



UNIVERSIDADE FEDERAL DO CEARÁ
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, ATUÁRIA, CONTABILIDADE E
SECRETARIADO EXECUTIVO
CURSO DE CIÊNCIAS ECONÔMICAS

MATEUS FERNANDES PESSOA

PREVENDO O DESEMPENHO ESCOLAR USANDO *MACHINE*
***LEARNING* E HABILIDADES SOCIOEMOCIONAIS**

FORTALEZA
2022

MATEUS FERNANDES PESSOA

PREVENDO O DESEMPENHO ESCOLAR USANDO *MACHINE*
LEARNING E HABILIDADES SOCIOEMOCIONAIS

Monografia apresentada ao Curso de Ciências Econômicas do Departamento de Teoria Econômica da Universidade Federal do Ceará, como requisito parcial para obtenção do Título de Bacharel em Ciências Econômicas.

Orientador: Prof. Dr. Rafael Barros
Barbosa

FORTALEZA
2022

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

P568p Pessoa, Mateus Fernandes.
Prevendo o desempenho escolar usando machine learning e habilidades socioemocionais / Mateus
Fernandes Pessoa. – 2022.
45 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Faculdade de Economia,
Administração, Atuária e Contabilidade, Curso de Ciências Econômicas, Fortaleza, 2022.
Orientação: Prof. Dr. Rafael Barros Barbosa.

1. Educação . 2. Habilidades socioemocionais . 3. Instrumento Senna . 4. Machine learning. I. Título.
CDD 330

MATEUS FERNANDES PESSOA

PREVENDO O DESEMPENHO ESCOLAR USANDO *MACHINE
LEARNING* E HABILIDADES SOCIOEMOCIONAIS

Monografia apresentada ao Curso de Ciências Econômicas do Departamento de Teoria Econômica da Universidade Federal do Ceará, como requisito parcial para obtenção do Título de Bacharel em Ciências Econômicas.

Aprovada em: __/__/____.

BANCA EXAMINADORA

Prof. Dr. Rafael Barros Barbosa (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Vitor Hugo Miro Couto Silva
Universidade Federal do Ceará (UFC)

Prof. Dr. Marcelo de Castro Callado
Universidade Federal do Ceará (UFC)

A Deus.

À Nossa Senhora.

Aos meus pais, Socorro e Marcondes.

AGRADECIMENTOS

A Deus primeiramente, pelo dom da vida e a graça de poder crescer acadêmica e profissionalmente e no conhecimento da Verdade.

À Universidade Federal do Ceará pela excelente formação e por ter sido minha segunda casa.

Ao meu orientador Prof. Dr. Rafael Barros Barbosa, pelo seu empenho e dedicação nas diversas análises e correções deste trabalho.

Aos professores participantes da banca examinadora Vitor Hugo Miro Couto Silva e Marcelo de Castro Callado pelo tempo, pelas valiosas colaborações e sugestões.

Aos colegas da turma e em especial a Ciro Lúcio e Mateus Holanda, pela amizade, apoio e presença durante esses anos na faculdade.

“Educação é aquilo que fica depois que
você esquece o que a escola ensinou.”

(Albert Einstein)

RESUMO

Existe uma diversidade de estudos que mostram as habilidades não cognitivas como importantes para o desenvolvimento dos alunos na escola. Porém, esses estudos não possuem uma abordagem que busque estudar as interações entre essas habilidades e a não linearidade presente entre elas. Neste trabalho, foi utilizado o instrumento SENNA, construído para mensurar as habilidades sociais e emocionais em escolas brasileiras. O instrumento foi aplicado pelo Instituto Ayrton Senna em 105.594 alunos do Ceará em 2015. Com métodos de *machine learning*, se tentou capturar os efeitos não lineares e prever o desempenho escolar nas disciplinas de português e matemática por meio das habilidades socioemocionais. Foi percebido que o modelo de *Random Forest* chegou a ser 16% melhor que o modelo linear na previsão do desempenho dos alunos e que a adição de variáveis interativas melhoram o modelo de previsão. Na previsão dos piores alunos na disciplina de português, o modelo de *Random Forest* teve uma acurácia de 76,4%. Foi possível concluir que os modelos de *machine learning* possuem propriedade para prever o desempenho dos alunos e também há espaço para se utilizar outros métodos de *machine learning* mais novos e que utilizam mais poder computacional.

Palavras-chave: educação; habilidades socioemocionais; instrumento Senna; machine learning.

ABSTRACT

There is a diversity of studies that show non-cognitive skills as important for the development of students in school. However, these studies do not have an approach that seeks to study the interactions between these skills and the non-linearity present between them. In this work, the SENNA instrument was used, built to measure social and emotional skills in Brazilian schools. The instrument was applied by the Ayrton Senna Institute to 105,594 students in Ceará in 2015. Using machine learning methods, an attempt was made to capture non-linear effects and predict school performance in Portuguese and mathematics through socio-emotional skills. It was noticed that the Random Forest model was 16% better than the linear model in predicting student performance and that the addition of interactive variables improves the prediction model. In predicting the worst students in the Portuguese subject, the Random Forest model had an accuracy of 76.4%. It was possible to conclude that machine learning models have the property to predict student performance and there is also room to use other newer machine learning methods that use more computational power.

Keywords: education; socio-emotional skills; Senna instrument; machine learning.

LISTA DE FIGURAS

Figura 1 – Nós da árvore de decisão	33
---	----

LISTA DE GRÁFICOS

Gráfico 1 – Perfil por gênero	28
Gráfico 2 – Perfil por presença da mãe	28
Gráfico 3 – Perfil por presença do pai	28
Gráfico 4 – Perfil por raça	28
Gráfico 5 – Perfil por creche	28
Gráfico 6 – Perfil por pré-escola	28
Gráfico 7 – Perfil por reprovação	29

LISTA DE TABELAS

Tabela 1 – Relação de variáveis	26
Tabela 2 – Estatística descritiva das variáveis	27
Tabela 3 – Resultado dos modelos de ML em matemática	37
Tabela 4 – Resultados dos modelos de ML em português	38
Tabela 5 – Matriz de confusão	40
Tabela 6 – Resultados de classificação em matemática	42
Tabela 7 – Resultados de classificação em português	42

LISTA DE ABREVIATURAS E SIGLAS

BF	Bolsa Família
EQM	Erro quadrático médio
IAS	Instituto Ayrton Senna
<i>Lasso</i>	<i>Least absolute shrinkage and selection operator</i>
ML	<i>Machine Learning</i>
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
QI	Quociente de inteligência
SAT	<i>Scholastic Assessment Test</i>
SEDUC - CE	Secretaria de Educação do estado do Ceará
SENNA	<i>Social and Emotional (or Noncognitive) Nationwide Assessment</i>
SPAECE	Sistema Permanente de Avaliação da Educação Básica do Ceará
SQR	Soma dos quadrados dos resíduos

SUMÁRIO

1	INTRODUÇÃO	14
2	REVISÃO DE LITERATURA	17
2.1	Aspectos não cognitivos	19
2.2	Relação entre aspectos não cognitivos e desempenho em testes padronizados	22
2.2.1	<i>Literatura internacional</i>	23
2.2.2	<i>Literatura nacional</i>	24
3	BASE DE DADOS	25
4	ESTRATÉGIA EMPÍRICA	29
4.1	Modelos econométricos tradicionais de previsão da relação entre habilidades não cognitivas e testes padronizados	29
4.2	Modelos de <i>Machine Learning</i>	30
4.2.1	<i>Lasso</i>	31
4.2.2	<i>Random Forest</i>	32
4.3	Abordagens econométricas	35
5	RESULTADOS	36
5.1	Qualidade da previsão	36
5.1.1	<i>Resultados da previsão</i>	37
5.2	Previsão da relação entre habilidades não cognitivas e desempenho em testes padronizados	39
5.2.1	<i>Definição de estudantes com elevado e baixo desempenho</i>	39
5.2.2	<i>Resultados do modelo de classificação</i>	41
6	CONCLUSÕES	42
	REFERÊNCIAS	45

1 INTRODUÇÃO

Educadores, psicólogos e economistas conduziram pesquisas nas últimas décadas debatendo a importância das habilidades não cognitivas para que um aluno tenha bons resultados na sua vida acadêmica e profissional. Essas habilidades socioemocionais estão associadas à auto estima, persistência, responsabilidade, desejo de cooperar entre outros (CARNEIRO; CRAWFORD; GOODMAN, 2007).

É comum pensar que o sucesso na vida, tem como um dos principais impulsionadores, o desempenho escolar do aluno. E esse sucesso está associado não só a bons resultados educacionais posteriores, mas em outros aspectos importantes da vida, como emprego, renda, saúde e bem-estar. Entretanto, alunos que possuem as mesmas notas na escola não necessariamente tiveram os mesmos resultados ao longo da vida.

Existe uma hipótese que diz que há alunos que possuem uma propensão maior a ter sucesso em comparação a outros devido o fato de terem desenvolvido melhor as habilidades sociais e emocionais, com isso esses alunos melhor respondem às tarefas de um século XXI cada vez mais volátil e incerto (OCDE, 2021).

Na escola é reconhecido o fato de que o aluno deve ser desenvolvido não só no seu nível cognitivo, mas também em um nível multidimensional envolvendo as habilidades sociais e emocionais. Porém, apesar de haver a necessidade de um desenvolvimento das competências não cognitivas, pouco esforço é voltado para o desenvolvimento dessas áreas e menos ainda voltado para a avaliação de intervenções que buscam esse amadurecimento das habilidades (SANTOS; PRIMI, 2022).

Em 2015, em conjunto com a Secretaria de Educação do estado do Ceará (SEDUC-CE), o Instituto Ayrton Senna (IAS) aplicou um instrumento, conhecido como SENNA, para 105.594 estudantes do 1º ano do ensino médio da rede estadual. Neste questionário, havia perguntas acerca das habilidades socioemocionais dos alunos, além de suas características socioeconômicas e de ambiente familiar.

Também há na base de dados as notas de português e matemática, pois o instrumento SENNA foi produzido conjuntamente com a prova do Sistema Permanente de Avaliação da Educação Básica do Ceará (SPAECE). Assim, pela

primeira vez pode-se ter um panorama socioemocional de um grupo de alunos do Ceará.

O IAS elaborou o instrumento *Social and Emotional (or Noncognitive) Nationwide Assessment* (SENNA) que mede o desenvolvimento da pessoa em cinco aspectos, seguindo a teoria do *Big Five inventory*, os cinco grandes domínios da personalidade. Esses aspectos são: conscienciosidade, abertura à novas idéias, extroversão, neuroticismo, também chamado de estabilidade emocional e amabilidade.

Estes aspectos são a junção de conceitos, desconhecidos inicialmente, que são obtidos por meio de análises de fatores realizada sobre os resultado de grandes questionários com diversas perguntas sobre comportamentos que representam todas as características relativamente estáveis de personalidade que um ser humano poderia manifestar (JOHN; SRIVASTAVA, 1999; ALMLUND *et al.*, 2011).

A literatura econômica tem encontrado evidências de que as habilidades socioemocionais dos estudantes são um bom preditor para o desempenho em português e matemática (CARNEIRO; CRAWFORD; GOODMAN, 2007), (DUNCAN; MAGNUSON *et al.*, 2011), (FERGUSON; HORWOOD, 1998), (CUNHA; HECKMAN; SCHENNACH, 2010). A maior parte dessas evidências baseiam-se em modelos lineares, que desconsideram potenciais efeitos interativos e não lineares entre as diferentes medidas das habilidades socioemocionais.

No entanto, não existe uma forma direta de introduzir efeitos interativos ou não lineares para prever o desempenho dos estudantes. A teoria não fornece estrutura suficiente que permita definir adequadamente uma forma funcional apropriada¹.

Nesse sentido, este trabalho busca adotar uma abordagem dirigida por dados para definir a forma funcional ótima para a relação entre habilidades socioemocionais e performance em matemática e português dos estudantes do ensino médio. Isto é, quer se responder a seguinte pergunta: é possível prever o rendimento escolar com base nas habilidades socioemocionais utilizando técnicas

¹ A maior parte da literatura teórica sobre o tema baseia-se na teoria de formação das habilidades cognitivas e não cognitivas de Cunha, Heckman e Schennach (2010) segundo o qual os autores definem a possibilidade de existência de efeito não lineares e complementares entre as diferentes habilidades.

computacionais de *machine learning* (ML) que definem a melhor forma funcional de acordo com o poder de previsão de diferentes modelos?

Assim, será verificado o desempenho em termos preditivos de diferentes modelos de ML contra a abordagem tradicional (modelo linear sem interações ou não linearidades). Caso se verifique que tais modelos de ML possuem um desempenho preditivo melhor que os modelos tradicionais, será estimado o quanto de variação no desempenho de português e matemática é devido a mudanças nas habilidades socioemocionais.

Modos de melhorar esse poder preditivo tem sido pesquisado, um exemplo é um recente estudo de Swift e Peterson (2019), concluem que “Simplesmente especificando um contexto para uma medida de personalidade, a validade preditiva de resultados congruentes com o contexto pode ser significativamente aumentada em até 30%.”

Este trabalho também tem como objetivo aumentar a qualidade da previsão feita pelas habilidades socioemocionais utilizando essas técnicas computacionais e estatísticas de ML.

Como destacado anteriormente as habilidades socioemocionais tem demonstrado correlação com os resultados escolares. Assim este trabalho busca construir um modelo de aprendizado de máquina que possa prever o desempenho escolar em matemática e em português de alunos da rede pública com base nas suas habilidades socioemocionais obtidas por meio do instrumento SENNA.

O presente trabalho está estruturado em seis seções, incluindo a introdução e a conclusão. Na primeira seção, encontra-se a introdução, onde são expostos a contextualização sobre o assunto, a metodologia do trabalho e seus objetivos. Na segunda seção há a revisão de literatura explorando trabalhos e teorias que fundamentam esse estudo. A terceira seção aborda a base de dados mostrando uma análise descritiva e diversos aspectos relacionados às características dos alunos quanto à sua vida individual, familiar, escolar e desempenho nas provas. Na quarta seção é mostrado a metodologia das análises produzidas, é explicado os modelos de ML e métodos estatísticos utilizados para a análise. Na quinta seção estão dispostos os resultados dos modelos de previsão e análise dos resultados. E, na última seção, encontram-se as conclusões.

2 REVISÃO DE LITERATURA

Desde o século passado o interesse no estudo da personalidade vem crescendo. Muitos psicólogos estavam empreendendo pesquisas sobre as diferenças individuais no comportamento e traços das pessoas. Junto com a personalidade também eram estudadas inteligência, interesses, motivação e outros.

Ao longo dos anos, diversos modos de tentar medir e tentar qualificar a educação foram criados, desde o teste de QI (Quociente de inteligência) com Alfred Binet no século XIX até as teorias discutidas na atualidade, buscando meios para tornar a educação cada vez mais eficiente, com o intuito de dar mais oportunidade às pessoas.

Também no desejo de contribuir para a educação, psicólogos que estudam a personalidade procuraram formas de mensurar a personalidade e como ela impacta nos acontecimentos e escolhas do indivíduo.

Alfred Binet, o grande idealizador do teste de QI, já naquele tempo percebia que o rendimento escolar não dependia apenas da habilidade cognitiva da pessoa, em um de seus escritos ele já dizia:

... admite outras coisas além da inteligência; para ter sucesso nos estudos é preciso ter qualidades que dependem de atenção, vontade e caráter; por exemplo, uma certa docilidade, uma regularidade de hábitos e, sobretudo, continuidade de esforços. Uma criança, mesmo que inteligente, aprenderá pouco na aula se nunca ouvir, se passar o tempo pregando peças, dando risadinhas, se vadiando (BINET; SIMON, 1916, p. 254).

Charles Spearman um psicólogo inglês que também estudava estatística e modelos para mensurar inteligência produziu um teste cognitivo que ficou conhecido como fator "g", que busca capturar a inteligência. Também ele relata em seus estudos como outros fatores além dos cognitivos impactam no sucesso ou falha nos muitos domínios da vida:

Quais são os principais traços de personalidade que, interagindo com g , se relacionam com as diferenças individuais de realização e sucesso vocacional? O traço de personalidade mais universal é a consciência, ou seja, ser responsável, confiável, atencioso, organizado e persistente (Jensen, 1998, p. 575).

Essa percepção de que habilidade não cognitivas impactam em resultados é estudado não só em rendimento escolar mas também em saúde, crime, resultados no mercado de trabalho e orientação política. Segundo Almlund *et al* (2011), “Um crescente grupo de evidências sugere que as medidas de personalidade, especialmente aquelas relacionadas à Conscienciosidade e, em menor grau, ao Neuroticismo, preveem uma ampla gama de resultados.” Mas também fazem a ressalva afirmando que o poder preditivo de qualquer medida de personalidade em particular tende a ser menor do que o poder preditivo do QI, mas em alguns casos rivaliza com ele.

A psicologia da personalidade tenta descrever a pessoa como um todo. Ele considera tanto os traços universais quanto as diferenças individuais. Examina as maneiras pelas quais as pessoas são únicas.(ALMLUND, 2011).

Dentro da psicologia da personalidade é definido traços da personalidade como “Padrões relativamente duradouros de pensamentos, sentimentos e comportamentos que refletem a tendência de responder de certas maneiras sob certas circunstâncias.” (ROBERTS, 2009).

Assim os traços são como as principais unidades da personalidade, quando alguém pergunta como é sua personalidade geralmente a resposta são os traços da personalidade, se a pessoa é honesta, alegre, se irrita fácil etc.

Logo, os traços da personalidade sendo um determinante da personalidade, se faz necessário, de forma a ser cuidadosamente distinguidos da expressão completa da personalidade.

Essa expressão completa da personalidade é gerada pela interação dos traços com outros fatores. A personalidade é vista como um sistema de comportamentos, pensamentos e sentimentos que emergem dos componentes que interagem. (ALMLUND, 2011).

Os psicólogos se juntaram e começaram a pegar do dicionário palavras que descrevem o comportamento humano e seus traços, eliminaram os sinônimos e começaram a clusterizar essas palavras conforme fossem de característica similar e apresentassem uma ligação (ALMLUND, 2011).

Desses longos anos de estudo surgiu uma taxonomia consensual aceita por muitos membros da área. O *Big Five* agrupa dentro de 5 traços principais várias características que são chamadas de facetas. É discutido mais do *Big Five* a frente.

Os estudos lexicais que produziram o modelo do *Big Five* quase sempre incluíram descritores de habilidades, e a personalidade é um conceito suficientemente largo para cobrir os dois: habilidades e traços da personalidade. (DEYOUNG *et al*, 2013)

2.1 Aspectos não cognitivos

Nesse sentido, mensurar a personalidade ou seus traços foi sendo pautado por esses padrões observados, segundo as interações das pessoas e suas respostas sob suas certas circunstâncias.

Conforme implicado pelo modelo teórico no trabalho de Thiel e Thomsen (2009), as habilidades cognitivas e não cognitivas evoluem e se cultivam em conjunto. No entanto, as habilidades cognitivas (em particular, o QI) provaram se estabilizar muito mais cedo no curso da vida. Esse fato acaba sendo extremamente útil ao concluir sobre a formação da personalidade usando dados que compreendem principalmente medidas de habilidades cognitivas e desempenho escolar. (THIEL; THOMSEN, 2009)

Embora os traços de personalidade sejam relativamente estáveis em todas as situações, eles não são imutáveis. Eles mudam ao longo do ciclo de vida. (HECKMAN; KAUTZ, 2012).

Na mensuração da personalidade duas questões principais se levantam: a primeira se refere ao fato de como elas são coletadas nas pesquisas, se são auto relatos dos indivíduos das pesquisas ou se os dados são coletados por meio de um especialista que faz a análise comportamental necessária para a mensuração dos traços da personalidade da pessoa. Essa questão é discutida mais detalhadamente à frente.

A segunda pergunta levantada sobre a mensuração da personalidade é como será a agregação dos traços da personalidade. Há uma variedade de taxonomias associadas ao mapeamento da personalidade como um todo. Neste trabalho é utilizado como já citado o modelo *Big Five* que é o mais aceito pelos estudiosos da área.

O instrumento psicológico que mapeia esses aspectos socioemocionais, é mais conhecido e utilizado para esse objetivo é chamado de *Big Five inventory*. Ela

contém a taxonomia dos traços da personalidade que é mais aceita, de forma ampla mas não universal (ALMLUND *et al.*, 2011).

Por meio de variáveis obtidas por questionários esses itens são agrupados em cinco domínios de personalidade presentes no *Big Five Inventory*, a seguir é descrito cada um deles e suas facetas:

Conscienciosidade é o constructo que representa a tendência a ser organizado, trabalhador duro e responsável. A pessoa conscienciosa é caracterizada como alguém com autocontrole, com foco, eficiente, organizado, autônomo, disciplinado, sem impulso e orientado para seus objetivos.

As facetas da conscienciosidade são: competência, capacidade de lidar com desafios e dificuldades da vida. Ordem: ter clareza e abordagem metódica das tarefas; clareza na vida. Senso de Dever: governado pela consciência; muito dogmático em seus valores. Realização-Esforço: disposição para trabalhar duro; fortemente motivado por metas. Autodisciplina: capacidade de seguir com as tarefas e limitar a distração. Deliberação: tendência a contemplar cuidadosamente as decisões antes de agir.

Extroversão é a tendência a se aproximar energicamente do mundo externo, das pessoas e das coisas (em vez do mundo interno da experiência subjetiva). O indivíduo extrovertido é caracterizado como sendo amigável, sociável, autoconfiante, enérgico, aventureiro e entusiasmado.

As facetas da extroversão são: Calorosidade: facilidade de aproximar-se das outras pessoas; carinhoso e amigável. Gregariedade: preferência pela companhia dos outros; evitar estar sozinho. Assertividade: tendência para liderar e dominar situações sociais. Atividade: disposição energética; estilo de vida acelerado e propensão para ocupação. Buscador de Entusiasmo: desejo de alegria e estimulação; preferência por ruído. Emoções Positivas: tendência a experimentar emoções positivas; inclinação para o otimismo.

Estabilidade Emocional (Neuroticismo) representa uma tendência a experimentar estados emocionais negativos. Indivíduos com baixa instabilidade emocional são mais difíceis de serem perturbados e são menos reativos emocionalmente. Eles tendem a ser calmos, emocionalmente estáveis, e livres de sentimentos negativos persistentes. No entanto, a escassez de sentimentos negativos não significa necessariamente que estes indivíduos experimentem muitos sentimentos positivos.

As seis facetas do neuroticismo são: Ansiedade: medo, tensão, inquietação. Hostilidade Enfurecida: tendência a experimentar frustração e amargura, bem como raiva. Depressão: propensão a experimentar sintomas depressivos, como perda de energia, dificuldade de concentração e problemas com o sono. Autoconsciência: desconforto em torno dos outros; frequentes experiências de vergonha e constrangimento. Impulsividade: incapacidade de controlar desejos ou impulsos. Vulnerabilidade: dificuldade em lidar com o estresse; dependência de outros para apoio.

O traço de Amabilidade representa uma orientação comunitária e pró-social, englobada por tendências à polidez e à compaixão. É a tendência para ser compassivo e cooperante em vez de suspeito e antagonista face aos outros. Indivíduos “amáveis” valorizam a boa relação com os outros. Eles são geralmente respeitosos, amigáveis, generosos, prestáveis e dispostos.

As facetas da amabilidade são: Confiança: inclinado a acreditar que os outros são honestos e bem-intencionados. Simplicidade e sinceridade na expressão de opiniões e pensamentos. Altruísmo: fortemente movido e dedicado à promoção do bem-estar dos outros; extremamente generoso. Conformidade: inibição da agressão, deferência aos outros em conflitos interpessoais. Modéstia: humildade em falar de realizações próprias. Empatia: altamente simpático e preocupado com os outros.

Abertura a Novas Experiências é a tendência de estar aberto a novas experiências estéticas, culturais e intelectuais. O indivíduo que está aberto a novas experiências é caracterizado como sendo imaginativo, artístico, excitante, curioso e não convencional, ao mesmo tempo que possui uma ampla gama de interesses.

As facetas da abertura à experiência incluem: Fantasia: vida mental ativa e forte imaginação; vida interior rica e criativa. Estética: forte valorização da arte e da beleza; movido pela poesia, arte e música. Sentimentos: receptivos aos próprios sentimentos e emoções; sentir emoções muito intensamente. Ações: dispostos a explorar novos lugares, experimentar novos alimentos e experimentar novas atividades. Ideias: curiosidade intelectual; gozo por argumentos filosóficos e quebra-cabeças. Valores: pronto para explorar e avaliar seus próprios valores sociais, políticos e religiosos.

Com essas habilidades mensuráveis também surgiu o debate sobre a confiança de uma auto avaliação, tendo em vista que o indivíduo tem um viés sobre as avaliações que faz de si mesmo.

Segundo Thiel e Thomsen (2009) “As medidas auto relatadas são convincentes devido à sua implementação simples, mas pressupõem implicitamente que a personalidade é conscientemente avaliável, o que geralmente não ocorre.”

Já no caso de avaliações feitas por especialista é possível ter mais confiança e como podem haver distorções na auto avaliação o esforço por um especialista pode valer a pena.

Roberts, Harms, Smith, Wood e Webb (2006) revisaram vários estudos nos quais a utilidade de ambos os tipos de classificação foi testada em vários domínios. Concluem que a resposta a esta questão depende da característica de interesse.

Para traços referentes a relações sociais típicas, como conhecer um estranho ou ter uma discussão, as classificações de observadores especialistas tendem a prever o comportamento melhor do que os auto-relatos, uma vez que o potencial de desordem na autopercepção é alto, podendo tornar a variância do experimento muito grande.

Além disso, no caso de análise de um especialista existe o potencial para fingir, ou seja, subestimar e superestimar a verdadeira situação. Isso acontece em maior proporção para medidas de personalidade do que no caso de testes padronizados para habilidades cognitivas, como o teste de QI (THIEL; THOMSEN, 2009).

Ou seja, a personalidade do indivíduo também tem uma relação responsiva com relação aos incentivos presentes em determinadas situações. E ele usará sua percepção aliada aos incentivos para parecer ao especialista conforme lhe for mais vantajoso.

2.2 Relação entre aspectos não cognitivos e desempenho em testes padronizados

Há uma variedade de pesquisas na literatura que encontraram importantes correlações entre as habilidades socioemocionais e os resultados educacionais. (ALMLUND *et al.* 2011). Muitos trabalhos têm mostrado como essas

habilidades socioemocionais possuem um poder explicativo (CARNEIRO; CRAWFORD; GOODMAN, 2007), (DUNCAN; MAGNUSON *et al.*, 2011), (FERGUSON; HORWOOD, 1998), (CUNHA; HECKMAN; SCHENNACH, 2010) sobre resultados escolares.

A seguir são explorados alguns trabalhos e suas descobertas sobre essa relação existente entre as habilidades socioemocionais e os resultados em testes padronizados.

2.2.1 Literatura internacional

Pesquisas como a de Mischel *et al* (1989) mostrou que crianças que possuem menor capacidade de postergar recompensa tal como medido no Teste do Marshmallow² obtiveram notas inferiores no exame padronizado SAT³, utilizado no ingresso à universidade nos Estados Unidos. A correlação entre essas variáveis é surpreendentemente elevada: 0,42 e 0,57 nos exames de linguagem e matemática, respectivamente.

Duckworth e Seligman (2005) mostram com seu estudo que a autodisciplina medida no início do ano letivo explica mais de duas vezes melhor a variância da nota de alunos da oitava série do que a explicada pela inteligência.

Nesse contexto diversas pesquisas têm explorado como as habilidades socioemocionais impactam no resultado escolar. Mostrando como essas habilidades não cognitivas têm efeitos sobre resultados não só na vida escolar mas também profissional.

Em uma pesquisa feita pela OCDE em vários países ao redor do mundo foi exposto que alguns dos resultados apontam para as habilidades sociais e emocionais como boas preditoras para as notas escolares. A pesquisa ainda ressalta que:

Ser intelectualmente curioso e persistente é a habilidade social e emocional mais fortemente relacionada às notas escolares tanto para os de 10 quanto para os de 15 anos de idade em todas as três disciplinas (matemática,

² No Teste do Marshmallow, o pesquisador oferece um doce à criança e sugere que se ela resistir à tentação de comê-lo enquanto ele se ausenta, ganhará um segundo doce. As reações da criança são filmadas e o tempo de espera da criança antes de comer o doce é uma medida de sua capacidade de postergar recompensa.

³ O SAT (Scholastic Assessment Test) é um teste padronizado aplicado a estudantes do ensino médio nos Estados Unidos, que serve de critério para admissão nas universidades norte-americanas.

leitura e artes). Em menor grau, mas ainda significativo, ser mais assertivo e responsável também está positivamente relacionado a melhores notas escolares. (OCDE, 2021, p. 65)

Outros resultados do trabalho ainda mostram que estudantes de 15 anos de idade que relataram ser mais resistentes a estresse, em média, têm notas de leitura baixas, e aqueles mais sociáveis têm notas de matemática mais baixas. Aqueles que relataram ser mais criativos têm, em média, notas mais baixas em matemática.

Um ponto que vale ressaltar nesse trabalho da OCDE é que, mesmo com o fato de que as habilidades socioemocionais diferem entre os alunos de acordo com a sua origem socioeconômica e gênero, os dados deles demonstraram que, alunos com o mesmo status socioeconômico e gênero que possuem melhores habilidades socioemocionais são mais propensos a obter melhores notas. Portanto, apesar do status socioeconômico e do gênero, as habilidades socioemocionais dos alunos desempenham um papel decisivo no desempenho escolar.

2.2.2 Literatura nacional

Estudos na área da educação mostram que abertura a novas ideias e amabilidade possuem um efeito na nota de português, enquanto que conscienciosidade e estabilidade emocional são os domínios que influenciam a nota de matemática (SANTOS; BERLINGERI; CASTILHO, 2017). Há trabalhos que mostram que alunos que têm o aspecto de ser mais aberto às novas experiências também escolhiam cursos mais difíceis, com maior carga de matemática e faltam menos as aulas.

A autodisciplina está dentro do aspecto da Conscienciosidade e é ela o domínio mais associado às diversas medidas de sucesso no aprendizado (SANTOS; BERLINGERI; CASTILHO, 2017).

O Instituto Ayrton Senna em 2013 fez a aplicação do instrumento SENNA no Rio de Janeiro, junto com o SAERJINHO, que é um sistema de avaliação bimestral do processo de ensino e aprendizagem nas escolas da rede estadual do Rio de Janeiro. Coletando dados de mais de 24 mil estudantes.

Em sua pesquisa eles mostraram que o impacto de Conscienciosidade sobre o desempenho de matemática é 4,5. Isso significa que, se for elevado a

conscienciosidade de um estudante que hoje está entre os que possuem menor nível desse atributo (isto é, um estudante no percentil 25), até que ele esteja entre os que possuem maior nível do mesmo (alguém no percentil 75), o desempenho desse estudante subiria o equivalente a 4,5 meses de aprendizado. (SANTOS; PRIMI, 2014).

No caso de português, o constructo do *Big Five* que se mostrou maior impacto foi o de Abertura a Novas Experiências, mostrando que o desempenho de um estudante subiria o equivalente a mais de 4 meses de aprendizado.

Segundo os resultados, a Extroversão tem impacto negativo no desempenho das notas. Segundo a literatura internacional a Extroversão apresenta efeitos difíceis de se captar em exercícios estatísticos, pois a relação deste atributo com resultados educacionais não parece ser monotônica, isto é, nem sempre ter mais (ou menos) deste atributo indica uma vantagem na sala de aula.

Os níveis de Amabilidade e Estabilidade Emocional não parecem estatisticamente associados ao desempenho educacional. A pesquisa do instituto faz ainda diversas análises mais profundas verificando a diferença dos efeitos entre as séries mais novas (alunos do 5º ano) e alunos mais velhos (1º e 3º ano do ensino médio), e mostram que o papel dos atributos socioemocionais é maior nas séries menores. Há também análises relacionadas ao gênero dos alunos e se o aluno é economicamente vulnerável.

3 BASE DE DADOS

Este trabalho analisou dados secundários que o Instituto Ayrton Senna, em parceria com a Seduc, obtiveram ao aplicar o instrumento SENNA em alunos do 1º ano do ensino médio na rede pública do Ceará. Por meio de questionários foram obtidas as variáveis que medem o nível de desenvolvimento das habilidades socioemocionais de cada aluno.

A aplicação foi feita em 2015 junto com a prova do SPAECE para 105.594 alunos espalhados em 643 escolas do estado. Além dos resultados do instrumento SENNA, na base ainda há as respostas dos estudantes a respeito de suas características individuais (sexo, idade e raça), do ambiente familiar em que vivem, características educacionais e comportamento de seus pais. Na base de dados

também há os resultados dos testes de português e matemática realizados como parte do sistema de avaliação da rede estadual do Ceará (SPAECE).

A tabela 1 traz uma descrição das variáveis utilizadas e a tabela 2 contém a estatística descritiva das variáveis, sobre as variáveis *dummys* é apresentado a proporção delas no campo média.

Tabela 1 - Relação de variáveis

Variável	Descrição
Constructos Socioemocionais (Não Cognitivos)	
Abertura a Nova Ideias	Análise fatorial
Consiensiodidade	
Extroversão	
Amabilidade	
Estabilidade Emocional (Neuroticismo)	
Características Individuais	
Gênero	Dummy (1 = feminino)
Idade	Inteiro
Raça	Dummy (1 = preto)
Ambiente Familiar	
Ajuda a mãe	Dummy (1 = sim)
Ajuda o pai	Dummy (1 = sim)
Total de pessoa na família	Inteiro
Irmãos mais novos	Inteiro
Irmãos mais velhos	Inteiro
Educação da mãe	7 níveis educacionais
Características Educacionais	
Fez Creche	Dummy (1 = sim)
Já foi reprovado	Dummy (1 = sim)
Fez Pré-escola	Dummy (1 = sim)
Desempenho Educacional	
Português	Nota padronizada (m = 0, v = 1)
Matemática	Nota padronizada (m = 0, v = 1)

Fonte: Elaborado pelo autor

Tabela 2 - Estatística descritiva das variáveis

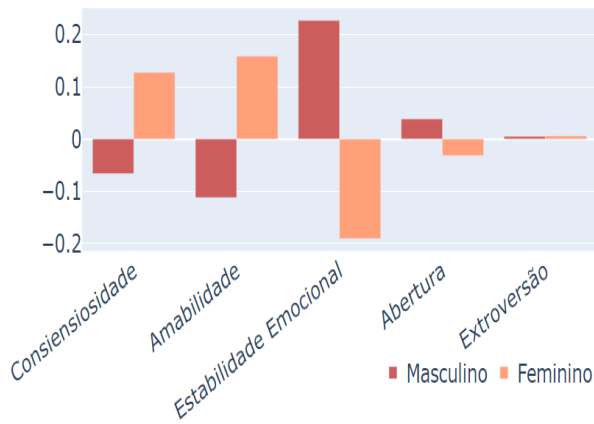
Variável	Média	Variância	Máximo	Mínimo
Constructos Socioemocionais (Não Cognitivos)				
Abertura	-0.427	1.293	5.949	-7.499
Conscienciosidade	-0.343	1.630	5.250	-7.082
Extroversão	-0.251	1.834	8.401	-10.149
Amabilidade	-0.672	3.403	8.491	-11.739
Estabilidade Emocional (Neuroticismo)	-0.212	1.557	8.195	-8.881
Características Individuais				
Gênero	0.515	0.249	-	-
Idade	16.96	3.206	21.00	8.00
Raça	0.071	0.066	-	-
Ambiente Familiar				
Ajuda a mãe	0.049	0.047	-	-
Ajuda o pai	0.132	0.115	-	-
Total de pessoa na família	4.064	3.399	10	1
Irmãos mais novos	2.013	1.924	6	1
Irmãos mais velhos	2.17	2.202	6	1
Características Educacionais				
Fez Creche	0.632	0.232	-	-
Já foi reprovado	0.349	0.227	-	-
Fez Pré-escola	0.815	0.15	-	-
Desempenho Educacional				
Português	255.446	2102.541	379.633	134.850
Matemática	257.753	2819.639	439.061	140.288

Fonte: Elaborado pelo autor

Os gráficos abaixo apresentam os dados das variáveis socioemocionais dos alunos de acordo com características familiares, individuais, educacionais. O gráfico 1 mostra que o perfil por gênero tem o resultado similar à literatura internacional (SOTO *et al.*, 2011), apresentando meninas com níveis maiores em conscienciosidade, amabilidade e extroversão e níveis menores de estabilidade emocional.

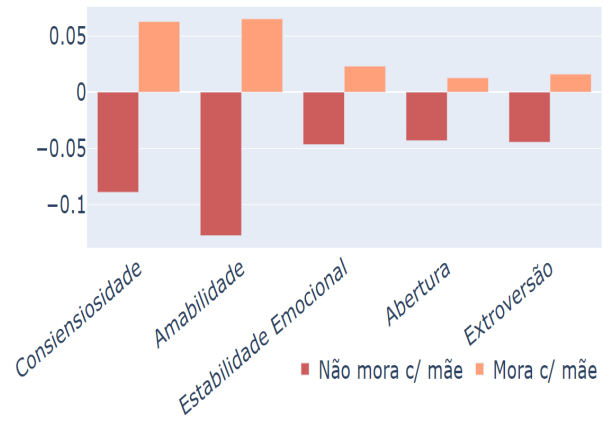
Os gráficos de 2 a 8 apresentam resultados de forma bem intuitiva mostrando que alunos com mais presença materna e paterna, que fizeram pré-escola, creche e não reprovaram apresentam níveis maiores de desenvolvimento das habilidades socioemocionais.

Gráfico 1 - Perfil por gênero



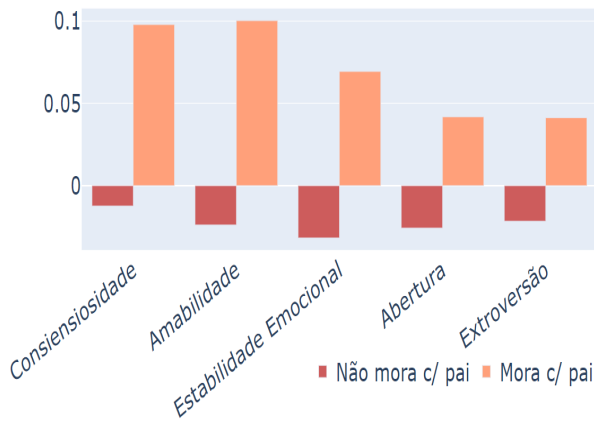
Fonte: Elaborado pelo autor

Gráfico 2 - Perfil por presença da mãe



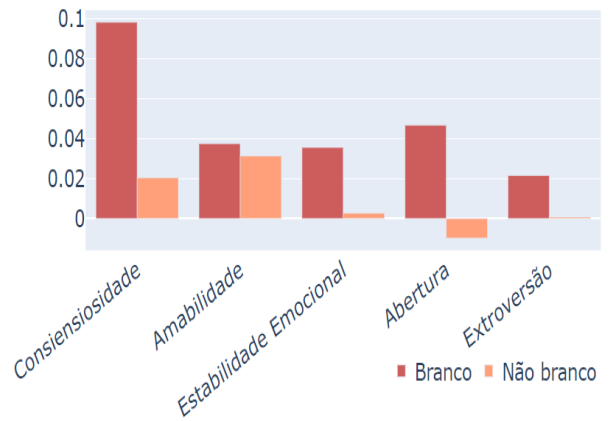
Fonte: Elaborado pelo autor

Gráfico 3 - Perfil por presença do pai



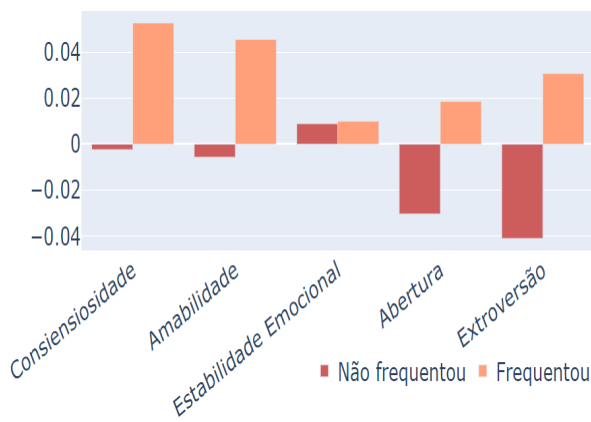
Fonte: Elaborado pelo autor

Gráfico 4 - Perfil por raça



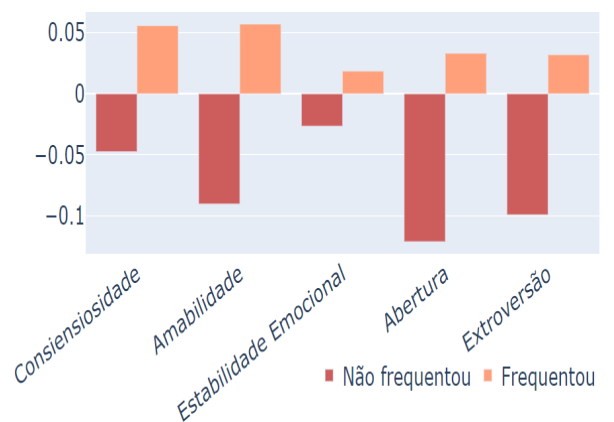
Fonte: Elaborado pelo autor

Gráfico 5 - Perfil por creche



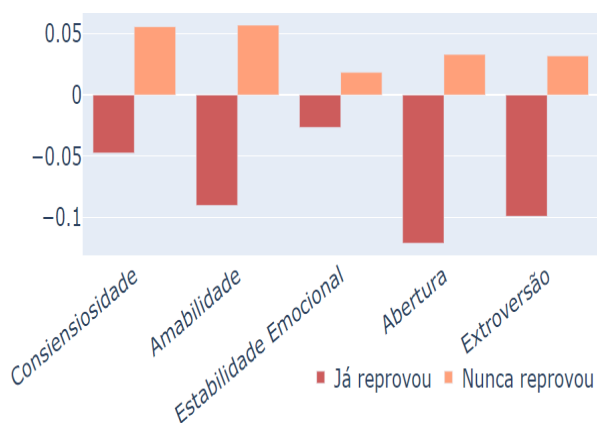
Fonte: Elaborado pelo autor

Gráfico 6 - Perfil por pré-escola



Fonte: Elaborado pelo autor

Gráfico 7 - Perfil por reprovação



Fonte: Elaborado pelo autor

4 ESTRATÉGIA EMPÍRICA

Para investigar o poder de previsão das habilidades socioemocionais sobre o desempenho escolar foram escolhidos três modelos de ML. O modelo linear representa as abordagens tradicionais e estudos que já mostraram a correlação entre as variáveis socioemocionais e os resultados em testes padronizados.

O segundo modelo de ML, o *Random Forest* foi escolhido com objetivos de capturar os efeitos interativos e não lineares entre as variáveis de habilidade socioemocional. E o terceiro modelo, *Lasso*, foi escolhido por sua característica de fazer seleção de variáveis buscando selecionar as melhores variáveis preditoras para a previsão dos valores. A seguir, com base no texto de James *et al* (2013), são explicados de forma mais profunda cada um dos modelos e as fórmulas utilizadas para a construção dos modelos de previsão.

4.1 Modelos econométricos tradicionais de previsão da relação entre habilidades não cognitivas e testes padronizados

O modelo linear é um modelo amplamente utilizado para fazer estimativas quantitativas. Os modelos lineares apesar de não serem os melhores previsores, são simples de se entender, de estimar e devolvem bons resultados. Ou seja, nem sempre vale a pena utilizar modelos de ML que são computacionalmente mais pesados. Embora haja algumas abordagens mais modernas de aprendizado

estatístico, a regressão linear ainda é um método de aprendizado estatístico útil e amplamente utilizado (JAMES *et al*, 2013).

Para investigar as relações das habilidades socioemocionais com o nível de aprendizado em português e matemática é utilizado o modelo linear. Utilizando apenas um constructo do *Big Five* mais as características dos estudantes como gênero, raça, idade.

$$y_i = x_i\beta + SE_i\gamma + \epsilon_i \quad (1)$$

as notas de matemática e português são indicadas por y_i , x_i é o vetor de características dos estudantes, características essas individuais, educacionais e familiares; β é o vetor de parâmetros que mede a sensibilidade da nota a estas características. A habilidade sócio emocional do aluno i é dada por SE_i e γ representa o parâmetro associado à habilidade sócio emocional. Por fim, ϵ_i representa o erro.

O mesmo modelo linear é utilizado por Santos, Berlingeri e Castilho (2017), porém o objetivo dele era estudar a causalidade das habilidades socioemocionais e para isso ele usa efeitos fixos. No estudo feito por esse trabalho se está preocupado com a previsão, por isso não é utilizado os efeitos fixos, é feito apenas os controles para as características dos alunos como também há no modelo citado.

Outras especificações também foram mensuradas, com objetivo de controlar para as outras habilidades sócio emocionais e para outras características como escolaridade da mãe, número de irmãos, se foi reprovado em alguma série, se esteve em creche e outros.

4.2 Modelos de *Machine Learning*

Com o surgimento do *Big Data* os métodos de estimação tornaram-se inadequados devido o número de variáveis terem se tornado maior que o número de observações $p > n$. Dessa forma outros métodos surgiram que buscaram reduzir o p com o objetivo de restar apenas as variáveis consideradas relevantes,

principalmente para previsão. Os métodos mais conhecidos são os de seleção de subconjunto, de regularização e de redução de dimensão.

Em muitos casos pode ser difícil escolher quais variáveis são as melhores preditoras para o problema que se está lidando. Uma alternativa a fazer um subconjunto das variáveis é utilizar métodos de regularização que restringem ou reduzem as estimativas dos coeficientes para zero ou muito próximo de zero, deixando com um valor que fará diferença somente as variáveis que apresentarem alguma relevância para o modelo.

Este trabalho comparará o desempenho preditivo dos modelos lineares a dois modelos de ML que são recorrentemente utilizados na literatura: o *Lasso* (*least absolute shrinkage and selection operator*) e o *Random Forest*. Ambos permitem a introdução de interações entre as variáveis de forma dirigida por dados, isto é, a definição da forma funcional final é obtida de acordo com a capacidade de previsão dos modelos.

4.2.1 Lasso

Dentre os modelos de *Shrinkage* existem o *lasso* e o *ridge*. O *lasso* diferentemente do modelo de *Ridge* tem a vantagem de que, depois da aplicação do método, sobram somente as variáveis que foram consideradas relevantes para o modelo. O modelo de *Ridge* reduz as variáveis com menos relevância até muito próximo de zero, mas elas nunca recebem o valor zero, dessa forma depois da aplicação desse método todos os parâmetros continuarão lá.

O método *lasso* para realizar a redução utiliza uma penalidade dada por: $\|\beta\|_1 = \sum |\beta|$. Assim o objetivo será otimizar:

$$\min \left(y_i - \beta_o - \sum_{j=1}^p \beta_j x_{ij} \right)^2 \quad (2)$$

sujeito a:

$$\sum_{j=1}^p |\beta| \leq s \quad (3)$$

O parâmetro s é o valor que mede o grau da restrição que acontecerá sobre as variáveis.

Pode-se ver essas equações da seguinte forma. O objetivo do método é encontrar o conjunto de estimativas dos coeficientes que levam ao menor SQR (soma do quadrado dos resíduos) sujeito à restrição s de tamanho $\sum_{j=1}^p |\beta_j|$. Quando s é extremamente grande então a restrição não é muito limitante. Se o s realmente for muito grande a solução produzida será igual a de mínimos quadrados ordinários. Em contraste, se s for pequeno, então $\sum_{j=1}^p |\beta_j|$ deve ser pequeno para evitar a violação da restrição.

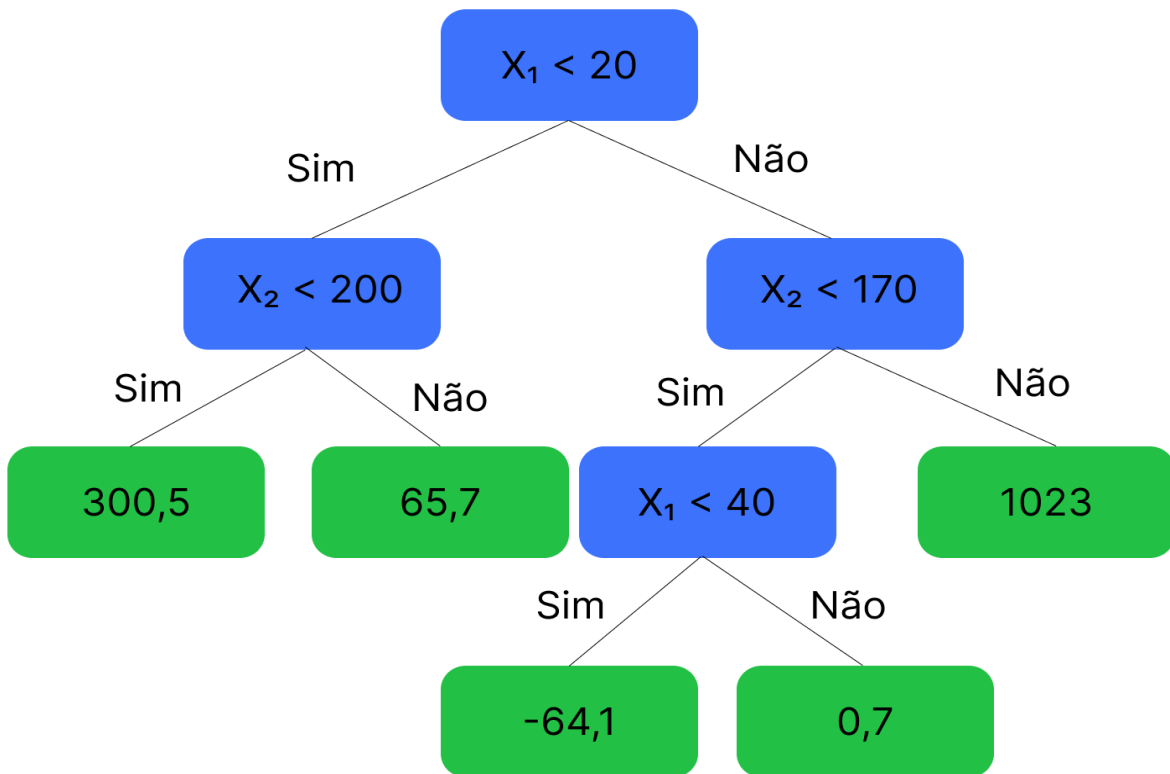
O método *lasso* além de realizar o encolhimento (*shrinkage*) das variáveis também realiza o que é chamado de *feature selection*, o modelo seleciona os melhores parâmetros forçando alguns deles, entendidos como não importantes, a zero. Dessa forma o número de variáveis será menor depois de se aplicar o método *lasso*, restando apenas a variáveis mais importantes.

4.2.2 Random Forest

O modelo de *Random Forest*, ou floresta aleatória, é a estimação de várias árvores de decisão (*decision trees*), e faz parte dos métodos *ensemble*. A característica dos métodos *ensemble* é que eles fazem a combinação de vários modelos para obter um único resultado, com isso o modelo fica mais robusto e complexo. Existe o lado negativo, de exigir um maior custo computacional, porém, isso geralmente vem acompanhado com melhores resultados.

O modelo de árvores de decisão consiste em uma série de regras de divisão, começando no topo da árvore. O algoritmo cria uma estrutura similar a um fluxograma, com 'nós' onde uma condição será verificada, e se atendida o fluxo segue por um ramo, se não atendida, por outro.

Figura 1 - Nós da árvore de decisão



Fonte: Elaborado pelo autor.

Na figura 1 é mostrado de forma simples como funciona o algoritmo de árvore de decisão de regressão. Cada nó possui dois ramos, assim conforme cada observação for sendo verificada, elas vão sendo divididas entre os ramos. Para cada observação que segue por um ramo é feita a previsão, que será a média dos valores de resposta para as observações de treinamento. O objetivo é encontrar ramos que minimizam a soma dos quadrados dos resíduos, SQR, dados por:

$$\sum_{j=1}^J \sum_{i \in R_j} \left(y_i - \hat{y}_{R_j} \right)^2 \quad (4)$$

onde \hat{y}_{R_j} é a resposta média para as observações de treinamento dentro do ramo.

Para fazer a divisão do ramo é utilizada uma abordagem conhecida como divisão binária recursiva, esse método busca fazer a melhor divisão possível no topo da

árvore (no ponto em que todas as observações pertencem a um único ramo), em vez de deixar para fazer uma escolha melhor mais para frente.

Para realizar a divisão binária recursiva primeiro é selecionado o preditor X_j e o ponto de corte s de modo que a divisão do espaço do preditor nas regiões $\{X|X_j < s\}$ e $\{X|X_j \geq s\}$ leve à maior redução de SQR. Para cada j se tem:

$$R_1(j, s) = \{X|X_j < s\} \text{ e } R_2(j, s) = \{X|X_j \geq s\} \quad (5)$$

R_1 e R_2 representam ramos que a árvore fez. Procura-se o s e j que minimiza a equação:

$$\sum_{i: x_i \in R_1(j,s)} (y_i - \hat{y}_{R_1})^2 + \sum_{i: x_i \in R_2(j,s)} (y_i - \hat{y}_{R_2})^2 \quad (6)$$

onde \hat{y}_{R_1} é a resposta média para as observações de treinamento em $R_1(j, s)$, e \hat{y}_{R_2} é a resposta média para as observações de treinamento em $R_2(j, s)$. Encontrar os valores de j e s que minimizam (6) pode ser feito rapidamente, especialmente quando o número de variáveis p não é muito grande. Porém também pode demorar muito se houver muitas variáveis, uma alternativa é construir árvores até o SQR não reduzir mais. Esta estratégia gera árvores menores, mas com menor variância o viés se apresenta maior.

A floresta aleatória (*Random Forest*) como explicado, é a junção de várias árvores de decisão, consistindo assim em um método que melhora o desempenho das árvores de regressão. Um ponto a ser ressaltado neste método é que ao invés de se utilizar todas as variáveis, p , para construir a árvore, utiliza-se $m \ll p$ variáveis, escolhidas de forma aleatória. Normalmente, $m \approx \sqrt{p}$.

O procedimento de *Random Forest* mostra uma grande consistência por fazer esse subconjunto dos preditores m , com essa limitação de $(p - m) / p$ nos previsores, o regressor mais forte não estará presente sempre, com isso cada árvore será bem diferente do outra, reduzindo a correlação entre elas, diminuindo assim a variância do modelo.

O algoritmo do modelo de *Random Forest* segue como o de árvore de decisão, calculando a média das árvores criadas a partir do subconjunto de previsores.

Na segunda parte deste trabalho temos um problema de classificação, árvores de classificação são muito semelhantes às árvores de regressão, a diferença é que o resultado será qualitativo. É percebida a região onde a observação é classificada como mais comum. O critério de divisão empregado no modelo de *random forest* para o processo de classificação foi o índice Gini, definido por:

$$G = \sum_{k=1}^K p_{mk}(1 - p_{mk}) \quad (7)$$

o índice Gini assume valor pequeno quando se todos os p_{mk} 's estiverem próximos de zero ou um, um valor pequeno indica que um nó possui predominantemente observações de uma única classe.

4.3 Abordagens econométricas

Foram utilizadas algumas fórmulas para cada um dos três modelos utilizados, modelo linear, *lasso* e *random forest*. As primeiras fórmulas continham apenas uma variável socioemocional e as variáveis de controle para os alunos como gênero, idade, escolaridade da mãe, raça. A fórmula com a conscienciosidade representa as abordagens tradicionais em que não há interações e nem se tenta capturar a não linearidade entre as habilidades socioemocionais.

Outra fórmula utilizada foi a que continha todos os cinco constructos do *Big Five*, junto com as variáveis de controle para os alunos. Nela se tenta deixar evidente as relações entre as variáveis não cognitivas. De forma a capturar efeitos que se podem ter entre os constructos do *Big Five*.

Por último, foi utilizado a fórmula que contém todas as variáveis socioemocionais, de controle para os alunos, e algumas variáveis feitas a partir da interação entre essas variáveis.

Cada variável não binária, como a idade, número de irmãos, escolaridade da mãe e as variáveis socioemocionais, foi adicionada ao quadrado e ao cubo. Outra transformação feita foi a adição de variáveis que foram feitas a partir da interação das variáveis binárias com as outras variáveis disponíveis no subconjunto,

variáveis como gênero, raça, se participou de creche, pré escola, se foi reprovado foram relacionadas com as variáveis socioemocionais. Com todas as transformações feitas, o conjunto de dados possui quase 90 variáveis.

Para o processo de treino e teste a base de dados foi dividida de forma aleatória em 65% de treino e 35% para teste. Também foi utilizado o método de validação cruzada *k-fold*.

O método de validação cruzada *k-fold* consiste em dividir o conjunto de variáveis aleatoriamente em k grupos de tamanho aproximadamente igual. O primeiro grupo é tratado como um conjunto de validação e o método é ajustado nos $k - 1$ grupos restantes. O erro quadrático médio, EQM_1 , é então calculado nas observações do grupo separado. Este procedimento é repetido k vezes; cada vez, um grupo diferente de observações é tratado como um conjunto de validação. Este processo resulta em k estimativas do erro de teste, EQM_1 , EQM_2 , . . . , EQM_k . A estimativa de CV *k-fold* é calculada pela média desses valores. Nos modelos descritos no trabalho se utilizou $k = 10$ e com apenas uma repetição.

5 RESULTADOS

Nesta seção é apresentado os resultados das previsões dos modelos de ML. Como mostrado no capítulo anterior os resultados são referentes aos modelos linear, *random forest* e *lasso* aplicando algumas fórmulas diferentes.

Primeiro é mostrado o poder de previsão de cada uma das habilidades socioemocionais de forma separada, elas continuam as variáveis de controle para as características dos alunos. O modelo linear com a conscienciosidade representa os modelos tradicionais da literatura.

É apresentado o resultado do *Big Five*, ou seja, os cinco constructos juntos mais as variáveis de controles ao nível do estudante. E por último, a fórmula com todas as variáveis socioemocionais, de controle, e as combinações explicadas na seção anterior.

5.1 Qualidade da previsão

Para medir a qualidade da previsão foi utilizada a razão entre os erros quadráticos médios EQM, utilizando a fórmula do modelo linear com o constructo da

Conscienciosidade, que representa os modelos tradicionais, como base para o cálculo dos resultados.

$$EQM = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (8)$$

$$RES_{K,C} = \frac{EQM_K}{EQM_C} \quad (9)$$

5.1.1 Resultados da previsão

A seguir as Tabelas mostram o $RES_{K,C}$ e como base se tem o modelo linear com a Conscienciosidade. Todos os modelos são estimados com os controles para as características dos alunos.

Tabela 3 - Resultados dos modelos de ML em matemática

Matemática			
	Linear	<i>Random Forest</i>	<i>Lasso</i>
Conscienciosidade	1	0,93582	0,99969
Extroversão	1,01439	0,94634	1,01403
Neuroticismo	1,00911	0,93818	1,00884
Abertura	1,00842	0,94249	1,00804
Amabilidade	0,99898	0,92636	0,99873
<i>Big Five</i>	0,98541	0,91564	0,98602
Todas as variáveis	0,90020	0,87912	0,90272

Fonte: Elaborado pelo autor

Tabela 4 - Resultados dos modelos de ML em português

Português			
	Linear	<i>Random Forest</i>	<i>Lasso</i>
Conscienciosidade	1	0,92894	1,00027
Extroversão	1,00494	0,92986	1,00514
Neuroticismo	1,00218	0,91902	1,00265
Abertura	0,98220	0,91659	0,98257
Amabilidade	0,97315	0,89682	0,97338
<i>Big Five</i>	0,96125	0,87079	0,96141
Todas as variáveis	0,86339	0,84011	0,86478

Fonte: Elaborado pelo autor

Nas Tabelas 3 e 4 apresentam-se os principais resultados do estudo. Como se nota, os modelos têm melhor previsão com a fórmula que contém todas as variáveis, ou seja, a inclusão de mais variáveis torna melhor a previsão tanto do modelo linear como dos outros modelos nesse caso.

Foi verificado que a adição das variáveis interativas melhoram a previsão, percebendo assim que há algumas interações entre as variáveis que se mostram relevantes para a previsão no modelo.

O modelo *random forest* foi o que teve melhor resultado na previsão dos valores. Olhando para o modelo com todas as variáveis o método de *random forest* chega a ser 16% melhor que o modelo linear na previsão das notas de português e quase 13% nas notas de matemática.

Se compararmos as duas tabelas é possível perceber que na tabela 4, de português, as previsões são melhores do que as da Tabela 3, de matemática. Apenas o modelo *Lasso* com o constructo da Conscienciosidade se mostrou melhor na previsão da nota de matemática em comparação com a nota de português.

Olhando para os construtos de forma individual se constata que a Amabilidade, no modelo de *Random Forest*, chegou a ser quase 11% melhor que o modelo linear em português e quase 8% melhor em matemática. O constructo da extroversão como dito antes, tem efeitos difíceis de se capturar estatisticamente e no nosso modelo se mostrou com a menor pontuação, sendo pior no modelo linear e *Lasso* e sendo apenas um pouco mais de 3% melhor no *Random Forest*.

Em português se observa o que é dito pela literatura, os constructos da Amabilidade e Abertura à novas ideias tem uma melhor força de previsão para a disciplina de português. Em matemática, a Amabilidade apareceu como melhor, mas os traços da conscienciosidade e neuroticismo (estabilidade emocional) também pontuaram bem no seu poder de previsão.

O modelo *Lasso* apresentou resultados bastante similares aos resultados produzidos pelo modelo linear, tanto para a previsão da nota de matemática como de português. Dando indícios que o processo de seleção de variáveis feito pelo modelo *Lasso* não se apresentou relevante para esse conjunto de dados.

Já o modelo de *Random Forest* apresentou resultados bem melhores em comparação com o modelo linear. A melhor pontuação do modelo de *Random Forest* confirma a nossa hipótese de capturar os efeitos de não linearidade.

Segundo os resultados do modelo com todas as variáveis e o modelo com o *Big Five*, se nota que os resultados são significativamente melhores devido ao fato de se poder capturar os efeitos não lineares entre os construtos socioemocionais e seus efeitos interativos. Mesmo contendo muitas variáveis no modelo, o resultado com todas as variáveis também mostrou resultados significantes em comparação com o modelo linear.

5.2 Previsão da relação entre habilidades não cognitivas e desempenho em testes padronizados

Na Tabela 2 se demonstra uma variância muito grande nos resultados dos testes de português e matemática. A média da nota em matemática para alunos acima do percentil 75 chega a ser 327 pontos, enquanto os abaixo do percentil 25 com apenas 192 pontos em média.

Em português a variância e o intervalo entre máximo e mínimo é menor mas ainda existe a diferença de 293 pontos em média para alunos acima do percentil 75 e 216 pontos para alunos abaixo do percentil 25.

5.2.1 Definição de estudantes com elevado e baixo desempenho

Para investigar isso foram produzidas duas variáveis binárias para classificar alunos com elevado desempenho e alunos com baixo desempenho.

Alunos com nota superior ou igual ao percentil 75 eram classificados como alunos de elevado desempenho, recebendo 1 na variável binária e 0 para os demais alunos. Alunos com nota abaixo ou igual ao percentil 25 foram classificados com baixo desempenho e recebiam 1 na variável binária de 0 para os demais alunos. A criação dessas variáveis foi realizada para ambos os assuntos: português e matemática. A intenção nessa segunda parte da monografia é verificar se os métodos de ML têm maior poder de prever quem são os estudantes nos extremos da distribuição.

Importante notar que esses estudantes, com elevado e baixo desempenho, são geralmente alvo de políticas públicas, portanto, sua identificação pode possibilitar a criação de políticas públicas mais focalizadas. Estudantes com baixo desempenho geralmente são alvos de políticas educacionais que visem elevar seu desempenho, evitando que sua performance em testes padronizados se reflita ao longo da sua vida em outros importantes resultados, como abandono escolar, menores salários, menores cuidados com saúde, entre outros. Por sua vez, os estudantes com elevado desempenho têm maiores chances de ingressar na universidade após o ensino médio e, portanto, geralmente são alvos de políticas que visem facilitar essa transição entre ensino médio e ensino superior.

Dado que as variáveis analisadas nesta seção são binários, então o método de análise passa a ser o de classificação. Para tanto foi utilizado a matriz de confusão para o cálculo da acurácia dos modelos de previsão. Como o objetivo é detectar os melhores e os piores desempenhos dos alunos, também foi utilizado a medida de sensibilidade, que é a proporção de acertos dos verdadeiros positivos.

Tabela 5 - Matriz de confusão

		Valores Preditos	
		0	1
Valores Verdadeiros	0	Verdadeiro Negativo (VN)	Falso Positivo (FP)
	1	Falso Negativo (FN)	Verdadeiro Positivo (VP)

Fonte: Elaborado pelo autor

$$Acurácia = \frac{VP + VN}{\text{Número total de observações}} \quad (10)$$

$$Sensitividade = \frac{VP}{VP + FN} \quad (11)$$

Para escolha dos modelos foi utilizado o modelo de *Random Forest* com todas as variáveis, o modelo que melhor pontuou nos resultados na seção anterior. Foi utilizado também o modelo linear binomial para a classificação, considerando apenas uma variável socioemocional, modelo tradicional tomado como *benchmark*, e as variáveis de controle para as características dos alunos.

5.2.2 Resultados do modelo de classificação

A seguir são mostrados os resultados da análise de classificação dos alunos com melhores e piores desempenhos. A acurácia do modelo de *Random Forest* foi superior em todos os casos, apenas a medida de sensibilidade no caso dos piores alunos em matemática que o modelo *Random Forest* não foi o melhor. Mas nos outros casos ele se mostrou muito melhor que os outros modelos na previsão dos verdadeiros positivos.

O melhor resultado do modelo de *Random Forest* foi na previsão dos piores alunos de português, tendo uma acurácia de 76,4% e uma sensibilidade de 24,7%.

Quanto aos modelos com as habilidades socioemocionais é possível perceber que neuroticismo e amabilidade mostram melhor acurácia na previsão dos melhores alunos de matemática, enquanto para os piores alunos os melhores preditores foram Amabilidade, Conscienciosidade e Abertura.

Olhando para a tabela 7, com os resultados de português, se observa que o melhor preditor para os alunos com ótimos resultados foi o constructo da Abertura, enquanto para os piores alunos os melhores preditores foram Amabilidade e Conscienciosidade. O constructo da Amabilidade foi o que chegou mais perto dos resultados de sensibilidade do modelo de *Random Forest* na previsão dos verdadeiros positivos dos piores desempenhos de português, chegando a 16,6%.

Apesar dos modelos acertarem bem na média, o resultado mostrado pela sensibilidade demonstra que é mais difícil acertar os verdadeiros positivos. Porém, de forma geral, o modelo de *Random Forest* pontua melhor que os modelos lineares na medida de sensibilidade.

Tabela 6 - Resultados de classificação em matemática

Matemática						
	Melhores			Piores		
	Acurácia	IC	Sensibilidade	Acurácia	IC	Sensibilidade
<i>Random Forest</i>	76,1%	[75,3%; 76,7%]	18,2%	75,5%	[74,8%; 76,2%]	6,1%
Conscienciosidade	75,0%	[74,3%; 75,7%]	3,0%	75,2%	[74,5%; 75,9%]	6,5%
Extroversão	75,0%	[74,3%; 75,7%]	5,5%	75,0%	[74,3%; 75,7%]	4,9%
Neuroticismo	75,2%	[74,5%; 75,9%]	4,6%	75,1%	[74,4%; 75,8%]	4,8%
Abertura	75,1%	[74,4%; 75,8%]	4,6%	75,2%	[74,5%; 75,9%]	5,2%
Amabilidade	75,2%	[74,5%; 75,9%]	5,2%	75,5%	[74,8%; 76,2%]	8,7%

Fonte: Elaborado pelo autor

Tabela 7 - Resultados de classificação em português

Português						
	Melhores			Piores		
	Acurácia	IC	Sensibilidade	Acurácia	IC	Sensibilidade
<i>Random Forest</i>	76,2%	[75,5%; 76,8%]	20,2%	76,4%	[75,7%; 77,1%]	24,7%
Conscienciosidade	74,9%	[74,2%; 75,6%]	2,7%	75,6%	[74,9%; 76,3%]	10,8%
Extroversão	75,1%	[74,4%; 75,8%]	6,4%	75,3%	[74,6%; 76,0%]	9,7%
Neuroticismo	74,9%	[74,3%; 75,7%]	5,9%	75,4%	[74,7%; 76,1%]	8,9%
Abertura	75,5%	[74,7%; 76,2%]	8,9%	75,4%	[74,7%; 76,1%]	9,2%
Amabilidade	75,0%	[74,3%; 75,7%]	4,1%	76,3%	[75,6%; 77,0%]	16,6%

Fonte: Elaborado pelo autor

6 CONCLUSÕES

É reconhecida por todos a importância de uma educação de qualidade para que os alunos tenham um sucesso individual e coletivo. Porém, percebe-se que

não há ainda um consenso sobre o que é uma educação de qualidade para os desafios do século 21. Com isso é perceptível que apesar de serem fundamentais, as competências cognitivas e os conhecimentos tradicionais dos currículos escolares já não são mais suficientes para garantir o sucesso acadêmico, profissional e pessoal nos dias de hoje (SANTOS; PRIMI, 2014).

Como as habilidades não cognitivas ou socioemocionais ou até *soft skills* como o mercado está chamando essas habilidades tem se mostrado importante no sucesso dos alunos ao longo de suas conquistas, esse trabalho buscou explorar o poder de previsão dos métodos de ML sobre os resultados escolares por meio dessas habilidades.

O objetivo principal deste trabalho foi analisar o potencial dos modelos de ML capturando assim efeitos não lineares e efeitos interativos entre as variáveis. Para fazer essa análise foi utilizado dados do IAS que aplicou o instrumento SENNA em 643 escolas do estado do Ceará, instrumento esse que mapeia as habilidades não cognitivas dos alunos, categorizado no nosso trabalho com a teoria do *Big Five inventory*.

A partir desse estudo, verificou-se que os modelos de ML apresentam resultados satisfatórios em relação ao modelo linear na previsão de desempenho dos alunos nos testes padronizados. O modelo de *Random Forest* se destacou como o melhor modelo para essa previsão, chegando a ser quase 11% melhor que o modelo linear em português e quase 8% melhor em matemática.

Como dito antes, o modelo *Lasso* apresentou resultados bastante similares ao modelo linear, deixando a entender que o processo de seleção de variáveis feito pelo modelo não se apresentou relevante para esse conjunto de dados.

Foi verificado também que a adição das variáveis interativas melhoram a previsão, assim os modelos que continham essas variáveis apresentaram melhores resultados. O modelo linear com a adição dessas variáveis chegou a pontuar quase 14% melhor do que o próprio modelo linear seguindo a fórmula tradicional.

O constructo da Amabilidade apresentou ótimos resultados em comparação com os outros, na previsão dos resultados das duas disciplinas. Como dito antes, a literatura mostra a Amabilidade junto com a Abertura como os constructos com efeitos no desempenho das notas de português.

O objetivo de se tentar capturar a não linearidade e os efeitos interativos entre os construtos de mostrou relevante com o fato de que o modelo do *Big Five*, com todos os constructos, ter mostrados resultados significantes, chegando a ser quase 13% melhor com o ML *Random Forest* para os resultados de português.

Há ainda espaço para se estudar os novos e diversos outros modelos de ML que tem surgido no mundo da computação. Pode-se também procurar utilizar técnicas mais avançadas como *deep learning* e redes neurais.

Na segunda análise feita por esse trabalho tentou-se prever os alunos nas extremidades da distribuição, ou seja, os alunos com melhores e piores desempenhos, que são os alunos muitas vezes alvos de políticas públicas.

Foi verificado que o modelo de *Random Forest* também apresentou melhores resultados neste teste de previsão com uma acurácia de 76,4% e uma sensibilidade de 24,7%.

Nesta parte do trabalho o constructo que mostrou melhor desempenho foi o da Abertura na previsão dos ótimos alunos. E Amabilidade e Conscienciosidade foram os melhores preditores dos piores alunos.

Quanto à medida de sensibilidade o melhor constructo a pontuar foi a Amabilidade com 16.6% na previsão dos verdadeiros positivos dos piores alunos na disciplina de português.

A previsão dos verdadeiros positivos se mostrou bastante difícil para os modelos de ML porém na média os modelos preveem bem chegando a uma acurácia de 76,4% na previsão dos piores alunos na disciplina de português.

Por fim, é possível concluir que os modelos de ML possuem um potencial para prever o desempenho dos alunos nas disciplinas de matemática e português. E que isso pode ser usado para fim de políticas públicas com o objetivo de incentivar e criar espaços para o desenvolvimento das habilidades sociais e emocionais de forma personalizada.

O melhor entendimento dos efeitos das habilidades sociais e emocionais por meio das técnicas de ML podem auxiliar na gestão e elaboração de atividades inovadoras que buscam aumentar o desempenho dos alunos da rede pública de ensino.

REFERÊNCIAS

ALMLUND, Mathilde. *et al.* **Personality psychology and economics**. NBER Working Paper No. 16822. Cambridge, 2011.

BINET, Alfred; SIMON, Theodore. **The Development of Intelligence in Children** (the Binet-Simon Scale). Baltimore, MD, Williams & Wilkins Co. 1916

CARNEIRO, P.; CRAWFORD, C.; GOODMAN, A. **The impact of early cognitive and non-cognitive skills on later outcomes**. Centre for Economics of Education, 2007.

CUNHA, F.; HECKMAN, J. J.; SCHENNACH, S. M. **Estimating the technology of cognitive and noncognitive skill formation**. *Econometrica*, Wiley Online Library, v. 78, n. 3, p. 883–931, 2010.

DEYOUNG, Colin. *et al.* Openness to Experience, Intellect, and Cognitive Ability. *In: Journal of personality assessment*. 96. 46-52. 10.1080/00223891.2013.806327. 2013

DUCKWORTH, A. L.; SELIGMAN, M. E. Self-discipline outdoes iq in predicting academic performance of adolescents. **Psychological science**, SAGE Publications Sage CA: Los Angeles, CA, v. 16, n. 12, p. 939–944, 2005.

DUNCAN, G. J.; MAGNUSON, K. *et al.* **The nature and impact of early achievement skills, attention skills, and behavior problems**. *Whither opportunity*, p. 47–70, 2011.

FERGUSON, D. M.; HORWOOD, L. J. Early conduct problems and later life opportunities. **The Journal of Child Psychology and Psychiatry and Allied Disciplines**, Cambridge University Press, v. 39, n. 8, p. 1097–1108, 1998.

HECKMAN, James J.;KAUTZ Tim D. **Hard Evidence on Soft Skills** NBER Working Paper No. 18121 June 2012 JEL No. D01,I20

JAMES, Gareth. *et al.* **Introduction to Statistical Learning**. New York: Springer, 2013

JENSEN, Arthur R.. **The g Factor: The Science of Mental Ability**. Westport, CT, Praeger. 1998

JOHN, O. P.; SRIVASTAVA, S. The big five trait taxonomy: History, measurement, and theoretical perspectives. *In: Handbook of personality: Theory and research*, Guilford, v. 2, n. 1999, p. 102–138, 1999.

Mischel, W.; Shoda, Yuichi; Rodriguez, Monica. **Delay of Gratification in Children**. *Science* (New York, N.Y.). 244. 933-8. 10.1126/science.2658056. 1989

OECD, *Beyond Academic Learning: First Results from the Survey of Social and Emotional Skills*, **OECD Publishing**, Paris, 2021. Disponível em: <<https://doi.org/10.1787/92a11084-en>> acesso em 30/04/2022

ROBERTS, B. Back to the Future: Personality and Assessment and Personality Development. **Journal of Research in Personality** 43(2): 137-145. (2009)

SANTOS, D.; BERLINGERI, M.; CASTILHO, R. **Habilidades Socioemocionais e Aprendizado Escolar**: evidências a partir de um estudo em larga escala. Ribeirão Preto, 2017.

SANTOS, Daniel; PRIMI, Ricardo. **Desenvolvimento socioemocional e aprendizado escolar**: uma proposta de mensuração para apoiar políticas públicas. Relatório sobre resultados preliminares do projeto de medição de competências socioemocionais no Rio de Janeiro. São Paulo: OCDE/SEEDUC/Instituto Ayrton Senna, 2014. Disponível em: <<https://institutoayrtonsenna.org.br/content/dam/institutoayrtonsenna/documentos/de-senvolvimento-socioemocional-e-aprendizado-escolar.pdf>>. Acesso em 01 maio 2022

SOTO, C. J. *et al.* Age differences in personality traits from 10 to 65: Big five domains and facets in a large cross-sectional sample. **Journal of personality and social psychology**, American Psychological Association, v. 100, n. 2, p. 330, 2011.

SWIFT, Victor; PETERSON, Jordan B. Contextualization as a means to improve the predictive validity of personality models. *In: Personality and Individual Differences*. University of Toronto, 27 King's College Circle, Toronto, Ontario M5S 1A1 Canada: 2019

THIEL, Hendrik; THOMSEN, Stephan L. **Noncognitive skills in economics: Models, measurement, and empirical evidence**, ZEW Discussion Papers, No. 09-076, Zentrum für Europäische Wirtschaftsforschung (ZEW): Mannheim, 2009