Mestrado em

MÉTODOS QUANTITATIVOS PARA A DECISÃO ECONÓMICA E EMPRESARIAL

Trabalho Final de Mestrado

RELATÓRIO DE ESTÁGIO

MODELO DE PREVISÃO DE VENDAS APLICADO AO SETOR DE PANIFICAÇÃO

MARIA DA CONCEIÇÃO PERES PACHECO

Orientação: Carlos Daniel Rodrigues De Assunção Santos

Documento especialmente elaborado para a obtenção do grau de mestre

JULHO - 2025

Aos meus avós,

que me inspiraram a chegar até aqui.

RESUMO

O seguinte Trabalho Final de Mestrado provém de um estágio na empresa Gleba-Nossa, Lda. em parceria com o Instituto Superior de Economia e Gestão, no âmbito do Mestrado de Métodos Quantitativos para a Decisão Económica e Empresarial.

O presente trabalho tem como objetivo o desenvolvimento e comparação de diferentes modelos preditivos para a previsão de vendas da empresa, com especial foco na redução de roturas de stock e minimização de desperdício alimentar. Para alcançar este objetivo, foi realizada uma extensa análise estatística, integrando técnicas tradicionais de regressão, modelos de séries temporais e redes neuronais. O estudo baseou-se em dados históricos diários de vendas segmentados por loja, produto e data, com variáveis explicativas.

Inicialmente, apliquei um modelo de regressão com efeitos fixos para captar a influência das variáveis categóricas e temporais sobre as vendas. Seguidamente, foi utilizado o modelo *Seasonal Autoregressive Integrated Moving Average with Exogenous Variable*, adequado para séries temporais com variáveis exógenas, que permite capturar dependências ao longo do tempo. Por fim, foi implementado um modelo de Redes Neuronais, para explorar a capacidade de algoritmos de *Machine Learning* na previsão de vendas. As previsões foram avaliadas com recurso a métricas estatísticas.

Os softwares utilizados foram *Microsoft Excel*, *Jupyter* e *CPI-Retail*, que é o sistema usado pela empresa.

PALAVRAS-CHAVE: PREVISÃO DE VENDAS; *PANEL OLS*; *SARIMAX*; REDES NEURONAIS.

ABSTRACT

This Master's Final Project results from an internship at the company Gleba-Nossa, Lda., in partnership with the Lisbon School of Economics and Management, within the Master's in Quantitative Methods for Economic and Business Decision Making.

The main objective of this project is the development and comparison of different predictive models for sales forecasting at the company, with a special focus on reducing stockouts and minimizing food waste. To achieve this goal, an extensive statistical analysis was conducted, integrating traditional regression techniques, time series models, and neural networks. The study was based on historical daily sales data, segmented by store, product, and date, including explanatory variables.

Initially, a fixed effects regression model was applied to capture the influence of categorical and temporal variables on sales. Subsequently, the SARIMAX model was used, suitable for time series with exogenous variables, allowing the capture of temporal dependencies. Finally, a Neural Network model was implemented to explore the capabilities of Machine Learning algorithms in sales forecasting. Forecast accuracy was evaluated using statistical performance metrics.

The software used throughout this project included Microsoft Excel, Jupyter Notebook, and CPI-Retail, the system used by the company.

KEYWORDS: SALES FORECASTING; PANEL OLS; SARIMAX; NEURAL NETWORKS

GLOSSÁRIO

ACF – Autocorrelation Function

ADAM - Adaptive Moment Estimation

ADF – Augmented Dickey-Fuller test

AIC – Akaike Information Criterion

AR – Auto Regressive

ARIMA – Auto Regressive Integrated Moving Average

B2B – Business to Business

B2C – Business to Consumer

CRISP-DM – Cross-Industry Standard Process for Data Mining

GARCH – Generalized Autoregressive Conditional Heteroskedasticity

I - Integrated

LSTM – Long Short-Term Memory

MA – Moving Average

MAE – Mean Absolute Error

MAPE – Mean Absolute Percentage Error

ML – Machine Learning

OLS – Ordinary Least Squares

PACF – Partial Autocorrelation Function

RMSE – Root Mean Square Error

RNN – Recurrent Neural Networks

SARIMAX – Seasonal Autoregressive Integrated Moving Average with Exogenous Variables

TFM – Trabalho Final de Mestrado

ÍNDICE

ÍNDICE DE TABELAS

ÍNDICE DE FIGURAS

AGRADECIMENTOS

Este foi sem dúvida, para mim, um grande desafio quer académico quer pessoal. Contudo, não seria possível sem todo o apoio e ajuda que tive ao longo desta jornada, apoio este indispensável.

Quero primeiramente agradecer ao meu orientador, Prof. Dr. Calos Daniel Santos, sem a ajuda dele, este projeto não era de todo possível. Fico grata por todas as sugestões, apoio e sabedoria que me foi transmitida.

De seguida quero agradecer à minha companheira de todas as horas, Lea Vaz, por me levantar em momentos que não conseguia sozinha e por me dar forças e manter a positividade que em certos dias era escassa. Este trabalho seria muito mais difícil de realizar sem ela e não posso deixar de dar um agradecimento sincero, obrigada por caminhares e cresceres ao meu lado.

Quero agradecer a todos os meus amigos que me proporcionaram bons momentos nesta fase, e me animaram em momentos mais desafiantes. Por me acompanharem em tardes na biblioteca e em pausas de café maiores do que gostaríamos.

Fico também agradecida aos meus pais, sem eles isto também não era concretizável, foi com eles que aprendi a ser resiliente, persistente e a nunca desistir por muito que a vida nos derrube.

Quero agradecer à empresa Gleba-Nossa por todo o carinho com que me recebeu e a aprendizagem que me proporcionou nestes meses, foram sem dúvida transformadores para mim e fizeram-me crescer muito. Um obrigado especial ao João Antunes por partilhar os seus conhecimentos comigo.

Quero por fim agradecer a todos os "iseguianos" que passaram por mim nesta jornada, esta universidade foi durante cinco anos a minha segunda casa e este foi um percurso que guardo com muito carinho e amor. Um obrigado também aos meus caros colegas tunantes, foi a Tuna Económicas que me deu as melhores formas de lidar com pessoas diferentes e a melhor forma de saber gerir o tempo. Todas as pessoas nos ensinam algo e muitas vezes "um simples gesto vale mais do que mil palavras", que a empatia entre o ser humano nunca se perca.

1. INTRODUÇÃO

A execução deste Trabalho Final de Mestrado (TFM) foi realizada no âmbito do mestrado de Métodos Quantitativos Para a decisão Económica e Empresarial (MQDEE). Procede da realização de um estágio em colaboração com a empresa Gleba Nossa, Lda. e com o Instituto Superior de Economia e Gestão (ISEG). Este estágio teve a duração de cinco meses, desde o dia dois de fevereiro a trinta de junho, do presente ano.

Este trabalho tem como principal propósito desenvolver um modelo preditivo de vendas, com o apoio de linguagem de programação, de forma a otimizar o plano de produção da empresa e a gestão de stock para reduzir o desperdício alimentar. De momento a empresa executa o plano de produção manualmente em ficheiros de *Microsoft Excel,* o que torna este processo demorado, pouco prático e nada eficiente.

Nos dias em que vivemos, encontramos uma enorme competitividade nos mercados, existe uma elevada quantidade de dados, mas nem sempre um bom método para os trabalhar. É, assim, responsabilidade de cada empresa retirar o melhor dos seus dados, utilizando-os como fonte de informação para criar valor.

Os métodos de previsão que existem são variados e a falta de previsões precisas pode resultar em desperdício, elevados custos, escassez de produtos, prejudicar a experiência do cliente e reduzir a fidelidade. Por isso, cabe à empresa escolher o que faz mais sentido explorar consoante as suas necessidades, mas tendo sempre em conta novas atualizações, uma vez que o mercado está em constante mudança.

Para realizar uma boa previsão devemos saber como e de que forma agir, para posteriormente tomarmos a decisão mais acertada. As informações sobre os clientes e o mercado são fundamentais porque criam valor para qualquer departamento dentro da empresa. Logo, uma boa previsão seguida de uma eficiente análise de dados é fulcral na vida de um gestor e pode tornar a sua empresa a melhor do mercado.

A Gleba Nossa, Lda. é uma padaria, cafetaria e pastelaria. Possui também exploração agrícola, armazenamento e moagem de cereais. Contudo, tem como sua principal atividade a panificação, fundada em 2016. Encontra-se com cerca de vinte e três lojas, parte de uma matéria-prima de origem 100% portuguesa. Fornece também cereais provenientes de pequenos agricultores com práticas de agricultura sustentável.

A sua visão rege-se por "Produzir o melhor pão, de forma natural e a partir dos melhores cereais portugueses". Esta empresa orgulha-se de ser nacional, a maior parte do seu pão é feito a partir de cereais locais e sustentáveis. O fator que impera é, portanto, o preço e, infelizmente, devido às suas dimensões, solo e clima, o nosso país não é competitivo no grande mercado cerealífero. No entanto, os nossos cereais são riquíssimos em variedade e qualidade. Apesar de não sermos tão produtivos como outros países, podemos apostar em produtos de excelência, baseados na nossa riquíssima tradição.

Os meus objetivos na realização deste estágio passam por analisar e reestruturar a base de dados da empresa, para facilitar a análise de desempenho das lojas e garantir consistência entre os sistemas para facilitar futuras análises de comportamento e planos de produção. Mais pormenorizadamente irei estudar o contexto da empresa, identificar o problema, organizar os dados a utilizar, estudar as séries temporais mais adequadas, aplicá-las e avaliá-las. Por fim, com os resultados obtidos através dos modelos, identificar o modelo com melhor desempenho de previsão de vendas para este projeto, ou seja, o que mais otimiza a gestão de stock e atende à procura de clientes com a mais eximia eficiência.

O meu TFM está dividido em sete capítulos. No começo deste relatório vou realizar uma breve introdução, uma apresentação da empresa e os objetivos do estágio em si. O segundo capítulo é constituído por uma revisão de literatura, onde é estudada a importância de usar modelos preditivos e *Machine Learning* (ML) no contexto da panificação. Também realizar um pequeno estudo sobre séries temporais e como estas são possíveis candidatos ideais para desenvolver modelos preditivos. No seguinte capítulo é explicado a metodologia usada na realização deste projeto. No quarto capítulo irei estudar os dados e realizar alguns testes aos mesmos. Seguidamente, no capítulo cinco irei apresentar e desenvolver o meu modelo. No penúltimo capítulo discutir os resultados obtidos. E por fim, terminar com uma pequena conclusão com propostas de melhorias para futuros trabalhos e limitações do presente trabalho.

2. REVISÃO DE LITERATURA

2.1.Previsão de Vendas para Reduzir Roturas e Desperdício Alimentar

Para elaborar a minha previsão, realizei uma revisão de literatura para estudar os métodos de previsão já existentes de produtos perecíveis e assim entender de que forma o preço é formado no mercado, a relevância de estudar este tema, o que já existe e o que pode ser melhorado. Importa, por isso, conhecer a importância da previsão de vendas em padarias. Existem desafios típicos, como um elevado desperdício, difícil controlo na incoerência das vendas, entre outros. Uma gestão eficiente da produção depende de uma boa previsão de vendas, especialmente devido à curta validade dos produtos. Previsões inadequadas podem gerar desperdício ou falta de produtos, impactando a rentabilidade e a satisfação dos clientes. Em padarias, onde muitos produtos têm um ciclo de vida de apenas um dia, falhas na previsão podem gerar roturas de stock, resultados em perdas de receita e insatisfação do cliente ou excesso de produção, conduzindo a desperdício alimentar. As taxas de rotura continuam a ser consideravelmente elevadas atualmente e uma luta constante de empresas neste meio (Corsten & Gruen, 2004).

Segundo o estudo de Mena et al. (2014), a previsibilidade da procura é um dos fatores-chave para reduzir o desperdício ao longo da cadeia de abastecimento alimentar. Quando a previsão é fraca, há uma tendência a produzir em excesso para "jogar pelo seguro", o que, no caso de padarias, resulta frequentemente em produtos não vendidos no final do dia, gerando um forte desperdício com impacto económico e ambiental significativo.

De acordo com IE (2024), práticas de previsão baseadas em modelos de ML, que utilizam dados históricos e fatores relevantes como o dia da semana, condições meteorológicas ou campanhas promocionais, podem reduzir o desperdício alimentar significativamente em negócios de retalho alimentar de pequena escala. Isso é ainda mais relevante no contexto de uma padaria com muitas lojas, onde a complexidade aumenta com a diversidade de produtos, localização das lojas, perfis de cliente e canais de venda. A utilização de previsões baseadas em dados contribui para solucionar estes problemas.

Além disso, Ciccullo et al. (2021) argumentam que o uso de técnicas avançadas de previsão de vendas está diretamente ligado à sustentabilidade dos sistemas alimentares urbanos, sendo um dos pilares para alcançar práticas mais circulares e reduzir externalidades ambientais associadas à sobreprodução. Estas tecnologias incluem sistemas de previsão de procura, monitorização de qualidade, agrupamento inteligente de produtos e soluções de extensão de vida útil.

A utilização de algoritmos de previsão como modelos de ML, destacam-se como uma ferramenta crítica para antecipar padrões de consumo e evitar roturas de stock ou excesso de produção. Estes ajustam-se a variações diárias sazonais e até comportamentos não lineares induzidos por promoções ou feriados. (Fries & Ludwig, 2024)

Neste trabalho, foram exploradas três abordagens complementares de previsão, regressão linear painel com efeitos fixos, o modelo *Ordinary Least Squares* (OLS), o modelo *Seasonal Autoregressive Integrated Moving Average with Exogenous Variables* (SARIMAX) e *Recurrent Neural Networks* (RNN), cada uma com motivações distintas e alinhadas com diferentes características dos dados.

2.2. Modelo de Regressão OLS

O modelo OLS baseia-se na premissa de que existe uma relação linear nos seus parâmetros. A sua forma mais simples, o modelo de regressão linear simples, é dado pela expressão:

(1) $$Y = \beta_0 + \beta_1 X + u$$

Em que Y representa as vendas, X uma variável explicativa, β_0 e β_1 os coeficientes de regressão, e u o termo de erro. Esta estrutura pode ser expandida para um modelo de regressão múltipla, dada pela seguinte expressão:

(2) $$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + u$$

Este consegue controlar múltiplos fatores que influenciam a variável dependente e não apenas um. É interessante adicionar mais variáveis à equação de forma a explorar e a explicar melhor a variável dependente.

Quando o modelo é estimado com efeitos fixos, permite controlar diferenças inobserváveis específicas dessas unidades. A estimação dos parâmetros é realizada pelo método dos mínimos quadrados, que consiste na minimização da soma dos quadrados dos resíduos. Este modelo oferece elevada interpretabilidade e é adequado para avaliar o impacto isolado de diferentes fatores sobre a variável dependente (Wooldridge, 2016).

2.3. Modelo ARIMA

Um modelo *Auto Regressive Integrated Moving Average* (ARIMA) é uma técnica estatística para prever séries temporais (Box et al., 2008). O ARIMA (p, d, q) com ausência de sazonalidade, caracteriza-se por três componentes, a componente *Auto Regressive* (AR), a de *Integrated* (I) que aborda a não estacionariedade, e a de *Moving Average* (MA). A primeira componente cria uma tendência a partir de valores passados para modelos preditivos.

A auto-regressão funciona como um modelo de regressão em que usamos as desfasagens dos valores passados da série como repressores. A segunda componente tem em mente que a nossa série temporal deve ser estacionária, isto é, que a média e a variação devem permanecer constantes ao longo de um certo período. Esta etapa é fundamental para ajudar o modelo a ajustar-se aos dados e não obter ruído.

A componente das médias móveis concentra-se na relação entre uma certa observação e um erro residual. Associa como as observações atuais estão ou não relacionadas aos erros do passado, pode esta ter uma possível tendência. O modelo ARIMA (p, d, q) é representado por:

$$(3) \qquad \left(1 - \phi_1 B - \cdots - \phi_p B^p\right)(1 - B)^d Y_t = c + \left(1 - \theta_1 B - \cdots - \theta_q B^q\right) a_t$$

Onde $(1-B)^d$ é o operador de diferenciação de ordem d, para d ≥ 1, $\phi_p(B)$ é um operador autorregressivo estacionário, com i = 1, …, p, θ_q(B) é um operador de média móvel invertível, com i = 1, …, q. Por fim a_t é um ruído branco de média zero e c é uma constante.

Sendo também que temos B que é um operador que expressa as relações temporais entre as observações de uma série. Isto faz com que este modelo modere a relação entre os resíduos e as observações do mesmo.

2.4. Modelo SARIMAX

O modelo SARIMAX combina a estrutura do modelo ARIMA, que capta padrões temporais e tendências nas séries, com a inclusão de variáveis exógenas. Esta abordagem é indicada para séries com sazonalidade e autocorrelação. De acordo com Athanasopoulo (2018), o SARIMAX permite uma modelação mais realista da evolução da variável dependente ao longo do tempo, incorporando tanto a inércia da série como fatores externos.

Segundo Arunraj & Ahrens (2015), a integração de variáveis exógenas é crucial e fatores externos influenciam as vendas. Este modelo destaca-se justamente por permitir incluir essas variáveis, o que modelos como o ARIMA não permitem. Este artigo mostra também como dados diários podem ser altamente voláteis, ao incorporar variáveis externas e desfasamentos, ajudamos a suavizar estas variações e a melhorar a previsão.

Este modelo é frequentemente denotado como SARIMAX (p, d, q) (P, D, Q, s) e é representado por:

$$\phi_p(B)\Phi_p(B^S)(1-B)^d(1-B^S)^D y_t = \theta_q(B)\Theta_Q(B^S)a_t + X_t\beta \quad (4)$$

Onde $\phi_p(B)$ e $\theta_q(B)$ são fatores autorregressivos regulares (não sazonais) e de média móvel, respetivamente, $\Phi_p(B^S)$ e $\Theta_Q(B^S)$ são fatores autorregressivos sazonais e de média móvel, o s é o período sazonal, o d e o D são as ordens de diferenciação, o a_t é o

ruído branco. O y_t é a série temporal dependente, o X_t é o vetor de variáveis exógenas no tempo t e por fim, o β é o vetor de coeficientes associados às variáveis X.

2.5. Redes Neuronais Recorrentes

As RNN, especialmente na sua variante *Long Short-Term Memory* (LSTM), são arquiteturas de aprendizagem profunda, *Deep Learning* projetadas para lidar com dados sequenciais. Inspiradas no funcionamento do cérebro humano, estas redes mantêm memória de estados anteriores ao longo do tempo, permitindo capturar dependências temporais de curto e longo prazo (Goldberg, 2017).

As LSTM superam as limitações das RNN tradicionais no tratamento de sequências longas, graças ao seu mecanismo de "portas" que regulam o fluxo de informação. São especialmente úteis quando a relação entre as variáveis explicativas e a variável alvo é altamente não linear ou envolve múltiplos efeitos acumulados.

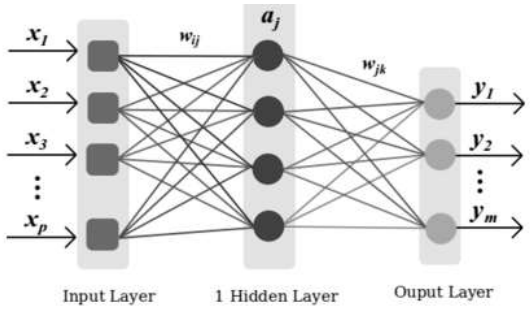

Figura 1 – Rede Neuronal de Camada Única
Fonte: Melli (2015)

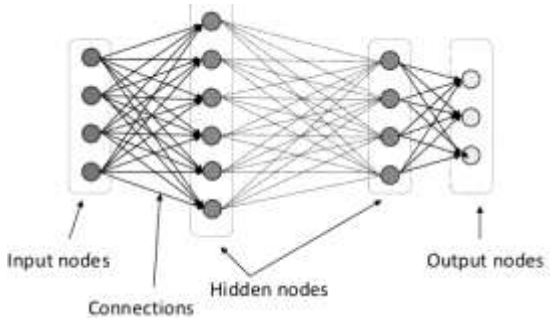

Figura 2 – Rede Neuronal de Múltiplas Camadas
Fonte: Melli (2015)

As redes neuronais com estrutura simples têm apenas um neurónio e uma *input layer,* onde ocorre a ativação dos neurónios da camada oculta. Os $(x_1, \ \ldots \ , x_p)$ são os vetores de entrada, e $(y_1, \ \ldots \ , y_m)$ são os vetores de saída. Por fim, o w_{ij} e w_{jk} são as matrizes

de peso de entrada e saída. Para esta rede ser mais precisa foi criada a rede neuronal com *multi-layer,* estas precisam de ter duas ou mais camadas.

2.6. Critérios de Validação e Avaliação

2.6.1. Erros Absolutos e Relativos

O *Mean Absolute Error* (MAE) caracteriza a alteração entre os valores originais e previstos e é extraído como a média de alteração total do conjunto de dados. Esta é medida como a média dos valores de erro absoluto, não eleva ao quadrado os erros, o que significa que dá o mesmo peso a todos os erros. Esta métrica é importante para medir a robustez de *Outliers*, uma vez que é menos sensível a valores extremos nos dados. Segundo Willmott & Matsuura (2005), esta é uma das métricas mais apropriadas para avaliar o desempenho médio de modelos preditivos, especialmente em contextos em que a presença de valores externos pode distorcer outras métricas. O MAE apresenta a vantagem de ser linear, trata todos os erros proporcionalmente à sua magnitude, o que facilita uma interpretação direta do erro médio em unidades da variável dependente. Esta métrica é representa por:

(5) $$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Onde, y_i é o valor real, $\hat{y}_i$ é o valor previsto e n é o número de observações.

2.6.2. Erro Quadrático Médio

O *Root Mean Squared Error* (RMSE) é uma métrica sensível a grandes erros, penalizando desajustes mais severos, o que a torna útil quando é necessário garantir previsões com baixa variabilidade. De acordo com Chai & Draxler (2014), esta é uma das métricas com mais capacidade de penalizar mais fortemente erros maiores. Esta é uma característica relevante, especialmente em ambientes de previsão onde há desvios

significativos que podem gerar custos operacionais relevantes. Para além disso, é sensível a grandes erros, o que o torna útil para identificar falhas na previsão que possam impactar negativamente o plano de produção em ambientes como o setor de panificação. Matematicamente, é definido como:

(6) $$\sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}$$

Onde, y_t representa o valor real e $\hat{y}_t$ o valor previsto. Valores mais baixos de RMSE indicam melhor desempenho preditivo (Willmott & Matsuura, 2005).

2.6.3. Erro Médio Percentual Absoluto

O *Mean Absolute Percentage Error* (MAPE) expressa o erro médio absoluto em termos percentuais, oferecendo uma medida interpretável da precisão relativa da previsão. É uma das métricas mais utilizadas na avaliação de modelos preditivos, especialmente pela sua interpretação simples em termos percentuais.

No entanto, Kim & Kim (2016) destacam algumas limitações desta métrica, nomeadamente a sua tendência para gerar valores excessivamente elevados em séries com baixos volumes de vendas, o que é comum em produtos perecíveis com forte sazonalidade. Os autores propõem novas abordagens para melhorar a robustez do MAPE, realçando a importância de utilizar esta métrica com precaução, sobretudo em contextos de procura intermitente. Ainda assim, a utilização do MAPE permanece relevante como métrica comparativa, sendo complementada por outras medidas de erro como o RMSE e o MAE. A fórmula utilizada é a seguinte:

(7) $$\frac{100}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right|$$

Onde, y_t representa o valor real e $\hat{y}_t$ o valor previsto. O MAPE é particularmente útil por permitir comparar o desempenho de modelos entre produtos ou períodos com escalas diferentes. No entanto, tem limitações quando $y_t = 0$ uma vez que a divisão por zero torna a métrica indefinida ou instável.

2.7. Ljung-Box Test

Box & Pierce (1970) introduziram um teste "portmanteu" para verificar a $H0: \rho 1 = \rho 2 = \cdots = \rho k = 0$. Este teste estuda a existência de autocorrelação numa série temporal, ou seja, se os resíduos do modelo se comportam como ruído branco. Contudo, (Ljung & Box (1978) propuseram uma versão modificada do teste estatístico. Este analisa se existe autocorrelação nos resíduos até um determinado número de *lags*. A hipótese nula do teste afirma que não existe autocorrelação significativa nos resíduos se a hipótese nula for rejeitada, isso sugere que o modelo não captou completamente a estrutura temporal da série, sendo necessário considerar ajustes no modelo.

Segundo Hyndman & Athanasopoulos (2018), a não rejeição da hipótese nula é um bom indicativo de que o modelo está adequadamente especificado em termos das dependências temporais.

3. METODOLOGIA

Existem múltiplas metodologias para realizar modelos preditivos. Na realização deste trabalho optei por utilizar a metodologia *Cross Industry Standard Process for Data Mining* (CRISP-DM) uma das mais utilizadas em projetos de análise de dados e ML, pela sua flexibilidade, estrutura lógica e aplicabilidade em diversos domínios. Esta metodologia é composta por seis fases: *Bussiness Understanding, Data Understanding, Data Preparation, Modeling, Evaluation e Deployment* (Chapman et al., 2000).

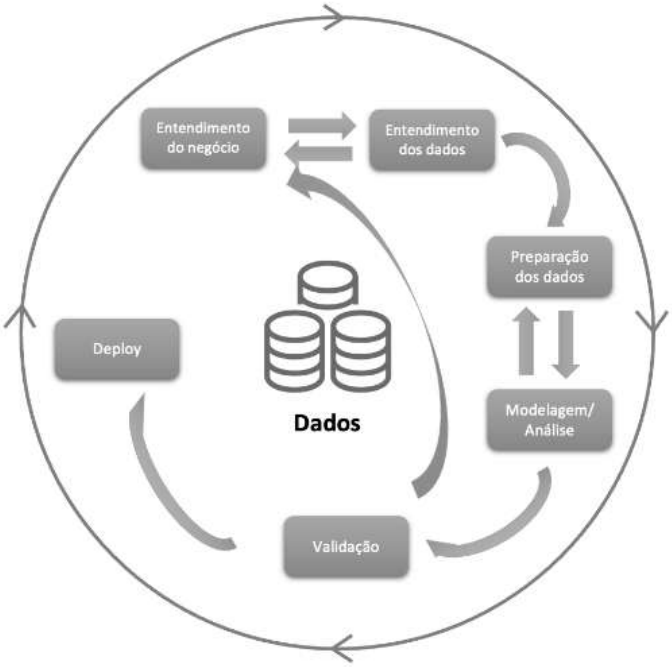

Figura 3 – Fases do Método CRISP – DM
Fonte: Preditiva.ai (2025)

A sua natureza interativa permite avançar e recuar entre etapas consoante os resultados obtidos, o que se revelou especialmente útil num contexto exploratório e preditivo como este. Para este projeto foram aplicadas as cinco primeiras fases, uma vez que a implementação final do modelo em ambiente produtivo não foi realizada.

Na primeira fase, *Business Understanding,* foi executada uma imersão no funcionamento da empresa em estudo, com especial atenção à dinâmica das lojas, canais de venda, sazonalidade do consumo e vários desafios operacionais, como as roturas de stock ou quebras. O objetivo central definido passou por realizar uma previsão de vendas

diárias por loja e produto, de forma a apoiar decisões de planeamento, produção e gestão de stock.

Em segundo, temos a fase de *Data Understanding,* esta passou por uma análise à base de dados fornecida pela empresa, que continham dados de carácter diário de vendas por loja e produto, com variáveis como canal de venda, data e existência de desconto. No decorrer desta fase foram utilizadas ferramentas como o *CPI-Retail* e *Microsoft Excel.*

A terceira fase, *Data Preparation*, envolveu forte precaução e múltiplas tarefas, como limpeza de dados, criação de variáveis, definição de variáveis binárias, geração de variáveis desfasadas, *lags*, transformação logarítmica das vendas e reestruturação dos dados em painel (loja-produto-data). Foram também tratados *missing values* e normalizados os formatos de data e categorização. Esta fase teve um papel crucial na viabilização dos diferentes modelos tratados. Nesta e na próxima fase, foi utilizada a ferramenta *Jupyter*.

Seguidamente, na fase *Modeling,* foram aplicados três modelos preditivos distintos. Modelos estes como regressão OLS, modelo SARIMAX e Redes Neuronais. Os dados foram divididos em conjunto de treino e teste, com previsão *out-of-sample* e *in-sample*.

Na quinta fase*, Evaluation,* os modelos foram avaliados com base em três métricas principais, RMSE, MAE, MAPE. Desta forma, conseguimos avaliar como os modelos reagem aos dados consoante o objetivo principal deste trabalho.

Por fim, *Deployment,* a fase onde devemos implementar o modelo e colocá-lo em prática, devendo ficar exposto para acesso e assim agregar valor para a empresa.

4. RESOLUÇÃO DO CASO EM ESTUDO

4.1 Interpretação e Preparação dos Dados

Para iniciar o meu caso em estudo recorri aos dados disponibilizados pela empresa Gleba-Nossa, Lda., no horizonte temporal entre 1 de janeiro de 2024 e 30 de abril de 2025, com carácter diário. Numa fase inicial, decidi realizar uma análise descritiva dos dados, para posteriormente encontrar o modelo de previsão ideal para o problema em concreto.

Os dados representam o número de vendas que são segmentados por loja e divididos por produto. A notar que duas destas lojas tiveram abertura no presente ano civil.

Para além disto, os dados estão também divididos por canal de venda. O canal de venda é constituído em venda *Business to Business* (B2B), em *Business to Consumer* (B2C), Funcionários (Colaboradores da Gleba-Nossa), em *Online* (compras feitas no site da loja), *Too Good To Go* (aplicação móvel que vende produtos que sobraram) e *Uber Eats* (plataforma de entrega de alimentos).

Por fim, os dados possuem também o facto de o cliente ter ou não desconto na compra de um produto, seja este, desconto de funcionário, promoções temporárias, descontos em alturas festivas, etc.

Atualmente, a empresa utiliza métodos de previsão de vendas tradicionais com a utilização da ferramenta *Microsoft Excel*. Estas são concebidas apenas com base em dados históricos de vendas. Ver exemplo em anexo, figura 17.

Este procedimento é demorado e bastante tendente a erros, especialmente com o crescimento exponencial da marca. Isto também leva a um potencial aumento nos custos devido ao desperdício de produtos de padaria e a uma potencial perda direta de 5% nos ganhos devido a roturas de stock. O desperdício mensal é cerca de 9%, o que realça a necessidade de melhorar a previsão da procura.

Com isto, a Gleba está a passar pela otimização deste procedimento com recurso a modelos estatísticos. O meu objetivo passa por usar alguns modelos preditivos com a ajuda de linguagem de programação. Para isso parti primeiramente de uma organização dos dados.

Após a recolha dos dados, agreguei-os num ficheiro de Excel para uma leitura mais eficiente no *Python*. A seguir a uma organização dos dados, comecei a definir as variáveis do meu modelo para começar a análise estatística.

O meu modelo passa por uma previsão de vendas diárias por loja, produto e dia. Estabeleci como variáveis explicativas os canais de venda, desconto, o dia da semana, a existência de feriados para precaver datas festivas e, por fim, adicionei uma variável com efeito fixo por loja-produto devido à diferente localização das mesmas, com possibilidade de ser localizada num centro comercial, na rua ou em formato quiosque. Criei variáveis desfasadas temporais das vendas, de forma a incorporar dependência temporal e padrões de repetição semanais. Por fim, recorri à criação de variáveis binárias para representar dias de semana, feriados e descontos.

Decidi incorporar estas variáveis independentes uma vez que, de forma teórica e empírica, afetam diretamente a procura e o volume de vendas. Fatores temporais, como dias da semana ou datas comemorativas, têm influência na procura de produto, como por exemplo, o Natal, a Páscoa, o Dia dos Namorados. Estes eventos tendem a aumentar positivamente a procura por serem datas especiais e com mais celebração. Variáveis relacionadas com o produto como presença de desconto e desfasagens das vendas, também foram incorporadas, refletindo o comportamento passado e a persistência da série.

Podemos analisar as variáveis utilizadas neste trabalho na tabela 1.

Tabela 1 – Descrição das Variáveis

Nome da Variável	Descrição da Variável
Vendas	Vendas agregadas por produto, loja e dia
Motivo de Desconto	Variável binária que toma o valor de 1 caso exista desconto
Feriado	Variável binária que toma o valor de 1 caso exista feriado
Canal B2C	Variável dummy que toma o valor de 1 caso a venda seja B2C
Canal Funcionário	Variável dummy que toma o valor de 1 caso a venda seja para um funcionário
Canal Online	Variável dummy que toma o valor de 1 caso a venda seja através do site
Canal TOGOODTOGO	Variável dummy que toma o valor de 1 caso a venda seja em parceria com a TOGOODTOGO
Canal Uber	Variável dummy que toma o valor de 1 caso a venda seja em parceria com a Uber
Dia da semana (Monday)	Variável dummy que toma o valor de 1 caso o dia da semana seja segunda-feira
Dia da semana (Tuesday)	Variável dummy que toma o valor de 1 caso o dia da semana seja terça-feira
Dia da semana (Wednesday)	Variável dummy que toma o valor de 1 caso o dia da semana seja quarta-feira
Dia da semana (Thursday)	Variável dummy que toma o valor de 1 caso o dia da semana seja quinta-feira
Dia da semana (Friday)	Variável dummy que toma o valor de 1 caso o dia da semana seja sexta-feira
Dia da semana (Saturday)	Variável dummy que toma o valor de 1 caso o dia da semana seja sábado
Dia da semana (Sunday)	Variável dummy que toma o valor de 1 caso o dia da semana seja domingo
Lag1	Vendas do mesmo produto na mesma loja no dia anterior
Lag7	Vendas do mesmo produto na mesma loja no mesmo dia da semana anterior

4.1.1. Tratamento de Valores Inconsistentes

Em seguida, identifiquei os valores nulos e valores que não estavam disponíveis em certas observações, *missing values* (NaN), e eliminei-os de forma a não influenciar negativamente os resultados. Foram também removidos registos inconsistentes e duplicados. Para este tratamento dos dados e também ao longo deste projeto usei a modelação que consta na tabela 2, representada de seguida.

Tabela 2 – Modelação

Nome da Biblioteca	Descrição
arch	Estimação de modelos GARCH
linearmodels.panel	Estimação de modelos de regressão de dados em painel com efeitos fixos de robustez
matplotlib	Geração de gráficos e visualizações personalizadas
numpy	Operações matemáticas e vetoriais
pandas	Manipulação e análise de dados estruturados
seaborn	Biblioteca de visualização
sklearn.metrics	Cálculo de métricas de erro
statsmodels	Estimação de modelos estatísticos
statsmodels.tsa.stattools	Ferramentas de séries temporais
statsmodels.tsa.statespace.sarimax	Estimação do modelo SARIMAX com variáveis exógenas
scipy.stats	Cálculos estatísticos complementares
tensorflow	Criação de Redes Neuronais

4.2. Teste de Estacionariedade

Para avaliar a estacionariedade da série temporal foi aplicado o teste de *Dickey-Fuller Augmented (*ADF) às séries de vendas agregadas por loja e produto. A análise foi realizada para todas as combinações loja-produto disponíveis.

Contudo, de acordo com os resultados obtidos em toda a amostra, verificou-se heterogeneidade entre as séries. Algumas séries de vendas revelaram-se estacionárias, enquanto outras não apresentaram características de estacionariedade, reforçando a necessidade de transformação ou diferenciação em casos específicos. De forma a garantir uma apresentação objetiva dos resultados é apresentado o exemplo relativo à "Loja Amoreiras" e ao produto "Pão de Água Médio".

Após este teste, a estatística do teste foi -2.76, com valor-p de 0.0637. Comparando com os valores críticos, observa-se que a hipótese nula de presença de raiz unitária não pode ser rejeitada ao nível de significância de 5%, mas poderia ser rejeitada ao nível de 10%. Isso indica que a série apresenta sinais de não estacionariedade, sendo recomendada a diferenciação da série antes da aplicação de modelos que assumem estacionariedade.

Após a identificação de possível não estacionariedade da série de vendas do produto "Pão de Água Médio" na "Loja Amoreiras", foi aplicada a primeira diferença da série. O teste ADF sobre a série diferenciada resultou de uma forte rejeição da hipótese nula de raiz unitária. Conclui-se, assim, que a série diferenciada é estacionária, permitindo a aplicação de modelos que exigem estacionariedade. A diferenciação também foi corroborada visualmente, com a série transformada com uma média e variância aproximadamente constantes ao longo do tempo. A notar que a empresa mudou de sistema em dezembro, daí haver uma quebra de série no padrão entre os anos.

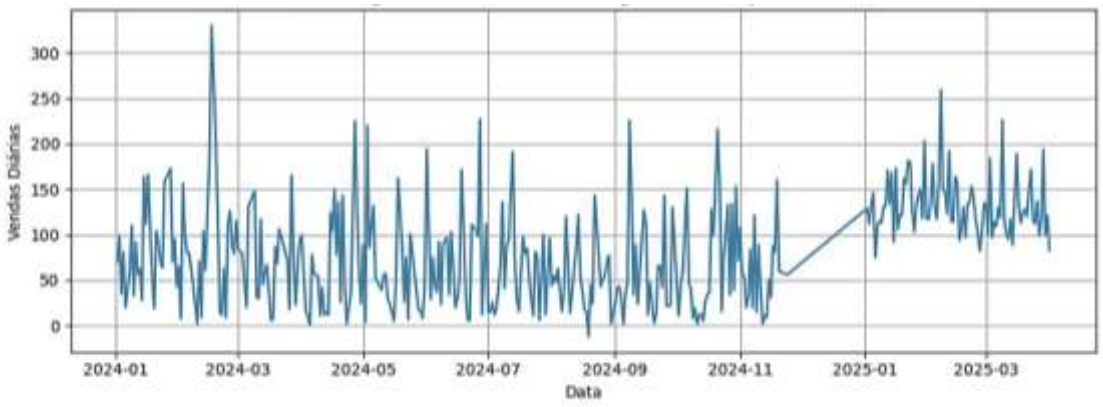

Figura 4 – Evolução Temporal das Vendas Diárias (Unidades) do Produto Pão de Água Médio na Loja Amoreiras

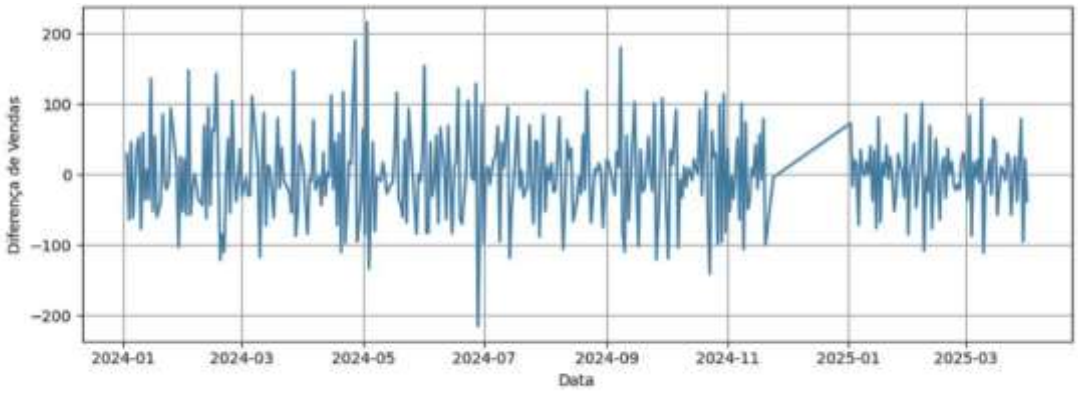

Figura 5 – Série de Vendas Após a Diferenciação de Primeira Ordem

Este resultado confirma a complexidade das séries de vendas de produtos perecíveis em contexto multiloja, uma vez que a estacionariedade varia entre produtos e lojas em função dos padrões de consumo e características específicas de cada unidade de venda. A elevada volatilidade observada nas séries temporais de vendas é uma característica também expectável neste contexto. Ocorreram algumas alterações estruturais que se verificou muito nesta série de vendas. A oscilação entre dias com 50 unidades vendidas e outros com mais de 200, observada na Figura 5, ilustra bem essa realidade, sendo amplificada pelo facto de os dados estarem ao nível diário e se referirem a produtos de consumo imediato, em certos dias as vendas são escassas e em outros dias não, podendo ocorrer, por exemplo, campanhas de marketing, onde as vendas podem disparar. Este tipo de comportamento é compatível com a natureza do produto e do setor, e reforça a necessidade de utilizar técnicas adequadas de suavização e transformação para modelação e previsão das vendas, como será abordado nas secções seguintes.

4.3. Teste à Autocorrelação

Adicionalmente, foram produzidos gráficos de *Autocorrelation Function* (ACF) e *Partial Autocorrelation Function* (PACF). O gráfico ACF da série original exibiu autocorrelação significativa em diversas *lags*, com decaimento lento, enquanto o PACF mostrou um corte acentuado após o primeiro *lag*. Estes padrões são característicos de

processos autorregressivos, sugerindo a adoção de um modelo com diferenciação de ordem 1 (d=1) e um componente autorregressivo de primeira ordem (p=1). A possibilidade de inclusão de um termo de média móvel (q=1) será considerada posteriormente.

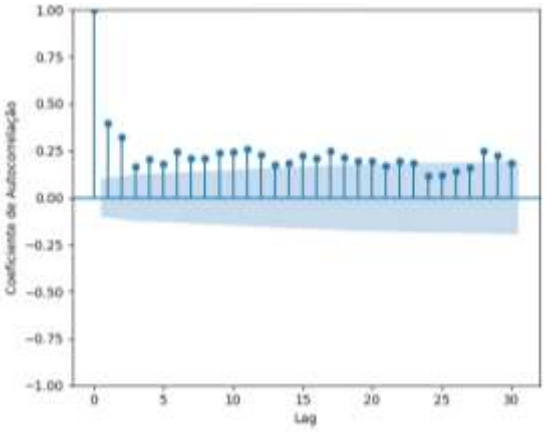

Figura 6 – Autocorrelação das Vendas

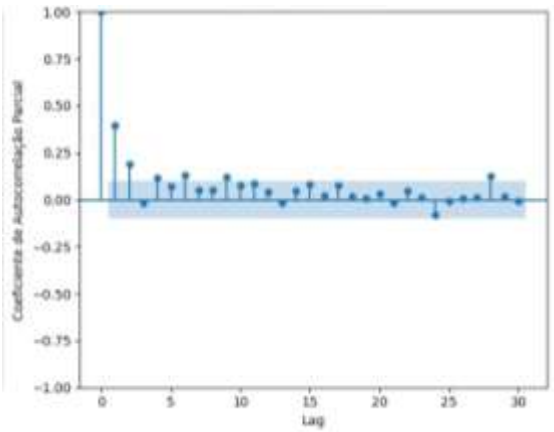

Figura 7 – Autocorrelação Parcial das Vendas

5. RESULTADOS DO CASO EM ESTUDO

5.1. Modelo de Regressão com Efeitos Fixos

5.1.1. Transformação da Variável Dependente

Primeiramente, foi efetuada uma análise da distribuição da variável dependente. Esta análise envolve o cálculo de duas estatísticas fundamentais, a assimetria (*skewness*) e curtose (*kurtosis*), que permitem avaliar, respetivamente, a simetria da distribuição e a concentração dos dados em torno da média.

Os resultados revelaram valores muito elevados tanto para a curtose como para a assimetria (134.8844; 5.5397, respetivamente), indicando uma distribuição altamente distorcida, com a presença de caudas longas e assimetria positiva. Em termos práticos, isto significa que a maior parte das observações se concentra em níveis baixos de vendas, com alguns dias excecionalmente elevados, o que é típico em dados de vendas e pode afetar a *performance* e os pressupostos dos modelos estatísticos.

Para resolver este problema, a solução passou por uma transformação logarítmica da variável dependente. Apliquei "Log (Vendas +1)", uma abordagem clássica e de fácil interpretação em modelos de previsão de séries temporais. Esta transformação suaviza a influência de *outliers*, reduz assimetrias e melhora significativamente a distribuição da variável dependente e aproxima-a da normalidade. Na modelação deste modelo, optei por não ignorar os zeros de modo a preservar a informação sem comprometer a estabilidade numérica do modelo e para não ignorar dados de dias de produtos sem venda, o que podia enviesar o modelo. Após este processo foram novamente analisadas as estatísticas de assimetria e curtose e observamos uma redução substancial de ambos os indicadores, 0.7673 para a curtose e -0.2276 para a assimetria.

5.2. Diagnóstico dos Resíduos e Heterocedasticidade

Para realizar a análise, é relevante caracterizar os dados. Tratam-se de dados temporais e categóricos que temos de estudar como se comportam, como por exemplo, a sua estacionariedade, análise de *outliers* e avaliação da distribuição das variáveis.

5.2.1. Teste White

Para garantir a validade estatística do modelo, procedeu-se à análise dos resíduos estimados. Esta previsão revelou uma concentração em torno de zero, com pouca assimetria à esquerda. Tem uma alta frequência perto de zero o que indica que muitas previsões do modelo estão muito próximas dos valores reais, o que sugere um bom desempenho do modelo. Isto indica uma tendência de superestimação dos valores reais, especialmente para dias com vendas baixas. Ver no anexo figura 18.

Para verificar a presença de heterocedasticidade nos resíduos do modelo estimado, foi aplicado o *Teste White* (White, 1980). Com este teste podemos observar fortes evidências da mesma nos resíduos do meu modelo. Com base no valor-p extremamente baixo, a hipótese nula do teste assume homocedasticidade, ou seja, que a variância dos erros é constante. Os resultados obtidos indicaram uma estatística de teste de 265584,81 com um valor-p próximo de zero. Estes resultados levam à rejeição da hipótese nula ao nível de significância de 1%, o que indica forte evidência de heterocedasticidade nos resíduos.

5.2.2. Modelação com Efeitos Fixos e Erros Robustos

Diante do problema de heterocedasticidade, optou-se por utilizar erros-padrão robustos nas estimações, para garantir maior confiabilidade nas inferências estatísticas realizadas.

O modelo foi estimado com a técnica de regressão de efeitos fixos, OLS, que utiliza a variável logarítmica como dependente. A opção por efeitos fixos permite capturar as heterogeneidades não observadas específicas de cada loja-produto. Para garantir a robustez dos resultados estatísticos, especialmente perante a presença de heterocedasticidade identificada nos resíduos, a estimação foi feita com erros-padrão robustos do tipo *HC3*.

Neste contexto, a presença de efeitos fixos mitiga parte da heterogeneidade entre unidades, mas não elimina a possibilidade de autocorrelação nos resíduos ao longo do tempo também, em parte, observada pelo teste ACF. Por esta razão, optou-se por utilizar erros padrão robustos na estimação principal, que ajustam a variância dos coeficientes

aos grupos definidos e asseguram uma inferência mais fiável mesmo na presença de heterocedasticidade e autocorrelação.

Assim, a estratégia adotada visa garantir a robustez das conclusões estatísticas, diminuindo o risco de decisões enviesadas resultantes da violação da hipótese de homocedasticidade.

Este ajustamento evita a subestimação dos erros-padrão e assegura que os testes de significância dos coeficientes sejam válidos mesmo em presença de variância não constante dos erros.

5.2.3. Modelo GARCH

Após a constatação de heterocedasticidade nos resíduos, evidenciada pelo *Teste White*, foi ajustado um modelo *Generalized Autoregressive Conditional Heteroskedasticity*, GARCH, conforme proposto por Bollerslev (1986). Esta abordagem permite modelar a variância condicional dos resíduos ao longo do tempo. O modelo foi estimado com o método de máxima verossimilhança, com distribuição normal para os erros. O modelo é definido por:

(8) $$\sigma_t^2 = \omega + \alpha \cdot \varepsilon_{t-1}^2 + \beta \cdot \sigma_{t-1}^2$$

Este foi estimado sobre os resíduos obtidos da regressão robusta, com erros-padrão *HC3* e aplicado um GARCH (1,1). Os coeficientes estimados foram ω = 0.5160, α = 0.0782 e β = 0.6882, onde ω representa a variância de longo prazo, α captura o impacto dos choques passados e β indica a persistência da volatilidade passada. A soma $\alpha + \beta$ = 0.7664 confirma a estacionariedade da variância condicional, sendo indicativa de uma volatilidade persistente, mas decrescente ao longo do tempo.

O coeficiente positivo e estatisticamente significativo de α evidencia que choques recentes nas vendas aumentam a variância futura, enquanto o elevado β confirma que esta volatilidade tende a ser persistente, embora decaia ao longo do tempo. Este

comportamento é típico de contextos onde há incerteza e flutuação de procura, como o setor alimentar.

A modelação GARCH revelou-se eficaz para capturar estas oscilações adaptando-se a períodos de maior ou menor volatilidade nas vendas e aumentou a robustez, particularmente relevante para decisões de produção.

Dado que os meus dados são constituídos por observações diárias de vendas por loja e produto numa determinada data, optei primeiro, por um modelo de regressão com efeitos fixos, OLS. Este modelo permite isolar efeitos não observáveis específicos de cada loja-produto, além de incluir múltiplas variáveis explicativas relevantes. Além disso, a transformação logarítmica das vendas facilita a interpretação dos coeficientes, que podem ser interpretados como elasticidades ou mudanças percentuais nas vendas em resposta às variáveis explicativas.

O modelo foi estimado com efeitos fixos por loja-produto e com erros padrão robustos. Foram incluídas desfasagens de vendas (*lags* de 1 e 7 dias), efeitos desfasados do canal de venda (T-7), *dummies* para dias da semana e variáveis binárias para feriados e descontos. Podemos ver a fórmula do modelo estatístico de seguida.

$$(10)\quad \log(Vendas_{i,t}) = \beta_0 + \beta_1 \cdot Desconto_{\,i,t} + \beta_2 \cdot Feriado_{\,i,t} + \beta_3 \cdot Lag1_{i,t} + \beta_4 \cdot Lag7_{i,t} + \sum_{c\in C} \gamma_c \cdot Canal_{c,i,t} + \sum_{d\in D} \delta_d \cdot DiaSemana_{\,d,i,t} + \alpha_i + \varepsilon_{i,t}$$

Para a estimativa do modelo de regressão linear, foi utilizado a totalidade dos dados completos, isto é, considerando todas as lojas, todos os produtos e o respetivo histórico diário de vendas. O objetivo principal foi garantir uma análise global e uma maior robustez dos coeficientes estimados, captando o comportamento das vendas ao nível geral da empresa e por fim irei apresentar os dados *in-sample* e *out-of-sample* com ajuda de métricas de avaliação, que representam as diferentes variações de cada loja-produto.

Para validar a capacidade preditiva, os dados foram divididos em amostra de treino (80%) e teste (20%) com base na distribuição temporal dos dados. Podemos ver os resultados na tabela 3, abaixo representada.

O ajuste *in-sample* revelou um R^2 *overall* de 79,79%, sugerindo boa capacidade explicativa considerando todas as lojas e produtos. O R^2 *within* é de 17,12%, refletindo a variabilidade entre loja-produto ao longo do tempo, e que é relativamente baixo, indicando que o modelo não captura completamente as variações entre a loja-produto. O MAPE apresentou valores elevados devido à existência de valores de vendas muito próximos de zero em várias observações, um fenómeno comum em vendas diárias de produtos perecíveis. Esse fenómeno resulta em erros percentuais muito altos, tornando o MAPE indefinido em algumas situações. A transformação logarítmica das vendas também exacerba esse efeito, uma vez que pequenas variações em vendas muito baixas podem gerar grandes variações no erro percentual. Assim, os valores extremos do MAPE podem ser um reflexo das características específicas dos dados, mais do que uma falha no modelo.

A forte discrepância entre métricas *in-sample* e *out-of-sample* é explicável pela transformação logarítmica aplicada à variável dependente e pela agregação heterogénea de produtos e lojas com diferentes escalas de vendas, dependendo da dimensão e da localização da loja. A transformação logarítmica tende a comprimir valores elevados, a expandir diferenças relativas em valores baixos, o que altera a distribuição dos resíduos e pode levar a uma subvalorização dos erros absolutos quando os valores reais são posteriormente reconvertidos à escala original.

Adicionalmente, a agregação de observações de lojas e produtos com escalas muito diferentes (ex.: lojas com grande volume vs. lojas pequenas, ou produtos com procura diária vs. semanal) faz com que os erros relativos assumam pesos desiguais nas métricas globais. Assim, uma previsão relativamente boa em produtos de grande volume pode ser "anulada" por erros em produtos com vendas baixas, sobretudo após reconversão do logaritmo.

Tabela 3 – Métricas In-Sample e Out-of Sample pelo OLS

Métrica	**In-Sample**	**Out-of-Sample**
RMSE	221167.58	47.66
MAE	621.04	8.26
$\boldsymbol{R^2}$	79.79%	-

No entanto, com o intuito de tornar a análise gráfica e a interpretação visual mais clara e objetiva, optei por apresentar exemplos específicos no corpo da tese. Assim, as previsões foram ilustradas com base num produto e numa loja concretos, que representa um caso exemplificativo, apenas para ajuda visual e melhor perceção.

Realizei uma previsão de trinta dias para o produto "Pão de Água Médio" para a "Loja Amoreiras".

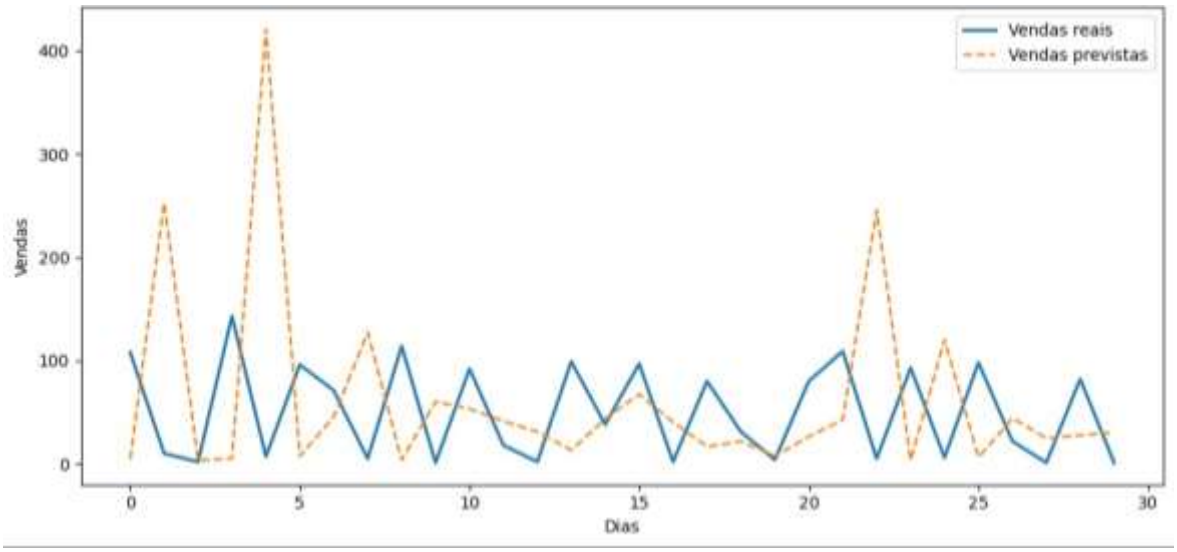

Figura 8 – Previsão pelo Modelo OLS, Vendas Diárias (Unidades)

5.3. Modelo SARIMAX com Variáveis Exógenas

Para complementar a abordagem baseada com efeitos fixos, foi implementado um modelo SARIMAX, com o objetivo de captar a dependência temporal e o efeito de variáveis exógenas no processo de previsão, enquanto acomoda tendências e autocorrelações na série temporal.

Tal como no modelo anterior para aferir a performance preditiva do modelo SARIMAX, foram calculadas métricas de erro tanto *in-sample* como *out-of-sample*, os dados foram divididos em amostra de treino (80%) e de teste (20%). Os resultados evidenciaram um bom ajuste do modelo aos dados. O RMSE e o MAE apresentaram valores reduzidos em ambas as amostras, sendo respetivamente 12.71 e 6.18 *in-sample*, e 9.03 e 4.64 *out-of-sample*. O SARIMAX modela séries temporais univariadas em sequência temporal continua e com isto perde-se parte da variabilidade associada às especificações de cada loja produto. Isso significa que o modelo pode não capturar completamente as diferenças que existem entre os diferentes produtos e lojas, assim é natural que os erros apresentem valores inferiores. Podemos ver estes resultados na tabela seguinte. Além das métricas de erro, também foi realizado o teste de *Ljung-Box* para verificar se os resíduos do modelo apresentam autocorrelação significativa. Este teste é uma análise importante para avaliar a adequação do modelo pois verifica se os resíduos do modelo estão independentes ao longo do tempo. O teste mostrou que não há autocorrelação significativa nos resíduos, o que indica que o modelo conseguiu capturar a estrutura temporal dos dados de forma eficaz. Podemos ver estes dados no anexo.

Tabela 4 – Métricas In-Sample e Out-of Sample pelo SARIMAX

Métrica	**In-Sample**	**Out-of-Sample**
RMSE	12.71	9.03
MAE	6.18	4.64
$\boldsymbol{R^2}$	-	-

Contudo, decidi, para ajuda visual, mostrar uma previsão para um produto e loja em concreto. Este foi aplicado no produto "Pão de Água Médio" na "Loja Amoreiras", sendo estimado com os seguintes parâmetros: SARIMAX (1,1,1) (0,0,0). Ou seja, p = 1, um termo autorregressivo, d = 1, primeira diferenciação da série, q = 1, um termo de média móvel e sem componente sazonal. Esta decisão baseou-se no facto de já existir variáveis explicativas como *dummies* para os dias da semana e feriados, por exemplo, que captam de forma eficaz variações sistemáticas ao longo do tempo. E também, por ser um período de trinta dias, optei por ajustar o modelo sem sazonalidade, focando apenas nos componentes ARIMA e nas variáveis exógenas.

Para calcular esta previsão usei código em Python que gerou posteriormente o resultado da figura 11.

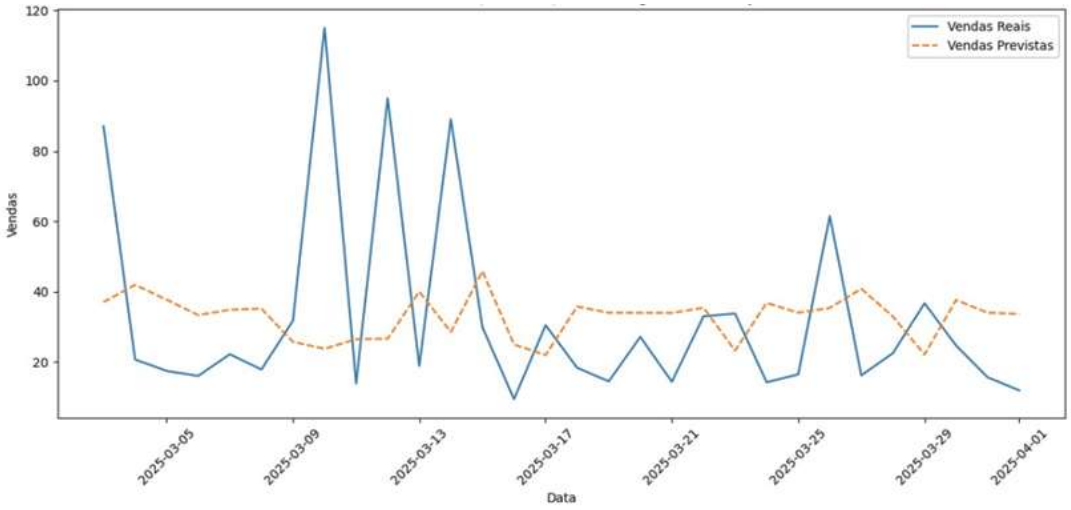

Figura 9 – Previsão do Modelo SARIMAX, Vendas Diárias (Unidades)

5.4. Redes Neuronais

Para complementar os modelos estatísticos tradicionais aplicados anteriormente, foi explorada uma abordagem baseada em redes neuronais recorrentes, através da arquitetura LSTM.

Contudo, para este modelo de RNN, de forma a manter a coerência na comparação entre modelos, optei por agregar os resultados obtidos. A RNN foi treinada individualmente para cada combinação de loja e produto. Esta abordagem foi adotada por razões práticas, permitindo à rede captar melhor os padrões temporais específicos de cada série. No entanto, esta estrutura implicou a produção de métricas de desempenho distintas para cada grupo, o que inviabilizaria uma comparação direta com os modelos estatísticos, se analisadas isoladamente. Assim, para garantir comparabilidade, foram calculadas as médias das métricas de erro obtidas em cada modelo individual de RNN. Estas médias representam uma estimativa agregada do desempenho da RNN e permitem alinhar os resultados com os critérios utilizados nos modelos anteriores, assegurando uma avaliação justa e consistente entre os diferentes métodos de previsão. Podemos ver esses resultados na tabela 5.

Tabela 5 – Métricas In-Sample e Out-of Sample pelas RNN para a Média de Todas as Métricas Criadas

Métrica	**In-Sample**	**Out-of-Sample**
RMSE	4.81	4.54
MAE	3.66	3.66
$\boldsymbol{R^2}$	-	-

Como estes valores são obtidos a partir da média de várias métricas individuais, esta abordagem tende a produzir valores mais baixos para as métricas RMSE e MAE, quando comparados com análises aplicadas a produtos específicos. Tal deve-se, em primeiro lugar, à presença de produtos com baixos volumes de venda, nos quais os erros absolutos, mesmo que pequenos, resultam em métricas reduzidas. Em segundo lugar, a média utilizada não é ponderada pelo volume de vendas, o que significa que produtos residuais têm o mesmo peso que produtos com elevada expressão comercial. Assim, estas métricas não devem ser interpretadas como representativas do desempenho nos produtos mais relevantes, mas sim, como uma medida agregada e geral.

Contudo, decidi calcular também métricas para um exemplo específico, o produto "Pão de Água Médio" na "Loja Amoreiras", utilizado exclusivamente o histórico de vendas sem considerar variáveis exógenas. Os dados foram normalizados com recurso a escalonamento *Min-Max,* garantindo que os valores da variável "Vendas" ficassem no intervalo [0,1], facilitando o processo de aprendizagem da rede. Esta opção deveu-se à elevada complexidade computacional associada ao treino de modelos baseados em redes neuronais, quando aplicados a todas as combinações de loja-produto, bem como à necessidade de uma análise visual clara e objetiva. A análise focada num único exemplo permite uma interpretação mais detalhada do desempenho do modelo em termos de ajustamento às vendas reais.

Os resultados indicaram um desempenho fraco da RNN para este caso específico, com um erro *in-sample* de RMSE igual a 38.29, MAE igual a 32.16 e um MAPE extremamente elevado, o que sugeriu dificuldades do modelo em captar padrões adequados nos dados históricos, mesmo nos dados de treino. O coeficiente de determinação R^2 não foi calculado pois este nem sempre tem a mesma interpretação direta que nos modelos lineares, dado que as redes não modelam uma relação linear entre variáveis.

Optou-se por uma janela temporal de trinta dias para gerar as sequências de treino, refletindo o padrão semanal de consumo. No período *Out-of Sample* observou-se uma deterioração adicional dos resultados com um RMSE igual a 47.44, MAE igual a 44.20 e um MAPE bastante elevado também, evidenciando uma fraca capacidade preditiva fora

da amostra, com erros absolutos e relativos consideráveis. Podemos analisar estes valores na tabela seguinte.

Tabela 6 – Métricas In-Sample e Out-of Sample pelas RNN para o Produto Pão de Água Médio e a Loja Amoreiras

Métrica	**In-Sample**	**Out-of-Sample**
RMSE	38.18	48.42
MAE	31.23	44.44
MAPE	$+\infty$	813.31
$\boldsymbol{R^2}$	-	-

Estes resultados indicam que, apesar da flexibilidade das RNN para captar padrões não lineares, a elevada volatilidade das vendas diárias, a curta série temporal disponível e o reduzido volume de dados por produto e loja tornam o modelo pouco adequado no contexto analisado. A apresentação dos resultados agregados e do exemplo individual permite uma avaliação crítica da adequação do modelo às características específicas do negócio, reforçando que as métricas médias podem ser influenciadas por séries com baixo volume de vendas.

A rede LSTM foi composta por uma única camada com cinquenta unidades e ativação ReLU *(Rectified Linear Unit),* seguida de uma camada *dense* (camada totalmente conectada) e treinada durante cinquenta épocas com o otimizador *Adaptive Moment Estimation,* ADAM.

Contudo, a LSTM foi capaz de captar parte do comportamento da série, mesmo com uma arquitetura simples e sem o uso de variáveis exógenas. É importante reconhecer que a abordagem adotada é introdutória, servindo como base para futuras extensões. Entre as possíveis melhorias, temos a inclusão de variáveis exógenas relevantes, aplicação de arquiteturas alternativas e ampliação da janela de previsão para um horizonte temporal maior. Para calcular esta previsão usei código em Python, que gerou posteriormente o resultado da figura 12.

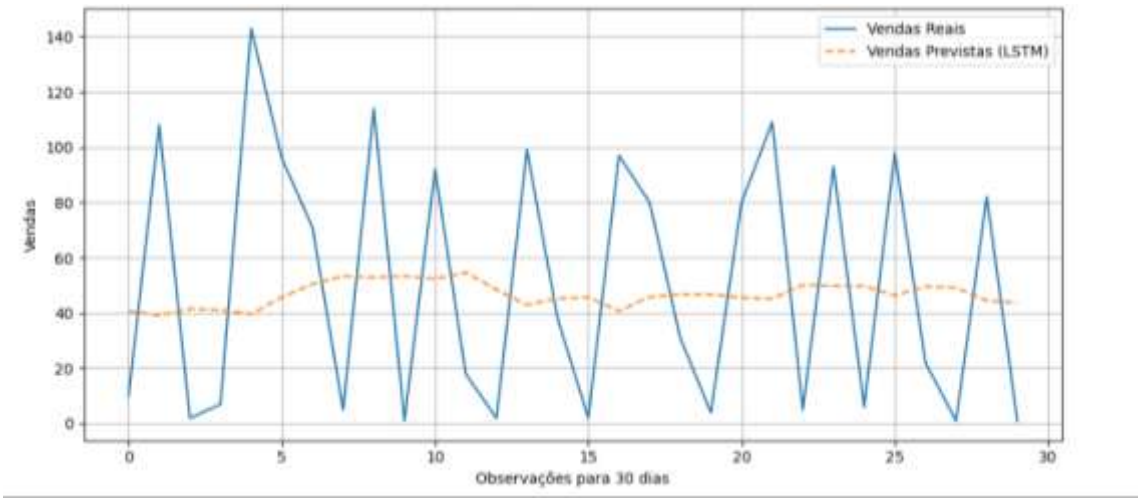

Figura 10 – Previsão Modelo com Redes Neuronais, Vendas Diárias (Unidades)

6. COMPARAÇÃO DE DESEMPENHO DOS MODELOS

Após a aplicação dos diferentes modelos de previsão, OLS com efeitos fixos, SARIMAX e RNN, foi possível observar diferenças claras nos desempenhos e adequação de cada abordagem. De notar que, embora a variável dependente tenha sido transformada por logaritmo natural para estimar os modelos, todas as previsões foram posteriormente reconvertidas à escala original de vendas. Esta reconversão foi fundamental para permitir que as métricas de avaliação fossem interpretadas diretamente em unidades vendidas, assegurando relevância prática e comparabilidade entre modelos. A tabela 6 apresenta as principais métricas de erro obtidas na previsão.

Tabela 7 – Comparação de Métricas de Avaliação dos Modelos

	Regressão OLS		**SARIMAX**		**RNN**	
Métricas	**In-Sample**	**Out-of-Sample**	**In-Sample**	**Out-of-Sample**	**In-Sample**	**Out-of-Sample**
Nº Obs.	424.332	106.084	292.194	73.049	296.900	89.070
RMSE	221167.58	47.66	12.71	9.03	4.81	4.54
MAE	621.04	8.26	6.18	4.64	3.66	3.66
MAPE	+∞	+∞	+∞	+∞	+∞	+∞

Após a implementação dos três modelos de previsão, OLS com efeitos fixos, SARIMAX e as RNN, foi possível identificar diferenças claras no desempenho e adequação de cada abordagem às características dos dados analisados.

O modelo OLS destacou-se pela sua facilidade de interpretação, permitindo compreender o efeito de variáveis explicativas de forma transparente. No entanto, apesar da simplicidade, o modelo revelou limitações importantes. Observou-se que, embora o desempenho *out-of-sample* fosse aceitável, os erros dentro da amostra tenderam a aumentar, especialmente em situações com elevada variabilidade de vendas entre produtos e lojas, refletindo a dificuldade deste modelo em capturar padrões temporais mais complexos. Observou-se também que existe uma acentuada discrepância nos dias com picos de vendas, onde o modelo subestima significativamente os valores reais. Este

comportamento decorre da limitação inerente ao modelo linear, que não considera a dinâmica temporal das vendas.

Em contraste, o modelo SARIMAX evidenciou uma melhor capacidade para modelar a dinâmica das séries temporais, integrando de forma eficiente componentes autorregressivas, de diferenciação e o impacto de variáveis exógenas, este modelo demonstrou robustez na previsão geral. Contudo, foram também identificadas limitações nos casos mais irregulares, como produtos com padrões de venda muito voláteis, onde o SARIMAX teve dificuldades em generalizar.

Já a aplicação das RNN, apesar de explorarem o potencial do *deep learning* para dados sequenciais, mostrou resultados menos consistentes. Devido à necessidade de treinar redes individualmente para cada combinação loja-produto, os resultados foram agregados através de médias das métricas, o que, embora necessário para garantir comparabilidade, pode ter atenuado erros em produtos de maior impacto. Um aspeto relevante a destacar foi a inconsistência na métrica MAPE, que atingiu valores infinitos ou extremamente elevados em todos os modelos. Tal comportamento é explicado pela presença de observações com vendas muito reduzidas ou nulas, distorcendo o denominador desta métrica. Por esse motivo, a análise concentrou-se nas métricas RMSE e MAE, consideradas mais fiáveis no contexto em análise. Esta rede não conseguiu identificar os picos, o que resultou num *underfitting* claro. Esta limitação poderá estar associada à complexidade dos dados, que exigiriam mais camadas ou até mesmo redes mais profundas.

Em suma, os resultados evidenciam que o modelo SARIMAX oferece o melhor compromisso entre precisão e robustez no contexto de previsão de vendas com elevada granularidade. O OLS continua útil pela sua interpretabilidade, enquanto a RNN apesar do seu potencial, exige mais dados e maior estabilidade das séries temporais para alcançar um desempenho competitivo.

Para concluir penso que estes modelos representam um avanço significativo face à metodologia de previsão utilizada atualmente pela empresa. Este processo, além de ser demorado e sujeito a erros, era limitado na sua abrangência, focando-se apenas nos produtos de padaria e desconsiderando variáveis relevantes. Com a implementação e comparação destes modelos foi possível desenvolver uma abordagem mais robusta,

automatizável e adaptável à complexidade do negócio. A notar que ao longo deste trabalho, estendi também os modelos para produtos de mercearia e não apenas de padaria, o que tornou sempre os dados a parecer incoerentes ou os modelos com pior *performance*, uma vez que muitos dos produtos de mercearia são vendidos raramente e em poucas quantidades.

7. CONCLUSÃO E TRABALHOS FUTUROS

O mundo está em constante mudança e inovação e se cada empresa não estiver constantemente a acompanhar esta inovação acaba por não conseguir combater a concorrência. O aumento da Inteligência Artificial no mercado de trabalho tem vindo a ter um crescimento colossal. Gerir uma empresa de bens perecíveis exige ainda uma melhor e mais cuidada previsão e é fundamental tornar esta previsão o mais eficiente, célere e eficaz para uma melhor gestão de stock e redução do desperdício alimentar.

Este estudo permitiu não apenas comparar a eficácia preditiva dos modelos, mas também, demonstrar o valor da complementaridade entre métodos estatísticos e técnicas de ML. As previsões obtidas têm uma aplicação direta e prática, podem auxiliar a tomada de decisão em ambiente de retalho alimentar, este como ajustamento de produção, planeamento de encomendas ou campanhas promocionais. Promove assim uma gestão mais eficiente, sustentável e orientada para os dados.

Durante a realização deste trabalho, a maior limitação que senti foi sem dúvida a organização dos dados. Estes não estavam de todo bem organizados, a empresa mudou de sistema e sem dúvida que estes não estão coesos nem a acompanhar o crescimento exponencial das lojas. Este crescimento foi sem dúvida uma limitação que encontrei também, uma vez que há dados de lojas que já fecharam, lojas que abriram há anos e outras que apenas abriram este ano. Esta falta de correlação afetou bastante este processo de organização de dados.

Para trabalhos futuros, será interessante voltar a realizar esta previsão com a integração de variáveis externas, como condições meteorológicas, eventos locais, etc., que podem ter impacto direto no comportamento do consumidor. Também se poderá utilizar uma aplicação de modelos *ensemble*, que combinem previsões de múltiplos modelos para obter maior robustez e precisão. Esta empresa divide a distribuição em manhã e tarde, seria importante fazer uma avaliação de granularidade distinta, como previsões por horas, de modo a otimizar a produção em turnos.

REFERÊNCIAS BIBLIOGRÁFICAS

Arunraj, N. S., & Ahrens, D. (2015). A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting. *International Journal of Production Economics , 170*, 321-335.

Athanasopoulo, R. J. (2018). *Forecasting: Principles and Practice* (2nd ed.). (OTexts, Ed.) Melbourne.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal, 31*(3), 307-327.

Box, G. E., & Pierce, D. A. (1970). Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association, 65*(332), 1509-1526.

Box, G., Jenkins, G., Reinsel, G., & Ljung, G. (2008). *Time Series Analysis: Forecasting and Control* .

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development, 7*(3), 1247-1250.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide.* CRISP-DM Consortium.

Ciccullo, F., Cagliano, R., Bartezzaghi, G., & Perego, A. (2021). Implementing the circular economy paradigm in the agri-food supply chain: The role of waste prevention technologies. *Resources, Conservation and Recycling, 164*, 105-114.

Corsten, D., & Gruen, T. (2004). Stock-Outs Cause Walkouts. *Harvard Business Review, 82*(5), 26-36.

Fries, M., & Ludwig, T. (2024). 'Why are the Sales Forecast so low?' Socio-Technical Challenges of Using Machine Learning for Forecasting Sales in a Bakery. *Comput Supported Coop Work, 33*, 253-293.

Goldberg, Y. (2017). *Neural Network Methods for Natural Language Processing.* Morgan & Claypool.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice.*

IE, O. (2024). Reducing Food Waste Using Machine Learning Models: Forecasting and Optimization Approaches. *Open Access Journal of Data Science and Artificial Intelligence, 2*(1).

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting, 32*(3), 669-679.

Ljung, G. M., & Box, G. E. (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika, 65*(2), 297-303.

Melli, G. (2015). *Multilayer Freedforward Network. Recuperado de:*

https://www.gabormelli.com/RKB/Multilayer_Feedforward_Neural_Network

Mena, C., Terry, L. A., Williams, A., & Ellram, L. (2014). Causes of waste across multi-tier supply networks: Cases in the UK food sector. *International Journal of Production Economics, 152*, 144-158.

Preditiva.ai. (2025). *Entenda o CRISP-DM, as suas etapas e como de fato gerar valor com esta metodologia.* Recuperado de:

https://www.preditiva.ai/blog/entenda-o-crisp-dm-suas-etapas-e-como-de-fato-gerar-valor-com-essa-metodologia

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica, 48*(4), 817-838.

Willmott, C., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Inter-Research, 30*(1), 79-82.

Wooldridge, J. M. (2016). *Introductory econometrics a modern approach* (6ª ed.).Mason, OH: South-Western Cengage Learning

ANEXOS

Só editar formulas, NÃO editar valores/numeros		04/04/2024				04/04/2024											
						Alcantara											
						28/03/2024	28/03/2024	28/03/2024	Manhã	Tarde	Total	Variação pedido e n-1		Manhã		Tarde	
Artigo	Artigo	Margem B2B	Margem Loja	Mult. Fábrica	Mult. Loja	Venda Manhã	Venda tarde	Quebras fim do dia	Venda Potencial manhã (venda+perdas)x(1+crescimento+margem	Venda Potencial tarde (venda+perdas)x(1+crescimento+margem				Fábrica	Loja	Fábrica	Loja
Pão de Trigo Barbela Grande	PA0008	5	6	1	5	22,00	19,00	14,50	29,46	20,40	50	9	22%	5	25	5	15
Pão Alongado de Mistura de Sementes	PA0110	0	0	0	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Pão Trigo do Alentejo Grande	PA0003	1	5	1	5	8,00	6,00	18,50	12,18	7,14	19	5	38%	2	10	3	5
Trigo, Centeio e Malte Grande	PA0068	1	1	1	5	4,00	3,00	8,00	6,10	4,08	10	3	45%	2	5	5	0
Pão Trigo do Alentejo - Forma Média	PA0005	1	2	1	5	1,00	2,00	9,00	4,04	3,06	7	4	137%	5	0	3	0
Pão Trigo do Alentejo - Forma Gigante	PA0080	0	0	0	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Pão de Espelta - Forma Média	PA0020	1	2	1	5	4,00	3,00	8,00	7,10	4,08	11	4	60%	2	5	5	0
Forma Média Quinoa	PA0098	1	1	1	5	1,00	2,00	6,00	3,04	3,06	6	3	103%	4	0	3	0
Forma Média Sementes	PA0099	1	2	1	5	0,00	2,00	8,00	3,02	3,06	6	4	204%	4	0	3	0
Fruit Bread	PA0713	1	1	1	5	0,00	1,00	8,00	2,02	2,04	4	3	306%	3	0	2	0
Ratatouille	PA0616	1	1	1	6	1,50	0,50	5,00	3,55	1,53	5	3	154%	4	0	3	0
Cenouras, Alperces & Amêndoas	PA0558	1	1	1	6	2,00	0,00	2,50	4,06	1,02	5	3	154%	4	0	3	0
Broa de Milho	PA0016	1	1	1	0	2	3	5	4	4	8	3	70%	8	0	0	0
Centeio Forma Média	PA0100	1	1	1	0	7	2	7	9	3	12	3	36%	13	0	0	0
Croissant	PA0534	1	1	20	0	33	13	10	36	14	50		8%	50	0	0	0
Brioche Médio (1un)	PA0084	1	5	20	0	10,00	28,00	23,00	16,22	27,54	44	8	22%	44	0	0	0
Vegan Alfarroba Médio (1UN)	PA0087	1	5	20	0	3,00	10,00	27,00	9,08	11,22	20	7	56%	21	0	0	0
Vegan Abóbora Médio (1UN)	PA0090	1	5	20	0	7,00	7,00	28,00	15,16	8,16	21	7	52%	22	0	0	0

Figura 11 – Exemplo de Método de *Forecast* Atual da Empresa

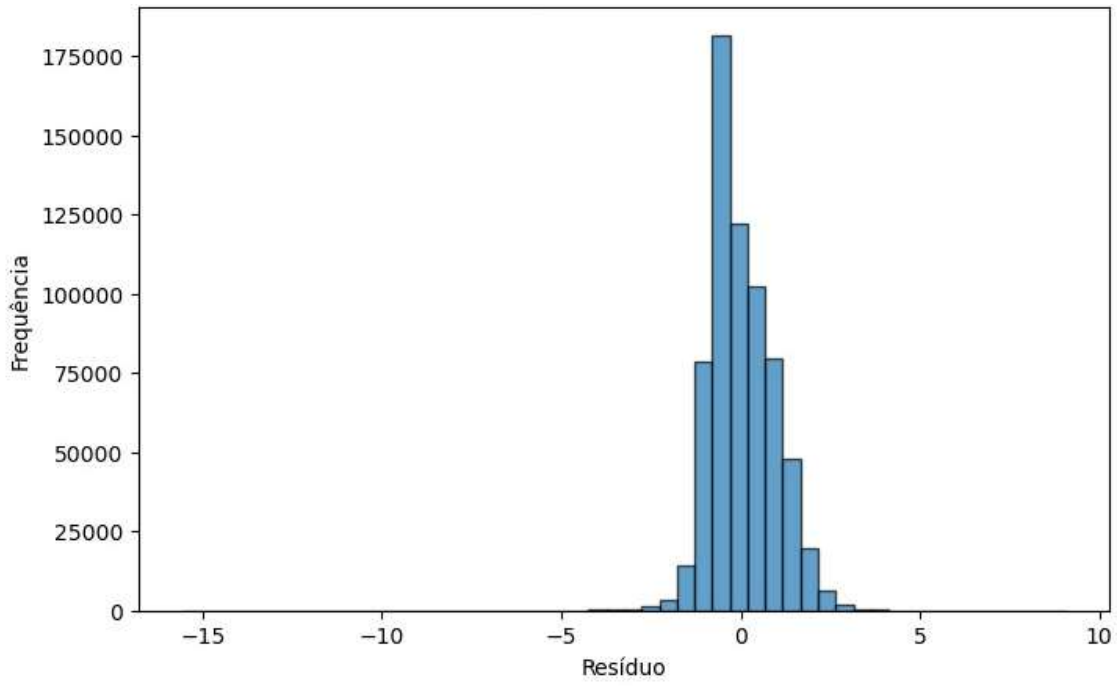

Figura 12 – Histograma dos Resíduos

```
Iteration:      5,   Func. Count:     55,   Neg. LLF: 6.588608660788324e+16
Optimization terminated successfully    (Exit mode 0)
            Current function value: 4320700.036730393
            Iterations: 12
            Function evaluations: 82
            Gradient evaluations: 8
                      Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:                    None   R-squared:                       0.000
Mean Model:             Constant Mean   Adj. R-squared:                  0.000
Vol Model:                      GARCH   Log-Likelihood:           -4.32070e+06
Distribution:                  Normal   AIC:                       8.64141e+06
Method:            Maximum Likelihood   BIC:                       8.64145e+06
                                        No. Observations:               660297
Date:                Sat, Jul 19 2025   Df Residuals:                   660296
Time:                        17:51:52   Df Model:                            1
                            Mean Model
============================================================================
                 coef    std err          t      P>|t|  95.0% Conf. Int.
----------------------------------------------------------------------------
mu          -75.0105      0.626   -119.921      0.000 [-76.236,-73.785]
                          Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega         0.5160      0.364      1.418      0.156     [ -0.197,  1.229]
alpha[1]      0.0782  6.341e-05   1233.240      0.000 [7.807e-02,7.832e-02]
beta[1]       0.6882  1.630e-05  4.223e+04      0.000     [  0.688,  0.688]
============================================================================

Covariance estimator: robust
```

Figura 13 – Modelação GARCH

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:              Log_Vendas   R-squared:                        0.7979
Estimator:                    PanelOLS   R-squared (Between):              0.9979
No. Observations:               530416   R-squared (Within):               0.1712
Date:                Thu, Jul 17 2025   R-squared (Overall):              0.7979
Time:                        18:38:56   Log-likelihood                -6.961e+05
Cov. Estimator:                Robust
                                        F-statistic:                   1.396e+05
Entities:                          24   P-value                           0.0000
Avg Obs:                     2.21e+04   Distribution:                F(15,530401)
Min Obs:                       730.00
Max Obs:                     4.151e+04   F-statistic (robust):          1.41e+05
                                        P-value                           0.0000
Time periods:                     368   Distribution:                F(15,530401)
Avg Obs:                       1441.3
Min Obs:                       1.0000
Max Obs:                       3477.0

                             Parameter Estimates
=============================================================================================
                            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
---------------------------------------------------------------------------------------------
Motivo__de__Desconto_Binario   -0.3702     0.0091    -40.476     0.0000     -0.3881     -0.3522
Feriado_Binario                 0.0479     0.0077     6.2541     0.0000      0.0329      0.0630
Lag1                            0.0193     0.0003     63.650     0.0000      0.0187      0.0199
Lag7                            0.0195     0.0003     64.968     0.0000      0.0190      0.0201
Canal_t7_B2C                    1.1550     0.0043     269.72     0.0000      1.1467      1.1634
Canal_t7_Funcionário            1.3219     0.0104     126.62     0.0000      1.3015      1.3424
Canal_t7_ONLINE                 1.3007     0.0089     145.60     0.0000      1.2832      1.3182
Canal_t7_TOGOODTOGO             1.4746     0.0065     227.03     0.0000      1.4618      1.4873
Canal_t7_UBER                   1.5158     0.0053     283.46     0.0000      1.5053      1.5263
Dia_Semana_Monday               0.1399     0.0049     28.456     0.0000      0.1302      0.1495
Dia_Semana_Saturday             0.2244     0.0049     45.605     0.0000      0.2148      0.2341
Dia_Semana_Sunday               0.1841     0.0049     37.955     0.0000      0.1746      0.1936
Dia_Semana_Thursday             0.1450     0.0049     29.828     0.0000      0.1354      0.1545
Dia_Semana_Tuesday              0.1427     0.0049     29.069     0.0000      0.1331      0.1524
Dia_Semana_Wednesday            0.1326     0.0048     27.365     0.0000      0.1231      0.1421
=============================================================================================
```

Figura 14 – Estimação do Modelo OLS

SARIMAX Results

Dep. Variable:	Vendas	No. Observations:	365243
Model:	SARIMAX(1, 1, 1)	Log Likelihood	-1446796.193
Date:	Sun, 20 Jul 2025	AIC	2893610.387
Time:	13:43:46	BIC	2893707.662
Sample:	0	HQIC	2893638.289
	- 365243		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
Motivo__de__Desconto_Binario	0.3925	0.116	3.382	0.001	0.165	0.620
Feriado_Binario	0.5024	0.084	5.986	0.000	0.338	0.667
Lag1	0.7100	0.000	3538.057	0.000	0.710	0.710
Lag7	0.1310	0.000	308.589	0.000	0.130	0.132
Canal_t7_B2C	-0.9807	0.027	-36.922	0.000	-1.033	-0.929
Dia_Semana_Saturday	1.0331	0.032	32.435	0.000	0.971	1.096
ar.L1	0.0026	0.002	1.683	0.092	-0.000	0.006
ma.L1	-0.9979	0.000	-9682.737	0.000	-0.998	-0.998
sigma2	161.5093	0.051	3148.461	0.000	161.409	161.610

Ljung-Box (L1) (Q):	0.00	Jarque-Bera (JB):	208317964.79
Prob(Q):	0.99	Prob(JB):	0.00
Heteroskedasticity (H):	1.01	Skew:	3.77
Prob(H) (two-sided):	0.02	Kurtosis:	119.75

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).

Figura 15 – Estimação do Modelo SARIMAX

Model: "sequential_1"

Layer (type)	Output Shape	Param #
lstm_1 (LSTM)	(None, 50)	10400
dense_1 (Dense)	(None, 1)	51

Total params: 10,451
Trainable params: 10,451
Non-trainable params: 0

IN-SAMPLE -> RMSE=37.94, MAE=31.15, MAPE=4401800579.75%, R^2=0.07
OUT-OF-SAMPLE -> RMSE=47.30, MAE=43.47, MAPE=820.08%, R^2=-0.06

Figura 16 – Estimação das RNN para um Produto Específico