

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
INSTITUTO DE CIÊNCIAS EXATAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

Análise das Tendências de Suicídio no Brasil: Características Sociodemográficas e o Efeito da Pandemia de COVID-19

Luciana Nascimento Santana Prachedes

JUIZ DE FORA
SETEMBRO, 2024

Análise das Tendências de Suicídio no Brasil: Características Sociodemográficas e o Efeito da Pandemia de COVID-19

LUCIANA NASCIMENTO SANTANA PRACHEDES

Universidade Federal de Juiz de Fora
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Bacharelado em Ciência da Computação

Orientador: Heder Soares Bernardino
Coorientador: Leonardo Fernandes Martins

JUIZ DE FORA
SETEMBRO, 2024

ANÁLISE DAS TENDÊNCIAS DE SUICÍDIO NO BRASIL:
CARACTERÍSTICAS SOCIODEMOGRÁFICAS E O EFEITO DA
PANDEMIA DE COVID-19

Luciana Nascimento Santana Prachedes

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS
EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA, COMO PARTE INTE-
GRANTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE
BACHAREL EM CIÊNCIA DA COMPUTAÇÃO.

Aprovada por:

Heder Soares Bernardino
Doutor em Modelagem Computacional

Leonardo Fernandes Martins
Doutor em Psicologia

Jairo Francisco de Souza
Doutor em Informática

Luciana Conceição Dias Campos
Doutora em Engenharia Elétrica

JUIZ DE FORA
20 DE SETEMBRO, 2024

Aos meus amigos e irmãos.

À minha mãe, pelo apoio e sustento.

Resumo

O suicídio é uma das principais causas de morte no mundo todo. Atrelado às condições de saúde mental de uma população, eventos globais estressantes podem afetar as tendências. Este trabalho analisou as tendências de suicídio no Brasil entre 2003 e 2022, com foco no período da pandemia de COVID-19. A análise utilizou dados do Sistema de Informação sobre Mortalidade (SIM) do DATASUS, acessados através da plataforma da Fiocruz, PC-DaS, totalizando 218.707 ocorrências de suicídio. Inicialmente, para identificar padrões de suicídio entre diferentes grupos sociodemográficos, técnicas de análise estatística descritiva foram aplicadas. Para prever as taxas de mortalidade por suicídio, avaliando o impacto da pandemia, modelos de séries temporais (Prophet, SARIMA e LSTM) foram utilizados. Para compreender os fatores associados à escolha dos métodos utilizados no momento do suicídio, algoritmos de aprendizado de máquina (XGBoost e Random Forest) foram empregados, e variáveis como sexo, idade e local de ocorrência se destacaram. Os resultados indicam um aumento gradual nas taxas de suicídio, especialmente entre homens, pessoas de 30 a 49 anos e pessoas brancas, com um incremento mais acentuado durante os anos da pandemia. A análise destacou a importância de políticas públicas voltadas para saúde mental e intervenções direcionadas aos grupos vulneráveis identificados. O estudo reforça a necessidade de ações preventivas, adaptadas a diferentes grupos demográficos, especialmente no contexto de crises como a COVID-19.

Palavras-chave: Suicídio, COVID-19, Aprendizado de Máquina, Séries Temporais.

Abstract

Suicide is one of the leading causes of death worldwide. Linked to the mental health conditions of a population, global stressful events can affect trends. This study analyzed suicide trends in Brazil between 2003 and 2022, focusing on the COVID-19 pandemic period. The analysis used data from the Mortality Information System (SIM) of DATASUS, accessed through Fiocruz's PCDaS platform, totaling 218,707 suicide occurrences. Initially, descriptive statistical analysis techniques were applied to identify suicide patterns among different sociodemographic groups. To predict suicide mortality rates, evaluating the impact of the pandemic, time series models (Prophet, SARIMA, and LSTM) were used. To understand the factors associated with the choice of methods used in suicide, machine learning algorithms (XGBoost and Random Forest) were employed, with variables such as gender, age, and location of occurrence standing out. The results indicate a gradual increase in suicide rates, especially among men, people aged 30 to 49 years, and white people, with a sharper increase during the pandemic years. The analysis highlighted the importance of public policies aimed at mental health and interventions targeted at the vulnerable groups identified. The study reinforces the need for preventive actions, adapted to different demographic groups, especially in the context of crises such as COVID-19.

Keywords: Suicide, COVID-19, Machine Learning, Time Series

Agradecimentos

À minha família, pelo encorajamento e apoio.

Ao professor Heder pela orientação, amizade e principalmente, pela paciência, sem a qual este trabalho não se realizaria.

À Fiocruz, por ceder o acesso aos dados tratados sem os quais este trabalho não seria desenvolvido.

Ao CNPQ, pelo incentivo à pesquisa e desenvolvimento científico.

Aos professores do Departamento de Ciência da Computação pelos seus ensinamentos e aos funcionários do curso, que durante esses anos, contribuíram de algum modo para o nosso enriquecimento pessoal e profissional.

“Eu fico achando que a vida é um quadro: e que a gente tem que ir pintando como se fosse um pintor. Se você não gosta de uma cor que você botou, ou de um risco que você fez, vai lá e pinta outra vez”.

Lygia Bojunga

Conteúdo

Lista de Figuras	7
Lista de Tabelas	8
Lista de Abreviações	9
1 Introdução	10
1.1 Apresentação do Tema	10
1.2 Motivação	10
1.3 Descrição do Problema	11
1.4 Objetivos	11
2 Fundamentação Teórica	13
2.1 Conceituação do Suicídio	13
2.2 Aprendizado de Máquina	14
2.2.1 Random Forest	15
2.2.2 XGBoost	17
2.2.3 Importância de Características	18
2.2.4 Métricas e Divisão dos Dados	20
2.3 Modelagem de Séries Temporais	21
2.3.1 Prophet	23
2.3.2 SARIMA	24
2.3.3 LSTM	25
2.3.4 Métricas	27
3 Trabalhos Relacionados	30
4 Metodologia	33
4.1 Fonte dos Dados	33
4.2 Código Fonte	34
4.3 Análise Estatística Descritiva	35
4.4 Modelagem de Séries Temporais	35
4.4.1 Características Específicas dos Modelos	37
4.5 Modelagem com Classificação	38
4.5.1 Características Específicas dos Modelos	38
5 Resultados e Discussão	40
5.1 Análise Descritiva	40
5.2 Análise de Séries Temporais	46
5.3 Modelagem de Classificação Binária	51
6 Conclusão e Trabalhos Futuros	54
Bibliografia	56

Lista de Figuras

5.1	Taxa de Suicídio no Brasil	40
5.2	Evolução da Taxa de Suicídio por Faixa Etária	42
5.3	Evolução da Taxa de Suicídio por Sexo	42
5.4	Evolução da Taxa de Suicídio por Raça	43
5.5	Evolução da Taxa de Suicídio por Tempo de Estudo	44
5.6	Evolução da Taxa de Suicídio por UF	44
5.7	Evolução da Taxa de Suicídio por Município	45
5.8	Evolução da Taxa de Suicídio por Causa Básica	45
5.9	Evolução da Taxa de Suicídio por Sexo e Causa Básica	46
5.10	Componentes do Modelo Prophet	47
5.11	Comparação das Previsões SARIMA vs Prophet vs LSTM	50
5.12	Curva ROC - Random Forest vs XGBoost	53

Lista de Tabelas

4.1	Descrição das variáveis selecionadas no banco de dados	34
5.1	Comparação de Métricas de Erro e Estatísticas	48
5.2	Comparação da Importância das Variáveis com Desvio Padrão	52
5.3	Comparação de Desempenho com Desvio Padrão	52

Lista de Abreviações

PCDaS	Plataforma de Ciência de Dados Aplicada à Saúde
DATASUS	Departamento de Informática do Sistema Único de Saúde
SIM	Sistema de Informação sobre Mortalidade
CID-10	Classificação Internacional de Doenças, décima revisão
IBGE	Instituto Brasileiro de Geografia e Estatística
SARIMA	Seasonal Autoregressive Integrated Moving Average
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
MAPE	Mean Absolute Percentage Error
ROC-AUC	Receiver Operating Characteristic - Area Under Curve
XGBoost	Extreme Gradient Boosting
UF	Unidade Federativa

1 Introdução

1.1 Apresentação do Tema

O suicídio é reconhecido globalmente como um grave problema de saúde pública que afeta indivíduos, famílias e comunidades inteiras. De acordo com a Organização Mundial da Saúde (OMS), aproximadamente 700 mil pessoas morreram por suicídio em 2019, representando uma das principais causas de morte em todo o mundo (WHO, 2021). No Brasil, as taxas de suicídio têm apresentado um aumento gradual na última década, evidenciando a necessidade de uma compreensão mais aprofundada desse fenômeno no contexto nacional (SOARES; STAHNKE; LEVANDOWSKI, 2023).

Diversos fatores estão associados ao suicídio, incluindo aspectos biológicos, psicológicos, sociais e culturais (LOPES et al., 2023). Fatores demográficos como sexo, idade, raça/cor e nível de escolaridade desempenham papéis importantes na determinação dos grupos mais vulneráveis (BRAUN et al., 2023). Além disso, eventos de grande impacto social e econômico, como a pandemia de COVID-19, podem exacerbar condições de risco e influenciar as tendências de suicídio em escala nacional (JUNIOR et al., 2023).

1.2 Motivação

A motivação para este estudo está enraizada na necessidade de explorar o potencial das técnicas avançadas de análise de dados, como aprendizado de máquina e mineração de dados, na compreensão de questões delicadas como o suicídio. Essas ferramentas têm sido amplamente utilizadas na literatura para examinar grandes volumes de dados e identificar padrões ocultos, oferecendo visões sobre relações entre variáveis que não seriam facilmente detectadas por análises tradicionais (SHARIFANI; AMINI, 2023; REEL et al., 2021; NURMI; LOHAN, 2021). Em um contexto como o do suicídio, a aplicação de modelos preditivos e algoritmos de aprendizado de máquina permite analisar grandes bases de dados, identificar fatores determinantes e propor intervenções personalizadas para

diferentes grupos sociodemográficos.

Essa abordagem é particularmente relevante no cenário brasileiro, onde há escassez de análises aprofundadas sobre os efeitos da pandemia de COVID-19 na saúde mental e nas taxas de suicídio. Com a pandemia elevando os riscos à saúde mental, é preciso investigar a maneira como variáveis sociodemográficas e econômicas influenciam as tendências suicidas. A utilização dessas técnicas não só permite preencher essa lacuna, mas também oferece a oportunidade de informar políticas públicas mais eficazes, além de apoiar a criação de campanhas preventivas direcionadas e baseadas em dados robustos.

1.3 Descrição do Problema

A pandemia de COVID-19, declarada em março de 2020, trouxe desafios únicos para a saúde pública global. Além das consequências diretas à saúde física, as medidas de contenção do vírus, como o isolamento social, a interrupção de atividades econômicas e educacionais, e o aumento do desemprego, tiveram impacto na saúde mental da população (BOSI; ALVES, 2023). Esse cenário pode ter contribuído para alterações nas taxas de suicídio, tornando essencial a investigação de possíveis mudanças nas tendências durante esse período. No Brasil, a falta de estudos abrangentes que analisem o impacto da pandemia nas taxas de suicídio dificulta a compreensão completa do problema. Além disso, há uma necessidade de explorar como diferentes fatores sociodemográficos estão associados não apenas à ocorrência de suicídio, mas também aos métodos utilizados. Essa compreensão pode auxiliar na implementação de estratégias de prevenção mais eficazes e direcionadas.

1.4 Objetivos

Como objetivo geral deste trabalho, coloca-se a análise de dados de mortalidade no Brasil, especificamente no contexto de suicídio, a fim de apoiar o desenvolvimento de políticas públicas. Para atingir o objetivo principal, propõe-se, especificamente:

- Realizar uma análise estatística descritiva dos dados de suicídio, explorando padrões relacionados a sexo, idade, raça/cor, escolaridade e outras variáveis relevantes.

-
- Analisar as tendências mensais de suicídio de 2003 até 2018 utilizando modelos de séries temporais, validá-los com dados de 2019 e prever as tendências para os anos de 2020 a 2022.
 - Avaliar o impacto da pandemia de COVID-19 nas tendências de suicídio, comparando as previsões dos modelos com os dados reais observados durante o período pandêmico.
 - Identificar os fatores mais influentes na determinação dos métodos empregados de suicídio, por meio da análise de importância de características em modelos de classificação.
 - Analisar as soluções dos diferentes modelos estatísticos e de aprendizado de máquina empregados, destacando suas vantagens e limitações quando resolvendo os problemas abordados aqui.

2 Fundamentação Teórica

A compreensão aprofundada dos fatores que influenciam o comportamento suicida é essencial para o desenvolvimento de estratégias eficazes de prevenção e intervenção. Este capítulo aborda os conceitos base relacionados ao suicídio, explorando suas dimensões psicológicas, socioculturais e epidemiológicas. Além disso, são discutidas as principais teorias que buscam explicar o suicídio. Por fim, apresentam-se os métodos estatísticos e de aprendizado de máquina utilizados na análise de séries temporais e na classificação de dados de mortalidade por suicídio, fundamentando as abordagens usadas neste trabalho.

2.1 Conceituação do Suicídio

O suicídio é definido como um ato deliberado iniciado e realizado pelo indivíduo, cuja intenção consciente é a morte, independentemente de ser ou não compreendida pelas pessoas ao seu redor (SHER; OQUENDO, 2023). Trata-se de um fenômeno complexo que resulta de uma interação de fatores biológicos, psicológicos, sociais, culturais e ambientais (FORTGANG; NOCK, 2021). A compreensão desses fatores é essencial para o desenvolvimento de estratégias eficazes de prevenção e intervenção.

Do ponto de vista psicológico, transtornos mentais como depressão, transtorno bipolar, esquizofrenia e abuso de substâncias estão fortemente associados ao aumento do risco de suicídio (ALI; REHNA, 2022). Além disso, fatores como desesperança, impulsividade, baixa autoestima e sentimentos de isolamento social desempenham papéis decisivos na vulnerabilidade individual (DOUPNIK et al., 2020).

Culturalmente, o estigma em torno da saúde mental e do suicídio pode impedir indivíduos em risco de buscar ajuda (OEXLE; MAYER; RÜSCH, 2020). Normas sociais que desencorajam a expressão de sofrimento emocional ou que promovem expectativas irrealistas de sucesso podem contribuir para o aumento do risco (EVANS; ABRAHAMSON, 2020). Questões socioeconômicas, como desemprego, pobreza e desigualdades sociais, também estão correlacionadas com maiores taxas de suicídio (SINYOR et al., 2024).

Globalmente, o suicídio é uma das principais causas de morte entre jovens, com mais óbitos devido ao suicídio do que por malária, HIV/AIDS, câncer de mama, guerra e homicídios (WHO, 2021). As taxas variam significativamente entre países e regiões, influenciadas por fatores como desenvolvimento econômico, conflitos sociais, acesso a meios letais e disponibilidade de serviços de saúde mental (ARENSMAN et al., 2020).

A Teoria Interpessoal do Suicídio sugere que o suicídio ocorre quando um indivíduo possui um desejo de morrer, resultado da percepção de ser um fardo para os outros e de não pertencer socialmente, aliado à capacidade adquirida de causar autolesão (ORDEN et al., 2010). O modelo de Estresse e Vulnerabilidade enfatiza a interação entre fatores estressantes externos e vulnerabilidades internas, como predisposições genéticas ou psicológicas (BOAVENTURA et al., 2023). Eventos estressantes marcantes podem desencadear comportamentos suicidas em indivíduos vulneráveis.

No Brasil, as taxas de suicídio têm mostrado um aumento alarmante na últimas décadas. Fatores como desigualdade social, violência, acesso limitado a serviços de saúde mental e questões culturais específicas contribuem para o cenário nacional (BARBOSA; TEIXEIRA, 2021). Regiões com menores índices de desenvolvimento humano tendem a apresentar taxas mais elevadas, indicando a influência dos determinantes sociais na saúde (REIS et al., 2023).

A compreensão dos fatores que influenciam a escolha dos métodos de suicídio é necessária para o desenvolvimento de estratégias de prevenção. Estudos indicam que a disponibilidade de meios letais podem influenciar essa escolha (CANTOR; BAUME, 1998). Medidas que restringem o acesso a certos meios, como armas de fogo ou pesticidas, têm demonstrado eficácia na redução das taxas de suicídio. Fatores como idade, sexo, raça/cor e nível de escolaridade podem influenciar na preferência por determinados métodos. Por exemplo, estudos mostram que homens tendem a utilizar métodos mais letais em comparação com mulheres (DENNING et al., 2000).

2.2 Aprendizado de Máquina

O aprendizado de máquina é um ramo da inteligência artificial que se dedica ao desenvolvimento de algoritmos capazes de aprender a partir de dados, permitindo que façam pre-

visões ou tomem decisões sem serem explicitamente programados para tarefas específicas (ZHOU, 2021). Esses algoritmos constroem modelos matemáticos baseados em dados de treinamento, o que possibilita a realização de previsões ou decisões sem depender de regras fixas (NAQA; MURPHY, 2015). Os problemas de aprendizado de máquina podem ser classificados em três tipos principais, dependendo da natureza dos dados de entrada e da tarefa a ser realizada.

No aprendizado supervisionado, o modelo é treinado em um conjunto de dados rotulados, onde cada exemplo de treinamento está associado a uma resposta desejada (CUNNINGHAM; CORD; DELANY, 2008). O objetivo é aprender uma função que mapeia entradas para saídas. Dentro desse tipo de aprendizado, existem tarefas de classificação, quando a saída é uma categoria discreta — como classificar e-mails como spam ou não spam ou reconhecer dígitos escritos à mão — e tarefas de regressão, quando a saída é um valor contínuo, como prever preços de casas ou estimar consumo de energia (HASTIE et al., 2009a).

No aprendizado não supervisionado, o modelo é treinado em dados não rotulados, buscando encontrar estruturas ou padrões ocultos nesses dados (HASTIE et al., 2009b). As principais tarefas incluem o agrupamento, que envolve segmentar dados em grupos baseados na similaridade dos dados, e a redução de dimensionalidade, que consiste em simplificar os dados ao reduzir o número de variáveis, mas preservando a informação essencial (CELEBI; AYDIN, 2016).

O aprendizado por reforço envolve treinar um agente para tomar decisões sequenciais, recebendo recompensas ou penalidades. Nesse cenário, o agente aprende a otimizar uma função de recompensa ao interagir com um ambiente dinâmico, ajustando suas ações para maximizar os ganhos ao longo do tempo (WIERING; OTTERLO, 2012).

2.2.1 Random Forest

O *Random Forest* é um algoritmo de aprendizado de máquina do tipo *ensemble* que utiliza múltiplas árvores de decisão para melhorar a precisão das previsões e controlar o *overfitting* (BIAU; SCORNET, 2016). Este método funciona construindo uma “floresta” de árvores de decisão durante a fase de treinamento e produzindo a classificação final (no

caso de problemas de classificação) ou a média das previsões (no caso de regressão) de todas as árvores individuais (LIU; WANG; ZHANG, 2012).

O funcionamento do *Random Forest* baseia-se na ideia de que uma combinação de modelos fracos pode formar um modelo forte. Cada árvore de decisão na floresta é construída a partir de uma amostra aleatória com reposição (*bootstrap*) do conjunto de dados original (OSHIRO; PEREZ; BARANAUSKAS, 2012). Além disso, durante a construção de cada nó da árvore, o algoritmo considera apenas um subconjunto aleatório de features, selecionando aquela que melhor divide os dados de acordo com uma determinada medida de impureza, como a entropia ou o índice Gini (GABER; ATWAL, 2013).

Essa introdução de aleatoriedade, tanto na seleção dos dados quanto na escolha das features, tem como objetivo reduzir a correlação entre as árvores individuais, aumentando a diversidade do conjunto (GREGORUTTI; MICHEL; SAINT-PIERRE, 2017). Como resultado, o modelo combinado geralmente apresenta um desempenho melhor do que qualquer uma das árvores individuais, especialmente em termos de capacidade de generalização para dados não vistos (PRAJWALA, 2015).

Uma das vantagens significativas do *Random Forest* é sua robustez a *overfitting*, que é um problema comum em árvores de decisão tradicionais (BRAMER, 2007). Como o modelo considera a média ou o voto majoritário das previsões individuais, os erros de *overfitting* tendem a ser cancelados. Além disso, o algoritmo é eficaz em lidar com dados com alta dimensionalidade e pode fornecer estimativas de importância das características, auxiliando na compreensão dos fatores mais influentes nas previsões (YUAN et al., 2023).

O *Random Forest* também calcula uma estimativa interna do erro de generalização conhecida como erro *Out-of-Bag* (OOB) (MATTHEW et al., 2011). Como cada árvore é treinada com uma amostra *bootstrap* do conjunto de dados, aproximadamente um terço dos dados não é utilizado no treinamento de uma determinada árvore. Esses dados excluídos podem ser usados para testar a árvore, proporcionando uma estimativa do erro sem a necessidade de um conjunto de validação separado (RAMOSAJ; PAULY, 2019).

Apesar de suas muitas vantagens, o *Random Forest* não está isento de limitações. O modelo pode ser computacionalmente intensivo, especialmente quando se lida com

grandes conjuntos de dados e um número elevado de árvores, o que pode aumentar o tempo de treinamento e os requisitos de memória (MISHINA et al., 2015). Além disso, a interpretabilidade do modelo pode ser comprometida devido à complexidade introduzida pela combinação de múltiplas árvores (ROBINSON et al., 2017).

Em aplicações práticas, o *Random Forest* tem sido amplamente utilizado em diversas áreas, como detecção de fraudes, diagnóstico médico, análise financeira e reconhecimento de padrões, devido à sua capacidade de lidar com dados complexos e produzir resultados precisos (XUAN et al., 2018; SHI; CHARLES; ZHU, 2024; CAI et al., 2023; DAI et al., 2018). Sua facilidade de uso e flexibilidade o tornam uma escolha popular entre profissionais de ciência de dados e acadêmicos que buscam um equilíbrio entre desempenho e simplicidade na construção de modelos preditivos.

2.2.2 XGBoost

O XGBoost (*Extreme Gradient Boosting*) é um algoritmo de aprendizado de máquina baseado em árvores de decisão que tem se destacado pela alta eficiência computacional e excelente desempenho preditivo (NIELSEN, 2016). O XGBoost aprimora o método tradicional de boosting gradiente ao incorporar otimizações avançadas e técnicas de regularização, tornando-o especialmente eficaz em problemas de aprendizado supervisionado (CHEN et al., 2019).

Fundamentalmente, o XGBoost constrói modelos através da adição sequencial de árvores de decisão que visam minimizar uma função de perda diferenciável (CHEN; GUESTRIN, 2016). Em cada iteração, o algoritmo ajusta uma nova árvore aos resíduos das previsões anteriores, permitindo que o modelo capture padrões complexos e não lineares nos dados. Essa abordagem iterativa resulta em uma convergência rápida e em uma redução significativa dos erros de previsão (BROWNLEE, 2016).

Uma das características distintivas do XGBoost é sua implementação otimizada, que inclui suporte para paralelização durante o treinamento e técnicas de processamento em bloco (NALLURI; PENTELA; ELURI, 2020). Essas otimizações permitem que o algoritmo lide eficientemente com grandes volumes de dados e aproveite os recursos computacionais modernos. O algoritmo também incorpora regularização L1 e L2 para controlar

a complexidade do modelo e prevenir o *overfitting* (SHENG; YU, 2022). A regularização L1 promove a esparsidade nas features, facilitando a seleção de atributos relevantes, enquanto a regularização L2 reduz a magnitude dos pesos, contribuindo para a estabilidade do modelo (CHEN; GUESTRIN, 2016). Essas técnicas de regularização são necessárias para melhorar a capacidade de generalização do modelo em dados não vistos.

Outra vantagem significativa do XGBoost é sua capacidade de lidar com dados esparsos e valores ausentes. O algoritmo trata automaticamente os valores faltantes durante o processo de construção das árvores, determinando de forma otimizada a direção das ramificações quando confrontado com dados incompletos (RUSDAH; MURFI, 2020). Essa funcionalidade é particularmente útil em aplicações práticas, onde conjuntos de dados frequentemente apresentam imperfeições e inconsistências.

O XGBoost tem demonstrado desempenho superior em uma ampla variedade de domínios, incluindo detecção de fraudes, previsão de risco de crédito, e análise de expressão genética (HAJEK; ABEDIN; SIVARAJAH, 2023; WANG et al., 2022; LI et al., 2019). Sua eficiência e alta capacidade preditiva o tornaram uma escolha recorrente em competições de aprendizado de máquina e em projetos industriais que requerem modelos robustos e escaláveis.

No entanto, para maximizar o potencial do XGBoost, é importante um ajuste cuidadoso de seus hiperparâmetros. Parâmetros como a taxa de aprendizado, o número de estimadores e a profundidade máxima das árvores precisam ser otimizados para equilibrar o viés e a variância do modelo (ZIVKOVIC et al., 2022). Técnicas como validação cruzada e métodos de busca em hiperparâmetros são comumente empregadas para identificar a configuração que proporciona o melhor desempenho (PUTATUNDA; RAMA, 2018).

2.2.3 Importância de Características

A importância de características ou importância de atributos (do inglês *feature importance*) é um conceito fundamental na aprendizagem de máquina, especialmente no contexto de interpretabilidade de modelos e análise de dados. Refere-se a técnicas que quantificam a contribuição de cada variável de entrada para o desempenho preditivo de um modelo (ZIEN et al., 2009). Compreender a importância dos atributos auxilia na inter-

pretação dos modelos, na redução da dimensionalidade e na melhoria do desempenho, concentrando-se nas variáveis mais influentes (HOOKER et al., 2018).

Diversos métodos são empregados para avaliar a importância de atributos, cada um com suas vantagens e limitações. Uma abordagem comum envolve modelos baseados em árvores de decisão, como *Random Forest* e *Gradient Boosting Machines*. Nesses modelos, a importância pode ser quantificada medindo a redução total no critério de divisão (por exemplo, impureza de Gini ou entropia) proporcionada por cada atributo ao longo de todas as árvores no ensemble (LUNDBERG, 2017). Atributos que resultam em maiores reduções são considerados mais importantes.

Adicionalmente, técnicas como os valores SHAP (SHapley Additive exPlanations) fornecem uma medida unificada de importância de atributos ao calcular a contribuição de cada atributo para a previsão, baseando-se em conceitos da teoria dos jogos cooperativos (BROECK et al., 2022). Os valores SHAP oferecem consistência e precisão local, tornando-os valiosos para interpretar previsões individuais e compreender a importância global dos atributos.

Análises de Componentes Principais (PCA, do inglês *Principal Component Analysis*) e outros métodos de redução de dimensionalidade também podem fornecer perspectivas sobre a importância de atributos ao identificar combinações de variáveis (componentes principais) que capturam a maior variância nos dados (MAATEN et al., 2009). Embora esses métodos transformem os atributos originais, eles ajudam a entender a estrutura subjacente e a importância relativa das combinações de atributos.

A avaliação da importância de atributos tem várias implicações práticas. Em domínios como finanças, saúde e ciências ambientais (YIN; BINGI, 2023; LIN; BAI, 2022; YAN et al., 2022), compreender quais variáveis influenciam mais as previsões do modelo pode fornecer entendimentos acionáveis e apoiar processos de tomada de decisão. Além disso, ao identificar e remover atributos menos importantes, é possível reduzir a complexidade do modelo, mitigar o *overfitting* e melhorar a eficiência computacional (COFFER et al., 2021).

No entanto, é importante considerar potenciais vieses e limitações ao interpretar as pontuações de importância de atributos. Por exemplo, atributos correlacionados

podem levar a medidas de importância enganosas, uma vez que a importância pode ser distribuída entre variáveis correlacionadas (SANDRI; ZUCCOLOTTO, 2008). Em modelos baseados em árvores, atributos com mais níveis ou maior cardinalidade podem parecer artificialmente mais importantes (DENG; RUNGER; TUV, 2011). Portanto, uma análise cuidadosa e, quando necessário, a aplicação de múltiplos métodos são recomendadas para obter uma compreensão robusta da importância dos atributos.

2.2.4 Métricas e Divisão dos Dados

Focaremos nossa atenção nos problemas de aprendizado supervisionado, especificamente em tarefas de classificação. As métricas utilizadas variam conforme o tipo de problema. Para problemas de classificação, métricas como acurácia, que é a proporção de previsões corretas sobre o total de previsões realizadas, são frequentemente utilizadas (NAIDU; ZUVA; SIBANDA, 2023). Outras métricas importantes incluem a precisão (*precision*), que é a proporção de verdadeiros positivos em relação ao total de positivos previstos, e a revocação (*recall*), que é a proporção de verdadeiros positivos em relação ao total de positivos reais (DAVIS; GOADRICH, 2006). O F1-Score, que é a média harmônica entre precisão e revocação, é útil para balancear ambos em casos de classes desbalanceadas (SAILUSHA et al., 2020). Além disso, a matriz de confusão é uma ferramenta que permite visualizar o desempenho do algoritmo, mostrando verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos (LIANG, 2022).

A curva ROC (do inglês *Receiver Operating Characteristic*) é uma ferramenta essencial na avaliação do desempenho de modelos de classificação binária em aprendizado de máquina e estatística. Ela representa graficamente a relação entre a taxa de verdadeiros positivos e a taxa de falsos positivos em diversos limiares de decisão (NARKHEDE, 2018). Ao traçar a curva ROC, é possível visualizar a capacidade discriminativa do modelo em distinguir entre as classes positivas e negativas em diferentes pontos de corte. A área sob a curva ROC (AUC - Area Under the Curve) fornece uma medida quantitativa do desempenho global do modelo, onde um valor de 1 indica uma classificação perfeita e um valor de 0,5 corresponde a um desempenho aleatório (HOO; CANDLISH; TEARE, 2017). A análise da curva ROC permite comparar modelos distintos e selecionar aquele que

apresenta melhor capacidade preditiva, além de auxiliar na escolha do limiar de decisão mais adequado para equilibrar as taxas de falsos positivos e falsos negativos conforme os requisitos específicos da aplicação (MARZBAN, 2004).

A divisão adequada dos dados é fundamental para a construção de modelos robustos e generalizáveis (TAN et al., 2021). O conjunto de treinamento é usado para treinar o modelo, permitindo que ele aprenda os padrões e relações nos dados. O conjunto de validação é utilizado para ajustar hiperparâmetros e realizar a seleção de modelos, ajudando a avaliar o desempenho do modelo durante o processo de treinamento e a prevenir o *overfitting*. Por fim, o conjunto de teste é usado somente após o treinamento completo para avaliar o desempenho final do modelo em dados não vistos, fornecendo uma estimativa imparcial de sua capacidade de generalização (SHEPARD; NAHEED, 2021). Essa separação em treino, validação e teste assegura que o modelo não apenas memorize os dados de treinamento, mas também seja capaz de performar bem em dados novos e desconhecidos (BAI et al., 2021). Este panorama geral do aprendizado de máquina estabelece as bases para explorar mais profundamente as técnicas e metodologias específicas que serão abordadas nos tópicos seguintes.

2.3 Modelagem de Séries Temporais

A série temporal é definida como uma sequência de observações de uma ou mais variáveis registradas em intervalos de tempo sucessivos, frequentemente espaçados de forma uniforme (ARAGHINEJAD, 2014). A análise de séries temporais é uma área estatística e de aprendizado de máquina que se concentra na modelagem e previsão desses dados ordenados cronologicamente (CRYER, 1986). A natureza sequencial das séries temporais permite a captura de padrões temporais, tendências, ciclos e outras estruturas que não são aparentes em dados não ordenados (CHATFIELD, 2000). A compreensão e a modelagem adequadas de séries temporais são fundamentais em diversos domínios, como finanças, economia, meteorologia, engenharia e ciências sociais, onde a previsão precisa de eventos futuros é importante para a tomada de decisões informadas (HAMILTON, 2020).

Uma característica distintiva das séries temporais é a autocorrelação, na qual as observações em diferentes momentos podem estar correlacionadas (LOMNICKI; ZA-

REMBA, 1957). Essa dependência temporal requer técnicas especializadas de modelagem que considerem a estrutura sequencial dos dados. Modelos estatísticos clássicos, como os modelos autorregressivos (AR), de médias móveis (MA) e autorregressivos de médias móveis (ARMA), têm sido amplamente utilizados para capturar padrões lineares em séries estacionárias (ZHANG, 2018). Para séries não estacionárias, onde as propriedades estatísticas mudam ao longo do tempo, modelos como o autorregressivo integrado de médias móveis (ARIMA) e suas extensões sazonais (SARIMA) são aplicados para modelar e diferenciar tendências e sazonalidades (SIRISHA; BELAVAGI; ATTIGERI, 2022).

Com o avanço das técnicas de aprendizado de máquina, modelos mais complexos foram introduzidos para lidar com não linearidades e interações sofisticadas em séries temporais. Redes neurais recorrentes (RNNs), particularmente aquelas com unidades de memória de curto e longo prazo (LSTM) e redes neurais recorrentes bidirecionais, têm demonstrado eficácia na captura de dependências de longo alcance e padrões complexos em dados temporais (DUBEY et al., 2021). Além disso, modelos baseados em atenção e transformadores têm sido explorados para melhorar a capacidade de modelar relacionamentos temporais de maneira mais flexível (FARSANI; PAZOUKI, 2020).

A avaliação do desempenho de modelos de séries temporais requer métricas específicas que considerem a natureza sequencial dos dados. Erros como o Erro Absoluto Médio (MAE), o Erro Quadrático Médio (MSE) e o Erro Percentual Absoluto Médio (MAPE) são comumente utilizados para quantificar a precisão das previsões (AIJAZ; AGARWAL, 2020). Além disso, a validação cruzada em séries temporais, que respeita a ordem temporal dos dados, é importante para evitar a introdução de vieses e fornecer estimativas realistas do desempenho preditivo (BERGMEIR; BENÍTEZ, 2012).

A modelagem eficaz de séries temporais proporciona benefícios substanciais. Na área financeira, por exemplo, previsões precisas de preços de ativos, taxas de câmbio e indicadores econômicos informam estratégias de investimento e gestão de riscos (CHAN, 2004). Na gestão de operações e cadeia de suprimentos, a previsão de demanda permite otimizar estoques e recursos (AVIV, 2003). Na meteorologia, modelos de séries temporais contribuem para previsões climáticas e alertas de eventos extremos (MUDELSEE, 2010).

2.3.1 Prophet

O Prophet é um modelo de séries temporais desenvolvido pela Meta com o objetivo de fornecer uma ferramenta robusta e de fácil utilização para a previsão de séries temporais (RAFFERTY, 2021). Projetado para lidar com dados que apresentam tendências não lineares, sazonalidades múltiplas e efeitos de feriados, o Prophet baseia-se em um modelo aditivo onde diferentes componentes são combinados para capturar padrões complexos nos dados temporais (SIVARAMAKRISHNAN et al., 2022).

A estrutura fundamental do Prophet consiste em decompor a série temporal em três componentes principais: tendência, sazonalidade e efeitos de feriados (TAYLOR; LETHAM, 2018). A tendência representa a evolução a longo prazo dos dados e pode ser modelada como uma função linear ou logística para acomodar situações onde o crescimento apresenta saturação. A sazonalidade captura padrões periódicos que ocorrem em ciclos fixos, como variações diárias, semanais ou anuais, e é modelada utilizando funções periódicas baseadas em séries de Fourier. Os efeitos de feriados incorporam impactos ocasionais causados por eventos irregulares, permitindo que o modelo considere datas específicas que afetam significativamente a série temporal.

Uma característica distintiva do Prophet é sua abordagem orientada ao usuário, permitindo que especialistas no domínio influenciem o processo de modelagem. Os usuários podem incorporar conhecimento específico, como datas de feriados relevantes, ajustes manuais nas sazonalidades ou alterações nos parâmetros de crescimento. Essa flexibilidade facilita a adaptação do modelo a diferentes contextos e melhora a precisão das previsões em cenários onde informações adicionais estão disponíveis (TAYLOR; LETHAM, 2018).

Em termos de aplicação prática, o Prophet tem sido amplamente utilizado em diversas áreas, como previsão de demanda, análise de tráfego web, planejamento de recursos e monitoramento de métricas de negócios (BORGES; NASCIMENTO, 2022; KATWAL; SHRESTHA; SHARMA, 2024; DARAGHMEH et al., 2021; NAVRATIL; KOLKOVA, 2019). Sua implementação está disponível em bibliotecas para linguagens de programação como Python e R, o que facilita sua integração em pipelines de ciência de dados e sistemas de produção (TAYLOR; LETHAM, 2018).

2.3.2 SARIMA

O modelo *Seasonal Autoregressive Integrated Moving Average* (SARIMA) é uma extensão do modelo ARIMA que incorpora componentes sazonais para modelar e prever séries temporais que exibem padrões sazonais regulares (DUBEY et al., 2021). O SARIMA é amplamente utilizado em estatística e econometria para capturar tanto a dinâmica de curto prazo quanto as flutuações sazonais de longo prazo em dados temporais (KACZMARCZYK, 2017).

A formulação do modelo SARIMA é denotada como $SARIMA(p, d, q)(P, D, Q)$, onde p representa a ordem autorregressiva não sazonal, d é o grau de diferenciação não sazonal e q denota a ordem de médias móveis não sazonais. Os termos P , D e Q correspondem às ordens autorregressiva, de diferenciação e de médias móveis sazonais, respectivamente, enquanto s indica o período sazonal (por exemplo, $s = 12$ para dados mensais com sazonalidade anual). Os componentes não sazonais (AR, I, MA) capturam a dependência linear de curto prazo entre observações em diferentes momentos, enquanto os componentes sazonais (SAR, SI, SMA) modelam padrões que se repetem em intervalos fixos ao longo do tempo. A diferenciação (d e D) é aplicada para tornar a série estacionária, removendo tendências e componentes sazonais de nível, o que é importante para a aplicação adequada de modelos autorregressivos (FENG et al., 2022).

A modelagem com SARIMA envolve várias etapas críticas. Inicialmente, realiza-se uma análise exploratória dos dados para visualizar a série temporal, identificar tendências, sazonalidades e possíveis outliers. A estacionariedade da série é verificada utilizando testes estatísticos como o teste de Dickey-Fuller aumentado (MUSHTAQ, 2011). Se a série não for estacionária, aplica-se diferenciação não sazonal (d) e/ou sazonal (D) para alcançar estacionariedade em média e variância (BOX et al., 2015).

A identificação dos parâmetros p, d, q, P, D, Q é guiada pela análise das funções de autocorrelação (ACF) e autocorrelação parcial (PACF), que ajudam a detectar autocorrelações em diferentes defasagens e a determinar a presença de componentes autorregressivos e de médias móveis (BOX et al., 2015). A estimação dos parâmetros é realizada por métodos como máxima verossimilhança ou mínimos quadrados, visando minimizar a diferença entre as previsões do modelo e os valores observados (GIKUNGU; WAITITU;

KIHORO, 2015).

O diagnóstico do modelo envolve a análise dos resíduos para verificar se eles se comportam como ruído branco, ou seja, se não apresentam autocorrelação significativa e possuem distribuição aproximadamente normal com média zero. Ferramentas como o teste de Ljung-Box são utilizadas para avaliar a independência dos resíduos (PEREIRA et al., 2021). Uma vez que o modelo é considerado adequado, ele pode ser utilizado para realizar previsões, incorporando tanto os componentes não sazonais quanto os sazonais identificados durante o processo de modelagem.

O modelo SARIMA é particularmente útil em situações onde a sazonalidade desempenha um papel significativo na dinâmica da série temporal, como na previsão de vendas, consumo de energia ou indicadores climáticos (CHOI; YU; AU, 2011; BLÁZQUEZ-GARCÍA et al., 2020; ZIA, 2021). Sua capacidade de modelar dependências em diferentes escalas temporais o torna adequado para uma ampla gama de aplicações que requerem previsões precisas e confiáveis.

O modelo SARIMA é uma ferramenta robusta e versátil para a análise e previsão de séries temporais com características sazonais. Ao integrar componentes sazonais à estrutura do ARIMA, o SARIMA fornece um meio estruturado para capturar dependências temporais de curto e longo prazo, auxiliando pesquisadores e profissionais na tomada de decisões informadas em diversos setores onde a dinâmica temporal é um fator crítico (VAGROPOULOS et al., 2016).

2.3.3 LSTM

As redes neurais artificiais são modelos computacionais inspirados no funcionamento do cérebro humano, projetadas para reconhecer padrões complexos e realizar tarefas como classificação, regressão e previsão (WALCZAK, 2019). Elas consistem em camadas de neurônios artificiais interconectados, onde cada neurônio processa as entradas recebidas, aplica uma função de ativação e transmite o resultado para os neurônios da próxima camada. O aprendizado em redes neurais é alcançado através do ajuste iterativo dos pesos sinápticos, geralmente utilizando algoritmos de otimização como o gradiente descendente e técnicas como a retropropagação do erro (WU; FENG, 2018).

As redes neurais recorrentes (RNNs) são uma classe especial de redes neurais projetadas para lidar com dados sequenciais ou temporais, onde a ordem das informações é importante (CATERINI et al., 2018). Diferentemente das redes *feedforward* tradicionais, as RNNs possuem conexões cíclicas que permitem que informações de estados anteriores influenciem o processamento atual (TSANTEKIDIS; PASSALIS; TEFAS, 2022). Essa característica confere às RNNs uma forma de memória interna, possibilitando a modelagem de dependências temporais e padrões sequenciais em tarefas como processamento de linguagem natural, reconhecimento de fala e análise de séries temporais (JELODAR et al., 2020; ORUH; VIRIRI; ADEGUN, 2022; HEWAMALAGE; BERGMER; BANDARA, 2021).

No entanto, as RNNs tradicionais enfrentam limitações significativas ao lidar com dependências de longo prazo devido ao problema do gradiente desaparecente ou explosivo durante o treinamento (YU et al., 2019). Esse problema ocorre quando os gradientes utilizados para atualizar os pesos se tornam extremamente pequenos ou grandes, dificultando o aprendizado de padrões que dependem de informações distantes na sequência (FADZISO, 2020). Como resultado, as RNNs podem ter desempenho limitado em tarefas que exigem a captura de relações de longo alcance nos dados temporais.

Para superar essas limitações, foram desenvolvidas arquiteturas recorrentes mais avançadas, sendo a Long Short-Term Memory (LSTM) uma das mais proeminentes. As LSTMs foram introduzidas em 1997 como uma solução para o problema do gradiente desaparecente, permitindo que a rede aprenda dependências de longo prazo de maneira eficaz (HOCHREITER, 1997). A arquitetura LSTM introduz uma estrutura de memória especializada, denominada célula LSTM, que utiliza mecanismos de portas para controlar o fluxo de informações.

A célula LSTM contém três portas principais: a porta de entrada, a porta de esquecimento e a porta de saída. A porta de entrada controla quais informações das novas entradas devem ser adicionadas ao estado interno da célula. A porta de esquecimento determina quais informações previamente armazenadas devem ser descartadas, permitindo que a rede atualize sua memória de forma seletiva. A porta de saída regula quais informações do estado interno serão utilizadas para gerar a saída atual e transmitidas para

o próximo estado (GRAVES; GRAVES, 2012). Essas portas são implementadas por meio de funções sigmoide que produzem valores entre zero e um, atuando como coeficientes de filtragem das informações.

O mecanismo de portas das LSTMs permite que a rede retenha ou ignore informações conforme necessário, adaptando-se dinamicamente às características dos dados sequenciais (LANDI et al., 2021). Isso confere à LSTM a capacidade de capturar dependências temporais de longo alcance, tornando-a altamente eficaz em tarefas onde o contexto histórico é fundamental. Além disso, as LSTMs são capazes de lidar com sequências de comprimentos variáveis e de processar dados em tempo real, o que amplia suas aplicações práticas (LI; ABDEL-ATY; YUAN, 2020).

O treinamento de redes LSTM é realizado utilizando algoritmos de otimização baseados em gradiente, como o *Backpropagation Through Time (BPTT)*, que estende o método de retropropagação para redes recorrentes (CHEN; HUO, 2016). Devido à complexidade da arquitetura e à necessidade de processar sequências inteiras, o treinamento pode ser computacionalmente intensivo e sensível à inicialização dos pesos e à escolha dos hiperparâmetros. Técnicas como regularização, normalização por lotes e otimização com gradiente estocástico adaptativo são frequentemente utilizadas para melhorar a convergência e o desempenho do modelo (MERITY; KESKAR; SOCHER, 2017).

2.3.4 Métricas

A avaliação de modelos de séries temporais é uma etapa importante no processo de modelagem e previsão, pois permite quantificar a precisão das previsões e a adequação do modelo aos dados observados. Para isso, utilizam-se métricas de erro que mensuram a discrepância entre os valores previstos pelo modelo e os valores reais da série temporal. Entre as métricas de erro mais comuns estão o Erro Absoluto Médio (MAE), o Erro Quadrático Médio (MSE), a Raiz do Erro Quadrático Médio (RMSE) e o Erro Percentual Absoluto Médio (MAPE) (HYNDMAN, 2018).

O Erro Absoluto Médio (MAE) é calculado como a média das diferenças absolutas entre as previsões e os valores observados. O MAE fornece uma medida da magnitude média dos erros, sem considerar a direção dos desvios. O Erro Quadrático Médio (MSE) é

a média dos quadrados das diferenças entre as previsões e os valores observados. O MSE penaliza erros maiores de forma mais intensa devido à elevação ao quadrado, o que o torna sensível a outliers. A Raiz do Erro Quadrático Médio (RMSE) é a raiz quadrada do MSE. O RMSE tem a vantagem de estar na mesma unidade dos dados originais, facilitando a interpretação. O Erro Percentual Absoluto Médio (MAPE) expressa o erro como uma porcentagem dos valores observados. O MAPE é útil para comparar o desempenho em diferentes escalas, mas pode ser problemático quando aproxima-se de zero, levando a valores extremos.

A escolha da métrica de erro apropriada depende das características da série temporal e dos objetivos da análise. É comum utilizar múltiplas métricas para obter uma avaliação abrangente do desempenho do modelo (SAIGAL; MEHROTRA, 2012). Além das métricas de erro, é fundamental avaliar a significância estatística dos parâmetros estimados e a adequação do modelo aos dados. O teste t de Student é amplamente utilizado para verificar se os coeficientes estimados em um modelo de regressão (incluindo modelos de séries temporais) são estatisticamente diferentes de zero (BUONOCORE; ASTE; MATTEO, 2016). O valor t é então comparado com a distribuição t de Student com graus de liberdade apropriados para determinar o p-valor associado. Se o p-valor for menor que um nível de significância pré-definido (por exemplo, 0,05), rejeita-se a hipótese nula, indicando que o coeficiente é estatisticamente significativo (WOOLDRIDGE, 2002).

Outro aspecto importante é a verificação dos pressupostos do modelo, como a normalidade dos resíduos. O teste de Shapiro-Wilk é um teste de aderência que avalia a hipótese nula de que uma amostra provém de uma distribuição normal (GONZÁLEZ-ESTRADA; COSMES, 2019). A estatística do teste é calculada com base na correlação entre os dados ordenados e os valores esperados se a distribuição fosse normal. Um p-valor menor que o nível de significância indica que se rejeita a hipótese de normalidade dos resíduos. A avaliação da normalidade dos resíduos é essencial, pois muitos procedimentos estatísticos, incluindo intervalos de confiança e testes de hipóteses, assumem que os erros são normalmente distribuídos. Desvios significativos dessa suposição podem indicar problemas na especificação do modelo ou a presença de outliers e heterocedasticidade.

A avaliação de modelos de séries temporais requer uma combinação de métricas

de erro para quantificar a precisão das previsões e testes estatísticos para verificar a significância dos parâmetros e a adequação dos pressupostos do modelo. O uso apropriado de métricas como MAE, MSE, RMSE e MAPE, aliado a testes como o teste t de Student e o teste de Shapiro-Wilk, ajuda a validação e refinamento de modelos preditivos, assegurando a confiabilidade das conclusões e decisões baseadas nas análises realizadas.

Em estudos sobre suicídio, a abordagem de utilizar várias análises é essencial para identificar padrões complexos e fatores preditivos. A aplicação de modelos estatísticos avançados e algoritmos de aprendizado de máquina permite uma análise mais profunda e precisa dos dados. Essas técnicas podem revelar padrões não detectáveis por métodos tradicionais e contribuir para a elaboração de estratégias de intervenção mais eficazes. Modelos como Prophet, SARIMA e LSTM são capazes de capturar tendências complexas e prever comportamentos futuros, fornecendo informações para o planejamento de políticas de saúde. Algoritmos como XGBoost e *Random Forest* permitem identificar fatores-chave associados a desfechos específicos, como a escolha do método de suicídio, possibilitando intervenções direcionadas e eficientes.

A fundamentação teórica apresentada fornece um embasamento introdutório e sólido para a compreensão do fenômeno do suicídio e das metodologias empregadas neste trabalho. Compreender os múltiplos fatores que influenciam o suicídio, bem como as ferramentas analíticas disponíveis, é essencial para abordar de forma eficaz esse problema de saúde pública. Este estudo busca contribuir com o tema ao fornecer clareza sobre tópicos que possam orientar políticas e ações preventivas no Brasil.

3 Trabalhos Relacionados

Diversos estudos têm explorado o uso de modelos preditivos e técnicas de aprendizado de máquina para estimar padrões e tendências em dados de saúde pública. Neste capítulo, serão destacados alguns trabalhos dentro do contexto em questão.

Feng et al. (2022) teve como objetivo desenvolver modelos preditivos de curto prazo para estimar o número de internações por acidentes de trânsito (RTIs) no Nordeste da China, e fazer uma análise comparativa entre os modelos SARIMA, LSTM e Prophet. Os modelos foram treinados com dados de 2015 a 2019, e a acurácia preditiva foi avaliada com base em dados de 2020. As métricas de desempenho utilizadas foram RMSE, MAE e MAPE. O modelo LSTM apresentou a maior precisão preditiva, seguido pelo Prophet, enquanto o SARIMA teve o pior desempenho. Os resultados indicam que o modelo LSTM é o mais robusto e eficaz na previsão de internações com padrões sazonais e variações abruptas, fornecendo uma base sólida para o planejamento e gestão de recursos na administração de saúde (FENG et al., 2022).

Soares et al. (2023) teve como objetivo descrever a tendência das taxas de suicídio no Brasil de 2011 a 2020 e investigar a associação dessas taxas com variáveis relacionadas à pandemia de COVID-19. O estudo utilizou um método ecológico de série temporal para analisar a mortalidade por suicídio, com foco em 2020, o primeiro ano da pandemia. A taxa de mortalidade por suicídio por 100 mil habitantes foi o desfecho principal. O estudo empregou a regressão generalizada de Prais-Winsten para calcular a tendência das taxas de suicídio ao longo do período e utilizou o modelo ARIMA para prever as taxas de suicídio em 2020.

Além disso, investigou a associação entre as taxas de suicídio e variáveis como mortalidade por COVID-19, desemprego, recebimento de auxílio emergencial e taxas de distanciamento social. Os resultados mostraram uma taxa de suicídio de 6,68 por 100 mil habitantes em 2020. As taxas de suicídio aumentaram em ambos os sexos e em todas as regiões do país, especialmente no Sul. No entanto, o estudo não encontrou evidências de um aumento considerável de suicídios em 2020 em relação ao número esperado durante o

primeiro ano da pandemia. O estudo concluiu que, embora as taxas de suicídio tenham aumentado de 2011 a 2020, especialmente entre homens e nas regiões Sul, não houve um aumento inesperado no primeiro ano da pandemia, com a tendência se mantendo com o crescimento dentro do esperado em 2020 (SOARES; STAHNKE; LEVANDOWSKI, 2023).

Um estudo de 2024 investigou padrões de tentativas e mortes por suicídio no Sistema Único de Saúde (SUS) do Brasil entre 2009 e 2020, focando nos estados do Sul do país (PEREIRA et al., 2024). O objetivo foi identificar como diferentes fatores de risco interagem e afetam o risco de suicídio, integrando esses dados com informações sobre a infraestrutura de saúde. O problema foi enquadrado como uma tarefa de classificação para prever a taxa de suicídio e re-tentativas (com as classes sendo baixa, moderada ou alta) em cada microrregião. Foram desenvolvidos e comparados três modelos de Machine Learning: *Árvore de Decisão*, *Random Forest* e *XGBoost*, com uma precisão de aproximadamente 97%. As variáveis mais importantes para prever as taxas de suicídio foram o número de unidades e clínicas de saúde mental, e para prever tanto suicídios quanto re-tentativas, o número de médicos e enfermeiros disponíveis. O estudo traz novas perspectivas sobre onde os governos devem investir para reduzir as taxas de suicídio.

No contexto da Coreia do Sul, um trabalho investigou o efeito da pandemia de COVID-19 nas tendências de suicídio na Coreia por meio de uma análise de séries temporais (RYU et al., 2022). Usaram o Prophet para gerar modelos de previsão com base no número mensal de mortes por suicídio na Coreia do Sul entre 1997 e 2018. Validaram os modelos comparando com os números de suicídios de 2019 e previram os números de suicídios para 2020. A comparação entre os números esperados e os observados durante a pandemia indicou que o total de suicídios durante a pandemia não se desviou significativamente das projeções feitas com base no período pré-pandemia. No entanto, o número de suicídios entre mulheres e pessoas com menos de 34 anos ultrapassou significativamente o nível esperado. O estudo concluiu que a pandemia de COVID-19 não aumentou significativamente a taxa geral de suicídios, mas que o aumento de suicídios entre mulheres e jovens sugere que a pandemia pode ter contribuído para o aumento de suicídios nesses grupos. O estudo recomenda mais pesquisas para verificar o impacto a longo prazo da pandemia na quantidade de ocorrências de suicídio.

Tanaka et al. (2021) analisou o impacto da pandemia de COVID-19 na mortalidade por suicídio no Japão, utilizando registros mensais de suicídios. Usando *difference-in-difference estimation*, os pesquisadores descobriram que as taxas mensais de suicídio caíram 14% durante os primeiros cinco meses da pandemia (fevereiro a junho de 2020). Essa queda pode ter ocorrido devido a fatores complexos, como subsídios generosos do governo, redução das horas de trabalho e fechamento das escolas. No entanto, as taxas mensais de suicídio aumentaram 16% durante a segunda onda da pandemia (julho a outubro de 2020), com um aumento maior entre mulheres (37%) e crianças e adolescentes (49%). O estudo sugere que, embora os impactos negativos da pandemia de COVID-19 possam persistir a longo prazo, modificadores como os subsídios governamentais podem não ser sustentáveis. Assim, a prevenção eficaz do suicídio, especialmente entre populações vulneráveis, deve ser uma importante consideração de saúde pública (TANAKA; OKAMOTO, 2021).

Apesar dos avanços, há lacunas importantes, como dados atualizados - muitos estudos utilizam dados até 2019 ou início de 2020, não capturando totalmente o impacto da pandemia. Há poucos estudos comparando diferentes modelos estatísticos e de aprendizado de máquina para previsão e classificação em um mesmo conjunto de dados. A maioria das pesquisas foca nas taxas gerais de suicídio, sem aprofundar nos métodos empregados e fatores associados a essa escolha. O presente trabalho busca preencher algumas dessas lacunas ao utilizar dados recentes - foram analisados dados até dezembro de 2022, fornecendo uma visão atualizada das tendências de suicídio no Brasil. Além disso, pretende-se aqui realizar uma análise comparativa dos modelos Prophet, SARIMA e LSTM para análise de séries temporais, além de XGBoost e *Random Forest* para classificação, permitindo um estudo abrangente. Ao reunir a causa básica em dois grandes grupos - auto-intoxicação e lesão autoprovocada, foi investigado fatores que se relacionam à escolha do meio, contribuindo para intervenções mais direcionadas.

4 Metodologia

Este capítulo descreve a metodologia empregada neste estudo para analisar as tendências de suicídio no Brasil entre 2003 e 2022, avaliar se houve impacto da pandemia de COVID-19 e identificar fatores associados aos métodos de suicídio. A abordagem adotada combina técnicas de análise estatística descritiva, modelos de séries temporais e algoritmos de classificação. A seguir, são detalhadas cada etapa do processo metodológico.

4.1 Fonte dos Dados

Os dados utilizados neste estudo foram extraídos da Plataforma de Dados de Pesquisa em Saúde (PCDAS)¹, que disponibiliza acesso a diversas bases de dados de saúde pública no Brasil. Especificamente, foram utilizados dados do Sistema de Informação sobre Mortalidade (SIM) do Departamento de Informática do SUS (DATASUS), que reúne informações sobre óbitos registrados em todo o país.

Foi realizada uma consulta SQL para selecionar os registros de óbitos por suicídio entre 2003 e 2022, com base nos códigos de causa básica relacionados ao suicídio. Seguindo o padrão de boletins epidemiológicos do Ministério da Saúde, consideraram-se suicídios os óbitos em indivíduos com 5 anos de idade ou mais, com causa básica registrada com códigos X60-X84 e Y87.0 da Classificação Internacional de Doenças, décima revisão (CID-10) (SAUDE, 2024).

Essa consulta resultou em 218.707 registros, incluindo informações sobre data do óbito, causa básica, sexo, raça/cor, escolaridade, local de ocorrência, município de residência, estado da ocorrência e idade do óbito. As variáveis escolhidas foram baseadas nas variáveis que aparecem em boletins epidemiológicos do governo e possuíam mais de 60% de preenchimento.

¹<https://pcdas.icict.fiocruz.br/sobre-nos/>

Variável	Descrição	Valores possíveis
ano_obito	Ano do óbito	2003 a 2022
data_obito	Data do óbito	01-01-2003 a 31-12-2022
CAUSABAS	Causa básica do óbito	X60 a X84 e Y87
SEXO	Sexo	0:“Ignorado”, 1:“Masculino”, 2:“Feminino”
RACACOR	Raça/Cor	1:“Branca”, 2:“Preta”, 3:“Amarela”, 4:“Parda”, 5:“Indígena”, 9:“Ignorado”
ESC	Escolaridade	1:“Nenhuma”,2:“1 a 3 anos”,3:“4 a 7 anos”,4:“8 a 11 anos”,5:“12 e mais”,9:“Ignorado”
def_loc_ocor	Local de ocorrência	Hospital; Outros estabelecimentos de saúde; Domicílio; Via pública; Outros; Ignorado
res_MUNNOME	Município de residência	Nome (acentuado, maiúsculas e minúsculas) do Município (padrão DOS, página de código 850) de residência da pessoa que foi à óbito
ocor_SIGLA_UF	UF da ocorrência	Sigla da unidade da federação de ocorrência do óbito
idade_obito_anos	Idade do óbito em anos	Valores maiores que 5 anos

Tabela 4.1: Descrição das variáveis selecionadas no banco de dados

A única variável que continha valores *NaN* era a `idade_obito_anos`, totalizando 519 instâncias nessa situação. A mediana de todas as idades foi utilizada para preencher os valores.

4.2 Código Fonte

O código fonte em *Python* utilizado para a realização do tratamento dos dados, das análises e dos modelos pode ser encontrado em um *notebook* do *Google Colab*². Os dados aqui utilizados foram cedidos pela Fiocruz e para solicitá-los deve ser feito contato pela PCDaS³.

²<https://colab.research.google.com/drive/1VqRCtDAHdJGvuyyEpX81VXoQnxSPHV92?usp=sharing>

³<https://pcdas.icict.fiocruz.br/contato/>

4.3 Análise Estatística Descritiva

A primeira etapa consistiu em realizar uma análise descritiva dos dados para identificar padrões e distribuições das variáveis sociodemográficas. Para analisar as tendências de suicídio sem que o aumento no número de suicídios seja apenas um reflexo do crescimento populacional, calculou-se as taxas de suicídio por 100 mil habitantes para cada ano, ajustando-as pela população correspondente de cada ano. Foi utilizado como numerador o número de suicídios em um determinado ano e, como denominador, o número total de indivíduos, essa divisão foi então multiplicada por 100 mil, como sugerido por Soares (SOARES; STAHNKE; LEVANDOWSKI, 2023).

As taxas de suicídio por 100 mil habitantes foram calculadas por período, estado, municípios com maior incidência, sexo, raça, faixa etária, escolaridade e causa básica da morte. As informações do denominador foram obtidas a partir do censo populacional de 2010 e 2022 e das projeções populacionais do Instituto Brasileiro de Geografia e Estatística (IBGE) (Instituto Brasileiro de Geografia e Estatística (IBGE), 2024). Além disso, foram gerados gráficos para comparação entre sexo e causa básica dividida em dois grandes grupos (lesão autoprovocada e auto-intoxicação). Os resultados serão apresentados na seção 5.

4.4 Modelagem de Séries Temporais

Para analisar as tendências temporais dos suicídios e avaliar se houve impacto da pandemia de COVID-19, foram utilizados três modelos de séries temporais. Primeiro foi feita uma agregação mensal onde os dados foram agregados por mês, resultando em uma série temporal mensal de número de suicídios de janeiro de 2003 a dezembro de 2022. Não haviam valores faltantes.

Utilizar dados mensais na modelagem de séries temporais para análise de suicídios é interessante porque permite identificar padrões sazonais e tendências de longo prazo com mais clareza. Essa granularidade ajuda a capturar variações mensais significativas, como as influenciadas por eventos sociais ou econômicos, e permite prever comportamentos futuros de forma mais detalhada. Além disso, os dados mensais oferecem um equilíbrio

entre a simplicidade dos dados anuais e o ruído excessivo dos dados diários, sendo ideais para detectar impactos mensuráveis, como o efeito da pandemia de COVID-19, sobre as taxas de suicídio.

Os três modelos utilizados (Prophet, SARIMA, LSTM) seguiram o mesmo padrão de treino com dados até 2018 e janela de informação de 12 meses antes. Uma janela de 12 meses para a modelagem de séries temporais é justificada pela captura de padrões sazonais anuais, comuns em diversos tipos de dados como saúde e economia (BOX et al., 2015). Essa janela equilibra a memória de curto e longo prazo, permitindo que o modelo identifique tanto tendências recentes quanto padrões sazonais recorrentes, sem sobrecarregar o processo com dados antigos ou irrelevantes.

Os modelos foram executados para gerarem previsões dos meses de 2019 a 2022. Com os dados de 2019, os modelos foram avaliados segundo as métricas MAE, MSE, RMSE, e MAPE. Cada uma dessas métricas oferece uma perspectiva diferente sobre o desempenho do modelo, variando na sensibilidade a erros grandes e na interpretação do resultado. Além disso, os resultados das previsões de 2019 de cada modelo passaram por um teste de significância estatística.

Duas análises estatísticas foram realizadas para avaliar a qualidade das previsões dos modelo em relação aos valores observados. Primeiro, os resíduos (diferença entre os valores observados e previstos) são calculados. Um teste de Shapiro-Wilk é aplicado para verificar se os resíduos seguem uma distribuição normal. Se o p-valor for maior que 0,05, não rejeita-se a hipótese nula de que os resíduos seguem uma distribuição normal.

Em seguida, um teste t-student é realizado para verificar se a média dos resíduos é significativamente diferente de zero, o que indicaria que as previsões são diferentes dos valores reais. Se o p-valor for menor que 0,05, a hipótese nula é rejeitada, significando que há uma diferença significativa. Após os modelos serem avaliados com os dados reais de 2019 e verificadas as suas respectivas capacidades preditivas com as métricas, as previsões para 2020, 2021 e 2022 foram comparadas com seus valores reais.

4.4.1 Características Específicas dos Modelos

O modelo Prophet foi inicializado com um intervalo de confiança de 95% (*interval_width=0.95*) para garantir uma margem de segurança razoável para lidar com a variabilidade inerente aos dados de séries temporais, sem ser excessivamente conservador. O método *plot_components()* do modelo também foi utilizado para visualizar os componentes da previsão, como a tendência e a sazonalidade, oferecendo uma análise detalhada dos fatores que influenciam as previsões do modelo.

O modelo SARIMA teve os dados primeiramente submetidos a um teste de dickey-fuller aumentado, para verificar se a série é estacionária. Depois, foi ajustado com dois conjuntos de parâmetros principais: *order=(p, d, q)* que são os parâmetros ARIMA não sazonais, onde *p* é a ordem do componente autoregressivo e tem valor 1, *d* é o grau de diferenciação e tem valor 1, e *q* é a ordem do componente de média móvel com valor 1; e *seasonal_order=(P, D, Q, m)* que refere-se aos parâmetros sazonais do SARIMA, com *P* sendo o componente autoregressivo sazonal com valor 1, *D* a diferenciação sazonal com valor 1, *Q* o componente de média móvel sazonal também com valor 1, e *m* com valor 12 indicando a sazonalidade anual (dados mensais). As opções *enforce_stationarity* e *enforce_invertibility* foram adicionadas como *false* e permitem que o modelo aceite estados estacionários ou invertíveis, mas não forçam essas condições. Após ajustado o teste de dickey-fuller aumentado foi novamente rodado e verificou-se a estacionariedade da série.

O modelo LSTM teve os dados primeiramente normalizados usando o *MinMaxScaler* para escalonar os números de suicídio entre 0 e 1. Diferente do Prophet e SARIMA, redes neurais, como o LSTM, tendem a funcionar melhor com dados escalonados porque os valores não escalonados podem dificultar a convergência durante o treinamento, devido ao modo como essas redes processam os dados (GOODFELLOW, 2016). Em seguida, foi criada uma função para gerar sequências de tempo, onde os últimos 12 meses de dados são usados para prever o mês seguinte. O modelo foi definido com 50 unidades ocultas e uma camada densa final para gerar a previsão de um valor. Foi compilado usando o otimizador *Adam* e a função de perda *mean squared error*. O treinamento do modelo foi realizado com 100 épocas, incluindo a técnica de *early stopping* para interromper o treinamento caso a perda no conjunto de validação não melhorasse por 5 épocas consecutivas,

evitando *overfitting*. Os parâmetros foram definidos tendo em vista o tamanho pequeno do conjunto de dados.

4.5 Modelagem com Classificação

A causa básica da morte foi dividida entre dois grupos, auto-intoxicação e lesão auto-provocada, baseado nos códigos CID-10. A causa básica da morte (método de suicídio utilizado) se tornou a variável alvo a ser predita. Em um primeiro momento, foi realizado o pré-processamento de dados e o balanceamento de classes para os modelos. As colunas categóricas foram convertidas em valores numéricos usando o *LabelEncoder*. Os dados foram divididos em treino e teste na proporção 70:30. Nos dados de treino foi aplicada a técnica de *undersampling* para lidar com o desbalanceamento das classes no conjunto de dados. A classe majoritária foi reduzida para ter o mesmo número de amostras que a classe minoritária (19798). Para garantir que ambas as classes ficassem equilibradas em tamanho, o resample da biblioteca sklearn foi utilizado para criar essa nova amostra balanceada, combinando as instâncias da classe majoritária subamostrada com as da classe minoritária.

4.5.1 Características Específicas dos Modelos

No modelo *Random Forest* foi usado o *RandomForestClassifier* do pacote *ensemble* da biblioteca sklearn, com 100 estimadores (árvores de decisão) definidos pelo parâmetro *n_estimators*. No modelo *XGBoost* foi usado o *XGBClassifier* da biblioteca *xgboost*, com o parâmetro *objective* como 'binary:logistic'. Depois de treinados, os modelos foram avaliados segundo métricas de desempenho usando o conjunto de dados de teste. As métricas incluem acurácia, precisão, *recall*, F1-Score, e ROC-AUC.

Além disso, foi extraída a importância das características de cada um dos modelos. Essa importância é usada para determinar o grau de relevância de cada variável independente em relação ao modelo preditivo. Em modelos de árvore de decisão, como *Random Forest* e *XGBoost*, a importância das características é medida com base em quanto cada variável contribui para reduzir a impureza nas divisões da árvore. Quanto

maior a redução de impureza proporcionada por uma variável, maior sua importância.

Foram realizadas 100 execuções de cada modelo posto que é uma prática recomendada para garantir a robustez e a estabilidade dos resultados, especialmente em cenários quando há aleatoriedade envolvida no processo de treinamento (KOHAVI et al., 1995). Ao rodar o modelo repetidamente, variando o *random_state*, é possível obter uma visão mais precisa das métricas de desempenho ao calcular suas médias e desvios padrão. Isso permite uma avaliação mais confiável do modelo, minimizando o impacto de flutuações ocasionais ou divisões específicas dos dados. Além disso, as múltiplas execuções ajudam a validar a consistência da importância das variáveis e a generalização do modelo, fornecendo uma base sólida para conclusões sobre a eficácia do modelo treinado.

5 Resultados e Discussão

Este capítulo apresenta os resultados obtidos a partir das análises realizadas e discute suas implicações no contexto das tendências de suicídio no Brasil entre 2003 e 2022, com ênfase no impacto da pandemia de COVID-19. Os resultados são organizados em três seções principais - Análise Descritiva, Análise de Séries Temporais e Modelagem de Classificação Binária.

5.1 Análise Descritiva

A Figura 5.1 apresenta a taxa de suicídio no Brasil entre os anos de 2003 e 2022, revelando uma tendência de crescimento, com algumas fases de estabilidade e outras de aumento mais acentuado. Entre 2003 e 2010, a taxa de suicídio permaneceu relativamente estável, variando entre 4.4 e 5.0 por 100.000 habitantes, com um crescimento lento e modesto. Esse período pode indicar um contexto em que os fatores de risco para suicídio, como condições socioeconômicas e saúde mental, estavam relativamente equilibrados, sem grandes mudanças que afetassem drasticamente os números.

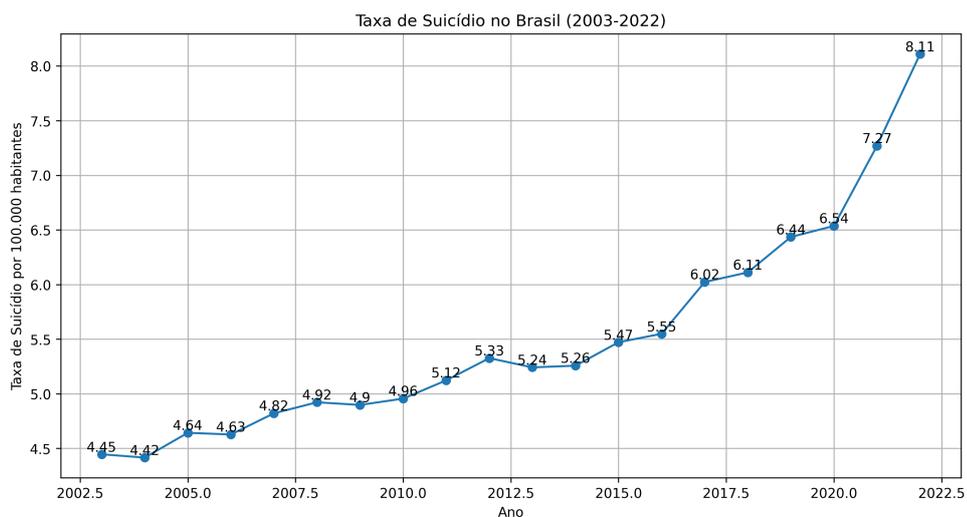


Figura 5.1: Taxa de Suicídio no Brasil

A partir de 2011, observa-se um crescimento gradual e constante na taxa de suicídio, que atinge 5.55 em 2016. Esse aumento pode ser explicado por diversas variáveis,

como mudanças econômicas, aumento de distúrbios mentais e sociais, ou melhorias no registro e notificação dos casos de suicídio. No entanto, foi a partir de 2017 que a taxa de crescimento se intensificou, com a taxa de suicídio subindo para 6.11 em 2019, antes da pandemia de COVID-19.

Após 2020, nota-se uma aceleração ainda mais acentuada, com a taxa subindo rapidamente de 6.54 para 8.11 por 100.000 habitantes até 2022. Esse aumento pode estar associado aos impactos negativos da pandemia, como o aumento do isolamento social, dificuldades econômicas e uma crise de saúde mental. O período entre 2020 e 2022 marca um ponto crítico, no qual os fatores desencadeados pela pandemia podem ter agravado os problemas já existentes, resultando em um aumento nas taxas de suicídio. Ao longo dos 20 anos analisados, a taxa de suicídio no Brasil praticamente dobrou, passando de 4.45 para 8.11 por 100.000 habitantes.

A Figura 5.2 é um gráfico que mostra a evolução da taxa de suicídio no Brasil entre 2003 e 2022, segmentada por faixa etária. Observa-se que a faixa etária de 30 a 49 anos apresenta a maior taxa ao longo do período, com um crescimento mais acentuado a partir de 2020. As faixas etárias de 50 a 69 anos e 70 anos ou mais também mostram um aumento gradual, com destaque para o crescimento mais pronunciado nos últimos anos. Já as faixas etárias mais jovens, de 5 a 14 anos e 15 a 19 anos, apresentam taxas de suicídio mais baixas e relativamente estáveis, com um ligeiro aumento após 2020. Embora as taxas de suicídio sejam mais elevadas entre adultos e idosos, é importante notar que o aumento recente em várias faixas etárias sugere um agravamento da situação geral.

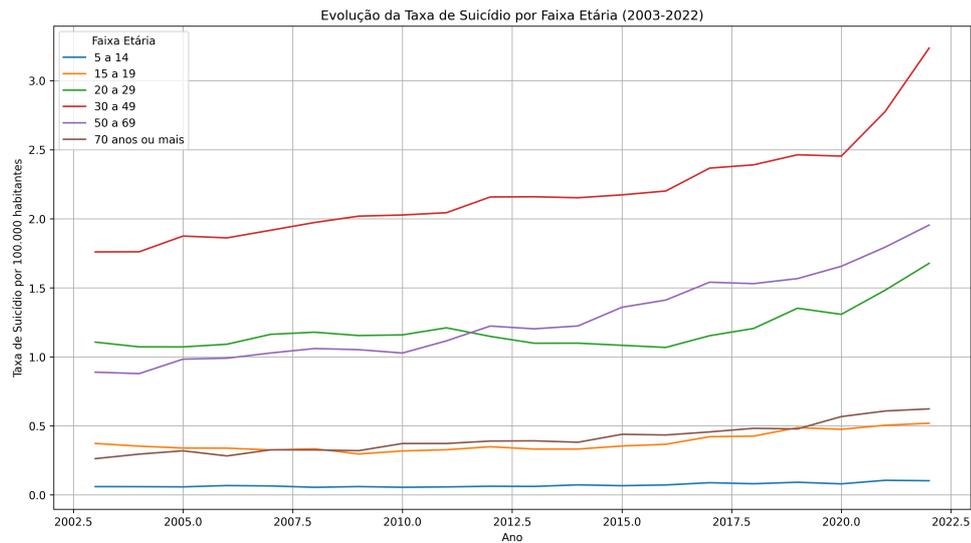


Figura 5.2: Evolução da Taxa de Suicídio por Faixa Etária

A Figura 5.3 mostra a evolução da taxa de suicídio segmentada por sexo. A linha azul representa o sexo masculino (1) e a linha laranja representa o sexo feminino (2). Revela-se uma disparidade significativa entre as taxas de suicídio de homens e mulheres no Brasil, com os homens apresentando taxas substancialmente mais altas e um aumento mais acentuado ao longo do tempo. Essa diferença de gênero reflete um padrão global conhecido, em que os homens são mais propensos a cometer suicídio, possivelmente devido a fatores como menor busca por ajuda, normas sociais e uso de métodos mais letais.

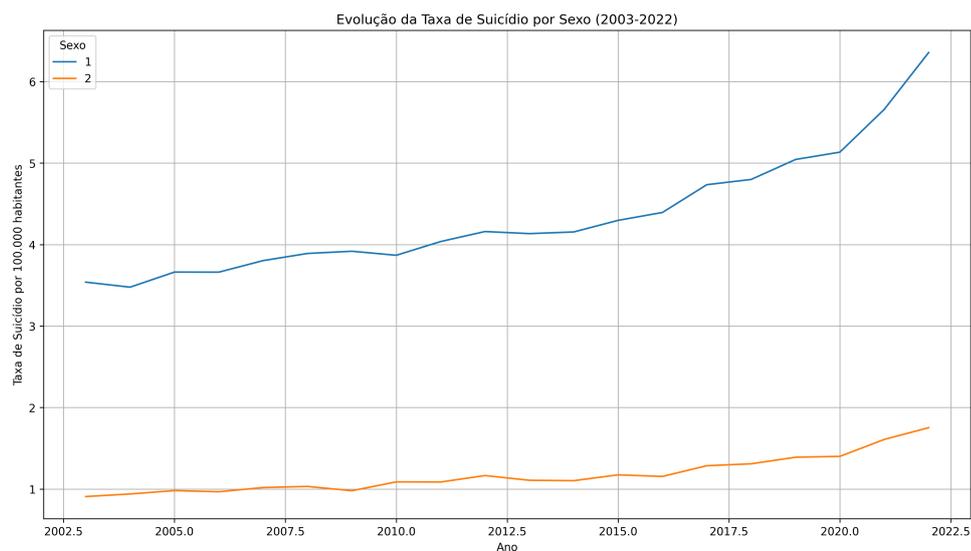


Figura 5.3: Evolução da Taxa de Suicídio por Sexo

O gráfico da Figura 5.4 mostra a evolução da taxa de suicídio por raça/cor. As pessoas de cor branca apresentam consistentemente as maiores taxas de suicídio ao longo

do período, subindo de 2.5 para mais de 3.5 por 100.000 habitantes em 2022, com um crescimento mais acentuado após 2017. A taxa de suicídio entre pessoas pardas também aumenta significativamente, passando de 1.5 para quase 3.0 por 100.000 habitantes em 2022. As pessoas pretas têm taxas mais baixas, mas apresentam um crescimento contínuo, subindo de 0.5 para cerca de 1.0 por 100.000 habitantes. As categorias amarela, indígena e ignorado mantêm taxas estáveis e relativamente baixas ao longo do tempo.

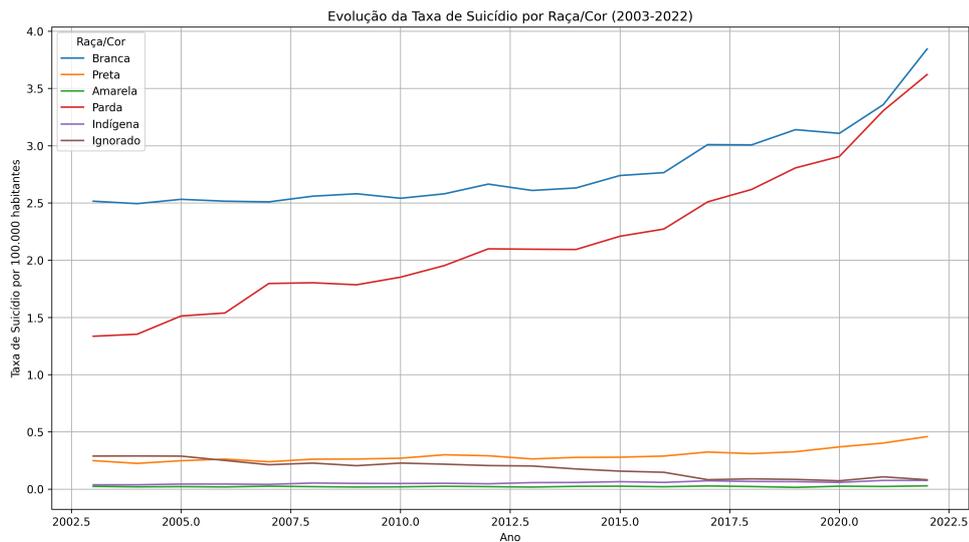


Figura 5.4: Evolução da Taxa de Suicídio por Raça

O gráfico da Figura 5.5 mostra a evolução da taxa de suicídio por tempo de estudo. A categoria de indivíduos com 8 a 11 anos de estudo apresenta o maior aumento ao longo do tempo, especialmente após 2015, com uma aceleração notável após 2020, ultrapassando 2 por 100.000 habitantes em 2022. Já a faixa com 4 a 7 anos de estudo também mostra um crescimento contínuo, seguido pela categoria nenhum estudo, que apresenta taxas mais baixas e relativamente estáveis. As faixas de 1 a 3 anos e 12 ou mais anos de estudo mantêm taxas mais estáveis e menores ao longo do período. Este gráfico indica grupos intermediários de escolaridade apresentando maior risco.

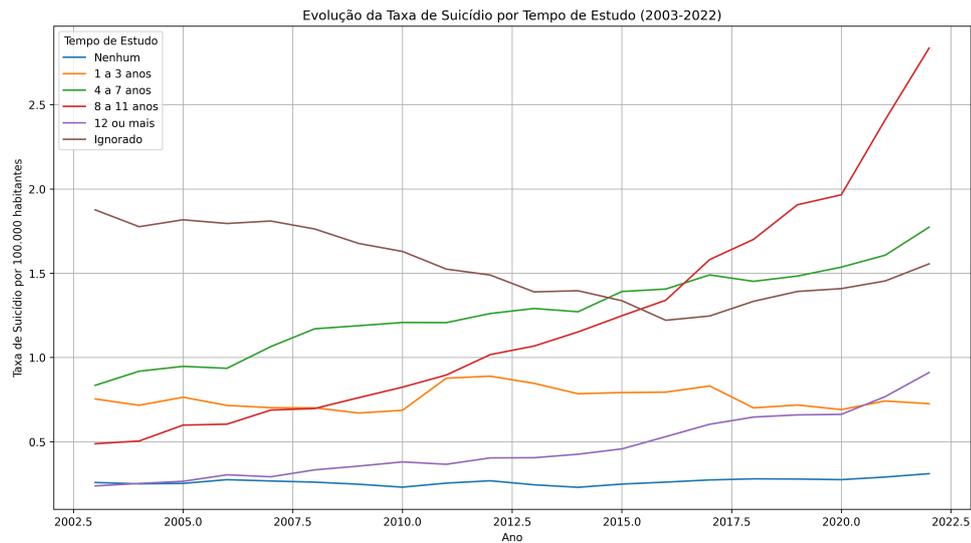


Figura 5.5: Evolução da Taxa de Suicídio por Tempo de Estudo

No gráfico da Figura 5.7 que mostra a evolução da taxa de suicídio por município, destaca-se que São Paulo apresentou as taxas mais altas durante a maior parte do período, com um crescimento acentuado até 2017 e uma queda drástica nos anos seguintes. Outros municípios, como Brasília e Belo Horizonte, mantêm taxas mais estáveis, com leve aumento após 2020. De forma geral, as cidades analisadas apresentam flutuações modestas nas taxas de suicídio, com algumas variações mais notáveis após 2015. No gráfico que apresenta a taxa de suicídio por estado (Figura 5.6), observamos que São Paulo mantém consistentemente a maior taxa de suicídio ao longo do período, com um aumento constante que se intensifica após 2020, chegando a 1.5.

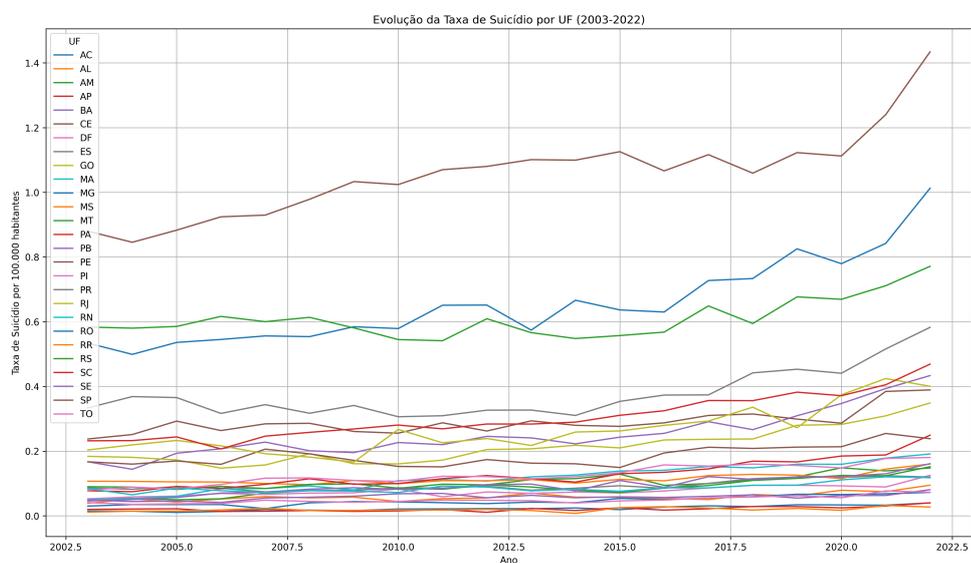


Figura 5.6: Evolução da Taxa de Suicídio por UF

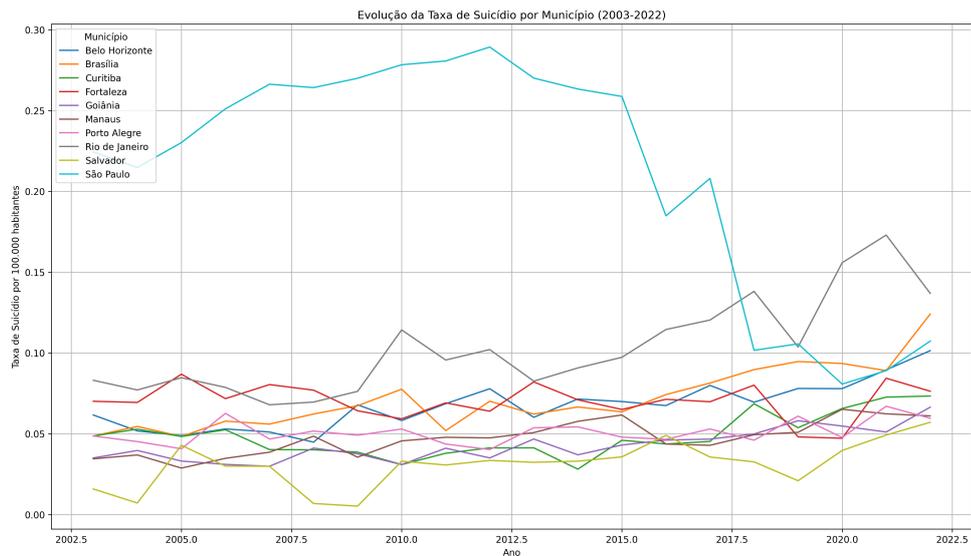


Figura 5.7: Evolução da Taxa de Suicídio por Município

O gráfico da Figura 5.8 mostra a evolução da taxa de suicídio por causa básica. A causa X70 (Lesão autoprovocada intencionalmente por enforcamento, estrangulamento e sufocação) é a mais prevalente, com um aumento consistente ao longo do tempo, passando de cerca de 2.5 em 2003 para mais de 5.0 em 2022. As demais causas permanecem com taxas relativamente baixas e estáveis ao longo de todo o período, sem variações significativas. O gráfico destaca que a causa X70 é a principal responsável pelo aumento geral das taxas de suicídio, com crescimento particularmente acelerado nos últimos anos, sugerindo que este método de suicídio merece atenção especial em políticas de prevenção e saúde pública.

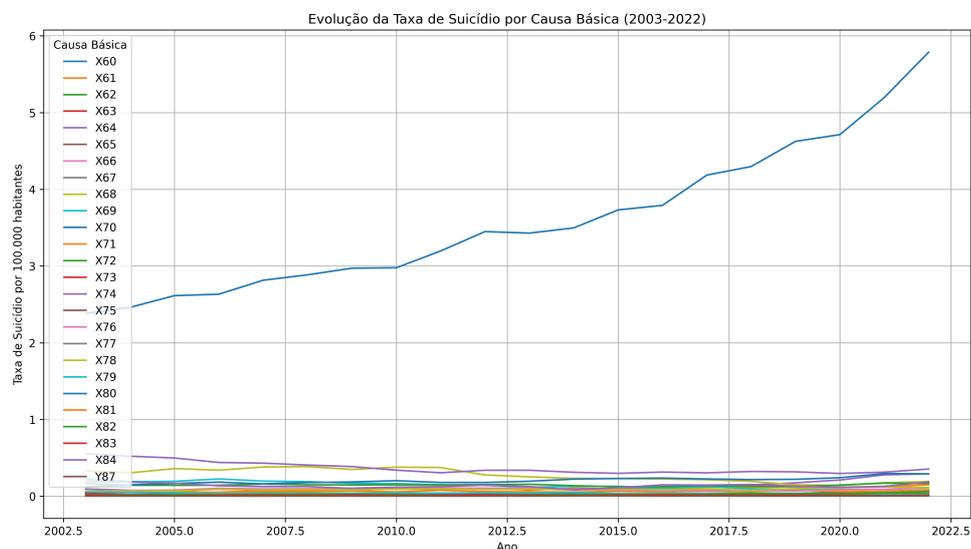


Figura 5.8: Evolução da Taxa de Suicídio por Causa Básica

A Figura 5.9 é um gráfico que mostra a evolução da taxa de suicídio por sexo e causa básica, separando os casos de envenenamento (auto-intoxicação) e lesão autoprovocada. O destaque é para a taxa de lesão autoprovocada entre os homens, que começa em torno de 3.0 em 2003 e cresce consistentemente, ultrapassando 5.0 em 2022. Este é o grupo com o aumento mais acentuado, especialmente após 2015. As mulheres com lesão autoprovocada apresentam uma taxa bem menor durante todo o período, embora também tenha havido um leve aumento nos últimos anos. Os casos de auto-intoxicação, tanto para homens quanto para mulheres, mostram uma estabilidade maior ao longo do tempo. A análise revela uma diferença marcante entre os métodos empregados de suicídio, com lesão autoprovocada sendo muito mais prevalente entre os homens, indicando uma tendência de uso de métodos mais letais. As mulheres, por outro lado, apresentam taxas mais equilibradas entre os dois métodos, mas em níveis gerais mais baixos.

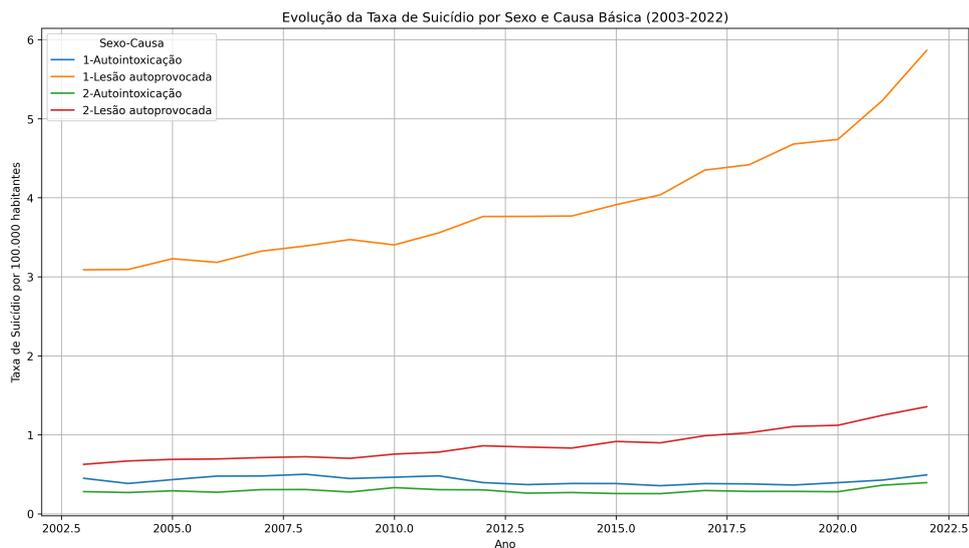


Figura 5.9: Evolução da Taxa de Suicídio por Sexo e Causa Básica

5.2 Análise de Séries Temporais

A Figura 5.10 é uma decomposição dos componentes do modelo Prophet utilizado. Este gráfico decompõe a série temporal em duas componentes principais, tendência e sazonalidade anual. A linha azul no gráfico superior representa a tendência subjacente dos dados ao longo do tempo, e a área sombreada ao redor indica o intervalo de confiança (95%). A tendência geral é claramente de crescimento contínuo e suave do número de

suicídios ao longo do período analisado (desde 2003 até 2022). Nota-se que o crescimento é praticamente linear até cerca de 2016, quando a inclinação da curva aumenta levemente, sugerindo um aumento mais acelerado no número de suicídios nos últimos anos do período. Isso pode ser interpretado como um possível reflexo de mudanças sociais ou econômicas, que contribuíram para esse aumento.

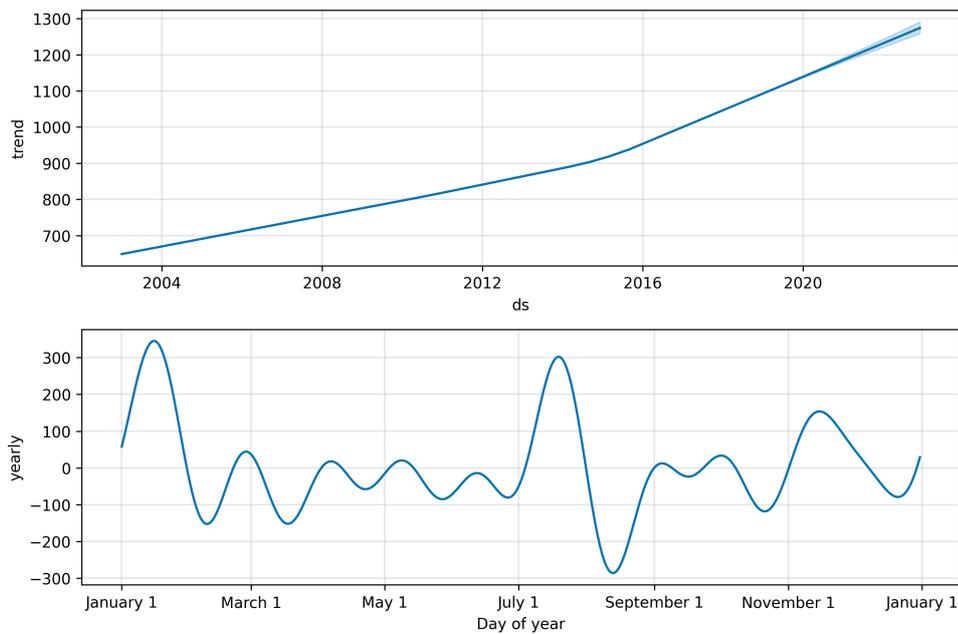


Figura 5.10: Componentes do Modelo Prophet

No final da série, após 2020, a tendência continua subindo, sugerindo que o modelo espera um crescimento contínuo no futuro, o que pode estar associado ao impacto de fatores externos, como a pandemia de COVID-19. O gráfico inferior representa a componente sazonal anual da série temporal. Ele mostra como o número de suicídios varia ao longo do ano, capturando padrões que se repetem em determinados meses ou estações. Observa-se que há picos claros no início do ano, particularmente em janeiro, onde a sazonalidade atinge seu valor mais alto. Isso pode indicar um aumento sazonal no número de suicídios após o período de festas de fim de ano. Há também um segundo pico notável por volta de julho, seguido por um vale acentuado em setembro, sugerindo que esses meses também têm um padrão sazonal repetido de aumento e diminuição dos casos de suicídio.

Em geral, os meses entre março e junho apresentam uma leve variação, sem grandes picos ou quedas, sugerindo que os valores tendem a se estabilizar nesse período. A tendência de crescimento contínuo reflete um aumento gradual no número de suicídios

ao longo do tempo. Este padrão pode estar relacionado a fatores de longo prazo, como mudanças econômicas, sociais e de saúde pública. A componente sazonal sugere que o número de suicídios segue um ciclo anual claro, com aumentos mais pronunciados no início e no meio do ano. Esses picos podem estar associados a fatores sazonais, como estresse após o período de festas e mudanças nos ciclos sociais e econômicos ao longo do ano.

A Tabela 5.1 compara os modelos de previsão Prophet, SARIMA e LSTM em termos de várias métricas de erro e testes estatísticos. Com base nas informações sobre o teste t e o teste de Shapiro-Wilk dos resíduos, podemos interpretar de forma mais completa o desempenho desses modelos. O modelo SARIMA obteve os menores valores de MAE (39.76), MSE (2173.45) e RMSE (46.62), sugerindo que ele teve um desempenho superior em termos de precisão absoluta e quadrática. Isso significa que, em média, as previsões do SARIMA são mais próximas dos valores reais do que as previsões do Prophet e do LSTM.

Tabela 5.1: Comparação de Métricas de Erro e Estatísticas

Métrica	Prophet	SARIMA	LSTM
MAE	44.64	39.76	45.65
MSE	2571.27	2173.45	3157.20
RMSE	50.71	46.62	56.19
MAPE	3.96%	3.49%	3.98%
Estatística t	1.04	2.64	1.99
p-valor (t-test)	0.32	0.02	0.07
Estatística Shapiro	0.93	0.96	0.97
p-valor (Shapiro)	0.34	0.78	0.93

O LSTM apresentou o maior MSE (3157.20) e RMSE (56.19), indicando maior variação nos erros de previsão. No entanto, sua MAE (45.65) e MAPE (3.98%) são semelhantes às do Prophet, o que sugere que o LSTM cometeu mais erros grandes, enquanto o Prophet errou de forma mais consistente, mas com menor impacto. O Prophet, embora não tenha o melhor desempenho em comparação com SARIMA, ainda apresentou métricas bastante próximas, indicando que é uma alternativa razoável com MAPE de 3.96%.

O SARIMA apresentou um p-valor de 0.02 no teste-t, indicando que há uma diferença significativa entre os resíduos e zero, sugerindo que o modelo fez previsões significativamente diferentes dos valores observados, mas em um sentido que pode ser considerado

adequado dada a baixa magnitude dos erros. O Prophet e o LSTM têm valores p de 0.32 e 0.07, respectivamente, o que implica que falhamos em rejeitar a hipótese nula. Isso sugere que, para esses modelos, as previsões não são significativamente diferentes dos valores observados. Portanto, os erros podem estar mais próximos de zero, o que, no contexto, é uma vantagem.

O teste de Shapiro-Wilk nos mostra que, para todos os modelos, os resíduos seguem uma distribuição normal, dado que todos os valores p são superiores a 0.05 (0.34 para Prophet, 0.78 para SARIMA e 0.93 para LSTM). Isso é um indicador de que os erros são bem comportados e que a distribuição dos resíduos segue o padrão esperado para um modelo bem ajustado. SARIMA se destaca como o melhor modelo em termos de precisão (menores valores de MAE, MSE e RMSE) e é estatisticamente significativo, embora seus resíduos sejam diferentes de zero.

Prophet e LSTM, embora tenham desempenho ligeiramente inferior, apresentam resíduos que não são significativamente diferentes de zero, o que indica previsões bastante estáveis. Em termos de normalidade dos resíduos, todos os modelos têm bom desempenho, o que sugere que os erros são distribuídos de maneira apropriada, reforçando a qualidade das previsões. Portanto, SARIMA se mostra superior em precisão, enquanto Prophet e LSTM têm uma estabilidade maior em termos de resíduos não significativos.

A Figura 5.11 é um gráfico comparativo das previsões feitas pelos modelos SARIMA, Prophet, e LSTM contra os valores reais de suicídios ao longo do tempo, especificamente de 2019 a 2022. O gráfico fornece visões importantes sobre o comportamento dos suicídios durante e após o início da pandemia de COVID-19, que começou em 2020. A partir do início de 2020, observa-se um aumento nos valores reais de suicídios, com picos acentuados em diferentes momentos. Isso pode estar relacionado aos impactos psicológicos e sociais causados pela pandemia, como isolamento social, incertezas econômicas e aumento de problemas de saúde mental. Esses picos sugerem que a pandemia exacerbou os fatores de risco associados ao suicídio, levando a um aumento abrupto dos casos em determinados períodos.

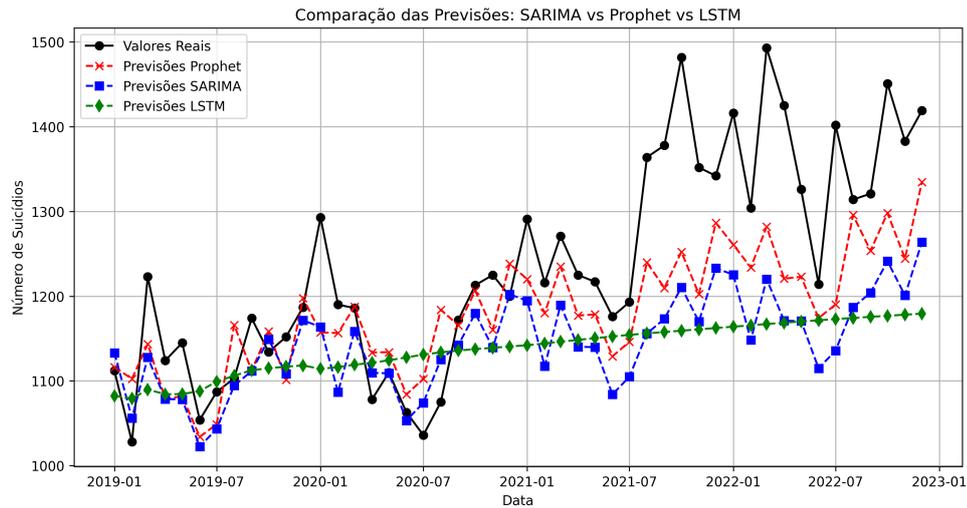


Figura 5.11: Comparação das Previsões SARIMA vs Prophet vs LSTM

Nenhum dos modelos preveu corretamente o aumento acentuado nos números de suicídios durante a pandemia. Embora o SARIMA se aproxime dos valores reais em alguns momentos, ele ainda subestima os picos que ocorrem durante 2020 e 2021. Prophet e LSTM também mostram dificuldades em prever os picos, e o LSTM, em particular, apresenta previsões mais suavizadas, falhando em capturar a volatilidade dos dados durante a pandemia. Durante 2020 e 2021, os dados reais exibem mais variações bruscas em comparação com os períodos anteriores, indicando que o comportamento dos suicídios durante a pandemia foi muito mais imprevisível.

Isso pode ser um reflexo das condições adversas e instáveis impostas pela pandemia, o que dificulta a previsão precisa por meio de modelos estatísticos tradicionais e até mesmo por redes neurais. No ano de 2022, os valores reais de suicídios parecem continuar elevados, e embora haja uma leve queda em alguns meses, o padrão de picos altos e frequentes continua. Isso sugere que os impactos da pandemia podem ter efeitos prolongados, mesmo após a fase crítica da pandemia. Isso pode ser atribuído a efeitos residuais, como o aumento de problemas de saúde mental e o esgotamento das redes de suporte social e econômico.

Após 2020, o crescimento populacional do Brasil desacelerou. Entre 2010 e 2022, a população brasileira cresceu apenas 6,5%, chegando a 203,1 milhões de habitantes em 2022, com uma taxa de crescimento anual média de 0,52%—a menor taxa já registrada na história do país ((IBGE), 2022). Portanto, o aumento das observações de suicídios após

a pandemia dificilmente pode ser explicado apenas pelo crescimento populacional, que foi bastante modesto nesse período.

Após 2021, o SARIMA e o Prophet conseguem capturar a tendência geral de alta nos números de suicídios, mas, novamente, falham em prever com precisão os picos mais acentuados. Esses picos podem refletir crises momentâneas ou eventos específicos, difíceis de serem capturados por modelos baseados apenas em dados históricos. O LSTM mantém sua tendência de previsões suavizadas, continuando a subestimar as flutuações reais, sugerindo que ele não foi capaz de se ajustar à mudança no comportamento dos dados após a pandemia.

Os dados reais indicam que a pandemia de COVID-19 pode ter tido um impacto no aumento das taxas de suicídio, levando a picos mais altos e uma maior volatilidade. Nenhum dos modelos (Prophet, SARIMA, ou LSTM) foi capaz de prever com precisão os aumentos bruscos durante a pandemia, refletindo a complexidade e o impacto inesperado que esse evento global teve na saúde mental. Mesmo após a fase crítica da pandemia, os dados sugerem que os efeitos prolongados da crise continuam a influenciar os números de suicídios, resultando em uma tendência geral de alta. Isso ressalta a necessidade de mais estudos e modelos que levem em consideração os efeitos prolongados da pandemia e suas influências em crises de saúde mental.

5.3 Modelagem de Classificação Binária

Na Tabela 5.2, observamos que tanto o XGBoost quanto o *Random Forest* identificam a variável `def_loc_ocor` como a mais importante, com uma importância muito alta no XGBoost (0.7007) e consideravelmente menor no *Random Forest* (0.2905). Isso sugere que a localização do óbito é um fator determinante na modelagem dos dados, especialmente no XGBoost, onde ela domina o modelo.

A variável `SEXO` também é considerada bastante relevante no XGBoost (0.1573), mas significativamente menos importante no *Random Forest* (0.0481), indicando que o XGBoost dá mais peso à variável `sexo` na classificação. Por outro lado, variáveis como `idade_obito_anos` (0.2342) e `res_MUNNOME` (0.2502) são muito mais importantes no *Random Forest*, o que pode sugerir que este modelo está dando maior importância à faixa

etária e à localidade de residência para a previsão.

Esses resultados mostram que cada modelo dá mais relevância a diferentes variáveis, refletindo as diferenças na forma como *Random Forest* e XGBoost constroem suas árvores de decisão. O XGBoost parece priorizar as variáveis mais diretamente ligadas ao evento de óbito e ao sexo, enquanto o *Random Forest* distribui a importância de maneira mais equilibrada entre variáveis relacionadas à idade e localidade.

Tabela 5.2: Comparação da Importância das Variáveis com Desvio Padrão

Variável	XGBoost (Média ± DP)	Random Forest (Média ± DP)
def_loc_ocor	0.7007 ± 0.0139	0.2905 ± 0.0037
SEXO	0.1573 ± 0.0122	0.0481 ± 0.0018
ocor_SIGLA_UF	0.0387 ± 0.0020	0.0933 ± 0.0017
ESC	0.0280 ± 0.0017	0.0545 ± 0.0012
idade_obito_anos	0.0258 ± 0.0013	0.2342 ± 0.0026
res_MUNNOME	0.0258 ± 0.0012	0.2502 ± 0.0017
RACACOR	0.0235 ± 0.0017	0.0292 ± 0.0009

Na Tabela 5.3, o XGBoost apresenta métricas ligeiramente melhores que o *Random Forest* em todas as áreas. A acurácia do XGBoost (0.8118) é superior à do *Random Forest* (0.7686), indicando que o XGBoost realiza um número maior de previsões corretas no geral. A precisão (0.9533 vs. 0.9500) e o *recall* (0.8245 vs. 0.7752) também são superiores no XGBoost, o que significa que este modelo é ligeiramente melhor em identificar corretamente os positivos reais e em evitar falsos positivos.

O F1-Score (0.8842 para XGBoost e 0.8537 para Random Forest) reflete a melhor harmonia entre precisão e *recall* no XGBoost, indicando que ele tem um desempenho mais equilibrado em termos de classificação correta. Finalmente, o ROC-AUC (0.8363 para XGBoost e 0.8092 para Random Forest) sugere que o XGBoost tem uma capacidade ligeiramente maior de discriminar entre as classes, o que pode ser decisivo em problemas de classificação.

Tabela 5.3: Comparação de Desempenho com Desvio Padrão

Métrica	Random Forest (Média ± DP)	XGBoost (Média ± DP)
Acurácia	0.7686 ± 0.0022	0.8118 ± 0.0024
Precisão	0.9500 ± 0.0009	0.9533 ± 0.0008
Recall	0.7752 ± 0.0027	0.8245 ± 0.0031
F1-Score	0.8537 ± 0.0016	0.8842 ± 0.0017
ROC-AUC	0.8092 ± 0.0022	0.8363 ± 0.0020

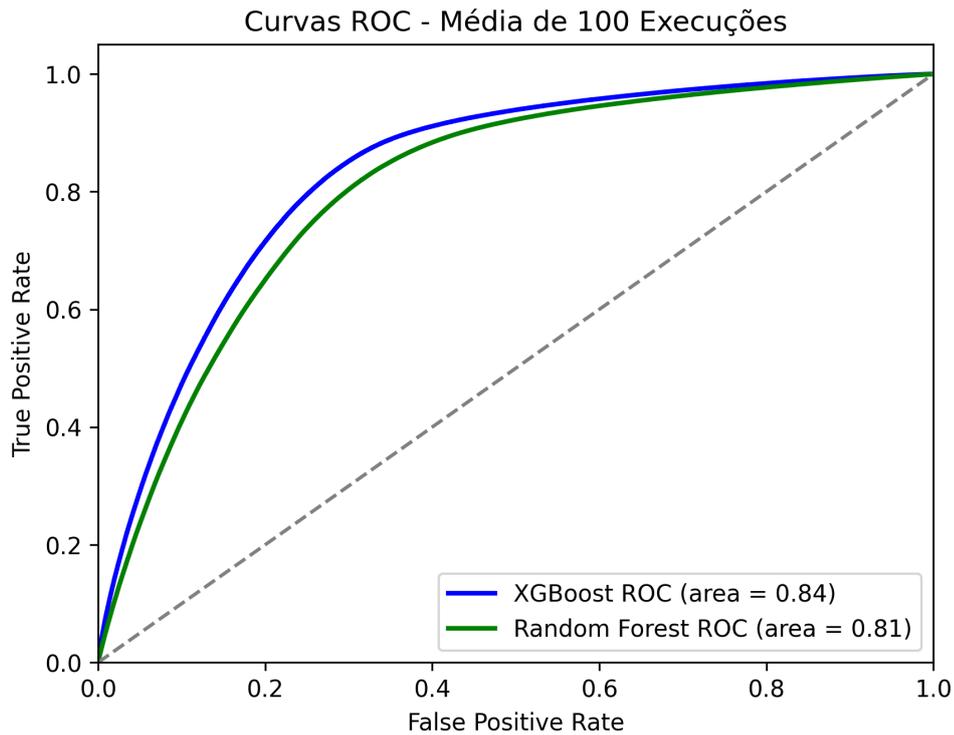


Figura 5.12: Curva ROC - Random Forest vs XGBoost

Em termos de desempenho, o XGBoost se sai um pouco melhor em todas as métricas comparadas ao Random Forest, o que o torna uma escolha superior para esse conjunto de dados. No entanto, a forma como os dois modelos tratam as variáveis também é interessante: enquanto o XGBoost dá muito mais importância a variáveis específicas como `def_loc_ocor` e `SEXO`, o *Random Forest* distribui a importância de forma mais uniforme, especialmente em variáveis como `idade_obito_anos` e `res_MUNNOME`. Isso indica que o XGBoost está mais focado em certas características preditivas, enquanto o *Random Forest* tem uma visão mais diversificada dos fatores influentes.

6 Conclusão e Trabalhos Futuros

Este trabalho realizou uma análise das tendências de suicídio no Brasil entre os anos de 2003 e 2022, com foco especial no impacto da pandemia de COVID-19 e na identificação de fatores associados aos métodos de suicídio. Utilizando dados do SIM/DATASUS e aplicando técnicas de análise estatística e aprendizado de máquina, foram obtidos padrões que contribuem para a compreensão deste problema no cenário brasileiro.

Observou-se um aumento gradual nas taxas de suicídio ao longo do período estudado, com uma concentração maior entre homens e indivíduos brancos na faixa etária de 30 a 49 anos. A análise geográfica identificou estados específicos, como São Paulo, com incidências mais elevadas. A causa lesão autoprovocada intencionalmente por enforcamento, estrangulamento e sufocação (X70) é a mais prevalente.

A pandemia parece ter exacerbado as tendências de suicídio, com desvios entre as previsões baseadas em dados históricos e os valores observados durante 2020 a 2022. Este aumento sugere que os efeitos adversos da pandemia na saúde mental, incluindo isolamento social, insegurança econômica e acesso limitado a serviços de saúde, contribuíram para o incremento nas taxas de suicídio. Além disso, o aumento das observações de suicídios após a pandemia dificilmente pode ser explicado apenas pelo crescimento populacional, pois foi sutil nesse período. A análise de classificação binária revelou que variáveis como idade, sexo, município de residência e local de ocorrência estão associadas à escolha do método de suicídio. Mulheres tendem a utilizar mais a auto-intoxicação, enquanto homens são mais propensos a utilizar lesões autoprovocadas.

Há a necessidade de ampliar o acesso a serviços de saúde mental, especialmente em períodos de crise como a pandemia. Programas de apoio psicológico e intervenções precoces podem mitigar os riscos associados ao suicídio. Estratégias específicas devem ser desenvolvidas para os grupos mais vulneráveis identificados, como homens adultos e residentes em regiões com altas taxas de suicídio. Campanhas de conscientização e programas comunitários podem ser eficazes nesses contextos. Políticas que visem restringir o acesso a meios comuns de suicídio, como certos medicamentos ou armas, podem contribuir para

a redução das taxas, especialmente entre os grupos identificados como mais propensos a utilizar esses métodos.

O suicídio é um fenômeno complexo que requer abordagens multidisciplinares e ações coordenadas entre governos, profissionais de saúde e sociedade civil. Este trabalho contribui para a compreensão das tendências e fatores associados ao suicídio no Brasil, destacando o impacto da pandemia de COVID-19. Espera-se que as perspectivas aqui apresentados possam orientar políticas públicas mais eficazes e sensibilizar a sociedade para a importância de abordar a saúde mental com a seriedade e urgência que o tema pede.

A dependência de registros oficiais sujeita o presente trabalho a subnotificações ou inconsistências. As conclusões refletem tendências gerais e podem não capturar todas as nuances regionais ou culturais presentes em todo o território brasileiro. Embora modelos avançados tenham sido utilizados, nenhum modelo pode capturar completamente a complexidade do comportamento humano, e os resultados devem ser interpretados dentro das limitações metodológicas. A inclusão de variáveis como desemprego, renda, acesso a serviços de saúde e histórico de transtornos mentais pode enriquecer as análises. Estudos focados em regiões específicas podem identificar fatores locais que influenciam as tendências de suicídio. A taxa seria melhor calculada se ajustada por idade, pois a estrutura etária da população pode mudar ao longo do tempo, e diferentes faixas etárias têm riscos de suicídio distintos.

Bibliografia

- AIJAZ, I.; AGARWAL, P. A study on time series forecasting using hybridization of time series models and neural networks. *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science)*, Bentham Science Publishers, v. 13, n. 5, p. 827–832, 2020.
- ALI, H.; REHNA, T. The psychology of suicide: From research understanding to intervention and treatment. *JPMA. The Journal of the Pakistan Medical Association*, v. 72, n. 6, p. 1175–1178, 2022.
- ARAGHINEJAD, S. *Time Series Modeling*. [S.l.]: Springer, 2014.
- ARENSMAN, E. et al. Suicide and suicide prevention from a global perspective. *Crisis*, Hogrefe Publishing, 2020.
- AVIV, Y. A time-series framework for supply-chain inventory management. *Operations Research*, INFORMS, v. 51, n. 2, p. 210–227, 2003.
- BAI, Y. et al. How important is the train-validation split in meta-learning? In: PMLR. *International Conference on Machine Learning*. [S.l.], 2021. p. 543–553.
- BARBOSA, B. de A.; TEIXEIRA, F. A. F. de C. Perfil epidemiológico e psicossocial do suicídio no brasil. *Research, society and development*, v. 10, n. 5, p. e32410515097–e32410515097, 2021.
- BERGMEIR, C.; BENÍTEZ, J. M. On the use of cross-validation for time series predictor evaluation. *Information Sciences*, Elsevier, v. 191, p. 192–213, 2012.
- BIAU, G.; SCORNET, E. A random forest guided tour. *Test*, Springer, v. 25, p. 197–227, 2016.
- BLÁZQUEZ-GARCÍA, A. et al. Short-term office building elevator energy consumption forecast using sarima. *Journal of Building Performance Simulation*, Taylor & Francis, v. 13, n. 1, p. 69–78, 2020.
- BOAVENTURA, S. F. et al. Epigenética e vulnerabilidade ao suicídio: como experiências traumáticas na infância afetam a expressão de genes relacionados à saúde mental. Universidade Estadual de Goiás, 2023.
- BORGES, D.; NASCIMENTO, M. C. Covid-19 icu demand forecasting: A two-stage prophet-lstm approach. *Applied soft computing*, Elsevier, v. 125, p. 109181, 2022.
- BOSI, M. L. M.; ALVES, E. D. Distanciamento social em contextos urbanos na pandemia de covid-19: desafios para o campo da saúde mental. *Physis: Revista de Saúde Coletiva*, SciELO Brasil, v. 33, p. e33007, 2023.
- BOX, G. E. et al. *Time series analysis: forecasting and control*. [S.l.]: John Wiley & Sons, 2015.

- BRAMER, M. Avoiding overfitting of decision trees. *Principles of data mining*, Springer, p. 119–134, 2007.
- BRAUN, B. F. et al. Perfil epidemiológico dos casos de tentativa de suicídio: revisão integrativa. *SMAD, Revista Eletrônica Saúde Mental Álcool e Drogas (Edição em Português)*, v. 19, n. 1, p. 112–122, 2023.
- BROECK, G. Van den et al. On the tractability of shap explanations. *Journal of Artificial Intelligence Research*, v. 74, p. 851–886, 2022.
- BROWNLEE, J. *XGBoost With python: Gradient boosted trees with XGBoost and scikit-learn*. [S.l.]: Machine Learning Mastery, 2016.
- BUONOCORE, R. J.; ASTE, T.; MATTEO, T. D. Measuring multiscaling in financial time-series. *Chaos, Solitons & Fractals*, Elsevier, v. 88, p. 38–47, 2016.
- CAI, C. et al. Human motion pattern recognition based on the fused random forest algorithm. *Measurement*, Elsevier, v. 222, p. 113540, 2023.
- CANTOR, C. H.; BAUME, P. J. Access to methods of suicide: what impact? *Australian and New Zealand Journal of Psychiatry*, Taylor & Francis, v. 32, n. 1, p. 8–14, 1998.
- CATERINI, A. L. et al. Recurrent neural networks. *Deep neural networks in a mathematical framework*, Springer, p. 59–79, 2018.
- CELEBI, M. E.; AYDIN, K. *Unsupervised learning algorithms*. [S.l.]: Springer, 2016. v. 9.
- CHAN, N. H. *Time series: applications to finance*. [S.l.]: John Wiley & Sons, 2004.
- CHATFIELD, C. *Time-series forecasting*. [S.l.]: Chapman and Hall/CRC, 2000.
- CHEN, K.; HUO, Q. Training deep bidirectional lstm acoustic model for lvcsr by a context-sensitive-chunk bptt approach. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, IEEE, v. 24, n. 7, p. 1185–1193, 2016.
- CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. [S.l.: s.n.], 2016. p. 785–794.
- CHEN, T. et al. Package ‘xgboost’. *R version*, The R Foundation Vienna, Austria, v. 90, n. 1-66, p. 40, 2019.
- CHOI, T.-M.; YU, Y.; AU, K.-F. A hybrid sarima wavelet transform method for sales forecasting. *Decision Support Systems*, Elsevier, v. 51, n. 1, p. 130–140, 2011.
- COFFER, B. et al. Machine learning with feature importance analysis for tornado prediction from environmental sounding data. *PAMM*, Wiley Online Library, v. 20, n. 1, p. e202000112, 2021.
- CRYER, J. D. *Time series analysis*. [S.l.]: Duxbury Press Boston, 1986. v. 286.
- CUNNINGHAM, P.; CORD, M.; DELANY, S. J. Supervised learning. In: *Machine learning techniques for multimedia: case studies on organization and retrieval*. [S.l.]: Springer, 2008. p. 21–49.

- DAI, B. et al. Using random forest algorithm for breast cancer diagnosis. In: IEEE. *2018 International symposium on computer, consumer and control (IS3C)*. [S.l.], 2018. p. 449–452.
- DARAGHMEH, M. et al. Time series forecasting using facebook prophet for cloud resource management. In: IEEE. *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*. [S.l.], 2021. p. 1–6.
- DAVIS, J.; GOADRICH, M. The relationship between precision-recall and roc curves. In: *Proceedings of the 23rd international conference on Machine learning*. [S.l.: s.n.], 2006. p. 233–240.
- DENG, H.; RUNGER, G.; TUV, E. Bias of importance measures for multi-valued attributes and solutions. In: SPRINGER. *Artificial Neural Networks and Machine Learning–ICANN 2011: 21st International Conference on Artificial Neural Networks, Espoo, Finland, June 14–17, 2011, Proceedings, Part II 21*. [S.l.], 2011. p. 293–300.
- DENNING, D. G. et al. Method choice, intent, and gender in completed suicide. *Suicide and Life-Threatening Behavior*, Wiley Online Library, v. 30, n. 3, p. 282–288, 2000.
- DOUPNIK, S. K. et al. Association of suicide prevention interventions with subsequent suicide attempts, linkage to follow-up care, and depression symptoms for acute care settings: a systematic review and meta-analysis. *JAMA psychiatry*, American Medical Association, v. 77, n. 10, p. 1021–1030, 2020.
- DUBEY, A. K. et al. Study and analysis of sarima and lstm in forecasting time series data. *Sustainable Energy Technologies and Assessments*, Elsevier, v. 47, p. 101474, 2021.
- EVANS, A.; ABRAHAMSON, K. The influence of stigma on suicide bereavement: A systematic review. *Journal of psychosocial nursing and mental health services*, SLACK Incorporated Thorofare, NJ, v. 58, n. 4, p. 21–27, 2020.
- FADZISO, T. Overcoming the vanishing gradient problem during learning recurrent neural nets (rnn). *Asian Journal of Applied Science and Engineering*, v. 9, n. 1, p. 197–208, 2020.
- FARSANI, R. M.; PAZOUKI, E. A transformer self-attention model for time series forecasting. *Journal of Electrical and Computer Engineering Innovations (JECEI)*, Shahid Rajaei Teacher Training University, v. 9, n. 1, p. 1–10, 2020.
- FENG, T. et al. The comparative analysis of sarima, facebook prophet, and lstm for road traffic injury prediction in northeast china. *Frontiers in public health*, Frontiers Media SA, v. 10, p. 946563, 2022.
- FORTGANG, R. G.; NOCK, M. K. Ringing the alarm on suicide prevention: A call to action. *Psychiatry*, Taylor & Francis, v. 84, n. 2, p. 192–195, 2021.
- GABER, M. M.; ATWAL, H. S. An entropy-based approach to enhancing random forests. *Intelligent Decision Technologies*, IOS Press, v. 7, n. 4, p. 319–327, 2013.
- GIKUNGU, S. W.; WAITITU, A. G.; KIHORO, J. M. Forecasting inflation rate in kenya using sarima model. *American Journal of Theoretical and Applied Statistics*, Science Publishing Group, v. 12, n. 3, p. 15–18, 2015.

- GONZÁLEZ-ESTRADA, E.; COSMES, W. Shapiro–wilk test for skew normal distributions based on data transformations. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 89, n. 17, p. 3258–3272, 2019.
- GOODFELLOW, I. *Deep learning*. [S.l.]: MIT press, 2016.
- GRAVES, A.; GRAVES, A. Long short-term memory. *Supervised sequence labelling with recurrent neural networks*, Springer, p. 37–45, 2012.
- GREGORUTTI, B.; MICHEL, B.; SAINT-PIERRE, P. Correlation and variable importance in random forests. *Statistics and Computing*, Springer, v. 27, p. 659–678, 2017.
- HAJEK, P.; ABEDIN, M. Z.; SIVARAJAH, U. Fraud detection in mobile payment systems using an xgboost-based framework. *Information Systems Frontiers*, Springer, v. 25, n. 5, p. 1985–2003, 2023.
- HAMILTON, J. D. *Time series analysis*. [S.l.]: Princeton university press, 2020.
- HASTIE, T. et al. Overview of supervised learning. *The elements of statistical learning: Data mining, inference, and prediction*, Springer, p. 9–41, 2009.
- HASTIE, T. et al. Unsupervised learning. *The elements of statistical learning: Data mining, inference, and prediction*, Springer, p. 485–585, 2009.
- HEWAMALAGE, H.; BERGMEIR, C.; BANDARA, K. Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, Elsevier, v. 37, n. 1, p. 388–427, 2021.
- HOCHREITER, S. Long short-term memory. *Neural Computation MIT-Press*, 1997.
- HOO, Z. H.; CANDLISH, J.; TEARE, D. *What is an ROC curve?* [S.l.]: BMJ Publishing Group Ltd and the British Association for Accident . . . , 2017. 357–359 p.
- HOOKER, S. et al. Evaluating feature importance estimates. *arXiv preprint arXiv:1806.10758*, v. 2, 2018.
- HYNDMAN, R. *Forecasting: principles and practice*. [S.l.]: OTexts, 2018.
- (IBGE), I. B. de Geografia e E. *De 2010 a 2022, população brasileira cresce 6,5% e chega a 203,1 milhões*. 2022. Acessado em setembro de 2024. Disponível em: (<https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/37237-de-2010-a-2022-populacao-brasileira-cresce-6-5-e-chega-a-203-1-milhoes>).
- Instituto Brasileiro de Geografia e Estatística (IBGE). *IBGE: Instituto Brasileiro de Geografia e Estatística*. 2024. Acessado em setembro de 2024. Disponível em: (<https://www.ibge.gov.br/>).
- JELODAR, H. et al. Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions: Nlp using lstm recurrent neural network approach. *IEEE Journal of Biomedical and Health Informatics*, IEEE, v. 24, n. 10, p. 2733–2742, 2020.
- JUNIOR, A. P. d. S. et al. Estratégias para prevenção e posvenção do suicídio em tempos de pandemia de covid-19. *Interface-Comunicação, Saúde, Educação*, SciELO Public Health, v. 27, p. e230181, 2023.

- KACZMARCZYK, P. Microeconomic analysis of telecommunication services market with the use of sarima models. *Dynamic Econometric Models*, v. 17, p. 41–57, 2017.
- KATWAL, S.; SHRESTHA, R.; SHARMA, G. Analysis of website traffic time series forecasting using arima, prophet, and lstm rnn. *Analysis of Website Traffic Time Series Forecasting using ARIMA, Prophet, and LSTM RNN*, v. 146, n. 1, p. 11–11, 2024.
- KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: MONTREAL, CANADA. *Ijcai*. [S.l.], 1995. v. 14, n. 2, p. 1137–1145.
- LANDI, F. et al. Working memory connections for lstm. *Neural Networks*, Elsevier, v. 144, p. 334–341, 2021.
- LI, P.; ABDEL-ATY, M.; YUAN, J. Real-time crash risk prediction on arterials based on lstm-cnn. *Accident Analysis & Prevention*, Elsevier, v. 135, p. 105371, 2020.
- LI, W. et al. Gene expression value prediction based on xgboost algorithm. *Frontiers in genetics*, Frontiers Media SA, v. 10, p. 1077, 2019.
- LIANG, J. Confusion matrix: Machine learning. *POGIL Activity Clearinghouse*, v. 3, n. 4, 2022.
- LIN, B.; BAI, R. Machine learning approaches for explaining determinants of the debt financing in heavy-polluting enterprises. *Finance Research Letters*, Elsevier, v. 44, p. 102094, 2022.
- LIU, Y.; WANG, Y.; ZHANG, J. New machine learning algorithm: Random forest. In: SPRINGER. *Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14-16, 2012. Proceedings 3*. [S.l.], 2012. p. 246–252.
- LOMNICKI, Z.; ZAREMBA, S. On the estimation of autocorrelation in time series. *The Annals of Mathematical Statistics*, JSTOR, v. 28, n. 1, p. 140–158, 1957.
- LOPES, L. O. R. et al. Fatores de risco e associados ao comportamento suicida no brasil: uma revisão sistemática. *Tempus Psicológico*, v. 6, n. 2, 2023.
- LUNDBERG, S. A unified approach to interpreting model predictions. *arXiv preprint arXiv:1705.07874*, 2017.
- MAATEN, L. V. D. et al. Dimensionality reduction: A comparative review. *Journal of Machine Learning Research*, v. 10, n. 66-71, p. 13, 2009.
- MARZBAN, C. The roc curve and the area under it as performance measures. *Weather and Forecasting*, American Meteorological Society, v. 19, n. 6, p. 1106–1114, 2004.
- MATTHEW, W. et al. Bias of the random forest out-of-bag (oob) error for certain input parameters. *Open Journal of Statistics*, Scientific Research Publishing, v. 2011, 2011.
- MERITY, S.; KESKAR, N. S.; SOCHER, R. Regularizing and optimizing lstm language models. *arXiv preprint arXiv:1708.02182*, 2017.
- MISHINA, Y. et al. Boosted random forest. *IEICE TRANSACTIONS on Information and Systems*, The Institute of Electronics, Information and Communication Engineers, v. 98, n. 9, p. 1630–1636, 2015.

- MUDELSEE, M. Climate time series analysis. *Atmospheric and*, Springer, v. 397, 2010.
- MUSHTAQ, R. Augmented dickey fuller test. 2011. Disponível em SSRN: <https://ssrn.com/abstract=1911068> or <http://dx.doi.org/10.2139/ssrn.1911068>.
- NAIDU, G.; ZUVA, T.; SIBANDA, E. M. A review of evaluation metrics in machine learning algorithms. In: SPRINGER. *Computer Science On-line Conference*. [S.l.], 2023. p. 15–25.
- NALLURI, M.; PENTELA, M.; ELURI, N. R. A scalable tree boosting system: Xg boost. *Int. J. Res. Stud. Sci. Eng. Technol*, v. 7, n. 12, p. 36–51, 2020.
- NAQA, I. E.; MURPHY, M. J. *What is machine learning?* [S.l.]: Springer, 2015.
- NARKHEDE, S. Understanding auc-roc curve. *Towards data science*, v. 26, n. 1, p. 220–227, 2018.
- NAVRATIL, M.; KOLKOVA, A. Decomposition and forecasting time series in the business economy using prophet forecasting model. *Central European Business Review*, University of Economics, Faculty of Business Administration, v. 8, n. 4, p. 26, 2019.
- NIELSEN, D. *Tree boosting with xgboost-why does xgboost win"every"machine learning competition?* Dissertação (Mestrado) — NTNU, 2016.
- NURMI, J.; LOHAN, E. S. Systematic review on machine-learning algorithms used in wearable-based ehealth data analysis. *IEEE Access*, IEEE, v. 9, p. 112221–112235, 2021.
- OEXLE, N.; MAYER, L.; RÜSCH, N. Suicide stigma and suicide prevention. *Der Nervenarzt*, Springer, v. 91, p. 779–784, 2020.
- ORDEN, K. A. V. et al. The interpersonal theory of suicide. *Psychological review*, American Psychological Association, v. 117, n. 2, p. 575, 2010.
- ORUH, J.; VIRIRI, S.; ADEGUN, A. Long short-term memory recurrent neural network for automatic speech recognition. *IEEE Access*, IEEE, v. 10, p. 30069–30079, 2022.
- OSHIRO, T. M.; PEREZ, P. S.; BARANAUSKAS, J. A. How many trees in a random forest? In: SPRINGER. *Machine Learning and Data Mining in Pattern Recognition: 8th International Conference, MLDM 2012, Berlin, Germany, July 13-20, 2012. Proceedings 8*. [S.l.], 2012. p. 154–168.
- PEREIRA, C. A. et al. Using data mining techniques to understand patterns of suicide and reattempt rates in southern brazil. In: *BIOSTEC (2)*. [S.l.: s.n.], 2024. p. 385–392.
- PEREIRA, T. P. et al. Análise de desempenho do teste ljung-box na aplicação do modelo sarima nos dados mensais da demanda de energia elétrica da classe consumidora comercial do rs. *Encontro nacional de modelagem computacional e encontro de ciência e tecnologia de materiais. Recife, Brasil*, p. 154013–24, 2021.
- PRAJWALA, T. A comparative study on decision tree and random forest using r tool. *International journal of advanced research in computer and communication engineering*, v. 4, n. 1, p. 196–199, 2015.

- PUTATUNDA, S.; RAMA, K. A comparative analysis of hyperopt as against other approaches for hyper-parameter optimization of xgboost. In: *Proceedings of the 2018 international conference on signal processing and machine learning*. [S.l.: s.n.], 2018. p. 6–10.
- RAFFERTY, G. *Forecasting Time Series Data with Facebook Prophet: Build, improve, and optimize time series forecasting models using the advanced forecasting tool*. [S.l.]: Packt Publishing Ltd, 2021.
- RAMOSAJ, B.; PAULY, M. Consistent estimation of residual variance with random forest out-of-bag errors. *Statistics & Probability Letters*, Elsevier, v. 151, p. 49–57, 2019.
- REEL, P. S. et al. Using machine learning approaches for multi-omics data analysis: A review. *Biotechnology advances*, Elsevier, v. 49, p. 107739, 2021.
- REIS, J. G. d. L. et al. Os determinantes socioeconômicos do suicídio nos municípios do brasil. Universidade Federal de Alagoas, 2023.
- ROBINSON, R. L. M. et al. Comparison of the predictive performance and interpretability of random forest and linear models on benchmark data sets. *Journal of chemical information and modeling*, ACS Publications, v. 57, n. 8, p. 1773–1792, 2017.
- RUSDAH, D. A.; MURFI, H. Xgboost in handling missing values for life insurance risk prediction. *SN Applied Sciences*, Springer, v. 2, n. 8, p. 1336, 2020.
- RYU, S. et al. Trends in suicide deaths before and after the covid-19 outbreak in korea. *PLoS One*, Public Library of Science San Francisco, CA USA, v. 17, n. 9, p. e0273637, 2022.
- SAIGAL, S.; MEHROTRA, D. Performance comparison of time series data using predictive data mining techniques. *Advances in Information Mining*, v. 4, n. 1, p. 57–66, 2012.
- SAILUSHA, R. et al. Credit card fraud detection using machine learning. In: IEEE. *2020 4th international conference on intelligent computing and control systems (ICICCS)*. [S.l.], 2020. p. 1264–1270.
- SANDRI, M.; ZUCCOLOTTO, P. A bias correction algorithm for the gini variable importance measure in classification trees. *Journal of Computational and Graphical Statistics*, Taylor & Francis, v. 17, n. 3, p. 611–628, 2008.
- SAUDE, S. de Vigilância em. *Boletim Epidemiológico, volume 55, nº 04*. 2024. Acessado em setembro de 2024. Disponível em: (<https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/boletins/epidemiologicos/edicoes/2024/boletim-epidemiologico-volume-55-no-04.pdf>).
- SHARIFANI, K.; AMINI, M. Machine learning and deep learning: A review of methods and applications. *World Information Technology and Engineering Journal*, v. 10, n. 07, p. 3897–3904, 2023.
- SHENG, C.; YU, H. An optimized prediction algorithm based on xgboost. In: IEEE. *2022 International Conference on Networking and Network Applications (NaNA)*. [S.l.], 2022. p. 1–6.

SHEPARD, A.; NAHEED, N. Application of data transformation techniques and train-test split. In: IEEE. *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*. [S.l.], 2021. p. 58–63.

SHER, L.; OQUENDO, M. A. Suicide: an overview for clinicians. *Medical Clinics*, Elsevier, v. 107, n. 1, p. 119–130, 2023.

SHI, Y.; CHARLES, V.; ZHU, J. Bank financial sustainability evaluation: Data envelopment analysis with random forest and shapley additive explanations. *European Journal of Operational Research*, Elsevier, 2024.

SINYOR, M. et al. The effect of economic downturn, financial hardship, unemployment, and relevant government responses on suicide. *The Lancet Public Health*, Elsevier, 2024.

SIRISHA, U. M.; BELAVAGI, M. C.; ATTIGERI, G. Profit prediction using arima, sarima and lstm models in time series forecasting: A comparison. *IEEE Access*, IEEE, v. 10, p. 124715–124727, 2022.

SIVARAMAKRISHNAN, S. et al. Forecasting time series data using arima and facebook prophet models. In: *Big data management in Sensing*. [S.l.]: River Publishers, 2022. p. 47–59.

SOARES, F. C.; STAHNKE, D. N.; LEVANDOWSKI, M. L. Tendência de suicídio no brasil de 2011 a 2020: foco especial na pandemia de covid-19. *Revista Panamericana de Salud Pública*, SciELO Public Health, v. 46, p. e212, 2023.

TAN, J. et al. A critical look at the current train/test split in machine learning. *arXiv preprint arXiv:2106.04525*, 2021.

TANAKA, T.; OKAMOTO, S. Increase in suicide following an initial decline during the covid-19 pandemic in japan. *Nature human behaviour*, Nature Publishing Group UK London, v. 5, n. 2, p. 229–238, 2021.

TAYLOR, S. J.; LETHAM, B. Forecasting at scale. *The American Statistician*, Taylor & Francis, v. 72, n. 1, p. 37–45, 2018.

TSANTEKIDIS, A.; PASSALIS, N.; TEFAS, A. Recurrent neural networks. In: *Deep learning for robot perception and cognition*. [S.l.]: Elsevier, 2022. p. 101–115.

VAGROPOULOS, S. I. et al. Comparison of sarimax, sarima, modified sarima and ann-based models for short-term pv generation forecasting. In: IEEE. *2016 IEEE international energy conference (ENERGYCON)*. [S.l.], 2016. p. 1–6.

WALCZAK, S. Artificial neural networks. In: *Advanced methodologies and technologies in artificial intelligence, computer simulation, and human-computer interaction*. [S.l.]: IGI global, 2019. p. 40–53.

WANG, K. et al. Research on personal credit risk evaluation based on xgboost. *Procedia computer science*, Elsevier, v. 199, p. 1128–1135, 2022.

WHO. Suicide worldwide in 2019: global health estimates. World Health Organization, 2021.

WIERING, M. A.; OTTERLO, M. V. Reinforcement learning. *Adaptation, learning, and optimization*, Springer, v. 12, n. 3, p. 729, 2012.

- WOOLDRIDGE, J. M. Introductory econometrics: a modern approach. *Language*, v. 27, n. 863p, p. 24cm, 2002.
- WU, Y.-c.; FENG, J.-w. Development and application of artificial neural network. *Wireless Personal Communications*, Springer, v. 102, p. 1645–1656, 2018.
- XUAN, S. et al. Random forest for credit card fraud detection. In: IEEE. *2018 IEEE 15th international conference on networking, sensing and control (ICNSC)*. [S.l.], 2018. p. 1–6.
- YAN, L. et al. Climate-informed monthly runoff prediction model using machine learning and feature importance analysis. *Frontiers in Environmental Science*, Frontiers Media SA, v. 10, p. 1049840, 2022.
- YIN, Y.; BINGI, Y. Using machine learning to classify human fetal health and analyze feature importance. *BioMedInformatics*, MDPI, v. 3, n. 2, p. 280–298, 2023.
- YU, Y. et al. A review of recurrent neural networks: Lstm cells and network architectures. *Neural computation*, MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , v. 31, n. 7, p. 1235–1270, 2019.
- YUAN, X. et al. Feature importance ranking of random forest-based end-to-end learning algorithm. *Remote Sensing*, MDPI, v. 15, n. 21, p. 5203, 2023.
- ZHANG, M. Time series: Autoregressive models ar, ma, arma, arima. *University of Pittsburgh*, 2018.
- ZHOU, Z.-H. *Machine learning*. [S.l.]: Springer nature, 2021.
- ZIA, S. Climate change forecasting using machine learning sarima model. *iRASD Journal of Computer Science and Information Technology*, v. 2, n. 1, p. 01–12, 2021.
- ZIEN, A. et al. The feature importance ranking measure. In: SPRINGER. *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2009, Bled, Slovenia, September 7-11, 2009, Proceedings, Part II 20*. [S.l.], 2009. p. 694–709.
- ZIVKOVIC, M. et al. Xgboost hyperparameters tuning by fitness-dependent optimizer for network intrusion detection. In: *Communication and intelligent systems: Proceedings of ICCIS 2021*. [S.l.]: Springer, 2022. p. 947–962.