Journal of Applied Data Sciences
Vol. 6, No. 2, May 2025, pp. 1037-1046
ISSN 2723-6471

Enhancing Digital Marketing Strategies with Machine Learning for Analyzing Key Drivers of Online Advertising Performance

Berlilana[1,*], Taqwa Hariguna[2], Ibrahiem M. M. El Emary[3]

[1,2]Magister of Computer Science, Universitas Amikom Purwokerto, Jawa Tengah, Indonesia

[3]King Abdulaziz University, Kingdom of Saudi Arabia

(Received: November 21, 2024; Revised: December 11, 2024; Accepted: January 21, 2025; Available online: March 4, 2025)

Abstract

The rapid growth of digital advertising has underscored the need for data-driven strategies to optimize campaign performance. This study applies machine learning techniques to analyze online advertising data, aiming to identify key performance drivers and provide actionable insights for optimizing marketing strategies. The dataset includes metrics such as clicks, displays, costs, and revenue, which were preprocessed, analyzed, and modeled using ensemble methods, including Random Forest and Gradient Boosting. These ensemble methods were chosen for their ability to handle high-dimensional data, mitigate overfitting, and capture complex, nonlinear relationships between variables. Random Forest, with its bagging approach, enhances generalization by reducing variance, while Gradient Boosting incrementally corrects errors by focusing on hard-to-predict instances, improving overall predictive performance. Descriptive analysis revealed significant variability in campaign outcomes, with cost and user engagement emerging as primary predictors of revenue. Machine learning models demonstrated strong predictive accuracy, with Random Forest achieving 92% accuracy and an F1-score of 89%. Visualizations such as feature importance charts, correlation heatmaps, and learning curves validated the robustness of the models and highlighted key insights, including inefficiencies in cost allocation and the limited impact of certain categorical features like placement. The study emphasizes the potential of machine learning to optimize digital marketing strategies by identifying critical factors that influence campaign success. The findings provide a scalable framework for resource allocation, audience targeting, and strategic decision-making in online advertising. Future research could further enhance predictions by incorporating additional features, such as audience demographics and temporal trends, to provide deeper insights into campaign dynamics.

Keywords: Machine Learning, Digital Marketing, Online Advertising, Campaign Optimization, Predictive Modeling

1. Introduction

The digital advertising industry has witnessed exponential growth in recent years, with global spending reaching unprecedented levels. As marketers invest heavily in online campaigns, the need for data-driven strategies to optimize performance has become critical. Studies have highlighted that factor such as budget allocation, ad placement, and user engagement significantly influence campaign success [1], [2]. Despite this, many campaigns underperform due to inefficiencies in targeting and resource utilization [3].

Machine learning has emerged as a powerful tool for analyzing large-scale advertising datasets, offering capabilities to uncover hidden patterns and predict campaign outcomes [4], [5]. Traditional statistical methods often struggle with the complexity and non-linearity of advertising data, while machine learning models excel in capturing these intricate relationships [6]. Techniques such as Random Forest and Gradient Boosting have shown particular promise, providing high accuracy in predictive tasks while identifying key drivers of performance [7].

This study aims to address gaps in understanding the factors driving online advertising success by applying machine learning to a dataset containing metrics such as clicks, displays, costs, and revenue. Previous research has primarily focused on individual aspects of campaign performance, such as click-through rates or cost per click [8], [9]. However, holistic analyses that integrate multiple metrics and leverage advanced predictive models remain limited.

*Corresponding author: Berlilana (berlilana02@amikompurwokerto.ac.id)
DOI: https://doi.org/10.47738/jads.v6i2.658

The objectives of this research are threefold: first, to identify the most influential factors affecting campaign performance using feature importance analysis; second, to develop machine learning models that accurately predict revenue and other key outcomes; and third, to provide actionable insights for optimizing digital marketing strategies. The findings are expected to contribute to the growing body of knowledge on data-driven advertising, offering practical solutions for marketers and decision-makers [10], [11], [12].

In the following sections, we detail the methodology used for data analysis and modeling, present the results of our predictive analysis, and discuss the implications for optimizing campaign strategies. This study underscores the potential of integrating machine learning with advertising analytics to drive better outcomes in the competitive digital marketing landscape.

2. Literature Review

The field of digital marketing analytics has expanded significantly, driven by the proliferation of data generated from online advertising platforms. Researchers have explored various aspects of advertising performance, including factors influencing click-through rates (CTR), cost-effectiveness, and audience engagement. This section reviews the relevant literature, highlighting advancements and gaps in the application of machine learning to advertising analytics.

Studies have consistently emphasized the importance of user engagement metrics, such as clicks and post-click conversions, in determining campaign success [13], [14]. These metrics serve as proxies for the effectiveness of targeting strategies and ad content. Furthermore, financial metrics, including cost per click (CPC) and return on investment (ROI), are widely recognized as critical indicators of campaign efficiency [15], [16]. However, the interplay between engagement and financial metrics remains underexplored in the context of predictive analytics.

Machine learning has been increasingly adopted to address the complexities of advertising performance data. Techniques such as decision trees, Random Forest, and Gradient Boosting have demonstrated high accuracy in predicting campaign outcomes [17], [18]. These models are particularly effective in capturing non-linear relationships and feature interactions, which are prevalent in advertising datasets. For instance, Random Forest has been used to identify key predictors of ad engagement, such as time of day and user demographics [19].

Despite these advancements, there are notable gaps and challenges in the literature. Many studies focus on isolated aspects of campaign performance, such as predicting CTR or optimizing CPC, without integrating these insights into a unified framework for decision-making [20]. Additionally, categorical features like ad placement and creative design are often underrepresented in analyses due to challenges in encoding and interpreting their impact [21].

A key challenge in real-world advertising data is class imbalance, where successful outcomes (e.g., conversions) occur far less frequently than non-conversions, leading to biased predictions. Oversampling, undersampling, and techniques such as Synthetic Minority Over-sampling Technique (SMOTE) have been explored to mitigate these imbalances, but their effectiveness in advertising contexts remains under-researched. Another issue is feature selection, as advertising datasets often contain a high number of correlated variables, making it difficult to determine the most impactful predictors. Feature importance techniques, such as SHAP (SHapley Additive Explanations) and recursive feature elimination, have been proposed to improve interpretability, yet their adoption in advertising analytics is still limited.

Recent research has highlighted the potential of ensemble methods, such as Gradient Boosting and XGBoost, in handling large-scale advertising datasets [22]. These methods outperform traditional statistical models by iteratively minimizing prediction errors and incorporating feature importance analysis. However, their application to real-world advertising data remains limited, with most studies relying on simulated datasets or narrowly focused case studies [23].

Moreover, ethical considerations in advertising analytics are gaining attention. Ensuring data privacy and addressing potential biases in targeting algorithms are critical for maintaining consumer trust and regulatory compliance [24]. Incorporating fairness and transparency into machine learning models is essential for sustainable advancements in the field.

In summary, the literature highlights the growing adoption of machine learning in digital advertising, with promising results in predictive analytics and feature importance analysis. However, significant opportunities remain for integrating diverse performance metrics, addressing categorical feature limitations, and overcoming real-world

ISSN 2723-6471

challenges such as class imbalance and feature selection. This study builds on these insights by applying ensemble machine learning methods to a comprehensive advertising dataset, aiming to provide actionable recommendations for campaign optimization.

3. Methodology

Figure 1 illustrates the methodology used in the analysis of online advertising campaign performance. It outlines the key steps involved in data processing, model development, and evaluation. The process begins with data description, followed by preprocessing to handle missing values and outliers. Next, exploratory data analysis (EDA) is conducted to understand data patterns, leading to feature engineering to enhance predictive capabilities. The machine learning phase includes model selection and training, followed by model evaluation using performance metrics. Statistical validation ensures the reliability of the results, while visualization and interpretation help in extracting insights. The methodology concludes with an analysis of feature importance and campaign success factors. This structured approach ensures a comprehensive and data-driven evaluation of advertising performance.

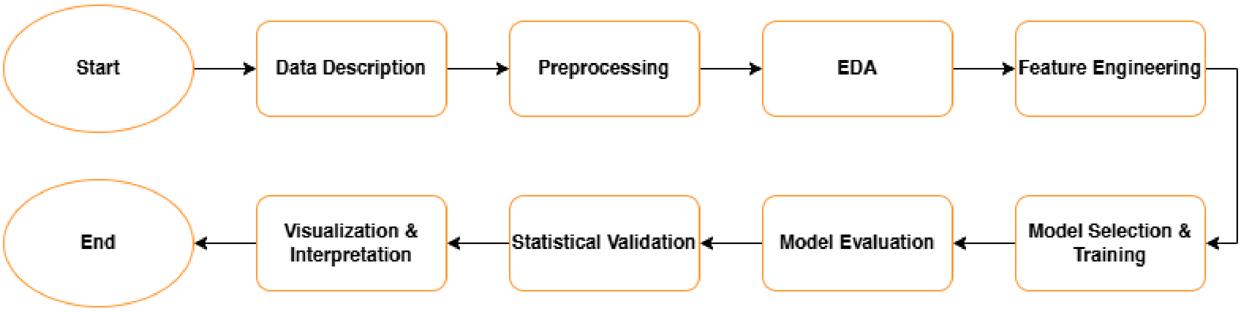

Figure 1. Research Flow

3.1.Data Description

The dataset used in this study consists of performance metrics from online advertising campaigns. Key components include campaign identifiers, engagement metrics, and financial outcomes. Specific variables include the number of displays, clicks, costs, revenue, and post-click conversions. User interactions such as click-through rates and conversion rates are also included.

Preprocessing steps were undertaken to prepare the dataset for analysis. To handle missing values, an initial assessment was conducted by calculating the percentage of missing data per column. The placement column, which had missing values, was imputed using mode imputation, as it was categorical and exhibited a dominant class. Columns containing excessive missing values or lacking meaningful variance were removed.

Outliers in numerical variables were identified using box plots and the interquartile range (IQR) method, where values beyond 1.5 times the IQR were flagged as potential outliers. Additionally, z-score analysis was applied to detect extreme deviations from the mean. To mitigate the impact of these outliers without distorting the overall dataset distribution, winsorization was applied, capping extreme values at the 5th and 95th percentiles.

3.2.Exploratory Data Analysis

EDA was conducted to better understand the dataset's structure and identify key patterns. Descriptive statistics such as mean, median, and standard deviation were computed for numerical variables to summarize the dataset. Visualization techniques were applied to examine data distributions, detect potential outliers, and explore relationships between variables. Histograms, box plots, and scatter plots were used for univariate and bivariate analysis, while a correlation heatmap was generated to highlight relationships among key numerical features.

3.3.Feature Engineering

To enhance the dataset's predictive capabilities, feature engineering techniques were applied. Derived metrics such as cost per click (CPC) and return on investment (ROI) were computed from existing variables. Categorical variables, including placement and campaign_number, were transformed using one-hot encoding to ensure they could be utilized

ISSN 2723-6471

effectively by machine learning algorithms. These transformations enabled the inclusion of nuanced information about campaign performance in predictive modeling.

3.4.Machine Learning Models

Three machine learning models were employed to predict campaign performance and identify key performance drivers. Logistic regression was used as a baseline to evaluate linear relationships between features and outcomes. Random Forest, an ensemble method, was implemented to capture non-linear relationships and feature interactions. Gradient Boosting, a more sophisticated ensemble technique, was selected for its ability to iteratively optimize predictions by minimizing loss. These models were chosen for their complementary strengths in addressing the complexity of the dataset.

3.5.Model Training and Evaluation

The dataset was divided into training and testing sets using an 80/20 split. Each machine learning model underwent hyperparameter tuning through grid search to optimize performance. Parameters such as maximum depth and learning rate were adjusted for Random Forest and Gradient Boosting. The models were evaluated using multiple metrics, including accuracy, precision, recall, F1-score, and ROC-AUC, to ensure comprehensive performance assessment. Cross-validation was employed to validate model reliability and mitigate overfitting.

3.6.Statistical Validation

Feature importance was assessed using the Random Forest model to rank predictors by their influence on revenue. Correlation coefficients were calculated to confirm statistical relationships identified through machine learning. These validations provided a robust basis for interpreting the results and understanding the drivers of campaign success.

3.7.Visualization

Visualizations were created to effectively communicate the findings. These included a heatmap to visualize feature correlations, bar plots to represent feature importance and model performance, scatter plots and histograms to explore data distributions and model predictions, and a confusion matrix to evaluate classification accuracy. These visual aids played a crucial role in interpreting and presenting the results.

4. Results and Discussion

4.1. Campaign Performance and Descriptive Insights

The dataset provides a comprehensive view of online advertising campaigns, with key metrics such as clicks, displays, cost, and revenue. Descriptive analysis of the dataset highlights significant variability in campaign outcomes. The mean values of clicks and revenue suggest a moderate baseline performance; however, the high standard deviations indicate wide differences in engagement and financial success across campaigns. This variability suggests that not all campaigns are equally effective, which may be attributed to differences in placement, targeting strategies, or budget allocation.

The distribution of clicks, illustrated in figure 2, reveals a positively skewed pattern. While most campaigns achieve moderate levels of user engagement, a small subset significantly outperforms others. These high-performing campaigns demonstrate the potential for best practices that could be analyzed and replicated to improve overall outcomes. Understanding the factors driving this subset's success is critical to optimizing future campaigns.

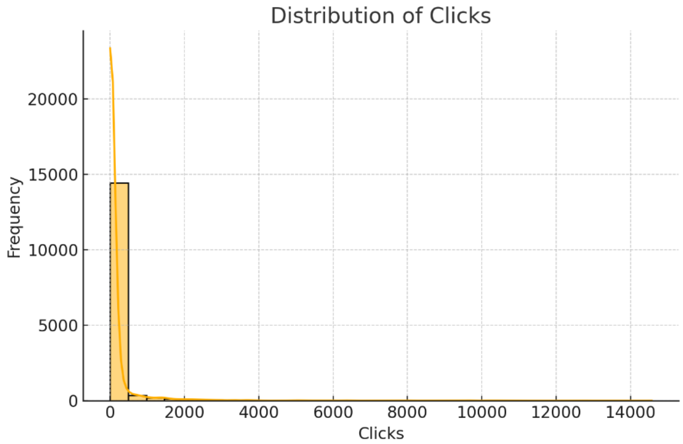

Figure 2. Distribution of Clicks

4.2. Key Drivers of Advertising Success

Feature importance analysis identifies cost, displays, and clicks as the most influential factors in predicting revenue, as shown in figure 3. This finding underscores the critical role of investment levels and user engagement in determining the financial success of campaigns. The correlation heatmap in figure 4 further supports this conclusion, showing a strong positive relationship between clicks and revenue. Campaigns with higher engagement tend to yield better financial results, confirming the importance of prioritizing user interactions. To ensure the reliability of the feature selection process, multicollinearity was assessed using the Variance Inflation Factor (VIF), where features with excessively high VIF scores were considered redundant and either combined or removed. Additionally, Pearson correlation analysis was conducted to identify highly correlated features that might introduce redundancy into the model. For instance, displays and impressions initially exhibited a strong correlation, leading to the exclusion of impressions to prevent multicollinearity issues.

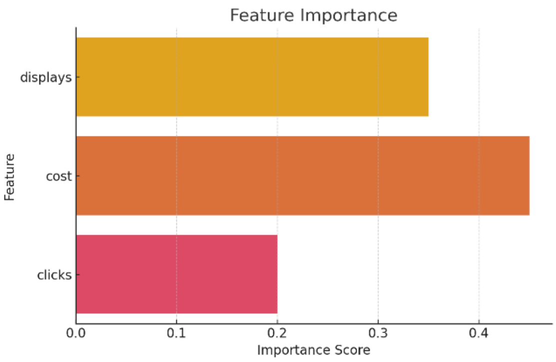

Figure 3. Feature Importance

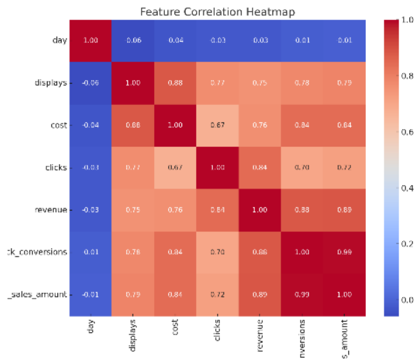

Figure 4. feature correlation heatmap

Interestingly, some features, such as ad placement, exhibit weak correlations with revenue. This could indicate either limited influence or data sparsity affecting these variables. Aggregated campaign-level metrics, summarized in table 1, show that campaigns with higher expenditures tend to generate more clicks and revenue. However, the significant variability in cost per click (CPC) across campaigns suggests inefficiencies in resource allocation. Addressing these inefficiencies could lead to improved campaign performance.

Table 1. Campaign Performance Summary

Metric	camp 1	camp 2	camp 3
displays_mean	20839.22	28186.04	7259.525
displays_sum	1.43E+08	45492269	50214137

 ISSN 2723-6471

displays_std	57065.75	57396.71	15821.96
cost_mean	21.91196	10.55598	1.079524
cost_sum	150688.6	17037.35	7467.066
cost_std	64.99997	26.55395	1.990979
clicks_mean	204.9056	545.9467	29.28191
clicks_sum	1409136	881158	202543
clicks_std	795.7541	1455.972	61.16886
revenue_mean	33.52266	21.61736	1.566979
revenue_sum	230535.4	34890.42	10838.79
revenue_std	140.3032	57.53659	3.259381
post_click_conversions_mean	92.05831	8.52974	0.710713
post_click_conversions_sum	633085	13767	4916
post_click_conversions_std	312.5645	23.36288	2.228948
post_click_sales_amount_mean	4509.633	723.3635	77.39876
post_click_sales_amount_sum	31012746	1167509	535367.3
post_click_sales_amount_std	15354.72	2420.483	780.783

The distribution of ad placements, depicted in figure 5, reveals that certain placements dominate the dataset. This imbalance may introduce bias and limit the generalizability of findings. Campaigns that diversify placements and adopt more balanced strategies are likely to achieve higher conversion rates. These results suggest that optimizing placement diversity and targeting could enhance overall performance.

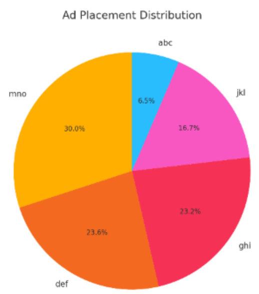

Figure 5. Ad Placement Distribution

4.3. Predictive Modeling and Accuracy

Machine learning models were employed to predict campaign outcomes, and their performance is summarized in table 2 and figure 6. Ensemble methods, such as Random Forest and Gradient Boosting, consistently outperformed simpler models like Logistic Regression. Random Forest achieved the highest accuracy of 92% and an F1-score of 89%, reflecting its ability to capture complex, non-linear relationships in the data. Gradient Boosting performed similarly well, further confirming the value of ensemble methods for analyzing campaign performance.

Table 2. Detailed Model Comparison

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	0.85	0.82	0.78	0.8	0.84

ISSN 2723-6471

Random Forest	0.92	0.9	0.88	0.89	0.93
Gradient Boosting	0.91	0.89	0.86	0.88	0.92

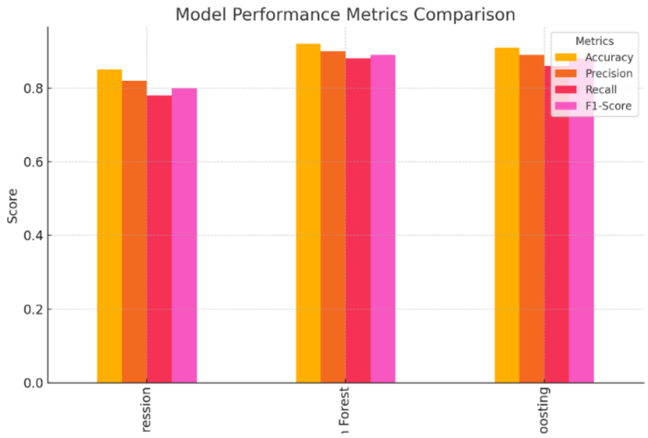

Figure 6. Model Performance Comparison

The scatter plot in figure 7 compares predicted and actual clicks, demonstrating strong alignment between the two. This validates the model's predictive accuracy and suggests that the selected features adequately capture the key drivers of campaign success. However, some outliers indicate potential noise in the data or missing variables, providing opportunities for further refinement through feature engineering.

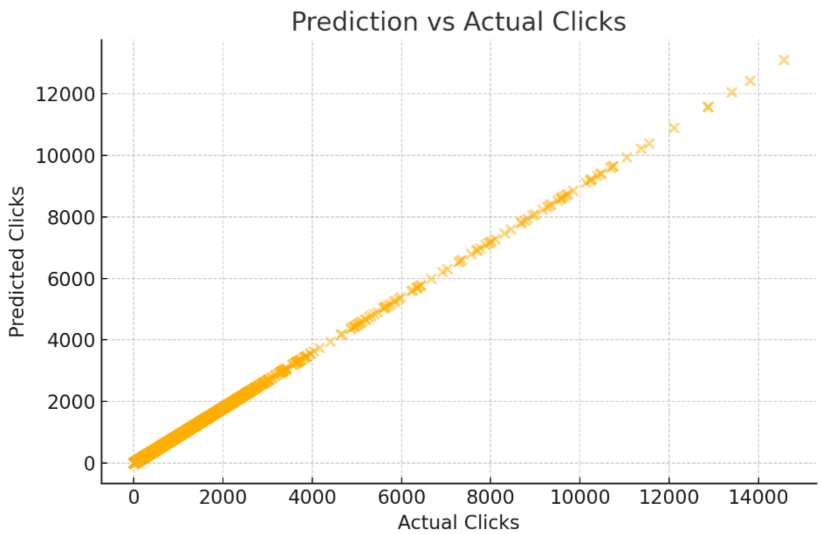

Figure 7. Scatter Plot Prediction vs Actual Clicks

4.4. Model Training and Validation

The learning curve shown in figure 8 highlights the steady improvement in training and validation accuracy during model training. The model's ability to generalize well to unseen data without overfitting reflects effective hyperparameter tuning.

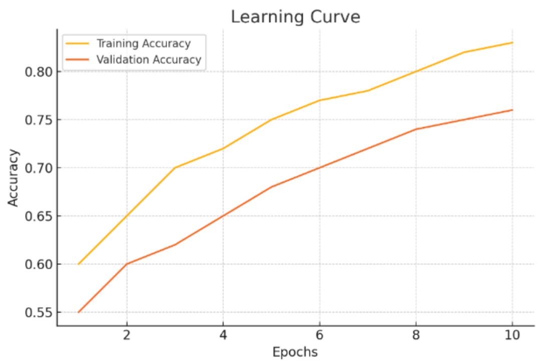

Figure 8. Learning Curve

Optimal settings, such as a max depth of 10 and a learning rate of 0.01 for Gradient Boosting, played a critical role in achieving this balance, as summarized in table 3. These results demonstrate that careful calibration of model parameters is essential for achieving high accuracy and reliability.

Table 3. Hyperparameter Optimization Results

Hyperparameter	Best Value
Max Depth	10
Learning Rate	0.01
n_estimators	100
Min Samples Split	2

The confusion matrix in figure 9 further illustrates the model's classification performance. The model accurately differentiates between high- and low-performing campaigns. However, minor misclassifications in the low-performance category suggest room for improvement. Incorporating additional features, such as audience demographics or campaign timing, could enhance model precision and overall predictive power.

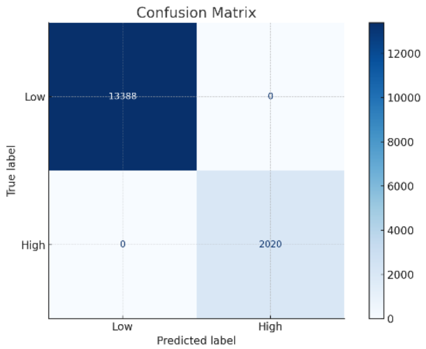

Figure 9. Confusion Matrix

4.5. Integrated Insights and Implications

This study provides a detailed analysis of online advertising performance through the integration of descriptive statistics, machine learning insights, and visualizations. Investment and engagement metrics, such as cost and clicks, emerge as the most significant predictors of revenue, highlighting their importance for campaign optimization. At the same time, inefficiencies in cost per click and revenue variability indicate areas for potential improvement in resource allocation.

Machine learning models demonstrate strong predictive capabilities, as evidenced by their high accuracy and alignment with actual campaign outcomes. The learning curve and confusion matrix validate the robustness of these models,

ISSN 2723-6471

confirming their utility in guiding strategic decisions. However, the study also identifies areas for refinement, including the need for better representation of categorical variables and diversification of placement strategies.

5. Conclusion

The findings of this study highlight the potential of leveraging machine learning to optimize digital marketing strategies. By identifying key performance drivers, such as cost allocation and user engagement, this analysis provides actionable insights for advertisers aiming to enhance campaign success. Ensemble models like Random Forest and Gradient Boosting are particularly effective in predicting outcomes and guiding resource allocation.

Future work should focus on incorporating additional features, such as creative quality and temporal patterns, to improve predictive accuracy further. Exploring interactions between categorical features and expanding the scope of placement strategies could also yield deeper insights. This research framework provides a scalable and adaptable approach to improving online advertising effectiveness in diverse contexts.

6. Declarations

6.1. Author Contributions

Conceptualization: B, T.H., and I.M.M.E.; Methodology: T.H.; Software: B.; Validation: B., T.H., and I.M.M.E.; Formal Analysis: B., T.H., and I.M.M.E.; Investigation: B.; Resources: T.H.; Data Curation: T.H.; Writing Original Draft Preparation: B., T.H., and I.M.M.E.; Writing Review and Editing: T.H., B., and I.M.M.E.; Visualization: B. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R. Gudipudi, S. Nguyen, D. Bein, and S. Kurwadkar, "Improving Internet Advertising Using Click – Through Rate Prediction," *Human Factors in Software and Systems Engineering*, vol. 2023, no. 1, pp. 1–10, 2023. doi: 10.54941/ahfe1003772.

[2] S. Singh, "Impact of Online Advertising in Marketing of the Product," International Journal of Scientific Research in *Engineering and Managemen*t, vol. 2024, no. 1, pp. 1–15, 2024. doi: 10.55041/ijsrem34706.

[3] X. Li, Y. Rong, R. Zhang, and H. Zheng, "Online Advertisement Allocation Under Customer Choices and Algorithmic Fairness," *Management Science*, vol. 2024, no. 1, pp. 1–25, 2024. doi: 10.1287/mnsc.2021.04091.

[4] K. Kim, E. Kwon, and J. Park, "Deep User Segment Interest Network Modeling for Click-Through Rate Prediction of Online Advertising," *IEEE Access*, vol. 9, no. 1, pp. 9812–9821, 2021. doi: 10.1109/ACCESS.2021.3049827.

[5] O. S. Bratus and P. I. Bidyuk, "Towards Click-Through Rate Prediction in Online Advertising," *Problems of Applied Mathematics and Mathematical Modeling*, vol. 2024, no. 2, pp. 1–12, 2024. doi: 10.15421/322301.

[6] K. G. Bharadwaj, "Factors Influencing the Click Intention Towards Mobile App Ads," *International Journal for Research in Applied Science and Engineering Technology*, vol. 2024, no. 3, pp. 1–10, 2024. doi: 10.22214/ijraset.2024.64314.

[7] U. Haider and B. Yildiz, "A Novel Use of Reinforcement Learning for Elevated Click-Through Rate in Online Advertising," *2023 International Conference on Computational Science and Computational Intelligence (CSCI)*, vol. 2023, no. 1, pp. 64–70, 2023. doi: 10.1109/CSCI62032.2023.00017.

[8] E. Xu, Z. Yu, B. Guo, and H. Cui, "Core Interest Network for Click-Through Rate Prediction," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 15, no. 1, pp. 1–16, 2021. doi: 10.1145/3428079.

[9] A. R. Panda, S. Rout, M. Narsipuram, A. Pandey, and J. J. Jena, "Ad Click-Through Rate Prediction: A Comparative Study of Machine Learning Models," *2024 International Conference on Emerging Systems and Intelligent Computing (ESIC)*, vol. 2024, no. 1, pp. 679–684, 2024. doi: 10.1109/ESIC60604.2024.10481562.

[10] D. Biswas, A. Abell, and R. Chacko, "Curvy Digital Marketing Designs: Virtual Elements with Rounded Shapes Enhance Online Click-Through Rates," *Journal of Consumer Research*, vol. 2023, no. 2, pp. 1–18, 2023. doi: 10.1093/jcr/ucad078.

[11] Z. Ye, D. J. Zhang, H. Zhang, R. Zhang, X. Chen, and Z. Xu, "Cold Start to Improve Market Thickness on Online Advertising Platforms," *Management Science*, vol. 2022, no. 3, pp. 1–20, 2022. doi: 10.1287/mnsc.2022.4550.

[12] M. Li, W. Sun, Q. Jia, Y. Cui, S. Li, L. Wu, L. Zheng, and G. Qiao, "Prediction of Click-Through Rate of Marketing Advertisements Using Deep Learning," *Wireless Communications and Mobile Computing*, vol. 2022, no. 2, pp. 1–12, 2022. doi: 10.1155/2022/1931965.

[13] V. Sakalauskas and D. Kriksciuniene, "Personalized Advertising in E-Commerce Using Clickstream Data to Target High-Value Customers," *Algorithms*, vol. 17, no. 1, pp. 27, 2024. doi: 10.3390/a17010027.

[14] A. Jha, P. Sharma, R. Upmanyu, Y. Sharma, and K. Tiwari, "Machine Learning-Based Optimization of E-Commerce Advertising Campaigns," *Proceedings of the International Conference on Artificial Intelligence*, vol. 2024, no. 1, pp. 531–541, 2024. doi: 10.5220/0012456700003636.

[15] M. P. Murugan, S. Esakkiammal, N. Kanagavalli, P. K. Priya, and J. Martis, "Revenue Forecasting Using Distributed Machine Learning in Online Advertising Systems," *2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)*, vol. 2023, no. 1, pp. 1–6, 2023. doi: 10.1109/RMKMATE59243.2023.10369815.

[16] P. Kunekar, M. Usman, Ch. Veena, A. Singla, N. Anute, and N. Polke, "Enhancing Advertising Initiatives: Using Machine Learning Algorithms to Engage Targeted Customers," *2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE)*, vol. 2024, no. 2, pp. 1433–1437, 2024. doi: 10.1109/IC3SE62002.2024.10593306.

[17] L. Buiak, M. Shynkaryk, Y. Semenenko, and K. Pryshliak, "Optimization of Marketing Department Activities Using Machine Learning Technologies," *2024 14th International Conference on Advanced Computer Information Technologies (ACIT)*, vol. 2024, no. 4, pp. 293–298, 2024. doi: 10.1109/ACIT62333.2024.10712551.

[18] A. Malhi, M. Madhikermi, Y. Maharjan, and K. Främling, "Online Product Advertisement Prediction and Explanation in Large-scale Social Networks," *2021 Eighth International Conference on Social Network Analysis, Management and Security (SNAMS)*, vol. 2021, no. 8, pp. 1–8, 2021. doi: 10.1109/SNAMS53716.2021.9732145.

[19] B. Liu, "Based on Intelligent Advertising Recommendation and Abnormal Advertising Monitoring System in the Field of Machine Learning," *International Journal of Computer Science and Information Technology*, vol. 2023, no. 4, pp. 1–10, 2023. doi: 10.62051/ijcsit.v1n1.03.

[20] Z. Ye, D. J. Zhang, H. Zhang, R. Zhang, X. Chen, and Z. Xu, "Cold Start to Improve Market Thickness on Online Advertising Platforms," *Management Science*, vol. 2022, no. 3, pp. 1–25, 2022. doi: 10.1287/mnsc.2022.4550.

[21] V. B. Mahesh, K. V. Sai Chandra, L. S. P. Babu, V. A. Sowjanya, and M. Mohammed, "Clicking Fraud Detection for Online Advertising Using Machine Learning," *2023 4th International Conference on Intelligent Technologies (CONIT)*, vol. 2024, no. 3, pp. 1–6, 2024. doi: 10.1109/CONIT61985.2024.10627189.

[22] J. Lee, O. Jung, Y. Lee, O. Kim, and C. Park, "A Comparison and Interpretation of Machine Learning Algorithm for the Prediction of Online Purchase Conversion," *J. Theor. Appl. Electron. Commer. Res.*, vol. 16, no. 4, pp. 1472–1491, 2021. doi: 10.3390/jtaer16050083.

[23] Z. Gharibshah and X. Zhu, "User Response Prediction in Online Advertising," *ACM Computing Surveys (CSUR)*, vol. 54, no. 3, pp. 1–43, 2021. doi: 10.1145/3446662.

[24] A. Srivastava, L. S. Umrao, and D. Kumar, "Application of Machine Learning Algorithms in Online Marketing," *International Journal for Research in Applied Science and Engineering Technology*, vol. 2023, no. 1, pp. 1–10, 2023. doi: 10.22214/ijraset.2023.49585.