AI Builder course
Title: Credit Card Fraud Detection

Course: AI Builder

Mentor name:
Abhishek Gagneja sir

Presented by:

Srijan Kumar Maheshwari (srijan25ai057@satiengg.in)

Kuldeep Rajak (kuldeep25ai024@satiengg.in)

Rupesh Namdev (rupesh25ai047@satiengg.in)

Vikas Dangi (vikas25ai064@satiengg.in)

Vishal Bhargava (vishal25ai066@satiengg.in)

Introduction

Credit card fraud has become a significant concern in the modern financial landscape, posing substantial risks to both consumers and financial institutions. With the increasing reliance on electronic payment systems, the frequency and sophistication of fraudulent activities continue to escalate, necessitating the development of robust detection mechanisms. In response to this challenge, the Credit Card Fraud Detection System project was conceived to address the pressing need for enhanced security measures in credit card transactions.

Methodology:-

1. Data Collection and Preprocessing
2. Feature Engineering
3. Model Selection and Training
4. Model Evaluation
5. Hyper-parameter Tuning and Optimization
6. Model Deployment and Integration
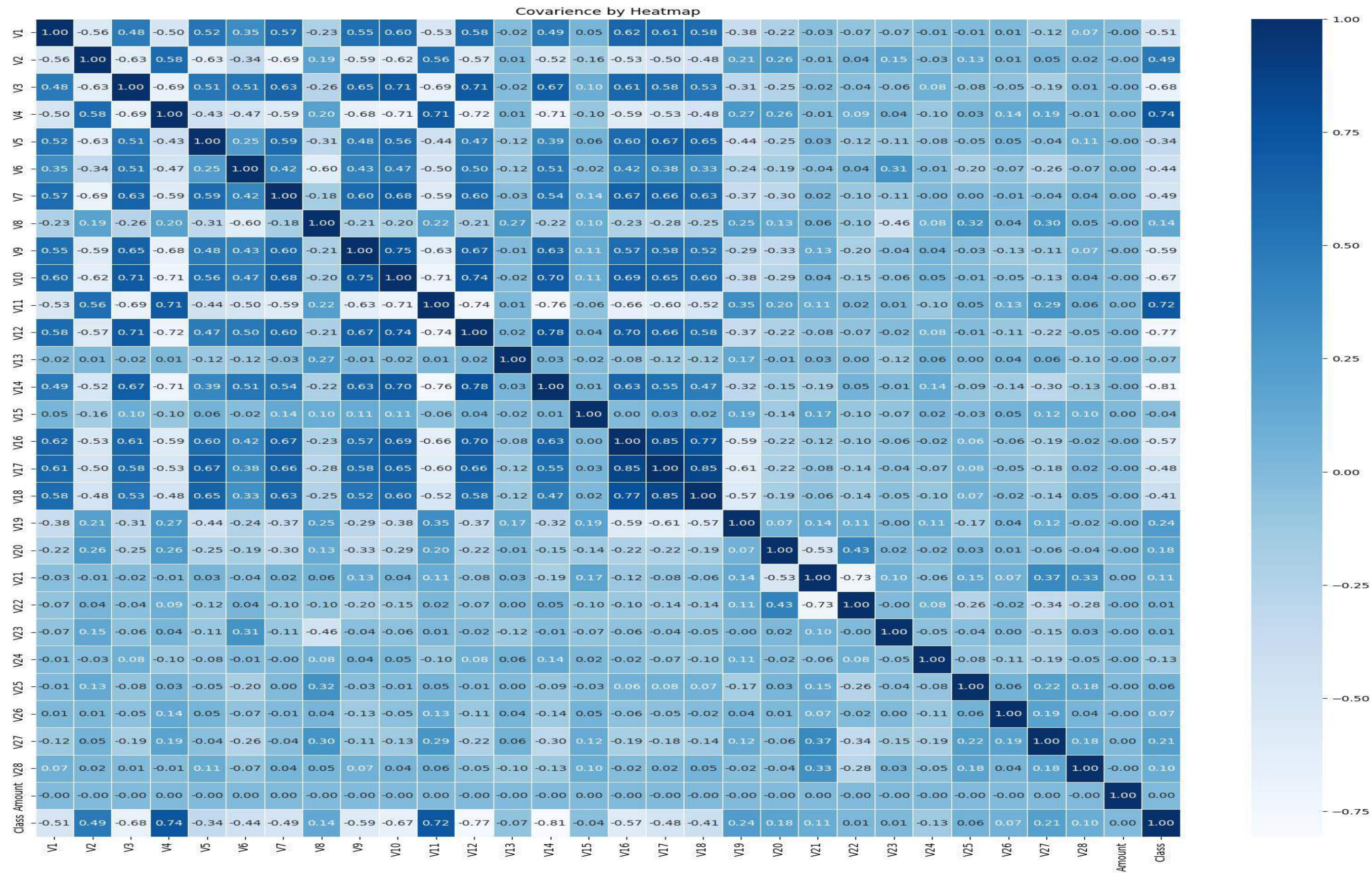
Architecture of Credit Card fraud Detection
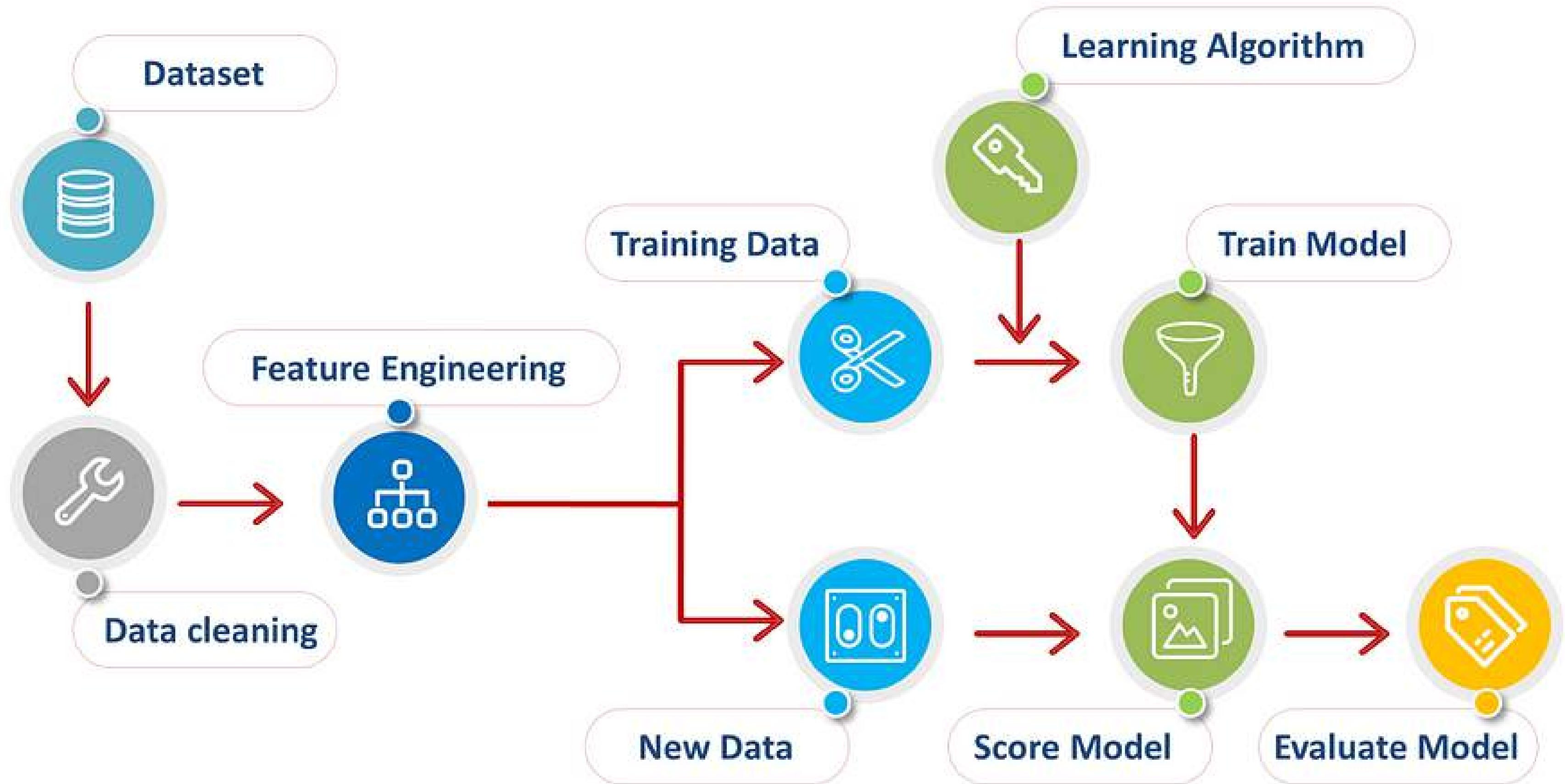
Data Preprocessing

Steps :-

- ➢ **Checking for null values**
- ➢ **Removing useless features**
- ➢ **Data understanding (describe function)**
- ➢ **Checking for unique values**
- ➢ **Scale the features**
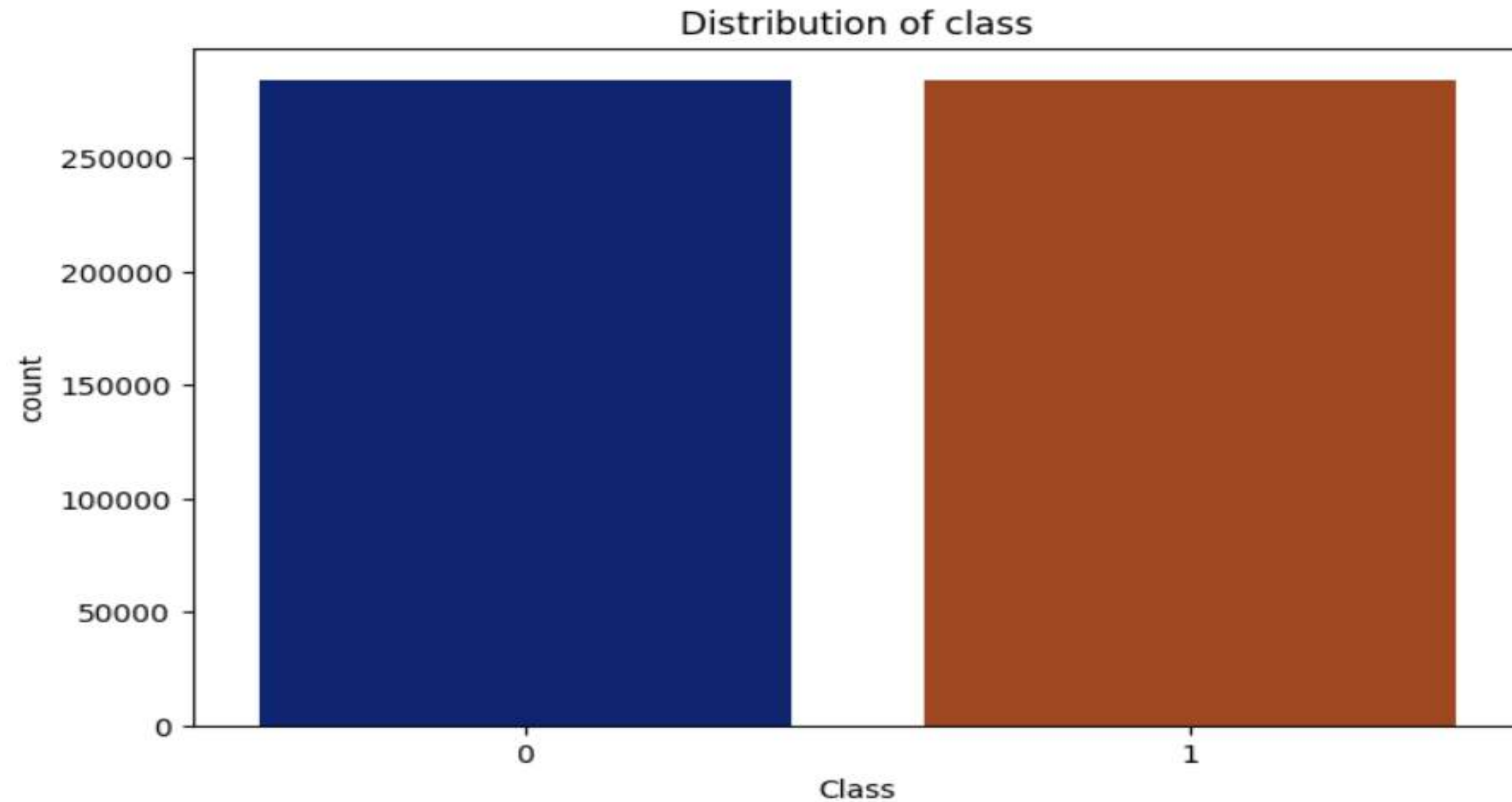- ➢ **Removing duplicate values (if any)**

Data visualization

Visualized the data with Heat map

Covarience by Heatmap

Distribute Data into Classes

Distribute Data into two classes 1. Fraud 2. Legit
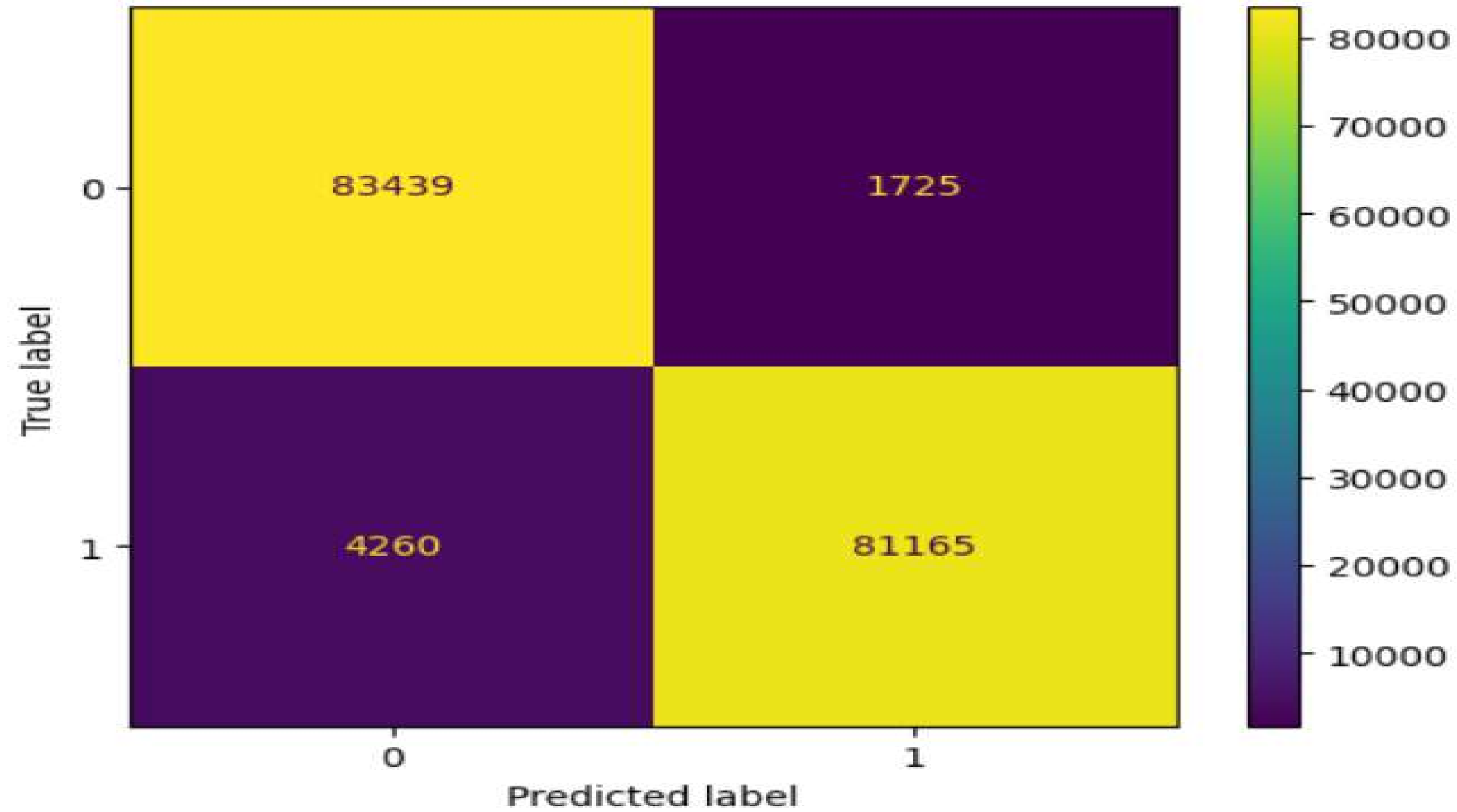

Distribution of class

Apply some machine learning models

1. Random Forest
2. Logistic Regression
3. Support Vector Machine (SVM)
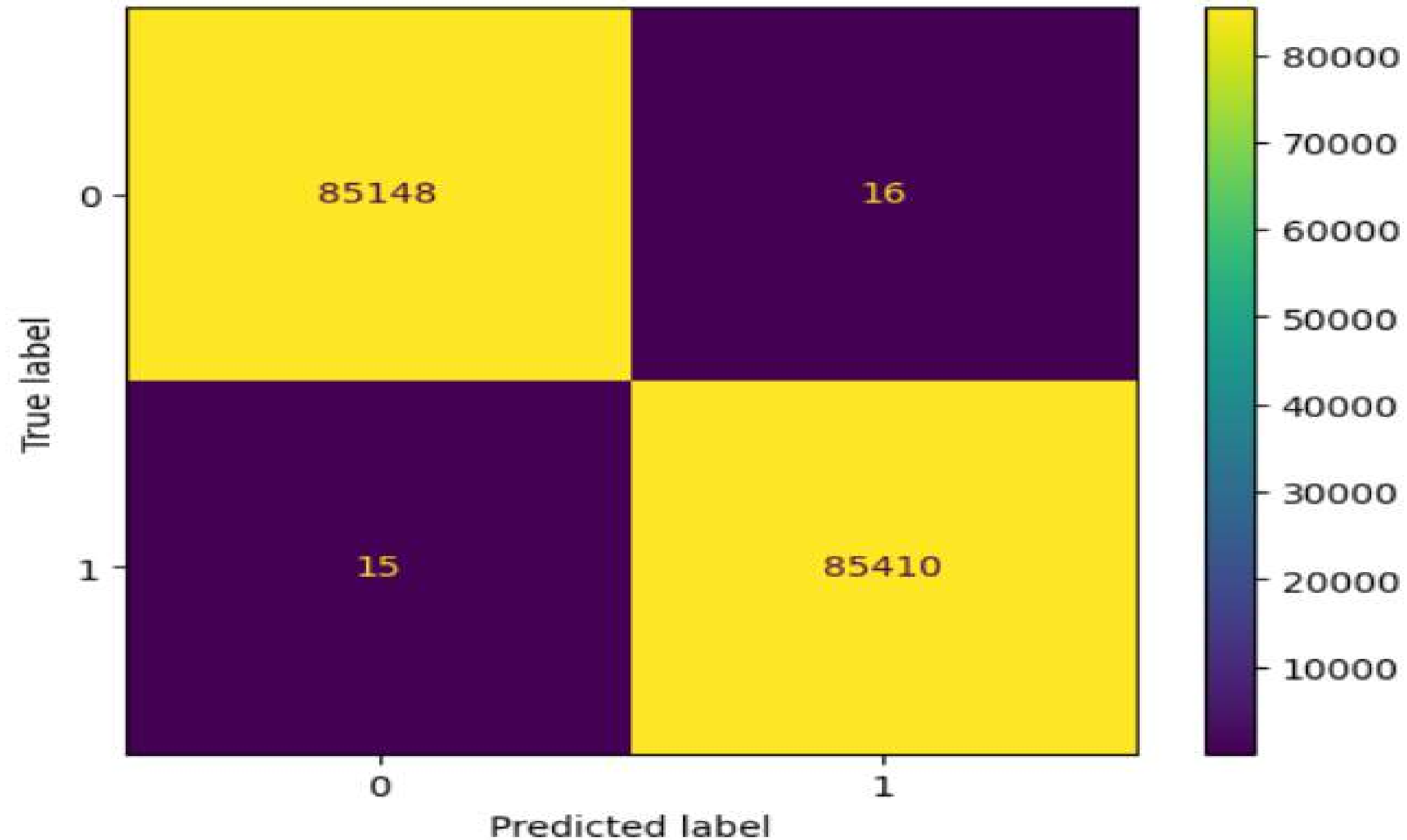4. K-nearest neighbours (KNN)
5. Sequential Model

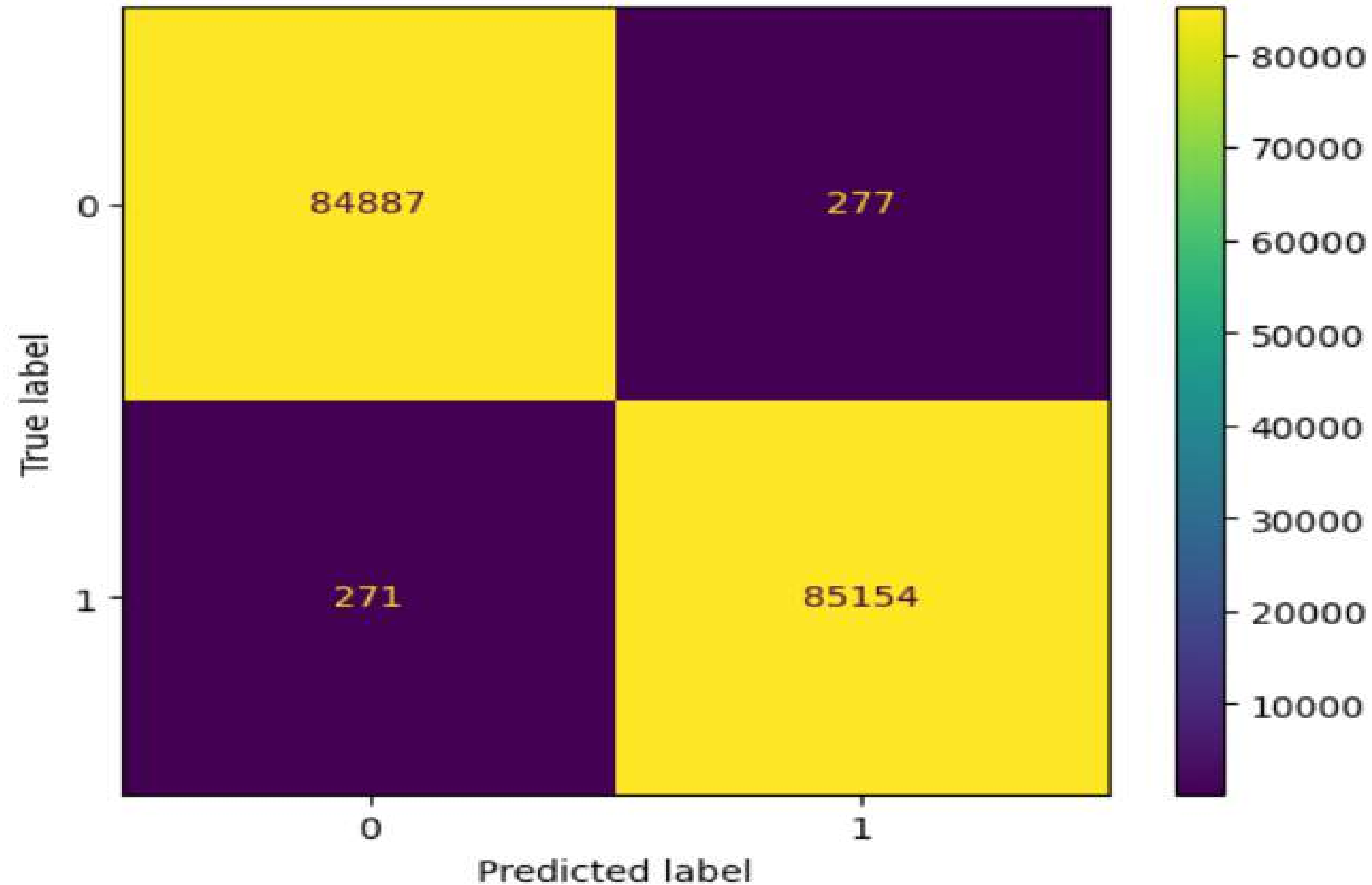# Logistic Regression
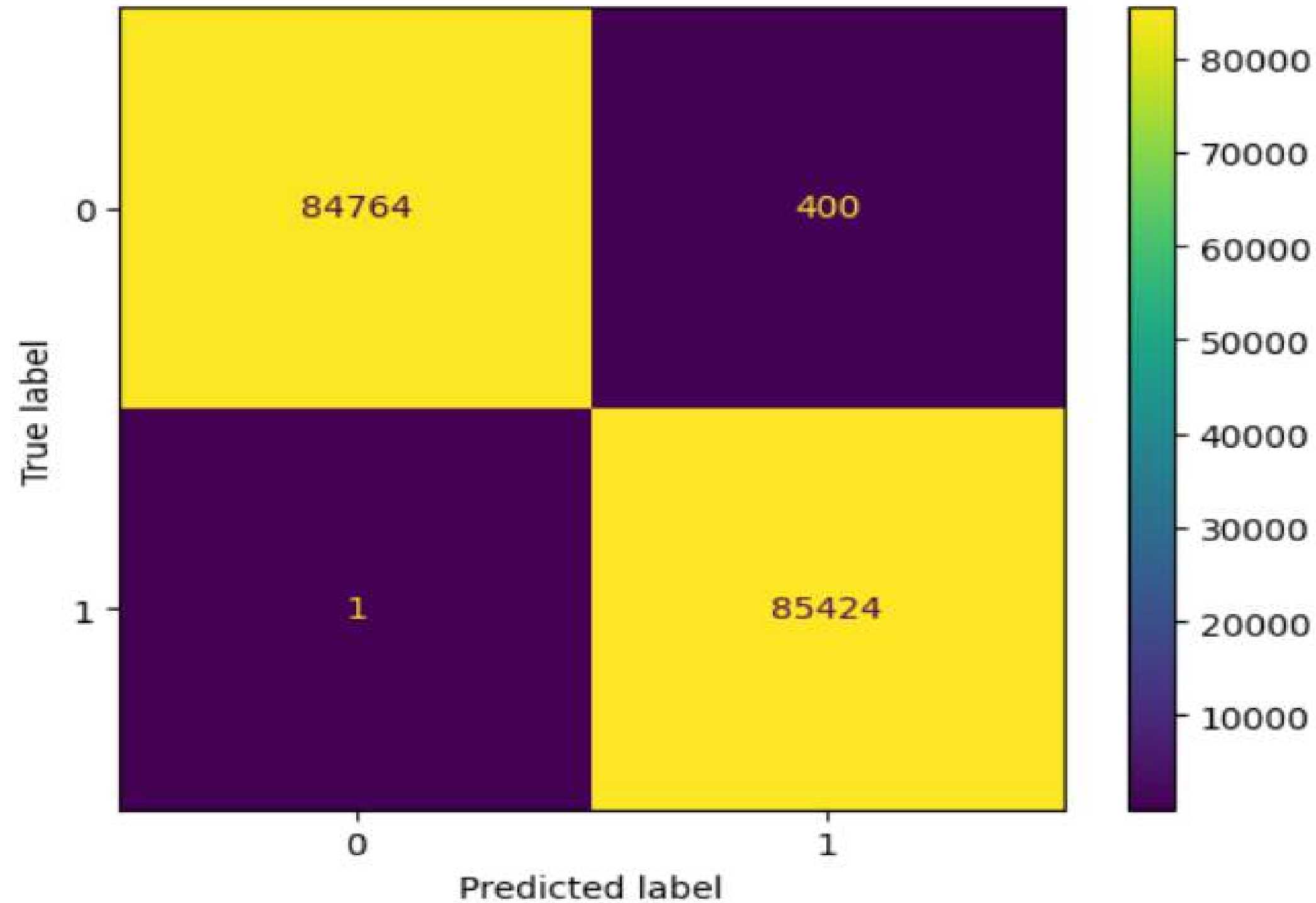
Random Forest

Support Vector Machine (SVM)

`<sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x7df81e4096f0>`

K-Neighbor Classifier

`<sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x7df81e40a740>`

Sequential model
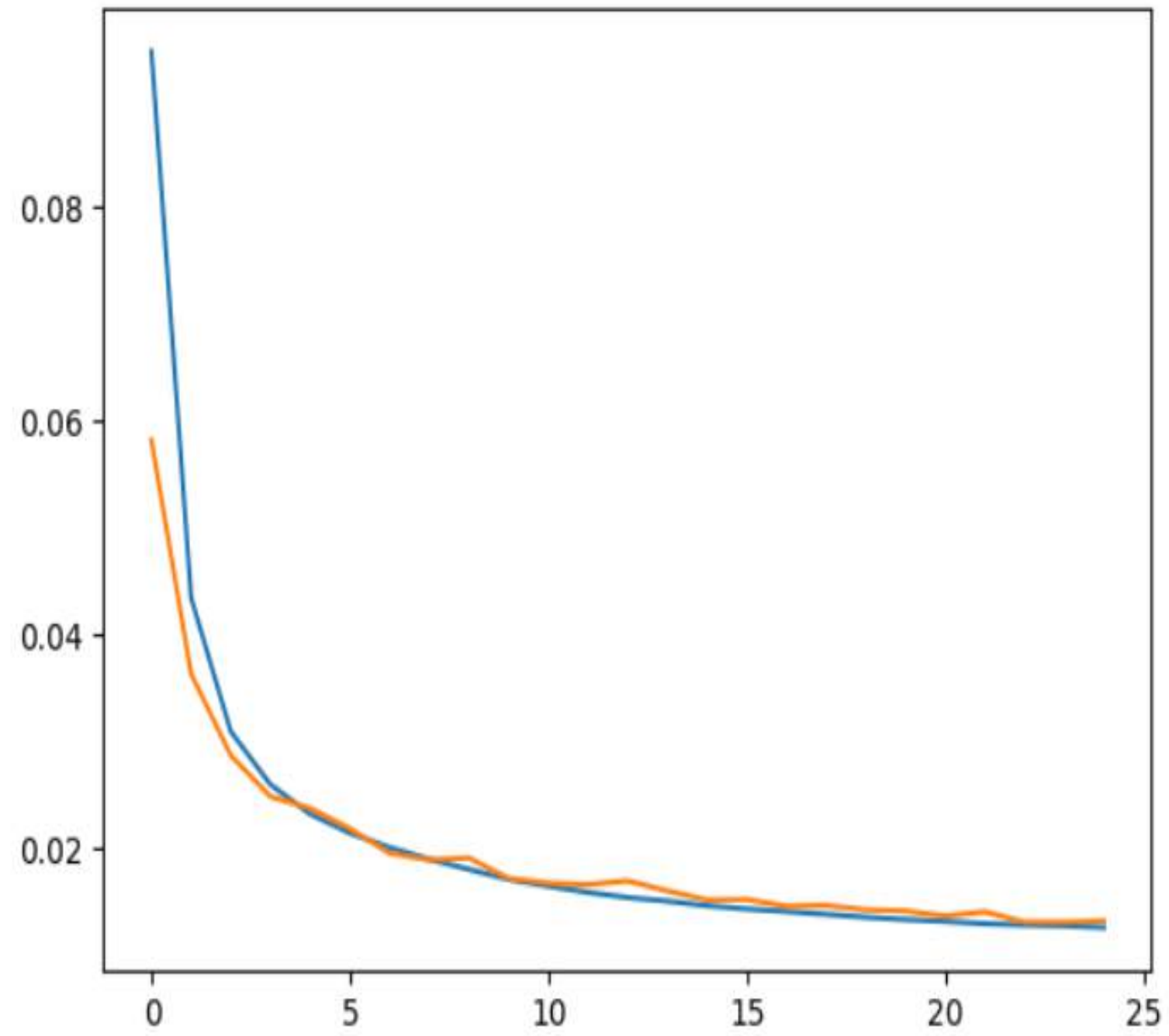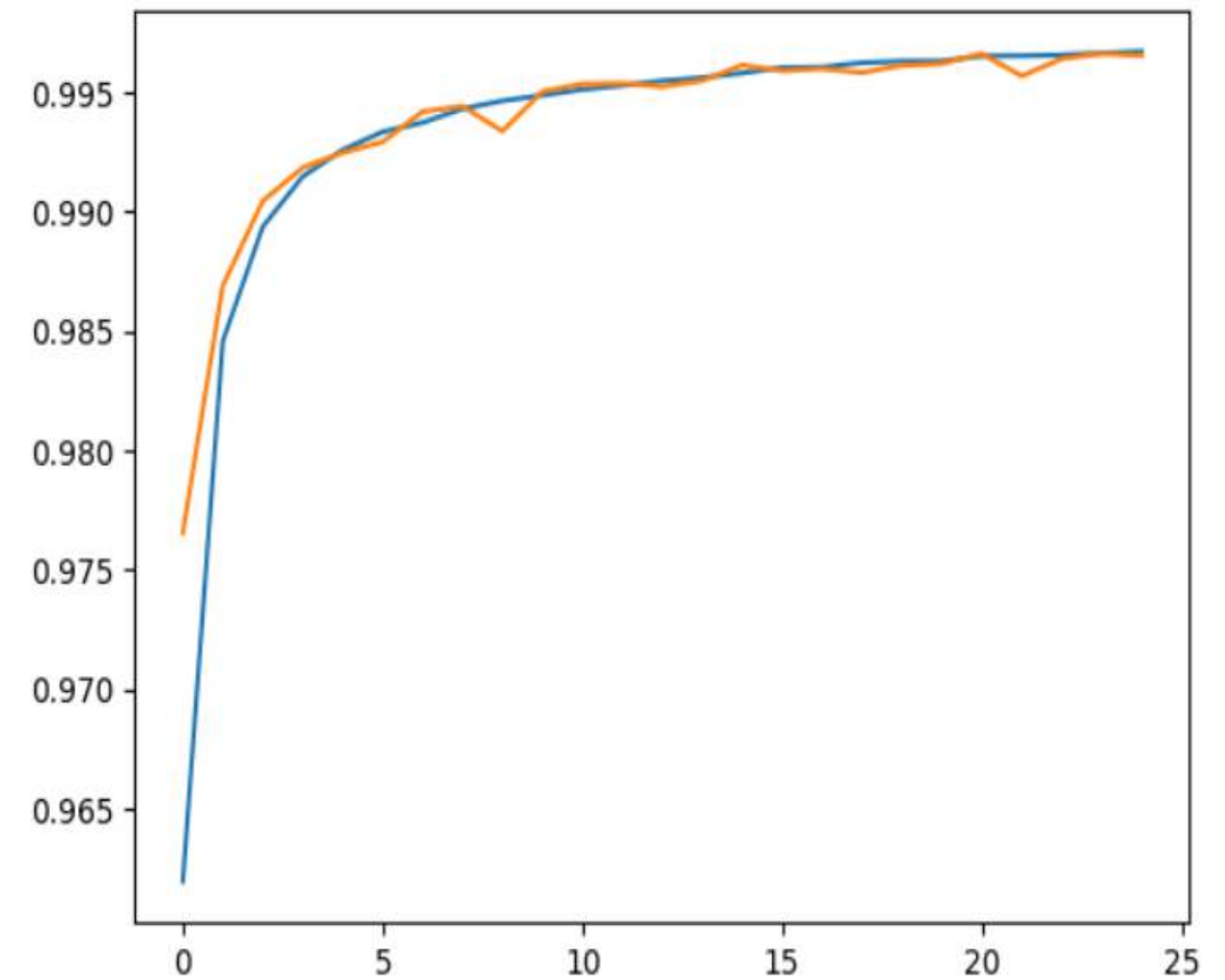

[<matplotlib.lines.Line2D at 0x7df81b7ef940>]

```
plt.plot(history.history["accuracy"])
plt.plot(history.history["val_accuracy"])
```

[<matplotlib.lines.Line2D at 0x7df81b86a080>]

Comparing model Accurary of all models

Comparing model Accuracy of all model

```python
# Comparing model
data = [["Random_forest",accuracy_score(y_test, y_rfc), f1_score(y_test, y_rfc)],
        ["Logistic_Regression",accuracy_score(y_test, y_pred_log), f1_score(y_test, y_pred_log)],
        ["SVM", accuracy_score(y_test, y_model_svm), f1_score(y_test, y_model_svm)],
        ["KNN",accuracy_score(y_test, y_model_knn), f1_score(y_test, y_model_knn)],
        ["Sequence_model",accuracy_score(y_test, y_pred_seq), f1_score(y_test, y_pred_seq)]]

model_accuracy = pd.DataFrame(data, columns=["Model_name","Acc_score", "F1-score"])
model_accuracy
```

	Model_name	Acc_score	F1-score
0	Random_forest	0.999818	0.999819
1	Logistic_Regression	0.964916	0.964442
2	SVM	0.996788	0.996793
3	KNN	0.997649	0.997658
4	Sequence_model	0.996477	0.996488

Conclusion & Future Improvements

In conclusion, the Credit Card Fraud Detection System project has yielded significant insights into the effectiveness of various machine learning and deep learning models in identifying fraudulent transactions. The top-performing models, such as the Random Forest Classifier and Sequential Neural Network, demonstrated robustness in accurately detecting fraudulent activities while minimizing false positives and negatives. Logistic Regression provided valuable interpretability, while K-Nearest Neighbors showcased scalability for real-time processing. These findings underscore the importance of leveraging ensemble methods and deep learning architectures in fraud detection systems.

References

Sequential model - [The Sequential model (keras.io)](#)

Dataset :-

[Credit Card Fraud Detection Dataset 2023 (kaggle.com](#)