UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA

ARTHUR SHINZO FERNANDES SHIMIZU

Estudo comparativo de métodos estatísticos e de Machine-Learning para previsão de demanda

São Paulo
2022

ARTHUR SHINZO FERNANDES SHIMIZU

Estudo comparativo de métodos estatísticos e de Machine-Learning para previsão de demanda

Trabalho de formatura apresentado à Escola Politécnica da Universidade de São Paulo para obtenção do Diploma de Engenheiro de Produção.

Orientador: Prof. Dr. Marco Aurélio de Mesquita

São Paulo

2022

RESUMO

O presente estudo propõe comparar métodos estatísticos tradicionais e métodos mais recentes de Machine Learning para a aplicação da previsão de demanda. São utilizadas duas bases de dados, uma univariada e previamente tratada, outra multivariada e sem tratamento prévio. A primeira é utilizada para referência ilustrativa da aplicação em séries temporais com clara sazonalidade e tendência, em uma situação de fácil modelagem. Já a segunda é utilizada para efetivamente comparar os métodos, simulando uma situação real onde as séries temporais não necessariamente seguem padrões de fácil previsibilidade. Pôde-se observar que métodos de Machine-Learning obtém melhor performance em um maior número de vezes do que os tradicionais. Porém, os métodos tradicionais ainda conseguem uma melhor performance em 29% dos casos.

ABSTRACT

The present study proposes to compare traditional statistical methods and more recent Machine Learning methods for demand forecasting. Two databases are used, one univariate and previously treated, the other multivariate and without previous treatment. The first is used as an illustrative reference of the application in time series with clear seasonality and trend, in a situation of simple modeling. The second is used to effectively compare the methods, simulating a real situation where the time series do not necessarily follow easily predictable patterns. It could be observed that Machine-Learning methods perform better in a larger number of times than traditional methods. However, traditional methods still perform better in 29% of the cases.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

LISTA DE SIGLAS

ARIMA	Autoregressive Integrated Moving Average
LSTM	Long-Short Therm Memory
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
RNN	Recurrent Neural Networks
SARIMA	Seasonal Autoregressive Integrated Moving Average
SKU	Stock Keeping Unit

SUMÁRIO

1 INTRODUÇÃO

1.1 Contexto

Estudos sobre aplicações de Machine Learning, ou aprendizado de máquina, têm ganhado cada vez mais espaço no meio acadêmico nos últimos anos. Alinhando estatística e grande poder computacional, estes estudos têm sido aplicados em diversas áreas, desde reconhecimento de imagem até interpretação de textos em linguagem natural. Um desses campos onde o aprendizado de máquina tem se mostrado relevante é na previsão de demanda.

A previsão de demanda é um campo de extrema importância do planejamento da produção. É essencial não só para o planejamento estratégico de uma empresa, mas também para as operações do dia-a-dia, como controle de estoque e custos da produção.

Tradicionalmente, métodos estatísticos como a Suavização Exponencial e o Auto-Regressive Integrated Moving Average (ARIMA) são conhecidos e utilizados tanto pelo bom desempenho quanto pela facilidade de implantação. Porém, com o crescimento da popularidade de temas como Big Data e Machine Learning nos últimos anos, muitos estudos têm sugerido abordagens alternativas para resolver problemas de previsão em séries temporais, mostrando resultados promissores e que são revisados no capítulo 2 deste estudo. Atualmente, muitas empresas já adotam modelos de Machine Learning para realizar suas previsões de demanda, como a Amazon que substituiu modelos tradicionais pelo Random Forest já em 2009.

Efetivamente, pode-se argumentar que, para grande parte dos casos, não há porquê recorrer ao aprendizado de máquina. Modelagens tradicionais já proveem métodos para contemplar tendência, sazonalidade e erro de séries temporais, componentes já bem conhecidos. Ainda assim , pode-se argumentar que ao recorrer ao aprendizado de máquina, podemos obter resultados melhores porque os algoritmos não se limitam a estes três componentes da demanda, podendo contemplar também mais componentes da série temporal, como por exemplo pontos de inflexão (como feriados e promoções) e incorporação de outliers de forma não enviesada.

Além disso, para casos de séries multivariadas, a aplicação de Machine Learning pode ser interessante por mapear não apenas auto-correlações das variáveis como também correlações entre as múltiplas variáveis que compõem a base de dados, com potencial para redução dos erros de previsão.

1.2 Objetivo do trabalho

O presente estudo compara métodos tradicionais de previsão de demanda, sendo eles Suavização Exponencial e o ARIMA, com métodos mais recentes de Machine Learning, sendo os escolhidos para o estudo o Prophet, o LSTM e o LightGBM.

As comparações são feitas utilizando uma base de dados multivariada com múltiplas séries temporais, medindo o desempenho dos diferentes métodos propostos e discutindo as vantagens e desvantagens de cada um.

1.3 Relevância

Apesar dos avanços recentes em Big Data e Machine Learning, ainda existem incertezas sobre o potencial de aplicação dos novos métodos e verificar se são vantajosos no contexto específico da previsão de demanda. Este estudo compara alguns métodos novos com métodos tradicionais, buscando entender o potencial de aplicação destes novos métodos.

1.4 Oportunidades de mudanças no transporte de automóveis novos

Este primeiro capítulo apresentou uma contextualização e os objetivos do presente trabalho.

O capítulo 2 revisa estudo de aplicações de métodos tradicionais e aplicações modernas de algoritmos de machine learning, assim como levanta referências bibliográficas que discutem aplicações da previsão de demanda.

O capítulo 3 descreve o método de trabalho, incluindo o planejamento dos experimentos.

O capítulo 4 apresenta uma análise exploratória das duas bases de dados utilizadas no trabalho, a primeira para fins ilustrativos e a segunda para comparação efetiva dos métodos. Além disso, descreve duas séries temporais selecionadas para demonstração do procedimento de previsão realizado.

O capítulo 5 testa diferentes algoritmos na primeira série temporal selecionada a fim de ilustrar o procedimento de previsão, finalizando com a discussão sobre os resultados obtidos.

O capítulo 6, analogamente ao 5, consiste na aplicação dos diferentes algoritmos, mas desta vez na segunda série temporal selecionada para demonstrar o procedimento de previsão. Por fim, apresenta a discussão sobre os resultados obtidos.

O capítulo 7 apresenta os resultados da aplicação do procedimento do capítulo 6 em uma amostra aleatória de 100 séries temporais selecionadas da segunda base de dados, a fim de comparar efetivamente os algoritmos estudados.

Por fim, o capítulo 8 encerra o trabalho com as conclusões e considerações finais.

2 REVISÃO BIBLIOGRÁFICA

2.1 Previsão de demanda

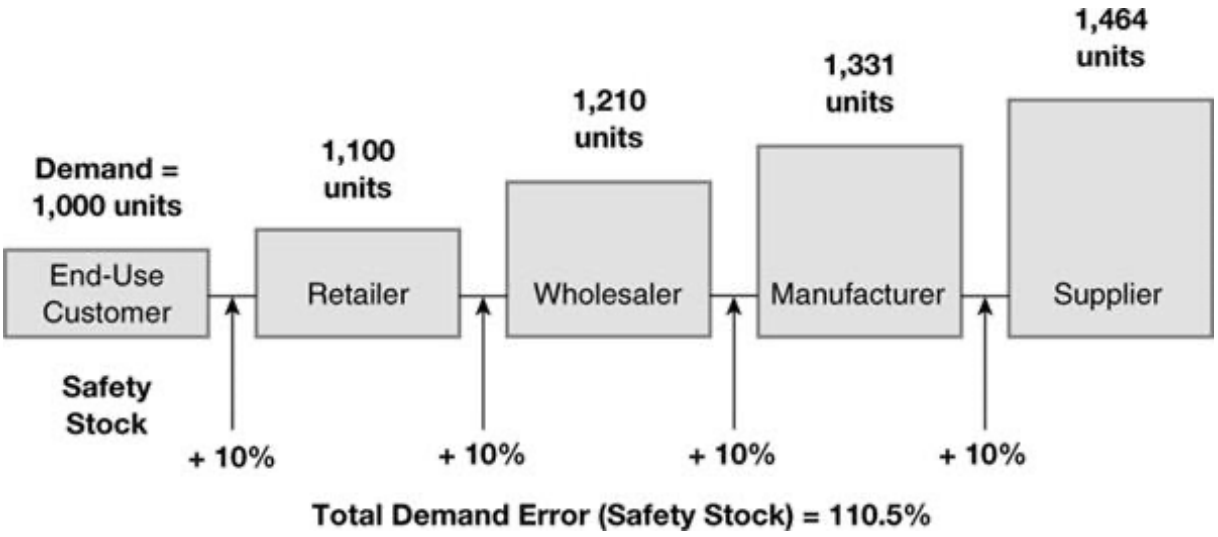
2.1.1 Relevância

Quando se trata de produção e controle de estoque, a previsão de vendas é um fator importante que tem implicações nos custos de estoque, nível de serviço, programação e produtividade dos funcionários, entre muitos outros aspectos importantes do desempenho operacional (GARDNER, 2006). De acordo com Chopra (2003), o primeiro passo que uma gestão da cadeia de suprimentos deve dar é implantar um sistema de previsão da demanda. As previsões de demanda são a base para a tomada de decisões estratégicas, táticas e operacionais. Além disso, a previsão de demanda não é só responsabilidade do especialista em logística, mas também requer a assistência de marketing, planejamento econômico ou um grupo que tenha sido especificamente estabelecido para este fim (BALLOU, 2001).

A forma pela qual uma empresa deve responder à demanda de seus produtos difere de acordo com a posição da empresa na cadeia de produção. O único ponto em que a demanda pode ser considerada completamente independente é quando ela é demandada pelo usuário final. O usuário final é o responsável por determinar a demanda que flui através da cadeia de produção (MENTZER; MOON, 2005).

O comportamento do erro de previsão ao longo da cadeia de produção pode ser observado a partir da Figura1, como mostrado na figura. Se assumirmos que há um estoque de segurança de 10% em cada etapa, então sabemos que se o cliente final fizer um pedido de 1000 unidades, haverá um fornecimento adicional de 1103 unidades, o que é igual a 110,3% da demanda do cliente.

Figura 1: Exemplo da propagação de erro na previsão de demanda.



Fonte: Mentzer e Moon (2005).

De acordo com Lustosa et al. (2008), as oscilações da demanda a montante nas cadeias de abastecimento são maiores que as oscilações observadas no varejo. Isto porque a política de compra de grandes lotes em um esforço para alcançar economias de escala faz com que as oscilações na demanda do varejo pareçam menores em comparação. Este fenômeno, conhecido como efeito "bullwhip", torna a gestão de estoques na cadeia de abastecimento particularmente difícil devido à complexidade que acrescenta ao processo.

Além disso, Lustosa et al. (2008) afirmam que o resultado econômico de uma empresa será diretamente impactado por sua capacidade de prever bem. Este é um pensamento compartilhado por Chopra (2003), que afirma que duas previsões de vendas diferentes, cada uma com diferentes níveis de erro, resultarão na adoção de uma política de gerenciamento de estoque diferente, o que resultará em última instância em resultados melhores ou piores.

2.1.2 Identificando características da demanda

O primeiro tipo de demanda que pode ser diferenciado dos demais é a demanda única, em oposição à demanda repetida. Uma demanda única é uma necessidade que surge de forma concentrada ou apenas uma vez, e depois desaparece completamente ou se reduz

rapidamente. Ao lidar com uma demanda desta natureza, o desafio da previsão reside na ausência de um histórico de demanda em que se possa confiar (LUSTOSA et al., 2008).

Há um segundo tipo de separação que pode ser feita sobre itens de demanda recorrente, e que é entre a demanda dependente e a demanda independente. De acordo com Slack (2007), a demanda dependente é uma forma de demanda previsível porque está relacionada com coisas que já são conhecidas. Uma excelente ilustração disto seria a demanda que um fabricante de automóveis tem por pneus. Como estamos cientes do número de automóveis fabricados, podemos prever que haverá cinco vezes mais pneus do que os automóveis criados. Portanto, a quantidade de pneus é diretamente proporcional ao número de automóveis.

Lustosa et al. (2008) distingue entre a demanda estacionária e a demanda de tendência para itens com demanda independente. A demanda estacionária é definida por um platô estável, com mudanças aleatórias identificadas ao seu redor. A demanda de tendências, por outro lado, é caracterizada por uma trajetória ascendente. Uma tendência de demanda crescente ou decrescente se correlaciona com um aumento ou diminuição correspondente nas vendas.

A sazonalidade é ainda outra faceta, e é exemplificada pelas diferenças que ocorrem em intervalos regulares durante um ciclo. Os protetores solares são um bom exemplo porque são vendidos durante todo o ano, mas atingem seu maior volume de vendas no verão.

Produzir projeções de alta qualidade não é um problema que possa ser facilmente resolvido por robôs ou pela maioria dos analistas. Assim, Taylor e Letham (2017) identificaram duas tendências principais que surgiram ao longo do processo de desenvolvimento de uma ampla gama de previsões da empresa:

a) As técnicas para previsões totalmente autônomas podem ser frágeis e freqüentemente rígidas demais para integrar suposições úteis ou heurísticas.

b) É bastante difícil encontrar analistas que sejam capazes de produzir previsões de alta qualidade porque a previsão é uma capacidade de ciência de dados especializada que requer uma quantidade significativa de experiência.

2.1.3 Métodos de previsão de demanda

Após muitos anos de pesquisa na área de previsão de vendas, foram criados vários métodos diferentes. De acordo com Ballou (2001) e Lustosa et al. (2008), estes métodos podem ser divididos em três categorias: qualitativa, projeções históricas e causal. As principais distinções entre estas abordagens são o quão assertivas são, o quão subjetivas são, e o quão simples são.

A organização enfrenta uma decisão desafiadora ao determinar a melhor abordagem a ser tomada, e vários estudos sugerem que para os melhores resultados, eles devem utilizar múltiplas abordagens e então combinar os resultados de cada uma (CHOPRA, 2003).

De acordo com Makridakis (1986), diferentes estudos chegaram a diferentes conclusões quanto ao desempenho dos métodos, e não é possível concluir que um método produz consistentemente melhores resultados do que outros. Isto porque diferentes estudos utilizaram critérios diferentes para avaliar a eficácia dos métodos.

A seguir, a tabela XX apresenta os métodos levantados na pesquisa bibliográfica para previsão de demanda. Além das categorias levantadas por Ballou (2001) e Lustosa et al. (2008), também está incluída uma seção para métodos de Inteligência Artificial que lista os algoritmos selecionados para o estudo.

Tabela 1: Métodos de previsão de demanda.

Classificação	Técnica
Métodos qualitativos	Pesquisa de mercado
	Estimativa de força de vendas
	Delphi
Projeções históricas	Média móvel
	Suavização exponencial
Causais	Regressão linear (simples/múltipla)
	ARIMA
Inteligência Artificial	Facebook Prophet
	LSTM
	LightGBM

Fonte: Adaptado de Ballou (2001).

Não serão analisados os métodos qualitativos por descompasso com a proposta. Para as demais categorias, foram selecionados os modelos mais sofisticados de cada uma para comparação, com exceção dos de Inteligência Artificial que terão análises para todos os métodos levantados. Assim, serão comparados a Suavização Exponencial, o ARIMA, o Facebook Prophet, o LSTM RNN e o LightGBM.

2.2 Suavização exponencial

2.2.1 Suavização exponencial simples

Esta estratégia funciona bem para uma demanda que não segue nenhuma tendência ou padrão sazonal. De acordo com Ballou (2001), este é provavelmente o método mais benéfico para previsões a curto prazo devido à sua simplicidade e ao fato de não haver necessidade de utilizar dados de uma qualidade particularmente alta. Ele oferece o resultado mais preciso em comparação com outros modelos concorrentes (simplicidade), e devido à ponderação, ele se ajusta automaticamente a diferenças significativas no ambiente.

O valor da constante de ponderação exponencial pode variar de 0 a 1, mas tenha em mente que quanto maior for seu valor, mais peso é dado em períodos mais recentes da série, o

que, por sua vez, faz com que o modelo reaja mais rapidamente às mudanças no ambiente. A previsão, por outro lado, se tornará mais volátil à medida que o valor constante aumentar, respondendo mais a flutuações aleatórias do que a mudanças fundamentais no estado de coisas (BALLOU, 2001).

2.2.2 Suavização exponencial com tendência (modelo de Holt)

Mesmo enquanto a simples suavização exponencial produz bons resultados, quando há uma grande tendência ou padrão sazonal nos dados, um atraso será inerente a resultados, o que resultará em erros de previsão inaceitáveis. Isto pode ser evitado com o uso de métodos de suavização mais complexos. (BALLOU, 2001)

Ao corrigir as tendências, uma segunda variável que reflete o aumento da demanda de um período para o outro é incluída na análise. De forma análoga à da previsão usando o método exponencial simples, esta variável recentemente introduzida será atualizada exponencialmente, com a introdução de uma nova constante (LUSTOSA et al., 2008).

Makridakis e Hibon (1991) demonstraram que a seleção dos valores iniciais não tem um impacto significativo sobre o resultado da previsão, e que é possível determinar estes valores usando métodos simples, ou mesmo defini-los iguais a zero na iteração inicial da previsão. Por outro lado, Lustosa et al. (2008) afirmam que os valores das constantes suavizantes, que são denotados pelos símbolos e, têm um impacto significativo no resultado, e a calibração do modelo está focada na determinação dos valores apropriados para estes parâmetros.

2.2.3 Suavização exponencial com tendência e sazonalidade (modelo Holt Winters)

Quando há componentes de tendência, bem como sazonalidade presentes, esta iteração particular da abordagem é utilizada. Neste cenário, um índice de sazonalidade é desenvolvido, e então este efeito é subtraído da demanda de base antes de ser adicionado à previsão no final da técnica (LUSTOSA et al., 2008).

Este modelo só deve ser utilizado no caso em que a demanda seja "estável, significativa e reconhecível a partir de variações aleatórias", como declarado por Ballou (2001). Em qualquer outro caso, as outras variantes do modelo, como aquela com uma constante de ponderação que tenha um alto valor atribuído a ela para eliminar o problema de defasagem, podem produzir resultados superiores.

O processo de solução da técnica funciona da mesma forma que os anteriores, mas agora há componentes adicionais que precisam ser computados.

De acordo com Lustosa et al. (2008), um dos desafios associados ao método é a exigência de séries históricas mais longas. No caso de anos, esta exigência se traduz em pelo menos dois ciclos sazonais completos.

2.3 Autoregressive Integrated Moving Average (ARIMA)

2.3.1 Componentes do modelo

Um modelo autoregressivo integrado de média móvel (ARIMA) é uma forma de análise de regressão que mede a força de uma variável dependente em relação a outras variáveis em mudança. O modelo é uma combinação do modelo autoregressivo diferenciado com o modelo de média móvel. (HYNDMAN; ATHANASOPOULOS, 2018)

Um modelo ARIMA pode ser entendido delineando cada um de seus componentes da seguinte forma:

a) Autoregressão (AR): refere-se a um modelo que mostra uma variável em mudança que regride sobre seus próprios valores defasados, ou anteriores;
b) Integrado (I): representa a diferenciação das observações brutas para permitir que a série temporal se torne estacionária (ou seja, os valores dos dados são substituídos pela diferença entre os valores dos dados e os valores anteriores);
c) Média móvel (MA): incorpora a dependência entre uma observação e um erro residual de um modelo de média móvel aplicado a observações defasadas.

2.3.2 Parâmetros

Para programar o ARIMA, é necessário definir os parâmetros p, d, e q. Eles podem ser definidos como:

a) p: o número de observações de atraso no modelo, também conhecido como a ordem de atraso (lag). Este parâmetro é responsável da modelagem da auto-regressão;
b) d: o número de vezes que as observações brutas são diferenciadas, também conhecido como o grau de divergência. Este parâmetro é responsável pela modelagem da diferenciação;
c) q: o tamanho da janela da média móvel. também conhecida como a ordem da média móvel.

2.3.3 Análise de estacionariedade

De acordo com Hyndman e Athanasopoulos (2018), para ingestão no ARIMA os dados são diferenciados a fim de torná-los estacionários. Um modelo que mostra a estacionariedade é aquele que mostra que há constância nos dados ao longo do tempo. A maioria dos dados econômicos e de mercado mostram tendências, portanto, o objetivo da diferenciação é remover quaisquer tendências ou estruturas sazonais.

A sazonalidade, ou quando os dados mostram padrões regulares e previsíveis que se repetem ao longo de um ano civil, pode afetar negativamente o modelo de regressão. Se uma tendência aparece e a estacionaridade não é evidente, muitos dos cálculos ao longo do processo não podem ser feitos com grande eficácia. Para mitigar este problema em situações onde há sazonalidade presente nos dados, é proposta a modelagem do Seasonal ARIMA (SARIMA).

2.3.4 Seasonal ARIMA (SARIMA)

O Seasonal ARIMA, ou SARIMA, é uma extensão da ARIMA que suporta explicitamente dados de séries temporais univariadas com um componente sazonal.

Ele adiciona três novos hiperparâmetros para especificar a autoregressão (AR), diferenciação (I) e média móvel (MA) para o componente sazonal da série, bem como um parâmetro adicional para o período da sazonalidade. (HYNDMAN; ATHANASOPOULOS, 2018)

Portanto, os parâmetros descritos anteriormente para modelagem do ARIMA se tornam os elementos da tendência da série temporal, enquanto surgem quatro novos parâmetros para modelagem da sazonalidade, sendo eles:

- P: a ordem de autoregressão, número de observações de atraso no modelo (lag), da sazonalidade;
- D: a ordem de diferenciação da sazonalidade;
- Q: o tamanho da janela de média móvel da sazonalidade;
- m: o número de unidades de tempo que completam um ciclo de sazonalidade.

Assim, a aplicação do SARIMA pode incluir potencialmente um grande número de combinações de parâmetros, sendo apropriado o teste de uma grande variedade de combinações para conseguir o melhor modelo. (HYNDMAN; ATHANASOPOULOS, 2018)

2.4 Machine Learning

2.4.1 Diferencial de métodos de Machine Learning

O aprendizado de máquinas, ou Machine Learning, é um ramo da inteligência artificial e da informática que se concentra no uso de dados e algoritmos para imitar a maneira como os humanos aprendem, melhorando gradualmente sua precisão.

De acordo com o site oficial da universidade de Berkeley [XX], o típico algoritmo de aprendizagem de máquina consiste de aproximadamente três componentes:

a) Um processo de decisão: Uma receita de cálculos ou outras etapas que leva nos dados e "adivinha" que tipo de padrão seu algoritmo está procurando encontrar.

b) Uma função de erro: Um método para medir quão bom foi o palpite, comparando-o com exemplos conhecidos (quando eles estão disponíveis). O processo de decisão deu certo? Se não, como você quantifica "o quão ruim" foi o erro.
c) Um processo de atualização ou otimização: Um método no qual o algoritmo analisa o erro e depois atualiza como o processo de decisão chega à decisão final, para que da próxima vez o erro não seja tão grande.

Uma diferença relevante de algoritmos de Machine Learning para outros métodos estatísticos é que, neles, o programador não precisa modelar/projetar como o algoritmo irá prever a variável objetivo. Basta entregar exemplos para o algoritmo aprender o que deve buscar, e ele aprenderá iterativamente. Isso é especialmente bom para problemas complexos e não-lineares, pois não é necessário detalhar cada etapa da solução e todas as relações entre as variáveis do problema

2.4.2 Aplicações de machine-learning em projeção de séries temporais

No campo da previsão, modelos de aprendizagem de máquinas surgiram na última década como concorrentes significativos aos modelos estatísticos mais tradicionais. A invenção do modelo de rede neural nos anos 80 marcou o início do projeto de pesquisa. Em seguida, a pesquisa expandiu a noção para incluir outros modelos, tais como máquinas vetoriais de suporte, árvores de decisão e outros, que são referidos coletivamente como modelos de aprendizagem de máquinas (Alpaydin, 2004; Hastie et al., 2001).

Há alguns destes modelos que podem ser rastreados até a literatura estatística inicial (ver Hastie et al., 2001 para uma discussão). Nos últimos anos, houve avanços significativos neste assunto, tanto na compreensão teórica dos modelos quanto na quantidade e variedade dos modelos que foram gerados. Estes avanços têm sido surpreendentes. Além do processo de desenvolvimento e análise dos modelos, é necessário um esforço paralelo de validação empírica dos muitos modelos diferentes que já estão em uso e contrastando os resultados destes modelos. Isto seria de tremendo valor para o praticante, pois o ajudaria a limitar suas escolhas prospectivas e lhe proporcionaria uma visão das áreas fortes e fracas dos muitos modelos que são acessíveis. Além disso, isto ajudaria a direcionar os esforços de pesquisa na direção das linhas de investigação mais promissoras.

Estudos em grande escala com o objetivo de comparar diferentes modelos de aprendizagem de máquinas se concentraram quase inteiramente no domínio da classificação (Caruana e Niculescu-Mizil, 2006). Não conseguimos localizar nenhuma pesquisa aprofundada sobre o tema das dificuldades de regressão. Quando se trata de previsões e outros tipos de questões econométricas, tem havido muitas pesquisas que comparam as redes neurais com as metodologias lineares padrão. Por exemplo, Sharda e Patil (1992) compararam as redes neurais com a ARIMA nos dados da série temporal M-competition. Eles descobriram que as redes neurais tiveram melhor desempenho. Hill et al. (1996) levaram em conta os dados da série M e compararam as redes neurais com as técnicas tradicionais de resolução de problemas. Em seu estudo de 1995, Swanson e White estenderam sua análise a nove variáveis macroeconômicas diferentes dos Estados Unidos. Sobre dados de vendas a varejo, Alon et al. (2001) compararam as redes neurais a outras abordagens tradicionais, incluindo suavização exponencial Winters, Box-Jenkins ARIMA e regressão multivariada. Este estudo foi conduzido comparando as redes neurais a métodos tradicionais. Em seu estudo de 1996, Callen e colegas examinaram o poder preditivo das redes neurais e modelos lineares para 296 séries temporais de ganhos diferentes. Zhang et al. (2004) consideraram uma situação muito semelhante, porém incluíram alguns fatores contábeis extras em sua análise. Em sua análise de 47 séries macroeconômicas diferentes, Terasvirta et al. (2005) levaram em conta redes neurais, autorregressões de transição suave e modelos lineares. Os resultados de todas essas pesquisas têm sido um tanto inconsistentes, mas, em média, as redes neurais têm demonstrado um desempenho melhor do que os métodos lineares tradicionais.

2.5 Facebook Prophet

O Facebook Prophet é um algoritmo desenvolvido pela equipe da Meta para facilitar o acesso à predição de séries temporais para especialistas e entusiastas. Porém, apesar de fácil implementação, o modelo é deveras sofisticado, tendo aplicação de técnicas diferentes para alcançar os melhores resultados, como o ARIMA e a suavização exponencial.

Em sua essência, o procedimento do Prophet é um modelo de regressão aditiva com quatro componentes principais:

a) Uma tendência linear ou de curva de crescimento logístico por partes. Prophet detecta automaticamente mudanças nas tendências, selecionando pontos de mudança a partir dos dados.
b) Um componente sazonal anual modelado utilizando a série Fourier.
c) Um componente sazonal semanal usando variáveis dummy.
d) Uma lista de feriados importantes, fornecida pelo usuário.

A idéia importante do Prophet é que ao fazer um trabalho melhor de ajuste ao componente tendência de forma muito flexível, é possível modelar com mais precisão a sazonalidade. Assim, ele é um modelo de regressão muito flexível comparado a modelos tradicionais de série temporal por facilitar o ajuste do modelo e lidar com dados ausentes ou outliers de forma mais sofisticada. Por padrão, Prophet fornece intervalos de incerteza para o componente de tendência, simulando futuras mudanças de tendência para suas séries temporais.

2.6 Long-Short Term Memory Recurrent Neural Network (LSTM)

2.6.1 Redes neurais

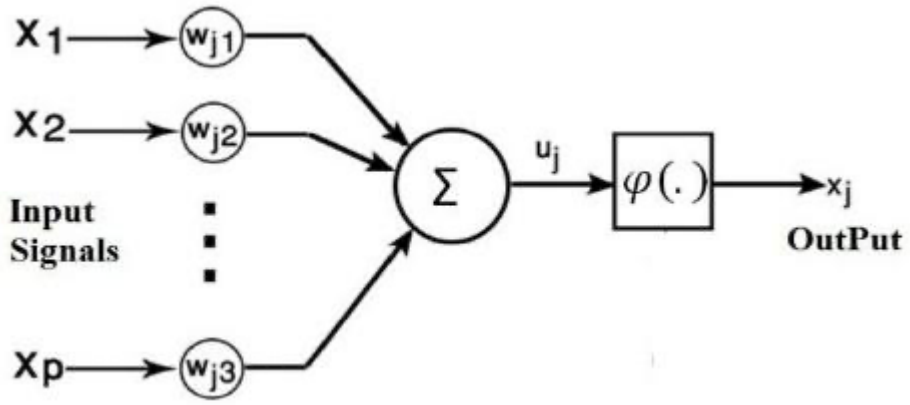
As Redes Neurais são algoritmos que buscam simular a capacidade do cérebro quando se trata de mudar a forma como processam as informações. Um modelo matemático e computacional que se inspira nos mecanismos de aprendizagem do cérebro humano e como ele realiza uma função específica é chamado de Rede Neural Artificial (ANN). Suas unidades de processamento são neurônios artificiais que estão interligados por meio de redes e são capazes de mudar a forma como processam, armazenam e transferem informações do ambiente externo. Estas unidades de processamento são acopladas por meio de redes (HAYKIN, 2009; SOARES, 2008).

Como podem ser aplicadas em uma variedade de contextos, as características das ANNs conhecidas como adaptabilidade e generalização da informação estão entre as mais valiosas de suas características. Estes modelos matemáticos têm a capacidade de modificar seus parâmetros internos de cálculo em resposta à adição de amostras externas que são introduzidas na rede. Com isto, eles serão capazes de aprender através de treinamento a maneira ideal de se adaptar ao modelo de dados que foi proposto. Além disso, o sistema cresce mais tolerante a erros de entrada como resultado do fato de que se a rede neural detectar algo incomum, o sistema é capaz

de escolher se deve ou não ignorar a entrada com base no treinamento que recebeu (DA SILVA; FLAUZINO; SPATTI, 2010).

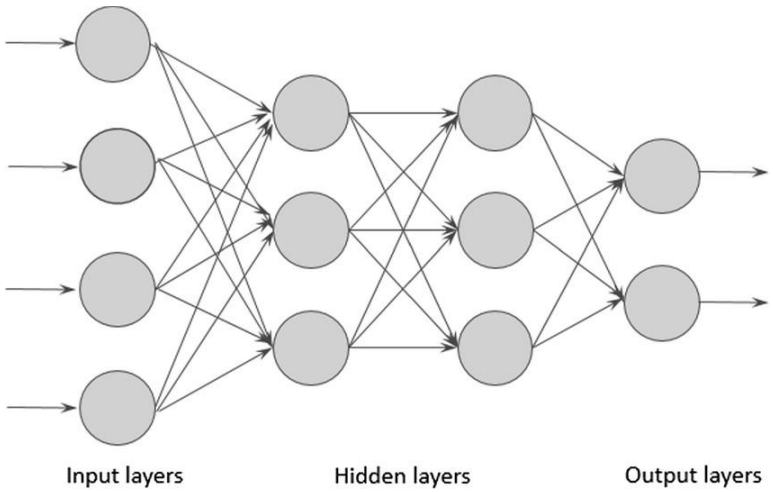
A figura 2 representa a arquitetura de um neurônio individual da rede. Já a figura 3 representa a arquitetura de uma rede neural, onde cada círculo representa um neurônio individual, e cada seta representa a direção da alimentação dos dados de input e output.

Figura 2: Arquitetura de Neurônios, chamados também de Perceptron.



Fonte: Haykin (2009).

Figura 3: Arquitetura de redes neurais.



Fonte: Haykin (2009).

2.6.2 Redes neurais recorrentes (RNN)

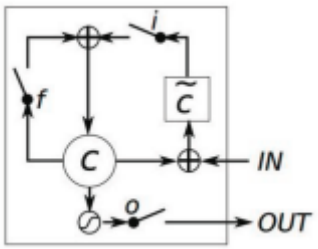
As redes neurais recorrentes (RNNs) são um tipo especial de rede neural desenvolvida para processar dados sequenciais. Os primeiros estudos foram realizados nos anos 80 por Hopfield (1984) e Rummelhart et al. (1985). Em estruturas tradicionais, cada amostra é treinada independentemente uma da outra, mas este treinamento não é um método suficiente para texto, som, imagem ou outros dados relacionados ao tempo como séries temporais.

As RNNs oferecem esta solução de sondagem. Ao contrário das redes neurais tradicionais, as RNNs têm conexões de feedback nas unidades de camada oculta. Com esta característica, eles podem realizar o processamento temporal e aprender sequências. A camada oculta atua como uma memória e pode armazenar estas informações.

2.6.3 LSTM

LSTM é um modelo RNN personalizado desenvolvido por Hochreiter e Schmidhuber (1997) que pode manter o rastro de informações a longo prazo através dos portões que contêm. Em uma unidade LSTM há basicamente três portões, que determinam quais informações armazenar, ou seja, o portão de entrada, o portão de esquecimento e o portão de saída. O portão de entrada especifica quais dos valores de entrada serão armazenados no próximo estado. O portal do esquecimento determina quais das informações do estado anterior não serão mais armazenadas. O portão de saída especifica qual das informações no novo estado será enviado como saída. Os portões que compõem a unidade LSTM são mostrados na Figura 4, onde **i**, **f**, e **o** representam o portão de entrada, o portão de esquecimento e o portão de saída, respectivamente. C representa o estado da unidade e Ĉrepresenta o próximo estado candidato.

Figura 4: Neurônio unitário de uma LSTM.



Fonte: Adaptado de Hochreiter e Schmidhuber (1997)

As equações usadas para calcular a saída e o estado de cada célula unitária do LSTM são as seguintes.

$$f_t = \sigma(W_f * [x(t), C(t-1), h(t-1)] + b_f) \quad (1)$$

$$i_t = \sigma(W_i * [x(t), C(t-1), h(t-1)] + b_i) \quad (2)$$

$$o_t = \sigma(W_o * [x(t), C(t-1), h(t-1)] + b_o) \quad (3)$$

$$C(t) = C(t-1) * f_t + \hat{C} * i_t \quad (4)$$

Nelas, σ é a função de ativação, x(t) é a entrada e h(t - 1) é a saída anterior. Já W_f, W_i, W_o e b_f, b_i, b_o são os pesos e os enviesamentos do portão de esquecimento, do portão de entrada e do portão de saída, respectivamente.

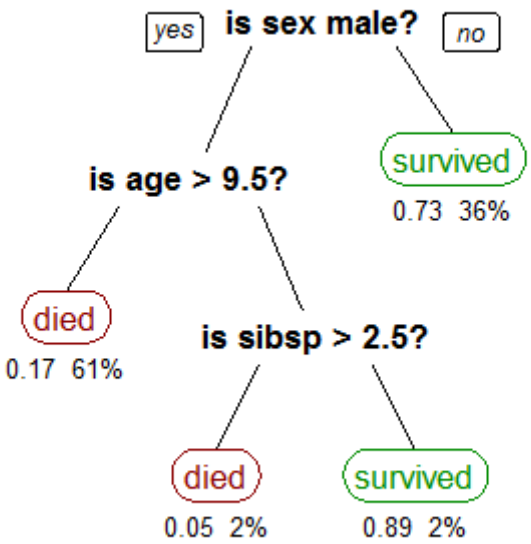
2.7 Light Gradient Boosting Machine (LightGBM)

2.7.1 Árvores de decisão

Um grupo de algoritmos relevantes no campo de machine learning é o dos algoritmos baseados em árvores de decisão. Árvores de decisão são uma forma de modelagem de simples interpretação, onde utiliza-se uma divisão binária recursiva dos atributos de um banco de dados a fim de construir um fluxograma no qual temos os componentes listados a seguir. Um exemplo de árvore de decisão se encontra na figura 5.

a) Nós: representando o teste em algum atributo do banco de dados;
b) Folhas: desdobramentos do teste, onde os dados são classificados com um valor ou classe;
c) Ramos: ligações entre nós e folhas.

Figura 5 Exemplo de árvore de decisão.



Fonte: Adaptado de James et al., 2013.

No exemplo acima, temos uma árvore de decisão baseada em três atributos: sexo (sex), idade (age) e número de irmãos/cônjuges à bordo (sibsp), todos para prever se um passageiro à bordo do titanic sobreviveu ou não ao naufrágio do navio. A árvore de decisão, portanto, assimila uma probabilidade de 36% de um paciente do sexo feminino sobreviver, enquanto pacientes do sexo masculino têm divisões mais detalhadas de sua chance de sobrevivência.

As divisões dos atributos são construídas a partir de uma função, como por exemplo a função de impureza de Gini, que calcula o quão bem um nó divide o banco de dados em duas classes distintas. (JAMES et al., 2013)

2.7.2 Bagging e boosting

Árvores de decisão, por si só, são algoritmos pouco confiáveis por terem uma grande variabilidade de resultados com pequenas mudanças nos valores dados. Para mitigar este problema, existem duas metodologias que são usadas para aprimorar este tipo de modelagem: o bagging e o boosting. (JAMES et al., 2013)

O termo bagging, ou bootstrap aggregating, propõe o treino de diversos modelos diferentes e independentes por meio da aplicação de bootstrapping em amostras do banco de dados original, e utilizando algum critério para ponderar os resultados de cada um para obter o resultado final de predição. Um exemplo de aplicação de bagging para algoritmos de árvores de decisão é o algoritmo Random Forests. (JAMES et al., 2013)

Já o boosting propõe aplicar métodos de forma sequencial, treinando iterativamente modelos com os resultados de modelos treinados anteriormente no banco de dados, alcançando a cada passo da iteração um modelo de predição mais robusto (JAMES et al., 2013). Um exemplo destes algoritmos é o LightGBM.

2.7.3 LightGBM

O LightGBM é um modelo de árvores de decisão com aplicação de boosting, onde o principal diferencial de outros métodos é que ele não cresce em nível de árvore - linha por linha - como a maioria das outras implementações. Em vez disso, ele cresce em nível de folha de árvore. Ele escolhe a folha que acredita produzir a maior redução de perdas. Além disso, ele implementa um algoritmo de aprendizado de árvore de decisão baseado em histograma altamente otimizado, que produz grandes vantagens tanto na eficiência quanto no consumo de memória. (GUOLIN, 2017)

3 MÉTODO

Para comparar os diferentes métodos de predição de demanda, propõe-se uma método de 3 etapas: a seleção da base de dados, experimentos e a análise de resultados. Cada uma delas será detalhada nas seções seguintes.

3.1 Seleção de base de dados

Para desenvolver o projeto, foram selecionadas duas bases de dados. Uma univariada e a outra multivariada.

A primeira foi selecionada com a finalidade de demonstrar a realização do procedimento de previsão de demanda em uma série temporal simples, com tendência e sazonalidade estáveis e sem grandes irregularidades ao longo do ano. Portanto, esta base foi selecionada para fins ilustrativos.

A segunda foi selecionada para real comparação dos métodos escolhidos para estudo. Esta base apresenta múltiplas séries temporais multivariadas de comportamentos distintos. A intenção de selecioná-la é testar diferentes algoritmos em situações diversas de volatilidade das séries temporais, além de entender também se a relação de múltiplas variáveis ajuda na previsão.

Entretanto, é importante garantir que as bases usadas sejam de fácil acesso e de qualidade, para garantir tanto a replicabilidade da análise de forma adequada.

Para isso, foram selecionadas duas bases da plataforma Kaggle, que será descrita a seguir.

3.1.1 Plataforma Kaggle

O Kaggle é uma plataforma de comunidade online para cientistas de dados e entusiastas do aprendizado de máquinas. Ele permite aos usuários colaborar uns com os outros, encontrar e publicar conjuntos de dados, usar notebooks integrados à GPU e competir com outros cientistas de dados para resolver desafios da área.

A plataforma já realizou centenas de concursos de aprendizagem de máquina desde que a empresa foi fundada. As competições têm variado desde a melhoria do reconhecimento de gestos para o Microsoft Kinect, da realização de uma IA de futebol para o Manchester City até uma aprimoração da busca pelo bóson de Higgs no CERN.

Uma chave para a produção de conhecimentos trazida pela plataforma é o efeito do quadro de líderes ao vivo, que incentiva os participantes a continuar inovando além das melhores práticas existentes para obter prestígio na comunidade.

3.1.2 Base de dados univariada

A base de dados selecionada para testar métodos de forma univariada é a base de dados da competição Store Item Demand Forecasting Challenge, cujo link está presente no capítulo de referências do trabalho.

A base é composta por 5 anos de dados de vendas de itens de 10 lojas diferentes, totalizando 50 itens categorizados na base. Esta base é previamente limpa pela moderação da competição do Kaggle, apresentando séries temporais que sejam relativamente limpas e consistentes. Portanto, foi selecionada apenas para ilustração do procedimento de previsão.

3.1.3 Base de dados multivariada

Para a análise multivariada, a base de dados selecionada foi o Walmart Recruiting - Store Sales Forecasting, que está presente no capítulo de referências do trabalho. A base contém vendas históricas de 45 lojas da rede Walmart nos Estados Unidos, cada loja contendo em média 74 departamentos, sendo no total 81 departamentos distintos. Por departamentos, se quer dizer a categorização de diferentes SKU's de uma mesma loja, análogo ao que é denominado item na base univariada. Além disso, são incluídas outras variáveis para ajudar na predição de forma indireta, como o índice de desemprego e temperatura na região. As variáveis contidas no banco serão melhor exploradas no capítulo 4.

Esta base de dados não contém tratamento prévio para evitar bases de dados de difícil previsão, e portanto foi escolhida para a comparação dos métodos dado que simula bem uma

situação real de previsão de demanda onde não há garantia de regularidade no histórico de vendas.

Ambas as bases de dados são públicas e podem ser acessadas através dos links disponibilizados na bibliografia.

3.2 Experimentos

Selecionadas as duas bases de dados utilizadas no estudo, é necessário definir como serão utilizadas para comparar os diferentes algoritmos de previsão de demanda. Para este fim, são definidos três experimentos principais que são realizados nos capítulos 5, 6 e 7 do estudo.

No capítulo 5 o método descrito nesta seção será aplicado em uma série temporal da base univariada para fins de ilustração. A ideia é demonstrar o procedimento de previsão de demanda em uma série temporal previamente tratada, onde houve uma preocupação pela adequação dos dados ao exercício de previsão, mas que não necessariamente traduz uma situação próxima do dia-a-dia de uma empresa.

No capítulo 6, o método será aplicado em uma série temporal da base multivariada para demonstrar de forma detalhada o procedimento de previsão nesta base que não contém tratamento prévio, e que é mais próxima de um caso real de aplicação. Assim, será possível aferir o contraste dos resultados obtidos em uma base previamente tratada para adequação ao exercício de previsão com uma base mais próxima do que se encontra no dia-a-dia da indústria.

Por fim, no capítulo 7, será repetido o procedimento realizado no capítulo 6 para uma amostra aleatória de 100 bases temporais advindas da base multivariada, a fim de obter resultados estatisticamente significativos para comparar os algoritmos selecionados para o estudo. A aplicação neste capítulo não será tão detalhada quanto nos capítulos 5 e 6 por envolver um volume muito maior de séries temporais. Por isso a preocupação de ilustrar de forma detalhada nestes dois capítulos anteriores ao 7 o procedimento de previsão que foi realizado em cada uma das séries temporais estudadas.

As subseções seguintes descrevem as etapas pelas quais cada experimento descrito acima passa.

3.2.1 Definição de bases de treino e teste

Nesta etapa, é necessário preparar os dois bancos de dados para aplicação dos modelos propostos. Para evitar o que chamamos de *overfitting*, que é quando o modelo estatístico se torna exageradamente adequado à base de dados usada para treiná-lo mas não consegue performar bem com dados fora dessa amostra, faremos o que é chamado de *train-test split*.

Este processo consiste em separar uma parte da base de dados para estar fora do treino dos algoritmos. Assim, podemos usá-la para testá-lo com amostras que ele nunca interagiu, e assim analisar sua performance de forma adequada.

Resumindo então, uma parte do banco de dados será usado para treino dos algoritmos e a outra será usada para testá-los e avaliá-los de forma comparativa.

Ademais, para que ambos os bancos de dados sejam postos no mesmo horizonte temporal, ambos são agrupados por semana.

3.2.1.1 Banco de dados univariado

Na base de dados univariada, a proporção escolhida de train-test split representa que teremos 1460 dias para treinar o modelo e 365 dias de teste. Sendo assim, a base de treino vai do dia 01/01/2013 até o dia 31/12/2016, enquanto a base de teste segue do dia 01/01/2017 até 31/12/2017. O banco de dados de dados originalmente tem granularidade diária, mas será tratado agrupando as vendas por semana para fins de comparação com a segunda base de dados, que tem granularidade semanal.

3.2.1.2 Banco de dados multivariado

Já na base de dados multivariada, a proporção escolhida de train-test split representa que teremos 726 dias para treinar o modelo e 268 dias de teste. Sendo assim, a base de treino vai do dia 05/02/2010 até o dia 01/02/2012, enquanto a base de teste segue do dia 02/02/2012 até 26/10/2012. Importante ressaltar que, apesar da contagem de dias, este banco de dados apresenta as vendas e demais variáveis agrupadas por semana.

3.2.2 Adaptação para aprendizado supervisionado

Para aplicação de algoritmos supervisionados de machine learning, é necessário que se adapte a série temporal para uma estrutura tabular, de forma que o algoritmo em questão entenda os dados. Isso se faz por meio de um deslocamento horizontal dos dados. Basicamente, definindo um número de lags (dado por n), criamos uma coluna para cada registro de vendas anterior à data presente. Assim, temos uma coluna para vendas em t, t-1, t-2…t-n. Para ambas as bases, usaremos um lag dado por n = 52, dado que ambas estão agrupadas por semana e portanto escolheu-se incluir um ciclo anual inteiro na tabela. Este mesmo lag é aplicado para as demais variáveis presentes no caso multivariado. Um exemplo da aplicação deste método na base univariada se encontra a seguir.

Além de utilizar as vendas deslocadas temporalmente, é possível extrair features separadas da data presente das vendas. Separa-se o ano, o mês, a semana, o dia e quaisquer outras variáveis temporais advindas da data como uma coluna separada, para que o algoritmo estude-as como variáveis independentes.

Os algoritmos alimentados pelas bases adaptadas dessa forma no presente trabalho são o LSTM e o LightGBM.

Tabela 2: Adaptação de série temporal para aprendizado supervisionado.

date	sales(t-52)	sales(t-51)	...	sales(t-3)	sales(t-2)	sales(t-1)	sales(t)
2014-01-10	447	426	...	492	491	514	492
2014-01-17	426	456	...	491	514	492	502
2014-01-24	456	478	...	514	492	502	499
2014-01-31	478	475	...	492	502	499	496
2014-02-07	475	543	...	502	499	496	573

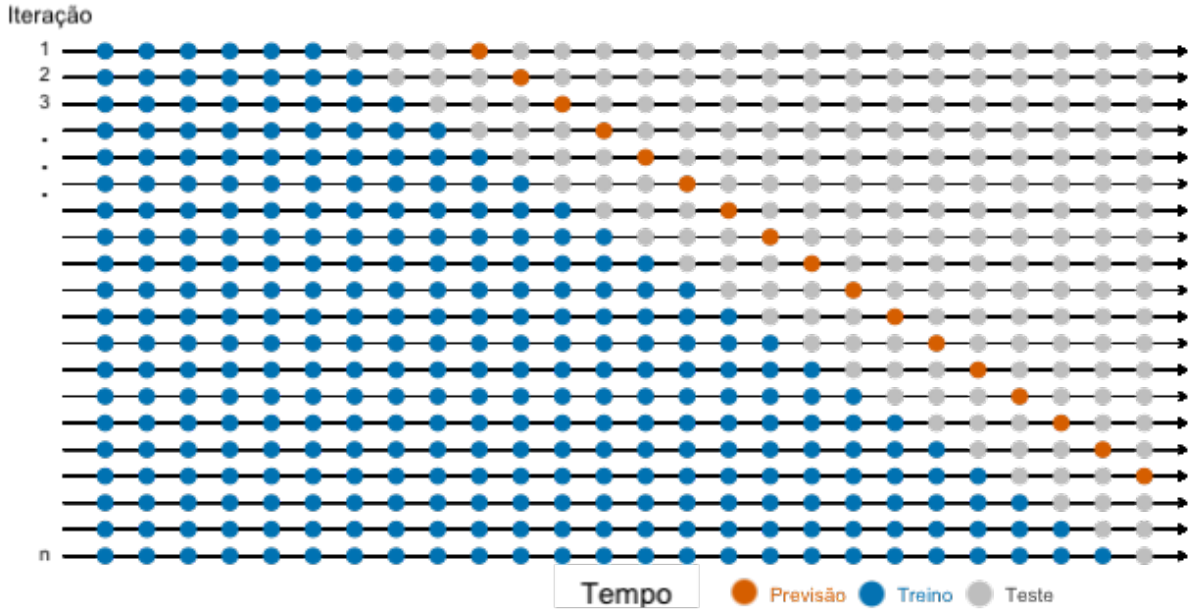
Fonte: Autor.

3.2.3 Aplicação dos métodos levantados

Nesta parte do experimento é realizada a aplicação dos algoritmos selecionados para o estudo, sendo eles a Suavização Exponencial, o ARIMA, o Prophet, o LSTM e o LightGBM.

Além disso, será aplicada a técnica de previsão contínua dos dados separados para teste. Isto significa que em cada semana, o modelo é re-treinado alimentando a base de treino com os novos dados observados da semana anterior (que anteriormente estavam separados na base de teste) de forma iterativa. Além disso, a cada iteração, o modelo prevê as vendas de 4 semanas à frente da semana atual, para evitar enviesamento por vazamento de dados futuros na base de dados presentes. Assim, simula-se uma situação real de previsão, onde prevê-se os dados com um mês de antecedência para preparo prévio da operação. Isto é importante não só para melhorar a previsão dos modelos mas também para equalizar o treino dos modelos estatísticos com os modelos de aprendizagem supervisionada, que são alimentados com os dados adaptados de forma tabular conforme observado na figura XX do item anterior. Este procedimento é representado pela figura XX a seguir.

Figura 6: Ilustração da modelagem de previsão contínua.



Fonte: Autor

3.3 Análise dos resultados

Por fim, a última etapa do trabalho visa comparar todos os algoritmos de forma que possamos aferir a qualidade da predição feita por cada um. Para isso, foi selecionada a métrica de Mean Absolute Percentage Error (MAPE), que será melhor descrita na subseção a seguir.

É importante ressaltar que a análise dos resultados é realizada no fim de cada um dos capítulos 5, 6 e 7.

3.3.1 Mean Absolute Percentage Error (MAPE)

O erro percentual médio absoluto (MAPE) mede a precisão de uma previsão como uma porcentagem, e pode ser calculado como o erro percentual absoluto médio para cada período de tempo menos os valores reais divididos pelos valores reais. Esta métrica é calculada segundo a equação 5.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \quad (5)$$

Aqui temos que $\hat{y}$ representa a predição e y consiste no valor observado da amostra, enquanto n representa o número de dados na amostra. A vantagem desta métrica é que as previsões podem ser comparadas entre algoritmos aplicados em séries temporais com volumes de vendas diferentes.

4 ANÁLISE EXPLORATÓRIA DAS BASES

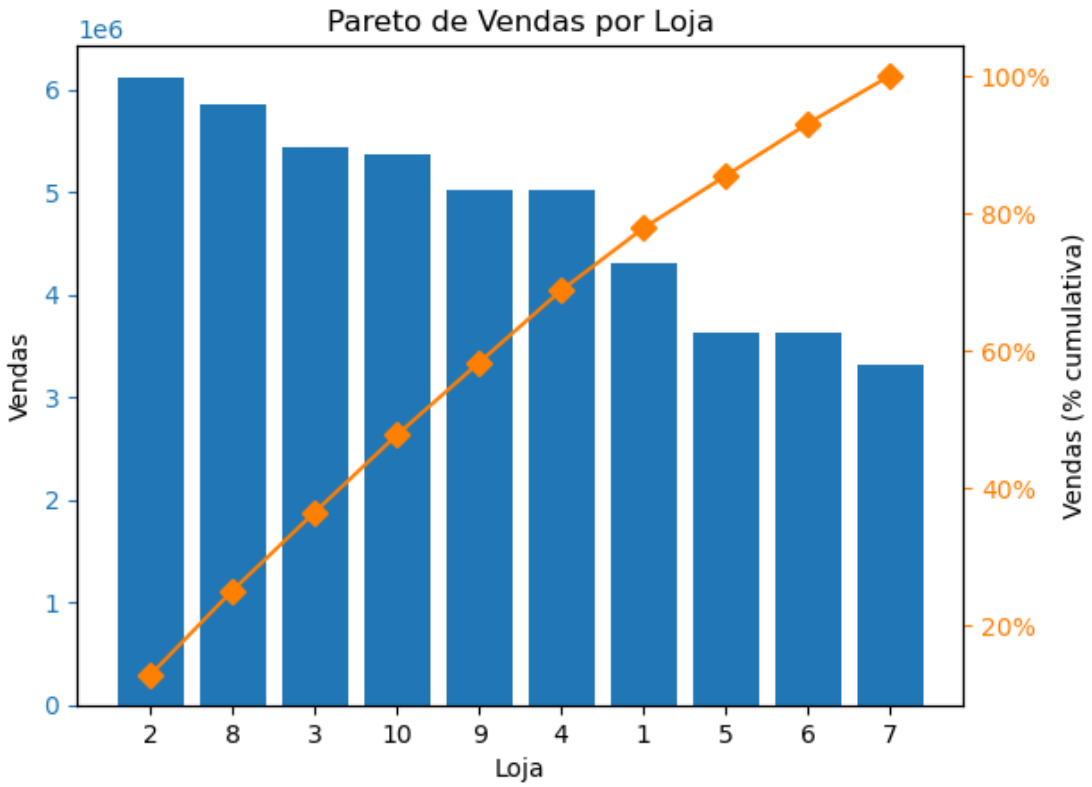
4.1 Base de dados univariada

4.1.1 Análise das séries de vendas por item e agregadas por loja

Primeiramente, é importante ressaltar que a base não contém nenhum valor nulo. Ela registra vendas de diversas lojas e itens do dia 01/01/2013 até o dia 31/12/2017.

Além disso, temos uma análise de Pareto agrupando os dados por loja e por item. Com isso, podemos mapear a distribuição de vendas entre essas duas categorias.

Figura 7: Pareto de vendas por loja da série univariada.



Fonte: Autor.

A partir da distribuição de vendas podemos ver que estas são relativamente bem distribuídas, de forma que a curva de porcentagem cumulativa nas duas se aproxima de uma reta (linear).

Importante ressaltar que esta base de dados apresenta vendas de 10 lojas, cada uma contendo registros de 50 itens diferentes. Portanto, na prática a aplicação de métodos

tradicionais tratará o problema como predição de 500 séries temporais em paralelo. Como a intenção com a base univariada é apenas ilustrar o processo de previsão em uma série temporal com poucas irregularidades, não é necessário utilizar todas as séries de vendas presentes na base.

Sendo assim, para simplificar a análise, foi selecionada a loja 2 pelo simples critério de ter o maior volume de vendas. A partir disso, foi selecionado um item de forma aleatória, resultando na escolha do item 15. Importante ressaltar que em todo o banco de dados há apenas um registro de 0 vendas e nenhum valor nulo, e portanto outras séries da base poderiam ser escolhidas para a demonstração.

4.1.2 Estatísticas descritivas

Analisando a série temporal selecionada (loja 2, departamento 15) e agrupando os valores de venda por semana, obtemos as seguintes estatísticas descritivas.

Tabela 3: Estatísticas descritivas das vendas.

Estatísticas descritivas: vendas	
Contagem	260
Soma	205.102
Média	788,9
Mediana	807
Desv. Padrão	165,6
Mínimo	426
Máximo	1.181

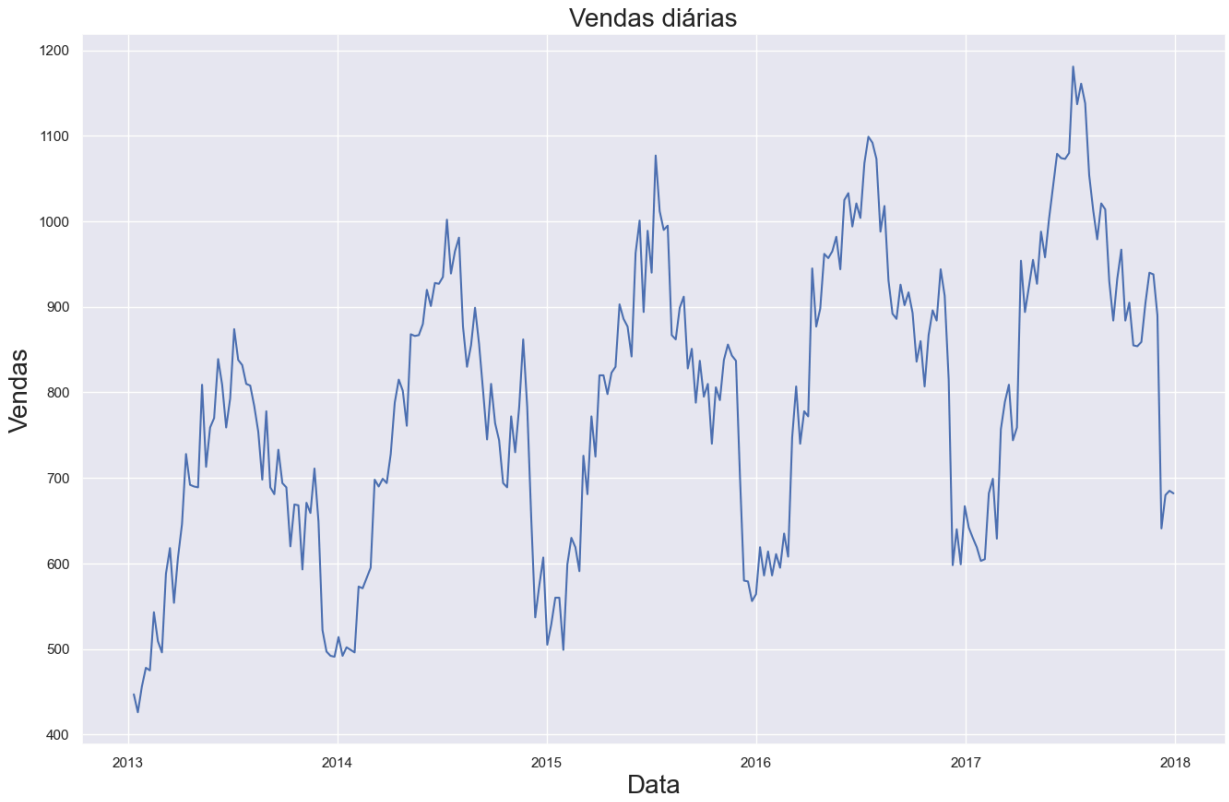
Fonte: Autor.

4.1.3 Descrição da série temporal escolhida para análise

Dado a seleção da loja 2 e do item 15, obtemos a seguinte série temporal. É possível notar um claro componente de sazonalidade anual, com uma leve tendência crescente ao longo dos anos. Importante ressaltar que os dados de vendas desta base originalmente têm granularidade diária, mas foram agrupados por semana para maior comparabilidade à base

multivariada efetivamente utilizada para o estudo de comparação dos métodos, que originalmente é agrupada por semana.

Figura 8: Série temporal de vendas da base univariada.

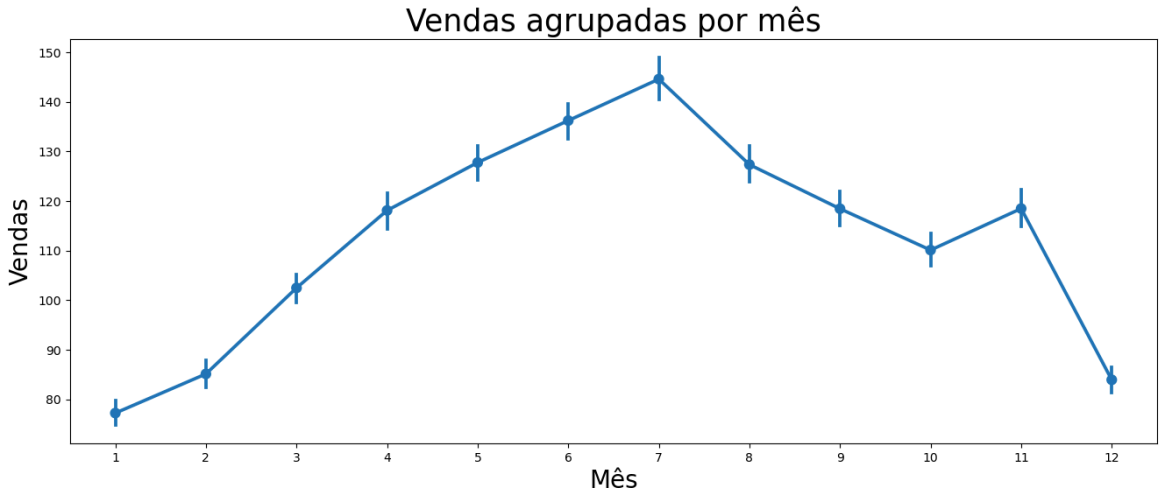


Fonte: Autor.

4.1.4 Sazonalidade ao longo do ano

A figura XX mostra a distribuição de vendas da amostra ao longo do ano, agrupado por mês. É possível observar que existem dois picos, um em Julho e outro em Novembro. Já o vale da curva se encontra em Janeiro, onde historicamente há tendência de haver queda do número de vendas.

Figura 9: Sazonalidade de vendas ao longo do ano.



Fonte: Autor.

4.2 Base de dados multivariada

Diferente da base de dados univariada, a base de dados multivariada apresenta séries de vendas com diferentes volatilidades e comportamentos, além de diversas variáveis que podem ser utilizadas para a previsão da demanda. Esta base será utilizada efetivamente para a comparação dos métodos selecionados, e ao longo deste capítulo será selecionada uma das séries temporais presentes nela para demonstração do procedimento de previsão de demanda.

Esta base contém dados de 45 lojas diferentes, cada uma contendo em média 74 departamentos (que é a denominação com a qual se categorizam os SKU's). Portanto, ela contém no total 3331 séries temporais multivariadas.

4.2.1 Descrição das variáveis

O banco de dados multivariado não contém variáveis nulas. As variáveis contidas no banco de dados são as seguintes:

a) **Date**: Campo contendo a data que referencia a semana de vendas. Portanto, a granularidade da base é semanal.
b) **Store**: Número de identificação da loja.
c) **Dept**: Número de identificação do departamento da loja. Uma loja contém diversos departamentos, e as vendas têm granularidade por departamento.

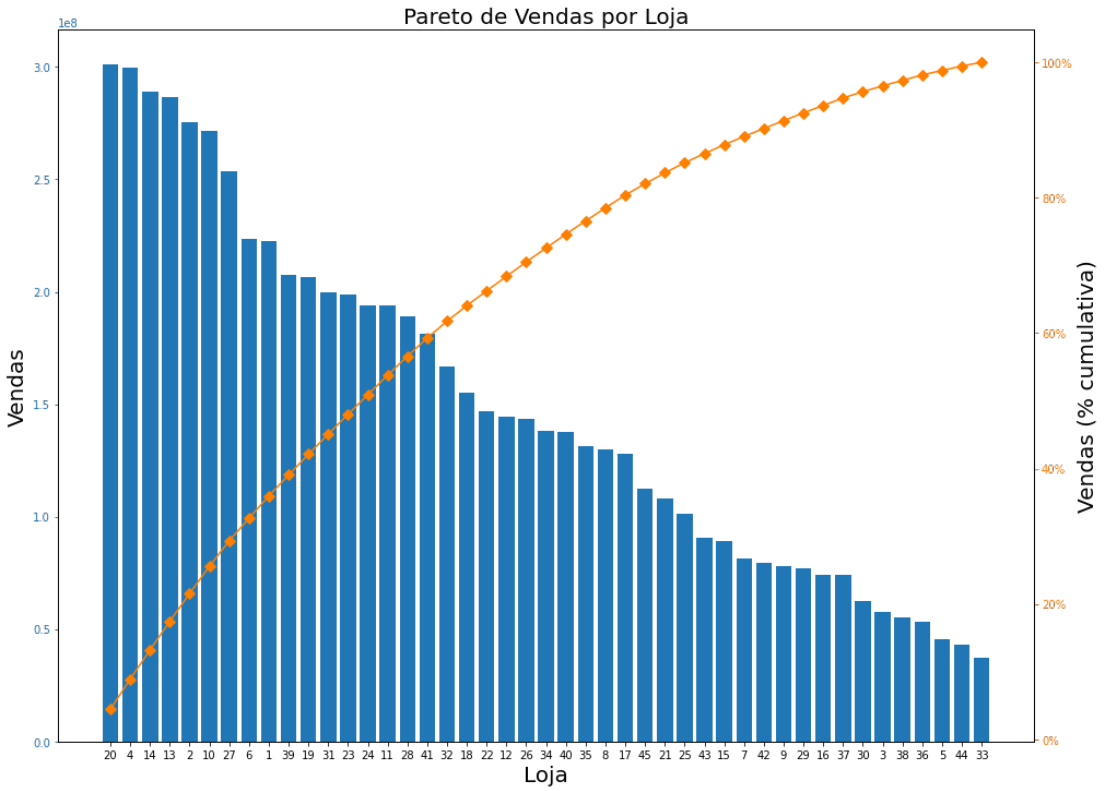
d) **Weekly_Sales**: Número de vendas na dada semana.
e) **Temperature**: Temperatura média na região em ºF.
f) **Fuel_Price**: Custo médio do combustível na região da loja.
g) **CPI**: Consumer price index, análogo ao IPCA no Brasil.
h) **Unemployment**: Taxa de desemprego.
i) **IsHoliday**: Campo booleano, onde 1 é feriado e 0 não.

4.2.2 Análise das séries de vendas por departamento e agregadas por loja

Analogamente à exploração da base de dados univariada, propõe-se selecionar uma das lojas e um dos departamentos da base de dados multivariada para que seja demonstrado o processo de previsão.

Isto é feito por dois motivos. Primeiro para demonstrar o contraste com o processo realizado na série temporal da base univariada, que apresenta um comportamento mais regular e fácil de ser previsto do que a base multivariada. Segundo, para demonstrar a aplicação do procedimento de previsão que será realizado no estudo em uma série de vendas mais próxima das que serão estudadas no capítulo 7. Para isso, temos a seguinte análise de Pareto da distribuição de vendas por cada loja.

Figura 10: Pareto de vendas por loja da série multivariada.



Fonte: Autor.

É possível constatar que ao longo das diferentes lojas, há uma boa diferença no volume de vendas. Para demonstração do processo de previsão, é escolhida a loja 20 pela simples razão de ter maior volume de vendas.

Já para a escolha do departamento, escolheu-se de forma aleatória. O resultado foi a escolha do departamento 1 da loja 20. Selecionando ambos, temos a série temporal da análise.

4.2.3 Estatísticas descritivas

O banco de dados contém entradas do dia 05/02/2010 até o dia 26/10/2012. Levando em conta todas as variáveis, com exceção da data, podemos levantar a seguinte tabela de estatísticas descritivas.

Tabela 4: Estatísticas descritivas da série multivariada

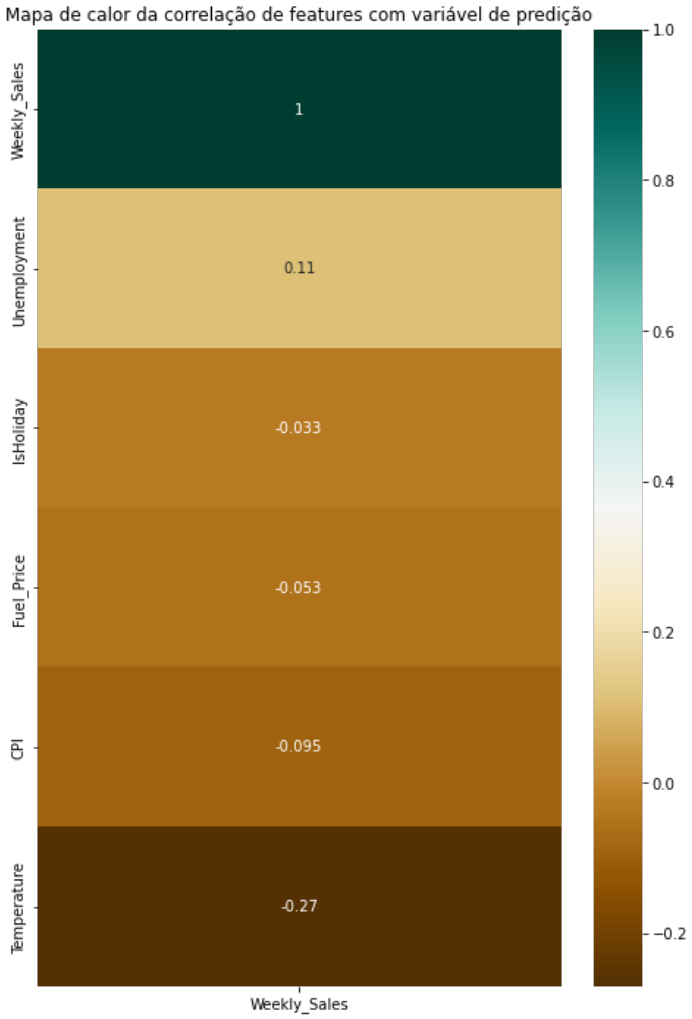
Métricas	Variáveis do banco de dados				
	Weekly_Sales	Temperature	Fuel_Price	CPI	Unemployment
Contagem	143	143	143	143	143
Média	40.545,5	68,3	3,2	216,0	7,6
Desvio padrão	21.554,2	14,3	0,4	4,4	0,4
Mínimo	24.174,6	35,4	2,5	210,3	6,6
Primeiro quartil (25%)	28.526,5	58,3	2,8	211,5	7,3
Mediana	33.117,4	69,6	3,3	215,5	7,8
Terceiro quartil (75%)	44.277,9	80,5	3,6	220,5	7,8
Máximo	156.234,3	91,7	3,9	223,4	8,1

4.2.4 Análise de correlações

Um dos principais motivos para usar uma base multivariada é poder usar variáveis complementares para ajudar a prever o número de vendas na semana especificada. Para que isso seja possível, um dos parâmetros interessantes que podemos analisar é a correlação entre as demais variáveis e a nossa variável de predição (no caso, *Weekly_Sales*).

À seguir, temos o mapa de calor que indica a correlação das diferentes variáveis do banco de dados com a variável principal.

Figura 11: Correlações das variáveis suporte com a variável de predição.



Fonte: Autor.

A partir da figura, podemos identificar que as variáveis de maior correlação com a variável de predição são respectivamente a temperatura, e o desemprego na região. As demais não aparentam ter relação estatisticamente significativa.

Importante ressaltar que, apesar de não ter uma correlação tão relevante quanto essas três variáveis, as demais variáveis de característica do banco de dados podem ter algum valor em modelos não-lineares que não é facilmente traduzido apenas pela correlação.

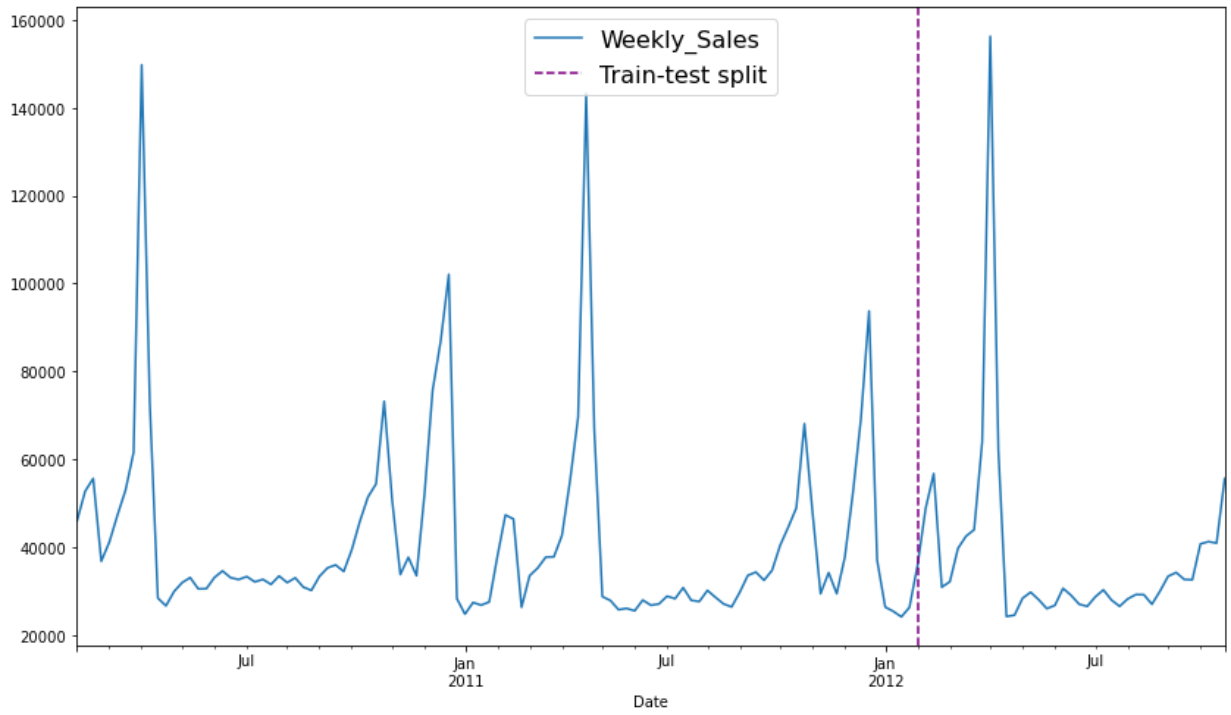
4.2.5 Descrição da série temporal escolhida para análise

Analisando apenas a variável de predição ao longo do tempo, obtemos a seguinte série temporal. Podemos perceber uma certa instabilidade da curva, com eventuais picos de demanda

ao longo do tempo. Além disso, aparenta ter uma certa sazonalidade de demanda, mas com ciclos pouco claros (não relacionados necessariamente a semana, mês ou ano, mas com pontos de inflexão aparentemente repetidos ao longo dos diferentes anos).

No gráfico, também foi indicada pela linha roxa pontilhada a separação entre base de treino e base de teste dos dados (ou train-test split). Para a esquerda dela, são dados para treino dos modelos, e à direita os dados de teste.

Figura 12: Vendas semanais da série multivariada.

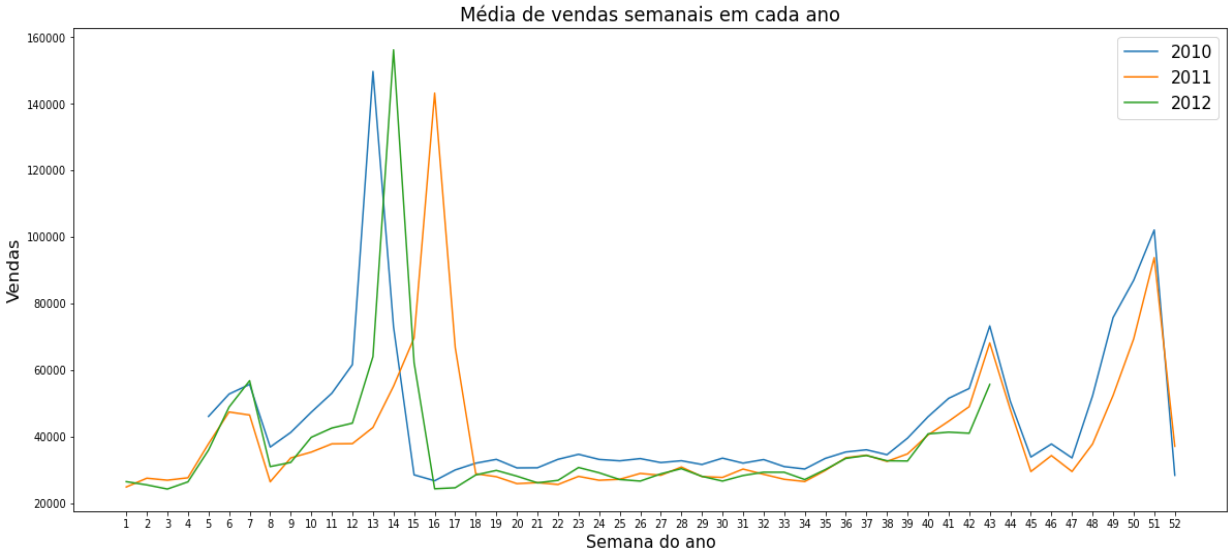


Fonte: Autor.

4.2.6 Sazonalidade ao longo do ano

Comparando o comportamento das vendas semanais na série temporal selecionada nos diferentes anos da base de dados, podemos aferir que os picos que aparecem na base apresentam certa ciclicidade. Porém, esta ciclicidade não é regular, dado que é percebido um deslocamento desses mesmos picos ao longo dos anos.

Figura 13: Sazonalidade anual das vendas da loja selecionada, série multivariada

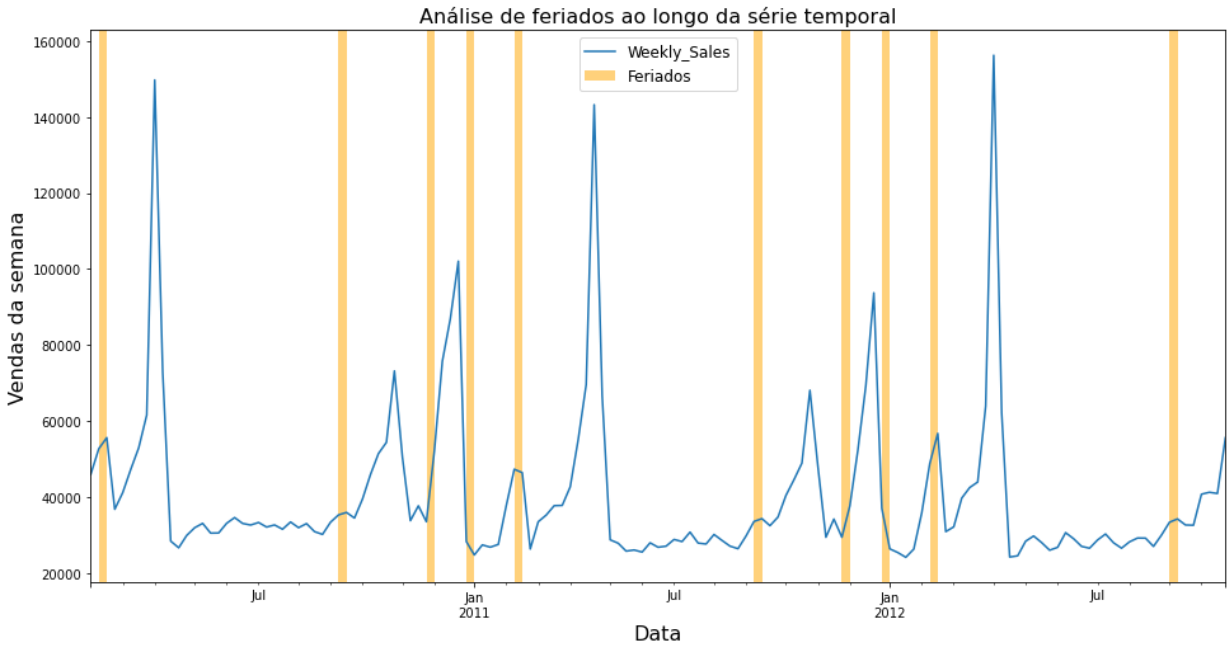


Fonte: Autor.

4.2.7 Análise do efeito de feriados

No gráfico à seguir, são destacados com faixas amarelas os feriados ao longo dos anos. Podemos ver que não há um padrão claro na relação destes feriados com os picos, talvez por conta de considerarmos todos os feriados como uma variável única.

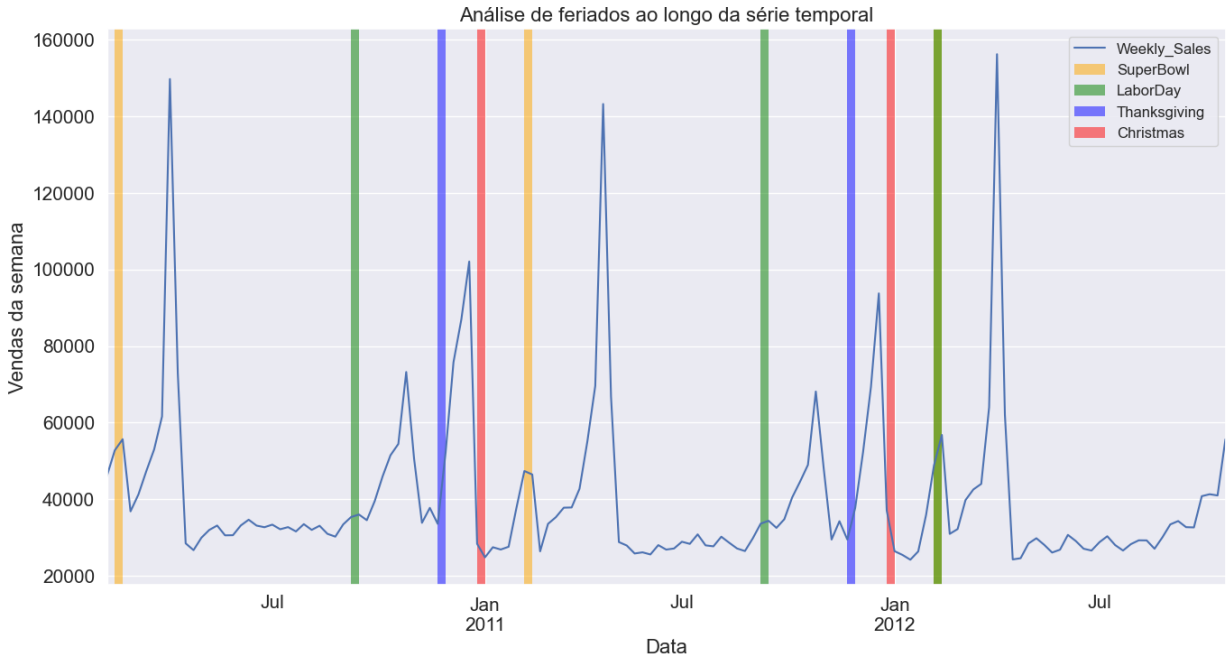
Figura 14: Série temporal multivariada com feriados destacados



Fonte: Autor.

Para melhor exploração dos dados, foram discriminados quais feriados são respectivos a quais datas. A partir dessa análise, foi possível distinguir o efeito de cada um nas vendas semanais. O resultado da análise se encontra na figura a seguir.

Figura 15: Série temporal multivariada com feriados destacados e discriminados.



Fonte: Autor.

5 APLICAÇÃO EM UMA SÉRIE TEMPORAL UNIVARIADA

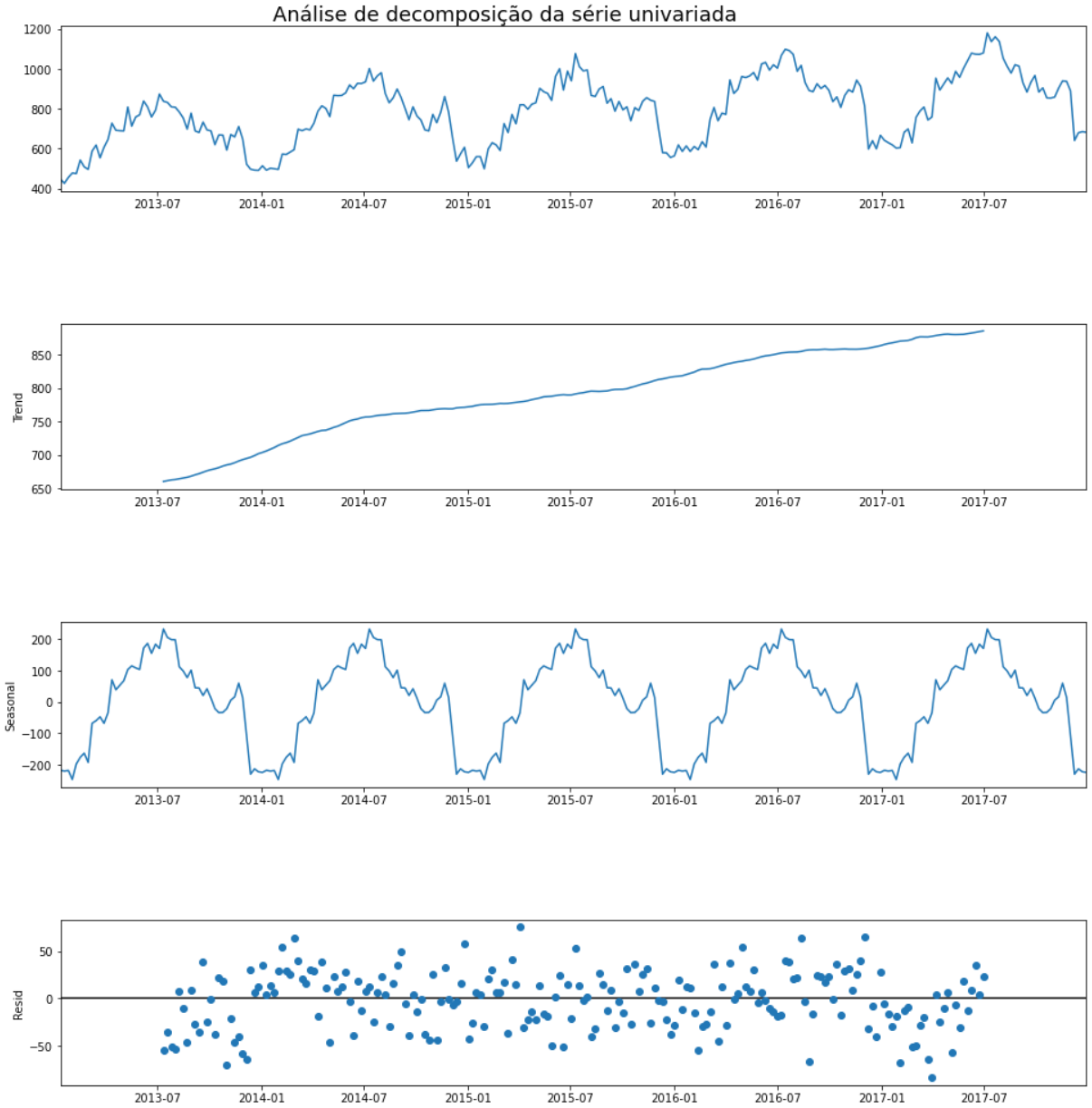
Para fins ilustrativos e de contraste com a base de dados efetivamente utilizada para o estudo, foi selecionada uma série temporal da base de dados univariada. Isto para demonstrar a aplicação dos métodos selecionados em uma situação de dados que apresentam regularidade na sazonalidade e na tendência. A aplicação dos diferentes métodos selecionados para estudo se dá a seguir.

5.1 Suavização exponencial

5.1.1 Análise de decomposição

Para realizar a modelagem da suavização exponencial com sazonalidade e tendência (Modelo Holt-Winters), primeiramente realizamos a decomposição da série temporal em suas componentes relevantes para o modelo. Essas são a tendência (trend), sazonalidade (seasonal) e o erro (resid). Uma representação gráfica destes componentes, calculada pela modelagem aditiva, é apresentada na figura 16.

Figura 16: Análise de decomposição da série univariada.



Fonte: Autor.

5.1.2 Aplicação dos métodos

Para modelar a série temporal, foram aplicados tanto o método aditivo quanto o multiplicativo. Desta forma, aferiu-se qual dos dois resultados são melhores para a previsão, e selecionou-se o melhor modelo para posterior comparação com outros métodos. Os resultados são apresentados a seguir.

Tabela 5: Métricas de avaliação do modelo Holt-Winters na série univariada.

Modelo	MAPE
Add_Holt-Winters	4,4%
Mul_Holt-Winters	4,3%

Fonte: Autor.

Figura 17: Previsão do método Holt-Winters comparado à base de teste da série univariada.

Fonte: Autor.

Podemos observar que ambas aplicações se adaptam bem tanto à sazonalidade quanto à tendência da série. Porém, vemos que a aplicação multiplicativa em sua maioria superestima a previsão. Isso é possível de aferir também visualmente na curva, onde a aplicação aditiva se mostra mais semelhante à real, principalmente na questão dos picos de demanda onde a aplicação aditiva se mostra mais conservadora na previsão.

5.2 ARIMA

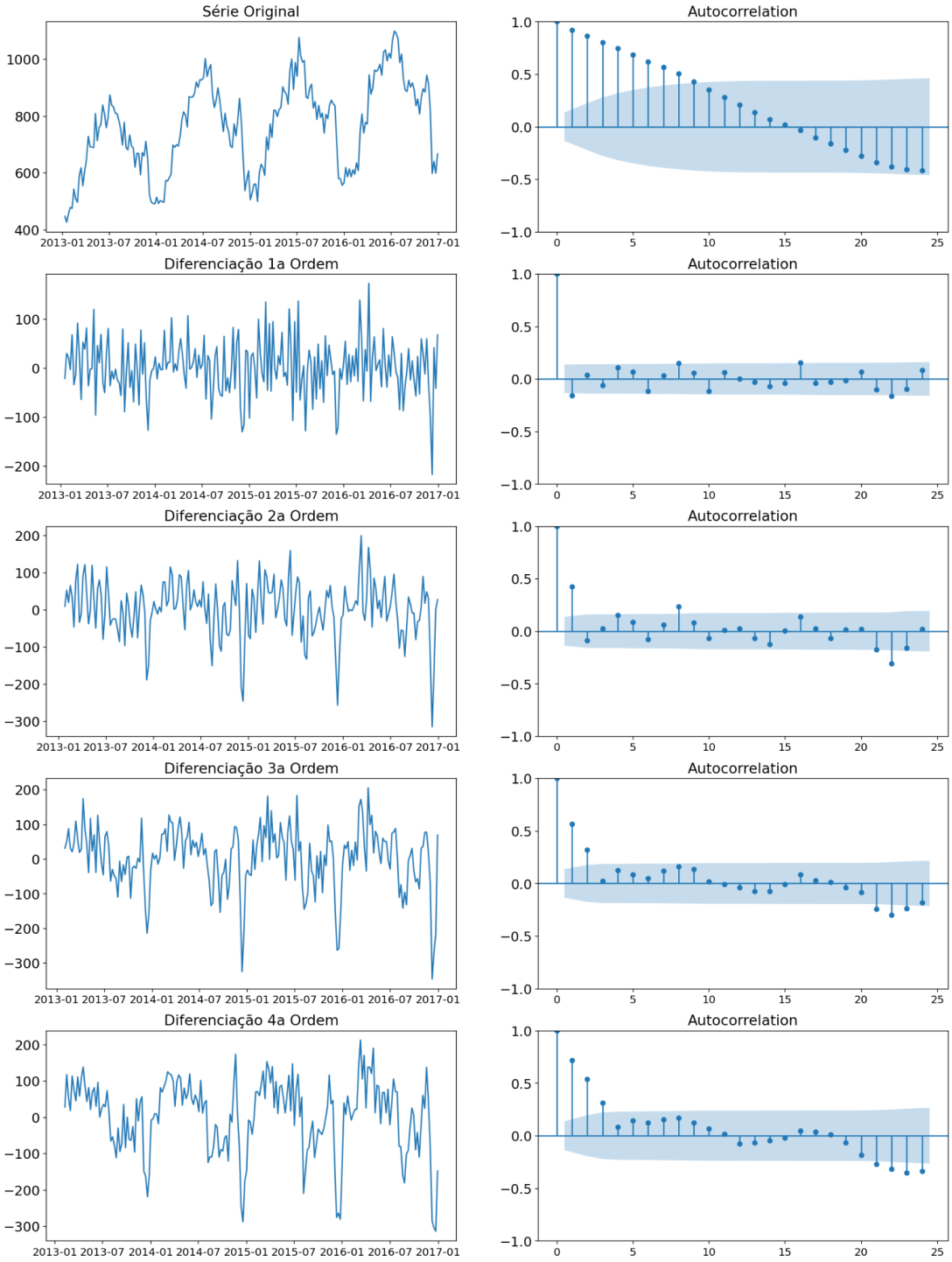
Para definir os parâmetros do ARIMA, precisamos passar por algumas análises respectivas a cada componente do modelo: a integração (diferenciação), a auto-regressão e a

média móvel. Cada uma dessas análises obtém respectivamente os parâmetros d, p e q do modelo ARIMA. Cada análise é explorada nas seções abaixo.

5.2.1 Análise de diferenciação

É necessário antes de tudo verificar se nossa série temporal é estacionária, e se não aplicar uma transformação para que seja. No caso, aplicamos a diferenciação dos valores em t com os valores em t-k, onde k é definido à partir do menor valor pelo qual conseguimos alcançar a estacionariedade. A seguir, temos a análise de autocorrelação da série original, assim como das suas versões diferenciadas de 1a até 4a ordem.

Figura 18: Análise de diferenciação e autocorrelação da série univariada.



Fonte: Autor.

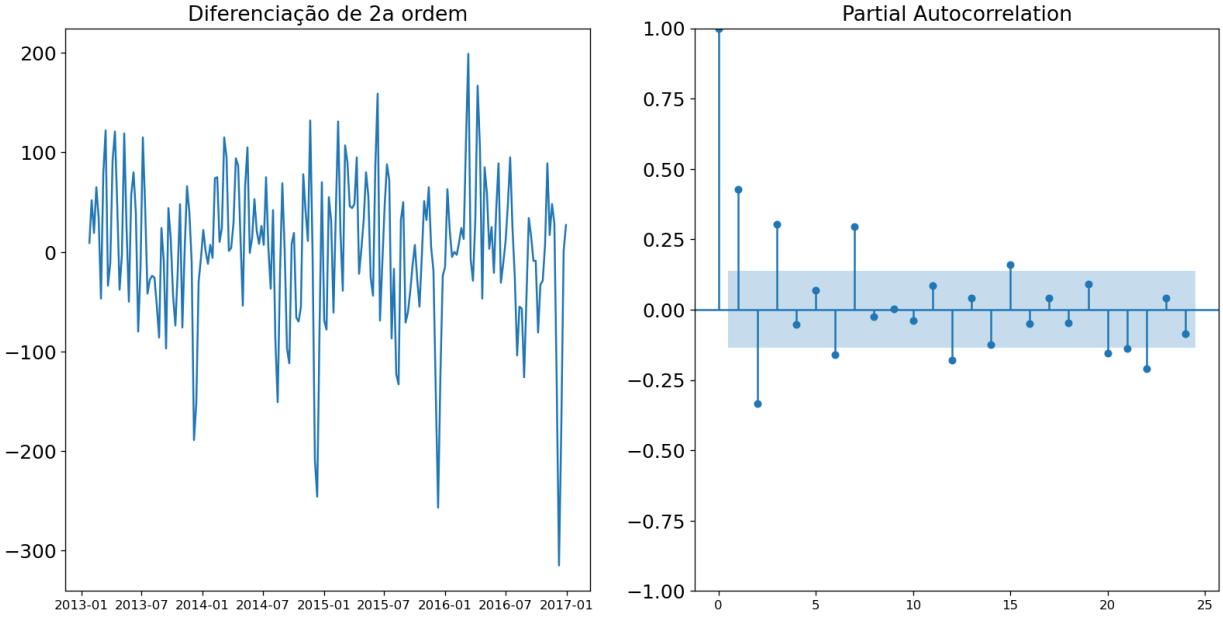
A partir da figura, podemos constatar que já na primeira diferenciação a série se aproxima estatisticamente de uma série estacionária. Porém, o efeito da diferenciação de primeira ordem parece eliminar também o efeito da autocorrelação no horizonte próximo. Por esse motivo, foi selecionada a diferenciação de segunda ordem, cujo resultado no Augmented

Dickey Fuller Test para estacionariedade alcançou um p-value de 0.041016, valor considerado estatisticamente significativo para aprovar a hipótese de que a série obtida é estacionária. Portanto, o valor escolhido para o parâmetro de diferenciação é d = 2.

5.2.2 Análise de auto-regressão

À seguir, precisamos considerar a autocorrelação parcial dos valores da série temporal. A figura a seguir apresenta a autocorrelação parcial da série temporal diferenciada em diferentes valores de lag.

Figura 19: Análise de auto-regressão e autocorrelação parcial da série univariada.



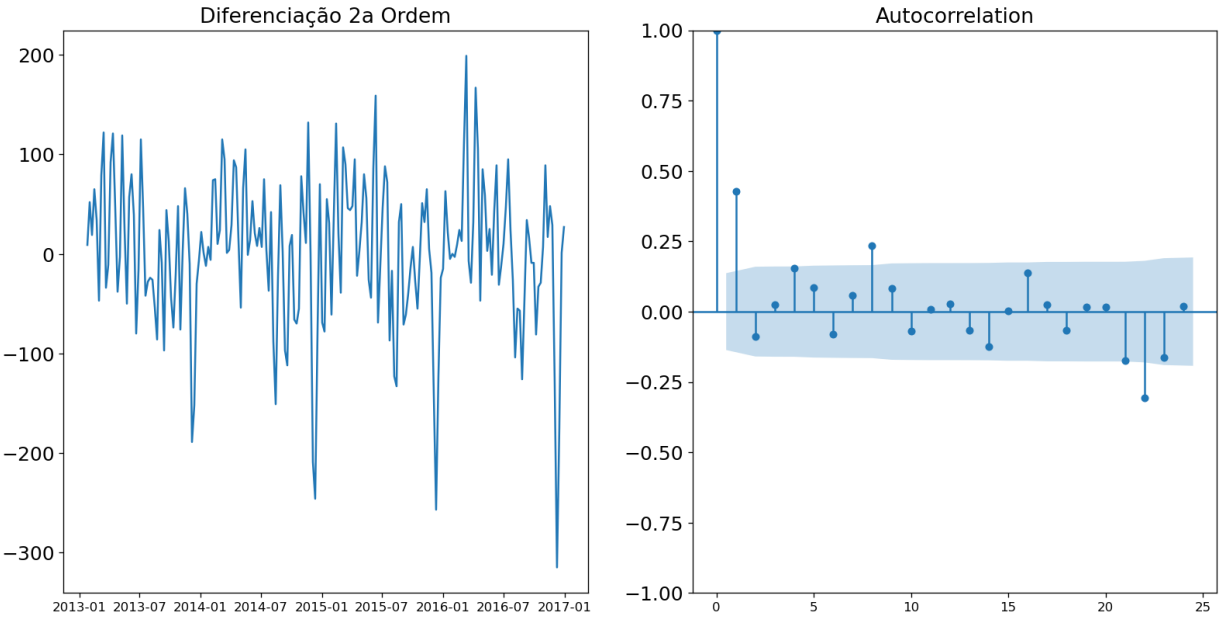
Fonte: Autor.

Podemos aferir que a autocorrelação parcial é estatisticamente significativa principalmente nos valores de p sendo 1, 2, 3 e 7. Porém, como temos dois pontos estatisticamente insignificantes e um ponto de pouca significância (os pontos 4, 5 e 6 respectivamente), e para reduzir a complexidade do modelo evitando o overfitting, foi decidido se bastar em escolher o valor do parâmetro de auto-regressão como p = 3.

5.2.3 Análise da média móvel

Por fim, voltamos ao gráfico de autocorrelação para decidir o alcance da média móvel. A seguir, temos o gráfico de autocorrelação da série temporal já diferenciada pelo parâmetro escolhido d = 2.

Figura 20: Análise de média móvel e autocorrelação da série univariada.



Fonte: Autor.

Vemos que a maior autocorrelação no horizonte próximo é com o lag de valor 1. Ou seja, o dia anterior é considerado o melhor preditor para o dia seguinte, apesar de parcialmente considerarmos também três dias antes (de acordo com o parâmetro p, escolhido na seção anterior). Portanto, decidiu-se por escolher como valor do parâmetro de alcance da média móvel como q = 1.

5.2.4 Aplicação do modelo ARIMA

A partir dos parâmetros definidos, foi aplicado o treino e teste do modelo ARIMA na base univariada. Porém, a predição não conseguiu se adaptar à sazonalidade anual presente na série temporal, como pode-se aferir na figura a seguir. O MAPE obtido foi de 11,0%, e a razão se mostra clara ao se desenhar a previsão contra a série temporal original. O ARIMA simples teve efeito aparente parecido com um simples deslocamento dos dados de t para t+1, e para que se melhore a performance se faz necessário modelar também a sazonalidade dos dados.

Figura 21: Previsão do método ARIMA comparado à base de teste da série univariada.

Fonte: Autor.

5.2.5 Aplicação do Seasonal ARIMA

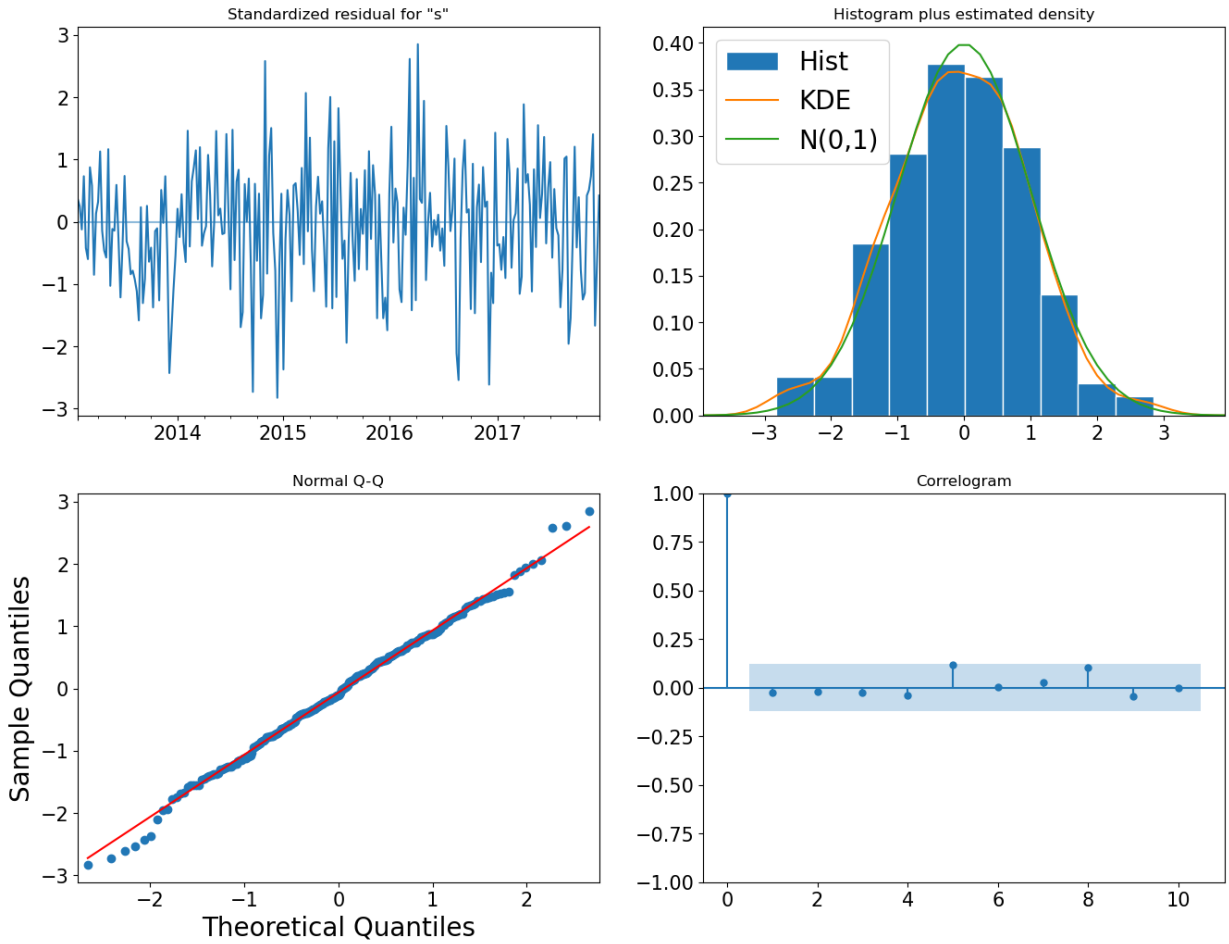
Por conta da sazonalidade presente nos dados previstos, é necessário uma adaptação do método ARIMA original para lidar com este fator. Para isso, existe o Seasonal ARIMA, ou SARIMA. Para encontrar os parâmetros P, D e Q do Seasonal ARIMA, foi aplicado um Grid Search para encontrar os parâmetros ótimos dado o modelo ARIMA encontrado anteriormente.

Como resultado, obteve-se os parâmetros P = 2, D = 0 e Q = 0, assim como a frequência de sazonalidade sendo 52 (dado que os dados são agrupados por semana e a sazonalidade é anual, temos 52 itens que compõem um ciclo).

5.2.6 Análise residual

Como fruto do modelo final, considerando o ARIMA original mais a sazonalidade captada pelo SARIMA, o resultado na base de dados de teste teve o seguinte diagnóstico residual.

Figura 22: Análise residual do SARIMA aplicado na série univariada.



Fonte: Autor.

Podemos ver que o resíduo se aproxima de uma distribuição normal de probabilidade, e que se encaixa de forma linearmente distribuída no Q-Q plot. Ademais, não há autocorrelação

estatisticamente significante no Correlograma do canto inferior direito, o que indica que a modelagem conseguiu captar bem as autocorrelações presentes na série temporal.

5.2.7 Análise dos resultados

A partir da modelagem da sazonalidade para a aplicação do ARIMA, foram obtidos os seguintes resultados.

Tabela 6: Métricas de avaliação dos modelos ARIMA e SARIMA na série univariada.

Modelo	MAPE
ARIMA	11,0%
SARIMA	4,4%

Fonte: Autor.

Figura 23: Previsão dos métodos ARIMA e SARIMA comparado à base de teste da série univariada.

Fonte: Autor.

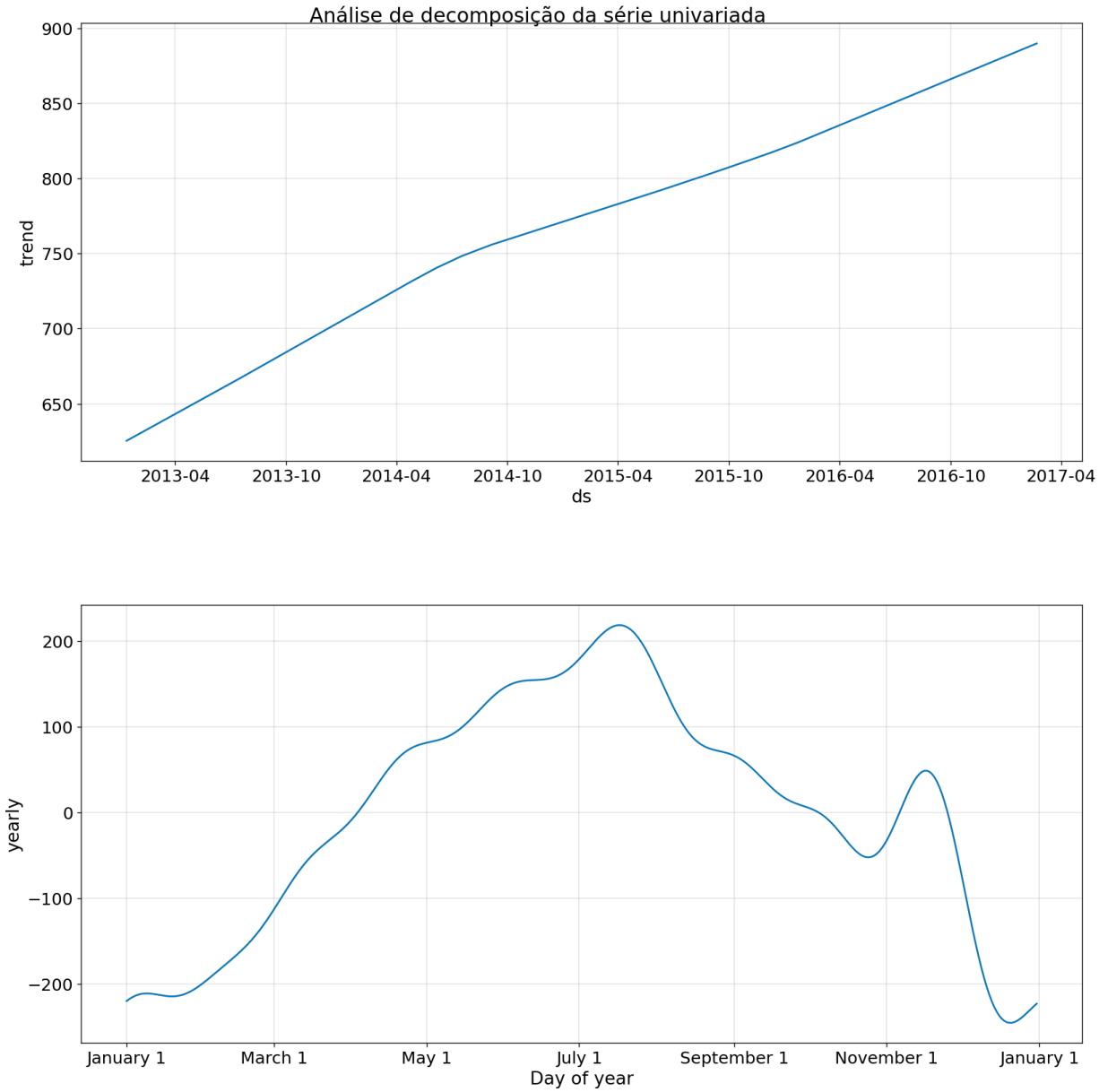
Podemos observar que agora a sazonalidade é adequadamente prevista pelo modelo e os resultados obtidos têm mais do que o dobro de precisão.

5.3 Prophet

5.3.1 Análise de decomposição

Para realizar suas análises, o Prophet produz diversos componentes de forma não-linear para analisar a série temporal. Temos a seguir esses componentes.

Figura 24: Análise de decomposição do Prophet na série univariada.



Fonte: Autor.

5.3.2 Aplicação do método

Os resultados da aplicação do Prophet são apresentados a seguir.

Tabela 7: Métricas de avaliação do modelo Prophet na série univariada.

Modelo	MAPE
Prophet	4,5%

Fonte: Autor.

Figura 25: Previsão do método Prophet comparado à base de teste na série univariada.

Fonte: Autor.

Apesar da simples aplicação comparado aos outros métodos, o Prophet apresenta um desempenho satisfatório, e pode-se perceber uma boa aproximação das vendas reais.

5.4 LSTM

5.4.1 Arquitetura do modelo

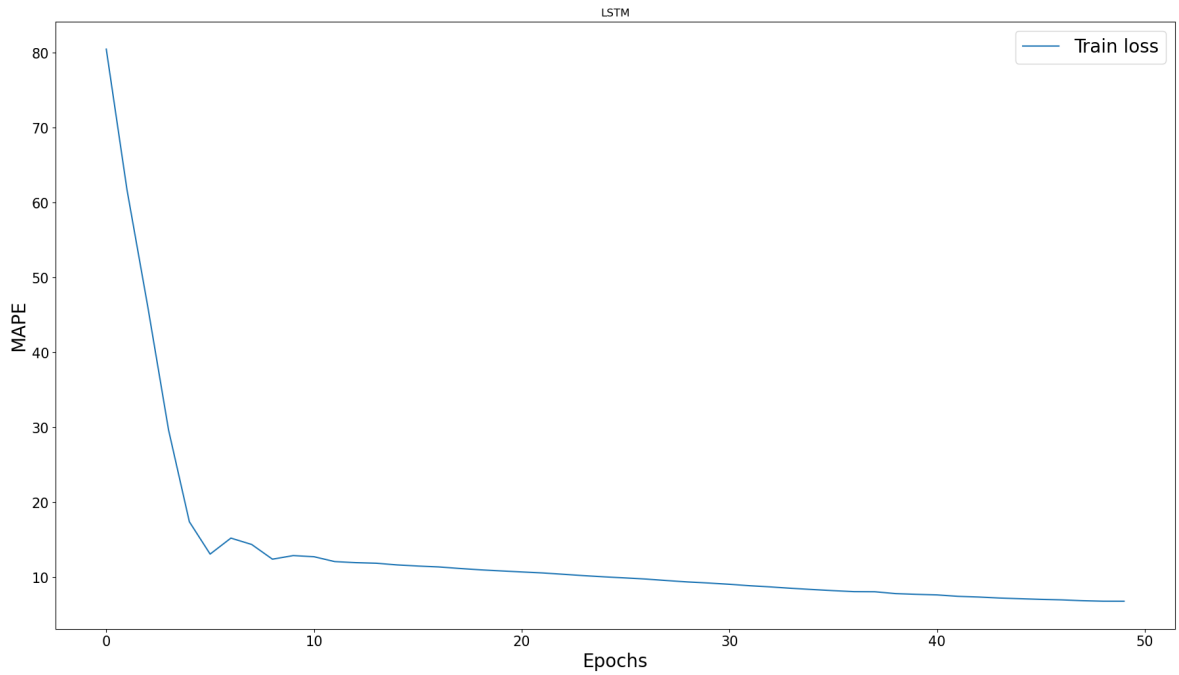
Para o estudo, foi montada uma rede neural com duas camadas. A primeira camada é composta por 50 células do tipo LSTM, conforme descritas na análise bibliográfica do capítulo 2. Portanto, estas células contêm no total 23.800 parâmetros (coeficientes) para treinar. A

função escolhida para estas células é a ReLu (rectified linear unit). A segunda camada é composta por uma camada chamada Dense layer, que condensa os resultados obtidos pela camada de células LSTM em um único output. Esta camada tem 51 parâmetros para treino, sendo 50 para cada saída da camada LSTM e 1 para o chamado *bias input*, que permite adicionar um valor constante para a camada evitando o overfitting.

5.4.2 Treino do modelo

Escolhendo o MAPE como métrica de otimização, foi realizado o treino do modelo com 50 iterações (chamadas epochs), com uma taxa de aprendizado de 0,01 e utilizando o otimizador Adam. O resultado do experimento é mostrado na figura a seguir.

Figura 26: Curva de aprendizado do LSTM para a série univariada.



Fonte: Autor.

5.4.3 Avaliação na base de teste

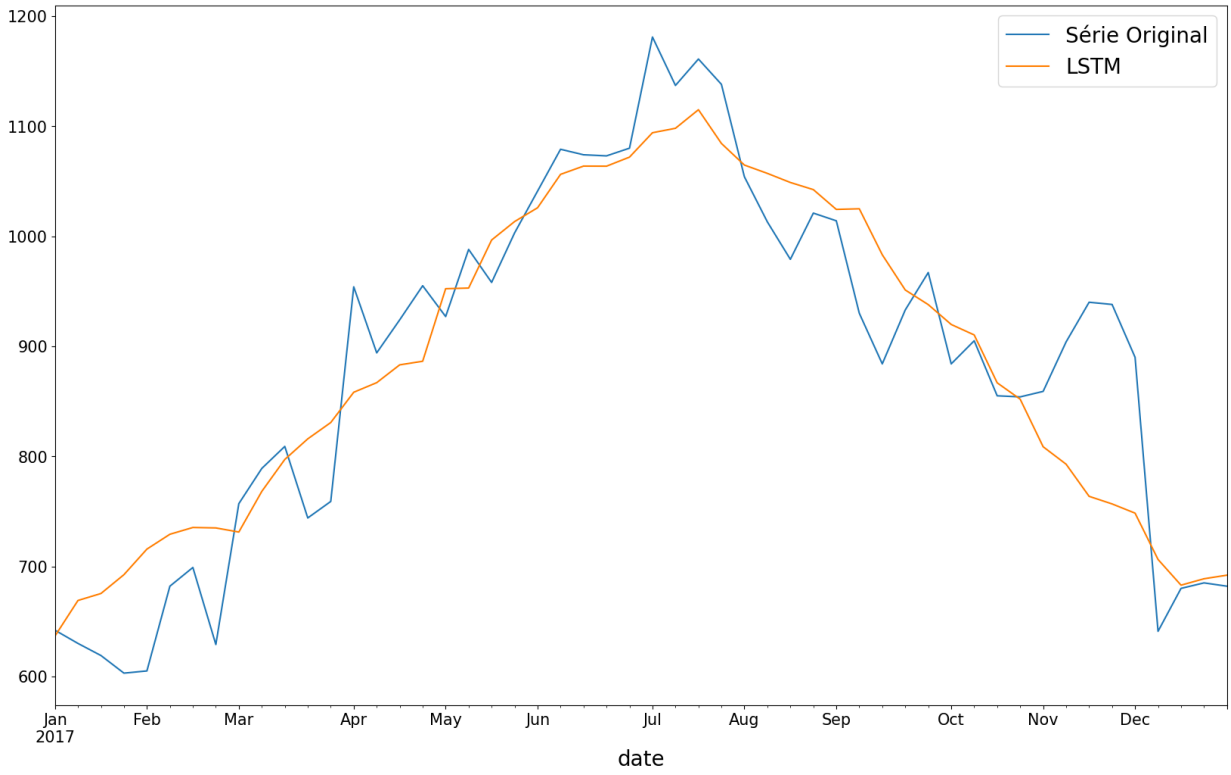
Os resultados da comparação do modelo treinado na base de testes é apresentado nas tabela e figura à seguir.

Tabela 8: Métricas de avaliação do modelo LSTM na série univariada.

Modelo	MAPE
LSTM	5,8%

Fonte: Autor.

Figura 27: Previsão do método LSTM comparado à base de teste da série univariada.



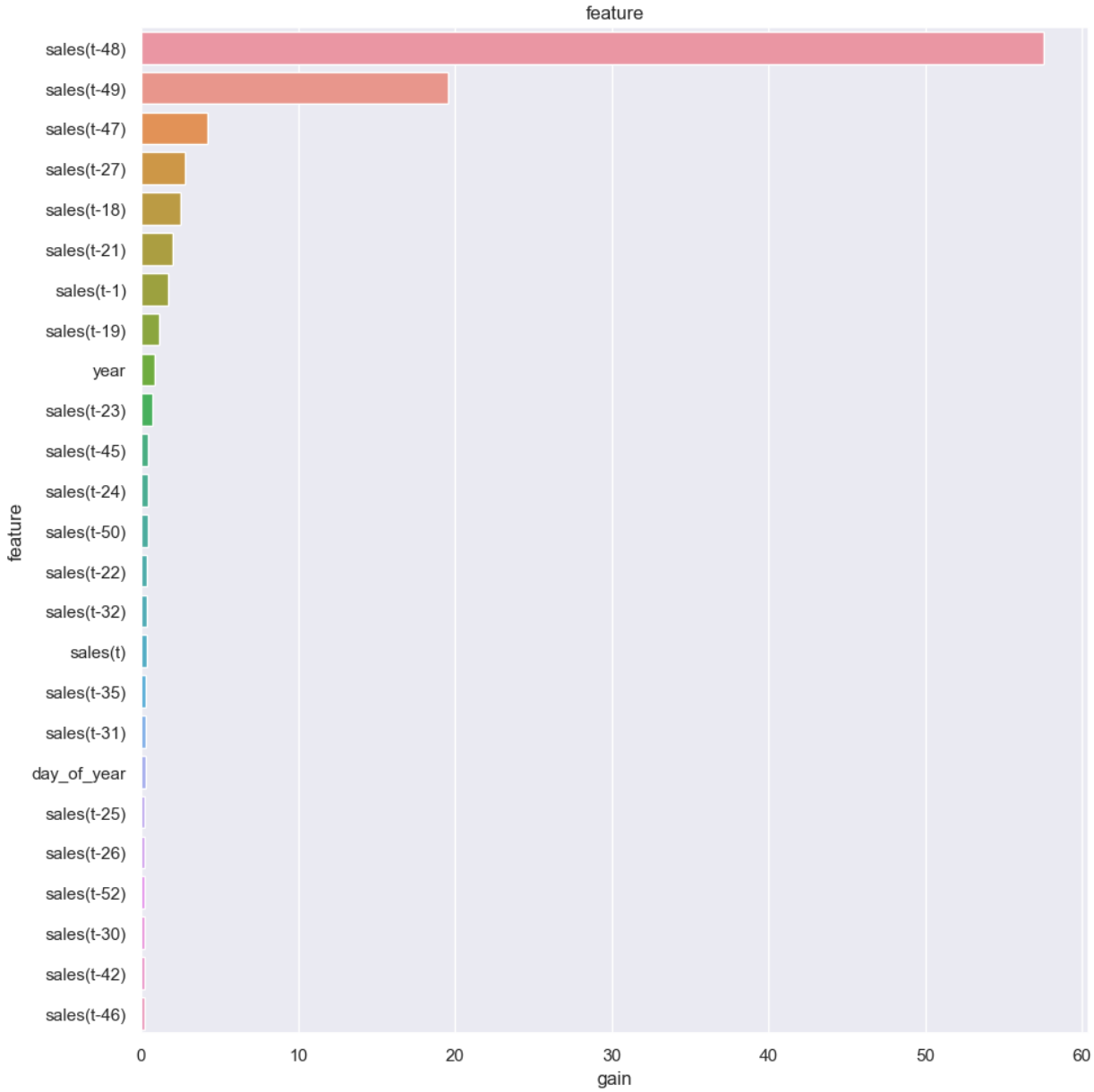
Fonte: Autor.

5.5 LightGBM

5.5.1 Análise de feature importance

Uma vantagem de algoritmos de árvore em machine learning é a possibilidade de avaliar a influência de cada variável preditora, também chamada de *feature*, do modelo treinado para a predição. A partir da análise feita no modelo treinado na série univariada, obtiveram-se os resultados na figura seguinte.

Figura 28: Análise de feature importance do LightGBM na série univariada.



Fonte: Autor.

Pode-se observar que as variáveis mais valiosas para o modelo são as vendas de 48 e 49 dias antes do momento presente, em ordem decrescente de importância. Interessante ver que de todas as variáveis temporais, a única destacada como importante é o ano da data presente.

5.5.2 Avaliação na base de teste

Os resultados da comparação do modelo treinado na base de testes é apresentado nas tabela e figura à seguir.

Tabela 9: Métricas de avaliação do modelo LSTM na série univariada.

Modelo	MAPE
LightGBM	4,7%

Fonte: Autor.

Figura 29: Previsão do método LightGBM comparado à base de teste da série univariada.

Fonte: Autor.

5.6 Análise dos resultados

Dado todos os algoritmos testados, temos o resultado da métrica escolhida, assim como o tempo de execução computacional de cada um dos algoritmos escolhidos para o experimento, na figura a seguir.

Tabela 10: Resultados obtidos na base univariada.

Modelo	MAPE	Tempo de execução (s)
Mul_Holt-Winters	4,25%	91,9
SARIMA	4,39%	1082
Add_Holt-Winters	4,45%	14,1
Prophet	4,52%	72
LightGBM	4,67%	0,1
LSTM	5,76%	1,9
ARIMA	11,02%	32,1

Fonte: Autor.

Como mencionado anteriormente, a escolha de testar os algoritmos selecionados para estudo na base univariada foi ilustrar o processo de previsão de demanda em uma série temporal estável, com clara tendência e sazonalidade que se comportam de forma regular. Isto é feito para mostrar o contraste com a segunda base de dados selecionada, a multivariada, que contém séries temporais com comportamentos diversos e mais próximos de uma situação real de previsão onde os dados presentes não apresentam garantia de estabilidade.

Além disso, podemos constatar que em séries temporais estáveis como a série univariada selecionada, métodos estatísticos tradicionais tem uma performance mais do que satisfatória ao prever a demanda. Isso se justifica por diversos motivos, mas principalmente pela constatação de que a base univariada tem uma tendência e sazonalidade claras e facilmente modeláveis. Ademais, a interpretabilidade desses métodos ajuda na adesão por parte de uma organização, facilitando a comunicação e explicação dos motivos pelos quais a previsão se dá.

Portanto, os resultados do experimento ilustram bem como existem casos onde a preferência pelos métodos tradicionais é justificada. Por mais que existam resultados promissores utilizando algoritmos mais recentes de aprendizado de máquina, o ganho em assertividade deles não chega a ser grande o bastante para contrapor à maior complexidade ou menor quantidade de validações acadêmicas para seu uso.

6 APLICAÇÃO EM UMA SÉRIE TEMPORAL MULTIVARIADA

Este capítulo visa ilustrar o procedimento de previsão aplicado na base efetivamente selecionada para comparação dos métodos no capítulo 7. Isto para poder demonstrar no nível detalhado a aplicação dos métodos selecionados para estudo em uma série mais próxima de uma situação real de vendas, sem garantias de regularidades na sazonalidade e na tendência. O mesmo procedimento aplicado neste capítulo é o mesmo que será repetido em cada uma das 100 séries temporais selecionadas no capítulo 7.

6.1 Suavização exponencial

6.1.1 Análise de decomposição

Analogamente à aplicação na base univariada, temos que decompor a série temporal escolhida considerando a tendência (trend), sazonalidade (seasonal) e o erro (resid). Uma representação gráfica destes componentes é apresentada na figura XX, referenciando a modelagem multiplicativa do método.

Figura 30: Análise de decomposição da série multivariada.

Fonte: Autor.

Podemos ver que, diferente da primeira aplicação do método, esta se comporta de forma mais irregular. A tendência é de queda, mas não é trivial distinguir pontos de inflexão como o de 10/2010 ou o de 07/2011. Além disso, a sazonalidade apresenta uma ciclicidade menos regular que a série univariada. Por fim, o resíduo contém pontos de maior variabilidade como podemos observar em 04/2010 e 04/2011.

6.1.2 Aplicação dos métodos

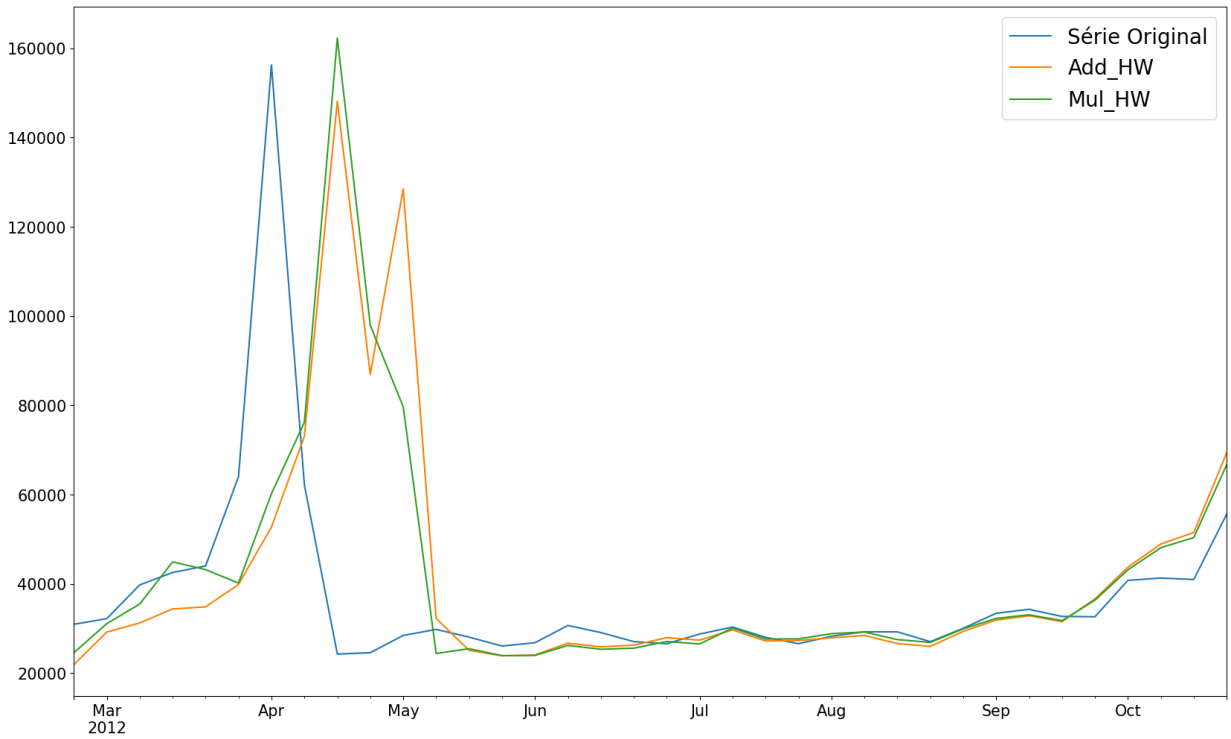
Os resultados da aplicação tanto do método aditivo quanto do multiplicativo são apresentados a seguir.

Tabela 11: Métricas de avaliação do modelo Holt-Winters na série multivariada.

Modelo	MAPE
Add_Holt-Winters	42,8%
Mul_Holt-Winters	38,9%

Fonte: Autor.

Figura 31: Previsão do método Holt-Winters comparado à base de teste da série multivariada.



Fonte: Autor.

É possível constatar que ambos modelos falham em modelar de forma precisa o deslocamento do pico, que como observamos no capítulo 4 ocorre em um mês diferente a cada ano.

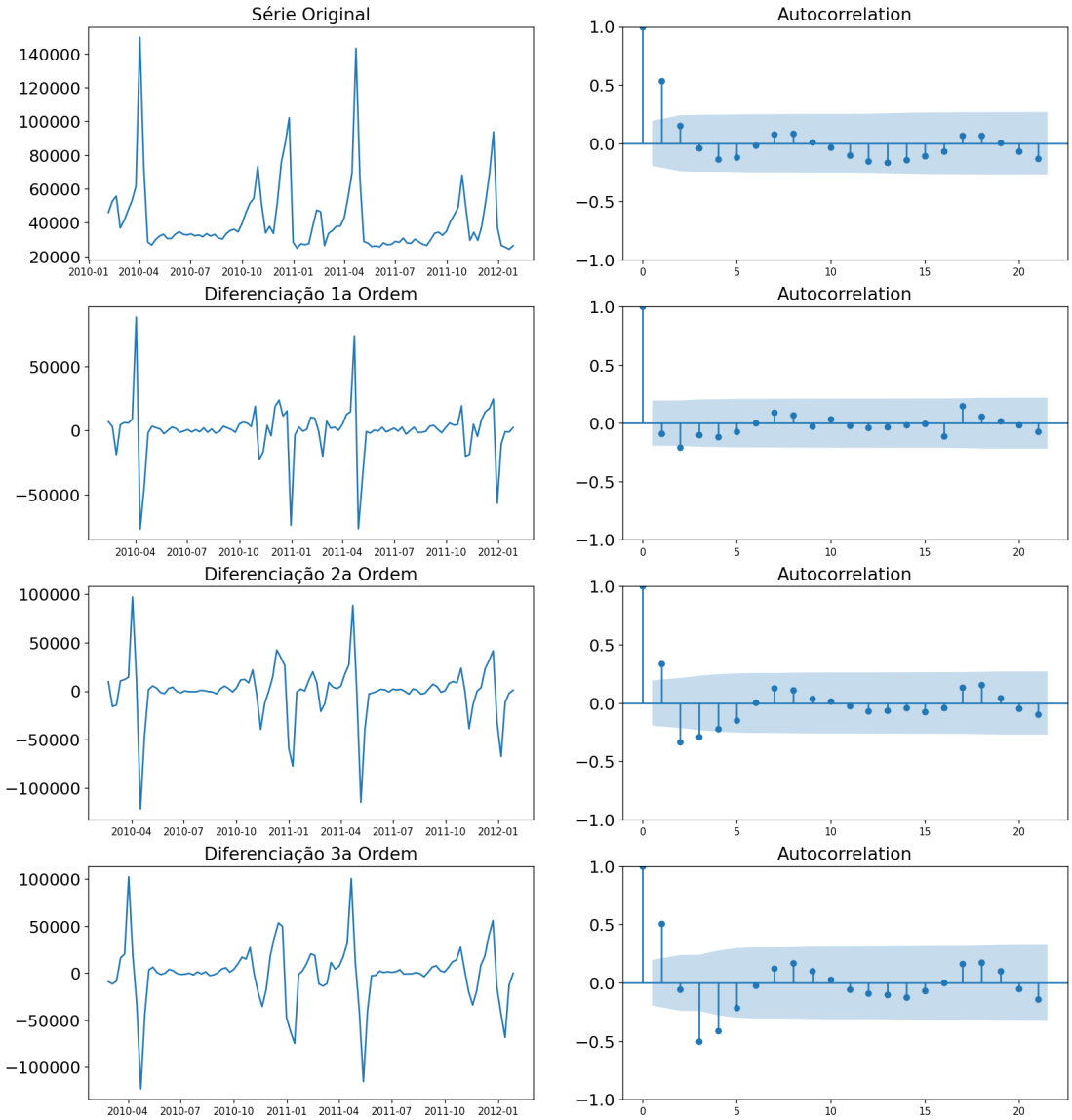
6.2 ARIMA

Precisamos definir os parâmetros do modelo passando por algumas análises respectivas a cada componente do modelo: a integração (diferenciação), a auto-regressão e a média móvel. Cada uma dessas análises obtém respectivamente os parâmetros d, p e q do modelo ARIMA. Cada análise é explorada nas seções abaixo.

6.2.1 Análise de diferenciação

Para verificar se a série temporal em análise é estacionária, e se não qual transformação é necessária que se aplique para que seja, temos a análise de autocorrelação da série original, assim como das suas versões diferenciadas de 1a até 3a ordem.

Figura 32: Análise de diferenciação e autocorrelação da série univariada.



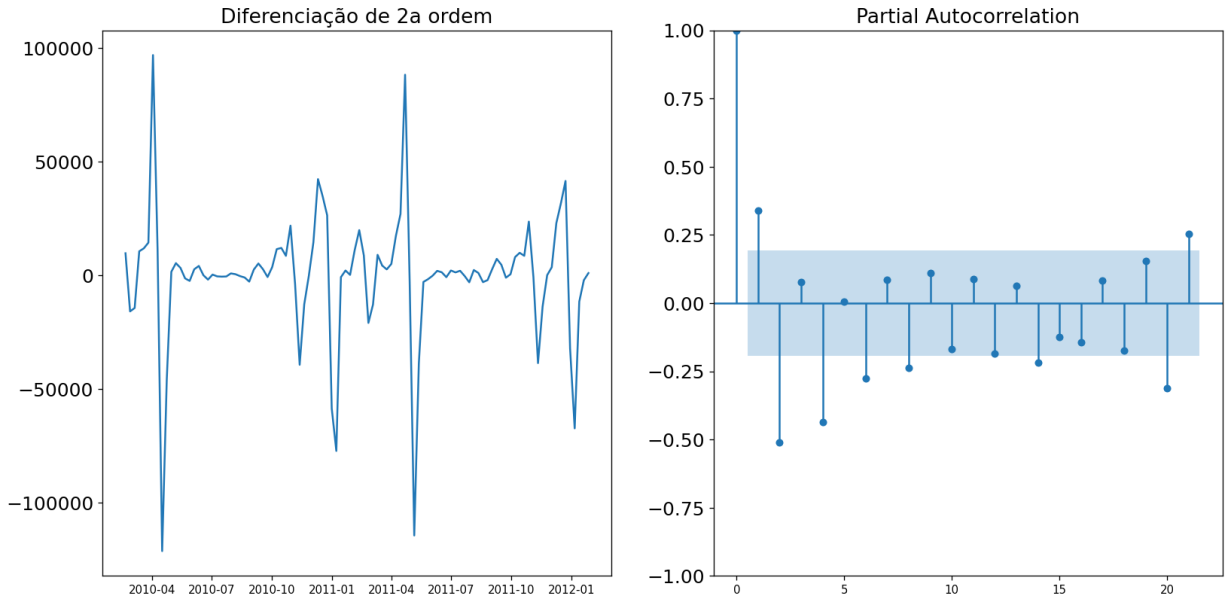
Fonte: Autor.

A partir da figura, podemos constatar que já na primeira diferenciação a série se aproxima estatisticamente de uma série estacionária. Porém, assim como no caso anterior de aplicação do ARIMA, o efeito da diferenciação de primeira ordem parece eliminar também o efeito da autocorrelação no horizonte próximo. Por esse motivo, foi selecionada a diferenciação de segunda ordem, cujo resultado no Augmented Dickey Fuller Test para estacionariedade alcançou um p-value de 0.000001, valor considerado estatisticamente significativo para aprovar a hipótese de que a série obtida é estacionária. Portanto, o valor escolhido para o parâmetro de diferenciação é d = 2.

6.2.2 Análise de auto-regressão

À seguir, precisamos considerar a autocorrelação parcial dos valores da série temporal. A figura a seguir apresenta a autocorrelação parcial da série temporal diferenciada em diferentes valores de lag.

Figura 33: Análise de auto-regressão e autocorrelação parcial da série multivariada.



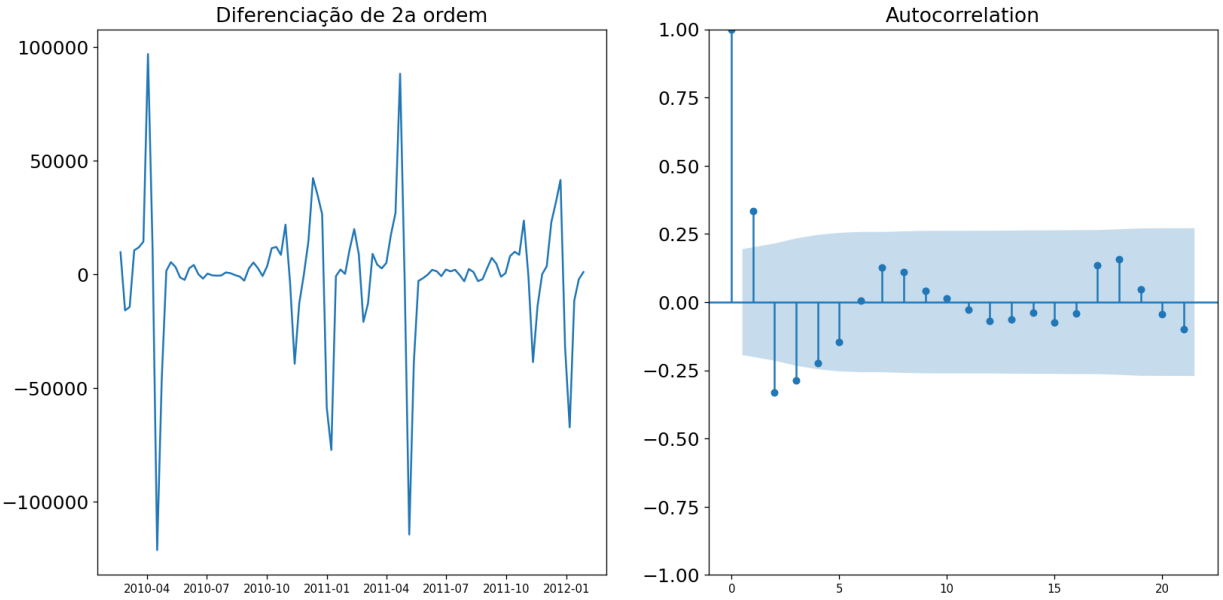
Fonte: Autor.

Podemos aferir que a autocorrelação parcial é estatisticamente significativa principalmente nos valores de p sendo 1 e 2. Porém, temos uma alternância à partir do ponto 3 tendo pontos estatisticamente insignificantes e pontos estatisticamente significativos com uma tendência de queda, até alcançarem insignificância no ponto 10 e voltando a obter valores de significância à partir daí. Para reduzir a complexidade do modelo evitando o overfitting, foi decidido se bastar em escolher o valor do parâmetro de auto-regressão como p = 2.

6.2.3 Análise da média móvel

Por fim, voltamos ao gráfico de autocorrelação para decidir o alcance da média móvel. A seguir, temos o gráfico de autocorrelação da série temporal já diferenciada pelo parâmetro escolhido d = 2.

Figura 34: Análise de média móvel e autocorrelação parcial da série multivariada.



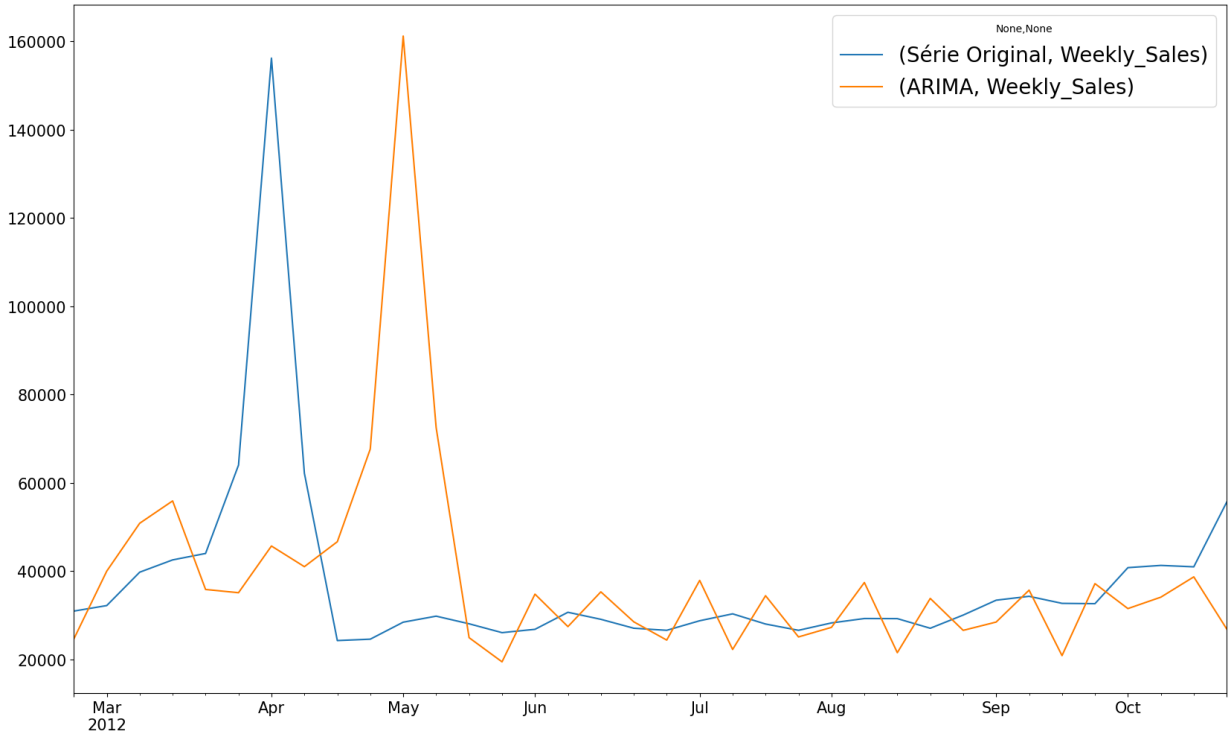
Fonte: Autor.

Vemos que até o ponto de lag igual a 3 ainda temos significância estatística para a autocorrelação. Portanto, decidiu-se por escolher como valor do parâmetro de alcance da média móvel como q = 3.

6.2.4 Aplicação do modelo ARIMA

A partir dos parâmetros definidos, foi aplicado o treino e teste do modelo ARIMA na base multivariada, apenas para a variável de vendas já que em sua forma pura o ARIMA é um modelo univariado. Porém, a predição não conseguiu se adaptar à sazonalidade presente na série temporal, como pode-se aferir na figura a seguir. O MAPE obtido foi de 44,8%, e a razão se mostra clara ao se desenhar a previsão contra a série temporal original.

Figura 35: Previsão do método ARIMA comparado à base de teste da série multivariada.



Fonte: Autor.

6.2.5 Aplicação do Seasonal ARIMA

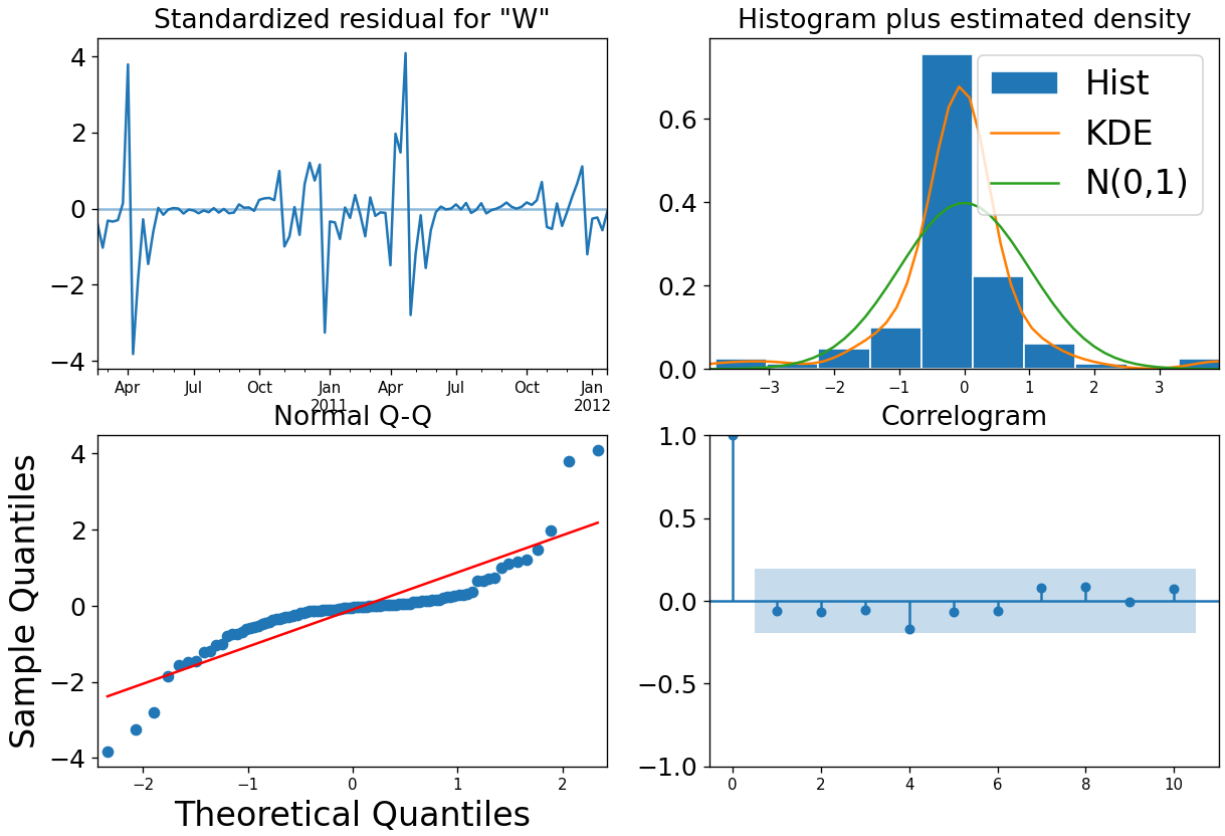
Por conta da sazonalidade presente nos dados previstos também na série multivariada, se faz necessário uma adaptação do método ARIMA original para lidar com este fator. Para encontrar os parâmetros P, D e Q do Seasonal ARIMA, foi aplicado um Grid Search para encontrar os parâmetros ótimos dado o modelo ARIMA encontrado anteriormente.

Como resultado, obteve-se os parâmetros P = 1, D = 0 e Q = 0, assim como a frequência de sazonalidade sendo 52 (dado que os dados são agrupados por semana e a sazonalidade é anual, temos 52 itens que compõem um ciclo).

6.2.6 Análise residual

Como fruto do modelo final, considerando o ARIMA original mais a sazonalidade captada pelo SARIMA, o resultado na base de dados de teste teve o seguinte diagnóstico residual.

Figura 36: Análise residual do SARIMA aplicado na série multivariada.



Fonte: Autor.

Podemos ver que o resíduo, apesar de se aproximar de uma curva normal de probabilidade, apresenta uma variância menor do que a esperada no modelo teórico, o que é reforçado pelo pouco encaixe no Q-Q plot. O correlograma, porém, constata que a autocorrelação foi em sua maioria captada pelo modelo, não se mostrando presente no resíduo.

6.2.7 Análise dos resultados

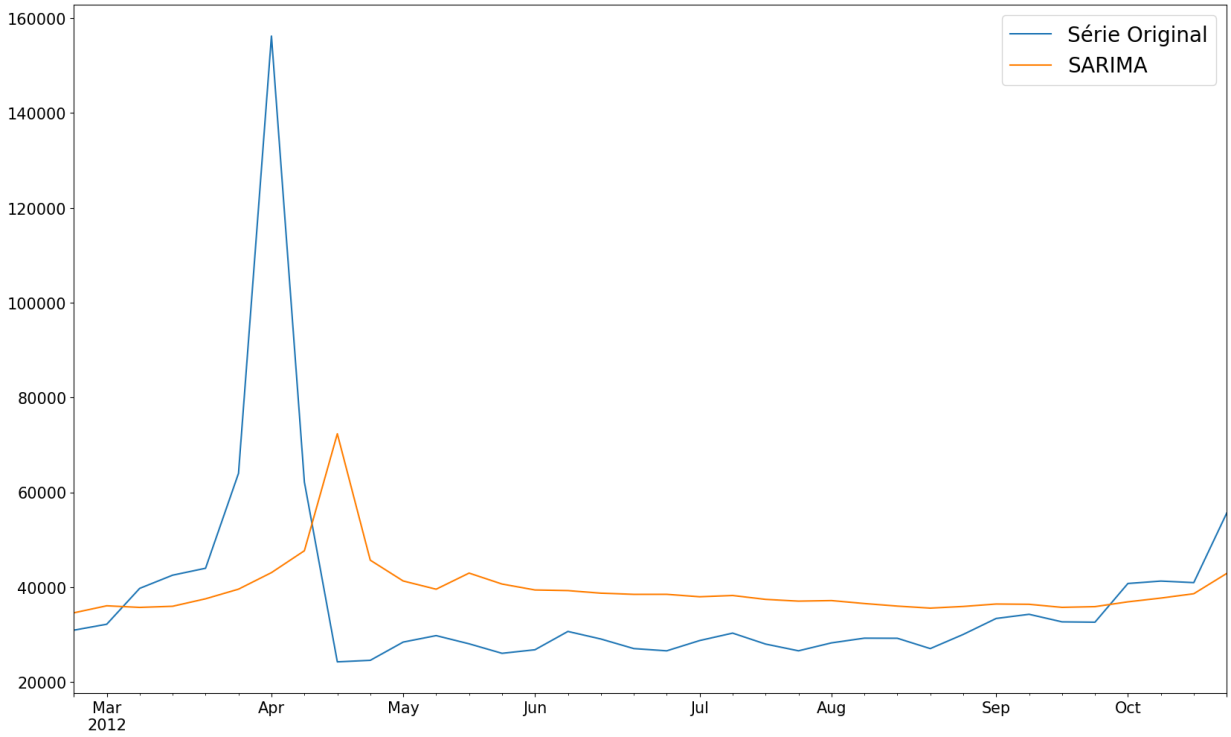
A partir da modelagem da sazonalidade para a aplicação do ARIMA, foram obtidos os seguintes resultados.

Tabela 12: Métricas de avaliação dos modelos ARIMA e SARIMA na série multivariada.

Modelo	MAPE
ARIMA	44,8%
SARIMA	33,6%

Fonte: Autor.

Figura 37: Previsão do método SARIMA comparado à base de teste da série multivariada.



Fonte: Autor.

Apesar de uma curva mais próxima da real, o SARIMA erra ao apontar que o pico de vendas seria no meio de Abril, quando na realidade ocorre no início do mês, assim como subestimar sua amplitude. Porém, por ser mais conservador quanto ao pico de vendas previsto, acaba com um resultado melhor do que o ARIMA tradicional.

6.3 Prophet

6.3.1 Análise de decomposição

Assim como na aplicação univariada, o Prophet produz diversos componentes de forma não-linear para analisar a série temporal. Como o Prophet é em sua maioria composto de análises univariadas na variável de resposta, não foi possível inserir as demais variáveis presentes no banco de dados. Porém, foram usados os feriados denotados pela variável *IsHoliday*, discriminados de acordo com o mencionado no capítulo 4.

Figura 38: Análise de decomposição do Prophet na série multivariada.

Fonte: Autor.

A modelagem do Prophet adicionou um fator negativo para a incidência dos feriados de natal, e uma menor incidência positiva nos demais. Além disso, consegue modelar com certa precisão os picos históricos da base de treino. Além disso, a tendência foi identificada de forma mais clara do que no método de Holt-Winters.

6.3.2 Aplicação do método

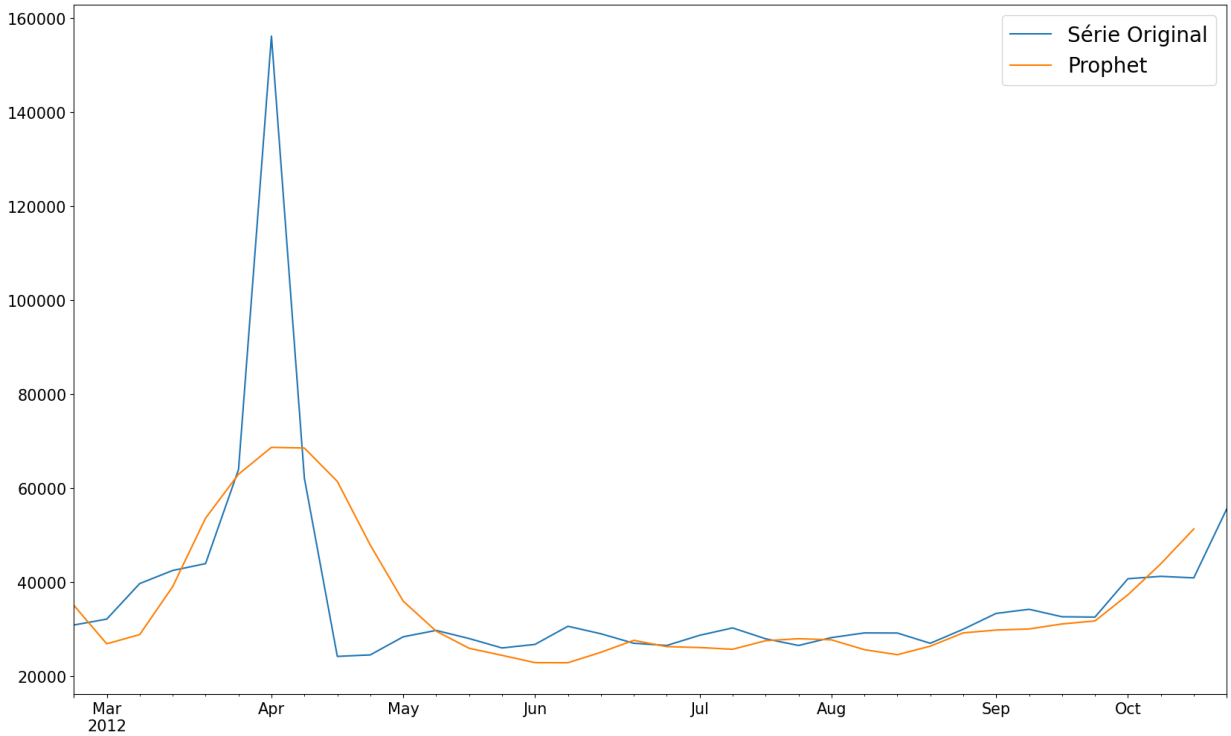
Os resultados da aplicação do Prophet são apresentados a seguir.

Tabela 13: Métricas de avaliação do modelo Prophet na base multivariada.

Modelo	MAPE
Prophet	23,5%

Fonte: Autor.

Figura 39: Previsão do método Prophet comparado à base de teste na série multivariada.



Fonte: Autor.

É possível constatar que, apesar de apresentar um MAPE melhor do que os demais modelos, o Prophet subestima sua previsão do pico ocorrido em Abril.

6.4 LSTM

6.4.1 Arquitetura do modelo

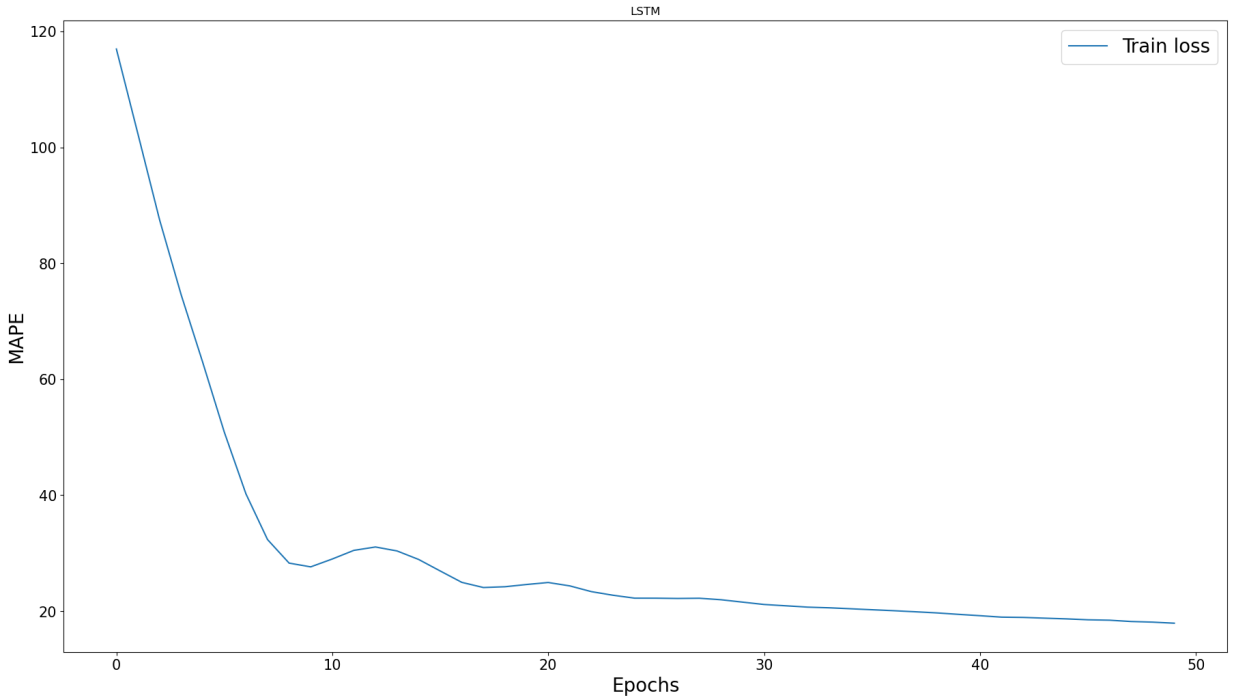
Assim como na base univariada, foi escolhida a mesma arquitetura de rede neural para o problema. A primeira camada é composta por 50 células do tipo LSTM. Portanto, estas células contêm no total 23.800 parâmetros (coeficientes) para treinar. A função escolhida para estas

células é a ReLu (rectified linear unit). A segunda camada é composta por uma camada chamada Dense layer, que condensa os resultados obtidos pela camada de células LSTM em um único output. Esta camada tem 51 parâmetros para treino, sendo 50 para cada saída da camada LSTM e 1 para o chamado *bias input*, que permite adicionar um valor constante para a camada evitando o overfitting.

6.4.2 Treino do modelo

Escolhendo o MAPE como métrica de otimização, foi realizado o treino do modelo com 50 iterações (chamadas epochs), com uma taxa de aprendizado de 0,01 e utilizando o otimizador Adam. O resultado do experimento é mostrado na figura a seguir.

Figura 40: Curva de aprendizado do LSTM para a série multivariada.



Fonte: Autor.

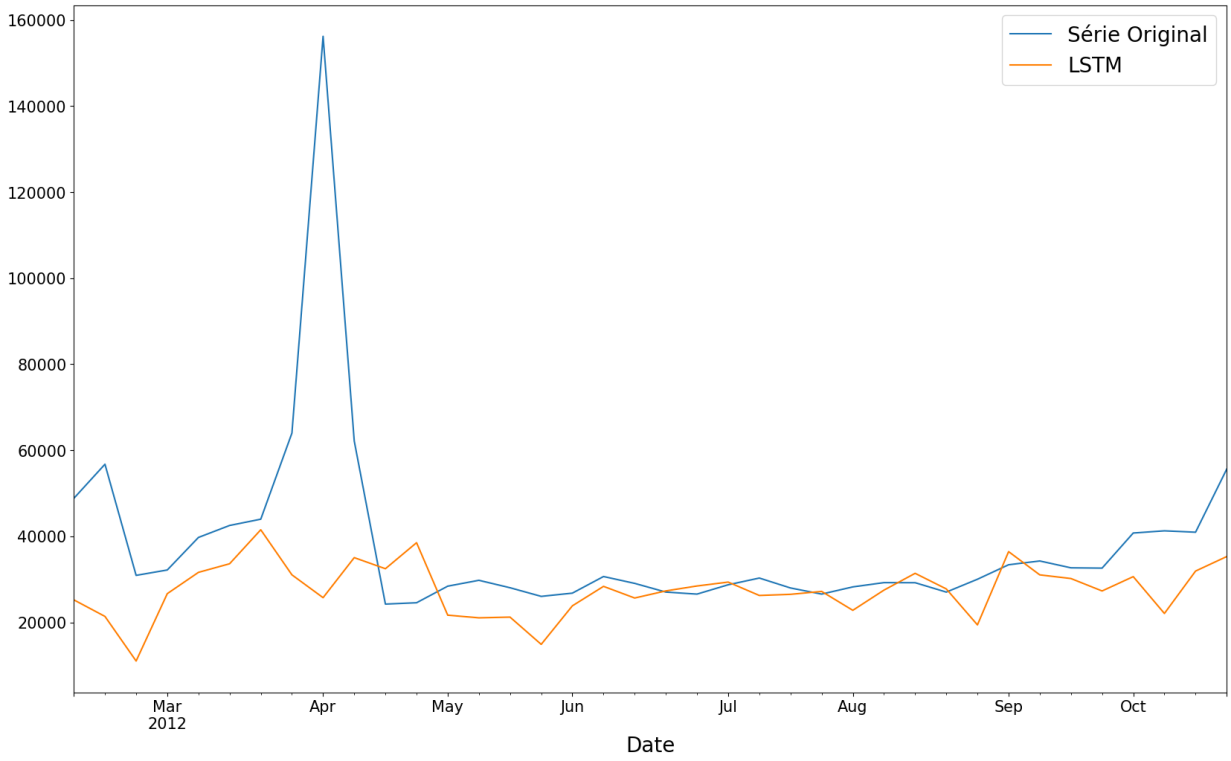
6.4.3 Avaliação na base de teste

Os resultados da comparação do modelo treinado na base de testes é apresentado nas tabela e figura à seguir.

Tabela 14: Métricas de avaliação do modelo LSTM na série multivariada. Fonte: Autor.

Modelo	MAPE
LSTM	24,6%

Figura 41: Previsão do método LSTM comparado à base de teste da série multivariada.



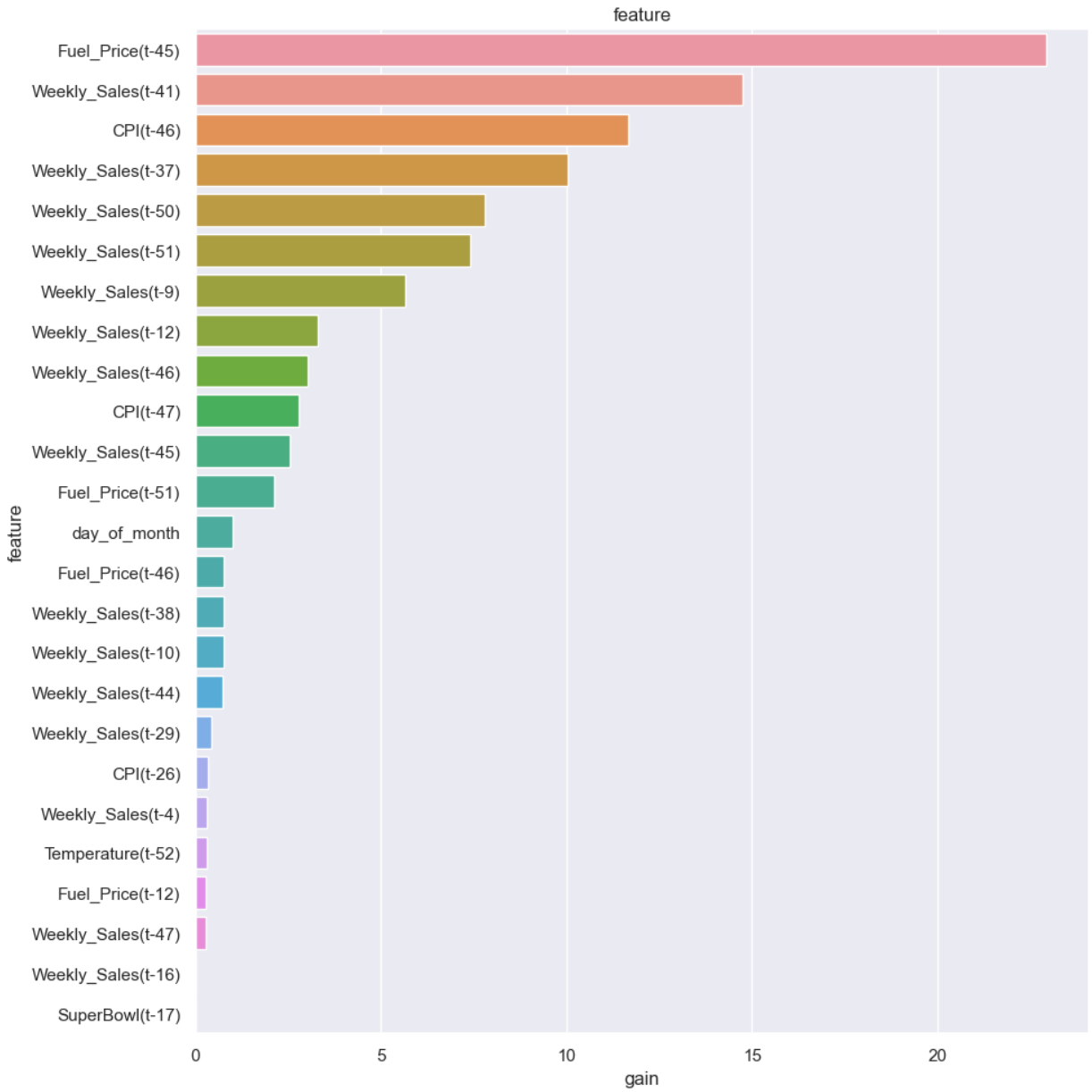
Fonte: Autor.

6.5 LightGBM

6.5.1 Análise de feature importance

Conforme já mencionado na aplicação do LightGBM na série univariada, podemos analisar as importâncias de cada variável preditora utilizada para o algoritmo, com a adição desta vez também da relevância das demais variáveis exógenas à série temporal de vendas. A partir da análise feita no modelo treinado na série univariada, obtiveram-se os resultados na figura seguinte.

Figura 42: Análise de feature importance do LightGBM na série multivariada.



Fonte: Autor.

É interessante constatar que variáveis como o CPI e Fuel Price são de grande importância para a predição, apesar de não constatarem correlação linear com a variável de vendas. Isso se explica pela natureza não-linear de modelos de árvores, como o LightGBM, que podem encontrar relações distintas não captadas pela correlação de Pearson.

6.5.2 Avaliação na base de teste

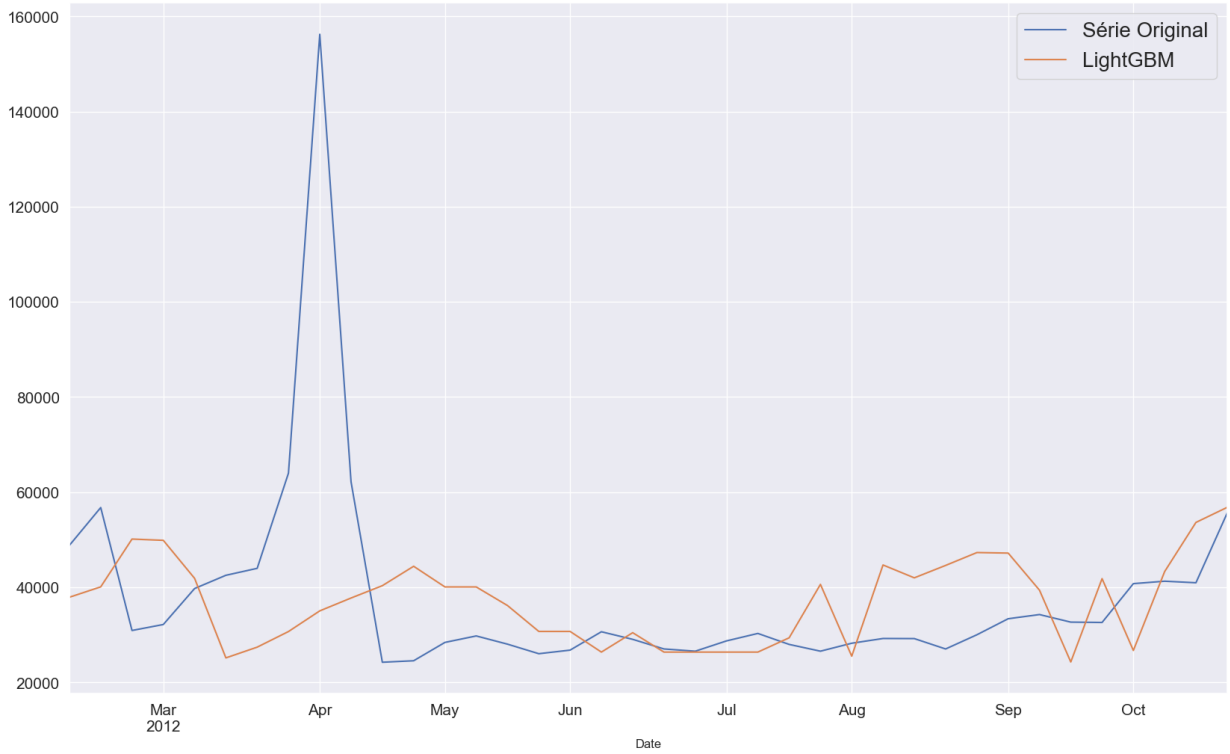
Os resultados da comparação do modelo treinado na base de testes é apresentado nas tabela e figura à seguir.

Tabela 15: Métricas de avaliação do modelo LightGBM na série multivariada.

Modelo	MAPE
LightGBM	34,4%

Fonte: Autor.

Figura 43: Previsão do método LightGBM comparado à base de teste da série multivariada.



Fonte: Autor.

6.6 Análise dos resultados

Da mesma forma que no capítulo anterior, temos a seguir os resultados obtidos na base multivariada ordenada de forma crescente pelo MAPE de cada algoritmo testado no experimento.

Tabela 16: Resultados obtidos na base multivariada.

Modelo	MAPE	Tempo de execução (s)
Prophet	23,52%	64,3
LSTM	24,56%	5,7
SARIMA	33,59%	654,7
LightGBM	34,43%	0,2
Mul_Holt-Winters	38,95%	17,3
Add_Holt-Winters	42,78%	7,4
ARIMA	44,76%	70,3

Fonte: Autor.

Diferente do teste do capítulo anterior, neste teste é utilizada a base multivariada, que conforme descrito no capítulo 3 foi selecionada para efetivamente comparar os algoritmos selecionados para o estudo.

É importante perceber o contraste nos resultados obtidos na série temporal escolhida da base de dados multivariada com a série temporal do capítulo anterior. Diferentemente da base univariada, a base multivariada não contém tratamento prévio para seleção de séries estáveis e consideradas adequadas para o experimento de previsão. Ao invés disso, a base multivariada contém séries temporais de diversas lojas da rede Walmart de varejo, expondo diversos casos onde existem séries de vendas com comportamentos instáveis e que acabam impondo maiores dificuldades no exercício.

Podemos observar que neste caso os métodos mais recentes se mostraram melhores para prever a demanda. Porém, é importante ressaltar que nenhum dos métodos obteve um resultado tão satisfatório quanto os resultados na base univariada para prever o comportamento da demanda observada.

Há um fator importante neste segundo experimento, que é a característica desbalanceada da série temporal de vendas. Há uma sazonalidade clara presente na base, porém ela se comporta de forma instável, tendo picos desproporcionais de vendas em momentos específicos do ano, o que dificulta para todos os métodos escolhidos alcançar uma previsão precisa.

Além disso, é importante ressaltar que os métodos tradicionais são alimentados por menos informações, dado que são modelados apenas na série temporal de vendas, enquanto todos os outros métodos comportam a possibilidade de utilizar outras variáveis presentes na base para ajudar na previsão.

Não obstante, vemos que de todos os métodos selecionados o Prophet foi o de melhor performance neste experimento, apesar de ter como única variável complementar além das vendas a marcação de feriados.

Este experimento foi realizado com fim ilustrativo de demonstrar o exercício detalhado da previsão de demanda, mas dessa vez num caso mais próximo do que é proposto para o experimento completo do capítulo seguinte. Deve-se ter em mente que o maior erro encontrado na aplicação dos algoritmos selecionados nesta série temporal não implica necessariamente na inviabilidade de utilizá-los para o planejamento operacional de um varejista. No caso, poderia simplesmente implicar em um maior estoque de segurança, que apesar de trazer maiores custos para a organização, ainda assim permitiria evitar problemas de falta de estoque para vendas futuras.

7 APLICAÇÃO EM AMOSTRA DE SÉRIES DA BASE MULTIVARIADA

7.1 Descrição do experimento

A fim de comparar os diferentes algoritmos selecionados para o estudo com relevância estatística, foi repetido o mesmo experimento do capítulo 6 em uma amostra de 100 séries temporais escolhida de forma aleatória das 3331 séries temporais presentes na base de dados multivariada. Em cada uma dessas séries selecionadas, foram aplicados os algoritmos estudados registrando o MAPE obtido por cada um.

Importante ressaltar que foram excluídos desta análise os modelos considerados redundantes para essa etapa do experimento. Estes são o ARIMA e o Holt-Winters aditivo. Isto porque ambos têm alternativas que apresentaram melhor precisão nos resultados obtidos até então, sendo as alternativas escolhidas para esta última análise o SARIMA e o Holt-Winters multiplicativo, respectivamente.

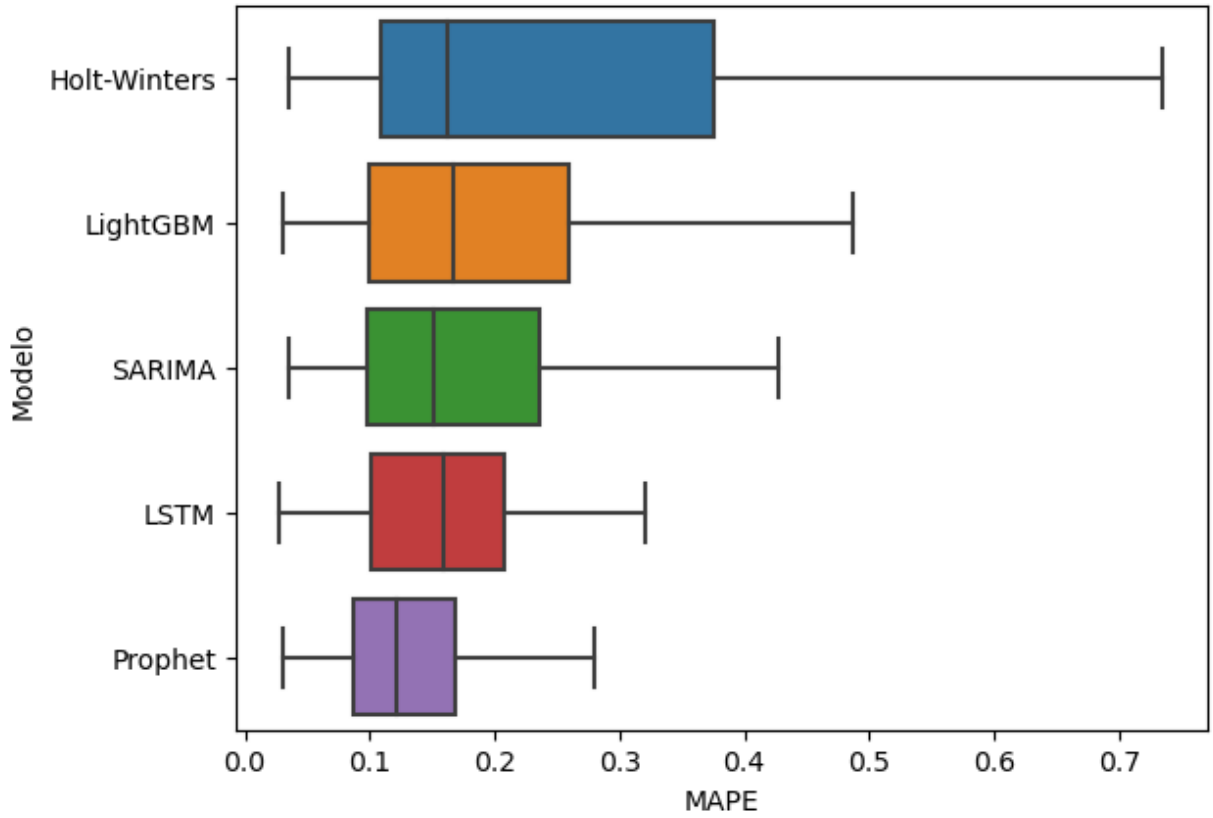
Além disso, todos os modelos são aplicados nas mesmas séries temporais. Ou seja, a amostragem aleatória é feita apenas uma vez, separando 100 séries temporais, e nestas mesmas séries são aplicados todos os métodos estudados a fim de obter maior comparabilidade. As séries temporais selecionadas para o estudo estão identificadas no apêndice.

7.2 Resultados

7.2.1 Visão geral

Aplicando os modelos em cada uma das 100 séries temporais escolhidas aleatoriamente, obtiveram-se os resultados apresentados na figura 44.

Figura 44: Resultado do experimento iterativo na base de dados multivariada.



Fonte: Autor.

Primeiramente, é possível aferir que as variâncias dos resultados obtidos por cada modelo proposto são diversas. Temos que métodos como o Prophet e o LSTM apresentam menor variância de MAPE obtidos, enquanto o Holt-Winters apresenta uma variância consideravelmente maior. Isto demonstra uma maior flexibilidade dos modelos de variância menor para aplicações em diversos cenários.

Segundamente, pode-se afirmar que em grande parte das séries temporais, o erro da previsão é consideravelmente grande. Escolhendo-se o valor de erro abaixo de 10% como um erro aceitável para operação, pode-se confirmar isso. A tabela 17 apresenta a porcentagem dos casos onde o MAPE obtido na série temporal escolhida é menor do que 10%.

Tabela 17: Porcentagem de previsões do experimento com MAPE menor que 10%.

Modelo	MAPE < 10%
Prophet	32%
SARIMA	27%
LightGBM	26%
LSTM	24%
Holt-Winters	22%

Fonte: Autor.

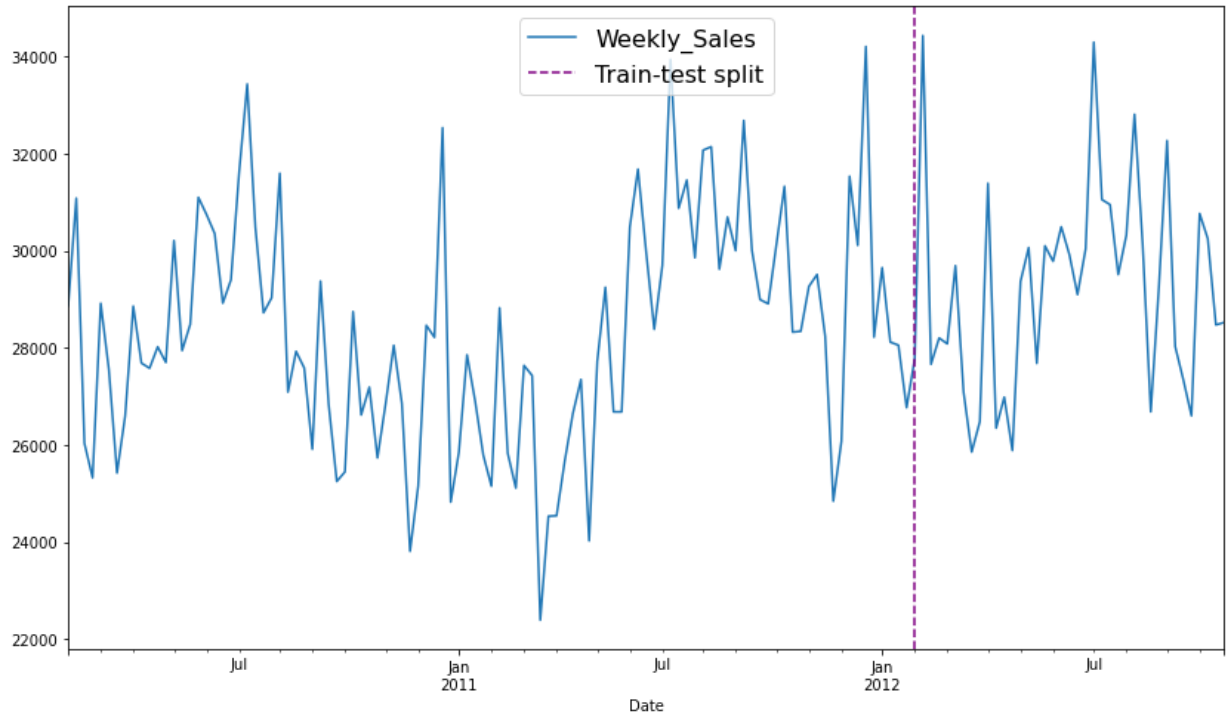
Para entender em quais casos de séries temporais os métodos tradicionais e os métodos de machine-learning se destacam, foram separados duas das 100 séries da amostra. A primeira série selecionada apresenta um caso onde os modelos estatísticos tradicionais obtiveram maior assertividade. A segunda, analogamente, foi selecionada para apresentar um caso onde modelos de machine-learning performaram melhor.

7.2.2 Caso de melhor performance de modelos estatísticos tradicionais

Como todos os modelos foram aplicados nas mesmas 100 séries temporais, é possível separar uma delas onde os resultados de algoritmos estatísticos tradicionais apresentaram maior precisão.

Para isso, foi separada a série temporal definida pela loja 24 e pelo departamento 80. A figura 45 apresenta seu comportamento ao longo do tempo.

Figura 45: Série temporal da loja 24, departamento 80.



Fonte: Autor.

É possível observar que o comportamento da série temporal em questão é muito próximo do que se observa na série univariada deste estudo. Isto fortalece a hipótese de que nos casos de comportamento estável, sem grandes outliers e variações na sazonalidade ao longo do tempo, métodos estatísticos tradicionais obtêm performance melhor do que os métodos de aprendizado de máquina.

Os resultados obtidos nesta série temporal são apresentados na tabela 18.

Tabela 18: MAPE's obtidos dos modelos na série temporal da loja 24, departamento 80.

Series	Holt-Winters	SARIMA	Prophet	LSTM	LightGBM
24-80	5,01%	6,22%	5,63%	8,63%	6,36%

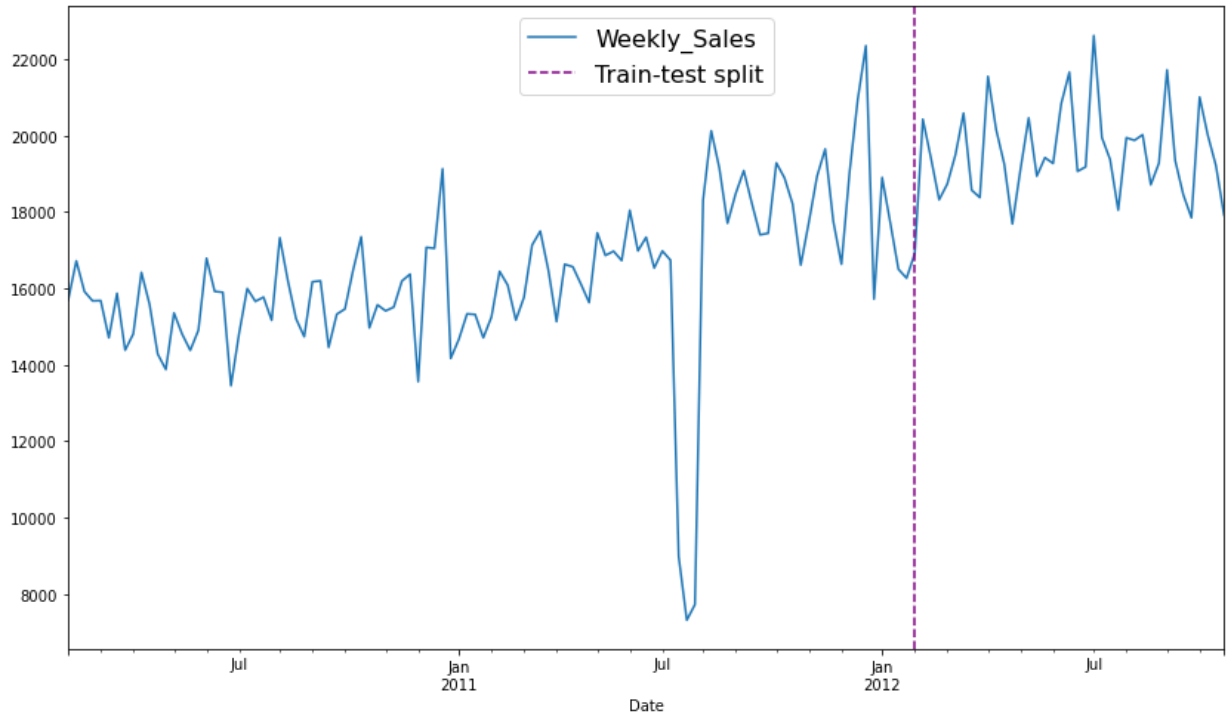
Fonte: Autor.

7.2.3 Caso de melhor performance de modelos de machine-learning

Da mesma forma que no item anterior, foi selecionada uma série temporal onde os métodos de machine-learning performaram melhor que os tradicionais para análise.

Para isso, foi separada a série temporal definida pela loja 1 e pelo departamento 80. A figura 46 apresenta seu comportamento ao longo do tempo.

Figura 46: Série temporal da loja 1, departamento 80.



Fonte: Autor.

Podemos verificar que, neste caso, a série temporal apresenta um claro outlier de vendas depois do mês de Julho. Por conta disso, métodos estatísticos tradicionais têm dificuldade em distinguir estes outliers do restante dos dados de vendas, e portanto apresentam uma performance menos satisfatória. Já os métodos de aprendizado de máquina obtiveram resultados melhores, com especial foco no LSTM.

Tabela 19: Resultados dos modelos na série temporal da loja 1, departamento 80.

Series	Holt-Winters	SARIMA	Prophet	LSTM	LightGBM
1-80	11,73%	15,53%	8,19%	5,45%	7,71%

Fonte: Autor.

8 CONCLUSÕES

A previsão de demanda é um tópico essencial para o planejamento estratégico e operacional para todo tipo de organização. Há anos são estudados métodos com diferentes abordagens para alcançar resultados mais precisos de previsão de valores futuros em séries temporais, e apesar dos avanços metodológicos e tecnológicos significativos nos últimos anos, ainda hoje modelos estatísticos tradicionalmente utilizados apresentam resultados satisfatórios em diversos casos.

Tanto no capítulo 5 quanto no capítulo 7, pudemos observar casos onde isto é verdade para modelos tradicionais como a Suavização Exponencial e o ARIMA. Estes são casos onde há uma tendência e sazonalidade claras presentes na série temporal, e portanto a auto-correlação dos valores presentes na série é uma boa base para prever valores futuros. Mais especificamente no capítulo 7, o algoritmo Holt-Winters obteve o menor erro de previsão em 16% dos casos. Analogamente, o SARIMA obteve o menor erro de previsão em 13% dos casos.

Já nos capítulos 6 e 7, foi possível entender em quais situações os métodos de machine-learning apresentam uma previsão mais satisfatória. Estes foram identificados como casos de séries temporais instáveis, onde a sazonalidade e a tendência não apresentam um comportamento regular em todo o período observado. Isto pode ser observado na série temporal analisada no capítulo 6, que contém picos de vendas que ocorrem em épocas diferentes nos anos registrados, e em volume muito maior do que as vendas usuais. Também foi possível constatar na série temporal separada para análise individual no capítulo 7, que contém um vale grande de vendas perto de Julho de 2011 cujo não apresenta ciclicidade ao longo do tempo. Estes casos se mostraram bastante presentes na amostra selecionada da base de dados multivariada, o que indica que não são desprezíveis no dia-a-dia da operação de uma organização.

Nos experimentos realizados na amostra aleatória, o Prophet apresentou o menor erro em 41% dos casos, enquanto o LSTM e o LightGBM apresentaram o menor erro em 19% e 11% dos casos respectivamente. Isto é interessante principalmente se analisado juntamente com a figura 44, onde percebe-se que, de todos os algoritmos estudados, o Prophet e o LSTM apresentam a menor variância nos resultados obtidos. Isto indica que esses dois métodos apresentam uma maior flexibilidade para se lidar com diferentes situações de séries temporais,

tanto nas que apresentam comportamento regular ao longo do tempo quanto nas que apresentam irregularidades como as descritas no parágrafo anterior. Isto é especialmente verdade para o Prophet, que de todos os algoritmos estudados apresentou a menor variância de resultados.

Naturalmente, em situações de séries temporais onde a sazonalidade, tendência e auto-correlação dos valores presentes não são regulares, todos os métodos apresentam dificuldades em prever a demanda. Porém, foi possível constatar que não só diferentes modelagem mas também o uso de dados exógenos à série temporal de previsão podem servir de ajuda para obter resultados mais satisfatórios do que os que poderiam ser alcançados com métodos estatísticos tradicionais.

É importante reconhecer que existem limitações para o estudo realizado.

Primeiramente, como a proposta se conteve em analisar os métodos quantitativos de previsão de demanda, não se pôs em prática a possível contribuição de métodos qualitativos, como o de Delphi, para a previsão. Além disso, análises qualitativas podem servir de ajuda para identificação e remoção de outliers da série temporal analisada, assim como inclusão de eventos que representem pontos de inflexão na distribuição de vendas presente, como por exemplo promoções e mudanças de comportamento dos clientes. Ademais, podem ser usadas para definir o melhor agrupamento temporal para cada série temporal à ser prevista, dado que existem itens que a previsão não necessariamente precisa ser semanal, e que teriam uma previsão mais assertiva se agrupados por outras unidades de tempo. Estas análises podem entrar não só como substituição dos algoritmos estudados, mas também como adição para melhorar o tratamento das séries temporais estudadas e melhor aplicação dos algoritmos.

Além disso, temos a limitação de dados escolhidos para a previsão da base multivariada. Por lidar-se com uma base de dados aberta com variáveis já selecionadas para o estudo, não passou-se pela importante etapa onde se realiza um estudo exploratório de relevância das variáveis exógenas que podem incrementar a previsão. Possivelmente existem muitas outras variáveis melhores para a previsão multivariada das vendas.

Ademais, existem desdobramentos possíveis para aprofundamento do presente estudo que não só a adição de métodos qualitativos e de outras variáveis exógenas.

Temos como possibilidade a exploração das demais séries temporais presentes na base multivariada, cada uma com um comportamento distinto e possível de descobertas adicionais.

Por fim, pode-se explorar diferentes formas de modelagem dos algoritmos estudados. Há por exemplo outras opções de hiperparâmetros para os algoritmos de aprendizagem de máquina, assim como outras arquiteturas de rede neural possíveis para a implementação do LSTM. Neste mesmo sentido, é possível expandir a análise adicionando outros algoritmos para fins de comparação.

REFERÊNCIAS

ADAM-BOURDARIOUS, Claire et al. The Higgs boson machine learning challenge. NIPS 2014 Workshop on High-energy Physics and Machine Learning, Dec 2014, Montreal, Canada. pp.37.

ALON, Ilan; QI, Min; SADOWSKI, Robert J. Forecasting aggregate retail sales:: a comparison of artificial neural networks and traditional methods. Journal of retailing and consumer services, v. 8, n. 3, p. 147-156, 2001.

ALPAYDIN, Ethem. Introduction to machine learning. MIT press, 2020.

BALLOU, R. H. Gerenciamento da cadeia de suprimentos: planejamento, organização e logística empresarial. 4. ed. Porto Alegre: Bookman, 2001.

BARBOSA, Anderson Henrique. Análise de Confiabilidade Estrutural Utilizando o Método de Monte Carlo e Redes Neurais. Ouro Preto, 2004. 143p Dissertação (Engenharia Civil) - Universidade Federal de Ouro Preto, 2004.

BARBOSA, C. L. R. Estudo de previsão de demanda de movimentação de diesel, gasolina e álcool, ver terminal petroquímico de Miramar, utilizando abordagem de séries temporais e séries causais. In: XXXIV ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 2014, Curitiba.

CALLEN, Jeffrey L. et al. Neural network forecasting of quarterly accounting earnings. International Journal of Forecasting, v. 12, n. 4, p. 475-482, 1996.
CARUANA, Rich; NICULESCU-MIZIL, Alexandru. An empirical comparison of supervised learning algorithms. In: Proceedings of the 23rd international conference on Machine learning. 2006.

CASULA, H. C. Aplicação de técnicas de previsão de demanda em manufatura: estudo de caso em uma indústria de laminados. 2012. 99 f. Dissertação (Mestrado em Engenharia Mecânica) - Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, Campinas, 2012.

CHOPRA, S. Gerenciamento da cadeia de suprimentos. São Paulo: Pearson Prentice Hall, 2003. controle da produção. Rio de Janeiro: Elsevier, 2008.

DA SILVA, Ivan Nunes; FLAUZINO, Rogério Andrade; SPATTI, Danilo Hernane. Redes Neurais Artificiais para Engenharia e Ciências Aplicadas: Curso Prático. Artliber, 2010.

FILDES, R.; HIBON, M.; MAKRIDAKIS, S.; MEADE, N. Generalising about univariate forecasting methods: Further empirical evidence. International Journal of Forecasting, v. 14, p. 339-358, 1998.

FLOWERS, A. D. A Simulation Study Of Smoothing Constant Limits For An Adaptive Forecasting System. Journal of Operations Management, v. 1, n. 2, p. 85-94, 1980.
FRAGOSO, B. B. Método para a criação de um processo de previsão de demanda de vendas. 2009.121 f. Dissertação (Mestrado em Engenharia Mecânica) - Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, Campinas, 2009.

GARDNER, E. S. Exponential smoothing: the state of the art-Part II. International Journal of Forecasting, v. 22, p. 637-666, 2006.

GUOLIN, et al. "Lightgbm: A highly efficient gradient boosting decision tree." Advances in neural information processing systems, 2017.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. Data mining, inference, and prediction. The elements of statistical learning Springer Series in Statistics. Springer-Verlag, New York, 2001.

HAYKIN, Simon. Neural Networks and Learning Machines. 3. ed. Hamilton, Ontario, Canada: Pearson Education, 2009. 823 p.

HILL, Tim; O'CONNOR, Marcus; REMUS, William. Neural network models for time series forecasts. Management science, v. 42, n. 7, p. 1082-1092, 1996.

HOCHREITER S., SCHMIDHUBER J. Long short-term memory. Neural Computation 1997.

HOPFIELD JJ. Neurons with graded response have collective computational properties like those of two-state neurons. Proceedings of the National Academy of Sciences 1984.

HYNDMAN, R. J.; ATHANASOPOULOS, G. Forecasting : Principles and Practice. 2. ed. Heathmont, Vic.: Otexts, 2018.

JAMES, G. et al. An introduction to statistical learning : with applications in R. [s.l.] Springer, 2013.

LEITE, Lucas de Oliveira Garrigós. Mineração de Dados Aplicada à Previsão do Preço de Ações de Concessionárias de Energia Elétrica do Estado de São Paulo. São Carlos - SP, 2016. 39 p TCC (Engenharia Elétrica) - UNIVERSIDADE DE SÃO PAULO, 2016.

LUSTOSA, L. J.; MESQUITA, M. A.; QUELHAS, O. L. G.; OLIVEIRA, R. J. Planejamento
e
MAKRIDAKIS, S. The art and science of forecasting: An assessment and future directions. International Journal of Forecasting, v. 2, p. 15-39, 1986.

MAKRIDAKIS, S.; HIBON, M. Exponential smoothing: The effect of initial values and loss functions on post-sample forecasting accuracy. International Journal of Forecasting, v. 7, p. 317–330, 1991.

MENTZER, J. T.; MOON, M. A. Sales Forecasting Management: a demand management approach. 2. ed. London: SAGE Publications, 2005.

RUMMELHART DE, Hinton GE, Williams RJ. Learning Internal Representations by Error Propagation. La Jolla, CA. USA: Institute for Cognitive Science, University of California, San Diego, 1985.

SHARDA, Ramesh; PATIL, Rajendra B. Connectionist approach to time series prediction: an empirical test. Journal of Intelligent Manufacturing, v. 3, n. 5, p. 317-323, 1992.

SILVA, Marcelino Pereira dos Santos. Mineração de Dados - Conceitos, Aplicações e Experimentos com Weka. Belo Horizonte - MG, 2004 Monografia - UNIVERSIDADE DO ESTADO DO RIO GRANDE DO NORTE.

SLACK, N. Administração da produção. 2. ed. São Paulo: Atlas, 2007.

SOARES, Anderson da Silva. Predição de Séries Temporais Econômicas por Meio de Redes Neurais Artificiais e Transformada Wavelet: Combinando Modelo Técnico e Fundamentalista. São Carlos - SP, 2008. 92 p Dissertação (Engenharia Elétrica) - Universidade de São Paulo, 2008.

Store Item Demand Forecasting Challenge. Disponível em: <https://www.kaggle.com/competitions/demand-forecasting-kernels-only/overview>. Acesso em: 11 out. 2022.

SWANSON, Norman R.; WHITE, Halbert. A model-selection approach to assessing the information in the term structure using linear models and artificial neural networks. Journal of Business & Economic Statistics, v. 13, n. 3, p. 265-275, 1995.

TAMIR, M. What Is Machine Learning? - I School Online. Disponível em: <https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/>. Acesso em: 11 out. 2022.

TAYLOR, Sean J.; LETHAM, Benjamin. Forecasting at scale. The American Statistician, v. 72, n. 1, p. 37-45, 2018.

TERÄSVIRTA, Timo; VAN DIJK, Dick; MEDEIROS, Marcelo C. Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. International Journal of Forecasting, v. 21, n. 4, p. 755-774, 2005.

The history of Amazon's forecasting algorithm. Disponível em: <https://www.amazon.science/latest-news/the-history-of-amazons-forecasting-algorithm>. Acesso em: 11 out. 2022.

Walmart Recruiting - Store Sales Forecasting. Disponível em: <https://www.kaggle.com/competitions/walmart-recruiting-store-sales-forecasting>. Acesso em: 11 out. 2022.

ZHANG, Wei; CAO, Qing; SCHNIEDERJANS, Marc J. Neural network earnings per share forecasting models: A comparative analysis of alternative methods. Decision Sciences, v. 35, n. 2, p. 205-237, 2004.

APÊNDICE

As 100 séries temporais selecionadas para compor a amostra aleatória estudada no capítulo 7 são identificadas segundo o modelo a seguir.

(Número da série). (Loja)-(Departamento)

Por exemplo, a primeira série da amostra é composta pela loja 24, departamento 50. Portanto, está identificada na lista da seguinte forma.

1. 24-50

Dito isto, segue a lista de séries temporais da amostra selecionada para o estudo.

1. 24-50
2. 34-90
3. 9-8
4. 22-79
5. 18-35
6. 38-3
7. 36-60
8. 42-5
9. 28-81
10. 39-9
11. 37-79
12. 30-13
13. 34-21
14. 33-80
15. 16-11
16. 32-83
17. 7-32
18. 10-72
19. 12-38
20. 24-80
21. 19-26
22. 1-41
23. 14-34
24. 22-3
25. 6-95
26. 19-72
27. 2-16
28. 29-34
29. 19-2
30. 39-35
31. 13-7
32. 30-94
33. 25-1
34. 3-23
35. 16-90
36. 1-80
37. 11-91
38. 14-10
39. 44-83
40. 9-7
41. 11-34
42. 44-4
43. 27-87
44. 20-87
45. 26-14
46. 45-17
47. 16-52
48. 1-92
49. 14-1
50. 11-1
51. 17-19
52. 31-21
53. 25-87
54. 44-7
55. 16-71
56. 26-31
57. 6-97
58. 34-29
59. 18-33
60. 3-5
61. 4-19
62. 16-6
63. 25-24
64. 30-95
65. 26-34
66. 40-67
67. 29-16
68. 7-16
69. 19-35
70. 26-27
71. 8-83
72. 45-56
73. 25-35
74. 43-16
75. 34-11
76. 21-10
77. 15-87
78. 22-83
79. 33-98
80. 40-92
81. 13-12
82. 25-93
83. 1-38
84. 13-85
85. 21-52
86. 3-38
87. 35-41
88. 22-59
89. 33-2
90. 30-90
91. 11-17
92. 18-27
93. 19-46
94. 7-8
95. 2-92
96. 33-3
97. 45-9
98. 16-87
99. 42-10
100. 16-72