



UNIVERSIDADE PRESBITERIANA MACKENZIE
FACULDADE DE COMPUTAÇÃO E INFORMÁTICA
CURSO DE SISTEMAS DE INFORMAÇÃO



OTIMIZAÇÃO DE VENDAS DE CAFÉ COM IA

EDUARDO BISPO FELIZARDO | 10400894@mackenzista.com.br
GUSTAVO AUGUSTO ALVES PIVATTO | 10401400@mackenzista.com.br
JEFFERSON DA CONCEICAO DANTAS | 10401327@mackenzista.com.br
MARINA MIKI SINZATO | 10401880@mackenzista.com.br

SÃO PAULO
2025

1. RESUMO

Este projeto propõe a utilização de técnicas de Machine Learning para prever a demanda de vendas de café em um determinado período. A previsão será baseada em uma base de dados contendo informações sobre as vendas diárias, incluindo data, horário, tipo de pagamento e nome do produto. O objetivo é auxiliar no planejamento de estoque e insumos, reduzindo desperdícios e otimizando o atendimento ao cliente.

2. INTRODUÇÃO

Empresas do setor alimentício enfrentam constantemente desafios relacionados ao planejamento de estoque e à gestão da demanda. A ausência de previsões precisas pode ocasionar o excesso de produtos armazenados, gerando desperdícios, ou a escassez de insumos em períodos de alta demanda, comprometendo o atendimento ao cliente e a eficiência operacional. Nesse contexto, o uso de modelos preditivos têm se mostrado uma alternativa eficaz para mitigar esses problemas, contribuindo para a tomada de decisões mais assertivas quanto às compras e ao controle de estoque, o que, por sua vez, resulta na redução de custos e no aumento da satisfação dos consumidores.

Um caso prático envolve um estabelecimento que necessita adquirir insumos para suas máquinas de café. A reposição inadequada desses insumos pode comprometer o fornecimento do produto, além de gerar custos desnecessários com estoques excessivos. Com a utilização de técnicas de *Machine Learning*, é possível prever e analisar o histórico de compras e consumo, otimizando a reposição de materiais e promovendo uma gestão mais eficiente e econômica dos recursos.

Dessa forma, este projeto tem como objetivo desenvolver um modelo preditivo baseado em séries temporais para prever as vendas de café, com o intuito de proporcionar uma gestão mais eficiente dos insumos e aprimorar o planejamento estratégico da empresa. A previsão da demanda de café se justifica por se tratar de um produto de consumo recorrente e sensível a variações sazonais, comportamentais e econômicas.

A abordagem adotada envolve a utilização de um framework de *Machine Learning*, com a aplicação de modelos específicos para séries temporais, como AutoARIMA, Regressão com XGBoost, Holt-Winters (Suavização Exponencial) e Média Móvel. Esses métodos serão comparados quanto à sua acurácia e aplicabilidade, de modo a identificar a solução mais adequada para o contexto da empresa analisada.

3. ÉTICA NO USO DE INTELIGÊNCIA ARTIFICIAL

A Inteligência Artificial (IA) tem sido amplamente utilizada em diversos setores, proporcionando avanços significativos. No entanto, sua aplicação levanta questões éticas fundamentais, como a transparência dos algoritmos, a privacidade dos dados e a equidade no acesso à tecnologia. Segundo Souza et al. (2022), a falta de transparência nos modelos de IA pode resultar em decisões enviesadas, afetando negativamente determinados grupos sociais. Além disso, a proteção da privacidade é um fator crucial, pois muitos sistemas de IA processam grandes volumes de dados sensíveis, exigindo regulamentações rigorosas para evitar usos inadequados (SANTOS; FERREIRA, 2023).

No contexto educacional, a IA pode ser utilizada para personalizar o ensino e melhorar o desempenho dos alunos. No entanto, segundo Silva (2021), sua implementação deve ser feita com cautela para evitar impactos negativos, como a exclusão digital de alunos sem acesso a recursos tecnológicos adequados. Dessa forma, a adoção de políticas que garantam a inclusão e a equidade no uso da IA torna-se essencial para que seus benefícios sejam amplamente aproveitados sem comprometer valores éticos fundamentais.

Portanto, o desenvolvimento e a aplicação da IA devem ser pautados por princípios éticos sólidos, garantindo que a tecnologia seja utilizada de maneira responsável e justa. O avanço da IA deve estar alinhado com regulamentações que protejam os usuários e garantam sua segurança, promovendo um equilíbrio entre inovação e responsabilidade social.

4. DESENVOLVIMENTO

4.1. DATASET

Coffee Sales, um Dataset de vendas de uma máquina de café em um estabelecimento, contendo informações sobre vendas de 2024 até 2025 e

campos como data, tipo de pagamento, valor, e nome do café. Com esse dataset, é possível analisar padrões de compra, tendências de vendas e preferências dos clientes relacionadas a produtos de café.

4.2. METODOLOGIA

Para o desenvolvimento deste estudo, foram seguidas as seguintes etapas metodológicas:

4.2.1. COLETA DE DADOS

Inicialmente, foi obtido um banco de dados contendo registros históricos das vendas diárias de café. As informações coletadas incluem a data e o horário das transações, o tipo de pagamento utilizado e o nome do produto vendido. Esses dados serviram de base para a construção das séries temporais por tipo de café.

4.2.2. PRÉ-PROCESSAMENTO DOS DADOS

Os dados foram submetidos a um processo de limpeza e tratamento, visando garantir consistência e qualidade. Foram removidos registros duplicados ou incompletos, e realizada a padronização dos campos de data e hora para um formato apropriado à análise temporal. Além disso, as vendas foram agregadas por mês, a fim de criar uma base coerente com o volume de dados disponíveis.

4.2.3. SELEÇÃO E APLICAÇÃO DE MODELOS DE INTELIGÊNCIA ARTIFICIAL

Com os dados devidamente preparados, foram aplicados diferentes modelos de previsão baseados em séries temporais e aprendizado de máquina. Os modelos testados incluíram:

- Holt-Winters (modelo aditivo com suavização de tendência e sazonalidade);
- Média Móvel Simples;
- ARIMA (AutoRegressive Integrated Moving Average);

- XGBoost com variáveis defasadas (lags).

A divisão entre treino e teste seguiu o formato de janelas deslizantes, com diferentes tamanhos de janela avaliados para lidar com a limitação de dados históricos. A métrica utilizada para comparação de desempenho foi o MAPE (Erro Percentual Absoluto Médio), que permite avaliar a precisão preditiva dos modelos.

4.2.4. ANÁLISE DE RESULTADOS E IMPLEMENTAÇÃO

Os modelos foram aplicados individualmente para cada tipo de café. A análise comparativa entre os valores previstos e os valores reais de vendas permitiu identificar o modelo mais adequado para cada caso. Com base nos resultados obtidos, foi possível gerar insights relevantes sobre o comportamento da demanda, além de recomendações práticas para auxiliar na tomada de decisão sobre produção e estoque.

4.3. RESULTADOS

Os resultados obtidos demonstram que o desempenho dos modelos de previsão variou de acordo com o tipo de café analisado. Para cada café, foi avaliado o MAPE de todos os modelos testados, e o modelo com o menor erro foi selecionado como o mais adequado para a previsão de vendas daquele produto específico. Modelos como Holt-Winters e XGBoost com defasagens se destacaram em séries com tendência ou variações mais complexas, enquanto modelos mais simples, como Média Móvel, mostraram bom desempenho em séries mais estáveis.

Tabela resumo com o melhor modelo para cada tipo de café:

Coffe_name	Melhor Modelo	MAPE (%)	Observação
Americano	Auto ARIMA	37,18	Modelo estatístico se destacou mesmo com dados limitados

Americano c/ Leite	XGBoost (com lags)	14,35	Modelo de ML captou bem o padrão de vendas
Cappuccino	Auto ARIMA	24,50	Melhor desempenho em série com possíveis variações suaves
Cocoa	XGBoost (com lags)	8,49	XGBoost foi mais eficaz que modelos estatísticos tradicionais
Cortado	Média Móvel	159,72	Todos os modelos tiveram alto erro, indicando série ruidosa, com alta variação e pouca tendência
Espresso	Holt Winters	56,51	Modelo com tendência e sazonalidade foi mais eficiente
Hot Chocolate	Auto ARIMA	39,68	Auto ARIMA lidou melhor com a série que apresentou sazonalidade
Latte	Média Móvel	27,30	Holt-Winters falhou (inf), modelo simples teve melhor precisão. Modelos com muita variabilidade, acabam reagindo melhor a uma média.

XGBoost (Machine Learning) teve ótimo desempenho em séries como *Cocoa* e *Americano com Leite*, demonstrando sua capacidade de capturar padrões complexos mesmo com poucos dados.

Auto ARIMA se mostrou consistente em vários cafés, sendo o melhor para *Americano*, *Cappuccino* e *Hot Chocolate*. Isso reforça sua adaptabilidade a diferentes padrões temporais.

Holt-Winters, embora útil para séries com tendência e sazonalidade, apresentou desempenho fraco em alguns casos, como no *Latte* (onde retornou inf no erro).

Média Móvel, embora simples, se destacou como melhor modelo para o *Cortado* e o *Latte*, o que pode indicar que séries mais ruidosas ou curtas se beneficiam de métodos menos complexos.

5. CONCLUSÃO

A análise preditiva realizada demonstrou que, mesmo com um volume limitado de dados históricos, foi possível identificar padrões relevantes de consumo para diferentes tipos de café. Utilizando diferentes abordagens de modelagem, como Holt-Winters, Média Móvel, Auto ARIMA e XGBoost com lags, foi possível avaliar o desempenho de cada técnica com base na métrica de erro percentual absoluto médio (MAPE), selecionando para cada tipo de café, o modelo com melhor acurácia.

No entanto, foi observado que o desempenho geral dos modelos foi impactado pela quantidade reduzida de dados disponíveis. O dataset utilizado abrangia poucos meses de registro, o que limita significativamente a capacidade dos modelos de capturar tendências e padrões sazonais mais robustos. Esse fator foi crítico para modelos que dependem de ciclos mais longos, como o Holt-Winters e o Auto ARIMA, além de dificultar o treinamento adequado de modelos mais complexos, como o XGBoost.

Apesar das limitações, os resultados obtidos permitem identificar caminhos promissores para previsões de demanda em ambientes com baixa variação temporal. Modelos mais simples como a Média Móvel apresentaram desempenho razoável em séries muito ruidosas, enquanto técnicas de aprendizado de máquina como o XGBoost mostraram bom potencial em séries com menor variabilidade.

Como possíveis próximos passos, seria interessante o uso de séries temporais mais longas e detalhadas, além do estudo de variáveis externas, como clima, sazonalidade de datas comemorativas e promoções (com modelos como o

Prophet) que podem ajudar a melhorar a acurácia dos modelos. Com a ampliação do conjunto de dados, espera-se que os modelos consigam capturar melhor os comportamentos de consumo e gerar previsões mais confiáveis para auxiliar na tomada de decisão estratégica.

6. ANEXOS

6.1. DATASETS

Link para o dataset de vendas de café:
<https://www.kaggle.com/datasets/ihelon/coffee-sales>

6.2. CÓDIGO GITHUB

Link para o repositório Github:
https://github.com/jeffersoncdantas/ml_coffee_sales

6.3. ENDEREÇO DO VÍDEO YOUTUBE

<https://www.youtube.com/watch?v=krelzmSu8C4>

7. REFERÊNCIAS BIBLIOGRÁFICAS

IHELON. *Coffee Sales*. Kaggle, 2023. Disponível em:
<https://www.kaggle.com/datasets/ihelon/coffee-sales>. Acesso em: 26 mar. 2025.

SANTOS, M.; FERREIRA, J. Ética na Inteligência Artificial: desafios e regulamentações. *Revista Brasileira de Tecnologia e Sociedade*, v. 5, n. 2, p. 45-62, 2023.

SILVA, R. A. O impacto da Inteligência Artificial na educação: desafios e perspectivas. *Revista de Educação e Tecnologia*, v. 8, n. 1, p. 20-35, 2021.

SOUZA, T. et al. Transparência e ética nos algoritmos de Inteligência Artificial. *Revista de Ciência e Computação Aplicada*, v. 10, n. 3, p. 78-95, 2022.