NUMBERS THAT MATTER: PREDICTING BANK CUSTOMER CHURN WITH MACHINE LEARNING AND EXPLAINABLE AI

ASHRAFUL BHUYAN SHUVON
Department of Electrical and Computer Engineering, North South University, Dhaka 1229, Bangladesh

MOHAMMDAD NAHID HOSSAIN
Department of Electrical and Computer Engineering, North South University, Dhaka 1229, Bangladesh

MARGUB HUSSAIN MUSAWI
Department of Electrical and Computer Engineering, North South University, Dhaka 1229, Bangladesh

***Abstract*— This study proposes an optimized approach using robust ML Models to predict outcomes from CRM data by employing ensemble methods and the synthetic minority oversampling technique (SMOTE) for data balancing, integrating SMOTE with k-Nearest Neighbors, Decision Trees, Logistic Regression, Adaboost, and Random Forest, where Random Forest achieves a accuracy of 84.5%, outperforming other models with enhanced interpretability using explainers such as LIME and SHAP as well as different Feature Engineering techniques.**

***Keywords*— SMOTE, Normalization, Machine Learning, Random Forest, LIME, SHAP**

I. INTRODUCTION

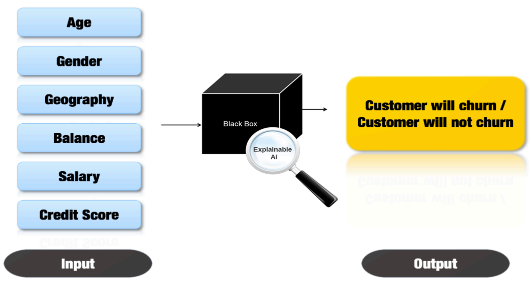

Fig. 1: Given the input, our Explainable AI model learns automatically who will churn and who will not by showing an output.

Customer retention has become a pivotal challenge for financial institutions in modern banking. Customer churn, or the rate at which clients stop doing business with a bank, not only threatens the stability of these institutions but also raises questions about profitability, client loyalty, and the competitiveness of the market as a whole. The problem of bank customer churn poses a significant concern in an era where financial institutions rely heavily on data-driven decision-making.

The gravity of the customer churn problem in the banking sector can be comprehended through alarming statistics. According to an annual report by the Banking and Payments Federation Ireland, the global banking industry experienced an average customer attrition rate of 21.9% in 2019. This figure signifies that nearly one-fifth of the customers severed ties with their banking institutions. Furthermore, according to [1] highlights the financial toll of customer churn, a bank may boost its profitability by up to 85% if it can enhance customer retention by only 5%. Consequently, predicting and preventing customer churn is not merely a matter of business continuity but is closely associated with maintaining banking institutions' financial viability and profitability.

By utilizing predictive analytics capabilities, machine learning (ML) has become a formidable tool for solving the problem of bank customer attrition. ML models can evaluate large amounts of consumer data to find trends, behaviors, and signs that anticipate client turnover. By recognizing these tendencies, banks can better prevent consumers from choosing to do business with a rival financial institution.

The available literature demonstrates that several approaches have already been suggested for the churn prediction problem [2]. For a highly imbalanced dataset, some algorithms, such as support vector machine (SVM), Naive Bayes (NB), k nearest neighbors (KNN), Etc., are significantly affected by the class imbalance and unilaterally favor the majority class [3]. In paper [4], they used SMOTE to deal with the imbalanced dataset, and appropriate data preprocessing is necessary to have the best model performance.

Existing methods to predict bank customer churn have drawbacks, such as having highly imbalanced datasets and heterogeneous data processing, class imbalance, and being limited by feature scales [4]. Lack of proper explanation of the AI/ML models about how they are making decisions, which attributes play a vital role in the entire decision-making process and making the models understandable to the customers, black box models, scalability issues for a large and dynamic dataset [5], and applying powerful algorithms like IBFR without using state of the art data balancing techniques such as SMOTE [6]. There are drawbacks to these techniques. For example, SMOTE is a widely used method in machine learning for dealing with imbalanced datasets, particularly in classification tasks where the classes are not represented equally. SMOTE oversamples the minority class in the training set but does not address class imbalance in the test set.

In our research, we tackle the challenge of imbalanced datasets by implementing diverse resampling techniques, including oversampling the minority class, under sampling the majority class, and employing SMOTE (Synthetic Minority Over-sampling Technique) to generate synthetic samples. We aim to empirically evaluate these methods to determine which proves most effective in addressing class imbalance. Furthermore, we recognize the importance of model interpretability. We will leverage Explainable AI (XAI) techniques, such as SHAP (Shapley Additive explanations) values and LIME (Local Interpretable Model-agnostic Explanations), to enhance the transparency and comprehensibility of our predictive models. This holistic approach seeks to optimize dataset balance and provide clear insights into the model's decision-making processes, offering a comprehensive solution to the limitations encountered in prior research.

This research seeks to address the limitations of existing bank customer churn prediction methods by empirically evaluating diverse resampling techniques for handling imbalanced datasets, enhancing model interpretability through Explainable AI (XAI) techniques, and providing a holistic solution to optimize dataset balance and transparency in predictive models.

1) *Comparison Between Resampling Techniques:* Our research offers an extensive comparison of oversampling, under sampling, and SMOTE for mitigating class imbalance, aiding in selecting the most effective resampling method for bank customer churn prediction.

2) *Enhanced Model Interpretability:* We employ XAI techniques, such as SHAP values and LIME, to make machine learning models transparent and comprehensible, bridging the gap between complex algorithms and practical use in the banking sector [7] [8].

3) *Holistic Solution for Imbalanced Data Challenges:* By combining resampling assessment and model interpretability, our research provides a well-rounded approach to tackle the limitations of prior methods, offering a practical framework for improving churn prediction models in the dynamic banking landscape.

4) *Develop an App:* We are poised to develop an accessible and user-friendly application that will extend the reach of our research to a broader audience. This innovative app empowers users from diverse backgrounds and industries, offering an intuitive platform for efficient and effective churn prediction.

II. RELATED WORKS

Customer Relationship Management (CRM) is a business strategy and technology-driven approach that focuses on managing and nurturing interactions with customers and potential customers. The primary goal of CRM is to build and maintain strong, long-lasting relationships with customers by understanding their needs and preferences to enhancing overall customer satisfaction. Predicting different possibilities by analyzing such data is a tedious task and requires robust AI, ML or Deep learning Models. Our findings for getting usable outcomes are stated here.

One paper used the KNN, SVM, Decision Tree and Random Forest classifiers with some feature selection methods on the churn modelling dataset from Kaggle [9]. In the data preprocessing, the method standardized the dataset and in the feature selection process, they used mRMR algorithm and Relief algorithm. Then they used oversampling before classifying the highly imbalanced dataset which has 7963 positive samples and 2037 negative churn. The Random Forest was set to 100 before training and the testing accuracy is 95.75% without using the feature selection algorithms and this is the best performance. It has been observed that for the Decision Tree and Random Forest the accuracy is high without using any feature selection algorithms. Another paper highlighted that it used repeated k-fold cross validation techniques where k=10 fold to select the best set of hyperparameters for each model on the training set comprising 90% of the dataset. Achieved a recall of 80.2% in identifying potential churners, with a 14.8% rate of misclassifying non-churners as potential churners [10]. Bayesian network, C5 tree, chi-square automatic interaction detection (CHAID) tree, classification and regression (CR) tree is used in another paper where two step clustering is implemented on the categorical variables. The C5 tree model outperformed other models by having accuracy 0.964, precision 0.914, recall 0.880, FOR 0.974, and F1 score 0.897 [11]. Another paper used SMOTE to balance the imbalance dataset before using fivefold cross validation and applying SVM, DT, ET, RF, NB etc. Among them the Random Forest tends to give the highest accuracy and F1 score of 0.86. We took this paper's performances as a benchmark since it used by far the better balancing algorithms during data preprocessing and also it showed such performance on the same dataset that is benign used in this paper [12]. Another paper received an 86% accuracy using Naïve Bayes model on a dataset used for experimentation containing 10 K records [13] consisting of 13 attributes and a class label. The classifier models are SVM, Logistic Regression, Naïve Bayes [13]. Also, another paper showed the differences between well-known classifiers performances with and without using SMOTE. It achieved a 91.90 F1 score and overall accuracy of 88.7% using RF. We also took this paper as the benchmark. In paper [14], proposed a study to find customer churn by introducing a new procedural approach. They normalize their data during data pre-processing. Then, a data cluster is formed by using a k-medoids method. The Davies-Bouldin index is used to assess clustering performance. Various neural networks (NN) were utilized to discover patterns within the data such as radial basis function (RBFNN), generalized regression (GRNN), multilayer perceptron (MLPNN), and SVM. According to the results, MLPNN and SVM models had higher precision and lower costs. In this research work, they use the CART algorithm to predict customer churning [15].

Different papers used different feature engineering algorithms and their own sets of hyperparameters for their identical datasets. Also, we tried to use Explainers like Lime and Shapley values to explain the learning characteristics of our models for the feature importance analysis. Explainable AI (XAI) refers to a set of techniques and methods used in artificial intelligence and machine learning to make the decisions and predictions of models understandable and interpretable by humans. The goal is to provide insights into how a model arrives at a particular conclusion, which is especially important in critical domains like finance where transparency and accountability are paramount. Shapley values involve considering all possible combinations of features and measuring the impact of each feature on the prediction. It takes into account both the presence and absence of each feature in different combinations and averages their contributions across all combinations. This ensures that contributions are fairly distributed among the features [8]. On the other hand, LIME (Local Interpretable Model-agnostic Explanations) is a framework used in machine learning for explaining the predictions of complex models. It's designed to provide interpretable explanations for individual predictions made by black-box models. The key idea behind LIME is to generate human-understandable explanations for why a model made a specific prediction for a given input [7].

III. MATERIALS AND METHODS

Banking datasets commonly feature skewed data, creating challenges for classifiers in accurately predicting consumer behavior. A combination of classifiers is employed in this approach to address these issues and improve performance. Moreover, to enhance interpretability, Explainable AI techniques are incorporated. An ensemble method is then applied to construct a classifier that outperforms individual classifiers. The proposed method, illustrated in Fig. 2's block diagram, encompasses multiple phases, focusing on accuracy and interpretability through Explainable AI.

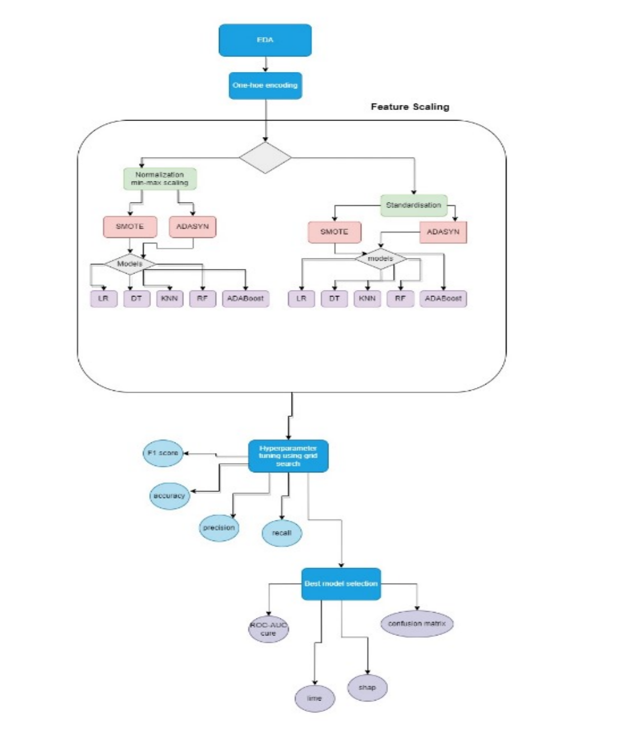

Fig.2 Block diagram, encompasses multiple phases, focusing on accuracy and interpretability through Explainable AI.

A. DATASET

The dataset used is publicly available for all in Kaggle Website [16]. The dataset consists of 14 different attributes and 10,000 instances. Out of those, the first three attributes do not contribute to the prediction of customer churn in Banking, so they have been dropped later on.

SL. NO.	Attribute	Description	Type	Role
1	RowNumber	Identifier for rows	Categorical	feature
2	IDclient	A unique identifier for each customer	Categorical	feature
3	Surname	The surname of the customer	Categorical	feature
4	CreditScore	This number is between 300 to 850 and depicts the creditworthness of a customer	Numerical	feature
5	Geography	The user's geographical location	Categorical	feature
6	Gender	User gender. Female (0) and Male (1)	Boolean	feature
7	Age	Client's age at the time of staring using the bank's service	Numerical	feature
8	Tenure	The number of years the client has attached with the bank	Numerical	feature
9	Balance	The actual bank balance of the client	Numerical	feature
10	NumOfProducts	The number of services used by the client	Numerical	feature
11	HasCrCard	The number of cradit card obtained from the bank by the client	Boolean	feature
12	IsActiveMember	Status of the client indicating whether he is using bank's services or not	Boolean	feature
13	EstimatedSalary	The estimated salary of the client	Numerical	feature
14	Exited	Binary flag indicating if the client closed an account or not	Boolean	target

Table I Details of the bank churn dataset.

B. EXPLANATORY DATA ANALYSIS (EDA)

The initial phase of this study involves meticulous preprocessing of the dataset to address challenges such as missing values and imbalanced classes. This is achieved by strategically dropping irrelevant features to prevent the non-churners' class from dominating the churners' class. The utilization of the sklearn. As detailed in Table I, the preprocessing package is integral, incorporating various utility functions and libraries for effective data preprocessing.

To delve deeper into the dataset and extract meaningful insights, Exploratory Data Analysis (EDA) is employed. This approach encompasses tasks such as handling missing values, normalizing and scaling data, and conducting univariate and bivariate analyses for numerical and categorical features. Bivariate Analysis involves using techniques like scatter plots, histograms, and heatmaps to discern relationships between variables and identify correlations among attributes.

In addition to the steps mentioned above, a thorough analysis of numerical features is conducted to assess skewness and kurtosis. Specifically, the z-score method is applied to address these statistical properties, focusing on the 'age' attribute.

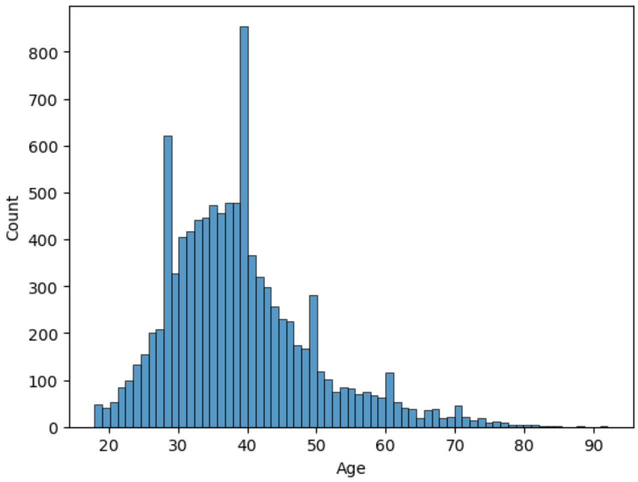

Fig. 3 Visualizing 'Age' variable

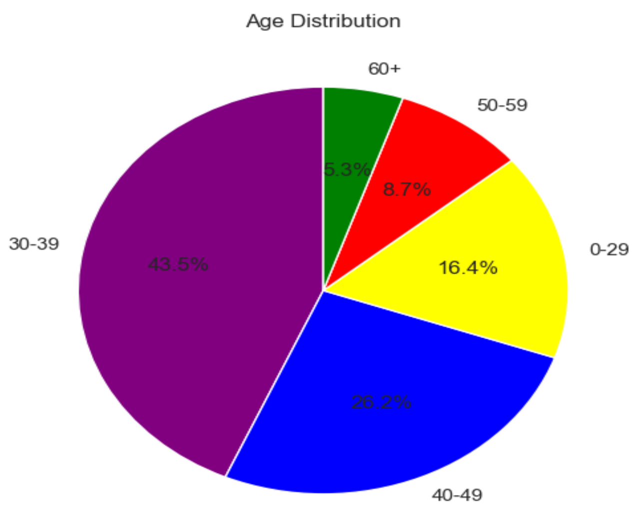

Fig. 4 Age distribution against churn customer.

Sl. No.	Libraries	Purpose
1	Pandas	Data manipulation and analysis library/ tool.
2	Numpy	Numerical computing library.
3	Matplotlib	Plotting library for visualizations.
4	Seaborn	Statistical data visualization enhancement.
5	StandardScaler	Feature standardization in machine learning.
6	get dummies()	Convert categorical variables to numerical.
7	Describe()	Descriptive statistics for DataFrame overview such as percentile, mean and so on.

Table I – Libraries and packages.

C. *DATA PREPROCESSING*

In this step, the dataset had 10,000 rows of heterogeneous data. The difference between churn (85%) and non-churn (15%) was extremely high which indicates that the dataset is highly unbalanced. As Machine Learning model needs to be trained on unbiased data, so two different types of data balancing techniques were used which are SMOTE & ADASYN. Both data balancing techniques fall under over-sampling method. Even though both techniques produced almost similar results, Smote Technique was observed to give slightly better results than ADASYN technique.

Also, two different scaling techniques were also used for Feature Scaling. One of them was Normalization which uses Min-Max Scaler to transform the values of numerical features in the range [-1-1]. The other feature scaling technique used was Standardization. It rescales the values of numerical features between 0-1.

D. *MODEL SELECTION*

For each of the data balancing techniques (i.e., SMOTE & ADASYN) and feature scaling techniques (i.e., Normalization & Standardization) we have used 5 different Machine Learning Algorithms to find out which method gives the best results. The models used were Linear Regression, K-Nearest Neighbor, Decision Tree, Random Forest, and AdaBoost.

The best accuracy obtained was from the data balancing technique, SMOTE, and the feature scaling technique, Normalization, with the machine learning algorithm, Random Forest.

Best hyperparameters for our model:

n_estimators: 100 – the greater the value, the less is the overfitting; higher n_estimators value means the more number of trees are used.

min_samples_leaf: 1 – it tells the minimum number of nodes required to be a leaf node.

max_depth: None – controls the maximum depth of each tree in the ensemble.

E. *MODEL DEPLOYMENT*

To use the model, a basic frontend is developed with html, CSS for the skeleton and JavaScript to interface with the Flask API to take input as string, converted it into the Json format like this:

{
  "CreditScore": 608,
  "Geography": "Spain",
  "Gender": "Female",
  "Age": 41,
  "Tenure": 3,
  "Balance": 89763.84,
  "NumOfProducts": 1,
  "HasCrCard": 0,
  "IsActiveMember": 0,
  "EstimatedSalary": 199304.74
}

and take the JSON inputs for scaling, dropping unnecessary columns and perform one-hot encoding before it is being fed to the Random Forest classifier to classify with the .pkl file which contains all the weightage of the trained model on the bank churn dataset. After that, the model shows the output which is shown to the frontend via Flask API and JavaScript.

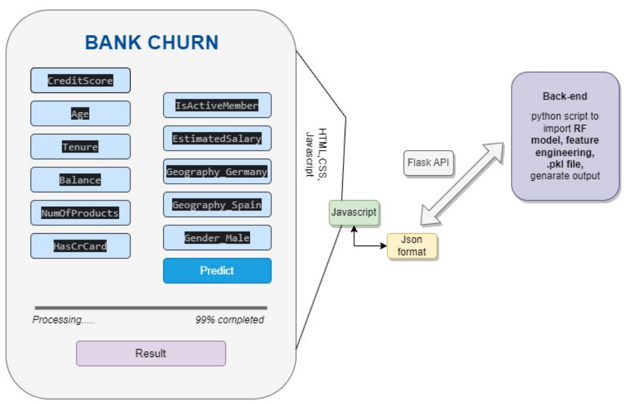

Fig. 5 APIs interface with Frontend, Backend diagram

Classifier	Hyperparameter
Linear Regression	'C': [0.1, 1.0, 10.0], 'penalty': ['l1', 'l2']
K-Nearest Neighbour	'n_neighbors': [3, 5, 7], 'weights': ['uniform', 'distance']
Random Forest	'n_estimators': [10, 50, 100], 'max_depth': [None, 10, 20], 'min_samples_leaf': [1, 2, 4]
Decision Tree	'max_depth': [None, 10, 20, 30] 'min_samples_split': [2, 5, 10]
AdaBoost	'n_estimators': [50, 100, 200], 'learning_rate': [0.01, 0.1, 1.0]

Table II Hyperparameter values for classifiers

How does Random Forest work?

A popular machine learning approach called random forest aggregates the output of several decision trees to get a single outcome. Since it uses an ensemble decision tree algorithm, the Random Forest approach consistently outperforms the Decision Tree algorithm in terms of outcomes.

IV. RESULTS AND DISCUSSION

Best accuracy was achieved from our Random Forest Model with Smote data balancing technique and standardization scaling which was 84.4%. The other models gave slightly lower accuracies, so we selected Random Forest as our best model. After accuracy, our next preferred evaluation technique was recall, as recall tells us how much of actual positives were correctly identified by the model.

Thus, higher recall helps to identify the customers which will be originally churning from the bank and not predict them as non-churn customer. From the perspective of our problem, predicting false negatives is costlier and expensive than false positive predictions. The recall achieved for our best model was 64.1%.

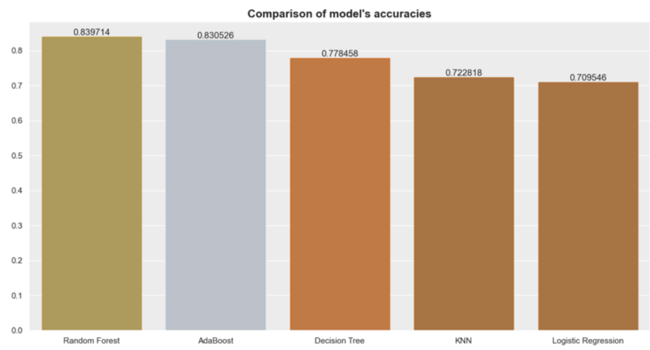

Fig. 5 Comparison between best model.

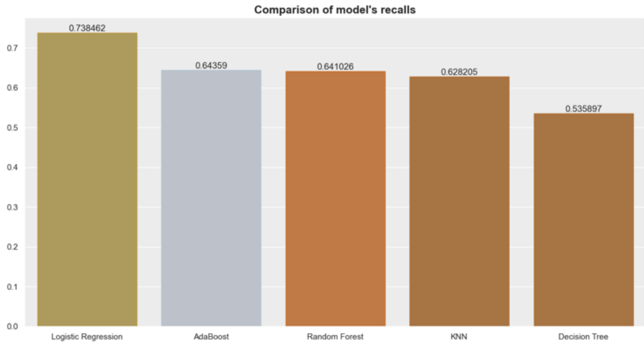

Fig. 6 Comparison of models recall values.

EVALUATION MEASURES

For the evaluation of our classification machine learning models, we have used the classification evaluation metrics which are discussed below:

A. ACCURACY

Accuracy gives the percentage of the number of correctly classifications done by the model out of all the predictions done by the model

$$Accuracy = \frac{(TP + TN)}{TP + FP + TN + FN}$$

where TP = True Positive, FP = False Positive, TN = True Negative, FN = False Negative

B. PRECISION

This metric quantifies the accuracy of positive predictions.

$$Presicion = \frac{TP}{TP + FP}$$

C. RECALL

It measures the proportion of actual positives that were correctly identified by the model.

$$Recall = \frac{TP}{TP + FN}$$

D. F-1 SCORE

It's the harmonic mean of precision and recall.

$$F1\ Score = \frac{Presicion * Recall}{Presicion + Recall}$$

E. CONFUSION MATRIX

A table that shows how well a categorization model performs. It has four components: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), and it shows the actual and projected classes.

These metrics and the confusion matrix help in understanding how well a model performs on a particular dataset. At different times, different metrics might be more important and crucial. For instance, in scenarios where false positives are costly, precision might be more important, while in other cases, recall might be the priority.

A confusion matrix provides a comprehensive view of a model's performance and aids in understanding where the model excels or requires improvement. Here, the actual negative is high compare to actual positive.

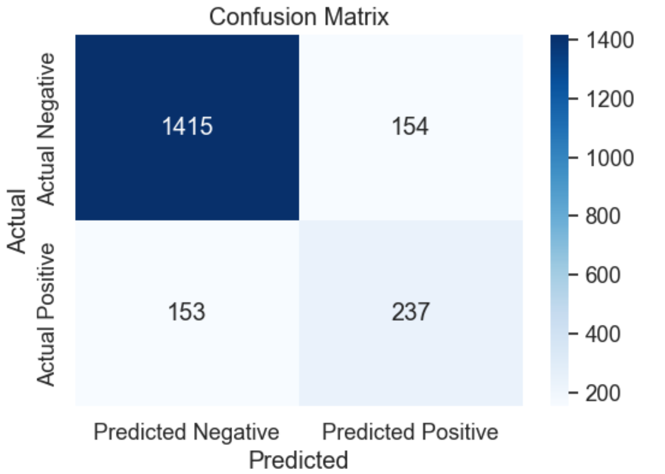

Fig. 7 Confusion matrix

F. ROC-AUC CURVE

ROC-CURVE:

It is a graphical depiction of a binary classifier's performance at different threshold values. The True Positive Rate and False Positive Rate are shown against one other.

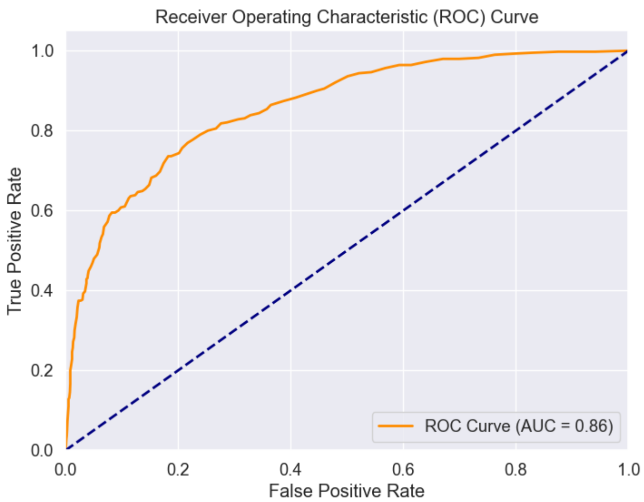

Fig. 8 ROC-Curve

AUC-CURVE:

It measures how well a binary classification model performs overall. The area under the ROC curve is denoted by the AUC. The model's ability to distinguish between positive and negative classes improves with a higher AUC.

	Accuracy	Precision	Recall	F1 Score
Normalization+ RF+ SMOTE	0.843287	0.606138	0.607692	0.607692
Normalization+ RF+ ADASYN	0.839714	0.589623	0.641026	0.641026
Standardization+ RF+ SMOTE	**0.844308**	**0.604938**	**0.628205**	**0.628205**
Standardization + RF+ ADASYN	0.831036	0.568765	0.625641	0.625641

Table III – Result comparison for different technique used with Random forest.

INTERPRETATION WITH EXPLAINABLE AI

A. LIME

LIME is a model-agnostic technique designed to provide local interpretability for individual predictions. It approximates the predictions of a complex model in a local neighborhood around the instance to be explained using a simpler, interpretable model.

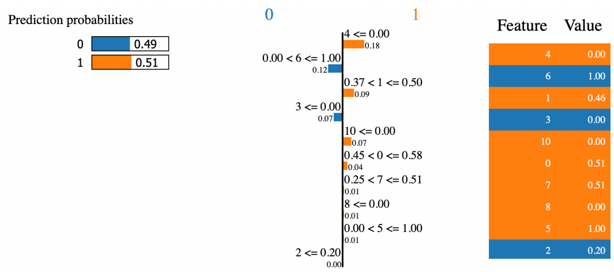

Fig. 9 LIME

Fig. 6 shows the Explainable Ai method Lime, showed the probability of which of the features would be likely to predict 0 & 1 based on the probability of similar certain events around the instance for which the prediction is to be done.

A. SHAP

SHAP values provide insights into how much each feature contributes to a prediction compared to the average prediction. They can be used to understand feature importance and individual feature effects on model predictions.

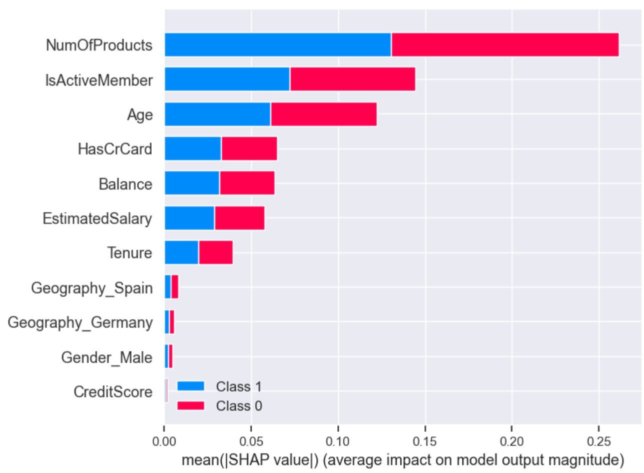

Fig. 10 SHAP

SHAP, on the other hand produced a graph which helped us to identify the features which played an important role in the model to give the predicted output. In our case, the features 'IsActiveMember', 'NumOfProducts' has the highest shap values as shown in Fig. 7, which means they have a high impact on predicting output values.

Both LIME and SHAP are valuable tools for model interpretability, but they have different scopes and approaches. LIME focuses on local interpretability for specific instances, while SHAP provides more global insights into feature importance and their effects on predictions.

WEB APPLICATION DEVELOPMENT

The developed application serves as a predictive tool within the banking sector, specifically focused on evaluating the likelihood of customer churn. Leveraging a Random Forest machine learning model, the app integrates seamlessly with a user-friendly web interface. Users input relevant information, including credit scores, demographic details, and banking behaviors, initiating a prediction request to the Flask backend. The application employs pre-trained models to generate predictions, indicating whether a customer is likely to churn or remain. This user-centric tool aims to provide valuable insights for decision-makers in the banking industry, facilitating proactive strategies for customer retention based on predictive analytics. The app's streamlined design and predictive capabilities make it a valuable asset for institutions aiming to optimize customer relationship management.

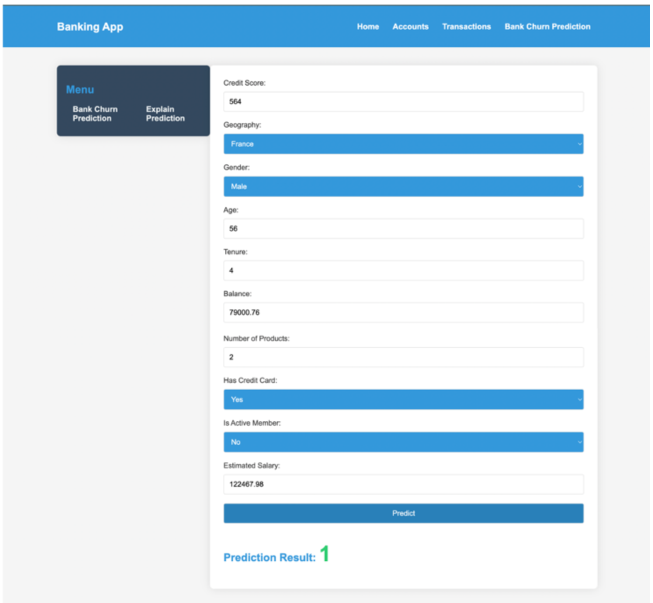

Fig. 11 User interface of Churn prediction Web App

V. CONCLUSION

For classifying the churner of bank customers, ensemble methods such as Logistic Regression, Random Forest, and Adaboost are performing well on the bank churn dataset. Still, Random Forest has a better overall performance matrix on the test dataset while applying SMOTE with standardization. Finally, explainable AI are applied to analyze and understand each input attribute's contributions in training each model being used on the dataset, allowing us to see each model's capabilities and sensitivity as well as their improvement areas.

VI. REFERENCES

[1] N. Guangli, R. Wei, Z. Lingling, T. Yingjie and S. Yong, "Credit card churn forecasting by logistic regression and decision tree," *Expert Systems with Applications,* vol. 38, no. 12, pp. 15273-15285, 2011.

[2] V. T., D. K.I., S. G. and C. K.Ch, "A comparison of machine learning techniques for customer churn prediction," *Simulation Modelling Practice and Theory,* vol. 55, pp. 1-9, 2015.

[3] K. T. Cédric Stéphane, W. Cherif and S. Hassan, "Optimizing the prediction of telemarketing target calls by a classification technique," *018 6th International Conference on Wireless Networks and Mobile Communications (WINCOM), Marrakesh, Morocco,* pp. 1-6, 2018.

[4] C. S. Koumetio Tekouabou, Ș. C. Gherghina, H. Toulni, P. Mata and J. Martins, "Towards Explainable Machine Learning for Bank Churn Prediction Using Data Balancing and Ensemble-Based Methods," *Mathematics,* vol. 10, no. 14, p. 2379, 2022.

[5] S. A and C. D, "A Survey on Customer Churn Prediction using Machine Learning Techniques," *International Journal of Computer Applications,* vol. 154, no. 10, pp. 13-16, 2016.

[6] X. Yaya, L. Xiu, N. E.W.T. and Y. Weiyun, "Customer churn prediction using improved balanced random forests," *Expert Systems with Applications,* vol. 36, no. 3, pp. 5445-5449, 2009.

[7] S. Lundberg and . S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," *NIPS,* 2017.

[8] M. T. Ribeiro, S. Singh and C. Guestrin, "Why Should I Trust You?": Explaining the Predictions of Any Classifier," *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16),* p. 1135–1144, 2016.

[9] I. Kaur and J. Kaur, "Customer Churn Analysis and Prediction in Banking Industry using Machine Learning," *2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC),* pp. 434-437, 2020.

[10] R. de Lima Lemos, T. Silva and B. Tabak, "Propension to customer churn in a financial institution: a machine learning approach," *Neural Comput & Applic,* vol. 34, p. 11751–11768, 2022.

[11] D. AL-Najjar, Nadia Al-Rousan and Hazem AL-Najjar, "Machine Learning to Develop Credit Card Customer Churn Prediction," *Journal of Theoretical and Applied Electronic Commerce Research,* vol. 17, no. 4, pp. 1529-1542, 2022.

[12] S. C. K. Tékouabou, C. G. Ștefan, T. Hamza, N. M. Pedro and M. M. José, "Towards Explainable Machine Learning for Bank Churn Prediction Using Data Balancing and Ensemble-Based Methods," *Mathematics,* vol. 10, no. 14, p. 2379, 2022.

[13] V. Koildurai, P. V. Rao, C. Selvan and R. Manuel, "Investigation on Customer Churn Prediction Using Machine Learning Techniques," *Proceedings of International Conference on Data Science and Applications,* 2022.

[14] D. Gholamiangonabadi, S. Nakhodchi, A. Jalalimanesh and A. Shahi, "Customer Churn Prediction Using a Meta-Classifier Approach," *A Case Study of Iranian Banking Industry.,* 2019.

[15] A. K. Ahmad, J. Assef and A. Kadan, "Customer Churn Prediction in Telecom Using Machine Learning in Big Data Platform," *Journal of Big Data,* vol. 6, no. 1, pp. 1-24, 2019.

[16] "Kaggle," [Online]. Available: https://www.kaggle.com/datasets/shubh0799/churn-modelling.