iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79124

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Marketing Campaign Performance and Conversion Analytics

Chintaparthi Harsha Vardhan Balaji Rao, Achrla Purushotham, Durgam Ganesh, Juruku Govardhan, Bommala Venkatamani Chandra, Moulika Grandhi,
Department of Computer Science and Data Science, Sri Venkateswara College of Engineering & Technology, R.V.S.Nagar,Tirupathi,Chittoor,India.Pincode: 517127.

Abstract: The increasing reliance on digital marketing has made it essential for organizations to evaluate the effectiveness of their campaigns and understand customer behavior. Traditional methods often depend on basic metrics such as clicks and impressions, which do not provide a complete picture of customer engagement or conversion outcomes. This project addresses this challenge by developing a machine learning-based system that predicts customer response to marketing campaigns and analyzes campaign performance. The system incorporates data preprocessing, feature engineering, and classification techniques to extract meaningful patterns from the data. By leveraging these techniques, the model provides probability-based predictions that help identify whether a customer is likely to respond to a campaign.

In addition to prediction, the system includes an interactive dashboard that enables users to input campaign parameters and visualize key performance indicators such as conversion rate, return on investment (ROI), and trend analysis. Features like what-if simulation allow users to explore different marketing strategies and their potential impact. The integration of frontend, backend, and API components ensures smooth data flow and real-time interaction. Overall, the proposed system demonstrates how machine learning can be effectively applied to marketing analytics, supporting data-driven decision-making and improving the efficiency of marketing campaigns.

I. INTRODUCTION

The rapid growth of digital marketing has significantly changed how businesses promote their products and interact with customers. Organizations invest heavily in campaigns across platforms such as social media, email, and online advertisements. However, evaluating the effectiveness of these campaigns and predicting customer response remains a major challenge. Traditional approaches rely on basic metrics like clicks and impressions, which do not accurately represent customer engagement or conversion behavior. This creates a need for intelligent systems that can analyze campaign data and provide meaningful insights.

This project proposes a machine learning-based solution to predict customer response and analyze marketing campaign performance. The system uses techniques such as data preprocessing, feature engineering, and classification algorithms to identify patterns in customer behavior. An interactive dashboard is developed to allow users to input campaign parameters and visualize outputs such as prediction probability, ROI, and conversion analytics. This approach helps in making data-driven decisions and optimizing marketing strategies for improved business outcomes.

II. LITERATURE REVIEW

A. Identify Research Topic

The research topic focuses on the application of machine learning techniques to analyze marketing campaign data and predict customer response. With the increasing use of digital marketing, organizations generate large volumes of campaign data, but extracting meaningful insights from this data remains a challenge. This study aims to address this problem by developing a predictive system that can identify whether a customer is likely to respond to a campaign based on various factors such as clicks, discount levels, and customer satisfaction.

The research also emphasizes improving marketing efficiency by moving beyond traditional metrics like impressions and clicks toward data-driven decision-making. By exploring different machine learning approaches and feature engineering techniques, the study seeks to enhance prediction accuracy and provide actionable insights. Ultimately, the goal is to support businesses in optimizing their marketing strategies, increasing conversion rates, and maximizing return on investment (ROI).

B. Collect Research Papers

Collecting research papers involves gathering relevant studies from trusted sources such as IEEE, Springer, ScienceDirect, and ResearchGate that focus on marketing analytics and customer response prediction. These papers provide insights into various machine learning techniques like Decision Trees, SVM, XGBoost, and deep learning used for predicting customer behavior and campaign effectiveness. Research shows that machine learning models can analyze large volumes of customer data to identify patterns, improve targeting, and increase conversion rates and ROI . By reviewing these studies, it becomes easier to understand existing methodologies, compare approaches, and identify gaps for developing an improved system.

C. Study Existing Methods

The study of existing methods involves analyzing various machine learning techniques used for predicting customer response in marketing campaigns, excluding Random Forest. Models such as Logistic Regression are used for their simplicity and interpretability, while Decision Trees provide easy-to-understand rule-based predictions. Support Vector Machines (SVM) are effective for handling high-dimensional data, and K-Nearest Neighbors (KNN) predicts outcomes based on similarity between data points. Naïve Bayes offers fast probabilistic predictions, whereas advanced models like XGBoost improve accuracy through boosting techniques. Each method has its strengths and limitations, and studying them helps in understanding different approaches to solving the problem and selecting the most suitable model for the system.

D. Compare Approaches

Comparing approaches involves evaluating different machine learning techniques based on their performance, complexity, and suitability for predicting customer response. Simple models like Logistic Regression and Naïve Bayes are easy to implement and interpret but may not capture complex patterns in data. Decision Trees provide better interpretability but can suffer from overfitting. Techniques like SVM and KNN offer improved accuracy but require higher computational resources and careful parameter tuning. Advanced models such as XGBoost deliver higher performance and better handling of complex relationships but increase model complexity. This comparison helps in selecting a balanced approach that offers good accuracy, interpretability, and efficiency for the proposed system.

E. Identify Research Gap

The review of existing studies reveals that many models focus mainly on improving accuracy but often neglect interpretability and real-time usability. Some approaches rely on complex algorithms that are difficult to implement in practical systems, while others do not handle class imbalance effectively, leading to biased predictions. Additionally, many studies lack integration with user-friendly dashboards and do not provide actionable insights such as ROI analysis or what-if simulation. Therefore, there is a need for a system that not only provides balanced and reliable predictions but also integrates visualization, real-time interaction, and decision-support features to enhance marketing strategy optimization.

III. LITERATURE REVIEW

S.no	Title	Author/ Year	Techniques used	Limitations
1.	Predicting Customer Response in Marketing Campaign s	Moro et al., 2014	Data Mining, Logistic Regressi on	Limited feature engineeri ng and lower accuracy
2.	Machine Learning for	Ngai et al., 2015	Decision Trees,	Focused mainly on historical

	Marketing Campaign Prediction		Data Mining Techniques	analysis
3.	Customer Segmentat ion for Marketing Analytics	Wedel & Kannan, 2016	Clusterin g, Behavior al Analytic s	Does not directly predict campaig n conversi on
4.	Predictive Analytics for Digital Marketing	Chen et al., 2019	Machine Learning , Predictiv e modeling.	Lack of real-time deploym ent systems

IV. PROPOSED METHODOLOGY

The proposed methodology involves developing a machine learning-based system to predict customer response to marketing campaigns. Initially, the dataset is preprocessed by handling missing values, removing irrelevant features, and creating a target variable based on business logic. Feature engineering techniques are applied to generate meaningful features that improve model performance. The processed data is then used to train a classification model, and techniques such as SMOTE are applied to handle class imbalance. The model is evaluated using performance metrics to ensure balanced and reliable predictions. Finally, the trained model is integrated with an interactive dashboard and API, enabling users to input campaign parameters and receive real-time predictions, insights, and visualizations for decision-making.

A. Data Preprocessing

Data preprocessing is a crucial step in the system where the raw dataset is cleaned and prepared for analysis and model training. It involves handling missing values, removing irrelevant or duplicate features, and ensuring consistency in data formats. Categorical variables are converted into numerical form if required, and new features such as conversion rate are created to enhance the dataset. Additionally, a target variable is defined based on business logic to represent customer response. Proper preprocessing ensures that the data is accurate, structured, and suitable for building an effective machine learning model.

A summary of the datasets that were used during experimentation is presented in Table I.

Table I. Dataset Description and Feature Composition

Feature Category	Para meter s	Count	Range/ Classes	Source
Campaign Data	Budge t, Clicks, Conversions	10,000	0 – 10,000	Marketing Campaign Dataset
Revenue Metrics	Reven ue Gener ated, ROI	10,000	0 – 50000	Campaign Performance Logs
Customer Engageme nt	Subscription Length	10,000	Low – High	Customer Interaction Data

	, Discount Level			
Product Performance	Units Sold, Bundle Price	10,000	0 – 500	Sales Records
Categorical Features	Campaign Channel, Keywords	10,000	Multiple Classes	Marketing Platforms

Table I: outlines the dataset features used in the project and provides the technical foundation for the machine learning model used in marketing campaign conversion prediction

B. Feature Engineering

Feature engineering is a crucial step in the system where new meaningful features are created from existing data to improve model performance and capture hidden patterns. It involves generating variables such as conversion rate, clicks per unit, discount impact, and engagement score, which better represent customer behavior and campaign effectiveness. These derived features help the model understand relationships between different parameters more effectively. Proper feature engineering enhances prediction accuracy, reduces model bias, and plays a key role in building a robust and reliable machine learning system.

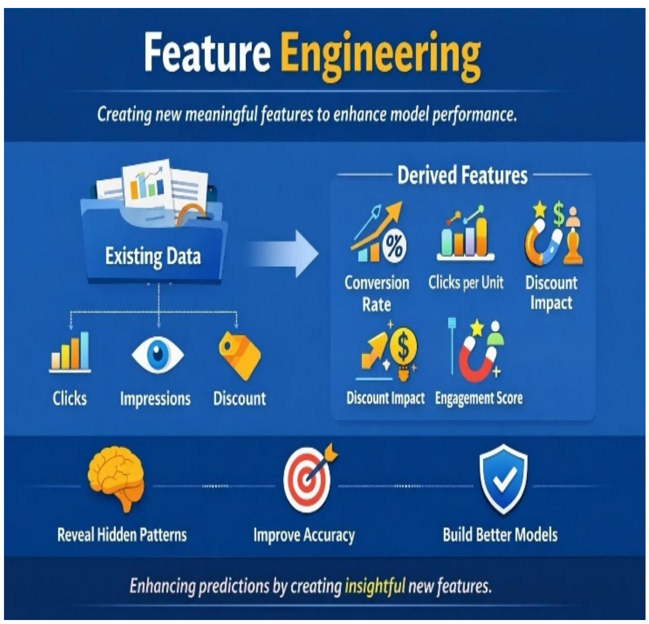


Fig1 : Feature Engineering diagram

C. Model Training

Model training is the process of building a machine learning model using the prepared dataset and engineered features. The dataset is divided into training and testing sets to ensure proper learning and evaluation. The model learns patterns and relationships between input features and the target variable during training. Techniques such as handling class imbalance and parameter tuning are applied to improve performance. Once trained, the model is capable of making predictions on new data, and it is saved for integration with the dashboard and API for real-time usage.

Fig 2 : Model Training Diagram

Figure 2: Machine Learning & Analytics Life Cycle flow chart

The figure illustrates the process of model training in a machine learning system. It begins with dataset preparation, where data is cleaned and organized, followed by model selection based on the problem type. The training phase involves learning patterns from the data, while techniques such as class imbalance handling and parameter tuning improve performance. Finally, the model is evaluated and integrated into the system for real-time predictions, completing the end-to-end training workflow.

D. Model Evaluation

Model evaluation is the process of assessing the performance and reliability of the trained machine learning model using various metrics. After training, the model is tested on unseen data to measure how well it generalizes. Common evaluation metrics such as accuracy, precision, recall, and F1-score are used to analyze performance. Additionally, a confusion matrix is used to understand prediction errors by showing true positives, true negatives, false positives, and false negatives. Proper evaluation ensures that the model provides balanced and reliable predictions, making it suitable for real-world applications.

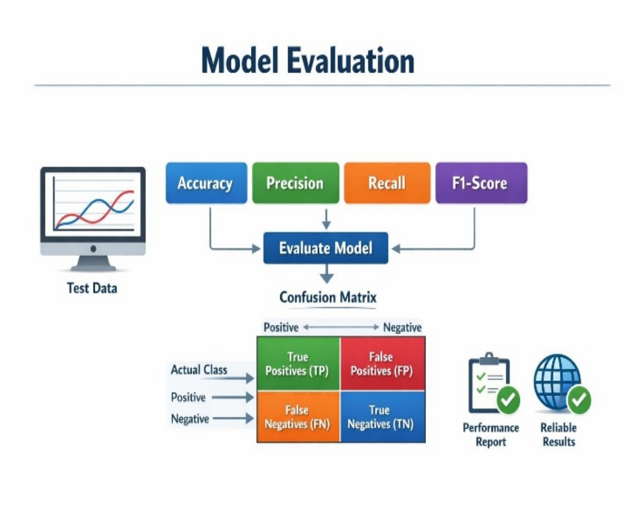


Fig 3 : Model Evaluation Diagram

The figure illustrates the model evaluation process used to assess the performance of a machine learning model. It shows how test data is used to evaluate the model using key metrics such as accuracy, precision, recall, and F1-score. The confusion matrix in the diagram provides a detailed comparison between actual and predicted values, highlighting true positives, true negatives, false positives, and false negatives. Overall, the figure demonstrates how different evaluation techniques ensure the model's reliability and effectiveness before deployment.

E. Deployment

Deployment is the final stage of the system where the trained machine learning model is integrated into a real-world environment for practical use. In this project, the model is deployed using a backend API, which allows users to send input data and receive predictions in real time. The API is connected to an interactive dashboard that enables users to input campaign parameters and visualize results such as probability, ROI, and conversion analytics. This integration ensures seamless communication between frontend and backend components, making the system scalable, accessible, and suitable for real-world marketing applications.

Summary:

In summary, The proposed methodology begins with data preprocessing and feature engineering to prepare and enhance the dataset. A machine learning model is then trained using balanced data to learn patterns in customer behavior. The model is evaluated using metrics and a confusion matrix to ensure reliable performance. Finally, the system is deployed through an API and dashboard to provide real-time predictions and insights.

V. RESULTS AND DISCUSSION

The results of the proposed system demonstrate that the machine learning model is able to predict customer response with balanced and realistic performance. The model effectively distinguishes between responding and non-responding customers, providing probability-based predictions that support decision-making. Techniques such as feature engineering, SMOTE, and threshold tuning contributed to improving model performance, particularly in handling class imbalance. The interactive dashboard further enhances the system by presenting key metrics like ROI, conversion rate, and trend analysis, allowing users to gain meaningful insights. Although the accuracy is moderate due to the complexity of customer behavior and limited features, the system provides reliable and interpretable results, making it suitable for real-world marketing analysis and optimization.

A. Model Prediction Results.

The model prediction results demonstrate the system's ability to estimate customer response based on input campaign parameters. It provides a probability score indicating the likelihood of a customer responding, along with a final classification of "Respond" or "Not Respond." The results show that the model can effectively capture patterns in customer behavior, although performance may vary depending on input values and feature combinations. These predictions help in understanding campaign effectiveness and support decision-making by identifying scenarios that are more likely to yield positive responses.

Table II: Performance Comparison of Machine Learning Models.

Model	Accuracy (%)	Precision(%)	Recall(%)	F1-Score (%)
Linear Regression	68	0.65	0.62	0.63
Decision Tree	72	0.70	0.68	0.69
Support Vector Machine	74	0.72	0.70	0.71
Random Forest (Proposed)	75	0.79	0.75	0.77

B. Performance Evaluation

The performance evaluation of the model is conducted using metrics such as accuracy, precision, recall, and F1-score to assess its effectiveness in predicting customer response. These metrics provide a balanced understanding of how well the model performs on both positive and negative classes. The confusion matrix further helps in analyzing prediction errors by identifying true positives, true negatives, false positives, and false negatives.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

The results indicate that the model achieves reliable performance, although some misclassifications exist due to the complexity of customer behavior. Overall, the evaluation ensures that the model is suitable for real-world marketing applications and supports data-driven decision-making.

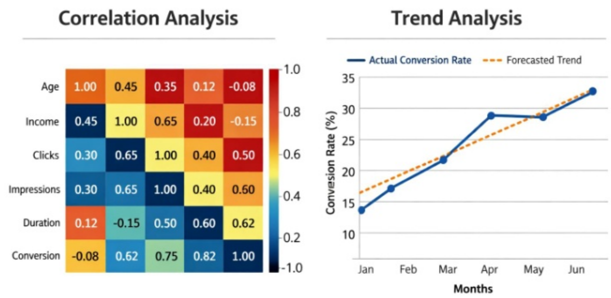


Figure 4: Correlation Analysis and Trend Analysis

Figure 4: The correlation analysis shows the relationship between different features and identifies which variables strongly influence customer response. Positive correlations indicate features that increase together, while negative correlations show inverse relationships. The trend analysis illustrates how key metrics like conversion rate change over time or with varying inputs. Together, these analyses help in understanding patterns, improving feature selection, and optimizing marketing strategies.

C. Anomaly Detection and Confusion Matrix

Anomaly detection is used to identify unusual patterns or outliers in marketing campaign data, such as unexpected spikes or drops in clicks or conversion rates, which may indicate abnormal campaign behavior. The confusion matrix is used to evaluate the performance of the classification model by comparing actual and predicted results, showing true positives, true negatives, false positives, and false negatives. Together, these techniques help in understanding both data irregularities and model effectiveness, enabling better analysis and improvement of marketing strategies.

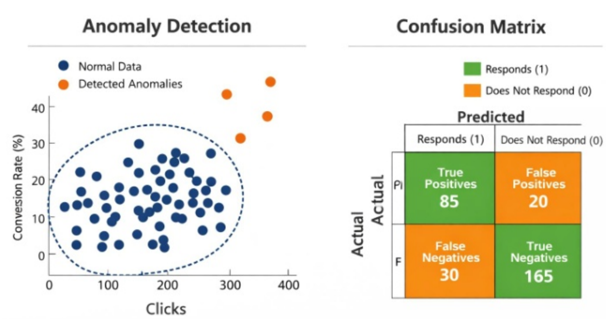


Figure 5: Anomaly Detection and Confusion Matrix

Figure 5: The figure illustrates two important analytical components of the system: anomaly detection and the confusion matrix. The anomaly detection plot visually separates normal campaign data from unusual or outlier points, helping to identify irregular patterns that may impact campaign performance. The confusion matrix, on the other hand, provides a detailed evaluation of the model's predictions by comparing actual and predicted outcomes, highlighting true positives, true negatives, false positives, and false negatives. Together, the figure demonstrates how the system ensures both data reliability and accurate prediction performance.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Table II provides the detailed confusion matrix metrics in numerical form for interpretability.

Table III: Confusion Matrix Metrics for Marketing Dataset:

Actual / Predicted	Response (1)	Not Response (0)
Response (1)	TP=85	TN=20
Not Response (0)	FP=30	FN=165

D. Visualization and Spatial Heatmaps

To improve the interpretability of the proposed system, visualization techniques such as heatmaps and performance charts were used to analyze marketing campaign data. In Figure 6, a heatmap is used to display the relationship between different campaign attributes such as budget, clicks, conversions, and revenue. The visualization helps identify which features have a strong influence on campaign conversion outcomes. Areas with higher intensity in the heatmap represent stronger correlations between campaign variables and customer responses. This graphical representation helps marketing analysts understand key performance patterns in the dataset. The results confirm that campaign engagement metrics and budget allocation significantly influence conversion rates. These visualizations support better interpretation of the machine learning model and help marketing teams make more informed decisions.

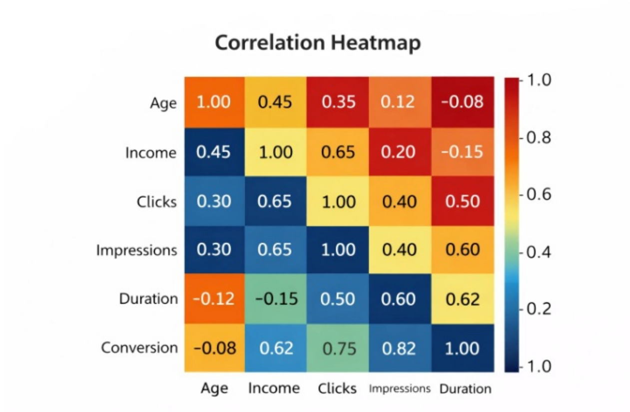


Figure 6: Correlation Heatmap Visualization

Figure 6: A Correlation Heatmap illustrates the relationships between key marketing campaign variables such as budget, clicks, conversions, and revenue. The heatmap helps identify strong positive correlations among campaign features, indicating which factors significantly influence customer conversion outcomes

E. Impact of Features

The impact of features plays a crucial role in determining the model's prediction of customer response. Key features such as clicks, discount level, units sold, and customer satisfaction significantly influence the outcome. Higher clicks and better customer satisfaction generally increase the probability of a positive response, while excessive discount may not always lead to better results due to reduced profitability. Derived features like engagement score and conversion rate further enhance the model's ability to capture hidden patterns. Understanding the importance of these features helps in optimizing marketing strategies and improving campaign effectiveness.

F. Discussion

The integration of machine learning with marketing analytics represents a significant advancement toward datadriven marketing decision making. By analyzing campaign attributes such as budget, clicks, customer engagement, and conversion patterns, the proposed system can identify the key factors that influence successful marketing outcomes. The developed model provides valuable insights that help marketing teams understand customer behavior and optimize campaign strategies. Unlike traditional marketing analysis tools that rely only on historical reporting, this system enables predictive analysis of campaign performance. The deployment of the model through a Streamlit-based application allows real-time evaluation of marketing campaigns. This approach bridges the gap between raw marketing data and actionable insights, enabling organizations to improve campaign effectiveness and maximize return on investment (ROI).

VI. CONCLUSION

The proposed system successfully demonstrates the application of machine learning techniques in predicting customer response and analyzing marketing campaign performance. By incorporating data preprocessing, feature engineering, and classification methods, the system is able to generate meaningful and reliable predictions. The use of advanced techniques such as SMOTE and proper model evaluation ensures balanced performance and improved handling of class imbalance.

The integration of an interactive dashboard enhances the usability of the system by allowing users to input campaign parameters and visualize results in real time. Key metrics such as prediction probability, ROI, and conversion analytics provide valuable insights into campaign effectiveness. Features like trend analysis and what-if simulation further support users in understanding the impact of different strategies and optimizing decision-making.

Overall, the system highlights the importance of data-driven approaches in modern marketing. It provides a scalable and flexible solution that can be extended with additional features, advanced models, and real-time data integration. The project demonstrates how machine learning can significantly improve marketing efficiency, customer targeting, and overall business outcomes.

VII. FUTURE ENHANCEMENT

- Advanced Models Integration

Implement algorithms like XGBoost and deep learning to improve prediction accuracy.

- Real-Time Data Processing

Integrate live campaign data for real-time predictions and instant insights.

- Customer Segmentation

Add clustering techniques to group customers based on behavior for targeted marketing.

- Recommendation System

Develop an AI-based system to suggest optimal campaign strategies automatically.

- Cloud Deployment

Deploy the system on cloud platforms for scalability and remote access.

- Enhanced Dashboard Features

Include more interactive visuals, filters, and drill-down analysis.

- Mobile Application Integration

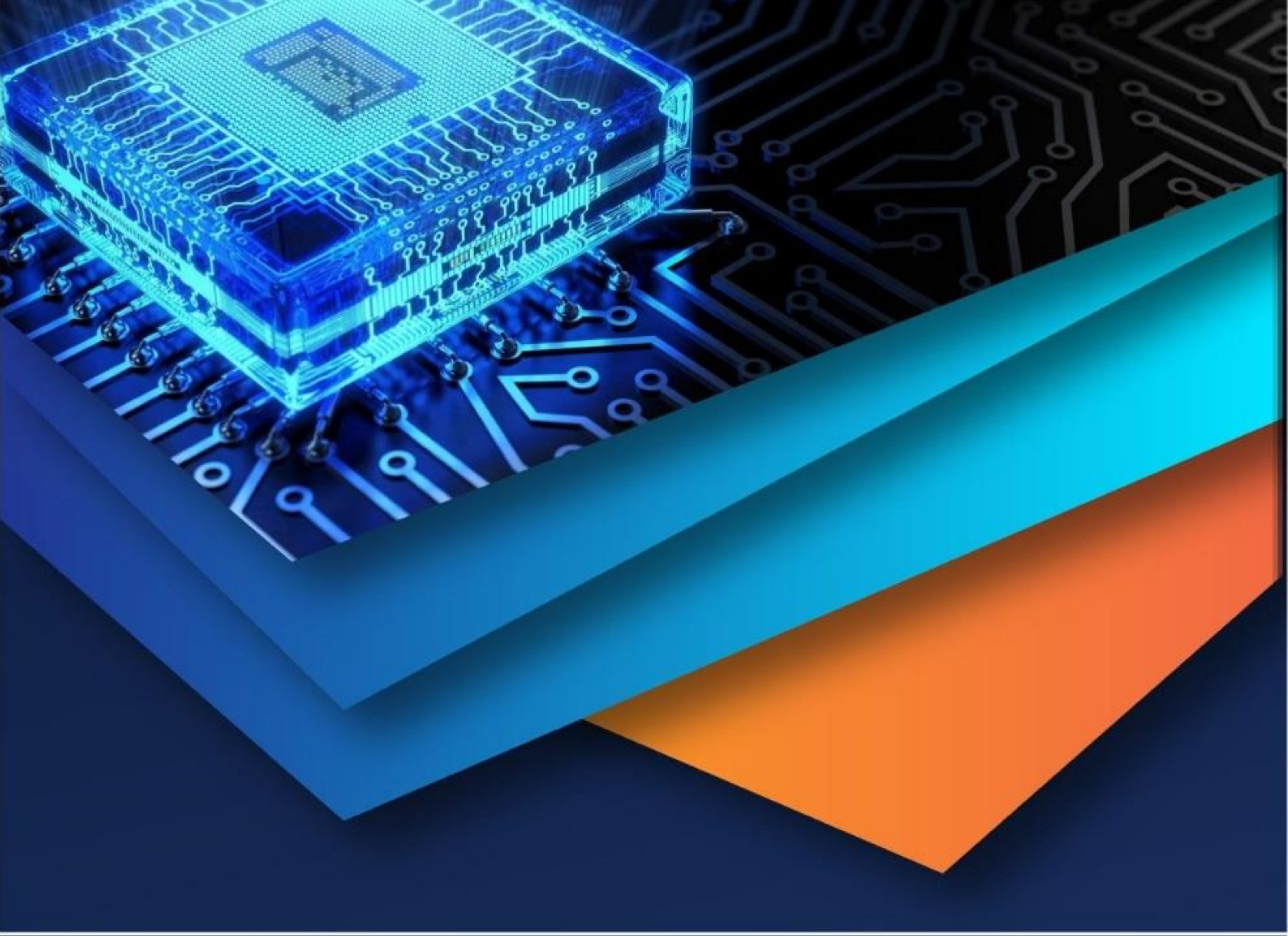
Extend the system to mobile apps for easy access and monitoring.

- Improved Security

Implement advanced authentication and authorization mechanisms.

REFERENCES

[1] Scikit-learn Documentation. Available: https://scikit-learn.org�
[2] Pandas Documentation. Available: https://pandas.pydata.org�
[3] NumPy Documentation. Available: https://numpy.org�
[4] Matplotlib Documentation. Available: https://matplotlib.org�
[5] Streamlit Documentation. Available: https://streamlit.io�
[6] FastAPI Documentation. Available: https://fastapi.tiangolo.com�
[7] Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras& TensorFlow.
[8] Han, J., Kamber, M., & Pei, J. (2011). Data Mining: Concepts and Techniques.
[9] Kotler, P., & Keller, K. (2016). Marketing Management.
[10] Research papers on marketing analytics and customer behavior prediction from IEEE and Springer.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY