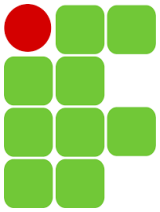


INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA BAIANO
Bacharelado em Sistemas de Informação

UTILIZAÇÃO DE REDES NEURAIS NA ANÁLISE E PREVISÃO DA TEMPERATURA MÉDIA EM ITAPETINGA-BA

Lucas Silva de Oliveira

Itapetinga - Bahia
5 de março de 2025

UTILIZAÇÃO DE REDES NEURAIS NA ANÁLISE E PREVISÃO DA TEMPERATURA MÉDIA EM ITAPETINGA-BA

Lucas Silva de Oliveira

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Orientador(a): Prof(a). Dr(a). Nome do(a) Orientador(a).

Itapetinga - Bahia
5 de março de 2025

UTILIZAÇÃO DE REDES NEURAIS NA ANÁLISE E PREVISÃO DA TEMPERATURA MÉDIA EM ITAPETINGA-BA

Lucas Silva de Oliveira

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

BANCA EXAMINADORA:

Prof(a). Dr(a). Nome do(a) Orientador(a) (Orientador(a))
IFBAIANO

Prof(a). Dr(a). Nome do(a) Avaliador(a)
Instituição

Prof(a). Dr(a). Nome do(a) Avaliador(a)
Instituição

Dedico este trabalho a...

Agradecimentos

Agradeço a...

"Coloca aqui a epígrafe."
— Nome do autor da epígrafe

Resumo

Aqui fica o resumo em português.
Palavras-chave: redes neurais, séries temporais, meteorologia.

Resumo

Aqui fica o resumo em inglês.
Keywords: neural networks, time series, meteorology.

Sumário

1 INTRODUÇÃO

Aqui é onde ficara a introdução da problemática

1.1 OBJETIVOS

Analisar e prever variações da temperatura média no município de Itapetinga-BA ao longo do tempo utilizando redes neurais artificiais.

1.2 OBJETIVOS ESPECÍFICOS

- Analisar séries temporais de temperatura média em Itapetinga-BA, utilizando redes neurais do tipo MLP.
- Identificar tendências de aquecimento e outros impactos ambientais, empregando um modelo de previsão baseado em aprendizado profundo.
- Avaliar o desempenho do modelo MLP considerando diferentes configurações, como número de camadas ocultas, neurônios por camada e funções de ativação.

1.3 JUSTIFICATIVA

Suspeita-se que as mudanças ambientais e climáticas tenham como principal responsável a ação humana, impulsionada pela intensa atividade industrial. A revolução industrial marcou o início dessa transformação, promovendo a adoção de novas fontes de energia e o fortalecimento do consumo de combustíveis fósseis, como o carvão mineral inicialmente e, posteriormente, o petróleo (MENDONÇA, 2006).

À vista disso, nas últimas décadas, os debates sobre as mudanças climáticas e a necessidade de uma sociedade mais consciente e participativa na preservação ambiental e no desenvolvimento sustentável se tornaram cada vez mais intensos. Em 1972, ao identificar a vulnerabilidade do planeta Terra, houve um esforço mundial conjunto em entender os problemas ambientais e ponderar medidas para previnir e amenizar determinados crises.

A Conferência das Nações Unidas para o Meio Ambiente Humano, também conhecida como Conferência de Estocolmo, realizada em 1972 na cidade de Estocolmo, na Suécia, foi um marco histórico por ser a primeira conferência global com foco no meio ambiente. Durante o evento, deu-se início à estruturação de mecanismos de proteção ambiental, que foram ampliados na Segunda Conferência das Nações Unidas sobre o Meio Ambiente e Desenvolvimento, realizada em 1992, conhecida como Rio-92. Nessas conferências, foram

estabelecidos diversos acordos para a proteção do meio ambiente, da biodiversidade e de outros aspectos relacionados à sustentabilidade, como a Agenda 21 (PASSOS, 2009).

Entretanto, de acordo com relatório especial publicado em 2020 pelo Painel Intergovernamental sobre Mudanças Climáticas (IPCC), desde o período pré-industrial, a temperatura média do ar na superfície da Terra quase dobrou em relação à média global registrada anteriormente. Além disso, estima-se que 23% das emissões antrópicas de gases de efeito estufa sejam provenientes de atividades relacionadas à agricultura, silvicultura e outras práticas agrícolas.

No Brasil, em 2019, logo após a transição para a nova gestão federal, as invasões às terras indígenas por grupos ilegais, como garimpeiros, foram retomadas. Como resultado, a taxa de desmatamento em junho daquele ano já apresentava um aumento alarmante de 60% em relação ao mesmo mês do ano anterior. Além disso, houve uma intensificação de atividades ilícitas, como a grilagem de terras, mineração clandestina e exploração madeireira na Amazônia (FILHO, 2020).

Filho (2020) afirma que esse aumento foi impulsionado por questionamentos feitos pelas autoridades governamentais naquele momento, que duvidavam da veracidade das informações fornecidas pelos órgãos responsáveis pelo monitoramento ambiental. Além disso, declarações sobre a possível flexibilização da regulamentação ambiental reforçaram, entre certos grupos, a percepção de uma "liberação total". Isso resultou na intensificação de práticas prejudiciais ao meio ambiente, à saúde pública e ao tecido social.

Segundo dados do MapBiomas, o Brasil já havia perdido cerca de 20% de suas áreas naturais até 1985. Entre 1985 e 2023, essa perda se intensificou, aumentando em mais 13%, atingindo um total de 33% do território nacional. A velocidade alarmante dessa transformação na cobertura e no uso do solo contribui significativamente para o agravamento dos riscos climáticos no país. No ano de 2023, a Bahia se destacou como o segundo estado com maior taxa de desmatamento. Em comparação com 2022, houve um aumento de 27% na área desmatada, sendo o Cerrado o bioma mais afetado, respondendo por 67% do total. Na sequência, aparecem a Caatinga e a Mata Atlântica como os biomas mais impactados pelo desmatamento no estado (POLCRI, 2024).

Dessa forma, torna-se fundamental dispor de instrumentos capazes de prever eventos climáticos com baixa margem de erro e antecedência suficiente para viabilizar a construção de soluções e estratégias eficazes na mitigação de danos. Nesse contexto, destacam-se os modelos de *machine learning*, que, ao utilizar dados históricos meteorológicos, podem realizar predições mais assertivas e robustas para auxiliar na tomada de decisão, permitindo um planejamento mais eficiente em setores como agricultura, energia e gestão de desastres naturais.

Nesse sentido, este trabalho tem como foco analisar as mudanças na temperatura média no município de Itapetinga-BA, cidade historicamente conhecida como a 'Capital da Pecuária'. Além de contar com um setor industrial significativo, Itapetinga abriga

um *campus* avançado da Universidade Estadual do Sudoeste da Bahia e um campus do Instituto Federal Baiano, consolidando-se como um importante polo educacional e econômico na região.

1.4 ORGANIZAÇÃO DOS CAPÍTULOS

Este trabalho está dividido da seguinte forma: No capítulo 2, é apresentado uma visão geral dos conceitos teóricos que fundamentam este projeto, como séries temporais, suas características e modelos de previsão. Além disso, é introduzido o conceito de redes neurais, abordando seus componentes, aprendizado, métricas e a arquitetura utilizada neste trabalho. O capítulo 3 detalha os métodos e técnicas utilizados na pesquisa. No capítulo 4, são mostrados os resultados obtidos. Por fim, no capítulo 5, são apresentamos as considerações finais, destacando as principais contribuições do trabalho e propondo direções para futuras pesquisas.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são apresentados os principais conceitos e fundamentos necessários para a compreensão, entendimento e progresso deste trabalho.

2.1 Séries Temporais

Muitas pessoas, em algum momento, já imaginaram como seria prever o futuro e ter acesso a informações sobre eventos ou situações de suas vidas. Essa curiosidade reflete um desejo universal, mas também uma necessidade presente em diversas áreas, como na gestão governamental, no setor financeiro e em contextos sociais. Nesse cenário, surge o conceito de Série Temporal, definido como um conjunto de observações organizadas sequencialmente no tempo, representadas por x_t, com cada valor correspondente a um instante específico t (BOX et al., 2015). O estudo de Séries Temporais permite não apenas compreender as características de fenômenos que evoluem ao longo do tempo, mas também desenvolver e ajustar modelos estatísticos capazes de explicar ou prever o comportamento dos dados observados.

De acordo com Brockwell e Davis (2002), séries temporais podem ser classificadas discretas e continuas, uma série temporal é discreta quando o conjunto t_0 de tempos em que as observações são feitas é um conjunto discreto, como o caso de observações que são realizadas em um determinado intervalo de tempo fixo. Sendo representada por:

$$T = \{t_1, \ldots, t_n\}, \quad \{X_t : t \in T\} \tag{2.1}$$

E seríes temporais continuas quando suas observações são obtidas continuamente no tempo. Sendo expressa por:

$$T = [t_1, t_2], \quad \{X(t) : t \in T\} \tag{2.2}$$

2.1.1 Estacionariedade

A estacionariedade refere-se ao comportamento dos valores de uma série temporal ao longo do tempo. Uma série é classificada como estacionária quando seus dados flutuam de maneira aleatória em torno de uma média fixa, mantendo-se em um estado de equilíbrio ao longo do período analisado. Contudo, na prática, muitas séries temporais do cotidiano não apresentam essa característica, exibindo tendências, sazonalidades ou outras formas de variação que indicam a presença de não estacionariedade (MORETTIN; TOLOI, 2018).

2.1.1.1 Teste de Estacionariedade

A maioria dos métodos comuns de análise estatística de séries temporais parte do princípio de que as séries a serem analisadas são estacionárias. Para verificar essa condição, é necessário aplicar testes de estacionariedade. Um dos mais conhecidos e utilizados é o teste de Dickey-Fuller Aumentado, que tem como objetivo identificar a presença de raízes unitárias nos operadores de retardos[1] (COSTA, 2019)

As hipóteses nula e alternativa do teste Dickey-Fuller Aumentado são:

$$H_0 : \text{a série possui raiz unitária}$$
$$H_1 : \text{a série não possui raiz unitária}$$

Caso a série tenha raiz unitária, ela não é estacionária; se não possuir raiz unitária, a série é estacionária.

2.1.2 Decomposição

De acordo com Costa (2019), a análise inicial de uma série temporal se beneficia significativamente do uso de gráficos que exibem os dados em ordem cronológica, pois essa abordagem facilita a identificação de padrões e características inerentes, tais como tendência, sazonalidade, ciclicidade e ruído (erro aleatório). Além disso, Corrêa (2024) esclarece o significado desses padrões, contribuindo para uma compreensão mais aprofundada dos comportamentos presentes nas séries.

A tendência (μ_t) representa o padrão de variação de uma grandeza ao longo do tempo. Em termos gerais, ela indica se uma série temporal segue um crescimento ou um declínio de forma consistente. Séries temporais que exibem esse comportamento são classificadas como não estacionárias.

A ciclicidade (ψ_t) refere-se a variações nos dados que ocorrem em intervalos regulares, com uma duração superior a um ano. Esses ciclos podem estar associados a fatores econômicos, políticos ou sociais, manifestando-se de forma recorrente ao longo do tempo.

A sazonalidade (γ_t) é uma variação periódica que se repete em intervalos regulares, mantendo a mesma frequência e intensidade, ocorrendo em um intervalo de tempo menor que um ano. Esse comportamento pode ser observado em vendas durante datas comemorativas anuais, como Páscoa, Natal e São João, além de em padrões climáticos.

O ruído (ϵ_t) representa a variabilidade imprevisível em uma série temporal, correspondendo a flutuações que não podem ser explicadas por outros componentes. Além disso, ele atua como uma fonte de erro nas previsões, afetando a precisão dos modelos.

[1] Raízes dos operadores de retardos vêm da equação característica de um sistema. Em séries temporais, elas determinam a estacionariedade.

2.1.3 Modelos

Morettin e Toloi (2018) afirmam que modelos probabilísticos ou estocásticos podem ser construídos no domínio temporal ou de frequências. Além disso, esses modelos devem ser simples e conter um número reduzido de parâmetros. Morettin e Toloi (2018) também destacam a existência de diferentes modelos que utilizam distintos métodos computacionais para calcular a mesma estimativa, especificamente a previsão de mínimos [2] de um valor futuro com base em combinações lineares de valores passados.

2.1.3.1 ARIMA

O modelo ARIMA (Autoregressive Integrated Moving Average) foi criado nos anos 1970 por George Box e Gwilym Jenkins, visando descrever as mudanças nas séries temporais utilizando a combinação de três componentes principais, conhecidos também como filtros (p, d, q), sendo eles:

- Autoregressão *(AR)*;
- Integração *(I)*;
- Média Móvel *(MA)*;

O componente AR, ou filtro p, tange à dependência das observações atuais em relação às observações passadasm, visto que ele toma como base a intuição de que o passado prediz o futuro. Com isso, ele pressupõe um processo de série temporal no qual o valor em um ponto no tempo t é uma função dos valores da série em pontos anteriores no tempo. É dada por (2.3).

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + e_t \tag{2.3}$$

y_t é o valor da série temporal no tempo t. Sendo a variável dependente, a qual queremos prever. ϕ_0 constante que representa o termo intercepto. $\phi_1, \phi_2, \ldots, \phi_p$ são os coeficientes do modelo que indicam a influência dos valores passados $y_{t-1}, y_{t-2}, \ldots, y_{t-p}$ no valor atual y_t da série. $y_{t-1}, y_{t-2}, \ldots, y_{t-p}$ são defasagens da variável y_t, ou seja, valores passados da série. Sendo que a ordem p indica até quantos períodos anteriores são considerados para prever o valor atual. E e_t é um termo de erro que varia com o tempo, pressupõe-se que esse termo de erro possui uma variância constante e uma média 0.

O componente MA, ou filtro q, baseia-se em um processo no qual o valor em cada instante é determinado por uma função dos termos de "erro" do passado recente, tratados como independentes entre si. De forma similar a um modelo de regressão, esse componente modela a relação entre as observações atuais e os erros anteriores (NIELSEN, 2021;

[2] O metódo dos minímos quadrados é uma técnica que busca encontrar o gráfico de melhor ajuste para um conjunto de amostras de dados (MIYASAKI, 2010).

CORRêA, 2024). Ele é definido pela equação (2.4), onde y_t representa o valor a ser previsto e $\theta_1, \theta_2, \ldots, \theta_q$ são os coeficientes da média móvel, os quais ponderam os termos de erro presentes e passados (NIELSEN, 2021).

$$y_t = \mu + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q} \tag{2.4}$$

Por fim, o componente I ou filtro d, refere-se à aplicação de diferenciação aos dados com o intuito de torná-los estacionários. Em termos práticos, isso significa subtrair cada a observação de sua predecessora, removendo assim tendências e sazonalidades (NIELSEN, 2021; CORRêA, 2024). Essa operação é descrita pela equação (2.5), onde y'_t epresenta a diferença entre os valores consecutivos da série temporal, y_t é o valor da série no tempo t e y_{t-1} corresponde ao período imediatamente anterior. Em essência, a diferenciação converte uma série temporal de valores em uma série de variações, facilitando a análise das mudanças ao longo do tempo.

$$y'_t = y_t - y_{t-1} \tag{2.5}$$

O ARIMA combina os modelos AR e MA, criando o modelo $ARMA$, por fim adicionando a diferenciação I. Morettin e Toloi (2018) recomenda que, antes da modelagem, seja realizado um teste de estacionariedade na série a ser analisada. Se a série for estacionária, o filtro d será 0, resultando em um modelo $ARMA$(p, 0, q). Todavia, pode-se utilizar o modelo $ARMA$ expresso na equação (2.6). Caso a série não seja estacionária, torna-se necessário aplicar diferenciações, conforme a equação (2.5), até que a série se estabilize e o modelo $ARMA$ possa ser empregado.

$$y_t = \phi_0 + \sum(\phi_i r_{t-i}) + e_t - \sum(\theta_i e_{t-i}) \tag{2.6}$$

2.1.3.2 SARIMA

Muitas séries temporais exibem padrões recorrentes que se repetem em intervalos regulares. O modelo ARIMA sazonal (SARIMA) assume uma estrutura multiplicativa para essa sazonalidade, possibilitando que o comportamento sazonal seja modelado como um processo ARIMA à parte. Assim, o modelo pode ser expresso pela notação

$$\text{ARIMA}(p, d, q) \times (P, D, Q)_m,$$

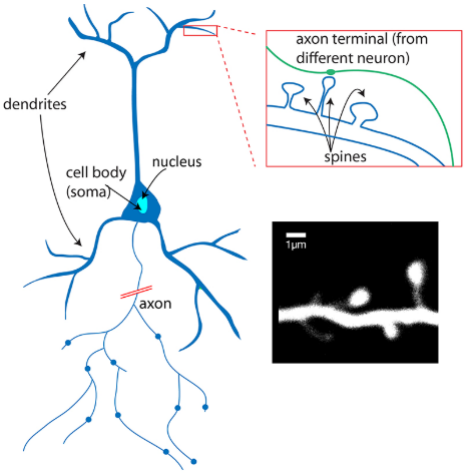
onde m indica o número de intervalos de tempo em cada ciclo sazonal (NIELSEN, 2021).

2.2 Redes Neurais

Analisando o cerebro, percebeu-se a constituição complexa, não linear e paralela dele, sendo composto por 10 bilhos de neurônios. Cada neurônio conectado, em média, a outros 10 bilhões de neurônios, formando uma vasta e sofistiada rede (HAYKIN, 2009).

Em uma rede neural, a comunicação é realizada através de sinais eletroquímicos, e esses sinais são transmitidos e processados através dos componentes presentes em sua estrutura. Os componentes presentes na estrutura de um neurônio são exibidos na Figura 1:

Figura 1 – Neurônio Biológico.



Fonte: The University Of Queensland.

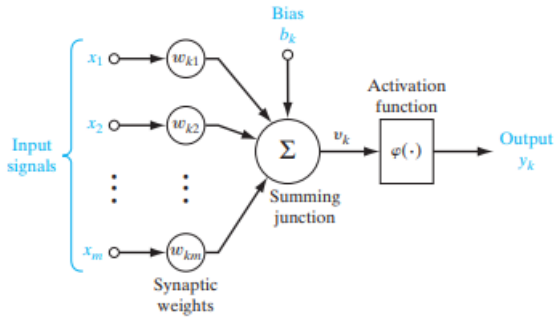
- **Corpo Celular ou Soma:** responsável por integrar os sinais recebidos de outros neurônios;
- **Dendritos:** responsáveis por receber informações transmitidas por outros neurônios. São consideradas as zonas receptivas;
- **Axônio:** responsável por transmitir as informações para outros neurônio, também chamado de linha de transmissão;

Quando o conjunto de sinais recebidos é suficientemente forte para ativar o neurônio, este gera um impulso elétrico que percorre seu axônio. Esse sinal eletroquímico coordena e organiza a atividade neuronal, permitindo ao cérebro realizar diversas formas de processamento de maneira extremamente eficiente, muitas vezes mais rápida do que os computadores digitais convencionais (HAYKIN, 2009). Ao entender de que forma o cerebro humano processa informações, cientistas buscaram reproduzir seu funcionamento de forma artificial.

Com isso, surgiram as Redes Neurais Artificiais (RNA's), que tenta mimetizar o sistema nervoso humano. RNA's são capazes de assimilar e reter conhecimento, possuíndo um alto grau de paralelsimo e extremamente conectada. As RNA's são utilizadas para desempenhar a mesma função após o treinamento. Sendo seu elemento constituinte chamado de neurônio artificial (ULINICK; SCHASTAI, 2019).

De acordo com Haykin (2009), um neurônio é uma unidade de processamento de informação que é fundamental para a operação de um rede neural. Na Figura 2 é ilustrado os componentes básicos da arquitetura mais simples de uma rede neural.

Figura 2 – Neurônio Artificial (Perceptron).



Fonte: Haykin (2009).

- **Camada de Entrada:** Responsável por receber as informações do ambiente externo e transmiti-las ao restante da rede para processamento.
- **Sinapses:** Representam as conexões entre os neurônios, simuladas por pesos $w_1, w_2, \ldots, w_n$, que determinam se o sinal terá um efeito estimulante (excitador) ou inibidor.
- **Bias:** É um valor adicional que ajusta a ativação do neurônio, permitindo que ele se "ajuste"de forma mais flexível. O viés pode ser positivo ou negativo e ajuda a rede a decidir quando ativar ou não o neurônio, independentemente das entradas. Ele atua como uma espécie de "ajuste fino", permitindo que a rede aprenda melhor a partir dos dados.
- **Somador:** Mecanismo que agrega todas as entradas, considerando seus respectivos pesos, para gerar um valor combinado.
- **Função de ativação:** Limita a amplitude do sinal de saída a um intervalo finito, determinando se o neurônio será ativado com base na soma ponderada das entradas.
- **Camada de Saída:** Fornece o resultado final do processamento do neurônio, representando a resposta da rede ao estímulo recebido.

2.2.1 Processo de Treinamento

Sistemas de aprendizado de máquina podem ser categorizados de acordo com o tipo de treinamento que eles recebem. O aprendizado supervisionado ocorre quando o modelo é treinado por meio de exemplos explícitos. Em contrapartida, no aprendizado não supervisionado, não há a definição de exemplos explícitos para orientar o modelo. Além disso, existem diversas boas práticas para garantir que o modelo consiga realizar um aprendizado satisfatório e métricas para avalia-lo

2.2.1.1 Aprendizado Supervisionado e Não Supervisionado

No aprendizado supervisionado, o modelo é treinado com pares de dados (x, y), onde x representa os dados de entrada e y o valor esperado (ou rótulo). Durante o treinamento, o modelo compara as previsões feitas com os valores reais utilizando uma função de perda, que mede o erro. Em seguida, seus parâmetros são ajustados iterativamente, geralmente por meio de métodos como o gradiente descendente, para minimizar esse erro e melhorar a precisão das previsões (ULINICK; SCHASTAI, 2019).

No aprendizado não supervisionado, o modelo não recebe o par de dados (x, y), mas apenas as entradas x. A partir disso, ele busca identificar padrões, estruturas ou associações presentes nos dados, ajustando os pesos de acordo com o objetivo do método utilizado (ULINICK; SCHASTAI, 2019).

2.2.1.2 Pre-Processamento dos Dados

Antes de treinar um modelo com algoritmos de aprendizado de máquina, é imprescindível realizar o pré-processamento dos dados. Esse processo assegura que os dados estejam padronizados, consistentes e adequados, permitindo que os modelos alcancem um desempenho superior e resultados confiáveis nas métricas de avaliação. Para isso, são utilizadas técnicas que evitam problemas como dados ausentes, inconsistências, valores conflitantes e incongruentes. Essas técnicas são geralmente divididas em quatro categorias principais: limpeza, integração, transformação e redução de dados (SILVA, 2021; OLIVEIRA, 2024).

2.2.1.2.1 Limpeza de Dados

Um problema comum em conjuntos de dados (*datasets*) é a presença de valores faltantes (ou nulos), que podem ocorrer devido a diferentes fatores. Esses fatores incluem registros manuais realizados de forma inadequada, falhas em sistemas de extração, transformação e carregamento de dados (*ETL*) ou até mesmo problemas em sensores de dispositivos autônomos. A presença de valores faltantes compromete tanto a qualidade dos dados quanto o desempenho do treinamento de modelos de *machine learning*, caso não seja tratada adequadamente. Sivakumar e Gunasundari (2017) apresentam algumas abordagens eficazes para lidar com essa problemática, como:

- Exclusão de linhas do *dataset* que contenham valores faltantes.
- Preenchimento dos valores ausentes utilizando métricas estatísticas, como a média ou mediana, para gerar estimativas aproximadas que mantenham a coerência do conjunto de dados.

2.2.1.2.2 Integração de Dados

É o processo de combinar dados provenientes de diversas fontes, ecossistemas e tecnologias, de maneira adequada e coerente. Durante esse processo de integração, podem surgir problemas, como inconsistências nos dados e redundâncias no conjunto de dados gerado (SIVAKUMAR; GUNASUNDARI, 2017; SILVA, 2021; OLIVEIRA, 2024).

2.2.1.2.3 Transformação de Dados

Nesse estágio, os dados são transformados em formatos adequados para utilização no modelo. Sivakumar e Gunasundari (2017) e Oliveira (2024) definem algumas atividades executadas nesta etapa:

- Uso de normalização para ajustar os valores dos dados a uma escala comum, permitindo fácil comparação entre diferentes atributos.
- Eliminação de ruídos com técnicas de suavização.
- Aplicação de técnicas de agregação para resumir dados complexos e detalhados.
- Generalização de valores específicos em categorias mais amplas, como, por exemplo, a generalização de faixas etárias.

2.2.1.2.4 Redução de Dados

Nessa etapa, são utilizadas metodos para reduzir o volume de dados que serão analisados, visando maior velocidade de processamento e melhora na eficiência do processo, mas sem comprometer a qualidade e integridade dos dados originais. Sivakumar e Gunasundari (2017) indicam algumas estrategias, sendo elas:

- Redução de dimensionalidade do *dataset*, removendo atributos que não melhoram a perfomance do modelo.
- Utilização de operações de agregação de dados para resumo de informações.
- Utilização de técnicas de *encoding* para compactação de dados.

2.2.1.3 Overfitting e Undefitting

Ao construir modelos de redes neurais, diversos desafios podem surgir, sendo dois dos mais comuns: o *Overfitting* e o *Underfitting*. O *Overfitting* ocorre quando o modelo aprende tão bem os padrões dos dados de treinamento que sua capacidade de generalização para novos dados é comprometida. Isso acontece porque o modelo não apenas captura as características relevantes dos dados, mas também absorve ruídos e padrões específicos

do conjunto de treinamento (Montesinos-LÓPEZ; Montesinos; Crossa, 2022). Como consequência, ao ser exposto a novos exemplos, seja do conjunto de teste ou de outros conjuntos de dados inéditos, o modelo tende a aplicar regras memorizadas, em vez de identificar padrões generalizáveis. Isso compromete sua capacidade de inferir corretamente a partir de dados não vistos, resultando em um desempenho insatisfatório.

Por outro lado, o *Underfitting* ocorre quando o modelo é excessivamente simples, muitas vezes devido à utilização de poucas variáveis de entrada. Isso impede que o modelo represente adequadamente os padrões predominantes no *dataset* e capture as características essenciais das amostras, resultando em um desempenho insatisfatório já durante o treinamento. Além disso, esse problema também pode surgir quando o conjunto de dados de treinamento é muito pequeno ou pouco representativo da população, comprometendo ainda mais a capacidade de aprendizagem do modelo (Montesinos-LÓPEZ; Montesinos; Crossa, 2022).

2.2.1.4 Técnicas de Validação

A validação é uma etapa crucial para avaliar a capacidade preditiva de um modelo de *machine learning*, além de ser fundamental para o ajuste de seus hiperparâmetros.

2.2.1.4.1 Hold-Out

É uma técnica simples de validação de modelos de *machine learning*, ela consiste em dividir o conjunto de dados em duas partes: uma para treino e outra para validação. Ou seja, parte dos dados é usada para treinar o modelo e a outra para testar a capacidade de previsão do modelo. Dado um conjunto de dados d, um subconjunto p é extraído dos dados, formando a amostra de treino d_t onde $t = n * (1 - p)$ e amostra de validação d_v com tamanho $v = n * p$. A técnica é muito simples, vide que divide os dados entre treino e validação, dessa forma, ele usa apenas uma parte dos dados para treinar o modelo. Dessa maneira, o modelo pode não generalizar bem para os dados não vistos presentes no conjunto de validação, acarretando um erro de previsão maior (CUNHA, 2019).

2.2.1.4.2 K-Fold

Nessa técnica de validação de modelo, a amostra d é dividida em K partes de tamanho semelhante. O processo de treino ocorre K vezes, usando K - 1 partes para treino e uma parte para validação, alternando-as partes a cada iteração. Dessa forma, ao final dos K passos, teríamos usado todos os dados tanto para treino e validação (CUNHA, 2019).

2.2.1.5 Métricas de Avaliação

Vide que a utilização de modelos de *machine learning* é com o foco de predizer determinados eventos atráves de métodos estáticos e probabilisticos. A exatidão da previsão é o fator crucial em avaliar a qualidade de um modelo. Sousa (2011) apresenta a Tabela 1 com as métricas mais comuns para avaliar as previsões dos modelos.

Tabela 1 – Métricas de Avaliação

Designação	**Fórmula**
Erro Absoluto Médio (MAE)	$\frac{1}{n}\sum_{t=1}^{n}\lvert e_t\rvert$
Erro Quadrático Médio (MSE)	$\frac{1}{n}\sum_{t=1}^{n}(e_t)^2$
Raiz do Erro Quadrático Médio (RMSE)	$\sqrt{\frac{1}{n}\sum_{t=1}^{n}(e_t)^2}$
Erro Percentual Absoluto Médio (MAPE)	$\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\lvert e_t\rvert}{\lvert y_t\rvert}\right)100$

Fonte: Sousa (2011)

2.2.2 Multi Layer Perceptron

Em 1958, o psicólogo Frank Rosenblatt publicou um artigo que, pela primeira vez, descreveu de forma algorítmica o funcionamento de um modelo de rede neural para aprendizagem supervisioanda. Essa publicação inspirou inúmeros pesquisadores a direcionarem seus esforços para estudos sobre redes neurais, explorando diversos aspectos dessa temática ao longo das décadas de 1960 e 1970 (HAYKIN, 2009).

Como apresentado na Figura 2, o perceptron consiste de um único neurônio com pesos sinápticos ajustáveis e um viés. Ele possui uma camada de entrada (a retina) conectada aos pesos e uma camada de saída. Seu funcionamento baseia-se em um combinador linear seguido por uma função de ativação que realiza uma função linear. Esse nó somador (o neurônio) calcula uma combinação linear das entradas aplicadas às suas sinapses, além de incorporar um viés aplicado externamente que ajusta a posição da função de ativação. O resultado dessa soma é passado à função de ativação, que produz uma saída de +1 se a entrada for positiva, ou -1, se for negativa.

O *Perceptron* é um classificador binário, pois resolve apenas problemas de classificação de padrões linearmente separáveis, ou seja, é capaz de lidar exclusivamente com problemas nos quais duas classes podem ser separadas por uma linha em um hiperplano (HAYKIN, 2009).

Com o objetivo de solucionar problemas não linearmente separáveis, que o *Perceptron* não consegue lidar, surge a *Multi-Layer Perceptron (MLP)*, uma generalização da rede

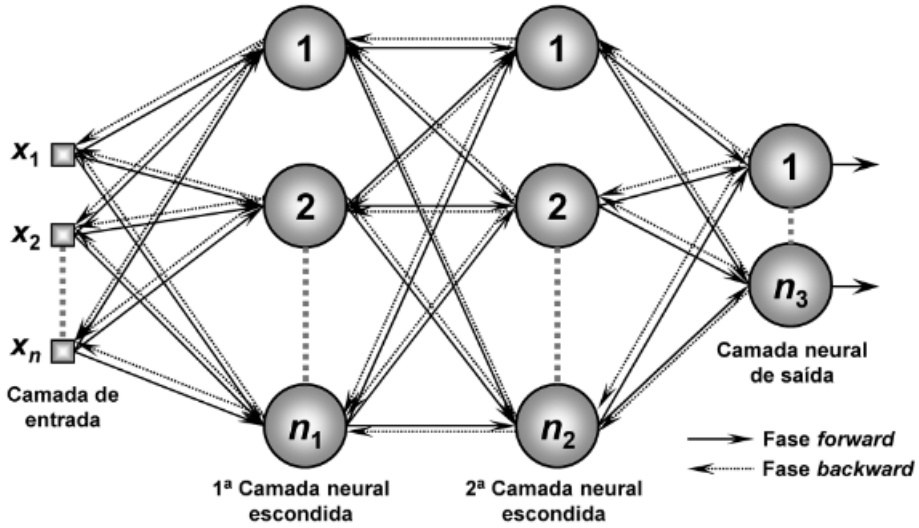
Perceptron, que adiciona camadas de neurônios intermediários, comumente chamadas de camadas ocultas, situadas entre a camada de entrada e a respectiva camada de saída. O treinamento dessa arquitetura também é supervisionado, e as funções de ativação mais usuais são a sigmoide logística e a tangente hiperbólica (SOARES, 2016).

Silva (2010) destaca que a rede *MLP* é uma das mais versáteis em termos de aplicação, sendo amplamente utilizada para tarefas como aproximação universal de funções, reconhecimento de padrões, identificação e controle de processos, previsão de séries temporais e otimização de sistemas.

Além disso, Silva (2010) ressalta uma diferença fundamental entre o *Perceptron* e a *MLP*: enquanto o *Perceptron* possui uma única saída, a *MLP* pode contar com múltiplos neurônios na camada de saída, permitindo que cada um deles represente uma variável distinta do problema modelado. Assim, se um processo apresenta m saídas, a *MLP* terá m neurônios em sua última camada. Outro aspecto essencial dessa arquitetura é o seu processo de aprendizado supervisionado, realizado por meio do algoritmo *backpropagation*, também conhecido como algoritmo de retropropagação.

O algoritmo de *backpropagation* ajusta os pesos da rede neural para reduzir o erro entre a saída prevista e o valor esperado. O erro é calculado e propagado pelas camadas até atingir um nível mínimo aceitável (MARANGONI, 2010).

Figura 3 – Backpropagation.



Fonte: Da Silva (2016).

Grus (2021) apresenta o funcionamento padrão do treinamento de uma rede neural utilizando o algoritmo de backpropagation como método de ajuste dos pesos. Considera-se que a rede possui n, os quais são ajustados de acordo com o seguinte procedimento:

1. Realiza-se o *feed-forward*, em que as entradas são processadas para produzir as saídas de todos os neurônios;

2. Como o algoritmo é supervisionado, os valores esperados das saídas são conhecidos. Assim, calcula-se uma função de perda, geralmente definida como a soma dos erros quadráticos entre as saídas reais e as esperadas;

3. O gradiente dessa função de perda é calculado em relação aos pesos dos neurônios de saída;

4. Os gradientes e os erros são propagados para trás com o objetivo de calcular os gradientes associados aos pesos dos neurônios ocultos;

5. Atualizam-se os pesos aplicando um passo em direção ao gradiente descendente, controlado por um parâmetro denominado learning rate (taxa de aprendizagem).

3 METODOLOGIA

Neste capítulo, são apresentados o ambiente e a abordagem metodológica que serão adotados para o desenvolvimento deste estudo, detalhando os procedimentos que serão seguidos desde a coleta e pré-processamento dos dados até a elaboração, treinamento e avaliação do modelo de rede neural.

3.1 Coleta dos Dados

Para a implementação dos modelos de previsão, será essencial a disponibilização de uma base de dados substancial para o treinamento, validação e teste do modelo, assim como para a realização das inferências sobre a população. No Brasil, o Instituto Nacional de Meteorologia (INMET) é o responsável pelo Banco de Dados Meteorológicos (BDMEP), que tem como finalidade a coleta, armazenamento, processamento e disponibilização de dados e informações sobre variáveis meteorológicas.

Esses dados poderão ser gerados localmente, por meio de estações meteorológicas convencionais ou automáticas, ou adquiridos remotamente, por sensores orbitais, radares, entre outros dispositivos (VIANNA et al., 2017). O Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP), especificamente, reunirá informações diárias provenientes das estações da rede do INMET, em conformidade com as normas da Organização Meteorológica Mundial (INMET, s.d.). Os dados utilizados neste estudo corresponderão ao período de 25/06/2009 a 31/12/2024, provenientes da estação meteorológica automática localizada na cidade de Itapetinga-BA.

3.2 Pré-processamento dos Dados

O pré-processamento dos dados terá início com a análise exploratória, com o objetivo de identificar padrões, anomalias e relações relevantes. De acordo com Tukey (1977), a análise exploratória de dados (AED) será caracterizada por um processo investigativo, análogo ao trabalho de um detetive, que buscará pistas e evidências, enquanto a análise confirmatória de dados se assemelhará ao trabalho judicial, onde as evidências serão testadas e verificadas.

Com base nesse conceito, será realizado o tratamento de inconsistências nos dados, incluindo a remoção ou correção de valores atípicos (*outliers*), o preenchimento de dados ausentes e a normalização das variáveis para garantir sua compatibilidade com o modelo de rede neural. Essas etapas serão essenciais para garantir que o modelo seja treinado com um conjunto de dados consistente e representativo.

3.3 Construção do Modelo de Rede Neural

A construção do modelo de rede neural será baseada na arquitetura *Multi-Layer Perceptron* (MLP), uma das mais utilizadas para previsão de séries temporais. Para a implementação do modelo, serão empregadas as bibliotecas *TensorFlow* e *Keras*, que fornecerão suporte para a construção e treinamento de redes neurais. Adicionalmente, serão aplicadas técnicas como *dropout*, regularização L2 e otimização com o algoritmo *Adam* para evitar sobreajuste e aprimorar a capacidade preditiva do modelo.

3.4 Treinamento e Validação do Modelo

O treinamento do modelo será realizado com o conjunto de dados previamente preparado, utilizando a técnica de validação cruzada e *early stopping* para prevenir sobreajuste. Os hiperparâmetros, como a taxa de aprendizado, o número de camadas ocultas, o número de neurônios por camada e o tamanho do lote (*batch size*), serão ajustados conforme necessário.

A divisão dos dados será realizada da seguinte forma:

- **Treinamento**: 70% dos dados, utilizados para ajuste dos pesos da rede neural.
- **Validação**: 15% dos dados, empregados para avaliação do desempenho do modelo durante o treinamento e ajuste dos hiperparâmetros.
- **Teste**: 15% dos dados, usados para avaliação da capacidade preditiva final do modelo.

3.5 Avaliação e Análise dos Resultados

A avaliação do modelo será realizada por meio de métricas amplamente adotadas na previsão de séries temporais, tais como:

- **Erro Quadrático Médio (MSE - Mean Squared Error)**: medida da média dos erros ao quadrado, penalizando desvios maiores.
- **Raiz do Erro Quadrático Médio (RMSE - Root Mean Squared Error)**: fornecendo uma interpretação mais intuitiva do erro, mantendo a mesma unidade da variável predita.
- **Erro Absoluto Médio (MAE - Mean Absolute Error)**: medida da média dos erros absolutos, sendo menos sensível a grandes *outliers* do que o MSE.
- **Coeficiente de Determinação (R2R2)**: medida da proporção da variabilidade explicada pelo modelo, indicando sua qualidade preditiva.

A análise dos resultados será realizada por meio da comparação do desempenho do modelo com abordagens tradicionais de previsão, como médias móveis e modelos estatísticos baseados em regressão linear. A partir dessa análise, será possível verificar a eficácia do modelo proposto na previsão da temperatura média da cidade de Itapetinga-BA.

4 RESULTADOS ESPERADOS

5 Cronograma

Etapas	**Out/2024**	**Nov/2024**	**Dez/2024**	**Jan/2025**	**Fev/2025**	**Mar/2025**
Escolha do tema	x					
Levantamento biblio.	x					
Elab. do anteprojeto		x				
Apres. do projeto			x			
Desenvolvimento				x	x	x
Org. do roteiro	x					
Redação					x	
Revisão final					x	x
Entrega					x	x

Tabela 2 – Cronograma de desenvolvimento do trabalho

6 CONSIDERAÇÕES FINAIS

Referências

BOX, G. E. P. et al. *Time Series Analysis: Forecasting and Control*. 5th. ed. Hoboken, NJ: John Wiley & Sons, 2015.

BROCKWELL, P. J.; DAVIS, R. A. *Introduction to Time Series and Forecasting*. Second. [S.l.]: Springer, 2002.

CLIMáTICAS, P. I. sobre M. *Mudança do clima e terra: sumário para formuladores de políticas*. Brasília: Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTI), 2020. Relatório especial sobre mudança do clima, desertificação, degradação da terra, manejo sustentável da terra, segurança alimentar e fluxos de gases de efeito estufa em ecossistemas terrestres. ISBN 978-92-9169-154-8. Disponível em: <https://repositorio.mcti.gov.br/handle/mctic/5301>.

CORRêA, N. M. Trabalho de Conclusão de Curso (TCC), *Análise de séries temporais e utilização de algoritmos de machine learning para a predição de casos de dengue em Santa Maria (RS)*. Santa Maria, RS: [s.n.], 2024.

COSTA, E. S. d. *Análise da Série Temporal de Precipitação Total Mensal do Município de Cruz das Almas-BA*. Brasil: [s.n.], 2019. Trabalho monográfico apresentado para obtenção do grau de bacharel em Ciências Exatas e Tecnológicas.

CUNHA, J. P. Z. *Um estudo comparativo das técnicas de validação cruzada aplicadas a modelos mistos*. Dissertação (Dissertação (Mestrado em Estatística)) — Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2019. Acesso em: 2025-02-23.

FILHO, H. B. Bolsonaro, meio ambiente, povos e terras indígenas e de comunidades tradicionais: Uma visada a partir da amazônia. *Cadernos de Campo (São Paulo - 1991)*, v. 29, p. e178663, 12 2020.

GRUS, J. *Data Science do Zero: Noções Fundamentais com Python*. 2. ed. Rio de Janeiro, Brasil: Alta Books, 2021. 416 p. ISBN 978-8550811765.

HAYKIN, S. *Neural Networks and Learning Machines*. Third. Upper Saddle River, NJ: Pearson Education, 2009.

MARANGONI, P. H. *Redes Neurais Artificiais para Previsão de Séries Temporais no Mercado Acionário*. Florianópolis: Universidade Federal de Santa Catarina - UFSC, 2010. Trabalho de Conclusão de Curso (Graduação em Ciências Econômicas).

MENDONÇA, F. Aquecimento global e suas manifestações regionais e locais: alguns indicadores da região sul do brasil. *Revista Brasileira de Climatologia*, v. 2, 2006.

MIYASAKI, C. A. H. Trabalho de Conclusão de Curso (Especialização), *Método dos mínimos quadrados: aspectos teóricos e suas aplicações*. Campo Mourão: [s.n.], 2010. 37 p.

Montesinos-LÓPEZ, O.; Montesinos, A.; Crossa, J. *Overfitting, Model Tuning, and Evaluation of Prediction Performance*. [S.l.: s.n.], 2022. 109-139 p. ISBN 978-3-030-89009-4.

MORETTIN, P. A.; TOLOI, C. M. C. *Análise de Séries Temporais: Modelos Lineares Univariados (Volume 1)*. 3. ed. São Paulo: Blucher, 2018. 474 p. ISBN 978-8521213512.

NIELSEN, A. *Análise Prática de Séries Temporais: Predição com Estatística e Aprendizado de Máquina*. 1. ed. Brasil: Alta Books, 2021. 480 p. ISBN 978-8550815626.

OLIVEIRA, J. V. de. *Análise comparativa de algoritmos de aprendizado de máquina aplicados ao Campeonato Brasileiro de Futebol*. 2024. Trabalho de Conclusão de Curso (TCC).

PASSOS, P. N. Calmon de. A conferÊncia de estocolmo como ponto de partida para a proteÇÃo internacional do meio ambiente. *Revista Direitos Fundamentais & Democracia*, v. 6, n. 6, 2009. Disponível em: <https://revistaeletronicardfd.unibrasil.com.br/index.php/rdfd/article/view/18>.

POLCRI, M. *Bahia desmata área equivalente a 737 campos de futebol por dia: o estado ocupa a segunda posição no ranking nacional, atrás apenas do Maranhão*. 2024. Disponível em: <https://www.correio24horas.com.br/minha-bahia/bahia-desmata-area-equivalente-a-737-campos-de-futebol-por-dia-0624>.

SILVA, D. F. B. F. d. *Pré-processamento de Dados e Comparação entre Algoritmos de Machine Learning para a Análise Preditiva de Falhas em Linhas de Produção para o Controlo*. Tese (Doutorado) — Instituto Superior de Engenharia do Porto, 2021.

SILVA, I. N. da. *Redes Neurais Artificiais para Engenharia e Ciências Aplicadas: Fundamentos Teóricos e Aspectos Práticos*. São Paulo: Artliber, 2010. 399 p. ISBN 978-8588098534.

SIVAKUMAR, A.; GUNASUNDARI, R. A survey on data preprocessing techniques for bioinformatics and web usage mining. *International Journal of Pure and Applied Mathematics*, v. 117, n. 20, 2017.

SOARES, E. S. de M. Trabalho de Conclusão de Curso, *MULTI-LAYER PERCEPTRON E RESERVOIR COMPUTING APLICADAS EM UM PROCESSO CHUVA x VAZÃO*. 2016.

SOUSA, J. A. V. *Aplicação de Redes Neuronais na Previsão de Vendas para Retalho*. Dissertação (Dissertação de Mestrado) — Faculdade de Engenharia da Universidade do Porto (FEUP), Porto, Portugal, 2011. Orientador na FEUP: Engº. Eduardo José Rego Gil Costa; Orientador no INESC Porto: Engº. Rui Diogo Rebelo.

TUKEY, J. W. *Exploratory Data Analysis*. 1. ed. [S.l.]: Pearson, 1977.

ULINICK, A. A. de Q.; SCHASTAI, B. *Previsão de demanda para controle de estoque: aplicação de redes neurais artificiais em séries temporais*. 2019. 58 p. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Elétrica) – Universidade Tecnológica Federal do Paraná, Ponta Grossa.

VIANNA, L. F. d. N. et al. Bancos de dados meteorológicos: Análise dos metadados das estações meteorológicas no estado de santa catarina, brasil. *Revista Brasileira de Meteorologia*, v. 32, n. 1, p. 53–64, jan 2017.