AI Course

Capstone Project   
Final Report

For students (instructor review required)

ⓒ2023 SAMSUNG. All rights reserved.

Samsung Electronics Corporate Citizenship Office holds the copyright of this document.

This document is a literary property protected by copyright law so reprint and reproduction without permission are prohibited.

To use this document other than the curriculum of Samsung Innovation Campus, you must receive written consent from copyright holder.

| Credit Card Fraud Detection |
| --- |

**Date (12/09/2025)**

**Visionaries**

1- Wasmiah Alharbi

2- Shumukh Alotaibi

3- Haneen Alsaeed

4- Abdulwahab Alammari

5 -Bandar Aljedani

6- Fatimah Almalki

|  |
| --- |

Content

1. Introduction

1.1. Background Information

1.2. Motivation and Objective

1.3. Members and Role Assignments

1.4. Schedule and Milestones

2. Project Execution

2.1. Data Acquisition

2.2. Training Methodology

2.3. Workflow

2.4. System Diagram

3. Results

3.1. Data Preprocessing

3.2. Exploratory Data Analysis (EDA)

3.3. Modeling

3.4. User Interface

3.5. Testing and Improvements

4. Projected Impact

4.1. Accomplishments and Benefits

4.2. Future Improvements

5. Team Member Review and Comment

6. Instructor Review and Comment

1. Introduction

**1.1. Background Information**

The credit card fraud detection data is a widely used benchmark for studying fraud detection in financial transactions. It contains over 280,000 transactions, of which less than 0.2% of fraudulent, making it highly imbalanced. Each transaction includes 28 anonymised numerical features obtained using principal component analysis (PCA ) , along with the transaction time amount and a class label indicating whether the transaction is fraudulent ,the Objective is to correctly tell the difference between normal and fraud transactions using machine learning techniques that can handle imbalance data.

**1.2. Motivation and Objective**

The primary motivation behind this project is to deal with the increasing problem of financial fraud, which lead a big losses for both banks and customers. traditional Ruba systems are limited in detecting new and changing fraud patterns, making machine learning models are more effective solution. The Goal is to build and evaluate a system that can find fraudulent transactions, focus on high recall, ensuring that many fraud cases as possible identified ,we try multiple classification models including logistic regression, decision tree, random forest, XGboost and LightGBM, also applying different data balancing techniques to handel the imbalance data

**1.3. Members and Role Assignments**

**• Haneen Alsaeed (Team Leader)**: Responsible for training and evaluating the Logistic Regression model.

**• Shumukh Alotaibi**: Responsible for applying data balancing techniques.

**• Wasmiah Alharbi**: Responsible for data cleaning, training and evaluating the XGBoost model.

**• Fatimah Almalki :**  training and evaluating the Decision Tree model.

**• Abdulwahab Alammari:** Responsible for exploratory data analysis (EDA), training and evaluating the Random Forest model.

**• Bandar Aljedani:** training and evaluation of the LightGBM model.

**1.4. Schedule and Milestones**

**• Week 1**:collect the data , cleaning, preform exploratory data analysis, and balancing techniques.

**• Week 2:** Train the classification : Logistic Regression, Decision Tree, Random Forest, XGBoost, and LightGBM models.

**• Week 3:** evaluate the Models, compare there results, select the best model, and prepare the final report and presentation.

**Milestones:**

* Dataset cleaned, balanced, and ready for training, end of Week 1.
* All models trained and tuned , end of Week 2.
* Final evaluation completed , presentation and final report by the end of Week 3.



2. Project Execution

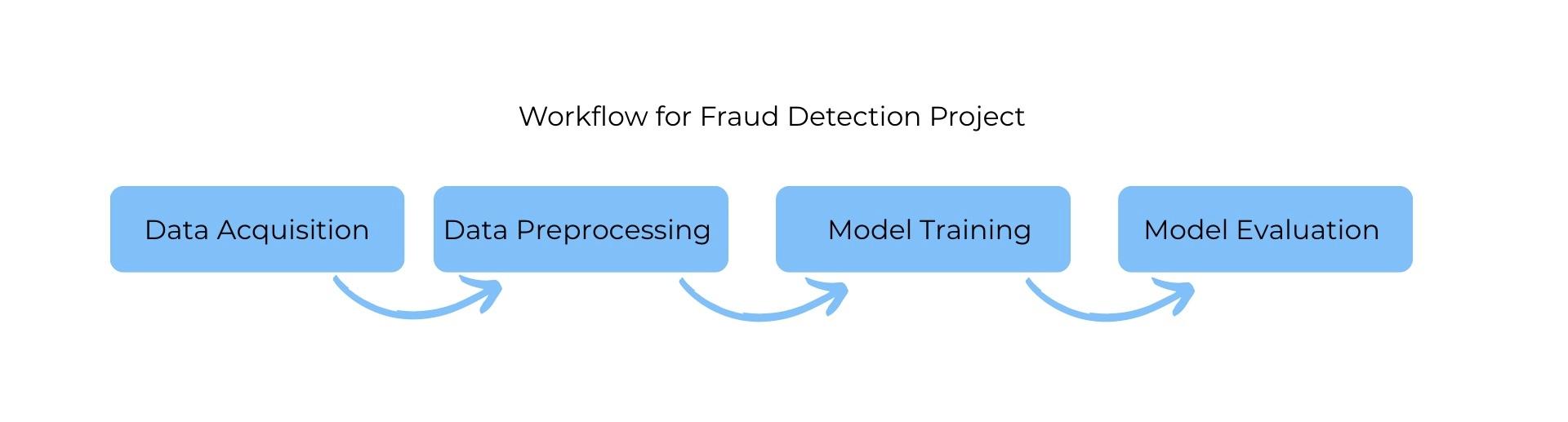
**2.1. Data Acquisition**

The dataset used in this study was obtained from the Kaggle Credit Card Fraud Detection dataset, which is widely recognized and publicly available for research purposes. It contains credit card transactions made by European cardholders in September 2013, collected and shared by the machine learning community to benchmark fraud detection models. The dataset consists of 284,807 transactions, of which 492 are labeled as fraudulent (≈0.17%). To ensure privacy and confidentiality, most of the original features were transformed using Principal Component Analysis (PCA), leaving only two untransformed attributes: Time (seconds elapsed between transactions) and Amount (transaction value). The target variable Class is binary, where 1 indicates fraud and 0 indicates a normal transaction.

**2.2. Training Methodology**

Five machine learning models were selected for this study: Logistic Regression, Decision Tree, Random Forest, XGBoost, and LightGBM. The dataset was preprocessed and balanced using Random Over Sampling (ROS), Random Under Sampling (RUS), and Synthetic Minority Over-sampling Technique (SMOTE) to address the severe class imbalance. Each model was trained on the processed dataset with hyperparameter tuning, including learning rate adjustment, maximum depth, number of trees, and batch size experimentation where applicable. The Adam optimizer was applied for gradient-based models, while ensemble models leveraged boosting and bagging strategies. Training was conducted over multiple epochs, and cross-validation was employed to ensure generalization and prevent overfitting. The models were evaluated using key metrics such as accuracy, precision, recall, and ROC-AUC to identify the best-performing approach for fraud detection.

**2.3. Workflow**



1. Data Acquisition

The dataset used in this project was obtained from the Kaggle Credit Card Fraud Detection dataset. It contains a total of 284,807 credit card transactions made by European cardholders in September 2013, of which only 492 transactions are labeled as fraud (≈0.17%). This dataset was chosen because it is real-world, highly imbalanced, and widely used as a benchmark for fraud detection research.

1. Data Preprocessing

In this stage, the raw dataset was cleaned and prepared for modeling. Duplicates and missing values were removed, and normalization was applied to improve the consistency of the features. Since the dataset is extremely imbalanced, balancing techniques such as Random Over Sampling (ROS), Random Under Sampling (RUS), and SMOTE were applied to create fairer training data and improve the detection of fraudulent transactions.

1. Model Training

Five machine learning models were trained on the preprocessed data: Logistic Regression, Decision Tree, Random Forest, XGBoost, and LightGBM. Hyperparameter tuning was performed to optimize the learning process, while cross-validation was applied to ensure that the models generalized well to unseen data. Each model was trained on multiple balanced versions of the dataset to compare their performance fairly.

1. Model Evaluation

The trained models were then evaluated on unseen test data. Performance was measured using key metrics such as ROC-AUC, Precision, and Recall, since accuracy alone is not reliable in imbalanced datasets. The results showed that Logistic Regression achieved the best ROC-AUC, Random Forest provided the highest Precision, and LightGBM and XGBoost offered the best balance between Precision and Recall. Additionally, applying SMOTE significantly improved Recall across all models, highlighting its effectiveness in handling class imbalance.

**2.4. System Design**

The system was designed as an offline, modular pipeline with four main components: data acquisition, preprocessing, model training, and evaluation. All experiments were implemented in Google Colab using Python notebooks. The workflow starts by loading the Kaggle credit-card fraud dataset, followed by data cleaning (removing duplicates and handling missing values), normalization when needed, and addressing the severe class imbalance using ROS, RUS, and SMOTE. Five machine learning models Logistic Regression, Decision Tree, Random Forest, XGBoost, and LightGBM were trained with cross-validation and hyperparameter tuning. Threshold adjustment was applied to improve the trade-off between precision and recall. Model evaluation was performed on a separate test set using ROC-AUC, precision, recall, and accuracy (secondary). All results, including confusion matrices and ROC/PR curves, were generated and documented directly in Colab notebooks. No web interface or deployment was included, as the focus of the project was on building and comparing models within the Colab environment.

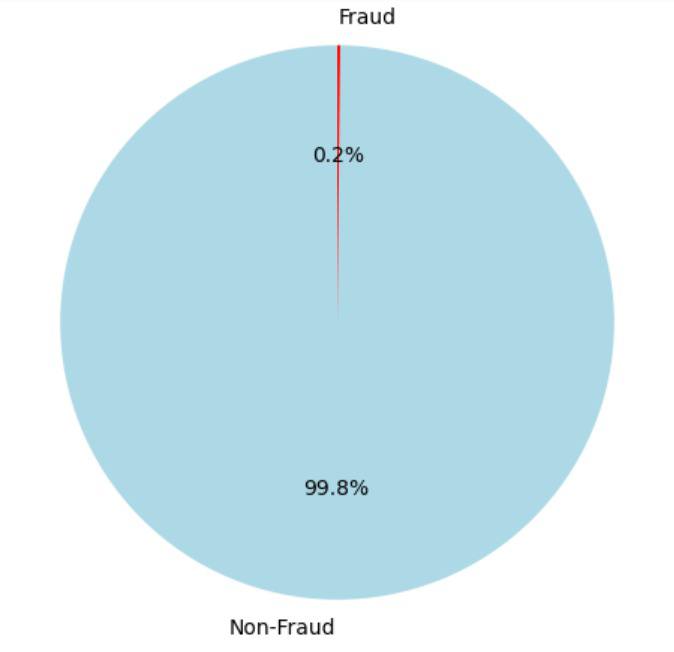
3. Results

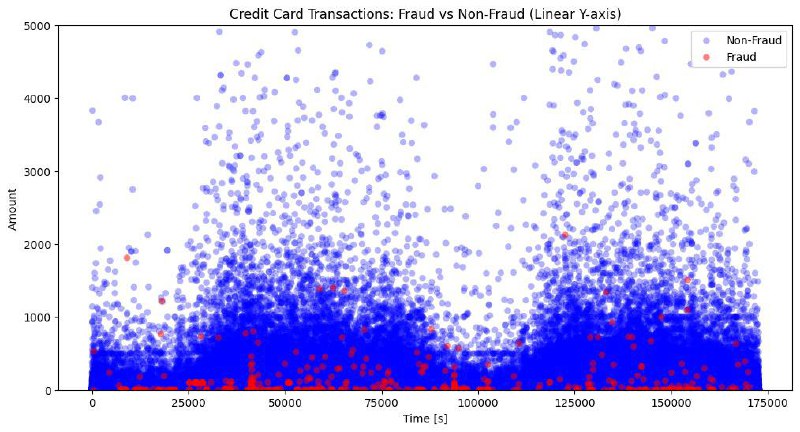
**3.1. Data Preprocessing**

The dataset was cleaned by removing duplicates and handling missing values. Due to the severe class imbalance (fraud cases ≈ 0.17%), different balancing techniques were applied, including Random Over Sampling (ROS), Random Under Sampling (RUS), and Synthetic Minority Over-sampling Technique (SMOTE). These methods helped prepare multiple balanced versions of the dataset for fair model training.

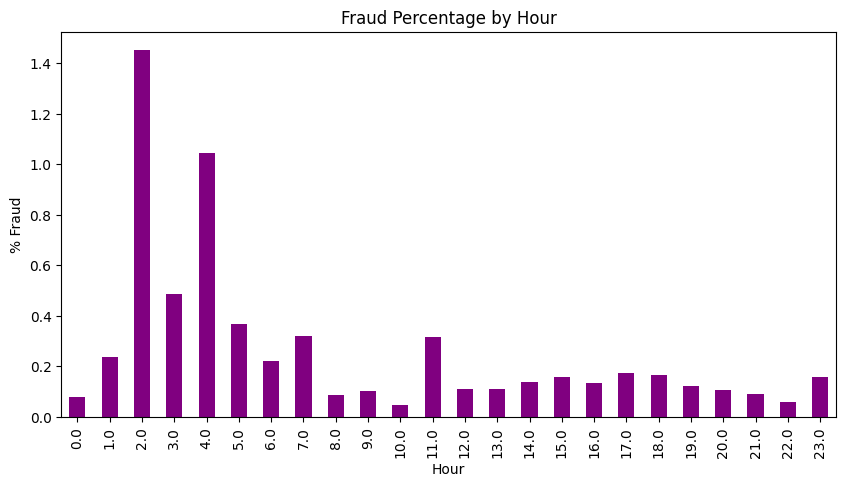
**3.2. Exploratory Data Analysis (EDA)**

The analysis showed that the dataset contains over 280,000 transactions, with only 492 labeled as fraud (≈0.17%). Fraudulent transactions are highly imbalanced compared to normal ones. Visualization of features confirmed that most fraud cases involve smaller amounts, and time distribution revealed no clear periodic pattern. Correlation analysis showed that certain PCA-transformed features are strongly related to fraud detection.





**Fraud vs Non-Fraud Transactions over Time and Amount**

****

**Fraud percentage peaks during early morning hours (around 2–4 AM), while remaining very low for the rest of the day.**

**3.3. Modeling**

Seven models Variants were : evaluated logistic regression,Random Forest ,XGboost( with different sampling strategies) and LightGBM. their performance vary across metrics:

**Logistic regression( tuned):**

achieve the highest ROC-AUC( 0.9528) confirming strong discrimination ability after threshold tuning recall reached (0.81 )but precision dropped to (0.20) leading to a modest F1 score (0.32).

**Random Forest (balanced):**

Delivered the highest precision (0.97) with recall (0.70 )achieving the strongest F1score( 0.82 )this makes it ideal when reducing false alarms is the priority.

**XGboost (scale pose weight):**

Balance trade -off with precision (0.62 )recall (0.82)PR-AUC (0.82) this indicates robust fraud detection while maintaining relatively fewer false positives.

**XGboost(Oversample):**

Similar recall (0.82) with improved precision (0.67), suitable when the objective is maximising fraud detection while keeping alerts reasonable,

**XGboost (Undersample):**

Very high recall (0.86) but extremely low precision (0.06) producing many false alarms.

**XGboost (SMOTE):**

Recall remained (0.82) ,but precision dropped to (0.40) still useful for scenarios emphasising fraud detection at the expensive alerts.

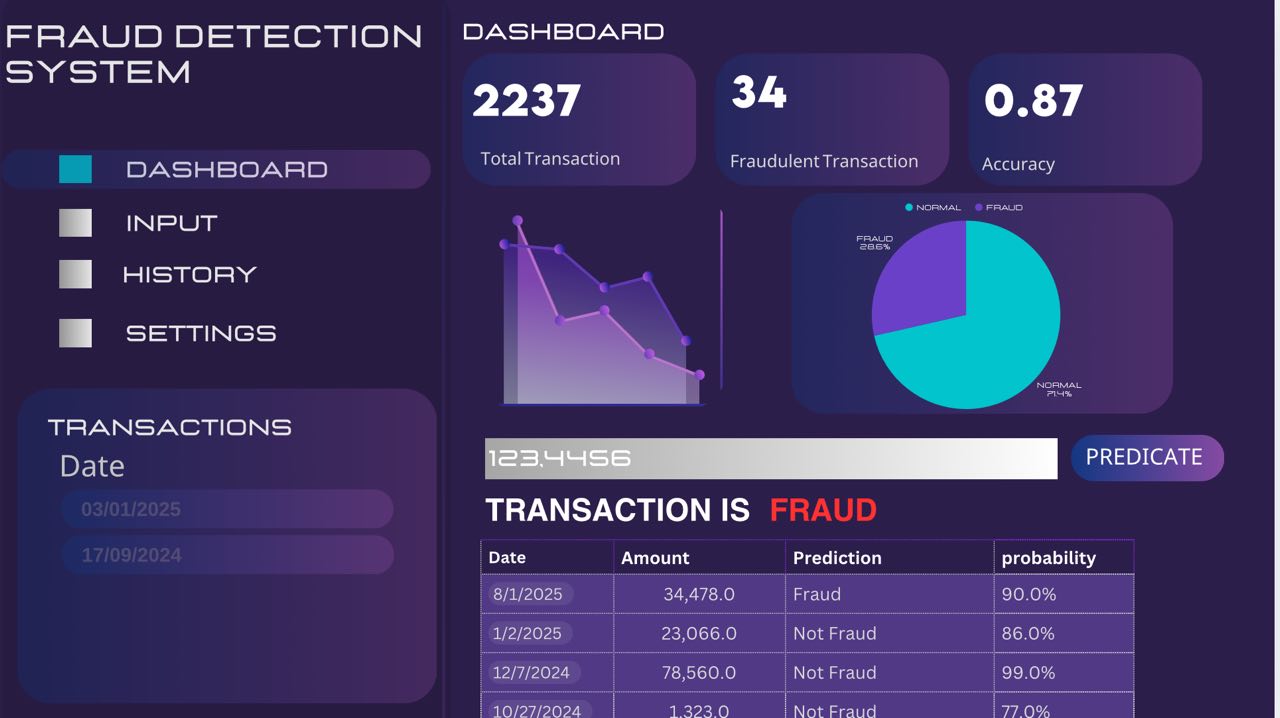
**LightGBM (scale\_pose\_weight):**

Achieved recall (0.76) with precision( 0.45 )but relatively weak PRAUC (0.35) showing it is less stable compared to XGboost variance

**Summary:**

* Random forest is best when minimising false positive is crucial
* XGBoost (over sample /scale\_pos\_weight) provides the best balance between recall and precision
* SMOTE strategies boosted recall but often reduced precision.

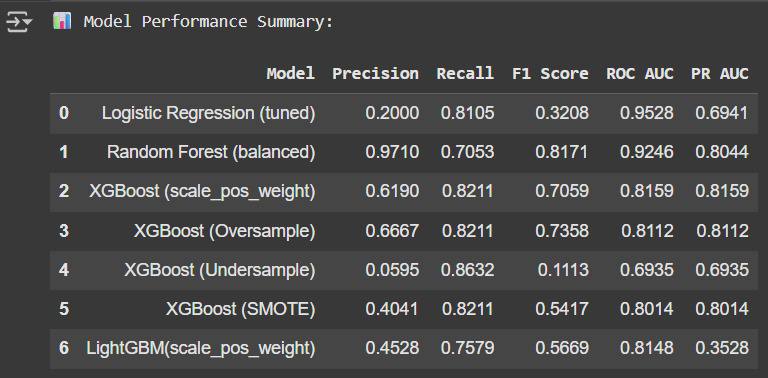
**3.4. User Interface**



**3.5. Testing and Improvements** The models were tested on unseen data to confirm generalization. Logistic Regression achieved ROC-AUC ≈0.95, but recall was limited (≈0.59) until threshold tuning raised it to ≈0.81 with lower precision (≈0.20). Random Forest achieved the highest precision (≈0.97) with recall around 0.70. LightGBM provided better balance, reaching recall ≈0.78 and precision ≈0.26, which improved to precision ≈0.45 and recall ≈0.76 at the best threshold. XGBoost also achieved balanced performance with ROC-AUC ≈0.90 and recall between 0.70–0.80.

Overall, the application of SMOTE significantly improved recall across models. Future improvements may include hyperparameter tuning, ensemble stacking, and real-time fraud detection integration.

And here Model Performance Summary :





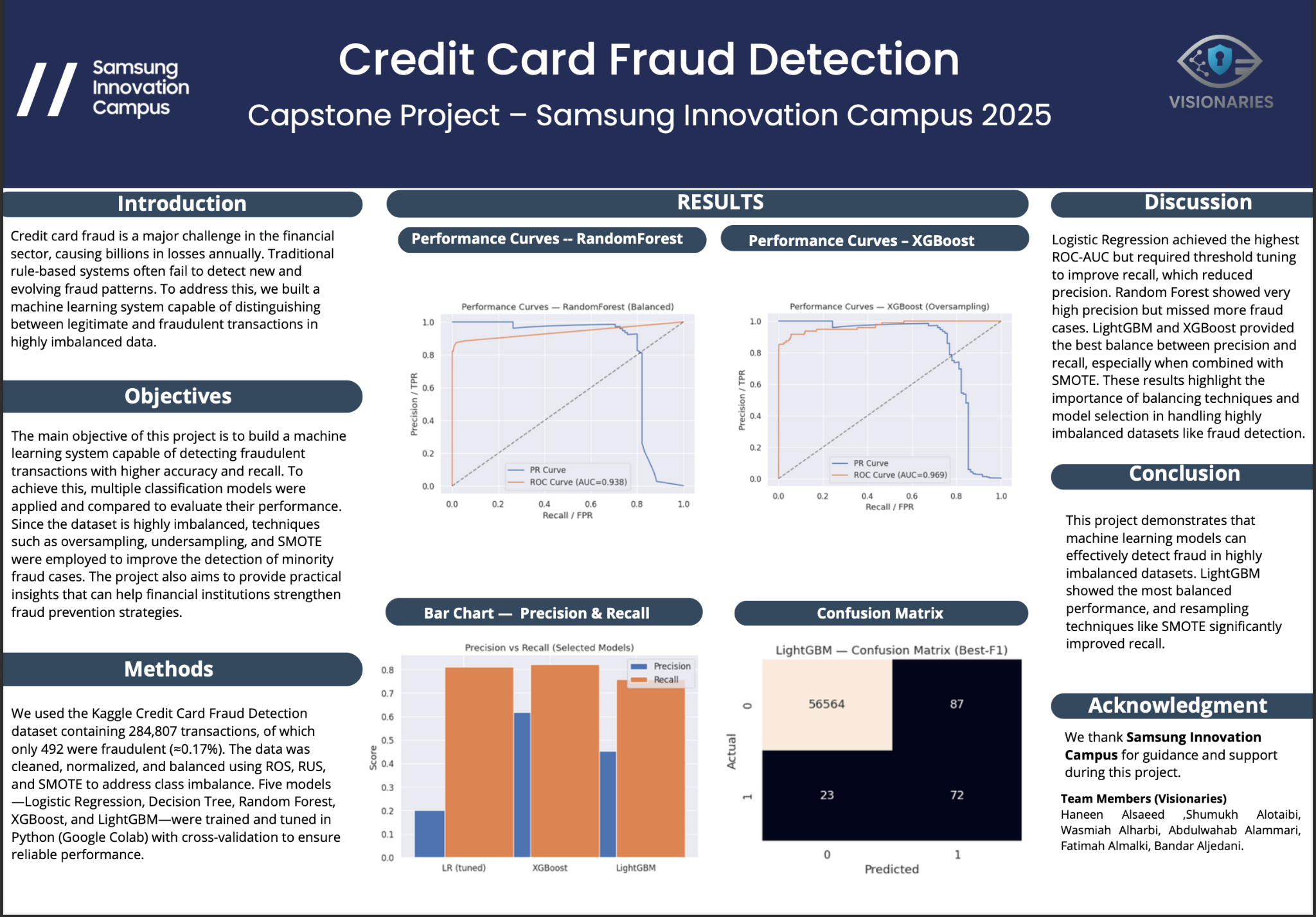
4. Projected Impact

**4.1. Accomplishments and Benefits**

* Successfully developed a machine learning system capable of detecting fraudulent credit card transactions with high performance.
* Applied multiple data balancing techniques (ROS, RUS, SMOTE) to handle severe class imbalance and improve fraud detection.
* Trained and compared five models — Logistic Regression, Decision Tree, Random Forest, XGBoost, and LightGBM — highlighting the strengths of each approach.
* Achieved a better trade-off between Precision and Recall using LightGBM and XGBoost, especially with SMOTE.
* Enhanced team members’ technical and collaborative skills, including:
  + Practical experience with Google Colab, Python, and cross-validation techniques.
  + Deeper understanding of imbalanced datasets and strategies to address them.
  + Strengthened teamwork, project management, and problem-solving abilities.

**4.2. Future Improvements**

* Perform more advanced hyperparameter tuning to further optimize model performance.
* Explore ensemble stacking methods to combine multiple models for higher accuracy.
* Extend the system into a **real-time fraud detection pipeline** rather than offline experiments.
* Design and implement a user-friendly interface (web or mobile) for practical usage.
* Expand the dataset to include larger and more recent transactions from diverse regions.
* Investigate advanced approaches such as **Neural Networks** or **AutoML** for potential performance gains.



5. Team Member Review and Comment 

|  |
| --- |

| NAME | REVIEW and COMMENT |
| --- | --- |
| Haneen | Leading this collaborative team was a valuable experience that enhanced my leadership and technical skills. |
| Bandar | This project helped me gain valuable knowledge, build meaningful friendships, and expand my network in the field |
| Fatimah | My skills improved and I learned a lot through the project and teamwork. |
| Abdulwahab | I benefited from discovering teamwork and also learned how to improve models using methods like SMOTE and balancing, which made this project a very valuable experience for me. |
| Wasmiah | This project helped me improve my technical skills, learn new concepts, and experience the value of teamwork |
| Shumukh | I gained new skills and enjoyed contributing to the project with my team |

6. Instructor Review and Comment

| CATEGORY | SCORE | REVIEW and COMMENT |
| --- | --- | --- |
| IDEA | \_\_/10 |  |
| APPLICATION | \_\_/30 |  |
| RESULT | \_\_/30 |  |
| PROJECT MANAGEMENT | \_\_/10 |  |
| PRESENTATION & REPORT | \_\_/20 |  |
| TOTAL | \_\_/100 |  |