

A MODELAGEM MATEMÁTICA COMO UMA FERRAMENTA DE OTIMIZAÇÃO NO TRABALHO DE ADOPS

ÉDITA MARTINS¹; MIGUEL STRELOW²; REGIS QUADROS³

¹*Universidade Federal de Pelotas – editagomes1878@gmail.com*

²*Universidade Federal de Pelotas – mdtstrelow@inf.ufpel.edu.br*

³*Universidade Federal de Pelotas – regis.quadros@ufpel.edu.br*

1. INTRODUÇÃO

Este trabalho foi pensado e desenvolvido por alunos da Universidade Federal de Pelotas, dos cursos de Licenciatura em Matemática e Engenharia da Computação, com o objetivo principal de analisar padrões relacionados à precificação de anúncios digitais em diferentes países, buscando otimizar métricas utilizadas na precificação e voltadas para a atividade de Adops (operações de anúncios).

A publicidade digital constitui uma das principais formas de geração de receita no ambiente online, tendo influência de vários fatores. Nesse contexto entender os elementos que interagem entre si e impactam nos resultados é crucial para desenvolver estratégias no trabalho de Adops. Assim como TRUZZI (2014) aponta em seu trabalho.

A publicidade digital opera com modelos de negócio online, baseado em leilões de anúncios, e no qual anunciantes firmam contratos com redes de anúncios. Neste contexto, campanhas são definidas por orçamento, tempo e modelo de precificação, como custo por ação (CPA) ou por clique (CPC), sendo estes os mais comuns. O processo caracteriza-se como uma decisão sequencial, em que a rede escolhe qual anúncio exibir para cada usuário, considerando orçamento, valor pago e perfil do público.

Com base nos aspectos apresentados, o trabalho propõe a construção de um modelo automatizado que coleta, processa, purifica e analisa os dados fornecidos. Essa abordagem verifica o algoritmo de dados com o melhor desempenho e também oferece percepções sobre o impacto das variáveis nas operações de anúncios.

O projeto apresentado neste documento está em fase de desenvolvimento e passando pelas primeiras fases de teste.

2. METODOLOGIA

A coleta de dados e análise ocorreu em etapas bem definidas, para garantir a consistência estatística e a qualidade das informações utilizadas no estudo. Inicialmente os dados foram coletados e armazenados em planilhas, depois foram avaliados e separados por grau de relevância no estudo.

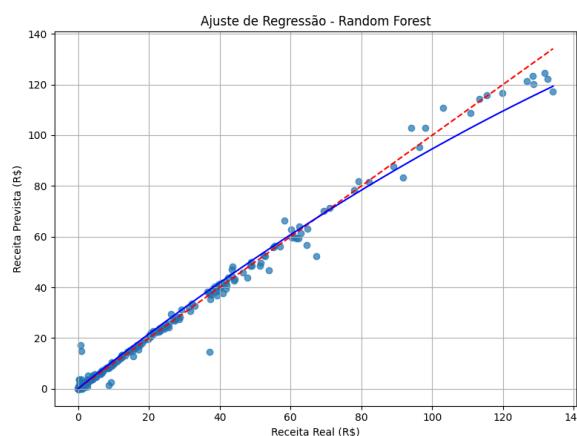
Em seguida, ocorreu o tratamento de valores ausentes e à detecção e remoção de outliers, com o método do intervalo interquartil (IQR) — medida estatística que identifica a dispersão dos dados e permite eliminar observações discrepantes que poderiam comprometer a análise.

Após a conclusão da etapa de seleção dos dados a serem utilizados, teve

início a etapa de seleção das variáveis que seriam independentes ou dependentes, estabelecendo a base para modelagem dos dados. Na sequência, foi aplicado o StandardScaler, para padronizar as variáveis e evitar que diferenças de escala influenciassem o desempenho dos algoritmos. Após a padronização, os dados foram divididos em dois subconjuntos: conjunto de treino (80%) e teste (20%).

Diversos modelos de regressão foram implementados e avaliados, entre eles a Ressagem Linear, o Random Forest Regressor e o Support Vector Regressor (SVR). Para fazer uma mensuração do desempenho foi adotada como métrica o coeficiente de determinação (R^2), que expressa a proporção da variação da variável dependente e é explicada pelas variáveis independentes do modelo. Essa métrica foi escolhida por ser bastante utilizada em problemas de regressão.

A partir desses resultados, foram elaborados gráficos de dispersão, (Figura 1) nos quais é possível visualizar a relação entre os valores reais e os valores previstos para as métricas adotadas. Essa visualização gráfica deixou evidente a qualidade do ajuste dos modelos e possibilitou a identificação de padrões de aderência e discrepância.



(Figura 1, autoria própria)

3. RESULTADOS E DISCUSSÃO

A execução do modelo trouxe a possibilidade de avaliar o desempenho de seis modelos diferentes. O critério de seleção indicou o modelo com maior valor de R^2 como o mais adequado para explicar a variação das métricas em função das variáveis selecionadas.

O gráfico de dispersão obtido evidenciou a proximidade entre valores reais e previstos, aproximando o cenário ideal e ajustes que precisavam ser feitos. Essa visualização permite identificar tanto os bons cenários quanto aqueles em que os modelos apresentaram mais erros.

Houve uma análise das cinco melhores previsões, onde o modelo mostrou que, embora os modelos consigam captar bem as tendências gerais, ainda existem erros nos valores absolutos, principalmente em cenários de receitas mais altas. Esse resultado reforça a necessidade de incorporar variáveis adicionais em estudos futuros, como sazonalidade, localização geográfica ou tipo de dispositivo pelo qual as pessoas assistem aos anúncios, para aumentar a margem de acerto

do modelo que virá a ser definitivo. Além disso, a comparação entre modelos mostrou que técnicas baseadas em árvores de decisão (como Random Forest e Gradient Boosting) apresentaram melhor desempenho que métodos lineares tradicionais em termos de ajuste, devido à capacidade de capturar relações não lineares entre variáveis.

4. CONCLUSÕES

O estudo se mostrou viável na aplicação de um modelo capaz de prever a receita de anúncios digitais a partir de métricas baseadas no trabalho de AdOps, pois em seu desenvolvimento se provou capaz de automatizar a preparação e limpeza de dados, avaliar e comparar diferentes algoritmos de regressão, e fornecer visualizações e relatórios baseados no comportamento dos dados fornecidos.

A principal contribuição do estudo é trazer integração de conceitos matemáticos e computacionais em um sistema capaz de amparar o trabalho de *adops*, auxiliando na identificação de padrões de precificação e no aperfeiçoamento de estratégias em diferentes contextos de operação.

Como perspectivas futuras, o projeto mostrou potencial para ir além da sua fase inicial. Assim, temos possíveis próximos passos para aprimorar os modelos e as análises, garantindo resultados mais precisos.

Também, se pensa em ampliar a base de dados analisada, incorporar variáveis de contexto adicionais como comportamento do usuário que assiste aos anúncios em diferentes países, fuso horário e dispositivos por onde o site é acessado, e testar métodos mais avançados, como redes neurais, a fim de melhorar a veracidade das previsões, previsões, assim como para aprimorar o projeto, o próximo passo será otimizar o modelo com o ajuste de hiperparâmetros e usar validação cruzada para garantir resultados mais confiáveis. Além do R^2 , será testada a possibilidade de incluir outras métricas para validação do modelo, e também experimentar outras formas de lidar com os outliers, que atualmente são removidos, para melhor adequar o modelo aos dados disponíveis.

5. REFERÊNCIAS BIBLIOGRÁFICAS

AHER, S. R.; GUPTA, T.; GURURAJ, H. S. A Machine Learning Based Approach for Price Estimation. In: INTERNATIONAL CONFERENCE ON SUSTAINABLE COMMUNICATION, NETWORKING AND APPLICATION, 3., Bangaluru, 2023. Anais... Bangaluru: Springer, 2023. p. 320-330.

BHATTACHARYYA, J. LASSO Regression – A Procedural Improvement. SSRN, 23 jan. 2025. Disponível em: <https://ssrn.com/abstract=5108591>. Acesso em: 29 ago. 2025.

HODGE, V. J.; AUSTIN, J. A survey of outlier detection methodologies. Artificial Intelligence Review, Dordrecht, v. 22, n. 2, p. 85-126, 2004.

ILYAS, M. S. B.; IKRAM, A.; BUTT, M. A.; TARIQ, I. Comparative Analysis of Regression Algorithms used to Predict the Sales of Big Marts. Journal of Innovative Computing and Emerging Technologies, [s. l.], v. 3, n. 1, p. 11-18, 2023.

JADHAV, A. Data-Driven Pricing Optimization: Using Machine Learning to Dynamically Adjust Prices Based on Market Conditions. SSRN, 2024. Disponível em: <https://ssrn.com/abstract=4793547>. Acesso em: 29 ago. 2025.

JADHAV, H. The Impact of Outliers in Machine Learning. Medium, 2 ago. 2023. Disponível em: <https://harshjadHAV100.medium.com/the-impact-of-outliers-in-machine-learning-6cf4e4d33f21>. Acesso em: 29 ago. 2025.

MATSUKI, K.; WANG, R.; BLAIS, C.; SUTHERLAND, M. The Random Forests statistical technique: an examination of its value for the study of reading. *Scientific Studies of Reading*, [s. l.], v. 20, n. 1, p. 20-33, 2016.

MODASSIR. Sales prediction: leveraging machine learning to enhance advertising effectiveness. Medium, 22 jan. 2024. Acessado em: 29 ago. 2025. Online. Disponível em: <https://medium.com/@mn.zamindar/sales-prediction-leveraging-machine-learning-to-enhance-advertising-effectiveness-0eeeb18a24eb>

YADAV, A. Linear Regression vs Random Forest. Medium, 20 jul. 2024. Disponível em: <https://medium.com/@amit25173/linear-regression-vs-random-forest-7288522be3aa>. Acesso em: 29 ago. 2025.

ZHAO, G.; CHENG, X. Building a Revenue Prediction Model Based on Machine Learning Tuning Methods. In: *International Journal of Machine Learning and Computing*, Singapura, v.13, n. 4, p. 1082-1090, 2023.

TRUZZI, Flávio Sales; REALI COSTA, Anna Helena. *Modelagem e soluções para redes de anúncios*. 2014.