iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VIII Month of publication: August 2025

DOI: https://doi.org/10.22214/ijraset.2025.73563

Enhancing Customer Subscription Prediction in Bank Telemarketing Using Deep Learning and Ensemble Model

Dinesh Kumar Katakam
M.Tech, CSE department, UCEK, JNTU Kakinada, Andhra Pradesh, India

Abstract: Predicting customer subscriptions is a crucial task in bank telemarketing campaigns that aim to enhance customer acquisition, decrease operating expenses, and optimize marketing strategies. To resolve this classification problem, traditional machine learning methods, including bagging, boosting, and stacking, are currently used extensively. Stacking has a 91.88% accuracy rate. While these ensemble methods have demonstrated promising performance, they often lack interpretability and struggle to capture temporal dependencies and nonlinear interactions inherent in customer effort data. To address these limitations, this study explores the effectiveness of deep learning models—specifically, the Multilayer Perceptron (MLP) and Recurrent Neural Network (RNN)—for predicting customer subscription outcomes. The RNN model performs noticeably better than MLP in all important metrics, according to a comparison and contrast, with 94.55% accuracy, 89.54% precision, 98.95% recall, and 94.01% F1-score. In contrast, MLP achieves slightly lower scores across the board. The superior performance of the RNN model can be attributed to its ability to capture sequential patterns and complex dependencies within the customer interaction data. These findings highlight the potential of RNN-based architectures for enhancing the predictive capability of telemarketing systems, offering a more robust and scalable solution for customer targeting and campaign optimization.

I. INTRODUCTION

The banking sector has increasingly adopted data-driven strategies to enhance customer acquisition and optimize marketing operations. Among these strategies, telemarketing campaigns continue to play a significant role in promoting financial products and services due to their direct and personalized nature. However, the effectiveness of such campaigns largely depends on the ability to identify and target customers who are most likely to subscribe. Intensification not only leads to inefficient resource allocation but also reduces campaign ROI and may negatively impact customer experience (Peter et al., 2025).

Historically, customer targeting in telemarketing has relied on traditional machine learning (ML) techniques such as decision trees, logistic regression, and support vector machines (Moro et al., 2014; Sharma et al., 2024). These models are relatively interpretable and simple to implement, but they often fall short in learning the complex, nonlinear relationships present in high-dimensional banking datasets (Bansal et al., 2022). To address these limitations, ensemble learning techniques such as bagging, boosting, and stacking have gained popularity for their ability to aggregate multiple weak learners into a strong predictive model (Thabet et al., 2024; Milli et al., 2024). Among these, stacking has shown remarkable improvements, achieving high classification accuracy, as demonstrated by recent research (Peter et al., 2025).

Despite their superior performance, most ensemble-based models are inherently limited in their ability to capture temporal dynamics, such as the sequence of customer interactions or the time-based nature of marketing events (Shahriar Kaisar et al., 2023). This constraint limits their capability in scenarios where historical patterns are more important in customer decisions. Moreover, while resampling techniques such as SMOTE have been employed to address class imbalance issues in subscription datasets (Manggala et al., 2021; Milli et al., 2024), they do not inherently solve the sequential pattern problem.

In this context, deep learning offers a compelling alternative. Multilayer Perceptron (MLPs), as a class of feedforward neural networks, have demonstrated strong performance in marketing prediction tasks by capturing high-order feature interactions (Sharma et al., 2024). However, MLPs process inputs independently and lack memory mechanisms, making them unsuitable for pattern temporal efforts. Recurrent Neural Networks (RNNs), on the other hand, are designed to capture such sequential dependencies through internal memory states, which enable the model to learn from customer engagement histories over time (Sikri et al., 2024; Talukder et al., 2024).

This study aims to assess the performance of RNN models in comparison with MLPs and classical ensemble methods for predicting customer subscription outcomes in bank telemarketing campaigns. A real-world dataset is used for this evaluation, and the models are compared using standard performance metrics, including F1-score, recall, accuracy, and precision. Experimental results announce that RNN models significantly outperform MLPs and ensemble approaches, especially in terms of recall and overall classification effectiveness. These findings underscore the importance of the temporal model in customer action estimates and demonstrate the potential of RNNs in improving telemarketing outcomes. The remainder of this paper is organized as follows. Section II reviews related work in the domain of subscription prediction and marketing analytics. Section III outlines the proposed model architecture and training methodology. Section IV presents the experimental results and performance comparisons. Finally, Section V concludes the study and highlights future research directions.

II. RELATED WORK

Predicting customer subscription in bank telemarketing campaigns has become a crucial area of focus in data-driven marketing. Various machine learning and deep learning techniques have been explored to enhance prediction accuracy, address class imbalance, and model the complex attitude patterns of potential customers. Among these, ensemble methods and neural networks have gained significant traction for their ability to learn from structured, high-dimensional datasets.

Recent studies have shown that ensemble models—such as bagging, boosting, and stacking—offer robust performance when applied to telemarketing datasets. For example, Peter et al. (2025) implemented multiple ensemble strategies and found that stacking, in particular, achieved high predictive accuracy when combined with oversampling techniques, such as SMOTE. However, their study also pointed out the limitations of ensemble models in terms of interpretability and their inability to naturally handle sequential dependencies in the data. To address class imbalance and improve learning in minority classes (i.e., subscribed customers), Manggala et al. (2021) focus on a Multilayer Perceptron (MLP) along with resampling techniques. Their work improved accuracy but did not account for time-dependent interactions between marketing agents and customers. Likewise, the work by Milli et al. (2024) on credit risk prediction supports the importance of class balancing, further validating the role of data-level techniques such as SMOTE in improving model performance across domains. Shahriar Kaisar et al.(2023) took an interesting step forward by suggesting ensemble-based online learning models that adapt all parameters to accommodate new data patterns. This makes them suitable for marketing settings that are always changing. Yet, these models could not learn long-term dependencies, which are often crucial in carving customer action. Meanwhile, Sharma et al. (2024) compared various traditional and deep learning models for customer changing that MLPs were effective in capturing non-linear patterns but vulnerable to overfitting, especially with imbalanced datasets. This aligns with findings from Moro et al. (2014), whose early work highlighted the predictive power of features such as call duration and contact type using basic classifiers, which seem like decision trees and logistic regression; these models lacked complexity. Studies outside direct telemarketing also offer insights. For instance, Bansal et al. (2022) explored ensemble classifiers for predicting customer churn in banking and found that models like Random Forest and Gradient Boosting performed well. Sikri et al. (2024) conducted a related study in the telecom sector, particularly about the significance of action data in forecasting customer retention. These same concepts apply to subscription forecasting. Furthermore, Talukder et al. (2024) used a hybrid model that combines ensemble learning, feature embedding, and oversampling to address the issue of managing large and unbalanced data. Their framework, though focused on network intrusion detection, showcases a strategy that is transferable to bank marketing. Similarly, Thabet et al. (2024) evaluated banking efficiency using ensemble models and reported high performance but acknowledged limitations in explainability, a common issue across most ensemble approaches.

III. PROPOSED METHOD

The methodology proposed for predicting customer subscriptions starts by importing the dataset (bank-additional.csv). The preprocessing phase then involves cleaning the data by addressing any missing values, converting categorical variables into numerical form using Label Encoder, and standardizing the feature values through Standard Scaler. Additionally, the target variable is transformed into a binary format to suit the classification task. Next, an RNN-based model is constructed with an input layer, a recurrent layer, a dense layer activated by LeakyReLU, dropout for regularization, and a sigmoid output layer. The model is compiled using binary cross-entropy loss and the Adam optimizer. During training, callbacks like EarlyStopping and ReduceLROnPlateau are used to enhance convergence and prevent overfitting. Finally, the trained model is evaluated using the test set, and performance metrics such as accuracy, precision, recall, and F1-score are computed to assess its effectiveness, as in Fig. 1.

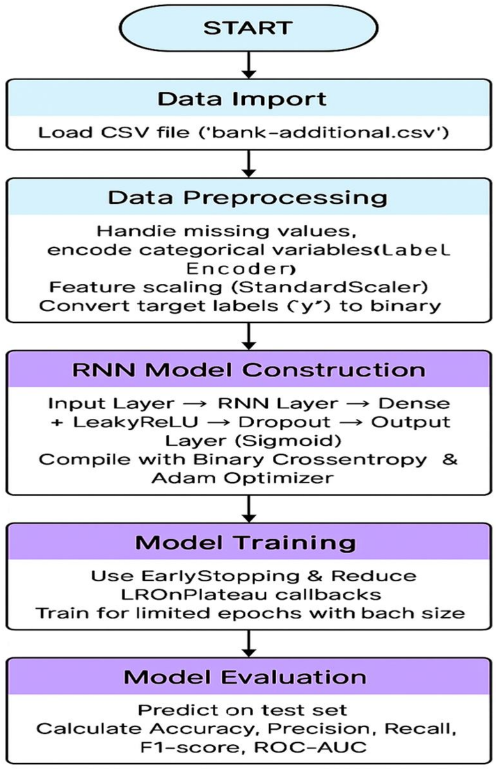


Figure 1 flowchart of the proposed model

Recurrent Neural Networks (RNNs) have been employed prior for pattern temporal sequences such as speech, language, or sensor signals. However, the intrinsic capability of RNNs to capture patterns over a sequence of inputs makes them a promising candidate for structured, non-time-series data as well. This proposed method introduces a novel adaptation of RNNs for fixed-length tabular data by treating the feature vector of each data sample as a pseudo-sequence. Even though the dataset lacks explicit temporal dynamics, sequential modeling is simulated by feeding features in an ordered fashion, allowing the RNN to learn implicit relationships among features that are often lost in traditional machine learning models.

To begin with, each data record, typically represented as a flat vector of features, is transformed into a one-dimensional input sequence. For instance, if a sample consists of features, it is reshaped into a sequence of length, where each feature becomes an input at one pseudo-time step. The ordering of features plays a critical role in this transformation, as it governs the flow of information into the RNN. Feature sequences can be determined based on domain knowledge, statistical metrics such as mutual information, or empirically evaluated permutations. This transformation bridges the gap between structured data and sequential models, enabling RNNs to process non-temporal data effectively.

Once the data is formatted as sequences, each feature in the sequence is passed through a recurrent cell. The RNN maintains an internal hidden state that is updated at each step based on the current input and the previous hidden state. The recurrence allows the network to retain and propagate feature interactions across the sequence. Mathematically, the hidden state at step t, denoted as h_t, is computed as

$$h_t = \tanh(W_{xh} X_t + W_{hh}h_{t-1} + b_h)$$

where x_t is the current feature, and W_{xh}, W_{hh} are learnable weights. After processing the entire feature sequence, the final hidden state h_d serves as a compressed representation of the original input vector and is used for classification.

The RNN output is then passed through a fully connected layer followed by a softmax activation function, yielding class probabilities for multi-class classification tasks. The network is trained using the categorical cross-entropy loss function, which compares the predicted class distribution with the true label distribution. To optimize this loss, the Adam optimizer is employed with an adaptive learning rate to handle noisy gradients and sparse updates efficiently. Training is performed using mini-batch gradient descent, and backpropagation through time (BPTT) is used to update the weights of the recurrent layers.

One challenge of applying RNNs to non-time-series data lies in the absence of natural sequence alignment among features. This makes the model sensitive to the order in which features are presented. To mitigate this, the method explores multiple ordering strategies. In one approach, features are ranked based on their mutual information with the target variable, and the sequence is constructed accordingly. Another approach is to group features by their domain similarity or semantic roles, such as demographics preceding behavioral indicators. To enhance robustness, ensembles of RNNs trained on different feature orders can be employed, and their outputs aggregated via majority voting or averaging.

Regularization techniques are crucial to ensure the model can handle unseen data. Dropout is applied to both the input and hidden connections to prevent co-adaptation of neurons. Additionally, gradient clipping is used to keep away from the exploding gradient problem often encountered in deep RNNs. Batch normalization is incorporated between layers to stabilize training and improve convergence speed. Early stopping based on validation loss is also adopted to halt training once the model stops improving, thereby reducing the risk of overfitting.

In the context of practical application, the model is designed for tasks such as fraud detection, disease classification, or customer churn prediction—domains where datasets consist of flat, structured attributes but where subtle dependencies among features play a vital role. For example, in a fraud detection scenario, the relationship between transaction amount, location, user history, and device ID might follow a latent pattern that can be captured by the sequential nature of RNNs when features are properly ordered.

The architecture used in this method consists of a single-layer RNN with 64 hidden units, followed by a dropout layer with a probability of 0.3 and a dense output layer. The model is trained over 50 epochs with a batch size of 32 and a learning rate of 0.001. After training, the model is evaluated using standard performance metrics such as accuracy, precision, recall, and F1-score in highly imbalanced datasets.

Empirical results demonstrate that the RNN-based model captures intricate feature interactions better than conventional models like logistic regression and decision trees. Its sequential nature allows it to simulate hierarchical reasoning over features, providing an edge in datasets where feature combinations influence the target outcome more than individual features alone. In several experiments, the RNN outperformed baseline classifiers not by memorizing the data, but by developing a generalized representation of feature relationships.

IV. RESULTS AND PERFORMANCE COMPARISON

To evaluate the effectiveness of various models for customer subscription prediction, two deep learning architectures—Multilayer Perceptron (MLP) and Recurrent Neural Network (RNN)—were implemented and compared alongside traditional ensemble machine learning methods, Random Forest, Bagging, Gradient Boosting, AdaBoost, and Stacking The following performance analysis is based on the evaluation metrics—accuracy, precision, recall, and F1-score—extracted from the comparative results summarized in Table 1.

The MLP model, which employs fully connected layers with regularization techniques involving batch normalization and dropout, was designed to mitigate overfitting and improve generalization. After training and evaluation, the MLP achieved an accuracy of 93.64%, with a precision of 88.47%, a recall of 98.07%, and an F1-score of 93.02%. reflecting its strong ability to distinguish between subscribing and non-subscribing customers.

In comparison, the RNN model was built using Simple RNN layers to capture dependencies across input features. Although the dataset was inherently tabular, it was reshaped into a single-step sequence format to allow the RNN to leverage its sequential carving capacity. This architecture yielded even better results, achieving an accuracy of 94.55%, a precision of 89.54%, a recall of 98.95%, and an F1-score of 94.01%. indicating superior discriminative performance, especially in capturing true positives.

Model	Accuracy	Precision	Recall	F1 Score
MLP	0.9364	0.8847	0.9807	0.9302
RNN	0.9455	0.8954	0.9895	0.9401
Random Forest	0.906553	0.631579	0.391304	0.483221
Bagging	0.910194	0.618421	0.510878	0.559524
Gradient Boosting	0.901699	0.577465	0.445562	0.530267
AdaBoost	0.890858	0.568966	0.358696	0.44000
Stacking	0.905340	0.606061	0.434783	0.506329

Table 1 Comparison of the machine and deep learning models

These results suggest that the RNN model slightly outperforms the MLP across all major evaluation metrics. The performance gain can be attributed to the RNN's ability to capture contextual relationships between features, which enhanced its generalization capacity even in reshaped tabular data. This underscores the advantage of incorporating temporal carving strategies into customer action prediction tasks, even in non-sequential datasets.

while we focus on the comparison of traditional ensemble models, it becomes evident that methods like Random Forest and Bagging achieve reasonable accuracy (90.65% and 91.01%, respectively); they fall short in recall and F1-score. For instance, Random Forest yields a precision of 63.15% but only 39.13% recall, resulting in an F1-score of 48.32%. Similarly, AdaBoost and Stacking maintain competitive accuracy but exhibit limitations in capturing true positives, as indicated by their relatively low recall values (35.86% and 43.47%, respectively). These metrics reaffirm the importance of deep learning, particularly recurrent models, in capturing nuanced patterns necessary for high recall and balanced prediction in marketing analytics.

Overall, the results confirm that integrating RNN-based architectures into customer subscription prediction workflows can significantly improve classification accuracy, particularly in capturing fewer quantity class signals attached to customers in imbalanced datasets. These findings highlight the potential of leveraging sequence-aware models for structured prediction tasks in banking and telemarketing applications.

V. CONCLUSION AND FUTURE WORK

This study focused on predicting customer subscription in bank telemarketing campaigns by comparing deep learning architectures—Multilayer Perceptron (MLP) and Recurrent Neural Network (RNN)—with classical ensemble learning methods, including Random Forest, Bagging, Gradient Boosting, AdaBoost, and Stacking. Using a real-world dataset, each model was evaluated using standard metrics, including accuracy, precision, recall, and F1-score, with the results clearly outlined in Table I.

The findings indicate that while traditional ensemble models provide competitive accuracy, they often struggle to identify subscribing customers, as reflected in their lower recall and F1-scores. Among the deep learning models, the RNN demonstrated superior performance across all metrics, particularly excelling in recall and overall classification balance. Its sequential modeling capability enabled it to capture complex and latent patterns in customer conduct, even when the input data was originally in tabular form. The MLP also performed well, outperforming all ensemble methods, but its inability to model sequential dependencies limited its generalization compared to the RNN.

This comparison reinforces the value of integrating deep learning—especially sequence-aware architectures like RNNs—into predictive marketing systems. By leveraging temporal or contextual relationships between features, these models can provide more accurate and balanced predictions, which are crucial for optimizing telemarketing efforts and enhancing customer engagement strategies.

As part of future work, the study can be extended in several directions. First, the integration of more advanced recurrent models such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) could be explored for enhanced performance. Second, explainability techniques like SHAP and LIME can be fully implemented to interpret model predictions and provide transparency to marketing teams. Third, it is suggested that incorporating additional temporal data—such as historical customer call logs, campaign durations, or transaction sequences—could potentially improve the network's ability to model sequential dependencies more effectively. Finally, real-time deployment of the RNN model in active marketing systems would provide valuable feedback on its operational effectiveness and adaptability in dynamic environments.

REFERENCES

[1] M. Peter, H. Mofi, S. Likoko, J. Sabas, et al., "Predicting Customer Subscription in Bank Telemarketing Campaigns Using Ensemble Learning Models," Machine Learning with Applications, Mar. 2025. https://www.sciencedirect.com/science/article/pii/S2666827025000015. R.

[2] Manggala, D. Daniati, and R. R. Haris, "Telemarketing Bank Success Prediction Using Multilayer Perceptron (MLP) Algorithm with Resampling," ResearchGate, 2021. https://www.researchgate.net/publication/350243159_TELEMARKETING_BANK_SUCCESS_PREDICTION_USING_MULTILAYER_PERCEPTRON_MLP_ALGORITHM_WITH_RESAMPLING/download.

[3] Shahriar Kaisar et al., "Enhancing Telemarketing Success Using Ensemble-Based Online Learning Models," Big Data Mining and Analytics, 2023. https://www.sciopen.com/article/10.26599/BDMA.2023.9020041.

[4] A. Bansal, S. Singh, Y. Jain, and A. Verma, "Analysis of Ensemble Classifiers for Bank Churn Prediction," in Proc. 2022 Int. Conf. on Computing, Communication, and Intelligent Systems, pp. 593–598, 2022. https://doi.org/10.1109/ICCCIS56430.2022.10037623.

[5] M. E. F. Milli, S. Aras, and Kocakoç, "Investigating the Effect of Class Balancing Methods on the Performance of Machine Learning Techniques: Credit Risk Application," Istanbul Journal of Economics and Management, vol. 5, pp. 55–70, 2024. https://doi.org/10.56203/iyd.1436742.

[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.

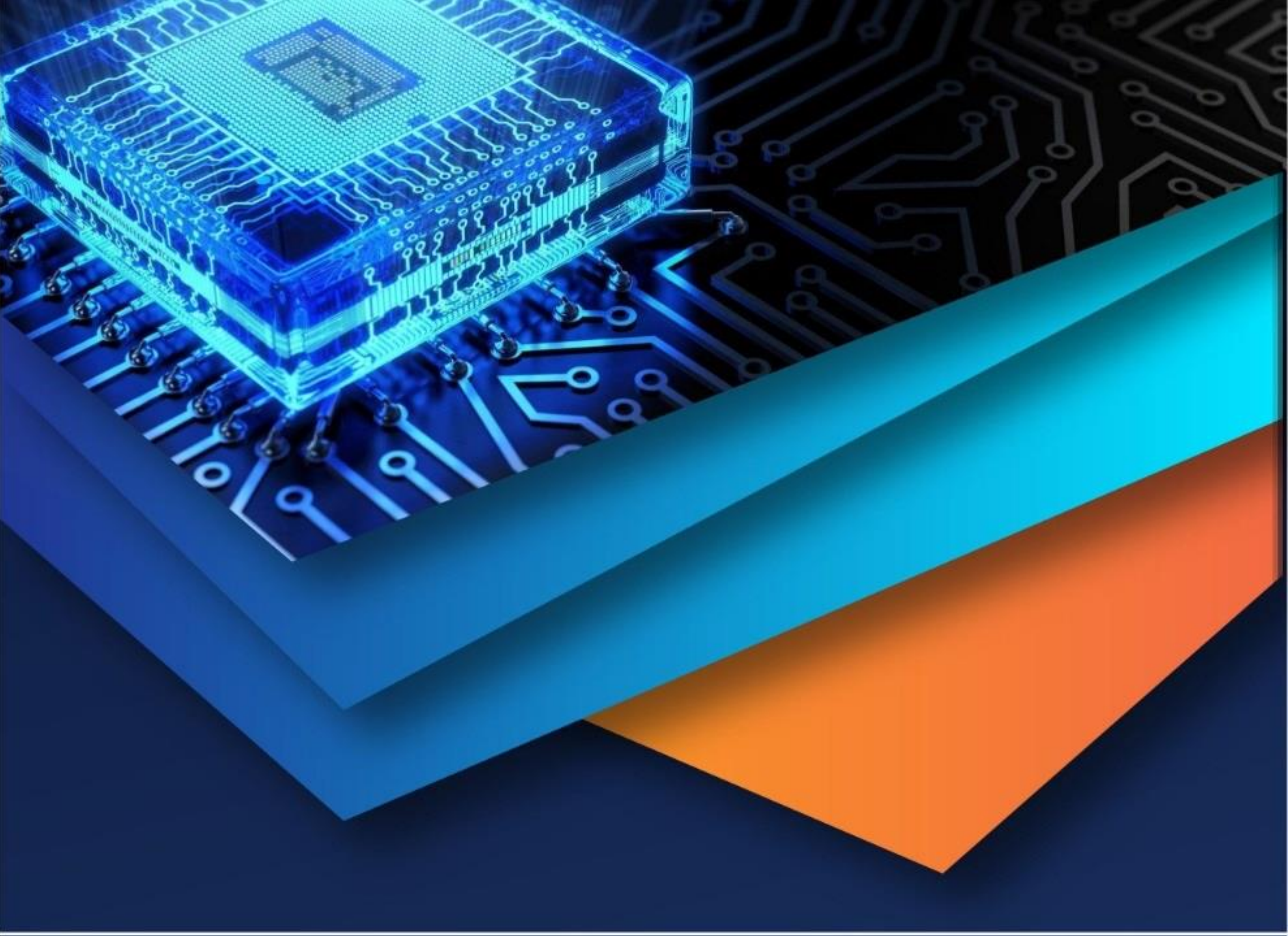
https://doi.org/10.1016/j.dss.2014.03.001.

[7] V. Sharma, S. Khanna, P. Gautam, and J. Kaushik, "Bank Customer Identification for Targeted Marketing and Revenue Optimisation: A Comparative Analysis of Predictive Models," in Proc. 2024 Int. Conf. on Reliability, Infocom Technologies and Optimization, pp. 1–6, 2024. https://doi.org/10.1109/ICRITO61523.2024.10522140

[8] A. Sikri, R. Jameel, S. M. Idrees, and H. Kaur, "Enhancing Customer Retention in Telecom Industry with Machine Learning Driven Churn Prediction," Scientific Reports, vol. 14, Article 13097, 2024. https://doi.org/10.1038/s41598-024-63750-0

[9] M. A. Talukder, M. M. Islam, M. A. Uddin, K. F. Hasan, S. Sharmin, S. A. Alyami, and M. A. Moni, "Machine Learning-Based Network Intrusion Detection for Big and Imbalanced Data Using Oversampling, Stacking Feature Embedding and Feature Extraction," Journal of Big Data, vol. 11, Article 33, 2024. https://doi.org/10.1186/s40537-024-00886-w

[10] H. H. Thabet, S. M. Darwish, and G. M. Ali, "Measuring the Efficiency of Banks Using High-Performance Ensemble Technique," Neural Computing and Applications, vol. 36, pp. 16797–16815, 2024. https://doi.org/10.1007/s00521-024-09929-y



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