

**UNIVERSIDADE DO VALE DO RIO DOS SINOS - UNISINOS
UNIDADE ACADÊMICA DE GRADUAÇÃO
CURSO DE ENGENHARIA DE PRODUÇÃO**

THAMIRIS MEDEIROS ARNECKE

**PREVISÃO DE DEMANDA APLICADA A SÉRIES TEMPORAIS DA
VITIVINICULTURA GAÚCHA**

Porto Alegre

2021

THAMIRIS MEDEIROS ARNECKE

**PREVISÃO DE DEMANDA APLICADA A SÉRIES TEMPORAIS DA
VITIVINICULTURA GAÚCHA**

Artigo apresentado como requisito parcial
para obtenção do título de Bacharel em
Engenharia de Produção, pelo Curso de
Engenharia de Produção da Universidade
do Vale do Rio dos Sinos - UNISINOS

Orientador: Prof. Ms. Marcos Leandro Hoffmann Souza

Porto Alegre

2021

Pela realização desse trabalho, agradeço a Deus por me dar forças e guiar meu caminho. Agradeço aos meus pais, João Batista e Márcia Helena, pela vida e por terem me ensinado a importância do conhecimento e o valor da disciplina. Agradeço ao meu esposo, Gustavo Antônio, por estar ao meu lado nos momentos difíceis e me apoiar em toda a minha transição de carreira. Agradeço também ao meu orientador, Marcos, pela paciência e ensinamentos nessa jornada.

PREVISÃO DE DEMANDA APLICADA A SÉRIES TEMPORAIS DA VITIVINICULTURA GAÚCHA

DEMAND FORECASTING APPLIED TO TIME SERIES OF RIO GRANDE DO SUL VINEYARDS

Thamiris Medeiros Arnecke*

Marcos Leandro Hoffmann Souza**

Resumo: A indústria agrícola caracteriza-se por um ambiente complexo e modelos preditivos podem auxiliar gestores e agricultores na organização e planejamento de seus negócios. No Brasil, a vitivinicultura tem ganhado espaço de forma estratégica nos últimos anos. O objetivo do presente estudo foi realizar a previsão de demanda para produção e comercialização de vinhos no Rio Grande do Sul. Para isso, utilizou-se *Cross Industry Standard Process for Data Mining* (CRISP-DM) e a plataforma *Amazon Forecast* que aplicou algoritmos convencionais e avançados para previsão de demanda de um período das séries temporais da vitivinicultura gaúcha. O método ARIMA demonstrou-se como o melhor algoritmo preditivo por atender aos critérios definidos para precisão. A demanda prevista para produção de vinhos foi semelhante à demanda real, enquanto a demanda prevista para comercialização de vinhos foi menor do que a demanda real o que se pode atribuir a efeitos inesperados no último ano tais como eventos climáticos favoráveis e pandemia. A abordagem utilizada nesse estudo é abrangente e compatível com diferentes contextos, podendo ser utilizada para previsão de demanda de qualquer produto que possua registros em um horizonte de tempo.

Palavras-chave: Modelagem, Previsão de Demanda, Séries Temporais, Vitivinicultura.

Abstract: The agricultural industry is characterized by a complex environment and predictive models can help managers and farmers in organizing and planning their businesses. In Brazil, vitiviniculture has gained strategic space in recent years. The aim of this study was to forecast demand for the production and sale of wines in state of Rio Grande do Sul. For this, we used the *Cross Industry Standard Process for Data Mining* (CRISP-DM) and the *Amazon Forecast* platform that applied conventional and advanced algorithms for forecasting for a period of the time series of vitiviniculture in Rio Grande do Sul. The ARIMA method proved to be the best predictive algorithm for meeting the defined criteria for precision. Forecast demand for wine production was similar to actual demand, while forecast demand for wine sale was lower than actual demand, which can be attributed to unexpected effects in the last year such as favorable weather events and COVID-19 pandemic. The approach used in this study

* Acadêmica em Engenharia de Produção, Universidade do Vale do Rio dos Sinos (UNISINOS), thamirisdemedeiros@gmail.com.

** Professor Orientador. Acadêmico em Doutorado em Computação Aplicada (PIPCA), Mestre e Graduado em Engenharia de Produção pela Universidade do Vale do Rio dos Sinos (UNISINOS), mlhoffmann@unisinobr.

is comprehensive and compatible with different contexts, and can be used to forecast demand for any product that has records in a time horizon.

Keywords: Modeling, Forecasting, Time Series, Vitiviniculture.

1 INTRODUÇÃO

A indústria agrícola caracteriza-se por um ambiente complexo que sofre influência de questões geográficas, climáticas, demográficas, econômicas e governamentais exigindo decisões dinâmicas, as quais terão efeitos observados em meses, anos, e por vezes, décadas (FISHER et al., 2000). Adicionalmente, o uso de dados é essencial para definição de políticas públicas do setor que influenciam liberações de crédito, planejamento de recursos, aquisição de insumos, direcionamento de investimentos, busca de mercado e desenvolvimento de tecnologias (FISHER et al., 2000; IFFT; KUHNS; PATRICK, 2018).

A aplicação de tecnologia avançada e o uso de dados para a tomada de decisão tem se tornado uma realidade no agronegócio, uma vez que contribuem para uma produção de alimentos mais eficiente, sustentável e que atenda às necessidades nutricionais da crescente população mundial (CULP, 2019; KOLTES et al., 2019).

Nesse sentido, a indústria vitivinícola faz parte desse cenário e traz características inerentes, entre elas produtividade, receita de vendas, rentabilidade, mercado marcado por comportamentos sociais e orientado pelo lucro (CZUPRYNA et al., 2018). Segundo a Organização Internacional do Vinho (2021), em 2020, o Brasil ocupava a 21ª posição mundial em área plantada em vinhedos (1,1%), 18ª posição em produção de vinhos (0,01%) e 13ª posição como consumidor (2,0%). Em 2019, a uva representou 0,09% da colheita brasileira e 0,93% do valor da produção nacional (IBGE, 2019). No mesmo ano, 90% da produção brasileira de uvas e vinhos concentrava-se no Rio Grande do Sul sendo a região da Serra Gaúcha onde localizam-se o maior número de produtores, em especial no município de Bento Gonçalves (FARIAS; TATSCH, 2014; MELLO; MACHADO, 2019).

Mundialmente, a vitivinicultura vivencia um paradoxo como atividade econômica, observando-se uma tendência de diminuição na área plantada de vinhedos enquanto no mesmo período há aumento na demanda pelo consumo de vinho, demonstrando a necessidade de operações produtivas, eficientes e lucrativas para que sejam competitivas (IBRAVIN, 2018; OIV, 2020). No Brasil, a indústria

vitivinícola tem recebido incentivos às certificações de indicação geográfica, desenvolvimento do enoturismo e apresentando produtos com maior nível de qualidade como diferenciais de mercado (PEIREIRA et al., 2020).

Embora não seja a bebida alcoólica mais consumida no país, tendo um consumo *per capita* anual menor de 3L, o consumo de vinho pelo brasileiro tem crescido nos últimos anos (ALMEIDA; BRAGAGNOLO; CHAGAS, 2015; OIV, 2021). Associado à melhoria da qualidade do vinho nacional, ambos tem impulsionado também importações e exportações e por consequência o aumento da demanda pelo produto (DE MELLO; MACHADO, 2020).

Perante esse contexto, modelos preditivos podem auxiliar gestores e agricultores na organização e planejamento de seus negócios. Uma vez que o volume de dados disponível é crescente, a combinação entre modelos estatísticos e técnicas computacionais propiciam um meio para melhor gerenciar os recursos e otimizar o desempenho da operação (VERMA; SHARMA; JINDAL, 2020).

É possível identificar na literatura a recomendação pela utilização de modelos combinados para previsão de demanda e diferentes autores sugerem a busca pelo aperfeiçoamento e otimização dos modelos existentes, bem como a validação e combinação com estratégias de aprendizado computacional (CHU et al., 2019; DE OLIVEIRA et al., 2020; JHA; SINHA, 2013; RIBEIRO; OLIVEIRA, 2011; WANG et al., 2018; ZHANG, 2003).

Sendo assim, o objetivo do presente estudo foi realizar a previsão de demanda para produção e comercialização de vinhos no Rio Grande do Sul. Ainda, apresentou-se um modelo preditivo, o qual oferece dados de previsão de demanda podendo auxiliar gestores da vitivinicultura na tomada de decisão tais como compra de insumos e negociação de venda do produto. Em suma, o presente trabalho encontrou uma previsão de redução na produção de vinhos e aumento na comercialização em um período, além de identificar o método ARIMA como o mais apropriado para a análise das séries estudadas.

Dessa forma, esse trabalho está organizado em 7 partes. Após a Introdução, os próximos capítulos encontrados são Fundamentação Teórica, onde são abordados conceitualmente modelos de aprendizagem profunda e suas aplicações; Materiais e Métodos, que descreve as etapas de desenvolvimento deste trabalho; Resultados, o qual apresenta os desfechos encontrados nessa pesquisa; Discussão, que traz a contextualização dos resultados perante a outras pesquisas descritas na literatura;

Considerações Finais, onde são abordadas as contribuições científicas do trabalho realizado bem como recomendações futuras; e por fim uma seção dedicada às referências utilizadas nesse estudo.

2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção, encontram-se a revisão de modelos convencionais e de aprendizagem profunda para previsão de demanda e suporte à tomada de decisão.

2.1 Previsão de Demanda Convencional

Segundo Cauchick-Miguel (2018), decisões estratégicas, táticas e operacionais são parte fundamental da gestão de sistemas de produção de produtos e serviços. Há pouco mais de uma década, Kerkkänen, Korpela e Huiskonen (2009) já apontavam em seu estudo mais de 70 diferentes técnicas de previsão de demanda descritas na literatura, o que torna a escolha do método adequado um desafio.

Entre as técnicas existentes está a previsão aplicada a séries temporais, a qual caracteriza-se por uma sequência de eventos ordenados em intervalos definidos, permitindo sua aplicação em qualquer variável que sofra alteração ao longo de um horizonte de tempo (TIWARI; VAIBHAW, 2020). Ainda, permite o agrupamento de dados registrados ao longo de períodos para construção e validação de modelos matemáticos capazes de estimar e prever valores futuros (ADHIKARI; AGRAWAL, 2013; KHANDELWAL; SATIJA; ADHIKARI, 2015; YUAN; SAN; LEONG, 2020). Tal método é largamente utilizado em diversas áreas de conhecimento tais como negócios, economia, finanças, ciência e engenharia (THOMAZI GONZALEZ, 2018).

Alguns métodos estatísticos clássicos para previsão de demanda em séries temporais são Média Móvel Simples, Média Móvel Dupla e Análise de Regressão (MOREIRA, 2011). Segundo Tiwari e Vaibhaw (2020), a ampla utilização de métodos lineares para previsão a partir de séries temporais se deve a sua fácil compreensão, baixo custo e eficácia na solução de problemas menos complexos.

Outro método amplamente utilizado, o modelo *Auto Regressive Integrated Moving Average* (ARIMA) utiliza médias móveis e pressupõe a correção linear entre os dados de séries temporais tendo como foco principal a redução de ruídos de frequência e extrair padrões para previsões futuras (PAPASTEFANOPOULOS;

LINARDATOS; KOTSIANTIS, 2020). Nessa lógica, o fundamento de auto-regressão avalia a correlação entre os eventos em estudo, enquanto a integração indica o grau de estacionariedade da série analisada e as médias móveis amenizam fatores desconhecidos, ainda sendo possível a análise de sazonalidade na variante SARIMA (PARMEZAN; SOUZA; BATISTA, 2019). Na mesma linha, Schutte (2018) utilizou ARIMA para prever demanda de produção de vinhos em estados norte-americanos pois o avaliou como um método completo por levar em consideração auto-regressão e técnica de média móvel.

Além desses, o método denominado *Exponential Smoothing* (ETS), tem em seus princípios a análise de dados com sazonalidade por meio de médias ponderadas das observações, diferenciando-se do Método de Média Móvel Simples por manter a suavização em declínio exponencialmente ao longo da linha de tempo (AMAZON WEB SERVICES, 2018). Por fim, Wang *et al.* (2018) destacam em seu estudo a dificuldade da obtenção de técnicas universais para previsão de demanda devido diferentes naturezas e cenários encontrados, demonstrando a necessidade de se compreender os dados em estudo para então elencar a técnica mais adequada.

2.2 Aprendizagem Profunda Aplicada

Com o volume de dados disponível nos dias de hoje, a interpretação dos mesmos torna-se um diferencial competitivo e estratégico para o mundo dos negócios uma vez que provém o embasamento para a tomada de decisão (KHAN *et al.*, 2020; ROMANOV; YARUSHKINA; FILIPPOV, 2020). Nesse sentido, a implementação de modelos de aprendizagem profunda proporciona uma ferramenta a fim de tornar mais robusto o processo de obtenção e interpretação da informação (MELNYK, 2020; SCHWARENTHORER *et al.*, 2020).

Nesse sentido, algoritmos que utilizam *deep learning* tem ganhado espaço na literatura recente, uma vez que são capazes de aprender automaticamente mapeamentos completos arbitrários e trabalhar com um expressivo volume de dados (CHEN; KUNG; HWANG, 2019; HUA *et al.*, 2019; KUREMOTO *et al.*, 2019; THOAZI GONZALEZ, 2018; TIWARI; VAIBHAW, 2020). Adicionalmente, entende-se que modelos de aprendizado contínuo superam as limitações presentes em modelos tradicionais para previsão em um grande número de observações (TIWARI; VAIBHAW, 2020). Dentre os modelos mais avançados encontram-se *Convolutional*

Neural Network (CNN), Non-Parametric Time Series (NPTS) e Prophet Forecasting Model (PFM).

2.1.1 *Convolutional Neural Network (CNN)*

Tem sua aplicação em 1D para séries temporais e 2D para base de dados por imagem e, entre uma camada de entrada e uma de saída, o modelo apresenta várias camadas ocultas responsáveis pelos cálculos de convolução, função de ativação e normalização (TIWARI; VAIBHAW, 2020). De maneira aplicada em séries temporais, os modelos CNN estão presentes em diferentes contextos.

Lee e Ahn (2018) utilizaram os princípios dessa metodologia para prever o direcionamento do mercado de ações e apresentaram o modelo CNN-FG em cinco etapas (*Convolutional Neural Network using Fluctuation Graph*) tendo como entrada gráficos de séries históricas. Os resultados demonstraram que o modelo proposto foi significativamente superior a *Logistic Regression (LOGIT)*, *Artificial Neural Networks (ANN)* e *Support Vector Machine (SVM)* gerando a recomendação dos autores para aplicação em problemas de negócios (LEE; AHN, 2018).

Pesquisadores da *Amazon* desenvolveram o algoritmo *Convolutional Neural Network - Quantile Regression (CNN-QR)* com base em conceitos de aprendizagem de máquina para previsão em séries temporais escalares em 1D (AMAZON WEB SERVICES, 2018). Outros pesquisadores também propuseram versões aprimoradas de CNN para previsões financeiras e encontraram melhores performances e resultados otimizados em comparação aos modelos convencionais (CAO; WANG, 2019; LIN et al., 2020; WU et al., 2020; XU et al., 2020b; ZHANG; WU; WANG, 2020).

Como diferencial estratégico no gerenciamento de energia elétrica, Kim et al. (2019) combinaram *Recurrent Neural Networks (RNN)* e CNN em um novo modelo para calibrar tempo e valores de previsão obtendo uma estrutura mais robusta para previsão de demanda no setor. De maneira similar, Kang e colaboradores (2020) comparam RNN e CNN separadamente e em conjunto e encontraram que o modelo isolado de CNN supera as demais propostas avaliadas para previsão de demanda de energia elétrica.

Em outra abordagem, Ju e colaboradores (2019) demonstraram um algoritmo baseado em CNN com bom desempenho em precisão e eficiência na previsão de energia eólica. Mujeeb et al. (2020) e Dou et al. (2018) desenvolveram modelos de

aprendizagem profunda fundamentados nos conceitos de CNN e validaram métodos com bons resultados também nessa área, trazendo alternativas para análises mais robustas. Ghimire *et al.* (2019) desenvolveram um modelo híbrido utilizando CNN e obtiveram bons prognósticos para previsão de radiação solar.

Além disso, outros pesquisadores aplicaram modelos híbridos de aprendizagem profunda baseados na lógica de CNN e obtiveram resultados promissores para previsões de temperatura da superfície marítima, eventos sísmicos, disponibilidade de oxigênio em aquicultura, qualidade do ar, mortalidade pós-operatória e velocidade de tráfego (FRITZ *et al.*, 2019; GENG *et al.*, 2019; GUO *et al.*, 2020; LV *et al.*, 2018; QI *et al.*, 2019; TA; WEI, 2018; XU *et al.*, 2020a; ZHAO *et al.*, 2019).

2.1.2 *Non-Parametric Time Series* (NPTS)

A lógica dos algoritmos de *Non-Parametric Time Series* (NPTS) consiste na utilização da amostragem de eventos anteriores para estatisticamente prever valores futuros, podendo ser aplicada em séries temporais que possuem lacunas e empregue em diferentes contextos (AMAZON WEB SERVICES, 2018). Estudos recentes baseados em aprendizado de máquina tem optado por premissas não-paramétricas pois realizam previsões sem a necessidade de definição prévia da característica da distribuição dos dados (GAUTAM; SINGH, 2020; PARMEZAN; SOUZA; BATISTA, 2019).

Modelos computacionais que utilizam esses princípios são *Artificial Neural Networks* (ANN) e *Support Vector Machines* (SVM) (PARMEZAN; SOUZA; BATISTA, 2019). Por definição, ANNs são modelos matemáticos para aproximação de função, classificação, reconhecimento de padrões e controle linear, entre outras características inspirados no processamento do cérebro humano e amplamente utilizados em estudos (KUREMOTO *et al.*, 2019; TEALAB, 2018). Já SVM é um modelo de regressão fundamentado na teoria do aprendizado estatístico e técnicas para minimização de erro de treinamento (PARMEZAN; SOUZA; BATISTA, 2019).

Wang *et al.* (2018) destacam a aplicabilidade de ANNs, devido a sua capacidade de modelagem flexível e característica auto adaptativa. Por outro lado, Adhikari e Agrawal (2013) sinalizaram limitações na estrutura de ANNs como taxa de convergência lenta, padrão complexo de superfícies de erro e deficiência na obtenção

de métodos de seleção de parâmetros de treinamento adequados. Para superá-las, desenvolveram e aplicaram em quatro séries temporais um esquema de combinação de ANNs, obtendo resultados positivos quanto ao aumento da precisão do modelo (ADHIKARI; AGRAWAL, 2013).

Em seu estudo, De Oliveira et al. (2020) recomendaram a combinação entre métodos paramétricos e não-paramétricos tais como *Auto Regressive Integrated Moving Average* (ARIMA) e *Artificial Neural Networks* (ANNs). Em revisão sistemática recente, autores destacaram que ainda não há consenso na literatura sobre qual das lógicas, lineares ou não-lineares, é mais eficiente pois ambas possuem fortalezas e limitações, recomendando-se a realização de mais estudos (GAUTAM; SINGH, 2020).

2.1.3 *Prophet Forecasting Model* (PFM)

O algoritmo denominado *Prophet Forecasting Model* (PFM) trata-se de um código aberto proposto por colaboradores do *Facebook* como uma alternativa aos métodos convencionais de previsão de demanda com base em séries temporais (TAYLOR; LETHAM, 2017). Tal método é capaz de adequar tendências não lineares a padrões de sazonalidade temporal em dias, semanas, anos e considerar feriados, bem como administrar dados faltantes de maneira acurada (VISHWAS et al., 2020). Segundo os autores, além da flexibilidade à sazonalidade e facilidade de ajuste a dados faltantes também são vantagens desse modelo a agilidade para editá-lo e parâmetros de fácil interpretação (TAYLOR; LETHAM, 2017).

Em diversos estudos encontra-se aplicado o método PFM para predição da qualidade e nível de poluição do ar demonstrando sua aplicabilidade em fenômenos ambientais (SAMAL et al., 2019; SHEN; VALAGOLAM; MCCALLA, 2020; SINGH, 2019; TEJASVINI et al., 2020; ZHAO et al., 2018). Oo e Phyu (2019) aliaram sensores inteligentes na coleta de dados reais de condições climáticas em estufa para validar as previsões de microclima obtidas com aplicação de PFM em séries históricas e obtiveram resultados significativos. Entre outros métodos, pesquisadores sugeriram que PFM apresenta-se como uma boa técnica para prever temperatura e precipitação (PAPACHARALAMPOUS; TYRALIS; KOUTSOYIANNIS, 2018). Entretanto, Asha et al. (2020) não corroboraram com essa definição, uma vez que compararam *Random Forest* e PFM para predição de temperatura e o primeiro modelo demonstrou resultados superiores quando comparado ao segundo.

Ying e colaboradores (2020) propuseram um modelo híbrido entre redes neurais de *Long Short-Term Memory* (LSTM) e PFM e recomendaram o segundo devido sua precisão e universalidade. Na área financeira, encontram-se estudos com bons resultados para aplicação de PFM em previsão de fluxo de caixa, valor de ações, Bitcoin, câmbio, entre outros (AHMED et al., 2020; CHAN, 2020; FANG et al., 2019; GONZÁLEZ MATA, 2020; MADHURI; CHINTA; KUMAR, 2020; VISHWAKARMA et al., 2020; WEYTJENS; LOHMANN; KLEINSTEUBER, 2019; YENIDOGAN et al., 2018).

Entre distintos exemplos que reforçam a importância do uso de aprendizagem profunda na tomada de decisão em negócios, ainda pode-se citar a pesquisa de Yurtsever (2020) que utilizou PFM para previsão de demanda de refeições em ambiente corporativo e ressalta a fácil aplicação do modelo. Da mesma forma, PFM também se destaca em artigos que utilizaram o modelo para previsão de vendas consolidando sua performance (NAVRATIL; KOLKOVA, 2019; YOO; OH, 2020; ZHAO; ZHANG, 2020; ZUNIC et al., 2020).

2.3 Aplicações na Agroindústria

Na cadeia da agricultura e negócios, modelos de previsão são aplicados há mais de um século com diferentes enfoques. A previsão de preços de *commodities* agrícolas no mercado é uma das principais aplicações na área (JHA; SINHA, 2013). Ribeiro e Oliveira (2011) aplicaram um modelo híbrido de previsão de preços para o setor de açúcar e álcool. Os autores utilizaram o método filtro de *Kalman* para descrever a evolução dos preços e conceitos de redes neurais artificiais para identificar variáveis que impactam no preço de ativos em análise.

Wang et al. (2019) propuseram uma abordagem acerca do preço de milho e soja com aplicação do método denominado *Artificial Bee Colony Algorithm* (ABC) que se baseia em otimização global. Diferentemente, Wu et al. (2017) propuseram a utilização combinada do modelo ARIMA e do método *Partial Least Squares Regression* (PLS) com base em fatores de tempo e espaço também para estimar preços de *commodities* agrícolas em cidade chinesa. Ainda na China, Chu e colaboradores (2019) desenvolveram um modelo híbrido com base na aplicação de inteligência artificial e dos conceitos de *Ensemble Empirical Mode Decomposition*

(EEMD) para previsão do preço da uva de mesa, o qual mostrou-se válido no contexto em estudo.

Chipanshi, Ripley e Lawford (1997) simularam a produção de trigo em três cidades canadenses utilizando dados climáticos e meteorológicos históricos. Após validar o modelo denominado *Crop Estimation Through Resource and Environment Synthesis-Wheat* (CERES), que obteve bons resultados comparado a dados reais, os autores recomendaram a utilização da técnica para auxiliar a tomada de decisão e minimizar perdas na produção em cenários não favoráveis. Akter e Rahman (2010) aplicaram modelos de séries temporais para prever a necessidade de leite em uma cooperativa de laticínios no Reino Unido. Os autores utilizaram modelos exponenciais ajustados à sazonalidade e obtiveram previsões com baixo percentual de erro, sugerindo a aplicação a cenários reais com características semelhantes.

No âmbito da indústria vitivinícola, Ye e Melkumian (2020) compararam três modelos para prever a demanda de vinho tinto na Austrália. Os pesquisadores testaram os modelos *Naive Forecast*, SARIMA e ANNs em seu estudo, no qual o segundo modelo apresentou melhores resultados. Além disso, De Oliveira et al. (2020) utilizaram dados de vinho tinto em séries temporais para testar e validar um modelo de previsão de demanda constituído por três etapas, sendo modelagem linear da série temporal, modelagem não linear da série de erros e combinação das previsões usando três abordagens distintas. Na mesma série de dados do presente estudo, Mareth et al. (2012) testou diferentes lógicas de programação matemática para previsão de demanda de vinho de mesa tinto e evidenciou a tendência linear dos dados.

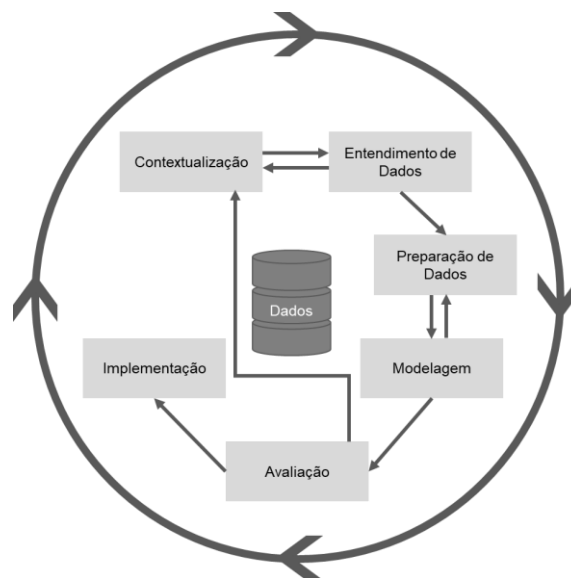
Em outra linha, Baidakov e colaboradores (2015) desenvolveram um modelo preditivo de negócio para aplicação na vitivinicultura, levando em consideração tendências para área de cultivo, custo de produção, fatores de impacto ambientais, interpretação de cálculos analíticos (análise de regressão e componente cíclica) e indicadores financeiros, permitindo a criação de cenários para avaliação de negócio. Na China, Ye, Li e Jia (2020) utilizaram algoritmos avançados de aprendizado de máquina para prever a qualidade de vinho tinto e obtiveram resultados positivos em um estudo inovador que propôs uma aliança entre essa tecnologia e o trabalho prático desenvolvido pelo enólogo.

3 MATERIAIS E MÉTODOS

Nesta seção, apresentam-se as etapas para construção e execução do presente estudo, assim como o detalhamento dos dados utilizados e métodos para obtenção dos resultados.

O procedimento metodológico utilizado neste trabalho foi *Cross Industry Standard Process for Data Mining* (CRISP-DM) que se trata de uma sequência de passos aplicáveis à mineração de dados em qualquer contexto de negócios e que catalisa dados em conhecimento para suporte à tomada de decisão (SHEARER, 2000). Baseado nos fundamentos apresentados por Shearer (2000), o estudo foi realizado nas etapas descritas na Figura 1 adaptadas para o contexto acadêmico:

Figura 1 - Fases do Modelo CRISP-DM



Fonte: Elaborada pela autora com base em Shearer (2000).

- i. Contextualização (*Business Understanding*): definição do contexto de negócio e objetivos, bem como a construção de um plano de ferramentas e técnicas a serem utilizadas;
- ii. Entendimento de Dados (*Data Understanding*): enfoque em uma visão inicial de dados, consistindo em coletar, descrever, explorar e verificar a qualidade dos mesmos;
- iii. Preparação de Dados (*Data Preparation*): preparação das informações para a versão final do banco de dados. Consiste em selecionar, limpar, construir, integrar e formatar os dados;

- iv. Modelagem (*Modeling*): parte do processo em que diferentes modelos e técnicas são selecionados, testadas e interpretados para identificar os melhores resultados;
- v. Avaliação (*Evaluation*): parte fundamental para avaliar o desempenho do modelo, a coerência dos resultados obtidos, bem como ajustar parâmetros, rever partes do processo e definir ajustes;
- vi. Implementação (*Deployment*): é a fase final de entrega de dados e aplicação no cenário corporativo, acompanhado de um plano de monitoramento, manutenção e revisão. Nesse estudo, tal etapa consiste na entrega e discussão de resultados.

A presente pesquisa analisou métodos de previsão de maneira independente, embora existam estudos que sugerem a combinação entre métodos de diferentes naturezas para modelos mais aprimorados (DE OLIVEIRA et al., 2020; GAUTAM; SINGH, 2020). O banco de dados utilizado para desenvolvimento deste trabalho trouxe registros históricos de produção e comercialização de vinhos no Rio Grande do Sul, entre os anos de 1970 e 2019, disponibilizados gratuitamente ao público pela Embrapa Uva e Vinho (DE MELLO; MACHADO, 2020).

O tratamento de dados foi feito em *Microsoft Excel 2013*. As variáveis disponíveis na base de dados formaram duas séries temporais principais, uma para produção e outra para comercialização de vinhos. Os *inputs* foram classe de vinho (de mesa, fino), tipo de vinho (branco, rosé, tinto), ano de registro (1970 a 2019), volume produzido (L) e volume comercializado (L).

Para a construção de uma série temporal complementar, buscou-se dados para representar dois principais fatores que influenciam o objeto em estudo, condições climáticas e econômicas. Assim, obteve-se valores anuais no mesmo período analisado para temperatura máxima (°C), temperatura mínima (°C) e volume de precipitação (mm) na região de Bento Gonçalves, RS, devido representatividade no contexto em estudo, bem como PIB per capita brasileiro (R\$), (BANCO CENTRAL, 2021; INSTITUTO NACIONAL DE METEOROLOGIA, 2021).

Na etapa de modelagem, iniciou-se com a criação de um cenário base para fins de controle, conforme indicado em estudos prévios (LV et al., 2018; TAYLOR; LETHAM, 2017; WAN et al., 2019). O cenário base, no termo em inglês denominado *baseline*, é uma importante parte da análise de dados pois consiste em criar uma referência de comparação de dados baseada em métodos matematicamente mais

simples e de eficiência reconhecida. Os registros históricos de produção e comercialização de vinhos foram analisados em *Microsoft Excel* 2013 utilizando como método preditivo de demanda a técnica de Média Móvel Dupla devido tendência linear existente. Para validar a precisão do modelo calculou-se *Weighted Average Percentage Error* (WAPE) e *Root Mean Squared Error* (RMSE), em acordo com estudos recentes (AMAZON WEB SERVICES, 2018; SEN; YU; DHILLON, 2019; SHEN; VALAGOLAM; MCCALLA, 2020; YOO; OH, 2020).

Posteriormente, outros modelos de previsão de demanda foram testados na funcionalidade *Amazon Forecast* da ferramenta *Amazon Machine Learning* da plataforma *Amazon Web Services* (AWS) (AMAZON WEB SERVICES, 2018). Para escolha da plataforma, levou-se em consideração outras ferramentas *Machine Learning as a Service* disponíveis, tais como *Microsoft Azure Machine Learning* e *Google AI Platform*, e ferramentas de código aberto, como por exemplo *Python* e *Software R*, porém a ferramenta *Amazon Machine Learning* demonstrou-se mais vantajosa nesse cenário. Entre os fatores positivos do uso da plataforma escolhida estão *interface* intuitiva, oferta de nível gratuito, utilização de algoritmos básicos e robustos em um único lugar e funcionalidade AutoML. O AutoML tem como característica realizar automaticamente a análise do comportamento dos dados e com base nisso ajustar hiper parâmetros dos algoritmos, indicando a seleção da melhor técnica de previsão e eliminando possíveis erros na parametrização das análises.

Por padrão, a funcionalidade *Amazon Forecast* testa 6 algoritmos: ARIMA, ETS, CNN-QR, NPTS, PFM e *Autoregressive Recurrent Neural Networks* (DeepAR). Por padrão, o modelo DeepAR exige maior frequência no horizonte de tempo das séries temporais, o que significa que somente avalia dados que estejam registrados em meses, dias ou horas, diferentemente dos dados em estudo que foram distribuídos em anos. Nesse caso, o método DeepAR não foi utilizado pela ferramenta para previsão de demanda.

A *Amazon Forecast* funciona em 3 etapas sequenciais, sendo importação de dados, treinamento de preditores e geração de previsões. As mesmas etapas foram reproduzidas duas vezes, uma para produção de vinhos com domínio para *Inventory* e outra para comercialização com domínio *Retail*. Para a primeira etapa, os dados foram armazenados na *Amazon Simple Storage Service* (*Amazon S3*) e importados para a *Amazon Forecast* para criação de um *dataset* de análise. Nesse momento, o domínio foi selecionado conforme a análise pretendida.

Na segunda etapa, gerou-se um preditor, onde o conjunto de dados foi automaticamente dividido em dois blocos, treinamento e teste, considerando uma entrada de dados e janela de teste determinada no ponto central da série temporal. Para seleção do melhor algoritmo, selecionou-se a função AutoML pelo enfoque em buscar uma solução no problema de negócio. A *Amazon Forecast* utiliza como critérios de seleção do melhor modelo, menores valores para *Weighted Quantile Loss* (wQL), WAPE e RMSE.

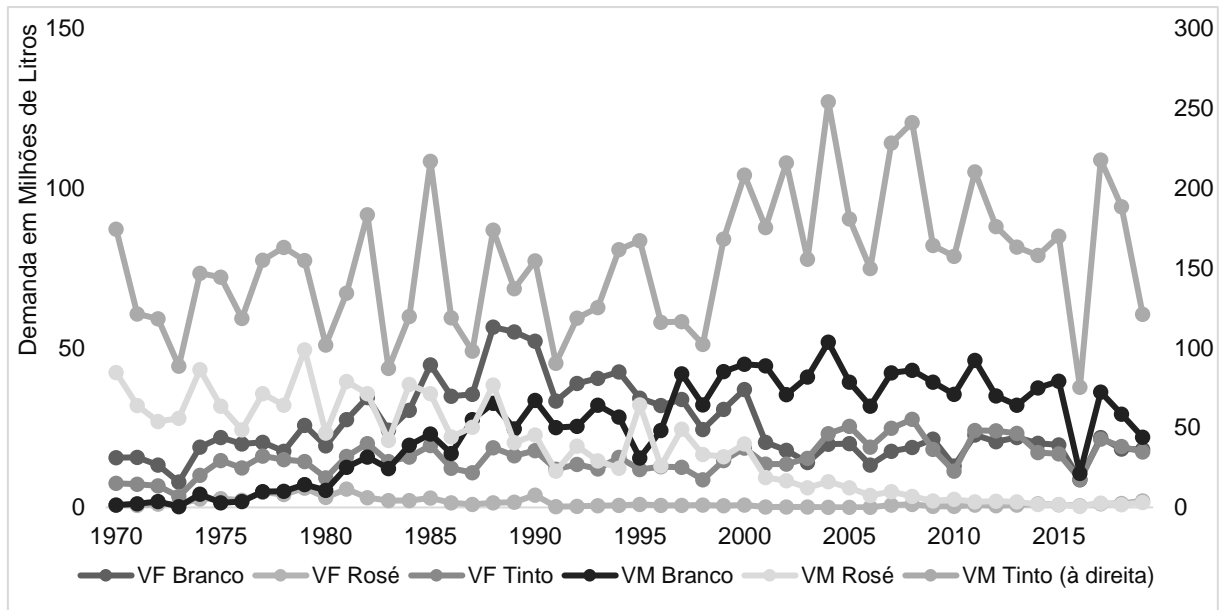
Na terceira etapa, gerou-se uma previsão de demanda para o ano seguinte, considerando quantis P10 (0,1), P50 (0,5) e P90 (0,9), o que garante um intervalo de confiança de 80%. Após, os resultados foram exportados para análise e a fase de implementação foi atingida com a descrição e a interpretação dos dados.

4 RESULTADOS

Nesta seção, apresentam-se os resultados do presente estudo, bem como o detalhamento das análises realizadas e dados obtidos ao longo do desenvolvimento. Complementarmente, imagens das etapas realizadas na *Amazon Forecast* são apresentadas no Anexo A.

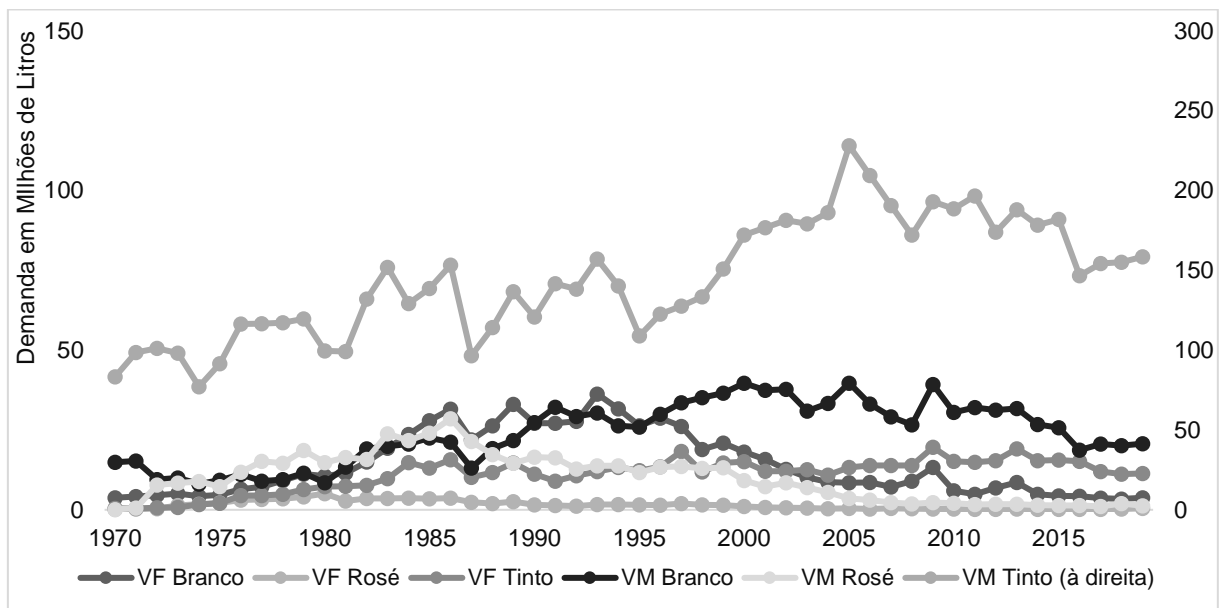
Inicialmente, os Gráficos 1 e 2 demonstram o comportamento das séries temporais de produção e comercialização de vinhos no Rio Grande do Sul. Todas as séries estão referenciadas na escala à esquerda, com exceção da série de vinho de mesa tinto, a qual deve ser interpretada pela escala à direita.

Gráfico 1 – Série Histórica de Produção de Vinhos no Rio Grande do Sul



Legenda: VF Vinho Fino, VM Vinho de Mesa. Fonte: Elaborado pela autora.

Gráfico 2 – Série Histórica de Comercialização de Vinhos no Rio Grande do Sul



Legenda: VF Vinho Fino VM Vinho de Mesa. Fonte: Elaborado pela autora.

As séries históricas de produção apresentaram tendência linear positiva para vinhos de mesa branco, tinto e vinho fino tinto, enquanto vinho de mesa rosé, vinhos finos branco e rosé apresentaram tendência linear negativa. As séries de comercialização apresentaram o mesmo padrão.

Nas Tabelas 1 e 2, são apresentados os valores de erro para o método base, elaborado de forma independente, e para cada modelo testado na *Amazon Forecast*

para as séries de produção e comercialização de vinhos, respectivamente. Após a análise do comportamento dos dados e ajuste dos parâmetros, a ferramenta de previsão dividiu os dados em 3 quantis e mediu a precisão dos valores por meio do indicador wQL.

Além disso, realizou o cálculo de WAPE e RMSE, elegendo como melhor algoritmo aquele que apresentou os menores valores. Nas Tabelas 1 e 2 é possível identificar que o método ARIMA demonstrou-se como o melhor algoritmo preditivo por atender aos critérios definidos, inclusive em comparação com o método base.

Tabela 1 – Performance de Modelos Preditivos Para Produção de Vinhos

	wQL [P10]	wQL [P50]	wQL [P90]	WAPE	RMSE
MMD	-	-	-	0,6106	33.169.952
ARIMA	0,0395	0,1925	0,1165	0,1925	9.733.162
CNN-QR	0,1013	0,2184	0,3364	0,2184	9.835.441
ETS	0,0388	0,2569	0,1269	0,2569	11.690.497
NPTS	0,0768	0,3108	0,1368	0,2890	13.515.722
PFM	0,1706	0,3485	0,1333	0,3454	14.553.528

Fonte: Elaborada pela autora.

Tabela 2 – Performance de Modelos Preditivos Para Comercialização de Vinhos

	wQL [P10]	wQL [P50]	wQL [P90]	WAPE	RMSE
MMD	-	-	-	0,2302	11.422.280
ARIMA	0,0641	0,1090	0,0205	0,1090	5.916.134
CNN-QR	0,0499	0,1146	0,2062	0,1146	5.679.216
ETS	0,0377	0,1339	0,0580	0,1339	8.142.851
NPTS	0,0914	0,1617	0,0577	0,1684	7.239.720
PFM	0,1488	0,2375	0,0757	0,2338	13.653.367

Fonte: Elaborada pela autora.

Na Tabela 3, são demonstrados os resultados de previsão de demanda para o período sequencial para produção e comercialização de vinhos, correspondente ao ano de 2020. Os dados apresentados correspondem aos resultados obtidos pelo método ARIMA, uma vez que o mesmo se demonstrou como o melhor modelo preditivo. Essa é uma característica inerente à *Amazon Forecast* que retorna ao pesquisador somente os resultados de previsão para o algoritmo vencedor.

Os números descritos para cada série nos quantis P10 e P90 indicam limite inferior e limite superior para as previsões demonstrando uma faixa de valores, ou seja, um intervalo de confiança de 80%. Os valores para P50 são a previsão mediana, o que significa que se espera que o valor real seja inferior ao número previsto em 50% das vezes. Os números negativos indicam uma demanda que tende a zero. Considerando os valores históricos e previstos em P50, observa-se que há uma tendência de queda para os volumes produzidos, ao contrário do que se pode observar para os volumes comercializados.

Tabela 3 – Previsão de Demanda de Produção e Comercialização de Vinhos Para Um Período

	Produção			Comercialização		
	P10	P50	P90	P10	P50	P90
Vinhos de Mesa (Milhões de Litros)						
Branco	1.71	12.54	23.37	15.36	20.66	25.96
Rosé	-11.66	-2.86	5.93	-2.16	1.27	4.69
Tinto	63.86	125.21	186.56	134.58	158.52	182.46
Vinhos Finos (Milhões de Litros)						
Branco	7.80	18.19	28.58	-0.78	3.75	8.27
Rosé	0.38	1.69	3.01	-0.20	0.50	1.21
Tinto	1.51	8.15	14.80	8.65	11.62	14.58

Fonte: Elaborada pela autora.

5 DISCUSSÃO

Nesta seção, os resultados obtidos com o desenvolvimento desse trabalho, cujo objetivo foi realizar a previsão de demanda para produção e comercialização de vinhos no Rio Grande do Sul, serão discutidos e interpretados com base na revisão bibliográfica realizada.

Observando a distribuição dos dados analisados, nota-se uma maior representatividade do vinho tinto frente aos demais tipos, tanto em produção, quanto em comercialização. Essa condição é diretamente influenciada pelo comportamento do consumo nacional (YAMAMOTO, 2008).

No presente estudo, a recomendação do método ARIMA entre os modelos testados para as séries temporais de produção e comercialização de vinhos se deu

com valores de erro de 10,9% e 19,2%, respectivamente. Em linhas gerais, valores de precisão escolhidos entre em estudos varia conforme característica dos dados e modelos elencados, mas sempre considerando o menor valor encontrado como o mais preciso. Taylor e Letham (2017) elegeram o melhor algoritmo de previsão de demanda em seu estudo com um valor de erro de aproximadamente 20%. Por outro lado, Lv et al. (2018) escolheram o método que apresentou erro de cerca de 5%.

Os resultados do presente estudo corroboram com os achados de Ye e Melkumian (2020), pesquisadores que consideraram ARIMA e SARIMA como melhores métodos preditivos para demanda de vinho tinto em comparação com métodos que utilizam redes neurais artificiais. Também, Paroissien (2020) utilizou princípios lógicos do método ARIMA em séries anuais e mensais para previsão do preço de barris de vinhos na França e recomendou a utilização das previsões como uma estratégia de negociação, destacando um potencial de aumento de 3,3% nos lucros. Por outro lado, há estudos que encontraram melhores resultados para algoritmos mais avançados para previsões financeiras e preço de *commodities* (CAO; WANG, 2019; LIN et al., 2020; WU et al., 2017, 2020; XU et al., 2020b; ZHANG; WU; WANG, 2020). Assim, percebe-se o potencial existente para gestores no conhecimento e a utilização adequada das técnicas de previsão de demanda.

Autores sugerem que modelos mais avançados de aprendizado contínuo são mais eficientes do que modelos convencionais nos casos de um número elevado de registros e observações (TIWARI; VAIBHAW, 2020). Entretanto, o volume e a frequência de dados existentes nas séries temporais de vitivinicultura gaúcha não caracterizam tal condição e assim pode-se entender essa característica como um indicativo para os resultados obtidos, uma vez que a ferramenta utilizada executa essa diferenciação para testar os algoritmos.

O objeto em estudo, um produto agrícola, tem seu comportamento fortemente influenciado por safras. Dessa forma, o problema de negócio analisado traz inerentemente consigo uma abertura em ciclos anuais. Embora não tenha sido o modelo recomendado, CNN-QR apresentou resultados de precisão muito próximos ao ARIMA, demonstrando sua natureza de operação linear. Diferentemente, ETS e NPTS apresentaram resultados mais imprecisos devido sua afinidade por dados sazonais e não-lineares. Em se tratando do modelo PFM, entende-se que o mesmo tenha uma melhor aplicação para dados de maior granularidade, devido sua lógica buscar uma valorização de períodos mais sensíveis como dias e feriados. Por fim, a diferença

observada entre os valores de precisão de MMD e ARIMA reforçam a aplicabilidade de modelos mais robustos, uma vez que o método base apresentou uma margem superior de imprecisão. Com isso, a utilização de ferramenta única que testa diferentes algoritmos e permite parametrização automática facilita a aplicação na solução de problemas de negócios.

Quanto aos dados de previsão gerados para o período sequencial, o período previsto nesta pesquisa foi o ano de 2020, para o qual até o presente momento estão disponíveis dados reais preliminares da produção de vinhos (OIV, 2021). O relatório da Organização Internacional do Vinho, indica uma queda na produção brasileira de 5% frente a 2019 e corrobora com a mesma tendência encontrada neste estudo.

Por outro lado, a Embrapa já divulgou dados atualizados para comercialização do produto no último ano (DE MELLO; MACHADO, 2020). Para esse caso, a demanda real foi maior que a prevista pelo método ARIMA em P50 para todos os tipos de vinhos, sendo cerca de 19% para vinhos de mesa branco, 10% para rosé, 20% para tinto e 36% para vinhos finos branco, 99% para rosé, 57% para tinto.

Nos últimos anos, a demanda por comercialização de vinhos tem se mantido estável, porém fortemente influenciada por condições climáticas favoráveis inesperadas, a safra de uva do ano passado apresentou elevados índices de qualidade e números recordes (ALVES; ZANUS; TONIETTO, 2020). Ainda, acredita-se que a atual pandemia tenha contribuído para um aumento de 18,4% no consumo de vinho pelos brasileiros no último ano, demonstrando-se uma atividade econômica estratégica no momento atual (OIV, 2021). Diante desse cenário, gestores e agricultores precisam estar preparados e amparados para atender às demandas de seus negócios, bem como estarem munidos de suporte para a tomada de decisão no âmbito da vitivinicultura.

6 CONSIDERAÇÕES FINAIS

Em suma, previsão de demanda traz em sua essência uma ferramenta que materializa direcionamentos empíricos e orienta estratégias conscientes. Nesse sentido, a escolha da melhor técnica de previsão é necessária para bons resultados e o papel do profissional que domina conceitos e técnicas é essencial para orientar e auxiliar as melhores escolhas para contextos aplicados.

Métodos convencionais e avançados possuem fortalezas e restrições sendo o principal desafio utilizar o método mais adequado para os dados no contexto em que se deseja prever a demanda. Em concordância com os resultados obtidos nesse estudo, os métodos convencionais se apresentaram como uma boa alternativa para análise de dados com frequência anual e com quantidade de registros inferiores a centenas. Por outro lado, técnicas avançadas podem ser úteis para um grande volume de dados como em séries com uma grande frequência de registros.

De maneira simples, pode-se obter a indicação do método ARIMA para as séries temporais de produção e comercialização de vinhos no Rio Grande do Sul, bem como a previsão de demanda para esses cenários. A abordagem utilizada nesse estudo é compatível com diferentes contextos, podendo ser utilizada para previsão de demanda de produtos que possuam registros em um horizonte de tempo. Os principais benefícios da metodologia utilizada consistem em consistência das etapas de análise, parametrização automática, agilidade e confiabilidade dos resultados obtidos.

O presente estudo apresenta limitações, entre elas a condição inerente do objeto de estudo que condiciona a frequência da base de dados estudada em anos, não permitindo uma visão mais ampla pelo teste do algoritmo DeepAR. Ainda, a utilização da ferramenta *Amazon Forecast* proporciona confiança nos resultados encontrados, porém fica-se limitado aos algoritmos pré-definidos da plataforma.

Ainda, recomenda-se para estudos futuros dois caminhos. Um deles seria a exploração da área de estudo buscando desenvolver um algoritmo específico para a vitivinicultura que permita a inclusão de mais variáveis que possam interferir nesse tipo de cultura, como por exemplo incentivos governamentais e eventos importantes. O outro seria explorar diferentes ferramentas para fins de comparação com a ferramenta utilizada nesse estudo. Enfim, o trabalho desenvolvido trouxe um olhar diferente para novas maneiras de executar previsões de demanda, desmistificando a ideia de que para fazer análises robustas são necessários conhecimento profundo de programação e elevado poder computacional.

REFERÊNCIAS

- ADHIKARI, R.; AGRAWAL, R. K. A Homogeneous Ensemble of Artificial Neural Networks for Time Series Forecasting. **International Journal of Computer Applications**, v. 32, n. 7, p. 1–8, 25 fev. 2013.
- AHMED, S. et al. FLF-LSTM: A novel prediction system using Forex Loss Function. **Applied Soft Computing Journal**, v. 97, p. 106780, 1 dez. 2020.
- AKTER, S.; RAHMAN, S. Agribusiness Forecasting with Univariate Time Series Modelling Techniques: the Case of a Dairy Cooperative in the UK. **Journal of Farm Management**, v. 13, n. 11, p. 747–764, 2010.
- ALMEIDA, A. N.; BRAGAGNOLO, C.; CHAGAS, A. L. S. Demanda por Vinho no Brasil: Elasticidades no consumo das famílias e determinantes da importação. **Revista de Economia e Sociologia Rural**, v. 53, n. 3, p. 433–454, 2015.
- ALVES, M. E. B.; ZANUS, M. C.; TONIETTO, J. Condições meteorológicas e sua influência na safra vitícola de 2020 em regiões produtoras de vinhos finos do Sul do Brasil. **Embrapa**, 2020.
- AMAZON WEB SERVICES. **Amazon Forecast Developer Guide**. [s.l: s.n.]. Disponível em: <<https://docs.aws.amazon.com/forecast/latest/dg/forecast.dg.pdf>>. Acesso em: 29 maio. 2021.
- ASHA, J. et al. Analysis of temperature prediction using random forest and facebook prophet algorithms. In: **Lecture Notes on Data Engineering and Communications Technologies**. [s.l.] Springer, 2020. v. 46p. 432–439.
- BAIDAKOV, A. N. et al. Methodical Bases for Developing Predictive Scenarios of Agribusiness. **Asian Social Science**, v. 11, n. 8, 2015.
- BANCO CENTRAL. **Sistema Gerenciador de Séries Temporais**. Disponível em: <<https://www4.bcb.gov.br/pec/series/port/aviso.asp?frame=1>>. Acesso em: 29 maio. 2021.
- CAO, J.; WANG, J. Stock price forecasting model based on modified convolution neural network and financial time series analysis. **International Journal of Communication Systems**, v. 32, n. 12, p. e3987, 1 ago. 2019.
- CAUCHICK-MIGUEL, P. **Metodologia de Pesquisa em Engenharia de Produção e Gestão de Operações**. 3. ed. Rio de Janeiro: [s.n.].
- CHAN, W. N. **Time Series Data Mining: Comparative Study of ARIMA and Prophet Methods for Forecasting Closing Prices of Myanmar Stock Exchange**. **Journal of Computer Applications and Research**. [s.l: s.n.].
- CHEN, C.-H.; KUNG, H.-Y.; HWANG, F.-J. Deep Learning Techniques for Agronomy Applications. **Agronomy**, v. 9, n. 3, p. 142, 20 mar. 2019.
- CHIPANSHI, A. C.; RIPLEY, E. A.; LAWFORDE, R. G. Early prediction of spring wheat

yields in saskatchewan from current and historical weather data using the ceres-wheat model. **Agricultural and Forest Meteorology**, v. 84, n. 3–4, p. 223–232, 1 abr. 1997.

CHU, X. et al. An optimized hybrid model based on artificial intelligence for grape price forecasting. **British Food Journal**, v. 121, n. 12, p. 3247–3265, 21 nov. 2019.

CULP, J. **Using Computer Science in Agribusiness**. New York: Rosen Publishing, 2019.

CZUPRYNA, M. et al. Agent-based modelling of viticulture development in emerging markets: The case of the Małopolska region. **JASSS**, v. 21, n. 3, 30 jun. 2018.

DE MELLO, L. M. R.; MACHADO, C. A. E. **Banco de Dados de Uva, Vinho e Derivados Embrapa Uva e Vinho**. [s.l.: s.n.]. Disponível em: <http://vitibrasil.cnpqv.embrapa.br/index.php?opcao=opt_01>. Acesso em: 29 nov. 2020.

DE OLIVEIRA, J. F. L. et al. A hybrid optimized error correction system for time series forecasting. **Applied Soft Computing Journal**, v. 87, p. 105970, 1 fev. 2020.

DOU, J.; LIU, C.; WANG, B. **Short-term Wind Power Forecasting Based on Convolutional Neural Networks** IOP Conference Series: Earth and Environmental Science. **Anais...**Institute of Physics Publishing, 17 jul. 2018 Disponível em: <<https://iopscience.iop.org/article/10.1088/1755-1315/170/4/042023>>. Acesso em: 21 nov. 2020

DRESCH, A.; LACERDA, D. P.; ANTUNES JÚNIOR, J. A. V. **Design Science Research: Método de Pesquisa para Avanço da Ciência e Tecnologia**. [s.l.: s.n.].

FANG, W. X. et al. **Combine Facebook Prophet and LSTM with BPNN Forecasting financial markets: The Morgan Taiwan Index** Proceedings - 2019 International Symposium on Intelligent Signal Processing and Communication Systems, ISPACS 2019. **Anais...**Institute of Electrical and Electronics Engineers Inc., 1 dez. 2019

FARIAS, C. V. S.; TATSCH, A. L. The Brazilian wine industry: A case study on geographical proximity and innovation dynamics. **Revista de Economia e Sociologia Rural**, v. 52, n. 3, p. 515–532, 2014.

FISHER, D. K. et al. Understanding technology adoption through system dynamics modeling: Implications for agribusiness management. **International Food and Agribusiness Management Review**, v. 3, n. 3, p. 281–296, 1 set. 2000.

FRITZ, B. A. et al. Deep-learning model for predicting 30-day postoperative mortality. **British Journal of Anaesthesia**, v. 123, n. 5, p. 688–695, 1 nov. 2019.

GAUTAM, A.; SINGH, V. **Parametric versus non-parametric time series forecasting methods: A review** **Journal of Engineering Science and Technology Review** Eastern Macedonia and Thrace Institute of Technology, , 2020.

GENG, Y. et al. Seismic Events Prediction Using Deep Temporal Convolution

Networks. **Journal of Electrical and Computer Engineering**, v. 2019, 2019.

GHIMIRE, S. et al. Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. **Applied Energy**, v. 253, p. 113541, 1 nov. 2019.

GONZÁLEZ MATA, A. **A comparison between LSTM and Facebook Prophet models: a financial forecasting case study**. [s.l.] Universitat Oberta de Catalunya (UOC), 2020. Disponível em: <<http://openaccess.uoc.edu/webapps/o2/handle/10609/107367>>. Acesso em: 22 nov. 2020.

GUO, K. et al. Optimized Graph Convolution Recurrent Neural Network for Traffic Prediction. **IEEE Transactions on Intelligent Transportation Systems**, p. 1–12, 14 jan. 2020.

HUA, Y. et al. Deep Learning with Long Short-Term Memory for Time Series Prediction. **IEEE Communications Magazine**, v. 57, n. 6, p. 114–119, 1 jun. 2019.

IBGE. **Tabela 5457: Área plantada ou destinada à colheita, área colhida, quantidade produzida, rendimento médio e valor da produção das lavouras temporárias e permanentes**. [s.l.: s.n.]. Disponível em: <<https://sidra.ibge.gov.br/tabela/5457#resultado>>. Acesso em: 2 nov. 2020.

IBRAVIN. **Comercialização de Vinhos**. [s.l.: s.n.]. Disponível em: <<https://www.ibravin.org.br/admin/arquivos/estatisticas/1561753574.pdf>>. Acesso em: 2 nov. 2020.

IFFT, J.; KUHNS, R.; PATRICK, K. Can machine learning improve prediction-an application with farm survey data International Food and Agribusiness Management Review. **International Food and Agribusiness Management Review**, v. 21, n. 8, p. 1083–1098, 2018.

INSTITUTO NACIONAL DE METEOROLOGIA. **Dados Históricos Anuais**. Disponível em: <<https://portal.inmet.gov.br/dadoshistoricos>>. Acesso em: 29 maio. 2021.

JHA, G. K.; SINHA, K. **Agricultural Price Forecasting Using Neural Network Model: An Innovative Information Delivery System** **Agricultural Economics Research Review**. [s.l.: s.n.]. Disponível em: <<https://ageconsearch.umn.edu/record/162150>>. Acesso em: 2 nov. 2020.

JU, Y. et al. A model combining convolutional neural network and lightgbm algorithm for ultra-short-term wind power forecasting. **IEEE Access**, v. 7, p. 28309–28318, 2019.

KANG, T. et al. Forecasting of Power Demands Using Deep Learning. **Applied Sciences**, v. 10, n. 20, p. 7241, 16 out. 2020.

KERKKÄNEN, A.; KORPELA, J.; HUISKONEN, J. Demand forecasting errors in industrial context: Measurement and impacts. **International Journal of Production Economics**, v. 118, n. 1, p. 43–48, 2009.

KHAN, M. A. et al. Effective Demand Forecasting Model Using Business Intelligence Empowered with Machine Learning. **IEEE Access**, v. 8, p. 116013–116023, 2020.

KHANDELWAL, I.; SATIJA, U.; ADHIKARI, R. **Forecasting seasonal time series with Functional Link Artificial Neural Network** 2nd International Conference on Signal Processing and Integrated Networks, SPIN 2015. **Anais...**Institute of Electrical and Electronics Engineers Inc., 24 abr. 2015

KIM, J. et al. Recurrent inception convolution neural network for multi short-term load forecasting. **Energy and Buildings**, v. 194, p. 328–341, 1 jul. 2019.

KOLTES, J. E. et al. **A Vision for Development and Utilization of High-Throughput Phenotyping and Big Data Analytics in Livestock** **Frontiers in Genetics** Frontiers Media S.A., , 17 dez. 2019. Disponível em: </pmc/articles/PMC6934059/?report=abstract>. Acesso em: 2 nov. 2020

KUREMOTO, T. et al. Training Deep Neural Networks with Reinforcement Learning for Time Series Forecasting. In: **Time Series Analysis - Data, Methods, and Applications**. [s.l.] IntechOpen, 2019.

LEE, M.-S.; AHN, H. A Time Series Graph based Convolutional Neural Network Model for Effective Input Variable Pattern Learning : Application to the Prediction of Stock Market. **Journal of Intelligence and Information Systems**, v. 24, n. 1, p. 167–181, 2018.

LIN, H. et al. Prediction model for stock price trend based on convolution neural network. **Journal of Intelligent & Fuzzy Systems**, v. 39, n. 4, p. 4999–5008, 10 jul. 2020.

LV, Z. et al. **LC-RNN: A deep learning model for traffic speed prediction** IJCAI International Joint Conference on Artificial Intelligence. **Anais...**International Joint Conferences on Artificial Intelligence, 2018

MADHURI, C. R.; CHINTA, M.; KUMAR, V. V. N. V. P. **Stock market prediction for time-series forecasting using prophet upon ARIMA** 2020 7th International Conference on Smart Structures and Systems, ICSSS 2020. **Anais...**Institute of Electrical and Electronics Engineers Inc., 1 jul. 2020

MARETH, T. et al. **PLANEJAMENTO E CONTROLE DA PRODUÇÃO DE VINHO UTILIZANDO MODELOS DE PREVISÃO DE DEMANDA E PROGRAMAÇÃO MATEMÁTICA**. Curitiba: [s.n.]. Disponível em: <<https://anaiscbc.emnuvens.com.br/anais/article/view/350>>. Acesso em: 18 out. 2020.

MELLO, L. M. R. DE; MACHADO, C. A. E. Comunicado Técnico Vitivinicultura Brasileira: Panorama 2019. **Embrapa**, 2019.

MELNYK, K. **The technology of short-term planning for resolving the problems with high level of uncertainty on an enterprise**. [s.l.: s.n.].

MOREIRA, D. A. **Administração da Produção e Operações**. 2. ed. São Paulo: [s.n.].

MUJEEB, S. et al. Wind Power Forecasting Based on Efficient Deep Convolution Neural Networks. In: **Lecture Notes in Networks and Systems**. [s.l.] Springer, 2020. v. 96p. 47–56.

NAVRATIL, M.; KOLKOVA, A. Decomposition and forecasting time series in business economy using prophet forecasting model. **Central European Business Review**, v. 8, n. 4, p. 26–39, 2019.

OIV. **State of the World Vitivinicultural Sector in 2019**. [s.l.: s.n.]. Disponível em: <<http://www.oiv.int/public/medias/7298/oiv-state-of-the-vitivinicultural-sector-in-2019.pdf>>. Acesso em: 2 nov. 2020.

OIV. **State of the World Vitivinicultural Sector in 2020**. Paris: [s.n.].

OO, Z. Z.; PHYU, S. **Microclimate Prediction Using Cloud Centric Model Based on IoT Technology for Sustainable Agriculture**. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/8821705>>. Acesso em: 22 nov. 2020.

PAPACHARALAMPOUS, G.; TYRALIS, H.; KOUTSOYIANNIS, D. Predictability of monthly temperature and precipitation using automatic time series forecasting methods. **Acta Geophysica**, v. 66, n. 4, p. 807–831, 1 ago. 2018.

PAPASTEFANOPOULOS, V.; LINARDATOS, P.; KOTSIANTIS, S. COVID-19: A comparison of time series methods to forecast percentage of active cases per population. **Applied Sciences (Switzerland)**, v. 10, n. 11, p. 3880, 1 jun. 2020.

PARMEZAN, A. R. S.; SOUZA, V. M. A.; BATISTA, G. E. A. P. A. Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model. **Information Sciences**, v. 484, p. 302–337, 2019.

PAROISSIEN, E. Forecasting bulk prices of Bordeaux wines using leading indicators. **International Journal of Forecasting**, v. 36, n. 2, p. 292–309, 1 abr. 2020.

PEIREIRA, G. E. et al. Panorama da produção e mercado nacional de vinhos espumantes. **Informe Agropecuário**, v. 41, n. 310, p. 7–18, 2020.

QI, Y. et al. A hybrid model for spatiotemporal forecasting of PM 2.5 based on graph convolutional neural network and long short-term memory. **Science of the Total Environment**, v. 664, p. 1–10, 10 maio 2019.

RIBEIRO, C. O.; OLIVEIRA, S. M. A hybrid commodity price-forecasting model applied to the sugar-alcohol sector. **Australian Journal of Agricultural and Resource Economics**, v. 55, n. 2, p. 180–198, 1 abr. 2011.

ROMANOV, A.; YARUSHKINA, N.; FILIPPOV, A. Application of Time Series Analysis and Forecasting Methods for Enterprise Decision-Management. In: [s.l.] Springer, Cham, 2020. p. 326–337.

SAMAL, K. K. R. et al. **Time series based air pollution forecasting using SARIMA and prophet model** ACM International Conference Proceeding Series. **Anais...** New York, New York, USA: Association for Computing Machinery, 16 ago. 2019 Disponível

em: <<http://dl.acm.org/citation.cfm?doid=3355402.3355417>>. Acesso em: 22 nov. 2020

SCHUTTE, M. **Time Series Analysis of U.S. Wine Production by Region An Honors Thesis**. [s.l: s.n.]. Disponível em: <<http://cardinalscholar.bsu.edu/handle/123456789/201551>>. Acesso em: 4 jun. 2021.

SCHWARENTHORER, P. et al. **INCREASED COMPANY PERFORMANCE THROUGH MACROECONOMICS SALES FORECASTING: A CASE STUDY**The **Journal of Japanese Operations Management and Strategy**. [s.l.] Japanese Operations Management and Strategy Association, 2020.

SEN, R.; YU, H.-F.; DHILLON, I. **Think Globally, Act Locally: A Deep Neural Network Approach to High-Dimensional Time Series Forecasting**Advances in **Neural Information Processing Systems**. [s.l: s.n.].

SHEARER, C. The CRISP-DM Model: The New Blueprint for Data Mining. **Journal of Data Warehousing**, v. 5, n. 4, p. 13–22, 2000.

SHEN, J.; VALAGOLAM, D.; MCCALLA, S. Prophet forecasting model: A machine learning approach to predict the concentration of air pollutants (PM2.5, PM10, O3, NO2, SO2, CO) in Seoul, South Korea. **PeerJ**, v. 8, p. e9961, 15 set. 2020.

SINGH, A. Air Pollution Forecasting and Performance Evaluation Using Advanced Time Series and Deep Learning Approach for Gurgaon. 2019.

TA, X.; WEI, Y. Research on a dissolved oxygen prediction method for recirculating aquaculture systems based on a convolution neural network. **Computers and Electronics in Agriculture**, v. 145, p. 302–310, 1 fev. 2018.

TAYLOR, S. J.; LETHAM, B. Forecasting at Scale. **PeerJ Preprints**, 27 set. 2017.

TEALAB, A. Time series forecasting using artificial neural networks methodologies: A systematic review. **Future Computing and Informatics Journal**, v. 3, n. 2, p. 334–340, 1 dez. 2018.

TEJASVINI, K. N. et al. **Air pollution forecasting using multiple time series approach**Advances in Intelligent Systems and Computing. **Anais...**Springer, 2020Disponível em: <https://link.springer.com/chapter/10.1007/978-981-15-2188-1_8>. Acesso em: 22 nov. 2020

THOMAZI GONZALEZ, R. **Using Deep Learning and Evolutionary Algorithms for Time Series Forecasting**. [s.l: s.n.]. Disponível em: <<https://lume.ufrgs.br/handle/10183/189125>>. Acesso em: 8 nov. 2020.

TIWARI, A.; VAIBHAW, V. **Comparative Study on Time Series Forecasting using Deep Learning Models**. [s.l: s.n.]. Disponível em: <https://www.researchgate.net/publication/339771117_Comparative_Study_on_Time_Series_Forecasting_using_Deep_learning_Models>. Acesso em: 9 nov. 2020.

VERMA, R.; SHARMA, J.; JINDAL, S. **Time Series Forecasting Using Machine Learning**Communications in Computer and Information Science. **Anais...**Springer,

2020

VISHWAKARMA, A. et al. Stock Price Prediction Using Sarima and Prophet Machine Learning Model. **ISSN (Print) ABCD-PQRS International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)**, v. 9, n. 1, 2020.

VISHWAS, B. V et al. Prophet. In: **Hands-on Time Series Analysis with Python**. [s.l.] Apress, 2020. p. 375–394.

WAN, R. et al. Multivariate Temporal Convolutional Network: A Deep Neural Networks Approach for Multivariate Time Series Forecasting. **Electronics**, v. 8, n. 8, p. 876, 7 ago. 2019.

WANG, J. et al. Artificial bee colony-based combination approach to forecasting agricultural commodity prices. **International Journal of Forecasting**, 12 dez. 2019.

WANG, L. et al. Optimal Forecast Combination Based on Neural Networks for Time Series Forecasting. **Applied Soft Computing Journal**, v. 66, p. 1–17, 1 maio 2018.

WEYTJENS, H.; LOHMANN, E.; KLEINSTEUBER, M. Cash flow prediction: MLP and LSTM compared to ARIMA and Prophet. **Electronic Commerce Research**, p. 1–21, 19 jul. 2019.

WU, H. et al. A new method of large-scale short-term forecasting of agricultural commodity prices: illustrated by the case of agricultural markets in Beijing. **Journal of Big Data**, v. 4, n. 1, p. 1, 1 dez. 2017.

WU, J. M. T. et al. **A new convolution neural network model for stock price prediction** Advances in Intelligent Systems and Computing. **Anais...**Springer, 1 nov. 2020Disponível em: <https://link.springer.com/chapter/10.1007/978-981-15-3308-2_64>. Acesso em: 21 nov. 2020

XU, L. et al. Prediction of sea surface temperature using a multiscale deep combination neural network. **Remote Sensing Letters**, v. 11, n. 7, p. 611–619, 2 jul. 2020a.

XU, Z. et al. Prediction research of financial time series based on deep learning. **Soft Computing**, v. 24, n. 11, p. 8295–8312, 1 jun. 2020b.

YAMAMOTO, C. H. **A demanda por bebidas alcoólicas no Brasil: 2008/2009**. [s.l.: s.n.]. Disponível em: <<http://bibliotecadigital.fgv.br/dspace/handle/10438/8612>>. Acesso em: 18 out. 2020.

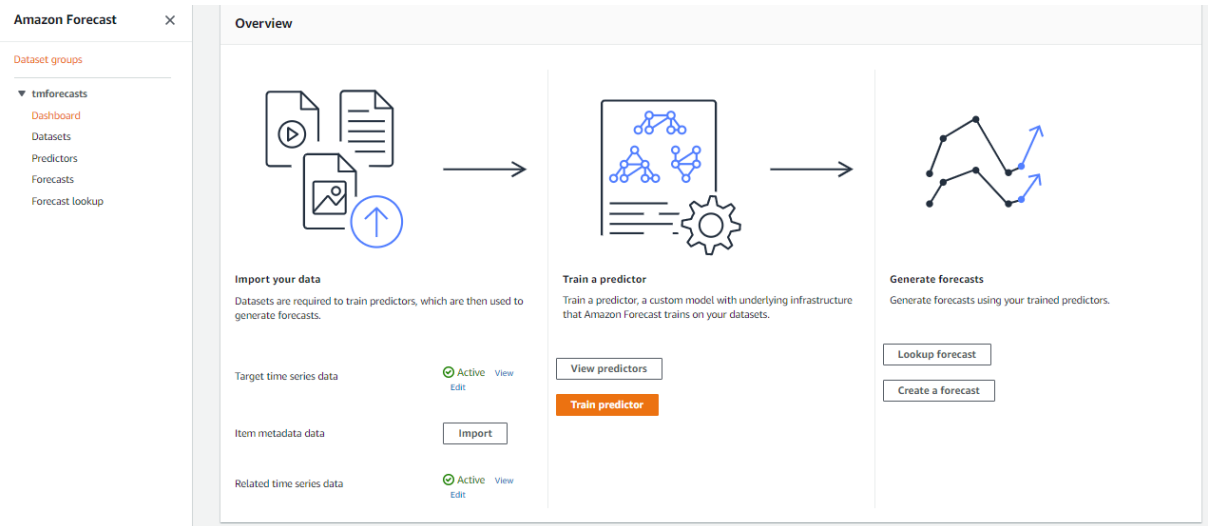
YE, C.; LI, K.; JIA, G. Z. **A new red wine prediction framework using machine learning** Journal of Physics: Conference Series. **Anais...**IOP Publishing Ltd, 30 nov. 2020Disponível em: <<https://iopscience.iop.org/article/10.1088/1742-6596/1684/1/012067>>. Acesso em: 4 jun. 2021

YE, W.; MELKUMIAN, A. V. **Forecasting Australian Red Wine Sales with SARIMA and ANNs**. [s.l.: s.n.]. Disponível em: <www.robjhyndman.com/TSDL>. Acesso em: 2 nov. 2020.

- YENIDOGAN, I. et al. **Bitcoin Forecasting Using ARIMA and PROPHETUBMK 2018 - 3rd International Conference on Computer Science and Engineering. Anais...**Institute of Electrical and Electronics Engineers Inc., 6 dez. 2018
- YING, Z. et al. Time Series Forecasting Model Based on LSTM-Prophet Nonlinear Combination. **Computer and Modernization**, v. 0, n. 09, p. 6, 24 set. 2020.
- YOO, T.-W.; OH, I.-S. Time Series Forecasting of Agricultural Products' Sales Volumes Based on Seasonal Long Short-Term Memory. **Applied Sciences**, v. 10, n. 22, p. 8169, 18 nov. 2020.
- YUAN, C. Z.; SAN, W. W.; LEONG, T. W. **Determining Optimal Lag Time Selection Function with Novel Machine Learning Strategies for Better Agricultural Commodity Prices Forecasting in Malaysia**ACM International Conference Proceeding Series. **Anais...**New York, NY, USA: Association for Computing Machinery, 12 ago. 2020Disponível em: <<https://dl.acm.org/doi/10.1145/3417473.3417480>>. Acesso em: 8 nov. 2020
- YURTSEVER, M.; TECIM, V. Forecasting Meal Requirements Using Time Series Methods in Organization. In: [s.l.] Springer, Cham, 2020. p. 243–254.
- ZHANG, P. G. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, v. 50, p. 159–175, 1 jan. 2003.
- ZHANG, R.; WU, Z.; WANG, S. **Prediction of Stock Based on Convolution Neural Network**Proceedings of the 32nd Chinese Control and Decision Conference, CCDC 2020. **Anais...**Institute of Electrical and Electronics Engineers Inc., 1 ago. 2020
- ZHAO, J.; ZHANG, C. **Research on sales forecast based on prophet-SARIMA combination model**Journal of Physics: Conference Series. **Anais...**Institute of Physics Publishing, 21 ago. 2020Disponível em: <<https://iopscience.iop.org/article/10.1088/1742-6596/1616/1/012069>>. Acesso em: 22 nov. 2020
- ZHAO, N. et al. Day-of-week and seasonal patterns of PM2.5 concentrations over the United States: Time-series analyses using the Prophet procedure. **Atmospheric Environment**, v. 192, p. 116–127, 1 nov. 2018.
- ZHAO, Y. et al. **Urban crowd flow forecasting based on cellular network**ACM International Conference Proceeding Series. **Anais...**New York, New York, USA: Association for Computing Machinery, 17 maio 2019Disponível em: <<http://dl.acm.org/citation.cfm?doid=3321408.3321579>>. Acesso em: 21 nov. 2020
- ZUNIC, E. et al. Application of Facebook's Prophet Algorithm for Successful Sales Forecasting Based on Real-world Data. **International Journal of Computer Science and Information Technology**, v. 12, n. 2, p. 23–36, 7 maio 2020.

ANEXO A – DETALHAMENTO DAS ANÁLISES

Figura 2 – Quadro de Construção de Previsão de Demanda de Comercialização de Vinhos



Fonte: Amazon Forecast.

Figura 3 – Séries de Dados Importadas de Previsão de Demanda de Comercialização de Vinhos

The screenshot shows the 'Datasets' page in Amazon Forecast. It displays a table of imported datasets and a section for tags.

Datasets (3) info						
	Dataset name	Dataset type	Status	Latest import status	Domain	Date created
<input type="radio"/>	tmforecasts_dataset	TARGET_TIME_SERIES	Active	Active	RETAIL	Fri, 04 Jun 2021 17:47:56 GMT
<input type="radio"/>	-	ITEM_METADATA	Not created	Not uploaded	-	-
<input type="radio"/>	tmforecasts_dataset2	RELATED_TIME_SERIES	Active	Active	RETAIL	Fri, 04 Jun 2021 17:51:59 GMT

Below the table, there is a 'Tags (0) info' section with a 'Manage tags' button. The tag section is currently empty, showing 'Empty tag' and 'No tags to display'.

Fonte: Amazon Forecast.

Figura 4 – Resultados de Treinamento e Testes de Algoritmos Para Previsão de Comercialização de Vinhos

The figure displays three screenshots of the Amazon Forecast console interface, showing the results of training and testing algorithms for wine sales prediction.

Screenshot 1: Predictor Overview

Predictor overview

WAPE 0.1090	AutoML True	Perform HPO -
Training status Active	AutoML optimization strategy Default	Date created Fri, 04 Jun 2021 18:13:34 GMT
Predictor ARN arn:aws:forecast:us-east-2:474482748348:predictor/tmpredictors	Winning autoML algorithm ARIMA	Total creation time 45 mins
Estimated time remaining -	Failed algorithm messages Refer to failed algorithm messages below	

Forecast configurations

Forecast horizon 1	Forecast dimensions -	Country for holidays -
Forecast frequency Y		Weather index enabled False

Screenshot 2: Accuracy Metrics

Accuracy metrics

Evaluation type	Test window start	Test window end	Items	wQL[0.1]	wQL[0.5]	wQL[0.9]	WAPE	RMSE
Backtest window	Mon, 01 Jan 1996 00:00:00 GMT	Wed, 01 Jan 1997 00:00:00 GMT	6	0.0641	0.1090	0.0205	0.1090	5916133.9632

Predictor backtest exports (1)

Export name	Status	Location	Date created
tmpredictors_backtest	Active	s3://tm-database-wine	Mon, 07 Jun 2021 11:42:46 GMT

Evaluation parameters

Number of backtest windows 1	Backtest window offset 24
---------------------------------	------------------------------

Screenshot 3: Failed Algorithm Messages

Failed algorithm messages

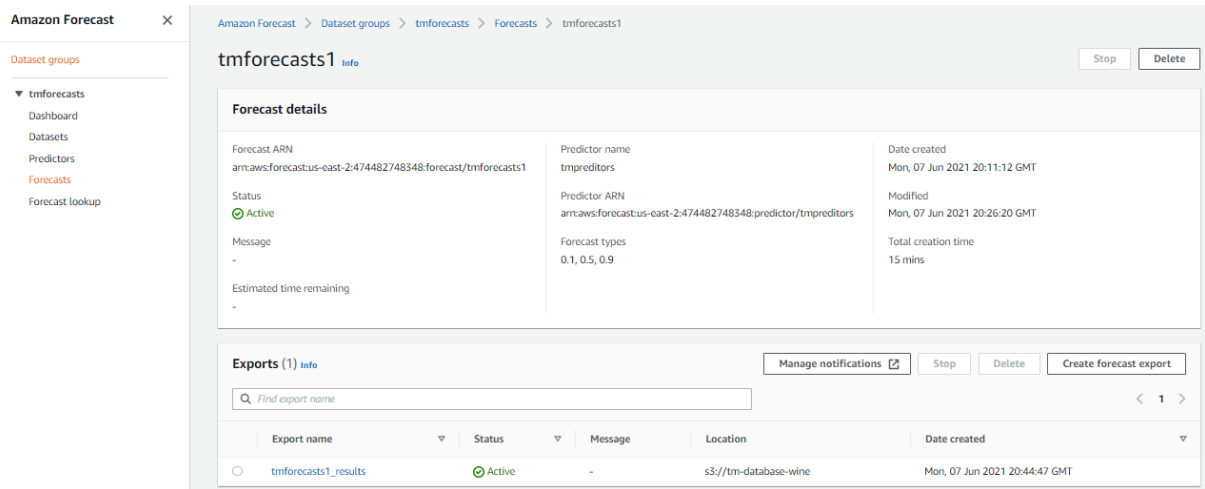
Test window start Mon, 01 Jan 1996 00:00:00 GMT	Test window end Wed, 01 Jan 1997 00:00:00 GMT
Status Create failed	Message Unknown HPO tuning job status: Failed. AlgorithmError: Customer Error: time_freq set to Y but should be one of min, H, D, W, M, optionally prefixed by an integer

Tags (0)

Key	Value
Empty tag No tags to display	

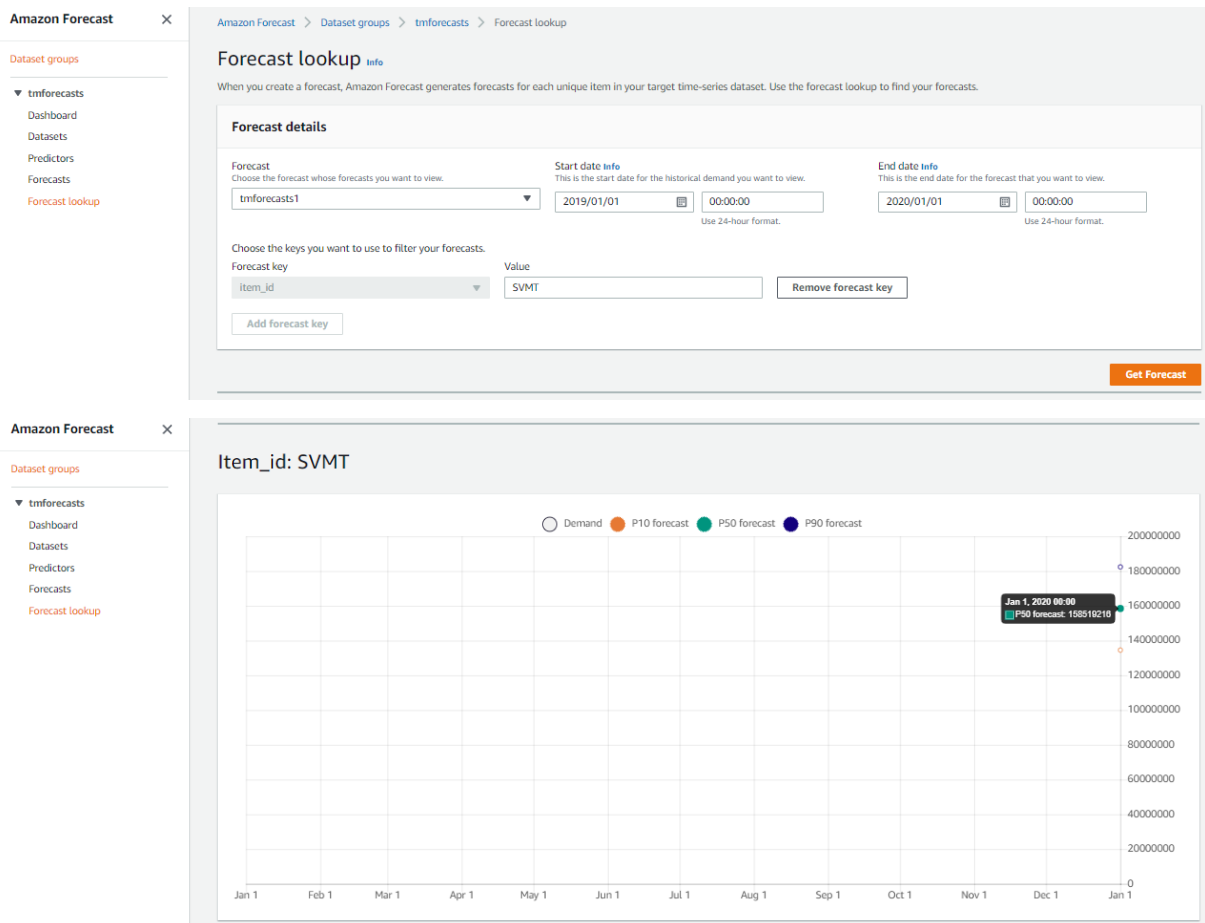
Fonte: Amazon Forecast.

Figura 5 – Resultado da Previsão de Demanda Para Comercialização de Vinhos



Fonte: Amazon Forecast.

Figura 6 – Valores Previstos Para a Série de Comercialização de Vinho de Mesa Tinto



Fonte: Amazon Forecast.