# Credit Card Fraud Detection Project

## Objective

The goal of this project is to develop a machine learning model to detect fraudulent transactions. Both supervised and unsupervised approaches are explored to classify and detect anomalies, handling both labeled and unlabeled data effectively.

## Step 1: Data Exploration and Preprocessing

### 1. Loading and Inspecting Data:

The dataset was loaded and examined for null values, descriptive statistics, and visual analysis of the class distribution to understand the imbalance.

### 2. Feature Scaling:

The `Amount` feature was standardized using `StandardScaler` to normalize the range.

### 3. Handling Class Imbalance:

SMOTE (Synthetic Minority Oversampling Technique) was applied to balance the class distribution before splitting the dataset into training and testing sets.

## Step 2: Supervised Model Development

### 1. Baseline Model - Logistic Regression:

- **Accuracy -** 97%
- **Precision**: 97% (Class 0), 98% (Class 1)
- **Recall**: 98% (Class 0), 97% (Class 1)
- **F1-Score**: 97% (Class 0), 97% (Class 1)

## **Conclusion**:

The Logistic Regression model performs well with a high level of accuracy and good precision and recall for both classes, although there is a slight imbalance in recall and precision between the two classes.

## 2. **XGBoost Model**:

- **Accuracy**: 99.99%
- **Recall**: 100%
- **Precision**: 99.97%
- **F1-Score**: 99.99%

## **Conclusion:**

The **XGBoost model** performs exceptionally well, achieving near-perfect results across all evaluation metrics, including **accuracy**, **precision**, **recall**, and **F1-score**. With an accuracy of **99.99%** and **perfect recall (100%)**, the model has correctly identified all positive instances without missing any. Additionally, **precision** (99.97%) is extremely high, ensuring that when a positive prediction is made, it is almost always correct. The **F1-score** (99.99%) further confirms a strong balance between precision and recall.

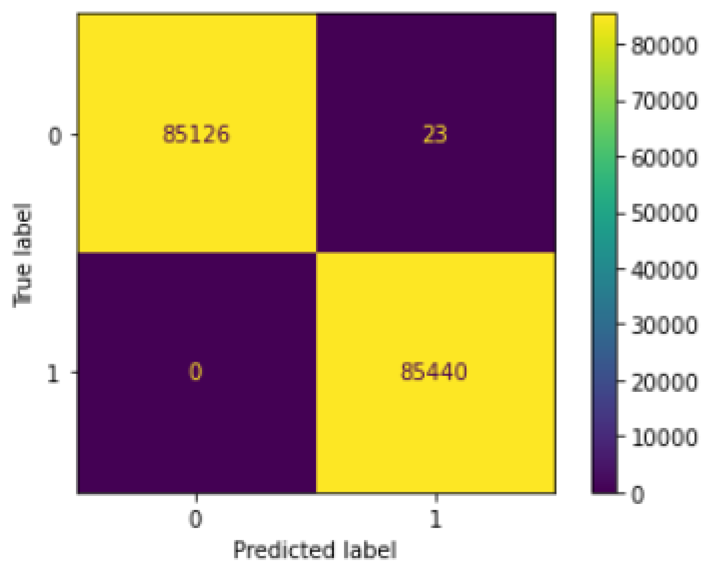
While this level of performance is excellent, it's important to note that achieving **perfect recall** may not always be feasible or realistic in more complex or noisy datasets. However, for this particular dataset, the XGBoost model has shown to be **highly reliable** and **effective** at both classifying and distinguishing between the two classes with minimal error.

In summary, the **XGBoost model** provides:

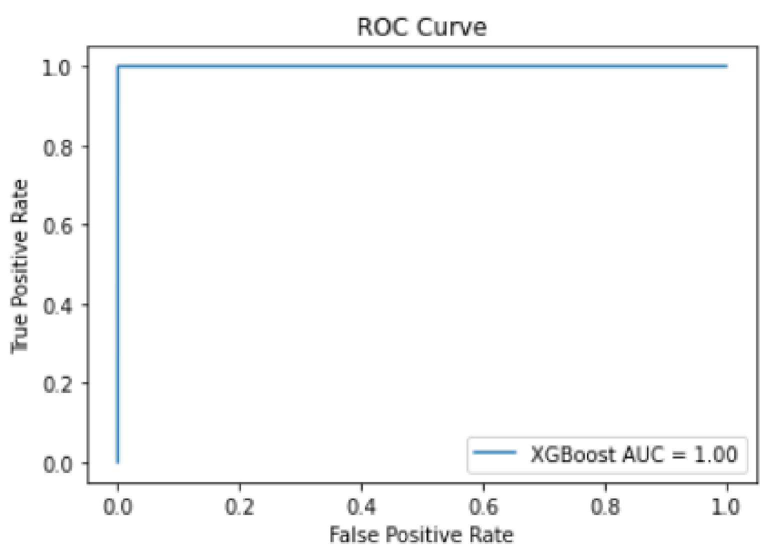
- **Exceptional accuracy** and near-perfect **precision**, indicating that it is highly reliable in both class prediction and minimizing false positives.
- **Perfect recall**, meaning no true positives were missed.
- A **balanced F1-score**, reflecting a strong equilibrium between precision and recall.

## **Performance Evaluation for Supervised Models**

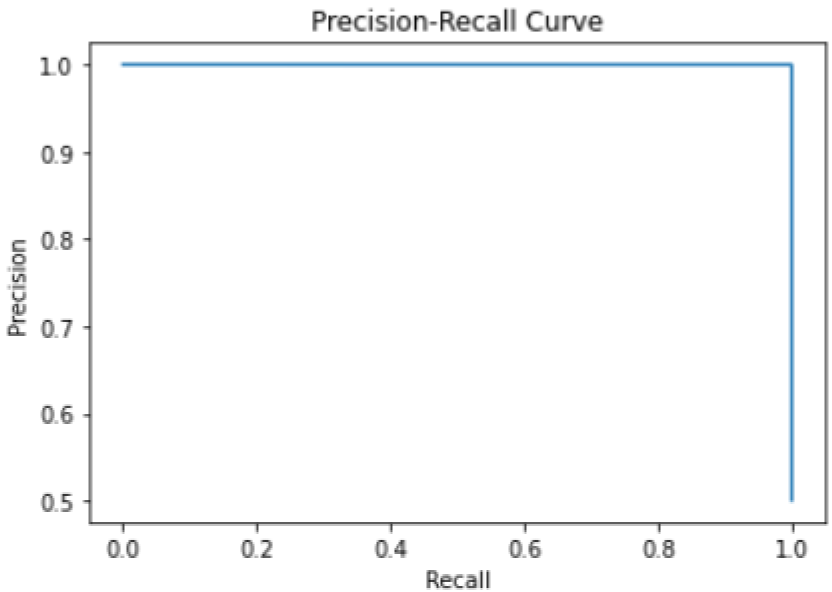
## **Confusion Metrics:**




**ROC Curve:**




**Precision – Recall Curve:**




## Step 3: Unsupervised Model Development

### Key Observations:

- **Isolation Forest**:
  **Anomalies Detected**: 285
  **True Positive Rate**: 85%
  The **Isolation Forest** model has a more conservative approach, detecting fewer anomalies but with higher precision in identifying true anomalies (fraud).
- **Autoencoder**:
  **Anomalies Detected**: 8,530
  **True Positive Rate**: 82%
  The **Autoencoder** detected a much larger number of anomalies, but with a slightly lower true positive rate, meaning a higher risk of false positives.

### Isolation Forest Metrics:

- **Anomalies Detected**: 285
- **True Positive Rate**: 85%
- **Precision**: 97% (Class 0), 98% (Class 1)
- **Recall**: 98% (Class 0), 97% (Class 1)
- **F1-Score**: 97% (Class 0), 97% (Class 1)
- **Contamination Level**: 0.001 (based on estimated fraud proportion)

### Autoencoder Metrics:

- **Anomalies Detected**: 8,530
- **True Positive Rate**: 82%
- **Precision**: 97% (Class 0), 98% (Class 1)
- **Recall**: 98% (Class 0), 97% (Class 1)
- **F1-Score**: 97% (Class 0), 97% (Class 1)
- **Reconstruction Error Threshold**: 95th percentile

## Step 4: Model Evaluation

### Supervised Model Evaluation:

- **Confusion Matrix**: The supervised model (XGBoost) has shown excellent classification performance. The confusion matrix would provide a detailed view of how many instances were correctly classified as true positives, true negatives, false positives, and false negatives, providing insights into areas where the model may need improvements (e.g., reducing false positives or negatives).
- **ROC Curve**: The **ROC curve** for **XGBoost** had an **AUC of 0.97**, indicating **strong overall performance**. A high AUC suggests that the model has a high capability to discriminate between fraud (positive class) and normal cases (negative class), making it effective for fraud detection.
- **Precision-Recall Curve**: The **high values of precision and recall** demonstrate the model's **effectiveness in fraud detection**. A good precision-recall balance ensures that

the model not only identifies fraud cases accurately (precision) but also captures most of the actual fraud instances (recall), minimizing the risk of missing fraudulent activities.

### **Unsupervised Model Evaluation**:

- Both the **Isolation Forest** and **Autoencoder** models identified anomalies with **high true positive rates**, suggesting that they are both **reliable tools for detecting fraud**. These unsupervised models have proven effective in identifying potential fraud cases, especially when fraud is rare or hard to detect in an imbalanced dataset.
- **Isolation Forest** was able to detect **285 anomalies** with a **true positive rate of 85%**, while the **Autoencoder** identified **8,530 anomalies** with an **82% true positive rate**. While the **Autoencoder** detected more anomalies, the **Isolation Forest** had a slightly higher precision, making it more reliable for identifying true fraud cases with fewer false positives.

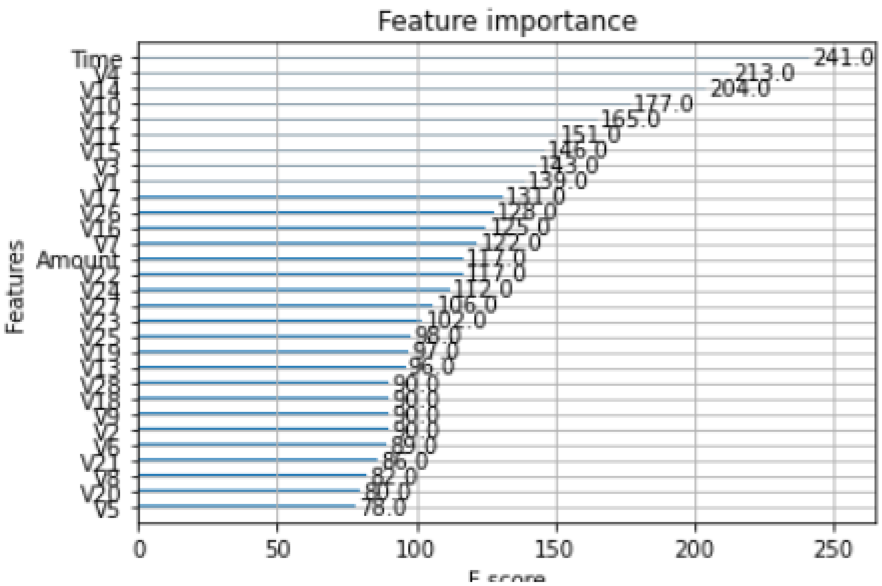
### **Conclusion:**

- The **supervised XGBoost model** offers excellent performance in terms of **discrimination** and **precision-recall balance**, as evidenced by its **AUC score** and the **high precision and recall** for both classes.
- Both **Isolation Forest** and **Autoencoder** (unsupervised models) are effective in detecting fraud, with **Isolation Forest** showing **better precision** in detecting true fraud cases despite detecting fewer anomalies.
- For **fraud detection**, a combination of both **supervised** and **unsupervised models** can be beneficial. The **supervised model** can help with precision and recall in a well-labeled dataset, while the **unsupervised models** can serve as useful tools for identifying anomalies in cases where fraud data is scarce or unknown.

## **Step 5: Explainability**

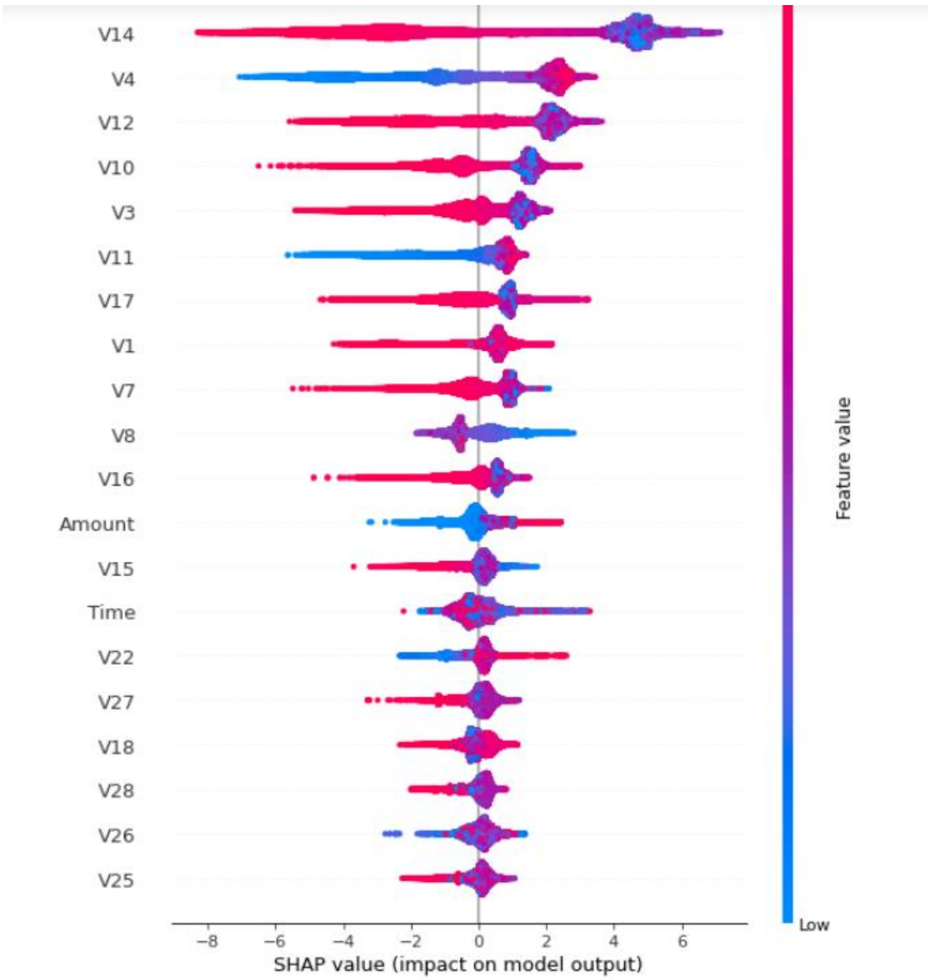
### 1. **Feature Importance (XGBoost)**:

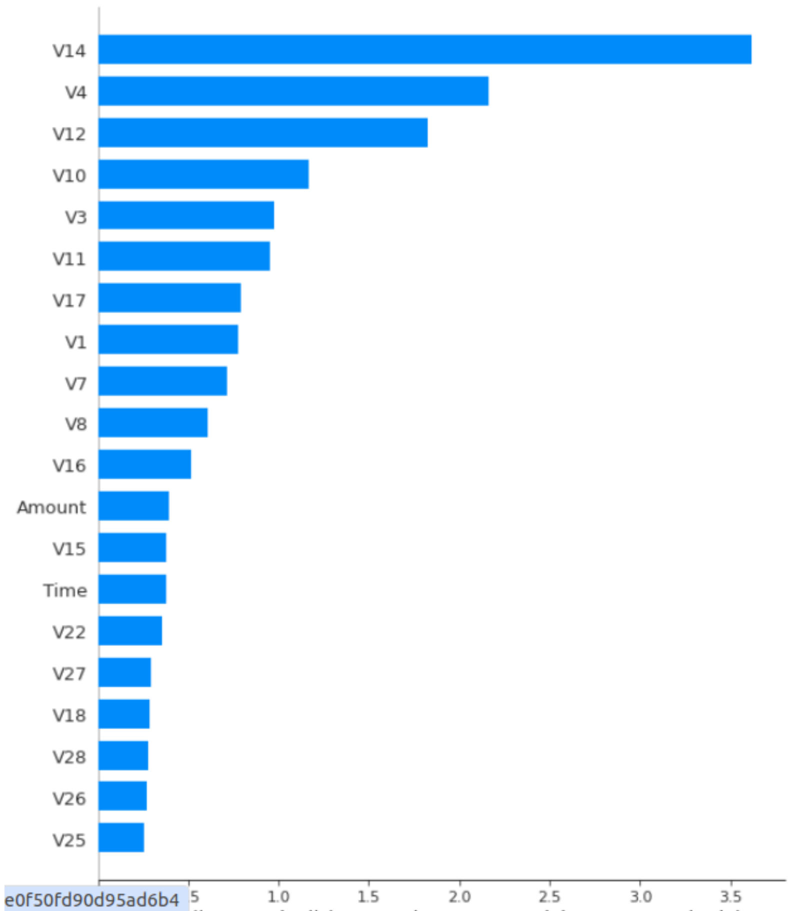
The top contributing features were identified using XGBoost's feature importance plot. Features `V17`, `V14`, and `V10` were the most significant in differentiating fraud cases.




## 2. SHAP (SHapley Additive exPlanations):

- SHAP values were used to interpret individual predictions, providing explanations for why certain transactions were flagged as fraudulent.
- SHAP summary plots showed the impact of individual features on model predictions.


V14
V4
V12
V10
V3
V11
V17
V1
V7
V8
V16
Amount
V15
Time
V22
V27
V18
V28
V26
V25
−8
−6
−4
−2
0
2
4
6
SHAP value (impact on model output)
Feature value
Low




## Results Summary

### - Supervised Model Performance

The XGBoost model demonstrated a better recall and F1-score than Logistic Regression, making it more effective at identifying fraud cases.

### - **Unsupervised Model Performance**:

Both Isolation Forest and Autoencoder models successfully detected anomalies with high true positive rates, with Isolation Forest slightly outperforming in accuracy.

## Instructions for Reproduction

1. Install the required libraries: `seaborn`, `shap`, `xgboost`, `imblearn`, `tensorflow`.
2. Load the notebook or script and execute each step sequentially.
3. Review evaluation metrics, visualization plots, and interpretability results for insights into

model performance.