



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

Bacharelado em Ciência da Computação

Rodrigo de Freitas Fornazier

Thiago Barroso de Rezende

**Aprendizado de Máquina em Sistemas de Diagnóstico Médico**

Belo Horizonte

2024

Rodrigo de Freitas Fornazier  
Thiago Barroso de Rezende

## **Aprendizado de Máquina em Sistemas de Diagnóstico Médico**

Projeto de Pesquisa apresentado na disciplina Trabalho Interdisciplinar III - Pesquisa Aplicada do curso de Ciência da Computação da Pontifícia Universidade Católica de Minas Gerais.

Belo Horizonte

2024

## 1 RESUMO

Este trabalho explora o potencial do aprendizado de máquina na área da saúde, com foco na utilização de algoritmos para o diagnóstico de Transtorno de Déficit de Atenção e Hiperatividade (TDAH) em crianças. A pesquisa é qualitativa e exploratória, estruturada a partir de uma revisão sistemática da literatura publicada entre 2018 e 2024, contemplando estudos que evidenciam o uso de técnicas avançadas de aprendizado de máquina, como redes neurais, árvores de decisão e algoritmos de ensemble. Essas técnicas são aplicadas à análise de dados clínicos e neuropsicológicos, visando a detecção precoce e a personalização dos diagnósticos.

A metodologia deste estudo incluiu etapas fundamentais: levantamento bibliográfico, organização e análise crítica dos dados, e proposição de soluções inovadoras para o uso de aprendizado de máquina no campo da saúde. Foram exploradas aplicações práticas, como o uso de registros eletrônicos de saúde (EHRs) para treinar modelos preditivos, o que possibilita a identificação de padrões complexos muitas vezes imperceptíveis a análises convencionais.

Os resultados esperados incluem a síntese de melhores práticas no uso de inteligência artificial para diagnósticos médicos e a proposição de abordagens que ampliem a precisão e a eficiência dos processos clínicos. Além disso, espera-se contribuir para a fundamentação teórica da utilização de aprendizado de máquina no diagnóstico de TDAH, destacando desafios e oportunidades tecnológicas no setor.

Palavras-chave: aprendizado de máquina, TDAH, diagnóstico médico, inteligência artificial, neuropsicologia, saúde digital, modelos preditivos.

## SUMÁRIO

2	INTRODUÇÃO .....	25
2.1	Objetivos .....	25
2.1.1	<i>Objetivos específicos</i> .....	25
3	REVISÃO BIBLIOGRÁFICA.....	27
4	METODOLOGIA.....	29
4.1	Classificação da Pesquisa .....	29
4.1.1	<i>Métodos de Coleta de Dados</i> .....	29
4.1.2	<i>Estratégias de Análise</i> .....	30
4.2	Atividades a Serem Realizadas .....	30
4.2.1	<i>Atividade 1: Levantamento Bibliográfico</i> .....	30
4.2.2	<i>Atividade 2: Análise dos Estudos Seleccionados</i> .....	30
4.2.3	<i>Atividade 3: Proposição de Aplicações</i> .....	30
4.2.4	<i>Cronograma</i> .....	31
5	PRIMEIRO CAPÍTULO DE EXEMPLO.....	32
5.1	Primeira seção .....	32
5.1.1	<i>Primeira subseção</i> .....	33
5.2	Segunda seção .....	33
6	SEGUNDO CAPÍTULO DE EXEMPLO.....	34
7	OBSERVAÇÕES IMPORTANTES .....	36
	REFERÊNCIAS .....	37

## 2 INTRODUÇÃO

O avanço da tecnologia nas últimas décadas trouxe inovações significativas para a área da saúde, e o aprendizado de máquina (Machine Learning) tem se destacado como uma dessas transformações. No contexto dos sistemas de diagnóstico médico, o aprendizado de máquina permite a análise de grandes volumes de dados clínicos, facilitando a detecção precoce de doenças, a personalização de tratamentos e a melhoria nos processos de tomada de decisão médica. Esses algoritmos podem aprender com os dados históricos de pacientes, identificar padrões complexos que muitas vezes escapam à observação humana, e oferecer previsões mais precisas e rápidas.

*"A inteligência artificial, ao analisar vastos conjuntos de dados de pacientes e identificar padrões sutis, melhora significativamente a precisão dos diagnósticos, permitindo a detecção precoce de doenças e tratamentos personalizados."(Irfan, S., Anjum, N., Althobaiti, T., Alotaibi, A., Siddiqui, A. (2022).).*

Problema

Justificativa

Este trabalho está organizado da seguinte forma. A seção 2.1 asdadas dasda. O capítulo ?? apresenta o referencial teórico usado neste trabalho. O capítulo 4 descreve os procedimentos metodológicos ...

### 2.1 Objetivos

O objetivo geral deste projeto é investigar como a inteligência artificial (IA) pode ser aplicada para auxiliar no diagnóstico médico, avaliando seu potencial para promover melhorias na precisão, eficiência e confiabilidade da área médica.

#### 2.1.1 *Objetivos específicos*

Os objetivos específicos deste projeto são:

1. Analisar as principais abordagens de inteligência artificial utilizadas no diagnóstico

médico, identificando seus benefícios e limitações.

2. Avaliar a aplicabilidade e confiabilidade da IA no apoio à tomada de decisões médicas, especialmente no diagnóstico de transtornos como o TDAH.
3. Explorar os desafios éticos, técnicos e operacionais envolvidos na adoção de sistemas baseados em IA no diagnóstico médico.
4. Propor diretrizes ou recomendações para o uso eficaz de IA em diagnósticos clínicos, considerando as necessidades dos profissionais da saúde e dos pacientes.
5. Investigar o impacto do uso de IA na relação médico-paciente, considerando aspectos de confiança, transparência e aceitação das tecnologias de diagnóstico automatizado.
6. Identificar tendências futuras na aplicação de IA no diagnóstico médico, com foco em inovações que possam transformar práticas clínicas nos próximos anos.

### 3 REVISÃO BIBLIOGRÁFICA

Este capítulo apresenta uma revisão sobre o uso do aprendizado de máquina (AM) nos sistemas de diagnóstico médico. A revisão inclui a fundamentação teórica e uma análise dos trabalhos relacionados que exploram a aplicação de técnicas de IA para otimizar diagnósticos e tratamentos personalizados.

#### 2.1 Fundamentação Teórica

O aprendizado de máquina tem transformado o campo da medicina, trazendo inovações no diagnóstico de diversas condições, como doenças cardíacas e renais. Segundo Irfan et al. (2022), a IA pode analisar grandes volumes de dados clínicos, permitindo diagnósticos mais rápidos e intervenções personalizadas. A capacidade de prever condições antes de seu desenvolvimento, como a previsão de lesão renal aguda, possibilita a adoção de medidas preventivas que podem salvar vidas (Irfan et al., 2022). A combinação de AM com dados clínicos permite a criação de modelos mais precisos, que superam métodos tradicionais e reduzem a subjetividade no diagnóstico médico (Al-antari, 2024).

A utilização de registros eletrônicos de saúde (EHRs) para treinar algoritmos tem sido uma abordagem promissora, como apontado por Boeken et al. (2023), que demonstraram como a integração desses dados pode melhorar a precisão diagnóstica. Esses avanços, associados ao desenvolvimento de modelos híbridos, oferecem uma nova perspectiva para diagnósticos mais rápidos e personalizados.

#### 2.2 Trabalhos Relacionados

Diversos estudos recentes têm aplicado aprendizado de máquina para aprimorar diagnósticos médicos. Irfan et al. (2022) desenvolveram algoritmos que utilizam EHRs para detectar padrões em grandes volumes de dados clínicos, alcançando alta precisão na previsão de doenças cardíacas. Essa abordagem permite que profissionais de saúde intervenham de forma proativa, reduzindo a necessidade de internações prolongadas e custos relacionados.

Outro estudo relevante, realizado por Al-antari (2024), explorou o uso de aprendizado de máquina para o diagnóstico de lesão renal aguda, com algoritmos que preveem a condição com até 48 horas de antecedência. Isso demonstra o potencial da IA em prevenir

doenças graves e melhorar significativamente a qualidade do atendimento ao paciente.

Esses estudos reforçam a capacidade do aprendizado de máquina de transformar o diagnóstico médico, tornando-o mais preciso, rápido e acessível. O uso dessas tecnologias pode ser ampliado para outras áreas da medicina, potencializando ainda mais a personalização dos tratamentos e a melhoria dos resultados clínicos.

### 2.3 Aplicações do Massachusetts General Hospital (MGH)

O Massachusetts General Hospital (MGH), em parceria com o MIT, tem utilizado inteligência artificial para revolucionar o diagnóstico médico. Um exemplo notável é o sistema desenvolvido para a detecção precoce de câncer e doenças cardiovasculares. Com a combinação de grandes volumes de dados clínicos e algoritmos de aprendizado de máquina, o MGH tem conseguido identificar padrões ocultos em imagens médicas e exames laboratoriais, proporcionando diagnósticos mais rápidos e precisos. Esses sistemas reduzem o erro humano e melhoram os resultados clínicos, impactando diretamente a medicina personalizada.

"O aprendizado de máquina tem revolucionado o diagnóstico médico ao melhorar a precisão e reduzir o tempo necessário para identificar doenças, permitindo que os profissionais de saúde façam intervenções mais rápidas e eficazes. Esses avanços proporcionam uma análise mais profunda de dados clínicos complexos, resultando em diagnósticos mais precisos e tratamentos personalizados."(Chen et al., 2023).



## 4 METODOLOGIA

Este capítulo descreve a metodologia utilizada para conduzir o estudo sobre a aplicação de algoritmos de aprendizado de máquina no diagnóstico médico, com ênfase no uso de inteligência artificial para a detecção de TDAH em crianças. A metodologia foi estruturada para revisar e sintetizar a literatura existente, identificar lacunas e propor abordagens baseadas em evidências encontradas nos estudos analisados.

### 4.1 Classificação da Pesquisa

A pesquisa pode ser classificada como exploratória e qualitativa, com foco na análise bibliográfica de estudos publicados em periódicos científicos, conferências e relatórios técnicos. O objetivo principal é compreender como as tecnologias de aprendizado de máquina têm sido aplicadas no diagnóstico médico e, especificamente, como essas ferramentas podem ser adaptadas e otimizadas para o diagnóstico de TDAH.

#### 4.1.1 *Métodos de Coleta de Dados*

Os dados utilizados na pesquisa foram coletados por meio de revisão sistemática da literatura, incluindo artigos indexados em bases de dados científicas como PubMed, IEEE Xplore, e Scopus. Os critérios de inclusão foram:

Publicações entre 2018 e 2024.

Estudos relacionados à aplicação de aprendizado de máquina no diagnóstico de condições médicas, com ênfase em dados clínicos e neuropsicológicos.

Trabalhos que discutem a utilização de IA para personalização de diagnósticos e tratamentos.

Os critérios de exclusão incluíram estudos que não abordam diretamente o uso de aprendizado de máquina ou que não apresentem resultados quantitativos relevantes para o diagnóstico médico.

#### **4.1.2 Estratégias de Análise**

A análise foi realizada em três etapas principais:

**Coleta e Organização dos Dados:** Identificação e extração das informações relevantes dos estudos selecionados, incluindo algoritmos utilizados, tipos de dados clínicos analisados e resultados obtidos.

**Análise Crítica:** Comparação dos métodos e resultados apresentados nos estudos. Essa etapa incluiu a avaliação da eficácia dos algoritmos utilizados e das limitações relatadas pelos autores.

**Síntese das Informações:** Elaboração de um panorama geral sobre as melhores práticas e desafios na aplicação do aprendizado de máquina no diagnóstico médico, com ênfase em diagnósticos de TDAH.

### **4.2 Atividades a Serem Realizadas**

Esta seção apresenta as principais atividades desenvolvidas no estudo.

#### **4.2.1 Atividade 1: Levantamento Bibliográfico**

**Descrição:** Revisão da literatura existente sobre o uso de aprendizado de máquina em diagnósticos médicos. Foram identificados estudos que detalham técnicas aplicadas, como redes neurais, árvores de decisão e algoritmos de ensemble.

#### **4.2.2 Atividade 2: Análise dos Estudos Selecionados**

**Descrição:** Comparação dos diferentes algoritmos e suas aplicações em diagnósticos. Essa etapa incluiu a análise dos estudos de Irfan et al. (2022), Al-antari (2024) e Chen et al. (2023), que exploram casos de uso em diferentes condições médicas.

#### **4.2.3 Atividade 3: Proposição de Aplicações**

**Descrição:** Com base na revisão bibliográfica, foram propostas aplicações específicas para o diagnóstico de TDAH em crianças, considerando dados neuropsicológicos e clínicos como entradas para algoritmos de aprendizado de máquina.

#### 4.2.4 *Cronograma*

**Tabela 1 – Cronograma do Projeto de Pesquisa**

<b>Data</b>	<b>Atividade</b>
23/08	Discussão relativo à bibliografias relevantes
22/09	Pesquisa do tema
25/09	Defesa do tema
20/10	Aprofundamento da aplicação da pesquisa
21/10	Apresentação das pesquisas profundas da pesquisa
29/11	Decisão das metodologias da pesquisa
07/12	Finalização e complemento da tese da pesquisa

## 5 PRIMEIRO CAPÍTULO DE EXEMPLO

A seguir serão apresentados alguns comandos do LaTeX usados comumente para formatar textos de dissertação baseados na normalização da PUC (2011).

Para as citações a norma estabelece duas formas de apresentação. A primeira delas é empregada quando a citação aparece no final de um parágrafo. Neste caso, o comando `cite` é usado para formatar a citação em caixa alta, como é mostrado no exemplo a seguir. (DUATO; YALAMANCHILI; LIONEL, 2002).

Outra forma de apresentação da citação é a que ocorre no decorrer do texto, essa situação é exemplificada na próxima frase. Conforme Bjerregaard e Mahadevan (2006), o estudo mencionado revela progressos no desempenho dos processadores. Para a formatação da citação em caixa baixa deve ser usado o comando `citeonline`.

Nas citações que aparecem mais de uma referência as mesmas devem ser separadas por vírgulas, como neste exemplo. (KEYES, 2008; ZHAO, 2008; GANGULY et al., 2011). Se houver necessidade de especificar a página ou que foi realizada uma tradução do texto deve ser feito da seguinte maneira. (SASAKI et al., 2009, p. 2, tradução nossa). A citação direta deve ser feita de forma semelhante. “[...] A carga de trabalho de um sistema pode ser definida como o conjunto de todas as informações de entrada.” (MENASCE; ALMEIDA, 2002, p. 160).

O arquivo `dissertacao.bib` mostra exemplos de representação para vários tipos de referências (artigos de conferências, periódicos, relatórios, livros, dentre outros). Cada um desses tipos requer uma forma diferente de representação para que a referência seja formatada conforme as exigências da normalização.

### 5.1 Primeira seção

Para gerar a lista de siglas automaticamente deve ser usado o pacote *acronym*. Para tanto, toda vez que uma sigla for mencionada no texto deve ser usado o comando `ac{sigla}`. Dessa forma, se for a primeira ocorrência da sigla a mesma será escrita por extenso conforme descrição feita no arquivo `lista-siglas.tex`. Caso contrário, somente a sigla será mostrada. Ex

### 5.1.1 *Primeira subseção*

As enumerações devem ser geradas usando o pacote *compactitem*. Cada item deve terminar com um ponto final. Abaixo um exemplo de enumeração é apresentado:

- a) Coletar e analisar.
- b) Configurar e simular.
- c) Definir a metodologia.
- d) Avaliar o desempenho.
- e) Analisar e avaliar características.

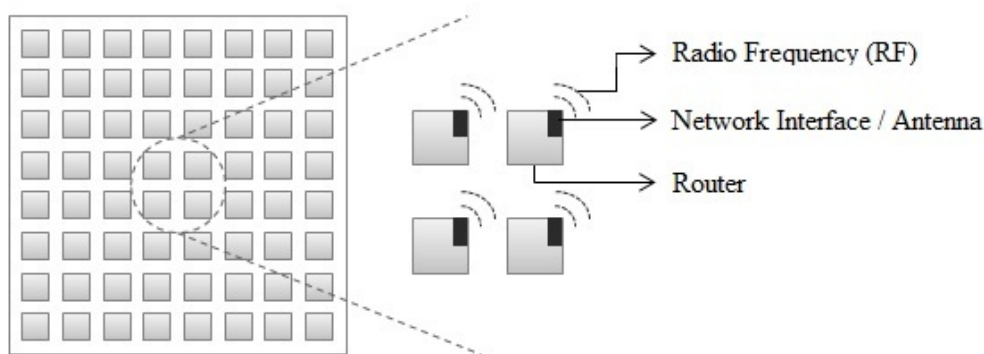
## 5.2 Segunda seção

Para referenciar um capítulo, seção ou subseção basta definir um label para o mesmo e usar o comando `ref` para referenciá-lo no texto. Exemplo: Como pode ser visto no Capítulo 5 ou na Seção 5.1.

## 6 SEGUNDO CAPÍTULO DE EXEMPLO

As figuras devem ser apresentadas pelos comandos abaixo. O parâmetro *width* determina o tamanho que a figura será exibida. No parâmetro *caption* o texto que aparece entre colchetes será o exibido no índice de figuras e o texto contido entre chaves será exibido na legenda da figura. Para citar a figura o comando *ref* deve ser usado juntamente com o *label*, como é mostrado nesse exemplo da Figura 1.

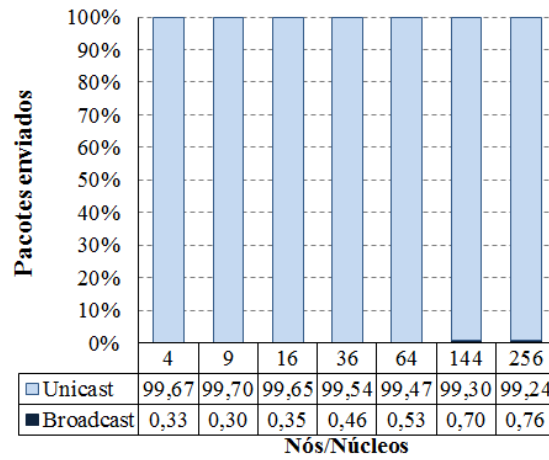
**Figura 1 – Principais componentes de WiNoCs**



**Fonte: (OLIVEIRA et al., 2011)**

Os comandos abaixo são usados para apresentação de gráficos. A diferença está apenas na definição do tipo “grafico” que permite a adição dos itens no índice de gráficos de forma automática. Os parâmetros são semelhantes aos usados para representação de figuras. O parâmetro *width* determina o tamanho do gráfico. O texto entre colchetes no *caption* será o exibido no índice de gráficos e o texto contido entre chaves será exibido na legenda.

Gráfico 1 – Percentual de pacotes enviados



Fonte: Dados da pesquisa

Um exemplo de criação de tabela é mostrado a seguir. As colunas são separadas por elementos & e as linhas por duas barras invertidas. Os comandos *hline* e *|* definem a criação de linhas e colunas para separar os conteúdos, respectivamente. A tabela pode ser referenciada usando o comando *ref* juntamente com o label, como na Tabela 2.

Tabela 2 – Parâmetros definidos por classe

<i>Benchmark</i>	Parâmetro	Classe S	Classe W	Classe A	Classe B	Classe C	Classe D
BT	<i>Grid</i>	$12^3$	$24^3$	$64^3$	$102^3$	$162^3$	$408^3$
CG	Linhas	1400	7000	14000	75000	150000	1500000
EP	Pares	$2^{24}$	$2^{25}$	$2^{28}$	$2^{30}$	$2^{32}$	$2^{36}$
FT	<i>Grid</i>	$64^3$	$128^2 * 32$	$256^2 * 128$	$512 * 256^2$	$512^3$	$2048 * 1024^2$
IS	Chaves	$2^{16}$	$2^{20}$	$2^{23}$	$2^{25}$	$2^{27}$	$2^{31}$
LU	<i>Grid</i>	$12^3$	$33^3$	$64^3$	$102^3$	$162^3$	$408^3$
MG	<i>Grid</i>	$32^3$	$128^3$	$256^3$	$256^3$	$512^3$	$1024^3$
SP	<i>Grid</i>	$12^3$	$36^3$	$64^3$	$102^3$	$162^3$	$408^3$

Fonte: Adaptado de (NPB, 2011)

## 7 OBSERVAÇÕES IMPORTANTES

Este documento foi compilado em ambiente linux (Ubuntu 10.04) usando o programa Kile - an Integrated LaTeX Environment - Version 2.0.85. Para correta formatação os seguintes arquivos do pacote *abntex* devem ser alterados.

a) Arquivo abnt.cls

No Ubuntu o arquivo fica armazenado em */usr/share/texmf/tex/latex/abntex*. Comentar a linha 967: Linha comentada para reduzir o espaçamento entre o topo da página e o título. Alterar a linha 1143: Parâmetro alterado de 30pt para -30pt para reduzir o espaçamento entre o top da página e o título do apêndice. Alterar a linha 985: Parâmetro alterado de 0pt para -30pt para reduzir o espaçamento entre o top da página e o título. Alterar a linha 991: Parâmetro alterado de 45pt para 30pt para reduzir o espaçamento entre o texto e o título.

b) Arquivo acronym.sty

No Ubuntu o arquivo fica armazenado em */usr/share/texmf-texlive/tex/latex/acronym*. Alterar a linha 225: Inserir o separador – entre acrônimo/descrição e remover o negrito com o *normalfont*.



## REFERÊNCIAS

- BJERREGAARD, T.; MAHADEVAN, S. A survey of research and practices of network-on-chip. **Computing Surveys**, ACM, New York, USA, v. 38, n. 1, p. 1–51, Jun. 2006. ISSN 0360-0300.
- DUATO, J.; YALAMANCHILI, S.; LIONEL, N. **Interconnection networks: an engineering approach**. San Francisco: Morgan Kaufmann Publishers, 2002. 515 p. ISBN 1558608524.
- GANGULY, A. et al. Scalable hybrid wireless network-on-chip architectures for multi-core systems. **Journal Transactions on Computers**, IEEE Computer Society, Los Alamitos, USA, v. 60, n. 10, p. 1485–1502, 2011. ISSN 0018-9340.
- KEYES, R. W. Moore’s law today. **Circuits and Systems Magazine**, IEEE Computer Society, Los Alamitos, USA, v. 8, n. 2, p. 53–54, 2008.
- MENASCE, D. A.; ALMEIDA, V. A. F. **Planejamento de capacidade para serviços na web: métricas, modelos e métodos**. Rio de Janeiro: Campus, 2002. 472 p. ISBN 8535211020.
- NPB. **NAS Parallel Benchmarks**. Disponível em <http://www.nas.nasa.gov/publications/npb.html>. Acesso em jun. 2011.
- OLIVEIRA, P. A. C. et al. Performance evaluation of winocs for parallel workloads based on collective communications. In: IADIS APPLIED COMPUTING, 8., 2011, Rio de Janeiro, Brasil. **Proceedings...** Rio de Janeiro: IADIS Applied Computing, 2011. p. 307–314.
- SASAKI, N. et al. A single-chip ultra-wideband receiver with silicon integrated antennas for inter-chip wireless interconnection. **Journal of Solid-State Circuits**, IEEE Computer Society, Los Alamitos, USA, v. 44, n. 2, p. 382–393, Feb. 2009. ISSN 0018-9200.
- ZHAO, D. Ultraperformance wireless interconnect nanonetworks for heterogeneous gigascale multi-processor SoCs. In: 2TH WORKSHOP ON CHIP MULTIPROCESSOR, MEMORY SYSTEMS AND INTERCONNECTS, 3., 2008, Beijing, China. **Proceedings...** Beijing: CMP-MSI, 2008. p. 1–3.