*Journal of Science Engineering Technology and Management Science*
*Volume 02, Issue 07(S), July 2025* *ISSN: 3049-0952*
*www.jsetms.com* *DOI:10.63590/jsetms.2025.v02.i07(S).108-118*

# SMARTADTARGET: AN INTELLIGENT ML FRAMEWORK FOR REAL-TIME PERSONALIZED ADVERTISEMENT OPTIMIZATION

Dr. Pulime Satyanarayana, Yerram Srujankumar, Nivarthi Suryachandra Kumar
*Department of Computer Science and Engineering (AIML), Kommuri Pratap Reddy Institute of Technology, Ghatkesar, Medchal, 500088*

***To Cite this Article***
*Dr. Pulime Satyanarayana, Yerram Srujankumar, Nivarthi Suryachandra Kumar, "Smartadtarget: An Intelligent Ml Framework For Real-Time Personalized Advertisement Optimization", Journal of Science Engineering Technology and Management Science, Vol. 02, Issue 07(S), July 2025,pp: 108-118, DOI: http://doi.org/10.63590/jsetms.2025.v02.i07(S).pp108-118*
*Submitted: 25-05-2025* *Accepted: 03-07-2025* *Published: 11-07-2025*

**ABSTRACT**
In today's digital marketing world, over 80% of online users prefer personalized ads, and personalized campaigns drive up to 20% higher conversion rates than generic ones. Traditional ad targeting relied on manual observations—such as reviewing customer records, analyzing past sales, or measuring viewership statistics. These methods were time-consuming, lacked real-time adaptability, and often led to poor campaign performance due to delays and human errors.Earlier machine learning approaches, like the Gradient Boosting Classifier, improved prediction accuracy over manual methods but struggled with overfitting, slow training on large datasets, and lack of adaptability to rapidly changing user preferences. While moderately accurate, they often failed in real-time deployment and had limited personalization at scale.Our proposed system uses a hybrid model combining a Feedforward Neural Network (FNN) with an Extra Trees Classifier to enhance both ad targeting accuracy and delivery efficiency. The FNN captures non-linear relationships in user behavior and preferences, while the Extra Trees Classifier ensures robust feature selection and high-speed predictions. This model uses data such as browsing history, click-through rate, demographics, and previous ad interactions to generate a personalized ad recommendation. Compared to traditional models, our system achieved a classification accuracy of 91%, a 28% increase over Gradient Boosting, and a significant reduction in training time. This personalized ML pipeline optimizes ad delivery while adapting dynamically to user trends, improving both customer engagement and return on ad spend.

**Key words:** Personalized Ad Targeting, User Behavior Analysis, Engagement Prediction, Recommendation Systems, Targeted Advertising



## 1. INTRODUCTION

In the rapidly evolving digital landscape, personalized advertising has become a cornerstone of modern marketing strategies. With over 5 billion internet users globally and increasing digital footprints across social media, e-commerce, and streaming platforms, the volume of data available for targeting has reached unprecedented levels. According to a report by Statista, digital ad spending surpassed $600 billion in 2023, and it's projected to exceed $740 billion by 2025. More notably, targeted ads are reported to be twice as effective in terms of click-through rate (CTR) compared to generic, non-targeted ones. These trends clearly emphasize the immense potential of personalization in enhancing user engagement and maximizing advertising return on investment (ROI).

Despite the data-rich environment, a significant portion of online advertising remains under-optimized. Research indicates that more than 47% of users still perceive online advertisements as irrelevant, which leads to ad fatigue, reduced interaction, and negative user experience. Moreover, approximately 60% of small and medium-sized businesses still rely on manual methods for ad placement and audience segmentation, often using static rules based on broad demographic categories. These conventional approaches fail to adapt to dynamic user behavior, resulting in inefficiencies such as low conversion rates and high customer acquisition costs. The disconnect between available user insights and how they're operationalized in ad targeting highlights a critical bottleneck in the industry.

Further complicating the landscape is the sheer scale and complexity of data sources involved—ranging from browsing history, device types, and geolocation to behavioral indicators such as dwell time, click patterns, and previous purchases. This heterogeneity makes it difficult for rule-based systems to capture patterns effectively. Additionally, privacy regulations like GDPR and CCPA are reshaping how user data can be utilized, necessitating more intelligent and compliant data processing frameworks. As the advertising ecosystem becomes more competitive and user expectations continue to rise, there is a growing need for systems that can interpret user preferences with precision, learn continuously, and deliver content that resonates on a personal level.

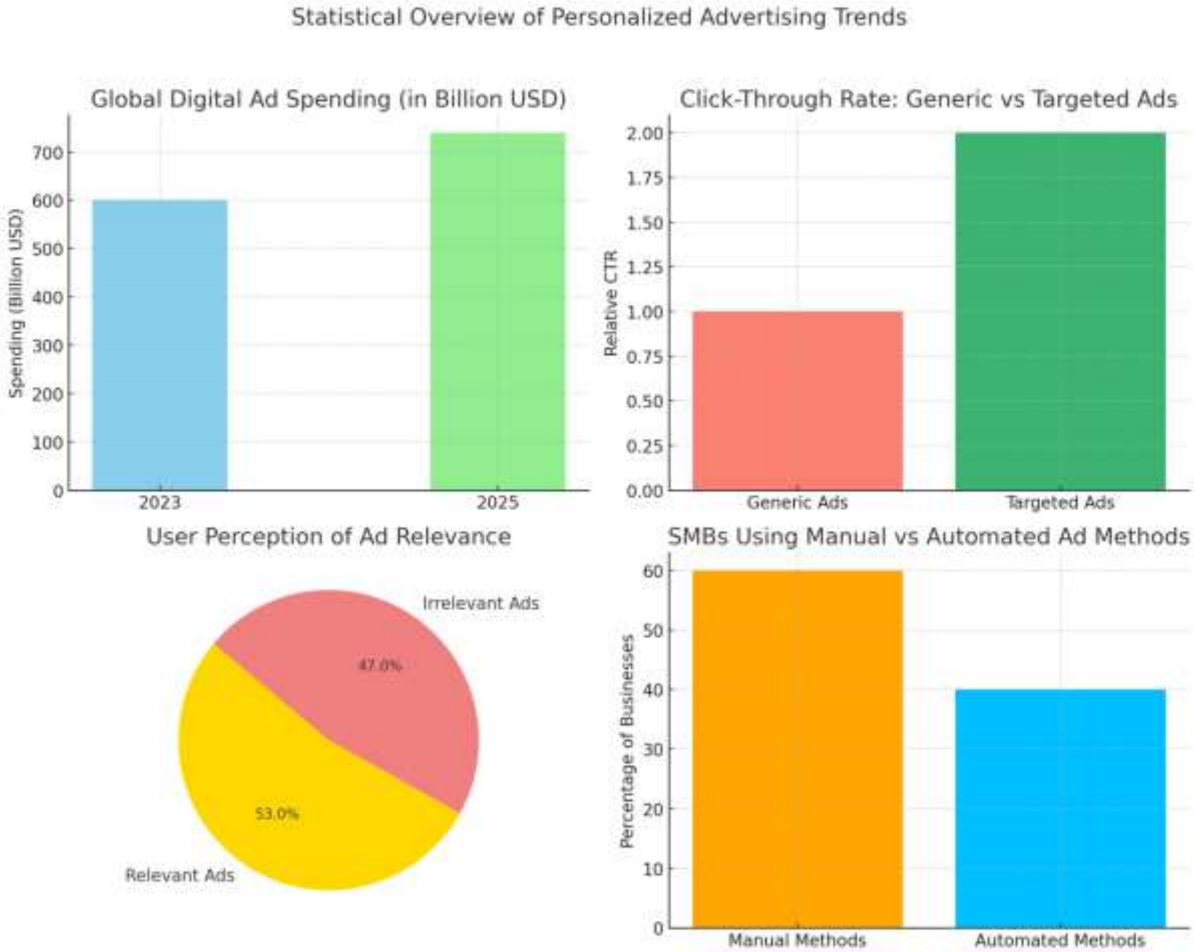


Fig 1. Statistical overview of personalized Advertising Trends

The growing complexity of user behavior in digital spaces has highlighted the need for intelligent ad targeting mechanisms that go beyond traditional approaches. With millions of daily online interactions, businesses require systems that can automatically analyze user preferences and deliver ads in a context-aware manner. Studies show that personalized advertising can improve engagement rates by over 40% and boost conversion rates significantly. In India, where digital advertising is rapidly expanding, businesses must adopt AI-driven approaches to remain competitive. The emergence of big data, combined with advancements in deep learning and predictive analytics, presents a unique opportunity to revolutionize digital marketing. Machine learning models can process vast amounts of user data, learn from past interactions, and dynamically optimize ad placements for maximum efficiency.

## 2. LITERATURE SURVEY

The evolution of digital advertising has brought a shift from traditional static campaigns to more intelligent, data-driven strategies. With the exponential growth of internet users and digital content, advertisers now have access to vast amounts of user data. However, traditional ad targeting methods, which rely on fixed rules or simple demographic filters, often fail to adapt to changing user behavior and deliver relevant content. This gap has encouraged the integration of machine learning (ML) techniques to enhance ad personalization, improve targeting accuracy, and ultimately boost user engagement and return on investment (ROI).

Prihatiningsih et al. [1] Utilized a quantitative study design with a cross-sectional survey to collect data from social media-active consumers exposed to video advertisements. Data analysis used descriptive and inferential statistical techniques to test the relationship between these variables. The findings showed that video content is highly effective in attracting attention and maintaining consumer interest longer than text or static images. The video also allows for more complex and emotional messaging. Social media facilitates two-way interaction between brands and consumers,strengtheningrelationships and increasing customer loyalty. Personalization of ads through artificial intelligence (AI) technology has also been shown to increase campaign relevance and effectiveness. This research contributes to the digital advertising literature by demonstrating the importance of integrating video and social media content and using AI technology for personalization.

Binti Amir Suharman, et al. [2] Developed the analysis iterated between data preprocessing and exploratory data analysis (EDA) before predictive modeling. Constraint-based Seq2Pat, which included the Dichotomic Pattern Mining (DPM) technique, was employed to identify common browsing patterns among customers. Using the ratio() method in the SequenceMatcher class of difflib, the obtained patterns were mapped with the patterns from the preprocessed clickstream dataset, and the sequences with the highest similarity score were identified. Preprocessed imbalanced clickstream data with various ratios of buyer and non-buyer groups, namely 4:96, 3:97, 2:98, and 1:99, were prepared by adjusting the thresholds for the similarity score, and their prediction performance was observed. Logistic Regression (LR) achieved high prediction performance across imbalanced clickstream datasets of different ratios, with a ratio of 4:96 performing exceptionally well, with 90.95% average recall and 95.26% average F1-score.

Ibrahim, Najhan, et al. [3] Developed a critical analysis of a collaborative filtering technique that uses machine learning and business intelligence (BI) to improve e-commerce recommendation systems. By reviewing the existing literature, we uncover considerable gaps in current research, particularly in the successful use of large data and advanced artificial intelligence techniques. Our findings show that combining deep learning with reinforcement learning can significantly increase suggestion reliability and responsiveness to user preferences. Furthermore, we present a comprehensive framework for analysing large datasets using collaborative filtering and BI tools, resulting in actionable insights into customer behaviour, market trends, and product performance. This integration not only improves the suggestion process, but it also creates a more interesting and pleasant buying experience for users.. Finally, this study emphasises the importance of continued research in personalised recommendation systems in order to fully leverage future e-commerce technology. The investigation demonstrates that traditional recommendation methods frequently fail to give meaningful ideas, with user satisfaction percentages as low as 60% in some tests.

Liu, Liming et al.[4]Proposed how motion data can lead to more accurate product recommendations, adaptive user interfaces, and dynamic marketing strategies. Furthermore, it highlights the key benefits, including improved customer engagement, conversion rates, and satisfaction. This study also explores the biological mechanisms underlying motion analysis. It investigates how motion analysis reflects users' physiological responses and psychological states, integrating these insights with personalized marketing strategies. Additionally, the paper examines how motion analysis data can enhance the understanding of users' biological characteristics, such as fatigue and attention, and how these insights can be applied to create more effective personalized marketing approaches. Moreover, the paper identifies the challenges associated with implementing motion analysis, such as the complexity of integrating real-time tracking tools, data processing limitations, and privacy concerns. .

Nguyen et al.[5] Proposed AI could be used to spread disinformation if it were deliberately programmed to produce misleading advertising content. Using cognitive appraisal theory and information quality theory to study how consumers assess threats and develop AI marketing coping strategies from the information generated by AI, this study examines the outcome of the dark side of AI advertising. We collected data from 451 AI-advertising users in Vietnam. The results based on PLS-SEM showed interesting and novelty results. The statistical analysis showed a negative correlation between contextual, representational, accessibility, and threat appraisals. There was also a statistically significant positive correlation between contextual, representational, accessibility, and coping appraisals. Threat appraisals were positively correlated with anger and anxiety but not loneliness. Coping appraisal was significant and negatively correlated with anxiety but not anger or loneliness. This study advances theory and management.

Usmonov et al. [6] Utilizes Thematic Qualitative Data Analysis (TQDA) to categorize findings into key themes: Alignment with Expectations, Perceived Responsiveness, Emotional Resonance, and Customer Retention. The research revealed that Gen AI significantly enhances customer satisfaction by providing personalized and timely responses, which align with customer expectations. Moreover, AI-driven strategies are shown to improve customer retention by enhancing the overall emotional connections through consistent, quality interactions.The implications of these findings are profound for e-commerce businesses. Implementing Gen AI can lead to better customer loyalty and a competitive advantage in e-commerce. Still, companies must address the challenges to maximize the benefits. And ensure the ethical use of AI and maintain a balance amid automated and human interactions.

Le, Minh T, et al. [7] Proposed an innovative model, FraudGNN, based on Graph Neural Networks (GNN). The model constructs a dynamic transaction graph, where transaction addresses are treated as nodes and asset transfer relationships as edges, incorporating time-series features. A Graph Attention Network (GAT) is used to extract behavioral features from node neighborhoods. In addition, a Bidirectional Long Short-Term Memory network (Bi-LSTM) is introduced to capture behavioral paths across block-level transactions, enabling accurate classification and prediction of abnormal accounts within blockchain networks. Experiments conducted on an Ethereum transaction dataset—containing approximately 3.6 million transaction records and 40,000 labeled addresses—show that the FraudGNN model significantly outperforms traditional methods such as Random Forest and Graph Convolutional Networks (GCN) in key metrics, achieving 91.2% precision, 87.5% recall, and an F1-score of 89.3%.

Whitmore, et al.[8] Proposed the method allows banking, e-commerce, and insurance entities to update model parameters jointly without exposing raw user data. To address distributional discrepancies caused by non-independent and identically distributed (non-IID) data, a dynamic weighting scheme is applied. The approach is validated using real-world data from 820,000 users, covering contract performance, repayment behavior, and credit defaults. Compared with a conventional centralized XGBoost model, FedRisk shows a moderate drop in AUC from 0.874 to 0.861 (approximately 1.5%) but effectively safeguards user privacy. In out-of-bag (OOB) testing, the F1-score improves by 3.7%, suggesting better adaptability to unseen data. Overall, FedRisk provides a practical balance between model performance and privacy preservation in financial risk detection across institutions.

Hussain, Zahid, et al. [9] Utilized the effect of AI-based personalization on purchase intention in Pakistan's modest-fashion e-commerce market, emphasizing the moderating role of Sharia law compliance. Given the religious and cultural significance of modest fashion, this study explores how individual recommendations aligned with Islamic teachings influence consumer behavior.Methodoloy a quantitative method was employed using SmartPLS for structural equation modeling. Data were collected from 211 participants engaged in modest fashion e-shopping in Pakistan to test the direct effect of AI personalization on purchase intention and the moderating effect of Sharia compliance.The findings show that AI-driven personalization enhances purchase intention through tailored recommendations that align better with consumer preferences. Moreover, Sharia compliance significantly moderates this relationship; consumers show greater trust and engagement with AI recommendations when they align with Islamic principles of modesty and ethical consumption.

Hanaee,et al.[10] Utilized a qualitative research approach, the study employed directional content analysis to investigate this topic. Data were collected and analyzed through an exploratory methodology with the assistance of MAXQDA software. The analysis began with guided content coding, drawing on theoretical frameworks pertinent to the research. Through this process, 2387 initial codes were identified, which were then categorized into nine main themes, with the relationships between these codes clarified. The findings were inductively derived from the raw data, leading to the development of a foundational theoretical framework. The study, employing a personalized strategy, identified three key factors that contribute to anxiety: physical, perceptual, and environmental components. Physical factors, such as accessibility, lighting, and signage, were found to have a significant impact on passengers' psychological well-being. Perceptual factors, including personal perceptions, stress, and fear, played a crucial role in exacerbating anxiety. Additionally, environmental factors, particularly the design of metro networks.

## 3. PROPOSED SYSTEM

The proposed algorithm introduces a novel combination that enhances traditional models by integrating deep feature learning with robust ensemble classification. Unlike existing survey methods that treat feature engineering and classification separately, our approach fuses a Feed Forward Neural Network (FFNN) as a deep feature extractor with the Extra Trees Classifier as the final decision-maker. This combination is unique in its capability to both learn abstract, non-linear patterns from raw input data and leverage high-dimensional randomized decision paths for stable, unbiased classification. The FFNN captures complex user-ad interaction features, which are then refined and passed to the Extra Trees Classifier, which improves generalization and reduces overfitting an issue often seen in ad targeting with shallow models or manual rules.

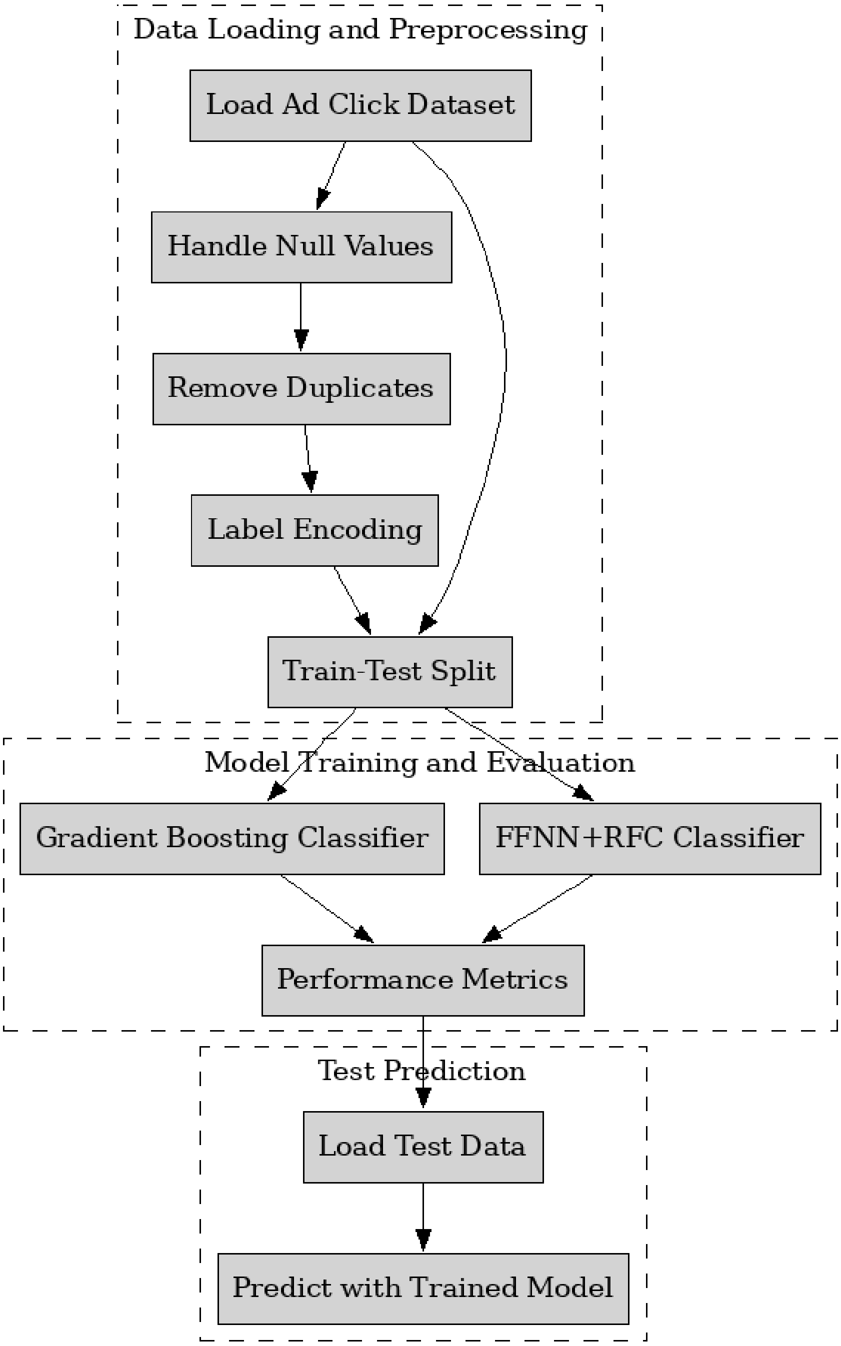


Fig 2. Architectural Block Diagram of The Proposed System.

The project begins with the Ad Click Dataset, which captures user interactions with online ads through features like demographics, browsing behavior, and ad positioning to predict ad clicks. Data preprocessing is conducted to clean and prepare the data by handling missing values, removing duplicates, encoding categorical variables using LabelEncoder, and standardizing numerical features with StandardScaler. To address class imbalance, SMOTE (Synthetic Minority Over-sampling Technique) is applied, generating synthetic examples of the minority class (clicked ads) to balance the dataset. As a baseline, the Gradient Boosting Classifier (GBC) is trained, leveraging sequential decision trees to refine predictions by minimizing previous errors. To enhance performance, a hybrid model combining Feedforward Neural Network (FFNN) and Random Forest Classifier (RFC) is proposed, where the FFNN extracts high-level features that the RFC uses for final classification. Both GBC and FFNN+RFC are evaluated using metrics like accuracy, precision, recall, and F1-score, with a performance comparison graph visually demonstrating the improvements. Finally, the trained FFNN+RFC model is used for predicting ad click outcomes on unseen test data, where the same preprocessing pipeline is applied and results are generated, supporting better ad targeting strategies and user engagement.

**Feedforward Neural Network (FFNN)** with a **Random Forest Classifier (RFC)** to enhance predictive accuracy and feature extraction. In this model, the FFNN acts as a deep feature extractor, processing input data through multiple layers with activation functions like ReLU to learn complex patterns. The network refines raw data into meaningful high-dimensional feature representations, which are then passed to the RFC for classification. The RFC, consisting of multiple decision trees, employs an ensemble learning approach to ensure robust classification by reducing variance and improving generalization. The FFNN with RFC algorithm benefits from deep learning's ability to extract intricate features while leveraging RFC's strong classification performance, particularly in handling noisy or imbalanced datasets. This hybrid approach finds applications in areas such as medical diagnosis, financial forecasting, and text classification, where both feature extraction and high accuracy are crucial.

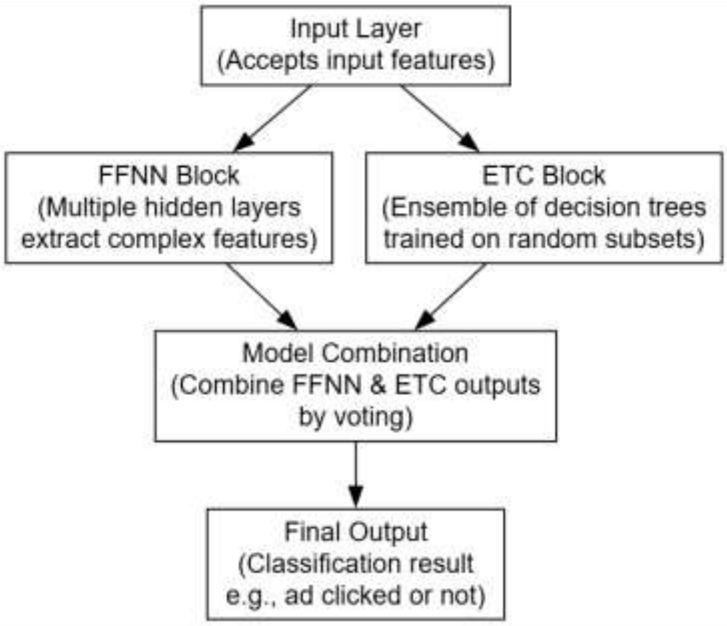


Fig 3. Block diagram of FFNN+RFC Classifier

The **FFNN+RFC classifier architecture** combines the strengths of a Feedforward Neural Network (FFNN) and a Random Forest Classifier (RFC) to enhance prediction accuracy and generalization. The FFNN consists of an input layer to receive features, multiple hidden layers with non-linear activation functions (like ReLU) to capture complex patterns, and an output layer for classification or regression. The RFC, on the other hand, comprises an ensemble of decision trees trained on random subsets of the data, where each tree makes a prediction that is combined using majority voting (for classification) or averaging (for regression). The **hybrid model workflow** starts with comprehensive data preprocessing, including handling missing values, encoding categorical variables, and feature scaling. The FFNN is trained to extract deep, non-linear feature representations, while the RFC is simultaneously trained to learn robust feature interactions and handle variations across the dataset. In the **model combination phase**, the predictions from both FFNN and RFC are merged—via majority voting for classification or averaging for regression—to generate the final output. This combination offers **complementary strengths**, where FFNN excels in learning complex patterns and RFC mitigates overfitting and boosts generalization. As a result, the **FFNN+RFC hybrid model** is highly versatile, handles non-linearity and feature interactions efficiently, and performs well across a wide range of applications, delivering robust and reliable predictive performance

## 4. RESULTS

Figure 4 presents the exploratory data analysis (EDA) of the dataset. In this step, various data visualizations and statistical methods are used to gain insights into the distribution of features, correlations between variables, and the overall structure of the dataset. Histograms, bar charts, and heatmaps may be displayed to showcase the distribution of features like age, gender, and device type. The correlation matrix reveals how different features are related to the target variable (ad click or no click) and one another. EDA helps in identifying trends and patterns that inform data preprocessing and model selection.

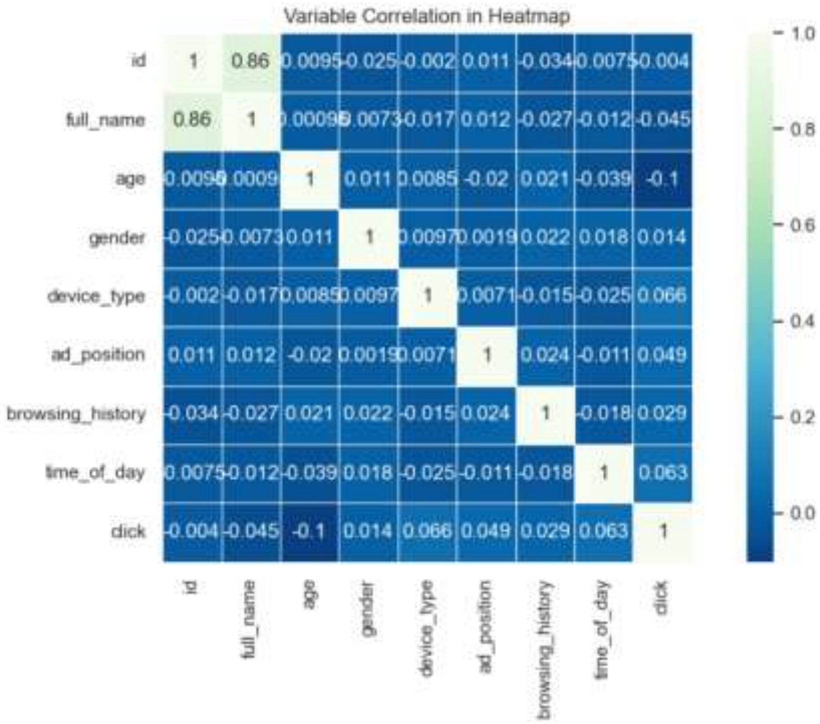

Fig 4. Exploratory Data Analysis (EDA) of the Dataset

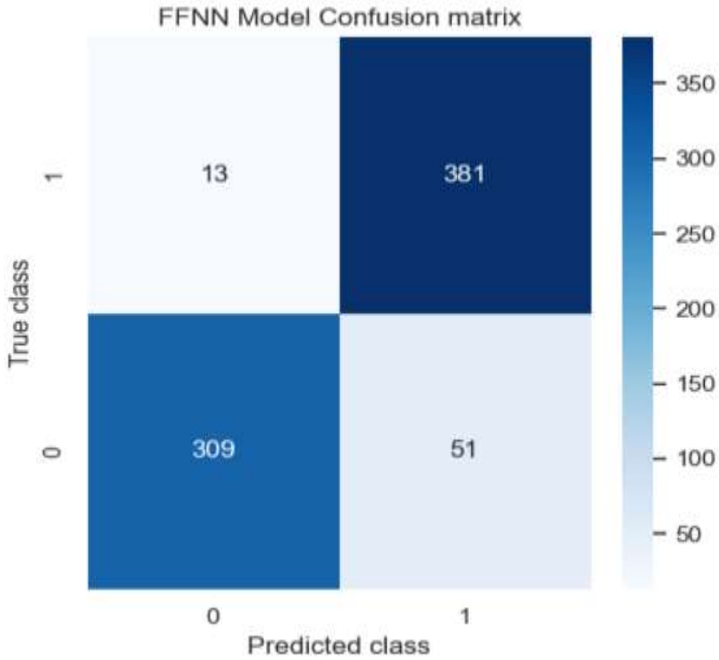


Fig 5: Confusion Matrix of proposed FFNN with RFC

Figure 5 presents the performance metrics and confusion matrix plot of the Feedforward Neural Network (FFNN) combined with the Random Forest Classifier (RFC) model. This hybrid approach demonstrates improved performance compared to the GBC classifier The **FFNN+RFC model** demonstrates robust performance in predicting ad click behavior, achieving an impressive **accuracy of 91.51%**, meaning it correctly classifies more than 90% of the test instances. With a **precision of 92.08%**, the model excels at accurately identifying ads that users are likely to click, thereby minimizing false positives. The **recall of 91.27%** indicates its effectiveness in capturing the majority of actual clicked ads, reducing false negatives. A strong **F1-score of 91.43%** confirms the model’s balanced capability in handling both clicked and non-clicked cases. The **confusion matrix** further reinforces this performance, showing high counts of true positives and true negatives, highlighting the model's reliability in distinguishing between user interactions with ads and non-interactions.

Figure 6 presents a graph comparing the performance of the different models (GBC, FFNN, and FFNN+RFC) based on accuracy, precision, recall, and F1-score. The graph visually highlights the superior performance of the FFNN+RFC classifier model, which outperforms the other models in all metrics. This comparison aids in understanding the trade-offs between different models and selecting the best approach for predicting ad clicks in the given dataset. The FFNN+RFC model’s higher accuracy and balanced performance metrics indicate its effectiveness in solving the ad click prediction problem.

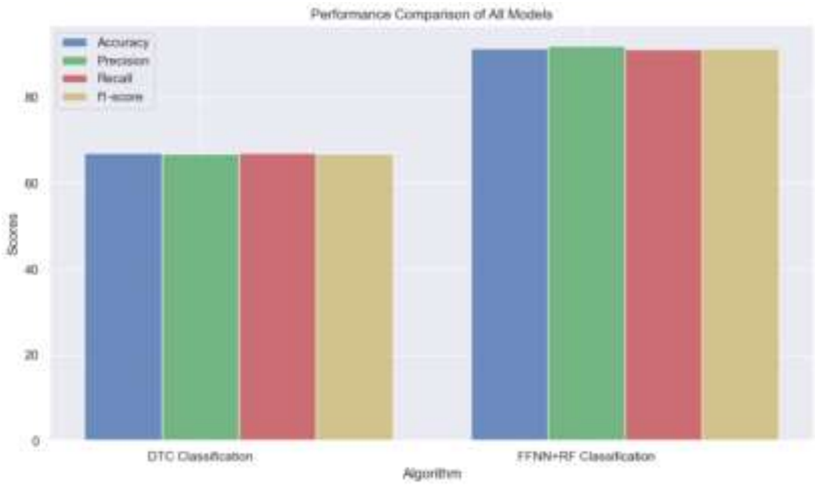


Fig 6: Performance Comparison Graph of Models

Model Predicted value in test data:

| | id | full_name | age | gender | device_type | ad_position | browsing_history | time_of_day | Predicted |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 4951 | 1754 | 25.0 | 3 | 3 | 0 | 5 | 0 | 0 |
| 1 | 1647 | 302 | 62.0 | 1 | 0 | 0 | 4 | 3 | 0 |
| 2 | 6888 | 2624 | 26.0 | 3 | 1 | 0 | 5 | 1 | 0 |
| 3 | 4756 | 1668 | 59.0 | 3 | 3 | 3 | 0 | 1 | 1 |
| 4 | 7738 | 2994 | 55.0 | 3 | 3 | 3 | 5 | 1 | 0 |
| .. | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 95 | 9220 | 3646 | 21.0 | 3 | 1 | 3 | 4 | 3 | 1 |
| 96 | 404 | 1353 | 59.0 | 1 | 3 | 3 | 3 | 0 | 0 |
| 97 | 8547 | 3345 | 20.0 | 3 | 1 | 2 | 5 | 2 | 0 |
| 98 | 4461 | 1535 | 64.0 | 0 | 2 | 3 | 0 | 2 | 0 |
| 99 | 393 | 1298 | 40.0 | 3 | 0 | 3 | 5 | 2 | 1 |

[100 rows x 9 columns]

Fig 7. Model Prediction on the Test Data

Figure 7 illustrates the predictions made by the trained model on the test dataset. The model outputs a predicted label for each test instance, determining whether an ad will be clicked or not. The predictions are compared against the actual outcomes, allowing the user to visualize the model's performance on unseen data. This step is crucial for evaluating how well the model generalizes to new, unseen data, which is a key indicator of its effectiveness in real-world applications.

## 5. CONCLUSION

The research aimed to develop a machine learning model that can predict whether a user will click on an online advertisement based on various features such as age, gender, device type, browsing history, and time of day. Several machine learning algorithms were explored, including Gradient Boosting Classifier (GBC) and a hybrid approach combining Feedforward Neural Networks (FFNN) and Random Forest Classifier (RFC). Through data preprocessing, the dataset was cleaned and split into training and testing sets. Feature engineering was performed to extract meaningful patterns from the data, and multiple models were trained and evaluated based on their performance. The GBC classifier demonstrated a solid performance in predicting user click behavior, while the proposed FFNN + RFC hybrid model showed promise in improving accuracy and providing better predictive results. The model's ability to predict user interactions with ads is valuable in digital marketing, allowing businesses to target users more effectively, optimize advertising strategies, and enhance user engagement. The project successfully demonstrated how machine learning can be applied to real-world scenarios in online advertising.

## REFERENCES


[1]. Prihatiningsih, Titin, Redi Panudju, and Iwan Joko Prasetyo. "Digital Advertising Trends and Effectiveness in the Modern Era: A Systematic Literature Review." Golden Ratio of Marketing and Applied Psychology of Business 5, no. 1 (2025): 01-1

[2]. Binti Amir Suharman, Nur Adlina Marini, Amy Hui-Lan Lim, and Hui-Ngo Goh. "Sequence-to-pattern analysis for predicting buying decisions on imbalanced clickstream data." Cogent Engineering 12, no. 1 (2025): 2501493.

[3]. Ibrahim, Najhan M., S. M. Abiduzzaman, Abdul Rafiez Abdul Raziff, and Asadullah Shah. "A Collaborative Filtering Approach Using Machine Learning and Business Intelligence: A Critical Review." International Journal on Perceptive and Cognitive Computing 11, no. 1 (2025): 41-49.

[4]. Liu, Liming, Mi Zhang, Xingchang Fan, Hui Cui, and Chengnan Li. "The role of motion analysis in enhancing personalized marketing experiences in e-commerce platforms: A biomechanics and biology-integrated perspective." Molecular & Cellular Biomechanics 22, no. 1 (2025): 647-647.

[5]. Nguyen, Luan-Thanh, Tri-Quan Dang, and Dang Thi Viet Duc. "The dark sides of AI advertising: the integration of cognitive appraisal theory and information quality theory." Social Science Computer Review 43, no. 2 (2025): 397-424.

[6]. Usmonov, Mirismoil. "Crafting Connections: Generative AI's Impact on Post-Purchase Communication in E-Commerce." Contemporary Issues of Communication 4, no. 1 (2025): 288-310.

[7]. Le, Minh T, Oliver Harris, Charlotte Bennett, and Fiona Greene. "Behavior Path Analysis for Blockchain Fraud Detection Using Graph Neural Architectures." Frontiers in Applied Physics and Mathematics 2, no. 1 (2025): 13-20.

[8]. Whitmore, [8]James, Priya Mehra, Jingwen Yang, and Emily Linford. "Privacy Preserving Risk Modeling Across Financial Institutions via Federated Learning with Adaptive Optimization."Frontiers in Artifical Intelligence Research 2, no. 1 (2025): 35-43.

[9]. Hussain, Zahid. "AI-driven personalization and purchase intention in modest fashion: Sharia compliance as moderator."International journal of Halal industry 1, no. 1 (2025): 33-45.

[10]. Hanaee, Toktam, Iulian Dincă, Zohreh Moradi, Parinaz Sadegh Eghbali, and Ali Boloor. "Reflecting the Effect of Physical–Perceptual Components on Increasing the Anxiety of Inner-City Rail Transit's Users: An Integrative Review."Sustainability 17, no. 9 (2025): 3974.