



UNIVERSIDADE FEDERAL DO CEARÁ
CÂMPUS DE QUIXADÁ
CURSO DE GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

AUGUSTO VASCONCELLOS NOGUEIRA DIOGENES

**UMA REVISÃO SISTEMÁTICA DA LITERATURA SOBRE O USO DE
APRENDIZADO DE MÁQUINA PARA O AUXÍLIO AO DIAGNÓSTICO DE EXAMES
DE MANOMETRIA ANORRETAL E ESOFÁGICA**

QUIXADÁ

2026

AUGUSTO VASCONCELLOS NOGUEIRA DIOGENES

UMA REVISÃO SISTEMÁTICA DA LITERATURA SOBRE O USO DE APRENDIZADO
DE MÁQUINA PARA O AUXÍLIO AO DIAGNÓSTICO DE EXAMES DE MANOMETRIA
ANORRETAL E ESOFÁGICA

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Ciência da Computação
do Câmpus de Quixadá da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Ciência da Computação.

Orientador: Prof. Dr. Paulo Armando
Cavalcante Aguilar.

QUIXADÁ

2026

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)

- D622r Diogenes, Augusto Vasconcellos Nogueira.
Uma Revisão Sistemática Da Literatura Sobre o Uso De Aprendizado De Máquina Para O Auxílio Ao Diagnóstico De Exames De Manometria Anorretal e Esofágica / Augusto Vasconcellos Nogueira
Diogenes. – 2026.
48 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Quixadá, Curso de Ciência da Computação, Quixadá, 2026.
Orientação: Prof. Dr. Paulo Armando Cavalcante Aguilár.
1. gastroenterologista. 2. aprendizado de máquina. 3. manometria anorretal. 4. manometria esofágica.
5. diagnóstico. I. Título.

CDD 004

AUGUSTO VASCONCELLOS NOGUEIRA DIOGENES

UMA REVISÃO SISTEMÁTICA DA LITERATURA SOBRE O USO DE APRENDIZADO
DE MÁQUINA PARA O AUXÍLIO AO DIAGNÓSTICO DE EXAMES DE MANOMETRIA
ANORRETAL E ESOFÁGICA

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Ciência da Computação
do Câmpus de Quixadá da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Ciência da Computação.

Aprovada em: 22/01/2026

BANCA EXAMINADORA

Prof. Dr. Paulo Armando Cavalcante
Aguilar (Orientador)
Universidade Federal do Ceará (UFC) - Câmpus de
Quixadá

Prof. Dr. Regis Pires Magalhães
Universidade Federal do Ceará (UFC) - Câmpus de
Quixadá

Dra. Lara Burlamaqui Veras
Médica Coloproctologista

AGRADECIMENTOS

Ao Prof. Dr. Paulo Armando Cavalcante Aguiar, pela excelente orientação.

Aos participantes da banca examinadora, Prof. Dr. Regis Pires Magalhães e Dra. Lara Burlamaqui Veras, pelo tempo, pelas valiosas colaborações e sugestões.

RESUMO

Exames de manometria são uma forma rápida e segura para detectar e diagnosticar distúrbios funcionais do trato digestivo com alta acurácia. Em função disso, esses exames são considerados úteis dentro da área da gastroenterologia, e existe interesse no ramo clínico em melhorar o procedimento desses exames com novas tecnologias - em particular o aprendizado de máquina. Esse trabalho procura contribuir a esse esforço acadêmico através de uma revisão sistemática da literatura existente. Os repositórios PubMed, Scopus e IEEE foram utilizados como fonte. Foram selecionados e analisados um total de 16 artigos, cada um detalhando a implementação de ferramentas de aprendizado de máquina para o auxílio ao diagnóstico de exames de manometria esofágica e anorretal. Concluiu-se que sistemas que utilizam redes neurais convolucionais pré-treinadas são o atual estado da arte, capazes de gerar diagnósticos completos. No entanto, para treinar essas redes, é necessário um conjunto de dados com milhares de imagens. Caso não se tenha acesso a um conjunto de dados desse tamanho, é recomendado utilizar modelos como perceptrons de múltiplas camadas, quantização de vetores de aprendizado, máquinas de vetor de suporte, k-ésimo vizinho mais próximo, regressão logística e florestas aleatórias. Esses sistemas não são capazes de gerar diagnósticos completos, mas podem identificar informações úteis que podem ajudar na avaliação de exames - como a detecção de movimentos normais e/ou anormais e a avaliação da pressão de relaxamento integrada.

Palavras-chave: proctologista; gastroenterologista; aprendizado de máquina; inteligência artificial; manometria anorretal; manometria esofágica; diagnóstico; automação.

ABSTRACT

Manometry exams are a fast and safe way to detect and diagnose functional disorders of the digestive tract with high accuracy. Because of this, these exams are considered useful in the field of gastroenterology, and there is interest in the clinical field in improving the procedure of these exams with new technologies - in particular machine learning. This work seeks to contribute to this academic effort through a systematic review of the existing literature. The PubMed, Scopus and IEEE repositories were used as sources. A total of 16 articles were selected and analyzed, each detailing the implementation of machine learning tools to aid in the diagnosis of esophageal and anorectal manometry exams. It was concluded that systems that use pre-trained convolutional neural networks are the current state of the art, capable of generating complete diagnoses. However, to train these networks, a dataset with thousands of images is required. If one does not have access to a dataset of this size, it is recommended to use models such as multilayer perceptrons, quantization of learning vectors, support vector machines, k-nearest neighbor, logistic regression, and random forests. These systems are not capable of generating complete diagnoses, but they can identify useful information which can aid in evaluating exams - such as the detection of normal and/or abnormal movements and the assessment of integrated relaxation pressure.

Keywords: proctologist; gastroenterologist; machine learning; artificial intelligence; anorectal manometry; esophageal manometry; diagnosis; automation.

LISTA DE FIGURAS

- Figura 1 – Exemplo dos dados gerados após um exame de manometria, na representação usual de grafo. O grafo de linhas à direita representa *pressão* \times *tempo* para cada sonda. O grafo à esquerda representa *pressão* \times *distância* \times *tempo*. . . 12
- Figura 2 – Ilustração dos passos de busca e seleção de artigos de uma revisão PRISMA. 20
- Figura 3 – Fluxograma PRISMA detalhando o processo de seleção de estudos. 27
- Figura 4 – Exemplo de sistema composto encontrado na literatura analisada, demonstrando como vários modelos treinados são usados tanto de forma sequencial quanto de forma paralela para processar os dados de entrada dados ao sistema. 36

LISTA DE QUADROS

Quadro 1 – Trabalhos relacionados	24
Quadro 2 – Questões de pesquisa	25
Quadro 3 – Critérios	26
Quadro 4 – Estudos inclusos	28
Quadro 5 – Tipos de dados utilizados na literatura	38
Quadro 6 – Maior Acurácia Alcançada Por Tipo de Modelo	39

SUMÁRIO

1	INTRODUÇÃO	10
1.1	Objetivos	11
1.1.1	<i>Objetivos Específicos</i>	11
2	EMBASAMENTO TEÓRICO	12
2.1	Manometria	12
2.1.1	<i>Manometria Esofágica</i>	13
2.1.2	<i>Manometria Anorretal</i>	13
2.2	Aprendizado de Máquina	14
2.2.1	<i>Perceptrons de Múltiplas Camadas</i>	15
2.2.2	<i>Redes Neurais Convolucionais</i>	15
2.2.3	<i>Variational Autoencoder</i>	16
2.2.4	<i>Long Short-Term Memory</i>	16
2.2.5	<i>Florestas Aleatórias</i>	16
2.2.6	<i>Quantização de Vetores de Aprendizado</i>	17
2.2.7	<i>K-Ésimo Vizinho Mais Próximo</i>	17
2.2.8	<i>Regressão Logística</i>	17
2.2.9	<i>Máquinas de Vetores de Suporte</i>	18
2.2.10	<i>Gradient Boosting</i>	18
2.3	Revisão Sistemática de Literatura	18
2.3.1	<i>Método PRISMA</i>	19
3	TRABALHOS RELACIONADOS	21
3.1	Aprendizado de máquina na cirurgia cardiovascular: uma revisão sistemática e avaliação crítica	21
3.2	O desempenho clínico de aprendizado de máquina automatizado: uma revisão sistemática	22
3.3	Revisões sistemáticas de aprendizado de máquina na assistência médica	22
3.4	Aplicações de algoritmos de aprendizado de máquina para prever resultados terapêuticos na depressão: uma meta-análise e revisão sistemática	23
3.5	Análise comparativa dos trabalhos relacionados	23
4	METODOLOGIA	25

4.1	Formulação das Questões de Pesquisa	25
4.2	Formulação de Estratégia de Busca	25
4.3	Busca e Seleção	26
5	RESULTADOS E DISCUSSÃO	27
5.1	Q1: Modelos Identificados Na Literatura	28
5.1.1	<i>Perceptrons de Múltiplas Camadas</i>	29
5.1.2	<i>Redes Neurais Convolucionais</i>	29
5.1.3	<i>Variational Autoencoder</i>	30
5.1.4	<i>Long Short-Term Memory</i>	31
5.1.5	<i>Florestas Aleatórias</i>	31
5.1.6	<i>Quantização de Vetores de Aprendizado</i>	32
5.1.7	<i>K-Ésimo Vizinho Mais Próximo</i>	32
5.1.8	<i>Regressão Logística</i>	32
5.1.9	<i>Máquinas de Vetores de Suporte</i>	33
5.1.10	<i>Gradient Boosting</i>	34
5.1.11	<i>Revisão dos Modelos</i>	35
5.2	Q2: Implementação dos Modelos Em Sistemas Decisores	36
5.3	Q3: Dados Para O Desenvolvimento dos Sistemas	37
5.4	Q4: Desempenho dos Modelos	38
6	CONCLUSÃO	41
	REFERÊNCIAS	43

1 INTRODUÇÃO

A manometria, dentro da área da proctologia e da gastroenterologia, se trata de um protocolo de aferição da pressão de várias partes do trato digestório. O procedimento envolve a inserção de um cateter contendo uma ou mais sondas de pressão através de um dado esfíncter do sistema digestivo. Qual esfíncter será sujeito a esse procedimento depende de qual parte do trato digestório está sendo examinada. Atualmente, exames desse tipo são considerados o padrão ouro para a avaliação de distúrbios funcionais anorretais (Bharucha *et al.*, 2022) e gastroesofágicos (Cohen; Shirin, 2023).

Devido a isso, existe um interesse dentro do campo da gastroenterologia em tecnologias que tenham o potencial de aprimorar a agilidade e a acurácia do processo de diagnóstico desses exames. Dentre essas tecnologias, uma que se destaca e que atrai uma atenção especial no meio acadêmico é o aprendizado de máquina (Saraiva *et al.*, 2023; Stefan-Lucian *et al.*, 2022). Através do uso de modelos de aprendizado de máquina dentro do protocolo de manometria esofágica e gastrointestinal, seria possível automatizar parcialmente ou totalmente o processo de diagnóstico, reduzindo a carga de trabalho dos médicos e permitindo que mais pacientes sejam atendidos.

No entanto, o desenvolvimento de modelos de aprendizado de máquina é um processo difícil e trabalhoso, onde vários ciclos de tentativa e erro são necessários. Isso apresenta um obstáculo para a criação e implementação prática de novas ferramentas de inteligência artificial que possam agilizar essa modalidade de exame. Seria interessante, para pesquisadores e desenvolvedores trabalhando neste tipo de ferramenta, uma forma de minimizar esses ciclos de tentativas e erro. Uma possível forma de alcançar esse objetivo seria através de um recurso, embasado no estado da arte da literatura, que mostra quais tipos de modelos podem ser utilizados para o problema da automatização da manometria, e dentre esses quais possuem a melhor chance de apresentar um bom desempenho.

Nesse contexto, a revisão sistemática da literatura (RSL) se apresenta como um recurso útil para responder a esta problemática. A RSL é um protocolo para responder a uma dada questão de pesquisa. Ela permite desenvolver uma busca ampla na literatura, onde informações específicas são condensadas sistematicamente por meio do mapeamento e avaliação dos estudos existentes, e a partir disso gerar conclusões fundamentadas no atual estado da arte (Synder, 2019). Dessa forma, uma RSL focada em identificar quais modelos de aprendizado de máquina podem ser utilizados para automatizar a manometria esofágica e anorretal seria capaz

de oferecer aos pesquisadores e desenvolvedores interessados na criação desses modelos um guia robusto para orientar seus trabalhos futuros.

Como uma RSL que atende a esse escopo não foi encontrada na literatura, esse trabalho foca em desenvolvê-la. Espera-se que, com acesso a essa revisão, pesquisadores e desenvolvedores possam agilizar o processo de desenvolvimento de ferramentas de aprendizado de máquina para o diagnóstico de exames de manometria. Efetivamente, isso irá ajudar na adoção prática dessas ferramentas no ambiente clínico. Isso, eventualmente, pode vir a aumentar a facilidade de administração e a acurácia da manometria anorretal e esofágica, o que beneficia tanto os profissionais de saúde que administram esses tipos de exames, quanto os pacientes que dependem deles.

1.1 Objetivos

O objetivo geral desse trabalho é, pelo uso do protocolo de RSL PRISMA (Page *et al.*, 2021), analisar sistematicamente o estado da arte do uso de aprendizado de máquina para apoiar o diagnóstico em exames de manometria esofágica e anorretal, identificando modelos, formas de implementação, tipos de dados, tarefas clínicas e desempenho reportado na literatura.

1.1.1 Objetivos Específicos

Os objetivos específicos dessa revisão são:

- Identificar os tipos de modelos de aprendizado de máquina utilizados na literatura.
- Identificar as tarefas clínicas abordadas pelos modelos de aprendizado de máquina nos exames de manometria esofágica e anorretal.
- Analisar as formas de implementação dos modelos de aprendizado de máquina em sistemas de apoio ao diagnóstico.
- Caracterizar os tipos de dados utilizados nos estudos analisados, considerando formato, origem, rotulagem e volume amostral.
- Identificar a correlação entre os tipos de dados manométricos e os modelos de aprendizado de máquina mais adequados.
- Identificar e comparar o desempenho dos modelos de aprendizado testados na literatura.
- Discutir quais desses modelos são mais promissores para futuras pesquisas.

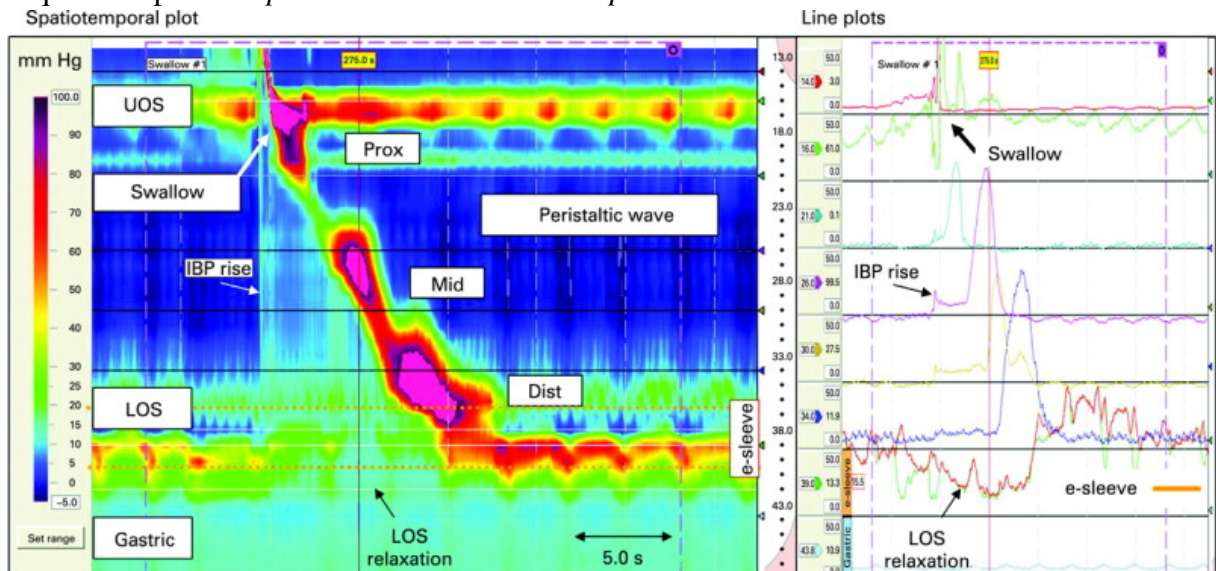
2 EMBASAMENTO TEÓRICO

Nesta seção, explicam-se os conceitos-chave deste trabalho.

2.1 Manometria

Exames de manometria utilizam cateteres com sondas de pressão distribuídas ao longo do comprimento do cateter para avaliar o funcionamento dos músculos de diversas partes do trato digestivo. O mesmo equipamento, com sondas diferentes, pode ser utilizado para examinar tanto o trato digestivo superior quanto o inferior (Basilisco; Bharucha, 2017; Fang *et al.*, 2004). Ao longo do exame, o equipamento cria um grafo de *pressão × tempo* para cada sonda. Esse grafo também pode ser representado como um grafo de *pressão × distância × tempo*, levando em consideração o intervalo de distância entre as sondas (Conklin, 2013). Exemplos desses grafos são demonstrados na figura 1. O uso principal desse tipo de exame é o diagnóstico de desordens funcionais e distúrbios de motilidade anorretais e esofágicos de forma rápida e segura para o paciente (Baldwin; Puckett, 2024; Heinrich; Misselwitz, 2018).

Figura 1 – Exemplo dos dados gerados após um exame de manometria, na representação usual de grafo. O grafo de linhas à direita representa *pressão × tempo* para cada sonda. O grafo à esquerda representa *pressão × distância × tempo*.



Fonte: Baldwin e Puckett (2024).

Atualmente, exames de manometria de alta resolução - que possuem um maior número de sondas de pressão por cateter - estão se tornando mais populares como uma ferramenta de diagnóstico, e seus resultados são considerados como padrão de ouro para o diagnóstico

de doenças de motilidade do trato digestivo (Hsiao, 2018). Dado que essas doenças estão se tornando mais comuns conforme o tempo - um resultado da crescente demografia de idosos em vários países desenvolvidos e em desenvolvimento - (Antunes *et al.*, 2024) há um interesse no ambiente clínico em como melhorar a acurácia, segurança e facilidade com que esse protocolo de exame é realizado.

2.1.1 Manometria Esofágica

A manometria esofágica é uma das modalidades mais comuns de manometria. Ela é usada para a detecção de anomalias na motilidade do trato digestivo superior, que indicam falhas no funcionamento do esôfago (Baldwin; Puckett, 2024). A Classificação de Chicago é considerada o padrão para a interpretação de resultados em grafos obtidos a partir da manometria esofágica (Baldwin; Puckett, 2024; Kahrilas *et al.*, 2015). Ela é um algoritmo que utiliza informações sobre padrões anormais de contração muscular durante o engolimento e sobre a pressão de relaxamento integrada - ambas obtidas a partir da análise do gráfico do exame - para diagnosticar desordens de fluxo de saída da junção esôfago-gástrica (JEG). Essas são desordens como: acalasia e obstrução da JEG, e distúrbios do peristaltismo, como ausência de contratilidade, espasmo esofágico distal, esôfago hipercontrátil, e distúrbios menores do peristaltismo caracterizados por trânsito do bolo alimentar prejudicado (Kahrilas *et al.*, 2015).

2.1.2 Manometria Anorretal

A manometria anorretal é um dos principais exames envolvidos na investigação de distúrbios da evacuação, medindo parâmetros pressóricos estáticos e dinâmicos que podem sugerir ou caracterizar anormalidades do esfíncter anal ou falhas na coordenação anorretal (Carrington *et al.*, 2018). Um dos principais componentes funcionais detectados pela manometria são os padrões de dissinergia anorretal. Existem quatro padrões de dissinergia. Esses padrões são diferentes combinações de contração paradoxal e relaxamento inadequado dos músculos do ânus e do reto durante o esforço evacuatório. Contextualizados com a história clínica e com exames físicos dos pacientes, esses padrões podem explicar quadros de distúrbios da defecação (Carrington *et al.*, 2018; Carrington *et al.*, 2020). A identificação desses padrões, porém, depende de uma boa execução do exame e de avaliação manual de parâmetros numéricos e gráficos por médicos experientes (Carrington *et al.*, 2020).

2.2 Aprendizado de Máquina

O aprendizado de máquina existe como parte da área da inteligência artificial (Awad; Khanna, 2015). A inteligência artificial procura criar algoritmos de computador que imitam e, em certos casos, excedem as capacidades da inteligência humana. Dentro do contexto da medicina diagnóstica, esses algoritmos são úteis, pois são capazes identificar padrões dentro de conjuntos de dados - nesse caso, nos resultados de exames médicos - que podem não ser aparentes para um ser humano, ou que são excessivamente complicados para análise manual (Fass *et al.*, 2024). Em sistemas de aprendizado de máquina, os padrões que são reconhecidos não são explicitamente programados dentro do algoritmo por seres humanos. Eles são derivados computacionalmente a partir de um conjunto de dados alimentado ao sistema por uma função de treinamento, e então utilizados para classificar novos conjuntos de dados. Existem diversos modelos de sistema que pertencem ao grupo do aprendizado de máquina, e dentre esses, existe um sub-grupo de modelos que pertencem a categoria do *deep learning*. Esses são modelos que possuem uma hierarquia de camadas, onde cada camada reconhece um padrão e passa sua saída como entrada para a próxima camada. Modelos de *deep learning* vem recebendo grande atenção nos últimos anos, pois eles são capazes de realizar tarefas fora do alcance de modelos de aprendizado tradicionais (Janiesch *et al.*, 2021).

É interessante, para os desenvolvedores de sistemas de aprendizado de máquina, avaliar a capacidade de seus programas de classificar de forma correta novos conjuntos de dados. Essa avaliação pode ser feita com diversos métodos. Um desses métodos é a divisão dos dados disponíveis em subconjuntos para treinamento, validação, e teste (GOOGLE FOR DEVELOPERS, 2025). O conjunto de treinamento é utilizado para ensinar padrões úteis ao sistema. O conjunto de validação é usado para avaliar o desempenho do sistema durante os ciclos de tentativa e erro em seu desenvolvimento, onde são mudanças na arquitetura do modelo, e o modelo é treinado e avaliado novamente. Finalmente, o conjunto de teste é usado para avaliar o desempenho do sistema em sua versão final. Outro método usado para avaliar o desempenho de modelos de aprendizado de máquina é a validação cruzada *k-fold* (Marcot; Hanea, 2021). Nesse método, o conjunto de dados disponíveis em k subconjuntos de tamanhos iguais. Para cada subconjunto c dentre os k subconjuntos, é realizada uma rodada de treinamento e validação do modelo de aprendizado. Em cada rodada o modelo é treinado em todos os k subconjuntos exceto c , e então é testado em c . Calcula-se, então, a média de todas k as rodadas, e essa média é dada como a estimativa do desempenho do modelo final, que será treinado com o conjunto de

dados completo.

No decorrer desse trabalho, diversos modelos específicos de aprendizado de máquina tradicionais e de *deep learning* serão discutidos, tendo seus desempenhos dentro de uma área de aplicação específica - o diagnóstico de exames de manometria - avaliadas. Esses algoritmos são: perceptrons de múltiplas camadas, quantização de vetores de aprendizado, redes neurais convolucionais, *variational autoencoder*, *long short-term memory*, florestas aleatórias, k-ésimo vizinho mais próximo, regressão logística, máquinas de vetores de suporte, e *gradient boosting*. O funcionamento de cada um desses algoritmos será brevemente elaborado nas próximas subseções: 2.2.1 a 2.2.10.

2.2.1 Perceptrons de Múltiplas Camadas

O perceptron de múltiplas camadas é um algoritmo de *deep learning*. Ele é composto por três ou mais camadas de neurônios artificiais: uma de entrada, uma de saída, e uma ou mais camadas escondidas (Baum, 1988). As camadas são conectadas umas às outras por funções não-lineares de ativação. Cada neurônio presente nessas camadas recebe o vetor de entrada de sua respectiva camada e multiplica-o por um vetor de pesos de mesmo tamanho, cujos valores são otimizados pela função de aprendizado. O vetor composto pela saída de cada neurônio artificial de uma camada representa a entrada da camada seguinte. Atualmente, a função sigmoide e a função ReLU são as funções mais populares, pois elas funcionam bem com o algoritmo de aprendizado de retropropagação de erro (Dubey *et al.*, 2022). Durante seu funcionamento, a rede recebe um vetor de entrada na sua primeira camada e retorna os valores de ativação dos neurônios da camada de saída. Em uma rede classificadora, cada neurônio na saída corresponde a uma classe de classificação.

2.2.2 Redes Neurais Convolucionais

A rede neural convolucional, também um tipo de rede neural artificial, é um modelo de *deep learning* cujo uso é popular - mas não estritamente limitado - para a análise de imagens e vídeos (LeCun *et al.*, 2015). Essas redes consistem em uma camada de entrada, várias camadas convolucionais com camadas de *pooling* distribuídas entre elas, e então camadas de neurônios artificiais. As camadas convolucionais aplicam um filtro de convolução à entrada, onde os valores da matriz do filtro são otimizados pelo algoritmo de aprendizado. As camadas de *pooling* reduzem a resolução da entrada, servindo para economizar recursos computacionais. Finalmente,

as camadas de neurônios artificiais operam da mesma forma que as camadas de um perceptron multicamadas. A entrada e a saída desse tipo de rede neural também operam de forma semelhante ao funcionamento de um perceptron multicamadas: a camada de entrada recebe um vetor de valores para ser classificado, e a saída retorna o valor de ativação de neurônios correspondentes a cada classe que a rede foi projetada para reconhecer.

2.2.3 Variational Autoencoder

O *Variational Autoencoder* (VAE) é um algoritmo construído a partir de duas redes neurais. Essas podem ser perceptrons multicamadas ou redes neurais convolucionais, por exemplo (Kingma; Welling, 2022). Uma rede opera de forma normal, e reduz uma entrada a um conjunto de variáveis representando as probabilidades dessas entradas corresponderem a certas classes ou possuir certas características. A outra rede opera de forma "reversa", recebendo uma distribuição de variáveis e gerando uma saída que corresponde a essa distribuição. O VAE não é útil, por si só, como um classificador de dados. Mas ele pode ser útil como um modelo de aprendizado generativo que gera mapas de características que outros sistemas podem usar para detectar padrões e classificar dados (Nishizaki, 2017).

2.2.4 Long Short-Term Memory

O algoritmo *Long Short-Term Memory* (LSTM) é um tipo de rede neural recorrente (RNR). A RNR é uma variação do perceptron de múltiplas camadas onde os neurônios artificiais nas camadas escondidas podem usar a sua própria saída em um passo de execução anterior como uma de suas entradas no passo atual, efetivamente adicionado um "loop" nas conexões do neurônio. Isso permite que a RNR analise dados sequenciais (Tealab, 2018). A LSTM expande sobre a RNR, introduzindo uma célula de memória a esse "loop", que pode preservar valores de passos anteriores, acessá-los quando necessário, e descartar eles quando não são mais úteis (Hochreiter; Schmidhuber, 1997).

2.2.5 Florestas Aleatórias

A floresta aleatória é um algoritmo composto por um conjunto de árvores de decisão (Ho, 1995). A árvore de decisão é um algoritmo de aprendizado de máquina baseado na estrutura de um fluxograma de árvore binária. Cada nó nesse fluxograma, exceto os nós folhas, representa

uma variável contida no vetor de entrada. Cada um desses nós possui um valor condicional, tal que se a variável correspondente a esse nó for menor que esse valor, o algoritmo que percorre a árvore durante o processo de decisão irá prosseguir para o seu primeiro nó filho e, caso contrário, ele irá prosseguir para o segundo nó filho. Esses valores condicionais são otimizados pelo algoritmo de aprendizado. Durante seu funcionamento, a árvore de decisão é percorrida, começando pelo nó raiz, por um algoritmo que segue o caminho ditado pelos valores condicionais até chegar a um nó folha, que corresponde a uma classe de classificação. Em um algoritmo de florestas aleatórias, várias árvores de decisão com estruturas diferentes são treinadas juntas e executadas independentemente, e a saída do algoritmo é a média das saídas de todas essas árvores.

2.2.6 Quantização de Vetores de Aprendizado

O algoritmo da quantização de vetores de aprendizado possui duas camadas de neurônios artificiais, uma de entrada e uma de saída, que operam de forma semelhante às camadas de entrada e saída de um perceptron multicamadas. No entanto, o treinamento dessa rede é diferente de um perceptron comum. Ao treinar essa rede de neurônios, para cada elemento no conjunto de treinamento, encontra-se o neurônio na primeira camada cuja saída nesse elemento mais se aproxima a saída desejada, e atualiza-se seus pesos (Hotel *et al.*, 2018).

2.2.7 K-Ésimo Vizinho Mais Próximo

O algoritmo do k-ésimo vizinho mais próximo é uma arquitetura relativamente simples de aprendizado de máquina usada para classificação e regressão (Peterson, 2009). Para classificar uma entrada, o modelo recebe um vetor de entrada e compara-o aos vetores já classificados em seu conjunto de dados de treinamento. A classe do vetor de entrada é determinada como sendo a média das classes dos vetores mais próximos dentro do conjunto de dados do modelo. O número de vetores usados para esse cálculo é determinado pelo valor da variável k, escolhido durante o desenvolvimento do modelo.

2.2.8 Regressão Logística

A regressão logística é um algoritmo de aprendizado que, dado um conjunto de variáveis, determina-se a probabilidade entre 0 e 1 de que esse conjunto pertence a uma certa

classe. Esse cálculo é feito realizando um somatório dessas variáveis, cada uma multiplicada por seu peso correspondente, e então aplica-se uma função sigmoide à saída desse somatório (Tolles; Meurer, 2016). Dessa forma, é possível considerar um algoritmo de regressão logística como um algoritmo neural contendo um único neurônio artificial.

2.2.9 Máquinas de Vetores de Suporte

A máquina de vetor de suporte é um algoritmo de aprendizado de máquina que recebe um conjunto de dados de vetores de treinamento de n-dimensões, os agrupa dentro de um espaço n-dimensional, e então desenha uma linha ou hiperplano que separa todos os vetores de classes diferentes e agrupa todos os vetores da mesma classe, criando "territórios" diferentes para cada classe (Cortes; Vapnik, 1995). Ao receber um novo vetor de entrada, o algoritmo o classifica com base em qual "território" dentro do espaço n-dimensional ele ocupa.

2.2.10 Gradient Boosting

Gradient boosting é uma arquitetura de aprendizado de máquina que combina vários modelos de aprendizado simples - tipicamente uma árvore de decisão - e de baixo desempenho em um modelo sequencial forte (Mason *et al.*, 1999). Isso é feito ao treinar os modelos de forma sucessiva. O primeiro modelo recebe uma entrada e é treinado para retornar uma previsão. Cada modelo após esse recebe a previsão do modelo anterior, e é treinado para retornar uma previsão que mais se aproxima às classificações reais presentes no conjunto de dados de treinamento.

2.3 Revisão Sistemática de Literatura

A revisão sistemática de literatura (RSL) é um tipo de pesquisa científica que reúne estudos relevantes sobre uma questão formulada, utilizando o banco de dados da literatura como fonte de artigos científicos. Este tipo de pesquisa utiliza um protocolo sistemático para formular uma estratégia de busca dentro desses bancos de dados, e selecionar, dentre os artigos levantados pela busca inicial, aqueles que são relevantes para a questão formulada inicial. A RSL é preferida no meio acadêmico por ser uma forma eficiente e efetiva de avaliar a literatura científica dentro de um dado tópico de interesse (Koutsos *et al.*, 2019), gerando conclusões úteis para futuras pesquisas dentro de sua área.

O processo para a realização de uma RSL é baseado na seguinte sequência de ações:

a formulação da questão ou questões de pesquisa, a formulação de uma estratégia de busca em repositórios digitais, a busca e seleção de artigos que são relevantes à questão ou questões de pesquisa, a análise e síntese de dados obtidos a partir dos artigos, e o desenvolvimento de conclusões a partir desses dados (Page *et al.*, 2021). A definição das questões centrais de pesquisa pode ser feita através do modelo PICO (Schardt *et al.*, 2007), que agrupa os conceitos centrais da pesquisa nas categorias de população (quem está sendo estudado), intervenção (o quê está sendo estudado), comparação (em que contexto a pesquisa se insere), e desfecho (qual é objetivo do que está sendo estudado).

É importante, para a reprodutibilidade da RSL, que os passos de pesquisa sejam descritos de forma concisa - especialmente o passo de formulação de estratégia de busca. A estratégia de busca deve incluir as bases de dados escolhidas e a *string* de busca utilizada. Essa *string* é composta das palavras-chave e operadores lógicos, de tal forma que sinônimos ou termos da mesma categoria são agrupados e unidos por OR, e esses grupos são unidos por operadores AND (UNIVERSITY OF MICHIGAN LIBRARY, 2024). Um exemplo de *string* de busca seguindo essa regra seria ("*termo a*"OR "*termo b*") AND ("*termo c*"OR "*termo d*"), onde "*termo a*" e "*termo b*" são sinônimos e "*termo c*" e "*termo d*" são outro grupo de sinônimos.

2.3.1 Método PRISMA

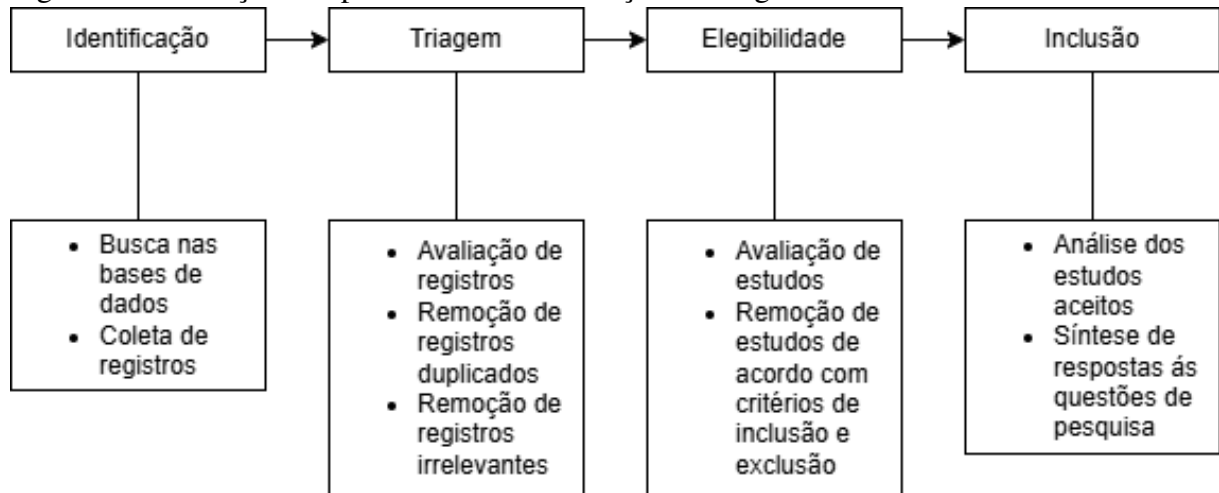
O método PRISMA, que é utilizado neste trabalho, é um conjunto de diretrizes que ditam boas práticas para a realização de uma RSL completa e transparente (Page *et al.*, 2021). As RSLs que seguem essas diretrizes possuem uma metodologia distinta, na qual devem buscar e selecionar artigos em quatro passos. Esses passos são:

1. Identificação, onde se realiza uma busca de artigos nas bases de dados selecionadas para a revisão. Esses artigos serão analisados primeiro como registros: ou seja, resumos e títulos.
2. Triagem, onde os registros são lidos e avaliados. Nesta avaliação, são eliminados os registros duplicados e os registros que não satisfazem os critérios de relevância da revisão. Os registros aceitos neste passo passam para o próximo como estudos: ou seja, artigos completos.
3. Elegibilidade, onde os estudos são lidos e avaliados, e aqueles que não satisfazem os critérios de inclusão e exclusão da revisão são eliminados.
4. Inclusão, onde as informações úteis dos estudos aceitos são extraídas e levadas à discussão para responder as perguntas de pesquisa.

Esses passos também são ilustrados na Figura 2.

Após a conclusão desses passos, é necessário documentar a sua execução na forma de um fluxograma. Esse fluxograma deve detalhar a execução de cada passo, destacando quantos registros e estudos foram aceitos e rejeitados por cada métrica de avaliação. A criação desse fluxograma é necessária para garantir a transparência de uma RSL do modelo PRISMA.

Figura 2 – Ilustração dos passos de busca e seleção de artigos de uma revisão PRISMA.



Fonte: compilação do autor.

3 TRABALHOS RELACIONADOS

Este trabalho se insere na literatura como parte de um grupo maior de revisões sistemáticas de literatura (RSL) que procuram responder questões relevantes ao uso do aprendizado de máquina dentro da medicina. Existem vários outros trabalhos que, assim como este, procuram esclarecer questões sobre o uso dessa tecnologia em suas áreas de interesse dentro da saúde. Alguns desses trabalhos foram selecionados para servir de auxílio ao desenvolvimento desta pesquisa, sendo modelos de inspiração para o formato e metodologia.

O ideal, para servir de auxílio ao desenvolvimento dessa pesquisa, seriam RSLs que buscam atender os mesmos tópicos de pesquisa deste trabalho. No entanto, revisões sistemáticas sobre o uso de aprendizado de máquina em exames de manometria não foram identificadas durante a produção deste artigo. Devido a isso, os trabalhos selecionados para servir de inspiração são revisões sistemáticas sobre o uso de aprendizado de máquina em várias áreas da saúde. Esses trabalhos estão listados nas seguintes subseções.

3.1 Aprendizado de máquina na cirurgia cardiovascular: uma revisão sistemática e avaliação crítica

Assim como este trabalho de conclusão de curso se trata de uma RSL de aprendizado de máquina dentro de uma área específica da medicina - os exames de manometria dentro da gastroenterologia - Li *et al.* (2022) foca no uso de tecnologias de aprendizado de máquina relevantes especificamente à cirurgia cardiovascular.

Os repositórios de literatura MEDLINE, Embase, e Cochrane CENTRAL foram utilizados para a busca de artigos. Todos os estudos originais envolvendo a aplicação de aprendizado de máquina na cirurgia vascular foram aceitos, totalizando 212 artigos aceitos dentro do grupo de 2235 artigos levantados pela busca inicial. Tendências de publicação, condições de doença, metodologias e resultados foram resumidos. A avaliação crítica foi realizada utilizando as ferramentas de risco de viés PROBAST, e a revisão seguiu o protocolo TRIPOD.

A partir desse processo de revisão, esta pesquisa concluiu que, embora o número de artigos dentro da sua área de interesse tenha aumentado nas últimas décadas, a qualidade dessas pesquisas ainda deixa a desejar. A maioria dos estudos foi retrospectiva e unicêntrica, sem ensaios clínicos randomizados. Os autores da pesquisa concluem recomendando que estudos futuros sobre o uso de aprendizado de máquina dentro da cirurgia cardiovascular utilizem ferramentas

como PROBAST e TRIPOD para melhorar sua qualidade e aplicabilidade no ambiente clínico.

3.2 O desempenho clínico de aprendizado de máquina automatizado: uma revisão sistemática

Thirunavukarasu *et al.* (2024) utiliza o protocolo PROSPERO para avaliar o estado da arte do uso de ferramentas de aprendizado de máquina automatizado (AutoML) na literatura clínica. As bibliotecas online Cochrane, Embase, MEDLINE e Scopus foram utilizadas para a busca de artigos.

A pesquisa identificou o uso de 26 plataformas de AutoML distintas utilizadas no grupo de 82 artigos aceitos. Doenças cerebrais e pulmonares eram as áreas de aplicação mais comuns para essas ferramentas. Nas comparações das plataformas de AutoML, o AutoPrognosis e o Amazon Rekognition tiveram melhor desempenho com dados estruturados e não estruturados, respectivamente. A qualidade dos relatórios foi fraca, com uma pontuação média do DECIDE-AI de 14 de 27.

Os autores concluíram, com base nessa revisão de literatura, que ferramentas de AutoML são promissoras devido a seu bom desempenho e facilidade de treinamento, e que mais trabalhos são necessários para melhorar a qualidade de estudos de validação dentro dessa área de interesse.

3.3 Revisões sistemáticas de aprendizado de máquina na assistência médica

Kolasa Bisrat Admassu e Perni (2024) é uma RSL sobre outras RSLs focadas no uso de aprendizado de máquina na área de saúde. Dessa forma, essa pesquisa oferece a visão mais completa sobre o contexto no qual esse trabalho de conclusão de curso e outras pesquisas semelhantes se encaixam.

As bases de dados Web of Science, EBSCO, Cochrane, PubMed, IEEE Xplore, e Scopus foram utilizadas. No total, 220 RSLs cobrindo 10.462 algoritmos de máquina de máquina foram aceitas pelo protocolo PRISMA usado nesta pesquisa. Foi identificado que a principal aplicação da IA na medicina está relacionada à predição clínica e ao prognóstico de doenças em oncologia e neurologia com o uso de dados de imagem. O modelo mais comum utilizado na literatura foi redes neurais. Máquinas de vetor de suporte ocupam o segundo lugar, e árvores de decisão o terceiro.

Essa pesquisa indicou lacunas consideráveis na prestação de informações em termos do desempenho do ML, tanto na validação interna como externa. Seus autores concluíram que maior acessibilidade aos dados de saúde para desenvolvedores pode garantir uma melhora de qualidade na literatura dentro dessa área, além da adoção mais rápida de algoritmos de ML na prática clínica.

3.4 Aplicações de algoritmos de aprendizado de máquina para prever resultados terapêuticos na depressão: uma meta-análise e revisão sistemática

Lee *et al.* (2018) utilizou o protocolo PRISMA para investigar qualitativamente as diferentes aplicações de aprendizado de máquina como um agente previsor no tratamento de distúrbios depressivos e transtornos de personalidade.

Foram identificados 639 registros a partir da base de dados MEDLINE Pubmed, 75 estudos completos foram avaliados por elegibilidade, e 20 estudos foram incluídos na análise qualitativa e quantitativa. A análise concluiu que algoritmos de classificação conseguem, em média, prever resultados terapêuticos com uma acurácia de 82%.

No entanto, o estudo também ressaltou suas limitações ligadas ao fato que a maioria dos estudos aceitos na revisão foram estudos retrospectivos, e a implementação heterogênea de sistemas de aprendizado entre diferentes estudos torna impossível a análise do impacto de variáveis individuais no desempenho do sistema.

3.5 Análise comparativa dos trabalhos relacionados

Os trabalhos escolhidos demonstram diferentes protocolos de RSL em uso. Dentre esses protocolos, o método PRISMA foi o mais utilizado. Também se observa que Pubmed foi a base de dados mais utilizada, seguida por Cochrane e Scopus. Essas informações são destacadas no quadro 1, que compara esses trabalhos.

Quadro 1 – Trabalhos relacionados

Estudo	Bases de dados	Protocolo de RSL
Li <i>et al.</i> (2022)	MEDLINE Pubmed, Embase, e Cochrane CENTRAL	TRIPOD
Thirunavukarasu <i>et al.</i> (2024)	MEDLINE Pubmed, Embase, Cochrane, Scopus	PROSPERO
Kolasa Bisrat Admassu e Perni (2024)	Web of Science, EBSCO, Cochrane, PubMed, IEEE Xplore, Scopus	PRISMA
Lee <i>et al.</i> (2018)	MEDLINE Pubmed	PRISMA

Fonte: compilação do autor.

4 METODOLOGIA

4.1 Formulação das Questões de Pesquisa

As questões centrais desta pesquisa foram definidas através das categorias *Population, Intervention, Comparison, Outcome* (PICO) (Schardt *et al.*, 2007). Essas categorias são, respectivamente, as populações e/ou grupos relevantes para a área de pesquisa, a intervenção sendo estudada na pesquisa, a comparação (área central de aplicação, a base da métrica de avaliação) sobre a qual a intervenção é aplicada e o desfecho e/ou objetivo da intervenção. Para os propósitos deste trabalho, essas categorias foram definidas como:

- População: Médico, paciente, gastroenterologista, proctologista.
- Intervenção: Aprendizado de máquina, inteligência artificial, visão computacional.
- Comparação: Manometria, manometria anorretal, manometria esofágica.
- Desfecho (*Outcome*): Identificação, classificação, diagnóstico.

A partir dessas categorias, as seguintes questões são definidas no quadro 2.

Quadro 2 – Questões de pesquisa

Q1: Quais modelos de aprendizado de máquina podem ser identificados na literatura?
Q2: Como esses modelos são implementados dentro de sistemas decisores?
Q3: Com que dados esses modelos são e treinados?
Q4: Quais tipos de modelo apresentaram o melhor desempenho?

Fonte: compilação do autor.

4.2 Formulação de Estratégia de Busca

Com base nos termos definidos nas categorias PICO, uma estratégia de busca é formulada. Estabelece-se primeiro a *string* de busca ("*artificial intelligence*" OR "*computer vision*" OR "*machine learning*") AND ("*manometry*" OR "*anorectal manometry*" OR "*esophageal manometry*") AND ("*identification*" OR "*classification*" OR "*diagnosis*"), que foi feita para a pesquisa de artigos em inglês nas bases de dados PubMed, Scopus e IEEE. Esta *string* utiliza os termos das categorias de intervenção, comparação e desfecho como palavras-chave. A categoria de população foi omitida nesta *string* para evitar um excesso de palavras chave na busca, aumentando o número de artigos que podem ser obtidos no passo inicial da identificação de registros.

4.3 Busca e Seleção

A busca e seleção dos artigos segue os passos de identificação, triagem, elegibilidade e inclusão especificados no método PRISMA (Page *et al.*, 2021). A identificação foi feita seguindo a estratégia de busca formulada na seção 4.2. Foram feitas várias buscas seguindo essa estratégia entre 06/2024 e 12/2025, e os registros obtidos foram agregados na ferramenta online *parsif.al*, preparando-os para a análise e seleção manual. A triagem foi feita com a remoção manual de registros duplicados e, então, com a remoção de registros que não atendiam a um ou mais dos critérios de relevância deste trabalho. Os estudos que passaram pela triagem tiveram sua elegibilidade determinada pelos critérios de inclusão e exclusão. Os estudos excluídos foram contados com base em qual critério de inclusão ou exclusão eles violaram, e os estudos aceitos foram incluídos na discussão do trabalho. Todos os critérios de relevância, inclusão e exclusão estão listados no quadro 3.

Quadro 3 – Critérios

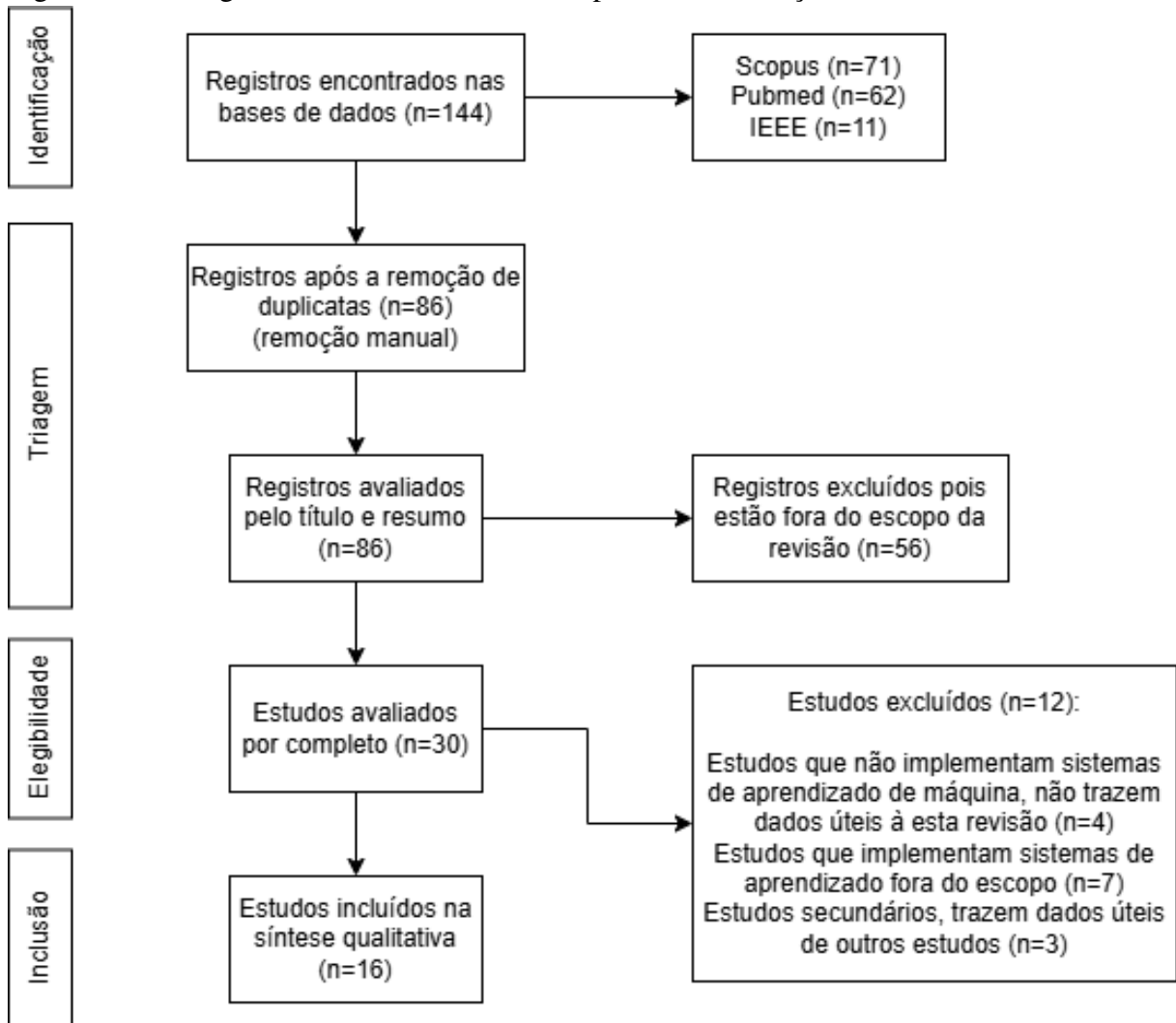
Critérios de relevância
Os artigos devem estar disponíveis em língua inglesa.
Os artigos devem ser publicados antes de 2026.
Os artigos devem ser publicações em jornais, conferências, ou <i>preprints</i> .
Os artigos devem focar em exames de manometria esofágica e anorretal, aplicados especificamente para a detecção de desordens e doenças do trato digestivo.
Os artigos devem discutir a implementação de tecnologias de aprendizado de máquina.
Critérios de inclusão
Os artigos devem implementar seus próprios sistemas de aprendizado e incluir dados sobre o desempenho destes sistemas.
Os sistemas de aprendizado discutidos no artigo devem servir para auxiliar na síntese laudos e diagnósticos a partir de dados manométricos.
Critério de exclusão
Os artigos não podem conter somente dados duplicados, referentes a modelos de aprendizado de máquina já discutidos em outro artigo.

Fonte: compilação do autor.

5 RESULTADOS E DISCUSSÃO

Uma busca inicial foi feita nas bases de dados PubMed, IEEE e Scopus, utilizando a *string* de busca gerada a partir das categorias PICO. A busca mais atual foi feita no dia 28/12/2025. Levantaram-se inicialmente 144 registros de artigos: 71 da Scopus, 62 da Pubmed e 11 da IEEE. Eliminando os duplicados, reduziu-se o número de registros para 86. Dentre esses, 30 foram aceitos por cumprirem os requisitos de relevância e foram então avaliados como estudos completos. 12 estudos foram excluídos: 4 por violar o primeiro critério de inclusão, 7 por violar o segundo critério de inclusão e 3 por violar o único critério de exclusão. No final, 16 estudos passaram para a síntese qualitativa na discussão. Todo esse processo de busca e seleção é ilustrado na figura 3.

Figura 3 – Fluxograma PRISMA detalhando o processo de seleção de estudos.



Fonte: compilação do autor.

Dentre os trabalhos inclusos, foram identificados 3 estudos focados em manometria

anorretal e 13 estudos focados em manometria esofágica. Percebe-se que a literatura sobre o uso de aprendizado de máquina em exames de manometria anorretal é menos extensa do que a literatura sobre o uso de aprendizado de máquina em manometria esofágica. Uma lista dos trabalhos aceitos, junto com as tarefas clínicas dos sistemas que esses trabalhos desenvolveram, está disponível no quadro 4.

Quadro 4 – Estudos inclusos

Estudo	Tarefa Clínica	Manometria
Mielens <i>et al.</i> (2012)	Identificar presença ou ausência de desordens	Esofágica
Schilling <i>et al.</i> (2015)	Estimar tempo de relaxamento do esfíncter esofágico	Esofágica
Kou <i>et al.</i> (2021)	Clusterizar datasets de exames	Esofágica
Stefan-Lucian <i>et al.</i> (2022)	Automação da Classificação de Chicago	Esofágica
Surdea-Blaga <i>et al.</i> (2022)	Automação da Classificação de Chicago	Esofágica
Czako <i>et al.</i> (2022)	Estimar pressão de relaxamento integrada	Esofágica
Kou <i>et al.</i> (2022a)	Automação da Classificação de Chicago	Esofágica
Kou <i>et al.</i> (2022b)	Identificar diferentes desordens de motilidade	Esofágica
Saraiva <i>et al.</i> (2023)	Identificar presença ou ausência de desordens	Anorretal
Zifan <i>et al.</i> (2023)	Identificar disfagia	Esofágica
Levy (2023)	Identificar presença ou ausência de desordens	Anorretal
Storonova <i>et al.</i> (2024)	Identificar acalásia	Esofágica
Popa <i>et al.</i> (2024)	Automação da Classificação de Chicago	Esofágica
Mascarenhas <i>et al.</i> (2025a)	Identificar diferentes desordens de motilidade, Identificar presença ou ausência de desordens	Anorretal
Mascarenhas <i>et al.</i> (2025b)	Identificar refluxo, Identificar presença ou ausência de desordens	Esofágica
Lin <i>et al.</i> (2025)	Identificar presença ou ausência de desordens	Esofágica

Fonte: compilação do autor.

As informações obtidas a partir desses estudos foram extraídas, resumidas e, a partir delas, são sintetizadas as respostas para as questões centrais desta pesquisa.

5.1 Q1: Modelos Identificados Na Literatura

Para responder a primeira pergunta de pesquisa, esta seção traz um resumo dos modelos de aprendizado de máquina desenvolvidos na literatura analisada, dividindo-os por tipo e destacando informações pertinentes sobre como foram desenvolvidos em seus estudos, às tarefas para as quais foram projetados e o desempenho que alcançaram. Esse resumo é concluído com uma revisão geral que discute como esses diferentes tipos de algoritmos foram aplicados na literatura.

5.1.1 *Perceptrons de Múltiplas Camadas*

Dois estudos utilizaram perceptrons de múltiplas camadas.

Mielens *et al.* (2012) desenvolveu um modelo de perceptron de múltiplas camadas para detectar engolimentos normais ou anormais em gráficos de exames de manometria esofágica de alta resolução. Esse modelo foi treinado e validado com dados obtidos de 23 exames, cada um oferecendo quatro exemplos de engolimento. O modelo foi então avaliado com uma validação cruzada *5-fold*, resultando em uma acurácia de $96,44\% \pm \%1,27$.

Kou *et al.* (2022a) empregou um modelo de perceptron de múltiplas camadas como parte de um sistema que automatiza a Classificação de Chicago para exames de manometria esofágica. O perceptron dentro desse sistema recebe as saídas de três redes neurais convolucionais, que identificam tipos de engolimento, pressão de engolimento e outra que identifica a pressão de relaxamento integrada. Ele então decide, em paralelo com um modelo de regras e um modelo de *gradient boosting*, o diagnóstico do exame. Usou-se um conjunto de dados de 1741 exames, que foi dividido em conjuntos de treinamento/teste/validação de acordo com a razão 70/15/15. A acurácia do sistema integrado foi de 81%.

5.1.2 *Redes Neurais Convolucionais*

Seis estudos utilizaram redes neurais convolucionais.

Surdea-Blaga *et al.* (2022) desenvolveu um sistema que usa duas redes neurais convolucionais pré-treinadas: uma baseada na arquitetura DenseNet201 e outra na arquitetura InceptionV3. A primeira rede classifica o tipo de desordem de engolimento, e a segunda classifica a pressão de relaxamento integrada. Dadas essas duas classificações, o exame é laudado de acordo com a Classificação de Chicago v3.0. 1535 imagens de exames foram usadas para treinar e testar a primeira rede, e 1079 para a segunda. O sistema, como um todo, apresentou uma acurácia de laudo de 86%.

Kou *et al.* (2022a) criou um sistema que combina três redes neurais convolucionais. Uma identifica tipos de engolimento, outra identifica a pressão de engolimento, e outra identifica a pressão de relaxamento integrada. As saídas dessas são passadas para três modelos que decidem, em paralelo, o diagnóstico seguindo as normas da Classificação de Chicago. Esses modelos são: uma rede neural artificial, um modelo de regras e um modelo de *gradient boosting*. Utilizou-se um conjunto de dados de 1741 exames, que foi dividido em conjuntos de treinamento/teste/validação

para todos os modelos de acordo com a razão 70/15/15. A acurácia do sistema integrado foi de 81%.

Stefan-Lucian *et al.* (2022) e Popa *et al.* (2024) usaram redes neurais convolucionais pré-treinadas: uma baseada na arquitetura InceptionV3 e outra baseada na arquitetura Efficient-Net, respectivamente. A primeira rede utilizou *um dataset* de 1570 imagens diagnosticadas de acordo com os critérios da Classificação de Chicago V3.0, e então foram separadas em conjuntos de treinamento/teste/validação na razão 70/15/15. Essa rede apresentou uma acurácia de 93%. A segunda rede utilizou um conjunto de dados de 926 imagens diagnosticadas da mesma forma, e então separadas em conjuntos de treinamento/teste/validação na razão 70/10/20. A acurácia desse sistema foi de 88%.

Em total, quatro estudos empregaram redes neurais convolucionais em sistemas feitos para automatizar o diagnóstico de exames de manometria esofágica seguindo a Classificação de Chicago. Dois desses estudos criaram sistemas complexos, que empregam mais de um modelo de aprendizado como decisores internos. Os outros dois são sistemas simples, que empregam somente um modelo de aprendizado - sendo esses ambos modelos pré-treinados.

Czako *et al.* (2022) também desenvolveu um modelo pré-treinado baseado na arquitetura InceptionV3, dessa vez para identificar a pressão de relaxamento integrada (PRI) e verificar o posicionamento correto da sonda de manometria de alta resolução. 2437 imagens representando o posicionamento correto/incorreto da sonda e 1079 imagens representando PRI normal/anormal foram usadas. Ambos esses conjuntos de dados foram divididos em conjuntos de treinamento/teste/validação na razão 70/15/15. O modelo demonstrou acurácia de 97% ao classificar PRI e 91% ao verificar o posicionamento da sonda.

Lin *et al.* (2025) desenvolveu um modelo de rede neural convolucional baseado na rede MobileNet. Esse modelo tem como meta ser uma alternativa mais computacionalmente leve que outras redes neurais convolucionais que foram desenvolvidas para diagnosticar exames de manometria esofágica. A rede tem 4,3 milhões de parâmetros, alcançou uma acurácia de 93,99%, e foi treinada com 2080 imagens de exames de manometria esofágica, classificadas como normais ou anormais. O conjunto de dados dessas imagens foi dividido em conjuntos de treinamento/teste/validação na razão 6/2/2.

5.1.3 Variational Autoencoder

Dois estudos usaram *variational autoencoders*.

Levy (2023) treinou modelos do tipo *variational autoencoder* que foram testados tanto por si só quanto em conjunto com regressão logística e análise discriminante (linear e quadrática), formando sistemas híbridos. 302 exames de manometria foram usados no estudo. Uma validação cruzada *5-fold* foi empregada, e o desempenho foi medida como a área sob a curva (ASC) característica de operação do receptor. Os sistemas híbridos tiveram uma ASC de 0,96, enquanto os sistemas que somente empregavam um modelo do tipo *variational autoencoder* somente alcançaram uma ASC de 0,91.

Kou *et al.* (2021) criou um modelo de aprendizado não supervisionado do tipo *variational autoencoder*. O modelo foi treinado com 32.415 imagens de engolimento oriundas de exames de manometria esofágica, e foi usado para identificar *clusters* em seu conjunto de dados que correspondem a tipos normais e anormais de engolimento. A capacidade do modelo de discriminar esses *clusters* foi avaliada com um modelo supervisionado de análise discriminante linear (ADL) aplicado sobre o modelo original, com o seu conjunto de dados dividido em conjuntos de treinamento/teste/validação na razão 70/15/15. A acurácia obtida pela ADL no conjunto de validação foi 86%.

5.1.4 Long Short-Term Memory

Somente um estudo utilizou um modelo de *long short-term memory*.

Esse estudo (Kou *et al.*, 2022b) treinou um modelo desse tipo para a identificação de diferentes tipos de engolimentos anormais em dados de manometria de alta resolução. Dados de 26.115 engolimentos obtidos a partir de 1.741 exames de manometria foram utilizados, sendo divididos em conjuntos de treinamento/teste/validação na razão 70/15/15. O sistema demonstrou uma acurácia de 81% no conjunto de validação.

5.1.5 Florestas Aleatórias

Quatro estudos utilizaram florestas aleatórias

Saraiva *et al.* (2023) e Mascarenhas *et al.* (2025a) treinaram modelos de florestas aleatórias para identificar padrões de motilidade em exames de manometria anorretal. O primeiro modelo usou 827 exames anorretais como fonte de dados. 90% desses dados foram usados para treinamento e 10% para validação, obtendo uma acurácia de $82,6\% \pm 4,8\%$. O segundo modelo usou 701 exames diagnosticados de acordo com a Classificação de Londres para seu treinamento, obtendo um *F1-score* de $83\% \pm 3,2\%$ ao detectar a presença ou ausência de distúrbios de

motilidade.

Storonova *et al.* (2024) testou um modelo de florestas na tarefa de diagnosticar acalásia a partir de dados de exames de manometria esofágica. O modelo foi treinado com os dados espaço-temporais de pressão de 750 engolimentos realizados sob exame de manometria, e conseguiu atingir um *F1-score* de 0,91%.

Zifan *et al.* (2023) treinou um modelo de florestas aleatórias para reconhecer casos de disfagia. Usaram-se dados de exames de manometria realizados em 60 pacientes, cada um realizando 18-20 engolimentos. Uma validação cruzada *5-fold* foi aplicada para avaliar o desempenho do modelo, o qual demonstrou uma acurácia de 90,5%.

5.1.6 Quantização de Vetores de Aprendizado

Somente um estudo utilizou modelos de quantização de vetores de aprendizado.

Esse estudo (Mielens *et al.*, 2012) aplicou esse tipo de modelo para detectar engolimentos normais ou anormais em gráficos de exames de manometria esofágica de alta resolução. Dados para treinamento e validação do modelo foram obtidos de 23 exames, cada um oferecendo quatro exemplos de engolimento. Uma validação cruzada *5-fold* foi aplicada para avaliar o modelo, rendendo uma acurácia de $91,03\% \pm 0,98\%$.

5.1.7 K-Ésimo Vizinho Mais Próximo

Dois estudos utilizaram modelos de k-ésimo vizinho mais próximo.

Saraiva *et al.* (2023) desenvolveu um desses modelos próximo para identificar padrões de motilidade em exames de manometria anorretal. 827 exames anorretal foram usados como fonte de dados. 90% desses dados foram usados para treinamento e 10% para validação. O exibiu uma acurácia de $74,4\% \pm 3,8\%$ a ser avaliado.

Zifan *et al.* (2023) treinou um modelo de k-ésimo vizinho mais próximo para reconhecer casos de disfagia. Usaram-se dados de exames de manometria realizados em 60 pacientes, cada um realizando 18-20 engolimentos. Uma validação cruzada *5-fold* foi aplicada sobre o modelo, demonstrando uma acurácia de 88,1%.

5.1.8 Regressão Logística

Dois estudos usaram regressão logística.

Levy (2023) treinou modelos de regressão logística que foram testados tanto por si só, quanto em conjunto com modelos de *deep learning*, formando sistemas híbridos. Tais modelos incluem: um *variational autoencoder* e um modelo que combina operações de convolução e atenção. 302 exames de manometria foram usados no estudo. Uma validação cruzada *5-fold* foi empregada, e o desempenho foi medida como a área sob a curva (ASC) característica de operação do receptor. Os sistemas híbridos tiveram uma ASC de 0,96, enquanto o modelo de regressão logística por si só atingiu uma ASC de 0,93.

Zifan *et al.* (2023) implementou regressão logística para reconhecer casos de disfagia. Usaram-se dados de exames de manometria realizados em 60 pacientes, cada um realizando 18-20 engolimentos. Uma validação cruzada *5-fold* foi aplicada, retornando uma acurácia de 90,5%.

5.1.9 Máquinas de Vetores de Suporte

Cinco estudos usaram máquinas de vetores de suporte.

Saraiva *et al.* (2023) treinou um modelo de máquina de vetores de suporte para identificar padrões de motilidade em exames de manometria anorretal. 827 exames anorretal foram usados como fonte de dados. 90% desses dados foram usados para treinamento e 10% para validação. Esse modelo demonstrou uma acurácia de $81\% \pm 8\%$.

Mielens *et al.* (2012) aplicou um modelo de máquina de vetores de suporte para detectar engolimentos normais ou anormais em gráficos de exames de manometria esofágica de alta resolução. Dados para treinamento e validação dos modelos foram obtidos de 23 exames, cada um oferecendo quatro exemplos de engolimento. Uma validação cruzada *5-fold* foi aplicada para avaliar o modelo, demonstrando uma acurácia de $85,39\% \pm 2,4\%$.

Schilling *et al.* (2015) testou um algoritmo de máquina de vetores de suporte para estimar o tempo em que o esfíncter esofágico superior irá retornar a seu estado de descanso após o engolimento durante um exame de manometria. Os dados espaço-temporais de pressão foram obtidos a partir de 10 pacientes, cada um realizando 10 engolimentos durante um exame de manometria, resultando em um conjunto de dados de tamanho 100 que foi dividido em conjuntos de treinamento/teste/validação na razão 70/10/20. A máquina de vetor de suporte exibiu uma acurácia de 90,26%.

Zifan *et al.* (2023) e Zifan *et al.* (2024) treinaram máquinas de vetores de suporte para reconhecer casos de disfagia. A primeira máquina foi treinada com dados de exames de

manometria realizados em 60 pacientes, cada um realizando 18-20 engolimentos. Uma validação cruzada *5-fold* foi aplicada para avaliar seu desempenho, retornando uma acurácia de 85,7%. A segunda máquina foi treinada com dados obtidos de exames de manometria esofágica realizados em 60 pacientes, cada um engolindo 16-20 vezes. Ela foi avaliada com uma validação *k-fold*, e seu desempenho foi medida como a área sob a curva (ASC) característica de operação do receptor. Nesse caso, a máquina de vetor de suporte teve uma ASC de 0,95.

5.1.10 *Gradient Boosting*

Cinco estudos usaram modelos de *gradient boosting*.

Saraiva *et al.* (2023) desenvolveu um desses modelos para identificar padrões de motilidade em exames de manometria anorretal. 827 exames anorretal foram usados como fonte de dados. 90% desses dados foram usados para treinamento e 10% para validação. A acurácia do modelo foi $84,6\% \pm 2,9\%$.

Kou *et al.* (2022a) empregou um modelo de *gradient boosting* como parte de um sistema que automatiza a Classificação de Chicago para exames de manometria esofágica. O perceptron dentro desse sistema recebe as saídas de três redes neurais convolucionais, que identificam tipos de engolimento, pressão de engolimento e outra que identifica a pressão de relaxamento integrada. Ele então decide, em paralelo com um modelo de regras e um modelo de perceptron de múltiplas camadas, o diagnóstico do exame. Usou-se um conjunto de dados de 1741 exames, que foi dividido em conjuntos de treinamento/teste/validação de acordo com a razão 70/15/15. A acurácia do sistema integrado foi de 81%.

Storonova *et al.* (2024) testou um modelo de *gradient boosting* - especificamente o CatBoostClassifier - na tarefa de diagnosticar acalásia a partir de dados de exames de manometria esofágica. O modelo foi treinado com os dados espaço-temporais de pressão de 750 engolimentos realizados sob exame de manometria, e conseguiu atingir um *F1-score* de 0,90%.

Mascarenhas *et al.* (2025a) treinou vários modelos de *gradient boosting* para identificar desordens de motilidade nos dados de exame de manometria anorretal. 701 exames diagnosticados de acordo com a Classificação de Londres foram usados para treinar os modelos. Dentre todos os modelos de *gradient boosting* avaliados, o modelo LGBM Classifier obteve a melhor acurácia: 87% na tarefa de reconhecer e identificar diferentes desordens de motilidade anorretais.

Mascarenhas *et al.* (2025b) também avaliou o desempenho de vários modelos do tipo

gradient boosting em um dataset de 618 exames de manometria esofágica. Os modelos foram treinados para duas tarefas: reconhecer engolimentos normais ou anormais e reconhecer os sinais de desordens de refluxo. Dentre os modelos avaliados, o modelo *GradientBoostingClassifier* foi o melhor em detectar refluxo e o modelo *XGBoost* foi o melhor em reconhecer engolimentos anormais. Os modelos demonstraram acurácias de 94,2% e 80,9% respectivamente.

5.1.11 Revisão dos Modelos

Ao analisar essas informações, o primeiro fato que pode ser destacado sobre a escolha de modelos na literatura selecionada é que os algoritmos utilizados são, predominantemente, modelos de aprendizado supervisionado. Isso é algo esperado, uma vez que a tarefa de analisar e diagnosticar exames de manometria é uma atividade na qual os alvos já são conhecidos. Isso, claro, não quer dizer que o aprendizado não-supervisionado seja inútil, uma vez que pode ser utilizado para auxiliar no desenvolvimento de algoritmos supervisionados (Kou *et al.*, 2021; Levy, 2023); mas essa estratégia não demonstra nem maior acurácia, nem maior economia de dados para treinamento do que o uso direto de algoritmos supervisionados.

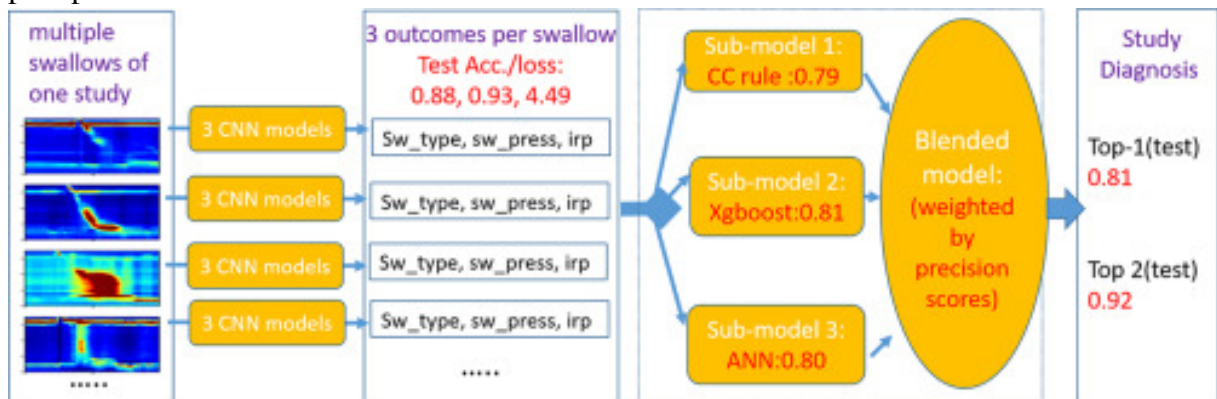
Dentre esses diversos algoritmos supervisionados que apareceram na literatura, as redes neurais convolucionais, as máquinas de vetor de suporte os modelos de gradient boosting se destacaram como os modelos mais utilizados em diferentes trabalhos. As redes neurais convolucionais também se destacam por serem o único tipo de modelo que, até então, foi usado como decisor dentro de todos os algoritmos capazes de gerar diagnósticos completos de várias doenças através da automação da Classificação de Chicago (Surdea-Blaga *et al.*, 2022; Kou *et al.*, 2022a; Popa *et al.*, 2024). Outros sistemas, em contraste, detectam somente uma doença ou identificam sintomas que podem se aplicar a várias doenças.

Vários estudos usaram modelos pré-treinados de redes neurais convolucionais (Stefan-Lucian *et al.*, 2022; Czako *et al.*, 2022; Popa *et al.*, 2024). Essa estratégia diminui o tamanho do conjunto de dados de treinamento necessário e melhora a acurácia em comparação aos estudos que utilizaram redes convolucionais que não foram pré-treinadas (Kou *et al.*, 2022a; Surdea-Blaga *et al.*, 2022).

5.2 Q2: Implementação dos Modelos Em Sistemas Decisores

Os modelos de aprendizado identificados na literatura são implementados em sistemas decisores que podem ser divididos em duas categorias: os sistemas simples, que utilizam apenas um modelo de aprendizado de máquina como decisor interno, e os sistemas compostos, que combinam mais de um modelo de aprendizado de máquina dentro de um algoritmo maior.

Figura 4 – Exemplo de sistema composto encontrado na literatura analisada, demonstrando como vários modelos treinados são usados tanto de forma sequencial quanto de forma paralela para processar os dados de entrada dados ao sistema.



Fonte: Kou *et al.* (2022a).

Somente 3 dos 16 estudos aceitos usaram arquiteturas compostas em seus sistemas decisores. Cada um desses sistemas possui uma arquitetura única. Eles serão, portanto, descritos nesta seção.

Dois estudos utilizaram redes neurais convolucionais para automatizar o protocolo da Classificação de Chicago para exames de manometria esofágica.

Um estudo (Surdea-Blaga *et al.*, 2022) utilizou uma arquitetura que combina duas redes neurais convolucionais pré-treinadas: uma rede DenseNet201 que classifica o tipo de desordem de engolimento e uma rede InceptionV3 que classifica a pressão de relaxamento integrada. As saídas dessas redes passam pelo algoritmo da Classificação de Chicago para gerar um diagnóstico. Esse sistema alcançou uma acurácia de 88% usando um conjunto de dados de 2614 imagens.

Outro estudo (Kou *et al.*, 2022a) utilizou uma arquitetura que combina três redes neurais convolucionais. Uma identifica tipos de engolimento, outra identifica a pressão de engolimento e outra identifica a pressão de relaxamento integrada. As saídas dessas são passadas para três modelos que decidem, em paralelo, o diagnóstico de acordo com a Classificação de Chicago. Esses são: uma rede neural artificial, um modelo de regras e um modelo de *gradient*

boosting. A arquitetura desse sistema é ilustrada na figura 4, servindo de exemplo para ilustrar a lógica de sistemas compostos. Esse sistema alcançou uma acurácia de 81% com um conjunto de dados de 1741 imagens.

Além desses, há um estudo que combinou modelos de regressão logística com um modelo de *deep learning* do tipo *variational autoencoder* para identificar movimentos anormais em exames de manometria anorretal (Levy, 2023). As saídas desses modelos são combinadas em paralelo para criar um sistema híbrido, capaz de alcançar uma área sob a curva (ASC) de 0,96 usando dados de 302 exames de manometria.

Esses sistemas compostos podem ser comparados a 3 sistemas simples encontrados na literatura que realizam as mesmas tarefas, mas usando uma arquitetura mais simples. Dois sistemas utilizaram redes neurais convolucionais para laudar exames de manometria esofágica de acordo com a Classificação de Chicago, um alcançando 93% de acurácia com um conjunto de dados de 1570 imagens e outro alcançando 88% de acurácia com um dataset de 926 imagens (Stefan-Lucian *et al.*, 2022; Popa *et al.*, 2024). Um sistema usou um modelo de *gradient boosting* para detectar movimentos anorretais anormais, alcançando uma área sob a curva (ASC) de 0,939 com um conjunto de dados de 827 elementos (Saraiva *et al.*, 2023).

5.3 Q3: Dados Para O Desenvolvimento dos Sistemas

Os artigos presentes na literatura analisada usam, por via de regra, dados obtidos pelos próprios pesquisadores ou por suas instituições. Não foram encontrados, nesta revisão, trabalhos que utilizam conjuntos de dados públicos.

Os dados utilizados na literatura representam os valores do gráfico de pressão e distância em função do tempo durante os momentos em que são realizados movimentos do trato digestivo. Para isolar o momento em que esses movimentos ocorrem, os dados devem ser editados manualmente e, para o treinamento de modelos de aprendizado supervisionado, também devem ser classificados manualmente de acordo com os critérios desejados para o sistema.

Há dois tipos de dados observados na literatura: dados gráficos exportados em formato de imagem e dados numéricos exportados em formato tabular. O quadro 5 demonstra quais trabalhos usaram quais tipos de dados, em conjunto com quais tipos de modelos de aprendizado utilizaram esses dados.

Ao analisar este quadro, identifica-se que a escolha de modelo está correlacionada com a escolha de formato de dados. Modelos de florestas aleatórias, k-ésimo vizinho mais

Quadro 5 – Tipos de dados utilizados na literatura

Tabulares	
trabalho	modelos
Saraiva <i>et al.</i> (2023)	Florestas Aleatórias, K-ésimo Vizinho Mais Próximo, Máquinas de Vetor de Suporte, Gradient Boosting
Mielens <i>et al.</i> (2012)	Perceptrons de Múltiplas Camadas, Quantização de Vetores de Aprendizado, Máquinas de Vetor de Suporte
Kou <i>et al.</i> (2022b)	Long Short-Term Memory
Schilling <i>et al.</i> (2015)	Regressão Logística, Máquinas de Vetor de Suporte
Storonova <i>et al.</i> (2024)	Florestas Aleatórias, Gradient Boosting
Zifan <i>et al.</i> (2023)	Florestas Aleatórias, K-ésimo Vizinho Mais Próximo, Regressão Logística, Máquinas de Vetor de Suporte
Mascarenhas <i>et al.</i> (2025a)	Florestas Aleatórias, Gradient Boosting
Mascarenhas <i>et al.</i> (2025b)	Gradient Boosting
Imagens	
trabalho	modelos
Surdea-Blaga <i>et al.</i> (2022)	Redes Neurais Convolucionais
Kou <i>et al.</i> (2022a)	Perceptrons de Múltiplas Camadas, Redes Neurais Convolucionais, Gradient Boosting
Levy (2023)	Variational Autoencoder, Regressão Logística
Stefan-Lucian <i>et al.</i> (2022)	Redes Neurais Convolucionais
Czako <i>et al.</i> (2022)	Redes Neurais Convolucionais
Kou <i>et al.</i> (2021)	Variational Autoencoder
Popa <i>et al.</i> (2024)	Redes Neurais Convolucionais
Lin <i>et al.</i> (2025)	Redes Neurais Convolucionais

Fonte: compilação do autor.

próximo, máquinas de vetor de suporte, quantização de vetores de aprendizado e *long short-term memory* foram utilizados somente com dados em formato tabular. Modelos de redes neurais convolucionais e *variational autoencoder* foram utilizados somente com dados em formato de imagem. Modelos de percéptrons de múltiplas camadas, *gradient boosting* e regressão logística foram utilizados com ambos os formatos de dados.

5.4 Q4: Desempenho dos Modelos

A resposta a essa pergunta começa com uma comparação entre as melhores acurácias alcançadas por cada tipo de modelo encontrado na literatura. O quadro 6 destaca, para cada tipo de modelo, o exemplo com maior acurácia encontrado na literatura, demonstrando sua acurácia, sua tarefa e o tamanho do conjunto de dados usado em sua produção.

Como pode ser visto no quadro 6, todos os tipos de modelos testados na literatura demonstraram a capacidade de alcançar altas acurácias (>80%) na tarefa para a qual foram

Quadro 6 – Maior Acurácia Alcançada Por Tipo de Modelo

Modelo	Acurácia	Dataset	Tarefa
Perceptrons de Múltiplas Camadas (Mielens <i>et al.</i> , 2012)	Acurácia: 96,44%	23 exames, 4 engolimentos por exame	Classificação de engolimentos normais e anormais
Redes Neurais Convolucionais (Lin <i>et al.</i> , 2025)	Acurácia: 93,99%	2080 imagens de exame	Classificação de engolimentos normais e anormais
Variational Autoencoder (Kou <i>et al.</i> , 2021)	Acurácia (avaliada por análise discriminante linear): 86%	32.412 imagens de exame	Clusterização de datasets de manometria esofágica
Long Short-Term Memory (Kou <i>et al.</i> , 2022b)	Acurácia 81%	26.115 imagens de exame	Classificação de engolimentos normais e anormais
Florestas Aleatórias (Zifan <i>et al.</i> , 2023)	Acurácia: 90,5%	60 exames, 8-10 engolimentos por exame	Diagnosticar disfagia
Quantização de Vetores de Aprendizado (Mielens <i>et al.</i> , 2012)	Acurácia: 85,39%	23 exames, 4 engolimentos por exame	Classificação de engolimentos normais e anormais
K-Ésimo Vizinho Mais Próximo (Zifan <i>et al.</i> , 2023)	Acurácia: 88,1%	60 exames, 8-10 engolimentos por exame	Diagnosticar Disfagia
Regressão Logística (Zifan <i>et al.</i> , 2023)	Acurácia: 90,5%	60 exames, 8-10 engolimentos por exame	Diagnosticar Disfagia
Máquinas de Vetor de Suporte (Schilling <i>et al.</i> , 2015)	Acurácia: 90,26%	10 exames, 10 engolimentos por exame	Estimar tempo de relaxamento do esôfago após contração.
Gradient Boosting (Mascarenhas <i>et al.</i> , 2025b)	Acurácia: 94,2%	618 exames, 15 engolimentos por exame	Diagnosticar refluxo

Fonte: compilação do autor.

desenvolvidos. Isso dificulta afirmar exatamente qual tipo de modelo de aprendizado apresenta o melhor desempenho com base somente em sua capacidade de classificar dados corretamente. Ainda assim, pode-se identificar que as três melhores acurácias foram alcançadas, respectivamente, pelos seguintes tipos de modelo: Perceptrons de múltiplas camadas (96,44%), *gradient boosting* (94,2%) e redes neurais convolucionais (93,99%).

No mesmo quadro, também é possível identificar as diferenças no tamanho dos conjuntos de dados utilizados pelos diferentes modelos destacados. Ao contrário das diferenças entre acurácias, esses valores demonstram grande variação entre diferentes tipos de modelos. Observa-se que os modelos do tipo *variational autoencoder* e *long short term memory* consumi-

ram os maiores conjuntos de dados em seu desenvolvimento, na ordem de dezenas de milhares de elementos. Enquanto isso, os modelos de perceptrons de múltiplas camadas, quantização de vetores de aprendizado, máquinas de vetor de suporte, k-ésimo vizinho mais próximo, regressão logística e florestas aleatórias demonstraram o melhor desempenho em dados limitados, sendo capazes de alcançar altas acurácias com conjuntos de dados de 92 a 600 elementos.

Existem também certos modelos que não apareceram no quadro 6, pois não apresentaram a melhor acurácia dentro de sua categoria, mas que ainda devem ser mencionados na discussão do desempenho dos diferentes tipos de modelos identificados na literatura. Estes são os modelos que foram utilizados como decisores centrais em sistemas capazes de gerar diagnósticos completos para exames de manometria esofágica através da automação da Classificação de Chicago, sendo, portanto, os sistemas mais completos que foram identificados na literatura (Surdea-Blaga *et al.*, 2022; Kou *et al.*, 2022a; Popa *et al.*, 2024; Stefan-Lucian *et al.*, 2022). Esses sistemas utilizaram redes neurais convolucionais, demonstraram acurácias entre 81% e 93% e requereram conjuntos de dados de 926 a 2614 elementos.

6 CONCLUSÃO

A revisão de literatura realizada neste trabalho identificou 16 pesquisas que implementaram ferramentas de aprendizado de máquina para o auxílio ao diagnóstico de exames de manometria esofágica e anorretal. Ao analisar os métodos e os resultados dessas pesquisas, é possível identificar os principais usos e características desses sistemas, servindo de auxílio no desenvolvimento de futuros trabalhos na área.

Primeiramente, destaca-se que redes neurais convolucionais foram o tipo de modelo mais adotado na literatura, sendo utilizado em 6 estudos. Além disso, 4 desses estudos demonstraram que redes neurais convolucionais são capazes de automatizar a Classificação de Chicago para exames de manometria esofágica: demonstrando, portanto, a capacidade de oferecer diagnósticos completos para distúrbios do trato digestivo superior (Surdea-Blaga *et al.*, 2022; Kou *et al.*, 2022a; Popa *et al.*, 2024; Stefan-Lucian *et al.*, 2022).

O treinamento das redes neurais convolucionais identificadas na literatura requereu conjuntos de dados com milhares de imagens de gráficos de exames de manometria, selecionadas e diagnosticadas manualmente. Isso é menos do que foi necessário para treinar outros modelos de *deep learning* identificados na literatura, como o *variational autoencoder* ou o *long short-term memory*, que ambos precisam de conjuntos de dados na ordem de dezenas de milhares de elementos (Kou *et al.*, 2022b; Kou *et al.*, 2021). Mas ainda assim, isso esteve fora do alcance de muitos pesquisadores.

Vários estudos não tiveram acesso a tamanha quantidade de dados, mas mesmo assim conseguiram desenvolver sistemas inteligentes capazes de ajudar no processo de diagnóstico com alta acurácia. Esses sistemas usaram outros modelos de aprendizado e tiveram escopos limitados. Em vez de oferecer um diagnóstico completo, eles focaram em tarefas mais simples que ajudam no processo de diagnóstico, como detectar movimentos e/ou classificar movimentos como normais ou anormais (Jell *et al.*, 2020; Schilling *et al.*, 2015; Zifan *et al.*, 2023; Zifan *et al.*, 2024).

Dentre esses sistemas, destacam-se aqueles que usam modelos de perceptrons de múltiplas camadas, quantização de vetores de aprendizado, máquinas de vetor de suporte, k-ésimo vizinho mais próximo, regressão logística e florestas aleatórias, pois esses demonstraram que podem ser treinados em datasets com menos de 600 elementos (Schilling *et al.*, 2015; Zifan *et al.*, 2023; Mielens *et al.*, 2012). Esses modelos costumam a usar dados tabulares ao invés de dados em formato de imagem. Esses dados também foram extraídos e classificados

manualmente.

No entanto, deve-se ressaltar as limitações desta pesquisa. Dentre todos os artigos identificados e analisados, somente 3 deles focam em exames de manometria anorretal. Portanto, as conclusões desta revisão não podem ser aplicadas com o mesmo nível de confiança a sistemas que analisam dados de exames anorretais. Além disso, identifica-se que os estudos são, em sua maioria, unicêntricos e sem validação externa. Nenhum dos estudos analisados avaliam o impacto real de seus sistemas de aprendizado de máquina em decisões médicas.

Existe, portanto, amplo espaço para futuras pesquisas sobre o desenvolvimento de novos sistemas de aprendizado para exames de manometria anorretal. Há também espaço para pesquisas que busquem validar sistemas de auxílio ao diagnóstico de exames de manometria esofágica e anorretal em um ambiente clínico real.

REFERÊNCIAS

- ANTUNES, C.; ALEEM, A.; CURTIS, S. A. Gastroesophageal reflux disease. In: **StatPearls**. Treasure Island (FL): StatPearls Publishing, 2024.
- AWAD, M.; KHANNA, R. Machine learning. In: **Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers**. Berkeley, CA: Apress, 2015. p. 1–18. ISBN 978-1-4302-5990-9. Disponível em: https://doi.org/10.1007/978-1-4302-5990-9_1. Acesso em: 6 fev. 2026.
- BALDWIN, D.; PUCKETT, Y. Esophageal manometry. In: **StatPearls**. Treasure Island (FL): StatPearls Publishing, 2024.
- BASILISCO, G.; BHARUCHA, A. E. High-resolution anorectal manometry: An expensive hobby or worth every penny? **Neurogastroenterol Motil**, England, v. 29, n. 8, ago. 2017.
- BAUM, E. B. On the capabilities of multilayer perceptrons. **Journal of Complexity**, v. 4, n. 3, p. 193–215, 1988. ISSN 0885-064X. Disponível em: <https://www.sciencedirect.com/science/article/pii/0885064X88900209>. Acesso em: 3 fev. 2026.
- BHARUCHA, A. E.; BASILISCO, G.; MALCOLM, A.; LEE, T. H.; HOY, M. B.; SCOTT, S. M.; RAO, S. S. C. Review of the indications, methods, and clinical utility of anorectal manometry and the rectal balloon expulsion test. **Neurogastroenterology & Motility**, v. 34, n. 9, p. e14335, 2022. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/nmo.14335>. Acesso em: 31 jan. 2026.
- CARRINGTON, E. V.; HEINRICH, H.; KNOWLES, C. H.; FOX, M.; RAO, S.; ALTOMARE, D. F.; BHARUCHA, A. E.; BURGELL, R.; CHEY, W. D.; CHIARIONI, G.; DINNING, P.; EMMANUEL, A.; FAROUK, R.; FELT-BERSMA, R. J. F.; JUNG, K. W.; LEMBO, A.; MALCOLM, A.; MITTAL, R. K.; MION, F.; MYUNG, S.-J.; O'CONNELL, P. R.; PEHL, C.; REMES-TROCHE, J. M.; REVEILLE, R. M.; VAIZEY, C. J.; VITTON, V.; WHITEHEAD, W. E.; WONG, R. K.; SCOTT, S. M.; GROUP, A. members of the I. A. P. W. The international anorectal physiology working group (iapwg) recommendations: Standardized testing protocol and the london classification for disorders of anorectal function. **Neurogastroenterology & Motility**, v. 32, n. 1, p. e13679, 2020. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/nmo.13679>. Acesso em: 4 fev. 2026.
- CARRINGTON, E. V.; SCOTT, S. M.; BHARUCHA, A.; MION, F.; REMES-TROCHE, J. M.; MALCOLM, A.; HEINRICH, H.; FOX, M.; RAO, S. S.; on behalf of the International Anorectal Physiology Working Group and the International Working Group for Disorders of Gastrointestinal Motility and Function. Advances in the evaluation of anorectal function. **Nature Reviews Gastroenterology & Hepatology**, v. 15, n. 5, p. 309–323, maio 2018.
- COHEN, D.; SHIRIN, H. Technical success in performing esophageal high-resolution manometry: a review of competency recommendations, predictors of failure, and alternative techniques. **Diseases of the esophagus** : official journal of the International Society for Diseases of the Esophagus, v. 36, 03 2023.
- CONKLIN, J. Evaluation of esophageal motor function with high-resolution manometry. **Journal of neurogastroenterology and motility**, v. 19, p. 281–94, 07 2013.

CORTES, C.; VAPNIK, V. N. Support-vector networks. **Machine Learning**, v. 20, p. 273–297, 1995. Disponível em: <https://api.semanticscholar.org/CorpusID:52874011>. Acesso em: 4 fev. 2026.

CZAKO, Z.; SURDEA-BLAGA, T.; SEBESTYEN, G.; HANGAN, A.; DUMITRASCU, D. L.; DAVID, L.; CHIARIONI, G.; SAVARINO, E.; POPA, S. L. Integrated relaxation pressure classification and probe positioning failure detection in high-resolution esophageal manometry using machine learning. **Sensors**, v. 22, n. 1, 2022. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/22/1/253>. Acesso em: 6 fev. 2026.

DUBEY, S. R.; SINGH, S. K.; CHAUDHURI, B. B. Activation functions in deep learning: A comprehensive survey and benchmark. **Neurocomputing**, v. 503, p. 92–108, 2022. ISSN 0925-2312. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0925231222008426>. Acesso em: 3 fev. 2026.

FANG, J. C.; HILDEN, K.; TUTEJA, A. K.; PETERSON, K. A. Comparison of air-coupled balloon esophageal and anorectal manometry catheters with solid-state esophageal manometry and water-perfused anorectal manometry catheters. **Dig Dis Sci**, United States, v. 49, n. 10, p. 1657–1663, out. 2004.

FASS, O.; ROGERS, B. D.; GYAWALI, C. P. Artificial intelligence tools for improving manometric diagnosis of esophageal dysmotility. **Current Gastroenterology Reports**, v. 26, n. 4, p. 115–123, abr. 2024.

GOOGLE FOR DEVELOPERS. **Datasets**: Dividing the original dataset. [S. l.], 2025. Disponível em: <https://developers.google.com/machine-learning/crash-course/overfitting/dividing-datasets>. Acesso em: 06 mar. 2025.

HEINRICH, H.; MISSELWITZ, B. High-Resolution anorectal manometry - new insights in the diagnostic assessment of functional anorectal disorders. **Visc Med**, Switzerland, v. 34, n. 2, p. 134–139, abr. 2018.

HO, T. K. Random decision forests. In: **Proceedings of 3rd International Conference on Document Analysis and Recognition**. [S. l.: s. n.], 1995. v. 1, p. 278–282 vol.1.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 11 1997. ISSN 0899-7667. Disponível em: <https://doi.org/10.1162/neco.1997.9.8.1735>. Acesso em: 6 fev. 2026.

HOTEL, O.; POLI, J.-P.; MER-CALFATI, C.; SCORSONE, E.; SAADA, S. A review of algorithms for saw sensors e-nose based volatile compound identification. **Sensors and Actuators B: Chemical**, v. 255, p. 2472–2482, 2018. ISSN 0925-4005. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0925400517317057>. Acesso em: 6 fev. 2026.

HSIAO, M.-Y. 3 - adult neurogenic communication and swallowing disorders. In: CIFU, D. X.; LEW, H. L. (Ed.). **Braddom's Rehabilitation Care: A Clinical Handbook**. Elsevier, 2018. p. 18–23.e9. ISBN 978-0-323-47904-2. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780323479042000039>. Acesso em: 1 fev. 2026.

JANIESCH, C.; ZSCHECH, P.; HEINRICH, K. Machine learning and deep learning. **Electronic Markets**, v. 31, n. 3, p. 685–695, set. 2021.

JELL, A.; KUTTLER, C.; OSTLER, D.; HÜSER, N. How to cope with big data in functional analysis of the esophagus. **Visceral Medicine**, v. 36, n. 6, p. 439–442, 11 2020. ISSN 2297-4725. Disponível em: <https://doi.org/10.1159/000511931>. Acesso em: 6 fev. 2026.

KAHRILAS, P. J.; BREDENOORD, A. J.; FOX, M.; GYAWALI, C. P.; ROMAN, S.; SMOUT, A. J. P. M.; PANDOLFINO, J. E.; GROUP, I. H. R. M. W. The chicao classification of esophageal motility disorders, v3.0. **Neurogastroenterology & Motility**, v. 27, n. 2, p. 160–174, 2015. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/nmo.12477>. Acesso em: 4 fev. 2026.

KINGMA, D. P.; WELLING, M. **Auto-Encoding Variational Bayes**. 2022. Disponível em: <https://arxiv.org/abs/1312.6114>. Acesso em: 6 fev. 2026.

KOLASA BISRAT ADMASSU, M. H.-V. K. J. K. J.-E. P. K.; PERNI, S. Systematic reviews of machine learning in healthcare: a literature review. **Expert Review of Pharmacoeconomics & Outcomes Research**, Taylor & Francis, v. 24, n. 1, p. 63–115, 2024. PMID: 37955147. Disponível em: <https://doi.org/10.1080/14737167.2023.2279107>. Acesso em: 1 fev. 2026.

KOU, W.; CARLSON, D. A.; BAUMANN, A. J.; DONNAN, E.; LUO, Y.; PANDOLFINO, J. E.; ETEMADI, M. A deep-learning-based unsupervised model on esophageal manometry using variational autoencoder. **Artificial Intelligence in Medicine**, v. 112, p. 102006, 2021. ISSN 0933-3657. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0933365720312719>. Acesso em: 6 fev. 2026.

KOU, W.; CARLSON, D. A.; BAUMANN, A. J.; DONNAN, E. N.; SCHAUER, J. M.; ETEMADI, M.; PANDOLFINO, J. E. A multi-stage machine learning model for diagnosis of esophageal manometry. **Artificial Intelligence in Medicine**, v. 124, p. 102233, 2022. ISSN 0933-3657. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0933365721002268>. Acesso em: 4 fev. 2026.

KOU, W.; GALAL, G. O.; KLUG, M. W.; MUKHIN, V.; CARLSON, D. A.; ETEMADI, M.; KAHRILAS, P. J.; PANDOLFINO, J. E. Deep learning-based artificial intelligence model for identifying swallow types in esophageal high-resolution manometry. **Neurogastroenterology & Motility**, v. 34, n. 7, p. e14290, 2022. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/nmo.14290>. Acesso em: 31 jan. 2026.

KOUTSOS, T. M.; MENEXES, G. C.; DORDAS, C. A. An efficient framework for conducting systematic literature reviews in agricultural sciences. **Sci Total Environ**, Netherlands, v. 682, p. 106–117, abr. 2019.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, England, v. 521, n. 7553, p. 436–444, maio 2015.

LEE, Y.; RAGGUETT, R.-M.; MANSUR, R. B.; BOUTILIER, J. J.; ROSENBLAT, J. D.; TREVIZOL, A.; BRIETZKE, E.; LIN, K.; PAN, Z.; SUBRAMANIAPILLAI, M.; CHAN, T. C.; FUS, D.; PARK, C.; MUSIAL, N.; ZUCKERMAN, H.; CHEN, V. C.-H.; HO, R.; RONG, C.; MCINTYRE, R. S. Applications of machine learning algorithms to predict therapeutic outcomes in depression: A meta-analysis and systematic review. **Journal of Affective Disorders**, v. 241, p. 519–532, 2018. ISSN 0165-0327. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0165032718304853>. Acesso em: 6 fev. 2026.

LEVY, J. J. Video-based deep learning to detect dyssynergic defecation with 3d high-definition anorectal manometry. **Digestive Diseases and Sciences**, 2023. Disponível em: <https://link.springer.com/article/10.1007/s10620-022-07759-3>. Acesso em: 01 mai. 2024.

LI, B.; FERIDOONI, T.; CUEN-OJEDA, C.; KISHIBE, T.; MESTRAL, C. de; MAMDANI, M.; AL-OMRAN, M. Machine learning in vascular surgery: a systematic review and critical appraisal. **npj Digital Medicine**, v. 5, n. 1, p. 7, jan. 2022.

LIN, J.; WU, X.; CHEN, Q.; GUO, C.; MA, L.; LUO, S.; HUANG, J. Portablenet: An automatic diagnosis method for esophageal motility disorders. **IEEE Access**, v. 13, p. 25806–25816, 2025.

MARCOT, B.; HANEA, A. What is an optimal value of k in k-fold cross-validation in discrete bayesian network analysis? **Computational Statistics**, v. 36, 09 2021.

MASCARENHAS, M.; MENDES, F.; MOTA, J.; RIBEIRO, T.; CARDOSO, P.; MARTINS, M.; ALMEIDA, M. J.; CORDEIRO, J. R.; FERREIRA, J.; MACEDO, G.; SANTANDER, C. Artificial intelligence as a transforming factor in motility disorders-automatic detection of motility patterns in high-resolution anorectal manometry. **Sci Rep**, England, v. 15, n. 1, p. 2061, jan. 2025.

MASCARENHAS, M.; MOTA, J.; CORDEIRO, J. R.; MENDES, F.; MARTINS, M.; CARDOSO, P.; ALMEIDA, M. J.; COSTA, A. Pinto da; MARTINEZ, I. H.; ROYO, V. M.; NILAND, B.; PALMA, J. D.; FERREIRA, J.; MACEDO, G.; SANTANDER, C. Artificial intelligence driven diagnosis of motility patterns in high-resolution esophageal manometry: A multicentric multidevice study. **Clinical and Translational Gastroenterology**, 2025. ISSN 2155-384X. Disponível em: https://journals.lww.com/ctg/fulltext/9900/artificial_intelligence_driven_diagnosis_of.481.aspx. Acesso em: 6 fev. 2026.

MASON, L.; BAXTER, J.; BARTLETT, P. L.; FREAN, M. Boosting algorithms as gradient descent. In: **Neural Information Processing Systems**. [S. n.], 1999. Disponível em: <https://api.semanticscholar.org/CorpusID:6101385>. Acesso em: 4 fev. 2026.

MIELENS, J. D.; HOFFMAN, M. R.; CIUCCI, M. R.; MCCULLOCH, T. M.; JIANG, J. J. Application of classification models to pharyngeal high-resolution manometry. **J Speech Lang Hear Res**, United States, v. 55, n. 3, p. 892–902, jan. 2012.

NISHIZAKI, H. Data augmentation and feature extraction using variational autoencoder for acoustic modeling. In: **2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)**. [S. l.: s. n.], 2017. p. 1222–1227.

PAGE, M. J.; MCKENZIE, J. E.; BOSSUYT, P. M.; BOUTRON, I.; HOFFMANN, T. C.; MULROW, C. D.; SHAMSEER, L.; TETZLAFF, J. M.; AKL, E. A.; BRENNAN, S. E.; CHOU, R.; GLANVILLE, J.; GRIMSHAW, J. M.; HRÓBJARTSSON, A.; LALU, M. M.; LI, T.; LODER, E. W.; MAYO-WILSON, E.; MCDONALD, S.; MCGUINNESS, L. A.; STEWART, L. A.; THOMAS, J.; TRICCO, A. C.; WELCH, V. A.; WHITING, P.; MOHER, D. The prisma 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, BMJ Publishing Group Ltd, v. 372, 2021. Disponível em: <https://www.bmj.com/content/372/bmj.n71>. Acesso em: 6 fev. 2026.

PETERSON, L. K-nearest neighbor. **Scholarpedia**, v. 4, p. 1883, 01 2009.

POPA, S. L.; SURDEA-BLAGA, T.; DUMITRASCU, D. L.; POP, A. V.; ISMAIEL, A.; DAVID, L.; BRATA, V. D.; TURTOI, D. C.; CHIARIONI, G.; SAVARINO, E. V.; ZSIGMOND, I.; CZAKO, Z.; LEUCUTA, D. C. Gemini-assisted deep learning classification model for automated diagnosis of high-resolution esophageal manometry images. **Medicina**, v. 60, n. 9, 2024. ISSN 1648-9144. Disponível em: <https://www.mdpi.com/1648-9144/60/9/1493>. Acesso em: 5 fev. 2026.

SARAIVA, M. M.; POUCA, M. V.; RIBEIRO, T.; AFONSO, J.; CARDOSO, H.; SOUSA, P.; FERREIRA, J.; MACEDO, G.; JUNIOR, I. F. Artificial intelligence and anorectal manometry: Automatic detection and differentiation of anorectal motility Patterns-A Proof-of-Concept study. **Clin Transl Gastroenterol**, United States, v. 14, n. 10, p. e00555, out. 2023.

SCHARDT, C.; ADAMS, M. B.; OWENS, T.; KEITZ, S.; FONTELO, P. Utilization of the PICO framework to improve searching PubMed for clinical questions. **BMC Med Inform Decis Mak**, England, v. 7, p. 16, jun. 2007.

SCHILLING, N.; BUSCHE, A.; MILLER, S.; JUNGHEIM, M.; PTOK, M.; SCHMIDT-THIEME, L. Event prediction in pharyngeal high-resolution manometry. In: LAUSEN, B.; KROLAK-SCHWERDT, S.; BÖHMER, M. (Ed.). **Data Science, Learning by Latent Structures, and Knowledge Discovery**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015. p. 341–352. ISBN 978-3-662-44983-7.

STEFAN-LUCIAN, P.; SURDEA-BLAGA, T.; DUMITRASCU, D.; CHIARIONI, G.; SAVARINO, E.; DAVID, L.; ISMAIEL, A.; DANIEL-CORNELIU, L.; ZSIGMOND, I.; SEBESTYEN, G.; HANGAN, A.; CZAKO, Z. Automatic diagnosis of high-resolution esophageal manometry using artificial intelligence. **Journal of Gastrointestinal and Liver Diseases**, v. 31, p. 383–389, 12 2022.

STORONOVA, O.; KANEVSKII, N.; TRUKHMANOV, A.; IVASHKIN, V. Own experience in the use of artificial intelligence technologies in the diagnosis of esophageal achalasia. **Russian Journal of Gastroenterology, Hepatology, Coloproctology**, v. 34, p. 32–39, 12 2024.

SURDEA-BLAGA, T.; SEBESTYEN, G.; CZAKO, Z.; HANGAN, A.; DUMITRASCU, D. L.; ISMAIEL, A.; DAVID, L.; ZSIGMOND, I.; CHIARIONI, G.; SAVARINO, E.; LEUCUTA, D. C.; POPA, S. L. Automated chicao classification for esophageal motility disorder diagnosis using machine learning. **Sensors**, v. 22, n. 14, 2022. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/22/14/5227>. Acesso em: 4 fev. 2026.

SYNDER, H. Literature review as a research methodology: An overview and guidelines. **Journal of Business Research**, 2019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0148296319304564>. Acesso em: 30 abr. 2024.

TEALAB, A. Time series forecasting using artificial neural networks methodologies: A systematic review. **Future Computing and Informatics Journal**, v. 3, n. 2, p. 334–340, 2018. ISSN 2314-7288. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2314728817300715>. Acesso em: 6 fev. 2026.

THIRUNAVUKARASU, A. J.; ELANGOVA, K.; GUTIERREZ, L.; HASSAN, R.; LI, Y.; TAN, T. F.; CHENG, H.; TEO, Z. L.; LIM, G.; TING, D. S. W. Clinical performance of automated machine learning: A systematic review. **Ann Acad Med Singap**, Singapore, v. 53, n. 3, p. 187–207, mar. 2024.

TOLLES, J.; MEURER, W. J. Logistic regression: Relating patient characteristics to outcomes. **JAMA**, v. 316, n. 5, p. 533–534, 08 2016. ISSN 0098-7484. Disponível em: <https://doi.org/10.1001/jama.2016.7653>. Acesso em: 6 fev. 2026.

UNIVERSITY OF MICHIGAN LIBRARY. **Evidence Syntheses (Scoping, systematic, other types of reviews)**. [S. l.], 2024. Disponível em: <https://guides.lib.umich.edu/c.php?g=283340&p=2126706>. Acesso em: 24 set. 2024.

ZIFAN, A.; LEE, J. M.; MITTAL, R. K. Enhancing the diagnostic yield of esophageal manometry using distension-contraction plots of peristalsis and artificial intelligence. **American Journal of Physiology-Gastrointestinal and Liver Physiology**, v. 327, n. 3, p. G405–G413, 2024. PMID: 38953836. Disponível em: <https://doi.org/10.1152/ajpgi.00139.2024>. Acesso em: 5 fev. 2026.

ZIFAN, A.; LIN, J.; PENG, Z.; BO, Y.; MITTAL, R. K. Unraveling functional dysphagia: A game-changing automated machine-learning diagnostic approach. **Applied Sciences**, v. 13, n. 18, 2023. ISSN 2076-3417. Disponível em: <https://www.mdpi.com/2076-3417/13/18/10116>. Acesso em: 5 fev. 2026.