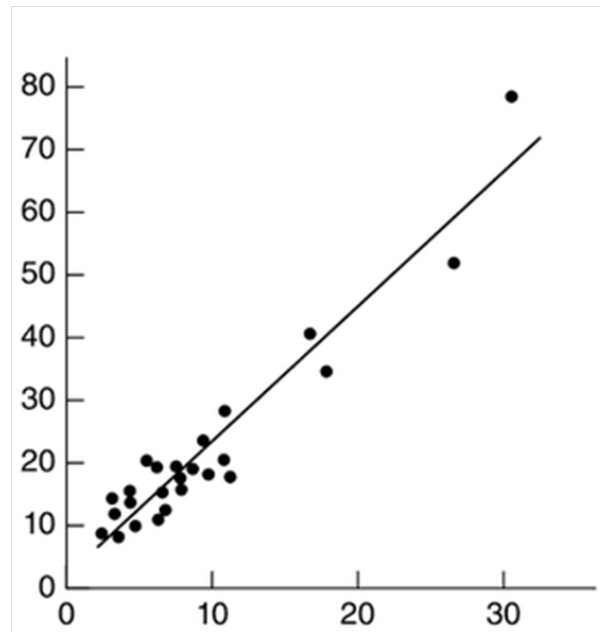
Figura 14 – Funcionamento do SVM e de seus vetores de Suporte.



Fonte – Própria

- **Regressão Linear:** consiste em uma das técnicas mais utilizadas no campo de estatística aplicada. Um modelo de regressão linear formaliza como uma variável dependente se comporta, de forma sistemática, diante de qualquer alteração sobre uma ou mais variáveis independentes. Desta forma, a técnica atua no estudo da relação entre uma variável dependente ( $y$ ) com uma ou mais variáveis independentes ( $x$ ). No exemplo de uma regressão linear simples, temos um par  $(x, y)$ . A ideia é que a partir do conhecimento do valor de  $x$ , pode-se descobrir o valor de uma variável  $y$ . Produz modelos simples, que minimiza os problemas de *overfitting* vistos em sistemas complexos. No entanto, seus dados devem conter poucos *outliers*, e não deve conter lacunas. Um exemplo de geração do modelo de regressão linear no contexto educacional, por exemplo, voltado a previsão de sucesso ou insucesso de um aluno, deve contar com dados de todos os alunos para que possa fazer uma previsão da turma. Se alguns alunos não tiverem registros que possam ser usados na entrada dos dados para compor o modelo, estes alunos não tem como serem contemplados pelo modelo preditivo.

Figura 15 – Representação de uma Regressão Linear.

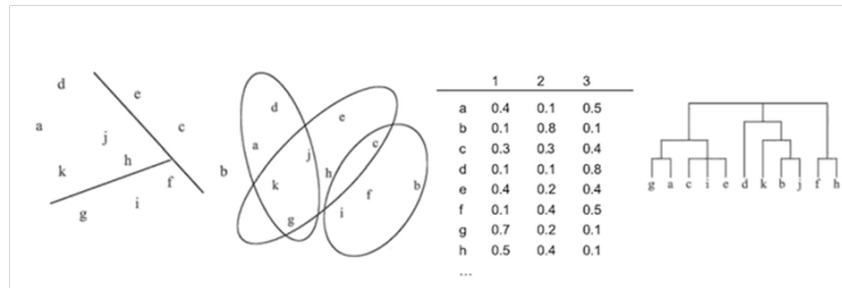


Fonte – Montgomery et al. (2015).

### 2.2.2 Agrupamento

O algoritmos para agrupamento de dados visam encontrar dados que tenham características semelhantes, gerando grupos ou categorias. O conhecimento sobre estes grupos não se dá de forma inicial (direta). As técnicas de agrupamento separam os grupos de acordo com as características dos dados, agrupando os similares entre si (BAKER et al., 2011). A técnica permite fazer previsões, que serão feitas mediante inferência. Para realizar a clusterização, é preciso estabelecer critérios adequados à classificação do conjunto de dados em subconjuntos, sendo os mais comuns os de homogeneidade e separação. O primeiro é a medida no interior de um próprio *cluster*, verificando o grau de similaridade; o segundo, mede os elementos *interclusters*, medindo o quão diversos são entre si (KAMPFF, 2009). A Figura 16 ilustra diversas formas de representar a técnica de *clusters*.

Figura 16 – Diferentes formas de representar *clusters*.



Fonte – Witten et al. (2016).

As técnicas de *clusters* tem sido vastamente adotadas por trabalhos de pesquisa em EDM. Elas tem sido utilizadas em aplicações para a classificação de alunos quanto a seu nível de aprendizado, análise de textos, recomendação futura de ações a estudantes em situação similar, etc.

Diversas aplicações das técnicas apresentadas nesta seção no domínio de ensino híbrido e mineração de dados educacionais são novamente discutidas no Capítulo 3, onde trabalhos relacionados ao tema da tese serão apresentados.

### 2.3 Sistemas de recomendação

Com o crescimento em todo o mundo, sobretudo na última década, de experiências de formação educacional que fazem amplo uso de tecnologias para acesso a conteúdos e interação entre usuários (alunos, professores e especialistas), emerge uma área que tem investigado e desenvolvido diversas soluções tecnológicas para educação, denominada Aprendizagem Incrementada por Tecnologia (*Technology Enhanced Learning-TEL*).

Aprendizagem Incrementada por Tecnologia representa uma área de estudos a qual investiga o uso de tecnologias como forma de potencializar o aprendizado de pessoas através da transformação de metodologias educacionais adotadas por indivíduos e instituições (MANOUSELIS et al., 2011). Nesse domínio de aplicação, as inovações tecnológicas suplantam uma diversidade de experiências educacionais, como atividades de ensino e o processo de aprendizagem do aluno.

Uma das principais ferramentas trabalhadas pela área TEL é o uso de sistemas de recomendação. Esses sistemas serão referenciados nesse trabalho como SRE (Sistema de Recomendação Educacional). O uso de SRE tem propiciado o surgimento de uma vasta diversidade de aplicações (DRACHSLER et al., 2008). Manouselis et al. (2011) apresentam

em seu estudo uma abrangente análise de trabalhos referentes a SRE. A maioria dos sistemas apresentados são responsáveis por sugerir aos interagentes fontes complementares de conteúdo ou informar quais outros usuários estão aptos a colaborar para a realização de determinadas atividades. Comumente esses sistemas utilizam técnicas de Inteligência Artificial, Mineração de Dados, Estatística, dentre outros. Os principais métodos adotados para a implementação das soluções de SR são:

- Filtragem baseada em conteúdo (PAZZANI; BILLSUS, 2007) – promove a recomendação de objetos (conteúdos) a um usuário de acordo com a avaliação feita por ele a objetos similares anteriormente. Os objetos de interesse nessa filtragem são definidos por suas características. O perfil de cada usuário, portanto, é definido a partir dos perfis dos itens bem avaliados por ele. Formalmente, essa filtragem estima a utilidade  $u(c, s)$  de um item  $s$  para um usuário  $c$  baseado nas utilidades  $u(c, s_i)$  indicadas pelo usuário  $c$  para os itens  $s_i \in S$  que são mais próximos (similares) ao item  $s$  (SOUZA, 2011), (ADOMAVICIUS; TUZHILIN, 2005);
- Filtragem Colaborativa (SCHAFER et al., 2007): esta técnica busca no histórico de interações entre usuários e itens relações que permitam identificar relacionamentos entre estes usuários. Nenhuma informação sobre os itens é levada em consideração. O que é levado em conta é o feedback oferecido através das interações dos usuários. A partir disto são estimadas as classificações de itens e objetos para a socialização entre o grupo de interagentes. Formalmente, essa técnica estima a utilidade  $u(c, s)$  de um item  $s$  para um usuário  $c$  baseado na utilidade  $u(c, s_i)$  do item  $s$  indicada pelos usuários  $c_j \in C$  que são similares ao usuário  $C$  (SOUZA, 2011), (ADOMAVICIUS; TUZHILIN, 2005);
- Filtragem Híbrida: abordagem que supera as limitações encontradas nas filtragens acima descritas, e comuta as virtudes de cada uma delas em uma única abordagem. Burke (2007) especificou sete mecanismos básicos de como pode se dar essa hibridização, são eles: ponderada, misturada, comutada, combinação de características, aumento de características, cascata e meta-nível. Os principais fatores que tem motivado pesquisadores a mesclar as abordagens de filtragem, são o problema da partida a frio (*cold-start*), esparsidade de dados e problemas de escalabilidade (LU et al., 2015).

As técnicas acima descritas são adotadas em razão das circunstâncias em que são aplicadas. Cada uma delas possui vantagens e desvantagens e sua adoção deve ser analisada de acordo com cada finalidade. A grande desvantagem da filtragem baseada em conteúdo é

sua dependência de informações acerca dos itens disponíveis. A cada entrada de novo item deve ser feita a inserção de metadados descritivos de suas características. Já a abordagem de filtragem colaborativa, se por um lado não se preocupa com essa alimentação de informações sobre os itens, tem um problema do *cold start* (partida a frio). Este problema diz respeito a toda vez que um novo usuário é inserido no sistema., pois como não se tem, comumente, nenhuma informação sobre suas preferências, fica inviável recomendar elementos alinhados à seu interesse. Outra questão que limita essa abordagem diz respeito a opção de muitos usuários de não compartilharem seus gostos/preferências, isto obstaculiza o princípio desta técnica.

### 2.3.1 *Sistemas de Recomendação Sensíveis ao Contexto (Context-aware Recommender Systems – CARS)*

Os Sistemas de Recomendação buscam classificar determinados itens a partir de estimativas geradas por elementos já avaliados pelos usuários. Os SR convencionais são modelados a partir de uma relação bidimensional relativa às informações sobre usuários e aos itens a serem recomendados. Dado que  $U$  e  $I$  sejam, respectivamente, o conjunto de informações sobre os usuários e o conjunto de informações sobre os itens a serem recomendados, temos  $f$  como a função utilidade ( $R : \text{Usuários} \times \text{Itens} \Rightarrow \text{classificação}$ ) (VERBERT *et al.*, 2012).

A função  $R$  obtém a classificação ( $r$ ), que identifica, por exemplo, o grau de interesse de um determinado usuário ( $u \in U; u = \text{Maria}$ ) em um item específico ( $i \in I; i = \text{Rico e Joana}$ ).  $R_{\text{livro}}(\text{Maria}, \text{Rico e Joana}) = 10$ .

No entanto, CARS são sistemas que oferecem um nível maior de informações sobre um determinado domínio. A definição de uma estrutura informacional mais abrangente é de essencial importância para a obtenção de recomendações mais eficientes aos usuários (BRYANet *al.*, 2008), pois proporciona, por exemplo, a oferta de recomendações baseadas nas necessidades e características dos usuários. No contexto educacional, essas informações complementares podem ser relativas aos objetivos do aprendizado, nível de conhecimento pré-existente, tempo disponível para estudo, etc (KANTOR *et al.*, 2011).

Ao contrário de SR que trabalha numa perspectiva de duas dimensões ( $U \times I$ ), CARS utiliza uma estrutura multidimensional. Dessa forma, sua função utilidade assume a formação da equação (2.1):

$$R(U \times \text{Itens} \times \text{Contexto}) \Rightarrow \text{Classificação} \quad (2.1)$$



Adomavicius e Tuzhilin (2005) foram os pioneiros no desenvolvimento de CARS. A pesquisa dos autores trabalhou na criação de abordagens para o mapeamento do contexto onde as recomendações são realizadas. Dessa forma, segundo VERBERT *et al.* (2012), um contexto pode ser obtido através das seguintes estratégias:

- Extração de forma explícita – o usuário informa dados sobre suas características e preferências. Ex. Um formulário, uma enquete, etc;
- Extração de forma implícita – as informações contextuais são obtidas de forma automática pelo sistema. Ex. Localização do usuário via GPS;
- Extração a partir de inferências – as características e opções do usuário são obtidas através de análises e cruzamentos de informações. Técnicas estatísticas e de mineração de dados são comumente utilizadas por essa abordagem. Ex.: a verificação da audiência de TV por parte de múltiplos membros de uma família a partir dos registros de visualização de programas.

A inserção de informações contextuais ao processo de recomendação pode ser realizada segundo duas diferentes abordagens. A primeira abordagem é consulta e pesquisa orientada ao contexto. Essa abordagem utiliza informações contextuais para consultar ou buscar em um repositório específico de itens um melhor resultado para o usuário. Um exemplo são os SR para dispositivos móveis que utilizam dados como geolocalização para informar a um turista as diversas opções de restaurantes próximos à sua localização de momento.

A segunda abordagem refere-se a estimativas e indução. Essa estratégia tem sido uma tendência na literatura de CARS (VERBERT *et al.*, 2012). Essa abordagem visa modelar e aprender sobre o contexto das preferências do usuário. Os CARS que utilizam essa abordagem são desenvolvidos a partir do conhecimento gerado parcialmente sobre os interesses do usuário sob um determinado contexto. A modelagem dos dados se dá, tipicamente, no formato <usuário, item, contexto, classificação>. Cada registro captura quanto um usuário aprecia um determinado item sob um determinado contexto. Por exemplo, quanto o usuário aprecia telejornal durante a semana <Paulo, Telejornal, Dia de semana, assiste>. O registro dessas observações pode se dar a partir de logs de participação do usuário com o sistema, ou a partir das interações com outros itens já recomendados ao usuário. São comumente utilizadas técnicas de mineração de dados, assim como técnicas de aprendizado de máquina (*Support Vector Machine* e classificadores bayesianos) (KANTOR *et al.*, 2011).

O processo de recomendação sensível ao contexto baseado na abordagem de estima-

tivas e indução, é implementado por um dos três paradigmas a seguir (VERBERT *et al.*, 2012), (KANTOR *et al.*, 2011):

- Pré-filtragem contextual (ou contextualização prévia à recomendação): nessa abordagem a informação sobre um determinado contexto é utilizada para filtrar e assim definir um conjunto de dados mais relevantes antes de executar o processo de recomendação tradicional do sistema (usuário  $\times$  item);
- Pós-filtragem contextual (ou contextualização a posteriori à recomendação): nessa abordagem o contexto é ignorado inicialmente, e as recomendações já são geradas na entrada de dados. O conjunto de recomendações resultante é ajustado (contextualizado) para cada usuário usando informações contextuais;
- Modelagem Contextual (ou contextualização da função de recomendação): essa abordagem utiliza as informações de contexto diretamente na função de recomendação. Ao contrário das abordagens anteriores, algoritmos de recomendação tradicional (2D) não são utilizados, e sim algoritmos de recomendação multidimensional.

Uma das principais questões a serem estabelecidas para a construção de um contexto é a especificação de características que venham a representar significativamente um determinado domínio. Dessa maneira, pesquisadores têm investigado e buscado definir de forma precisa e consistente itens informacionais para caracterizar um determinado contexto (BEALE; LONSDALE, 2004) (SCHMIDT, 2005) (METTOURIS; PAPADOPOULOS, 2016). Adomavicius e Tuzhilin (2005) definiram uma classificação simples de informação de contexto que é relevante para aplicações em SRE. Dentre os principais itens de informação definidos pelos autores estão:

- Dados Computacionais – corresponde à infraestrutura computacional utilizada pelo usuário, tanto em nível de software, como de hardware;
- Localização – representa o registro de informações sobre a localização do usuário, que por sua vez pode ser obtido de forma implícita ou explícita. A primeira dá-se pelo uso de sensores ou mesmo mecanismos de geolocalização (ex.: GPS). A segunda ocorre pela indicação direta do usuário;
- Tempo – consiste em dados relativos a eventos registrados desde instantes específicos (hora, turno), até períodos (meses, ano). Esse dado é comumente associado a outros atributos para composição de informações mais abrangentes sobre um determinado contexto;
- Usuários – a modelagem do perfil de usuários tem sido objeto de muitas pesquisas, como Dolog e Najdl (2003), Brusilovsky e Millan (2007), Specht (2000). Diversas

especificações têm sido desenvolvidas no sentido de padronizar a modelagem de um contexto de usuário, como IMS LIP, IMS ePortfolio, IMS Enterprise, IEEE RCD, FOAF (friend-of-a-friend), e HR-XML (ADOMAVICIUS; TUZHILIN, 2005). Considerando o contexto de SRE, o trabalho desenvolvido por Dolog e Nejdl (2003), oferece uma arquitetura que conta com os seguintes elementos: dados pessoais (nome, endereço, telefone etc.), objetivos do usuário (descrição, prioridade, etc.), preferências (linguagem, proficiência, etc.), desempenho em estudos (portfólio, experiência, desempenho estudantil, bagagem cultural etc.)

- **Relações Sociais** - corresponde às relações estabelecidas entre dois ou mais indivíduos e as diferentes conexões que são estabelecidas em diferentes ciclos sociais (amizade, família, trabalho, ambiente acadêmico, etc). Para modelar essas conexões podem ser utilizadas linguagens de especificação como FOAF que, comumente, tem sido utilizada para descrever ligações entre pessoas e as respectivas informações sobre elas a partir da WEB (BRICKLEY, 2014).

### 2.3.2 *Sistemas de recomendação educacionais – SRE*

Como descrito anteriormente, TEL tem sido uma área em que tem sido tratado o desenvolvimento de aplicações tecnológicas educacionais em apoio a professores e alunos. Esse campo de pesquisa tem atraído pesquisadores para a investigação de estruturas computacionais que venham a potencializar o aprendizado, projetando, desenvolvendo e experimentando inovações tecnológicas em contribuição a práticas educacionais.

Dentro desse domínio de aplicações, tem-se destacado o desenvolvimento sistemas de recomendação (Sistemas de Recomendação Educacionais - SRE).

Como uma aplicação tecnológica voltada ao contexto educacional, SRE têm atendido a uma série de tarefas em seu domínio de atuação. A seguir são listadas algumas das aplicações de SRE disponíveis (BUDER; SCHWIND, 2012):

- **Sugestão de cursos** – baseado nos cursos que determinado usuário tem realizado com maior frequência, o sistema recomenda cursos de assuntos correlatos, inferindo que esses cursos tem forte grau de interesse por parte do usuário. Ex.: MOOC Coursera;
- **Busca de itens/objetos significativos** – promove recomendação de conteúdos educacionais como objetos de aprendizagem, vídeos, sites (conteúdo), etc, alinhados a determinado tópico de interesse do usuário. Ex.: sugestão de blogs e sites afins às leituras comumente

desenvolvidas pelo usuário na WEB (DRACHSLER et al., 2008);

- Recomendação de sequências de navegação – usuário recebe um roteiro de navegação a partir do registro feito de usuários que tiveram sucesso;
- Recomendação de outras pessoas com o mesmo interesse de estudo – usuário recebe a indicação de outros indivíduos que apresentam similaridade de dificuldades, interesse, etc;
- Sugestão de leitura complementar – dado determinado curso, ou conteúdo trabalhado em uma lista de exercícios, o SRE recomenda ao usuário exercícios correlatos e conteúdos aderentes à temática em trabalho.

Como apontado acima, as aplicações no contexto de SRE são bem diversas, havendo uma quantidade considerável de finalidades. Outras aplicações serão abordadas na seção 2.3.2.3.

Os SR convencionais possuem duas principais entidades, usuários e elementos. Dessa forma, eles oferecem recomendações aos usuários baseados nas informações que detém sobre eles, assim como pelas informações que possuem sobre os elementos que serão recomendados. Já os SRE, além dessas informações devem possuir as seguintes particularidades (RIFON et al., 2013):

- Registrar informações sobre os usuários levando em consideração não somente suas preferências e interesses, mas, também, seu conhecimento e desempenho prévio em questões referentes ao domínio tratado;
- Utilizar informações sobre os recursos educacionais que devem contemplar a qualidade, complexidade, nível de interatividade, público-alvo, etc;
- Fazer recomendações a partir de informações do contexto em que o usuário está inserido (ex.: sugestão de conteúdos no formato de mídia mais adequado ao dispositivo que o usuário esteja utilizando).

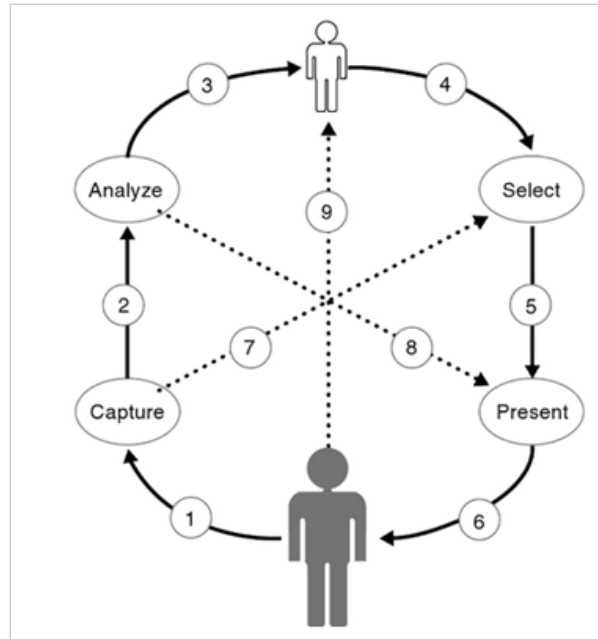
Uma das principais deficiências encontradas em sistemas educacionais é a sua dificuldade em promover experiências personalizadas aos objetivos, tarefas, enfim interesses cada usuário (*one-size-fits-all*). Em razão desta limitação, uma nova categoria de sistemas adaptativos voltados ao contexto educacional vem sendo trabalhada, essa área é chamada sistemas adaptativos de hipermídia educacional (*AEH – Adaptive Educational Hypermedia Systems*) (RICCI, 2011).

#### 2.3.2.1 *Sistemas Adaptativos de Hipermídia Educacional (AEH)*

Sistemas adaptativos voltados ao contexto educacional caracterizam-se por sua flexibilidade em atender aos diferentes aspectos de preferências e condições de exploração dos

usuários como forma de promover uma experiência de interação mais qualificada. Shute e Zapata-Riveira (2012) definiram um ciclo adaptativo em quatro etapas: captura, análise, seleção e apresentação (Figura 17).

Figura 17 – Fluxo de processo em quatro etapas AEH.



Fonte – Adaptado de Shute e Zapata-Riveira (2012).

**Captura:** essa etapa do processo requer a obtenção de informações sobre o aprendiz (usuário), seja a partir de informações obtidas pela interação dele com o sistema (ex.: resolução de problemas), seja por informações inferidas (motivação, engajamento, etc.).

**Análise:** requer a criação e manutenção de um modelo do aluno em relação ao domínio, representando informações em termos de inferências sobre o cenário vigente. Em outras palavras, consiste na verificação sistemática da condição de aprendizagem do aluno sobre temáticas que foram trabalhadas por ele. **Seleção:** de acordo com a análise feita, é selecionado um elemento educacional para a sequência de interação do aluno.

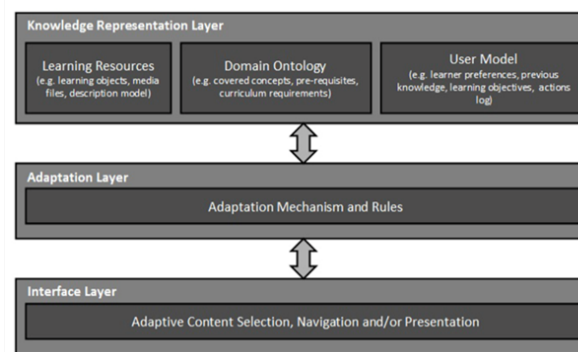
**Apresentação:** os objetos selecionados na etapa anterior são então apresentados para o usuário. Cada sistema AEH possui seu(s) nível(is) de adaptação. Esta adaptação pode ser feita em termos de conteúdo, interface, processo de navegação, etc.

Brusilovsky (2016) descreve uma arquitetura genérica para AEH (Figura 18) que inclui: uma camada de representação e organização de conhecimento sobre conteúdo educacional (recursos de aprendizagem), o domínio (ontologia de domínio) e o usuário (modelo de usuário); uma camada que inclui a adaptação de mecanismos e regras; e uma camada que provê a

adaptação em tempo real de resultados ao usuário.

O desenvolvimento de sistemas AEH vem se dando desde década de 1990 e teve forte crescimento na linha de pesquisa de Tutores Inteligentes (ITS- *intelligent tutoring systems*). Outro campo que cresceu nesse período foi o de sistemas hipermídia de educação, que promoveram a adaptação de sistemas aos estudantes individualmente (MANOUSELIS et al., 2010). Para mais detalhes, segue no Apêndice B a Tabela 34.

Figura 18 – Arquitetura Genérica de AEH.



Fonte – (RICCI et al. , 2011), (KARAMPIPERIS; SAMPSON, 2005), (BRUSILOVSKY, 2016).

### 2.3.2.2 Técnicas utilizadas em SRE'S

Diversas são as técnicas adotadas para a implementação de soluções de sistemas de recomendação educacionais. As aplicações de SRE tem feito forte uso de soluções do domínio de aprendizagem de máquina. Drachsler et al.(2015) elaboraram uma extensa revisão de sistemas de recomendação em TEL analisando trabalhos no intervalo de quinze anos (2000 a 2014). Os autores organizaram os trabalhos em sete grupos, a saber:

1. SRE utilizando abordagens de filtragens colaborativas em outros domínios;
2. SRE com propósito de melhorar abordagens de filtragem colaborativa levando em conta as particularidades de TEL;
3. SRE que consideram explicitamente as restrições do domínio educacional como uma fonte de informação para o processo de recomendação;
4. SRE que explora técnicas diferentes de filtragem colaborativa;
5. SRE que considera a informação de contexto educacional para o melhoramento do processo de recomendação;

6. SRE que avaliam o impacto educacional frente a partir da realização de recomendações;
7. SRE que recomendam cursos aos usuários.

Considerando a presente seção que versa sobre técnicas adotadas para a construção de SRE, serão apresentados exemplos de técnicas que se enquadram, de acordo com Drachsler *et al.* (2015), no agrupamento 4. Zaiane (2002) utilizou a técnica de regras de associação para a construção de um modelo que representa o comportamento dos estudantes e, a partir deste modelo, recomendar atividades ou atalhos que possam ajudar os aprendizes quanto a uma melhor navegação dos materiais digitais. Chen *et al.* (2008) utiliza lógica *fuzzy* e teoria de resposta ao item para sugerir materiais de apoio de acordo com graus de dificuldade específicos, levando em consideração os feedbacks oferecidos pelos aprendizes. Em Wang (2012), o autor utiliza lógica *fuzzy* para a recomendação personalizada de conhecimento através da descoberta de caminhos de aprendizado efetivos a partir de experiências de aprendizagem passadas, utilizando para isso um modelo de otimização de colônia de formigas.

Hsieh *et al.* (2013) utilizou meta-regras derivadas do modelo de cadeias de Markov para a realização do cálculo de probabilidades frente a apresentação de novos objetos educacionais no âmbito de um curso sequenciado, permitindo a descoberta de rotas de aprendizagem.

Sweeney *et al.* (2016) utilizaram técnicas clássicas e outras mais contemporâneas comumente adotadas no domínio de e-commerce para o contexto educacional. Neste trabalho foram utilizadas as técnicas de Factorization Machines (FM), Random Forest (RF) e *Personalized Linear Multiple Regression*. Os autores utilizaram de forma combinada (híbrida) FM e RF para a previsão de notas de novos alunos e alunos repetentes. A previsão deste desempenho, segundo os pesquisadores, pode contribuir com a melhora no índice de repesamento e abandono, uma vez que a instituição e professores podem com antecedência dimensionar ações de correção das dificuldades encontradas no aprendizado dos alunos.

Thai-Nghe *et al.* (2012) também utilizam técnicas de fatoração de matrizes (SVD) para gerar avaliações e previsões eficazes. Além disso, foi utilizada a técnica de gradiente descendente estocástico com o propósito de reduzir o erro. Os autores consideram fatores temporais em sua implementação e a técnica de tensores foi utilizada para decair a taxa de aprendizagem do estudante ao longo do tempo (TFW – Tensor Factorization Weighting).

Pardos e Heffernan (2010) utilizaram Modelos Ocultos de Markov (Hidden Markov Model) e Random Forest para prever a performance de estudantes frente a novos exercícios. Os autores desenvolveram um modelo bayesiano HMM que prevê a probabilidade de conhecimento

para cada estudante e para cada atividade, assim como determina a probabilidade de resposta correta para cada etapa de atividades. O modelo calibra parâmetros específicos relativos à estudantes de forma individualizada (taxa de aprendizado, suposição e lapso). O modelo resultante considera a composição entre usuário e parâmetros de habilidade o que supera os modelos que somente consideram parâmetros de habilidade.

No Capítulo 3 serão apresentados mais trabalhos que informam sobre técnicas inovadoras para a predição de desempenho estudantil que, por sua vez, corresponde à uma das aplicações mais em evidência na área geral de tratamento computacional de dados.



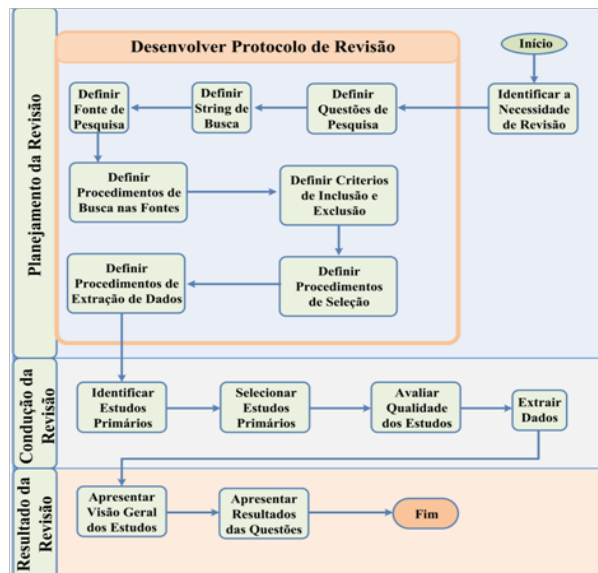
### 3 TRABALHOS RELACIONADOS

O presente capítulo apresenta a revisão sistemática realizada de trabalhos relacionados à temática da tese. O objetivo deste levantamento vem no sentido de apresentar o estado da arte de sistemas de recomendação aplicados à educação, assim como trabalhos que utilizam técnicas de EDM e aprendizado de máquina para previsão de desempenho e personalização de atendimento ao estudante, assim como experiências de ensino híbrido por instituições de ensino.

#### 3.1 Processo de revisão sistemática

Com o intuito de promover um levantamento abrangente e atualizado dos principais trabalhos que vem sendo desenvolvidos nos últimos anos pela comunidade científica nas áreas de Ensino Híbrido, Mineração de Dados Educacionais e Sistemas de Recomendação, foi desenvolvida uma revisão sistemática orientada sob um processo de busca organizada, com o objetivo de obter resultados de tópicos mais pesquisados, correlação entre trabalhos, tendências investigativas, desafios etc. Para isto, foi utilizado o método de revisão sistemática de Kitchenham (2004).

Figura 19 – Fluxo do Protocolo de Revisão



Fonte – Adaptado de KITCHENHAM (2004).

### 3.1.1 Planejamento e revisão

**1 – Identificação da necessidade de revisão:** o uso de tecnologias da informação no contexto educacional dá-se de forma difusa. Considerando que esta tese reúne em uma mesma solução três grandes temas que, na maioria das vezes, não são discutidos conjuntamente na literatura, é importante buscar trabalhos em que a associação destas áreas ocorra, ainda que não simultaneamente.

**2 – Desenvolvimento de protocolo de revisão:** compõem esta etapa do processo, sete sub-etapas:

- (i) Definir questões de pesquisa: Esta revisão sistemática tem como questões a se trabalhar: Questão de Pesquisa Principal(QP): Qual o estado da arte nas áreas de sistemas de recomendação educacionais, mineração de dados educacionais e ensino híbrido?
  - QP01:** De que forma a tecnologia tem propiciado a implantação de inovações metodológicas para a melhoria do processo de ensino e aprendizagem em cenários de *Blended Learning*?
  - QP02:** Quais aspectos do campo educacional podem ser diretamente trabalhados e aprimorados pela utilização de TEL no contexto de *Blended Learning*?
  - QP03:** Como viabilizar um processo de ensino e aprendizagem personalizado?
  - QP04:** Como atender aos principais requisitos necessários à implantação de uma solução tecnológica de *Blended Learning*?
- (ii) Definir *string* de busca: Foram utilizadas as seguintes palavras-chave: *educational data mining, predicting student performance, student classifying, predicting failure, personalized learning, blended learning, flipped classroom, virtual learning environment, recommender systems, survey*, ensino híbrido, sistemas de recomendação educacionais, classificação de estudantes.
- (iii) Definir fontes de pesquisa: Foram pesquisados os indexadores ACM *Digital Library*, *Science Direct* e *IEEEExplore*.
- (iv) Definir procedimentos de busca nas fontes: Foi utilizada a opção de busca avançada em cada um dos serviços indexadores, utilizando as strings apresentadas anteriormente.

- (v) Definir critérios de inclusão e exclusão: Algumas restrições foram estabelecidas, como o período de publicação dos artigos, a partir de 2010 em um primeiro momento. Depois foi definido de 2013 aos tempos atuais. Uma vez que nesse intervalo mais contemporâneo não eram obtidos bons resultados, seja em número de trabalhos, seja na qualidade dos mesmos, era então ampliado para buscar trabalhos mais antigos.
- (vi) Definir procedimento de seleção:
  - Etapa 1:** a estratégia de busca foi aplicada nas fontes.
  - Etapa 2:** durante a seleção de um conjunto primário de estudos, todos os artigos obtidos foram lidos e analisados segundo os critérios de inclusão e exclusão.
  - Etapa 3:** Todos os artigos selecionados na etapa 2 foram lidos em sua íntegra e analisados novamente sob o prisma dos critérios adotados em (v). Os arquivos incluídos foram encaminhados para a etapa de extração de dados.
- (vii) Definir procedimento de extração de dados: O processo de extração de dados foi realizado mediante a composição de uma planilha contendo os seguintes itens a serem preenchidos para cada trabalho: título, ano de publicação, veículo de publicação, autores, palavras-chave, tema, tipos de dados (reais ou fictícios), técnica, método(s) de avaliação, pontos positivos e pontos negativos do trabalho.

### **3.1.2 Condução da revisão**

- (i) Identificar Estudos Primários: A coleta dessa revisão sistemática ocorreu em momentos distintos. O primeiro período foi em abril de 2015. O segundo, em fevereiro de 2016. As demais etapas ocorreram no segundo semestre de 2016, com atualização no primeiro semestre de 2017. Nesta etapa, foram encontrados 261 trabalhos.
- (ii) Selecionar Estudos Primários: Nesta etapa, foram lidos os resumos dos 261 trabalhos. Destes, 81 são alinhados com as temáticas que são objeto da tese.
- (iii) Avaliar a Qualidade dos Estudos: A qualidade dos trabalhos lidos se deu a partir do critério de aderência à temática da tese e se eram trabalhos com clareza na descrição de suas técnicas e escopo de trabalho.
- (iv) Extrair Dados: O processo de extração de dados caracterizou-se pela coleção de informações como ano e veículo da publicação, título, autores, tema e palavras-chave, assim como

observações e técnicas de aprendizado de máquina utilizadas (para os artigos na área de EDM e aprendizagem de máquina).

### 3.1.3 Resultado da revisão

Esta ação foi subdividida em dois estágios:

- (i) Visão Geral dos Estudos: Oferece uma visão geral dos estudos pesquisados. Estes resultados mostram aspectos gerais que objetivam contextualizar os assuntos trabalhados na tese.
- (ii) Resultado das Questões Apresenta o resultado das questões de pesquisa definidas para presente o trabalho (principal e secundárias).

## 3.2 Visão geral dos estudos

O uso de tecnologias computacionais no contexto educacional tem se dado ao longo das últimas décadas em diferentes aplicações. O presente trabalho, por sua vez, objetiva propor uma solução tecnológica inserida no domínio de Ensino Híbrido (*Blended Learning*-BL). Dado que a pesquisa por artigos na área de BL resultam, muitas vezes, em abordagens que extrapolam os objetivos desta tese, foi necessário filtrar e excluir as ocorrências de trabalhos voltados à gestão de processos educacionais ou mesmo à aspectos eminentemente pedagógicos.

Por outro lado, foi buscado outros campos de investigação que atuam no uso e desenvolvimento computacional de tecnologias (sistemas de recomendação, plataformas *web*, técnicas e algoritmos de aprendizagem de máquina, etc.), mesmo que na quase totalidade dos casos não haja, necessariamente, uma aplicação dessas propostas ao contexto de ensino híbrido, mas em um contexto correlato educacional (educação presencial e/ou a distância).

Na última década tem crescido o interesse de instituições educacionais pela modalidade de ensino híbrido. A necessidade de inovar suas práticas educacionais, sobretudo a partir da consolidação de recursos multimidiáticos e tecnológicos na sociedade, tem proporcionado o surgimento de uma série de iniciativas de BL em todo o mundo. Essas iniciativas são de diversas ordens de complexidade, desde as mais tímidas (mais simples) até mais ousadas (mais estruturadas).

Diversos autores tem se esforçado no sentido de estabelecer definições e modelos que caracterizem *Blended Learning*. Littlejohn e Pegler (2007) definiram os seguintes tipos de

BL: *space blend* (presencial ou *online*), *time blend* (geografia e disponibilidade), *media blend* (ferramentas, tecnologias e recursos) e *activity blend* (atividades de ensino e aprendizagem, individual ou em grupo).

Bonk e Ghaham (2005) propõe as seguintes categorias:

- *Enabling Blends*: foca na resolução de problemas de acesso e conveniência. Como exemplo, esta categoria provê uma flexibilidade adicional ao aprendiz ou pode prover as mesmas oportunidades ou experiência de aprendizado através de diferentes modalidades.
- *Enhancing Blend*: permite mudanças incrementais à abordagem pedagógica, porém não oferece profundas alterações nas formas de ensinar e aprender. Um exemplo, consiste na adoção complementar ao ensino presencial de recursos didáticos disponibilizados em ambiente *online*.
- *Transforming Blends*: corresponde à experiência híbrida de promover profundas mudanças na condição metodológica existente. Esta mudança pode representar a transformação da postura do estudante, que passa de um indivíduo passivo, sem espaço e sem estímulo inventivo, para um modelo em que o aprendiz abandona essa passividade para assumir um comportamento proativo através de interações dinâmicas. Estes tipos de mudanças não seriam possíveis sem o emprego da tecnologia (BONK e GRAHAM, 2005).

Macdonald (2006) definiu a existência de três abordagens de BL. A primeira, a mais comumente encontrada, porém a mais simplificada, consiste em proporcionar encontros entre os estudantes fisicamente no campus associado a atividades assíncronas *online*. A segunda, representa um cenário comum a cursos a distância, os quais são dimensionados em ambiente virtual, possuindo atividades, encontros síncronos, disponibilização de conteúdos e tendo ainda, esporadicamente, encontros presenciais (aulas, testes, práticas laboratoriais, etc). A terceira abordagem considera tanto alunos com matrículas na instituição presenciais e a distância em uma mesma oferta de uma disciplina, por exemplo. Eles passam a interagir conjuntamente via ambiente virtual.

Como pode ser observado nas citações acima, os autores sempre apontam para a integração entre tecnologia e educação. Isso é facilmente observado uma vez que as ferramentas WEB que proporcionaram a expansão da educação a distância no mundo, hoje também são expandidas ao contexto presencial, seja em uma perspectiva institucional (*sites*, *wikis*, ambientes virtuais de aprendizagem) seja livremente por conta de estudantes (pesquisas *online*, troca de informações via redes sociais, etc). O que se vê comumente é uma adoção tímida e minimamente

planejada de soluções tecnológicas, uma vez que são tratadas como um apêndice ao que se trabalha nos cursos presenciais. Muitas instituições utilizam um AVA basicamente como repositório de notas de aula, exercícios, artigos e áreas de submissão de trabalhos. O planejamento da disciplina permanece praticamente imutável, pois a inserção da tecnologia não vem acompanhada de mudanças sob o prisma metodológico. Assim o professor continua se desgastando ao realizar suas aulas expositivas, que são exaustivas para ele e seus alunos. A tecnologia é, portanto, subutilizada, uma vez que sua aplicação não é dimensionada para transformar a prática de ensino e aprendizagem. Muitas experiências que se intitulam como BL, cometem esse mesmo equívoco.

A questão que se coloca é, qual o melhor cenário de BL? Quais componentes devem ser abordados?

A resposta à questão acima formulada não pode e nem deve ser feita de forma única, uma vez que não existe, pois as realidades sócio-culturais são díspares entre países e regiões como, por exemplo, no Brasil.

Ainda que não se possa ser restrito e determinístico na definição de elementos que constituiriam uma base de regras de ouro para BL, pode-se no entanto recomendar algumas características. McGee e Reis (2012) apresentam as seguintes recomendações:

- **Planejamento:** a preparação de uma disciplina sob o contexto híbrido é completamente diferente do que simplesmente fazer enxertos pontuais de atividades à uma disciplina tradicional (presencial). Portanto, a definição de objetivos bem definidos no plano de aula é de fundamental importância para a composição de uma disciplina no formato de BL, pois o professor definirá os meios de distribuição de informação (em classe ou *online*), o modelo pedagógico (definição de quais atividades serão feitas *online* e quais presencialmente), momentos e locais de encontros presenciais e *online*, além de quais atividades de avaliação serão desenvolvidas e como se dará a composição das avaliações.
- **Estratégias pedagógicas:** um dos pontos mais críticos para o sucesso de um curso BL consiste na definição de ações pedagógicas que equacione harmonicamente as atividades para as perspectivas *online* e presencial. A metodologia a ser adotada deve contemplar atividades formais (atividades conduzidas pelo professor ou reconhecidamente promovidas pela instituição de ensino) e não formais (redes sociais, visitas a museus, feiras de ciências externas, olimpíadas de conhecimento, etc.). A aprendizagem ativa deve ser, em qualquer situação, estimulada pelo professor à seus alunos. Logo, o protagonismo discente deve ser caracterizado, dentre outros aspectos, de uma ação autônoma e de autoconhecimento. Esse

autoconhecimento pode ser exemplificado pela percepção que o aluno tem de como ele se encontra em termos de domínio de assuntos em uma disciplina. Esse monitoramento é comumente encontrado em plataformas voltadas à BL (SHIBLEY, 2009).

- Utilização de Tecnologia presencialmente e virtualmente: a escolha por tecnologias específicas para a realização das ações educacionais deve estar alinhada às estratégias definidas para o curso para a obtenção de um resultado mais eficaz. As soluções a serem adotadas devem ser simples e atrativas, como forma de promover o engajamento dos estudantes (ALBERTS et al., 2010; CRUMET et al., 2010). Os Ambientes Virtuais de Aprendizagem (AVA's) são os principais sistemas utilizados por cursos a distância e experiências de ensino híbrido. A comodidade oferecida por essa categoria de sistemas é atrativa a professores e alunos, mesmo aqueles sem profundos conhecimentos quanto ao manuseio de recursos computacionais. Por outro lado, a aderência das ferramentas aos objetivos do curso deve ser forte, do contrário os alunos podem incorrer em um forte processo de desmotivação e abandono da plataforma. Daí, observa-se que a escolha da plataforma deve ser feita de forma alinhada aos objetivos metodológicos definidos na disciplina (HARNETT, 2009). A maioria das plataformas adotadas pelas experiências de BL são AVA's tradicionais, utilizados em ead. Eles atuam de forma estática, sem capacidade de ação proativa ou autônoma de se moldar ao aluno e seu contexto. Essa é, sem dúvida, uma das grandes limitações que as experiências de BL tem ao fazer uso destas plataformas.
- Estratégias de Avaliação: ainda que as políticas institucionais, muitas vezes, engessem as possibilidades de inovação de novas abordagens de avaliação, o contexto de BL tecnicamente pode estabelecer novas possibilidades avaliativas, como a bonificação auferida pela participação e desempenho em atividades virtuais. É possível migrar a ênfase que o modelo tradicional de ensino dá a avaliações somativas, para as avaliações formativas. Considerando que plataformas *online* podem aplicar quizzes, questionários, exercícios, etc, o professor poderá semanalmente ter um retorno de como seus estão se desenvolvendo ao longo da disciplina e, com isso, poder estabelecer atitudes que venham a melhorar a condição cognitiva dos cursistas.
- Comunicação para a implantação do curso: a implantação de um curso de BL deve ser criteriosa e muito clara aos envolvidos, sobretudo alunos. Os objetivos, formas de distribuição da informação, natureza dos encontros presenciais, dinâmicas presenciais em

sala de aula e/ou laboratório, prazos e avaliações devem ser esclarecidos rigorosamente aos cursistas. Ao aluno deve ser dada a liberdade para que ele possa exercer sua autonomia quanto a seu processo de ensino aprendizagem, estudar em seu ritmo, poder rever os materiais relativos à assuntos já trabalhados na disciplina, etc. O professor deve ter pleno domínio sobre a implantação do plano que elaborou, ser solícito à seus estudantes, dado sobretudo o ineditismo desta natureza de ofertas. Ele ainda deve elaborar sua disciplina com bastante critério, disponibilizando e indicando materiais didáticos, além de exercícios. Tudo isso, como forma de possibilitar que o aluno tenha em mãos um vasto e rico material didático aderente à seu curso (TAN et al., 2010).

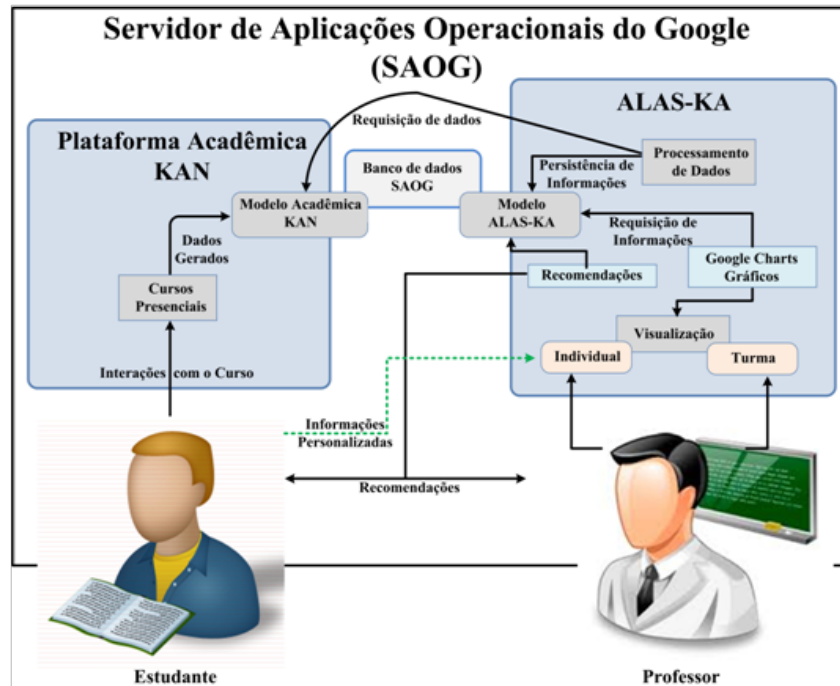
Uma varredura pelos principais trabalhos sobre BL no mundo, demonstra que as adoções tecnológicas por parte de instituições dá-se, prioritariamente, por meio de AVA's. Como já citado anteriormente, estes sistemas, projetados originalmente para atendimento de cursos a distância, são adotados em razão de serem os mais próximos das demandas estabelecidas para o contexto de cursos híbridos.

Contrariando a tendência acima indicada, a plataforma Khan Academy é hoje reconhecida como a principal plataforma de BL no mundo. Esta plataforma é caracterizada pela possibilidade de um cursista desenvolver suas habilidades através de exercícios, vídeo-aulas, dicas, etc. Ela oferece à professores e cursistas um conjunto de instrumentos para análise do comportamento sobre o desempenho do aluno (MURPHY et al., 2014). Mesmo possuindo conteúdos como Matemática, Física, Ciências, Biologia, é no primeiro onde mais se destaca o seu acervo de conteúdos.

Um dos pontos fortes da Plataforma Khan é sua estrutura de relatórios/gráficos (learning analytics) sobre dados de interação do estudante. No entanto, ainda que possua ricos recursos nessa seara, ainda há muitas melhorias que podem ser implantadas. Ruipérez-Valiente et al. (2015), apresentam a arquitetura ALAS-KA que expande os recursos disponíveis pela configuração padrão Khan Academy. Os autores, desenvolveram um plugin que se integra à plataforma Khan. A API (*Application Programming Interface*) Khan foi instalada nos servidores da Universidade Carlos III de Madri. Foi montado um curso de pós-graduação em física que contou com a participação de 100 estudantes. A Figura 20 ilustra a Arquitetura ALAS-KA.



Figura 20 – Arquitetura ALAS-KA



Fonte – Adaptado de Ruipérez-Valiente et al. (2015).

A solução desenvolvida pelos autores amplia a capacidade analítica da plataforma Khan, oferecendo ao professor novos indicadores quanto ao desempenho acadêmico de seus estudantes. Além disso, uma ferramenta de recomendação foi desenvolvida para atuar junto a cursistas e professores na indicação de atitudes, ou alertas que venham a informar esses atores para que possam reagir a determinados cenários controlados por regras definidas por especialistas.

O trabalho de Ruipérez-Valiente et al. (2015) aprimora uma ferramenta já existente na Plataforma Khan (learning analytics) e traz uma nova funcionalidade de sistemas de recomendação. No entanto, esta solução não consegue modificar a experiência de utilização do usuário, ficando limitada ao modus operandi original da Khan Academy.

Morrison & DiSalvo (2014) destacam a característica de jogos presente na plataforma Khan. Com o objetivo de promover uma maior motivação por parte dos cursistas, o processo de interação que considera a superação de fases, desafios, com a obtenção de pontuação/premiação, vem sendo utilizado pela plataforma e corresponde a um dos seus fatores de destaque, sobretudo com o público jovem. Os autores, mesmo destacando a importância do processo gamificado para a promoção de maior engajamento, apontam alguns pontos negativos na plataforma, como a superficialidade da abordagem que atua apenas no aspecto de atribuição de pontos a sucessos em

respostas, quando poderia apresentar aos interagentes soluções mais robustas e efetivas de jogos. A Tabela 1 apresenta a associação de processos motivacionais e como podem ser modelados sob a perspectiva de gamificação.

Tabela 1 – Exemplos de aplicação de gamificação em processos educativos

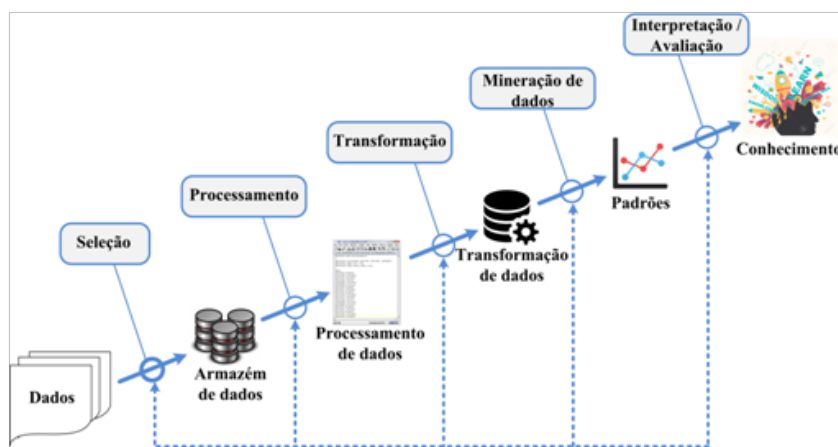
<b>Processos Motivacionais</b>	<b>Gamificação no contexto educacional</b>
Estabelecer objetivos específicos ao invés de generalidades	Especificar claramente sobre o que um aluno precisa fazer para a obtenção de pontuações e medalhas
Indicativo de esforço	Permitir diagnóstico de nível do usuário, para posicioná-lo no estágio apropriado à sua condição cognitiva.
Estabelecer os próximos objetivos, que sejam de fácil compreensão e que possibilitem maior motivação	Premiação com medalhas / pontuação sobre controles individuais, tópicos etc.
Definir dificuldade das metas de longo prazo.	Deve indicar aos alunos que para alcançar um objetivo difícil ou de longo prazo, demandará tempo e esforço que reforçará seu aprendizado.
Focar no processo e não no desempenho final	Premiações para projetos de pares de alunos que encoragem a reflexão, desenvolvimento de processo orientado à elementos de jogos associados à objetivos de aprendizagem.
Prover comparações entre pessoas similares	Painéis (dashboards) mostram o progresso de alunos em relação à outros cursistas de uma turma, quantidade de tempo gasto no sistema por cada cursista para se completar um nível, etc.

Percebe-se que há uma grande limitação quanto ao número de plataformas voltadas a BL existentes hoje no mundo. Mesmo considerando a existência de expoentes como o Khan Academy e Duolingo, essas ferramentas impõem limitações aos professores no sentido de serem restritas à seus programas pré-estabelecidos, dificultando assim a associação direta dos objetivos curriculares às dinâmicas oferecidas pelas plataformas. A partir da apresentação das características das plataformas listadas anteriormente, é possível verificar que boa parte destas iniciativas não incorporam tecnologias que sejam efetivamente voltadas ao contexto de BL nas instituições de ensino.

Uma alternativa à essa situação é o emprego de técnicas de aprendizado de máquina às práticas pedagógicas, oferecendo novas possibilidades a professores quanto à implantação de

curso híbrido. Por outro lado, a proeminente área de mineração de dados que utiliza diversos algoritmos de aprendizado de máquina, tem trazido nos últimos anos uma rica gama de aplicações para a área educacional. Mineração de dados está inserida no contexto do processo de descobrimento do conhecimento a partir de bases de dados (Knowledge Discovery in Databases - KDD). Costa et al. (2012) descreve KDD como um processo geral de descoberta de conhecimento, composto pelas etapas seleção de dados, pré-processamento e limpeza de dados, transformação dos dados, mineração dos dados, interpretação e avaliação.

Figura 21 – Processo KDD



Fonte – Adaptado de Costa et al. (2012).

Mineração de dados educacionais (educational data mining) procura utilizar técnicas e algoritmos para a obtenção de uma melhor compreensão dos dados. Esses dados são obtidos a partir de sistemas educacionais como AVA's, sistemas de tutores inteligentes, MOOC's, plataformas de BL, etc.

Zacharis (2016) apresenta um trabalho de predição de desempenho de estudantes a partir do emprego da técnica de Redes Neurais Artificiais. O pesquisador utilizou dados obtidos de um AVA Moodle em um curso universitário no paradigma de BL, para identificar as possibilidades de sucesso e insucesso dos alunos ao término do curso. Através do algoritmo back-propagation, foi treinada uma rede neural perceptron de múltiplas camadas. A taxa de precisão obtida foi bem elevada, chegando a 98,3%.

Sorour et al. (2015) propõe uma nova abordagem baseada em técnicas de mineração de textos para prever o desempenho estudantil usando análise semântica latente (*latent semantic analysis*) e o método de agrupamento (clusterização) *k-means*. Os experimentos apontaram para uma média de acurácia de 66,4%. Com a aplicação do *k-means* associado ao método *overlap* e

medida de similaridade, os resultados melhoraram e foram obtidos, respectivamente, 73,6% e 78,5%.

Os trabalhos pesquisados demonstram um diversificado cenário de aplicações no campo de TEL. Diante das carências observadas em plataformas de BL, as abordagens como Mineração de Dados Educacionais e Sistemas de Recomendação pode contribuir substantivamente para a qualificação e inovação das práticas pedagógicas. A seguir serão discutidos como essas técnicas podem auxiliar na resposta das questões definidas para a presente pesquisa.

### 3.3 Resultado das questões da pesquisa

#### 3.3.1 *Questão principal: qual o estado da arte de modelos de blended learning?*

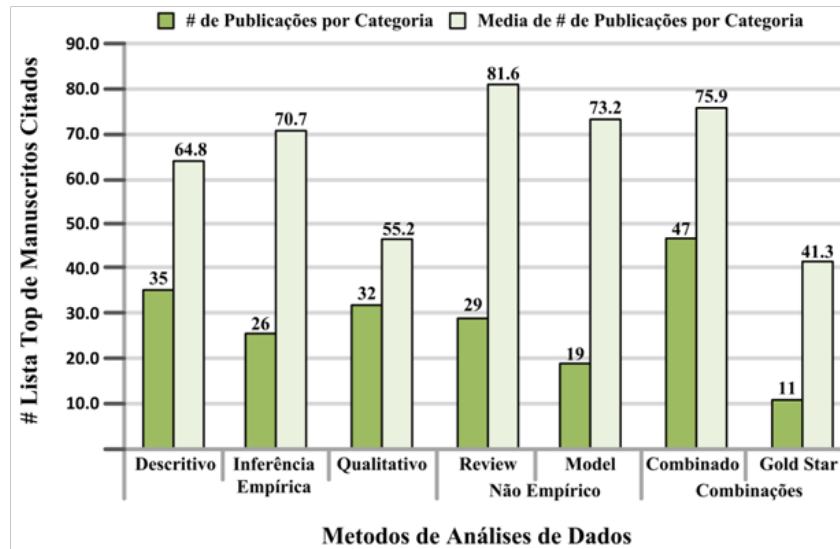
Halverson et al. (2015) realizaram um trabalho de revisão compreendendo as publicações mais impactantes ao longo de uma década na temática de BL. Foram catalogados 60 artigos, 25 capítulos de livros editados, 10 livros e 15 publicações não acadêmicas. Esses trabalhos foram ranqueadas por contagem de citações. Os autores organizaram os trabalhos nas seguintes áreas: metodologias (técnicas de análise de dados), agenda (pesquisa por questões ou afirmações propostas), e frameworks teóricos.

A classificação metodológica distinguiu os trabalhos entre métodos empíricos e não-empíricos. Os estudos empíricos foram divididos em: estatística descritiva, inferência estatística e análise qualitativa. Os não-empíricos foram classificados como revisão da literatura e modelo/tratamento teórico. Publicações que utilizaram mais de um tipo de análise de dados foram classificadas como combinadas. As publicações as quais usaram pesquisa empírica para o desenvolvimento ou aplicação de um modelo/framework teórico foram consideradas gold star e consiste no cenário ideal para pesquisas em BL. A Figura 22 apresenta o gráfico de publicações analisadas de acordo com o método de análise.

Como visto na Figura 22 a categoria de modelos não corresponde à mais desenvolvida pelos pesquisadores de BL segundo Halverson et al. (2015). No entanto, a seguir serão apresentados alguns trabalhos nesta linha.

Hoic-Bozic et al. (2009) desenvolveram um modelo de BL em que misturam a metodologia de aprendizagem baseada em problema (*problem based learning-PBL*), aprendizado independente e discussões colaborativas *online*. Através de um AVA próprio (AHyco), os autores integraram atividades presenciais (face-to-face) com atividades de aprendizado *online*.

Figura 22 – Número de publicações usando cada tipo de análise de dados.

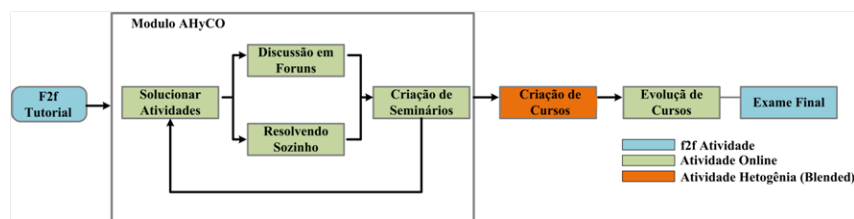


Fonte – Adpatado de Halverson et al. (2015).

A experiência para validação da proposta foi realizada em um curso de pós-graduação da Universidade de Rijeka, Croácia. O curso oferecia de início (dois primeiros assuntos) de forma *online*, e os demais assuntos no ambiente virtual. Como atividade, o curso demandava que cada cursista elaborasse um artigo e enviasse para o AVA para posterior correção pelo professor.

Alguns desses artigos também eram discutidos via fórum. Os alunos também desenvolviam trabalhos em grupo. A Figura 23 ilustra o fluxo deste trabalho.

Figura 23 – Fluxo de atividades.



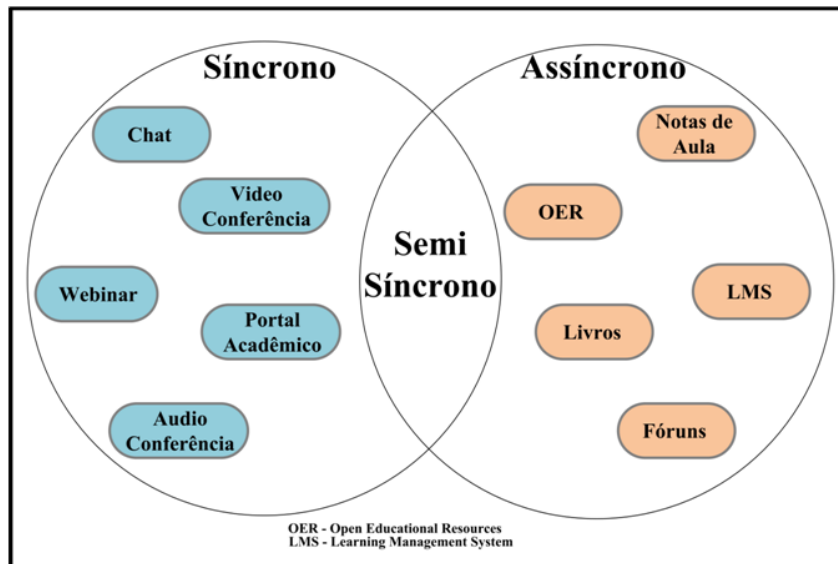
Fonte – HOIC-BOZIC (2009).

O trabalho anterior assemelha-se a experiências de ensino a distância que possui encontros presenciais (semipresenciais). A utilização de conteúdos didáticos hipermedia, o uso de testes de múltipla escolha, são opções comumente encontradas em cursos na modalidade de ead. Nesse cenário, a distribuição da informação dá-se uniformemente aos alunos, sem oferecer uma possibilidade de experiência individualizada aos alunos. A plataforma AHyco é proprietária da

universidade e possui funcionalidades comuns às demais plataformas AVA encontradas. Apesar de possuir um módulo de autoria, na verdade seria um módulo de edição do curso. Onde se organiza quais conteúdos e atividades deverão compor o curso.

Norberg et al.(2011) apresenta a discussão de um cenário de BL não pelo prisma de atividades presenciais e a distância (virtual/*online*), mas pela consideração do aspecto temporal (síncrono e assíncrono). A Figura 24 apresenta algumas possibilidades do modelo BL baseado no tempo.

Figura 24 – Relação de Atividades e Objetos segundo a condição temporal.

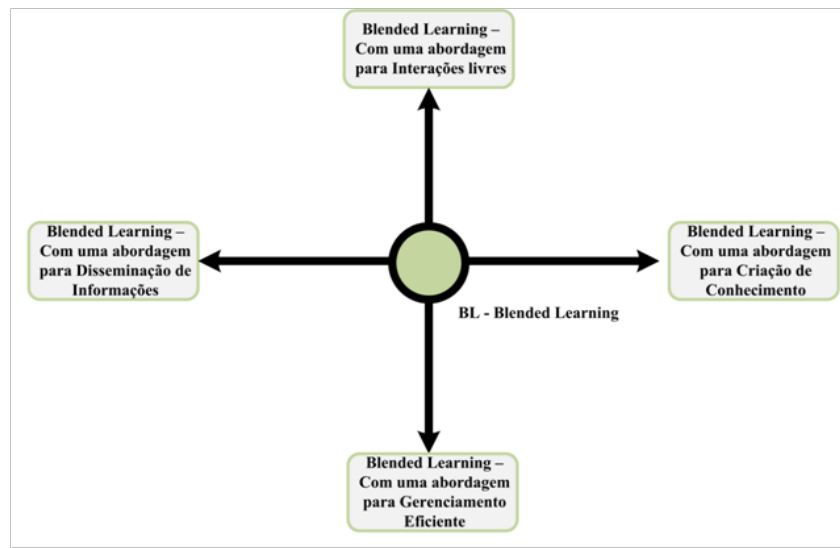


Fonte – Adaptado de Norberg et al.(2011).

Segundo os autores, é importante fazer um bom uso combinado de mídias digitais e diferentes recursos de comunicação para que, associados à diferentes momentos e locais de encontros, se possa fazer um novo processo de construção de aprendizagem em contraponto à sala de aula tradicional. Pela perspectiva de sistemas sociais (WALBERG; ANDERSON, 1968; GETZELS;THELEN,1960), *blended learning* baseado no tempo sugere mudanças fundamentais na postura de professores e alunos os quais, respectivamente, passam a ser mediadores e atores motivados e integrados com seus pares. Os autores, portanto, apontam que a adoção de BL deve representar uma efetiva transformação metodológica. Jung e Suzuki (2006) desenvolveram um framework voltado à análise de abordagens instrucionais de aprendizado, o qual categoriza tipos de modelos BL a partir de quatro perspectivas instrucionais, a saber: interação aberta, criação do conhecimento, gerenciamento da eficiência e disseminação da informação. Em sua aplicação, os

autores definiram que a cada encontro presencial, seriam trabalhadas três sessões de chat, além de fórum específico com especialistas na área. A Figura 25 apresenta as quatro categorias de análise deste *framework*.

Figura 25 – Um *framework* para análise de abordagens instrucionais em BL.



Fonte – Adaptado de Jung e Suzuki (2006).

Dziuban et al. (2006) apresenta o modelo M Model a ser aplicado às unidades curriculares de um curso. Nestas unidades curriculares, as atividades são divididas entre *online* e presencial. Os autores exemplificam a realização de um encontro presencial semanal e as demais atividades serem feitas *online*.

De acordo com Ross e Gage (2006), os modelos de BL no ensino superior podem se encaixar dentro de três perspectivas: 1) adição de atividades *online* ao curso sem que haja uma redução do tempo em sala de aula (*web-supplemented courses ou technology-enhanced courses*); 2) parte do tempo em sala de aula é substituído por atividades virtuais e 3) o estudante determina qual parte do curso quer fazer em regime presencial, em regime de blended-learning e, ainda, totalmente *online*.

Marques (2012) propôs uma solução que se situa na segunda categoria apontada por Ross e Gage (2006) e está no nível de atividade curricular. O pesquisador reforça a importância do professor realizar um planejamento estruturado de sua aplicação de BL e fazer suas escolhas pela intensidade de virtualização de seus processos e atividades de forma a fazer um correto e oportuno uso, no intuito de potencializar o atingimento dos objetivos de um curso. O autor optou pelo modelo de Múltiplas Perspectivas. Esse modelo desenvolve junto aos alunos um processo

de construção do aprendizado em nível avançado. Com o uso de mídias educacionais (texto, áudio, vídeo, animações, etc.). Este modelo é constituído de dois componentes: o processo de desconstrução e o apoio *online* com reflexão (Figura 26). O processo de desconstrução visa explorar casos disponibilizados em um AVA sob a forma de objetos de aprendizagem. As discussões sobre um caso podem se dar a partir de atuação direta do professor ou através da inserção de elementos externos (vídeos, podcasts, etc.). Nesta etapa, o professor deve planejar quais eventos/tarefas constarão no AVA, qual o tempo necessário para que os objetivos de aprendizagem sejam atingidos pelo grupo de aprendizes, e montar um agendamento de sessões de chat ou videoconferência para tirar dúvidas dos alunos. O componente apoio *online* e reflexão centram-se na aprendizagem do aluno com o apoio do docente. São disponibilizados fóruns específicos contendo casos em que os alunos deverão discutir e fazer associações com exemplos de outros contextos e aplicações. Esse espaço de discussão assíncrono é muito rico, pois permite, segundo o autor, o aprofundamento de temas e forte interação entre os cursistas. O Modelo de Múltiplas Perspectivas, adotado pelo autor, foi desenvolvido para o contexto de *e-learning*. No entanto, sua utilização de forma conjugada à educação presencial, possibilita um rico cenário de *blended learning* para a aquisição de conhecimentos.

Figura 26 – Componentes do Modelo de Múltiplas Perspectivas.



Fonte – Própria

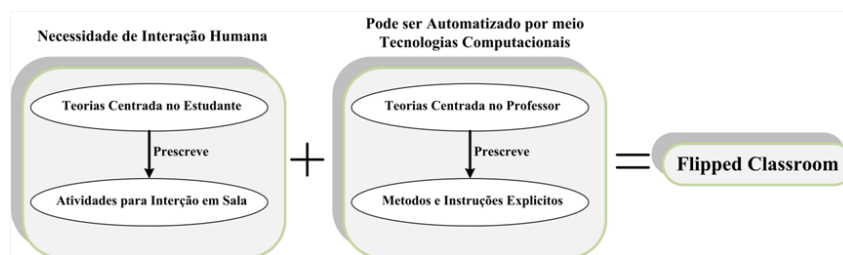
Associado ao conjunto de discussões em TEL e BL está a categoria de *flipped*



*classroom* – FC (sala de aula invertida). Mesmo não sendo uma ideia nova, foi com os últimos avanços tecnológicos que o assunto obteve uma maior evidência (DAVIES e WEST, 2013). Lage et al. (2000) conceberam o termo “*inverted classroom*” que, para os autores, significava que as atividades comumente trabalhadas em sala de aula passavam a ser realizadas fora da sala de aula e vice-versa. Os autores modelaram uma disciplina de Microeconomia, onde os estudantes tinham que ler capítulos de livros, realizar a visualização de vídeos de palestras e fazer a leitura de notas de aula em apresentações multimídia narradas antes dos encontros presenciais. Durante os encontros presenciais, os alunos eram estimulados a aplicarem os conceitos já estudados de economia nas discussões e os professores organizam mini-palestras para responder as dúvidas dos alunos. O desempenho desse grupo de alunos foi comparado com outra turma que cursou a mesma disciplina sob o formato tradicional de ensino, e foram constatados resultados positivos, evidenciando uma maior motivação por parte dos alunos pertencentes ao experimento de sala de aula invertida (VALENTE, 2014).

Bishop e Verleger (2013) caracterizaram um modelo de FC, definindo-o como uma técnica educacional que possui duas partes: atividades de aprendizagem interativa em grupo realizadas em sala de aula e instrução individual por meio de recursos computacionais (Figura 27).

Figura 27 – Modelo de *Flipped Classroom*



Fonte – Adaptado de Bishop e Verleger (2013).

Bergman & Sams (2012) apresentam o modelo “*flipped mastery*”. Neste modelo, os alunos avançam sobre novos conteúdos/objetivos a medida que eles alcançam um determinado domínio sobre um assunto. Nesse modelo, os alunos podem assumir ritmos diferentes. Os autores ponderam que sua abordagem se difere das abordagens mais tradicionais de FC que, por exemplo, oferecem aos alunos em seus momentos *online* (fora de sala de aula) materiais didáticos idênticos a todo o corpo discente para que possam nas aulas presenciais discorrerem sobre o mesmo objeto. Os autores questionam essa abordagem única. Eles colocam a importância de

que cada aluno interaja com os objetos educacionais que desejarem, oferecendo assim uma perspectiva de aprendizado personalizado, garantindo que cada aluno siga na disciplina sob seu ritmo e interesse. As aulas no formato de *Flipped Mastery* oferecem aos estudantes uma grande flexibilidade de construir seu aprendizado, permitindo a eles a chance de ver, rever, escolher temas de seu interesse e de acordo com a necessidade de cumprimento de objetivos curriculares (LANE-KELSO,2014).

Davies e West (2013) destacam que o uso de uma das plataformas mais consolidadas no contexto de BL e FC, a *Khan Academy*, dá-se de forma trivial (KHAN, 2012). Com a adoção da plataforma é possível que o professor elimine ou reduza seu esforço de realização de aulas expositivas, sobrando mais tempo para o exercício de atividades que sejam mais importantes à seus alunos. Uma desvantagem, porém dessa adoção, é que o professor acaba subjugando seu planejamento à oferta da plataforma de questões, conteúdos e atividades. Um professor não consegue criar aulas complementares, criar exercícios, etc. Esta falta de autonomia do docente impõe restrições para que um leque maior de disciplinas façam uso da solução.

### ***3.3.2 QP01: De que forma a tecnologia tem propiciado a implantação de inovações metodológicas para a melhoria do processo de ensino e aprendizagem em cenários de Blended Learning?***

*Blended Learning* representa um dos temas mais debatidos nos últimos 15 anos em TEL. Drysdale et al. (2013) destacam como uma importante oportunidade de pesquisa a discussão sobre ações que proporcionem um maior grau de engajamento e motivação de estudantes em cursos modelados em BL. Os autores reforçam que mesmo já existindo uma série de iniciativas que buscam atender às preferências dos interagentes, este campo ainda precisa se aprofundar em novas possibilidades de projeto, para despertar nos aprendizes um maior interesse e engajamento. Neste sentido, Halverson (2016) propõe um framework teórico para o desenvolvimento de engajamento do aprendiz que inclui tanto indicadores cognitivos quanto emocionais, e investiga sua relevância para o contexto de BL.

Boelens et al. (2017) argumentam que o emprego de tecnologia em um contexto de BL aumenta as oportunidades de personalização do aprendizado ao estudante. Em seu trabalho de revisão são apontados apenas cinco trabalhos que abordam a questão da adaptação de conteúdos ou tarefas considerando as diferenças individuais dos estudantes. Nos cinco casos encontrados, os autores argumentam que a tentativa pela personalização do aprendizado foi

realizada através de um AVA e que tanto a mediação do professor, quanto a adaptação pelo AVA, foram projetados para permitir realinhamentos individuais quanto à oferta de conteúdos e atividades de diferentes níveis de dificuldade, oferecidos de acordo com as necessidades de cada estudante (BRODERSEN e MELUZZO, 2017).

Phillips (2016) apresenta um experimento de aprendizagem personalizada (AP). A autora destaca que a implantação desta abordagem resulta em mudanças na postura de professores e alunos. Cabe aos professores adotarem ferramentas adaptadas que permitam o aprendizado flexível em um ambiente estruturado, e ao aprendiz cabe encarregar-se do seu processo de aprendizagem, tornando-se assim responsável pelo seu direcionamento de estudos, desenvolvendo e exercendo maior autonomia. Desta forma, o professor deixa uma postura tradicional de transmissor de conhecimento para um facilitador da construção do conhecimento, enquanto o aprendiz torna-se mais engajado ao processo de aprendizagem, desempenhando maior protagonismo para a construção de seu saber. Outra mudança observada a partir da implantação de AP está na relação entre professor e aluno. A relação tradicional oferece condição hierarquizada, em que o professor está acima do aluno. Já com a abordagem A.P, professor e aluno estão em condição de parceria, numa relação horizontal (PHILIPS, 2016; ZMUDA et al., 2015).

Wongwatkit et al. (2016) propõem uma abordagem baseada na avaliação formativa com o objetivo de melhorar o desempenho estudantil através de um ambiente de aprendizado personalizado. Os autores desenvolveram um sistema de aprendizado web baseado na integração das abordagens diagnóstica e formativa. Um experimento foi realizado com alunos de uma escola primária na Tailândia, sendo dividido em quatro etapas. A primeira consistia em um pré-teste, onde questões sobre sete conceitos de matemática básica no tema de Área do Círculo foram apresentados aos alunos. Esse teste foi aplicado de forma *online*. O resultado desse teste alimentou a decisão de definição de caminhos de aprendizado, em que uma sequência de assuntos foi definida para cada estudante de acordo com seu nível de dificuldade no assunto. A segunda etapa analisou o estilo de aprendizado do aluno. Com o intuito de identificar se os alunos eram adeptos à abordagem visual (A) ou verbal (B), 11 questões foram apresentadas para que essa classificação fosse possível. A partir desta separação, o sistema passava a apresentar ao estudante materiais educativos no formato associado à sua maior preferência. A terceira etapa compreende a avaliação do estudante durante a instrução. Após as etapas 1 e 2 que, respectivamente, identificaram o caminho de estudos necessário a determinado estudante, assim

como o formato mais adequado de apresentação dos conteúdos, os aprendizes passaram então a receber materiais didáticos específicos. Transcorridas essas etapas, as questões foram listadas aos alunos como forma de avaliar a condição de aprendizado de cada um. A quarta e última etapa consistiu na aplicação do pós-teste que é disponibilizado aos alunos assim que cada um finaliza o seu roteiro de estudos. A Figura 28 apresenta o fluxo estabelecido pelos autores do trabalho.

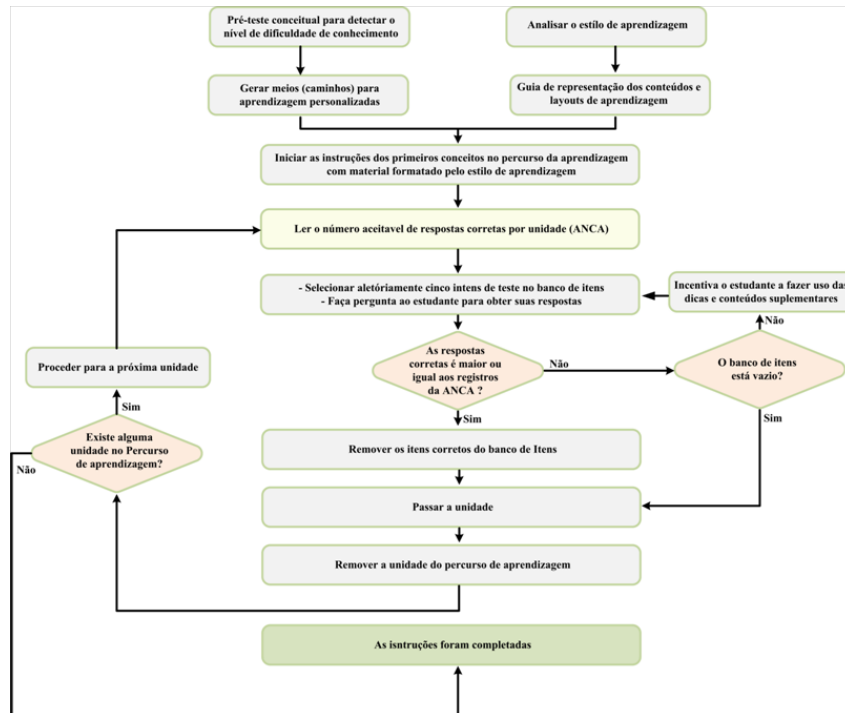
O trabalho utilizou dois grupos de estudantes. Um utilizando a solução proposta, e outro atuando no formato convencional, sem sistema e processo personalizado. Os resultados deste trabalho revelaram que os estudantes do grupo de aplicação tiveram resultados superiores do que os de proposta tradicional e ensino.

Outro importante campo de investigação que tem ganho relevância nos últimos anos é a previsão de desempenho estudantil (*predicting student performance*). Estas pesquisas visam conhecer de forma mais antecipada possível a condição de desenvolvimento acadêmico dos estudantes frente a conteúdos e atividades. O conhecimento prévio da condição do estudante pode ensejar ações por parte de alunos, professores e instituição de ensino para a correção de rumos, reduzindo o impacto de represamentos em cursos e evasão, além de proporcionar direcionamentos mais específicos de estudos, dentro outras possibilidades (ZAFRA et al., 2011).

Sweeney et al. (2015) desenvolveram um sistema para prever as notas de alunos para os próximos componentes curriculares a serem matriculados. Neste trabalho, os estudantes em cada disciplina recebiam notas na escala de A a F. Através do histórico de notas, o sistema previa o desempenho das próximas disciplinas a serem cursadas pelo estudante. Os autores utilizaram a técnica de Máquina de Fatorização, implementando um algoritmo baseado na proposta geral de Fatoração de Matrizes. O experimento utilizou, além da técnica de Máquina de Fatorização, os métodos SVD (*singular value decomposition*), SVD + knn (*k-nearest neighbors*), tendo o primeiro obtido melhor desempenho frente aos demais.

Ashenafi et al. (2016) demonstram como os dados de um ambiente de avaliação entre pares pode ser usado para construir modelos de predição de progresso estudantil. Os autores demonstram o potencial de predição a partir da construção de um modelo de regressão linear para a predição do progresso e performance semanal de estudantes. O experimento realizado neste trabalho utilizou alunos matriculados em dois cursos de pós-graduação ao longo de oito semanas. Destes dados, segundo os autores, foram gerados modelos de previsão de acordo com duas interpretações distintas de predição de performance. A primeira interpretação se

Figura 28 – Procedimento utilizado para caracterização do sistema de aprendizado web personalizado.



Fonte – Adaptado de Wongwatkit et al. (2016).

ateve a comparação de performance do estudante em uma das semanas de duração do curso, considerando os níveis de desempenho passado dos estudantes obtido na mesma semana de um mesmo curso. A segunda observou a medida de quão distante um estudante está de alcançar o nível desejado de desempenho ao término do curso. Os autores argumentam que embora as notas previstas não foram exatas, não produzindo níveis satisfatórios de desempenho para ambas as abordagens, altos níveis de desempenho foram obtidos para ambas as interpretações de progresso de estudante quando as previsões se deram dentro de um nível de mesmo grau. Neste trabalho, os autores não foram conclusivos quanto a discussão das razões que justifiquem um modelo ter um desempenho melhor que o outro, considerando haver fatores latentes que precisavam ser analisados de forma mais detida em futuros trabalhos.

O desenvolvimento da pesquisa na área de predição de desempenho estudantil tem se desenvolvido fortemente graças ao emprego de técnicas de aprendizado de máquina e mineração de dados. No enquadramento de estudos de *Eduming (Educational Data Mining)* temos uma vasta gama de exemplos do uso de algoritmos utilizados para esta finalidade preditiva. Osmanbegović e Siljić (2012) utilizaram três algoritmos classificadores para preverem o resultado final de alunos em um curso (aprovação ou reprovação). Os métodos foram avaliados em relação a seu índice de

acerto de classificação (*accuracy*), facilidade de aprendizado e facilidade de entendimento de seus resultados por parte dos usuários. Os resultados deste trabalho mostraram que Naive Bayes foi o método de melhor desempenho, pois, segundo os autores, além de superar dos demais em termos de precisão, também é uma técnica de mais fácil entendimento por parte dos professores.

Os variados trabalhos apresentados nesta seção não exaurem o contemporâneo e diverso leque de pesquisas e aplicações de TEL. A intenção foi apresentar os principais trabalhos associados ao contexto desta tese.

### 3.3.3 *QP02: Quais aspectos do campo educacional podem ser diretamente trabalhados e aprimorados pela utilização de TEL no contexto de Blended Learning?*

A seguir são descritas algumas questões nas quais as pesquisas discutidas anteriormente podem ser utilizadas com o objetivo de aprimorar as experiências de BL:

- Processo de aprendizagem personalizado: ainda que discutir a personalização de aprendizagem seja uma questão afeita ao campo metodológico (motivação, engajamento, etc.) (BRAY e MCCLASKEY, 2013), faz-se necessária a discussão sobre sua viabilização mediante o emprego de tecnologias da informação disponíveis. As tecnologias devem propiciar ao aluno um atendimento mais adequado à suas necessidades, pois o formato tradicional de educação ignora a heterogeneidade do corpo de aprendizes, promovendo uma homogeneização na ação pedagógica, deixando pelo caminho os indivíduos que, de alguma forma, não acompanham o ritmo ou não se identificam com a abordagem estabelecida pela instituição e/ou professor (WONGWATKIT et al. , 2016);
- Evolução do engajamento de estudantes: a falta de motivação discente é um dos fatores que corroboram para o desânimo de alunos e conseqüente evasão e/ou represamento. A adoção de tecnologias, acesso a múltiplas mídias digitais, dinâmicas diferenciadas para evolução na disciplina, gamificação etc, podem contribuir para um maior interesse dos alunos a um curso (CHEN e HWANG, 2014);
- Autoconhecimento do aluno: o processo de melhoramento da qualidade da educação requer, dentre suas diversas frentes de atuação, que haja uma postura de maior autonomia por parte do aluno (BACICH e MORAN, 2015). Esta autonomia não se dará sem um processo de autoconhecimento. Portanto, é importante a disponibilização de uma ação contínua de avaliação do aluno, para que esse possa ser informado quanto a seu nível de conhecimento sobre os diversos temas abordados em um curso. Além disso, esse *feedback*

deve situá-lo de como ele está diante dos objetivos de um curso e em relação à seu grupo de colegas (CORRIN e DE BARBA, 2014);

- Conhecimento em tempo real da turma: o mesmo processo avaliativo que apresenta ao aluno a sua condição de aprendizado, deve ser oferecido ao professor em tempo real. Desta forma, o docente poderá estabelecer medidas direcionadas a grupos de alunos ou até a nível individual. Toda esta possibilidade deve se dar a partir do emprego de uma avaliação formativa em seu curso. A utilização de instrumentos avaliativos e de análise de dados implícitos gerados a partir da interação dos cursistas em um ambiente virtual de aprendizagem, por exemplo, apresenta-se como uma rica fonte de informação ao professor. De posse desta informação em tempo real será possível ter uma maior eficácia da ação corretiva a ser empreendida (DUVAL, 2011; VERBERT et al., 2013).
- Transformar a postura do professor: o professor que outrora atuava como um ator superior, distante e, muitas vezes, indiferente das questões de cada aluno, passa a atuar mais próximo do aluno, o conhecendo melhor e se permitindo mais ao diálogo. A possibilidade de reduzir o tempo em sala de aula com a atividade de expor conteúdo, permite a abertura de possibilidades mais ricas de interação, como grupos de estudo, resolução de problemas, elaboração de projetos, etc (VALENTE, 2014; BACICH e MORAN, 2015).
- Predição de desempenho: a possibilidade de desenvolver cursos híbridos com a utilização de técnicas de aprendizagem de máquina, viabiliza a oferta inovadora de recursos educacionais. Uma das áreas mais ativas em TEL diz respeito à previsibilidade de desempenho do aluno. Antever condições críticas permite que os diversos atores envolvidos no processo educacional tenham a oportunidade de estabelecer medidas que venham a impedir que situações como abandono e reprovação se concretizem. A utilização de modelos preditivos para este fim tem sido amplamente desenvolvida no meio acadêmico e deve constituir-se como um dos temas mais importantes para o enriquecimento de BL (ZAFRA et al., 2011; ZACHARIS, 2016).

### ***3.3.4 QP03: Como viabilizarum processo de ensino e aprendizagem personalizado?***

A personalização do processo de aprendizado do aluno consiste em uma das principais inovações trazidas ao processo de ensino e aprendizagem. Considerando que o cenário educacional, sobretudo brasileiro, é caracterizado por turmas numerosas de alunos, esta possibilidade de atendimento individualizado por parte do professor mostra-se improvável. No entanto,

a partir da adoção de tecnologias da informação é possível projetar experiências em que um determinado nível de personalização possa ser estabelecido, apoiando professores e alunos nesse processo. A personalização pode ser estabelecida sob determinadas perspectivas, como ritmo, conteúdo, orientação, etc. A personalização por ritmo leva em consideração que cada indivíduo possui um ritmo ou uma condição de momento diferenciada para desenvolver cognitivamente determinados assuntos (BRAY e MCCLASKEY, 2013). Sabedores de que o mais importante é que o máximo possível de alunos cheguem ao final da disciplina aprovados, ou seja, tendo atingido aos objetivos de um curso, pouco importará se um aluno foi mais célere ou não no cumprimento de exercícios e requisitos em geral.

Uma das técnicas computacionais que tem sido vastamente adotadas para diversos domínios é a de Sistemas de Recomendação. Esta técnica tem se mostrado como um promissor recurso para viabilizar a personalização do processo de ensino e aprendizagem. Desta forma, o aprofundamento de pesquisas com a exploração desta técnica aplicada a BL tende a contribuir significativamente para a construção de ambientes virtuais verdadeiramente personalizáveis.

### ***3.3.5 QP04: Como atender aos principais requisitos necessários à implantação de uma solução tecnológica de Blended Learning?***

As experiências trazidas pelos trabalhos apresentados até o momento nesta tese, via de regra, contemplam aspectos pontuais no âmbito de BL. Desta forma, observa-se uma ausência de uma solução metodológica e/ou computacional que contemple os principais elementos que qualificam um curso BL. São pontos importantes que devem ser cobertos por um curso BL:

1. Ambiente para redução de necessidade de aulas expositivas in loco.
2. Personalização de Processo de Aprendizado do Aluno.
3. Visão individual e geral quanto ao desenvolvimento cognitivo dos alunos frente ao cumprimento dos objetivos de um curso.
4. Planejamento integral de um curso BL, permitindo ao professor a utilização de recursos pedagógicos em pleno acordo com os objetivos e elementos curriculares de sua disciplina.
5. Antecipação de análises sobre desempenho, permitindo que professores, alunos e instituição estabeleçam planos de contingência, minorando eventuais danos.
6. Avaliação formativa, como forma de prover a qualquer tempo uma análise sobre o rendimento dos alunos.
- 7 . Promoção de maior engajamento dos estudantes com a disponibilização de uma estru-



rada solução que seja tecnológica e multimidiática.

Se considerarmos que um cenário ideal de BL deva contemplar todos os pontos acima elencados, não há no presente momento desta pesquisa, nenhuma solução que ofereça tal situação.

Dada esta constatação, este trabalho apresenta uma solução metodológica que abrange todos esses elementos. Sobre este modelo, foi desenvolvido uma plataforma tecnológica que integra em módulos todas as funcionalidades necessárias para oferecer ao professor e ao aluno, uma ampla possibilidade de trabalho sob a égide do tema de cursos híbridos.

## **4 MODELO CONCEITUAL PARA ELABORAÇÃO DE CURSOS EM *BLENDED LEARNING* – DIGITAL CLASS**

As pesquisas no campo de BL tem demonstrado avanços quanto a oferta diferenciada de cursos nesta modalidade em comparação com o regime tradicional de ensino. Nos capítulos anteriores, vimos experiências que promoviam a melhoria da motivação e engajamento estudantil, a personalização do processo de aprendizagem, a transformação do papel do professor e aluno, a ressignificação da sala de aula, etc. O atendimento a estas e outras questões inseridas no contexto de cursos BL tem se dado na literatura de forma muito segmentada. Portanto, a integração destes frentes de atendimento é uma questão a se explorar no campo investigativo de BL.

O comportamento isolacionista e muito especializado do cenário de BL evidenciado na literatura científica, motivou que essa tese contemplasse os principais temas atualmente trabalhados nesse domínio, concebendo um modelo teórico que oferecesse a professores e alunos uma nova perspectiva de modelagem de cursos em BL.

Na seção 4.1 será descrito o Modelo Conceitual *Digital Class*. Na seção 4.2 será apresentada a Arquitetura Computacional *Digital Class* que corresponde à implementação tecnológica do Modelo conceitual.

### **4.1 Apresentação do Modelo Conceitual *Digital Class***

A promoção de mudanças paradigmáticas, sobretudo as que se projetam como disruptivas são, via de regra, de alta complexidade de implantação. Tal fato dá-se sob diversas circunstâncias, como burocracia institucional, conservadorismo docente, desconhecimento de pais e alunos, etc. Desta forma, a implantação de *Blended Learning* nas instituições segue uma série de empecilhos inerentes à toda mudança, sobretudo as que exigem transformações comportamentais de instituições, profissionais e aprendizes. Por outro lado, o descrédito dado à educação tradicional tem sido ponto pacífico de entendimento de que algo precisa ser feito para que a educação esteja mais associada aos diversos aspectos inerentes à contemporaneidade tecnológica e midiática.

Ao observar-se na literatura especializada os inúmeros trabalhos nesta seara, percebe-se que sua implantação ampla e efetiva requer um acolhimento de frentes segmentadas. Diante deste cenário, a presente tese define um modelo que colige os principais avanços em BL. Além disso, modernas e avançadas técnicas de aprendizado de máquina e sistemas de recomendação são inseridas no contexto de BL, trazendo uma inovação aos experimentos nessa temática. É

importante ressaltar, como já demonstrado nos capítulos anteriores, que o processo de gamificação tem tido destaque em alguns trabalhos na promoção de motivação estudantil. No entanto, conforme apontado por Morrison & DiSalvo (2014), há na maioria das experiências observadas no campo científico um trabalho superficial quanto à gamificação. Os autores destacam que boa parte das iniciativas alusivas à essa abordagem são resumidas à atribuição de pontos e medalhas, e carecem de um estudo aprofundado sobre o potencial real do uso de técnicas de jogos aplicados à práticas educacionais. Em virtude desta pertinente crítica, optou-se nesta tese por deixar a inserção de processos gamificados para um segundo momento em trabalhos futuros, já sob uma perspectiva evolutiva do Modelo *Digital Class*. O Modelo *Digital Class* objetiva auxiliar um professor na modelagem de um curso/disciplina a ser oferecido na modalidade de BL. Foram definidas seis dimensões que, uma vez adotadas, permitirão a professores e alunos a realização de uma estruturada experiência de BL. As seções seguintes descreverão cada dimensão prevista.

#### **4.1.1 Adequabilidade a objetivos do curso**

A criação de um curso em BL requer do Professor uma reflexão bem estruturada. Em primeiro lugar, o Professor deverá se certificar das questões de caráter legal e regramentos institucionais para não ter que desenhar um curso que venha a ser impedido de ser realizado. Um exemplo de um ponto a ser verificado é quanto à aula expositiva. Elemento inerente ao formato tradicional de ensino, em BL esta exposição pode e deve ser reduzida no face-a-face (presencial) e transposta para o ambiente virtual. Muitas instituições podem se opor à essa mudança, muitos pais podem reclamar e, com muitas vezes, os alunos também. Isto se dá pelo ineditismo da ação e pela falta de informação aos envolvidos. Daí, emerge a necessidade de um trabalho prévio de sensibilização à toda a comunidade escolar de forma a obter o apoio de todos os envolvidos, evitando assim desgastes e desencontros futuros.

Superada essa etapa de esclarecimentos, apoios e chancelas, será importante a definição de um plano de aulas que permita realizar todo o dimensionamento do curso. Assim, o Professor deverá definir quais as atividades irá desenvolver presencial e virtualmente, quais aulas serão presenciais e quais não. Quais materiais didáticos deverá adotar, quais itens/questões alimentará no banco de questões para as avaliações formativas, quais ferramentas tecnológicas serão utilizadas para a realização do curso, como serão feitas as avaliações, qual o cronograma da disciplina e, sobretudo, o Professor deverá deixar seu plano flexível para ajustes de percurso, uma vez que ele deverá ter uma série de informações para ir acompanhando o desenvolvimento

dos estudantes. Se o professor for trabalhar com conteúdos próprios, deverá antes reservar um bom tempo para produzir aulas, questões e atividades. Se for utilizar acervos públicos da internet, deverá mapeá-los para compartilhamento com os alunos.

O professor deve ter muito cuidado ao utilizar soluções tecnológicas que não permitam um controle amplo de sua dinâmica de fruição de conteúdos e metodologia de operação. Esta observação é colocada em razão da existência de plataformas de BL na web que possuem uma estrutura própria de funcionamento e uso de escopos de conteúdos previamente definidos. Isto faz com que professores que queiram utilizar tais ferramentas se subjuguem por completo ao que é ofertado. A capacidade de personalização dos recursos ao que se quer efetivamente trabalhar é mínima e, em alguns casos, essa opção não é dada. O impacto ao curso é que um aluno poderá trabalhar uma série de assuntos que estejam fora do escopo da disciplina, ou uma abordagem que não seja a melhor, uma linguagem que esteja distante da realidade a ser cobrada na disciplina da escola, etc.

No Modelo *Digital Class* a recomendação é que sejam adotados recursos tecnológicos que permitam uma experiência sob medida aos objetivos de um curso. Cabendo ao professor projetar com alto grau de liberdade o seu curso.

#### **4.1.2 Avaliação contínua**

O conhecimento sobre cada aluno ao longo de um curso é uma condição necessária para se estabelecer uma postura de atendimento individualizado ou agrupado, de forma a promover uma maior eficácia no esforço de ensino na busca pelo aprendizado do aluno. Bloom et al. (1971) foram os primeiros a discutir múltiplas estratégias de avaliação, divididas em três etapas, diagnóstica, formativa e somativa. A avaliação diagnóstica objetiva verificar se os aprendizes possuem conhecimentos e habilidades necessários à novas aprendizagens, ocorrendo no início de um curso. A formativa ocorre ao longo do processo de aprendizagem e objetiva a correção de rumos no processo educacional com a sugestão de ações alternativas para a solução de problemas. Por fim, temos a avaliação somativa que ocorre ao final de um curso, com o objetivo de mensurar resultados. De posse da informação advinda da avaliação, tanto professores, instituição e pais de alunos poderão tomar decisões sobre quais medidas precisarão ser tomadas para contornar obstáculos a tempo de não inviabilizar a continuidade de um aluno em um curso, contribuindo para uma possível redução de represamento em uma disciplina.

A presente dimensão na arquitetura *Digital Class* é de suma importância para dar

início aos processos automáticos de recomendação, pois não é possível estabelecer nenhum tipo de sugestão sem saber conhecer o interagente, suas habilidades, preferências e o contexto que está inserido. Desta forma, o conhecimento logo nas primeiras semanas do curso a partir de um teste diagnóstico é um pontapé inicial para que todo o processo possa se desenvolver. Além disso, a avaliação formativa proporcionada pela arquitetura, é feita de forma contínua, diária, a cada interação dos usuários com a solução. De forma automática via sistema de recomendação, ou humana, a partir da intervenção do professor, ações de conserto são empreendidas no sentido de contornar as dificuldades encontradas na avaliação. Ao considerarmos que a realidade escolar brasileira é exaustiva aos docentes que acumulam excessiva carga-horária de aulas, com turmas com grande número de alunos, é praticamente impossível que o professor tenha condições de atentar-se para as dificuldades de percurso de seus alunos e, muito menos, que organize planos individuais de estudo considerando o estágio cognitivo de cada cursista. O processo de avaliação no Modelo *Digital Class* é complementado pela sistemática de previsão do desempenho estudantil, que será descrito mais adiante.

#### **4.1.3 Paralelização de estudos**

Todo curso ou disciplina possui um data de início e fim, assim como um conteúdo programático a ser cumprido. Porém, como lidar com situações como a necessidade de que alguns alunos façam revisão de conteúdos anteriores à presente disciplina? Como reforçar assuntos vistos no curso se o calendário segue a flecha do tempo que é de única direção sem possibilidade de retorno? Não seria possível viabilizar que o ritmo particular de cada aluno fosse respeitado, assim como um professor não conseguiria terminar seu programa, sem que fossem respeitados a possibilidade de paralelização de trabalho de temas pelos cursistas. Assim, um professor poderá dar seguimento à sua disciplina sem abandonar pelo caminho os alunos. O Modelo *Digital Class* indica a importância dos cursos poderem disponibilizar aos cursistas formas de realizar o trabalho concomitante de temas até que atinja desempenho satisfatório. Além disso, esse recurso deverá permitir a verificação ao longo do tempo de como está a memória do estudante. Portanto, o reconhecimento da compreensão do aluno para alguma temática não é uma informação estabelecida *ad aeternum*, ela é um retrato no tempo. Deve-se, portanto, estabelecer um processo para essa verificação que, preferencialmente, pode se dar via meio tecnológico.

Uma outra contribuição trazida por esta dimensão diz respeito a alunos que cumprem

dependência em disciplinas. Assim um aluno pode dar seguimento à sua vida acadêmica e ter a seu dispor, paralelamente, todo um programa específico aos temas em débito.

#### **4.1.4 Ritmo próprio**

Desde a pré-escola até os bancos acadêmicos, as turmas de alunos são fortemente caracterizadas pela heterogeneidade de pensamentos, preferências, aptidões, etc. No entanto, é também comum vermos que as trajetórias de formação são realizadas de forma uníssona, onde essas particularidades dos formandos são praticamente ignoradas. Essa “pasteurização” de mentes é uma característica forte do ensino tradicional. Urge então a necessidade de um ensino personalizado. Essa é uma discussão antiga. O Sistema Personalizado de Ensino (Personalized System of Instruction-PSI), proposto em 1963 pelos professores Fred Keller, Carolina Martuscelli Bori, John Gilmour Sherman e Rodolpho Azzi, já apresentava um modelo flexível e disruptivo para a mudança do paradigma comportamental do professor e estudante (Todorov et al., 2009). Nesse contexto, o professor deixa o papel de expositor de conteúdos para um mediador do conhecimento, acompanhando e gerenciando o processo de aprendizado de seus alunos (Todorov & Tristão, 1975). O PSI possui cinco características principais, a saber:

- Domínio sequencial de conteúdo
- Ênfase na palavra escrita
- Ritmo próprio
- O papel indispensável do Professor
- Aulas de demonstrações como veículo de motivação

Das cinco características tradicionais acima descritas para PSI, uma delas será acrescida pelo Modelo *Digital Class*, “ênfase na palavra escrita” que, em nosso modelo, passa a ser “ênfase na informação multimídia”. Esta alteração é perfeitamente aderente ao novo contexto em que estamos imersos, em que as múltiplas mídias educacionais, fartamente oferecidas no ambiente web, são apresentadas de diferentes maneiras. Caberá ao estudante, através de sua preferência (gosto), acessar os materiais no formato que melhor lhe aprouver.

O caminhar sequencial é importante para que o aluno possa galgar progressos, sendo orientado a resolver questões e visualizar assuntos do nível básico ao avançado. No entanto, é preciso permitir que o aprendiz, a qualquer tempo, possa trabalhar com os temas que lhes são de interesse. Na *Digital Class* esse trabalho pode ser individualizado, não isolado. A importância de oferecer uma condução ao aluno para que ele atinja o objetivo definido pelo planejamento

do professor é um grande diferencial em relação a outras experiências em que o self-paced é estimulado.

No Modelo *Digital Class* os alunos podem estar em momentos distintos em termos de conclusão de estudos dos temas atinentes à seu curso. Serão, portanto, respeitados as diversas idiosincrasias presentes na turma de alunos. Com isso, respeita-se outro importante aspecto exigido de um bom aluno em qualquer modalidade de ensino, a autonomia. O comportamento autônomo deve ser estimulado a todo o instante, pois forja o aluno para uma construção do aprendizado mais transformadora. Esse perfil autônomo de tão importante, transpassa o contexto acadêmico, sendo exigido nos mais diversos ambientes como, por exemplo, o laboral. Como o termo autonomia não deve ser confundido com autodidatismo, caberá ao professor estabelecer as bases para esse desenvolvimento do aluno. A disponibilização de ferramentas computacionais que auxiliem o aluno em sua busca pelo conhecimento deve ser realizada.

#### **4.1.5 *Predição de desempenho***

Identificar com antecedência se um aluno terá problemas ao término do curso é uma das tarefas mais importantes para se poder atuar para a resolução do problema. Cursos tradicionais que operam sobretudo na perspectiva de avaliações somativas que são feitas no final do curso, acabam por não ter os melhores elementos para se tomar decisões de correção de rumo que, se fossem tomadas a mais tempo, mudariam a história para um grande número de estudantes.

O Modelo *Digital Class* indica a criação de um modelo preditivo para o desempenho de um aluno que utilize as técnicas mais modernas de inteligência computacional como aprendizagem de máquina, fatoração de matrizes e até técnicas estatísticas. Como já apresentado no Capítulo 3, esta área tem suscitado uma série de investigações na comunidade científica de TEL. Porém, é inédito a determinação explícita de seu uso a um cenário de aplicação de um curso em BL.

O Modelo *Digital Class* prevê uma avaliação contínua, ou seja, um plano de avaliação formativa. Portanto, uma quantidade representativa de dados tende a ser produzida a partir das interações dos cursistas com as atividades planejadas pelo professor. A partir destes dados, é gerado um modelo preditivo que permitirá ao professor saber com antecedência qual será o provável resultado de determinado aluno para uma determinada configuração de teste. Assim, uma vez que o professor indique qual a composição da prova em termos de tema(s) e até a

complexidade das questões, o modelo deverá indicar com alto grau de confiabilidade qual a nota que cada estudante terá. Isto possibilitará que o professor, em aulas subsequentes à sua análise e antes da prova, possa estabelecer aulas especiais, modeladas de acordo com a maior precariedade identificada pelo modelo. Além disso, a previsão servirá de orientação para dividir sua turma em grupos de alunos em que poderão ser trabalhados, para cada grupo, assuntos voltados a dificuldades específicas. O professor nestas oportunidades poderá utilizar monitores oficiais ou mesmo alunos que o sistema reconheceu como aptos em determinados assuntos para agirem como monitores da turma.

Permitir que o Professor possa prever o desempenho estudantil é uma inovação trazida ao contexto de *blended learning*. Via de regra, as plataformas utilizadas para a oferta de cursos híbridos não possuem mecanismo igual. Portanto, esta é uma das contribuições deste trabalho ao contexto de ensino híbrido.

#### **4.1.6 Personalização**

Uma das características mais importantes de BL é a possibilidade do aluno ter uma atenção pedagógica personalizada. Esta, porém, é uma questão de alta dificuldade de ser operacionalizada, dada a condição de turmas de alunos de grande tamanho.

O Modelo *Digital Class* considera altamente relevante a dimensão da personalização. Em nosso modelo, esse processo precisa do apoio tecnológico para que possa se desenvolver de forma mais ampla. Um dos aspectos é a oferta de uma avaliação que permita que o professor de forma rápida e intuitiva tenha acesso aos agrupamentos dos alunos quanto a sua condição de entendimento nos temas. Esses agrupamentos poderão ser feitos a partir de técnicas de mineração de dados como *clustering*. O acesso facilitado do professor às informações sobre o estágio de dificuldade de cada aluno permitirá que o professor tome medidas específicas às necessidades de um aluno ou um grupo. Isto já se configura como um primeiro movimento na busca de um atendimento personalizado. Além disso, o professor poderá contar com o apoio de um sistema que automatize a sugestão de conteúdos e/ou orientações didáticas. Esses sistemas são comumente estruturados como sistemas de recomendação ou tutores inteligentes e deverão guiar os alunos por caminhos de estudo e ofertar materiais específicos para sanar suas dificuldades. Desta forma, o professor terá auxiliares em seu processo de acompanhamento do corpo discente e os alunos, por sua vez, terão um apoio garantido de forma a superar suas principais dificuldades.



## 4.2 Arquitetura *Digital Class*

A literatura tem apresentado um conjunto de experiências relativas ao uso de tecnologias no contexto de *Blended Learning*, as quais, em sua maioria, são realizadas por meio de Ambientes Virtuais de Aprendizagem originalmente projetados para aplicações em cursos na modalidade de educação a distância. Outras plataformas, apresentadas em capítulos anteriores, são classificadas como soluções de BL. Porém são plataformas WEB estruturadas à revelia do planejamento pedagógico de escolas e professores. Em relação aos AVAs, eles permitem que professores e instituições o modelem em conformidade com o planejamento de seu curso. No entanto, seus recursos não permitem, por exemplo, que um dos principais requisitos de BL seja atendido, a personalização quanto a experiência do aluno. Desta forma, tudo que for projetado para uma turma será explorado de forma homogênea por todo o grupo de alunos. Já as plataformas comumente associadas ao cenário de BL, como Khan Academy, possuem um certo nível de navegação personalizada e gamificação. No entanto, a ferramenta não oferece um maior grau de liberdade ao professor para trabalhar melhor determinados conteúdos. O Professor não pode, por exemplo, produzir nenhuma aula e complementar os materiais da Khan, a não ser que utilize uma segunda plataforma, como um AVA.

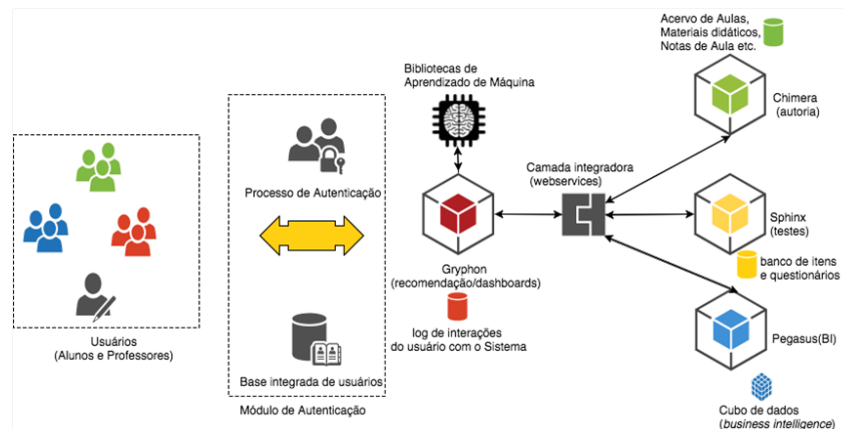
Diante da circunstância acima apresentada, nenhuma plataforma existente na literatura atualmente permitiria a validação do modelo conceitual *Digital Class*, pois algumas das dimensões definidas pelo Modelo não são encontradas nos principais ambientes virtuais de aprendizagem disponíveis como, por exemplo, a predição de desempenho. Para estabelecer todas as seis dimensões previstas no Modelo em um curso, é necessário que o professor adote ferramentas computacionais múltiplas para sua implementação. Desta forma, o professor teria que possuir conhecimentos de programação para implementar alguns dos mecanismos previstos no Modelo, como um Sistema de Recomendação e um modelo preditivo de desempenho estudantil. No entanto, sabemos que essa exigência seria proibitiva à uma ampla adoção do Modelo pela grande maioria dos professores.

Com o intuito de viabilizar a validação de toda a estrutura dimensionada para o Modelo *Digital Class*, foi desenvolvida uma arquitetura computacional que materializa seus preceitos. Esta arquitetura, homônima ao Modelo, oferece a professores e alunos a possibilidade de realizarem uma experiência de um curso em BL contendo todas as dimensões propostas por esse trabalho.

A arquitetura *Digital Class* está dividida em quatro módulos de viés acadêmico:

Chimera (autoria), Sphinx (itens e testes), Gryphon (sistemas de recomendação) e Pegasus (business intelligence). Como forma de evidenciar que a solução proposta é voltada ao ensino híbrido, cada módulo recebeu o nome de um animal híbrido presente na mitologia grega. Associado aos quatro módulos citados, foi estruturado um módulo mais técnico, para centralizar a estrutura de autenticação. A Figura 29 apresenta a arquitetura *Digital Class* e a integração de seus módulos.

Figura 29 – Arquitetura *Digital Class*.



Fonte – Própria.

#### 4.2.1 Módulo de autenticação

Controla o acesso a toda a suíte de serviços disponíveis na arquitetura. Mesmo tendo como padrão principal de uso da *Digital Class* através do Gryphon, os usuários, sobretudo professores, podem acessar diretamente os módulos Chimera e Sphinx para, respectivamente, produzir conteúdos e construir itens. O módulo de autenticação contém todas as informações relativas aos dados do usuário e controla as sessões dos usuários em todos os sistemas. Uma vez que um usuário faça um login, ele poderá acessar a qualquer um dos módulos sem a necessidade de se identificar novamente. Da mesma forma, ao efetuar um logout todas as sessões são encerradas.

#### 4.2.2 Gryphon

Corresponde ao principal módulo da arquitetura. Nele são concentrados os diversos serviços (módulos). Para os usuários, boa parte dos módulos nem existem (Sphinx e Pegasus), pois a interação com os recursos trazidos por esses módulos são realizados pela interface do

Gryphon, ficando assim despercebido, a eles, quais os responsáveis pelo serviço.

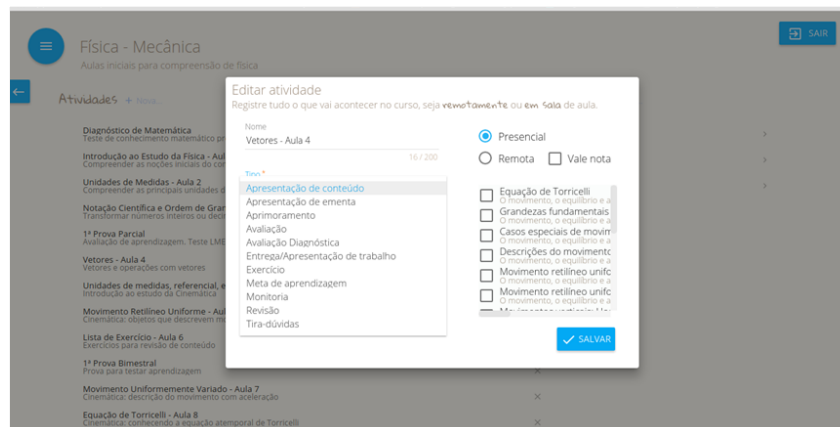
Gryphon é a ferramenta responsável armazenar a estrutura de cursos, manter estimativas da situação atual do aprendizado de cada aluno e, por fim, criar e fornecer planos de estudo e recomendações de materiais para cada aprendiz em tempo real e de modo individualizado.

A interface gráfica do módulo serve de gateway para o aluno, de modo que as outras ferramentas da solução são integradas visualmente de maneira transparente. Assim, os materiais e as funcionalidades da *Digital Class* como um todo são disponibilizadas sempre no momento oportuno sem que o usuário perceba transições bruscas nem necessite realizar operações complexas de navegação.

Este módulo permite a configuração de dois perfis professor e aluno. A seguir eles serão apresentados:

- **Perfil Professor:** oferece as funções de edição do sistema. Nele o Professor poderá configurar todo o seu curso/disciplina, indicando o seu conteúdo programático e todo o seu planejamento didático (como datas, eventos, etc.). A Figura 30 apresenta a visão do professor quanto ao lançamento das informações necessárias para a realização da disciplina no sistema.

Figura 30 – Criação de Atividades de um Curso



Fonte – própria

O sistema disponibiliza ao professor uma interface de cadastro de elementos didáticos. Estes elementos podem ser vídeo-aulas, textos (artigos e livros), simulações, etc. O cadastro de questões é feito no Sphynx, e será discutido em detalhes em seção posterior. Os materiais didáticos devem ser cadastrados de acordo com os temas indicados para serem trabalhados no curso. Além disso, cabe ao professor durante o processo de cadastro, apontar o tema

(conteúdo) da mídia e o seu nível de complexidade. Os módulos da arquitetura trabalham com a indexação dos elementos didáticos em razão de seu tema e grau de complexidade. Esse grau de complexidade se dá em três níveis, básico, intermediário e avançado. Esta classificação mais refinada ocorre no sentido de prover à solução uma condição analítica mais precisa, a fim de prover em seguida uma recomendação pedagógica mais apropriada ao aluno. Esta questão será melhor discutida nas seções subsequentes deste capítulo.

O professor possui uma outra importante funcionalidade a seu dispor, que é a seção de relatórios (*dashboards*). Esta seção oferece ao Professor os principais relatórios acadêmicos para que ele possa identificar em tempo real o desempenho de cada aluno de sua turma nos mais variados temas do curso. Com este recurso, o professor poderá tomar decisões importantes e em tempo hábil para correção de rumos, assim como otimizar os recursos disponíveis (monitoria, tutoria, reforço etc.). Como o processo de diagnóstico é uma ação formativa e caráter continuado, será possível a qualquer tempo que o professor observe o comportamento de cada estudante e de que forma suas turmas estão diante de todas as questões curriculares.

O cenário acima demonstra uma grande mudança do que se observa nas principais plataformas de BL no mundo, sobretudo as que operaram como um serviço *online* (ex.: Khan e Duolingo), que é o fato do professor poder configurar a experiência no sistema de acordo com sua necessidade. Na solução proposta, os alunos só irão trabalhar sobre assuntos que o professor deseja que ele trabalhe, assuntos estes presentes no projeto do curso, em sua construção curricular. Muitas vezes os alunos utilizam as ferramentas na internet acabam por perder um pouco o foco, pois estão sujeitos à trabalharem com um conjunto de temas definidos pelas plataformas, não pelo professor. É evidente que para a implantação de uma disciplina/curso na *Digital Class*, será demandado do professor um grande esforço inicial, pois ele deverá cadastrar questões, atividades, vídeos, artigos, enfim, todo um conjunto de elementos didáticos necessários à realização do curso. No entanto, esse esforço de cadastro tende a ser reduzido ao longo das reedições, ficando a cargo do professor fazer atualizações.

No contexto de BL, os relatórios assumem uma importância crucial para que o professor possa definir as configurações de sala de aula que ele fará para trabalhar de maneira mais específica junto seus alunos. Assim, para auxiliar o docente na tomada de decisões rápidas e corretas, é preciso munir o professor de informações fidedignas sobre a condição de cada aluno ao longo do curso. O professor pode cadastrar vários cursos, assim como várias turmas de um mesmo curso. Ele conta com um mecanismo de aproveitamento de conteúdos e configurações,

assim é evitada a repetição de esforços (retrabalho).

Na *Digital Class*, um curso é definido por uma coleção de temas a serem trabalhados (ementa), uma série de atividades que tratam desses temas e, por fim, um conjunto de turmas, cada uma com seu calendário para as atividades. O planejamento do curso e a alimentação do sistema com as informações pertinentes deve ser o passo inicial para a preparação de um curso na solução. Sugere-se que o planejamento do professor seja realizado no sistema de acordo com as três etapas a seguir:

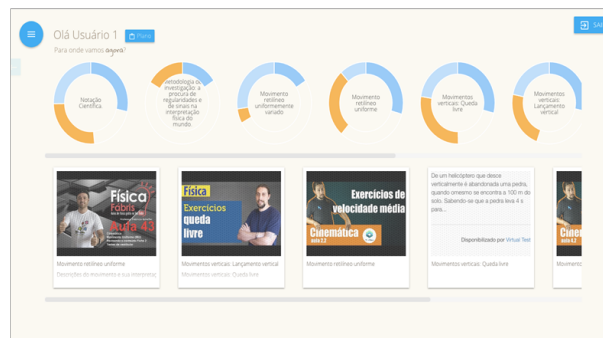
1. Seleção de temas - O sistema disponibiliza para o professor uma estrutura hierárquica com diversos temas facilitando sua busca e seleção. Para o estudo experimental foram disponibilizados conteúdos de ensino de física listados na matriz do INEP (Instituto Nacional de Estudos e Pesquisas Educacionais) bem como outros provenientes das orientações da SEDUC-CE (Secretaria da Educação do Estado do Ceará). Inicialmente, a hierarquia de competências e habilidades do INEP também foi ofertada, sendo removida a pedido do professor instituição envolvidos.
2. Plano de atividades - Uma vez definida a ementa do curso, o professor deve registrar cada atividade planejada. Como tratamos de um modelo de Ensino Híbrido, o ideal é que se encontrem no sistema tanto as atividades presenciais - sejam aulas, provas, encontro, etc. - quanto as virtuais. Na composição de cada atividade deve constar também cada tema abordado por ela. A importância desse detalhamento será elucidada adiante quando for tratada a geração de planos de estudos personalizados.
- 3 . Definição de turmas e calendários—um dos objetivos do Modelo *Digital Class* é estender o potencial do professor e atender com qualidade um número amplo de alunos. É imperativo então que se possa aproveitar as informações e definições de um curso para as turmas. Para tanto, o módulo Gryphon disponibiliza a criação de turmas com calendários. Criar uma nova turma então reduz-se à simples tarefa de nomeá-la e atribuir datas para cada atividade do curso.

Após a execução dessas três tarefas o sistema já tem fundamentação suficiente para iniciar o acompanhamento dos alunos e gerar planos de estudo orientados à sua biblioteca de materiais.

- Perfil Aluno: os alunos de um curso ao entrarem no sistema tem em sua primeira visão do sistema um conjunto de materiais para ele interagir. Essa exibição não é feita aleatória, ela segue os objetivos traçados pelo professor para serem trabalhados ao longo do período do

curso. Outra questão que é levada em consideração é a condição cognitiva de cada aluno. Se eles já tiverem passado por um teste de diagnóstico pelo Sphinx, cada aluno já receberá um conjunto de materiais personalizados. No entanto, se for a primeira vez que estejam entrando no sistema, sem ter passado por nenhum diagnóstico, então serão exibidos os assuntos que estão presentes na linha cronológica definida pelo professor. A Figura 31 apresenta a tela de entrada no Gryphon (visão aluno).

Figura 31 – Visão do aluno no módulo Gryphon.



Fonte – Própria

Um dos principais recursos oferecidos ao aluno, refere-se ao Plano (Plano de Estudos). Este plano tem como proposta indicar ao aluno quais assuntos ele precisa trabalhar e quando novos assuntos devem ser vistos (vinculação com cronograma da disciplina), sempre levando em consideração o desempenho do aluno frente à realização de exercícios.

Uma vez que o *software* mantém em sua base de dados a ementa do curso, os calendários de atividades e o registro detalhado de consumo dos materiais (objetos de aprendizagem) por parte dos alunos, a geração de um plano de estudos personalizado para cada estudante se torna plenamente viável. O processo pode ser dividido nas seguintes operações:

## 1 - Estimativa contínua de desempenho

Por meio do Gryphon, a *Digital Class* realiza uma avaliação formativa do estudante, mantendo uma estimativa de aproveitamento dele sobre cada um dos temas abordados em um curso e revisando essa estimativa constantemente, dadas as interações do usuário com o sistema. A maior parte dos materiais disponibilizados na biblioteca da solução deve ser composta por itens de pergunta/resposta. Diariamente, todos os dados históricos de preenchimento de questões, considerando os temas abordados e o grau de dificuldade de cada uma, são utilizados como

dados de entrada para treinar o sistema. O objetivo é que tenhamos uma função que, a cada vez que o usuário preencha um novo item, o resultado desse preenchimento (acerto ou erro) seja passado como parâmetro e, como retorno, temos a probabilidade de acerto do usuário em uma futura questão resolvida para o mesmo tema com um mesmo nível de complexidade. O algoritmo tratado foi escrito na linguagem R e integrada à api Node.js por meio dos streams de entrada e saída dos processos.

## **2 - Cálculo de desempenho esperado**

O plano de estudos corresponde a um guia, um caminho sugerido para que o aluno trilhe, levando-o de sua condição atual de aprendizado à condição esperada de acordo com o plano do curso. Define-se nessa arquitetura, então, o plano como uma estratégia futura, individual e dinâmica. Futura porque se atém sempre a especificar os próximos passos para um aluno. Trata do que deve ser feito partindo apenas de agora para que os objetivos sejam alcançados. Individual porque, tomando como ponto de partida sempre a situação corrente do aluno, tende a produzir orientações muito diversas para estudantes de perfis diferentes. Por último, dinâmico, posto que o cálculo contínuo da estimativa de desempenho do aluno leva o plano a ser atualizado e corrigido constantemente.

Considerando o cenário acima exposto, a operação seguinte para a definição de um plano de estudos é o cálculo do desempenho esperado do aluno. Além da data no calendário, podemos destacar em uma atividade a lista de temas abordados, se ela será utilizada como avaliação somativa ou o tipo da prática desempenhada. Analisando esses dados, podem ser geradas uma série de expectativas em relação ao aprendizado do aluno. Por exemplo, espera-se que no dia de uma atividade de revisão o aprendiz já tenha conhecimento intermediário sobre o tema. No dia de uma prova, por outro lado, espera-se que o aluno já tenha domínio. São definidas também regras temporais sobre as informações. Digamos, por exemplo, que um aluno tenha a previsão de uma avaliação em seu cronograma. O sistema então apresentará atividades de revisão em seu plano com dias de antecedência como forma de reforçar a preparação para o teste. Abaixo são listadas as regras utilizadas para a geração do plano de estudos:

### **Quanto ao domínio prévio do conteúdo**

- Nenhuma expectativa - Atividades do tipo “Apresentação de ementa”, “Avaliação Diagnóstica” e “Apresentação de conteúdo” não geram nenhuma expectativa anterior. Em outras palavras, não é esperado que os alunos cheguem ao dia da atividade tendo qualquer tipo de domínio sobre os temas. Portanto, os temas são abordados no plano de estudos com prioridade baixíssima e apenas após a apresentação.
- Progresso Inicial - Atividades do tipo “Aprimoramento” e “Aula de Monitoria” geram expectativa de que o aluno já tenha iniciado seus estudos e consiga resolver itens básicos sobre os temas abordados. Assim, os temas são tratados com prioridade baixa no plano de estudos.
- Progresso Avançado - Atividades do tipo “Tira-dúvidas”, “Exercícios” e “Revisão” geram uma expectativa de que o aluno seja capaz de resolver sem maiores dificuldades itens de complexidade intermediária sobre os temas tratados. Os temas são incluídos no plano de estudos com prioridade normal.
- Atingimento - Atividades classificadas como “Avaliação”, “Entrega / Apresentação de trabalho” ou qualquer outro tipo sendo utilizada como avaliação somativa geram uma expectativa de domínio acerca do assunto tratado, de modo que o estudante deve chegar ao dia da atividade sendo capaz de resolver sem dificuldade questões avançadas sobre os temas avaliados.

### **Quanto à temporalidade**

Na solução proposta, o plano de estudos é gerado sempre em tempo real. Além disso, qualquer atividade passada que gere uma expectativa não atingida também é considerada, garantindo que, enquanto algum objetivo do curso não for atingido ele deve ser trabalhado. Respeitando o andamento normal do calendário, as expectativas de desempenho listadas anteriormente só passam a ter impacto no plano na medida em que as atividades que as originam aproximam-se. Além disso, considerou-se que, quanto mais alta a expectativa, maior a antecedência com que o tema deve ser tratado no plano de estudos, de acordo com o seguinte mapeamento:

- Nenhuma expectativa - Os temas só passam a compor o plano de estudos na data da realização da atividade.
- Progresso inicial - 4 dias antes da atividade os temas passam a ser considerados no plano.
- Progresso avançado - 5 dias antes da atividade os temas passam a ser considerados no



plano.

- Atingimento - 7 dias antes da atividade os temas passam a ser considerados no plano.

Quanto às antecipações É importante estimular o interesse dos alunos e dar a liberdade e a possibilidade de ampliar seu conhecimento em seu próprio ritmo. Assim, uma vez atingida uma expectativa do plano de estudos, o tema não é removido. Pelo contrário, é criada uma expectativa mais avançada sendo que a prioridade passa a ser baixíssima. Assim, alunos avançados continuam a ser estimulados a desenvolver seu conhecimento sem, por outro lado, comprometer o andamento do resto do calendário.

Ainda, caso o plano de estudos para o dia corrente tenha sido concluído e todos os temas dominados com sucesso, o plano de estudos passa a considerar os dias subsequentes.

### **Quanto à confiabilidade da estimativa**

O atingimento das expectativas é verificado sempre em relação à estimativa de performance do aprendiz em um tema específico, no modelo descrito anteriormente. Não se pode descartar, no entanto, que a estimativa representa um momento no tempo. Embora o valor passe por constantes alterações, ele pode perder sua precisão com o passar do tempo. Seja por fatores de esquecimento ou pela falta de prática, a probabilidade de acerto de um aluno em uma questão de um tema específico é afetada após um intervalo longo sem atualizações. Tratando desse problema, o modelo proposto traz junto a cada estimativa a ideia de “grau de confiabilidade”. Segundo a regra adotada na aplicação piloto, uma estimativa só pode ser considerada confiável se o usuário preencheu um mínimo de três itens sobre o tema na complexidade dada e se, além disso, a idade do cálculo não for superior a 45 dias.

Como consequência desse conceito, o plano de estudos do aluno sofre pequenas interferências de revisão de conteúdo ao longo do tempo. Por exemplo, se um usuário dominou um conteúdo em fevereiro, em abril a solução volta a incluir em seu plano de estudos a resolução de um item desse mesmo conteúdo. Assim a estimativa de desempenho dele é atualizada e, apenas se necessário, mais materiais similares voltam a atuar em seu plano. A Figura 32 ilustra um plano.

Figura 32 – Plano de Estudos

Fonte – Própria

Ao clicar em algum dos temas do plano de estudos ou mesmo clicando em algum dos thumbs na tela principal, o aluno terá acesso ao objeto (vídeo-aula, exercício, simulação, etc.). É pedido ao aluno que ranqueeie o objeto em seu encerramento. Este ranqueamento será contabilizado para que o processo de recomendação possa ser trabalhado junto aos demais usuários. A Figura 33 apresenta a tela de exibição que permite a interação do aluno com os elementos didáticos.

Figura 33 – Tela de exibição de mídias educacionais.

Fonte – Própria

Vale saber que a medida em que o usuário for interagindo com o sistema, ou seja, a medida que ele responde a uma questão de um determinado tema, o sistema analisa em tempo

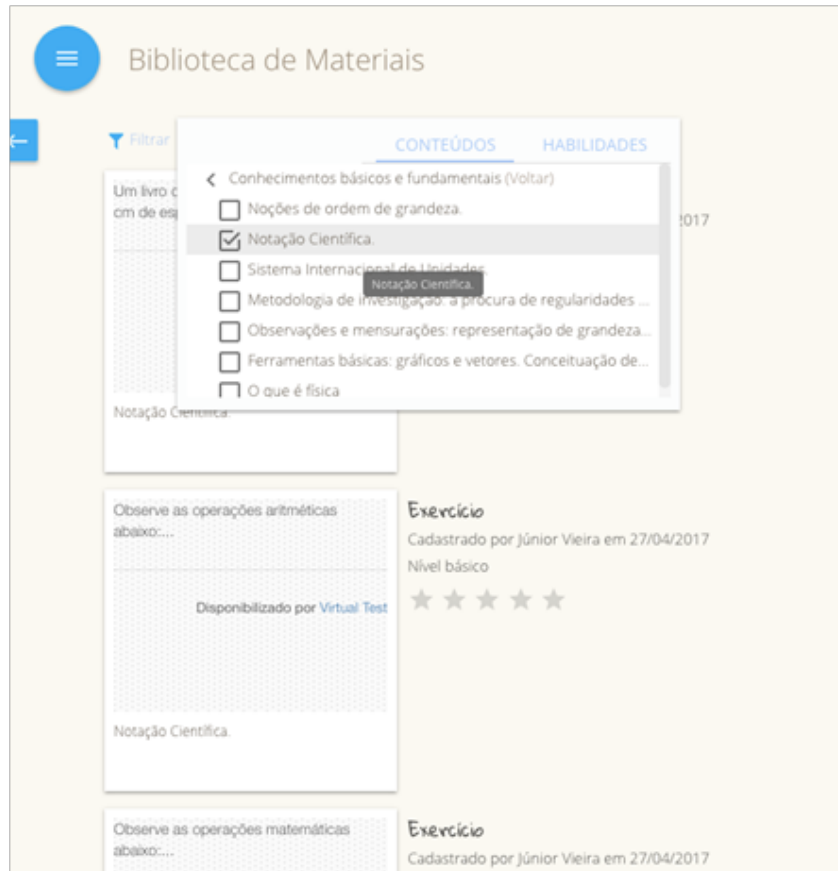
real se a resposta foi correta ou não e dispara um evento de atualização dos itens a serem exibidos na sequência. Por exemplo, a medida que um usuário responde corretamente uma questão sobre um assunto ele continua a receber questões daquele assunto até ele completar o quantitativo de respostas necessárias para que avance de nível ou mude de tema. Caso ele passe a errar as respostas, então o sistema começa a indicar vídeo-aulas naquele assunto para que o aluno possa se preparar melhor antes de voltar a ser testado pelos itens. Este processo se dá de forma individual e cada aluno e, portanto, o sistema se comporta de acordo com o padrão de interação de cada usuário. Esta sistemática atende à um dos temas mais cativos de BL, a personalização de aprendizagem. Esse comportamento não é observado nas tecnologias dos AVAs tradicionais, sendo uma das limitações impostas por essas ferramentas à implantação de cursos no formato BL.

Caso o aluno não queira seguir a condução feita pelo sistema, ele pode realizar buscas sobre o material de forma arbitrária. Assim, na **Seção Biblioteca**, todos os materiais didáticos cadastrados poderão ser acessados pelo aluno. Esta seção é importante, pois um aluno pode ter o interesse de reforçar por sua conta determinado assunto antes de uma prova, ou por qualquer outra razão. Mesmo considerando a importância da recomendação viabilizada pela solução, não é possível o desenvolvimento de nenhuma solução que não respeite a autonomia do aluno.

No Gryphon, o acompanhamento do andamento dos alunos se dá por meio da recomendação, apresentação e aplicação de objetos de aprendizagem organizados e classificados internamente em uma biblioteca de materiais. A coleção é composta por apresentações multimídia, vídeos, itens de pergunta/resposta, questionários, jogos interativos entre outros objetos, todos disponibilizados por meio de plugins integrando módulos da *Digital Class* ou mesmo ferramentas externas. São exemplos o site de *streaming Youtube*, a Wikipédia e até o repositório de simulações virtuais interativas PhET da Universidade do Colorado (EUA).

Para que o acompanhamento de um aluno no Gryphon seja bem sucedido em um curso, é fundamental que a biblioteca seja ricamente alimentada. Para cada tema explorado na ementa precisa existir um número razoável de materiais expositivos e questões de cada nível de complexidade. É também importante que os materiais abordem os temas de forma diversificada tanto em relação ao conteúdo quanto à abordagem ou à linguagem adotada. Essa heterogeneidade é fundamental para o viés de personalização que a ferramenta busca obter na geração de sugestões para cada estudante. A Figura 34 apresenta o recurso de busca nesta seção.

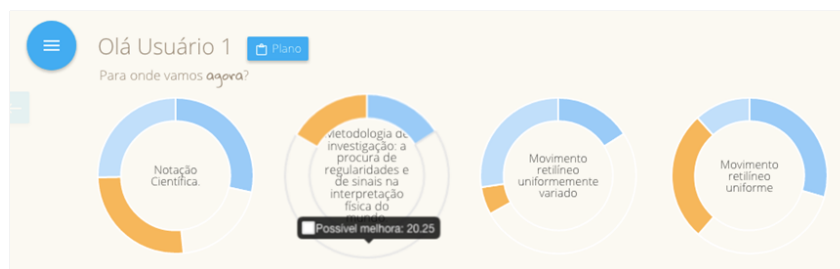
Figura 34 – Filtro de Itens na Seção Biblioteca



Fonte – Própria

Uma das questões principais para o aluno é se situar sobre seu rendimento em uma disciplina. Assim, logo na primeira tela do Gryphon temos a exibição de gráficos que indicam a condição do aluno de acordo com cada tema disponibilizado. Desta forma, o aluno sabe de maneira concisa sobre seu progresso e sobre o que precisa fazer para melhorar seus índices de desempenho. Mais importante de tudo é fazer com que o estudante e o seu professor tenham ciência da sua condição de facilidade ou dificuldade de entendimento, de tal sorte que possam ser empreendidos esforços no sentido de contornar essa situação a tempo de não haver prejuízos irreparáveis à continuidade do aprendiz no curso. A Figura 35 ilustra esses gráficos situacionais.

Figura 35 – Gráficos de situação do desenvolvimento do estudante nos temas do curso.



Fonte – Própria

### 4.2.3 Sistemas de recomendação

Como visto nos capítulos anteriores desta tese, os sistemas de recomendação voltados ao contexto educacional tem auxiliado professores e alunos quanto à uma série de questões, como prevenção da evasão, recomendação de trajetórias de estudos, etc. Esta tese utiliza sistemas de recomendação para apresentar aos alunos planos de estudos voltados à suas condições de aprendizagem de momento e materiais didáticos apropriados. Em outras palavras, o SRE apresenta a cada estudante um roteiro de estudos que ele deve cumprir para atingir os objetivos de aprendizagem definidos pelo professor quando do planejamento da disciplina. Os alunos, portanto, além de receberem um roteiro de estudos baseado no planejamento da disciplina e segundo um cronograma (linha de tempo), também recebem recomendações de elementos educacionais mais adequados ao momento de aprendizagem do estudante. Estes elementos podem ser vídeo-aulas, exercícios, simulações, artigos, etc. Para essa função, a biblioteca de materiais é utilizada como fonte de dados para um serviço de recomendação que utiliza uma abordagem híbrida com características tanto de filtro colaborativo quanto de recomendação baseada em conteúdo.

O SRE desenvolvido utiliza o paradigma colaborativo. O problema do cold start (partida a frio) comum a essa abordagem é resolvido tomando como base a informação de contexto. Essa informação indica o grau de conhecimento de cada aluno de acordo com os temas abordados em um teste diagnóstico. Diante desta informação, o sistema busca, por exemplo, recomendar elementos que sejam associados aos assuntos em dificuldade. A medida que o sistema vai sendo utilizado pelo estudante, o mesmo vai indicando qual seu grau de satisfação com determinado elemento (*ranking*). Com isso, o sistema tende a atender mais adequadamente os diversos perfis de alunos existentes, provendo uma personalização apropriada para cada tipo

de aprendiz. O Gryphon estimula o aluno, após explorar um material recomendado, a atribuir uma classificação (1 a 5 estrelas) ao objeto. Assim, a posteriori, o sistema passará a priorizar os elementos com melhor avaliação, continuando a levar em consideração a informação de contexto, ou seja, o que o aluno precisa para atender às suas dificuldades frente à necessidade de cumprimento dos objetivos do curso.

Com base no histórico de classificação de cada usuário, é disponibilizado um método que considera o interagente, um tema e um nível de complexidade, para retornar uma sugestão de materiais a consumir.

#### *4.2.3.1 Questões técnicas*

Para a implementação do sistema de recomendação foram utilizadas técnicas de filtragem colaborativa, mais especificamente, que utilizam a fatoração de matrizes como abordagem. Por se tratar de uma técnica sensível à detecção de informações latentes presentes em uma matriz de avaliação (utilidade), foi utilizado o algoritmo iterativo mínimos quadrados alternados (alternating-least squares with weighted-regularization). O código da solução de recomendação foi implementado utilizando a técnica SVD e teve como parâmetros para sua geração: fatoração com 10 atributos (features), taxa de regularização 0.05 e 10 interações. O cálculo implantado considera itens avaliados anteriormente pelo usuário. Além disso, é aplicada à pontuação (score) resultante uma ponderação (rescore) impulsionando os itens do tipo questão e penalizando os demais.

Por questões de performance, delegou-se o processamento das recomendações a um sub-módulo escrito em Java com acesso livre de leitura à base de dados. A comunicação entre o core da aplicação Gryphon e esse sub-módulo se dá por meio de requisições HTTP simples e troca de dados estruturados em JSON. Para a implementação das heurísticas de treinamento e recomendação foi utilizada a biblioteca de código livre Apache Mahout.

Para gerar as recomendações adequadamente, são utilizadas informações de usuários, materiais, temas, relações entre temas e materiais bem como scores atribuídos por usuários a materiais. Mesmo após as etapas de treinamento das matrizes, essas informações são acessadas exaustivamente, principalmente durante etapas de filtragem e rescore dos resultados. Portanto, depender de consultas à base de dados impactaria a performance da solução de modo proibitivo. Por outro lado, o volume relativamente reduzido de dados deste domínio - uma ordem de milhares de materiais e alunos, apenas - viabiliza o armazenamento em memória dos dados

necessários sem causar qualquer questão à escalabilidade. Portanto, o sub-módulo cria em memória estruturas hash para mapeamento das entidades de interesse à solução, confiando em mecanismos de atualização descritos mais a frente.

É responsabilidade do software em questão fornecer dois tipos de recomendações: a recomendação de material por tema e complexidade, empregada majoritariamente no plano de estudos, descrita anteriormente, e a recomendação de material relacionado.

A recomendação de materiais relacionados está atrelada a uma característica muito importante do Gryphon, que é a proatividade. Ao adentrar no sistema, o aluno é confrontado imediatamente com uma coleção de materiais de seu plano de estudos; de acordo com seu desempenho para aquela questão (sucesso ou insucesso), o sistema carrega um novo item do mesmo assunto, carrega um próximo assunto, ou mesmo um material didático para auxílio na resolução do tema associado à questão. Tudo é feito para manter o estudo em movimento, propiciando um processo contínuo de aprendizado do aluno, de forma prática, intuitiva, orientada, desafiadora e que promova autonomia do aprendiz. Prestar essa assistência prezando pela individualidade do aprendiz e mantendo uma interface extremamente simples é o objetivo da recomendação de materiais relacionados. Trata-se de um método simples que, dados um aluno e um material, retorna uma recomendação de próximo passo levando em consideração sua estimativa de desempenho. Por exemplo, enquanto o usuário abre um vídeo no sistema já são exibidas ao lado miniaturas de questões do mesmo tema e em um nível semelhante ao tratado no material.

Quando se solicitam materiais relacionados a um item avaliativo (questão), os itens expositivos (como vídeos, artigos, simulações) são impulsionados (têm seus scores multiplicados por 2.0) e vice e versa. O comportamento resultante dessa abordagem é que, no momento que o usuário está assistindo a um vídeo, por exemplo, é mais provável que ele seja guiado a resolver questões a respeito do tema em seguida, testando seu conhecimento. Por outro lado, enquanto preenche questões, é mais provável que sejam sugeridos materiais que possam tirar suas dúvidas se ainda não tiverem sido explorados.

A todo momento os alunos podem utilizar o sistema, progredindo em seu aprendizado e fornecendo novos dados sobre suas preferências. Do mesmo modo como é fundamental rever constantemente as estimativas de desempenho para ajustes no plano de estudos, também é crítico que o sub-módulo de recomendação seja atualizado tão constantemente quanto possível. Esse requisito é atendido através de mensagens enviadas pelo módulo principal do Gryphon ao

sub-módulo de recomendação sempre que são registradas alterações como inclusão de materiais na biblioteca, ajustes na relação entre materiais e temas ou mesmo atribuição de scores por usuários a materiais. Recebendo as mensagens, o sub-módulo agenda uma recarga de seus dados e o início de um novo treinamento de sua engine para os próximos segundos. Como o universo de usuários e materiais empregados em uma instituição é comumente pequeno, não foi notado qualquer impacto negativo na performance da solução durante a aplicação piloto.

Como exposto até aqui, muito da funcionalidade e riqueza do Gryphon provém de suas capacidades de aprendizado de máquina, desde o cálculo de estimativas até a geração de recomendações ao usuário. Embora a utilização dessas técnicas sejam adequadas, ela depende da resolução de uma questão muito relevante para alcançar eficácia: o chamado cold-start. Trata-se do estado da solução em que não existe ainda um acúmulo de dados suficiente para que os algoritmos de aprendizado sejam executados de modo satisfatório. Sem uma coleção histórica de informações sobre respostas de alunos a questões, não é possível treinar sistema para realizar uma previsão da probabilidade de acerto dele a outros itens semelhantes; sem dados de preferências de alunos por materiais, não se pode treinar um motor de recomendação para que ele estabeleça similaridades entre os mesmos e assim por diante. A concepção do Gryphon enquanto arquitetura precisa considerar tais requisitos e limitações de modo que tenha sucesso desde a aplicação piloto, sem perder de vista os ganhos de eficácia que o passar do tempo e o acúmulo de dados fomenta.

O processo de conhecimento sobre os alunos pode se dar de forma implícita (análise de dados comportamentais quanto ao uso de sistemas) ou explícita (informação oferecida diretamente pelo usuário como questionários, entrevistas, etc.). No modelo *Digital Class*, o processo de aplicação de um curso BL tem início com uma avaliação diagnóstica. Esta avaliação diagnóstica representa o marco para que as demais etapas e funcionalidades do sistema Gryphon possa se desenvolver. Desta forma, foi aplicado um teste online composto por questões sobre os temas basilares para o demais pontos a serem trabalhados ao longo do curso (pré-requisitos). Além de sua importância pedagógica, a prática é uma estratégia simples que, quando adotada, permite a geração de planos de estudos individualizados desde os primeiros dias, baseando-se em estimativas de desempenho personalizadas.

Ademais, a aplicação das avaliações no piloto permitiu que, já de partida, o *software* contasse com volume suficiente de resoluções que serviram de ponto de partida para o primeiro treinamento de seus algoritmos de estimativa e recomendação.



O Módulo Sphinx é parte componente da solução *Digital Class* encarregada de possibilitar a manipulação e o preenchimentos de questões e questionários. O módulo é plenamente integrado ao Gryphon e foi adotado para a aplicação das avaliações diagnósticas, de modo que os dados foram disponibilizados de modo transparente e em tempo real.

#### 4.2.3.2 Modelo matemático

Souza (2011) modelou um sistema de recomendação que permitia a recomendação de vídeos a assinantes de um portal. Inspirado em sua modelagem, foi feita a associação entre usuários-estudantes e vídeos-itens. Desta feita, seja o conjunto de estudantes  $E = e_1, e_2, \dots, e_n$ , e o conjunto de itens  $I = i_1, i_2, \dots, i_n$ , onde os itens correspondem aos objetos didáticos disponíveis no acervo e relacionados aos temas da disciplina.  $R$  representa a avaliação do estudante à um objeto recomendado (*rating*). O conjunto  $T_r$  representa o conjunto treinamento, onde:  $T_r \in E \times I \times R$ .

Uma função  $f$  mapeia a relação do estudante ( $e_m \in E$ ) com um item ( $i_n \in I$ ) para uma avaliação  $r_{mn}$ .

$$f : E \times I \rightarrow R$$

$$f(e_m, i_n) = r_{mn}$$

A função  $f$  é estimada por um modelo  $\hat{f}$  que prevê a avaliação para a relação estudante-item.

$$\hat{f}(e_m, i_n) = f(e_m, i_n) + e_{rr}$$

onde  $e_{rr}$  corresponde ao erro entre a predição e a avaliação real.

$$e_{rr} \sim N(\mu, \sigma^2),$$

onde  $e_{rr}$  segue uma distribuição normal, com média  $\mu$  e variância  $\sigma^2$ .

$$e_{rr} = (r - \hat{r}), \text{ onde } \hat{r} \text{ é a predição estimada.}$$

A matriz de avaliações  $R$  é dada por  $R = (r_{mn}), m = 1, \dots, p$  e  $n = 1, \dots, q$ .

$$R : User \times Items \times Contexts \rightarrow Ratings$$

$R$  é a matriz com todas as avaliações. Geralmente tem-se apenas um subconjunto desta matriz em diversos exemplos de sistemas de recomendação. Comumente, sobretudo no contexto de sistemas educacionais, a quantidade de dados para se fazer o treinamento de um modelo é bem pequena. Portanto, comumente as matrizes contendo a avaliação de alunos a objetos educacionais são compostas de poucas observações. Esta característica impõe dificuldades para se fazer um modelo preditivo com baixo erro de predição. Em razão dessa característica de grande esparsidade de dados, optou-se nessa tese pela utilização do algoritmo de filtragem colaborativa desenvolvido por Zhou et al. (2008), *alternating-least-squares with weighted- $\lambda$ -regularization* (ALS-WR). Esta técnica é de simples operação, uma vez que requer o uso de poucos parâmetros. Além disso, com poucas iterações já é possível obter bons resultados, sem riscos de *overfitting*. Outra característica é que enquanto boa parte das técnicas de recomendação, seja as de filtragem colaborativa, seja as de conteúdo, processam a similaridade, respectivamente, entre usuários e entre itens para fazer uma recomendação, o algoritmo ALS revela os fatores latentes existentes quanto a avaliação dos usuários aos itens, e busca encontrar fatores ótimos para minimizar o erro de predição. Em outras palavras, a técnica visa reduzir a função custo associada ao problema de erro médio quadrático (CHAN et al., 2013). Portanto, o ALS-WR busca reduzir o RMSE (*Root Mean Square Error*). RMSE é uma métrica utilizada para medir a qualidade da predição de modelo (quanto menos melhor). Neste trabalho, foram utilizadas o RMSE (SWEENEY et al., 2016) para avaliar a qualidade da previsão.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n err_i^2}$$

A função custo, questão a ser minimizada pela solução, é definida por uma função de cálculo de erro quadrático:

$$\mathcal{L}^2(r, e, i) = (r - \hat{r})$$

Em seguida, é possível definir a função empírica (custo total), como o somatório de todas as avaliações conhecidas.

$$\mathcal{L}^{emp}(R, E, I) = \frac{1}{n} \sum_{(m,k) \in T} \mathcal{L}^2(r_{m,k}, e_m, i_k)$$

onde  $T$  é o conjunto de elementos que contém os elementos avaliados, sendo  $n$  o tamanho deste conjunto. Portanto, o problema de aproximação de matrizes com baixo número de avaliações pode ser formulada a seguir:

$$(E, I) = \operatorname{argmin}_{(E, I)} \mathcal{L}^{emp}(R, E, I)$$

Zhou et al. (2008) em sua implementação apresentou um processo paralelizado. Essa condição melhora o desempenho da técnica, sendo aconselhável para domínios de grande quantidade de dados. Foi adotada a regularização de Tikhonov para eliminar diversos parâmetros livres, que provocam o efeito de *overfitting* no conjunto de dados de teste.

$$\mathcal{L}_\lambda^{reg}(R, E, I) = \mathcal{L}^{emp}(R, E, I) + \lambda(\|E\Gamma_E\|^2 + \|I\Gamma_I\|^2)$$

em que:

$\lambda$ : termo de regularização.

$E\Gamma_E$ : matriz de regularização dos estudantes.

$I\Gamma_I$ : matriz de regularização dos estudantes.

Com o intuito de encontrar a melhor matriz de regularização, os autores chegaram até uma composição que, empiricamente, não apresentou nenhum *overfit* com os dados de teste mesmo aumentando o número de atributos e iterações.

$$f(E, I) = \sum_{(p,j) \in I} (r_{p,j} - e_p^T i_j)^2 + \lambda \left( \sum_p n_{ep} \|e_p\|^2 + \sum_j n_{ij} \|i_j\|^2 \right)$$

em que  $n_{ep}$  e  $n_{ij}$  representam, respectivamente, o número de avaliações do usuário  $p$  e item  $j$ . Isto corresponde à regularização de Tikhonov, onde  $\Gamma_E = \operatorname{diag}(\sqrt{n_{ep}})$  e  $\Gamma_I = \operatorname{diag}(\sqrt{n_{ij}})$ .

A utilização dessa solução foi feita a partir da ferramenta Apache Mahout.

#### 4.2.3.3 Modelo Preditivo

Um dos principais recursos trazidos pelo Modelo *Digital Class* refere-se à disponibilização de modelos preditivos sobre o desempenho acadêmico do aluno.

Para a geração destes modelos são utilizados dados referentes a interação dos usuários com a arquitetura *Digital Class*. A cada acesso do usuário no sistema são registrados dados como tempo de acesso, sequência de atividades desenvolvidas, tempo em cada questão, quantidade de respostas corretas, quantidade de tentativas, quantidade de saltos, etc. Todas estas informações são utilizadas por duas técnicas distintas de predição, que definirão a probabilidade de um aluno vir a ter sucesso ou insucesso para determinado questão em um teste futuro. Este recurso é inovador no contexto de BL, sobretudo ao ser disponibilizado dentro de uma infraestrutura tecnológica única, sem a necessidade de aplicação de softwares estatísticos ou de mineração de dados por parte do professor em momentos diferentes. Para isso, foi realizada a integração à arquitetura das bibliotecas R e Apache Mahout.

Foram gerados dois modelos preditivos. O primeiro, utiliza a técnica de Modelo Linear de Regressão Logística e o segundo com  $SVD^{++}$ .

### Modelo Linear de Regressão Logística

Nos modelos de regressão linear, sejam eles simples ou múltiplos, a variável dependente  $Y$  é uma variável aleatória de natureza contínua. No entanto, em outros contextos a variável dependente pode ser de natureza qualitativa, assumindo dois ou mais categorias ou valores (Figueira, 2006). Neste trabalho, o que se deseja obter com um modelo preditivo é a indicação de qual a probabilidade de sucesso de um estudante frente a uma questão de um determinado tema. Portanto, optou-se pela regressão logística, que permite o uso de um modelo de regressão para a realização do cálculo da probabilidade de ocorrência de um determinado evento.

Um Modelo Linear de Regressão pode ser usado para avaliar se uma variável  $Y$  é influenciada por alguma outra variável, diga-se  $X$ . No contexto deste trabalho,  $Y_i$  é a resposta na questão de um futuro teste do  $i$  – simo aluno, ou seja, é caracterizada por

$$Y_i = \begin{cases} 1, \text{ se } t \text{ (true), ou seja, se o aluno acertar a questão.} \\ 0, \text{ se } f \text{ (false), ou seja, se o aluno errar a questão.} \end{cases}, i = 1, 2, \dots, n.$$

tem-se que a variável resposta é binária, o que sugere o uso do Modelo de Regressão Linear Logística Binária, que pode ser definido por:

$$\ln \frac{\pi_i}{1 - \pi_i} = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}, \text{ para } i = 1, \dots, n.$$

em que

- $\pi_i$  é uma função de ligação e indica a chance de ocorrência do evento de interesse, ou seja, a probabilidade de que a variável  $Y$  assumo o valor 1, indicando que o aluno acertou a questão da avaliação.

$$\pi_i = \frac{e^{(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i})}}{1 + e^{(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i})}}, i = 1, \dots, n.$$

- $X_1$ : representa o valor da variável explicativa Número de Acertos na  $i$ -ésima observação ( $i$ -ésimo aluno).
- $X_2$ : representa o valor da variável explicativa Número de Tentativas na  $i$ -ésima observação ( $i$ -ésimo aluno).
- $\beta_0, \beta_1, \beta_2$ : são os parâmetros do modelo, que serão estimados e que definem a reta de regressão.
- $n$ : o tamanho da amostra (número de alunos).

A discussões mais detalhadas sobre a modelagem e resultados quanto ao uso deste modelo de regressão poderão ser observadas, respectivamente, nos capítulos 5 e 6 desta tese.

### Modelo de Fatoração Matricial Incremental (SVD<sup>++</sup>)

A abordagem de Fatoração de Matrizes tem se notabilizado dentre as técnicas de filtragem colaborativa utilizadas para soluções de sistemas de recomendação (Koren, 2010).

Após vencer o concurso promovido pela Netflix, que premiou o algoritmo que melhorasse em no mínimo 10% o desempenho do algoritmo de recomendação utilizado pela empresa, Brandyn Webb (conhecido como Simon Funk) viu sua técnica ficar famosa. Baseada na técnica SVD (*Singular Value Decomposition*), SVD<sup>++</sup> utiliza o método estocástico denominado de gradiente descendente para minimizar o erro de predição. Enquanto que no SVD tradicional a matriz utilidade que contém as avaliações de usuários aos itens é decomposta em três matrizes menores ( $P = A\Lambda B$ ), o SVD<sup>++</sup> opera sobre duas matrizes características (SOUZA, 2011; THAI-NGHE et al., 2012):

$$\hat{p} = \sum_{k=1}^K w_{uk} h_{ik} = (\mathbf{WH}^T)_{u,i}$$

em que:

- $\mathbf{W} \in \mathbb{R}^{U \times K}$  corresponde a matriz onde cada linha  $u$  é um vetor contendo os  $K$  fatores latentes que caracterizam o estudante;
- $\mathbf{H} \in \mathbb{R}^{I \times K}$  corresponde a matriz onde cada linha  $i$  é um vetor contendo os  $K$  fatores latentes que descrevem uma atividade  $i$ ;

A seguir será apresentado a fundamentação científica assim como a notação usada para o cálculo da previsão de desempenho a partir do uso de sistemas de recomendação e técnicas de filtragem colaborativa agregadas à fatoração de matrizes.

Inicialmente definem-se os objetos que farão parte do sistema de recomendação.

Sendo:

- $E$  o conjunto de estudantes.
- $C$  o conjunto de conteúdos e
- $P \subseteq R$  o conjunto de possíveis performances de desempenhos (notas) dos estudantes em cada conteúdo.

Dessa forma, pode-se interpretar que a população de interesse é o conjunto completo de dados que descrevem as performances de desempenho de cada um dos  $n$  estudantes em todos os  $m$  conteúdos. À título de ilustração, segue a matriz  $\mathbf{X}$  como sendo a matriz que contém a informação de interesse completa.

$$\mathbf{X} = \begin{bmatrix} p_{e_1c_1} & p_{e_1c_2} & \cdots & p_{e_1c_m} \\ p_{e_2c_1} & p_{e_2c_2} & \cdots & p_{e_2c_m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{e_nc_1} & p_{e_nc_2} & \cdots & p_{e_nc_m} \end{bmatrix}$$

Definem-se também os seguintes objetos:

- $M_E$ : O conjunto de dados latentes que descrevem todos os estudantes, sendo  $m_E : E \rightarrow M_E$  o vetor dados latentes para cada estudante.
- $M_C$ : O conjunto de dados latentes que descrevem todos os conteúdos, sendo  $m_C : C \rightarrow M_C$  o vetor de dados latentes para cada conteúdo.

A matriz  $\mathbf{X}$  contendo as performances de desempenho de todos os estudantes em todos os conteúdos não é, comumente, completamente observada, ou seja, na prática não é possível obter a performance de desempenho de todos os alunos em todos os conteúdos. Isso acontece por diversos fatores como, por exemplo, o aluno usar o sistema virtual de aprendizagem

apenas em alguns conteúdos, ou mesmo, nem todos os alunos se dedicarem a fazer uso do sistema.

Dessa forma, tem-se no uso das técnicas de fatoração de matrizes uma forma de aproximar tal matriz  $\mathbf{X}$  pelo produto de duas outras matrizes  $\mathbf{W}$  e  $\mathbf{H}$  tal que

$$\mathbf{X} \approx \mathbf{W}\mathbf{H}^T.$$

em que:

$\mathbf{W} \in \mathbb{R}^{(U \times K)}$ : É uma matriz onde cada linha  $u$  é um vetor contendo  $K$  fatores latentes descrevendo o estudante  $u$  (usuário  $u$ , na linguagem dos sistemas de recomendação).

$$\mathbf{W} = \begin{bmatrix} w_{u_1,1} & w_{u_1,2} & \dots & w_{u_1,k} \\ w_{u_2,1} & w_{u_2,2} & \dots & w_{u_2,k} \\ \vdots & \vdots & \ddots & \vdots \\ w_{u_n,1} & w_{u_n,1} & \dots & w_{u_n,k} \end{bmatrix}$$

$\mathbf{H} \in \mathbb{R}^{I \times K}$ : É uma matriz onde cada linha  $i$  é um vetor contendo  $K$  fatores latentes descrevendo o conteúdo  $i$  (item  $i$ , na linguagem dos sistemas de recomendação).

$$\mathbf{H} = \begin{bmatrix} h_{i_1,1} & h_{i_1,2} & \dots & h_{i_1,k} \\ h_{i_2,1} & h_{i_2,2} & \dots & h_{i_2,k} \\ \vdots & \vdots & \ddots & \vdots \\ h_{i_n,1} & h_{i_n,1} & \dots & h_{i_n,k} \end{bmatrix}$$

Pode-se interpretar que a matriz  $\mathbf{W}$  contém, de forma resumida, em  $k$  fatores latentes a informação contida no conjunto de dados latentes sobre todos os estudantes  $M_E$ . De forma semelhante, pode-se interpretar que a matriz  $\mathbf{H}$  contém, de forma resumida, em  $k$  fatores latentes a informação contida no conjunto de dados latentes sobre todos os conteúdos  $M_C$ , ambos definidos anteriormente.

### Modelo Básico de Fatoração de Matrizes

O modelo básico de previsão de performance de desempenho escolar  $\hat{p}$  é definido da seguinte forma: Seja  $w_{uk}$  e  $h_{uk}$  os elementos de  $W$  e  $H$ , respectivamente. Então, a performance  $\hat{p}$  de determinado estudante  $u$  em um conteúdo  $i$  é dada por

$$\hat{p}_{ui} = \sum_{k=1}^K w_{uk} h_{ik} = \mathbf{W} \mathbf{H}_{u,i}^T$$

### Modelo com Fatoração de Matrizes com Vieses (*Biased Matrix Factorization*)

Ao modelo básico de fatoração de matrizes, podem ser incluídos o viés causado pela variação de cada estudante (*user effect*) e o viés causado pela variação de cada conteúdo abordado (*item effect*). Isso ajuda a modelar o quão bom um estudante é, a chance dele responder um conteúdo corretamente, e a modelar a dificuldade de cada tema no sentido de medir a chance de uma questão/conteúdo ser respondida corretamente.

Primeiramente, dado que a matriz  $\mathbf{X}$  contendo todas as performances de desempenhos dos estudantes versus conteúdos não é completamente observada, faz-se relevante a definição dos seguintes subgrupos. Sendo eles:

- $D^{\text{treino}} \subseteq (E \times C \times P)$ : É uma sequência de desempenhos escolares de estudantes por conteúdos observados (nem todos os estudantes, ou mesmo, nem todos os conteúdos, são observados na prática).
- $D^{\text{validação}} \subseteq (E \times C \times P)$ : É uma sequência, disjunta de  $D^{\text{treino}}$ , de desempenhos escolares de estudantes por conteúdos observados que será usada para validar as estimativas feitas a partir de  $D^{\text{treino}}$ .
- $D^{\text{teste}} \subseteq (E \times C \times P)$ : É a sequência a ser estimada, contendo os desempenhos escolares de estudantes por conteúdos não observados.

Assim, a nova função de previsão de desempenho do aluno  $u$  no conteúdo  $i$  fica definida pela seguinte equação

$$\hat{p}_{ui} = \mu + b_u + b_i + \sum_{k=1}^K w_{uk} h_{ik}$$

em que :

- $\mu$ : Representa a média global dos desempenhos observados.
- $b_u = \frac{\sum_i p_{ui} - \mu}{|D_u^{\text{treino}}|}$ : Representa a variação da média dos desempenhos observados em todos os conteúdos do estudante  $u$  em relação à media global  $\mu$ .



- $b_i = \frac{\sum p_{ui} - \mu}{\mathbb{D}^{treino}_i}$  : Representa a variação da média dos desempenhos observados em todos os alunos no conteúdo  $i$  em relação à média global  $\mu$ .
- $\sum_{k=1}^K w_{uk} h_{ik}$ : Representa a contribuição da informação contida nos  $K$  fatores latentes atribuídos ao estudante  $u$  e ao conteúdo  $i$ .

## Fase de Treinamento: Implementando o Modelo com Vieses

### Definições iniciais

A fase de treinamento resume-se, principalmente, em encontrar os melhores parâmetros para compor  $\mathbf{W}$  e  $\mathbf{H}$ . Inicialmente, a previsão de desempenho

$$p_{ui} = \mu + b_u + b_i + \sum_{k=1}^K w_{uk} h_{ik}$$

é estimada com base nos dados observados em  $\mathbb{D}^{treino}$  e com as matrizes de fatores latentes  $\mathbf{W}$  e  $\mathbf{H}$  iniciadas com valores aleatórios simulados a partir de uma distribuição Normal com média zero e variância 0.001,  $N(0, 0, 01)$ .

A partir da estimativa calculada, obtém-se o erro de estimação, calculado de acordo com a equação abaixo

$$erro = \sum_{(u,i) \in \mathbb{D}^{treino}} e_{u,i}^2 = \sum_{u,i \in \mathbb{D}^{treino}} (p_{ui} - \hat{p}_{ui})^2 = \sum_{(u,i) \in \mathbb{D}^{treino}} (p_{ui} - \sum_{k=1}^k w_{uk} h_{ik})^2$$

Tal erro de estimação pode ser minimizado através da atualização dos valores de  $\mathbf{W}$  e  $\mathbf{H}$  iterativamente através do Método do Gradiente Descendente.

### Método do Gradiente Descendente

O método do gradiente descendente é um algoritmo de otimização iterativa usado para encontrar o valor mínimo de uma função custo. Basicamente, tem-se que começando com um valor inicial  $x_0$ , o ponto de mínimo da função  $F$  é encontrado através de sucessivas atualizações calculadas através da equação abaixo:

$$x_{n+1} = x_n - \beta \nabla F(X_n).$$

em que

- $x_{n+1}$  : Valor atual.
- $x_n$  : Valor anterior.
- $\beta$  : Taxa de aprendizagem do algoritmo (o tamanho de cada passo em cada iteração).
- $\nabla F(x_n)$ : Valor do gradiente da função  $F$ .
- $-\beta\nabla F(x_n)$ : Termo que indica a direção decrescente do algoritmo no sentido de encontrar o mínimo (usando o sinal positivo, encontra-se o máximo).

Partindo da função erro definida anteriormente, calcula-se o valor da função gradiente em cada um dos pontos  $w_{uk}$  e  $h_{ik}$

$$\nabla F(w_{uk}) = \frac{\partial e_{ui}^2}{\partial w_{uk}} = -2e_{ui}h_{ik} = -2(p_{ui} - \hat{p}_{ui})h_{ik}$$

$$\nabla F(h_{ik}) = \frac{\partial e_{ui}^2}{\partial h_{ik}} = -2e_{ui}w_{uk} = -2(p_{ui} - \hat{p}_{ui})w_{uk}$$

Após definir os gradientes, os valores iniciais de  $w_{uk}$  e  $h_{ik}$  são atualizados na direção oposta (negativa) do gradiente para que os valores de mínimo sejam alcançados. Assim, em cada iteração do algoritmo, os valores, inicialmente gerados para  $\mathbf{W}$  e  $\mathbf{H}$  serão atualizados de acordo com a seguinte equação

$$w'_{uk} = w_{uk} - \beta\nabla F(w_{uk}) = w_{uk} - \beta(-2e_{ui}h_{ik}) = w_{uk} + 2\beta e_{ui}h_{ik}$$

$$h'_{ik} = h_{ik} - \beta\nabla F(h_{ik}) = h_{ik} - \beta(-2e_{ui}w_{uk}) = h_{ik} + 2\beta e_{ui}w_{uk}$$

### Termo de Regularização

Para prevenir a ocorrência de sobre-ajuste (*overfitting*) das previsões de desempenho, Thai-Nghe et al.(2012) mostram que a função erro pode ser modificada com a adição de um termo que controle a magnitude dos vetores de fatores latentes de forma que  $\mathbf{W}$  e  $\mathbf{H}$  sejam uma boa aproximação para  $\mathbf{X}$  sem precisarem conter grandes números. Assim, a função torna-se:

$$erro = \sum_{(u,i) \in \mathbb{D}^{treino}} e_{u,i}^2 = \sum_{(u,i) \in \mathbb{D}^{treino}} (p_{ui} - \hat{p}_{ui})^2 + \lambda(\|\mathbf{W}\|^2 + \|\mathbf{H}\|^2)$$

em que  $\lambda$  é o termo de regularização.

Dessa forma, os valores de  $w$  e  $h$  são atualizados de acordo com as seguintes equações:

$$w'_{uk} = w_{uk} + \beta(2e_{ui}h_{ik} - \lambda w_{uk})$$

$$h'_{ik} = h_{ik} + \beta(2e_{ui}w_{uk} - \lambda h_{ik})$$

Thai-Nghe et al (2012) lembram que a fatoração de matrizes se mostra como uma técnica adequada para a predição de desempenho de estudantes, principalmente, por capturar a variação provocada pelo “chute” e “erro” (*guess* e *slip*, respectivamente). Pois os fatores que definem a chance de um estudante “chutar uma resposta” e acertar (chute) ou errar uma resposta sabendo o conteúdo (erro) estão inclusos nas matrizes de fatores latentes  $W$  e  $H$ .

### Escolha do Modelo Final Estimado

O objetivo principal da predição de desempenho do estudante é: Dado os dados observados  $\mathbb{D}^{treino}$ , o pares de estudantes-conteúdos dos dados não observados  $(e, c)_{(\mathbb{D}^{teste})}$ , o conjunto de dados latentes descrevendo cada estudante  $W$  e o conjunto de dados latentes descrevendo cada conteúdo  $H$ , encontrar

$$p = \hat{p}_1, \hat{p}_2, \dots, \hat{p}_{|\mathbb{D}^{teste}|}$$

tal que:

$$erro(p, \hat{p}) = \sum_{i=1}^{|\mathbb{D}^{validação}|} (p_i - \hat{p}_i)^2$$

seja mínimo com  $p := \pi_p(\mathbb{D}^{validação})$ . Atentando que  $\mathbb{D}^{validação}$  é o tamanho  $n$  de observações contidas na sequência de desempenhos escolares de estudantes por conteúdos em  $\mathbb{D}^{validação}$ . O principal objetivo da técnica  $SVD^{++}$  é encontrar os fatores ótimos para  $W$  e  $H$ , de forma a minimizar o erro médio quadrático, que calcula o erro quadrático entre a estimativa do modelo e a avaliação de fato.

Assim, o modelo final definido será o modelo que contenha os melhores parâmetros de acordo com algum critério. No presente estudo, o critério usado foi o erro médio quadrático RMSE (*Root Mean Squared Error*).

$$RMSE = \sqrt{\frac{\sum_{ui \in \mathbb{D} \text{validação}} (p_{ui} - \hat{p}_{ui})^2}{|\mathbb{D} \text{validação}|}},$$

entende-se como parâmetros as variáveis:

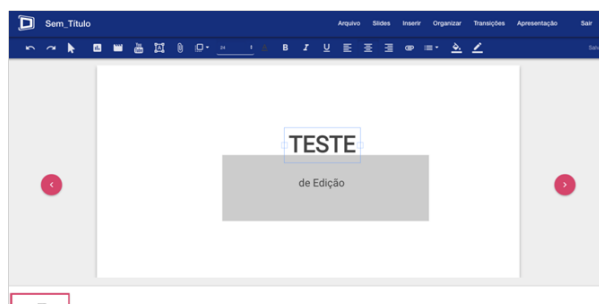
- *Iter*: O número de iterações do algoritmo
- *K*: O número de variáveis latentes usadas para compor *W* e *H*.
- $\beta$ : O valor atribuído à taxa de aprendizagem do algoritmo
- $\lambda$ : O valor atribuído ao termo de regularização do modelo.

Tais parâmetros influenciam fortemente a atualização dos valores iniciais atribuídos a *W* e *H*, o que por fim, definem a previsão das performances de desempenho.

#### 4.2.4 Chimera

O módulo Chimera permite que professores e alunos possam elaborar materiais didáticos autorais. Esse módulo permite a criação de notas de aula, trabalhos, seminários, aulas, etc. É possível criar do início todo um material, com um editor HTML 5, assim como permite a importação de materiais já existentes (documentos, imagens, vídeos, animações, etc.). Além disso, é possível realizar uma explanação sincronizada com o material didático. Assim, um professor poderá gravar suas aulas previamente e disponibilizá-las a seus cursistas. Este exemplo, demonstra a importância deste recurso à todos aqueles preocupados em implantar sala de aula invertida (*flipped classroom*). Os objetos são todos indexados através de *tags* livremente definidas pelo autor, de forma a facilitar a criação e recuperação de acervo.

Figura 36 – Ferramenta de Autoria de Slides (HTML 5)

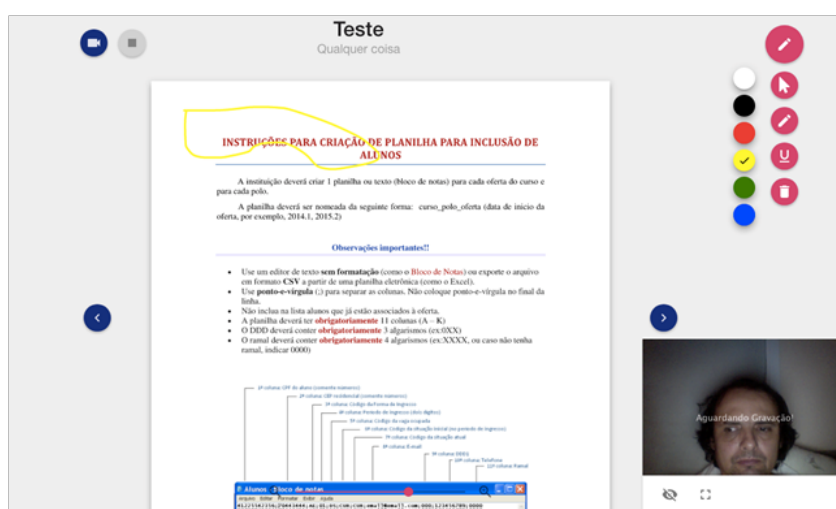


Fonte – Pesquisa do trabalho

Neste sistema, um autor tem a opção de deixar o material criado de caráter restrito ou de livre acesso. Caso opte pela segunda opção, este usuário poderá, também, ter acesso a materiais de outros interagentes. Com esta perspectiva, o sistema propicia a criação de um amplo acervo colaborativo, podendo a ferramenta tornar-se em muito pouco tempo, uma rica plataforma com diversificado repositório multimídia. Outra facilidade oferecida pelo sistema diz respeito à facilidade de atividades didáticas criadas serem compartilhadas com usuários que não tem acesso ao sistema, através de compartilhamento de *link* de acesso.

Uma aula criada na ferramenta poderá ser compartilhada a qualquer tempo que o professor desejar com seu grupo de alunos. A criação da aula pode ser feito de sua casa, gabinete, de sua sala de aula, etc. A plataforma pode ser associada à alguma solução de lousa digital. Neste caso, o professor poderá realizar uma aula presencial e ao mesmo tempo compor seu acervo. A ferramenta possui uma versão *online* e outra *offline*.

Figura 37 – Gravação de uma aula no Chimera.



Fonte – Pesquisa do trabalho

### 4.2.5 *Sphinx*

O módulo Sphinx consiste em um sistema online que aplica instrumentos avaliativos, propiciando um registro minucioso de informações sobre o comportamento do cursista durante a resolução dos itens de um instrumento. Essa ferramenta possui a versatilidade para se constituir como um simulado, um questionário investigativo, como listas de exercícios, etc. O Sistema também constitui-se como um repositório de itens. Ele foi desenvolvido para ser um sistema colaborativo para a construção e compartilhamento de itens e aplicação de instrumentos de avaliação. Na Arquitetura *Digital Class*, seu papel é destacado em alimentar o módulo Gryphon de um acervo indexado às temáticas de um curso. É opção do professor utilizar a matriz de habilidades e competências do Exame Nacional do Ensino Médio (ENEM). No cadastro, o professor pode indexar apenas pelo conteúdo da questão. Para a realização do experimento, foi solicitado do professor que indicasse o nível de complexidade da questão (básico, intermediário e avançado).

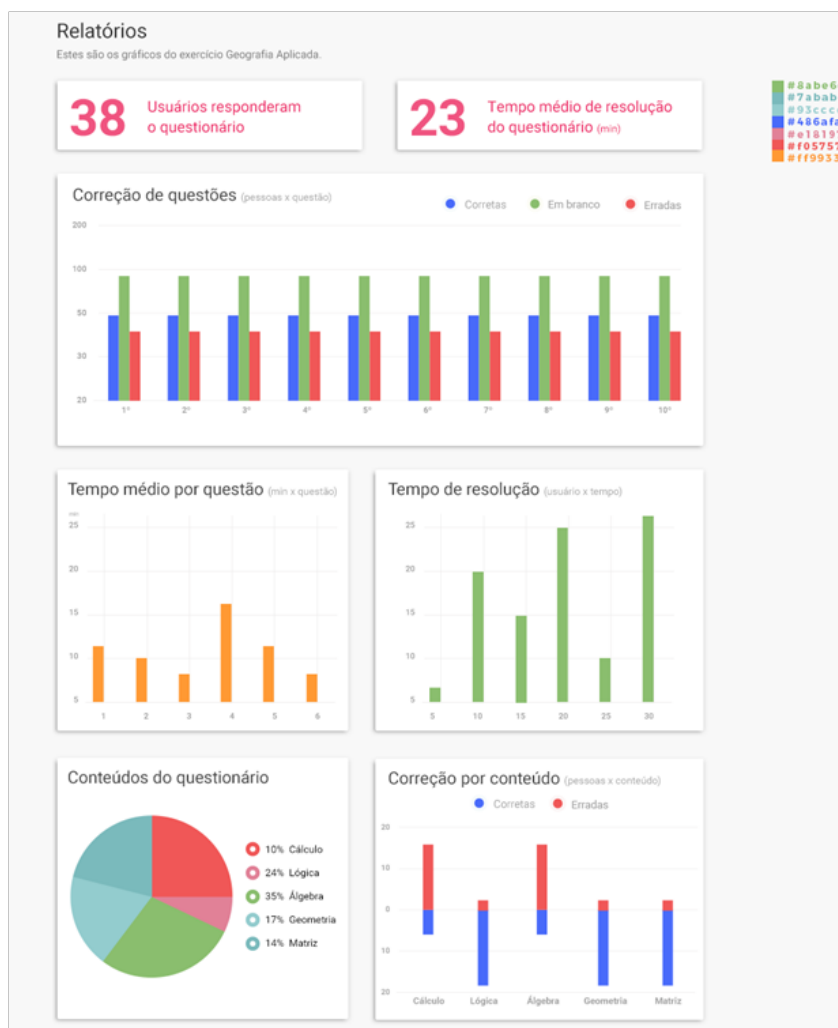
Figura 38 – Aplicação de um teste *online*.



Fonte – Pesquisa do trabalho

Esta ferramenta também possui uma estrutura própria de relatórios para que o professor possa visualizar o resultado da aplicação de seus instrumentos.

Figura 39 – Relatório pormenorizado após a aplicação de um teste online.



Fonte – Pesquisa do trabalho

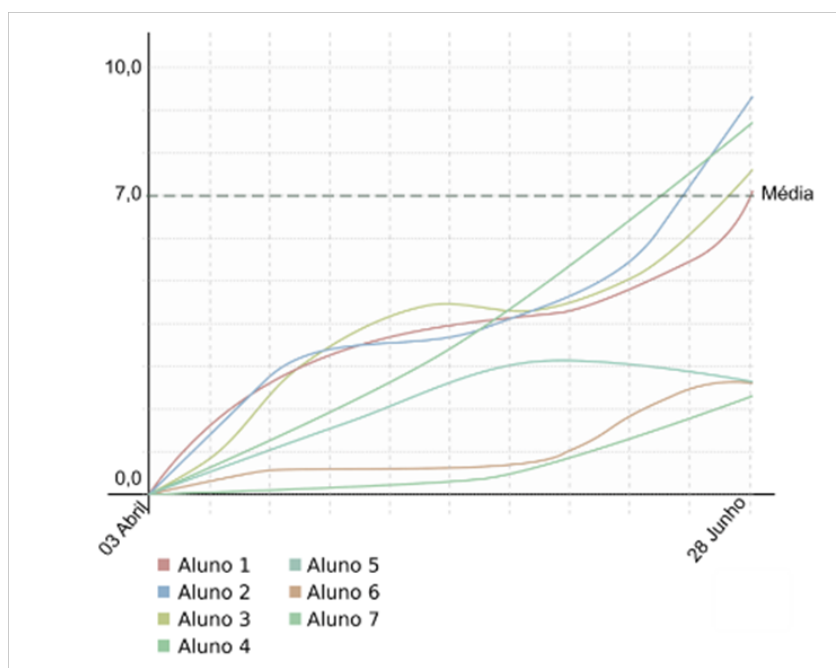
#### 4.2.6 Pegasus

Para o devido acompanhamento dos eventos existentes em um curso, seja por alunos, professores e gestores, a Arquitetura *Digital Class* disponibiliza um módulo de *dashboards* para a visualização facilitada de dados. A obtenção de informações precisas é de fundamental importância para todos os partícipes do processo de aprendizagem, dos alunos aos professores e gestores. Portanto, para a consolidação de dados oriundos da interação dos usuários com todos os módulos da Arquitetura *Digital Class*, é fundamental que a arquitetura consolide informações provenientes de seus diversos elementos de modo simplificado e transparente. O módulo **Pegasus** propõe-se a satisfazer essas necessidades de modo integrado e transparente para o usuário.

Pegasus é, na verdade, um módulo de *Business Intelligence*, implementando um ETL projetado para agregar dados provenientes dos diferentes módulos da Arquitetura *Digital Class* em um *data warehouse*. A base é disponibilizada para as outras ferramentas da solução por meio de uma *API Rest*. Esse comportamento dá ao software que consome as informações provenientes do Pegasus total liberdade sobre questões de exibição, navegação e manipulação de seus gráficos e relatórios. A ferramenta foi desenvolvida utilizando a tecnologia Node.js, utilizando PostgreSQL para o armazenamento do *data warehouse* bem como das bases intermediárias do ETL.

A Figura 40 apresenta um exemplo de gráfico construído a partir de informações fornecidas pela API do Pegasus. O diagrama apresenta a evolução do desempenho de alunos selecionados em um tema específico ao longo de um período de tempo. O gráfico explora a tabela fato performance do *data warehouse*, filtrada nas dimensões usuário, tema e tempo.

Figura 40 – Gráfico de Evolução do Desempenho Discente no tempo.



Fonte – Pesquisa do trabalho



## 5 METODOLOGIA DA VALIDAÇÃO DO EXPERIMENTO

Este capítulo tem como objetivo apresentar como foi estruturado o processo de validação do Modelo Conceitual *Digital Class*. Esta validação foi realizada em duas etapas. A primeira, diz respeito a implementação da Arquitetura *Digital Class*. A segunda, refere-se à realização do experimento em que a arquitetura *Digital Class* foi utilizada em uma disciplina real.

### 5.1 Primeira Etapa da Validação

#### 5.1.1 Construção da arquitetura

A construção da arquitetura foi orientada ao cumprimento das premissas teóricas definidas pelo Modelo Conceitual *Digital Class*.

Foi definido a criação de quatro módulos WEB, que promovem uma facilidade para que alunos e professores não tenham que se preocupar em fazer a instalação de nenhum software. Ao invés de uma estrutura única, a arquitetura foi dividida em módulos que, por sua vez, desempenham funções distintas, porém complementares. Em um cenário evolutivo, por exemplo, qualquer módulo pode ser desconectado da arquitetura e um outro assumir o posto sem maiores dificuldades. Um exemplo disso, é que se um módulo de autoria de material didático for mais atrativo do que o atual Chimera, este poderá dar espaço sem que sejam reimplementados todos os módulos restantes. Outro efeito positivo desta escolha é para o caso da existência de sistemas legados. Muitas instituições, por exemplo, já operaram plataformas próprias de avaliação. Desta forma, o módulo Sphinx seria substituído por um sistema legado. No entanto, a adoção de módulos que não os oferecidos pela arquitetura demandariam um mínimo esforço de integração, uma vez que a integração utiliza *webservices*, facilitando assim a integração entre os módulos.

O núcleo da arquitetura, porém, corresponde à “alma” da proposta e é um módulo obrigatório para a solução *Digital Class*. Sua natureza inovadora, composta, por exemplo, pelos mecanismos de recomendação e predição de desempenho provê no presente uma experiência única de apoio a cursos na modalidade de ensino híbrido. Os módulos são integrados através de chamadas a *webservices* que promovem a comunicação sistêmica necessária ao funcionamento pleno dos componentes da arquitetura.

A caracterização dos módulos componentes da arquitetura foi apresentada detalhadamente no Capítulo 4.4. Porém, vale lembrar que o Modelo conceitual definido neste trabalho é composto por seis dimensões. Como discutido no capítulo anterior, essas dimensões foram representadas através dos módulos que compõem a Arquitetura Digital Class e, portanto, a validação das dimensões deu-se a partir da análise do comportamento da arquitetura quanto a sua utilização por alunos e professor. As Tabelas 32 e 33, disponíveis no Apêndice A, descrevem como cada dimensão pode ser observada na arquitetura.

Uma das dimensões do Modelo, a Predição de Desempenho Estudantil, demandou a construção de dois modelos preditivos, modelados segundo duas técnicas diferentes (Modelo Linear de Regressão Logística e  $SVD^{++}$ ). Desta forma, será descrita a seguir a modelagem do problema e como foram desenvolvidos esses dois modelos.

### 5.1.2 Métodos Utilizados para a Previsão de Desempenho dos Estudantes

Um dos principais recursos trazidos pelo Modelo e que é implementado pela Arquitetura Digital Class é a predição de desempenho do estudante. Nesse estudo foram utilizadas duas técnicas para se obter esse resultado. A primeira, Regressão Logística Linear e a segunda  $SVD^{++}$ .

Foram utilizados dados que registraram a interação dos estudantes com o sistema. Inicialmente foram coletados os dados de acesso e uso do sistema de ensino virtual de 71 estudantes, do início do semestre letivo 2017.1 até a segunda avaliação bimestral (final do semestre). Ao todo foram abordados 10 temas, que para fins de praticidade na análise, foram codificados na sequência de 1 a 10.

Tabela 2 – Conteúdos abordados.

Notação	Conteúdo
1	Notação Científica.
2	Metodologia de Investigação: A procura de regularidades e de sinais na interpretação física do mundo.
3	Leis de Newton
4	Movimento retilíneo uniformemente variado.
5	Movimento retilíneo uniforme.
6	Movimentos verticais: Queda livre.
7	Movimentos verticais: Lançamento vertical.
8	Movimentos verticais: Horizontal e oblíquo.
9	Equação de Torricelli.
10	O que é física.

Os estudantes foram codificados de acordo com o ID usado pelo sistema de ensino, no anexo segue a tabela contendo todos os IDs e nomes dos estudantes que participaram do estudo.

Inicialmente, é necessário calcular uma performance de desempenho com base nos dados observados que o sistema de ensino fornece. Para cada estudante em um determinado conteúdo  $i$ , o sistema de ensino fornece três informações básicas:

- $q$ : O número de questões resolvidas no conteúdo  $i$ .
- $t$ : Quantidade de tentativas feitas para resolver as  $n$  questões no conteúdo  $i$ .
- $a$ : Quantidade de acertos dentro das  $n$  questões resolvidas desse mesmo conteúdo  $i$ .

Segue a Tabela 3 temos o exemplo dos estudantes  $a$ ,  $b$  e  $c$  no *Conteúdo 2*:

Tabela 3 – Relação entre estudante  $\times$  tentativa  $\times$  acerto de questões

IDestudante	quesT2	tentC2	acertosC2
a	5	5	5
b	5	9	3
c	2	2	1

A Tabela 3 indica que o estudante  $a$  resolveu 5 questões com 5 tentativas e acertou todas as 5 questões, o que indica que ele acertou cada uma das questões na primeira tentativa. Já o estudante  $b$  precisou de 9 tentativas para resolver o mesmo numero de questões que o estudante  $a$  e, mesmo assim, só acertou 3 questões das 5 totais. Isso indica que o nível de proficiência do estudante  $b$  é menor que o do estudante  $a$ , no conteúdo em questão. Porém, os aluno  $a$  e  $b$  indicam ter maior nível de proficiência no conteúdo 2 do que estudante  $c$ , que resolveu apenas 2 questões.

Dessa forma, para cada aluno, a performance de desempenho em cada conteúdo foi calculada da seguinte forma

$$p = \left( \frac{a}{t} + \frac{q}{\mu_q} \right)$$

em que  $\mu_q$  é a média geral de questões resolvidas pelos estudantes.

No intuito de padronizar a performance de desempenho de cada estudante dentro do intervalo  $(0, 1)$ , o valor  $p$  definido acima foi dividido pelo valor  $p$  máximo encontrado dentro dos dados observados  $\mathbb{D}^{treino}$ .

Voltando ao exemplo definido acima, supondo que a média geral  $\mu_q$  tenha sido de 6 questões por aluno, as performances de desempenhos dos alunos  $a$ ,  $b$  e  $c$  no conteúdo 2 ficariam:

Tabela 4 – Relação entre estudante x tentativa x acerto de questões x nota Geral x Nota Padronizada

IDestudante	quesT2	tentC2	acertosC2	notaGeral	notaPadr
a	5	5	5	1,83	1
b	5	9	3	1,17	0,64
c	2	2	1	0,83	0,45

Assim sendo, segue a Tabela 5 contendo a matriz  $X$  de performances observadas e não observadas:

Tabela 5 – Matriz esparsa com as 10 primeiras observações das performances observadas.

ID estud.	nota <sub>t2</sub>	nota <sub>t4</sub>	nota <sub>t15</sub>	nota <sub>t85</sub>	nota <sub>t86</sub>	nota <sub>t87</sub>	nota <sub>t88</sub>	nota <sub>t89</sub>	nota <sub>t90</sub>	nota <sub>t91</sub>
6	0,15	0,1					0,31	0,05	0,11	
11	0,38			0,15	0,54		0,05			
12	0,37	0,36		0,49	0,72		0,23			
13	0,36									
14				0,22	0,35		0,15	0,21		
15	0,48			0,05			0,05			
16	0,42									
17	0,4									
19	0,27									
24	0,44									
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Tem-se que 65,77% dos dados observados são desconhecidos.

## A Técnica SVD<sup>++</sup>

Para a construção do modelo de previsão, a implementação do algoritmo foi feita com os dados observados e importados em forma de lista. Para o conjunto de treinamento foram usadas 196 observações (performances observadas). Do total de 75 estudantes considerados no estudo, foi verificado que, para cada tema, há no mínimo 2 observações de desempenho e no máximo 62 observações.

Tabela 6 – Número de performances observadas por cada conteúdo

<b>Temas</b>	<b>c1</b>	<b>c2</b>	<b>c3</b>	<b>c4</b>	<b>c5</b>	<b>c6</b>	<b>c7</b>	<b>c8</b>	<b>c9</b>	<b>c10</b>
Nº de obs	62	24	15	28	32	1	26	1	5	2

A Tabela 7 exibe algumas tuplas do conjunto de dados de treinamento,  $\mathbb{D}^{treino}$ .

Tabela 7 – Amostra de tuplas dos dados do conjunto de treinamento  $\mathbb{D}^{treino}$ 

<b>Observação (tupla)</b>	<b>IDestudante</b>	<b>IDconteudo</b>	<b>notaPad</b>
1	13	1	0,36184211
2	16	1	0,41596639
3	17	1	0,39731682
4	19	1	0,27244582
5	24	1	0,43785935
6	27	1	0,30650155
⋮	⋮	⋮	⋮
196	94	10	0,48529412

De forma semelhante, 54 observações foram designadas para compor o conjunto de validação. A Tabela 8 exibe algumas observações do conjunto  $\mathbb{D}^{validação}$ .

Tabela 8 – Amostra de tuplas dos dados do conjunto de validação  $\mathbb{D}^{validação}$ 

<b>Observação (tupla)</b>	<b>IDestudante</b>	<b>IDconteudo</b>	<b>notaPad</b>
1	6	1	0,15
2	11	1	0,38
3	12	1	0,37
4	15	1	0,48
5	31	1	0,05
6	34	1	0,22
⋮	⋮	⋮	⋮
54	91	10	0,34

A implementação do algoritmo de previsão de desempenho escolar foi feita no *Software* Livre R, versão 3.4.0. O algoritmo segue a sequência abaixo:

- Linha 01 à linha 30: São definidos os parâmetros do modelo com valores iniciais calculados à partir dos dados observados em  $\mathbb{D}^{treino}$ .
- Linha 32 à linha 47: os valores iniciais dos parâmetros são atualizados, de forma que em cada iteração, tem-se:
  - Linha 34: Sorteia-se aleatoriamente uma tupla de  $\mathbb{D}^{treino}$ . Dessa tupla extrai-se os valores que preencheram as linhas 35,36 e 37.
  - Linha 35: Define-se aleatoriamente um estudante  $u$ .
  - Linha 36: Define-se aleatoriamente um conteúdo  $i$ .
  - Linha 37: Define-se aleatoriamente um performance de desempenho observada  $p_{ui}$ .
  - Linha 38: Estima-se uma performance  $\hat{p}_{ui}$  com base na média geral  $\mu$ , o viés do aluno  $u$  e do conteúdo  $i$ , definidos nas linhas 35 e 36 e do produto interno da  $u$ -ésima linha de  $W$  com a  $i$ -ésima linha de  $H$ . Tal que  $\hat{p}_{ui} = \mu + b_u + b_i + \sum_{k=1}^K w_{uk}h_{ik}$
  - Linha 39: Calcula o valor o erro de estimação entre  $p_{ui}$  e  $\hat{p}_{ui}$ .
  - Linha 40: Atualiza o valor da média geral  $\mu$ .
  - Linha 41: Atualiza o valor do viés de variação do estudante  $u$  sorteado.
  - Linha 42: Atualiza o valor do viés de variação do conteúdo  $i$  sorteado
  - Linhas 43-46: Atualiza o valor dos elementos  $w_{ui}$  e  $h_{ui}$  de  $W$  e  $H$ , respectivamente.
  - Linhas 48-52: Após o processo acima repetir-se  $iter$  vezes até a ocorrência da convergência do valor do erro de estimação, o algoritmo fornece os valores dos parâmetros ótimos para a previsão das performances de desempenho.

A Figura 41 apresenta o código do algoritmo SVD<sup>++</sup> escrito em linguagem R.

Figura 41 – Algoritmo SVD<sup>++</sup> utilizado para modelar a sistemática de predição.

```

1 Dtrain <- data.frame(d[,c(1,2,8)]) # sequencia de performances observadas
2 iter <- 20000 # numero de iterações
3 k <- 100 # numero de variáveis latentes
4 B <- 0.001 # taxa de aprendizado
5 lambda <- 0.02 # termo de regularização
6
7 mu <- mean(Dtrain[,3]) # média geral dos desempenhos observados
8
9 bu <- numeric(0) # definindo a variável vies do estudante u
10 usu <- names(table(Dtrain$IDaluna)) # definindo a variável estudantes
11 nusu <- length(usu); nusu # definindo a variável total de estudantes
12 for(i in 1:nusu){
13   bu[i] <- sum( Dtrain[which(Dtrain$IDaluna==usu[i]),3] - mu )/nusu
14 }
15 bu <- data.frame(bu); rownames(bu) <- usu # vetor com os vieses de cada aluno
16
17 bi <- numeric(0) # definindo a variável vies do conteúdo i
18 temma <- names(table(Dtrain$tema)); temma # definindo a variável conteúdos
19 ntemma <- length(temma); ntemma # definindo a variável total de conteúdos
20 for(i in 1:nTemma){
21   bi[i] <- sum(Dtrain[which(Dtrain$tema==temma[i]),3] - mu)/ntemma
22 }
23 bi
24 # vetor com os vieses de cada conteúdo
25
26 W <- matrix(data=rnorm(nusu*k,0,0.01),nco=k); rownames(W) <- usu
27 H <- matrix(data=rnorm(ntemma*k,0,0.01),nco=k)
28
29 p_ui <- numeric(0) # definindo a variável performance observada
30 p_ui_est <- numeric(0) # definindo a variável performance estimada
31 e_ui <- numeric(0) # definindo a variável erro
32 for(l in 1:(iter)){
33
34   sample <- Dtrain[ sample(1:dim(Dtrain)[1],1) ,]
35   u <- sample[ , 1]
36   i <- sample[ , 2]
37   p_ui <- as.numeric( Dtrain[ which( Dtrain$IDaluna==u & Dtrain$tema==i) , 3] )
38   p_ui_est <- mu + bu[which(rownames(bu)==u),] + bi[i] + w[which(rownames(bu)==u),]%%H[i,]
39   e_ui[l] <- (p_ui - p_ui_est); e_ui[l]
40   mu <- mu + (B*e_ui[l])
41   bu[which(rownames(bu)==u),] <- bu[which(rownames(bu)==u),] + B*(e_ui[l] - lambda*bu[which(rownames(bu)==u),])
42   bi[i] <- bi[i] + B*(e_ui[l] - lambda*bi[i])
43   for(j in 1:k){
44     w[which(rownames(bu)==u),j] <- w[which(rownames(bu)==u),j] + B*(e_ui[l]*H[i,j] - lambda*w[which(rownames(bu)==u),j] )
45     H[i,j] <- H[i,j] + B*(e_ui[l]*w[which(rownames(bu)==u),j] - lambda*H[i,j] )
46   }
47 }
48 mu # media geral final
49 bu # vetor final com vieses de cada estudante
50 bi # vetor final com vieses de cada conteúdo
51 W # matriz final de fatores latentes dos estudantes
52 H # matriz final de fatores latentes dos conteúdos

```

Fonte – Pesquisa do trabalho

## 5.2 Segunda Etapa da Validação

### 5.2.1 Processo de Aplicação do Modelo Digital Class

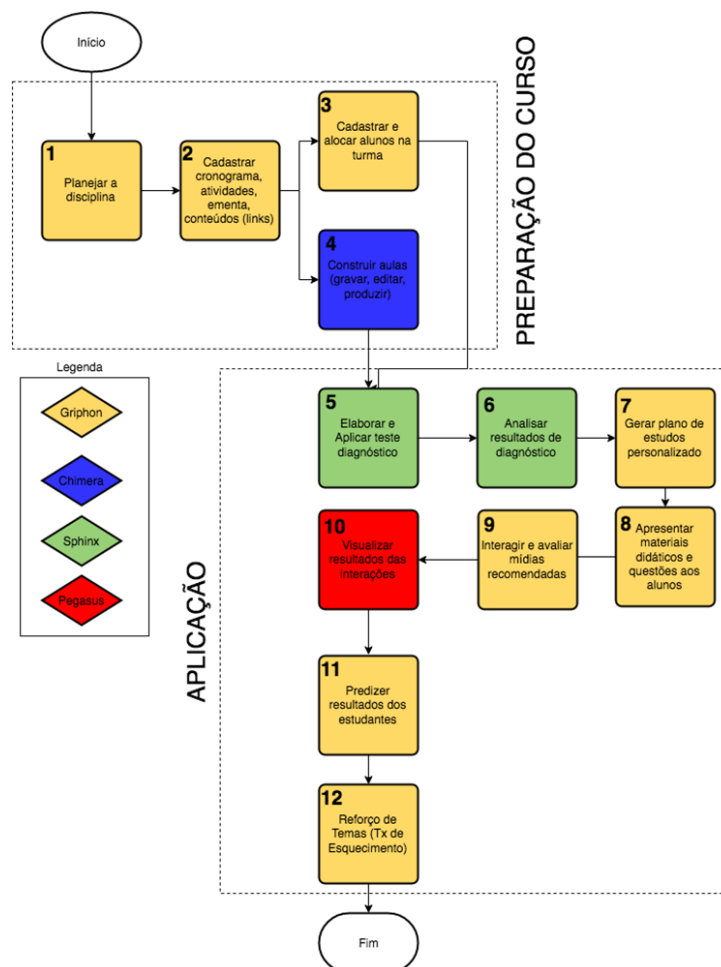
Para realizar o experimento de validação do Modelo Digital Class, foi especificado um fluxo de atividades que contemplou todas as dimensões e características da proposta.

O Modelo Digital Class permite ao professor moldar sua disciplina como considerar mais apropriado. Essa, inclusive, é uma das características que distinguem o Modelo e a Arquitetura de outras soluções em BL, comumente inflexíveis e não personalizáveis.

O fluxo de atividades apresentado na Figura 42 contempla todas as características previstas nas dimensões do Modelo Digital Class. Assim, qualquer professor que deseje utilizar amplamente o potencial do Modelo pode seguir estes passos de acordo com o ordenamento proposto.

A Figura 42 apresenta o fluxo de atividades realizado no experimento para a execução do Modelo Digital Class em uma disciplina regular. A prática ocorreu utilizando a Arquitetura

Figura 42 – Fluxo de atividades sugerido para aplicação do Modelo Digital Class.



Fonte – própria

Digital Class e todos os seus módulos.

A cor dos elementos informa em qual dos módulos criados na arquitetura será realizada cada ação. A seguir, serão descritos cada componente ilustrado na Figura 42 :

- 1 Planejamento da Disciplina: o Professor dimensiona sua disciplina informando ao sistema o conteúdo programático, o calendário de aulas, as atividades propostas e a natureza das atividades.
- 2 Cadastro de elementos do curso: o Professor realiza o cadastro de itens (questões) e conteúdos (aulas, vídeos externos, animações, etc.). O professor poderá indicar conteúdos externos ou próprios. O Módulo de autoria da Digital Class (Chimera) pode ser utilizado para essa tarefa. Em um primeiro momento, toda essa etapa será muito custosa ao Professor. No entanto, uma vez produzida, todo esse acervo poderá ser reutilizado em futuras edições de seu curso.



- 3 Cadastro de interagentes: o Professor ou administrador cadastra e aloca os alunos no sistema, separando as turmas e cursos.
- 4 Autoria de conteúdo: o Professor poderá criar notas de aulas, organizar materiais didáticos e realizar uma aula com áudio e vídeo sobre esse material. Esse conteúdo pode substituir o esforço expositivo tradicional feito em sala de aula, abrindo espaço para atividades mais ricas de acompanhamento dos alunos. Esta ferramenta fortalece o modelo de *flipped classroom*, possibilitando que os alunos assistam a aula expositiva, tenham acesso à conteúdos multimídia sobre um determinado assunto, e permitam que o tempo presencial seja utilizado para outros fins, como tirar dúvidas, realização de projetos, etc.
- 5 Elaboração e aplicação de teste diagnóstico: não é possível estabelecer qualquer estratégia de personalização sem que antes se saiba minimamente sobre as condições de aprendizado dos alunos. Desta forma, no início do curso é aplicado um teste de diagnóstico com cada aluno. No experimento, esse teste utilizou o módulo Sphinx.
- 6 Análise de Resultados do Teste Diagnóstico: após a aplicação do teste diagnóstico, a plataforma analisa o desempenho dos alunos para cada tema abordado, agrupando (clusters) os alunos de acordo com seu nível de conhecimento observado nos temas trabalhados pela avaliação.
- 7 Geração de Planos de Estudos Personalizados: o resultado do teste diagnóstico atua como fonte de entrada de dados para que, junto com o planejamento da disciplina (cronograma e objetivos), o sistema possa gerar planos de estudos contextualizados e personalizados para cada aluno.
- 8 Apresentação de Objetos Educacionais: cada aluno passa a interagir com o sistema para a realização contínua de itens que são sugeridos de acordo com o seu grau de dificuldade e objetivos a serem atingidos na disciplina. Os alunos têm acesso a materiais didáticos (vídeos, textos, animações, etc) e a uma gama de exercícios que são utilizados como termômetros para auferir a proficiência dos alunos com relação aos temas trabalhados.
- 9 Interação e avaliação dos objetos recomendados: todos os materiais que são apresentados ao estudante devem ser avaliados por eles. Esta avaliação corresponde a um *feedback* importante para o aprendizado do motor de recomendação, que poderá estabelecer uma hierarquia entre os objetos do banco, priorizando em futuras recomendações aqueles objetos de melhor avaliação.
- 10 Visualização de dados: alunos e professores possuem a seu dispor painéis de visualização

de dados que os auxiliam quanto ao entendimento da condição de aprendizado dos cursistas nos temas, o progresso/ritmo dos mesmos, o desempenho individual frente a média da turma, etc. Estas interfaces são de grande relevância, pois o formato visual de exibição dos dados facilita e motiva a percepção por parte de alunos e professores. Para o aluno, é muito importante saber se ele está progredindo, quais pontos tem tido maior sucesso, quais tem tido maior dificuldade, como está sua produção acadêmica em relação ao restante dos colegas, etc. Para o professor, além de todas as questões colocadas para o aluno, ele ainda pode tomar ciência em tempo real da quantidade e qualidade da participação de seus estudantes, podendo, caso queira, tomar atitudes a ponto de melhorar indicadores de rendimento escolar.

- 11 Predição de resultados: o sistema disponibiliza ao professor uma ferramenta de predição do desempenho estudantil. Essa funcionalidade corresponde a um modelo preditivo que indicará ao professor a condição de momento de cada aluno e da turma diante em relação a uma configuração de teste a ser aplicado no futuro. Esta funcionalidade permitirá ao professor traçar novos rumos didáticos para melhorar a condição de aprendizado dos alunos de um curso. A compreensão sobre o estado de cada estudante é obtida de forma contínua neste processo avaliativo, que tem início com a aplicação do teste diagnóstico e é alimentado, incrementalmente, a cada ação do usuário no sistema (questão resolvida). O conhecimento prévio de como deverá ser o comportamento do aluno diante de uma determinada configuração de teste, indica ao professor rumos de trabalho para tentar minorar os problemas cognitivos apontados pelo preditor.
- 12 Reforço: a plataforma avalia sistematicamente a condição do aprendizado dos estudantes nos temas, reaplicando de tempos em tempos itens de conteúdos já trabalhados. Com isso, um índice temporal (fator de esquecimento) é calculado e aplicado ao modelo que atualiza o nível de conhecimento do aluno nos temas. Considerar que um professor possa fazer retornos de tudo que tem trabalhado conceitualmente com seus alunos em um curso, é impraticável. No entanto, a ferramenta tem como princípio a verificação sistemática dentro de um intervalo de tempo programado pelo professor em que os temas são revistos, ainda que minimamente, como forma de verificar se houve um processo de esquecimento dos assuntos.

### 5.3 Caracterização Metodológica do Estudo

A estratégia de pesquisa adotada para o presente trabalho foi a modalidade de estudo de caso, com coleta de dados obtida através da interação dos estudantes com a arquitetura disponibilizada, Digital Class.

O estudo de caso foi realizado na Escola Pública de Ensino Médio Profissionalizante Jaime Alencar de Oliveira, localizada em Fortaleza-CE. As turmas de alunos trabalhadas foram as que estão no primeiro ano do ensino médio, e a disciplina escolhida foi a de Física. Esta escolha se estabeleceu em razão da maior disponibilidade do docente para colaborar com o presente estudo.

As duas turmas em que foi realizado o experimento são de cursos profissionalizantes distintos: Informática (45 alunos) e Produção de Áudio e Vídeo (47 alunos).

A seleção das turmas foi feita pelo próprio Professor. Ainda que tenha sido escolhida a mesma disciplina para o estudo de caso, destacam-se dois eixos de competência dentre os cursistas: um na área de conhecimento de exatas (Informática) e outro na área de comunicação social (Produção de Áudio e Vídeo).

O planejamento feito pelo professor para que a disciplina fosse oferecida sob a metodologia de ensino híbrido, especificou a abordagem de laboratório rotacional (STAKER E HORN, 2012). Assim, laboratórios de informática com acesso a internet foram disponibilizados para que os alunos pudessem praticar a ferramenta proposta. Em dias específicos, as turmas de alunos eram divididas, onde uma parte ficava em sala de aula e a outra se dirigia ao laboratório. Semanalmente, havia uma alternância.

Antes de iniciar o trabalho prático com os alunos junto à ferramenta em laboratório, foi feita uma apresentação da pesquisa às duas turmas. Ao lado do professor da disciplina e da Diretora da Escola, foram discutidos os objetivos que a pesquisa almejava alcançar e o que se esperava dos estudantes. Esta fala inicial foi muito oportuna, pois foi momento de tirar dúvidas, abrandar a desconfiança de muitos e motivar o corpo discente para a importância do estudo.

Foram realizadas idas aos laboratórios em um dia específico, com duração de 50 minutos. Os alunos de cada turma, eram divididos em dois laboratórios, um acompanhado pelo pesquisador e outro pelo próprio professor. Um laboratório continha *desktops* e outro *tablets*, todos com a acesso a internet e adequados à realização do experimento.

Foram dadas instruções iniciais aos alunos sobre o uso da ferramenta e suas funcionalidades. Também foi explicado sobre a importância da avaliação do aluno para cada objeto

utilizado que foi apresentado pela ferramenta (ranqueamento), como forma de socializar essa interação, para a melhoria da recomendação a outros colegas.

O trabalho teve início com o teste diagnóstico no início da disciplina. Esse diagnóstico foi realizado utilizando o módulo Sphinx. O experimento teve continuidade nos meses de maio e junho. Nesse intervalo, os alunos foram para o laboratório por seis semanas.

No último dia de trabalho em laboratório foi aplicado um questionário com quinze itens para medir a percepção e satisfação sobre a ferramenta por parte dos estudantes. A Seção 5.3.2 apresenta maiores detalhes destes instrumentos. Os resultados desse questionário e da experimentação como um todo serão apresentados e discutidos no Capítulo 6.

### ***5.3.1 Descrição da Avaliação do Modelo***

Uma vez desenvolvidos os módulos da arquitetura, a próxima etapa para a validação do Modelo Digital Class foi a realização da aplicação do modelo em uma disciplina real. Nesta aplicação, as dimensões definidas no modelo foram empregadas. A seguir, a Tabela 9 explica como cada dimensão foi avaliada.

Os dados que alimentaram o processo de análise e discussão do presente estudo, foram originários das interações dos alunos com a Arquitetura Digital Class. Foi realizada uma Análise Quantitativa dos dados, em que técnicas de estatística descritiva foram adotadas para observar possíveis correlações entre as variáveis. Foram aplicados testes para verificar a significância estatística de alguns resultados, utilizando o Teste Não Paramétrico Wilcoxon.

### ***5.3.2 Análise de Satisfação dos Usuários com a Arquitetura***

Um dos pontos relevantes para a validação da solução proposta é a identificação do grau de satisfação dos usuários (estudantes). Para isso, foi aplicado aos alunos um questionário com quinze questões. O instrumento do aluno foi aplicado no final da aula do último dia de utilização da plataforma em laboratório e continha itens que abordavam aspectos como: aceitação de tecnologia em educação (3 questões), utilidade dos elementos disponibilizados (4 questões), importância da ferramenta para o apoio ao aprendizado (3 questões), dificuldade de uso (1), etc. Além das questões objetivas, o instrumento apresentou duas questões subjetivas.

No capítulo seguinte (Capítulo 6) serão apresentados os resultados e discussões relativos à aplicação da arquitetura na experiência de campo.

Tabela 9 – Avaliação de Dimensões do Modelo

<b>Dimensão</b>	<b>Como foi avaliada?</b>
Adequabilidade a objetivos do curso	Foi observado o índice de aprovação dos alunos às recomendações realizadas de materiais educacionais.
Avaliação contínua	Foram analisados os logs com os registros da aplicação de itens às turmas ao longo do tempo. A cada aplicação foram geradas uma nota geral de cada aluno. Essas notas foram organizadas em dois grupos e, foi aplicado o teste de significância de Wilcoxon para saber se a diferença entre esses dois grupos é significativa. O objetivo foi verificar se a avaliação continuada contribui para a aquisição de conhecimento do alunos e se contribuem para a avaliação dos alunos.
Paralelização de estudos	Foi feita a análise da variabilidade de temas trabalhados concomitantemente pelos cursistas. Foi calculado, para cada aluno, a média de temas realizados ao longo das seis aplicações. Novamente foram utilizados os dois grupos de alunos. Foi verificado qual dos grupos teve maior índice de paralelização e de que forma isso impactou no desempenho dos alunos.
Ritmo próprio	Foi realizada uma análise dos perfis dos usuários considerando o ritmo da resolução de itens, utilizando a Nota Média do aluno ao longo das aplicações. O objetivo é verificar se foram observados diversos ritmos ao longo da aplicação e de que forma essas variações de comportavam em termos de obtenção de desempenho.
Predição de desempenho	Foi avaliado o desempenho dos modelos preditivos desenvolvidos. Foram utilizadas duas técnicas para esta ação preditiva: um modelo Regressor Linear Logístico e a técnica de filtragem colaborativa, no campo de fatoração de matrizes, SVD++. Foi adotada a métrica RMSE para medir a eficiência dos Modelos preditivos.
Personalização	Foi utilizada a separação de grupos de alunos de acordo com sua característica de evolução nos planos de estudo. Foi calculada uma nota geral para cada aluno por aplicação. Graficamente foram mostrados o histórico de planos de estudos de dois alunos, para verificar a personalização das recomendações do plano.

## 6 RESULTADOS E DISCUSSÕES

O presente capítulo apresenta os resultados obtidos com a realização do estudo de campo em que foram realizadas sessões de utilização da Arquitetura *Digital Class* pelos alunos.

O capítulo está assim organizado. A Seção 6.1 apresenta os resultados obtidos da aplicação em campo e discute de que forma eles validam cada dimensão do Modelo *Digital Class*. A Seção 6.2 discute o instrumento de satisfação aplicado com o corpo de estudantes.

### 6.1 Análise da Aplicação da Arquitetura *Digital Class* e Validação do Modelo

Nesta seção serão discutidos os resultados alcançados quanto à aplicação da Arquitetura *Digital Class*, assim como cada dimensão pode ser validada.

#### 6.1.1 Validação da dimensão adequabilidade a objetivos do curso

Essa dimensão possui uma grande importância em razão da necessidade de se prover ao docente uma condição absoluta de modelar os recursos didáticos presentes na arquitetura à sua turma de alunos. Como já discutido em capítulos anteriores, muitas plataformas de ensino híbrido restringem essa possibilidade de personalização, o que prejudica a realização de alguns cursos nesse formato.

A Arquitetura *Digital Class* foi construída de forma a permitir que o professor tenha um amplo controle sobre tudo o que será apresentado ao aluno, seja em termos de materiais didáticos, seja em relação a exercícios, testes avaliativos, planejamento, etc. Além disso, o professor poderá elaborar o seu próprio material didático, fazendo notas de aulas, gravando vídeos, etc. Os alunos na *Digital Class* tem acesso a vídeos, exercícios e atividades previstos pelo professor, porém de forma personalizada à sua condição de aprendizado. Logo, potencializa-se as chances de que um aluno que realize as atividades oferecidas pela arquitetura tenha uma preparação mais alinhada à disciplina. É comum observar que muitos alunos que utilizam ferramentas na internet quando vão para as avaliações do curso não confirmam o bom desempenho das plataformas. Aliando o cadastro qualificado do professor de materiais didáticos com o reconhecimento automático da necessidade de se fazer uma recomendação, faz-se com que as atividades sejam mais coerentes ao que o aluno precisa a cada instante. A Tabela 10 exibe as médias de avaliações feitas pelos alunos para os materiais apresentados pelo sistema de recomendação da *Digital Class*. Os usuários atribuíam estrelas aos objetos educacionais, na

escala de 1 a 5 estrelas. Vale lembrar que todos os materiais didáticos que foram recomendados foram inicialmente cadastrados pelo Professor.

Tabela 10 – Avaliação média dos alunos aos objetos didáticos recomendados (*rating*).

Tipo de Objeto	Média de Avaliações
Vídeo Sphinx	3,6
Vídeo Externo (Youtube)	4,06
Animação (PHET)	4

A Tabela 10 mostra, portanto, que os objetos recomendados foram bem avaliados pelos estudantes, tendo uma leve vantagem os vídeos didáticos externos que estão publicados no site *Youtube*.

### 6.1.2 Validação da dimensão avaliação contínua

Ao longo de todo o período de acompanhamento do uso da arquitetura pelos alunos, foi registrado em cada data de aplicação informações como a quantidade de exercícios resolvidos por cada aluno em cada tema, que engloba a quantidade de temas trabalhados e que fornece a nota de cada aluno em cada tema (calculada de acordo como já explicadona técnica SVD<sup>++</sup>).

A partir desses dados, para cada aplicação foi calculada uma nota geral de cada aluno, que leva em consideração os exercícios resolvidos por cada um. Assim, pode-se estabelecer que, supondo que na aplicação  $j$  o aluno  $i$  resolva exercícios nos temas  $\{a, b, \dots, k\}$ , sua nota geral na aplicação será:

$$\text{nota}_{\text{aluno } i, j} = \frac{\text{nota}_{\text{tema } a} + \text{nota}_{\text{tema } b} + \dots + \text{nota}_{\text{tema } k}}{\text{maior número de temas trabalhados por um aluno na aplicação } j}$$

$$i = 1, \dots, 92; j = 1, \dots, 6 .$$

A Tabela 11 elenca as seis aplicações realizadas no estudo, apresentando dados sobre os temas trabalhados e notas obtidas pelos alunos.

Tabela 11 – Relação de dados obtidos das aplicações (temas trabalhados × nota).

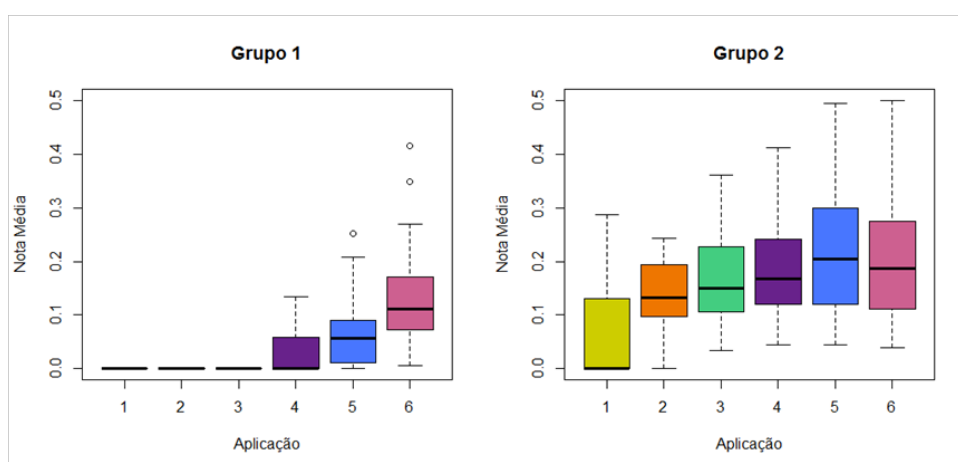
Aluno	Aplicação 1		Aplicação 2		Aplicação 3		Aplicação 4		Aplicação 5		Aplicação 6	
	temas trab.	nota	temas trab.	nota	temas trab.	nota	temas trab.	nota	temas trab.	nota	temas trab.	nota
56	2	0,39	2	0,16	3	0,29	3	0,31	5	0,45	5	0,39
57	1	0,17	2	0,22	3	0,19	3	0,16	3	0,16	3	0,14
58	0	0	0	0	0	0	1	0,05	1	0,05	1	0,04
59	0	0	0	0	0	0	2	0,12	2	0,12	2	0,11
60	0	0	3	0,25	3	0,25	3	0,26	3	0,26	3	0,22
61	1	0,17	5	0,32	5	0,28	5	0,24	5	0,24	5	0,21
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Observou-se que dos 75 alunos acompanhados pelo sistema de aprendizagem virtual, 36 só passaram a usar o sistema para resolução de exercícios na 4ª aplicação em diante. Dessa forma, foi realizado um agrupamento dos alunos em dois grupos, sendo:

- Grupo 1: Alunos que acessaram o sistema para estudos somente na 4ª, 5ª e 6ª aplicação (36 alunos);
- Grupo 2: Alunos que acessaram o sistema para estudos desde a 1ª aplicação em diante (19 alunos) e alunos que acessaram o sistema para estudos desde a 2ª aplicação em diante (20 alunos).

A Figura 43 exibe o gráfico que mostra a disposição das notas dos alunos de cada grupo ao longo das aplicações. Analisando o quesito avaliação contínua, foi obtido:

Figura 43 – Nota dos alunos obtida pelos grupos.



Fonte – da pesquisa

Com o propósito de comparar as notas de desempenho do Grupo 2, que fez uso do sistema como um ferramenta de avaliação contínua, com as notas de desempenho do Grupo



1 que não usou o sistema de forma plena (menos ocorrências), pelo menos no que se refere ao quesito avaliação contínua, foi realizado um teste estatístico não paramétrico para testar as seguintes hipóteses:

$H_0$  : Pop.grupo 1 = Pop.grupo 2 (As amostras Grupo 1 e Grupo 2 provém da mesma população)

$H_1$  : Pop.grupo 1  $\neq$  Pop.grupo 2 (As amostras Grupo 1 e Grupo 2 não provém da mesma população)

Entende-se que em cada aplicação, tem-se uma amostra de notas de desempenho que provém de uma determinada população, que se comporta de acordo com alguma distribuição de probabilidade.

O objetivo é testar se a amostra de notas de desempenho do Grupo 1 provém da mesma população das amostras de desempenho do Grupo 2. Uma vez que as amostras sejam de populações diferentes, teríamos como provar que o uso do sistema com a ferramenta de avaliação contínua influencia na obtenção de melhor desempenho dos alunos.

Ao analisar a igualdade das notas de desempenho de cada grupo, realizou-se o teste não paramétrico Wilcoxon, com um nível significância de 5%, e observou-se que em todas as aplicações as notas de desempenho do Grupo 1 são estatisticamente diferentes das notas de desempenho do Grupo 2. A Tabela 12 apresenta os p-valores encontrados.

Tabela 12 – Teste não paramétrico para a comparação das notas de desempenho dos grupos 1 e 2 em cada uma das aplicações.

<b>Teste Não Paramétrico Wilcoxon</b>			
<b>Aplicação</b>	Aplicação 4	Aplicação 5	Aplicação 6
<b>p-valor</b>	4,41e-12	1,67E-05	0,002217

Como é observado na Figura 44, os valores amostrais do Grupo 2 (alunos que fizeram uso da ferramenta de avaliação contínua de forma mais prolongada) são, em distribuição, maiores que os valores amostrais do Grupo 1, e essa diferença foi estatisticamente significativa (Tabela 12).

Isso indica que o valor esperado da nota média de desempenho dos alunos do Grupo 2 é maior que o valor esperado da nota média de um aluno pertencente ao Grupo 1. Reforçando a informação contida nos gráficos da Figura 43, as Tabelas 13, 14 e 15 apresentam as estatísticas básicas das notas desempenho dos dois grupos.

Tabela 13 – Estatísticas básicas das notas médias de desempenho de cada aluno na aplicação 6.

<b>Estatísticas básicas</b>							
	<b>Grupos</b>	Min	1º Qu.	Mediana	Média	3º Qu.	Max.
<b>Aplicação 6</b>	Grupo 1	0,006	0,073	0,111	0,130	0,166	0,416
	Grupo 2	0,038	0,111	0,186	0,203	0,275	0,501

Tabela 14 – Estatísticas básicas das notas médias de desempenho de cada aluno na aplicação 5.

<b>Estatísticas básicas</b>							
	<b>Grupos</b>	Min	1º Qu.	Mediana	Média	3º Qu.	Max.
<b>Aplicação 5</b>	Grupo 1	0	0,013	0,057	0,068	0,089	0,253
	Grupo 2	0,044	0,119	0,204	0,216	0,300	0,494

Tabela 15 – Estatísticas básicas das notas médias de desempenho de cada aluno na aplicação 4.

<b>Estatísticas básicas</b>							
	<b>Grupos</b>	Min	1º Qu.	Mediana	Média	3º Qu.	Max.
<b>Aplicação 5</b>	Grupo 1	0	0	0	0,030	0,059	0,134
	Grupo 2	0,044	0,119	0,168	0,185	0,242	0,412

Como pode-se observar, durante as três aplicações, os valores das estatísticas básicas sobre as notas médias de desempenho são sempre maiores no Grupo 2.

Os resultados demonstram que uma avaliação contínua além de fornecer ao professor informações sobre a condição de aprendizado do aluno, representa também uma importante atividade de exercício para eles. Os dados apontam que quanto mais exercícios e temas os alunos se submeteram a resolver, melhor passou a ser o seu desempenho.

### **6.1.3 Validação da dimensão paralelização de estudos**

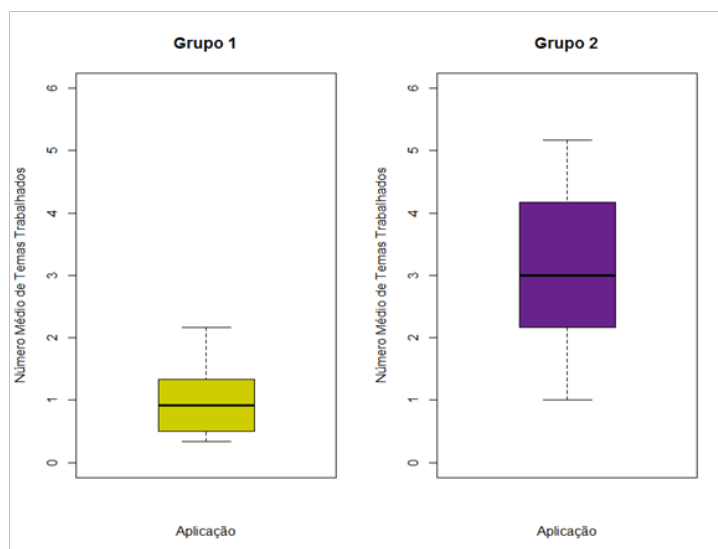
A Arquitetura *Digital Class* oferece ao estudante a oportunidade de realizar seus estudos de forma concomitante, ou seja, realizando seus estudos em mais de um tema simultaneamente. Desta forma, para se observar de que maneira os cursistas interagiram com os temas, foi observado o quantitativo de temas trabalhados por cada aluno durante as aplicações e, em seguida, calculada a média de temas realizados por cada aluno ao longo das aplicações. A Tabela 16 ilustra como foram organizados esses dados.

Tabela 16 – Média de Temas Trabalhados por alunos durante as aplicações.

ID alunos	Temas trabalhados						Valor médio por aplicação
	Aplicação 1	Aplicação 2	Aplicação 3	Aplicação 4	Aplicação 5	Aplicação 6	
56	2	2	3	3	5	5	3,33
57	1	2	3	3	3	3	2,50
58	0	0	0	1	1	1	0,50
59	0	0	0	2	2	2	1,00
60	0	3	3	3	3	3	2,50
61	1	5	5	5	5	5	4,33
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Continuando com os dois grupos definidos na sessão anterior, foi analisado o valor médio de temas trabalhados por cada aluno em cada grupo nas aplicações. A Figura 44 apresenta, para cada grupo, a relação entre o número médio de temas trabalhados pelas aplicações.

Figura 44 – Relação entre o número médio de temas trabalhados pelos grupos ao longo das aplicações.



Fonte – da pesquisa.

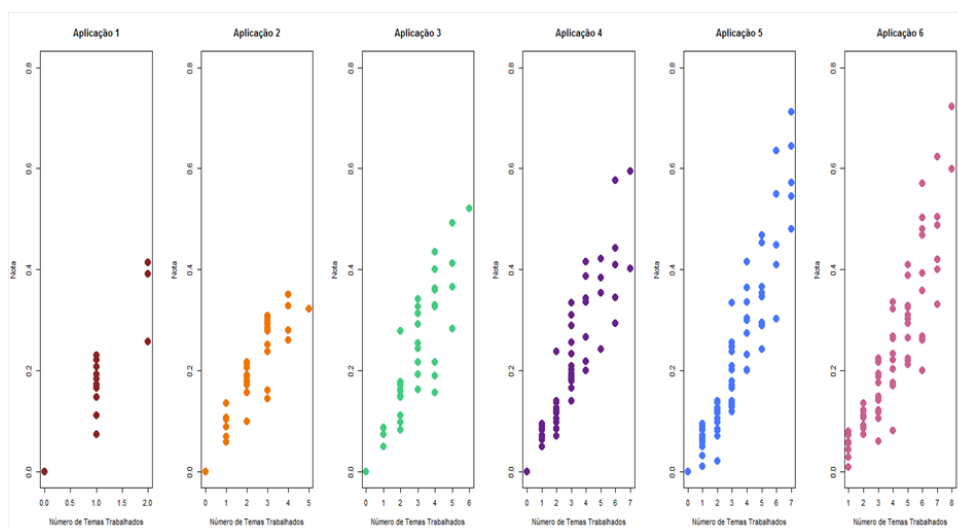
Percebe-se, que o número médio de temas trabalhados nas aplicações apresenta maiores valores no Grupo 2, em relação ao Grupo 1.

Pela Figura 44, percebe-se que 75% dos alunos do Grupo 2 realizaram, em média, de 2 a 5 temas, enquanto que no Grupo 1, o número médio de temas abordados foi de no máximo 2 temas.

Os resultados apresentados na Figura 44 demonstram claramente que a paralelização dos estudos feita pelo Grupo 2 foi superior ao Grupo 1. A discussão da seção anterior (Tabela 12) observou também que o desempenho dos alunos do Grupo 2 foi maior do que o Grupo 1.

Associando estes resultados, podemos sugerir que os alunos que trabalharam um maior número de temas simultaneamente, ou seja, que tiveram uma maior paralisação, obtiveram um melhor desempenho. Essa situação foi observada em todas as seis aplicações. Neste sentido, a Figura 45 exibe o gráfico de dispersão com as notas dos alunos relacionadas ao número de temas trabalhados por cada aplicação. Os dados corroboram com a conclusão anterior. Portanto, alunos com maior intensidade de estudos, apresentam de fato melhor desempenho.

Figura 45 – Gráficos de Dispersão das Aplicações (Relação de Notas e Número de Temas Trabalhados)



Fonte – da pesquisa

#### 6.1.4 Validação da dimensão ritmo próprio

A dimensão ritmo próprio pode ser validada a partir das discussões anteriores realizadas para as dimensões: avaliação contínua e paralelização de estudos. Isto se dá, pois um aluno com plena liberdade para resolver exercícios, que são a base da avaliação contínua, pode exercer sua autonomia para aprofundar estudos resolvendo o número de questões que quiser. Em relação à paralelização de estudos, este aluno pode realizar o cumprimento de um plano de aula em múltiplos temas. Uns de maneira mais célere que outros.

Como foi visto anteriormente, o indicador de desempenho de cada aluno ao longo das aplicações, aqui chamado de Nota Média do Aluno, tem forte relação com o número de questões resolvidas em cada aplicação, assim como o número de acertos, tentativas e o número de temas abordados.

No tópico sobre Avaliação Contínua foi visto que a maior frequência de uso do

sistema provoca maior aprendizado fazendo com que os alunos tenham maiores notas médias de desempenho no final do semestre. No tópico Paralelização dos Estudos foi visto que a interação concomitante em múltiplos temas pelos alunos, ou seja, os alunos que exploravam mais de um tema por aplicação, tiveram forte influência quanto ao aumento da nota média de desempenho.

Com isso, de forma geral, a nota média do aluno reflete o ritmo de uso de cada aluno no sistema, ao longo de cada aplicação. Seja pela frequência de acessos e número de temas abordados, seja pelo número de questões resolvidas e o número de acertos obtidos.

Analisando a questão ritmo de uso do sistema, percebe-se que o sistema permite certa flexibilidade ao aluno.

No tópico anterior foi definido o Grupo 2 como o grupo de alunos que usaram o sistema para a resolução dos exercícios de forma contínua, ou seja, fizeram uso do sistema em quase todas aplicações (alguns alunos deste grupo não acessaram na aplicação 1). De toda forma, foi comprovado com nível de significância de 5% que tais alunos tinham em distribuição uma nota média de desempenho esperada maior que os demais alunos.

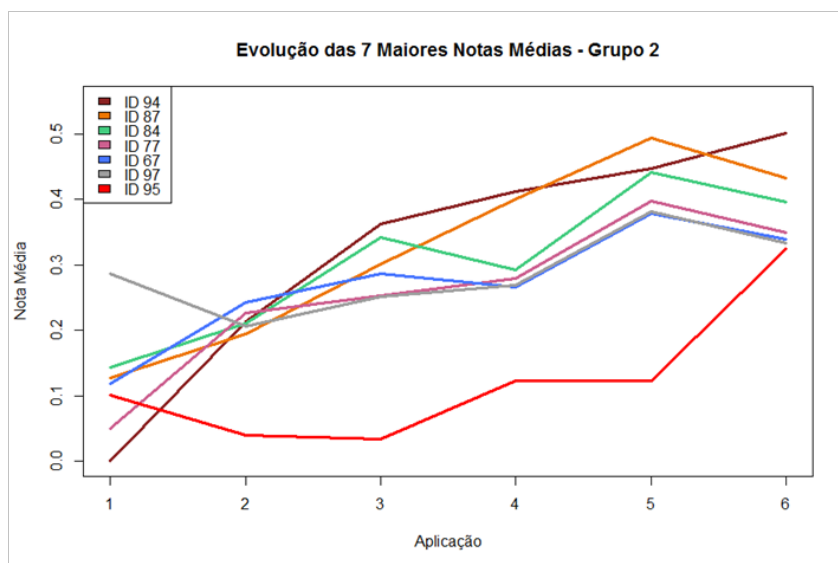
Analisando o ritmo de uso do sistema desses alunos, observa-se que a quantidade de exercícios resolvidos por aplicação, assim como a quantidade de acertos e temas abordados de um mesmo aluno, varia ao longo das aplicações de acordo com o ritmo deste.

A Figura 46 exibe a evolução das notas dos 7 alunos que na aplicação 6 (última aplicação) ficaram com as maiores notas médias de desempenho. Observou-se que cada aluno evoluiu de acordo com o seu ritmo próprio, o que não o impediu de obter boas notas médias finais.

Na Figura 46 é possível observar que há uma certa oscilação no ritmo de uso do sistema de alguns alunos durante as aplicações. O aluno de ID 95, por exemplo, teve uma queda no rendimento do uso do sistema nas aplicações 2 e 3, mas obteve melhora nas aplicações 4, 5 e, na aplicação 6, obteve uma nota bem próxima às notas dos alunos de ID 67 e 97, que ao longo das aplicações usaram o sistema com um ritmo maior.

Analisando de forma geral (não considerando apenas os alunos do Grupo 2 definido anteriormente), observa-se que alguns alunos que faziam parte do Grupo 1 conseguiram obter boas notas médias finais na última aplicação, chegando a ficar entre as 7 melhores notas.

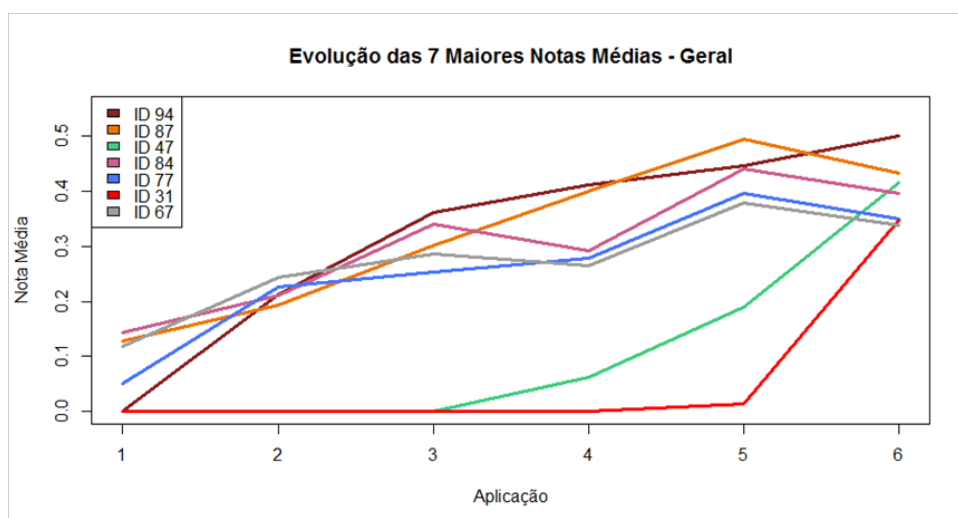
Figura 46 – Gráficos de Evolução dos sete alunos com maiores médias.



Fonte – da pesquisa.

Isso comprova que o sistema respeita o ritmo de cada aluno e que flexibilizar essa condição contribui para a personalização do curso. É importante observar que mesmo possuindo alunos com trajetórias diversificadas, uns mais rápidos, outros nem tanto, foram obtidos bons resultados no final da aplicação (Figura 47).

Figura 47 – Gráficos de Evolução das notas médias.



Fonte – Fonte: da pesquisa.

Na Figura 47 vale destacar os alunos de ID 47 e 31 que, apesar de não terem usado o sistema em todas as aplicações, conseguiram, de acordo com seus ritmos, ficar entre as 7 maiores

notas na última aplicação. Assim, mesmo o aluno de ID 47 tendo feito uso do sistema somente a partir da terceira aplicação, em um ritmo mais lento, obteve a 3<sup>a</sup> maior nota média final na última aplicação.

### **6.1.5 Validação da dimensão predição de desempenho**

Esta dimensão tem como objetivo prover ao professor um mecanismo preditivo que permita a averiguação antecipada e aproximada da condição de sucesso do aluno frente às futuras avaliações. Este conhecimento prévio permite que tanto a arquitetura, quanto o professor, possam estabelecer ações específicas de contingenciamento para minorar o déficit cognitivo de cada aluno em um tema. Como discutido em capítulos anteriores, a estratégia desse trabalho foi desenvolver duas abordagens para modelar o problema de predição: Modelo de Regressão Linear Logística e SVD<sup>++</sup> (*Singular Value Decomposition*).

#### **6.1.5.1 Modelo de regressão linear logística**

Considerando que o nível de proficiência de um aluno em determinado tema pode estar associado com a quantidade e a qualidade de exercícios que esse aluno faz, foi então feito uso da técnica de regressão linear logística para avaliar o nível de tal relação.

Para tanto, partindo de um determinado tema e nível de complexidade foram coletadas todas as respostas de todos os exercícios feitos por cada aluno, sendo que o último exercício realizado foi utilizado como questão-prova em função dos exercícios anteriores, ou seja, o último exercício foi a variável resposta que dependeu de exercícios anteriores, de forma a responder a seguinte pergunta:

*De que forma os exercícios realizados por um aluno em determinado tema/nível influenciam esse aluno quanto ao acerto de uma questão de mesmo tema/nível na avaliação bimestral?*

Dos dez temas abordados no curso (vide Tabela 2), apenas dois foram objeto de avaliação na segunda prova bimestral (aplicada no final do primeiro semestre de 2017.1), a saber, Movimento Retilíneo Uniforme e Movimento Retilíneo Uniformemente Variado. Desta forma, os modelos gerados concentraram-se nos dois temas citados.

### **Analisando o Tema 85: Movimento Retilíneo Uniformemente Variado.**

Analisando os dados que contêm as respostas dos exercícios dos alunos no tema 85: Movimento retilíneo uniformemente variado para os níveis básico, intermediário e avançado, apenas a variável **Número de Tentativas** se mostrou significativa para o modelo, ao nível de significância de 5%. Assim, obteve-se o seguinte modelo final ajustado:

$$(\text{Chance de acerto})_i = 0,13779 * (\text{N}^\circ \text{tentativas})_i, \quad i = 1, \dots, n.$$

A Tabela 17 apresenta a análise de estimativas realizada.

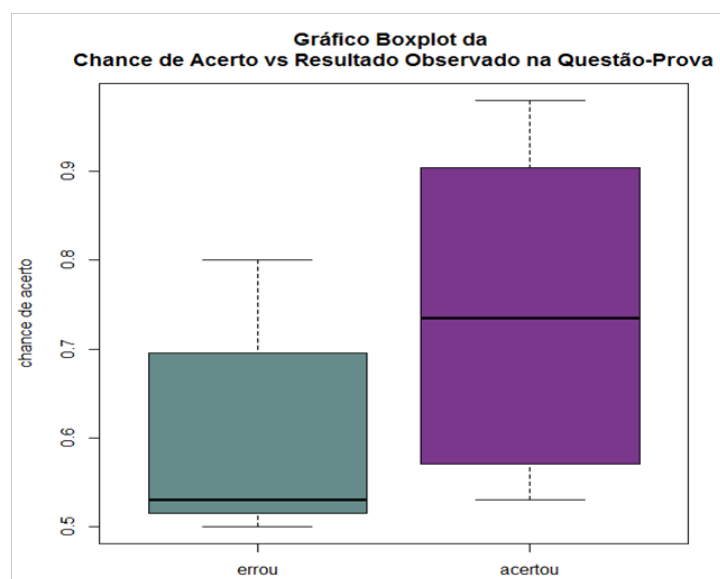
Tabela 17 – Análise de Estimativas

Variável	Estimativa	p-valor
Nº de tentativas	0,1378	0,0157

Observando os dados de cada aluno, como as variáveis número de acertos e tentativas de resolução nos exercícios no tema 85, foram obtidas as chances de acerto desses alunos numa suposta questão-prova sobre o mesmo tema.

A Figura 48 apresenta um gráfico boxplot que aponta as chances de acerto calculadas com base nos exercícios resolvidos.

Figura 48 – Fonte de Acerto × Resultado Observado na Questão-Prova



Fonte – da pesquisa.



Observando a Figura 48 percebe-se que todos os alunos que erraram a questão-prova tinham chance de acerto bem menor que os alunos que acertaram a questão-prova, ou seja, a quantidade de exercícios resolvidos mesmo sem considerar a quantidade de acertos, influencia no bom desempenho do aluno nas avaliações futuras (questão-prova, por exemplo).

A Tabela 18 traz uma comparação entre as chance de acerto na questão-prova com as notas obtidas pelos alunos na avaliação final bimestral.

Tabela 18 – Comparação entre a chance de acerto dimensionada com a nota obtida na bimestral.

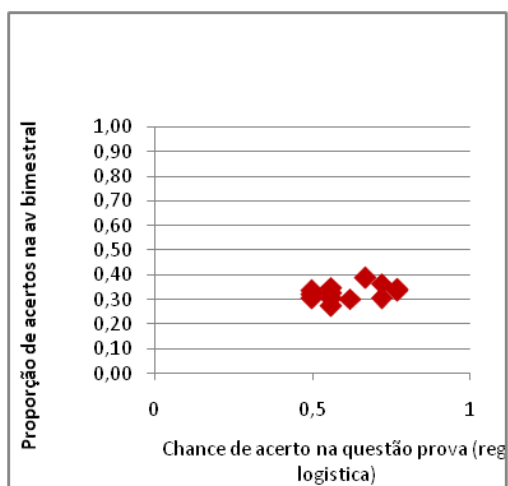
ID_aluno	N_quest	Acertos	Tentativas	Chance de Acerto (RL)	Questão-prova	Nota_t85_bim
15	0	0	0	0,5	0	0,29
81	0	0	0	0,5	0	0,43
85	0	0	0	0,5	0	0,29
26	1	0	1	0,56	0	0,57
48	1	1	1	0,56	0	0,29
83	1	1	1	0,56	0	0,14
99	1	0	1	0,56	0	0,43
56	3	3	4	0,67	0	0,14
47	4	4	10	0,72	0	0,29
90	5	3	5	0,77	0	0,86
44	1	0	1	0,56	1	0,71
54	1	0	1	0,56	1	0,29
55	1	1	1	0,56	1	0,29
61	1	1	2	0,56	1	1
62	1	1	2	0,56	1	0
92	2	1	3	0,62	1	0,57
72	3	3	8	0,67	1	0
96	3	1	4	0,67	1	0,29
12	5	4	6	0,77	1	0,43
91	5	2	7	0,77	1	0,43
⋮	⋮	⋮	⋮	⋮	⋮	⋮

A coluna Chance de Acerto indica a chance do aluno acertar uma questão-prova, baseada em seus exercícios anteriores. Enquanto que na coluna seguinte, temos a real resposta do aluno na última questão-prova (último exercício realizado pelo aluno antes da prova bimestral). Em seguida tem-se a nota obtida pelo aluno na avaliação final bimestral. Percebe-se nas primeiras três linhas da tabela, por exemplo, que mesmo sem ter resolvido nenhum exercício,

esses alunos tem 50% de chance de acertar uma questão-prova.

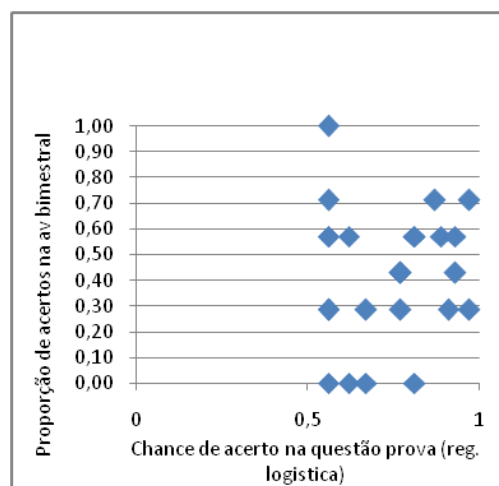
A Figura 49 apresenta um gráfico de dispersão dos alunos que erraram na questão-prova (último exercício realizado) com a nota obtida na avaliação bimestral. A Figura 50 mostra a mesma a relação com os alunos que acertaram a última questão-prova.

Figura 49 – Chance de Acerto x Proporção de Acerto na Bimestral (Alunos que erraram a questão-prova)



Fonte – da pesquisa.

Figura 50 – Chance de Acerto x Proporção de Acerto na AV Bimestral (alunos que acertaram a questão-prova)



Fonte – da pesquisa.

De forma geral, percebe-se que os alunos que acertaram a questão-prova (Figura 50), além de terem chances de acerto maiores por parte do modelo de regressão logística, grande parte obteve notas elevadas na avaliação bimestral, em comparação com os alunos que não acertaram na questão prova.

No entanto, percebe-se muito que as chances de acerto no geral são altas, maiores que 50% para todos os alunos, o que indica que o modelo de regressão logística para previsão de acerto na avaliação é pouco informativo.

### **Analisando o Tema 86: Movimento Retilíneo Uniforme**

Analisando os dados que contém as respostas dos exercícios do alunos no tema 86: Movimento Retilíneo Uniforme (níveis básico, intermediário e avançado)foi verificado que as variáveis **Número de Acertos** e **Número de Tentativas** se mostraram significativas para o modelo, ao nível de significância de 5%. Assim obteve-se seguinte modelo final ajustado:

$$(\text{Chance de acerto})_i = 0,6218 * (\text{N}^\circ \text{ acertos})_i - 0,4534 * (\text{N}^\circ \text{ tentativas})_i, \quad i = 1, \dots, n.$$

A Tabela 19 apresenta as estimativas para o modelo.

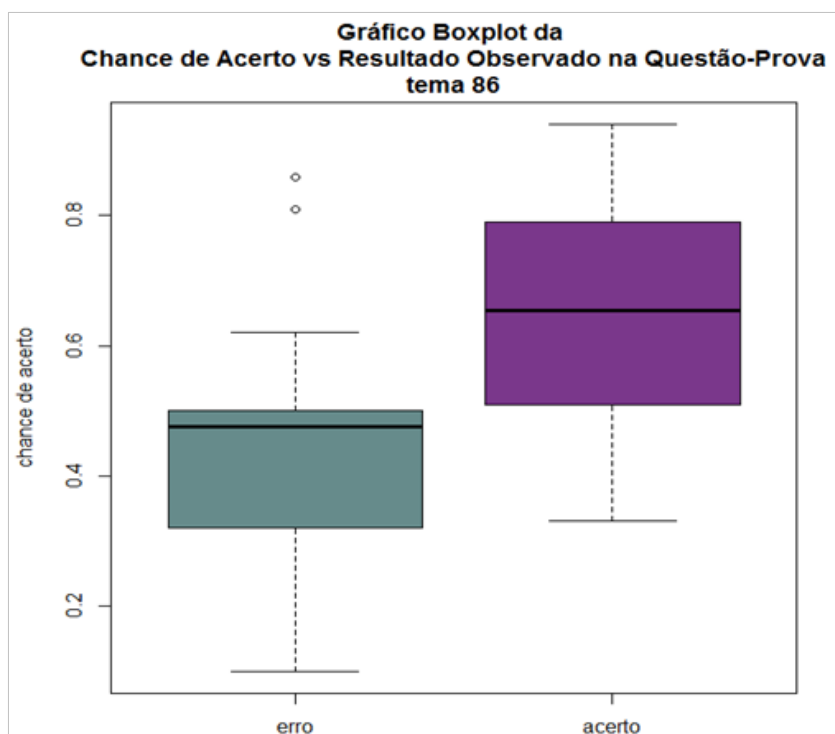
Tabela 19 – Análise das estimativas

Variável	Estimativa	p-valor
Nº de acertos	0,6218	0,021
Nº de tentativas	-0,4534	0,0296

Aplicando no modelo os dados observados de cada aluno (número de acertos e número de tentativas de resolução nos exercícios no tema 86: Movimento retilíneo uniforme), obteve-se as chances de acerto desses alunos numa suposta questão-prova sobre o mesmo tema.

A Figura 51 mostra as chances de acerto calculadas com base nos exercícios resolvidos dispostas em um gráfico boxplot.

Figura 51 – Chance de Acerto vs Resultado da Questão-prova (Tema 86).



Fonte – da pesquisa

Mais acentuado que no modelo do tema 85, percebe-se que as chances de acerto na questão-prova são maiores para os alunos do grupo que realmente acertou. O que indica um bom

desempenho do modelo em indicar a chance de acerto. A melhora na eficiência desse modelo em comparação ao anterior é em razão do maior número de observações disponíveis.

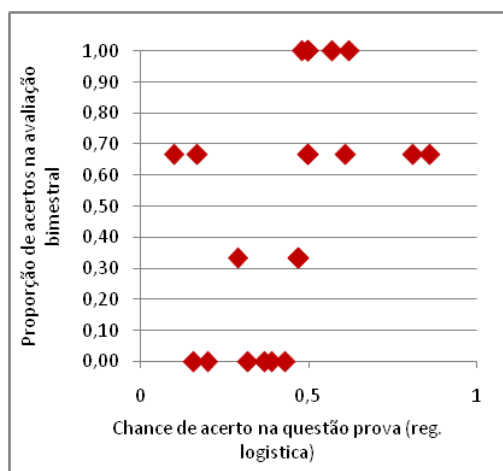
A Tabela 20 exibe os dados obtidos a partir da interação dos alunos com a arquitetura.

Tabela 20 – Dados da interação dos alunos e previsão de acerto do modelo.

id_aluno	n_quest	acertos	tentativas	chance_de_acerto (reg. logística)	resposta	
					na última questão dos exercícios	nota_t86_bim
85	13	6	16	0,10	0	0,67
60	9	4	10	0,17	0	0,67
59	2	1	4	0,43	0	0,00
56	14	10	21	0,47	0	0,33
91	7	5	8	0,48	0	1,00
⋮	⋮	⋮	⋮	⋮	⋮	⋮

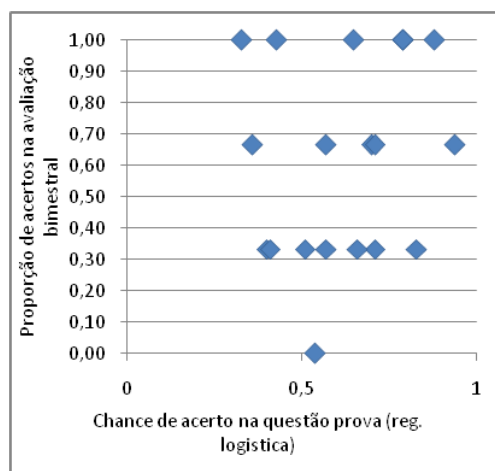
Analisando por meio de gráfico de dispersão, percebe-se que ao comparar com a nota obtida na avaliação bimestral, muitos alunos que tinham baixa chance de acerto numa questão-prova, tiveram um bom desempenho na avaliação bimestral (Figura 52). Sugerindo, mais uma vez, a importância de se adotar um outro método de previsão de desempenho que tenha melhor desempenho.

Figura 52 – Chance de Acerto x Proporção de Acerto na Bimestral (alunos que erraram)



Fonte – da pesquisa.

Figura 53 – Chance de Acerto x Proporção de Acerto na AV Bimestral (alunos que acertaram a questão-prova)



Fonte – da pesquisa.

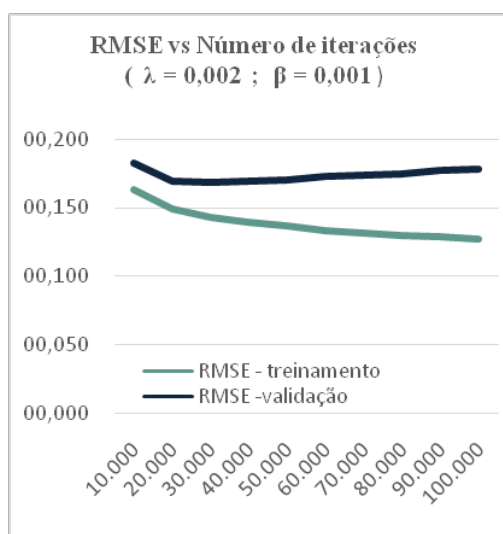
### 6.1.5.2 Modelo SVD<sup>++</sup>

A segunda abordagem utilizada no presente trabalho para modelar a predição do desempenho de estudantes foi o SVD<sup>++</sup>. Para uma maior eficiência do modelo de previsão é necessário escolher bem os valores a serem atribuídos ao número de fatores latentes  $K$ , à taxa de aprendizado  $\beta$  e ao termo de regularização  $\lambda$ . Para tanto, foram testadas várias configurações de parâmetros, a fim de reduzir o valor do RMSE calculado com os dados do conjunto validação (quanto menos melhor).

#### Número de iterações

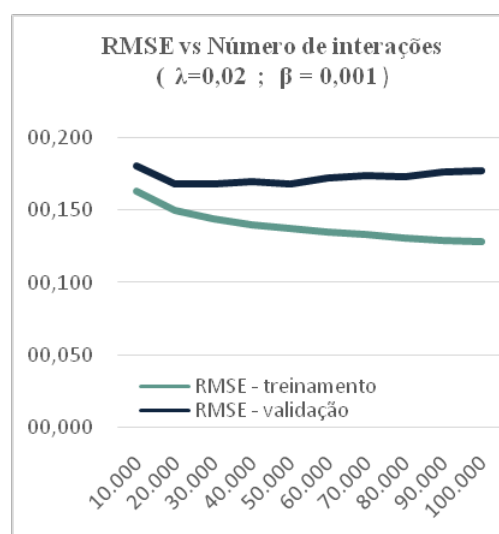
Considerando o baixo número de dados para se gerar o modelo, foram realizados tanto nas etapas de treinamento, quanto na de validação, um número elevado de iterações, variando de 10.000 a 100.000. Ao simular o modelo de previsão para diferentes tamanhos do número de iterações, conforme observado pelas Figuras 54 e 55, percebe-se pela comparação dos valores de RMSE que acima de 20.000 iterações passa a ocorrer um aumento constante do RMSE de validação em contraste com o RMSE de treinamento que diminui progressivamente. Isso indica que acima de 20.000 iterações, o modelo fica muito ajustado aos dados de treinamento (*overfitting*), passando a diminuir sua eficiência ao passo que o número de iterações aumenta.

Figura 54 – RMSE x Número de iterações



Fonte – da pesquisa.

Figura 55 – RMSE x Número de iterações



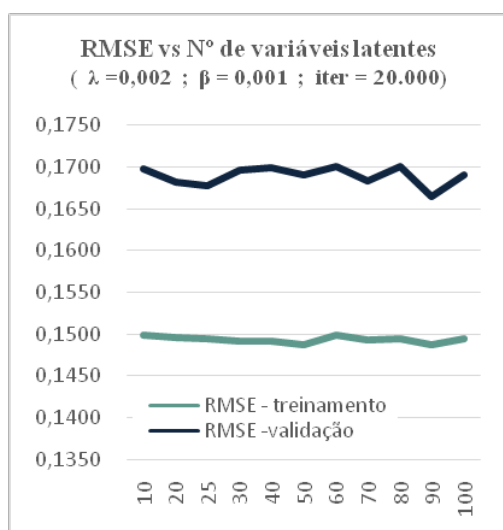
Fonte – da pesquisa.

## Número de Variáveis Latentes

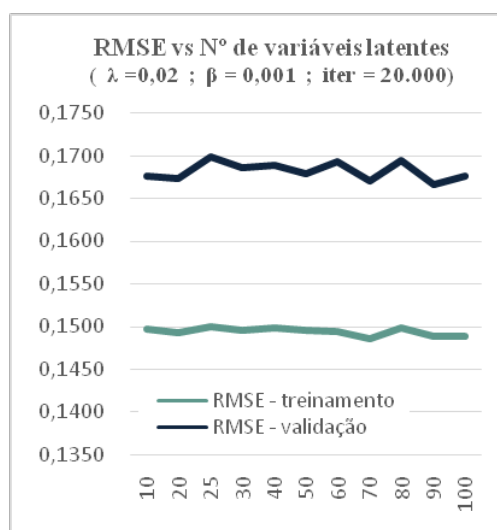
Os fatores latentes no modelo são informações contidas nos vetores de informações dos usuários (alunos) e dos itens (atividades, exercícios, etc.).

Para a obtenção de uma boa predição, a escolha apropriada da quantidade de variáveis latentes é muito importante. Persegue-se a obtenção de um valor que seja abrangente o suficiente para captar as principais características presentes no conjunto de dados, e que não comprometa a construção do modelo em termos de custo computacional. Para isso, foram observados os resultados do RMSE de acordo com a escala de 10 a 100 variáveis latentes. Os resultados são exibidos nas Figuras 56 e 57.

Figura 56 – RMSE x Número de variáveis latentes      Figura 57 – RMSE x Número de variáveis latentes



Fonte – da pesquisa.



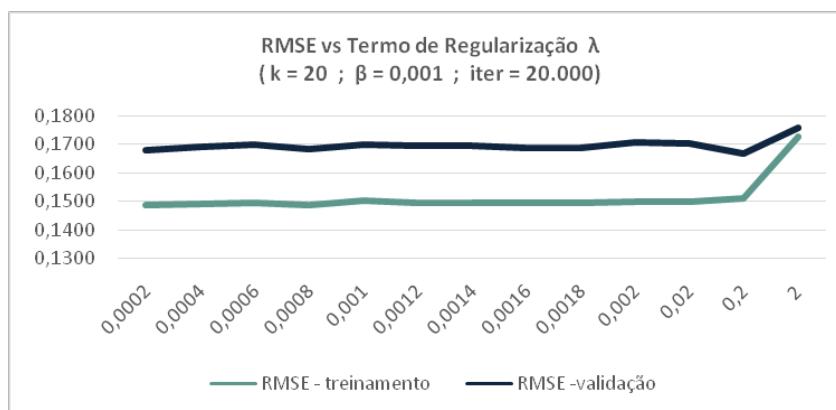
Fonte – da pesquisa.

Como pode ser observado no gráfico da Figura 56 o RMSE de treinamento tem um aprimoramento de seu desempenho com 50 variáveis latentes. A partir deste número o modelo passa a oscilar, voltando a ter valor menor com 90 variáveis latentes. Neste mesmo gráfico o RMSE de validação também obteve o melhor desempenho com 90 variáveis. No entanto, vale observar, também, o custo computacional. Para o processamento com o uso de 50 variáveis o tempo de processamento foi de 2,38 minutos e com 90 variáveis de 3,93 minutos.

## Valor do termo de regularização $\lambda$

O parâmetro termo de regularização está presente na equação do gradiente descendente para evitar o sobre-ajuste do modelo (*overfitting*) quanto à seus dados de treinamento. Um modelo sobre-ajustado possui alta precisão quanto a seu conjunto de dados de treinamento. Porém, em relação aos dados de teste apresenta oscilação e conseqüente elevação de valor do cálculo do erro, como, em nosso experimento, o RMSE.

Figura 58 – RMSE x Termo de Regularização ( $\lambda$ )



Fonte – da pesquisa

A Figura 58 utiliza como taxa de aprendizagem o valor 0,001 ( $\beta = 0,001$ ) e 20 variáveis latentes ( $k = 20$ ) e com termo de regularização  $\lambda$  variando de 0,0002 a 2. Observa-se que para valores muito pequenos de  $\lambda$ , não há, praticamente, alteração nos valores de RMSE, obtendo-se o melhor RMSE - validação (0,1667) em  $\lambda = 0,2$ . A partir deste valor, o RMSE tem forte crescimento (vide  $\lambda = 2$ ).

### A Melhor Configuração Encontrada

A Tabela 21 apresenta os dados utilizados nos experimentos para a obtenção da melhor configuração de parâmetros do modelo.

É possível destacar como melhor conjunto de valores para os parâmetros a primeira coluna da tabela (RMSE-validação = 0,1677, 20.000 iterações,  $k = 10$ ,  $\beta = 0,001$ ,  $\lambda = 0,02$ , tempo de processamento 0,75 min). Mesmo o conjunto de experimentos trabalhando com  $k = 70$  que obteve RMSE = 0,1671, seu tempo de processamento atinge 3,02 min, mais do que três vezes o processamento para a conjunto anterior.

Em todo o experimento foram utilizados no conjunto de treinamento, 196 tuplas. Já para o conjunto de validação, foram utilizados 54 registros. Em razão do baixo número de

dados, não foi feita uma separação das turmas para geração do modelo, sendo ambos os registros utilizados no processo.

Tabela 21 – Tabela de valores de parâmetros utilizados para geração do modelo

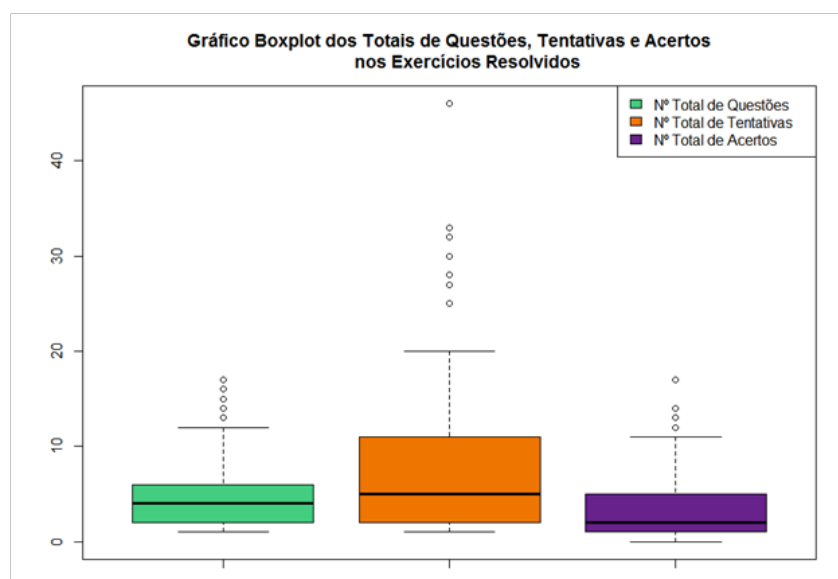
	RMSE segundo o número $k$ de variáveis latentes					
Nº de obs - Treinamento	196	196	196	196	196	196
Nº de obs - Validação	54	54	54	54	54	54
Interações	20.000	20.000	20.000	20.000	20.000	20.000
$K$	10	20	25	30	40	50
$\beta$ (taxa de aprendizado)	0,001	0,001	0,001	0,001	0,001	0,001
$\lambda$ (fator de regularização)	0,02	0,02	0,02	0,02	0,02	0,02
tempo de proc (minutos)	0,75	1,2	1,32	1,61	2,51	2,15
$\mu_{inicial}$	0,3399	0,3399	0,3399	0,3399	0,3399	0,3399
$\mu_{final}$	0,2989	0,2953	0,2986	0,2986	0,2997	0,2968
RMSE - treinamento	0,1496	0,1493	0,15	0,1495	0,1498	0,1495
RMSE - validação	0,1677	0,1673	0,1698	0,1685	0,1688	0,1678

Analisando inicialmente os dados, percebe-se que de forma geral, durante todo o período de acompanhamento, os alunos não acertaram as questões na primeira tentativa. A Figura 59 apresenta um gráfico *boxplot* com os valores totais das variáveis: Número de Questões Resolvidas, Número de Tentativas e Número de Acertos Obtidos.

Pode-se ver que amplitude do número de tentativas geral é bem maior que a amplitude do número de questões resolvidas no geral. Percebe-se, ainda, que 50% das observações tiveram, no máximo, apenas 2 acertos em todo o período de exercícios, o que é um baixo número de acertos comparado com o número mediano de tentativas, que indica que 50% das observações realizaram, no máximo, 5 tentativas.



Figura 59 – Boxplot dos Totais de Questões Resolvidas, Tentativas e Acertos obtidos.



Fonte – da pesquisa.

A Tabela 22 apresenta os dados contendo as estatísticas descritivas do gráfico anterior (Figura 59).

Tabela 22 – Tabela de valores de parâmetros utilizados para geração do modelo.

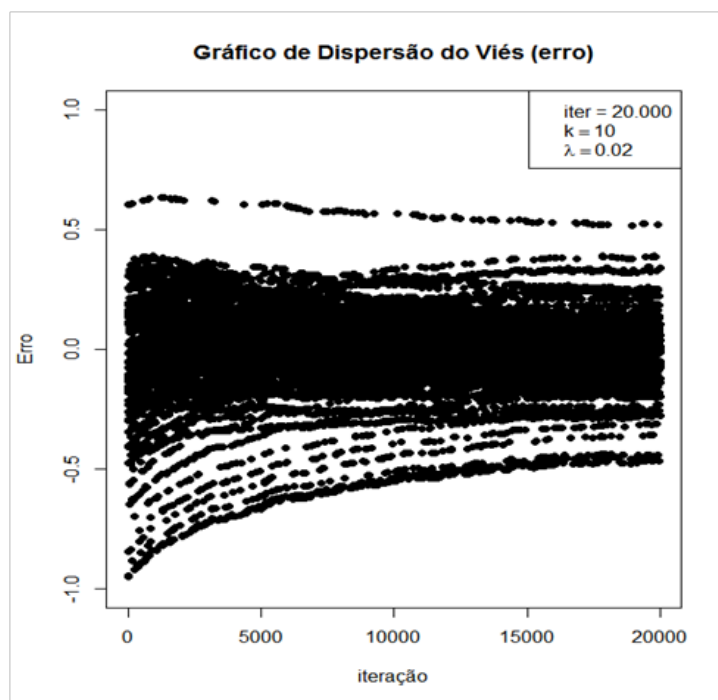
Variáveis	Min	1º Qu.	Mediana	Média	3º Qu.	Max.
Total de Questões Resolvidas	1	2	4	4,4	6	17
Total de Tentativas	1	2	5	7,3	11	46
Total de Acertos	0	1	2	3,3	5	17

De forma geral, percebe-se também um baixo nível de uso do sistema para a resolução dos exercícios, onde em 75% das observações os alunos resolveram ou tentaram resolver, no máximo, 6 questões em todo o período de acompanhamento.

Aplicando o algoritmo de previsão de desempenho  $SVD^{++}$  com o conjunto treinamento com 196 observações e, após 20.000 iterações, foi obtido o gráfico abaixo (Figura 60) que mostra a evolução do erro de estimação ao longo das iterações.

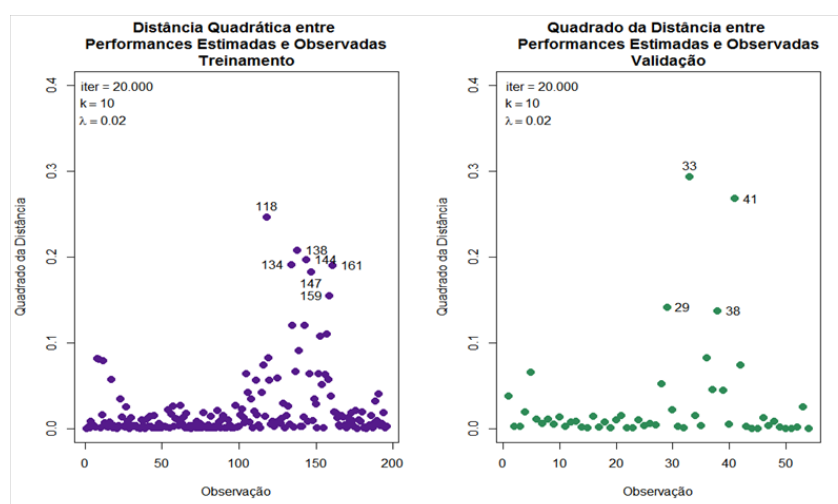
Através de outros testes, percebe-se uma convergência maior dos valores dos erros de estimação quando se tem um conjunto de treinamento maior. No entanto, só foi possível dispor de 196 observações para o conjunto treinamento, sendo que as 54 observações restantes foram usadas para o conjunto de validação.

Figura 60 – Dispersão do viés (erro) ao longo das iterações.



Analisando a diferença quadrática entre os valores preditivos e observados, foram obtidos os gráficos de dispersão da Figura 61, onde foi possível observar alguns valores em destaque.

Figura 61 – Dispersão do viés (erro) ao longo das iterações.



Fonte – da pesquisa

As Tabelas 23 e 24 analisam os pontos em destaque no conjunto treinamento e validação.

Tabela 23 – Dados de Treinamento

Dados de Treinamento				
linha	ID aluno	conteúdo	nota exerc	nota estimada
118	84	85	0,95	0,45
134	31	86	0,05	0,49
138	57	86	0,05	0,51
144	65	86	0,05	0,49
147	80	86	0,1	0,53
159	96	86	0,99	0,59
161	99	86	0,05	0,49

Tabela 24 – Dados de Validação

Dados de validação				
linha	ID aluno	conteúdo	nota exerc	nota estimada
29	67	85	0,73	0,35
33	77	85	0,89	0,35
38	76	86	0,2	0,57
41	41	86	0,05	0,57

Observando a Tabela 23, é possível perceber que os pontos discrepantes no conjunto treinamento foram os pontos referentes à alunos que tinham notas extremas em relação ao restante da turma. O que é perfeitamente admissível, dada a restrição do número de observações no conjunto treinamento. Dentre os alunos que utilizaram o sistema, obteve-se uma baixa quantidade de resoluções, o que implica em um cenário de pequena variação nos possíveis valores de notas. No entanto, como o algoritmo possui parâmetros baseados na média geral das notas, então, qualquer valor muito acima ou muito abaixo do valor médio, torna-se um ponto discrepante na predição.

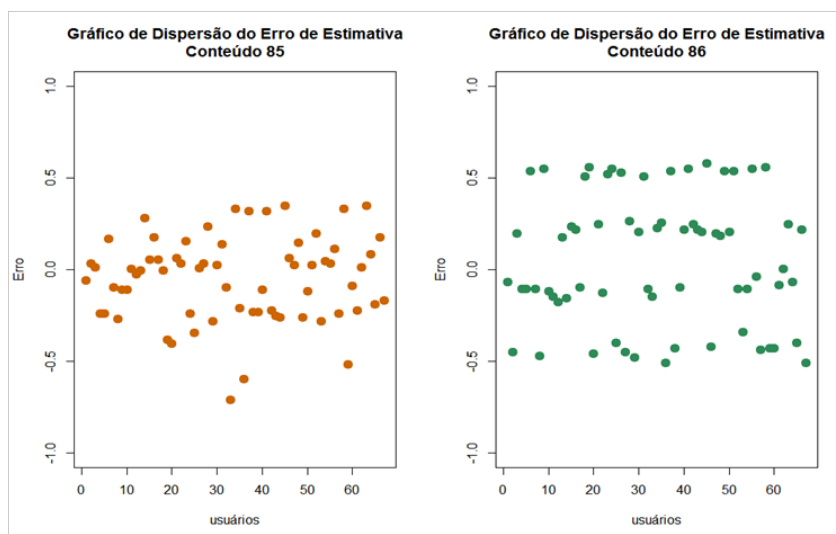
Tabela 25 – Notas observadas dos exercícios (conjunto treinamento).

Conteúdos	Min	1º Qu.	Mediana	Média	3º Qu.	Max.
85	0,051	0,179	0,294	0,336	0,469	0,945
86	0,051	0,329	0,596	0,549	0,799	1

A Tabela 25 demonstra, por meio de estatística básica, como o desempenho dos alunos no Tema 86 foi melhor em relação ao tema 85.

Dando sequência à análise entre as notas previstas e observadas, a Figura 62 exibe gráficos de dispersão referentes ao erro de estimativa dos conteúdos 85 e 86. Tal erro é definido como a diferença entre o valor previsto e observado.

Figura 62 – Dispersão do viés (erro) ao por usuário (Temas 85 e 86).



Fonte – da pesquisa.

Observa-se no gráfico acima (Figura 62) que houve uma dispersão maior do erro para o Tema 86, o que significa que o modelo teve pior resultado preditivo em relação ao Tema 85.

As Tabelas 26 e 27 exibem os resultados via estatística descritiva em relação aos valores previstos e observados na bimestral.

Tabela 26 – Estimativa do modelo e resultado efetivo na bimestral (Temas 85).

TEMA 85						
Notas	Min	1º Qu.	Mediana	Média	3º Qu.	Max.
SVD	0,26	0,31	0,32	0,33	0,34	0,44
BIMEST.	0	0,29	0,29	0,38	0,57	1

Tabela 27 – Estimativa do modelo e resultado efetivo na bimestral (Temas 86).

TEMA 86						
Notas	Min	1º Qu.	Mediana	Média	3º Qu.	Max.
SVD	0,49	0,54	0,55	0,56	0,57	0,67
BIMEST.	0	0,33	0,67	0,5	0,67	1

Ressalte-se que as análises apresentadas estão sendo restritas aos Temas 85 e 86, em razão de que foram os únicos conteúdos abordados pela avaliação bimestral.

No intuito de comparar as previsões de desempenho do SVD com os valores obser-

vados na avaliação bimestral, foram realizados testes paramétricos e não paramétricos para testar a igualdade dos dois grupos. Temos então a seguinte organização de grupos:

- SVD : grupo com as previsões de notas a serem obtidas pelos alunos na avaliação bimestral.
- BIM : grupo com desempenho real obtido pelos alunos na avaliação bimestral.

Uma forma de comparar dois grupos de notas é considerá-los como duas amostras específicas e testar se essas duas amostras advém da mesma população, ou seja, testar as seguintes hipóteses:

$H_0$  : População<sub>SVD</sub> = População<sub>BIM</sub> (As amostras SVD e BIM provém da mesma população).

$H_1$  : População<sub>SVD</sub>  $\neq$  População<sub>BIM</sub> (As amostras SVD e BIM não provém da mesma população).

Para comparação das duas amostras é comum o uso do Teste T de Student-Pareado para comparação das médias dos dois grupos. No entanto, esse teste paramétrico requer que as amostras tenham distribuição Normal. Analisando os dados, percebe-se que tal condição não é atendida. Assim, as amostras SVD e BIM, tanto do conteúdo 85 quanto do conteúdo 86, não provém de uma distribuição Normal ao nível de significância de 5%.

Tabela 28 – Teste de Normalidade.

Teste	Conteúdo 85		Conteúdo 86	
	SVD	BIM	SVD	BIM
	p-valor		p-valor	
Teste de Normalidade shapiro.test	0,004729	0,005434	0,003704	0,0000039

Uma alternativa ao uso do Teste T Student Pareado é o Teste não paramétrico de Wilcoxon. Tal teste permite analisar se as duas amostras (SVD e BIM) provém da mesma população sem precisar assumir que elas sigam a distribuição normal. Uma vez que as amostras pertençam à mesma população, podemos considerar que os valores das notas previstas pelo SVD<sup>++</sup> são estatisticamente iguais aos valores observados na avaliação, indicando, assim, bom desempenho preditivo.

Aplicando o teste, não rejeitou-se a hipótese  $H_0$  de que as amostras SVD e BIM provém da mesma população, tanto no conteúdo 85 quanto no conteúdo 86, com um nível de significância de 5%. A Tabela 29 informa o valor- $p$  obtido no teste para os dois temas:

Tabela 29 – Teste Não Paramétrico Wilcoxon.

Teste	Conteúdo 85	Conteúdo 86
	p-valor	p-valor
Teste Não Paramétrico -Wilcoxon: wilcox.test.	0,1552	0,08864

Com isso, ao nível de significância de 5%, pode-se concluir que as estimativas de desempenho feitas pelo algoritmo SVD<sup>++</sup> são iguais aos valores observados na avaliação, ou seja, o algoritmo apresenta bom desempenho preditivo.

A Tabela 30 apresenta o expressivo resultado do Modelo SVD<sup>++</sup> para a previsão de notas na avaliação bimestral para os temas 85 e 86.

Tabela 30 – do RMSE (Temas 85 e 86).

TEMAS	RMSE
85	0,05
86	0,12

### 6.1.5.3 *Discussão sobre os modelos preditivos adotados*

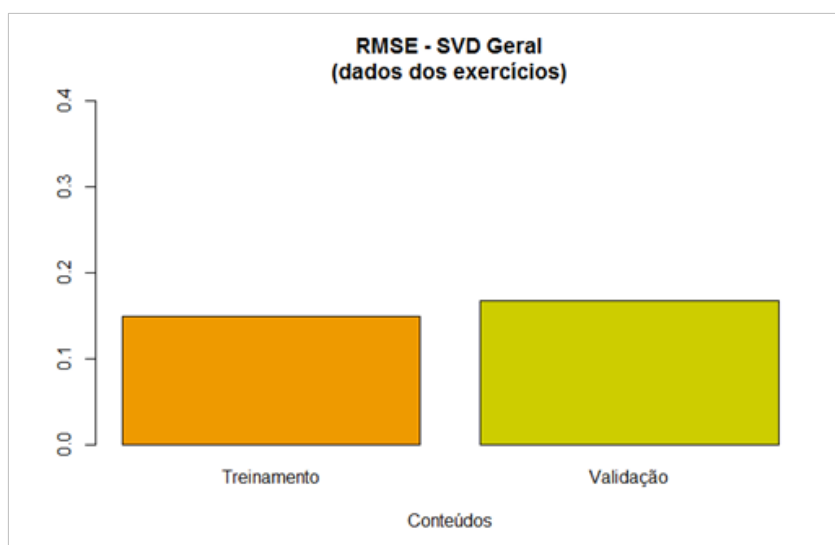
Os modelos preditivos utilizados nesta tese foram escolhidos dadas suas ocorrências frequentes em trabalhos que os utilizam para a finalidade de predição. Para a predição do desempenho de um determinado aluno em um determinado tema, o Regressor Linear Logístico leva em consideração apenas os dados observados em que o aluno respondeu a questões para esse tema específico (fornecendo assim variáveis explicativas), não sendo possível gerar predições para esse mesmo usuário em outro tema, nem sendo possível gerar predições para usuários que não possuam variáveis explicativas no tema em questão. Já o SVD<sup>++</sup> realiza a predição de desempenho para todos os usuários e temas, sendo menos sensível a ausência de informações. Sendo possível assim estabelecer uma predição para qualquer usuário em qualquer tema.

O Modelo Linear Logístico utiliza a informação restrita ao tema para gerar as chances de acerto do estudante, enquanto o  $SVD^{++}$  utiliza a média global, o viés dos estudantes (variação da média do estudante em relação a média global) e o viés dos temas (variação da média do tema em relação a média global). A característica colaborativa do  $SVD^{++}$  mostrou-se bem mais adequada ao presente contexto educacional, uma vez que tem alta precisão mesmo diante de poucos registros.

Analisando as previsões de desempenho obtidas pelo método  $SVD^{++}$ , foram obtidos os valores de RMSE ilustrados no gráfico abaixo para os dados de treinamento e validação, respectivamente.

É importante ressaltar que tais valores foram obtidos comparando as notas observadas no exercícios com os valores estimados pelo método  $SVD^{++}$ . Parte dos valores das notas observadas nos exercícios foram usadas como dados de treinamento e a outra parte restante foi usada como dados de validação.

Figura 63 –  $SVD^{++}$  (Treinamento  $\times$  Validação)

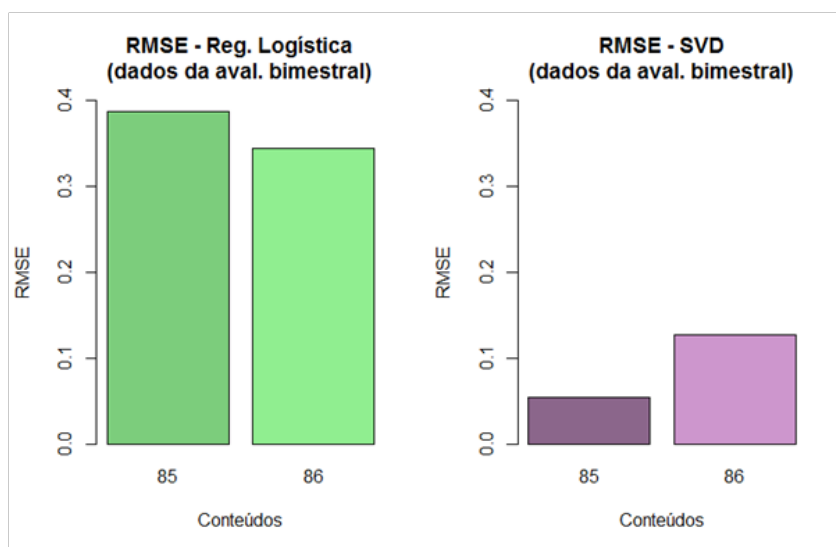


No intuito de comparar o desempenho das estimativas obtidas pelo método da regressão logística com as estimativas obtidas pelo método  $SVD^{++}$ , calculou-se o RMSE de cada método com base no valor obtido pelos alunos na avaliação bimestral (Tabela 31).

Tabela 31 – Valores de RMSE obtidos através das comparação das notas observadas na avaliação bimestral com as previsões fornecidas pelos métodos SVD<sup>++</sup> e Regressão Logística.

Conteúdos	Conteúdo 85	Conteúdo 86
RMSE - SVD	0,0546	0,1265
RMSE – Regressão Logística	0,3867	0,3441

Figura 64 – RMSE Regressor Logístico x SVD<sup>++</sup> (Temas 85 e 86)



Fonte – Fonte: da pesquisa

Analisando a Tabela 31 e a Figura 64 é possível observar que há uma maior eficiência do método SVD<sup>++</sup> para previsão das notas de desempenho dos alunos ao compará-lo com o Modelo de Regressão Logística.

A técnica de regressão logística se ressentida da falta de dados (alunos que não tem nota no determinado tema), pois o tamanho amostral é um importante fator para boas estimativas. Tal método utiliza apenas os dados observados em cada tema de forma individual para fazer a estimativa da previsão de nota. Isso faz com que ocorra um maior viés dos valores preditivos em relação aos observados.

O tamanho amostral usado para fazer as estimativas pelo método da regressão logística foram apenas 36 observações no tema 85 e 39 observações no tema 86.

Por outro lado, o método SVD<sup>++</sup>, contorna o problema de ter uma matriz de dados esparsa, utilizando em cálculo da previsão de desempenho de um determinado aluno a num determinado tema  $b$ :

- A variação ocorrida em relação a média geral do desempenho observado do aluno nos



demais temas;

- A variação ocorrida em relação a média geral do desempenho observado de todos os alunos no tema.

Os pontos acima elencados oferecem um melhor embasamento para as estimativas preditivas, fazendo do  $SVD^{++}$  uma ferramenta mais adequada para a situação de estimativas com dados faltantes e, conseqüentemente, de pequeno tamanho amostral.

O tamanho amostral utilizado para calcular as estimativas pelo método  $SVD^{++}$  foi de 196 valores observados no geral (valores observados nos temas onde houve resolução de exercícios).

Percebe-se, ainda, que o RMSE observado pelo método  $SVD^{++}$  no tema 86 apresentou-se um pouco mais elevado que o RMSE observado no tema 85. Isso deve-se ao fato que esse cálculo foi feito comparando as estimativas com as notas observadas na avaliação bimestral. As notas das estimativas foram calculadas com base nas notas observadas pelos exercícios. Tais notas variam no intervalo de  $[0, 1]$ .

As notas observadas na avaliação bimestral foram obtidas com base na proporção de acertos nas questões do Tema 86 que, por só terem caído 3 questões desse tema na prova, teve notas registradas apenas para os valores  $\{0, 0.33, 0.67, 1\}$ .

Essa diferença de variação do espaço amostral das notas fez com que o RMSE calculado pelo  $SVD^{++}$  para o Tema 86 ficasse mais elevado. Ainda assim, o erro apresenta-se bem menor quando comparado com o RMSE obtido pelo método da regressão logística.

Nghai et al. (2012) apresentam resultados similares ao encontrado neste trabalho, em que uma técnica de fatoração de matrizes tem menor RMSE que um regressor logístico para a predição de desempenho estudantil.

### **6.1.6 Validação da Dimensão Personalização**

Avaliando o quesito personalização, o sistema possui um plano de estudos personalizado de acordo com o perfil de cada aluno (ritmo de estudo, nível de aprendizado, quantidade de exercícios e temas abordados, entre outros).

Dos 75 alunos analisados, pode-se perceber que 45 destes só resolveram questões de nível básico e os 30 restantes evoluíram no plano de estudos indo para os níveis intermediário e avançado. Definindo, assim os grupos:

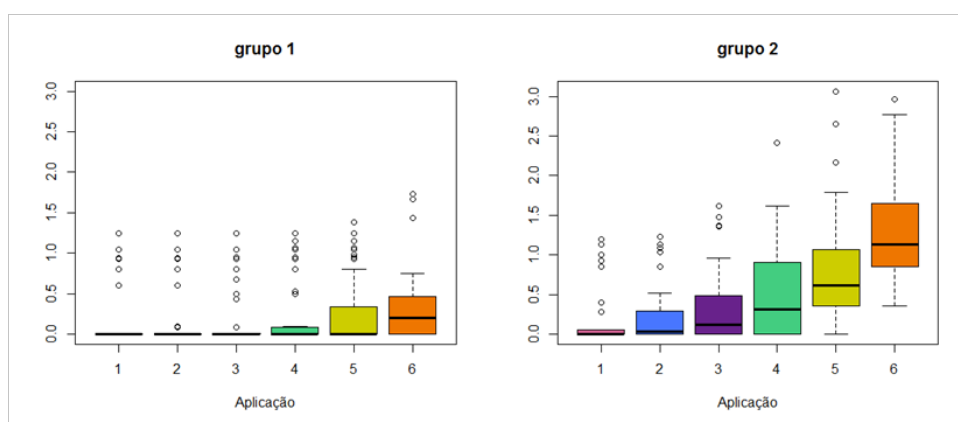
- Grupo 1: Alunos que não evoluíram no plano de estudos, resolvendo apenas questões de

nível básico.

- Grupo 2: Alunos que evoluíram no plano de estudos, resolvendo questões de nível básico e avançado.

Em cada grupo e aplicação, foi calculado para cada aluno o somatório das notas obtidas por esse aluno em cada um dos temas abordados por este, tendo assim uma nota geral para esse aluno em cada aplicação. Obteve-se, assim, os gráficos *boxplots* abaixo com as notas gerais de cada aluno em cada aplicação.

Figura 65 – Evolução de notas de cada grupo de alunos ao longo das aplicações.



Fonte – da pesquisa

Percebe-se que, em geral, o somatório das notas dos alunos que evoluíram no plano de estudos (Grupo 2), são maiores que o dos alunos que não evoluíram (Grupo 1). Isso é esperado, pois quanto mais se avança nos níveis intermediário e avançado, mais chances o aluno tem de aumentar sua nota.

De forma mais detalhada, as Figuras 66 e 67 contêm o plano de estudos sugeridos para dois alunos (ID 67 e 94).

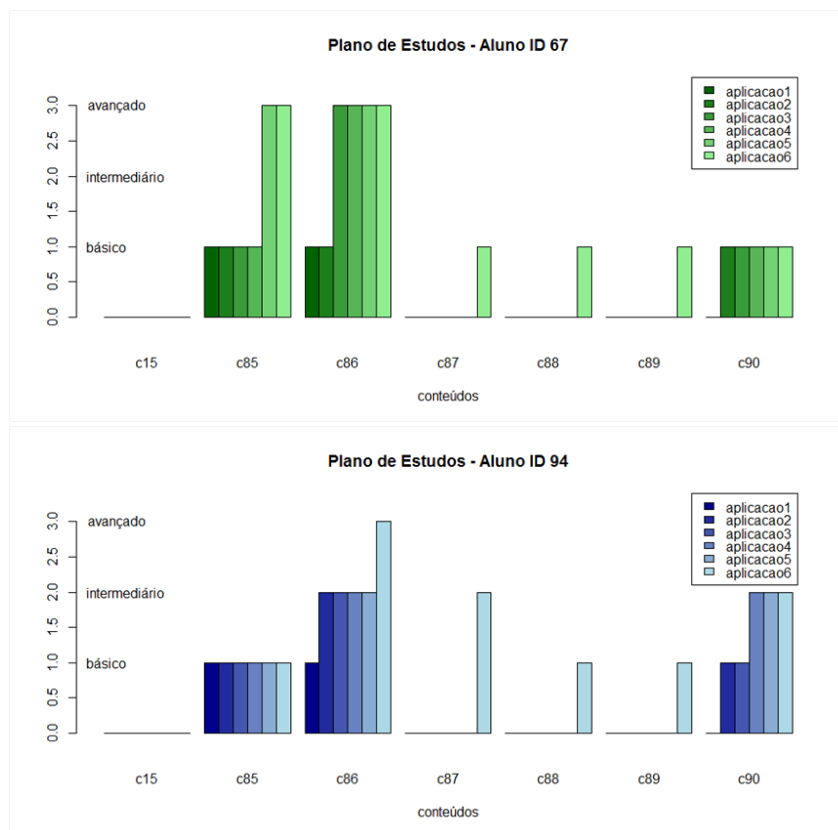
No eixo *x*, cada gráfico possui 7 blocos, que foram 7 temas abordados por esses alunos. Em cada bloco tem-se 6 colunas, que representam as 6 aplicações.

No eixo *y*, tem-se as alturas que representam os níveis de estudo em cada tema.

Sendo:

- 0: Não há plano de estudo para esse aluno.
- 1: Iniciar nível básico.
- 2: Iniciar nível intermediário.
- 3: Iniciar nível avançado.

Figura 66 – Planos de estudo (Aluno 67 e 94).



Fonte – da pesquisa

Esses dois exemplos ilustram a personalização que o sistema faz para cada aluno no plano de estudos a ser seguido.

Entre outras diferenças, percebe-se que no conteúdo 86, o plano de estudos do aluno de ID 67 já sugeria que este iniciasse o nível avançado na data da terceira aplicação, enquanto que para o aluno ID 94, o nível avançado só foi sugerido na data da sexta aplicação.

A comparação entre esses dois alunos quanto à seus planos de estudo, demonstra que a Arquitetura efetivamente implementou corretamente o requisito conceitual de personalização definido pelo Modelo *Digital Class*. Com essa última análise, todas as dimensões previstas pelo Modelo *Digital Class* foram efetivamente validadas por meio da sua arquitetura, cumprindo com o propósito deste trabalho.

## 6.2 Discussão do instrumento de satisfação aplicado aos alunos

Ao final da sexta e última sessão em laboratório, foi aplicado um questionário em que se pretendia auferir a receptividade dos estudantes à solução tecnológica a qual eles tiveram acesso nas últimas semanas, ao longo do período de aplicação do experimento. O instrumento

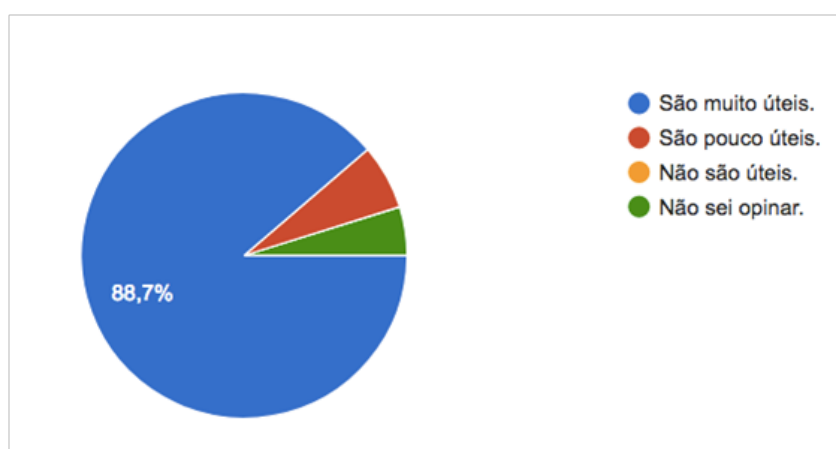
contava com 15 questões, sendo 13 delas de natureza objetiva e 2 subjetivas. Nas duas questões iniciais, os alunos identificaram a qual curso pertencia e qual seu sexo. Responderam ao questionário 62 alunos, sendo 36 da turma de Produção de Áudio e Vídeo (PAV) e 26 de Informática.

Dos respondentes, 67,7% (42) foram do sexo feminino e 32,3% (20) do sexo masculino.

As demais questões foram destinadas à obtenção de informações sobre a utilidade da ferramenta, como sua usabilidade, dinâmica de apresentação de conteúdos, dentre outros. Dentre as duas questões subjetivas, uma delas foi destinada para as sugestões dos cursistas.

**Questão 3: Como você avalia a utilidade do uso de ferramentas tecnológicas na escola (por exemplo: *Digital Class* e *Khan Academy*)?**

Figura 67 – Visão dos alunos sobre a utilidade de ferramentas tecnológicas na escola.



Fonte – da pesquisa.

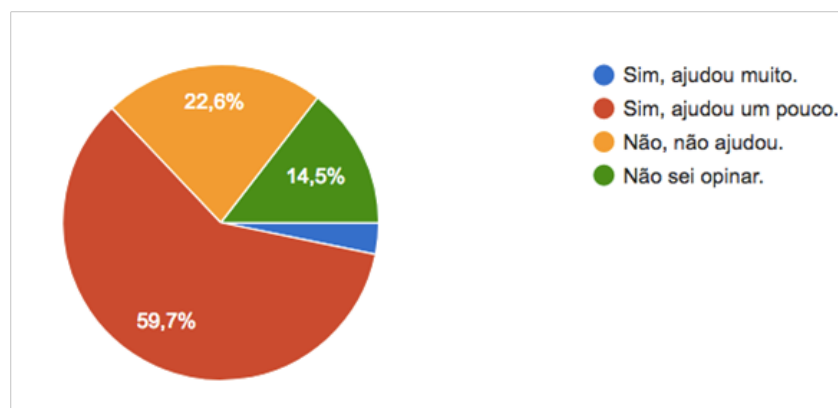
O objetivo desta questão foi identificar a receptividade dos alunos quanto ao uso de ferramentas computacionais na escola. Um grupo expressivo de alunos (88,7%), apontou que as ferramentas são úteis. Apenas 6,5% informaram que são pouco úteis e 4,8% não souberam opinar. Não houve respostas para a opção “não são úteis”. Estes resultados mostram que a grande maioria do grupo de alunos participantes deste estudo é plenamente aberto à inserção de práticas inovadoras na escola mediada por tecnologias, como os ambientes virtuais de aprendizagem.

A receptividade do corpo discente à inserção de inovações nas práticas de ensino

e aprendizagem é uma condição importante para que qualquer inovação possa se desenvolver, como por exemplo, a *Digital Class*. Um dos fatores que podem gerar resistências à inserção de novas práticas, seria a falta de compreensão dos alunos sobre os propósitos e benefícios buscados por essas mudanças. Visando minorar tal problemática, no presente experimento, foram feitas explanações iniciais que mostravam aos alunos a importância do ensino híbrido e suas potencialidades, o que pode ter contribuído para a elevada receptividade demonstrada pelos alunos.

#### Questão 4: A *Digital Class* ajudou na organização das suas atividades na disciplina?

Figura 68 – Auxílio da *Digital Class* na organização das atividades.

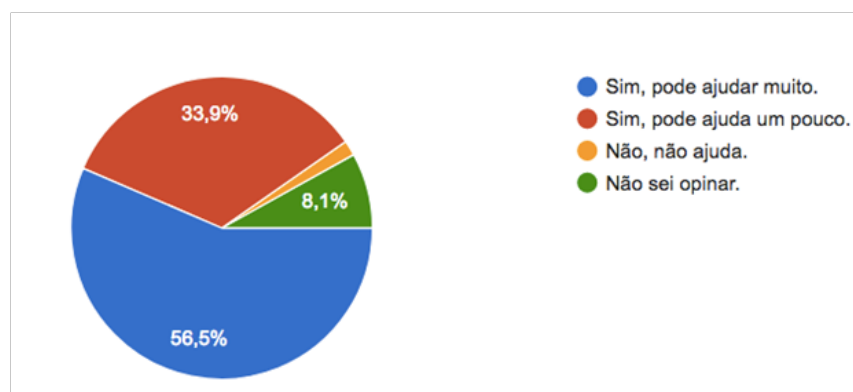


Fonte – da pesquisa.

As respostas positivas para essa questão foram de 62,9% (3,2% Sim, ajudou muito e 59,7% sim, ajudou um pouco). O percentual que afirmou que não ajudou foi de 22,6%, e 14,5% não souberam afirmar. Uma das finalidades da Arquitetura *Digital Class* é oferecer ao aluno um guia de estudos para que este possa saber, a todo o instante, quais passos são mais importantes cumprir de acordo com o planejamento do curso feito pelo professor e qual o seu estado de aprendizado. Esta orientação se dá pelos planos de estudos gerados continuamente para cada cursista. Desta forma, o aluno não fica com o sentimento de estar sozinho, perdido ou mesmo, deixado para trás, pois ele recebe do sistema importante apoio para o cumprimento dos objetivos traçados para o curso.

#### Questão 5: Você considera que o uso da *Digital Class* pode auxiliar no reforço de aprendizado de temas de uma disciplina?

Figura 69 – Auxílio ao Aprendizado



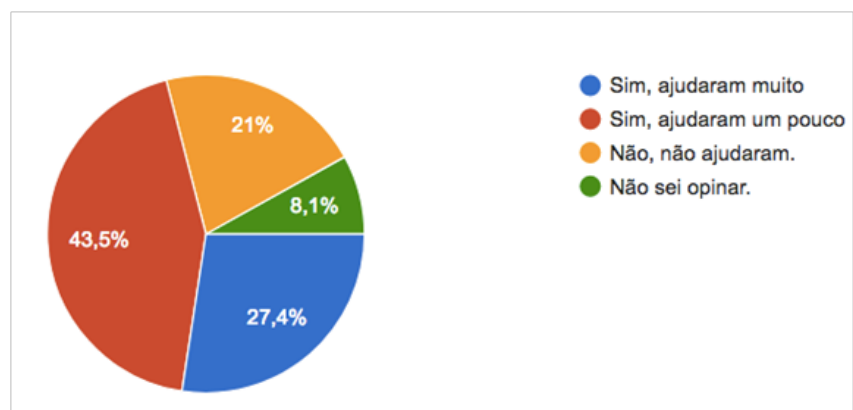
Fonte – da pesquisa

Para 90,4% (56,5% acham que pode ajudar muito e 33,9% acham que pode ajudar um pouco) dos alunos, a *Digital Class* tem potencial para auxiliar no reforço de assuntos para a construção do aprendizado. Apenas 1,6% considera que não ajuda, enquanto 8,1% não souberam opinar.

O reconhecimento, pela grande maioria dos alunos, da *Digital Class* como uma oportuna ferramenta de apoio ao aprendizado, enaltece o sucesso da proposta, pois esse é sem dúvida um dos propósitos da solução.

**Questão 6: As questões recomendadas para você na *Digital Class* ajudaram no seu aprendizado?**

Figura 70 – Auxílio ao Aprendizado



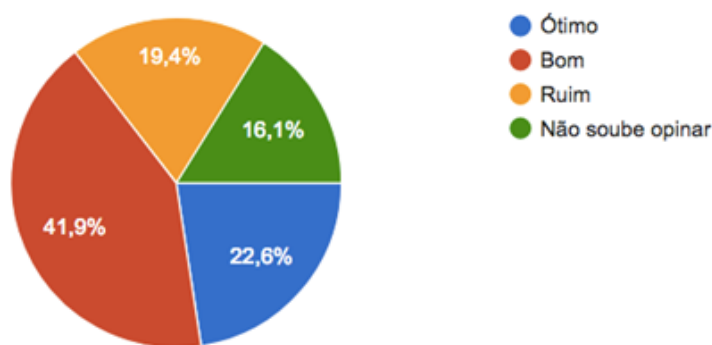
Fonte – da pesquisa

Para a maioria dos cursistas, 70,9% (27,4% informaram que sim, as questões ajudaram muito em seu aprendizado e 43,5% que sim, as questões ajudaram um pouco), as questões contribuíram para o aprendizado. Para 21% não foi percebida qualquer ajuda e 8,1% não souberam opinar.

Na Seção 6.1.2 a avaliação contínua foi discutida. Esta avaliação é composta, justamente, de exercícios (questões). Foi verificado que quanto maior a interação dos estudantes na resolução de questões, maiores eram os índices de desempenho dos alunos. Os resultados do presente item do questionário, portanto, se alinham ao já observado na discussão de avaliação contínua, ou seja, eles demonstram que uma maior resolução no número de questões está associada à obtenção de um maior aprendizado.

**Questão 7: Como você avalia a quantidade de questões recomendadas pela *Digital Class* para o seu aprendizado?**

Figura 71 – Opinião do aluno com relação à quantidade de questões recomendadas pela *Digital Class*.



Fonte – da pesquisa.

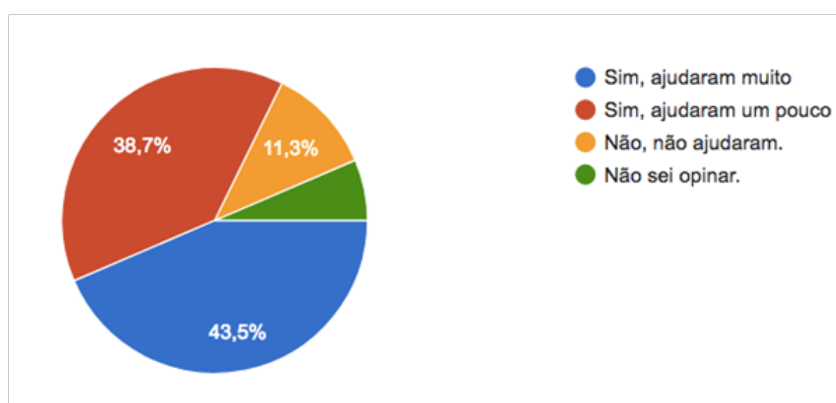
Para 22,6% o número de questões atendeu à necessidade de aprendizado (ótimo). Para 41,9% a quantidade de questões utilizada foi boa. Já para 19,4% o número de questões foi inadequado (ruim). Não souberam opinar ficou em 16,1%. Esses dados são importantes, pois a limitação de itens em uma plataforma de BL pode frustrar o desejo de um aluno em se aprofundar em um assunto.

As questões aplicadas aos alunos foram todas cadastradas pelo professor. Cada tema

continha três níveis de complexidade para a questão (básico, intermediário e avançado). Foi solicitado ao professor que, para cada nível de um tema, fossem cadastradas, ao menos, três questões. Então, em média, cada tema possuía nove itens a serem respondidos pelos alunos. Com as reedições de um curso, a tendência é que se consolide um bom acervo de questões, o que irá amortizar o trabalho do professor ao longo do tempo. Além disso, o módulo *Sphinx* que armazena e indexa essas aulas é colaborativo e, com o tempo, pode tornar-se um grande repositório de questões nas mais variadas áreas de conhecimento, o que possibilitará aos alunos uma oferta mais diversificada de itens.

**Questão 8: Os conteúdos dos vídeos recomendados para você na *Digital Class* ajudaram no seu aprendizado?**

Figura 72 – Utilidade dos vídeos para o aprendizado.



Fonte – da pesquisa.

Para 82,2% dos alunos, os vídeos disponibilizados pela *Digital Class* foram úteis no aprendizado, sendo que 43,5% consideraram que ajudaram muito e 38,7% que ajudaram um pouco. Já para 11,3% dos alunos, os vídeos não ofereceram nenhuma ajuda e 6,5% não souberam opinar.

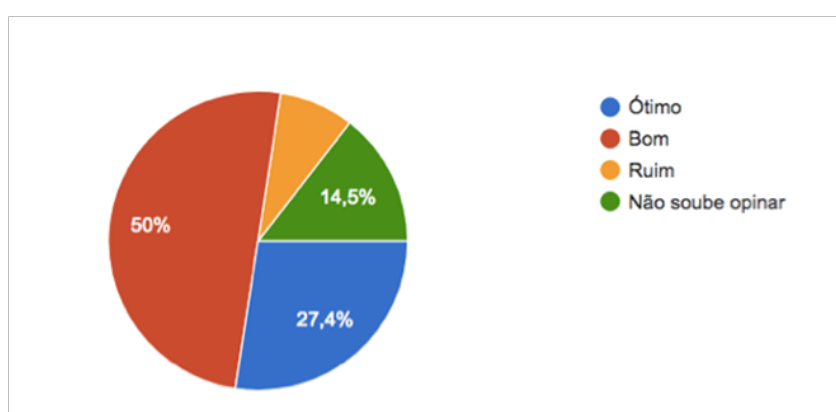
Uma outra fonte de informação que se coaduna com a aprovação diretamente apontada pelos alunos neste item, foi apresentada na Tabela 10. Nela, observa-se que as médias de avaliações feitas pelos estudantes aos vídeos recomendados foram muito positivas. Os vídeos cadastrados na plataforma (arquivos enviados ao servidor da aplicação) ficaram com média de avaliação 3,6 (escala de 1 a 5), enquanto os referenciados no Sistema (publicados externamente), ficaram com 4,06.



O resultado deste item destaca dois pontos positivos. O primeiro, diz respeito a importância dos vídeos como instrumentos de auxílio ao processo cognitivo do aluno. O segundo, evidencia a eficiência do sistema de recomendação da *Digital Class*, o qual sugeriu vídeos aos alunos de forma adequada. Tal constatação, indica a professores a importância de ampliar o uso desta mídia em seus cursos. Em relação aos alunos, é oportuno alertá-los para a importância de avaliar os objetos recomendados, pois este *feedback* é de suma importância para que, através desta ação colaborativa, o sistema venha a fazer cada vez mais recomendações apropriadas. O Professor utilizou vídeos em que ele já possuía em seu acervo pessoal, publicando-os na ferramenta, e complementou com referências externas a portais multimídia (ex.: *Youtube*). De toda forma, foi feita uma criteriosa seleção deste material de forma a fazê-lo completamente aderente aos objetivos do curso. Em futuras edições, espera-se que o professor possa produzir suas próprias aulas no módulo Chimera, complementando o material hoje disponibilizado e reduzindo a necessidade de aulas expositivas presenciais, para assim destinar maior tempo para atividades de projeto, atendimento personalizado, etc.

**Questão 9: Como você avalia a quantidade dos vídeos recomendados pela *Digital Class* para o seu aprendizado?**

Figura 73 – Visão do aluno em relação à quantidade de vídeos disponibilizados.



Fonte – da pesquisa

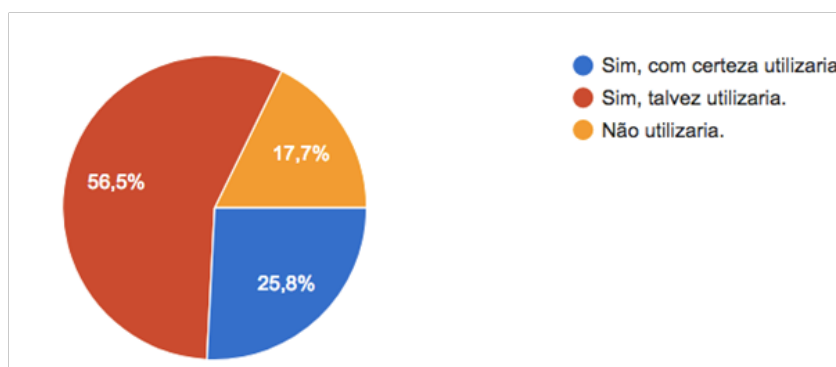
Para 77,4% a quantidade de vídeos foi adequada. Para 8,1% não foi adequada. E 14,5% não souberam opinar.

Os índices obtidos são coerentes com a questão anterior, em que cerca de 80% dos

alunos apontou como positiva a utilização dos vídeos na Arquitetura.

**Questão 10: Você utilizaria a *Digital Class* fora da escola (casa, *lan house*, celular, etc.)?**

Figura 74 – Utilização da *Digital Class* fora do ambiente escolar.



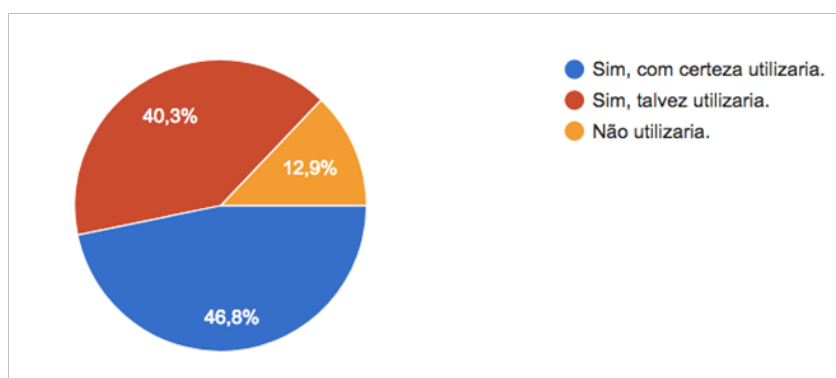
Fonte – da pesquisa

Nesta questão, 82,3% dos estudantes responderam que Sim, utilizariam a *Digital Class* em espaços fora da escola, enquanto 17,7% responderam que não utilizariam em outros espaços que não o escolar. Dos 82,3%, 56,5% disseram que Sim, porém com possíveis condicionantes que podem estar ligados a vários fatores, como a necessidade de viabilização de acesso a Internet. Essa é uma informação importante, pois indica à escola a possibilidade de realizar um trabalho mais estruturado de ensino híbrido a partir do uso da arquitetura. Diante dessa possibilidade de trabalho para momentos extra sala de aula, o Professor poderá realizar em sua disciplina um planejamento pedagógico em que reduza a necessidade de aulas expositivas. Para isso, o Professor poderá realizar gravações de vídeo-aulas e disponibilizá-las no sistema para que os alunos possam assistir às aulas antes do encontro presencial (sala de aula invertida). O resultado desta questão se sobressai ao encontrado por Jia et al. (2012), em que 30% dos alunos que participaram de seu experimento utilizando o AVA Moodle, disseram que não utilizariam a plataforma fora da aula.

**Questão 11: Você utilizaria a *Digital Class* como revisão para provas?**

Para 87,1 a Arquitetura *Digital Class* pode ser utilizada para a revisão de temas para uma prova/avaliação. A minoria, 12,9%, informou que não faria uso da ferramenta para esse

Figura 75 – Utilização da *Digital Class* para revisão de conteúdos.

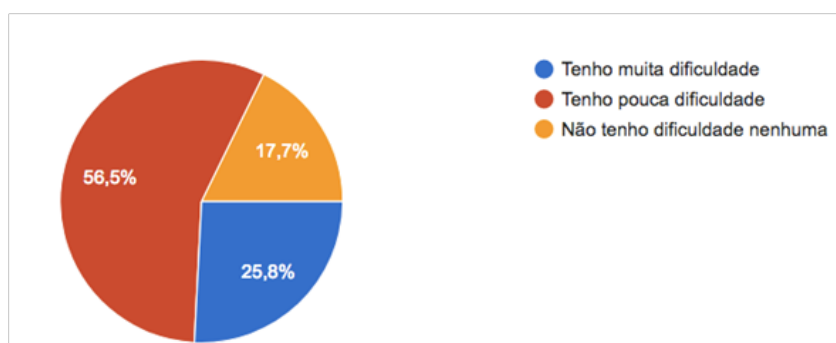


Fonte – da pesquisa.

propósito. Assim como já indicado pelas respostas dadas para a questão 5, os resultados da presente questão sugerem que os alunos reconhecem na solução *Digital Class* um instrumento de apoio efetivo para seus estudos, sendo efetiva para o reforço do aprendizado e preparação para avaliações.

#### Questão 12: Qual o seu grau de dificuldade em utilizar a plataforma?

Figura 76 – Dificuldade de utilização da plataforma.



Fonte – da pesquisa.

Para 25,8% dos alunos, a arquitetura impôs muitas dificuldades de utilização. Já 56,5% afirmaram que tiveram pouca dificuldade. Para 17,7% não foi verificada nenhuma dificuldade de utilização. Boa parte dos problemas observados durante a aplicação, e que podem ter contribuído para elevação dos índices dos que apontam alguma dificuldade, foram de ordem de acesso (problema apontado pelos alunos na questão 15). Vale ressaltar que, em relação a esse problema de acesso, durante a aplicação foi comum observar que alguns alunos esqueciam

suas senhas, ou confundiam com as de outros sistemas da escola. As respostas acima sinalizam que a maioria dos interagentes conseguiu utilizar a arquitetura sem maiores dificuldades. No entanto,  $\frac{1}{4}$  dos respondentes apontaram muita dificuldade de uso. Como forma de melhorar esse índice, cabe, no futuro, uma análise pormenorizada sobre todos os aspectos que impactam nesse resultado.

**Questão 13: Na sua opinião, qual o objetivo do uso da *Digital Class*?**

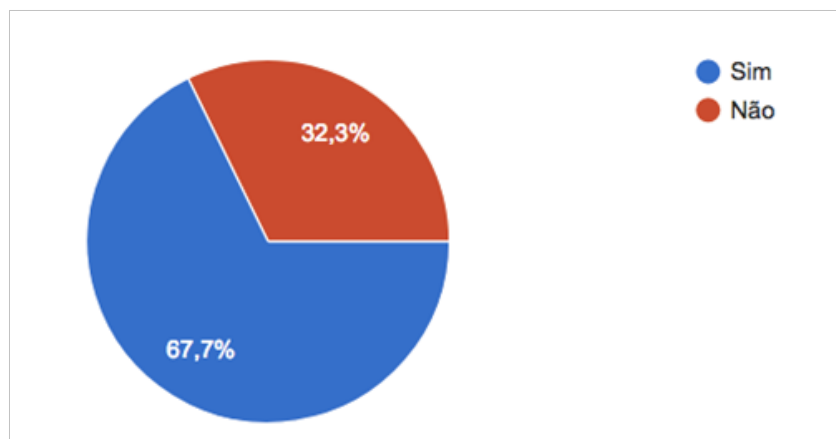
Foram dadas 62 respostas pelos alunos. Dada a similaridade de algumas respostas, elas foram agrupadas e são listadas a seguir:

- Promover a melhoria do aprendizado do aluno;
- Reforçar (apoiar) o aprendizado no seu curso;
- Disponibilizar uma ferramenta para ensino e revisão de conteúdos;
- Oferecer um aprendizado mais evoluído fugindo da repetição tradicional de sala de aula;
- Auxiliar os alunos em disciplinas de maneira geral;
- Levar o aprendizado também para momentos fora da escola;
- Atrair o interesse dos alunos nos estudos por meio da tecnologia digital;
- Permitir o reforço e/ou o aprofundamento de conhecimentos;
- Melhorar o desempenho escolar;
- Orientar os alunos sobre como aprender física;
- Oferecer uma nova forma de aprender;
- Orientar o aluno dentro da disciplina;
- Reforçar assuntos vistos em sala de aula;
- Orientar os alunos complementando o trabalho do professor.

As respostas apresentadas apontam para uma compreensão positiva dos objetivos do experimento. A percepção por parte dos alunos de que a solução *Digital Class* caracteriza-se como uma ação de melhoramento do processo de ensino e aprendizagem, oferecendo um maior apoio didático em um curso, em formato inovador, por meio de novas tecnologias, e ampliando o espaço de aprendizado para além da escola, mostram que os alunos compreenderam o propósito da solução.

**Questão 14: Você gostaria que outras disciplinas utilizassem a *Digital Class*?**

Figura 77 – Recomendação de adoção da *Digital Class* a outras disciplinas



Fonte – Fonte: da pesquisa.

Nesta questão, 67,7% dos respondentes disseram que gostariam de fazer uso da *Digital Class* em outras disciplinas e 32,3% disseram que não. Mesmo tratando-se de uma tecnologia ainda em validação, o resultado obtido nesta questão (quase 70% de aprovação) mostra um alto nível de aceitação, que, somado aos outros indicadores já discutidos, sinalizam para um elevado potencial de sucesso da ferramenta.

#### **Questão 15: Quais sugestões você daria para a melhoria da *Digital Class*?**

Foram oferecidas 56 respostas à essa questão. Assim como na questão 13, as respostas foram agrupadas nas seguintes sugestões:

- Resolução de problemas de acesso;
- Redução da complexidade das questões;
- Ampliação do uso da ferramenta por outras disciplinas como química, biologia e o tema de lógica matemática;
- Melhorias nos gráficos indicativos de desempenho do aluno;
- Vincular a cada erro de questão um vídeo específico;
- Disponibilização de mais vídeos com legendas;
- Aumentar a diversidade de questões e vídeos;
- Ampliar a personalização para que o aluno tenha um avatar próprio;

- Melhorar o layout e por mais cores;
- Permitir personalização do sistema (imagens, menus, etc.);
- Disponibilizar dicas;
- Tornar o sistema mais gamificado (com pontuações, pontos de conquista, etc.);
- Local de reclamações e sugestões;
- Espaço para desenhar gráficos;
- Manter a simplicidade e objetividade do sistema.

As respostas apontadas pelos alunos foram de suma importância, ao indicar o que se precisa fazer para atingir um maior índice de satisfação dos estudantes, afinal, é sobretudo para este grupo que a metodologia e, conseqüentemente, a plataforma foram desenvolvidos. As colocações dos estudantes são coerentes e de baixo grau de complexidade para que sejam resolvidas em um curto período de tempo. No entanto, essas melhorias não serão observadas no escopo desta tese.

## 7 CONSIDERAÇÕES FINAIS

Nesta pesquisa, foi proposto o Modelo *Digital Class* com o objetivo de oferecer uma solução para a modelagem de cursos voltados à metodologia de ensino híbrido. O modelo proposto foi caracterizado por seis dimensões: adequabilidade aos objetivos do curso, avaliação contínua, paralelização dos estudos, predição de desempenho, ritmo próprio e personalização de atendimento.

Foi definida inicialmente nesta tese a seguinte hipótese: *é possível atingir um estado de satisfação de professores e alunos de um curso, em formato de ensino híbrido, através da oferta de uma solução metodológica, instrumentada por tecnologia, que respeita os objetivos do curso e oferece aos alunos uma condição de personalização de estudos, com apoio automático para a sua preparação, planos de estudos individualizados e a garantia de um estudo com ritmo próprio, além de auxiliar o professor na avaliação contínua e previsão de desempenho de seus alunos.*

Esta hipótese foi comprovada integralmente a partir das validações realizadas. Foram definidas uma questão principal e quatro outras questões de pesquisa.

**Questão principal:** Qual o estado da arte de modelos de *blended learning*?

Essa questão foi abordada nas seções 3.2 e 3.3 em que diversos modelos de BL foram apresentados. Estes modelos são bem variados, pois não há nesta área de investigação uma padronização entre as estruturas que devem compor uma iniciativa em BL. Portanto, os estudos mostraram diversas abordagens metodológicas voltadas a cursos híbridos. Tais estudos auxiliaram esse trabalho quanto ao mapeamento de requisitos que fossem verdadeiramente inovadores à área de pesquisa de BL.

**QP01:** De que forma a tecnologia tem propiciado a implantação de inovações metodológicas para a melhoria do processo de ensino e aprendizagem em cenários de *Blended Learning*?

O processo de revisão sistemática realizado neste trabalho possibilitou mapear modelos atuais de BL e com isso nortear os elementos que vieram a compor as dimensões do modelo teórico de ensino híbrido proposto por este trabalho. Na seção 3.3.1, foram observados trabalhos que promovem o atendimento personalizado do aluno. Tais trabalhos fazem uso de avaliação

formativa, que promovem predição de desempenho estudantil, etc.

**QP02:** Quais aspectos do campo educacional podem ser diretamente trabalhados e aprimorados pela utilização de TEL no contexto de *Blended Learning*?

Na seção 3.3.3 são apresentadas questões que são tratadas por diversos autores que vem trabalhando na temática de TEL e que podem contribuir para a melhoria das experiências de BL. Foram citados trabalhos voltados ao processo de aprendizado personalizado, de avaliação do conhecimento em tempo real da turma, assim como de geração de modelos preditivos sobre o desempenho acadêmico do aluno.

**QP03:** Como viabilizar um processo de ensino e aprendizagem personalizado?

Na seção 3.3.4, foi discutida mais uma perspectiva de promoção de personalização, que é o respeito o ritmo do aluno. Esta possibilidade respeita a heterogeneidade das turmas de alunos e, mesmo não sendo um tema facilmente encontrado na literatura de TEL, é extremamente pertinente ao tema de BL. Em razão disto, esse elemento está presente na proposta do Modelo *Digital Class*. Os resultados foram apresentados e discutidos na seção 6.1.4. Os gráficos desta seção mostram a ocorrência de ritmos diversos dos alunos no uso da solução proposta. Também evidenciam de que forma isso foi oportuno para que um aluno, mesmo tendo iniciado o curso com atraso, não se prejudique ao final.

**QP04:** Como atender aos principais requisitos necessários à implantação de uma solução tecnológica de Blended Learning?

O levantamento bibliográfico realizado neste trabalho mostrou uma rica diversidade de trabalhos sobre as mais diversas finalidades. No entanto, não há em uma mesma solução a reunião dos principais assuntos de pesquisa observados em TEL. O Modelo *Digital Class* define um conjunto de dimensões que figuram, seguramente, na vanguarda de TEL. Desta forma, tanto o modelo conceitual quanto o tecnológico (arquitetura) consistem em uma solução integrada em que são reunidas relevantes dimensões que atuam na linha de frente das pesquisas em TEL. O posicionamento teórico desta solução frente ao estado da arte de TEL e BL, a definição teórica



de um modelo estruturado, a criação de uma plataforma tecnológica e a validação desta solução em um cenário real de um curso de ensino médio consistem nas principais contribuições desta tese à área de BL.

Para validar a pertinência de cada uma das dimensões citadas na hipótese, foi desenvolvida uma arquitetura computacional em que os aspectos teóricos definidos no modelo foram transformados em funcionalidades presentes na Arquitetura. O desenvolvimento dos módulos da Arquitetura correspondeu a um primeiro estágio de validação do Modelo *Digital Class*.

Dando seguimento ao processo de validação, um segundo estágio foi cumprido a partir da aplicação do modelo em um curso real, utilizando os fundamentos previstos pela solução. Desta forma, um estudo prático foi realizado com duas turmas do primeiro ano de Ensino Médio de uma escola pública profissionalizante.

Os resultados desta segunda etapa foram discutidos em três momentos. O primeiro buscou analisar os dados de interação dos alunos com a arquitetura. Foram realizados testes estatísticos não-paramétricos para identificar a significância estatística de alguns dados apresentados na discussão de cada dimensão. O segundo verificou diretamente com os alunos partícipes do experimento qual a satisfação dos mesmos em relação às diversas funcionalidades oferecidas pela Arquitetura. Para isto, foi utilizado um questionário contendo quinze questões para colher as impressões do corpo discente. Os dois momentos acima descritos foram discutidos e analisados no Capítulo 6.

Os resultados obtidos validam a proposta do modelo conceitual e, conseqüentemente, todas as suas dimensões que, reunidas, fortalecem uma das premissas de ensino híbrido que é a personalização do processo de ensino-aprendizagem. Mesmo colocando esta premissa em evidência, muitos teóricos de BL não estabelecem de que forma pode se dar essa personalização. No caso da realidade brasileira, um professor via de regra possui exaustiva carga-horária e turmas com alto número de alunos. Desta maneira, as ferramentas computacionais que materializaram os conceitos estabelecidos pela proposta obtiveram dados expressivos em termos de aceitação pelos usuários e pela contribuição para que os alunos obtivessem melhores rendimentos acadêmicos.

A arquitetura computacional desenvolvida foi exitosa em validar as dimensões definidas pelo modelo. Com isso, tem-se uma robusta ferramenta tecnológica (plataforma virtual de aprendizagem) que oferece a professores e alunos um conjunto de funcionalidades peculiares em apoio a realização de um curso em ensino híbrido.

A implementação de ferramentas automáticas, como sistemas de recomendação, reforçou o atendimento ao aluno durante o semestre. O professor atuou de forma associada à solução proposta, cadastrando o planejamento do curso, com datas, atividades e materiais didáticos. Coube aos módulos de recomendação identificar os momentos de agir, recomendando conteúdos específicos, exibindo sequências de exercícios, redefinindo planos de estudos, tudo isso em razão do aprendizado individual de cada aluno, respeitando sempre os objetivos e prioridades do curso.

O processo de realização de exercícios/questões do sistema mostrou-se fortemente associado aos melhores rendimentos obtidos pelos cursistas. Esses mesmos alunos reconheceram a solução *Digital Class* como apropriada ao apoio ao processo de ensino e aprendizagem, considerando ainda a importância de sua adoção por outras disciplinas de seu curso.

Dentre as principais inovações trazidas pelo trabalho ao contexto de EH, está o módulo de predição de desempenho. Foram desenvolvidos dois modelos para previsão de desempenho estudantil. Estes modelos são importantes para oferecer ao docente uma condição privilegiada de análise de seus alunos ao longo de um curso. Dentre os modelos, o  $SVD_{++}$  teve desempenho muito positivo. Já o modelo Regressor Linear Logístico não apresentou um bom resultado, uma vez que os dados para treinamento foram limitados. O  $SVD_{++}$ , dada sua característica colaborativa, possui maior robustez para lidar com dados faltantes e, por isso, foi mais eficiente na tarefa de prever o desempenho do estudante.

A experiência foi exitosa, apesar dos problemas inerentes a uma experimentação piloto, que foram de pequena ordem e não comprometeram o planejamento da validação da pesquisa. Professores e alunos manifestaram aprovação à solução proposta neste trabalho, seja de forma explícita (questionário e entrevista), seja de forma implícita (*logs* de interação em banco de dados).

Em razão da validação realizada, as soluções produzidas neste trabalho, desde o modelo conceitual até a arquitetura (*softwares*), podem ser adotadas por outras disciplinas da escola em que se está desenvolvendo a pesquisa. É esperado que esta solução possa, além de continuar seu desenvolvimento na escola, ser adotada por cursos da própria Universidade Federal do Ceará, promovendo a integração entre cursos presenciais e a distância na instituição.

Como trabalhos futuros, espera-se o desenvolvimento de um módulo bem estruturado de gamificação. Este módulo deverá possibilitar ao professor a elaboração de estratégias de promoção efetiva de motivação à participação do estudante, além de oferecer jogos que condicionem

melhor o cursista ao processo de aprendizagem, promovendo dinâmicas lógicas, exercícios de memória e concentração, etc. Para promover maior engajamento do estudante, deve ser realizada uma análise de usabilidade, para dotar os módulos da arquitetura de uma melhor experiência de usuário, com elementos gráficos e navegacionais atrativos e com forte apelo multimídia. Esta, inclusive, foi uma demanda postulada pelos alunos em diversas oportunidades.

Quanto aos sistemas de recomendação, estes podem passar a recomendar o estabelecimento de pares, associando colegas para, em determinadas circunstâncias, se unirem para a superação de dificuldades. Outra possível recomendação será a de rotas de estudo. Uma vez que são registrados todos os passos dos usuários no sistema, as trajetórias de alunos com melhor desempenho poderão ser recomendadas para alunos que tenham conhecimento coerente para cumprir tal direcionamento. Outra recomendação possível seria a de orientação do professor. Uma vez estabelecido um mecanismo de comunicação do professor com o(s) aluno(s), o sistema poderia aprender com as intervenções do professor frente a um dado problema, ou circunstância do aluno em um curso e, futuramente, uma vez que compreenda certa similaridade de circunstâncias, reutilizar orientações outrora realizadas pelo professor. Espera-se a geração de novos modelos preditivos com o uso de técnicas diversas como SVM, Fatorização de Tensores, *Restricted Boltzmann Machine* (RBM), etc.

## REFERÊNCIAS

- ADOMAVICIUS G., TUZHILIN A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. **IEEE Transactions on Knowledge and Data Engineering** 17, 734–749, June. 2005.
- ALBERTS, P.P., MURRAY, L.A., and STEPHENSON, J.E., Eight *In: Educational Considerations for Hybrid Learning*. *In: Wang, F.L., Fong, J., and Kwan, R.C. (Eds.) Handbook of Research on Hybrid Learning Models: Advanced Tools, Technologies, and Applications*. Information Sciences Reference: Hershey, PA, 185-202. 2010.
- ALEVEN, V., BAKER, R, WANG, Y., SEWALL, J., POPESCU, O. Bringing Non-programmer Authoring of Intelligent Tutors to MOOCs. *In: Proceedings of the Third (2016) ACM Conference on Learning@ Scale*. ACM, 2016. p. 313-316.
- ALL, A., CASTELLAR, E. P. N., VAN LOOY, J. Assessing the effectiveness of digital game-based learning: Best practices. **Computers & Education**, v. 92, p. 90-103, 2016.
- ALONSO, F., LÓPEZ, G., MANRIQUE, D., VIÑES, J. M. An instructional model for web-based e-learning education with a blended learning process approach. **British Journal of educational technology**, 36(2), 217-235, 2005.
- ASHENAFI, M. M., RONCHETTI, M., RICCARDI, G. Predicting Student Progress from Peer-Assessment Data. *In: EDM*. 2016. p. 270-275.
- BACICH, L., MORAN, J. Aprender e ensinar com foco na educação híbrida. **Revista Pátio**, Edição 25, Junho de 2015.
- BACICH, L., Neto, TREVISANI, A. T., MELLO, F. Ensino Híbrido: Personalização e Tecnologia na Educação. **Editora Penso**. 2015.
- BADER R., NEUFELD, E., WOERNDL, W., PRINZ, V. Context-Aware POI Recommendations in an Automotive Scenario Using Multi- Criteria Decision Making Methods. Proc. **Workshop Context- Awareness in Retrieval and Recommendation**, pp. 23-30, 2011.
- BAKER, R. S. J. d. *In: Data Mining for Education*. (Eds). McGAW, B., PETERSON, P., BAKER, E. **International Encyclopedia of Education**. 3rd edition. Oxford, UK: Elsevier. 2010.
- BAKER, R. S. J.; ISOTANI, S.; CARVALHO, A. M. J. B. Mineração de Dados Educacionais: Oportunidades para o Brasil. **Revista Brasileira de Informática na Educação**, Volume 19, Número 2. 2011.
- BARATA, G., GAMA, S., JORGE, J., GONÇALVES, D. Improving participation and learning with gamification. *In: Proceedings of the First International Conference on gameful design*, research and applications. ACM, 2013. p. 10-17.
- BEALE R., LONSDALE, P. Mobile Context Aware Systems: The Intelligence to Support Tasks and Effectively Utilise Resources. **Mobile Human-Computer Interaction**, S. Brewster and M. Dunlop, eds., vol. 3160, pp. 573-576, Springer, 2004.

- BERGMANN, J., SAMS, A. Flipping for mastery. **Educational Leadership**, v. 71, n. 4, p. 24-29, 2014.
- BERGMANN, J., SAMS, A. Our Story: Creating the Flipped Classroom. **Flip Your Classroom**, p. 11, 2012.
- BERNATEK, B., COHEN, J., HANLON, J., Wilka, M. Blended Learning in Practice: Case Studies from Leading Schools. **KIPP Empower Academy**,” Michael & Susan Dell Foundation, Setembro de 2012.
- BERSIN, J. The Blended Learning Handbook. New York: **Wiley**. 2004
- BISHOP, J. L., VERLEGER, M. A. The flipped classroom: A survey of the research. *In: ASEE National Conference Proceedings*, ATLANTA, GA. p. 1-18. 2013.
- BLOOM, B. S., HASTINGS, J. T., MADAUS, G. F., FLOREZ, M. C. F., PATTO, M. H. S., QUINTÃO, L. R., VANZOLINI, M. E. Manual de avaliação formativa e somativa do aprendizado escolar. São Paulo: **Pioneira**, 1971.
- BOELEN, R., WEVER, B., VOET, M. Four key challenges to the design of blended learning: A systematic literature review. **Educational Research Review**, 2017.
- BONK, C., GRAHAM, C. Handbook of blended learning: Global perspectives, local designs. San Francisco, CA: **Pfeiffer Publishing**. 2005.
- BRAMER, M. Principles of Data Mining (Undergraduate Topics in Computer Science). **Springer**; 2nd ed. 2013 edition (February 25, 2013).
- BRAY, B., MCCLASKEY, K. A Step-by-Step Guide to Personalize Learning. **Learning & Leading with Technology**, v. 40, n. 7, p. 12-19, 2013.
- BRICKLEY, D., MILLER, L. FOAF Vocabulary Specification 0.98. **Namespace Document**, Disponível em: <http://xmlns.com/foaf/spec>. Acessado em novembro de 2014.
- BRODERSEN, R. M., MELLUZZO, D. Summary of Research on Online and Blended Learning Programs That Offer Differentiated Learning Options. REL 2017-228. **Regional Educational Laboratory Central**, 2017.
- BRUSILOVSKY, P., MILLA'N, E., User Models for Adaptive Hypermedia and Adaptive Educational Systems. *In: The Adaptive Web*. (Eds), Brusilovsky, P., Kobsa, A., Nejdl, W. vol. 4321, pp. 3- 53, Springer, 2007.
- BRUSILOVSKY, P. **Domain Modeling for Personalized Guidance**. 2016.
- BRYAN, K., O'MAHONY, M., CUNNINGHAM, P. Unsupervised retrieval of attack profiles in collaborative recommender systems. *In: RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*, pp. 155–162. ACM, New York, NY, USA. DOI <http://doi.acm.org/10.1145/1454008.1454034>. 2008
- BUDER, J.; SCHWIND, C. Learning with personalized recommender systems: A psychological view. **Computers in Human Behavior**, v. 28, n. 1, p. 207-216, 2012.
- BURKE, R., Hybrid web recommender systems, *In: Brusilovsky, P., Kobsa, A., Nejdl, W. (Eds.). The Adaptive Web*. **Springer-Verlag**, Berlin Heidelberg. pp. 377–408, 2007.

CASTRO, E. A., RIBEIRO, V. C., SOARES, R., DE SOUSA, L. K. S., PEQUENO, J. O. M., MOREIRA, J. R. ENSINO HÍBRIDO: DESAFIO DA CONTEMPORANEIDADE?. **Projeção e Docência**, v. 6, n. 2, p. 47-58, 2015.

CASTRO, L. N., ZUBEN, F. J. V. **Inteligência Computacional Aplicada**. 2001. Disponível DOI: <http://dx.doi.org/10.1142/S1469026801000238> .

CHAN, S.; TRELEAVEN, P.; CAPRA, L. Continuous hyperparameter optimization for large-scale recommender systems. *In: Big Data, 2013 IEEE International Conference on IEEE*, 2013. p. 350-358.

CHEN, C. M. , DUH, L. J. Personalized web-based tutoring system based on fuzzy item response theory. **Expert Systems with Applications**, Volume 34, Issue 4, May 2008, pp. 2298-2315, ISSN 0957-4174. 2008. Disponível em: <http://dx.doi.org/10.1016/j.eswa.2007.03.010> .2008

CHEN, N. S., HWANG, G. J. Transforming the classrooms: innovative digital game-based learning designs and applications. **Educational Technology Research and Development**, v. 62, n. 2, p. 125-128, 2014.

CHRISTENSEN, C. Inovação na sala de aula: como a inovação disruptiva muda a forma de aprender. Porto Alegre: **Bookman**, 2012.

CORRIN, L.; DE BARBA, P. Exploring students' interpretation of feedback delivered through learning analytics dashboards. *In: Proceedings of the ascilite 2014 conference*. 2014. p. 629-633.

COSTA, E., BAKER, R. S., AMORIM, L., MAGALHÃES, J., MARINHO, T. Mineração de dados educacionais: Conceitos, técnicas, ferramentas e aplicações. **Jornada de Atualização em Informática na Educação**, 1(1):1-29.

CRUMMETT, C., MAZOUÉ, J., ANDERSON, R., DAUGHTREY, T., LOVE, W., MACDONALD, S., STOLOFF, M., WILCOX, D., and ZEMLIANSKY, P. Best Practices for Online and Hybrid Course Delivery. **James Madison University**. 2010. Disponível em: [http://www.jmu.edu/dl/wm\\_library/JMU\\_Best\\_Practices.pdf](http://www.jmu.edu/dl/wm_library/JMU_Best_Practices.pdf)

DAVIES, R. S., DEAN, D. L., Ball N. Flipping the classroom and instructional technology integration in a college-level information systems. Spreadsheet Course. **Educational Technology Research and Development (ETR&D)**, 61:4,pp. 563-580. DOI 10.1007/s11423-013-9305-6, 2013

DAVIES, R., WEST, R. Technology integration in school settings. *In: SPECTOR, M., MERRILL, D., ELEN, J., BISHOP, M. J. (Eds.) Handbook of research on educational communications and technology (4th ed.)*. New York: **Taylor & Francis Ltd**. 2013.

DE-MARCOS, L., DOMÍNGUEZ, A., SAENZ-DE-NAVARRETE, J., PAGÉS, C. An empirical study comparing gamification and social networking on e-learning. **Computers & Education**, v. 75, p. 82-91, 2014.

DETERDING, S., DIXON, D., KHALED, R., & NACKE, L. From game design elements to gamefulness: defining gamification. **Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments**. ACM, 2011. p. 9-15.

DOLOG, P., Nejd, W. Challenges and Benefits of the Semantic Web for User Modelling. **Proc. Workshop Adaptive Hypermedia and Adaptive Web-Based Systems (AH '03)**, pp. 99-112, 2003.

DRACHSLER, H., HUMMEL, H. G. K., KOPER, R. Identifying the goal, user model and conditions of recommender systems for formal and informal learning. **J. Digit. Inform.**, vol. 10, no. 2, pp. 4–24. 2009.

DRACHSLER, H., HUMMEL, H. G. K., KOPER, R. Personal recommender systems for learners in lifelong learning: requirements, techniques and model. **International Journal of Learning Technology** 3(4), 404–423. 2008.

DRACHSLER, H., VERBERT, K., SANTOS, O. C., e MANOUSELIS, N. Panorama of recommender systems to support learning. *In: Recommender systems handbook* (pp. 421-451). Springer US, 2015.

DRYSDALE, J. S., GRAHAM, C. R., SPRING, K. J., HALVERSON, L. R. An analysis of research trends in dissertations and theses studying blended learning. **The Internet and Higher Education**, 17, 90-100, 2013.

DUARTE, G. B., ALDA, L. S.; LEFFA, V. J. Gamificação e o feedback corretivo: considerações sobre a aprendizagem de línguas estrangeiras pelo Duolingo. **Raído**, v. 10, n. 23, p. 114-128. 2017.

DUOLINGO. **Aulas de Idiomas**. Site Oficial. Disponível em: <<https://www.duolingo.com>>. Acesso em: 30 mar. 2017.

DUVAL, E. Attention please!: learning analytics for visualization and recommendation. *In: Proceedings of the 1st international conference on learning analytics and knowledge*. ACM, 2011. p. 9-17.

DZIUBAN, C., HARTMAN, J., JUGE, F., MOSKAL, P., SORG, S. Blended learning enters mainstream. *In: BONK, C. J., GRAHAM, C. R. (Eds.). The handbook of blended learning. Global perspectives, local designs* (pp. 195-208). San Francisco, CA: John Wiley & Sons, 2006.

ELMARI, R.; NAVATHE, S. B.; **Sistemas de Banco de Dados**. 6ª Edição. Editora Pearson. (pp 698 – 731). 2010.

ERDT, M., FERNANDEZ, A., RENSING, C. Evaluating recommender systems for technology enhanced learning: A quantitative survey. **IEEE Transactions on Learning Technologies**, v. 8, n. 4, p. 326-344, 2015.

ERHEL, S., JAMET, E. Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. **Computers & Education**, v. 67, p. 156-167, 2013.

FARDO, M. L. A gamificação aplicada em ambientes de aprendizagem. **RENOTE**, v. 11, n. 1, 2013.

FARHAN S. M., RIZWAN, U. R. Web based Multimedia Recommendation System for e-Learning Website, **Int. J. of Advanced Networking and Applications**, Volume: 01, Issue: 04, pp. 217-223, 2010.

FERGUSON, R; CLOW, D. Examining engagement: analysing learner subpopulations in massive open online courses (MOOCs). *In: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. ACM, 2015. p. 51-58

FERREIRA, V. H., RAABE, A. L. LORSys – Um Sistema de Recomendação de Objetos de Aprendizagem SCORM. *RENOTE - Revista Novas Tecnologias na Educação*, CINTED-UFRGS, ISSN 1679-1916, v. 8, n.2 , 2010.

FIGUEIRA, C. V. Modelos de Regressão Logística. Dissertação (mestrado em matemática) **Programa de Pós-graduação em matemática do Instituto de Matemática Universidade Federal do Rio Grande do Sul**. Porto Alegre, 2006.

GARCÍA, E., ROMERO, C., VENTURA, S., CASTRO, C. An architecture for making recommendations to courseware authors using association rule mining and collaborative filtering. *Use Model. User-Adap.* 19, 99–132, 2009.

GETZELS, J. W., THELEN, H. A. The classroom as a unique social system. **National Society for the Study of Education Yearbook**, 1960, 59, 53-81, 1960.

GODOY, D., AMANDI, A. Link Recommendation in E- learning Systems based on Content-based Student Profiles. *In: ROMERO, C., VENTURA, S., PECHENIZKIY, M., BAKER, R. (Eds.), Handbook of Educational Data Mining, Data Mining and Knowledge Discovery Series. Chapman & Hall/CRC Press, 273-286, 2010.*

GOMES, P. Entenda como funcionam as plataformas adaptativas. 2013. Disponível em: <<http://porvir.org/porcriar/entenda-como-funcionam-plataformas-adaptativas/20130328>>. Acesso em: 01 mai. 2017.

GRAHAM, C. BLENDED, R. Learning systems: Definition, current trends, and future directions. *In: C. J. Bonk & C. R. Graham (Eds.), The handbook of blended learning: Global perspectives, local designs (pp. 3–21). San Francisco, CA: Pfeiffer Publishing. 2006.*

HALVERSON, L. R. Conceptualizing Blended Learning Engagement. Tese de Doutorado. **Brigham Young University**, 2016.

HALVERSON, L. R., GRAHAM, C. R., SPRING, K. J., DRYSDALE, J. S. An analysis of high impact scholarship and publication trends in blended learning. **Distance Education**, 33(3), 381-413, 2012.

HALVERSON, L. R., GRAHAM, C. R., SPRING, K. J., DRYSDALE, J. S., HENRIE, C. R. A thematic analysis of the most highly cited scholarship in the first decade of blended learning research. **The Internet and Higher Education**, 20, 20-34. 2014.

HÄMÄLÄINEN , W & VINNI , M. Classifiers for educational technology . *In: C Romero , S Ventura , M Pechenizkiy & R S J Baker (eds) , Handbook on Educational Data Mining . Chapman & Hall/CRC Data Mining and Knowledge Discovery Series , CRC Press , pp. 54-74. 2011.*

HAN, J., KAMBER, M., PEI, J. Data Mining: Concepts and Techniques. **Morgan Kaufmann**, San Francisco, CA. 2011.

HARIRI, N., MOBASHER, B., BURKE, R. Query-Driven Context Aware Recommendation. *In: Proceedings of the 7th ACM Conference on Recommender Systems (RecSys 2013)*, pp. 9–16. ACM, Nova Iorque, 2013.



- HARTNETT, M. Factors Undermining Motivation in Place-based Blended Learning. **Proceedings of ASCILITE 2009**, Auckland, NZ, 443-449. 2009.
- HELOU, S. E., SALZMANN, C., GILLET, D. The 3a Personalized, Contextual and Relation-Based Recommender System. **J. Universal Computer Science**, vol. 16, no. 16, pp. 2179-2195, Ago, 2010.
- HILL, P. Online educational delivery models: A descriptive view. **Educause review**, v. 47, n. 6, p. 84, 2012.
- HOIC-BOZIC, N., MORNAR, V., BOTICKI, I. A blended learning approach to course design and implementation. **IEEE transactions on education**, v. 52, n. 1, p. 19-30, 2009.
- HSIEH, T. C., LEE, M. C., SU, C. Y. Designing and implementing a personalized remedial learning system for enhancing the programming learning. **Educational Technology & Society** 16(4): 32-46, 2013
- HUANG, Y. M., HUANG, T. C., WANG, K. T., HWANG, W. Y. A Markov-based Recommendation Model for Exploring the Transfer of Learning on the Web. **Educational Technology & Society**, 12(2), 144-162.
- HUIYI, T., JUNFEI, G., YONG, L. E-Learning Recommendation System. **International Conference on Computer Science and Software Engineering**, 2008.
- HUMMEL, H. G. K., VAN DEN BERG, B., BERLANGA, A. J., DRACHSLER, H., JANSSEN, J., NADOLSKI, R. J., KOPER, E. J. R. Combining Social- and Information-based Approaches for Personalised Recommendation on Sequencing Learning Activities. **International Journal of Learning Technology**, Vol. 3, No.2, pp. 152 – 168. 2005.
- HUNG, C. M., HUANG, I., HWANG, G. J. Effects of digital game-based learning on students' self-efficacy, motivation, anxiety, and achievements in learning mathematics. **Journal of Computers in Education**, v. 1, n. 2-3, p. 151-166, 2014.
- IMBERNÓN, F. *Pedagogia Freinet: a atualidade das invariantes pedagógicas*. Porto Alegre: Editora **Penso**. 2010.
- JÁCOME J., L.; MENDES NETO, F. M.; FLORES, C. D.; SILVA, L. C. N.; SOMBRA, E. L.; COSTA, A. A. L. Uma extensão do moodle para recomendação ubíqua de objetos de aprendizagem. Revista **Renote**, v. 10, n. 3, 2012.
- JUNG, I., SUZUKI, K. Blended learning in Japan and its application in liberal arts education. **The handbook of blended learning**, 267-280. 2006.
- KAMPFF, A. Mineração de Dados Educacionais para Geração de Alertas em Ambientes Virtuais de Aprendizagem como Apoio à Prática Docente. **Programa de Pós-Graduação em Informática na Educação, UFRGS**. Porto Alegre, RS. 2009.
- KANTOR, P. B., ROKACH, L., RICCI, F., SHAPIRA, B. Recommender systems handbook. **Springer**, 2011.
- KARAMPIPERIS, P., SAMPSON, D. Adaptive Learning Resources Sequencing in Educational Hypermedia Systems. **Educational Technology & Society** 8(4), 128–147. 2005.

- KHAN, S. *The One World Schoolhouse – Education Reimagined*. Hodder and Stoughton, London, 2012 and Twelwe, **Boston & New York** 2012. 272 pages. ISBN 978 1 444 75577 0. 2012
- KHAN, Z. R. Using innovative tools to teach computer application to business students—A Hawthorne effect or successful implementation here to stay. **J. Univ. Teach. Learn. Pract.**, vol. 11, no. 1, pp. 1–13, 2014.
- KHRIBI, M. K., JEMNI, M., NASRAOUI, O. Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval. **Educational Technology & Society**, 12(4), 30-42, 2009.
- KITCHENHAM, B. Procedures for Performing Systematic Reviews. **Keele University Technical Report** TR/SE-0401 ISSN:1353-7776. 2004. Disponível em: <http://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf>
- KLOCK, A. C. T., CARVALHO, M. F., ROSA, B. E., GASPARINI, I. Análise das técnicas de Gamificação em Ambientes Virtuais de Aprendizagem. **RENOTE - Revista Novas Tecnologias na Educação**, v. 12, n. 2, p. 1–10. 2014.
- KONG, S. C., SONG, Y. An experience of personalized learning hub initiative embedding BYOD for reflective engagement in higher education. **Computers & Education**, v. 88, p. 227-240, 2015.
- KOPEINIK, S., KOWALD, D., LEX, E. Which Algorithms Suit Which Learning Environments? A Comparative Study of Recommender Systems in TEL. *In: EC-TEL*, p. 124-138, 2016.
- KOREN, Y. Factor in the neighbors: Scalable and accurate collaborative filtering. **ACM Transaction on Knowledge Discovery Data**, 4(1):1-24, 2010.
- LAGE, M. J.; PLATT, GLENN, J.; TREGLIA, M. Inverting the classroom: A gateway to creating an inclusive learning environment. **The Journal of Economic Education**, v. 31, n. 1, p. 30-43, 2000.
- LANE-KELSO, M. Mobiles and Flipping in Oman: Bloom blays a smart phone. **Journal of Information Tecnology and Application in Education**, 3(2), 67-76, 2014.
- LAUMAKIS, M., GRAHAM, C. R., DZIUBAN, C. D. The Sloan-C Pillars and Boundary Objects As a Framework for Evaluating Blended Learning. **Journal of Asynchronous Learning Networks**, 13. pp 75-87, 2009.
- LITTLEJOHN, A., PEGLER, C. Preparing for blended E-learning: Understanding blended and online learning (connecting with E-learning). London: **Routledge**. 2007.
- LU, J., WU, D., MAO, M., WANG, W., ZHANG, G. Recommender system application developments: a survey. **Decision Support Systems**, 74, 12-32, 2015.
- MACDONALD, J. **Blended Learning and Online Tutoring: A Good Practice Guide**, Aldershot, UK: Gower Publishing Co., 2006.
- MAIA, A., LOPES, J.B., MARTINS, P., PESSOA, T. Authoring Tools As Instruments For A New Approach Of Educational Planning. *In: INTED2015 Proceedings*. IATED, 2015. p. 5149-5158.

- MCCULLOCH, W. S., PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, vol. 5, pp. 115-133, 1943.
- MANOUSELIS, N., COSTOPOULOU, C. Experimental Analysis of Design Choices in Multi-Attribute Utility Collaborative Filtering. **International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI)**, Special Issue on Personalization Techniques for Recommender Systems and Intelligent User Interfaces, 21(2), 311-331, 2007.
- MANOUSELIS, N., DRACHSLER, H., VERBERT, K., DUVAL, E. Recommender Systems for Learning. New York, NY, USA: **Springer**, 2013.
- MANOUSELIS, N., DRACHSLER, H., VUORIKARI, R., HUMMEL, H., KOPER, R. Recommender Systems in Technology Enhanced Learning. *In: Recommender Systems Handbook: A Complete Guide for Research Scientists and Practitioners.* (Eds) ROKACH, L., SHAPIRA, B., KANTOR, P., RICCI, F. pp. 387-409, Springer, 2011.
- MANOUSELIS, N., DRACHSLER, H., VUORIKARI, R., HUMMEL, H., KOPER, R. Recommender systems in technology enhanced learning. **Recommender Systems Handbook**, Springer, pp. 387-415, 2010
- MARGARYAN, A., BIANCO, M., LITTLEJOHN, A. Instructional quality of Massive Open Online Courses (MOOCs). **Computers & Education**, Volume 80, January 2015, Pages 77-83, ISSN 0360-1315.
- MARQUES, C. G. Desenvolvimento e implementação de um modelo de blended-learning com objectos de aprendizagem no ensino superior. Tese em Ciências da Educação (Tecnologia Educativa), **Universidade do Minho**, Registro: <http://hdl.handle.net/1822/19700>. 2012.
- MCCALLA, G. The ecological approach to the design of e-learning environments. **Journal of Interactive Media in Education**, Vol. 7, 2004.
- MCGEE P., REIS A. Blended course design: a synthesis of best practices. **Journal of Asynchronous Learning Networks**, 16 (4) .(2012), pp. 7–22
- MERINO, P. J. M., RUIPÉREZ-VALIENTE, J. A., and DELGADO, C. Inferring higher level learning information from low level data for the Khan Academy platform. Proceeding **LAK '13 Proceedings of the Third International Conference on Learning Analytics and Knowledge**, 112-116. 2013.
- METTOURIS, C.; PAPADOPOULOS, G. A. Using appropriate context models for CARS context modelling. *In: Knowledge, Information and Creativity Support Systems.* **Springer**, Cham, 2016. p. 65-79.
- MONTGOMERY, D. C.; PECK, Elizabeth, A.; VINING, G. G. Introduction to linear regression analysis. **John Wiley & Sons**, 2015.
- MORAN, J. M. Mudando a educação com metodologias ativas. Coleção Mídias Contemporâneas. **Convergências Midiáticas, Educação e Cidadania: aproximações jovens**, v. 2, 2015.
- MORRISON, B. B., DISALVO, B., Khan academy gamifies computer science. **Proceedings of the 45th ACM technical symposium on Computer science education**, March 05-08. 2014.

- MOSKAL, P., DZIUBAN, C., HARTMAN, J. Blended learning: A dangerous idea Blended Learning in Higher Education. **Policy and Implementation Issues**, 15 -23. doi:10.1016/j.iheduc.2012.12.001. 2013.
- MURPHY, R., GALLAGHER, L., KRUMM, A., MISSLEVY, J., HAFTER, A. Research on the Use of Khan Academy in Schools. Menlo Park, CA: **SRI Education**. 2014.
- NEDUNGADI, P., SMRUTHY, T. K. Personalized multi-relational matrix factorization model for predicting student performance. **Intelligent Systems Technologies and Applications**, v. 1, p. 163, 2015.
- NG, A. Supervised learning. CS 229 **Machine Learning Course Materials**. Acesso em 21 de julho de 2017.
- NORBERG, A., DZIUBAN, C. D., MOSKAL, P. D. A time-based blended learning model. **On the Horizon**, v. 19, n. 3, p. 207-216, 2011.
- OSMANBEGOVIC, E., SULJIC, M. Data Mining Approach For Predicting Student Performance Economic Review. **Journal of Economics and Business X**. 2012.
- PARDOS, Z. A. HEFFERNAN, N. T. Using hmms and bagged decision trees to leverage rich features of user and skill from an intelligent tutoring system dataset. **KDD Cup 2010: Improving Cognitive Models with Educational Data Mining**. 2010.
- PARK, H. S., YOO, J. O., CHO, S. B. A Context-Aware Music Recommendation System Using Fuzzy Bayesian Networks with Utility Theory. **Fuzzy Systems and Knowledge Discovery**, L. Wang, L. Jiao, G. Shi, X. Li, and J. Liu, eds., vol. 4223, pp. 970-979, Springer, 2006.
- PAZZANI, Michael J.; BILLSUS, Daniel. Content-based recommendation systems. *In: The adaptive web*. **Springer**, Berlin, Heidelberg, 2007. p. 325-341.
- PEQUENO, P. A. L. Um sistema de recomendação com filtragem híbrida de conteúdos para ambientes virtuais de aprendizagem como instrumento de suporte a alunos e acompanhamento de turmas numerosas. Dissertação (mestrado) – **Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós- Graduação em Engenharia de Teleinformática**, Fortaleza, 2014.
- PHILLIPS, B. Beyond Classroom Learning: Personalized Learning Through Digital Technologies. **Neue Lehr – und Lernkulturen in der technologiegestutzten Lehre**. Disponível em: [http://ffhoarep.fh-ooe.at/bitstream/123456789/664/1/120\\_214\\_Phillips\\_FullPaper\\_en\\_Final.pdf](http://ffhoarep.fh-ooe.at/bitstream/123456789/664/1/120_214_Phillips_FullPaper_en_Final.pdf). 2016.
- R Core Team. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. URL <https://www.Rproject.org/>. (2017)
- REAL, L. M. C.; TAVARES, M.N.R.; PICETTI, J. Formação de Professores para o Uso Educacional de Tablets no Ensino Médio: possíveis mudanças na prática pedagógica. *In: Anais dos Workshops do Congresso Brasileiro de Informática na Educação*. 2013.
- RENDLE, S. Factorization Machines with libFM. **ACM Transactions on Intelligent Systems and Technology (TIST)**, v.3 n.3, p.1-22, Maio de 2012.

- REVISTABW. Classificadores Bayesianos. **Revista Brasileira de Web: Tecnologia**. Disponível em <http://www.revistabw.com.br/revistabw/classificadores-bayesianos/>. Criado em: 14/02/2015. Última atualização: 24/07/2015. Visitado em: 28/04/2017.
- RICCI, F., ROKACH, L., SHAPIRA, B. Introduction to recommender systems handbook. **Springer US**, 2011.
- RIFON, L. E. A., RODRIGUEZ, A. C., RORIS, V. M. A., GAGO, J. M. S. IGLESIAS, M. J. F. A Recommender System for Educational Resources in Specific Learning Contexts. **The 8th International Conference on Computer Science & Education (ICCSE 2013)**. Colombo, Sri Lanka. April 26-28. 2013.
- ROMERO, C., VENTURA, S.; Educational Data Mining: A Survey from 1995 to 2005. **Expert Systems with Applications**, 33(1), 135-146. 2007.
- ROSS, B., GAGE, K. Global perspectives on blended learning. Insight from WebCT and our customers in higher education. *In*: BONK, C. J., GRAHAM, C. R. (Eds.), **The handbook of blended learning**. Global perspectives, local designs (pp. 155-168). San Francisco, CA: John Wiley & Sons, 2006.
- RAUBER, T.W. Redes neurais artificiais. **Universidade Federal do Espírito Santo**, 2005.
- RUIPÉREZ-VALIENTE, J. A., MERINO, P. J. M. , LEONY, D., Kloos, C. D. ALAS-KA: A Learning Analytics Extension for Better Understanding the Learning Process in the Khan Academy Platform. **Computers in Human Behavior**, v.47 n.C, p.139-148. 2014.
- SANTOS, O. C., BOTICARIO, J. G., PÉREZ-MARÍN, D. Extending web-based educational systems with personalised support through User Centred Designed recommendations along the e-learning life cycle. **Science of Computer Programming**, v. 88, p. 92-109, 2014.
- SCHAFER, J., FRANKOWSKI, D., HERLOCKER, J., SEN, S. Collaborative Filtering Recommender Systems The Adaptive Web. *In*: **The Adaptive Web: Methods and Strategies of Web Personalization, Lecture Notes in Computer Science**. (Eds) BRUSILOVSKY, P., vol. 4321. Berlin, Heidelberg: Springer Berlin 1 Heidelberg. pp. 291-324. 2007.
- SCHMIDT, A. Potentials and Challenges of Context-awareness for Learning Solutions. Proc. **LWA—LernenWissensentdeckung— Adaptivita't**, 13th Ann. Workshop SIG Adaptivity and User Modeling in Interactive Systems, pp. 63-68, 2005.
- SHANI, G., GUNAWARDANA, A. Evaluating recommendation systems. *In*: **Recommender Systems Handbook**. New York, NY, USA: Springer, ch. 8, pp. 257–297, 2011.
- SHARPE, R., BENFIELD, G., ROBERTS, G., FRANCIS, R. The undergraduate experience of blended e-learning: A review of UK literature and practice. Executive summary. **Higher Education Academy**. Disponível em: [https://www.heacademy.ac.uk/sites/default/files/sharpe\\_benfield\\_roberts\\_francis\\_0.pdf](https://www.heacademy.ac.uk/sites/default/files/sharpe_benfield_roberts_francis_0.pdf). 2006.
- SHEN R., WANG, M., GAO, W., NOVAK, D., and Tang, L. Mobile learning in a large blended computer science classroom: System function, pedagogies, and their impact on learning. **IEEE Trans. Educ.**, vol. 52, no. 4, pp. 538–546, Nov. 2009.
- SHIBLEY, I. Blended course design. **Magna Publications** (2009). Disponível em: <http://www.magnapubs.com/catalog/blended-learning-course-design-whitepaper/>.

SHUTE, V. J.; ZAPATA-RIVERA, D. Adaptive educational systems. **Adaptive technologies for training and education**, v. 7, p. 27, 2012.

SOROUR, S. E., MINE, T., GODA, K., HIROKAWA, S. A predictive model to evaluate student performance. **Journal of Information Processing**, 23(2), 192-201, 2015.

SOUZA, B. F. M. Modelos de fatoração matricial para recomendação de vídeos. **Pontifícia Universidade Católica do Rio de Janeiro**. Certificação Digital 0812538CA. 2011.

SPECHT, M. ACE - Adaptive Courseware Environment. **Proc. Int'l Conf. Adaptive Hypermedia and Adaptive Web-Based Systems (AH '00)**, pp. 380-383, 2000.

STAKER, H.; HORN, M. B. Classifying K–12 blended learning. **Mountain View, CA: Innosight Institute**, Inc. 2012. Disponível em: <<http://files.eric.ed.gov/fulltext/ED535180.pdf>>. Acesso em: 29 de abr. 2017.

SWEENEY, M., LESTER, J., RANGWALA, H. Next-term student grade prediction. *In: Big Data (Big Data)*. **IEEE International Conference - IEEE**, 2015. p. 970-975.

SWEENEY, M., RANGWALA, H., LESTER, J., JOHRI, A. Next-term student performance prediction: A recommender systems approach. **Computers e Society**. preprint arXiv:1604.01840. 2016. Disponível em: <https://arxiv.org/pdf/1604.01840.pdf>

TAN, L., WANG, M., XIAO, J. Best Practices in Teaching Online or Hybrid Courses: A Synthesis of Principles. **Lecture Notes in Computer Science** 6248: 117-12. 2010.

TANG, T. Y., MCCALLA, G. I. Smart Recommendation for an Evolving E-Learning System: Architecture and Experiment. **International Journal on E-Learning**, 4 (1), 105-129, 2005.

TANKELEVICIENE L., DAMASEVICIUS, R. Towards a Conceptual Model of Learning Context in E-Learning. **Proc. IEEE Int'l Conf. Advanced Learning Technologies**, pp. 645-646, 2009.

TAYLOR, J. A., NEWTON, D. Beyond blended learning: A case study of institutional change at an Australian regional university. *The Internet and Higher Education*. Volume 18, July 2013, Pages 54-60, ISSN 1096-7516, <http://dx.doi.org/10.1016/j.iheduc.2012.10.003>. <http://www.sciencedirect.com/science/article/pii/S1096751612000668>

TENG, D. C., CHEN, N. S., LEE, C. H. Enhancing English Reading Comprehension by Integrating Direct Access to Digital Materials and Scaffolded Questionings in Paper Prints. **Proc. IEEE 11th Int'l Conf. Advanced Learning Technologies (ICALT)**, pp. 244-248, July 2011.

THAI-NGHE, N. DRUMOND, L. HORVÁTH, T. KROHN-GRIBERGHE, A. NANOPOULOS, A. SCHMIDT-THIEME, L. Factorization Techniques for Predicting Student Performance. **Universidade de Hildesheim**. Alemanha, 2011.

THAI-NGHE, N., DRUMOND, L., HORVITH, T., KROHN-GRIMBERGHE, A., NANOPOULOS, A., SCHMIDT-THIEME, L. Factorization Techniques for Predicting Student Performance. *In: SANTOS, O., BOTICARIO, J. (Eds) Educational Recommender Systems and Technologies: Practices and Challenges*, pp. 129-153. doi:10.4018/978-1-61350-489-5.ch006. 2012.

- THÜS, H, CHATTI, M. A., YALCIN, E, PALLASCH, C, KYRYLIUK, B, MAGERAMOV, T, SCHROEDER, U. Mobile Learning in Context. *In: International Journal of Technology Enhanced Learning*, Vol. 4 No. 5/6, 2012.
- TODOROV, J, C., TRISTÃO, G. Sistema personalizado de ensino: bases psicológicas e abordagem administrativa. *Cadernos de Psicologia Aplicada*, 3, 65-71, 1975.
- TODOROV, J. C., MOREIRA, M. B., MARTONE, R. C. Sistema personalizado de ensino, educação a distância e aprendizagem centrada no aluno. *Psicol. teor. pesqui*, v. 25, n. 3, p. 289-296, 2009.
- TRATTNER, C., KOWALD, D., SEITLINGER, P., LEY, T., KOPEINIK, S. Modeling Activation Processes in Human Memory to Predict the Use of Tags in Social Bookmarking Systems. *J. Web Science*, 2(1), 1-16. 2016.
- VALENTE, J. A. Blended learning e as mudanças no ensino superior: a proposta da sala de aula invertida. *Educar em Revista*, n. 4, 2014.
- VERBERT, K. MANOUSELIS, N., OCHOA, X., WOLPERS, M. DRACHSLER, H., BOSNIC, I., DUVAL, E. Context-Aware Recommender Systems for Learning: A Survey and Future Challenges. *IEEE Transactions on Learning Technologies*, v.5 n.4, p.318-335, Janeiro de 2012.
- VERBERT, K., DRACHSLER, H., MANOUSELIS, N., WOLPERS, M., VUORIKARI, R., DUVAL, E. Dataset-driven research for improving recommender systems for learning. *In: Proc. 1st Int. Conf. Learn. Analytics Knowl.*, 2011, pp. 44–53.
- VERBERT, K., DUVAL, E., KLERKX, J., GOVAERTS, S., SANTOS, J. L. Learning analytics dashboard applications. *American Behavioral Scientist*, v. 57, n. 10, p. 1500-1509, 2013.
- WALBERG, H.J.; ANDERSON, G. J. Classroom climate and individual learning. *Journal of educational Psychology*, v. 59, n. 6p1, p. 414, 1968.
- WANG, F. H. On extracting recommendation knowledge for personalized web-based learning based on ant colony optimization with segmented-goal and meta-control strategies. *Expert Syst. Appl.* 39(7), pp. 6446-6453. 2012
- WANNER, T., PALMER, E. Personalising learning: Exploring student and teacher perceptions about flexible learning and assessment in a flipped university course. *Computers & Education*, v. 88, p. 354-369, 2015.
- WERBACH, K., HUNTER, D. For The Win: How Game Thinking Can Revolutionize Your Business. Filadélfia, Pensilvânia: **Wharton Digital Press**, 2012.
- WITTEN, Ian H., FRANK, E., HALL, M. A., PAL, C. J. Data Mining: Practical machine learning tools and techniques. **Morgan Kaufmann**, 2016.
- WONGWATKIT, C., SRISAWASDI, N., HWANG, G. J., PANJABUREE, P. Influence of an integrated learning diagnosis and formative assessment-based personalized web learning approach on students learning performances and perceptions. *Interactive Learning Environments*, 1-15, 2016.

- XU, D., WANG, Z., CHEN, K., HUANG, W. Personalized Learning Path Recommender Based on User Profile Using Social Tags. **Computational Intelligence and Design (ISCID)**, 2012 Fifth International Symposium, v. 1, doi: 10.1109/ISCID.2012.133, p. 511 – 514, 2012.
- YU, Z., ZHOU, X., ZHANG, D., CHIN, C. Y., WANG, X., MEN, J. Supporting Context-Aware Media Recommendations for Smart Phones. **IEEE Pervasive Computing**, vol. 5, no. 3, pp. 68-75, Jul-Set., 2006.
- ZACHARIS, N. Z. Predicting Student Academic Performance in Blended Learning Using Artificial Neural Networks. **International Journal of Artificial Intelligence and Applications**, 7 (5), p. 17-29, 2016.
- ZAERI, N. Blended Learning System Performance Evaluation. **International Journal of Computer Applications** 76(4). pp 33-39, August 2013.
- ZAFRA, A., ROMERO, C., VENTURA, S. Multiple instance learning for classifying students in learning management systems. **Expert Systems with Applications**, v. 38, n. 12, p. 15020-15031, 2011.
- ZAIANE, O. R. Building a recommender agent for e-learning systems. **Proceedings of 2002 International Conference on Computers in Education**, vol. 51 pp. 55–59. 2002.
- ZAKRZEWSKA, D. *In*: Building group recommendations in E-learning systems. (Eds.) JEDRZEJOWICZ, P., NGUYEN, N. T., HOWLET, R. J., JAIN, L. C. KES-AMSTA 2010. LNCS, vol. 6070, pp. 391–400. Springer, Heidelberg, 2010.
- ZAMPA, M. P., MENDES, L. F. C. Gamificação: uma proposta para redução da evasão e reprovação em disciplinas finais da graduação. **Caderno de Estudos em Sistemas de Informação**, v. 3, n. 2. 2017.
- ZHENG, Y., BURKE, R., Mobasher, B. The role of emotions in context-aware recommendation. *In*: **ACM RecSys**, the 3rd Workshop on Human Decision Making in Recommender Systems. ACM, 2013.
- ZHOU, Y., Zhou, Y., WILKINSON, D., SCHREIBER, R., PAN, R. Large-scale parallel collaborative filtering for the netflix prize. **Lecture Notes in Computer Science**, v. 5034, p. 337-348, 2008.
- ZMUDA, A. ULLMAN, D., CURTIS, G. Learning personalized: The evolution of the contemporary classroom. **John Wiley & Sons**, 2015.



**APÊNDICE A** – Mapeamento de componentes entre Modelo Conceitual e Arquitetura Digital  
Class.

Tabela 32 – Mapeamento de componentes entre Modelo Conceitual e Arquitetura Digital Class

Dimensão no Modelo Digital Class	Descrição de Características da dimensão	Módulo(s) da Arquitetura Digital Class que representa a dimensão	Descrição de Características do Módulo
Adequabilidade a objetivos do curso	- Planejamento do curso (atividades, conteúdos, materiais didáticos, objetivos, etc.)	- Gryphon  - Sphinx  - Chimera	- Gryphon: verificação da configuração do curso e seus componentes por parte do professor - Sphinx: disponibilização do cadastro de questões, simulados, questionários, etc. - Chimera: permissão para gravação de aulas, compartilhamento de materiais da Internet, etc.
Avaliação contínua	- disponibilização diária de questões avaliativas	- Gryphon  - Sphinx	- Gryphon: apresenta as questões vinculadas ao contexto em que o aluno está inserido, considerando para isso os objetivos a serem cumpridos no curso e o seu grau de conhecimento sobre cada assunto trabalhado. - Sphinx: a plataforma corresponde a um acervo colaborativo em que professores podem cadastrar e compartilhar os seus itens.
Paralelização de estudos	- possibilidade do aluno de estudar simultaneamente mais de um componente curricular (acesso a aulas, vídeos, textos, questões, etc.)	- Gryphon  - Chimera  - Sphinx	- Gryphon: possibilita que um aluno esteja trabalhando com mais de um tema ao mesmo tempo, mesmo que a disciplina esteja em um determinado ponto de seu cronograma, o aluno pode continuar trabalhando temas já superados pelo professor, como forma de se preparar continuamente nesses temas em que apresenta maior fragilidade. - Chimera: os materiais armazenados neste módulo poderão ser acessados de forma reiterada, sempre que o aluno desejar. - Sphinx: o acervo de exercícios permite que o aluno vá testando seu conhecimento e evolução nos assuntos.

Tabela 33 – Mapeamento de componentes entre Modelo Conceitual e Arquitetura Digital Class (continuação)

Dimensão no Modelo Digital Class	Descrição de Características da dimensão	Módulo(s) da Arquitetura Digital Class que representa a dimensão	Descrição de Características do Módulo
Ritmo próprio	Possibilidade do aluno avançar sobre novos conteúdos, reforçar assuntos já vistos. Individualidades são respeitadas.	- Gryphon - Chimera	Com os materiais e planos de estudos sempre a disposição dos alunos e de forma personalizada, os estudantes poderão em seu ritmo acelerar o processo de aprendizado, antecipando temas do curso, ou mesmo trabalhando em dias que lhe são mais convenientes.
Predição de desempenho	- Modelo para prever o desempenho do estudante para cada assunto do curso.	- Gryphon	As informações colhidas quanto a resolução de testes e questões na plataforma, permitem a geração de um modelo preditivo que auxiliará o professor na identificação de pontos em que cada aluno terá notória dificuldade. Esta identificação precoce de problemas, orienta o professor quanto a tomada de medidas que venham a superar determinadas limitações do curso, podendo, mais uma vez, estabelecer trabalhos individualizados de acordo com a carência de cada estudante.
Personalização	- Cada aluno deve receber de forma reiterada em um curso, uma sequência de atividades a seguir (plano de aula). Esta sequência deve ser específica à condição de momento do aluno em termos de dificuldades dependências existentes.  - Os materiais didáticos oferecidos respeitam as individualidades e contexto do usuário.	- Gryphon	Diversos mecanismos avaliativos são disponibilizados na plataforma para que o conhecimento sobre a condição cognitiva do aluno seja conhecida em tempo real ao longo de toda a realização de um curso. A partir disto, diversas ações personalizadas poderão ser empreendidas pelo sistema (geração de planos de estudos, geração de lista de conteúdos e materiais didáticos) e pelo professor (montagem de grupos de estudos, aulas de reforço com temáticas definidas para cada grupo de alunos, monitorias, etc.)

**APÊNDICE B – Arquitetura Genérica de AEH**

Tabela 34 – Principais técnicas de filtragem contribuem com sistemas AEH.

Filtro	Nome	Descrição	Vantagens	Desvantagens	Aplicação em TEL
FC (FC - Colaborativa / FBC - Filtragem Baseada em Conteúdo )		Usuários que classificaram o mesmo item similarmente e provavelmente possuem o mesmo gosto. Baseado nessa premissa, o sistema recomenda itens ainda não vistos por um usuário, porém vistos por usuários similares.	- Não analisa a natureza de conteúdos. - Independente de domínio - A qualidade melhora com o tempo - Abordagem bottom-up	- Novos usuários, - Escalabilidade, - Esparsidade	- Benefícios da experiência, - aprendizes podem ser alocados em grupos de acordo com suas, similaridades)
FC	Baseado em Usuário	- Foco nos itens, assumindo que itens avaliados similarmente são provavelmente similares. Eles recomendam itens com alta correlação (baseado na correlação dos itens)	- Não analisa a natureza de conteúdos - Independente de domínio - A qualidade melhora com o tempo - Abordagem bottom-up.	- Novo item é um problema, - preferências populares, - possui problema de cold-start, - Esparsidade	- beneficia-se da experiência
FC	Estereótipos ou demográficos	- são encontrados usuários com atributos similares, para recomendar itens que são preferidos por usuários similares (baseado no uso de dados e não na classificação deles)	- Sem problema de cold-start, - independente de domínio.	- informação insuficiente, - somente gosto popular, - obtenção de metadados, - manutenção de ontologia	- aloca aprendizes a grupos, - beneficia a experiência, - recomendação de início de RS

Continuação					
Filtro					
	Nome	Descrição	Vantagens	Desvantagens	Aplicação em TEL
	FC - Colaborativa / FBC - Filtragem Baseada em Conteúdo )				
RBC (raciocínio baseado em casos)	Raciocínio baseado em casos	Assume que se um usuário gosta de um determinado item, ele provavelmente também gostará de um similar	- sem análise de conteúdo - qualidade melhora com o tempo. - independente de domínio.	- problema com a entrada de novos usuários. - super-especialização. - Esparsidade. - problema de cold-start	- mantém o aprendiz informado sobre o objetivo da aprendizagem. - útil para sistema de recomendação híbrido.
CB	Técnicas baseadas no atributo	Recomenda itens baseados na associação entre os atributos e o perfil de um usuário. Atributos podem ser ponderados de acordo com sua importância ao usuário	- sem problema de cold-start. - problema a cada novo item cadastrado. - sensível a mudanças de preferências	- só trabalha com categorias de itens. - requer definição de ontologies. - super-especialização.	Útil para uma abordagem híbrida de RS.

Fonte – Fonte: DRACHSLER et al.(2008).