

Aplicação de Modelos de Machine Learning para a Previsão da Demanda de Fardamento na Marinha do Brasil

Autoria: C-ApA-IM 2024 - DAbM-04

RESUMO

Este estudo analisa a viabilidade de aplicação de modelos de Machine Learning no processo logístico de previsão da demanda de fardamento da Marinha do Brasil. Esta pesquisa enfatizou o desenvolvimento de um código computacional para mensurar a eficiência da demanda estimada dos itens da cadeia de suprimento de fardamento que são gerenciados por meio do Sistema de Informações Gerenciais do Abastecimento da Marinha do Brasil (SINGRA). A pesquisa adota uma abordagem exploratória, utilizando referenciais teóricos, bem como análise dos dados coletados por meio de consultas extraídas do Banco de Dados do SINGRA. A aplicação da Machine Learning foi realizada por meio de Redes Neurais Artificiais e Árvores de Decisão e os resultados foram comparados com os modelos tradicionais de Regressão Linear, Média Móvel e Amortecimento Exponencial Simples.

Palavras-chave: Machine Learning, Métodos de Previsão de Demanda, Modelo Computacional, Redes Neurais Artificiais, Árvore de Decisão.

1. INTRODUÇÃO

A Marinha do Brasil (MB) possui uma estrutura complexa de abastecimento que garante a disponibilidade dos itens de fardamento necessários para os militares da MB em serviço ativo (Brasil, 2020). A gestão desses itens é realizada por meio do Sistema de Abastecimento da Marinha (SAbM), que tem como um de seus objetivos o controle de estoques e a previsão de demanda. Essa previsão é gerenciada pelo Centro de Suprimentos do Abastecimento através do Subsistema de Planejamento do Sistema de Informações Gerenciais do Abastecimento (SINGRA). A eficácia da previsão é crucial, pois permite o atendimento contínuo e eficaz das necessidades de fardamento da instituição.

O sistema SINGRA, conforme delineado pelo Manual do SINGRA (Brasil, 2007), atualmente utiliza métodos tradicionais de previsão, como Amortecimento Exponencial Simples, Média Móvel e Regressão Linear, que são aplicados de forma generalizada para todas as cadeias de suprimento, sem distinção de comportamento específico da demanda. Embora esses métodos ofereçam uma base sólida, eles podem ser limitados em termos de flexibilidade e precisão para demandas que apresentam padrões específicos ou mudanças sazonais.

Diante dessa necessidade de aprimoramento, está em andamento o projeto de modernização do sistema atual para uma nova solução chamada SINGRA-GCV, uma plataforma Enterprise Resource Planning (ERP) personalizada para a MB. Este sistema trará um conjunto de funcionalidades avançadas, proporcionando uma gestão mais integrada e automatizada dos processos logísticos, otimizando o planejamento e, consequentemente, melhorando a previsão de demanda dos itens de fardamento.

A aplicação de tecnologias de ML surge, nesse contexto, como uma abordagem promissora para suprir as limitações dos métodos de previsão atuais. Modelos como Redes Neurais Artificiais (RNA) e Árvores de Decisão são capazes de captar padrões

complexos nos dados e responder de forma mais eficaz às variações da demanda. Assim, este estudo visa avaliar a viabilidade de implementação de modelos de ML no sistema logístico de previsão de demanda, utilizando dados extraídos do SINGRA, e compará-los com os métodos tradicionais em termos de precisão e eficiência.

Isto posto, este trabalho acadêmico salienta o seguinte problema de pesquisa: A aplicação dos modelos de Machine Learning de RNA e árvore de decisão podem otimizar o processo logístico de previsão da demanda de fardamento na MB?

O objetivo geral desta pesquisa é analisar a acurácia de modelos de RNA e árvores de decisão regressiva de ML na otimização do processo logístico de previsão de demanda. Para alcançar esse objetivo geral, foram definidos os seguintes objetivos específicos: criar modelos de RNA e árvores de decisão com a capacidade de prever a demanda de fardamento da MB e, em seguida, avaliar a acurácia desses modelos por meio da comparação das métricas MASE e RMSSE com métodos quantitativos tradicionais de previsão de demanda.

2. REFERENCIAL TEÓRICO

Com o intuito de dirimir dúvidas sobre os significados e conceitos-chave ao longo da pesquisa, utiliza-se os referenciais teóricos (Cooper; Schindler, 2003). Esses significados e conceitos são bem característicos, ajustados às peculiaridades do tema específico e possui a diretriz de aferir as variáveis englobadas (Richardson, 1999). A seguir estão elencados os referenciais teóricos praticados no decorrer desta pesquisa:

2.1 Previsão da demanda

Segundo Kourentzes (2014), previsões precisas de demanda são essenciais para apoiar as decisões organizacionais relacionadas à manutenção e reposição de estoques. Na ausência dessas previsões, as organizações carecem de informações fundamentais para um planejamento eficaz diante de eventos imprevistos, ficando limitadas a reagir a essas ocorrências (Pellegrini e Fogliatto, 2001).

Existem diversos métodos de previsão de demanda disponíveis para identificar o modelo mais apropriado a cada contexto (Makridakis, Wheelwright e Hyndman, 1998). No entanto, quando a demanda possui um comportamento intermitente, torna-se fundamental o uso de técnicas de previsão específicas para lidar com essa particularidade (Pellegrini e Fogliatto, 2001).

Segundo Veríssimo, Henning, Amaral, Cruz e Alves (2012), a previsão de demanda é uma das atividades de gestão mais cruciais para a tomada de decisões, pois o sucesso futuro de qualquer empresa depende amplamente da experiência do gestor em identificar tendências e desenvolver estratégias adequadas. A precisão diminui conforme o período de previsão se amplia, elevando o risco de imprecisões. Assim, a escolha criteriosa da técnica de previsão adequada às características específicas de cada demanda é essencial (Ballou, 2009).

A busca por métodos estatísticos de previsão e a avaliação de sua qualidade tornaram-se diferenciais para empresas anteciparem cenários futuros. Assim, é possível planejar recursos e reduzir custos de decisões equivocadas (Veríssimo et AL., 2012). Kourentzes (2014) acrescenta que previsões precisas são essenciais para decisões relacionadas à manutenção e reposição de estoques, fornecendo suporte ao planejamento organizacional e à resposta a eventos imprevistos.

Makridakis et. al (1998) apontam que métodos mais elaborados nem sempre proporcionam resultados superiores, ressaltando que os métodos quantitativos

tradicionais continuam eficazes e amplamente aplicados. Kourentzes (2014) reforça que, apesar das novas técnicas oferecerem maior precisão, sua utilização ainda é restrita, o que evidencia a relevância dos métodos tradicionais estabelecidos nas previsões de demanda.

2.2 Métodos quantitativos de previsão de demanda

Os métodos quantitativos de previsão de demanda utilizam modelos matemáticos para projetar a demanda futura com base em dados históricos. Estas previsões podem ser elaboradas por meio de algoritmos, que estruturam os modelos a partir das informações passadas sobre a demanda (Calsing, 2015).

A utilização de múltiplos métodos quantitativos é essencial para aprimorar a precisão das previsões, pois diferentes modelos podem capturar distintas nuances e padrões no comportamento da demanda. Assim, a definição do modelo mais ideal é determinada preferencialmente pelo padrão de comportamento da série temporal em análise (Pellegrini, 2000).

Segundo Slack, Jones e Johnston (2018), as séries temporais monitoram uma variável ao longo do tempo, considerando as causas das variações de tendência, com o objetivo de utilizar a extrapolação para prever comportamentos futuros.

No Quadro 1, são indicadas as notações dos parâmetros empregados nas equações dos métodos de Amortização Exponencial Simples, Regressão Linear e Média Móvel, além das métricas de erros de previsão de MASE e RMSSE. Esses métodos fundamentam o código computacional adotado para a previsão de demanda.

Quadro 1 - Notação das equações de previsão de demanda e estimativa de erros

Notação	Definição	Equação
$\hat{y}_t$	Previsão para o período t	AES, Média Móvel, RMSSE e MASE
y_t	Demanda no período t	AES, Média Móvel, RMSSE e MASE
α	Parâmetro de amortização	AES
Y	Variável dependente	Regressão Linear
X	Variável independente	Regressão Linear
β_0	Intercepto	Regressão Linear
β_1	Coeficiente angular	Regressão Linear
n	Número de observações	Média Móvel, RMSSE e MASE
ϵ	Termo de erro	Regressão Linear

Fonte: Elaborado pelo autor (2024)

2.2.1 Amortização Exponencial Simples (AES)

O modelo de amortização exponencial atribui um peso a cada valor da série temporal, de modo que os dados mais recentes recebem uma maior ponderação (Pellegrini e Fogliatto, 2001). Conforme Makridakis et. al (1998), as principais vantagens desse modelo são sua simplicidade e baixo custo, características que contribuem para seu uso frequente em diversas aplicações.

De acordo com Souza, Samohyl e Miranda (2008), o método de amortização exponencial foi um dos primeiros modelos desenvolvidos para previsão de séries temporais. No modelo computacional desenvolvido, o parâmetro foi otimizado individualmente para cada item, com o objetivo de aprimorar a precisão das previsões. A notação matemática utilizada é apresentada a seguir:

$$\hat{y}_{t+1} = \hat{y}_t + \alpha \times (y_t + \hat{y}_t)$$

2.2.2 Regressão Linear

Segundo Fávero, Belfiore, Takamatsu e Suzart (2014), a regressão linear é uma técnica que estima o valor esperado de uma variável dependente com base nas variações de uma ou mais variáveis explicativas, considerando que a dependente é uma função linear dessas explicativas.

No entanto, é importante considerar que a inclusão do termo de erro na expressão é legítima pelo fato de que qualquer relação proposta dificilmente será perfeita. Em outras palavras, é muito provável que o fenômeno em estudo, representado pela variável Y, tenha relação com alguma outra variável X não incluída no modelo proposto e que, portanto, necessita ser representada pelo termo de erro ϵ (Fávero e Belfiore, 2017). Seu cálculo utiliza as seguinte equação:

$$y = \beta_0 + \beta_1 X + \epsilon$$

2.2.3 Média Móvel

Segundo Syntetos (2001), a média móvel é um estimador que emprega um modelo de médias constantes por período. Esse estimador é utilizado para suavizar séries temporais, reduzindo flutuações de acordo com o número de observações utilizadas na média, ou seja, o valor de n (Morettin e Toloi, 2004).

$$\hat{y}_{t+1} = \frac{1}{n} \times \sum_{t=1}^n y_t$$

A partir da fórmula matemática apresentada, é possível observar que valores menores de n tornam a resposta às mudanças da série temporal mais rápida, enquanto valores maiores resultam em respostas mais lentas. Neste modelo, o cálculo é realizado a cada período, substituindo a observação mais antiga pela mais recente (Morettin e Toloi, 2004).

2.3 Inteligência Artificial, Machine Learning e Redes Neurais Artificiais

A Inteligência Artificial (IA) é um ramo da Ciência da Computação que busca desenvolver computadores capazes de pensar e agir de forma inteligente (Gomes, 2010). Segundo Russell e Norvig (2013), a IA, quando associada ao processo de raciocínio,

pode ser entendida como a capacidade de “fazer a coisa certa” com base no conhecimento disponível.

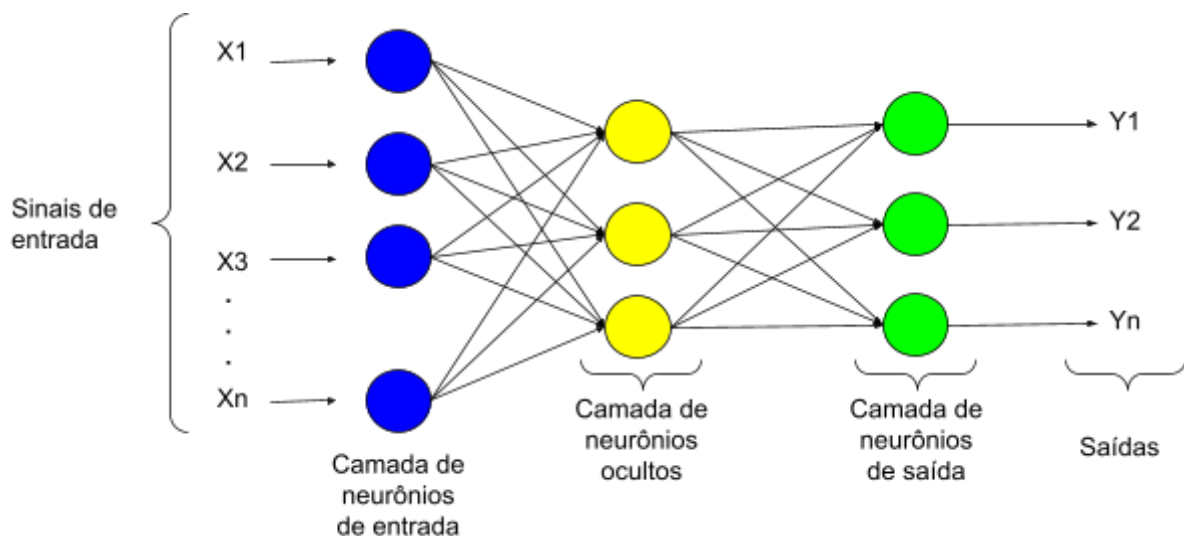
Algoritmos de aprendizado de máquina podem ser entendidos como um processo de busca em um vasto conjunto de modelos possíveis, orientado pela experiência adquirida no treinamento, com o objetivo de identificar um modelo que maximize a métrica de desempenho (Jordan e Mitchell, 2015).

A função principal desta ferramenta é criar e desenvolver algoritmos que capacitem sistemas a evoluir e se ajustar às mudanças dos dados monitorados. Um aspecto essencial na pesquisa de aprendizado de máquina é a geração automática de modelos, como regras e padrões, com base nos dados de treinamento analisados (Pascual, 2015).

As redes neurais artificiais (RNAs) são sistemas computacionais inspirados na estrutura e funcionamento do cérebro humano, capazes de aprender e generalizar a partir de dados fornecidos (Haykin, 2001). As RNAs consistem em unidades de processamento interconectadas, denominadas neurônios artificiais, que trabalham em conjunto para resolver problemas complexos (Rumelhart, McClelland, 1986).

Esses neurônios estão organizados em camadas: uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Cada neurônio recebe sinais de múltiplos neurônios da camada anterior, processa esses sinais através de uma função de ativação e transmite o resultado para os neurônios da próxima camada (Bishop, 2006). A Figura 1 ilustra um esboço simplificado de uma rede neural artificial com uma única camada oculta.

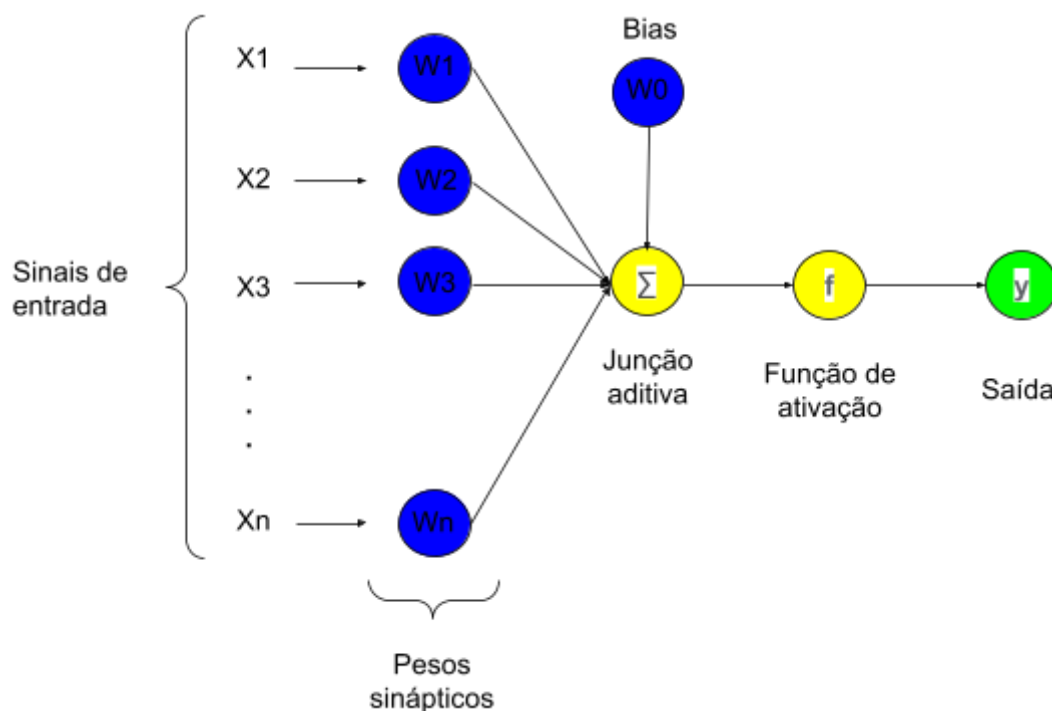
Figura 1 - DNN com 1 camada oculta



Fonte: Adaptado de Haykin (2001)

Segundo Haykin (2001), cada neurônio das camadas mencionadas anteriormente é uma unidade fundamental de processamento de informações, sendo essencial para o funcionamento de uma rede neural. O processamento inicial dos sinais de entrada consiste em multiplicá-los pelos respectivos pesos sinápticos. Em seguida, o neurônio soma todos esses produtos, e o resultado dessa soma é passado a função de ativação. Ao final é apresentado o resultado (Daibert, Braga, Oliveira e Rocha, 2009). Este neurônio pode ser representado pela Figura 2 a seguir.

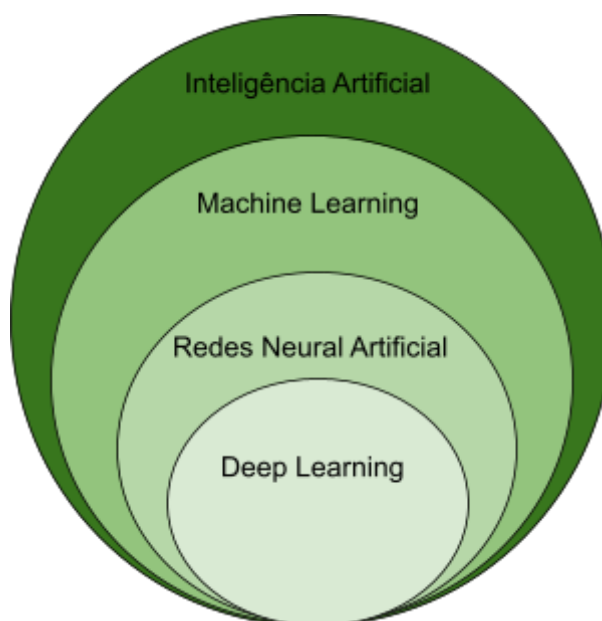
Figura 2 - Esboço de uma rede neural artificial simples



Fonte: Adaptada de Haykin (2001)

Os conceitos apresentados nesta seção são de grande relevância, pois estabelecem a base para o entendimento da área e suas inter-relações fundamentais. A estrutura e a interdependência desses conceitos podem ser visualizadas e compreendidas por meio da teoria unionista dos conjuntos, conforme ilustrado na Figura 3.

Figura 3 - Teoria Unionista

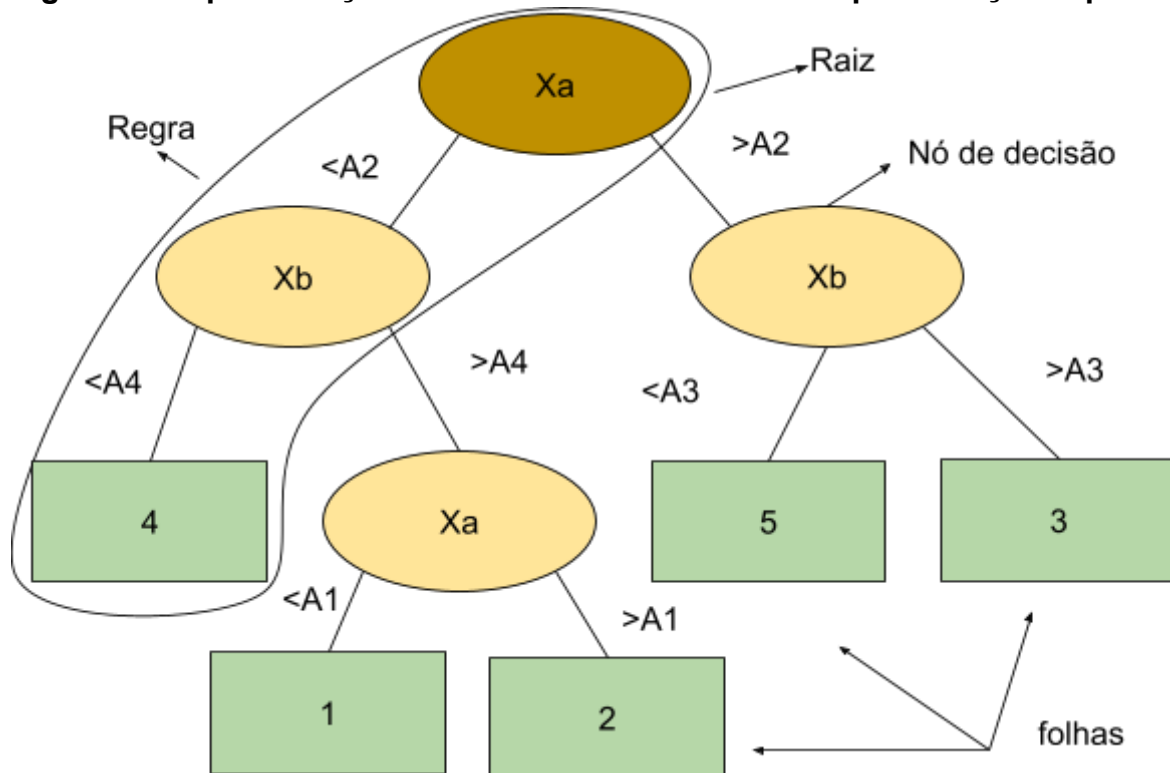


Fonte: Adaptado de Pinkowski (2023)

2.4 Árvore de decisão

Segundo Loh (2011), árvores de decisão são modelos de Machine Learning utilizados para criar modelos de previsão a partir de dados. Esses modelos são construídos por meio de uma divisão recursiva do espaço de dados, aplicando um modelo de previsão simples em cada subdivisão.

Figura 4 - Representação de árvore de decisão e sua representação espacial



Fonte: Adaptado de Gama (2002)

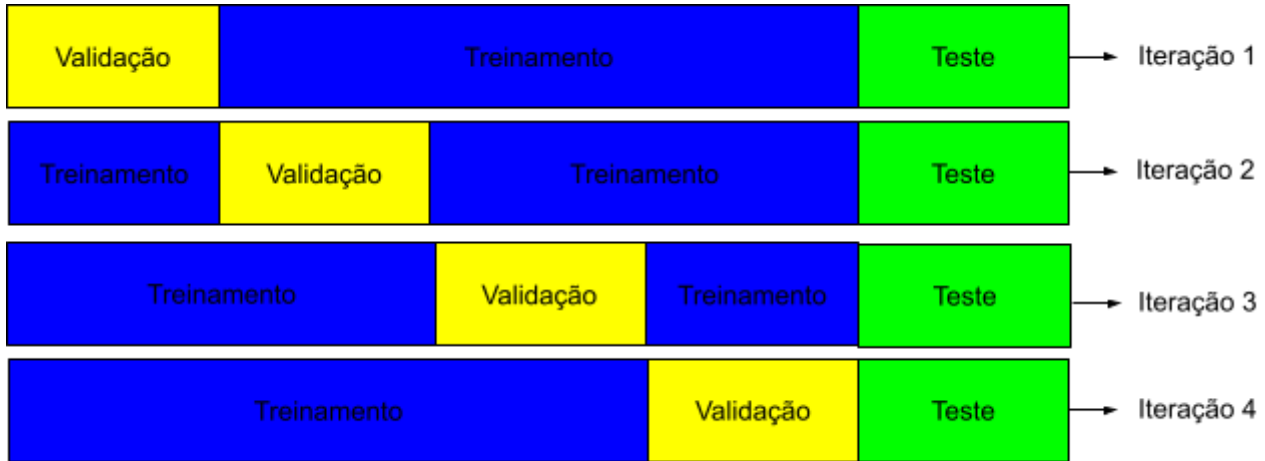
A figura acima ilustra uma árvore de decisão constituída das seguintes características: (i) raiz é o ponto inicial da árvore, onde ocorre a primeira divisão dos dados com base no atributo mais informativo; (ii) nó de decisão representa testes de atributos específicos; (iii) folhas são os pontos terminais da árvore que fornecem a previsão para os dados que chegam a esse ponto; (iv) regra de decisão é o caminho percorrido na árvore a partir da raiz até a folha (Tan, Steinbach, Kumar, 2006).

2.5 Validação Cruzada

A validação cruzada é uma técnica para realizar amostragem iterativa, na qual, a cada repetição, o conjunto de dados iniciais são separados em dados de treinamento e dados de teste. Essa divisão pode ser realizada de maneira aleatória ou seguindo um padrão sistemático, conforme ilustrado na Figura 5.

Ao longo das iterações utilizando um subconjunto de teste distinto são calculados as métricas de acurácia do modelo de previsão avaliado. A avaliação final do desempenho do modelo é dada pela média desses valores ao longo de todas as repetições (Licort, León, Agarwal, Sahelices, Torre, González e Amoon, 2024).

Figura 5 - Validação Cruzada



Fonte: Adaptado de Licort et. AL.(2024)

2.6 Erro de previsão de demanda

A previsão de demanda é essencial para a otimização logística, exigindo alta precisão nas estimativas. Para alcançar este objetivo, é fundamental avaliar corretamente a diferença entre a previsão e a demanda real utilizando uma métrica adequada, pois métricas inadequadas podem comprometer a eficiência dos modelos e resultar em previsões imprecisas (Martin, Spitzer e Kühl, 2020).

Segundo Shcherbakov, Brebels, Shcherbakova, Tyukin e Kamaev (2013), cada métrica de erro possui desvantagens que podem levar a uma avaliação imprecisa dos resultados da previsão, tornando inviável a escolha de uma única métrica. Nesse sentido, foram utilizadas as métricas Mean Absolute Scaled Error (MASE) e Root Mean Squared Scaled Erro (RMSSE), ou seja Erro Absoluto Médio Escalado e o Raiz do Erro Quadrático Médio Escalado, que, conforme ressaltam os autores, são resistentes a valores discrepantes. A seguir, são elucidadas as representações matemáticas das métricas selecionadas:

$$RMSSE = \sqrt{\frac{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{\frac{1}{n-1} \sum_{t=2}^n (Y_t - Y_{t-1})^2}}$$

$$MASE = \frac{\frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|}{\frac{1}{n-1} \sum_{t=2}^n |Y_t - Y_{t-1}|}$$

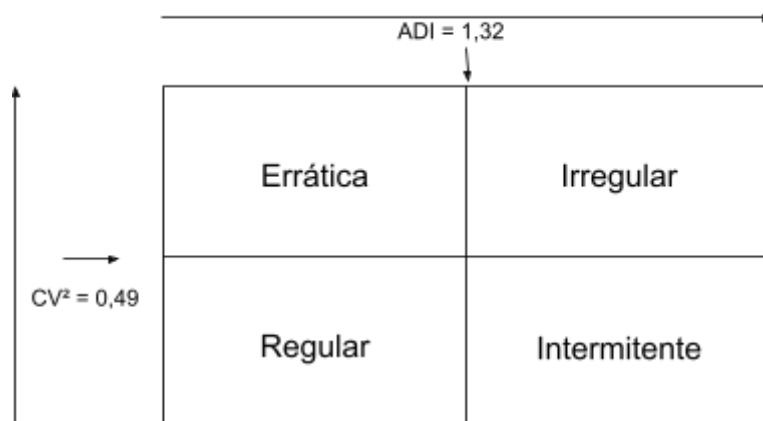
2.7 Classificação da demanda

As diretrizes de categorização da demanda auxiliam a determinar quais métodos de previsão de demanda devem ser aplicados a cada tipo específico de item. Como

resultado, essas regras têm implicações significativas na previsão de demanda, nos níveis de estoque mantidos e na satisfação dos clientes (Boylan, Syntetos e Karakostas, 2008).

Dessa forma, Syntetos, Boylan e Croston (2005) propõem uma classificação utilizando parâmetros e valores de corte para essas medidas, conforme ilustrado na Figura 5. Os parâmetros utilizados são o quadrado do coeficiente de variação (CV^2), calculado por meio do quadrado da divisão do desvio-padrão da demanda pela demanda média e o intervalos médios entre demandas (ADI - Average Demand Interval), equivalente ao número médio de períodos de tempo entre duas demandas sucessivas.

Figura 6 - Classificação da demanda dos itens de fardamento



Fonte: Adaptado de Syntetos, Boylan e Croston (2005)

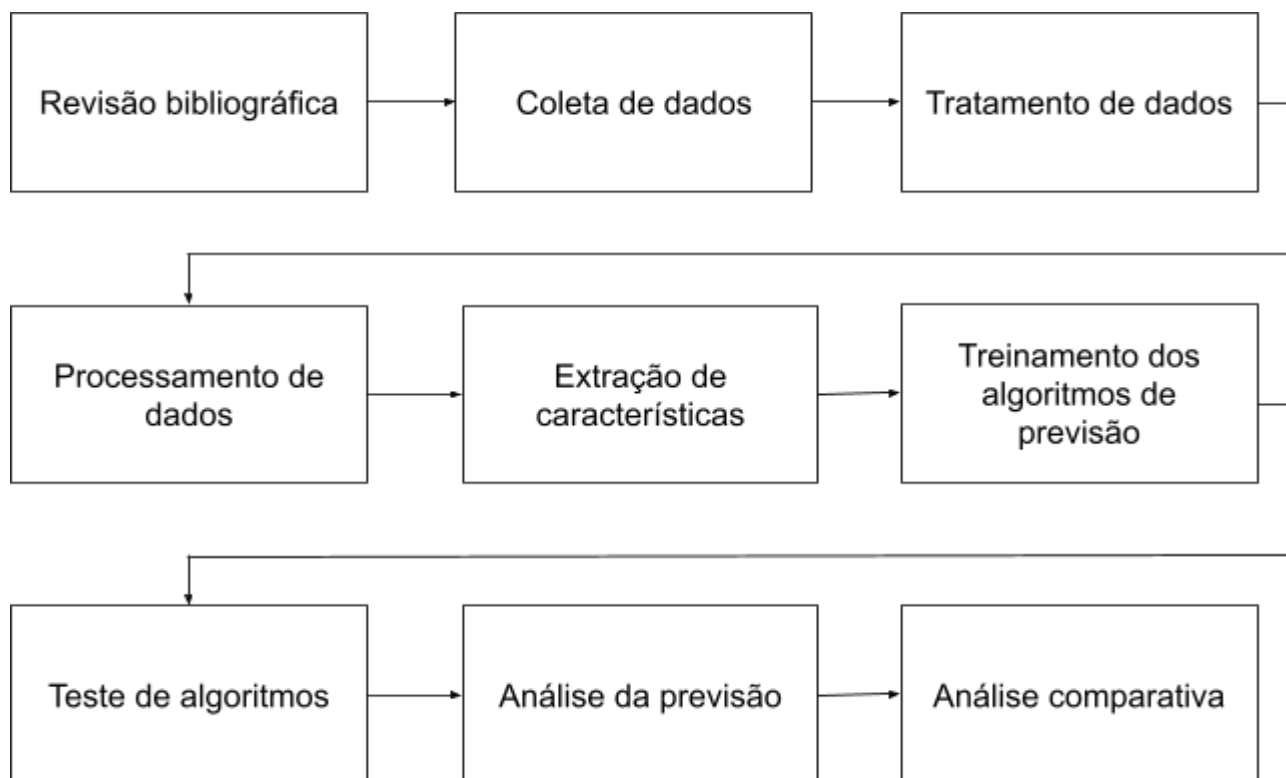
Com base na Figura 6, podem ser identificadas quatro categorias de demanda: (i) regular, que corresponde a itens de alta rotatividade com baixa variação na quantidade demandada; (ii) intermitente, referente a itens que têm vários períodos sem demanda mas apresentam baixa variação na quantidade quando demandados; (iii) errática, que engloba itens com alta rotatividade e alta variabilidade na quantidade solicitada; e (iv) irregular, caracterizada por itens que passam por muitos períodos sem demanda e possuem alta variabilidade na quantidade quando há demanda. (Syntetos et. AL, 2005).

3 METODOLOGIA DA PESQUISA

A terceira seção foi dividida em quatro etapas para facilitar a análise e a organização do estudo. A primeira etapa concentra-se nos tipos de pesquisa utilizados, detalhando as abordagens metodológicas adotadas para compreender o fenômeno em questão. A segunda etapa descreve o processo de coleta e tratamento dos dados, abordando as técnicas empregadas para assegurar a qualidade e a precisão das informações coletadas, com vistas a obter resultados consistentes para a posterior análise.

Na terceira etapa, detalhamos as bibliotecas e ferramentas computacionais empregadas no desenvolvimento do modelo. Na última etapa, descrevemos o processo de desenvolvimento do modelo computacional, desde a preparação dos dados até a validação dos resultados. A Figura 7 apresenta um resumo das etapas da metodologia empregada.

Figura 7 - Resumo da metodologia utilizada.



Fonte: Elaborado pelo autor (2024)

3.1 Tipo de Pesquisa

Segundo Gil (2017), esta pesquisa é caracterizada por sua natureza quantitativa e pode ser classificada como aplicada, pois busca gerar conhecimento com o objetivo de resolver problemas específicos. Além disso, segundo o mesmo autor, pesquisas quantitativas têm como objetivo analisar, de forma numérica, as tendências, atitudes ou opiniões de uma população com base no estudo de uma amostra.

A pesquisa quantitativa considera que tudo pode ser quantificável, traduzindo em números opiniões e informações para classificá-las e analisá-las. Essa abordagem requer o uso de recursos e técnicas estatísticas, como percentuais e médias, garantindo a precisão dos resultados e evitando contradições no processo de análise e interpretação (Prodanov e Freitas, 2013).

De acordo com Creswell (2021), a pesquisa quantitativa é uma abordagem que procura testar teorias objetivas examinando a relação entre variáveis, as quais são medidas geralmente com instrumentos para que os dados numéricos possam ser analisados com procedimentos estatísticos.

3.2 Coleta e Tratamento de Dados

A coleta de dados foi realizada por meio de um levantamento, modalidade que, segundo Gil (2017), pode servir para o alcance de objetivos tanto acadêmicos quanto profissionais. Desta forma, as análises para comparação entre os métodos de previsão de demanda para uniformes militares foram desenvolvidas por meio de dados coletados da série histórica de fornecimento dos itens da cadeia de fornecimento em estudo. Estes dados foram disponibilizados pela Diretoria de Abastecimento da Marinha, órgão

responsável pelo SAbM e administrador do SINGRA, e refere-se aos registros de demanda de, em média, 2.266 itens de fardamento durante o período de janeiro de 2020 a dezembro de 2023.

Por meio da aplicação da curva ABC, verificou-se que os itens podem ser divididos em três grupos com diferentes graus de impacto no valor total da demanda. O Grupo A, que corresponde a 50% do valor total, é composto por 61 itens, destacando-se como o grupo de maior relevância econômica e, portanto, exigindo atenção especial. O Grupo B, representando 30% do valor total, abrange 182 itens, enquanto o Grupo C, composto pelos 2.852 itens restantes, corresponde a apenas 20% do valor total da demanda. Esses dados evidenciam a concentração de valor em um número reduzido de produtos, reforçando a importância de focar os esforços de gestão nos itens de maior impacto.

Dentro desse contexto, os 61 itens do Grupo A, que representam metade do valor da demanda, merecem uma análise detalhada e cuidadosa. Focar nesses itens permite priorizar recursos e estratégias de maneira mais eficiente, garantindo maior controle sobre os estoques e otimizando o planejamento logístico. O grupo de 182 itens do Grupo B também requer atenção, embora em menor escala, enquanto os itens do Grupo C, com menor valor agregado, podem ser geridos com menor frequência e investimento, garantindo assim uma administração de estoque proporcional ao impacto econômico de cada grupo.

A análise mais específica do Grupo A revelou padrões de consumo distintos entre os itens, conforme sua frequência e regularidade de demanda. Dentre os 61 itens, 5 apresentaram uma demanda regular, caracterizando-se por um consumo constante e previsível, o que facilita o planejamento e reduz a necessidade de ajustes frequentes no estoque. Em contrapartida, os demais itens possuem demandas menos previsíveis, indicando a necessidade de estratégias de gestão diferenciadas para atender de forma eficaz às variabilidades de consumo.

Em contrapartida, a análise mostrou que a maioria dos itens, especificamente 48, possui uma demanda irregular, enquanto 8 itens se classificam como de demanda esporádica. Essa irregularidade indica oscilações significativas no consumo, o que dificulta o planejamento eficiente e pode gerar tanto excesso quanto falta de estoque, dependendo das condições e variabilidades de uso.

3.3 Bibliotecas Utilizadas

Segundo Borges (2010), bibliotecas são compostas por pacotes ou módulos, que são ferramentas para implantar inúmeros soluções computacionais em Python, conforme descrito no Quadro 2.

Quadro 2 - Bibliotecas utilizadas no algoritmo produzido na linguagem Python

Biblioteca	Função
Numpy	Oferece suporte para grandes matrizes, vetores multidimensionais e funções matemáticas de alto nível
Pandas	Oferece estruturas de dados eficientes para facilitar a análise prática de dados rotulados ou relacionais
Scikit-learn	Fornecer ferramentas para análise e modelagem de aprendizado de ML

Biblioteca	Função
Matplotlib	Cria visualizações estáticas, animadas e interativas
Pytorch	Cria redes neurais e aplicações de inteligência artificial
Neuralprophet	Une rede neurais e métodos clássicos para previsões interpretáveis de séries temporais com base em Pytorch
Optuna	Otimiza automaticamente hiperparâmetros de algoritmos de ML
Uuid	Gera um identificadores únicos universais

Fonte: Adaptado de <https://pypi.org/> (2024)

3.4 Construção do Modelo Computacional

Inicialmente, são instaladas e importadas as bibliotecas necessárias para o ambiente de desenvolvimento em Python. Esse procedimento é essencial para garantir que as dependências dos pacotes de terceiros sejam compatíveis com a versão do Python utilizada.

Após procedimentos de instalação e importação iniciais foi definida uma semente aleatória para garantir a reprodutibilidade dos resultados nos diversos modelos. Ademais, foi carregado o arquivo em excel contendo os dados de fornecimento de fardamento. Este arquivo era estruturado com dados distribuídos em três colunas: o identificador do item (PI), o valor total fornecido do item e a data de fornecimento.

Inicialmente, foi realizada a soma da variável alvo para cada data, seguida de uma reamostragem para frequência mensal, com o objetivo de identificar padrões de consumo ao longo do tempo.

Durante a análise e classificação da série histórica mensal dos itens de fardamento foram verificadas como de demanda irregular, refletindo um consumo instável e de difícil previsão em intervalos curtos. No entanto, ao expandir a análise para períodos trimestrais, foi possível categorizar a maioria dos itens de forma mais detalhada, embora uma parcela permanecesse sem classificação, evidenciando a complexidade e variabilidade desse tipo de demanda.

A próxima etapa é a decomposição dos dados em conjunto de treinamento e teste para os modelos utilizados para previsão. O conjunto de treinamento foi composto de 12 períodos trimestrais e o conjunto de teste é formado pelos últimos 4 períodos da série histórica trimestral. Durante essa fase é utilizada a técnica de validação cruzada a fim garantir uma melhor precisão por meio da iteração dos dados de validação, treinamento e teste.

Esta seção se inicia com a configuração dos modelos para posterior execução. O primeiro a ser definido foi a Rede Neural Artificial, Árvore de decisão, regressão linear, amortecimento exponencial e média móvel.

A próxima etapa é iniciada com a otimização dos modelos de amortecimento exponencial e rede neurais artificiais por meio de iteração para melhor escolha do parâmetro α e dos pesos sinápticos, respectivamente.

Os cinco modelos são treinados e testados com os dados segmentados na etapa de decomposição dos dados e são salvos os valores apresentados para posterior aplicação na medida de erro que são calculados por meio de MASE e RMSSE.

Ao término do modelo computacional, foram apresentados os resultados dos cinco modelos preditivos durante o período de teste, utilizando o MASE e o RMSSE como métricas de acurácia para os itens que apresentaram maior volume de vendas. A visualização desta comparação é consolidada por meio de uma tabela em que são apresentadas a média das métricas de erro, o desvio-médio, o máximo e mínimo para cada um dos modelos adotados.

4 ANÁLISE DOS RESULTADOS

Esta seção está organizada em duas etapas para simplificar a análise e a interpretação dos dados. A primeira etapa foi realizada um estudo dos dados para entender a classificação da demanda. A segunda etapa avalia a acurácia dos métodos de previsão de demanda: média móvel, regressão linear, amortecimento exponencial, redes neurais artificiais e árvore de decisão regressiva.

4.1 Classificação da Demanda

A análise e a classificação de demanda são essenciais para a seleção adequada do método de previsão de séries temporais, especialmente em ambientes onde a variabilidade de consumo e a previsibilidade de uso dos itens representam desafios constantes. Essas práticas permitem identificar padrões e características específicas da demanda, facilitando a escolha de modelos que melhor se adaptem ao comportamento dos itens, resultando em previsões mais precisas e um gerenciamento de estoques mais eficiente.

Durante a análise mensal dos 3.095 itens de fardamento, observou-se que todos foram inicialmente classificados como de demanda irregular, representando uma característica geral de consumo instável e difícil de prever em períodos curtos. No entanto, quando analisados em períodos trimestrais, foi possível categorizar 2.733 itens de forma mais detalhada, enquanto 364 itens, cerca de 11,7% dos itens, não puderam ser classificados segundo o método utilizado. Entre os itens classificados, 1.599, aproximadamente 51,6%, foram categorizados como de demanda irregular, sendo a categoria predominante. Em seguida, 805 itens, correspondente a 26,0% foram classificados como de demanda errática, 277 itens, em torno de 8,9% como intermitentes e 50 itens, cerca de 1,6% dos itens, como regulares. Essa distribuição destaca a predominância de itens com comportamentos de consumo pouco previsíveis e variáveis.

A análise trimestral permitiu identificar variações no comportamento da demanda, possibilitando uma classificação mais detalhada para a maioria dos itens. Com 51,6% dos itens classificados como de demanda irregular, ficou claro que a maior parte do estoque analisado requer uma atenção especial para ajustes de reposição e prevenção de faltas ou excessos de estoque. A distribuição dos itens entre as categorias, com 26% erráticos, 8,9% intermitentes e 1,6% regulares, somada aos 11,7% de itens não classificados, evidencia a necessidade de políticas de controle de estoque adaptativas e flexíveis para melhor atender a cada perfil de demanda.

A partir dos resultados da classificação trimestral e da identificação de padrões de demanda, é possível estabelecer políticas de estoque mais eficientes e customizadas para cada tipo de item. Com base na periodicidade e no comportamento identificado para cada categoria de demanda, podem-se implementar estratégias diferenciadas, com foco nos itens de demanda irregular, que representam 51,6% do total classificado.

4.2 Análise dos Resultados

A estatística descritiva apresentada no Quadro 3, revela uma análise detalhada do desempenho dos modelos avaliados, utilizando as métricas RMSSE e MASE para

comparar a precisão e a consistência das previsões. Cada modelo possui valores médios, desvios padrão, mínimos e máximos, fornecendo uma visão abrangente de suas características de erro e variabilidade. Esses dados permitem identificar quais modelos têm um desempenho mais consistente, oferecendo previsões mais precisas. Essa análise é essencial para guiar a escolha do modelo mais adequado às necessidades específicas do estudo.

Quadro 3 - Estatística Descritiva dos Modelos

Modelo de Previsão	RMSSE		MASE	
	Média	Desvio padrão	Média	Desvio padrão
Amortização Exponencial	0,43	0,14	0,48	0,12
Média Móvel	0,46	0,16	0,45	0,12
Neural Prophet	0,43	0,15	0,45	0,12
Regressão Linear	0,45	0,15	0,43	0,11
Árvore de Decisão	0,39	0,14	0,45	0,11

Fonte: elaborado pelo autor (2024)

A análise da média de erro entre os modelos Neural Prophet e Árvore de Decisão em comparação com os modelos Regressão Linear, Média Móvel e Amortização Exponencial revelou diferenças não significativas, indicando desempenhos semelhantes em termos de precisão nas previsões. O Neural Prophet apresentou uma média de erro RMSSE de 0,43 e uma média de erro absoluto MASE de 0,45. Por sua vez, a Árvore de Decisão obteve uma média de 0,39 para RMSSE e 0,45 para MASE.

Esses valores são bastante próximos das médias dos erros dos demais modelos, como a Regressão Linear que apresentou o valor de 0,45 para RMSSE e 0,43 para MASE. O modelo de Média Móvel apresentou o número de 0,46 para RMSSE e 0,45 para MASE. Por último, o modelo de Amortização Exponencial apresentou a grandeza de 0,43 para RMSSE e 0,48 para MASE. Essa proximidade sugere que, em termos de média de erro, não houve uma diferença substancial entre os modelos, indicando uma precisão comparável nas previsões de todos eles.

Além disso, ao observar o desvio padrão das métricas RMSSE e MASE, percebe-se que os modelos mantiveram uma consistência semelhante na variabilidade de erros. O desvio padrão do RMSSE no Neural Prophet foi de 0,15 e de 0,12 para o MASE, enquanto a Árvore de Decisão apresentou desvio padrão de 0,14 para RMSSE e 0,11 para MASE.

Esses números estão em linha com os desvios padrões dos modelos restantes. O modelo de Regressão Linear apresentou valor de 0,15 para RMSSE e 0,11 para MASE, a Média Móvel a quantia de 0,16 para RMSSE e 0,12 para MASE e Amortização Exponencial o número de 0,14 para RMSSE e 0,12 para MASE. Dessa forma, os desvios padrão próximos sugerem uma dispersão de erros parecida, evidenciando uma estabilidade uniforme nas previsões de cada modelo.

5 CONSIDERAÇÕES FINAIS

Esta pesquisa revelou que o modelo computacional proposto neste trabalho foi efetivamente desenvolvido e implementado conforme estabelecido nos objetivos específicos. O desenvolvimento do modelo possibilitou o cumprimento das etapas metodológicas planejadas e a realização de testes para análise de desempenho em termos de acurácia. Apesar de o modelo ter sido executado conforme o esperado, deve-se enfatizar que o ambiente de execução do algoritmo não foi adequado, o que contribuiu para que os resultados não evidenciam diferenças expressivas em termos de acurácia.

Essas limitações nos resultados podem ser atribuídas a fatores como o número restrito de iterações no processo de otimização, causado pela insuficiência de recursos computacionais disponíveis, e ao pequeno volume de dados utilizados. Essas condições restringiram a capacidade do modelo de machine learning de realizar ajustes mais complexos e, consequentemente, de atingir uma acurácia elevada. A limitação de dados também comprometeu a diversidade dos cenários de teste, reduzindo a generalização dos resultados.

Cabe ressaltar que os modelos de machine learning estão em constante evolução, e essa evolução exige ambientes com maior massa de dados e poder de processamento para explorar seu pleno potencial. Esses fatores são cruciais para que os modelos possam realizar iterações mais detalhadas e otimizações mais profundas, aumentando sua capacidade de acurácia e sua aplicação em contextos variados.

Como sugestão para pesquisas futuras, recomenda-se que o modelo computacional desenvolvido seja aplicado em ambientes com maior poder computacional, como é o caso do servidor que hospeda o atual SINGRA, o que permitiria uma quantidade maior de iterações no processo de otimização e o uso de uma base de dados mais ampla. Com esses recursos, o modelo teria potencial para alcançar resultados mais robustos e generalizáveis, aprimorando sua capacidade de acurácia e, assim, contribuindo de forma mais efetiva para a resolução de problemas específicos na área de estudo.

NOTAS

Link para acesso aos dados utilizados na pesquisa e scripts implementados:

https://drive.google.com/drive/folders/16A0gn_owKz50PbREXPn2plunD8thfgQp?usp=drive_link

REFERÊNCIAS

BALLOU, Ronald H. **Gerenciamento da Cadeia de Suprimentos**: Logística Empresarial. Bookman editora, 2009.

BISHOP, Christopher M. **Pattern Recognition and Machine Learning**. 1. ed. New York: Springer, 2006. 738 p. (Information Science and Statistics). ISBN 978-0387310732.

BORGES, Luiz Eduardo. **Python para desenvolvedores**. 2. ed. Rio de Janeiro: Edição do Autor, 2010. ISBN 978-85-909451-1-6.

BOYLAN, John E.; SYNTETOS, Aris A.; KARAKOSTAS, George C. **Classification for forecasting and stock control: a case study**. *Journal of the Operational Research*

Society, v. 59, n. 4, p. 473-481, abr. 2008. Disponível em:
<<https://doi.org/10.1057/palgrave.jors.2602312>> Acesso em: 01 nov. 2024.

BRASIL. Marinha do Brasil. Diretoria de Abastecimento da Marinha. **Manual do Sistema de Informações Gerenciais do Abastecimento (SINGRA)**. Rio de Janeiro, 2007.

_____. Marinha do Brasil. Secretaria-Geral da Marinha. **Normas para Execução do Abastecimento (SGM-201)**. 7. Revisão. Brasília, 2020.

CALSING, Luciana Cristina. **Previsão de demanda combinada a partir de métodos quantitativos e opinião de especialistas**, 2015. Dissertação (Mestrado em Engenharia) - Universidade Federal do Rio Grande do Sul, Porto Alegre.

COOPER, Donald R.; SCHINDLER, Pamela S. **Métodos de Pesquisa em Administração**. 7. ed. Porto Alegre: Bookman, 2003.

CRESWELL, John W.; CRESWELL, J. David. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto** 5. ed. Porto Alegre: Penso, 2021.

DAIBERT, M. S.; BRAGA, J. L.; OLIVEIRA, A. P.; ROCHA, M. N. **Monitoramento de Riscos em Projetos de Software: Uma Abordagem Baseada em Dinâmica de Sistemas e Redes Neurais Artificiais**. Viçosa: Departamento de Informática, Universidade Federal de Viçosa, 2009.

FÁVERO, L. P.; BELFIORE, P. **Manual de análise de dados**. Rio de Janeiro: Elsevier Brasil, 2017.

FÁVERO, L. P.; BELFIORE, P.; TAKAMATSU, R. T.; SUZART, J. **Métodos quantitativos com Stata®**. Rio de Janeiro: Campus Elsevier, 2014

GALAR PASCUAL, D. **Artificial intelligence tools: Decision support systems in condition monitoring and Diagnosis**. Londres, England: CRC Press, 2020

GAMA, João – **Árvores de decisões** –
http://www.dcc.fc.up.pt/~ines/aulas/MIM/arvores_de_decisao.pdf (Acesso em 19 de outubro de 2024).

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 6. ed. São Paulo: Atlas, 2017.

GOMES, Dennis dos Santos. **Inteligência Artificial: Conceitos e Aplicações**. Revista Olhar Científico, Faculdades Associadas de Ariquemes, v. 01, n. 2, p. 234-246, ago./dez. 2010.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001.

JORDAN, M. I.; MITCHELL, T. M. **Machine learning: Trends, perspectives, and prospects**. *Science*, v. 349, n. 6245, p. 255-260, 2015.

KOURENTZES, Nikolaos. **On intermittent demand model optimisation and selection**. *International Journal of Production Economics*, v. 156, p. 180-190, 2014

LOH, Wei-Yin. **Classification and Regression Trees**. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, jan. 2011. Disponível em:
<https://www.researchgate.net/publication/227658748>. Acesso em: 31 out. 2024.

MAKRIDAKIS, Spyros; WHEELWRIGHT, Steven C.; HYNDMAN, Rob J. **Forecasting Methods and Applications**. 3 ed. New York: John Wiley & Sons, 1998.

MARTIN, D.; SPITZER, P.; KÜHL, N. **A new metric for lumpy and intermittent demand forecasts: Stock-keeping-oriented prediction error costs**. arXiv preprint

arXiv:2004.10537, 2020. Disponível em: <https://arxiv.org/pdf/2004.10537>. Acesso em: 31 out. 2024.

MARTÍNEZ-LICORT, Rosmeri; DE LA CRUZ LEÓN, Carlos; AGARWAL, Deevyankar; SAHELICES, Benjamín; DE LA TORRE, Isabel; MIRAMONTES-GONZÁLEZ, José Pablo; AMOON, Mohammed. **Breast Carcinoma Prediction through Integration of Machine Learning Models**. IEEE Access, 2024. DOI: 10.1109/ACCESS.2024.0429000.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. São Paulo: Edusp, 2004.

PELLEGRINI, Fernando Rezende. **Metodologia para implementação de sistemas de previsão de demanda**. 2000. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal do Rio Grande do Sul, Porto Alegre.

PELLEGRINI, Fernando R.; FOGLIATTO, Flávio S. **Passos para Implantação de Sistemas de Previsão de Demanda - Técnicas e Estudo de Caso**. Revista Produção, v. 11, n. 1, p. 43-64, nov. 2001. Disponível em: . Acesso em: 01 nov. 2024.

PINKOWSKI, Márcio Luiz. **Aplicação da inteligência artificial na gestão de projetos: aderência, impacto e grau de desenvolvimento**. Dissertação (Mestrado) - Universidade Nove de Julho, São Paulo, 2023.

PRODANOV, C. C.; FREITAS, E. C. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico**. Novo Hamburgo: Feevale, 2013.

PYTHON PACKAGE INDEX. PyPI – **The Python Package Index**. Disponível em: <https://pypi.org/>. Acesso em: 1 nov. 2024.

RICHARDSON, R. J. **Pesquisa Social: métodos e técnicas**. 3. ed. São Paulo: Atlas, 1999.

RUMELHART, D. E.; MCCLELLAND, J. L. **Parallel distributed processing: explorations in the microstructure of cognition**. Cambridge: MIT Press, 1986.

RUSSELL, Stuart; NORVIG, Peter. **Inteligência Artificial**. Tradução de Regina Célia Simille. 3. ed. Rio de Janeiro: Elsevier, 2013.

SHCHERBAKOV, M. V.; BREBELS, A.; SHCHERBAKOVA, N.; TYUKIN, I.; KAMAEV, V. **A survey of forecast error measures**. World Applied Sciences Journal, v. 24, n. 24, p. 171-176, 2013. Disponível em: <https://www.researchgate.net/publication/281718517>. Acesso em: 31 out. 2024.

SLACK, Nigel; BRANDON-JONES, Alistair; JOHNSTON, Robert. **Administração da Produção**. 8. ed. São Paulo: Atlas, 2018.

SOUZA, G. P.; SAMOBYL, R. W.; MIRANDA, R. G. **Métodos simplificados de previsão empresarial**. Rio de Janeiro: Editora Ciência Moderna, 2008, 180p.

SYNTETOS, Argyrios. **Forecasting of intermittent demand**. 2001. Tese (Doutorado em Filosofia) – Business School, Buckinghamshire Chilterns University College, Brunel University, 2001.

SYNTETOS, Aris A.; BOYLAN, John E.; CROSTON, J. D. **On the categorization of demand patterns**. Journal of the Operational Research Society, v. 56, n.5, p. 495-503,

mai. 2005. Disponível em: < <https://doi.org/10.1057/palgrave.jors.2601841> > Acesso em: 01 nov. 2024.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. **Introduction to Data Mining**. Boston: Pearson Addison Wesley, 2006.

VERÍSSIMO, Andrey; HENNING, Elisa; AMARAL, Claiton; CRUZ, Altair; ALVES, Custodio. **Métodos estatísticos de suavização exponencial Holt-Winters para previsão de demanda em uma empresa do setor metal mecânico**. Revista Gestão Industrial, v. 8, n. 4, 2012. DOI: 10.3895/S1808-04482012000400009.