

Previsão Inteligente de Gastos Governamentais: Proposta de uma Arquitetura Baseada em Aprendizado de Máquina para Eficiência e Transparência

Kelly Santos De Oliveira¹, Felipe Barreto de Oliveira¹, Valerio Aymoré Martins¹,
Vinícius Coutinho Guimarães Coelho¹, Georges Daniel Amvame Nze¹,
Fábio Lúcio Lopes de Mendonça¹.

kelly.santos@redes.unb.br; felipe.oliveira@redes.unb.br;
valerioaymoremartins@gmail.com; vinicius.coelho@redes.unb.br;
georges@unb.br; fabio.mendonca@redes.unb.br

¹ Programa de Pós-Graduação Profissional em Engenharia Elétrica – PPEE - Universidade de Brasília,
Faculdade de Tecnologia, Departamento de Engenharia Elétrica, Brasília, Brasil - Zipcode 70910-900

Pages: 146-159

Resumo: A gestão eficiente dos gastos é essencial para a sustentabilidade financeira de organizações públicas e privadas. Entretanto, a complexidade na alocação de recursos e a quantidade de dados envolvidos dificultam a tomada de decisão. Neste contexto, este estudo propõe uma metodologia para previsão de gastos orçamentários utilizando algoritmos de aprendizado de máquina. A proposta é avaliada a partir de dados extraídos de um Data Warehouse especializado em finanças públicas brasileiras, contemplando variáveis quantitativas e qualitativas. Os objetivos incluem: identificar padrões de despesas; selecionar as melhores variáveis para previsão; avaliar os modelos com base em métricas de desempenho; e comparar os resultados com os gastos do exercício vigente. Os resultados visam apoiar decisões estratégicas, contribuindo para o uso mais inteligente dos recursos.

Palavras-chave: Qualidade do gasto; previsão orçamentária; aprendizado de máquina; finanças públicas; análise preditiva

Intelligent Prediction of Government Expenditure: An ML Architecture Proposal for Efficiency and Transparency

Abstract: Efficient expenditure management is essential for the financial sustainability of both public and private organizations. However, the complexity of resource allocation and data volume make decision-making challenging. In this context, this study proposes a methodology for budget expenditure forecasting using machine learning algorithms. The proposed approach is evaluated using data extracted from a specialized Data Warehouse containing Brazilian public finance information, encompassing quantitative and qualitative variables. The objectives include identifying spending patterns, selecting the most relevant features for prediction, evaluating models based on performance metrics, and comparing the

predictions with the actual expenditures of the current fiscal year. The results aim to support strategic decision-making and contribute to more intelligent use of available resources.

Keywords: Expenditure Quality; Budget Forecasting; Machine Learning; Public Finance; Predictive Analysis

1. Introdução

A gestão eficiente dos recursos financeiros é um dos pilares fundamentais para o crescimento sustentável de organizações públicas e privadas. O conceito de qualidade dos gastos visa maximizar os resultados e minimizar desperdícios. No entanto, a complexidade de distribuição e logística, aliada à grande quantidade de dados de cada despesa, torna a gestão de gastos um desafio considerável. A qualidade dos gastos implica em avaliar não apenas a quantidade de recursos empregados, mas também sua adequação e eficiência. Em ambientes governamentais, por exemplo, a má gestão dos gastos pode resultar em impactos financeiros negativos, afetando a sustentabilidade a longo prazo. Alocações inadequadas de recursos, podem afetar diretamente o desempenho institucional e aumentar o risco de crises financeiras. Assim, garantir a qualidade dos gastos requer uma análise criteriosa da sua distribuição, alinhada às prioridades estratégicas da organização, visando um uso mais inteligente dos recursos disponíveis. Entre os principais desafios enfrentados para manter a qualidade dos gastos, destacam-se: a falta de visibilidade clara sobre as despesas e a dificuldade em identificar padrões de comportamento financeiro ao longo do tempo. Em muitas organizações, a análise financeira tradicional se mostra limitada pela quantidade de dados e pela complexidade das variáveis envolvidas.

Diante desse cenário, técnicas de aprendizado de máquina (ML) têm se destacado como ferramentas promissoras para melhorar a previsão de gastos e apoiar a tomada de decisões estratégicas. O uso dessas técnicas tem se mostrado inovador e eficiente, permitindo que algoritmos aprendam com grandes volumes de dados, identificando padrões, tendências e correlações que seriam difíceis de detectar por métodos convencionais. Essa capacidade preditiva torna o aprendizado de máquina uma ferramenta valiosa para gerenciar e otimizar gastos em diferentes esferas, desde o setor privado até a administração pública.

Dessa forma, este trabalho propõe uma metodologia para previsão de gastos orçamentários, independentemente de sua natureza, com base em métodos de aprendizado de máquina. A metodologia será testada com dados extraídos de um Data Warehouse especializado em informações financeiras do setor público brasileiro, contendo tanto variáveis quantitativas quanto qualitativas.

2. Trabalhos Correlatos

Diversos estudos propuseram a aplicação de técnicas de aprendizado de máquina para a gestão e previsão de gastos no setor público. A seguir, são apresentados alguns dos trabalhos mais relevantes nessa linha de pesquisa.

O estudo realizado por Awariefe e Ogumeyo (Awariefe e Ogumeyo, 2023) explora o uso de métodos de aprendizado de máquina para prever os gastos militares da Nigéria.

Foram avaliados três modelos: multiple linear regression (MLR), autoregressive integrated moving average with exogenous variables (ARIMAX), e artificial neural networks (ANN). Os modelos foram treinados com dados históricos dos gastos militares, PIB e indicadores econômicos relevantes. Os resultados mostraram que o modelo ANN superou significativamente o MLR em termos de precisão, enquanto o ARIMAX apresentou desempenho promissor, mas ainda inferior ao ANN.

Hao e Xie (Hao e Xie, 2025) propuseram um framework baseado em múltiplos agentes LLM (Large Language Models), para apoiar a tomada de decisão em problemas de alocação de recursos no setor público. As soluções sugeridas por cinco LLMs distintos foram testadas em dois cenários simulados. Os resultados indicaram que o framework é capaz de modelar impactos políticos em agentes heterogêneos, revelando um caminho promissor para a análise econômica e a formulação de políticas públicas, explorando a capacidade de raciocínio e o poder computacional dos LLMs.

Por sua vez, Zarzà (Zarzà et al., 2024) apresentaram novas abordagens para orçamento financeiro individual e cooperativo, para maximizar a economia pessoal por meio da alocação eficiente entre diferentes categorias de despesa. Utilizando modelos LLM, o estudo oferece soluções iniciais viáveis para indivíduos e famílias com pouca familiaridade em planejamento financeiro. Os resultados preliminares sugerem que as soluções propostas produzem planos orçamentários sólidos, alinhados a metas e preferências dos usuários, contribuindo para uma gestão financeira mais acessível e eficiente.

O trabalho de Ajagbe (Ajagbe et al., 2021) aplicou o modelo ARIMAX para analisar a influência de gastos públicos, custos com mão de obra e preços ao produtor na previsão da produção anual de cacau na Nigéria, no período de 1981 a 2016. Por meio do teste de Dickey-Fuller, observaram que todas as variáveis eram estacionárias. O modelo que obteve melhor desempenho foi o ARIMAX (1,1,2), indicando que os gastos públicos e a mão de obra tiveram um efeito negativo, enquanto o preço ao produtor, com defasagem de dois anos, teve efeito positivo. As projeções indicaram tendência de crescimento na produção de cacau entre 2017 e 2031.

Por fim, o trabalho de Pimentel (Pimentel et al., 2023) investigou a previsão de receitas públicas utilizando métodos estatísticos e técnicas de ML, com foco na seleção de variáveis exógenas relevantes para estimativas orçamentárias trimestrais. A pesquisa seguiu a metodologia CRISP-DM e incluiu estudos de caso com experimentos comparativos entre modelos, visando identificar o mais adequado para o contexto analisado. O estudo também sugere o uso de técnicas como stacking e testes estatísticos para aprimorar a avaliação dos resultados. A Tabela 1 apresenta uma comparação entre o escopo desses trabalhos e a proposta desenvolvida neste estudo.

Critério	(Awariefe e Ogumeyo, 2023)	(Hao e Xie, 2025)	(Zarzà et al., 2024)	(Ajagbe et al., 2021)	(Pimentel et al., 2023)	Este trabalho
Predição de Gastos Públicos	X	-	-	X	X	X

Critério	(Awariefe e Ogumeyo, 2023)	(Hao e Xie, 2025)	(Zarzà et al., 2024)	(Ajagbe et al., 2021)	(Pimentel et al., 2023)	Este trabalho
<i>Deep Learning</i>	-	X	X	-	-	X
<i>Auxílio na tomada de decisão</i>	-	X	-	X	X	X
<i>Alocação Eficiente Automática</i>	-	X	-	-	-	-
<i>Tratamento dos dados</i>	X	-	-	X	X	X

Tabela 1 – Comparação entre os trabalhos correlatos e a proposta deste estudo

3. Metodologia

Para superar os possíveis desafios, a equipe de pesquisa foi composta por pesquisadores de Tecnologia da Informação e Orçamento Público, proporcionando uma multidisciplinaridade e fortalecendo a pesquisa.

Os dados utilizados neste estudo foram extraídos de um Data Warehouse especializado em informações financeiras do setor público brasileiro, contendo tanto dados quantitativos quanto qualitativos. O uso de Data Warehouses no setor público tem se mostrado eficaz para integrar grandes volumes de dados, permitindo análises detalhadas e integradas que facilitam a tomada de decisões financeiras e orçamentárias (Inmon, 2020). A seleção das variáveis quantitativas foi realizada com base no conhecimento técnico dos pesquisadores, resultando na escolha de 17 variáveis essenciais que cobrem todas as etapas do ciclo orçamentário, como dotação, empenho, liquidação e pagamento, além de variáveis transversais, como disponibilidades de crédito e restos a pagar, segundo as normas da Lei nº 4.320/64 (Brasil, 1964).

Entre as variáveis quantitativas, as relacionadas à liquidação da despesa se destacam, pois essa etapa representa o momento em que a despesa é efetivamente realizada, seguindo o princípio da competência contábil. A liquidação é crucial para prever a execução orçamentária e avaliar a real necessidade de recursos públicos, auxiliando os gestores na alocação eficiente de recursos (Gomes et al., 2019). Além das variáveis quantitativas, foram selecionadas 13 variáveis qualitativas, sendo que quatro delas desempenham papel de destaque no modelo de análise. Entre essas variáveis estão a unidade orçamentária, unidade executora e unidade responsável, que indicam a origem do orçamento, o órgão responsável pela execução e o gestor direto da aplicação dos recursos, respectivamente. A segregação dessas variáveis organizacionais é essencial para garantir a transparência e a accountability no uso dos recursos públicos (Peppard; Ward, 2021).

A análise dessas variáveis permitiu identificar padrões de comportamento de gastos, avaliar a origem dos recursos e, a partir do histórico de despesas, verificar se os recursos

alocados são suficientes para atender às necessidades futuras. A projeção orçamentária, nesse contexto, é uma ferramenta poderosa que melhora a eficiência do planejamento financeiro, permitindo que os gestores antecipem déficits ou subutilização de recursos (Gao; Liu; Wang, 2022). No entanto, a qualidade dos dados utilizados é um fator crítico que pode impactar diretamente a confiabilidade das previsões. A integração de múltiplas fontes de dados requer uma governança eficaz para assegurar a integridade e a precisão dos resultados (Korhonen; Ritala, 2021). Além disso, a capacitação técnica dos gestores públicos é essencial para que essas ferramentas sejam utilizadas de maneira eficaz e estratégica (Pina; Liddle, 2020).

O uso de tecnologias preditivas no setor público também levanta questões políticas e éticas. A introdução de algoritmos pode afetar a autonomia dos gestores na alocação de recursos, deslocando a discricionariedade política para decisões baseadas em dados e modelos automatizados, o que pode gerar resistências institucionais (Muir; Ostergaard, 2023). Assim, a implementação bem-sucedida dessas ferramentas depende de uma abordagem que equilibre a eficiência técnica com a necessidade de manter a transparência e a accountability na gestão pública.

3.1. Arquitetura Geral

O modelo em geral é idealizado para o acesso a uma infraestrutura de serviços baseada no Python Unicorn+FastAPI como ilustrado na Figura 1.

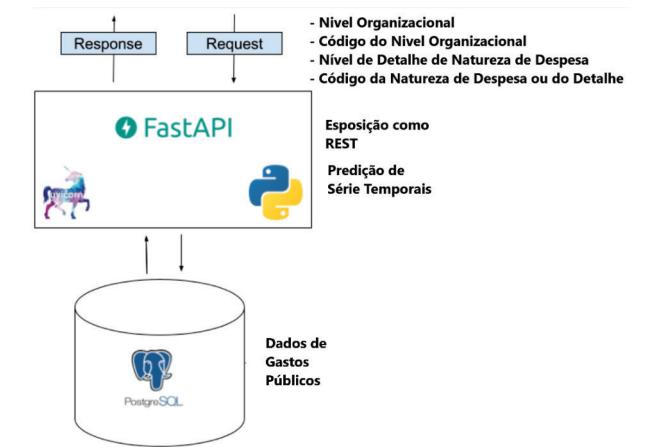


Figura 1 – Arquitetura Geral

A arquitetura pode ser organizada em quatro camadas principais, sendo elas: camada de disponibilização (Load Balancer), camada de aplicação (FastAPI) e camada de persistência (Banco de Dados). Cada uma dessas camadas desempenha um papel específico para garantir uma infraestrutura robusta e eficiente.

A camada de disponibilização utiliza um balanceador de carga NGINX, onde sua função é distribuir as requisições entre múltiplas instâncias do FastAPI para escalabilidade

horizontal e balancear a carga conforme métricas como número de conexões (round-robin). A camada de aplicação utiliza a ferramenta FastAPI, onde atua como a camada de serviço, lidando com a lógica de negócios e o acesso ao banco de dados. O acesso a FastAPI é garantido por limitação do CORS configurado, para permitir acesso de origens confiáveis. Por último temos a camada de persistência que é implementado um banco de dados relacional Postgresql, onde há a normalização para garantir integridade dos dados, otimização de índices de consulta e particionamento de tabelas para grandes volumes de dados.

Com isso, a metodologia proposta foi subdividida em três grandes etapas. A primeira delas é a responsável pela análise inicial dos dados a serem preditos. Nesta etapa foi observada as características dos dados em questão, para identificar padrões e particularidades nos dados. Nesta etapa há também o pré-processamento dos dados, onde os mesmos são ajustados e normalizados para a melhor eficiência nas previsões. Na etapa seguinte temos a definição e aplicação de modelos de previsão, apresentando a justificativa de sua seleção. Por fim temos a etapa de seleção e apresentação de indicadores, métricas e estimadores, com o intuito de apresentar a qualidade da previsão realizada. Cada uma destas etapas será explanada com mais detalhes nas seções ao longo do trabalho.

3.2. Análise dos dados

3.2.1. Agregação dos dados

Como etapa inicial, é importante selecionar níveis de agregação ou granularidade que permitam uma análise, evitando níveis com pouca informação.

Para iniciar as abordagens de verificação de agregação de dados, foi realizada uma análise de suficiência de dados, verificando, assim, o número de pontos de dados (mínimo de 30-50 pontos de dados é geralmente considerado suficiente para aplicar modelos de séries temporais tradicionais (como ARIMA) com maior precisão) e a distribuição temporal dos dados (se os dados em cada nível são distribuídos ao longo do tempo de forma uniforme ou se há lacunas significativas. Séries com muitas lacunas, ou que estejam concentradas em períodos curtos, podem não ser adequadas para modelagem preditiva). Em seguida, foram realizadas as análises de contribuição e variabilidade de cada nível dos dados. Níveis que apresentam pouca variabilidade ou contribuição podem não ser relevantes para uma previsão detalhada e, portanto, podem ser agregados a outros níveis.

3.2.2. Dados de anos intermediários com valores incongruentes (COVID)

Para lidar com valores incongruentes que fogem da tendência, ou outliers, em anos intermediários, existem várias abordagens que podem ser usadas para corrigir esses desvios e melhorar a precisão de previsões e análises (A Figura 2 mostra um exemplo de despesa com comportamento atípico em 2020). Para tal foram implementadas 4 abordagens diferentes, com o objetivo de verificar qual a que melhor se enquadra na situação enfrentada. Na escolha dessa, a solução sistêmica foi desenvolvida considerando um parâmetro definido como ajuste.

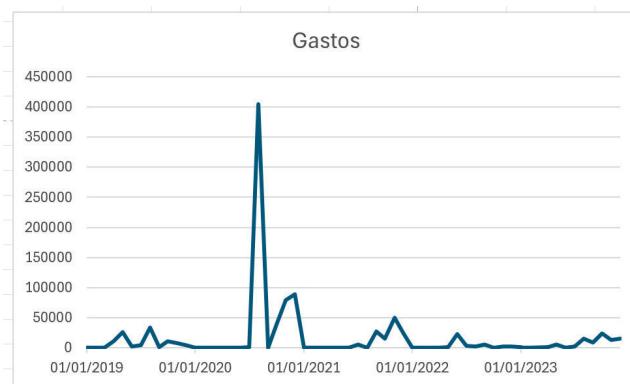


Figura 2 – Como exemplo mostramos o comportamento das séries de gastos de 2019 a 2023 da “Natureza Despesa Detalhada” de “Materiais Hospitalares”

Após análises das situações e seus valores, foi tomada a decisão de interpor diversos modelos, para agregar nas avaliações. Estes são os possíveis valores:

- **ajuste=0:** Não contabilizar os dados dos anos atingidos pela pandemia na predição dos meses futuros.
- **ajuste=1:** Mover os dados anteriores para “cobrir” os anos incongruentes, ajustando diversas questões de médias móveis.
- **ajuste=2:** Realização de interpolação linear
- **ajuste=3:** Realizar predição dos anos afetados pela pandemia, levando em consideração os anos anteriores. Após a primeira predição, realiza a predição dos anos futuros, com os valores preditos nos anos de pandemia.

3.2.3. Resto a Pagar

No decorrer da análise dos dados foi identificado o desafio de valores pertencentes a um ano constar em um mês qualquer do ano seguinte. Dessa forma, tornou-se necessário corrigir essa alocação incorreta, garantindo que o valor seja movido para o período correto. Esse problema pode ocorrer por erros de registro, processamento ou por convenções de datas diferentes.

Para atender os limites desse parâmetro e de forma resolver o problema do RAP apontado, foram elencadas 3 soluções possíveis:

- **rap=0:** Ignora o valor do RAP (a princípio esse entendimento já era descartado devido à natureza importânciia do valor do RAP e sua influênciia nas decisões de gastos futuras)
- **rap=1:** deslocar o valor de RAP para o último mês do ano anterior no entendimento que já era um compromisso de gasto dentro do ano fiscal correspondente (o que já era entendido que poderia impor uma tendênciia de gastos “excessiva” no último mês do seu ano fiscal).
- **rap=2:** distribuir ponderadamente o RAP em meses, no qual estava presente um gasto no ano anterior (o que de inicio era o entendimento de uma medida mais eficaz).

3.3. Técnicas de Predição

Inicialmente, para a definição das técnicas de predições a serem utilizadas, foram gerados gráficos representativos dos dados, de modo a responder alguns questionamentos pertinentes. São eles: “Existem padrões consistentes?”, “Existe uma tendência significativa?”, “A sazonalidade é importante para o modelo?”, “Há evidências da presença de ciclos econômicos (no caso de gastos)?”, “Existem valores discrepantes nos dados que precisam ser explicados por pessoas com conhecimento especializado?”, “Quão fortes são as relações entre as variáveis disponíveis para análise?”. Vale ressaltar que cada modelo é em si uma construção artificial que se baseia num conjunto de pressupostos (explícitos e implícitos) e geralmente envolve um ou mais parâmetros que devem ser estimados utilizando os dados históricos conhecidos. O modelo deve prever mecanismos que permitam a operação de predição de 12 meses após o mês corrente e de validação dos últimos 12 meses ocorridos.

Analisando os dados disponíveis, foram definidos dois modelos para realização das predições esperadas: (i) **Modelo ARIMA**: Modelos estatísticos/estocásticos autorregressivos integrados de médias móveis, tradicionalmente utilizado para séries temporais e; (ii) **Modelo LLM**: Utilização de ferramentas de rede neural LLM (Large Language Model), com capacidade de interpretar dados complexos por meio do processamento de linguagem. Neste trabalho, foram utilizados o ChatGPT-40 mini e a biblioteca PandasAI, explorando o potencial dos LLMs na identificação de padrões e geração de predições com base em linguagem estruturada e não estruturada.

3.3.1. ARIMA

Modelo clássico de séries temporais baseado em três componentes principais: a parte autorregressiva (AR), a parte de médias móveis (MA) e a integração (I) que lida com a estacionariedade dos dados. O modelo requer pré-processamento dos dados para verificar se são estacionários. Se os dados não forem estacionários, a diferenciação (*integration*) é aplicada para torná-los adequados ao modelo. Vale ressaltar que a sazonalidade pode ser incorporada com variações como SARIMA/SARIMAX (Seasonal ARIMA).

Entre as principais vantagens do ARIMA, destaca-se sua interpretação estatística clara, o que o torna um modelo bem compreendido na literatura. Ele apresenta bom desempenho em séries temporais com padrões regulares, tendências suaves ou sazonalidade simples. A flexibilidade na escolha dos parâmetros (AR, I e MA) permite adaptá-lo a diferentes características de séries temporais.

No entanto, o ARIMA apresenta limitações importantes: (i) baixo desempenho em séries com padrões não lineares ou alta complexidade temporal; (ii) sensibilidade à estacionariedade, exigindo testes e ajustes manuais prévios e; (iv) necessidade de calibração cuidadosa dos parâmetros, o que pode demandar tempo e conhecimento especializado.

3.3.2. LLM

Os modelos LLMs, como GPT-3 ou GPT-4, são baseados em arquiteturas de aprendizado profundo treinadas com grandes volumes de texto para executar tarefas variadas, assim

como, a inclusão da análise de séries temporais. Entre as principais vantagens dos LLMs, destacam-se: (i) capacidade de capturar padrões sutis e relações não lineares; (ii) menor dependência de pré-processamento estatístico, como verificação de estacionaridade; (iii) interação por linguagem natural, que simplifica o uso por não especialistas, como na plataforma OpenAI Completion e; (iv) flexibilidade e automação ao integrar com ferramentas como PandasAI, que permite a manipulação de DataFrames por comandos em linguagem natural, facilitando análises temporais mesmo sem domínio avançado da linguagem Python.

Por outro lado, os LLMs apresentam algumas limitações relevantes: (i) operam como “caixas-pretas”, dificultando a interpretação dos parâmetros internos e das relações inferidas; (ii) demandam maior capacidade computacional, tanto para treinamento quanto para execução; (iii) o desempenho está diretamente relacionado à qualidade e quantidade dos dados de entrada e; (iv) o PandasAI, embora facilite o uso, não é um modelo de predição por si só, mas sim uma interface que amplia a acessibilidade na manipulação dos dados.

3.4. Avaliação do Modelo de previsão

Uma vez selecionado um modelo e estimados seus parâmetros, o modelo é usado para fazer previsões. O desempenho do modelo só poderá ser devidamente avaliado após a disponibilização dos dados do período de previsão. Para gerar estimadores que permitam comparar a eficácia de diferentes previsões em séries temporais, utilizam-se diversas métricas estatísticas que avaliam a qualidade das previsões em relação aos valores observados. Essas métricas podem ser aplicadas para comparar diferentes modelos e identificar qual deles fornece previsões mais precisas. A seguir estão os principais estimadores (ou métricas de erro) usados para comparar resultados de previsões: (i) **Erro Médio Absoluto (MAE - Mean Absolute Error)**: O MAE mede a magnitude média dos erros em um conjunto de previsões, sem levar em consideração a direção do erro (positivo ou negativo); (ii) **Erro Médio Quadrático (MSE - Mean Squared Error)**: O MSE é uma métrica que penaliza erros maiores mais severamente, já que os erros são elevados ao quadrado; (iii) **Raiz do Erro Médio Quadrático (RMSE - Root Mean Squared Error)**: O RMSE é a raiz quadrada do MSE, trazendo a métrica de volta à unidade original dos dados e; (iv) **Erro Percentual Absoluto Médio (MAPE - Mean Absolute Percentage Error)**: O MAPE calcula o erro em termos percentuais, sendo útil para comparar séries temporais com escalas diferentes.

4. Resultados

Durante os testes, foi observado que, ao se acionar um prompt para o modelo GPT solicitando previsão com base em aspectos como estacionalidade e sazonalidade, o modelo frequentemente recomendava o uso da função ExponentialSmoothing da biblioteca statsmodels, especificamente com o modelo Holt-Winters, conforme exemplificado na Figura 3.

```

# Criando a série temporal
dates = pd.date_range(start="2021-01-01", periods=36, freq='MS')
values = [
    0.00, 0.00, 0.00, 0.00, 49126.79, 700.15, 295968.75, 201518.40, 11762.15, 713815.75, 269079.71, 308341.91,
    0.00, 0.00, 10281.93, 10531.00, 207287.00, 493906.01, 99067.86, 279138.27, 336258.60, 511359.65, 293863.37, 355836.01,
    0.00, 0.00, 111999.58, 547936.24, 97344.27, 341436.87, 322296.96, 336232.90, 163373.13, 574823.29, 601538.49, 83515.88
]

# Criando um DataFrame
df = pd.DataFrame({'Date': dates, 'Value': values})
df.set_index('Date', inplace=True)

# Aplicando Holt-Winters para previsão
model = ExponentialSmoothing(df['Value'], trend='add', seasonal='add', seasonal_periods=12)
fit = model.fit()
forecast = fit.forecast(steps=12)

# Exibindo a previsão
forecast_df = forecast.to_frame(name="Forecast")
import ace_tools as tools
tools.display_dataframe_to_user(name="Previsão para 2024", dataframe=forecast_df)

```

Figura 3 – Exemplo modelo Holt-Winters

O modelo Holt-Winters, embora seja uma das técnicas de previsão mais populares para séries temporais, comumente aplicado em problemas como detecção de anomalias e previsão meteorológica, demonstrou limitações para este estudo. Esse modelo é uma extensão do modelo de Holt, desenvolvido por Winter e, sua lógica baseada em suavização exponencial tende a amortecer picos de gastos, o que compromete a acurácia das previsões em contextos onde essas variações são relevantes. Para contornar essa limitação, foi incluída no prompt uma instrução explícita: "..., porém sem usar nenhum processo de suavização, isto é, mantendo os picos de valores de gastos". Com esses ajustes, foram realizadas previsões de despesas para o ano de 2024, posteriormente comparadas com os valores reais observados.

A Figura 4 apresenta os resultados da predição para uma despesa típica da universidade, onde nos períodos de férias escolares há uma redução de gastos e no período letivo há um crescimento da despesa. É possível observar que o modelo obteve uma previsão correta, diferenciando os períodos de férias escolares e períodos letivos, gerando previsões consideráveis da despesa ao longo do ano.

Para as despesas de natureza discricionária, sua demanda é intempestiva junto a diversos anos de análise, provocando uma previsão que pode ficar inadequada, não em valor, mas em momento junto aos meses. Esse comportamento foi claramente identificado nos resultados, conforme Figura 5.

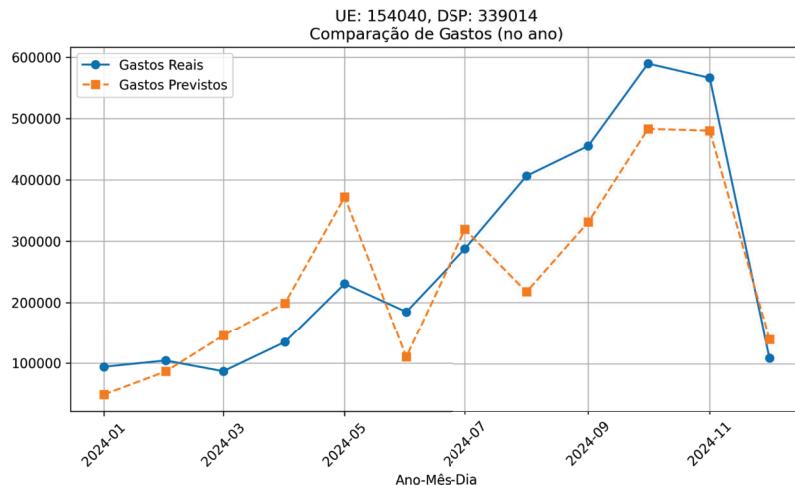


Figura 4 – Resultado predição despesa típica.

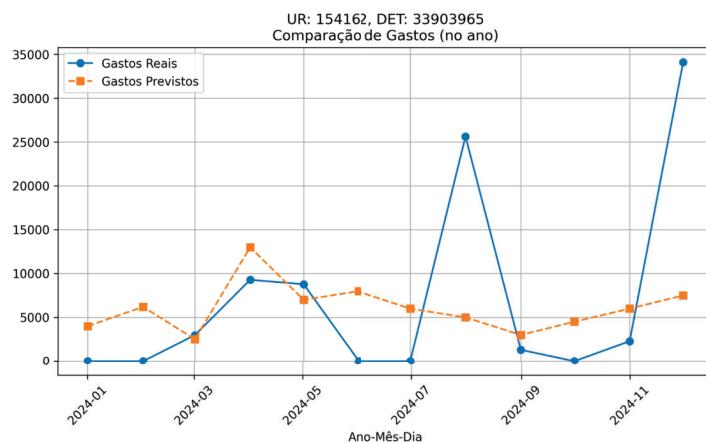


Figura 5 – Resultado predição despesa discricionária.

Por fim, foram identificadas ocorrências de deslocamento (conforme apresentado na Figura 6 entre o mês 1 e o mês 7). Isto ocorre dada a natureza específica da Unidade Executora específica, que tende a utilizar seus recursos associados à abertura de períodos letivos de ensino. Dessa forma, a “natureza” da unidade, tal como a citada, pode influir em aspectos “negativos” na capacidade de previsão de uso de recursos (“gastos”). Esse tipo de deslocamento impede uma simples verificação de erro por MSRE (*Mean Square Root Error*).

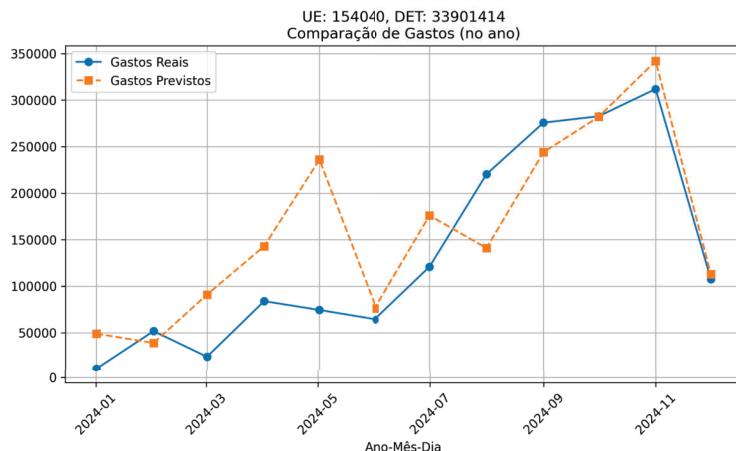


Figura 6 – Resultado predição com deslocamento.

5. Conclusão

Este trabalho propôs uma metodologia para a predição de gastos orçamentários, independentemente de sua natureza, com base em técnicas de aprendizado de máquina. Para isso, o desenvolvimento foi estruturado em três grandes etapas: (i) análise exploratória dos dados, com identificação de padrões e características específicas das séries temporais; (ii) seleção e aplicação de modelos preditivos, com justificativas técnicas para sua escolha; (iii) avaliação do desempenho preditivo, por meio da aplicação de métricas e estimadores estatísticos. Os resultados obtidos indicam que a metodologia apresentou bom desempenho na previsão de despesas típicas, especialmente aquelas relacionadas ao funcionamento regular da universidade, como períodos letivos e de recesso. A variação sazonal foi corretamente identificada, refletindo a adequação da abordagem proposta.

Entretanto, observaram-se limitações na previsão de despesas discricionárias, cuja demanda é mais irregular ao longo do tempo e, em despesas que apresentam efeitos de deslocamento temporal, comprometendo a precisão do modelo em determinadas situações. Durante os testes realizados no modelo, foram identificadas possibilidades de aprimoramento, que poderão ser exploradas em trabalhos futuros: (i) a quantidade de dados históricos influencia diretamente a qualidade da predição. Verificou-se que séries com menos de 36 meses resultam em erros superiores a 20%; (ii) o tempo de execução médio para previsão de uma única despesa foi de aproximadamente 12 segundos, o que pode ser otimizado com pré-processamento mais eficiente e paralelização.

Agradecimentos

Os autores agradecem ao apoio técnico e computacional do Laboratório de Tecnologias para Tomada de Decisão - LATITUDE, da Universidade de Brasília, que conta com

apoio do CNPq - Conselho Nacional de Pesquisa (Outorgas 312180/2019-5 PQ-2 e 465741/2014-2 INCT em Cibersegurança), da Advocacia Geral da União (Outorga AGU 697.935/2019), do Fundação de Apoio à Pesquisa do Distrito Federal, da Polícia Federal (Outorga PF 03/2020), do Mestrado Profissional em Engenharia Elétrica, na área de concentração: Segurança Cibernética – 1ª Turma para Profissionais do Setor de Inteligência (Outorga ABIN 01/2019) ao Decanatos de Pesquisa e Inovação e de Pós-Graduação da Universidade de Brasília (Outorga 7129 FUB/EMENDA/DPI/COPEI/AMORIS), do Projeto SISTER City –Sistemas Inteligentes Seguros e em Tempo Efetivo Real para Cidades Inteligentes (Outorga 625/2022) e do Decanato de Administração da Universidade de Brasília.

Referências

- BRASIL. Lei nº 4.320, de 17 de março de 1964. Estatui normas gerais de direito financeiro para elaboração e controle dos orçamentos e balanços da União, dos Estados, dos Municípios e do Distrito Federal. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l4320.htm. Acesso em: 23 set. 2024.
- GAO, J.; LIU, Z.; WANG, X. Big data and predictive modeling in public financial management: Opportunities and challenges. International Journal of Public Administration, v. 45, n. 3, p. 321-338, 2022.
- GOMES, R. C.; FERNANDES, A. C.; CARVALHO, E. Gestão orçamentária no setor público. Revista de Administração Pública, v. 53, n. 4, p. 6 17-640, 2019.
- INMON, W. Building the data warehouse. 5. ed. New York: John Wiley & Sons, 2020.
- KORHONEN, J.; RITALA, P. Data governance in public organizations: Ensuring data quality and trust in digital transformation. Journal of Digital Government, v. 15, n. 1, p. 23-36, 2021.
- MUIR, K.; OSTERGAARD, M. Algorithmic decision-making in government: Balancing innovation and accountability. Journal of Public Policy, v. 43, n. 2, p. 211-231, 2023.
- PEPPARD, J.; WARD, J. The strategic management of information systems: Building a digital strategy. 4. ed. London: Wiley, 2021.
- PINA, V.; LIDDLE, J. Data-driven governance: Challenges for integrating predictive analytics in public management. International Journal of Public Sector Management, v. 33, n. 7, p. 863-879, 2020.
- AWARIEFE C.; OGOMEYO, S. A. Prediction Accuracy of Nigerian Military Expenditure: MLR, ARIMAX, and ANN Models in Statistical and Machine Learning Frameworks. FUDMA Journal of Sciences. v. 7, n. 6, p. 149-156, 2023.
- HAO Y.; XIE, D. A. Multi-LLM-Agent-Based Framework for Economic and Public Policy Analysis. Department of Economics, The Hong Kong University of Science and Technology, 2025.
- DE ZARZÀ I. at al. Optimized Financial Planning: Integrating Individual and Cooperative Budgeting Models with LLM Recommendations. AI (MDPI). v. 5, p. 91-114, 2024.

AJAGBE S. A et al. The Role of Public Expenditure, Labour and Producer Price in Forecasting Cocoa Production in Nigeria Using ARIMAX Model. Direct Research Journal of Agriculture and Food Science. v. 9, p. 385-395, 2021.

PIMENTEL C. P. S. et al. Aplicação de Técnicas de Aprendizagem de Máquina com Seleção de Variáveis na Previsão de Receitas Públicas de 8 Capitais. XIV Computer on the Beach, 2023.