

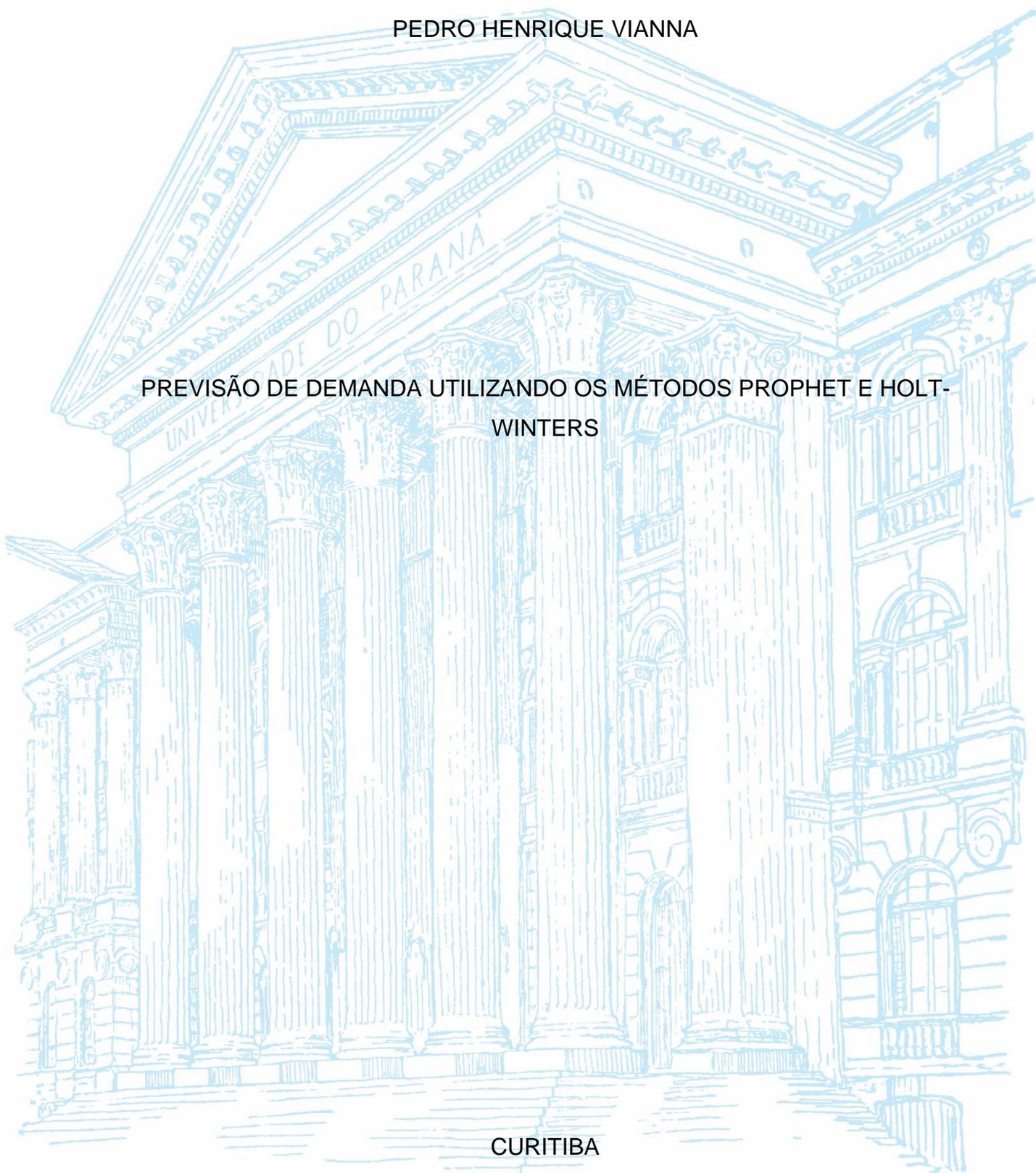
UNIVERSIDADE FEDERAL DO PARANÁ

PEDRO HENRIQUE VIANNA

PREVISÃO DE DEMANDA UTILIZANDO OS MÉTODOS PROPHET E HOLT-
WINTERS

CURITIBA

2022



PEDRO HENRIQUE VIANNA

PREVISÃO DE DEMANDA UTILIZANDO OS MÉTODOS PROPHET E HOLT-
WINTERS

TCC apresentado ao curso de Graduação de Engenharia de Produção, Setor de Tecnologia, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Bacharel em Engenharia de Produção

Orientador(a): Prof(a). Dr(a). Mariana Kleina

CURITIBA

2022

AGRADECIMENTOS

Agradeço minha família, em especial minha mãe, Rita de Cácia Vianna, que nunca mediu esforços para que eu tivesse uma educação de qualidade. Mesmo nos momentos mais adversos esse tema sempre foi prioridade.

Agradeço à Professora Dr. Mariana Kleina por toda sua dedicação a docência. A maestria com que exerce a profissão despertou em mim um interesse ainda maior por suas disciplinas e inspirou em mim o interesse pelo tema dessa pesquisa.

Por fim, agradeço aos meus amigos e a todos que de alguma forma me apoiaram nessa jornada, colegas de faculdade, colegas de trabalho e demais professores.

RESUMO

O acesso cada vez mais disseminado ao mundo digital e as mudanças nos hábitos proporcionadas pela pandemia de Covid-19 são duas das principais razões para o aumento no consumo no meio digital. A previsão de demanda dentro desse mercado é essencial para que as companhias sejam cada vez mais competitivas no mercado, entregando seus produtos de forma ágil sem que isso comprometa seus custos operacionais. Esse trabalho investiga a performance do modelo de previsão Prophet, desenvolvido recentemente pelo Facebook, comparando-o com o modelo clássico Holt-Winters na previsão de demanda de um e-commerce filipino. Ambos os modelos são utilizados para previsão de séries temporais que possuam tendência e sazonalidade. Os modelos foram implementados no software R a fim de determinar qual deles possuiria o menor erro percentual médio de previsão para um conjunto de dados de teste. O Holt-Winters obteve uma performance um pouco superior a performance obtida pelo método Prophet. O erro percentual médio absoluto do primeiro foi de 12%, enquanto a previsão do segundo foi de 18%.

Palavras-chave: Previsão de demanda, Prophet, Holt-Winters; Séries Temporais; Dados de e-commerce .

ABSTRACT

The increasingly widespread access to the digital world and the changes in habits brought about by the Covid-19 pandemic are two of the main reasons for the increase in consumption in the digital environment. Demand forecasting within this market is essential for companies to be increasingly competitive in the market, delivering their products quickly without compromising their operating costs. This work investigates the performance of the Prophet forecast model, recently developed by Facebook, comparing it with the classic Holt-Winters model in the demand forecast of a Philippine e-commerce. Both models are used to forecast time series that have trend and seasonality. The models were implemented in the R software in order to determine which one would have the smallest mean percentage forecast error for a test dataset. The Holt-Winters obtained a performance slightly superior to the performance obtained by the Prophet method. The mean absolute percentage error of the first was 12%, while the second prediction was 18%.

Key words: Demand Forecast, Prophet, Holt-Winters, Time series, E-commerce data

LISTA DE FIGURAS

FIGURA 1 - Ilustração da curva de crescimento logístico	13
Figura 2 - Etapas do Trabalho	18
Figura 3 - Representação gráfica da quantidade de vendas (<i>orders</i>) por dia do e-commerce	20
Figura 4 - Gráfico da Componente de Previsão Tendência	22
Figura 5 - Gráfico da Componente de Previsão Feriados	23
Figura 6 - Gráfico da Componente de Previsão Sazonalidade Anual	23
Figura 7 - Gráfico da Componente de Previsão Sazonalidade Semanal	24
Figura 8 - Representação Gráfica da Previsão Executada pelo método Prophet	24
Figura 9 - Regressor aplicado a função de previsão atrelado as variáveis exógenas	25
Figura 10 - Previsão gerada com a inclusão de variáveis exógenas.....	26
Figura 11 - Representação gráfica da demanda prevista pelo método Holt-Winters	27

LISTA DE TABELAS

Tabela 1 - Tamanho do passo para cada parâmetro em cada etapa na busca dos parâmetros mais apropriados.....	20
Tabela 2 - Métricas de Performance e Parâmetros mais apropriados obtidos no Prophet	22
Tabela 3 - Parâmetros de suavização mais apropriados obtidos.....	27
Tabela 4 - Métricas de Performance obtidas pelo método Holt-Winters	27

SUMÁRIO

1 INTRODUÇÃO	9
1.1 JUSTIFICATIVA	10
1.2 OBJETIVOS	11
1.2.1 Objetivo Geral	11
1.2.2 Objetivos Específicos	11
2 REVISÃO DE LITERATURA	13
2.1 SÉRIES TEMPORAIS	13
2.2 MODELO PROPHET	13
2.2 MODELO HOLT-WINTERS.....	15
2.3 MÉTRICAS DE PERFORMANCE	17
3 METODOLOGIA	18
4 RESULTADOS	20
4.1 PRÉ-TRATAMENTO DOS DADOS.....	20
4.2 PROPHET	21
4.2.1 Parâmetros e Regressores.....	21
4.2.2 Resultados	22
4.2.3 Resultados com a inclusão de variáveis exógenas	24
4.3 HOLT-WINTERS	26
4.3.1 Parâmetros.....	26
4.3.2 Resultados	27
5 CONSIDERAÇÕES FINAIS	28
REFERÊNCIAS	29

1 INTRODUÇÃO

O conceito de demanda está estreitamente ligado à quantidade de bens ou de serviços que um determinado mercado pretende consumir em um certo intervalo de tempo. A sua previsão, de acordo com Silva (2020), é uma tarefa importante pois é necessária para várias decisões operacionais.

Segundo Huber (2020), um dos principais desafios é prever a demanda em dias especiais que estão sujeitos a padrões de demanda muito diferentes dos dias normais. Esses dias podem estar atrelados a feriados comerciais e aos mais diversos fatores que de alguma forma possam impactar na necessidade de determinado produto ou serviço a ser consumido.

De acordo com o *Digital Global Overview Report* (2022), publicado pela *Hootsuite* e pela *We Are Social*, 58% da população mundial usuária da Internet com idade entre 18 e 64 anos realiza alguma compra online por semana. Segundo Guthrie et al. (2021), a disseminação do acesso à Internet e a conjuntura econômica-social provocada pela pandemia de Covid-19 corroboram para o crescimento desse número nos últimos anos.

Dado isso, pretende-se realizar a previsão de demanda de um *e-commerce* de uma empresa no sudeste asiático. Essa região em específico apresenta um percentual ainda maior de usuários da internet que realizaram alguma compra pela internet. Nesse contexto em particular, a previsão de demanda se torna importante para que a companhia consiga atender seus clientes com entregas ágeis e de maneira eficiente, isto é, sem onerar os custos logísticos.

Os últimos anos têm sido marcados por evoluções e ampliação nas possibilidades de se prever a demanda. Neste trabalho, a previsão será calculada por meio da comparação entre dois métodos: um clássico, desenvolvido na década de 70, *Holt-Winters* e um mais recente, o *Prophet*.

1.1 JUSTIFICATIVA

A competitividade no mercado e a necessidade de proporcionar uma melhor experiência aos clientes também está presente no meio digital. Martins (2012) aponta que os consumidores são ainda mais impacientes quando estão realizando suas compras no varejo virtual.

Dessa forma, é importante que os clientes tenham uma experiência positiva tanto na pré-venda como na pós-venda.

O valor de frete alto, de acordo com pesquisa realizada por Lemos (2015), é o principal motivo pelo qual um cliente desiste de realizar uma compra no e-commerce. De outro lado, o senso de urgência do consumidor tem aumentado na mesma proporção que o setor tem evoluído nos últimos anos, assim, ao se deparar com uma previsão de entrega demorada, o cliente tende a procurar outros meios de adquirir aquele produto.

Assim, para que as empresas tenham cada vez mais sucesso nesse mercado é importante que elas proporcionem uma boa experiência aos seus clientes, realizando as entregas de forma ágil e cumprindo os prazos propostos, bem como, consigam entregar essa experiência sem operar valores exorbitantes na venda de frete.

Nisso, a previsão de demanda das empresas nesse mercado pode ser uma importante aliada de modo que as corporações possam tomar decisões de forma mais assertiva proporcionando melhorias dentro do seu processo.

1.2 OBJETIVOS

A seguir serão apresentados os objetivos deste trabalho.

1.2.1 Objetivo Geral

O objetivo deste trabalho é fornecer a previsão de demanda em um intervalo de tempo delimitado por meio dos métodos Prophet e Holt-Winters de modo que a empresa de e-commerce tenha maior previsibilidade a respeito de um comportamento futuro de vendas.

É esperado que com os dados gerados a empresa possa tomar decisões mais assertivas em tudo aquilo que tange os seus processos e a cadeia de suprimentos.

1.2.2 Objetivos Específicos

Os objetivos específicos do trabalho são:

- Realizar a previsão de demanda no horizonte de um ano utilizando os algoritmos de Holt-Winters e Prophet com o suporte da linguagem de programação R e os pacotes e bibliotecas pré-existentes.
- Comparar o resultado da previsão proporcionado pelo método Prophet com somente os dados da série histórica com os resultados da previsão com a inclusão de variáveis exógenas.
- Determinar o método com melhor resultado, aplicando o cálculo do erro percentual absoluto médio (MAPE).

2 REVISÃO DE LITERATURA

2.1 SÉRIES TEMPORAIS

De acordo com Ehlers (2003), uma série temporal consiste na coleção de observações feitas sequencialmente ao longo do tempo. A característica mais relevante deste tipo de dado é que as observações vizinhas são dependentes e a modelagem desta dependência é um dos principais objetivos de seu estudo.

Ao se analisar uma ou mais séries temporais a representação gráfica dos dados sequencialmente ao longo do tempo é um dos elementos principais na análise e pode revelar padrões de comportamentos importantes, tais como tendências de crescimento (ou decrescimento), padrões cíclicos, alterações estruturais, *outliers*.

Desse modo, o gráfico temporal deve ser sempre o primeiro passo e antecede qualquer análise.

2.2 MODELO PROPHET

O Prophet é um modelo de previsão de séries temporais *open source* desenvolvido pelo Facebook e que foi disponibilizado ao público em 2017 para uso nas linguagens de programação R e Python.

Segundo Taylor e Letham (2018), o modelo foi criado com a intenção de facilitar a atividade de se realizar previsões. O Prophet trabalha com séries temporais decompostas em tendências, sazonalidade e feriados. A soma dessas previsões decompostas resulta na previsão completa realizada pelo modelo que é representada em 1.

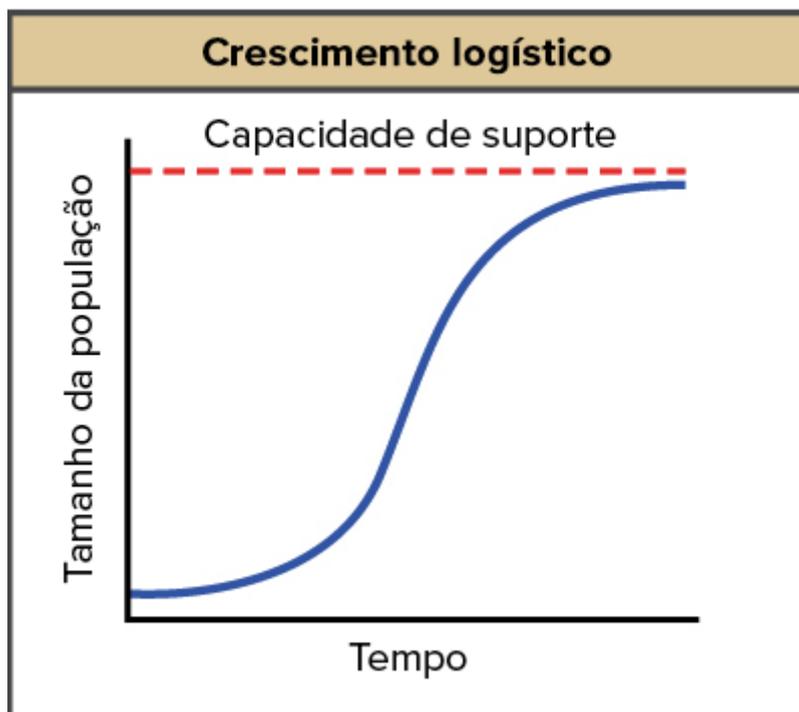
$$y(t) = g(t) + s(t) + h(t) + e \quad (1)$$

Na equação, $g(t)$ é a função de tendência que modela mudanças no valor da série temporal, $s(t)$ representa mudanças relacionadas à sazonalidade e $h(t)$ representa os efeitos dos feriados que ocorrem em um ou mais dias da série. O termo de erro representa quaisquer mudanças idiossincráticas que não sejam acomodadas pelo modelo.

A função tendência pode descrever dois tipos de comportamento: o crescimento logístico e o crescimento linear.

O crescimento logístico ou sigmoidal envolve o crescimento exponencial seguido por uma redução constante no crescimento até que o volume se estabilize, assumindo uma curva em forma de “S”, como mostra a FIGURA 1.

FIGURA 1 - Ilustração da curva de crescimento logístico



Fonte: Herold e Souza (2015)

Por sua vez, o crescimento linear é utilizado em problemas de previsão que não apresentam crescimento saturante e que pode ser bem descrito por uma taxa de crescimento constante.

Ainda de acordo com Taylor e Letham (2018), as séries temporais que descrevem os resultados dos mais variados tipos de negócios geralmente têm sazonalidade de vários períodos como um resultado dos comportamentos humanos que eles representam. Por exemplo, a distinção de comportamento entre dias úteis e finais de semana pode produzir efeitos em uma série temporal que se repete a cada semana, enquanto os períodos de férias e os intervalos escolares podem produzir efeitos que se repetem a cada ano. Para ajustar e prever esses efeitos, deve-se especificar modelos de sazonalidade que são funções periódicas de t .

Feriados e eventos fornecem mudanças grandes e um tanto quanto previsíveis para muitas séries temporais de negócios e muitas vezes não seguem um padrão, assim sendo seus efeitos não são bem modelados por um ciclo suave. O impacto de um determinado feriado na série temporal é muitas vezes semelhante ano após ano, por isso é importante incorporá-lo à previsão.

Uma das principais qualidades do método Prophet é sua flexibilidade e a possibilidade de incluir novas variáveis que possam ser utilizadas pelo modelo para que a previsão seja mais fidedigna.

Dado isso, a documentação no software R do modelo Prophet aponta alguns parâmetros que podem ser ajustados na definição do modelo de modo que ele possa se adaptar melhor a cada contexto.

O *changepoint_prior_scale* é provavelmente o parâmetro mais impactante. Ele determina o quanto a tendência é flexível, isto é, o quanto ela muda nos pontos de mudança de tendência, que podem ser capturados automaticamente pelo modelo ou inseridos manualmente pelos usuários. Se ele for muito pequeno, a tendência será sub ajustada e a variação que deveria ter sido modelada com mudanças de tendência acabará sendo tratada com o termo de ruído. Por outro lado, se ele for muito grande, a tendência se ajustará demais e, no caso mais extremo, pode-se acabar com a tendência capturando a sazonalidade anual.

O *seasonality_prior_scale* controla a flexibilidade da sazonalidade. Da mesma forma, um valor grande permite que a sazonalidade se ajuste a grandes flutuações, um valor pequeno diminui a magnitude dela. O padrão é 10, que basicamente não aplica nenhuma regularização. Um intervalo razoável para ajuste provavelmente seria entre 0,01 e 10. Quando definido para 0,01 a magnitude da sazonalidade é forçada a ser muito pequena.

O *holidays_prior_scale* controla a flexibilidade para ajustar os efeitos de feriados e datas especiais. Semelhante ao *seasonality_prior_scale*, o padrão é 10, que basicamente não aplica nenhuma regularização.

O *seasonality_mode* apresenta duas opções: aditivo ou multiplicativo. O padrão é 'aditivo', mas muitas séries temporais de negócios terão sazonalidade multiplicativa.

2.2 MODELO HOLT-WINTERS

O algoritmo Holt-Winters para previsão foi desenvolvido por Charles Holt e Peter Winters. De acordo com Veríssimo et al. (2012), o modelo costuma ser empregado quando a série apresenta componentes de nível: sazonalidade e tendência.

Ele é uma expansão do modelo de suavização exponencial simples o qual se despreza a tendência e o efeito sazonal, o mais comum, no entanto é que uma série possua tendência e efeito sazonal, por isso o Holt-Winters se torna um modelo de previsão mais completo.

De acordo com Araújo et al. (2018), os modelos de Holt-Winters são classificados em dois grupos: aditivo, que considera a amplitude da variação sazonal constante ao longo do tempo, e o multiplicativo que considera a amplitude da variação sazonal aumentando ou diminuindo como função do tempo.

Utilizam-se as equações 2, 3, 4 e 5 para o modelo de Holt-Winters multiplicativo, alvo desse trabalho (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998):

$$L_t = \alpha \frac{y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (2)$$

$$T_t = \beta(L_t + L_{t-1}) + (1 - \beta)T_{t-1} \quad (3)$$

$$S_t = \gamma \frac{y_t}{L_t} + (1 - \gamma)S_{t-s} \quad (4)$$

Onde, os subíndices t e $t-1$, indicam o tempo atual e o tempo imediatamente passado, y é o valor observado (real) na série temporal; L é a componente de nível; T é a componente de tendência; s é o comprimento da

sazonalidade; S é a componente de sazonalidade. α , β e γ e são valores entre 0 e 1, e são chamadas de constantes de suavização do modelo.

É importante definir os valores iniciais para as componentes. É comum selecionar $L_1 = x_1$, $T_1 = 0$ e efeitos sazonais $S_1 \dots S_s$ como a média das observações de uma estação.

A previsão utilizando o modelo de Holt-Winters para um horizonte de tempo k pode ser calculado por 5.

$$\hat{y}_{t+k} = (L_t + kT_t)S_{t-s+k} \quad (5)$$

2.3 MÉTRICAS DE PERFORMANCE

O erro médio absoluto (MAE) mede o quanto o valor estimado ou previsto difere dos valores reais (XU et al., 2017). Deve ser olhado com atenção já que é uma medida dependente de escala, então não deve ser usado quando as séries que serão comparadas estão em escalas diferentes (HYNDNAM E KOEHLER, 2006).

O erro percentual absoluto médio (MAPE) é a medida estatística de precisão da previsão de um método de previsão (DU et al., 2018).

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t| \quad (6)$$

Em 6, e_t é obtido pela diferença entre o valor real de série e o valor previsto e n é a quantidade de valores da série que foi estudada.

$$MAPE = \frac{1}{n} \sum_{t=1}^n |re_t| \quad (7)$$

Em 7, re_t é obtido pela razão entre e_t e y_t

3 METODOLOGIA

Este estudo se trata de uma pesquisa aplicada, uma vez que produz conhecimentos com o objetivo de utilizá-los para fins práticos. No caso desse trabalho, a determinação de qual método de previsão pode ser mais assertivo na previsão de demanda de uma empresa. Tem natureza exploratória, porque foram captados dados, referentes ao desempenho histórico da empresa e que foram utilizados para obter os resultados da pesquisa. É caracterizada como quantitativa já que os dados coletados eram variáveis numéricas que foram utilizadas com o objetivo de gerar informações claras e objetivas a respeito do futuro dessas vendas. Quanto aos procedimentos de pesquisa foram realizados uma pesquisa bibliográfica e um estudo de caso.

O trabalho pode ser descrito pela execução de quatro etapas ilustradas na FIGURA 2.

Figura 2 - Etapas do Trabalho



Fonte: O autor (2022)

A primeira etapa da pesquisa consistiu na busca por referências e entendimento dos resultados de trabalhos semelhantes a fim de embasar e fornecer insumos para a realização do estudo. Essa etapa forneceu informações para a seleção dos métodos de previsão que seriam utilizados na execução do trabalho.

A segunda etapa da pesquisa consistiu na coleta e pré-tratamento dos dados que seriam explorados no trabalho. Os dados de um e-commerce do sudeste asiático foram obtidos por meio da plataforma *Kaggle* e foram importados na ferramenta R Studio. O pré-tratamento dos dados consistiu na verificação de dados faltantes no período em questão, a conversão para o formato de Data Frame e a separação em conjunto de verificação e conjunto de validação.

A terceira etapa consistiu na execução do modelo de previsão para os dois métodos de previsão selecionados. Para o método Prophet, foi utilizado o pacote de mesmo nome na aplicação. A aplicação do modelo está sujeita a determinação de fatores que podem tornar a previsão mais ou menos aceitável. Foi criado um laço com o objetivo de aplicar o modelo para a maior combinação de parâmetros possível. Para o método Holt-Winters, foi utilizado o pacote "Forecast" e o procedimento de utilização foi o mesmo.

A quarta etapa consistiu na avaliação dos modelos tendo como referência as métricas de performance selecionadas.

4 RESULTADOS

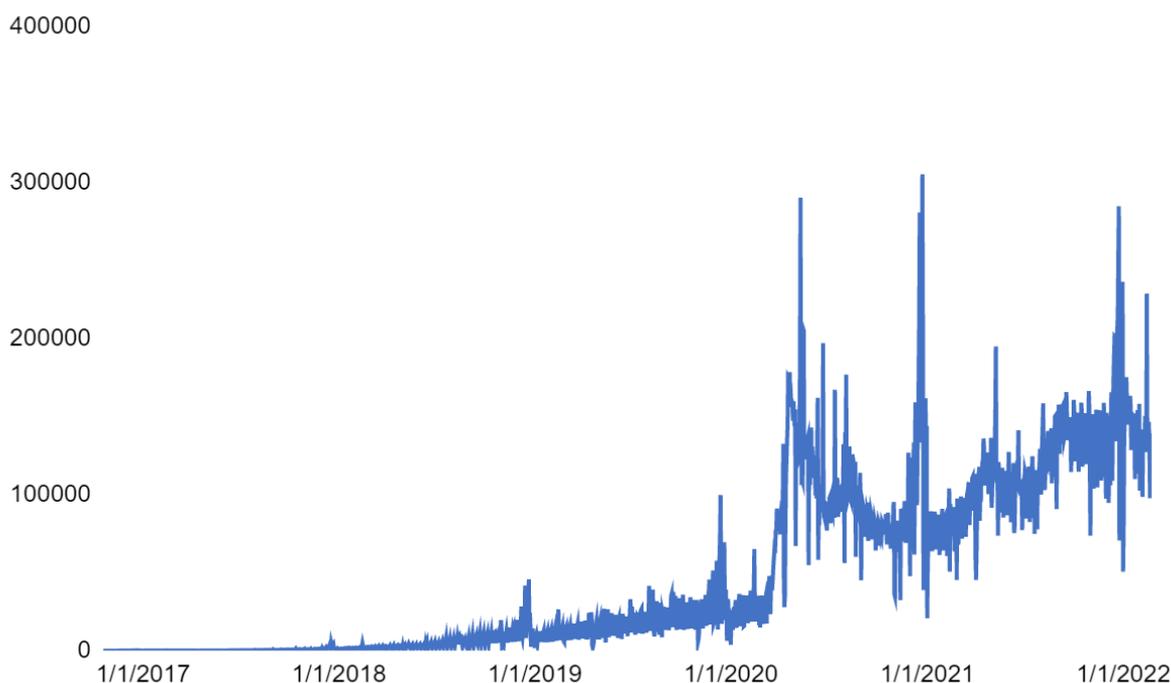
4.1 PRÉ-TRATAMENTO DOS DADOS

A base selecionada apresentava duas colunas contendo o registro de vendas por dia no período entre 19 de outubro de 2016 e 23 de fevereiro de 2022, contemplando assim um total de 1955 registros. O comportamento da demanda observado nessa empresa está ilustrado na FIGURA 3.

O conjunto foi importado ao ambiente de trabalho no R e separado em grupo de validação, primeiros 80% dos registros, e grupo de teste, os demais 20%.

Observa-se que os primeiros anos de medição registraram demandas com baixa ordem de grandeza, mostrando provavelmente que se tratava do início desse negócio e retratando um período cuja influência dos *e-commerces* era menor.

Figura 3 - Representação gráfica da quantidade de vendas (*orders*) por dia do e-commerce



Fonte: O autor (2022)

4.2 PROPHET

4.2.1 Parâmetros e Regressores

O Prophet se destaca em relação a outros modelos de previsão por sua flexibilidade e sua adaptabilidade aos mais diversos contextos. Dentro da aplicação R é possível incluir no modelo todas as datas comemorativas de determinado país para que o modelo leve em conta durante sua previsão e possa determinar se esses feriados geram ou não algum tipo de impacto.

Por fim é imprescindível que se determine para o conjunto que será analisado qual o valor mais apropriado a ser utilizado dentre os parâmetros apresentados anteriormente. Para isso, foram feitos laços de repetição encadeados, ou seja, dentro de um primeiro laço que variava um dos parâmetros, foi incluído outro laço para variação de outro parâmetro, a fim de se cobrir a maior quantidade de combinações possíveis e calcular para cada uma delas o MAPE.

O intervalo utilizado para cada parâmetro cobriu justamente o intervalo possível dos valores de cada um deles no Prophet e foi realizado em etapas de modo a se economizar tempo computacional, iniciando com passos mais largos e a partir dos resultados entendendo em qual parâmetro faria sentido ter um passo cobrindo uma ordem de grandeza menor, o que pode ser observado na TABELA 1.

Tabela 1 - Tamanho do passo para cada parâmetro em cada etapa na busca dos parâmetros mais apropriados

Etapa	<i>Changepoint prior scale (CPS)</i>	<i>Seasonal prior scale (SPS)</i>	<i>Holiday prior scale (HPS)</i>
1	0,1	1	1
2	0,01	0,01	0,01
3	-	0,01	0,01

Fonte: O Autor (2022)

Os dados apontam que para a base em questão, o modelo Prophet se adequa melhor com os valores mínimos de cada parâmetro. Os parâmetros mais

apropriados, além do erro absoluto médio e do erro percentual absoluto médio podem ser observados na TABELA 2.

Tabela 2 - Métricas de Performance e Parâmetros mais apropriados obtidos no Prophet

Erro Absoluto Médio (MAE)	Erro Percentual Absoluto Médio (MAPE)	<i>Changepoint prior</i> <i>scale (CPS)</i>	<i>Seasonal prior</i> <i>scale (SPS)</i>	<i>Holiday prior</i> <i>scale (HPS)</i>
18805.08	18,23%	0,001	0,01	0,01

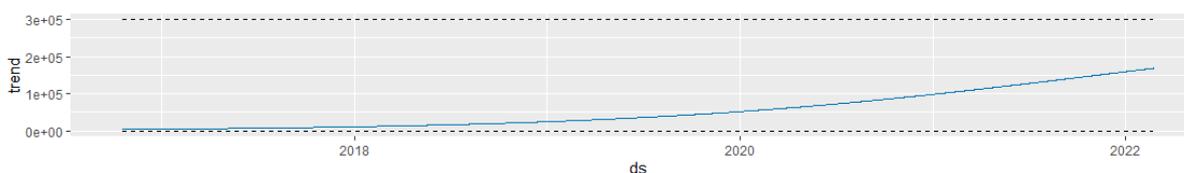
Fonte: O Autor (2022)

Os parâmetros aplicados resultaram num erro percentual médio absoluto de 18,23%, enquanto o erro absoluto médio foi de 18805.

4.2.2 Resultados

Nota-se o impacto na componente de previsão tendência na FIGURA 4, em que nos primeiros anos de previsão, o componente tendência tem valores com baixa ordem de grandeza, contudo, com o passar do tempo, com a proximidade dos últimos anos de previsão nota-se o impacto e como esse componente se torna cada vez mais expressivo.

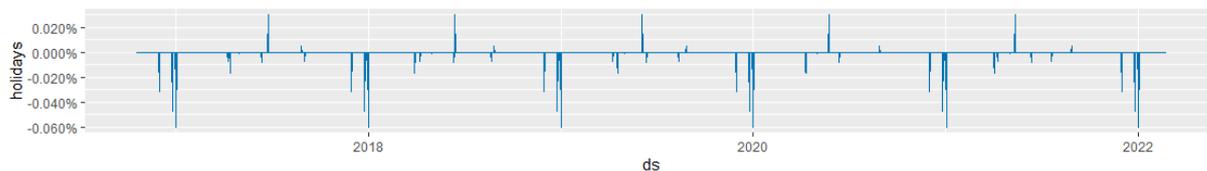
Figura 4 - Gráfico da Componente de Previsão Tendência



Fonte: O autor (2022)

A maioria dos feriados apresenta um impacto negativo na demanda, mas como o parâmetro que se relaciona com esse componente que resulta nos melhores resultados de previsão é pequeno e, portanto, torna-o menos sensível a previsão, faz com que a oscilação que essa componente provoca nos números está em um pequeno intervalo entre -0,5% e 0,5%.

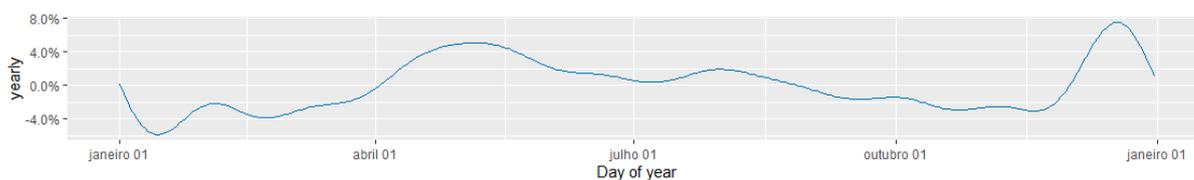
Figura 5 - Gráfico da Componente de Previsão Feriados



Fonte: O Autor (2022)

A sazonalidade anual por sua vez mostra que a demanda é bem expressiva no mês de dezembro, gerando um acréscimo na demanda de até 8% nesse período. A Filipina é um país asiático majoritariamente católico. Nesse mês é comemorado o Natal, data que comemora o nascimento de Jesus Cristo e aquece o comércio já que as pessoas costumam se presentear e se reunir para celebrar esse momento. Por outro lado, o mês de janeiro costuma trazer um impacto negativo de até 5% nas vendas.

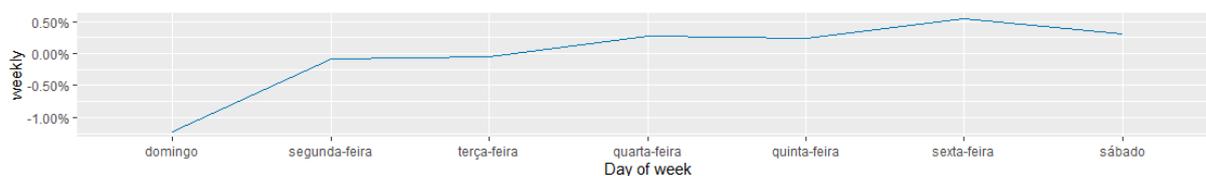
Figura 6 - Gráfico da Componente de Previsão Sazonalidade Anual



Fonte: O Autor (2022)

Quando a componente semanal é visualizada (FIGURA 7) nota-se que domingo provoca um impacto negativo na demanda, diferente dos demais dias da semana que acabam tendo um comportamento mais semelhante uns com os outros.

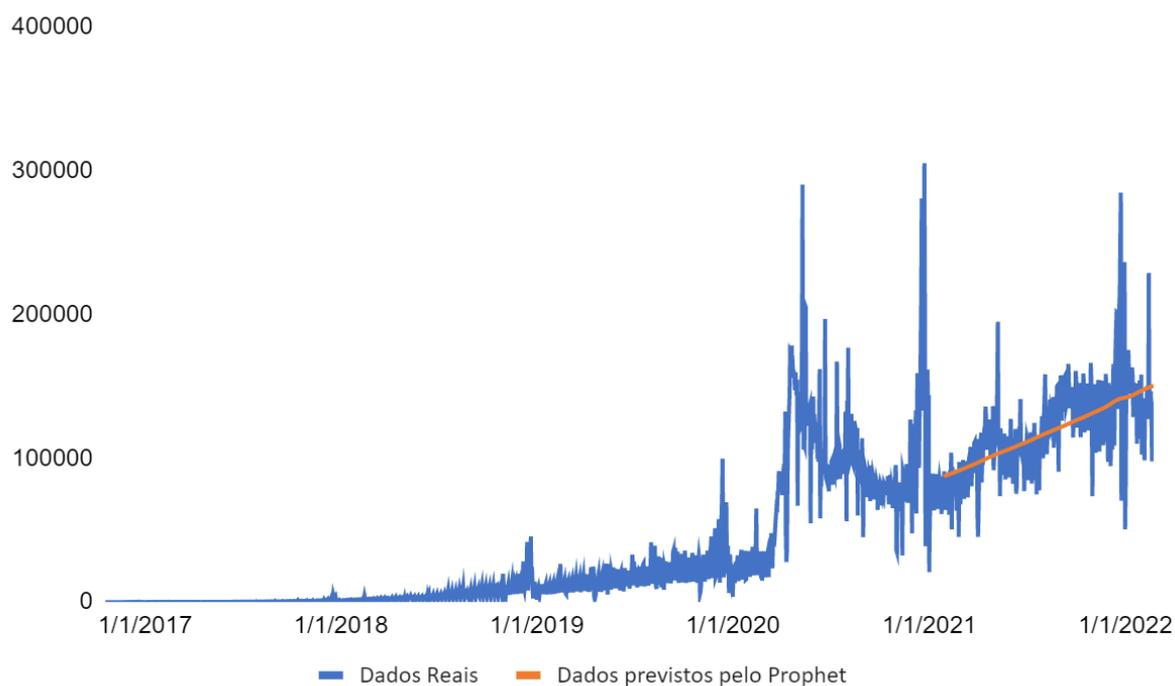
Figura 7 - Gráfico da Componente de Previsão Sazonalidade Semanal



Fonte: O Autor (2022)

A plotagem dos dados previstos pode ser visualizada na FIGURA 8.

Figura 8 - Representação Gráfica da Previsão Executada pelo método Prophet



Fonte: O autor (2022)

4.2.3 Resultados com a inclusão de variáveis exógenas

Neste trabalho, fez-se uso de duas variáveis exógenas: a média de temperatura diária e a quantidade de precipitação.

Hoje, o comportamento do consumidor pode ser bastante afetado pelas condições citadas anteriormente. Por exemplo, se o negócio em questão se trata de

um *delivery* de *fast-food* espera-se que a demanda seja maior em dias com baixa temperatura já que essas condições climáticas podem reduzir a vontade do cliente sair de casa.

Por outro lado, no varejo têxtil que confecciona casacos e roupas para o frio, por exemplo, a demanda pode ser impactada positivamente em dias com a temperatura mais baixa.

Como não se sabe com exatidão o nicho de mercado da empresa que gerou a base de dados utilizada, entende-se que a inclusão dessas variáveis na pesquisa e o comportamento da previsão gerado podem até nos gerar mais insumos sobre ela.

As variáveis exógenas foram incluídas no modelo que tinha como base os parâmetros mais apropriados obtidos anteriormente. A componente delas geradas pela previsão pode ser visualizada na FIGURA 9 e a previsão dos dados pode ser visualizada na FIGURA 10.

Figura 9 - Regressor aplicado a função de previsão atrelado as variáveis exógenas

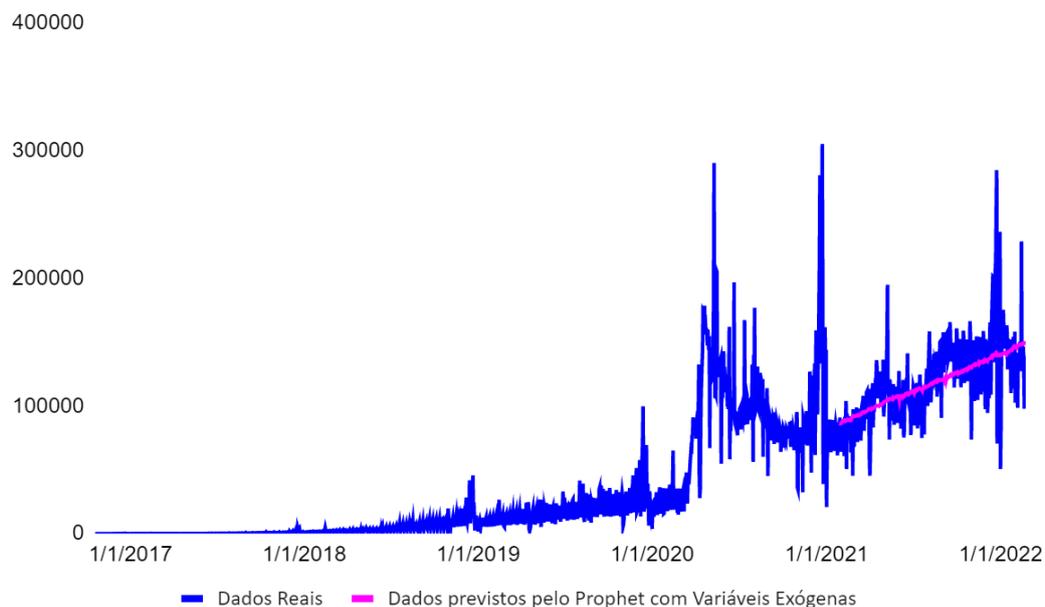


Fonte: O Autor (2022)

Nota-se que a inclusão das variáveis exógenas no modelo pode provocar uma variação entre -3% e 2% na previsão do modelo Prophet ao longo dos anos de estudo.

A inclusão dessas variáveis não provocou variação significativa no MAPE mensurado anteriormente, isso traz indícios que em alguns momentos elas podem ter contribuído na melhora do modelo e em outros terem prejudicado a previsão inicial.

Figura 10 - Previsão gerada com a inclusão de variáveis exógenas



Fonte: O Autor (2022)

4.3 HOLT-WINTERS

4.3.1 Parâmetros

Da mesma forma que no método Prophet, os parâmetros mais apropriados para a função Holt-Winters puderam ser obtidos a partir da execução de um laço de repetição, em que os parâmetros de suavização variaram no intervalo entre 0,0001 e 1 ao passo de 0,001, a previsão era obtida e o erro percentual médio absoluto era calculado a fim de se obter a melhor combinação desses parâmetros. Os resultados obtidos podem ser visualizados na TABELA 3.

Foi testada também a normalização dos dados da série histórica, mas sem resultados consideráveis para a pesquisa.

Tabela 3 - Parâmetros de suavização mais apropriados obtidos

α	β	γ
0,0001	0,0401	0,0121

Fonte: O autor (2022)

4.3.2 Resultados

A partir da seleção de parâmetros estabelecida como a mais apropriada para a previsão calculou-se as métricas de performance assim como no modelo Prophet (TABELA 4).

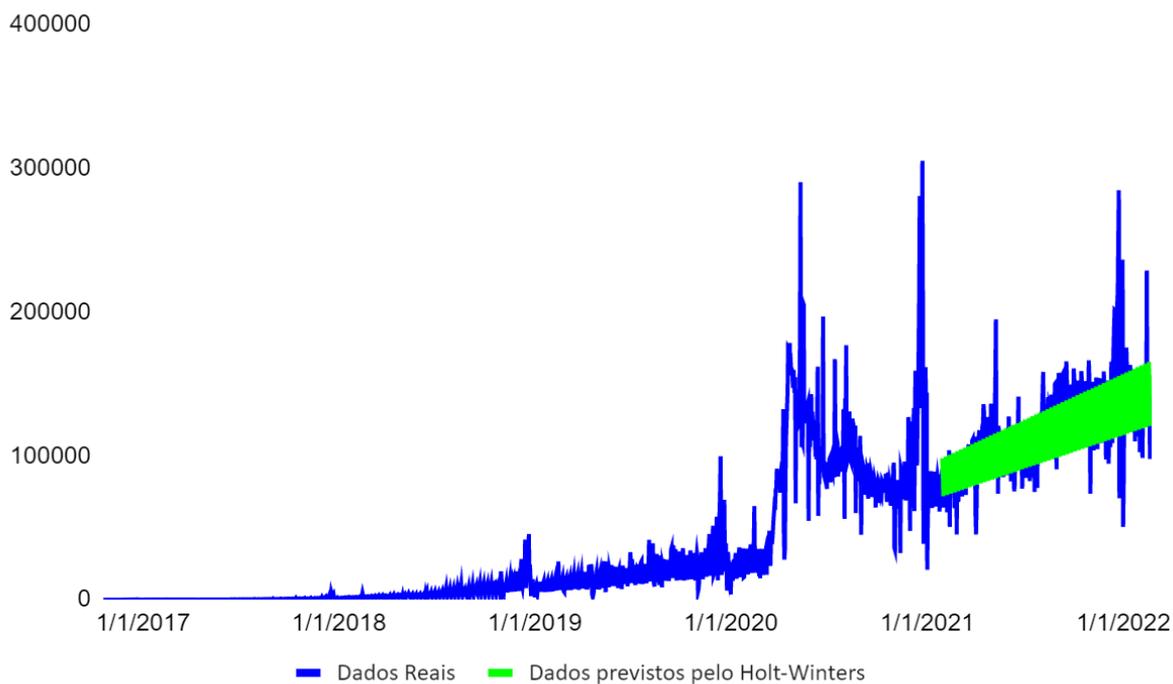
Tabela 4 - Métricas de Performance obtidas pelo método Holt-Winters

Erro Absoluto Médio (MAE)	Erro Percentual Absoluto Médio (MAPE)
14066	12%

Fonte: O autor (2022)

O erro percentual absoluto médio (MAPE) foi de 12% que representa para o conjunto de dados em questão um erro absoluto médio de 14066. A representação dos dados previstos pode ser visualizada na FIGURA 11.

Figura 11 - Representação gráfica da demanda prevista pelo método Holt-Winters



Fonte: O autor (2022)

Nota-se que os dados previstos formam uma banda em que os valores mais baixos dela representam a previsão semanal de vendas realizadas aos domingos, momento da semana responsável pela pior performance do *e-commerce*.

5 CONSIDERAÇÕES FINAIS

O objetivo principal deste trabalho consistia na comparação entre os resultados obtidos na previsão de vendas de um *e-commerce* pelos métodos Prophet e Holt-Winters.

O método clássico desenvolvido no início da década de 70 teve uma performance ligeiramente superior ao método desenvolvido pelo Facebook. A melhor previsão realizada pelo Holt-Winters resultou em um erro percentual absoluto médio de 12%, enquanto o método Prophet apresentou um desempenho aquém com esse mesmo erro na casa de 18%.

Entende-se que a aplicação dessa previsão no contexto do negócio, utilizando ela como suporte para tomada de decisão dentro dos mais diversos setores da companhia, ainda que com o método Holt-Winters, poderia ser prejudicial em vista da magnitude dos erros obtidos.

Nota-se uma certa dificuldade nos modelos em reconhecer de forma mais assertiva os padrões da série temporal. Uma alternativa sugerida para futuros trabalhos consiste no tratamento prévio e estudo mais aprofundado das propriedades das séries temporais, buscando fornecer insumos que permitam que os modelos preditivos tenham uma performance melhor e sejam ainda mais assertivos que o apresentado nesse trabalho.

Outra possibilidade a ser explorada consiste na realização da previsão em intervalos menores do grupo de teste, de modo que após cada os elementos previstos sejam utilizados como dados de validação nas previsões dos intervalos subsequentes.

A experiência com a inclusão de variáveis exógenas trouxe resultados interessantes na previsão realizada pelo modelo Prophet ainda que não tenham provocado grandes variações nos resultados, o próprio MAPE permaneceu praticamente constante quando comparadas as duas previsões. Esse ponto pode ser melhor explorado em futuros trabalhos buscando selecionar novas variáveis que gerem maior impacto positivo no modelo de previsão. Existem algoritmos de *Machine Learning* como árvores de decisão que podem contribuir nessa seleção.

REFERÊNCIAS

- ARAÚJO, G. C.; SILVA, J. P. Z.; SOUZA, L. R.; LOUREIRO, M. B.; FERONI, R. C. **Previsão de Demanda e Análise Simplificada da Gestão de Estoque Aplicada à uma Empresa do Setor Alimentício**. Brazilian Journal of Production Engineering, v. 4, no. 2, p. 48-64, 2018
- DU, P.; WANG, J.; GUO, Z.; YANG, W.; **Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting**. Energy Conversion and Management, v. 150, p. 90-107, 2017
- GUTHRIE, C.; FOSSO-WAMBA, S.; ARNAUD J. B. **Online consumer resilience during a pandemic: An exploratory study of e-commerce behavior before, during and after a COVID-19 lockdown**. Journal of Retailing and Consumer Services, v.61, 102570. 2021.
- EHLERS, R. S. **Análise de Séries Temporais**. Laboratório de Estatística e Geoinformação da Universidade Federal do Paraná, Curitiba, 2003.
- HEROLD, L. S.; SOUZA, M.L.H. **Previsão de Vendas com Machine Learning: implementação do algoritmo Prophet em Linguagem R**, 2015.
- HYNDMAN, R. J.; KOEHLER, A. B. **Another look at measures of forecast accuracy**. International Journal of Forecasting, v. 22, p. 679-688. 2006
- HUBER, J.; STUCKENSCHMIDT, H. **Daily retail demand forecasting using machine learning with emphasis on calendric special days**. International Journal of Forecasting, v.36, c. 4, p. 1420-1328. 2020.
- MAKRIDAKIS, S.; WHEELWRIGHT, S.;HYNDMAN, R. J..**Forecasting methods and applications**. 3ed. New York: John Wiley & Sons, 1998
- SILVA, F. H. **Previsão de demanda e planejamento comercial**. São Paulo: Senac São Paulo, 2020.
- TAYLOR, J. S.; LETHAM, B. **Forecasting at Scale**. The American Statistician v.72, no. 1, p. 37-45. 2018.
- VERÍSSIMO, A. J.; ALVES, C.C.; HENNING, E.; AMARAL, C. E.; CRUZ, A. C. **MÉTODOS ESTATÍSTICOS DE SUAVIZAÇÃO EXPONENCIAL HOLTWINTERS PARA PREVISÃO DE DEMANDA EM UMA EMPRESA DO SETOR METAL MECÂNICO**. Revista Gestão Industrial v. 08, n. 04: p. 154-171, 2012
- XU, Y.; YANG, W.; WANG, J.; YANG. **Air quality early-warming system for cities in China**. Atmospheric Environment, v. 148, p. 239-257, 2017