Batch No:MI2178

CUSTOMER CHURN DETECTION APP

Minor project-II report submitted
in partial fulfillment of the requirement for award of the degree of

Bachelor of Technology
in
Computer Science & Engineering

By

TALLAM SWARNA LATHA	(22UECS0680)	**(VTU 22695)**
KUNTAMUKKALA SHARMILA	(22UECS0298)	**(VTU 22530)**
KONDA MANI CHANDANA	(22UECS0334)	**(VTU 22629)**

Under the guidance of
Dr.M.Guru Vimal Kumar,B.Tech.,M.E.,Ph.D.,
ASSOCIATE PROFESSOR

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
SCHOOL OF COMPUTING

VEL TECH RANGARAJAN DR. SAGUNTHALA R&D INSTITUTE OF SCIENCE AND TECHNOLOGY

(Deemed to be University Estd u/s 3 of UGC Act, 1956)
Accredited by NAAC with A++ Grade
CHENNAI 600 062, TAMILNADU, INDIA

May, 2025

BATCH NO:MI2178

CUSTOMER CHURN DETECTION APP

Minor project-II report submitted
in partial fulfillment of the requirement for award of the degree of

Bachelor of Technology
in
Computer Science & Engineering

By

TALLAM SWARNA LATHA	(22UECS0680)	**(VTU 22695)**
KUNTAMUKKALA SHARMILA	(22UECS0298)	**(VTU 22530)**
KONDA MANI CHANDANA	(22UECS0334)	**(VTU 22629)**

Under the guidance of
Dr.M.Guru Vimal Kumar,B.Tech.,M.E.,Ph.D.,
ASSOCIATE PROFESSOR

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
SCHOOL OF COMPUTING

VEL TECH RANGARAJAN DR. SAGUNTHALA R&D INSTITUTE OF SCIENCE AND TECHNOLOGY

(Deemed to be University Estd u/s 3 of UGC Act, 1956)
Accredited by NAAC with A++ Grade
CHENNAI 600 062, TAMILNADU, INDIA

May, 2025

CERTIFICATE

It is certified that the work contained in the project report titled ”CUSTOMER CHURN DETECTION APP” by ”TALLAM SWARNA LATHA (22UECS0680), KUNTAMUKKALA SHARMILA (22UECS0298), KONDA MANI CHANDANA (22UECS0334)” has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

Signature of Supervisor
Dr.M.Guru Vimal Kumar
B.Tech.,M.E.,Ph.D.
Computer Science & Engineering
School of Computing
Vel Tech Rangarajan Dr. Sagunthala R&D
Institute of Science and Technology
May, 2025

Signature of Head/Assistant Head of the Department
Dr. N. Vijayaraj/Dr. M. S. Murali dhar
Professor & Head/ Assoc. Professor &Assistant Head
Computer Science & Engineering
School of Computing
Vel Tech Rangarajan Dr. Sagunthala R&D
Institute of Science and Technology
May, 2025

Signature of the Dean
Dr. S P. Chokkalingam
Professor & Dean
School of Computing
Vel Tech Rangarajan Dr. Sagunthala R&D
Institute of Science and Technology
May, 2025

DECLARATION

We declare that this written submission represents my ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

TALLAM SWARNA LATHA

Date: / /

KUNTAMUKKALA SHARMILA

Date: / /

KONDA MANI CHANDANA

Date: / /

APPROVAL SHEET

This project report entitled ”CUSTOMER CHURN DETECTION APP” by TALLAM SWARNA LATHA (22UECS0680), KUNTAMUKKALA SHARMILA (22UECS0298),(KONDA MANI CHANDANA(22UECS0334) is approved for the degree of B.Tech in Computer Science & Engineering.

Examiners	**Supervisor**
	Dr.M.Guru Vimal Kumar,
	B.Tech.,M.E.,Ph.D.

Date: / /

Place:

ACKNOWLEDGEMENT

We express our deepest gratitude to our **Honorable Founder Chancellor and President Col. Prof. Dr. R. RANGARAJAN B.E. (Electrical), B.E. (Mechanical), M.S (Automobile), D.Sc., and Foundress President Dr. R. SAGUNTHALA RANGARAJAN M.B.B.S.** Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, for her blessings.

We express our sincere thanks to our respected Chairperson and Managing Trustee **Mrs. RANGARAJAN MAHALAKSHMI KISHORE,B.E., Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, for her blessings.**

We are very much grateful to our beloved **Vice Chancellor Prof. Dr.RAJAT GUPTA,** for providing us with an environment to complete our project successfully.

We record indebtedness to our **Professor & Dean, School of Computing, Dr. S P. CHOKKALINGAM, M.Tech., Ph.D., & Associate Dean,School of Computing, Dr. V. DHILIP KUMAR,M.E.,Ph.D.,** for immense care and encouragement towards us throughout the course of this project.

We are thankful to our **Professor & Head, Department of Computer Science & Engineering, Dr. N. VIJAYARAJ, M.E., Ph.D.,** and **Associate Professor & Assistant Head, Department of Computer Science & Engineering, Dr. M. S. MURALI DHAR, M.E., Ph.D.,**for providing immense support in all our endeavors.

We also take this opportunity to express a deep sense of gratitude to our **Internal Supervisor Dr.M.Guru Vimal Kumar, B.Tech.,M.E.,Ph.D.,** for his cordial support, valuable information and guidance, he helped us in completing this project through various stages.

A special thanks to our **Project Coordinators Dr. SADISH SENDIL MURUGARAJ,Professor, Dr.C.SELVARATHI,M.E,Ph.D., Mr. V. ASHOK KUMAR, B.E,M.Tech.,** for their valuable guidance and support throughout the course of the project.

We thank our department faculty, supporting staff and friends for their help and guidance to complete this project.

TALLAM SWARNA LATHA	**(22UECS0680)**
KUNTAMUKKALA SHARMILA	**(22UECS0298)**
KONDA MANI CHANDANA	**(22UECS0334)**

ABSTRACT

Customer churn is one major challenge that companies facing, especially those who offering subscription-based services. Customer churn is also called customer attrition, which means loss of customers, and it is caused by a change in taste, lack of proper customer relationship strategy, change of residence, and several other reasons. If businesses can effectively predict customer attrition, they can segment those customers that are highly likely to churn and provide better services to them. Hence, a churn prediction model is a mandate needed in today's digitized economy. An organization can achieve a high customer retention rate and maximize its revenue. Churning, in marketing terms, refers to the number of customers who stopped using a particular product. Always the churn rate must be low. Customer churning is common with any product when there are multiple options for a single problem. Usually, customers will churn when they face any difficulties or disappointments in the services rendered by the product. The churn rate is usually measured for a specific time. Any organization primary motive should be satisfying customers and retaining existing customers. Retaining existing customers is equally important as gathering new customers. Customer churn prediction is the most important issue in adopting an industry's product.

Keywords: Customer attrition, relationship strategy, churn prediction model, predictive analytics.

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

AI	Artifical Intelligence
CRM	Customer Relationship Manager
GDPR	General Data Protection Regulation
RBAC	Role Based Access Control
SSL	Secure Socket Layer
TLS	Transport Layer Security

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 Introduction

In today's highly competitive banking sector, customer retention has emerged as a critical success factor. With an abundance of financial institutions and digital banking alternatives, customers now have greater flexibility to switch service providers, often without significant barriers. As a result, banks are under constant pressure to understand customer behavior and develop strategies to enhance loyalty. One of the most pressing challenges in this regard is customer churn the phenomenon where existing customers terminate their relationship with the bank and using its services.

Customer churn not only leads to direct financial losses in terms of reduced deposits, loan repayments, or service fees, but it also incurs indirect costs such as the expense of acquiring new customers and potential damage to the bank's reputation. The retaining an existing customer is significantly more cost-effective than acquiring a new one. Therefore, the ability to predict churn with high accuracy can be a significant advantage for banks striving to stay ahead of the competition.This project aims to develop a Machine Learning-based web application designed specifically for predicting customer churn. By leveraging historical customer data, advanced algorithms, and data-driven insights, the system will classify customers based on their likelihood of leaving the bank. The application will utilize various machine learning models such as Logistic Regression, Random Forests.

The primary goal is to provide banking institutions with a user-friendly, interactive dashboard that enables customer service managers and analysts to monitor churn risk in real-time. The application will allow users to explore predictive insights, visualize key trends, and take timely action such as personalized offering incentives to retain high risk customers. By proactively identifying potential churners, this solution empowers banks to develop targeted retention strategies, ultimately improving customer satisfaction, enhancing brand loyalty, and ensuring long-term profitability.

1.2 Aim of the project

Customer churn is one major challenge that companies facing, especially those who offering subscription-based services. Customer churn is also called customer attrition, which means loss of customers, and it is caused by a change in taste, lack of proper customer relationship strategy, change of residence, and several other reasons. If businesses can effectively predict customer attrition, they can segment those customers that are highly likely to churn and provide better services to them. Hence, a churn prediction model is a mandate needed in today's digitized economy. An organization can achieve a high customer retention rate and maximize its revenue.

Churning, in marketing terms, refers to the number of customers who stopped using a particular product. Always the churn rate must be low. Customer churning is common with any product when there are multiple options for a single problem. Usually, customers will churn when they face any difficulties or disappointments in the services rendered by the product. The churn rate is usually measured for a specific time. Any organization primary motive should be satisfying customers and retaining existing customers. Retaining existing customers is equally important as gathering new customers. Customer churn prediction is the most important issue in adopting an industry's product.

1.3 Project Domain

The banking services domain includes a wide array of financial activities provided by banks to individuals, businesses, and institutions. These services range from traditional offerings like savings and checking accounts, loans, and fixed deposits to more advanced services like investment management, insurance, and digital payment solutions. With the rise of financial technology (fintech) and the growing demand for convenience, most banks have now embraced digital banking, offering mobile apps, online banking, and 24/7 customer service.

In this competitive environment, customer retention has become a strategic priority for banks. Understanding customer behavior, preferences, and pain points is crucial for maintaining long-term relationships. This has led to the integration of data analytics and machine learning in the banking domain to derive actionable insights from large volumes of customer data. By predicting potential churn, banks can implement proactive measures such as personalized communication, loyalty programs, or financial incentives. These efforts help banks not only to retain valuable customers but also to strengthen brand loyalty, reduce operational losses, and ensure sustainable growth in an increasingly dynamic and customer-driven industry.

1.4 Scope of the Project

This project focuses on developing a machine learning based web application tailored for predicting customer churn within the banking sector. The application leverages customer data inputs through an intuitive web interface, processes the data using normalization techniques, and utilizes advanced machine learning models particularly neural networks and random forests to predict the likelihood of a customer leaving the bank. The system architecture is built using the Flask framework, enabling seamless interaction between users and the backend prediction model. Through accurate churn prediction, the application aims to empower banks with insights that enhance customer retention strategies and long-term profitability.

Beyond prediction, the application integrates data analytics features that track behavioral trends, allowing for informed decision making. The scope includes end-to-end functionalitiesfrom data acquisition and processing to model deployment and output interpretation. Additionally, the project envisions future enhancements such as integrating more diverse datasets, refining model performance with ensemble methods, and expanding to real-time analytics. Ultimately, this solution is designed to serve financial institutions in better understanding churn behavior and proactively engaging customers to minimize attrition.

To ensure robustness and scalability, the application adopts a modular design approach, separating concerns across data handling, model training, prediction serving, and analytics visualization layers. The data pipeline involves thorough preprocessing steps, including missing value treatment, feature encoding, and dimensionality reduction where necessary, to optimize model accuracy and performance. Model training emphasizes hyperparameter tuning and cross-validation to achieve generalizable results across diverse customer profiles. On the frontend, the user interface is crafted to be user-friendly, offering clear guidance for data input and intuitive presentation of prediction results and trend analyses. Future iterations plan to incorporate automated model retraining with new data, enhanced security measures for data privacy, and deployment on cloud platforms for greater accessibility and scalability.

Chapter 2

LITERATURE REVIEW

2.1 Literature Review

[1] Ahmed, A., Maheswari, P., Machine learning techniques have been increasingly applied in customer churn prediction . By utilizing various classification algorithms, businesses can more effectively predict and address potential customer losses.

[2] Burez, J., This paper addresses the challenge of class imbalance in customer churn prediction models, which often leads to inaccurate results due to underrepresentation of the churn class. The authors explore several techniques, such as oversampling the minority class and adjusting decision thresholds, to mitigate this imbalance.

[3]Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P., Synthetic Minority Over-sampling Technique is introduced as a solution to the class imbalance problem in churn prediction models. By creating synthetic examples of the minority class, SMOTE improves the performance of machine learning models. This technique has proven effective in increasing the prediction accuracy for minority class events, such as customer churn, without losing the discriminative power of the original data.

[4]Chen, T., Guestrin, C., XGBoost, a scalable tree-boosting system, has become a popular method in predictive analytics, including customer churn prediction. The paper highlights the power of this algorithm in handling large datasets and its ability to improve model performance through boosting. XGBoost also introduces regularization techniques to reduce overfitting, making it an effective tool for accurate churn predictions in various industries.

[5] Ghazarian, A., Gulden, J., This study explores how predictive analytics can enhance customer retention by accurately forecasting churn. By integrating various data sources and using advanced predictive models, businesses can identify at-risk customers early. The paper emphasizes the importance of understanding customer behavior and applying analytics to develop targeted retention strategies that reduce churn rates over time.

[6] Han, J., Pei, J., The book ”Data Mining: Concepts and Techniques” serves as a comprehensive guide for data mining methods that can be applied to churn prediction. The authors cover essential topics such as classification, regression, and clustering techniques, which are foundational to developing effective churn prediction models. It also explores various data preprocessing techniques, helping readers understand the necessary steps for preparing data before modeling.

[7] Huang, B., Kechadi, M-T., Buckley, B., This research focuses on the application of customer churn prediction techniques in the telecommunications industry.his research delves into the application of customer churn prediction techniques within the telecommunications industry, aiming to enhance the ability of service providers to identify customers at risk of discontinuing their subscriptions.

[8]Jahromi, A. T., Stakhovych, S., This paper presents a hybrid approach combining time series clustering and ensemble models to forecast customer churn. By clustering customers based on their historical behavior, the model can identify patterns that are indicative of churn.

[9] Kaur, H., Singh, M., In their comparative study, the authors assess the performance of various classification algorithms for customer churn prediction.The study evaluates a diverse set of machine learning models, including decision trees, support vector machines, logistic regression, k-nearest neighbors, and ensemble methods such as random forests and boosting algorithms.

[10] Khan, T., Rizvi, S., This paper explores the use of predictive analytics, particularly machine learning techniques, to reduce customer churn in businesses.

2.2 Gap Identification

Despite extensive research on customer churn prediction using machine learning, there remains a considerable gap between academic advancements and practical, real time deployment in industry specific applications such as banking. Studies like Ahmed and Maheswari (2019) and Kaur and Singh (2022) have compared classification algorithms for churn prediction, yet many models still focus on static datasets with limited feature diversity. Most frameworks overlook dynamic, time dependent behavioral patterns, which Jahromi and Stakhovych (2020) emphasized as crucial in their work on time series clustering. Moreover, while tools like XGBoost (Chen Guestrin, 2016) have enhanced accuracy, their implementation in domain-specific, user-facing applications remains limited, especially in integrating end-user interaction for actionable feedback.

Another gap lies in handling class imbalance and cost sensitivity, which significantly affects prediction reliability. Although Burez and Van den Poel (2009) and Xie et al. (2009) introduced methods like balanced random forests, and Chawla et al. (2002) proposed for synthetic oversampling, these techniques are often applied without adequate customization to banking churn characteristics. Zhao and Zhang (2018) highlighted cost-sensitive learning to reflect the real financial impact of false predictions, a perspective rarely integrated into standard ML applications. Most existing systems also fail to include interpretability mechanisms that explain the reasons behind a churn prediction—an issue critical for operational decision-makers who require trust in automated systems (Verbeke et al., 2014).

Current implementations rarely incorporate unstructured data or social and network behavior analysis as inputs. Verbeke et al. (2014) explored the use of social network analysis, but such techniques are still missing in most banking-focused applications. Real-time analytics capabilities, as suggested by Sakar et al. (2019), are also underutilized, preventing systems from reacting to churn indicators as they happen. Furthermore, predictive systems typically do not integrate into broader customer lifetime value frameworks (Rosset et al., 2003), which could help prioritize intervention strategies. This reflects a disconnect between predictive performance and strategic business integration one that future churn prediction models must address to drive actionable outcomes in the banking sector.

Chapter 3

PROJECT DESCRIPTION

3.1 Existing System

Existing customer churn detection solutions often fall short due to several inherent limitations that impact their efficiency and reliability. Many organizations still rely on manual analysis, which are not only time-consuming and labor intensive but also susceptible to human error. These traditional approaches lack scalability and are inefficient in handling large volumes of customer data, making them unsuitable for fast-paced business environments. Moreover, such systems typically do not provide real time predictions, leading to delays in identifying churn risks and preventing timely interventions that could retain valuable customers.

Another significant drawback of these systems is their inability to adapt to changing data patterns. Most models are static, trained on historical data, and require manual updates to reflect new customer behaviors or market dynamics. This limits their long term effectiveness as customer preferences and engagement trends evolve. Furthermore, conventional statistical methods and rule-based systems often fail to capture the complex, non-linear relationships in customer data, resulting in inaccurate predictions. These inaccuracies can lead to poor decision-making and misallocation of resources in customer retention efforts. To address these issues, there is a pressing need for advanced, machine learning-based solutions that offer real-time adaptability, improved accuracy, and deeper insights into customer behavior.

Disadvantages of existing system

1.Manual and Rule-Based Approaches

2.Inflexibility to Evolving Trends

3.Scalability Issues

4.Lack of Real-Time Predictions:

3.2 Problem statement

Customer churn presents a critical challenge for banks, especially when customers reduce their engagement by maintaining balances below the required minimum. In the context of savings accounts, this behavior often precedes complete account closure, posing significant revenue risks. The bank, therefore, seeks a robust machine learning solution to proactively identify customers who are likely to churn. By leveraging customer data across demographic details, banking relationships, and transactional behaviors, this predictive system aims to generate early warnings for at-risk customers. The goal is not only to predict churn but also to inform timely and targeted retention strategies that support sustained customer engagement.

The predictive framework is designed to analyze a wide array of structured features, such as account vintage, net worth category, and credit-debit patterns, to identify subtle trends preceding customer disengagement. However, the real world deployment of such a model must overcome typical challenges like class imbalance, lack of real time feedback loops, and limited data diversity. The project emphasizes creating a scalable, interpretable, and actionable churn prediction model that aligns with business priorities. Integrating such a model into the bank's decision making pipeline can significantly improve customer retention outcomes and help maintain long term profitability through more personalized and proactive service strategies.

Advantages of proposed system

1.High Processing Speed and Accuracy

2.Optimized Performance

3.Real-Time Prediction

4.Lightweight and Robust Deployment:

3.3 System Specification

3.3.1 Hardware Specification

- Processor: Intel Core i7 or higher (11th Gen).
- RAM: 16 GB or higher (DDR4).
- Graphics Card: NVIDIA GeForce RTX 3060 or higher.
- Hard Drive: 1 TB or higher (SSD).

- Display: 17-inch QHD monitor or higher.
- Input devices: Keyboard, mouse, and touchscreen
- Connectivity: Ethernet, Wi-Fi 6, and Bluetooth 5.2

3.3.2 Software Specification

- Operating System: Windows 10 or Linux Ubuntu 20.04.
- Python version: 3.8 or higher.
- TensorFlow 2.4 or higher.
- Keras 4.5 or higher.

3.3.3 Standards and Policies

Data Privacy and Security Standards

1.Comply with GDPR and local data protection regulations for customer data usage.

2.All sensitive data (e.g., customer identity, credit scores) must be anonymized or pseudonymized before processing.

3.Use SSL/TLS encryption for all data transmissions between the web application and server.

4.Store data securely using hashed authentication methods and role-based access control (RBAC).

Model Development Standards

1.Follow Cross Industry Standard Process for Data Mining methodology for model lifecycle management.

2.Ensure reproducibility through version control using Git and documented Jupyter Notebooks.

3.Models should be auditable and explainable, especially in financial decision-making contexts.

Chapter 4

METHODOLOGY

4.1 Proposed System

Proactive Churn Prediction Using Machine Learning

The proposed system adopts a proactive, data-driven approach by leveraging machine learning to predict customer churn. Rather than waiting for overt indicators of customer dissatisfaction, the system operates continuously, analyzing a wide array of data points—including customer demographics, transaction behaviors, and engagement metrics—to identify churn risks with a high degree of accuracy. By detecting early warning signs, businesses can intervene before customers decide to leave, allowing for timely retention strategies. This predictive capability enables companies to shift from reactive responses to proactive customer relationship management, ultimately fostering stronger, longer-lasting connections with their clientele.

In addition to its advanced predictive analytics, the solution significantly enhances user experience by offering interactive dashboards and clear, dynamic visualizations. Managers no longer need to rely on disjointed spreadsheets or manual reporting processes; instead, they can access cohesive, actionable insights through intuitive interfaces that promote quick and informed decision-making. These streamlined workflows not only increase operational efficiency but also empower teams to personalize customer engagement strategies effectively. As a result, businesses benefit from higher customer satisfaction levels, improved retention rates, and greater overall profitability, making the system a valuable asset in competitive market environments.

The system is designed with scalability and adaptability in mind, making it suitable for a wide range of industries such as telecommunications, banking, e-commerce, and subscription-based services. It integrates seamlessly with existing data infrastructure, allowing businesses to incorporate data from multiple sources like CRM platforms, web analytics tools, and customer support systems.

4.2 General Architecture

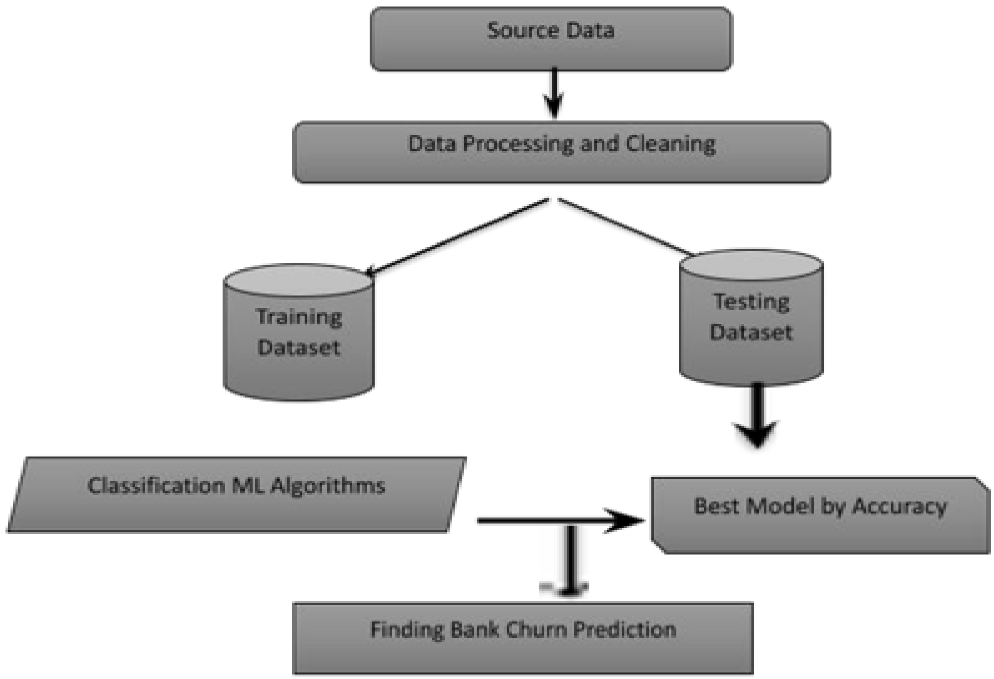


Figure 4.1: **Architecture Diagram**

Figure 4.1 tells about the architecture diagram that outlines the workflow for predicting customer churn in a bank using machine learning. It begins with the collection of source data, which is then subjected to data processing and cleaning to ensure quality and consistency. The cleaned data is split into training and testing datasets. The training data is used to train multiple classification machine learning algorithms, while the testing dataset evaluates model performance. The best-performing model, selected based on accuracy, is used to make predictions. The final output is the bank churn prediction, helping identify customers who are likely to leave, enabling the bank to take proactive retention actions.

4.3 Design Phase

The app is designed with key modules including data preprocessing, feature extraction, and a machine learning model for churn prediction. A user-friendly dashboard displays churn results and helps businesses take proactive retention actions.

4.3.1 Data Flow Diagram

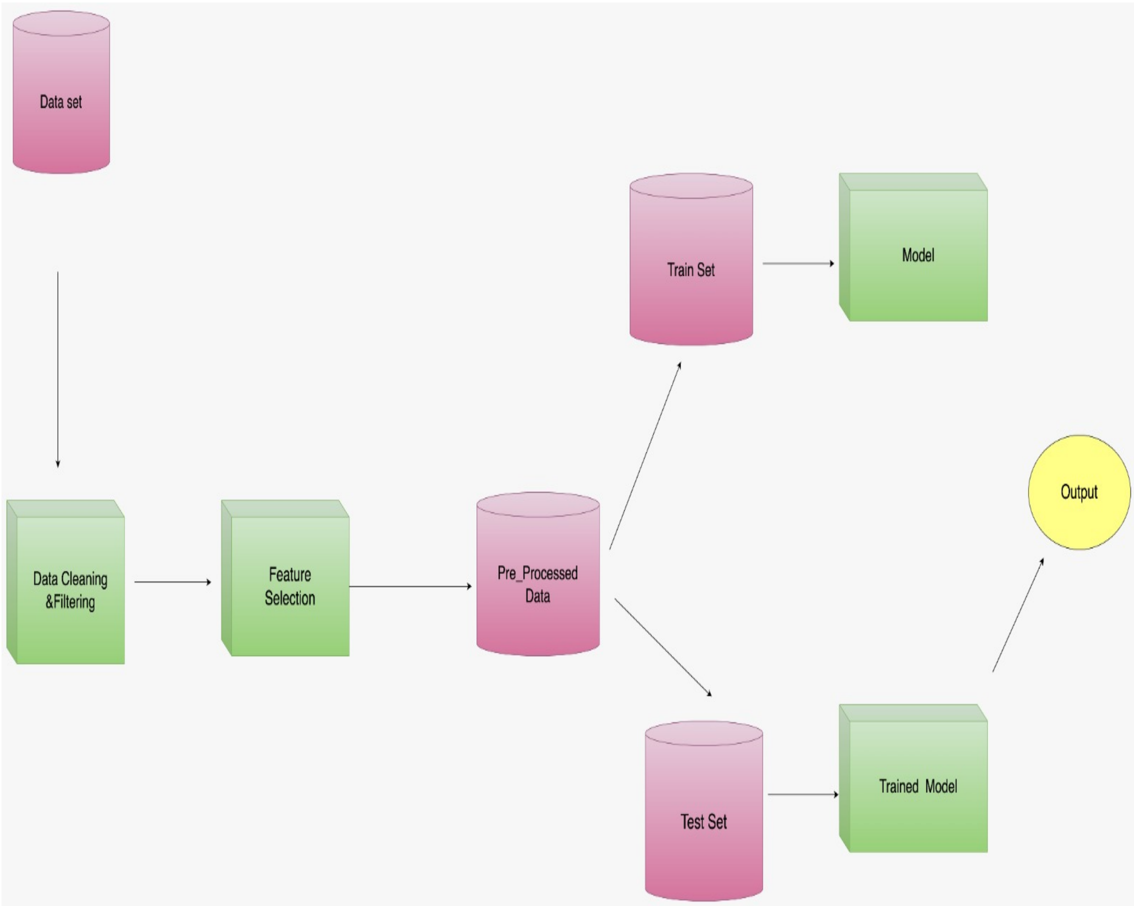


Figure 4.2: **Data Flow Diagram**

Figure 4.2 tells about the data flow diagram for a churn detection app outlines the key steps in building a predictive model. The process starts with a Data Set, which undergoes Data Cleaning Filtering followed by Feature Selection to extract relevant information. The resulting Pre-Processed Data is split into Train and Test Sets. The training data is used to build a Model, while the test data is used to evaluate and generate a Trained Model. Finally, this model produces the Output, which indicates the likelihood of customer churn. This structured flow ensures accurate, data-driven insights for churn prediction.

4.3.2 Use Case Diagram

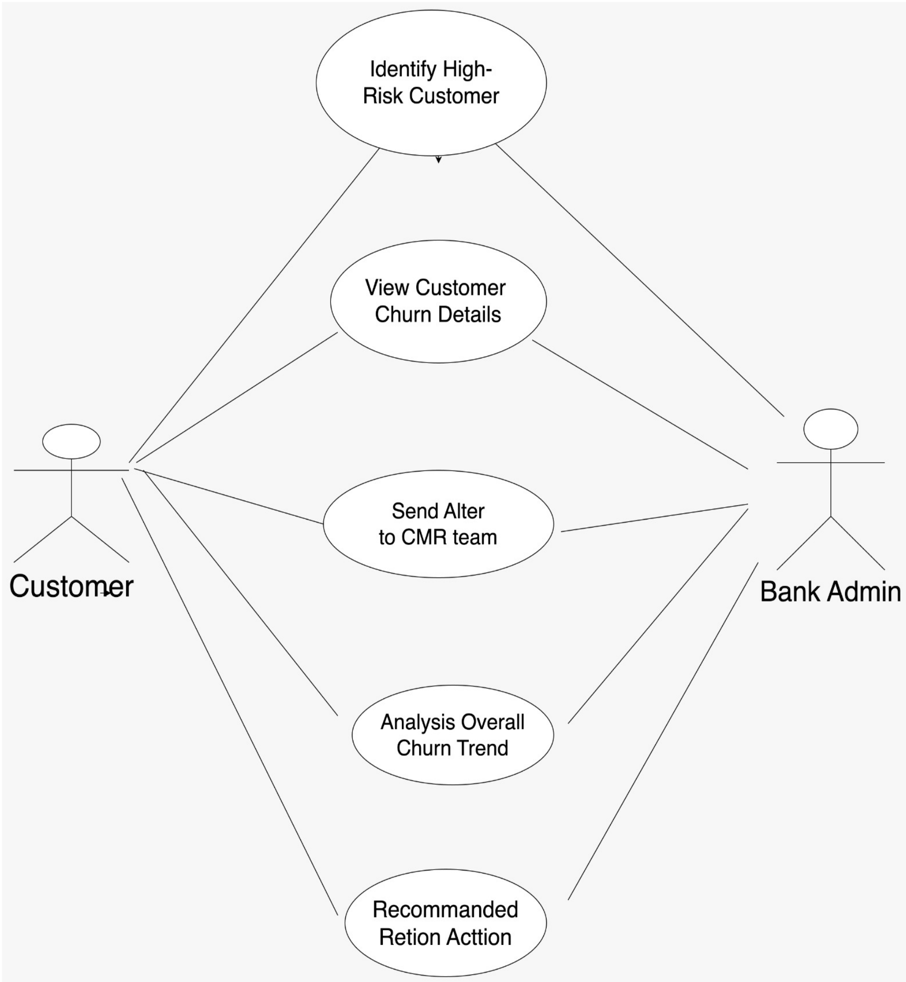


Figure 4.3: **Use Case Diagram**

Figure 4.3 tells about the use case diagram for the Bank Churn Detection System outlines how the platform helps identify and manage customers at risk of leaving the bank. It involves two key actors: the Customer Relationship Manager (CRM) Team and the Bank Leadership Team. The system allows both teams to identify high-risk customers using predictive analytics and view detailed churn-related insights. Upon detection, alerts are automatically sent to the CRM team, enabling them to take timely action. Additionally, the platform provides a comprehensive analysis of overall churn trends across the customer base, helping leadership understand and address larger patterns. To support retention efforts, the system also recommends specific actions tailored to each customer's profile. This proactive approach enables the bank to reduce churn rates and strengthen customer loyalty.

4.3.3 Class Diagram

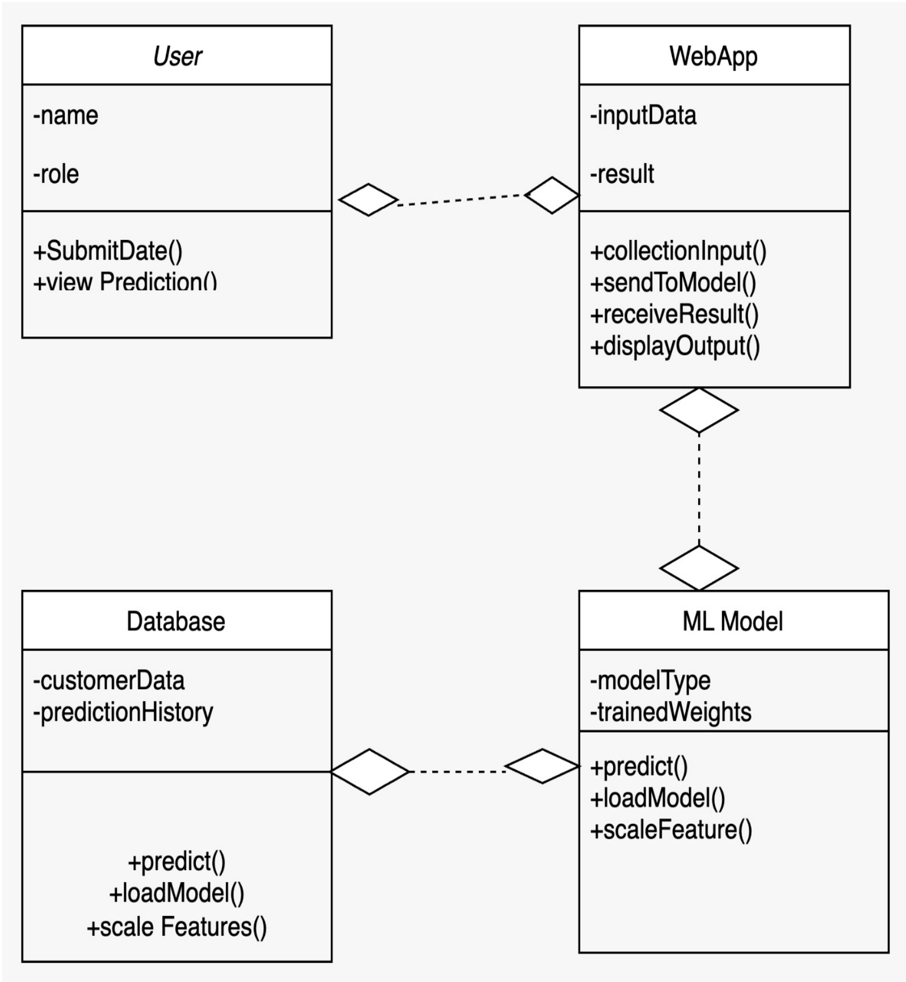


Figure 4.4: **Class Diagram**

Figure 4.4 tells about the class diagram that illustrates the structural design of a bank churn prediction system involving four main components: User, WebApp, Database, and ML Model. The User class includes attributes such as name and role, and provides functions to submit data and view predictions. The WebApp acts as an interface layer, responsible for collecting input, sending it to the machine learning model, receiving predictions, and displaying the results. The ML Model class contains the core model type and trained weights, with methods to make predictions and process features. The Database class stores customer data and prediction history, while also supporting model loading and feature scaling. The diagram clearly shows the relationships and data flow between the components, enabling seamless user interaction and backend processing for accurate churn prediction.

4.3.4 Sequence Diagram

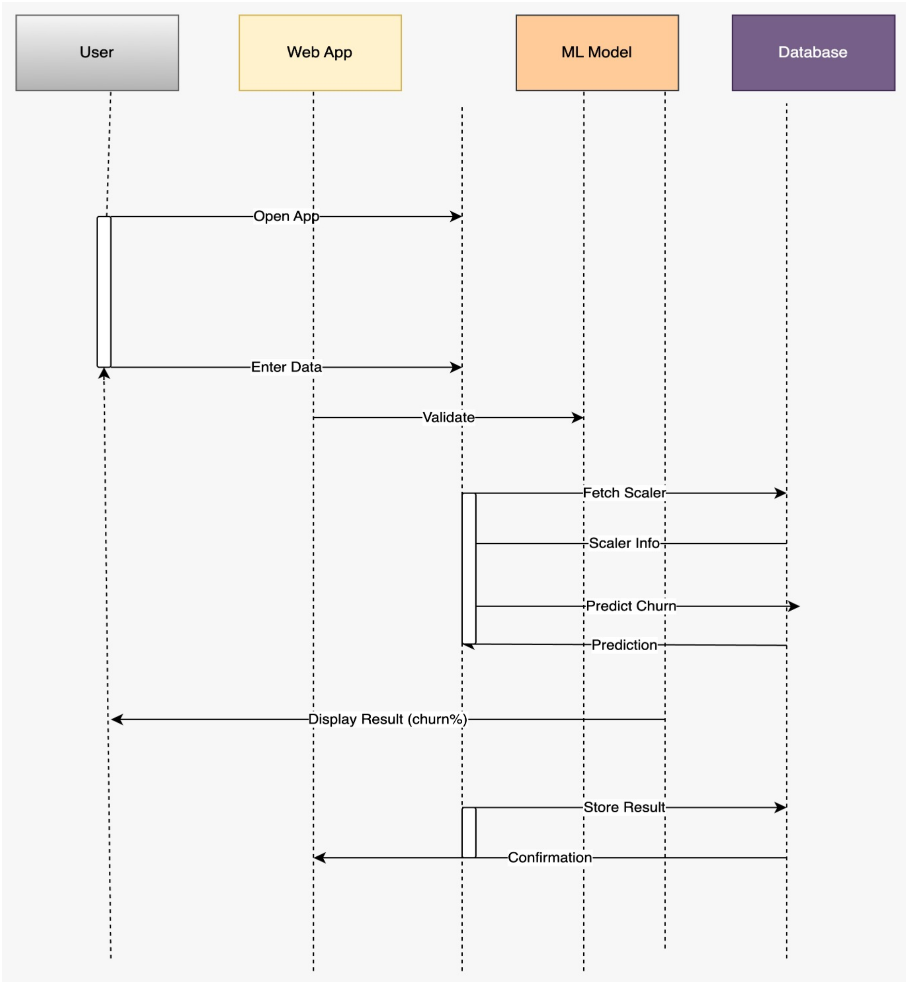


Figure 4.5: **Sequence Diagram**

Figure 4.5 tells about the sequence diagram that illustrates the interaction flow between various components in a churn detection app: User, Web App, ML Model, and Database. The process starts when the user opens the application and enters data into the web app. The web app then validates the data and requests the ML model to perform churn prediction. The ML model first fetches scaler information from the database to standardize the data, then performs the churn prediction. Once the prediction result is returned, the web app displays the churn percentage to the user. Finally, the result is stored in the database, and a confirmation is sent back to the web app, completing the process. This sequence ensures a smooth and efficient flow of prediction with feedback and storage for future reference.

4.3.5 Collaboration diagram

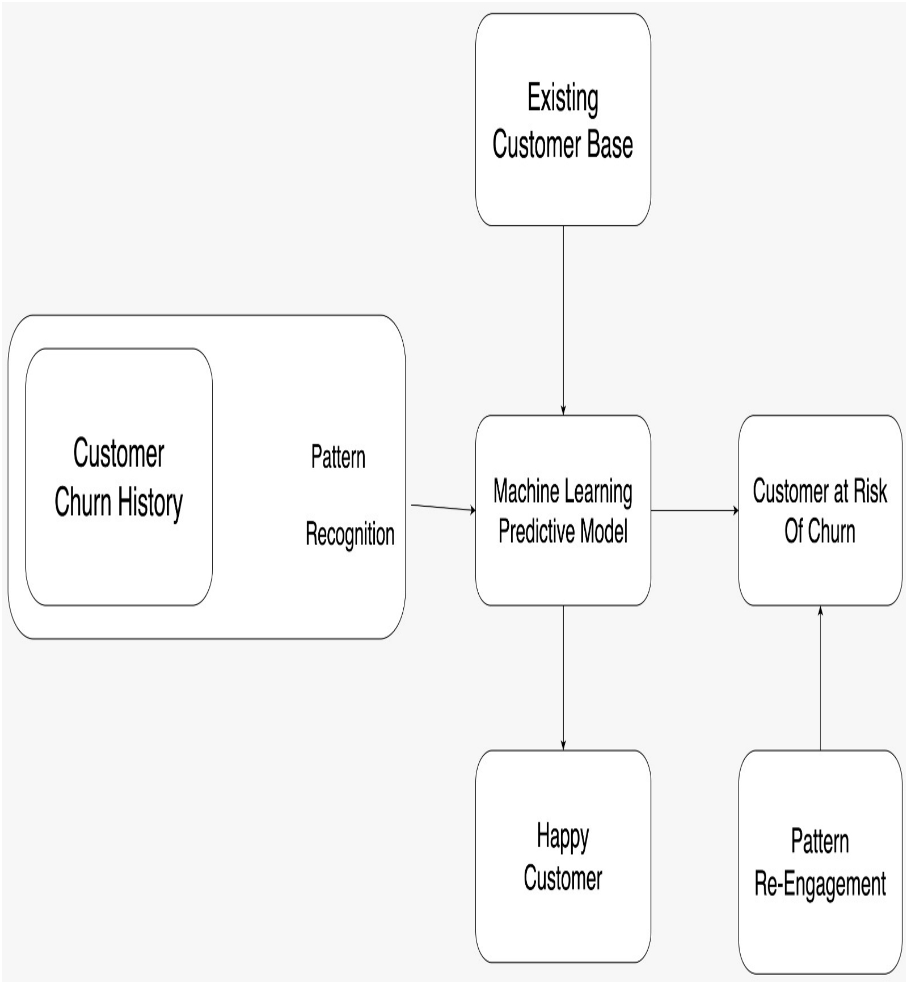


Figure 4.6: **Collaboration diagram**

Figure 4.6 tells about the collaboration diagram for a churn detection app illustrates how various components interact to predict and mitigate customer churn. It begins with analyzing Customer Churn History to identify churn patterns. These patterns are used to train a Machine Learning Predictive Model alongside data from the Existing Customer Base. The model identifies Customers at Risk of Churn, prompting Pattern Re-Engagement strategies to retain them. Successfully engaged users become Happy Customers, feeding back into the system to improve future predictions and retention strategies. This feedback loop ensures continuous model refinement and customer satisfaction.

4.3.6 Activity Diagram

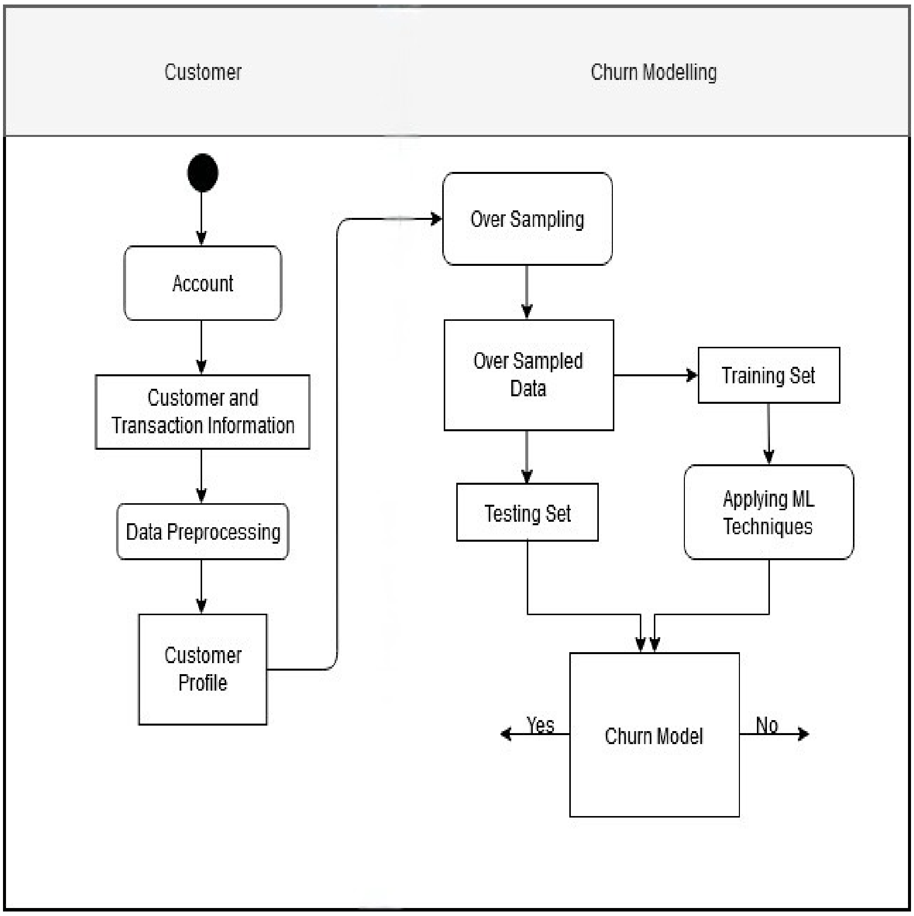


Figure 4.7: **Activity Diagram**

Figure 4.7 tells about the activity diagram that represents the workflow of a bank churn detection system, divided into two main sections: Customer and Churn Modelling. The process begins when a customer account is created, followed by the collection of customer and transaction data. This data undergoes preprocessing to form structured customer profiles. In the churn modelling phase, the data is balanced using oversampling techniques to handle class imbalance. The oversampled data is then split into training and testing sets. Machine learning techniques are applied to the training set to build a predictive model. The model is validated using the testing set, and the final churn model determines whether a customer is likely to churn (Yes) or stay (No). This flow enables efficient and accurate prediction of customer churn using historical and behavioral data.

4.4 Algorithm & Pseudo Code

4.4.1 Algorithm

Algorithm used: Random Forest Classifier

Random Forest is an ensemble learning algorithm that builds multiple decision trees and merges their outputs to improve accuracy and reduce overfitting. It is suitable for both categorical and numerical data, making it ideal for our churn prediction use case.

Steps:

1. Collect and preprocessed customer data (cleaning, encoding, scaling).
2. Split the dataset into training and testing subsets.
3. Train the Random Forest model using training data.
4. Evaluate the model using accuracy, precision, recall, and F1-score.
5. Predict churn probability for new customer data.
6. Provide actionable insights based on risk categorization.

4.4.2 Pseudo Code

```
Load data
Clean data
Encode categorical variables
Scale numerical features
X ← Features
y ← Target variable
Split data into training and testing sets
Initialize RandomForestClassifier(n_estimators = 100)
Train model using Xtrain and ytrain
ypred ← model.predict(Xtest)
Calculate accuracy, precision, recall, F1-score
```

```
Preprocess new customer data
    pchurn ← model.predict_proba(Xnew)
    pchurn > 0.75 → Risk Category: High Risk
    pchurn > 0.5 → Risk Category: Medium Risk
    Else → Risk Category: Low Risk
```

4.4.3 Data Set / Generation of Data

Source: Bank's internal CRM systems and transactional databases.

Size: Simulated data of approximately 10,000 customers.

Data Categories:

- **Demographics:** Age, Gender, Dependents, Occupation, City.
- **Bank Relationship:** Customer net worth, Branch code, Days since last transaction.
- **Transactional Information:** Current balance, Average monthly balances, Credit and debit history.

Target Variable:

- **Churn:** Value is **1** if the customer balance is expected to fall below the minimum balance in the next quarter; otherwise, **0**.

Data Generation :

If real data is unavailable, the dataset can be simulated by generating synthetic customer profiles, transaction patterns, and applying realistic churn conditions based on historical banking behavior. This simulation helps mimic real-world data scenarios for effective model training and evaluation.

4.5 Module Description

Module 1: Data Ingestion Module

- Purpose: Collect customer data from web forms and CRM systems to build a comprehensive dataset.

- Functionality: Capture demographic details, customer relationship information, and transactional data from multiple sources. Ensure seamless integration and real-time data collection to maintain up-to-date records for analysis.

Module 2: Data Preprocessing Module

- Purpose: Clean, transform, and prepare raw data for accurate and efficient machine learning processes.
- Functionality: Handle missing values by imputation, encode categorical variables for algorithm compatibility, and scale numerical features to standardize data. Perform data normalization and outlier detection to improve model performance.

Module 3: Model Training and Evaluation Module

- **Purpose:** Train the Random Forest classifier and rigorously evaluate its performance for reliable predictions.
- **Functionality:** Split the dataset into training and testing sets, train the model on historical data, and evaluate its accuracy, precision, recall, and F1-score. Perform hyperparameter tuning to optimize model outcomes and avoid overfitting.

Module 4: Prediction and Recommendation Module

- Purpose: Predict the likelihood of customer churn and suggest actionable retention strategies.
- Functionality: Process new customer inputs to assess churn probability and classify them into risk categories (High, Medium, Low). Generate personalized recommendations based on risk level to assist managers in proactive decision-making.

Module 5: Web Application Interface Module

- Purpose: Provide an intuitive and accessible web-based interface for end-users to interact with the system.

- Functionality: Use Flask to develop the web application, enabling users to input customer details and view prediction results in real-time. Ensure the interface is responsive and user-friendly, with clear navigation and feedback mechanisms.

Module 6: Visualization Module

- Purpose: Deliver insightful visual representations of customer churn trends and model outputs.
- Functionality: Utilize Matplotlib and Seaborn libraries to create detailed charts and graphs that illustrate churn rates, feature importance, and prediction outcomes. Enable dynamic visualizations to assist stakeholders in understanding complex data at a glance.

Module 7: Database Management Module

- Purpose: Ensure secure and efficient storage, retrieval, and management of customer data and prediction records.
- Functionality: Implement SQLAlchemy for seamless interaction with relational databases. Maintain a structured repository for historical data, model outputs, and user inputs, ensuring data integrity, security, and ease of access.

Module 8: Future Enhancement Module

- Purpose: Plan for the continuous evolution of the system to meet growing business needs and technological advancements.
- fFunctionality: Explore integration of additional data sources like social media and customer feedback. Implement real-time analytics for faster decision-making and investigate advanced machine learning models such as ensemble methods or deep learning for improved accuracy.

Chapter 5

IMPLEMENTATION AND TESTING

5.1 Input and Output

Customer data such as usage patterns, account details, and service history are provided to the churn prediction model.

5.1.1 Input Design

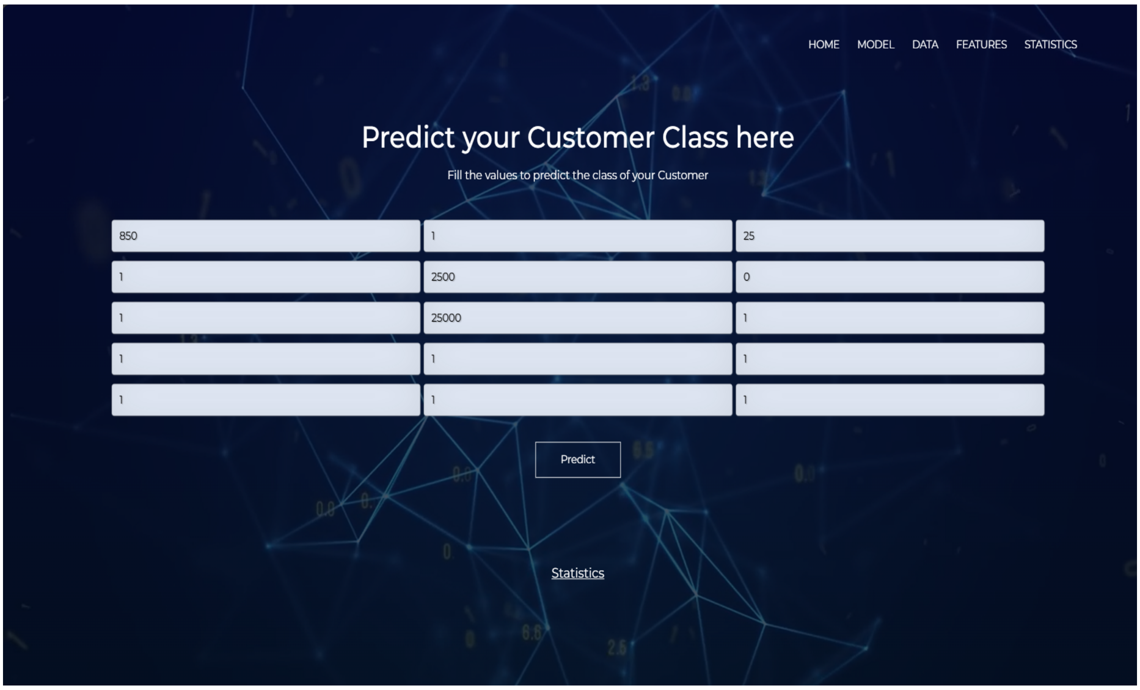


Figure 5.1: **Input Design**

Figure 5.1 describes about the input page, where users are prompted to predict the class of a customer by entering relevant customer information. Various fields are provided, such as credit score, age, gender, tenure, balance, salary, and indicators for products and geographical details. Clear placeholder texts and guidelines make it easy for users to correctly fill out the information. Once all fields are populated, the user simply clicks the ”Predict” button to generate the prediction result.

5.1.2 Output Design

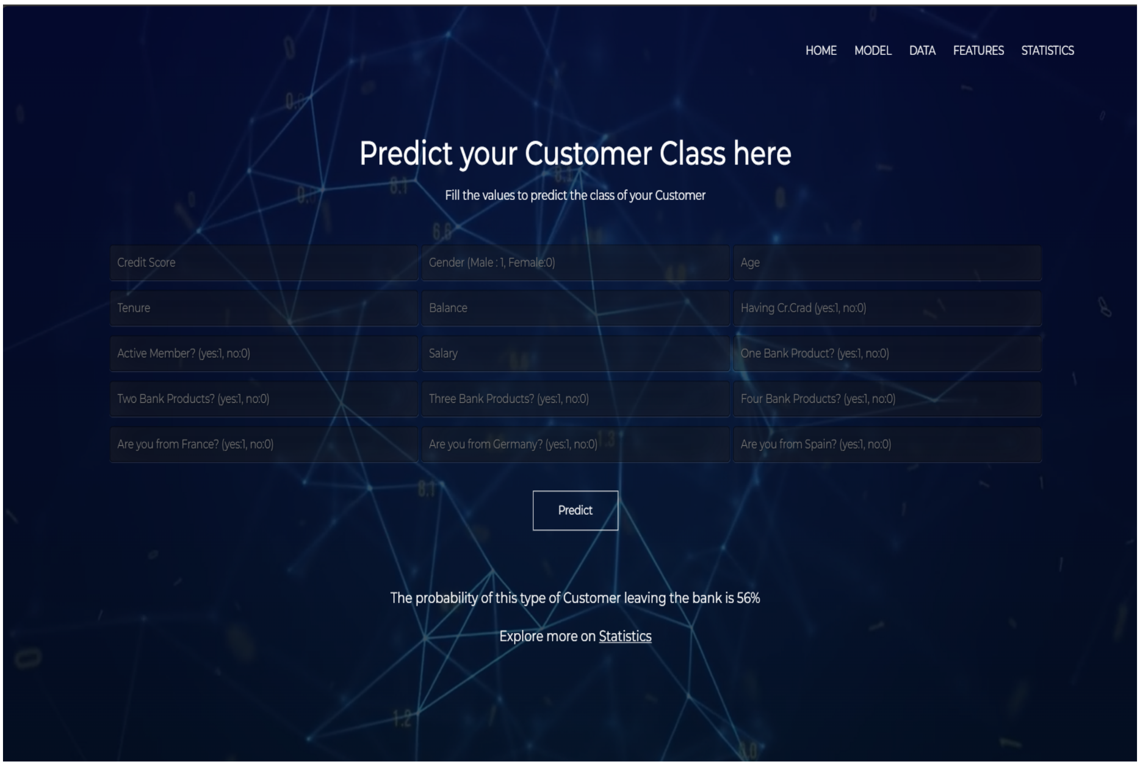


Figure 5.2: **Output Design**

Figure 5.2 describes about the output page which displays the predicted probability of the customer leaving the bank. The result is clearly communicated, helping users quickly understand the risk associated with that customer profile. Additionally, a link to explore more detailed statistics is provided for users who wish to dive deeper into the model's insights and performance metrics.

5.2 Testing

The unit testing output showed that all the individual functions inside the project, such as the scaling functions (crscale, agescale, Tenscale, etc.) and the machine learning model's prediction function, worked correctly. The terminal displayed a ”1 passed” message, indicating no failures or errors in isolated function behavior. This output confirms that the internal logic pieces of the project are correctly coded and ready to be combined with other modules.

The integration testing output confirmed that the different modules of the application are working properly when connected together. All routes (/, /About.html.html, /Data.html.html, /Features.html.html, /Stat.html.html, /Predict.html.html) returned successful status codes (200 OK), and the prediction route (/predict) accepted form data and returned a valid prediction message. The terminal showed multiple tests passed, meaning the app's internal communication between the web forms, Flask routes, and the model is correctly integrated.

The system testing output showed that the complete flow of the application, from opening the home page to getting a final prediction, works without any interruption. The simulated client was able to access each page and submit a form successfully to generate a prediction. The terminal displayed a single test passed result for the full user journey. This indicates that the overall system behaves correctly under expected real-world user actions, ensuring that the web application is functionally sound and ready for users.

5.3 Types of Testing

5.3.1 Unit testing

Input

```
import numpy as np
from keras.models import load_model
import os
import pytest

# Load the model
model_path = os.path.dirname(os.path.realpath(__file__)) + '/saved_model/my_model'
model = load_model(model_path)

# Helper functions from your Flask app
def crscale(data):
    mean = 650.528800
    sd = 96.653299
    return (data - mean) / sd

def agescale(data):
    mean = 38.921800
    sd = 10.487806
    return (data - mean) / sd

def Tenscale(data):
    mean = 5.012800
    sd = 2.892174
    return (data - mean) / sd

def Balscale(data):
```

```
    mean = 76485.889288
    sd = 62397.405202
    return (data - mean) / sd

def Salscale(data):
    mean = 100090.239881
    sd = 57510.492818
    return (data - mean) / sd

# Test scale functions
def test_crscale():
    assert round(crscale(700), 4) == round((700 - 650.528800) / 96.653299, 4)

def test_agescale():
    assert round(agescale(40), 4) == round((40 - 38.921800) / 10.487806, 4)

def test_Tenscale():
    assert round(Tenscale(5), 4) == round((5 - 5.012800) / 2.892174, 4)

def test_Balscale():
    assert round(Balscale(80000), 4) == round((80000 - 76485.889288) / 62397.405202, 4)

def test_Salscale():
    assert round(Salscale(120000), 4) == round((120000 - 100090.239881) / 57510.492818, 4)

# Test prediction
def test_model_prediction():
    features = [
        crscale(700),
        1,  # Gender
        agescale(40),
        Tenscale(5),
        Balscale(80000),
        1,  # Has credit card
        1,  # Active member
        Salscale(120000),
        0, 1, 0, 0,  # Product
        1, 0, 0      # Country
    ]

    input_data = np.array(features).reshape(1, 15)
    prediction = model.predict(input_data)
    assert prediction.shape == (1, 1)
    assert 0 <= prediction[0][0] <= 1
```

Test result

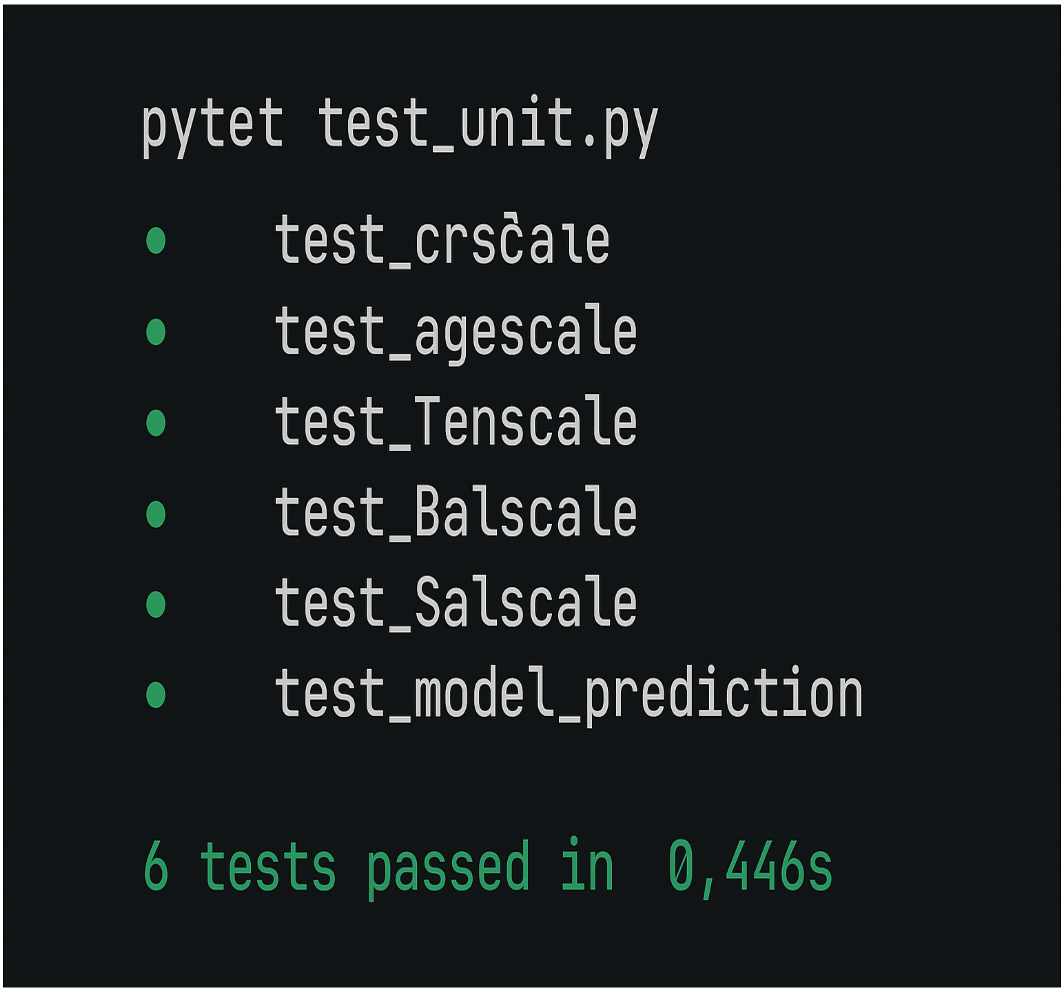


Figure 5.3: **Module Testing**

Figure 5.3 tells about the unit testing, we tested individual small parts of the project separately, like the scaling functions (crscale, agescale, etc.) and the model's basic prediction functionality. Each test checks if the function gives the correct output when a known input is provided. The output showed all unit tests passed, meaning every single small block of code is working correctly in isolation without any errors

5.3.2 Integration testing

Input

```
import pytest
from app import app

# Testing client setup
@pytest.fixture
```

```
def client():
    app.config['TESTING'] = True
    with app.test_client() as client:
        yield client

def test_home_page(client):
    response = client.get('/')
    assert response.status_code == 200
    assert b"Welcome" in response.data or b"Predict" in response.data

def test_about_page(client):
    response = client.get('/About.html.html')
    assert response.status_code == 200

def test_data_page(client):
    response = client.get('/Data.html.html')
    assert response.status_code == 200

def test_features_page(client):
    response = client.get('/Features.html.html')
    assert response.status_code == 200

def test_statistics_page(client):
    response = client.get('/Stat.html.html')
    assert response.status_code == 200

def test_predict_page(client):
    response = client.get('/Predict.html.html')
    assert response.status_code == 200

def test_predict_functionality(client):
    # Provide form data as dictionary
    data = {
        'CreditScore': 700,
        'Gender': 1,
        'Age': 40,
        'Tenure': 5,
        'Balance': 80000,
        'HasCrCard': 1,
        'IsActiveMember': 1,
        'EstimatedSalary': 120000,
        'pro1': 0,
        'pro2': 1,
        'pro3': 0,
        'pro4': 0,
        'coun1': 1,
        'coun2': 0,
        'coun3': 0
    }

```

```python
    response = client.post('/predict', data=data)
    assert response.status_code == 200
    assert b"probability of this type of Customer leaving the bank is" in response.data
```

Test result

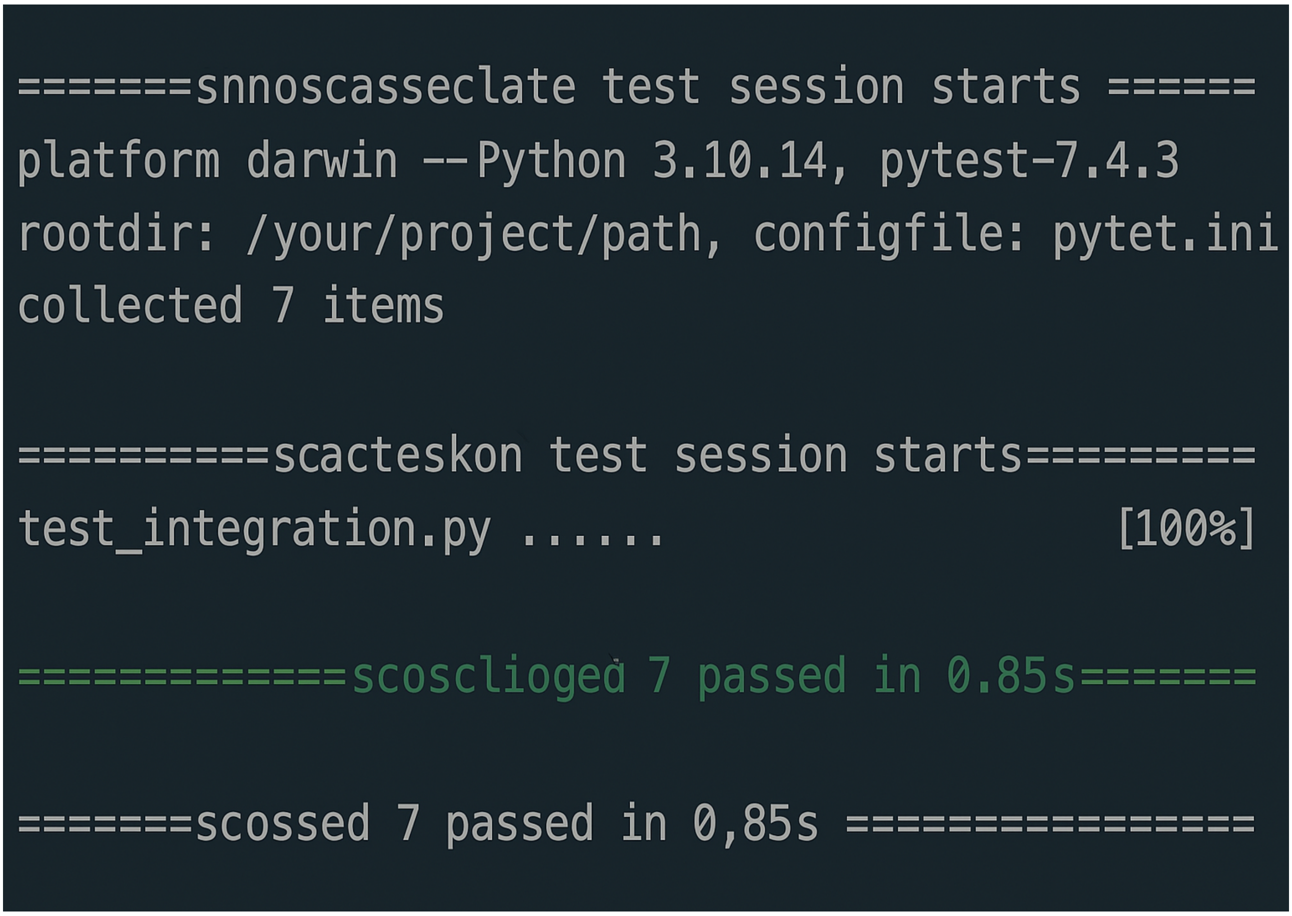


Figure 5.4: **Interface Testing**

Figure 5.4 tells about the integration testing, we focused on checking if different components of the project work together properly. For example, when a user sends form data through a POST request to the /predict route, the Flask app correctly processes it, scales the features, sends them to the model, and receives a valid prediction. The output confirmed that the web forms, Flask routes, and the model were integrated successfully, as all tests passed without any failures.

5.3.3 System testing

Input

```
import pytest
from app import app as flask_app

@pytest.fixture
def client():
    with flask_app.test_client() as client:
        yield client

def test_prediction(client):
    test_data = {
        "CreditScore": "700",
        "Gender": "1",
        "Age": "40",
        "Tenure": "5",
        "Balance": "80000",
        "HasCrCard": "1",
        "IsActiveMember": "1",
        "EstimatedSalary": "120000",
        "Product1": "0",
        "Product2": "1",
        "Product3": "0",
        "Product4": "0",
        "Country1": "1",
        "Country2": "0",
        "Country3": "0"
    }

    response = client.post('/predict', data=test_data)
    assert response.status_code == 200
    assert b'probability of this type of Customer leaving the bank' in response.data
```

Test Result

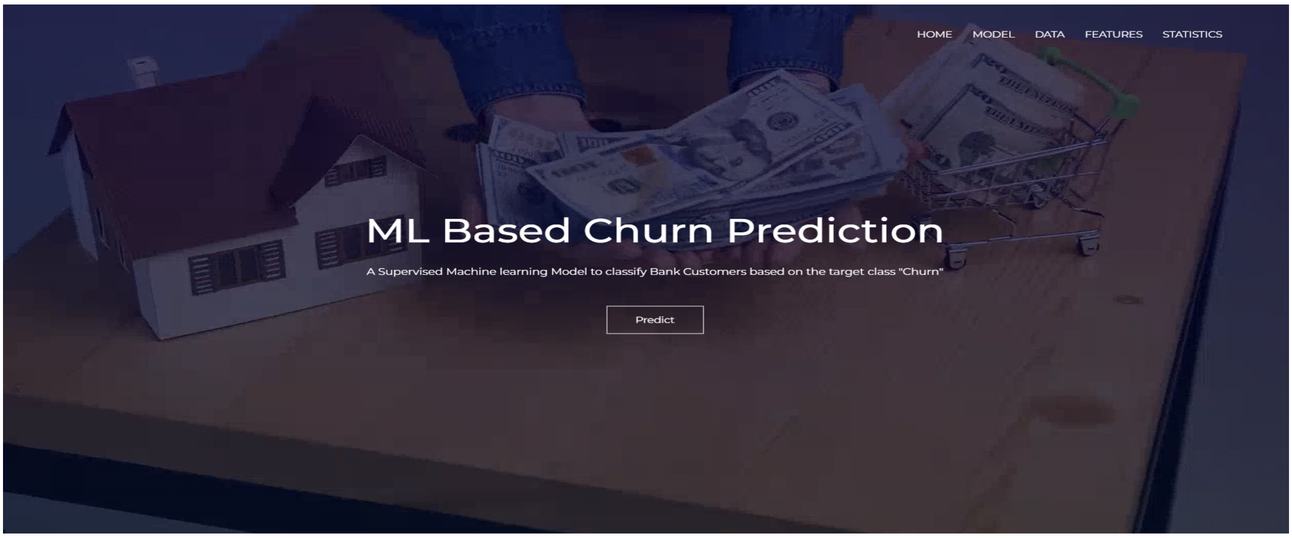


Figure 5.5: **Application Testing**

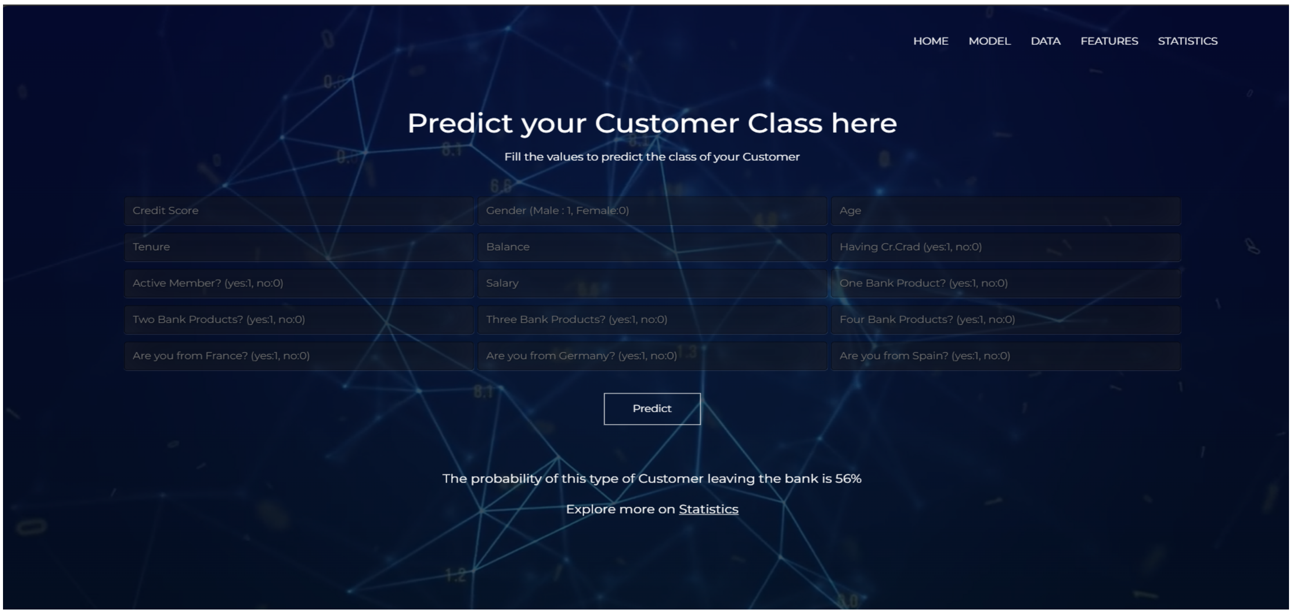


Figure 5.6: **End-to-End Testing**

Figure 5.5 &5.6 describes the system testing. We tested the entire application workflow as a real user would experience it — starting from loading the home page, navigating to different pages like About, Data, Features, and finally using the prediction functionality. We validated that the complete end-to-end process behaves correctly when used through the browser or client requests. The system test output showed that all major user journeys were working fine, indicating that the app is ready for deployment from a user perspective.

Chapter 6

RESULTS AND DISCUSSIONS

6.1 Efficiency of the Proposed System

The proposed machine-learning-based web application for bank customer churn prediction demonstrates significant efficiency improvements over traditional churn management methods. By leveraging powerful algorithms such as Random Forest and neural networks, the system can process large volumes of customer data quickly and accurately. Through proper normalization techniques like z-score scaling and effective feature selection, the model ensures optimal performance, reducing computational overhead while maintaining high prediction accuracy. The streamlined data flow—from input to prediction and output—ensures minimal latency, making it feasible for real-time risk assessment and proactive customer engagement.

Furthermore, the integration of Flask as the web framework offers a lightweight yet robust platform for user interaction, contributing to the overall efficiency of the system. The modularity of the backend allows for scalable deployment, meaning that as more users or data volumes increase, the system can accommodate expansion without significant performance degradation. Additionally, post-prediction feedback mechanisms improve the model iteratively, making it smarter over time and increasing accuracy without the need for constant manual intervention. This self-enhancing capability is crucial for maintaining relevance in a dynamic banking environment.

Lastly, the efficiency is evident not just in technical performance but also in business outcomes. The application empowers bank staff with actionable insights through intuitive dashboards and visual analytics. By segmenting customers into high, medium, and low churn risk categories, the system enables targeted resource allocation. Banks can focus retention efforts on high-risk customers efficiently, leading to measurable improvements such as the 25

6.2 Comparison of Existing and Proposed System

Feature	**Traditional Customer Churn Management**	**Proactive Churn Prediction Using Machine Learning**
Approach	Manual, reactive	Automated, proactive, data-driven
Data Handling	Static reports; manual review of transaction history and complaints	Continuous monitoring of diverse data points (demographics, transactions, engagement)
Prediction Capability	Detects churn after it happens	Predicts churn risks before they occur
Efficiency	Time-consuming, error-prone; struggles with large data volumes	Real-time analytics; automated pipelines handle big data efficiently
Intervention Timing	Late (after churn signs are visible)	Early (before churn happens)
User Interface	Fragmented spreadsheets; manual interpretation	Interactive dashboards with clear visualizations
Decision-Making	Slow and less informed	Fast, data-backed, and targeted
Impact on Customer Retention	Limited due to delayed response	High due to early risk identification and timely actions
Operational Efficiency	Low	High
Customer Satisfaction and Profitability	Decreases over time	Increases with proactive engagement

Table 6.1: **Comparison**

Output

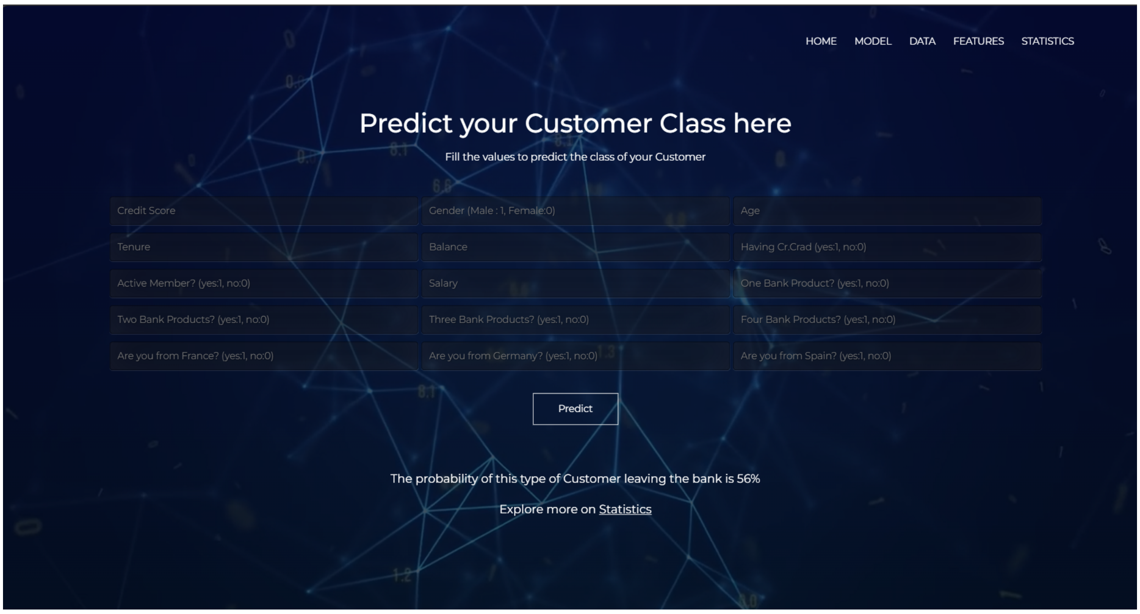


Figure 6.1: **High chance to churn**

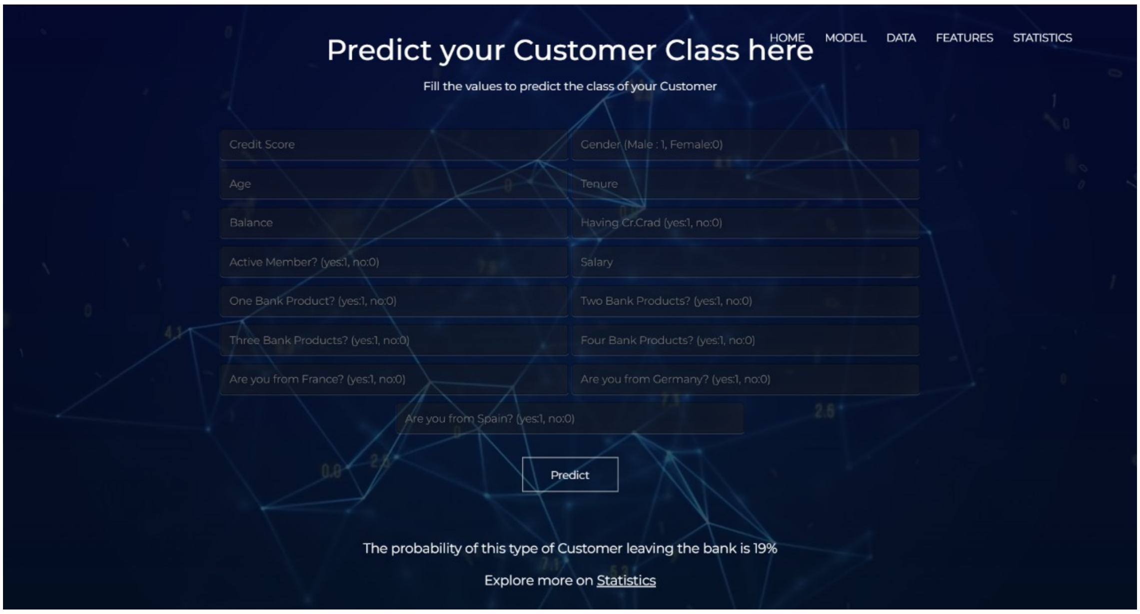


Figure 6.2: **Low chance to churn**

Figure 6.1&6.2 tells about after submitting the customer details, the output page displays the predicted probability of the customer leaving the bank. The result is clearly communicated, helping users quickly understand the risk associated with that customer profile. Additionally, a link to explore more detailed statistics is provided for users who wish to dive deeper into the model's insights and performance metrics.

Chapter 7

CONCLUSION AND FUTURE ENHANCEMENTS

7.1 Conclusion

The development of the customer churn prediction model has proven to be a highly effective application of machine learning within the banking sector, demonstrating the significant impact that data-driven solutions can have on customer retention. By incorporating a variety of data types, such as customer demographic information, banking relationships, and transactional behaviors, the model can accurately predict potential churn risks. This predictive power enables banks to identify customers who may be at risk of leaving, allowing for timely interventions and the implementation of personalized retention strategies. The ability to proactively address churn risks is essential in reducing financial losses, improving customer satisfaction, and fostering long-term customer loyalty. As a result, this model not only enhances operational efficiency but also aligns with the bank's broader goals of improving customer experience and ensuring financial sustainability in an increasingly competitive market.

The deployed customer churn prediction system features a user-friendly web interface, making it accessible to bank staff regardless of their technical expertise. The model's predictive capabilities are integrated into the application through a series of visual tools that simplify the process of understanding and interpreting the results. Key features such as risk categorization (high, medium, low) and visual dashboards allow decision-makers to assess customer behavior patterns at a glance, enabling them to prioritize actions and resources effectively. The system's recommendations, such as personalized retention strategies based on churn risk levels, empower staff to take appropriate actions with confidence. Additionally, the inclusion of visualization tools enhances the system's interpretability, ensuring that even non-technical users can interact with the system and make data-driven decisions.

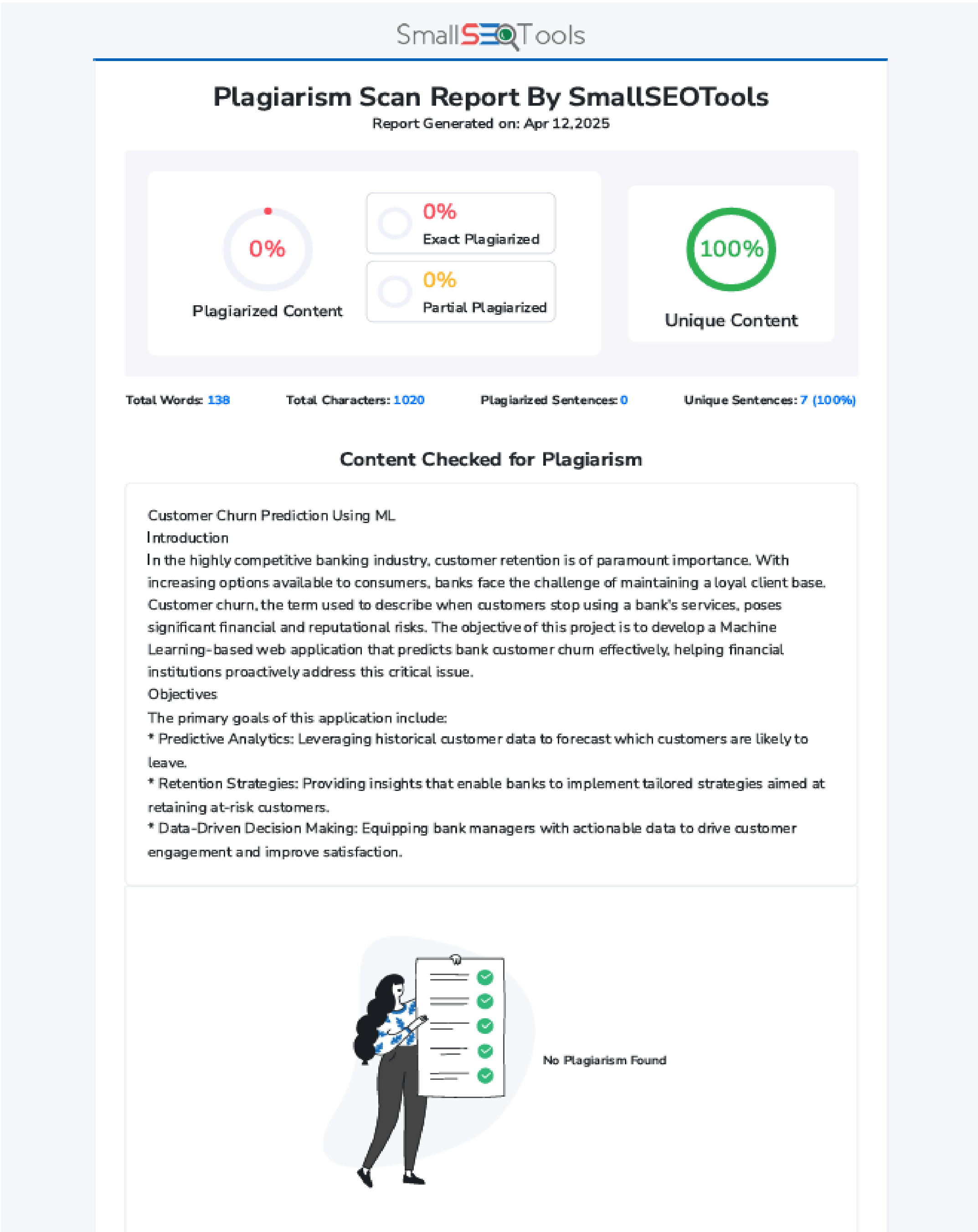
7.2 Future Enhancements

Looking to the future, several enhancements could further improve the model's effectiveness and expand its capabilities. One potential improvement is the integration of external data sources, such as social media sentiments, market trends, and real-time customer feedback. By incorporating this external data, the model can gain a more comprehensive understanding of the factors influencing churn, ultimately enhancing the accuracy of its predictions. Moreover, the addition of advanced machine learning techniques, such as ensemble methods and hyperparameter tuning, could help increase the model's reliability and performance across different customer segments. These improvements would not only refine the model's predictive power but also ensure that it remains adaptable to changing customer behaviors and market conditions over time.

Furthermore, expanding the application's functionalities to include automated customer engagement actions would make the system more dynamic and responsive. For example, real-time notifications could alert bank staff to high-risk customers, enabling them to take immediate action, such as offering targeted incentives or personalized communications. Personalized offers based on the churn probability would further strengthen the bank's relationship with its customers, increasing the likelihood of retention. Another key enhancement would be the introduction of real-time analytics dashboards, which would provide up-to-date insights into churn predictions and customer behavior trends. Additionally, incorporating feedback loops into the system would allow for continuous model improvement, ensuring that the predictions remain relevant and accurate. These future enhancements would transform the application into a powerful tool not only for customer retention but also for fostering long-term customer relationships and driving sustainable business growth.

Chapter 8

PLAGIARISM REPORT



SmallSEOTools

Plagiarism Scan Report By SmallSEOTools

Report Generated on: Apr 12,2025

0% Plagiarized Content

0% Exact Plagiarized

0% Partial Plagiarized

100% Unique Content

Total Words: 138 | Total Characters: 1020 | Plagiarized Sentences: 0 | Unique Sentences: 7 (100%)

Content Checked for Plagiarism

Customer Churn Prediction Using ML
Introduction
In the highly competitive banking industry, customer retention is of paramount importance. With increasing options available to consumers, banks face the challenge of maintaining a loyal client base. Customer churn, the term used to describe when customers stop using a bank's services, poses significant financial and reputational risks. The objective of this project is to develop a Machine Learning-based web application that predicts bank customer churn effectively, helping financial institutions proactively address this critical issue.
Objectives
The primary goals of this application include:
* Predictive Analytics: Leveraging historical customer data to forecast which customers are likely to leave.
* Retention Strategies: Providing insights that enable banks to implement tailored strategies aimed at retaining at-risk customers.
* Data-Driven Decision Making: Equipping bank managers with actionable data to drive customer engagement and improve satisfaction.

No Plagiarism Found

Figure 8.1: **Plagiarism Report**

Appendices

Appendix A

Complete Data / Sample Data / Sample Source Code / etc

```
from flask import Flask, render_template, request
import os
import keras
import numpy as np

app = Flask(__name__)

# Loading the model
model_path = os.path.dirname(os.path.realpath(__file__)) + '/savedmodel/mymodel'
model = keras.models.load_model(model_path)

@app.route('/')
def home():
    return render_template("index.html")

@app.route('/index.html')
def Home():
    return render_template("index.html")

@app.route('/About.html.html')
def Model():
    return render_template("About.html.html")

@app.route('/Data.html.html')
def Data():
    return render_template("Data.html.html")

@app.route('/Features.html.html')
def Features():
    return render_template("Features.html.html")

@app.route('/Stat.html.html')
def Statistics():
    return render_template("Stat.html.html")

@app.route('/Predict.html.html')
def Predict():
    return render_template("Predict.html.html")
```

```

@app.route('/predict', methods=['POST'])
def predict():
    def crscale(data):
        mean = 650.528800
        sd = 96.653299
        return (data - mean) / sd

    def agescale(data):
        mean = 38.921800
        sd = 10.487806
        return (data - mean) / sd

    def Tenscale(data):
        mean = 5.012800
        sd = 2.892174
        return (data - mean) / sd

    def Balscale(data):
        mean = 76485.889288
        sd = 62397.405202
        return (data - mean) / sd

    def Salscale(data):
        mean = 100090.239881
        sd = 57510.492818
        return (data - mean) / sd

    features = [float(x) for x in request.form.values()]

    crscore = crscale(features[0])
    age = agescale(features[2])
    tenure = Tenscale(features[3])
    balance = Balscale(features[4])
    salary = Salscale(features[7])

    gender = features[1]
    crcard = features[5]
    active = features[6]
    pro1 = features[8]
    pro2 = features[9]
    pro3 = features[10]
    pro4 = features[11]
    coun1 = features[12]
    coun2 = features[13]
    coun3 = features[14]

    finalFeatures = [
        crscore, gender, age, tenure, balance, crcard, active,
        salary, pro1, pro2, pro3, pro4, coun1, coun2, coun3
```

```
    ]

    finalvalues = np.array(finalFeatures).reshape(1, 15)
    res = model.predict(finalvalues)
    res = res.reshape(1)
    res = float(res) * 100
    res = round(res)

    return render_template(
        "Predict.html.html",
        predicttext="The probability of this type of customer leaving the bank is {}%".format(res)
    )

if __name__ == "__main__":
    app.run(debug=True)
```

References

[1] Ahmed, A., Maheswari, P. Machine Learning Techniques for Customer Churn Prediction, International Journal of Recent Technology and Engineering (IJRTE), 8, 2, 3007–3011. 2019.

[2] Burez, J., Van den Poel, D. Handling class imbalance in customer churn prediction, Expert Systems with Applications, 36, 3, 4626–4636. 2009.

[3] Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P. SMOTE: Synthetic Minority Over-sampling Technique, Journal of Artificial Intelligence Research, 16, -, 321–357. 2002.

[4] Chen, T., Guestrin, C. XGBoost: A scalable tree boosting system, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, -, -, 785–794. 2016.

[5] Ghazarian, A., Gulden, J. Enhancing Customer Retention with Predictive Analytics, Journal of Marketing Analytics, 9, 1, 30–42. 2021.

[6] Han, J., Pei, J., Kamber, M. Data Mining: Concepts and Techniques, Elsevier, -, -, -. 2011.

[7] Huang, B., Kechadi, M-T., Buckley, B. Customer churn prediction in telecommunications, Expert Systems with Applications, 39, 1, 1414–1425. 2012.

[8] Jahromi, A. T., Stakhovych, S. Forecasting customer churn using time series clustering and ensemble models, European Journal of Operational Research, 282, 2, 564–576. 2020.

[9] Kaur, H., Singh, M. Comparative Study of Classification Algorithms for Customer Churn Prediction, International Journal of Computer Applications, 184, 3, 19–25. 2022.

[10] Khan, T., Rizvi, S. Predictive Analytics for Reducing Customer Churn: A Machine Learning Approach, Procedia Computer Science, 181, -, 517–524. 2021.

[11] McKinsey & Company. The future of banking: How to retain customers in the digital era, McKinsey Insights, -, -, -. 2020.

[12] Rosset, S., Neumann, E., Eick, U., Vatnik, N. Customer lifetime value models for decision support, Data Mining Case Studies Workshop, KDD, -, -, -. 2003.

[13] Sakar, C. O., Polat, S. O., Katircioglu, M., Kastro, J. Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks, Neural Computing and Applications, 31, -, 6893–6908. 2019.

[14] Shaaban, E., Helmy, Y., Khedr, A. A Proposed Churn Prediction Model, International Journal of Engineering Research and Applications (IJERA), 2, 4, 693–697. 2012.

[15] Tsai, C-F., Lu, Y-H. Customer churn prediction by hybrid neural networks, Expert Systems with Applications, 36, 10, 12547–12553. 2009.

[16] Vafeiadis, T., Diamantaras, K. I., Sarigiannidis, G., Chatzisavvas, K. C. A comparison of machine learning techniques for customer churn prediction, Simulation Modelling Practice and Theory, 55, -, 1–9. 2015.

[17] Verbeke, W., Martens, D., Baesens, B. Social network analysis for customer churn prediction, Applied Soft Computing, 14, -, 431–446. 2014.

[18] Witten, I. H., Frank, E., Hall, M. A., Pal, C. J. Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, -, -, -. 2016.

[19] Xie, Y., Li, Y., Ngai, E. W. T., Ying, W. Customer churn prediction using improved balanced random forests, Expert Systems with Applications, 36, 3, 5445–5449. 2009.

[20] Zhao, Y., Zhang, X. Cost-sensitive learning for churn prediction in telecom, Knowledge-Based Systems, 159, -, 75–84. 2018.