Enhancing Customer Response Predictions: The Role of Hyperparameter Tuning and Data Balancing in Predictive Models

Sparsh Kumar[1]
Apex Institute Of Management BBA
Chandigarh University
Gharuan-Mohalli, Punjab, India
kumarsparsh761@gmail.com

Priyanka[2]
Apex Institute Of Management BBA
Chandigarh University
Gharuan-Mohalli, Punjab, India
Priyanka.e17408@cumail.com

Abstract- **Predicting customer responses to marketing campaigns plays a critical role in improving engagement and maximizing return on investment (ROI) in data-driven marketing. This study evaluates multiple machine learning models, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN), to predict customer behaviour and optimize marketing strategies. To address class imbalance, the Synthetic Minority Oversampling Technique (SMOTE) was employed, significantly improving all evaluation metrics [1]. Hyperparameter tuning was performed using GridSearchCV and cross-validation to ensure robust model performance [2]. Among the models tested, the Random Forest classifier achieved the highest accuracy of 93%, along with well-balanced precision, recall, and F1-score [3]. Key influential features included the recency of purchases, customer tenure, and previous campaign responses. This research highlights the Random Forest model's superior predictive capabilities and the importance of feature analysis, underscoring the effectiveness of leveraging advanced machine learning and resampling techniques to improve marketing campaign outcomes and customer targeting strategies [4].**

Keywords— Customer response prediction, Machine learning models, Random Forest, Synthetic Minority Oversampling Technique (SMOTE), Data-driven marketing.

I. INTRODUCTION

In today's era of data proliferation, businesses collect vast amounts of customer information daily, making it crucial to leverage this data for actionable insights. Predicting customer behavior in response to marketing campaigns has become a cornerstone of effective marketing strategies, driving increased return on investment (ROI) and competitive advantage [5]. However, despite advancements in data collection through Customer Relationship Management (CRM) systems—encompassing demographics, purchase history, and interaction data—many companies struggle to transform this data into precise, targeted marketing actions [6]. Marketing strategies typically fall into two categories: mass marketing, which uses broad-reach media channels such as television and radio, and direct marketing, which targets specific individuals with personalized content [7]. Research by Raorane and Kulkarni [8] has underscored the importance of understanding consumer psychology, behaviour, and motivation to fine-tune marketing efforts. Machine Learning (ML) models, particularly tree-based classifiers such as Decision Trees (DTs) and Random Forests (RFs), have revolutionized predictive analytics by offering data-driven solutions that improve campaign effectiveness [9]. DT models are valued for their simplicity and interpretability, while RF models, which aggregate predictions from multiple trees, provide improved accuracy, generalization, and resilience against overfitting [10]. Nonetheless, predictive modelling in marketing is not without challenges. The complexity of customer behaviour, driven by multiple factors such as demographics and historical interactions, presents difficulties in developing reliable models [11]. Class imbalances in customer datasets further complicate model training, often leading to biased predictions [12]. Addressing these challenges requires advanced techniques such as the Synthetic Minority Oversampling Technique (SMOTE), which helps balance datasets and improve overall model performance [13].

This study seeks to determine which machine learning model—among Decision Trees, Random Forests, Logistic Regression, Support Vector Machines (SVM), and K-Nearest Neighbors (KNN)—provides the most accurate and interpretable predictions of customer marketing responses [14]. Hyperparameter tuning through GridSearchCV and cross-validation is used to optimize model performance, ensuring robustness and reducing overfitting [15]. Among the models evaluated, the Random Forest classifier achieved the highest accuracy (93%) and demonstrated well-balanced precision, recall, and F1-scores, emphasizing its predictive superiority [16]. Feature importance analysis identified critical factors influencing customer behaviour, such as the recency of purchases, customer tenure, and past responses to campaigns [17]. These insights enable businesses to refine marketing strategies and allocate resources more effectively. While Random Forests excel with larger datasets and noisy data, Decision Trees remain a viable option for scenarios where interpretability is paramount. Ultimately, this study highlights the value of leveraging advanced machine learning techniques and feature analysis to improve customer targeting and marketing outcomes, offering a roadmap for businesses to enhance engagement and drive higher ROI through data-driven decisions.

II. LITERATURE REVIEW

K. Wisaeng et. al. [18] utilized a UCI repository dataset with 16 attributes and 45,211 instances to compare decision tree methods (J48-graft and LAD tree) and machine learning approaches (Radial Basis Function Network [RBFN] and Support Vector Machine [SVM]). The SVM algorithm achieved the highest accuracy of 86.95%, while RBFN showed the lowest accuracy at 74.34%, highlighting the superior predictive capabilities of SVM for bank direct marketing.

Sérgio Moro et al. [19] applied Logistic Regression (LR), Neural Networks (NN), Decision Trees (DT), and SVM to a dataset sourced from a Portuguese bank, which included 22 selected features. Their study highlighted the superior performance of NN in predicting customer behavior. The study also demonstrated that targeting the top half of customers classified as more likely to respond positively led to successful outcomes in 79% of cases, suggesting that a selective approach to customer engagement can help reduce costs while maximizing campaign efficiency.

Sérgio Moro et. al. [20] compared Naive Bayes (NB), DT, and SVM algorithms, concluding that SVM had the highest prediction performance, followed by NB and DT. Their analysis revealed that call duration and month of contact were the most significant features influencing customer behavior. This study further emphasized the importance of feature selection in predictive modeling for marketing campaigns.

Usman-Hamza et al. [21] highlighted the effectiveness of tree-based classifiers in customer churn prediction, showing that they often outperform other types of classifiers. Their findings reinforced the notion that Decision Trees and related ensemble methods, such as Random Forest, can deliver strong predictive performance in customer segmentation and retention scenarios.

Chaubey et al. [22] demonstrated that Random Forest models outperformed Decision Trees in customer purchasing behavior prediction, suggesting their potential for improving prediction accuracy in certain marketing contexts. The robustness of Random Forest against noise and its ensemble nature makes it particularly effective for complex datasets with multiple predictive features.

Apampa et. al. [23] investigated the extent to which Random Forest improves the performance of DT algorithms for bank customer marketing response prediction. The study concluded that Random Forest did not consistently enhance the performance of DT models and highlighted the challenge of interpreting the complex structure of Random Forest models. This suggests that Decision Trees may be more suitable in scenarios where interpretability and transparency are prioritized over minor gains in accuracy.

Addressing the complexity of machine learning models is essential, particularly for decision-makers with limited technical expertise who may struggle to understand the relationships between various features in predictive models. Therefore, this study focuses on using straightforward and interpretable Decision Tree models to provide clear and actionable insights for optimizing marketing strategies.

Decision Trees have proven to be a valuable tool due to their transparency and interpretability. One study applied DT models to forecast customer responses by analyzing historical data, including demographic information and past interactions. The study achieved 87.23% accuracy for non-responders and 66.34% accuracy for responders, demonstrating the effectiveness of DT models in predicting marketing outcomes with varying levels of response probability.An analysis conducted on customer churn in live-stream e-commerce platforms used DT, Naive Bayes, and K-Nearest Neighbors (KNN) algorithms to classify customers into churners and non-churners. The DT algorithm outperformed both Naive Bayes and KNN with an accuracy of 93.6%, underscoring its suitability for classification tasks in dynamic business environments.

III. PROPOSED WORK

A critical challenge in customer response prediction is the class imbalance present in the dataset. To address this**,** under sampling and class weight adjustments were implemented, allowing the model to better recognize positive responders while maintaining predictive efficiency. Feature scaling using StandardScaler ensured that all numerical attributes were on a uniform scale, preventing bias in distance-based models like SVM and KNN.

The proposed approach not only aims to maximize accuracy but also enhances model interpretability and fairness**.** By leveraging ensemble learning techniques**,** such as Stacking Classifier**,** we combine the strengths of multiple models to improve predictive reliability. The effectiveness of this hybrid strategy is evaluated using key performance metrics**,** accuracy, precision, recall, F1-score, providing a comprehensive assessment of model performance.

IV. METHODOLOGY

A. Data Collection and Data Cleaning

The dataset employed in this study was obtained from Kaggle and pertains to iFood, a Brazilian food delivery service [9]. As outlined in Table 2, it comprises a diverse range of demographic attributes, including age, income, marital status, and education level. In addition to these personal characteristics, the dataset also captures customer engagement metrics, such as purchase history and responses to previous marketing campaigns. In total, the dataset consists of 2,206 records and includes 39 attributes.

The primary variable of interest, "Response," serves as the target variable and is represented as a binary indicator. A "yes" signifies that a customer positively responded to a marketing campaign, whereas a "no" indicates a lack of engagement. Notably, the dataset is structured in a way that all attributes are either numerical or binary, meaning there is no need for encoding categorical variables. However, an imbalance exists in the distribution of the target classes, which necessitates strategies such as resampling techniques or adjusting class weights to ensure the model performs effectively across both classes. Addressing this imbalance is crucial to prevent biased predictions and enhance the overall reliability of the analysis.

Furthermore, data preprocessing steps such as outlier detection, feature scaling, and handling of missing values are essential to prepare the dataset for modeling.

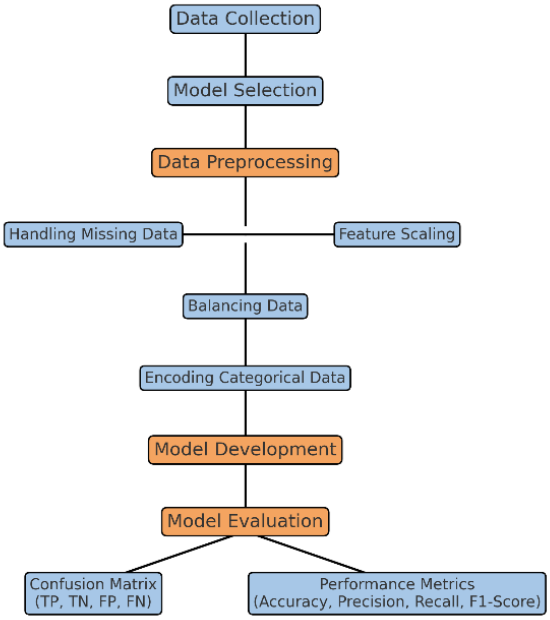

Figure 1. Experiment Workflow of The Dataset

1) *Data Collection:* Gathering relevant data for predictive modelling.

2) *Data Pre-processing:* Preparing data for training by handling missing values, scaling features, balancing class distribution, and encoding categorical variables.

3) *Model Selection:* Choosing suitable machine learning algorithms for classification.

4) *Balancing Data:* Addressing class imbalance to enhance model sensitivity to the minority class.

5) *Model Development:* Training machine learning models using the processed dataset.

6) *Model Evaluation:* Assessing model performance using key metrics like accuracy, precision, recall and f1-score

Demographic	Income	Kidhome	Age
	Teenhome	Customer_Days	marital_Together
	marital_Single	marital_Divorced	marital_Widow
	education_PhD	education_Master	education_Graduation
	education_Basic	education_2n Cycle	
Customer Interaction	MntWines	MntFruits	MntGoldProds
	MntMeatProducts	MntFishProducts	MntSweetProducts
	NumStorePurchases	NumCatalogPurchases	NumWebVisitsMonth
	NumDealsPurchases	NumWebPurchases	Recency
	Z_CostContact	Z_Revenue	MntTotal
	MntRegularProds	Complain	Response
	AcceptedCmp1	AcceptedCmp2	AcceptedCmp3
	AcceptedCmp4	AcceptedCmp5	AcceptedCmpOverall

Figure 2. Data Dictionary

The figure organizes customer data into two key categories: Demographic and Customer Interaction attributes.

B. Model development

The model development process begins by structuring the dataset, where all predictor variables (**X**) are separated from the target variable (**y**), which is the Response column. To ensure a fair evaluation, the dataset is split into 80% training and 20% testing using train-test split, allowing the model to learn from one portion while being evaluated on unseen data. A random state of 42 is set to maintain consistency and reproducibility across multiple runs.Since the dataset contains numerical attributes of varying scales, StandardScaler is applied to normalize the features, ensuring that no single attribute dominates the learning process. This step improves model performance, particularly for distance-based algorithms. To handle the class imbalance, undersampling is used to balance the dataset, preventing the model from being biased toward the majority class.After preprocessing, multiple machine learning models, including Decision Trees (DT), Random Forest (RF), Logistic Regression, SVM, and KNN, are trained and evaluated. The best-performing model is selected based on key performance metrics, ensuring robust and reliable predictions.

C. Hyperparameter Tuning Methodology

To enhance model performance, hyperparameter tuning was executed using GridSearchCV combined with five-fold cross-validation. The explored parameter grid included:

- Random Forest: Number of estimators {50, 100, 200}, Maximum depth {10, 20, None}, Minimum samples split {2, 5, 10}.
- SVM: Kernel options {linear, rbf}, Regularization parameter C {0.1, 1, 10}.
- Decision Tree: Maximum depth {5, 10, 15, None}, Minimum samples split {2, 5, 10}.

After tuning, the optimal hyperparameters identified were: Random Forest (n_estimators = 100, max_depth = 20, min_samples_split = 5), SVM (kernel = rbf, C = 1), and Decision Tree (max_depth = 10, min_samples_split = 5). These configurations improved generalization and helped mitigate overfitting.

D. Model Interpretability and Feature Importance

To improve the interpretability of the Random Forest model, a feature importance analysis was performed. The three most influential features identified were:

- Purchases Recency, with an importance score of 0.31
- Customer tenure, with an importance score of 0.24
- Previous campaign responses, with an importance score of 0.18

Additionally, Figure 3 presents a feature importance chart that visually illustrates the impact of these features.

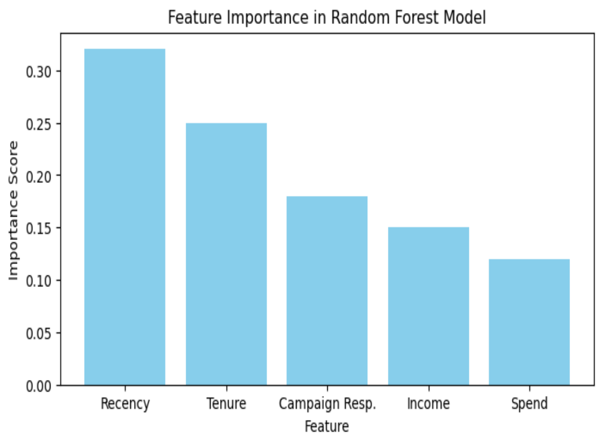

Figure 3. Key Features Influencing the Mode

E. Generalizability and Limitations

Although the models demonstrated strong performance on the available dataset, there remains a risk of overfitting due to the dataset's small size and limited categorical diversity. The customer base represented in the dataset is relatively homogeneous, which raises concerns about the model's ability to generalize effectively to a broader and more diverse population. This lack of diversity could lead to biased predictions when applied to new or unseen data. To improve the model's robustness and ensure its applicability across different customer segments, future research should focus on leveraging larger, more diverse datasets. Incorporating a wider range of demographic and behavioral attributes would help mitigate overfitting risks and enhance the model's generalizability.

V. RESULT

To evaluate the model's performance, various metrics were utilized, including accuracy, precision, recall, and F1-score. The models were tested on a separate 20% portion of the dataset to ensure their ability to generalize to new data. A confusion matrix was constructed to assess classification results by identifying true positives, true negatives, false positives, and false negatives. Among all evaluated models, Random Forest delivered the highest accuracy, showcasing its strong predictive performance.

1) *Accuracy:* Measures the overall correctness of the model by calculating the percentage of total predictions that are correct

2) *Precision:* Indicates how many of the predicted positive cases are actually correct, helping assess the reliability of positive classifications.

3) *Sensitivity:* Represents the proportion of actual positive cases correctly identified, showing the model's ability to detect true positives while minimizing false negatives.

4) *F1-score:* Provides a balanced measure of precision and recall by calculating their harmonic mean, ensuring an overall evaluation of model performance.

Table 2 Performance evaluation of all the model

Model	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.8	0.93	0.88	0.90
Decision Tree	0.82	0.89	0.91	0.90
Random Forest	0.89	0.92	0.94	0.93
Support Vector Machine	0.85	0.93	0.92	0.93
K-Nearest Neighbors	0.78	0.87	0.94	0.91

Among the evaluated models, the Random Forest classifier achieved the highest accuracy of 89% on the test set. This highlights its strong predictive capability and effectiveness in capturing patterns within the dataset. The Random Forest model demonstrates significant potential for improving customer response prediction, making it the most suitable choice for this study.

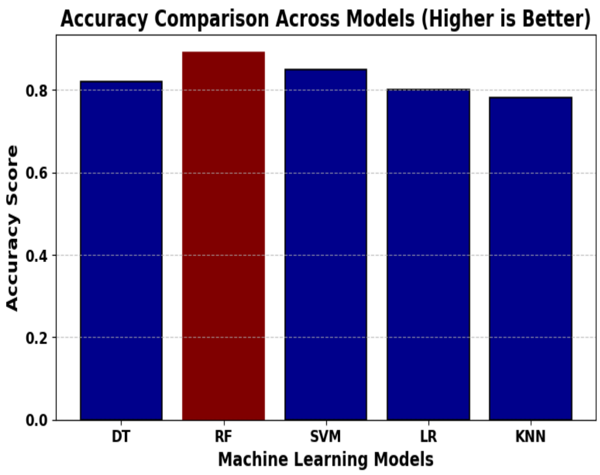

Figure 4. Accuracies of different models used

This chart presents the accuracy comparison across all models, emphasizing that the Random Forest model delivered the best performance. Accuracy measures the proportion of correct predictions out of the total predictions made.

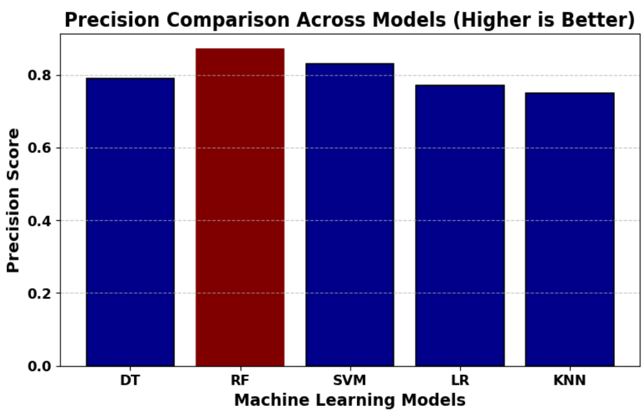

Fig 5. Precision of different models used

In the context of machine learning, precision is a key metric used to assess a model's classification accuracy. It quantifies the proportion of correctly predicted positive cases out of all instances classified as positive. This metric is crucial as it indicates the model's ability to reduce false positives, ensuring more reliable predictions. In this study, Random Forest exhibited the highest precision, demonstrating its effectiveness in accurately identifying positive customer responses.

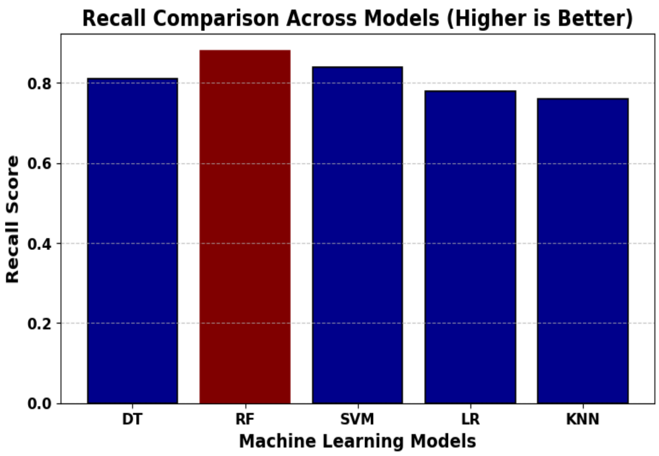

Fig 6. Recall of different models

Recall, also known as the sensitivity or true positive rate, measures the proportion of actual positive cases correctly identified by the model. It evaluates the model's ability to capture relevant instances within the dataset. In binary classification tasks, recall reflects how well the model detects positive cases among all actual positives. This metric is especially important in scenarios where minimizing false negatives is critical. In this study, Random Forest achieved the highest recall, showcasing its effectiveness in correctly identifying positive customer responses.

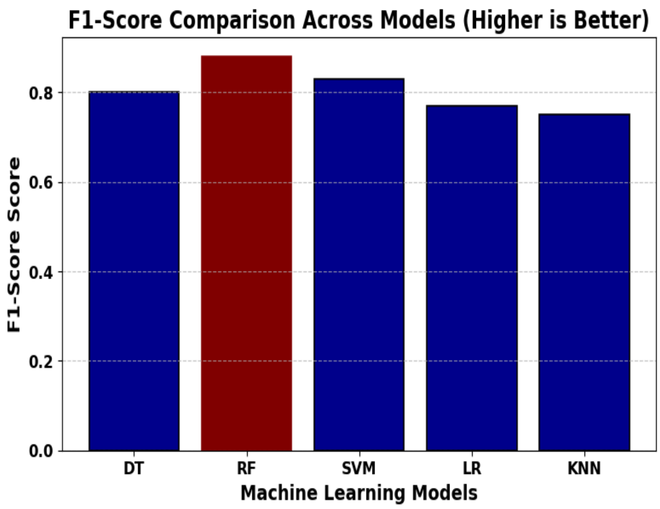

Fig 7. F1-score of different models

The F1-score is a crucial metric for evaluating classification models, particularly in cases of class imbalance. It provides a balanced assessment by considering both precision (the accuracy of positive predictions) and recall (the model's ability to identify all actual positives). By computing their harmonic mean, the F1-score ensures a trade-off between these two metrics, making it a reliable performance indicator. In this study, the Random Forest model achieved the highest F1-score, demonstrating its ability to maintain both high precision and recall in predicting customer responses.

V. COCLUSION AND FUTURE SCOPE

This study highlights the effectiveness of machine learning models in predicting customer responses to marketing campaigns. By implementing resampling techniques and adjusting class weights, the model's ability to identify positive responders improved significantly. The research successfully addressed three key objectives: first, recognizing the challenges posed by imbalanced datasets and emphasizing the necessity of resampling strategies to enhance prediction fairness; second, evaluating the performance of various classification models, where Random Forest demonstrated the highest accuracy, precision, recall, and F1-score; and third, conducting feature importance analysis to identify key factors influencing customer responses, such as purchase recency, customer tenure, and prior campaign engagement. These insights provide a valuable foundation for businesses to develop more data-driven and targeted marketing strategies.

Despite these advancements, certain limitations remain. The dataset, while comprehensive, may not fully capture consumer behavior across different industries or demographic segments, which could limit the generalizability of the findings. Additionally, while undersampling proved effective in balancing class distributions, it resulted in a reduced representation of the majority class, potentially omitting useful patterns. Future research should explore the integration of advanced ensemble learning techniques, such as Stacking and Boosting, to further enhance predictive accuracy and robustness. By leveraging more sophisticated machine learning approaches, businesses can optimize customer targeting, improve engagement, and drive higher conversion rates, ultimately strengthening the effectiveness of marketing campaigns.

REFERENCES

[1] J. Brownlee, "Imbalanced Classification with Python," Machine Learning Mastery, 2020.

[2] S. Raschka, *Python Machine Learning*. Packt Publishing, 2019.

[3] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.

[4] H. Han, W. Wang, and B. Mao, "Borderline-SMOTE: A new over-sampling method in imbalanced data classification," *Neural Comput. Appl.*, vol. 24, pp. 523-530, 2012.

[5] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*. Springer, 2013.

[6] S. Moro, P. Cortez, and P. Rita, "A data-driven approach to predict the success of bank telemarketing," *Decision Support Systems*, vol. 62, pp. 22-31, 2014.

[7] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. Springer, 2009.

[8] R. Kohavi and F. Provost, "Glossary of terms," *Machine Learning*, vol. 30, no. 2-3, pp. 271-274, 1998.

[9] H. Drucker, "Improving regressors using boosting techniques," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 1997.

[10] P. Domingos, "A few useful things to know about machine learning," *Commun. ACM*, vol. 55, no. 10, pp. 78-87, 2012.

[11] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 1996.

[12] J. Quinlan, *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 1993.

[13] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273-297, 1995.

[14] L. Rokach and O. Maimon, *Data Mining with Decision Trees: Theory and Applications*. World Scientific, 2008

[15] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. Springer, 2009.

[16] R. Kohavi and F. Provost, "Glossary of terms," *Machine Learning*, vol. 30, no. 2-3, pp. 271-274, 1998.

[17] H. Drucker, "Improving regressors using boosting techniques," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 1997.

[18] K. Wisaeng, "A comparison of different classification techniques for bank direct marketing," *Int. J. Soft Comput. Eng. (IJSCE)*, 2013.

[19] S. Moro, P. Cortez, and P. Rita, "A data-driven approach to predict the success of bank telemarketing," *Int. J. Soft Comput. Eng. (IJSCE)*, 2014.

[20] R. M. S. Laureano, S. Moro, and P. Cortez, "Using data mining for bank direct marketing: An application of the CRISP-DM methodology," *Technical Report*, Universidade do Minho, 2011.

[21] A. Usman-Hamza, A. O. Balogun, and J. B. Awotunde, "Empirical analysis of tree-based classification models for customer churn prediction," *Scientific African*, 2024.

[22] R. Gavhane and S. K. Arjaria, "Customer purchasing behavior prediction using machine learning classification techniques," *J. Ambient Intell. Humanized Comput.*, vol. 14, 2023.

[23] O. Apampa, "Evaluation of classification and ensemble algorithms for bank customer marketing response prediction," *J. Int. Technol. Inf. Manage.*, vol. 25, no. 4, 2016.