

**UNIVERSIDADE DE SÃO PAULO**

Instituto de Ciências Matemáticas e de Computação

**Metodologia para prever a reinternação hospitalar  
de pacientes baseada em Inteligência Artificial  
Explicável**

**Matheus Mendes dos Santos**

Monografia - MBA em Ciência de Dados (CEMEAI)



SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: \_\_\_\_\_

**Matheus Mendes dos Santos**

## **Metodologia para prever a reinternação hospitalar de pacientes baseada em Inteligência Artificial Explicável**

Monografia apresentada ao Centro de Ciências Matemáticas Aplicadas à Indústria do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Ciências de Dados.

Área de concentração: Ciências de Dados

Orientador: Prof. Dr. Rodrigo Colnago Contreras

**Versão original**

**São Carlos**

**2023**

É possível elaborar a ficha catalográfica em LaTeX ou incluir a fornecida pela Biblioteca. Para tanto observe a programação contida nos arquivos USPSC-modelo.tex e fichacatalografica.tex e/ou gere o arquivo fichacatalografica.pdf.

A biblioteca da sua Unidade lhe fornecerá um arquivo PDF com a ficha catalográfica definitiva, que deverá ser salvo como fichacatalografica.pdf no diretório do seu projeto.

**Matheus Mendes dos Santos**

# **Methodology for Predicting Hospital Readmission of Patients Based on Explainable Artificial Intelligence**

Monograph presented to the Centro de Ciências Matemáticas Aplicadas à Indústria do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, as part of the requirements for obtaining the title of Specialist in Data Science.

Concentration area: Data Science

Advisor: Prof. Dr. Rodrigo Colnago Contreras

**Original version**

**São Carlos**

**2023**



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>9</b>
<b>1.1</b>	<b>Contextualização e Motivação</b>	<b>9</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>11</b>
<b>2.1</b>	<b>Aprendizado de Máquina</b>	<b>11</b>
2.1.1	K-Vizinhos Mais Próximos	11
2.1.2	Regressão Logística	12
2.1.3	Máquinas de Vetores de Suporte	13
2.1.4	Florestas Aleatórias e <i>Gradient Boosted Trees</i>	14
<b>2.2</b>	<b>Trabalhos de classificação na área médica</b>	<b>16</b>
<b>2.3</b>	<b>Técnicas de explicação aplicadas a trabalhos de classificação na área médica</b>	<b>18</b>
<b>3</b>	<b>DESENVOLVIMENTO</b>	<b>21</b>
<b>3.1</b>	<b>Metodologia</b>	<b>21</b>
3.1.1	Definição de domínio do problema	21
3.1.2	Dados	21
3.1.3	Ferramentas utilizadas	22
3.1.4	Limpeza, análise exploratória e preparação dos dados	22
3.1.5	Treinamento dos classificadores	23
3.1.6	Análise de variáveis mais importantes	23
3.1.7	Apresentação dos resultados	23
	<b>REFERÊNCIAS</b>	<b>25</b>





# 1 INTRODUÇÃO

## 1.1 Contextualização e Motivação

Nos últimos anos, houve um aumento dramático na quantidade de dados criados e compartilhados através da internet. Grande parte desse conteúdo está disponível publicamente e sem custo, representando um suprimento virtualmente ilimitado de dados de treinamento para várias tarefas de modelagem estatística e rotulagem. Essa "enxurrada de dados" apresenta desafios significativos, mas também oportunidades revolucionárias para o desenvolvimento de sistemas baseados em dados (SELTZER; ZHANG, 2009).

O crescente volume de dados gerados a partir dessas diversas fontes tem impulsionado uma popularização da inteligência artificial (IA), sobretudo com o uso de aprendizado de máquina, que, de forma sucinta, trata da criação de algoritmos que respondam e se adaptem automaticamente aos dados sem a necessidade de intervenção humana de forma contínua (FILHO, 2015). Aliado a ferramentas poderosas para o processamento massivo e paralelo de dados, esses avanços têm permitido tanto à indústria quanto à comunidade científica desenvolver modelos altamente eficazes e realizar análises mais sofisticadas. A adoção generalizada de técnicas de aprendizado de máquina tem se mostrado fundamental para desvendar o potencial dos dados disponíveis, resultando em descobertas inovadoras e aplicações transformadoras em diversas áreas, incluindo o setor de saúde (BATISTA; FILHO, 2019).

O número de estudos médicos de IA cresceu de forma exponencial no período de 2005 a 2019 (MESKO; GOROG, 2020), revelando o forte interesse da comunidade científica em buscar métodos cada vez mais eficazes para aprimorar o cuidado e a qualidade de vida dos pacientes. A aplicação de modelos preditivos de aprendizado de máquina possui um potencial significativo para auxiliar na tomada de decisões em diversas etapas do cuidado à saúde, principalmente no diagnóstico, intervenção e acompanhamento de problemas de saúde (OBERMEYER; LEE, 2017).

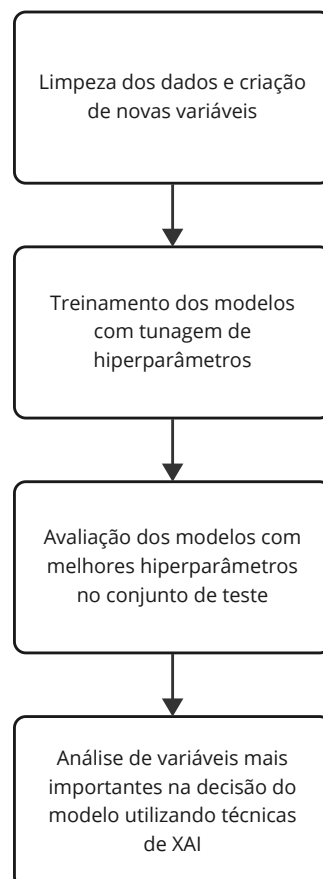
Na área da saúde, as decisões tomadas pelos sistemas de IA podem ter um impacto significativo na vida das pessoas. Se os usuários não conseguem compreender o processo de tomada de decisão, podem não confiar no sistema, chegando até a rejeitá-lo. Portanto, a capacidade de fornecer uma explicação sobre como e por que uma decisão específica foi feita tornou-se uma qualidade essencial em sistemas desse tipo (CONFALONIERI *et al.*, 2021). Além disso, a explicabilidade também pode auxiliar os desenvolvedores na identificação e correção de erros ou vies nos sistemas de IA, tornando-os mais justos e precisos.

Diante desse contexto, este trabalho se propõe a avaliar diversos modelos de clas-

sificação com o objetivo de prever a reinternação hospitalar de pacientes, utilizando um conjunto de dados de cuidados de saúde mental coletados pelo sistema de informação da Coordenação de Internações em Ribeirão Preto, Brasil, de julho de 2012 a dezembro de 2017. A relevância deste tipo de trabalho se dá, pois, por meio dessas previsões, medidas preventivas podem ser adotadas antecipadamente, evitando a necessidade de reinternação e, conseqüentemente, reduzindo os custos hospitalares, visto que o custo médio das internações é significativamente maior que o custo médio dos atendimentos ambulatoriais (CESCONETTO; LAPA; CALVO, 2008). Além da predição, o estudo também utilizará técnicas de inteligência artificial explicável (do inglês: *explainable artificial intelligence* - XAI) a fim de identificar quais variáveis estão mais associadas a casos de reinternação, fornecendo informações relevantes para o desenvolvimento de políticas e medidas preventivas mais eficazes.

A Figura 1 apresenta um fluxo geral deste trabalho:

Figura 1 – Fluxo geral do trabalho



Fonte: Autor (2023)

## 2 REVISÃO BIBLIOGRÁFICA

Neste capítulo, serão apresentadas as fundamentações teóricas das técnicas de aprendizado de máquina utilizadas neste trabalho, juntamente com um levantamento de estudos correlatos na área da saúde que empregaram técnicas semelhantes.

### 2.1 Aprendizado de Máquina

A crescente complexidade nos desafios computacionais e o volume massivo de dados gerados por diversas fontes impulsionaram a necessidade de ferramentas computacionais mais avançadas. Neste contexto, o aprendizado de máquina define-se como um campo de estudo da Inteligência Artificial que dá aos computadores a habilidade de aprender sem serem explicitamente programados. Para tal, no aprendizado de máquina, os computadores são treinados para aprender com dados passados, utilizando técnicas que os possibilitem a derivar uma função ou hipótese capaz de solucionar problemas a partir de observações específicas, oferecendo soluções baseadas em dados históricos (FACELI *et al.*, 2011).

Os sistemas de aprendizado de máquina podem seguir vários paradigmas de aprendizado. Existem sistemas supervisionados, não supervisionados, semi-supervisionados, auto-supervisionados e de reforço. O aprendizado supervisionado é usado quando o modelo pode ser treinado com exemplos rotulados, enquanto o aprendizado não supervisionado é usado quando não há rótulos disponíveis. O aprendizado semi-supervisionado é uma combinação dos dois, enquanto o auto-supervisionado é usado quando o modelo pode ser treinado com exemplos gerados automaticamente. O aprendizado por reforço é usado quando o modelo pode aprender a tomar decisões com base em recompensas e punições (GÉRON, 2022).

Dentro do paradigma de aprendizado supervisionado, os algoritmos podem ser usados em tarefas de classificação ou regressão. A classificação é usada para prever classes, enquanto a regressão é usada para prever valores. Por exemplo, a classificação pode ser usada para prever se um e-mail é spam ou não, enquanto a regressão pode ser usada para prever o preço de uma casa com base em suas características (GÉRON, 2022).

A seguir será apresentada uma breve introdução aos modelos de classificação utilizados no desenvolvimento deste projeto.

#### 2.1.1 K-Vizinhos Mais Próximos

O algoritmo K-Vizinhos Mais Próximos (ou KNN, do inglês *K-Nearest Neighbors*) é um método de classificação que opera com base na proximidade dos vizinhos mais próximos de um ponto de dados. Quando um novo ponto de dados precisa ser classificado, o algoritmo encontra os  $k$  pontos mais próximos a ele a partir do conjunto de dados

de treinamento. Em seguida, ele classifica o novo ponto de dados com base na classe mais frequente entre esses  $k$  vizinhos mais próximos. O algoritmo KNN assume que todas as observações correspondem a pontos em um espaço  $n$ -dimensional e não requer um processo de aprendizado. Ele apenas prevê a categoria do novo ponto de dados com base nas categorias dos pontos conhecidos no momento da classificação (WANG, 2019).

Os vizinhos mais próximos de uma instância são mensurados por métricas de distância, tais como a distância Euclidiana ou de Manhattan. A título de exemplo, será apresentada a distância Euclidiana, definida pela Equação 2.1:

$$d(x, x') = \sqrt{\sum_{i=1}^n (x_i - x'_i)^2} \quad (2.1)$$

### 2.1.2 Regressão Logística

A regressão logística é um modelo estatístico usado para modelar a probabilidade de um evento ocorrer em termos de variáveis independentes (KLEINBAUM; KLEIN, 2010). O modelo logístico é baseado na função logística, definida pela Equação 2.2. A Figura 2 apresenta um exemplo da curva logística.

$$f(z) = \frac{1}{1 + e^{-z}} \quad (2.2)$$

De forma simplificada, o modelo logístico pode ser definido da seguinte maneira:

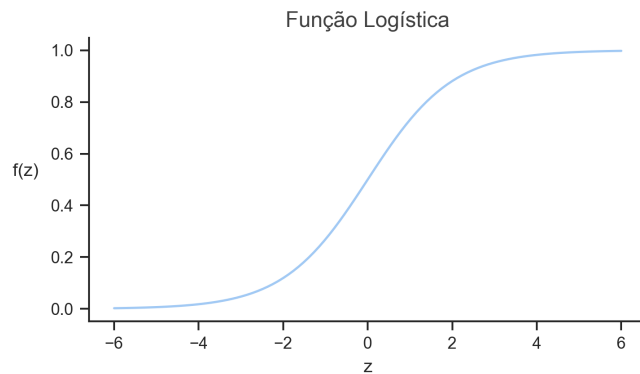
$$\hat{y} = f(Z) = P(y = 1|Z) = \frac{1}{1 + e^{-Z}} \quad (2.3)$$

onde  $P(y = 1|Z)$  é a probabilidade de ocorrência da classe positiva dado  $Z$ . Sendo  $Z$ :

$$Z = \ln\left(\frac{p}{1-p}\right) = \alpha + \sum_{j=1}^k \beta_j X_{jk} \quad (2.4)$$

onde  $p$  representa a probabilidade de ocorrência do evento de interesse, enquanto  $X$  denota o conjunto de variáveis preditoras. Os parâmetros do modelo,  $\alpha$  e  $\beta$ , são estimados por meio da técnica de máxima verossimilhança (FÁVERO *et al.*, 2009). Essa abordagem visa encontrar uma combinação de coeficientes que maximize a probabilidade de ocorrência do evento de interesse

Figura 2 – Exemplo de função logística



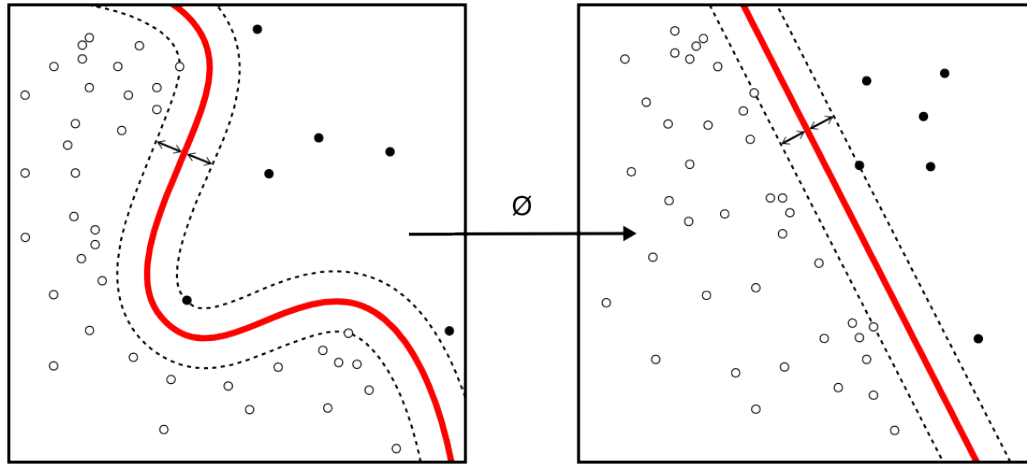
Fonte: Autor (2023)

### 2.1.3 Máquinas de Vetores de Suporte

As Máquinas de Vetores de Suporte (ou SVM, do inglês *Support Vector Machines*) são um método de aprendizado supervisionado que busca encontrar o hiperplano de separação que maximiza a margem entre as classes. Elas são usadas em problemas de classificação para encontrar a melhor fronteira de decisão entre as classes, permitindo sobreposição entre elas, mas minimizando essa sobreposição. As SVMs são particularmente úteis em problemas de classificação com muitas características, onde outras técnicas podem sofrer de sobreajuste, o que as torna robustas em relação a ruídos e *outliers* (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

As SVMs podem lidar com conjuntos de dados não lineares por meio do uso de *kernels*, que mapeiam os dados para um espaço de maior dimensão onde eles podem ser separados linearmente. A Figura 3 mostra um exemplo de um espaço de características que foi transformado.

Figura 3 – Exemplo de transformação utilizando *kernel*: Convertendo um problema de classificação não linear em um problema de classificação linear em uma dimensão superior.



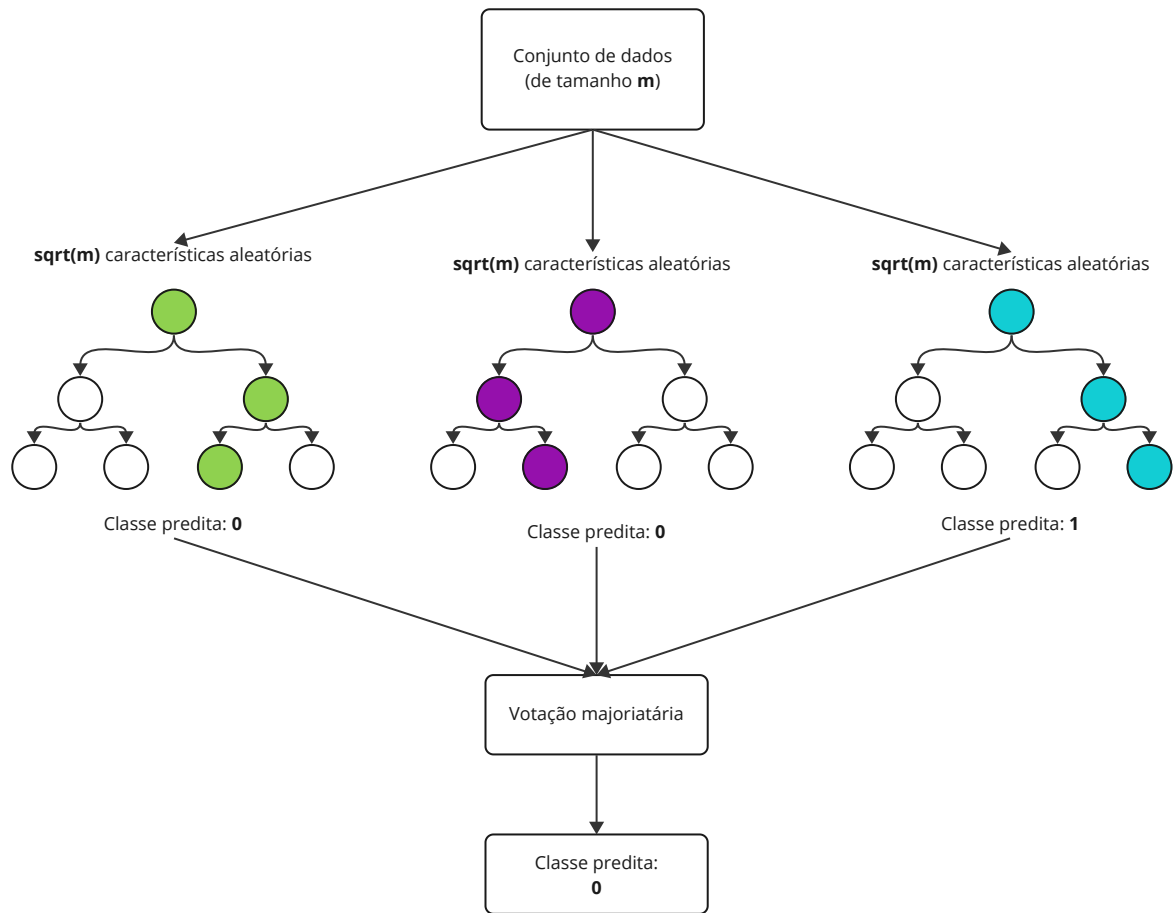
Fonte: (Wikipedia contributors, 2023)

Embora as SVMs sejam amplamente utilizadas e tenham várias vantagens, também apresentam algumas desvantagens. Uma das principais desvantagens das SVMs é a necessidade de ajuste cuidadoso dos parâmetros do modelo, como o parâmetro de regularização e a escolha do *kernel*, o que pode ser uma tarefa desafiadora e exigir conhecimento especializado. Além disso, as SVMs podem ser computacionalmente intensivas, especialmente em conjuntos de dados muito grandes, o que pode tornar seu treinamento demorado. Outra limitação das SVMs é a dificuldade em lidar com conjuntos de dados com muitas classes, uma vez que a abordagem de um-contra-um para problemas multiclasse pode levar a um grande número de classificadores binários (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

#### 2.1.4 Florestas Aleatórias e *Gradient Boosted Trees*

As Florestas Aleatórias (do inglês *Random Forests*) e *Gradient Boosted Trees* são exemplos de comitês (*ensembles*) de algoritmos de aprendizado de máquina. Resumidamente, o método consiste na construção de múltiplas árvores de decisão em subespaços aleatórios do espaço de características, utilizando amostras aleatórias dos dados de treinamento e subconjuntos aleatórios das características disponíveis. Cada árvore de decisão é capaz de generalizar sua classificação de forma única, e a combinação das previsões de todas as árvores resulta em uma melhoria monotônica na precisão da classificação (HO, 1995). Estudos recentes têm indicado que a aplicação desses algoritmos tem apresentado desempenho superior na previsão em comparação com outros métodos de aprendizado de máquina (GÉRON, 2022; RASCHKA; MIRJALILI, 2017). A Figura 4 ilustra a construção de três árvores de decisão para um problema de classificação, construídas pelo algoritmo *Random Forest*.

Figura 4 – Ilustração do algoritmo *Random Forest*.



Fonte: Autor (2023)

Com o objetivo de melhorar a precisão das previsões das árvores de decisão, o algoritmo *Gradient Boosted Trees* (FRIEDMAN, 2000) utiliza a estratégia de *boosting*. Em termos mais detalhados, o processo de *boosting* geralmente começa com um modelo base simples, como uma árvore de decisão rasa. Este modelo é treinado no conjunto de dados e usado para fazer previsões. Em seguida, os exemplos que foram classificados incorretamente ou para os quais as previsões tiveram grandes erros são ponderados mais fortemente no próximo modelo a ser treinado. Este próximo modelo é treinado para se concentrar nos exemplos que foram mal classificados pelo modelo anterior, e o processo é repetido (FRIEDMAN, 2000).

Ainda segundo Friedman (2000), a técnica de *Gradient Boosting* proposta pelo autor permitiu avanços importantes na capacidade de lidar com problemas complexos de previsão e classificação. Essa abordagem influenciou diretamente o desenvolvimento de algoritmos populares, como *XGBoost* e *LightGBM*, que são amplamente utilizados em competições de ciência de dados e em aplicações do mundo real devido à sua eficácia e desempenho.

## 2.2 Trabalhos de classificação na área médica

A utilização de algoritmos de aprendizado de máquina para análises preditivas é uma tarefa que envolve uma série de etapas que vão desde a preparação dos dados, até o lançamento e monitoramento da solução (GÉRON, 2022). O trabalho de Santos *et al.* (2019) se propõe a exemplificar as etapas envolvidas na utilização desses algoritmos para análises preditivas em saúde, com um exemplo de aplicação para prever óbito em idosos de São Paulo. A metodologia empregada envolveu a utilização de diferentes algoritmos de aprendizado de máquina para treinar modelos preditivos a partir de dados do estudo SABE, e a avaliação desses modelos em termos de acurácia, sensibilidade e especificidade. Os resultados mostraram que os modelos tiveram desempenho razoável na predição de óbito em idosos, com destaque para o algoritmo Random Forest. No entanto, os autores ressaltam que a amostra utilizada foi relativamente pequena e que a performance dos modelos pode ser melhorada com mais dados e refinamento dos algoritmos.

O trabalho de Loreto, Lisboa and Moreira (2020) aborda a predição de reinternações em UTI usando algoritmos de classificação. A proposta é investigar se as características basais e informações coletadas no momento da admissão do paciente podem possibilitar previsões precisas de reinternação na UTI. A metodologia empregada envolveu a análise de um conjunto de dados de 11.805 pacientes adultos de três UTIs em um hospital universitário brasileiro, com a exclusão de pacientes falecidos durante a primeira admissão na UTI. Foram testados oito algoritmos de classificação e avaliados com base em seis métricas. Os resultados mostraram que as características do paciente registradas na admissão na UTI foram capazes de prever o risco de reinternação, superando modelos anteriores que utilizavam apenas dados disponíveis na alta da UTI.

Wang *et al.* (2021) apresenta um estudo sobre a estimativa da probabilidade de reinternação à UTI para pacientes transferidos da UTI para a enfermaria geral, abordando os desafios de dados desbalanceados e esparsos. A pesquisa visa melhorar a gestão médica e reduzir as taxas de reinternação à UTI, levando a melhores resultados para os pacientes e redução de custos de saúde. A metodologia empregada inclui a análise de valores ausentes, o teste de razão de verossimilhança e a utilização de um modelo de floresta aleatória com decaimento de peso. Os resultados mostram que o modelo proposto supera outros métodos tradicionais em sete indicadores de desempenho diferentes.

O artigo de Stiglic *et al.* (2014) propõe uma abordagem para a classificação de reinternações pediátricas usando modelos de diferentes hospitais, permitindo alta performance e compreensibilidade dos resultados. A metodologia empregada é baseada em *deep learning* e *stacked generalization*, e foi avaliada usando dados de 54 hospitais na Califórnia para demonstrar as possibilidades de implantação em larga escala. Os resultados mostraram melhorias significativas no desempenho de classificação e interpretabilidade dos resultados.



O trabalho de Yin *et al.* (2023) tem como objetivo avaliar se a classificação de feridas do *Centers for Disease Control and Prevention* (CDC) pode ser usada como um indicador de risco para reinternação hospitalar em 30 dias após a cirurgia. A metodologia empregada foi uma revisão retrospectiva de pacientes submetidos a cirurgias eletivas em um hospital universitário. Os resultados mostraram que a classificação de feridas do CDC foi um preditor significativo de reinternação hospitalar em 30 dias após a cirurgia. Os autores concluem que a classificação de feridas do CDC pode ser uma ferramenta útil para identificar pacientes com maior risco de reinternação hospitalar após a cirurgia.

Fernandes *et al.* (2021) apresenta uma abordagem multipropósito de aprendizado de máquina para prever o prognóstico negativo da COVID-19 em São Paulo, Brasil. O trabalho envolveu o treinamento de cinco algoritmos de aprendizado de máquina com dados estruturados de pacientes com COVID-19, utilizando técnicas de validação cruzada e otimização bayesiana. O objetivo era prever o risco de desenvolver condições críticas em pacientes com COVID-19. Os resultados mostraram que a abordagem proposta teve um desempenho satisfatório na previsão do prognóstico negativo da COVID-19, o que pode ajudar na tomada de decisões clínicas e na alocação efetiva de recursos de saúde.

O trabalho de Asif *et al.* (2018) propõe uma metodologia baseada em aprendizado de máquina para identificar genes relacionados a doenças complexas. O modelo de aprendizado de máquina proposto utiliza uma matriz de similaridade funcional entre genes para prever doenças genéticas complexas. Essa matriz é construída a partir de medidas de similaridade entre genes, como perfis de expressão gênica, redes de interação proteína-proteína e Gene Ontology. Em seguida, um classificador de aprendizado de máquina é treinado com essa matriz para identificar genes relevantes para a doença em questão. O modelo proposto foi avaliado em relação à sua capacidade de identificar genes relacionados ao Transtorno do Espectro Autista e mostrou melhorias significativas em relação a outros métodos de identificação de genes. Foi desenvolvido um fluxo de trabalho automatizado para tornar a metodologia acessível a pesquisadores sem conhecimento extenso em programação e aprendizado de máquina.

Rosa *et al.* (2023) apresenta um estudo que utiliza técnicas de aprendizado de máquina para classificar fatores que influenciam a ocorrência de dermatites ocupacionais. A metodologia empregada envolveu a avaliação de diferentes algoritmos de aprendizado de máquina, a comparação dos resultados obtidos por cada técnica e a identificação dos fatores mais influentes na ocorrência da lesão ocupacional. Os resultados indicaram que todas as técnicas apresentaram acuracidade entre 55% e 69,4%, sensibilidade entre 49,1% e 80,7% e especificidade entre 50% e 66,7%. Os fatores mais influentes identificados foram a exposição a produtos químicos, o uso de equipamentos de proteção individual inadequados e a falta de treinamento adequado.

## 2.3 Técnicas de explicação aplicadas a trabalhos de classificação na área médica

A explicabilidade já foi identificada como um fator chave para a adoção de sistemas de IA numa vasta gama de contextos (DOSHI-VELEZ; KIM, 2017; LIPTON, 2017; RIBEIRO; SINGH; GUESTRIN, 2016). No estudo conduzido por Abououf *et al.* (2023) é apresentado um sistema de monitoramento de saúde que utiliza inteligência artificial para detectar e classificar eventos e anomalias médicas. A metodologia empregada envolve a utilização de um modelo de detecção de anomalias e eventos, um modelo de classificação e uma técnica de explicabilidade chamada *KernelSHAP*. O sistema foi avaliado em um conjunto de dados de pacientes e obteve resultados promissores na detecção e classificação de eventos e anomalias médicas. Além disso, a técnica de explicabilidade permitiu que os médicos entendessem melhor as decisões tomadas pelo sistema de inteligência artificial.

Com o uso de técnicas de aprendizado profundo e transferência de aprendizado para detectar a COVID-19 a partir de radiografias de tórax, o estudo de Brunese *et al.* (2020) utiliza as camadas de ativação da rede, ou seja, as áreas da radiografia de tórax que o modelo considerou para gerar a predição, para fornecer explicabilidade sobre a predição. Ainda segundo os autores, isso pode representar uma sugestão para o radiologista localizar imediatamente as áreas da radiografia de tórax que podem ser de interesse. Os resultados mostraram que o modelo proposto alcançou uma acurácia de 98,08% na detecção de COVID-19 em radiografias de tórax. O estudo também sugere que essa tecnologia pode ser usada como uma ferramenta de triagem para ajudar no diagnóstico da COVID-19 em configurações clínicas do mundo real.

Em (KIM; KIM, 2022) foi desenvolvida uma técnica de previsão de mortalidade relacionada ao calor em uma unidade espacial detalhada dentro de uma cidade, utilizando um modelo baseado em floresta aleatória e a técnica de explicação SHAP. A metodologia empregada incluiu a coleta de dados meteorológicos, demográficos e socioeconômicos, pré-processamento de dados, divisão em conjuntos de treinamento e teste, construção do modelo, otimização de hiperparâmetros, avaliação de desempenho e interpretação dos resultados. Os resultados mostraram que o modelo proposto é capaz de prever com precisão a mortalidade relacionada ao calor em uma área específica da cidade de Daegu, na Coreia do Sul. A técnica SHAP permitiu uma interpretação global e local dos resultados, identificando as variáveis mais importantes para a previsão.

Também utilizando modelos SHAP, o trabalho de GhoshRoy, Alvi and Santosh (2023) utiliza técnicas para prever a fertilidade masculina, com o intuito de melhorar a transparência, responsabilidade e explicabilidade de sete modelos de IA padrão da indústria. A metodologia empregada consistiu em coletar dados de pacientes com problemas de fertilidade masculina e aplicar os modelos de IA para prever a fertilidade. Em seguida, foram utilizadas técnicas de amostragem e validação cruzada para avaliar a robustez e estabilidade de cada modelo. Por fim, a técnica de XAI foi utilizada para explicar o desem-

penho de cada modelo e os resultados mostraram que o SHAP foi capaz de interpretar as previsões dos sistemas de IA e fornecer explicações claras e compreensíveis para os usuários.



### 3 DESENVOLVIMENTO

Neste capítulo, será apresentada uma exposição ordenada e detalhada do desenvolvimento do presente trabalho. O início deste capítulo se dá com uma explanação da metodologia empregada, delineando os passos e abordagens utilizados para alcançar os objetivos propostos.

#### 3.1 Metodologia

##### 3.1.1 Definição de domínio do problema

O problema abordado neste estudo refere-se à previsão de reinternação hospitalar de pacientes que receberam cuidados de saúde mental no sistema de informação da Coordenação de Internações em Ribeirão Preto, Brasil, no período de julho de 2012 a dezembro de 2017. Este domínio envolve a aplicação de modelos de classificação de aprendizado de máquina em um conjunto de dados específico, com o propósito de antecipar eventos de reinternação hospitalar. Além disso, o estudo incorpora técnicas de inteligência artificial explicável, notadamente o SHapley Additive exPlanations (SHAP), para identificar e quantificar as variáveis que apresentam maior influência nos casos de reinternação, contribuindo assim para uma compreensão mais profunda dos fatores associados a esses eventos. A análise deste domínio visa aprimorar a capacidade de tomada de decisões e políticas de saúde, com ênfase na prevenção eficaz e na otimização dos recursos hospitalares.

##### 3.1.2 Dados

O conjunto de dados empregado neste estudo já foi explorado em outras pesquisas (BARROS *et al.*, 2016; MIYOSHI *et al.*, 2018; KHARRAT *et al.*, 2020). Ele abrange informações de 8.755 pacientes, com uma média de idade de 37,6 anos. As características do conjunto de dados englobam aspectos sociodemográficos, dados sobre internações hospitalares, diagnósticos, utilização de serviços médicos, informações de alta hospitalar e registros temporais, como datas de admissão, alta e óbito. A Tabela 1 descreve as variáveis utilizadas neste estudo.

Tabela 1 – Variáveis utilizadas do conjunto de dados

Nome da variável	Tipo da variável
Arranjo domiciliar	Categórica
AVC	Booleana
Convulsão	Booleana
Dia da semana na 1ª internação	Numérica (inteira)
Diabetes	Booleana
Diagnóstico (CID10)	Categórica
Doença infecto	Booleana
Estado civil	Categórica
Etnia	Categórica
HAS	Booleana
Idade na 1ª internação	Numérica (inteira)
Mês da 1ª internação	Numérica (inteira)
Problemas respiratórios	Booleana
Quantidade de problemas na 1ª internação	Numérico (inteira)
Sexo	Categórica
Situação profissão	Categórica
Tempo de internação (em horas)	Numérica (contínua)
Traumatismo	Booleano

Fonte: Autor (2023)

É importante ressaltar que essas não são todas as variáveis disponíveis no conjunto de dados. Mais detalhes sobre como a seleção de atributos foi feita estão disponíveis na subseção 3.1.4.

### 3.1.3 Ferramentas utilizadas

É crescente o aumento da popularidade da linguagem Python em aplicações científicas. Boa parte deste aumento se dá pela diversidade de bibliotecas científicas como Numpy e Scipy. Na área de Machine Learning, a linguagem Python apresenta diversas bibliotecas que permitem a execução de modelos de ML, com destaque para a biblioteca Scikit-Learn (BUTINCK *et al.*, 2013)

### 3.1.4 Limpeza, análise exploratória e preparação dos dados

Inicialmente, a limpeza dos dados será realizada para garantir a consistência e integridade do conjunto de dados. Em seguida, uma análise exploratória detalhada dos

dados será conduzida, com a possibilidade de criação de novas variáveis que possam proporcionar insights valiosos. Um exemplo é a variável "tempo de internação", definida como a diferença entre a data de entrada e a data de saída, que poderá ser criada para uma compreensão mais precisa do tempo de permanência dos pacientes.

### 3.1.5 Treinamento dos classificadores

Posteriormente, a fase de treinamento dos classificadores será iniciada, explorando diferentes algoritmos de aprendizado de máquina, ajustando seus hiperparâmetros e empregando a validação cruzada para garantir que os modelos sejam robustos e generalizáveis.

### 3.1.6 Análise de variáveis mais importantes

Por fim, será utilizada a biblioteca `shap` para avaliar o impacto das variáveis no processo de tomada de decisões dos modelos. Isso permitirá a identificação de quais características estão mais fortemente associadas a casos de reinternação hospitalar, fornecendo uma visão detalhada das relações entre variáveis e resultados.

### 3.1.7 Apresentação dos resultados

Os resultados obtidos após o término dos processos mencionados na seção TAL serão apresentados de maneira clara e concisa, utilizando tabelas e gráficos informativos. As tabelas destacarão métricas de desempenho, como precisão, recall, F1-score e área sob a curva ROC, para cada modelo testado. Além disso, serão fornecidos gráficos que ilustrarão a importância relativa das variáveis no processo de previsão, com base nas análises SHAP.

[Incluir imagem com o fluxograma do processo descrito na metodologia]





## REFERÊNCIAS

- ABOUOUF, M. *et al.* Explainable ai for event and anomaly detection and classification in healthcare monitoring systems. **IEEE Internet of Things Journal**, p. 1–1, 2023.
- ASIF, M. *et al.* Identifying disease genes using machine learning and gene functional similarities, assessed through gene ontology. **PLOS ONE**, Public Library of Science, v. 13, n. 12, p. 1–15, 12 2018. Available at: <https://doi.org/10.1371/journal.pone.0208626>.
- BARROS, R. E. M. *et al.* Impact of length of stay for first psychiatric admissions on the ratio of readmissions in subsequent years in a large Brazilian catchment area. **Social Psychiatry and Psychiatric Epidemiology**, v. 51, n. 4, p. 575–587, abr. 2016. ISSN 1433-9285. Available at: <https://doi.org/10.1007/s00127-016-1175-x>.
- BATISTA, A. F. de M.; FILHO, A. D. P. C. Machine learning aplicado à saúde. Sociedade Brasileira de Computação, 2019. Available at: <https://sol.sbc.org.br/livros/index.php/sbc/catalog/download/29/95/245-1?inline=1>.
- BRUNESE, L. *et al.* Explainable deep learning for pulmonary disease and coronavirus covid-19 detection from x-rays. **Computer Methods and Programs in Biomedicine**, v. 196, p. 105608, 2020. ISSN 0169-2607. Available at: <https://www.sciencedirect.com/science/article/pii/S0169260720314413>.
- BUITINCK, L. *et al.* API design for machine learning software: experiences from the scikit-learn project. **CoRR**, abs/1309.0238, 2013. Available at: <http://arxiv.org/abs/1309.0238>.
- CESCONETTO, A.; LAPA, J. dos S.; CALVO, M. C. M. Avaliação da eficiência produtiva de hospitais do sus de santa catarina, brasil. **Revista Saúde Pública**, 2008.
- CONFALONIERI, R. *et al.* A historical perspective of explainable artificial intelligence. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, John Wiley & Sons Ltd., v. 11, n. 1, p. 21, 2021. ISSN 1942-4787.
- DOSHI-VELEZ, F.; KIM, B. **Towards A Rigorous Science of Interpretable Machine Learning**. 2017.
- FACELI, K. *et al.* **Inteligência artificial: uma abordagem de aprendizado de máquina**. [S.l.: s.n.]: LTC, 2011.
- FÁVERO, L. P. L. *et al.* **Análise de dados: modelagem multivariada para tomada de decisões**. [S.l.: s.n.]: Elsevier, 2009.
- FERNANDES, F. T. *et al.* A multipurpose machine learning approach to predict covid-19 negative prognosis in são paulo, brazil. **Scientific reports**, Nature Publishing Group, v. 11, n. 1, p. 1–11, 2021.
- FILHO, A. D. P. gatto. Uso de big data em saúde no brasil: perspectivas para um futuro próximo. **Epidemiologia e Serviços de Saúde**, SciELO Public Health, v. 24, p. 325–332, 2015.
- FRIEDMAN, J. Greedy function approximation: A gradient boosting machine. **The Annals of Statistics**, v. 29, 11 2000.

GÉRON, A. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 3rd Edition**. [*S.l.: s.n.*]: O'Reilly Media, Inc., 2022. ISBN 9781098125974.

GHOSHROY, D.; ALVI, P. A.; SANTOSH, K. Unboxing industry-standard ai models for male fertility prediction with shap. **Healthcare**, v. 11, n. 7, 2023. ISSN 2227-9032. Available at: <https://www.mdpi.com/2227-9032/11/7/929>.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning**. 2. ed. [*S.l.: s.n.*]: Springer New York Inc., 2009. ISBN 9780387848587.

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.

KHARRAT, F. G. Z. *et al.* Feature sensitivity criterion-based sampling strategy from the optimization based on phylogram analysis (fs-opa) and cox regression applied to mental disorder datasets. **PLOS ONE**, Public Library of Science, v. 15, n. 7, p. 1–25, 07 2020. Available at: <https://doi.org/10.1371/journal.pone.0235147>.

KIM, Y.; KIM, Y. Explainable heat-related mortality with random forest and shapley additive explanations (shap) models. **Sustainable Cities and Society**, v. 79, p. 103677, 2022. ISSN 2210-6707. Available at: <https://www.sciencedirect.com/science/article/pii/S2210670722000117>.

KLEINBAUM, D. G.; KLEIN, M. **Logistic Regression**. 3. ed. [*S.l.: s.n.*]: Springer New York, NY, 2010. ISBN 9781441917416.

LIPTON, Z. C. **The Mythos of Model Interpretability**. 2017.

LORETO, M.; LISBOA, T.; MOREIRA, V. P. Early prediction of icu readmissions using classification algorithms. **Computers in Biology and Medicine**, v. 118, p. 103636, 2020. ISSN 0010-4825. Available at: <https://www.sciencedirect.com/science/article/pii/S0010482520300329>.

MESKO, B.; GOROG, M. A short guide for medical professionals in the era of artificial intelligence. **npj Digital Medicine**, Nature Publishing Group, v. 3, p. 126, 2020.

MIYOSHI, N. S. B. *et al.* An ehealth platform for the support of a brazilian regional network of mental health care (ehealth-interop): development of an interoperability platform for mental care integration. **JMIR Mental Health**, 2018.

OBERMEYER, Z.; LEE, T. H. Lost in thoughtthe limits of the human mind and the future of medicine. **New England Journal of Medicine**, Mass Medical Soc, v. 377, n. 13, p. 1209–1211, 2017.

RASCHKA, S.; MIRJALILI, V. **Python Machine Learning**. 2. ed. [*S.l.: s.n.*]: Packt, 2017. ISBN 9781787125933.

RIBEIRO, M. T.; SINGH, S.; GUESTIN, C. **Model-Agnostic Interpretability of Machine Learning**. 2016.

ROSA, A. C. F. d. *et al.* Uso de técnicas de aprendizado de máquina para classificação de fatores que influenciam a ocorrência de dermatites ocupacionais. **Revista Brasileira de Saúde Ocupacional**, Fundação Jorge Duprat Figueiredo de Segurança e Medicina do Trabalho - FUNDACENTRO, v. 48, 2023. ISSN 0303-7657. Available at: <https://doi.org/10.1590/2317-6369/31620pt2023v48e4>.

SANTOS, H. G. d. *et al.* Machine learning para análises preditivas em saúde: exemplo de aplicação para prever óbito em idosos de São Paulo, Brasil. **Cadernos de Saúde Pública**, Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz, v. 35, n. 7, 2019. ISSN 0102-311X.

SELTZER, M. L.; ZHANG, L. The data deluge: Challenges and opportunities of unlimited data in statistical signal processing. *IEEE*, p. 3701–3704, 2009. Available at: <https://doi.org/10.1109/ICASSP.2009.4960430>.

STIGLIC, G. *et al.* Pediatric readmission classification using stacked regularized logistic regression models. **AMIA Annu Symp Proc**, p. 1072–1081, 2014.

WANG, B. *et al.* Predictive classification of ICU readmission using weight decay random forest. **Future Generation Computer Systems**, v. 124, p. 351–360, 2021. ISSN 0167-739X. Available at: <https://www.sciencedirect.com/science/article/pii/S0167739X21002065>.

WANG, L. Research and implementation of machine learning classifier based on KNN. **IOP Conference Series: Materials Science and Engineering**, IOP Publishing, v. 677, n. 5, p. 052038, dec 2019. Available at: <https://dx.doi.org/10.1088/1757-899X/677/5/052038>.

Wikipedia contributors. **Support vector machine — Wikipedia, The Free Encyclopedia**. 2023. [Online; accessed 6-January-2024]. Available at: [https://en.wikipedia.org/w/index.php?title=Support\\_vector\\_machine&oldid=1190739318](https://en.wikipedia.org/w/index.php?title=Support_vector_machine&oldid=1190739318).

YIN, V. *et al.* Centers for disease control (CDC) wound classification is prognostic of 30-day readmission following surgery. **World Journal of Surgery**, v. 47, n. 10, p. 2392–2400, 2023. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1007/s00268-023-07093-3>.