Optimizing Bank Marketing Using Machine Learning and Predictive Models

Amiya Ranjan Panda
School of Computer Engineering
KIIT Deemed to be University
Bhubaneswar, Odisha, India
Email: amiya87@gmail.com

Manoj Kumar Mishra
School of Computer Engineering
KIIT Deemed to be University
Bhubaneswar, Odisha, India
Email: manojku.mishra05@gmail.com

Sneha Kashyap
School of Computer Engineering
KIIT Deemed to be University
Bhubaneswar, Odisha, India
Email: snehakashyap020704@gmail.com

Vaibhav Poddar
School of Computer Engineering
KIIT Deemed to be University
Bhubaneswar, Odisha, India
Email: vaibhavpoddar09@gmail.com

Abhay Rathore
School of Computer Engineering
KIIT Deemed to be University
Bhubaneswar, Odisha, India
Email: abhayrathore036@gmail.com

***Abstract*—In particular, they put to the test all the machine learning methods able to gauge customer attitudinal responses to various marketing campaigns. More specifically, LightGBM showed the best accuracy result of (90.66%), succeeded by Random Forest and SVM with (89.33%), Gradient Boosting, and another Adaboost with (89.38%), Logistic Regression with (90.15%), XGBoost(90.54%), and Decision Tree with (87.44%), and finally Naive Bayes with (85.08%). This dataset includes some level of preprocessing to handle missing data, encode categorical variables, and carry out feature engineering most likely to improve model performance. In terms of metrics for evaluation, accuracy and precision, recall, and F1 score were analyzed. The results conclude that ensemble models, particularly Random Forest and Gradient Boosting, produce the best performance overall.**

***Index Terms*—Machine Learning, Bank Marketing, Customer Response Prediction, Random Forest, Gradient Boosting, XGBoost, Logistic Regression, Predictive Analytics, Customer Segmentation, Feature Engineering, Model Evaluation, Targeted Marketing, Financial Data Analysis, Real-Time Prediction, Deep Learning, Hybrid Models, Decision-Making, Banking Industry.**

I. INTRODUCTION

Over the years, the banking sector has significantly changed due to data-driven decision-making, especially in bank marketing, where analytics help with customer response prediction regarding campaigns. Many financial institutions want to boost customer interaction and make marketing efforts more effective. Machine learning models and statistical techniques are highly effective tools for analyzing vast amounts of customer data and gaining meaningful insights that could guide outreach efforts. The aim of this study is to analyze a bank-marketing dataset for predictive models for estimating customer likelihood for term-deposit subscription within. The dataset includes customer demographics, previous campaign interactions, and economic indicators. Machine learning algorithms like Random Forest, Gradient Boosting, and Logistic Regression will be applied in order to be able to make accurate predictions, thus enhancing customer targeting. The banking sector marketing campaigns have long been resource-intensive activities; so much has been invested, but still, a good number of campaigns do not yield desired results due to, among other things, ineffective targeting strategies. By implementing machine-learning techniques and exploratory data analysis, banking institutions will be able to unearth insights into consumer preferences, trends around engagement with campaigns, and the factors that drive conversion rates.

A. LITERATURE SURVEY

In recent years, marketing in the banking sector has undergone a massive transformation. The previous demographic and credit history-based strategies do not find any standing now given the capabilities ML brought in. The blend of unstructured and structured data with ML will, through clustering, segmentation, and predictive analytics, empower the banks to analyze data and predict a customer's behavior to target him/her effectively with better marketing campaigns. Feature engineering, which includes variables like customer occupation, age, level of past interactions, etc., entails finding the factors responsible for the success of marketing campaigns. The key features so created will help banks, through efficient marketing models, target the right customer with the right offers at the right time. This again results in higher conversion rates, better customer engagement, and improved marketing efficiency. Further improvements in the customer experience come through automated decision-making and real-time engagement. Because of advanced analytics, banks can target high-potential customers, reduce the marketing budget, increase retention rates, and ultimately lead to profit. By evolving their predictive models, banks can get ahead with AI-driven marketing strategies in their own domains. The future of bank marketing rests on a continuing development of dynamic predictive models, ensuring customer-centric campaigns that encourage maximum engagement and build long-lasting relationships with the bank.

II. Basic Concepts

A. Grid Search CV: In machine learning model tuning, Grid Search CV allows the users to choose the most favorable set of hyper-parameters concerning a specified grid of parameters. The algorithm then tests a given set of hyper-parameters and applies cross-validation to find an optimal set.

B. Randomized Search CV: Randomized search CV refers to a hyperparameter search method in which a random selection of subset hyperparameters from specified ranges is done rather than exhaustive GridSearch CV trials. This method, therefore, maintains very good performance with lower computational cost.

C. Bayesian Optimization: Bayesian Optimization serves as a black-box function optimization method that produces GP surrogate models during iterative processes without requiring gradient inputs. A balance between exploration and exploitation emerges from using mean predictions and uncertainty values inside its acquisition function.

D. Random Forest: Random Forest operates as an ensemble learning solution capable of handling both classification and regression tasks. Random Forest builds many decision trees at training time before it unites their predictive results to achieve higher accuracy in predictions. The training of each tree happens using a random selection of bootstrap data while it chooses features from a random subset during splitting procedures. The algorithm benefits from enhanced generalizability because of this feature.

E. Gradient Boosting: Gradient Boosting is a powerful machine-learning technique used for both classification and regression tasks. It builds models sequentially, where each new tree corrects the errors of the previous ones. Unlike Random Forest, which builds trees independently, Gradient Boosting optimizes performance by minimizing loss using gradient descent.

F. XGBoost: XGBoost functions as an open-source, high-performance gradient-boosting tool that handles classification and regression jobs. The system operates based on gradient boosting and provides maximum speed alongside superior performance functionality. Through the decision tree ensemble strategy, XGBoost runs an iterative process that reduces errors using gradient descent minimization.

G. K-Nearest Neighbour: K-Nearest Neighbour is when a new, unknown data point is found, KNN makes use of a chosen distance metric, like Euclidean distance, to select the k nearest data points that is, its neighbors, from the training set. In the case of predicting class labels or any other kind of classification problems, KNN predicts the most frequent class that it finds amongst its k nearest neighbors. A majority vote of the neighbors decides whether there is a fire or not.

H. Stacking: Ensemble learning utilizes stacking as one of its techniques by combining various models to obtain optimum predictive outcomes. A metamodel establishes its learning through the combination of outputs produced by its base models, which may include decision trees, SVMs, and neural networks. Precision enhancement occurs through stacking because different algorithms team up to discover optimal generalized results.

I. Bagging: Bagging, otherwise known as Bootstrap Aggregating, is an ensemble learning approach that improves the stability and accuracy of the model by training models on different random subsets of a dataset. The technique does reduce variance and overfitting. A major example of such an ensemble learning methodology includes Random Forest, which combines multiple decision trees to boost the predictive performance and robustness.

J. Extra Trees: Extra Trees is an extremely Randomized Trees (Extra Trees) ensemble learning method similar to Random Forest but introduces additional randomness. It selects split points randomly instead of choosing the best split, reducing variance and improving generalization. Extra Trees is faster and more robust, making it effective for high-dimensional datasets and classification tasks.

K. SVM (Support Vector Machine): The supervised learning model SVM operates as a well-known approach that addresses classification problems together with regression tasks. SVM selects the hyperplane that maintains the highest possible distance between classification groups because this yields excellent generalization capabilities. Linear or nonlinear classifications become possible through SVM models whose kernel functions exist in polynomial as well as RBF and sigmoid formats.

L. Naive Bayes: Naive Bayes operates as a straightforward classification system that depends on Bayes' Theorem. Naive Bayes functions by working with probabilistic calculations instead of executing direct feature to class label associations. The algorithm measures the probability that a particular data point belongs to each class group among its other characteristics. Naive Bayes adopts a basic rule that all features, regardless of their class names, exist independently from one another.

M. AdaBoost: Abbreviation for Adaptive Boosting, a kind of ensemble-type machine learning. Strong learning results from combining multiple weak learners through an aggregation process to enhance classification accuracy. The key feature of AdaBoost is its adaptive operation. Each iteration reads the training data point weights when updating their weights throughout the algorithm.

N. Quadratic Discriminant Analysis (QDA): In particular, they put to the test all the machine learning methods able to gauge customer attitudinal responses to various marketing campaigns undertaken. More specifically, XGBoost showed the best accuracy result of (91%, succeeded by Random Forest and SVM with 90% , Gradient Boosting and another Adaboost with 89% , Logistic Regression with 89% , KNN and DT with 88% , and finally Naive Bayes with 85% . This dataset includes some level of preprocessing to handle missing data, encode categorical variables, and carry out feature engineering most likely to improve model performance. In terms of metrics for evaluation, accuracy and precision, recall, and F1 score were analyzed. The results conclude that ensemble models, particularly Random Forest and Gradient Boosting, produce the best performance overall.

A. METHODLOGY

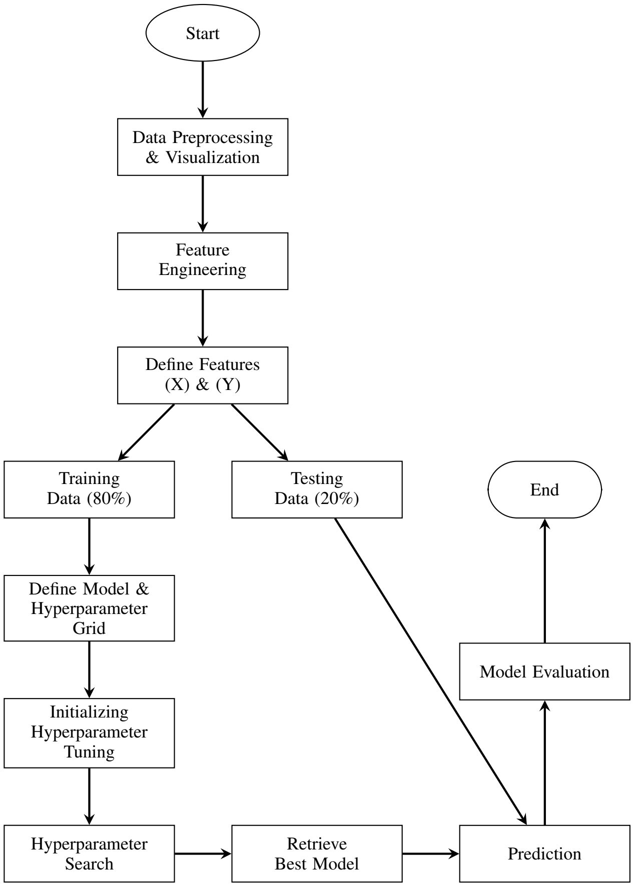

Fig. 1. Enter Caption

1) EDA: Exploratory Data Analysis reveals that the dataset is divided into a total of 45,211 rows that contain 17 features with both numerical and categorical characteristics. Residue data analysis validated the data quality results by producing correlation matrices that appeared as heat maps for interpretation. The implementation identified co-dependent features when their correlation reached 0.8, and these features should be removed. To preserve the 40,151 rows with 17 original features in the dataset, the IQR method required the removal of 4,060 data entries.

Model	Accuracy (%)	Precision	Recall	F1 Score
LightGBM	90.6668	0.6512	0.4876	0.5577
Stacking	90.6336	0.6514	0.4812	0.5535
XGBoost	90.5452	0.6418	0.4895	0.5554
Random Forest	90.4788	0.6808	0.3969	0.5014
Gradient Boosting	90.3572	0.6562	0.4216	0.5134
Logistic Regression	90.1581	0.6811	0.3465	0.4593
Extra Trees	90.1360	0.6799	0.3446	0.4574
Bagging	89.8153	0.6171	0.4106	0.4931
AdaBoost	89.3841	0.6202	0.3098	0.4132
Support Vector Machine	89.3398	0.6942	0.2081	0.3202
Decision Tree	87.4489	0.4806	0.4986	0.4894
QDA	86.3983	0.4362	0.4354	0.4358
Naive Bayes	85.0824	0.3927	0.4326	0.4117

TABLE I
PERFORMANCE COMPARISON OF CLASSIFICATION MODELS

2) HYPERPARAMETER TUNING: The model benefited from a random search together with Bayesian optimization to optimize the determination of key parameters, including learning rate, number of leaves, and tree depth, to reach superior bank marketing prediction outcomes. The scalability of random search strategies extended to all available hyperparameters, but Bayesian optimization improved the optimization process through its iterative probabilistic approach. Model accuracy, along with efficiency, increased due to these methods supported by cross-validation.

3) LIGHTGBM: The LightGBM Algorithm is a gradient-boosting ensemble method that relies on decision trees. It can be used for both classification and regression by developing decision trees that expand leaf-wise, meaning that for any given scenario, only one leaf will split based on the gain. In some cases, especially with smaller dataset sizes, these trees tend to overfit. LightGBM also uses a distribution histogram to bucket data into bins.

Method	Accuracy	ROC AUC	Precision	Recall	F1 Score
Default	0.9649	0.9934	0.9589	0.9859	0.9722
GridSearchCV	0.9737	0.9928	0.9722	0.9859	0.9790
RandomizedSearchCV	0.9737	0.9912	0.9722	0.9859	0.9790
BayesianSearchCV	0.9737	0.9921	0.9722	0.9859	0.9790

TABLE II
SUMMARY OF MODEL PERFORMANCE

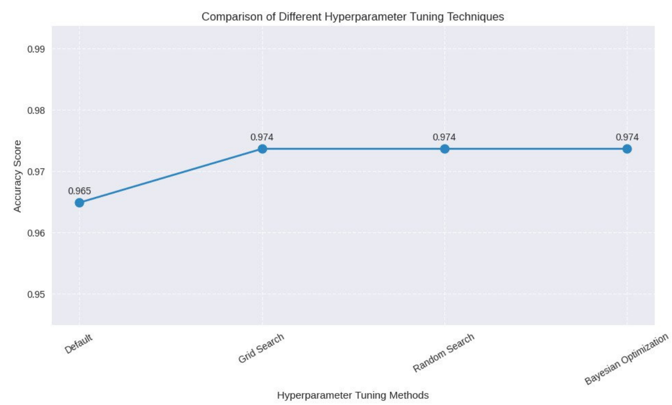

Fig. 2. Comparison of Different Hyperparameter Tuning Techniques

a) Best Hyperparameter Configurations: The optimal hyperparameters obtained using different search methods are listed below:

GridSearchCV

```
{'colsample_bytree': 0.8,
'learning_rate': 0.1,
'n_estimators': 300, 'num_leaves':
30, 'subsample': 0.6}
```

RandomizedSearchCV

```
{'subsample': 0.625, 'num_leaves':
120, 'n_estimators': 450,
'learning_rate': 0.464,
'colsample_bytree': 0.875}
```

BayesianSearchCV

```
OrderedDict([('colsample_bytree',
0.5), ('learning_rate',
1.0), ('n_estimators', 319),
```

```
('num_leaves', 200), ('subsample',
0.5)])
```

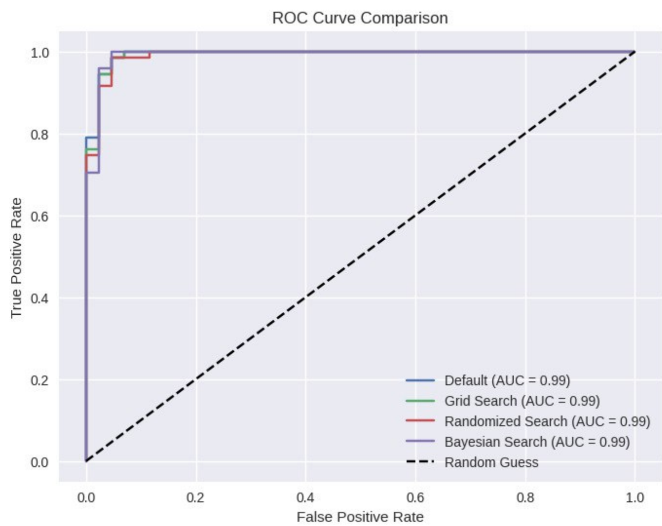

Fig. 3. ROC Curve Comparison for Different Hyperparameter Tuning Techniques

b) AUC-ROC Curve Analysis: The ROC curve demonstrates how LightGBM models define the relationship between their True Positive Rate (TPR) and their False Positive Rate (FPR) during optimization with various hyperparameter tuning approaches.

- **True Positive Rate (TPR) / Recall:** Measures how accurately the model detects genuine positive target variables.
- **False Positive Rate (FPR):** The metric calculates how often negative cases get misidentified as positive instances.
- **High Model Performance:** All models display an AUC value of approximately 0.99, which indicates outstanding classification capability.
- **Close Overlapping Curves:** All ROC curves from Grid Search, Random Search and Bayesian Optimization show no significant variation between their performance outcomes.
- **Steep Initial Rise:** A progressive reduction in FPR combined with high TPR shows that the model works efficiently at spotting positives early without many false findings.
- **Flattening at the Top:** The TPR approaches 1.0 when FPR values rise, which demonstrates that maximizing recall will inevitably produce additional false positive results.
- **Random Baseline:** All tuned models display outstanding performance compared to a random classifier (AUC = 0.5), as shown by the diagonal dashed line.

This evidence demonstrates that hyperparameter adjustment generates minimal performance improvements because the default model performs equally well during classification.

4) LOGISTIC REGRESSION: The statistical and machine learning algorithm Logistic Regression serves as a popular method for performing binary and multi-class classification operations. The logistic function serves as a mapping method to compute the estimated values that range between 0 and 1 in order to predict category membership probabilities for a given input. The model utilizes L1 (Lasso) and L2 (Ridge) regularization techniques for overfitting prevention that enhances generalization abilities.

Method	Accuracy	ROC AUC	Precision	Recall	F1 Score
Default	0.9016	0.9045	0.6811	0.3465	0.4593
GridSearchCV	0.9019	0.9044	0.6753	0.3602	0.4698
RandomizedSearchCV	0.9019	0.9044	0.6753	0.3602	0.4698
BayesianSearchCV	0.9021	0.9044	0.6770	0.3611	0.4710

TABLE III
SUMMARY OF LOGISTIC REGRESSION PERFORMANCE

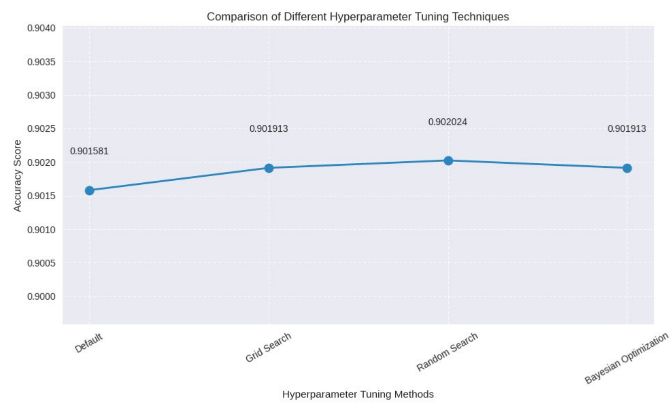

Fig. 4. Comparison of Different Hyperparameter Tuning Techniques

a) Best Hyperparameter Configurations: The optimal hyperparameters obtained using different search methods are listed below:

GridSearchCV
```
{'C': 100, 'penalty': 'l1',
'solver': 'liblinear'}
```
RandomizedSearchCV
```
{'C': 12.74, 'penalty': 'l1',
'solver': 'liblinear'}
```
BayesianSearchCV
```
{'C': 59.89, 'penalty': 'l2',
'solver': 'liblinear'}
```

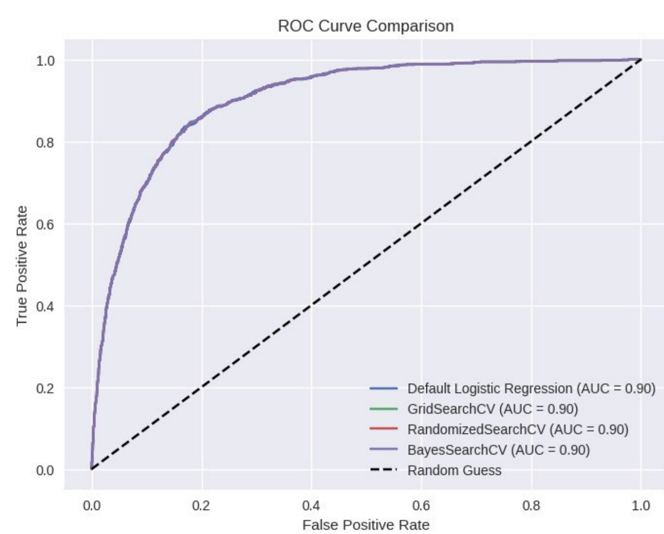

Fig. 5. ROC Curve Comparison for Different Hyperparameter Tuning Techniques

b) AUC-ROC Curve Analysis:

- **Consistent Model Performance:** The classification models reach a level of 0.90 AUC accuracy.
- **Close Overlapping Curves:** The ROC curves from different hyperparameter tuning techniques show minimal variation.
- **Steep Initial Rise:** The curve demonstrates efficient positive classification by displaying a rapid rise of TPR at low FPR levels.
- **Flattening at the Top:** The measurement of FPR shows that TPR saturation occurs when the process of increasing recall numbers in detections produces additional false positive results.
- **Random Baseline Comparison:** Strength of prediction is confirmed by the superior performance of tuned models which surpasses the random classifier (AUC = 0.5).

5) DECISION TREE: Decision Trees: The supervised learning algorithm known as Decision Trees allows users to tackle both classification and regression problems efficiently. The method uses feature conditions to split data groups into separate branches that form an organization like a tree. The decision rules that make up the tree structure are stored at internal nodes while the outcomes are found in leaf nodes. Decision Trees possess straightforward interpretability features that can deal with both numerical and categorical data types.

Method	Accuracy	ROC AUC	Precision	Recall	F1 Score
Default	0.8757	0.7142	0.4854	0.5014	0.4932
GridSearchCV	0.8975	0.8437	0.6459	0.3323	0.4392
RandomizedSearchCV	0.8974	0.8436	0.6453	0.3318	0.4383

TABLE IV
SUMMARY OF DECISION TREE MODEL PERFORMANCE

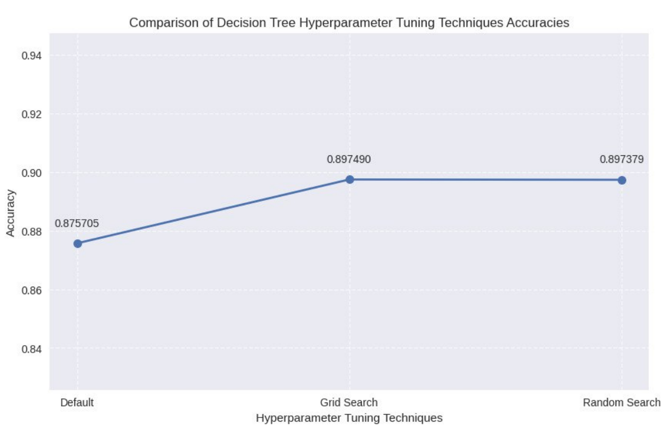

Fig. 6. Comparison of Different Hyperparameter Tuning Techniques for Decision Tree

a) Best Hyperparameter Configurations: The optimal hyperparameters obtained using different search methods are listed below:

GridSearchCV

```
{'criterion': 'gini', 'max_depth':
5, 'min_samples_leaf': 5,
'min_samples_split': 2}
```

RandomizedSearchCV

```
{'min_samples_split': 3,
'min_samples_leaf': 4, 'max_depth':
5, 'criterion': 'gini'}
```

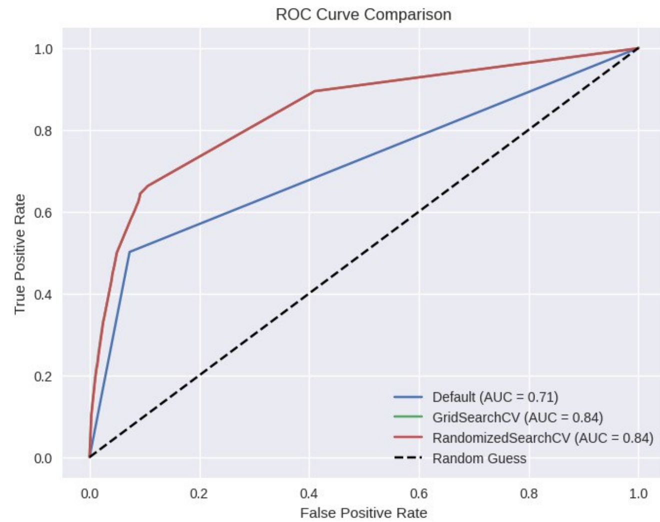

Fig. 7. ROC Curve Comparison for Decision Tree Hyperparameter Tuning Techniques

b) AUC-ROC Curve Analysis:

- **Model Performance:** The optimized models achieved an AUC value of 0.84 which provided better performance than default Decision Tree model AUC of 0.71.
- **Close ROC Curves:** The performance results from GridSearchCV models match those of RandomizedSearchCV models.
- **Early Rise in Curve:** Rapid improvement in TPR with minimal FPR increase shows that the tuned models can detect positives efficiently.
- **Comparison with Random Baseline:** All Decision Tree models demonstrate superior performance than the baseline random classifier with an AUC value of 0.5, which substantiates their effectiveness.

III. FIGURES

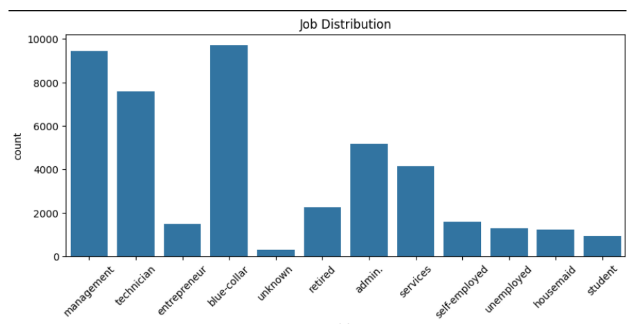

Fig. 8. JOB DISTRIBUTION

Job distribution forms an important aspect of bank marketing campaigns; people with various jobs show interest in different forms of financial products. In the database, customers of different occupations come under management, blue-collar, technician, and administrative jobs. Some jobs, such as management and technicians, might exhibit higher engagement towards bank products, and others, like blue-collar workers, tend to show a lower response rate.

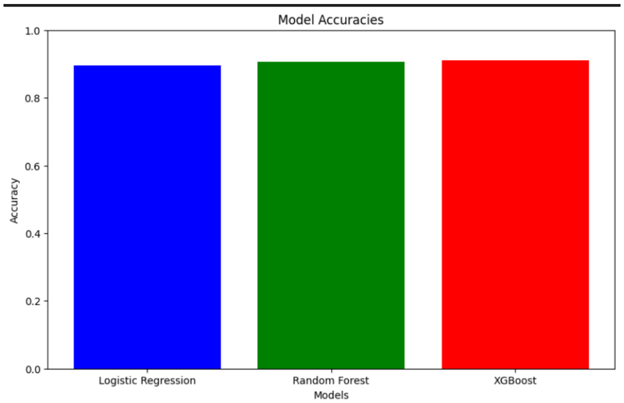

Fig. 9. MODEL ACCURACY

Logistic Regression (blue) has high baseline accuracy but is limited to capturing the more complex relationships within the dataset. Random Forest (green) is an ensemble learning technique that improves the accuracy of prediction by reducing the variance from the multiple decision trees. XGBoost (red) had the highest accuracy with gradient-boosting techniques to refine predictions iteratively.

IV. RESULTS

According to the study, the best models for forecasting consumer reactions in bank marketing are LightGBM, Stacking Classifier, and XGBoost, with respective accuracies of 90.67%, 90.63%, and 90.54%. These models perform very well because they can handle sizable and intricate datasets and identify significant patterns in consumer behavior. The accuracy of these models was greatly increased by employing sophisticated machine learning approaches, such as feature engineering and hyperparameter tuning, guaranteeing more trustworthy predictions. While hyperparameter tuning modifies the model's internal settings to maximize performance, feature engineering entails developing new input features or converting preexisting ones to provide the model with more meaningful data.

Banks can benefit greatly from machine learning approaches in terms of focused marketing and customer segmentation. Banks can divide their clientele into discrete categories according to their propensity to interact with particular offers by precisely forecasting customer reactions. This degree of accuracy makes highly targeted marketing efforts possible, in which offers and messaging are customized to each segment's requirements and preferences. Marketing initiatives become more effective as a result of increasing conversion rates and enhancing consumer interaction.

V. CONCLUSION

This study confirms that machine learning is a game-changer in bank branding by demonstrating a high ability to predict customer responses. Among the various models tested, LightGBM, Stacking Classifier, and XGBoost ranked the highest in accuracy compared to traditional methods such as Logistic Regression, with 90.67%, 90.63%, and 90.54% accuracy, respectively. The incorporation of feature engineering, data preprocessing, and hyperparameter tuning proved to be critical for enhancing model performance. Such advanced techniques enable better customer segmentation, allowing banks to identify target audiences with greater potential and design more effective marketing strategies. Using real-time predictive modeling, banks can reduce costs and increase customer engagement through better decision making. Machine learning for targeted marketing therefore represents a data-driven approach that facilitates better resource allocation and higher campaign success rates. In summary, AI-powered models pave the way for banks to operate more efficiently with a customer-centric and data-driven focus. Machine learning provides a powerful tool for modern banking marketing strategies, improving marketing outcomes, and driving business growth.

REFERENCES

[1] Hossain, Md Riyad, Timmer, Douglas, and Moya, Hiram. (2021). "Machine learning model optimization with hyper-parameter tuning approach."

[2] Shipinkoski, S., and Gjoreski, H. (2024). "Analyzing Telecom Customer Churn Using Machine Learning."

[3] Ramachandran, R., Dharani, J.S., and Lanso, J. (2024). "Comparative Analysis of Machine Learning Algorithms for Breast Cancer Detection and Classification."

[4] Panda, A. R., Rout, S., Narsipuram, M., Pandey, A., and Jena, J. J. (2024). "Ad Click-Through Rate Prediction: A Comparative Study of Machine Learning Models," 2024 International Conference on Emerging Systems and Intelligent Computing (ESIC), Bhubaneswar, India, 2024, pp. 679-684, doi: 10.1109/ESIC60604.2024.10481562.

[5] American Diabetes Association, "Diagnosis and classification of diabetes mellitus", Diabetes Care, vol. 43, no. Suppl. 1, pp. S14-S31, 2020.

[6] Shaw, J. E., Cho, N. H., Huang, Y., da Rocha Ohlrogge, A. W., Fernandes, J. D., and Malanda, B. (2018). "IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045", Diabetes Res. Clin. Pract., vol. 138, pp. 271-281.

[7] Obiniyi, A. A., Olokoba, A. B., and Olokoba, L. B. "Diabetes mellitus in Nigeria: The past, present, and future", Diabetes Mellitus- Insights and Perspectives, IntechOpen.

[8] Mehmood, M., and Mehmood, N. (2023). "Predictive analysis of diabetes mellitus using machine learning algorithms", Proc. Int. Conf. Data Eng. (ICDE'23), pp. 220-229.

[9] Mangukiya, M., Vaghani, A., and Savani, M. (2022). "Breast cancer detection with machine learning", International Journal for Research in Applied Science and Engineering Technology, vol. 10, no. 2, pp. 141-145.

[10] Apte, S. S., and Dangare, C. S. (2012). "Enhanced research on the data mining classification strategies for heart disease prediction systems", Int. J. Comput. Appl, vol. 47, no. 10, pp. 44-48.

[11] Nayak,Amlan, et al. "Language Detection Using Machine Learning." 2023 OITS International Conference on Information Technology (OCIT), 2023, pp. 732–37.

[12] Chakraborty, Srijita, et al. "Predicting Diabetes: A Comparative Study of Machine Learning Models." 2023 OITS International Conference on Information Technology (OCIT), 2023, pp. 743–48.