



PREVISÃO DE DEMANDA DE LONGO PRAZO APLICADA A UMA EMPRESA DO VAREJO DE COSMÉTICOS UTILIZANDO O PROPHET

LONG-TERM DEMAND FORECASTING APPLIED TO A RETAIL COMPANY USING PROPHET

PREVISIÓN DE LA DEMANDA A LARGO PLAZO APLICADA A UNA EMPRESA MINORISTA MEDIANTE PROPHET

Ana Clara Chaves Sousa ^{1*}, Yuri de Almeida Malheiros Barbosa ², Telmo de Menezes e Silva Filho ³, & Thaís Gaudencio do Rêgo ⁴

^{1,2,4} Universidade Federal da Paraíba (UFPB), Centro de Informática

³ University of Bristol, UK

¹ana.chaves@academico.ufpb.br ²yuri@ci.ufpb.br ³telmo.silvafilho@bristol.ac.uk ⁴gaudenciothais@ci.ufpb.br

ARTIGO INFO.

Recebido: 09.07.2024

Aprovado: 19.08.2024

Disponibilizado: 24.08.2024

PALAVRAS-CHAVE: Gestão da Cadeia de Abastecimento; Previsão de Demanda; Séries Temporais; Prophet.

KEYWORDS: Supply Chain Management; Demand Forecast; Time Series; Prophet.

PALABRAS CLAVE: Gestión de la cadena de suministro; Pronóstico de Demanda; Series Temporales; Prophet.

*Autor Correspondente: Sousa, A. C. C.

RESUMO

A disseminação da tomada de decisão baseada em dados tem sido impulsionada pela abundância de informações e pelo aumento da capacidade de processamento computacional. Para apoiar esse processo decisório, é viável extrair conhecimento e realizar previsões por meio da Ciência de Dados. Dentro da Gestão da Cadeia de Abastecimento, um desafio comum é fazer a previsão de demanda utilizando dados históricos. Prever o volume de vendas de uma empresa é complexo. Superestimar a demanda leva a desperdícios de estoque, enquanto subestimar causa ruptura. Neste estudo de caso, será feita a previsão de demanda de longo prazo (30 semanas) para dois canais de vendas diferentes de uma empresa do segmento da indústria e do varejo do mercado de Higiene Pessoal, Perfumaria e Cosméticos. Foi utilizado o algoritmo Prophet. Após implementada a metodologia, os resultados mostraram que a semana 30 apresentou um WAPE de 4% e 5% para os canais Loja e Venda Direta, respectivamente. Ao analisar o erro na trigésima semana para as três categorias de produtos mais comercializados, observou-se que o maior erro registrado foi de 9,36%. Esse resultado sugere que a metodologia empregada alcançou desempenho satisfatório.

ABSTRACT

The spread of data-driven decision making has been driven by the abundance of information and the increase in computer processing capacity. To support this decision-making process, it is feasible to extract knowledge and make forecasts using Data Science. Within Supply Chain Management, a common challenge is to forecast demand using historical data. Forecasting a company's sales volume is complex. Overestimating demand leads to wasted stock, while underestimating causes stock-outs. In this case study, long-term demand forecasting (30 weeks) will be carried out for two different sales channels of a company in the industry and retail segment of the Cosmetic, Fragrance and Toiletry market. The Prophet algorithm was used. After implementing the methodology, the results showed that week 30 had a WAPE of 4% and 5% for the Store and Direct Sales channels, respectively. When analyzing the error in the thirtieth week for the three most marketed product categories, the highest error recorded was 9.36%. This result suggests that the methodology employed achieved satisfactory performance.

RESUMEN

La difusión de la toma de decisiones basada en datos se ha visto impulsada por la abundancia de información y el aumento de la capacidad de procesamiento informático. Para apoyar este proceso de toma de decisiones, es factible extraer conocimientos y hacer previsiones utilizando la Ciencia de Datos. Dentro de la Gestión de la Cadena de Suministro, un reto habitual es la previsión de la demanda a partir de datos históricos. Prever el volumen de ventas de una empresa es complejo. Si se sobreestima la demanda, se desperdician existencias, y si se subestima, se agotan. En este estudio de caso, se realizará una previsión de la demanda a largo plazo (30 semanas) para dos canales de venta diferentes de una empresa de los segmentos industrial y minorista del mercado de Cuidado Personal, Perfumería y Cosmética. Se utilizó el algoritmo Prophet. Después de aplicar la metodología, los resultados mostraron que la semana 30 tenía un WAPE del 4% y del 5% para los canales Tienda y Venta Directa, respectivamente. Al analizar el error en la semana 30 para las tres categorías de productos más comercializadas, el error más elevado registrado fue del 9,36%. Este resultado sugiere que la metodología empleada obtuvo resultados satisfactorios.

1. INTRODUÇÃO

A disponibilidade de grandes volumes e variabilidade de dados, juntamente com a ampliada capacidade de processamento computacional, propiciou o que é denominado como tomada de decisão baseada em dados (do inglês, *data-driven decision making*) (Provost; Fawcett, 2013). *Data-driven decision making* é a prática de realizar escolhas fundamentadas na análise de dados, em contraposição à intuição (Provost; Fawcett, 2013).

Dentre as várias esferas em que decisões fundamentadas em dados podem ser aplicadas, a Gestão da Cadeia de Abastecimento (do inglês, *Supply Chain Management - SCM*) é um exemplo notável. A abrangência da área de SCM compreende logística, transporte, gestão de operações, compras, engenharia, pesquisa e desenvolvimento (Schoenherr; Speier-Peró, 2015).

Dentro desse domínio, inúmeras oportunidades se apresentam, como a análise do comportamento do cliente, a avaliação de tendências e a previsão de demanda (Seyedan; Mafakheri, 2020). Schoenherr e Speier-Peró (2015) realizaram uma pesquisa com o propósito de enfatizar as vantagens derivadas da aplicação da análise preditiva na área, utilizando dados obtidos de profissionais de SCM. Alguns exemplos desses benefícios são: aprimoramento na tomada de decisões, aumento da visibilidade, aprimoramento no gerenciamento de riscos na cadeia de suprimentos, aperfeiçoamento na eficiência da cadeia de suprimentos, otimização dos custos na cadeia de suprimentos, aprimoramento de recursos para o planejamento de demanda, aprimoramento de recursos para o planejamento de vendas e operações.

No entanto, prever o volume de vendas de uma empresa não é uma tarefa trivial. Quando uma previsão superestima a demanda, isso pode afetar a gestão de estoque e causar desperdícios. Por outro lado, quando a demanda é subestimada, ocorre ruptura de estoque (Andrade; Cunha, 2022). Em empresas que atuam simultaneamente nos segmentos de varejo e indústria, o problema é ainda mais complexo. É imperativo determinar quanto e quando adquirir a matéria-prima para a produção, o tempo necessário para fabricação e acabamento, além do volume a ser armazenado e transportado. A previsão de demanda impacta toda a cadeia de suprimentos. Nessas circunstâncias, prever apenas as vendas do próximo mês, por exemplo, é insuficiente.

Com base nisso, o tema deste trabalho é a previsão de demanda de longo prazo (30 semanas à frente), utilizando um método de séries temporais chamado Prophet (Taylor; Letham, 2018). Serão utilizados dados de uma grande empresa de varejo e indústria do mercado de Higiene Pessoal, Perfumaria e Cosméticos (do inglês, *Cosmetic, Fragrance and Toiletry - CFT*). As previsões serão feitas para dois canais de venda: Loja e Venda Direta. A fim de preservar a privacidade e integridade da empresa em questão, seu nome será mantido em sigilo, assim, será referida como Empresa X.

Este trabalho é complementar a um estudo publicado pelos autores, que mostra um comparativo entre algoritmos de *Gradient Boosting* em múltiplos horizontes temporais utilizando os mesmos dados. De acordo com os autores, o modelo que demonstrou maior estabilidade nas análises foi o LightGBM, embora o XGBoost também tenha apresentado um

bom desempenho geral (Sousa et al., 2023). No primeiro estudo, os métodos de *Gradient Boosting* foram utilizados devido à complexidade dos dados disponíveis e à sua flexibilidade para lidar com atributos categóricos e *outliers*. Como os métodos estudados anteriormente são de regressão, a pesquisa atual tem como foco estudar um algoritmo de natureza distinta. No presente estudo, o método Prophet foi utilizado devido à sua capacidade parametrizável de detecção de sazonalidade e à sua versatilidade para fazer previsões em diferentes janelas de horizonte temporal.

A estrutura deste estudo será dividida em seis seções, incluindo a Introdução. A Seção 2 fundamenta o Prophet. Na Seção 3, são apresentados alguns trabalhos relacionados. A Seção 4 detalha a metodologia empregada. Na Seção 5, são discutidos os resultados obtidos. Por fim, a Seção 6 expõe as considerações finais.

2. FUNDAMENTAÇÃO TEÓRICA

O Prophet emprega um modelo de série temporal decomponível, composto por três componentes principais: tendência, sazonalidade e feriados, que são combinados conforme a Equação 1 (Taylor; Letham, 2018):

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t, (2)$$

onde $g(t)$ corresponde à função de tendência que modela as variações não periódicas no valor da série temporal, $s(t)$ representa as alterações periódicas, $h(t)$ representa os efeitos dos feriados e o termo de erro ϵ_t engloba quaisquer mudanças idiossincráticas que não são capturadas pelo modelo (Taylor; Letham, 2018).

No Prophet, a tarefa de previsão é concebida através do ajuste de curva, distinguindo-se dos modelos de séries temporais, que consideram explicitamente a estrutura de dependência temporal nos dados (Taylor; Letham, 2018). O ajuste de curva é feito ajustando uma função à série temporal para prever valores futuros. O modelo estima os componentes de tendência, sazonalidade e feriados de forma separada e, em sequência, faz uma combinação para formar a previsão. Isso é feito minimizando a diferença entre os valores previstos e os valores reais históricos.

Em resumo, o Prophet é um modelo de regressão simples e modular, geralmente eficaz com parâmetros padrão. Ele oferece a capacidade de medir e monitorar a precisão das previsões, permitindo a identificação de melhorias incrementais. A combinação de modelos simples e ajustáveis, juntamente com um monitoramento de desempenho escalável, possibilita que um grande número de analistas realize previsões para diversas séries temporais. Por essa razão, é considerado um modelo de previsão em larga escala (Taylor; Letham, 2018).

3. TRABALHOS RELACIONADOS

A robustez representa um desafio significativo na previsão de demanda de longo prazo, conferindo complexidade substancial a essa tarefa, como destacado por Zhou et al. (2022). A detecção precisa da sazonalidade pelo modelo pode ser comprometida por padrões em dados ruidosos. Essa é a razão da complexidade em obter robustez nesse contexto, como explicado por Wu et al. (2021).

Hasan et al. (2022) utilizou dados de vendas do Walmart para realizar uma análise comparativa de diversos modelos de previsão para os 28 dias seguintes. O estudo investigou modelos tanto paramétricos, quanto não paramétricos, incluindo, ARIMA (Média Móvel Integrada Autoregressiva), LightGBM e Prophet. A análise e comparação de desempenho revelaram que o ARIMA se destacou, apresentando o menor Erro Quadrático Médio (RMSE) registrado, fixado em 1,09 (Hasan et al., 2022).

Foi conduzido um estudo por Lopes (2022) com o intuito de comparar o desempenho de diversos modelos em quatro diferentes horizontes de tempo: o dia subsequente, sete dias à frente, trinta dias à frente e noventa dias à frente. Os modelos tradicionais de previsão de séries temporais, como Naïve, HoltWinters, TBATS e SARIMA, foram avaliados em conjunto com modelos estabelecidos de aprendizado de máquina e aprendizado profundo, tais como MLP (do inglês, *Multilayer Perceptron*), LSTM (do inglês, *Long-Short Term Memory*), XGBoost e Prophet. Entre todos os modelos considerados, o Prophet demonstrou o menor erro no horizonte temporal mais distante, apresentando um Erro Quadrático Médio (RMSE) de 17,2 (Lopes, 2022).

Em uma pesquisa destinada a avaliar a confiabilidade da análise preditiva, o algoritmo XGBoost foi aplicado para gerar previsões diárias de volume de vendas ao longo de uma semana. O Erro Absoluto Médio (MAE) foi equivalente a 2,44 (Baržic et. al., 2022).

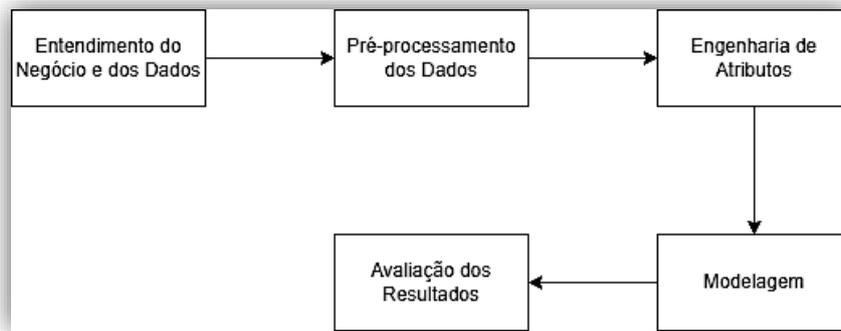
O ponto central de divergência entre a presente pesquisa e estudos anteriores reside no horizonte temporal de predição, sendo que o presente estudo apresenta uma previsão temporal mais longa que as pesquisas mencionadas. Outra diferenciação significativa entre o estudo atual e as pesquisas anteriores é a inclusão de fatores contextuais e características de dados distintos. O estudo abrange diversas campanhas de atração de consumidores e considera produtos que não são classificados como *commodities* essenciais.

4. METODOLOGIA

A abordagem metodológica adotada neste estudo fundamentou-se na Metodologia CRISP-DM (*Cross Industry Standard Process for Data Mining*), que se traduz como Processo Padrão entre Indústrias para Mineração de Dados. Amplamente empregada em projetos de Ciência de Dados, essa metodologia é caracterizada por sua independência em relação a *softwares* ou técnicas específicas (Kelleher; Tierney, 2018).

Originalmente, a estrutura da metodologia CRISP-DM compreende seis fases distintas (entendimento do negócio, entendimento dos dados, preparação dos dados, modelagem, avaliação e implantação) (Kelleher; Tierney, 2018). Entretanto, adaptações foram realizadas para atender às especificidades deste trabalho, visto que não houve implementação até o presente momento (Figura 1).

Figura 1. Metodologia geral aplicada baseada no CRISP-DM



Fonte: Autores.

- **Entendimento do Negócio e dos Dados:** Na Seção 1, o problema a ser solucionado foi estabelecido em conformidade com o contexto empresarial. Subsequentemente, na Seção 3.1, os dados disponíveis para conduzir a previsão de demanda foram apresentados.
- **Pré-processamento dos Dados:** Nesta fase, efetuou-se a preparação completa dos dados, conforme detalhado na Seção 4.4.
- **Engenharia de Atributos:** A Seção 4.5 detalha esta etapa.
- **Modelagem:** Nesta fase, descrita na Seção 4.6, realiza-se a modelagem utilizando o Prophet.
- **Avaliação dos Resultados:** Por fim, conduz-se uma análise para observar os resultados, conforme a Seção 5.

A Seção 4.2 apresenta detalhes mais específicos acerca das ferramentas e do hardware utilizados na realização deste estudo, englobando elementos como linguagens de programação, bibliotecas e ambiente em nuvem.

4.1. BASE DE DADOS

A base de dados empregada nesta pesquisa foi fornecida pela Empresa X no mercado de CFT, abrangendo 22 variáveis e 93,03 milhões de instâncias. Os dados correspondem às vendas no período compreendido entre os anos de 2018 e 2023. Os dados de 2018 foram exclusivamente utilizados na criação de características de *lag* (atraso), enquanto os dados de 2023 foram reservados para a fase de teste (Tabela 1).

Tabela 1. Variáveis disponíveis nos dados originais

Variável	Descrição
Data de Venda	Data em que o produto foi vendido.
Ciclo	Período do ano, sendo que cada ano é composto por 17 ciclos.
Descrição do Ciclo	Definição do tipo de ciclo.
Data do início do ciclo	Data em que o ciclo foi iniciado.
Data do fim do ciclo	Data em que o ciclo termina.
Canal	Canal de vendas, podendo ser Venda Direta ou Loja.
UF	Unidade Federativa do Brasil.
Regional	Regional de campanha.
Categoria	Categoria do produto.
Subcategoria	Subcategoria do produto.
Marca	Marca do produto.
Código Produto	Código do Produto/SKU.
Descrição Produto	Descrição do Produto/SKU.
Volume	Variável Alvo. Quantidade de produtos vendidos.

Valor Praticado	Valor de venda com descontos.
Valor Tabela	Valor de venda sem descontos.
Valor de Desconto	Valor do desconto.
Porcentagem de Desconto	Porcentagem do desconto.
Preço Unitário	Preço unitário considerando os descontos.
Promoção	Indica se o produto está promocionado ou não.
Foco	Indica se o produto está em foco de campanha ou não.
Subfoco	Indica se o produto está em subfoco de campanha ou não.

Fonte: Sousa, A. et al. (2023).

4.2. TECNOLOGIAS UTILIZADAS

Para realizar os experimentos, o ambiente escolhido foi o *JupyterLab* na plataforma *workbench* do GCP (*Google Cloud Platform*). O *hardware* utilizado possui 16 vCPUs e 60 GB de RAM.

A linguagem de programação empregada foi o Python 3.7 para pré-processamento dos dados, análises e construção do modelo, em conjunto com a linguagem de consulta estruturada SQL para importação, exportação e também para pré-processamento dos dados. No ambiente Python, foram utilizadas as seguintes bibliotecas:

- **pandas 1.3.5**, para manipulação dos dados por meio de *dataframes*;
- **numpy 1.21.6**, para execução de operações matemáticas e utilização de *arrays*;
- **matplotlib.pyplot 3.5.3**, para visualização dos dados por meio de gráficos;
- **google.cloud 3.3.3**, para conexão dos dados com o Google BigQuery;
- **sklearn 1.0.2**, para aplicação geral de técnicas de Aprendizado de Máquina;
- **prophet 1.1.1**, para modelagem por meio do método homônimo;
- **statsmodels 0.13.5**, para análise de séries temporais.

4.3. PRÉ-PROCESSAMENTO DOS DADOS

Nesta fase, os dados foram inicialmente formatados conforme as características individuais de cada variável. Após a formatação, procedeu-se à eliminação de dados nulos, resultando em uma redução de 0,22% na base de dados. Adicionalmente, realizou-se uma depuração para excluir produtos, categorias, subcategorias e marcas desativadas, considerando que não é pertinente prever a demanda de produtos que não integram mais o portfólio da empresa. Essa depuração resultou na remoção de 2,12% da base de dados.

Em seguida, os dados foram agrupados no nível temporal de semana, no nível de produto categoria/subcategoria/marca, no nível espacial de regional (sendo que são, no total, seis regionais) e no nível de canal de vendas. Após o agrupamento, a base de dados apresentou aproximadamente 7,34 milhões de instâncias.

4.4. ENGENHARIA DE ATRIBUTOS

Na fase de Engenharia de Atributos, diversas variáveis foram desenvolvidas para servirem como entrada no processo de modelagem preditiva. A Tabela 2 oferece uma descrição detalhada das variáveis criadas, levando em consideração os produtos e suas características, assim como promoções e campanhas.

Tabela 2. Atributos criados considerando produtos, campanhas e promoções

Variáveis de origem	Variáveis criadas	Descrição
Produto, Ciclo, Categoria, Subcategoria, Marca	Quantidade de produtos no portfólio (1 variável)	Quantidade de produtos naquela categoria/subcategoria/marca.
Promoção, Produto, Ciclo, Categoria, Subcategoria, Marca, Regional	Produtos em promoção (1 variável)	Quantidade de produtos em promoção dentro da granularidade.
Porcentagem de Desconto, Ciclo, Produto, Categoria, Subcategoria, Marca, Regional	Quantidade de produtos sem desconto (1 variável criada)	Quantidade de produtos que estão fora de promoção.
Porcentagem de Desconto, Ciclo, Produto, Categoria, Subcategoria, Marca, Regional	Faixa de desconto (3 variáveis criadas)	Quantidade de produtos que pertencem a cada faixa de desconto (10%, 20%, 30%).
Foco, Ciclo, Produto, Categoria, Subcategoria, Marca, Regional	Produtos em foco (1 variável)	Quantidade de produtos em foco dentro da granularidade.
Subfoco, Ciclo, Produto, Categoria, Subcategoria, Marca, Regional	Produtos em subfoco (1 variável)	Quantidade de produtos em subfoco dentro da granularidade.

Fonte: Sousa, A. et al. (2023)

Na Tabela 3, são mostradas as variáveis que calculam a fatia de mercado no ano anterior (do inglês, *market share*). O cálculo foi feito considerando as 52 semanas anteriores (duração de um ano) mais 30 semanas. As 30 semanas são somadas para garantir que somente dados já disponíveis no momento da predição estão sendo utilizados.

Tabela 3. Atributos criados calculando o *market share*

Variáveis de origem	Variáveis criadas	Descrição
Categoria, Volume	<i>Market share</i> de volume da categoria (1 variável)	Volume de vendas de uma categoria em relação ao volume total no período do ano imediatamente anterior.
Categoria, Valor Tabela	<i>Market share</i> do valor tabela da categoria (1 variável)	Valor tabela de uma categoria em relação ao valor tabela total no período do ano imediatamente anterior.
Subcategoria, Volume	<i>Market share</i> de volume da subcategoria (1 variável)	Volume de vendas de uma subcategoria em relação ao volume total no período do ano imediatamente anterior.
Subcategoria, Valor Tabela	<i>Market share</i> do valor tabela da subcategoria (1 variável)	Valor tabela de uma subcategoria em relação ao valor tabela total no período do ano imediatamente anterior.
Marca, Volume	<i>Market share</i> de volume da marca (1 variável)	Volume de vendas de uma marca em relação ao volume total no período do ano imediatamente anterior.
Marca, Valor Tabela	<i>Market share</i> do valor tabela da marca (1 variável)	Valor tabela de uma marca em relação ao valor tabela total no período do ano imediatamente anterior.
Regional, Volume	<i>Market share</i> de volume da regional (1 variável)	Volume de vendas de uma regional em relação ao volume total no período do ano imediatamente anterior.
Regional, Valor Tabela	<i>Market share</i> do valor tabela da regional (1 variável)	Valor tabela de uma regional em relação ao valor tabela total no período do ano imediatamente anterior.

Fonte: Autores.

A Tabela 4 contém 21 variáveis de *lag* que foram criadas. As *lags* são as variáveis que permitem que o algoritmo de previsão considere a variável tempo, ou seja, elas extraem propriedades que podem ordenar os dados (Bergmeir; Benítez, 2012). Em termos mais precisos, esses são atributos que evidenciam retardos temporais. Por exemplo, no presente estudo que emprega dados de vendas em distintos períodos temporais, um exemplo de atributo de defasagem seria o volume de vendas de um mês anterior.

Na Tabela 5, são apresentadas as variáveis de ciclo que foram incorporadas ao método de séries temporais. Outro atributo vinculado às informações do ciclo, utilizado como entrada para o modelo, é a descrição do ciclo; contudo, essa variável não está listada na Tabela 5, uma vez que se trata de um elemento original dos dados.

Tabela 4. Atributos de *lag* construídos

Variáveis de origem	Variáveis criadas	Descrição
Categoria, Volume	<i>Market share</i> de volume da categoria (1 variável)	Volume de vendas de uma categoria em relação ao volume total no período do ano imediatamente anterior.
Categoria, Valor Tabela	<i>Market share</i> do valor tabela da categoria (1 variável)	Valor tabela de uma categoria em relação ao valor tabela total no período do ano imediatamente anterior.
Subcategoria, Volume	<i>Market share</i> de volume da subcategoria (1 variável)	Volume de vendas de uma subcategoria em relação ao volume total no período do ano imediatamente anterior.
Subcategoria, Valor Tabela	<i>Market share</i> do valor tabela da subcategoria (1 variável)	Valor tabela de uma subcategoria em relação ao valor tabela total no período do ano imediatamente anterior.
Marca, Volume	<i>Market share</i> de volume da marca (1 variável)	Volume de vendas de uma marca em relação ao volume total no período do ano imediatamente anterior.
Marca, Valor Tabela	<i>Market share</i> do valor tabela da marca (1 variável)	Valor tabela de uma marca em relação ao valor tabela total no período do ano imediatamente anterior.
Regional, Volume	<i>Market share</i> de volume da regional (1 variável)	Volume de vendas de uma regional em relação ao volume total no período do ano imediatamente anterior.
Regional, Valor Tabela	<i>Market share</i> do valor tabela da regional (1 variável)	Valor tabela de uma regional em relação ao valor tabela total no período do ano imediatamente anterior.

Fonte: Autores.

Tabela 5. Atributos construídos com as características dos ciclos

Variáveis de origem	Variáveis criadas	Descrição
Data de início de ciclo, Data de fim de ciclo	Duração do ciclo (1 variável)	Duração do ciclo daquela semana em dias, comumente 21 dias.
Ciclo	Ano do ciclo, Número do ciclo (2 variáveis)	Informação sobre o ciclo daquela semana.

Fonte: Autores.

Também foram gerados atributos automáticos por meio do método *PolynomialFeatures* da biblioteca *scikit-learn*, com o parâmetro *degree* estabelecido como 2. Algumas dessas variáveis são destacadas na Tabela 6. As variáveis de origem utilizadas para a criação desses atributos automáticos compreenderam: quantidade de produtos no portfólio, quantidade de produtos em promoção, quantidade de produtos sem desconto, quantidade de produtos com 10% de desconto, quantidade de produtos com 20% de desconto, quantidade de produtos com 30% de desconto, quantidade de produtos em destaque e quantidade de produtos em subfoco.

Tabela 6. Atributos automáticos construídos utilizando a biblioteca *scikit-learn*

Variáveis de origem	Variáveis criadas	Descrição
Quantidade de produtos em portfólio	Quadrado de produtos em portfólio (1 variável)	Quantidade de produtos do portfólio ao quadrado.
Quantidade de produtos em promoção	Quadrado de produtos em promoção (1 variável)	Quantidade de produtos em promoção ao quadrado.
Quantidade de produtos sem desconto	Quadrado de produtos fora da promoção (1 variável)	Quantidade de produtos fora da promoção ao quadrado.
Quantidade de produtos nas faixas de 10%, 20%, 30% de desconto	Quadrado de produtos na faixa de 10%, 20%, 30% de desconto (3 variáveis)	Quantidade de produtos na faixa de 10%, 20%, 30% de desconto ao quadrado.
Quantidade de produtos em foco	Quadrado de produtos em foco (1 variável)	Quantidade de produtos em foco ao quadrado.
Quantidade de produtos em subfoco	Quadrado de produtos em subfoco (1 variável)	Quantidade de produtos em subfoco ao quadrado.

Fonte: Autores.

Adicionalmente às variáveis apresentadas na Tabela 6, foram geradas mais 28 variáveis automáticas, utilizando as mesmas características e o parâmetro previamente mencionado do método *PolynomialFeatures*. Essas 28 variáveis consistem nos produtos de cada uma dessas variáveis em pares, como exemplificado pela multiplicação da quantidade de produtos no portfólio pela quantidade de produtos em destaque. Dessa forma, o total resultante foi de 36 atributos automáticos construídos.

4.6. MODELAGEM

Foi implementado um modelo para cada série temporal, sendo que cada série é composta por uma categoria, subcategoria, marca, região e canal específicos. Por exemplo, foi executado um modelo para a categoria C1, subcategoria S1, marca M1, região R1 e canal Venda Direta. Sendo assim, cada combinação é um modelo distinto, totalizando em 1.556 modelos. Os atributos de entrada são:

- 7 variáveis da granularidade da previsão (semana, ciclo, regional, canal, categoria, subcategoria e marca);
- 1 variável original contendo a descrição do ciclo;
- 8 variáveis criadas considerando produtos, campanhas e promoções (Tabela 2);
- 8 variáveis de *market share* construídas (Tabela 3);
- 13 variáveis de *lag* (Tabela 4);
- 3 variáveis construídas considerando informações de ciclo (Tabela 5);
- 6 variáveis criadas de maneira automática (Tabela 6).

Foi feito um treinamento para cada modelo com os dados de 2019 até 2022, utilizando os hiperparâmetros, conforme configuração padrão da biblioteca do Prophet. Os dados do ano de 2023 foram utilizados como teste. O conjunto de teste tem o tamanho de 30 semanas e a *performance* do modelo foi avaliada utilizando esses resultados.

A métrica utilizada para a avaliação dos resultados foi o erro percentual absoluto ponderado (do inglês, *Weighted Mean Absolute Percentage Error* - WAPE). Devido à ponderação, os erros de previsão em produtos com maior volume de vendas seriam superiores aos erros em produtos com menor volume de vendas. O WAPE pode variar de 0% a qualquer número positivo em porcentagem, sendo desejável um valor menor para uma melhor precisão. O cálculo do WAPE (Equação 2):

$$WAPE = \frac{\sum_{t=1}^n |A_t - F_t|}{\sum_{t=1}^n |A_t|}, \quad (2)$$

onde A_t é o valor observado no instante t , F_t é o valor predito no instante t e n é o instante máximo de tempo.

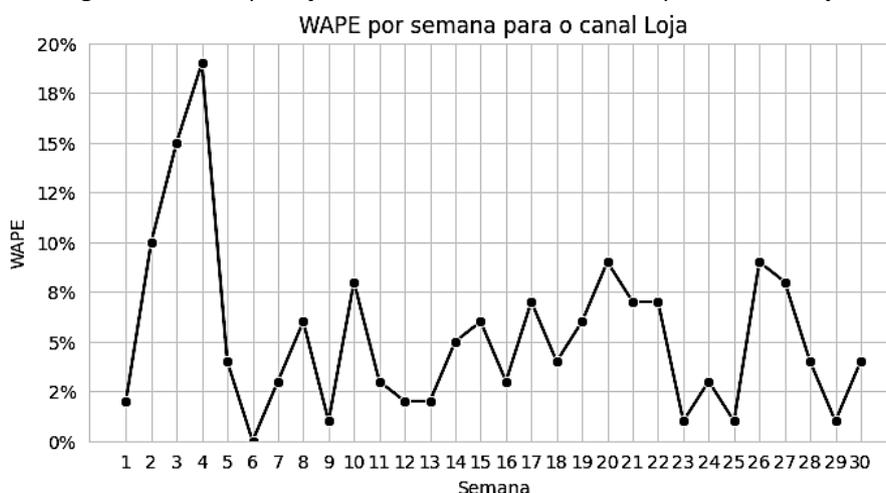
5. RESULTADOS E DISCUSSÃO

Nesta etapa, os resultados serão avaliados utilizando a métrica de erro WAPE, mencionada anteriormente na Seção 3.6. Por motivos de confidencialidade, o volume de vendas realizado e o volume predito não serão informados. Dessa forma, as análises serão restritas aos erros de modelagem. Como a previsão do Prophet foi de até 30 semanas, os primeiros resultados a serem analisados é o quanto o erro pode variar ao longo dessas semanas para os canais Loja e Venda Direta.

Em seguida, serão avaliados os erros na semana 30 (longo prazo) para as 3 categorias, subcategorias e marcas mais vendidas, para cada um dos 2 canais de venda. Novamente, por questão de sigilo dos dados, os nomes reais das categorias, subcategorias e marcas não serão identificados. Sendo assim, as categorias serão nomeadas como C1, C2 e C3, as subcategorias como S1, S2 e S3 e as marcas como M1, M2 e M3.

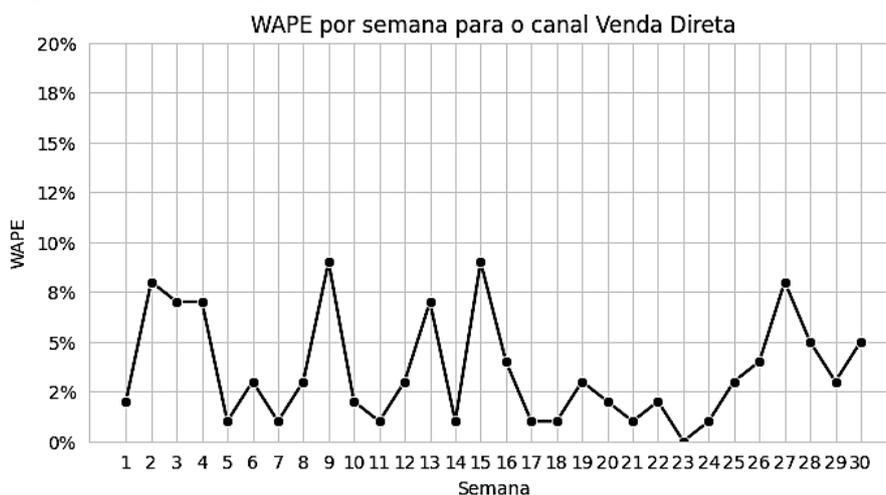
As Figuras 2 e 3 ilustram o WAPE ao longo das 30 semanas de previsão para o canal Loja e para o canal Venda Direta, respectivamente. No canal Loja (Figura 2), os erros em todas as semanas estão abaixo de 10%, com exceção das semanas 2, 3 e 4 que apresentaram os erros de 10%, 15% e 19%, respectivamente. Para o canal Venda Direta, todos os erros, sem exceção, estão abaixo dos 10%. Isso mostra que os resultados para o canal Venda Direta são melhores.

Figura 2. Erro de predição no horizonte de 30 semanas para o canal Loja



Fonte: Autores.

Figura 3. Erro de predição no horizonte de 30 semanas para o canal Venda Direta.



Fonte: Autores.

A Tabela 7 exibe os erros para as 3 categorias mais vendidas na trigésima semana. É possível notar que todos os erros estão abaixo de 10% e a categoria que apresentou os menores erros para ambos os canais foi a C2.

A Tabela 8 mostra os erros para as 3 subcategorias mais vendidas na última semana de predição. Diferente do que foi observado na Tabela 7, quando o recorte de subcategoria é observado, a subcategoria S3 apresenta erros entre 10% e 12%. A segunda subcategoria mais vendida (S2) apresentou o menor WAPE, tanto em Loja, quanto em Venda Direta, assim como aconteceu com a categoria C2.

Tabela 7. Erro de predição para as 3 categorias mais vendidas nos dois canais de venda na semana 30

Canal	Categoria	WAPE
Loja	C1	9,36%
Venda Direta	C1	5,20%
Loja	C2	2,19%
Venda Direta	C2	1,43%
Loja	C3	6,81%
Venda Direta	C3	9,13%

Fonte: Autores.

Tabela 8. Erro de predição para as 3 subcategorias mais vendidas nos dois canais de venda na semana 30

Canal	Categoria	WAPE
Loja	S1	7,89%
Venda Direta	S1	2,65%
Loja	S2	0,70%
Venda Direta	S2	1,82%
Loja	S3	11,71%
Venda Direta	S3	10,35%

Fonte: Autores.

A Tabela 9 ilustra o WAPE para as 3 marcas com maior volume de vendas no horizonte de predição mais distante. Nessa visão, o Prophet apresentou menor estabilidade do que nos recortes de categoria e subcategoria, com todos os erros entre 5% e 16%.

Tabela 9. Erro de predição para as 3 marcas mais vendidas nos dois canais de venda

Canal	Categoria	WAPE
Loja	M1	11,60%
Venda Direta	M1	10,00%
Loja	M2	5,03%
Venda Direta	M2	15,13%
Loja	M3	7,46%
Venda Direta	M3	9,90%

Fonte: Autores.

De maneira geral, pode-se dizer que o modelo apresenta previsões satisfatórias no longo prazo (30 semanas, aproximadamente 7 meses), visto que o erro mínimo é de 0%. Na visão semana a semana, um pequeno pico no canal Loja foi identificado nas semanas 3 e 4, mas a semana 30 apresentou um erro igual a 4%. No canal Venda Direta, o WAPE na última semana de predição foi equivalente a 5%. Ao analisar os erros utilizando as características dos produtos, apesar de ter sido identificado alguns erros acima de 10%, o maior erro encontrado foi de 15,13%, para a marca M2, no canal Venda Direta.

6. CONSIDERAÇÕES FINAIS

Na área da Gestão da Cadeia de Abastecimento, que está associada à Engenharia de Produção, um desafio frequente é realizar previsões de demanda com base em dados históricos. Essa tarefa não é trivial e, especialmente em previsão de demanda de longo prazo, é ainda mais

complexo devido à robustez (Zhou et al., 2022). Este estudo teve como objetivo fazer a previsão de demanda utilizando dados reais de uma empresa parceira em um período de 30 semanas à frente, utilizando o algoritmo Prophet. Após a implementação da metodologia em dois canais de venda distintos, pôde-se observar que os resultados foram satisfatórios. O WAPE do canal Loja foi de apenas 5% e, no canal Venda Direta, foi ainda menor: 4%. Ao comparar os resultados deste estudo com os obtidos no horizonte temporal de longo prazo do trabalho previamente publicado pelos autores, observa-se que, no canal Loja, apenas o XGBoost apresentou um WAPE menor que o Prophet (2,62%). No canal Venda Direta, por outro lado, nenhum dos métodos de *Gradient Boosting* (CatBoost, LightGBM e XGBoost) apresentou um erro menor que 4% (Sousa et al., 2023).

A principal limitação deste estudo consiste na restrição das análises realizadas à avaliação exclusiva do desempenho do Prophet em diversos cenários, utilizando a métrica WAPE. Dessa forma, é necessário realizar outras avaliações para determinar qual modelo seria o mais adequado para a implementação prática. Entre os aspectos cruciais para a compreensão da viabilidade da solução estão a análise do tempo de execução e o custo computacional.

Em futuras pesquisas, melhorias podem ser obtidas ao aplicar o estudo apresentado em conjuntos de dados distintos, proporcionando uma validação adicional. Além disso, executar a otimização de hiperparâmetros a partir de uma faixa de valores e de um número adequado de tentativas pode proporcionar resultados mais satisfatórios. Outra possibilidade é a realização de experimentos com outros algoritmos, incluindo técnicas de aprendizagem profunda, até mesmo uma combinação de modelos, que pode enriquecer as análises e a compreensão do problema abordado

6. REFERÊNCIAS

- Andrade, L. & Cunha, C. B. (2022). *Disaggregated retail forecasting: A gradient boosting approach*.
- Baržic, M., Munitic, N.F., Bronic, F., Jelic, L., & Lešic, V. (2022). Forecasting Sales in Retail with XGBoost and Iterated Multi-step Ahead Method. In *2022 International Conference on Smart Systems and Technologies (SST)* (pp. 153-158).
- Bergmeir, C. & Benítez, J. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191, 192-213.
- Hasan, M., Kabir, M., Shuvro, R., & Das, P. (2022). A Comparative Study on Forecasting of Retail Sales. *arXiv preprint arXiv:2203.06848*.
- Kelleher, J. & Tierney, B. (2018). *Data Science*. MIT Press.
- Lopes, G. (2022). The wisdom of crowds in forecasting at high-frequency for multiple time horizons: A case study of the Brazilian retail sales. *Brazilian Review of Finance*, 20(2), 77-115.
- Provost, F. & Fawcett, T. (2013). *Data Science for Business*. O'Reilly.
- Saltz, J. & Stanton, J. (2017). *An Introduction to Data Science*. SAGE Publications.
- Schoenherr, T. & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential. *Journal of Business Logistics*, 36(1), 120-132.
- Seyedan, M. & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data*, 7(1), 1-22.
- Sousa, A., Rego, T., Barbosa, Y., Menezes, T. (2023). Comparing Gradient Boosting Algorithms to Forecast Sales in Retail. In: *Anais do XX Encontro Nacional de Inteligência Artificial e Computacional* (pp. 596-609).
- Wu, H., Xu, J., Wang, J., & Long, M. (2021). Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Advances in Neural Information Processing Systems*, 34, 22419-22430.
- Zhou, T., Zhu, J., Wang, X., Ma, Z., Wen, Q., Sun, L., & Jin, R. (2022). TreeDRNet: A Robust Deep Model for Long Term Time Series Forecasting. *arXiv preprint arXiv:2206.12106*