CREDIT CARD FRAUD DETECTION USING DATA SCIENCE

\*\*Project Title:\*\* Credit Card Fraud Detection using Data Science

\*\*Objective:\*\* The primary aim of this project is to develop a machine learning model that can effectively identify and prevent fraudulent credit card transactions, ultimately improving the security of financial transactions.

\*\*Key Steps in the Project:\*\*

\*\*1. Data Collection:\*\*

- Gather historical credit card transaction data, which includes both legitimate and fraudulent transactions.

- The dataset should include features such as transaction amount, time, and various other transaction-related attributes.

\*\*2. Data Preprocessing:\*\*

- Handle missing values and outliers.

- Normalize or scale numerical features.

- Encode categorical features if applicable.

\*\*3. Exploratory Data Analysis (EDA):\*\*

- Visualize the data to understand patterns and relationships.

- Explore the distribution of legitimate and fraudulent transactions.

- Identify any potential correlations between variables.

\*\*4. Feature Selection and Engineering:\*\*

- Select relevant features that contribute to fraud detection.

- Create new features if they provide valuable information, e.g., transaction frequency or transaction velocity.

\*\*5. Model Selection:\*\*

- Choose appropriate machine learning algorithms for classification, such as logistic regression, random forests, or support vector machines.

- Consider deep learning models like neural networks for more complex patterns.

\*\*6. Data Splitting:\*\*

- Divide the dataset into training and testing sets to evaluate model performance.

- Consider using techniques like stratified sampling to maintain the fraud-to-non-fraud class balance.

\*\*7. Model Training:\*\*

- Train the selected models on the training dataset.

- Tune hyperparameters to optimize model performance.

\*\*8. Model Evaluation:\*\*

- Evaluate model performance using metrics such as accuracy, precision, recall, F1-score, and ROC AUC.

- Assess the model's ability to detect fraudulent transactions while minimizing false positives.

\*\*9. Deployment:\*\*

- Deploy the trained model to monitor real-time credit card transactions.

- Implement an alert system that triggers when potential fraud is detected.

\*\*10. Continuous Monitoring and Improvement:\*\*

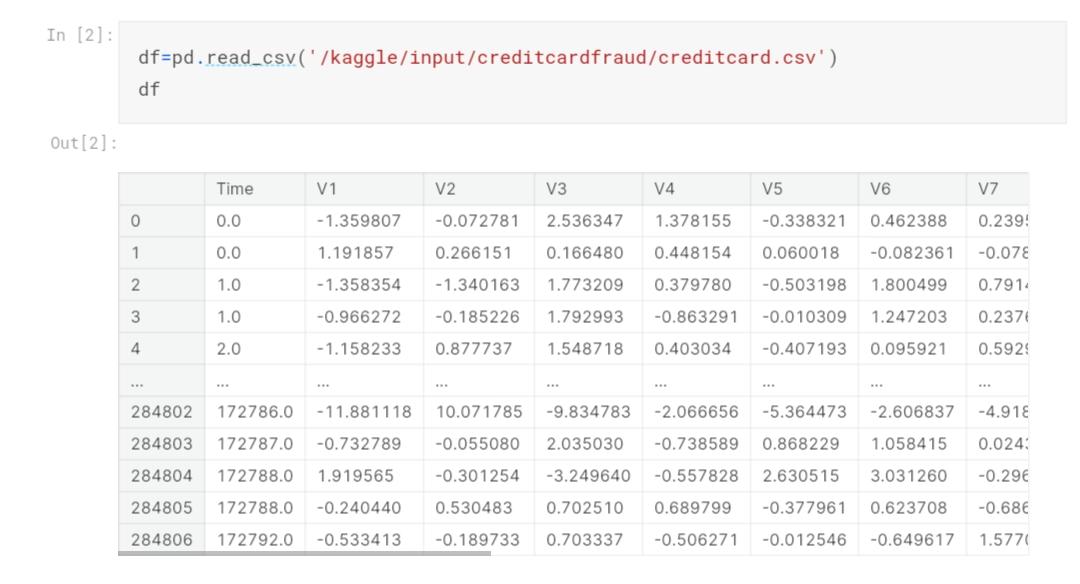
- Continuously monitor the model's performance.

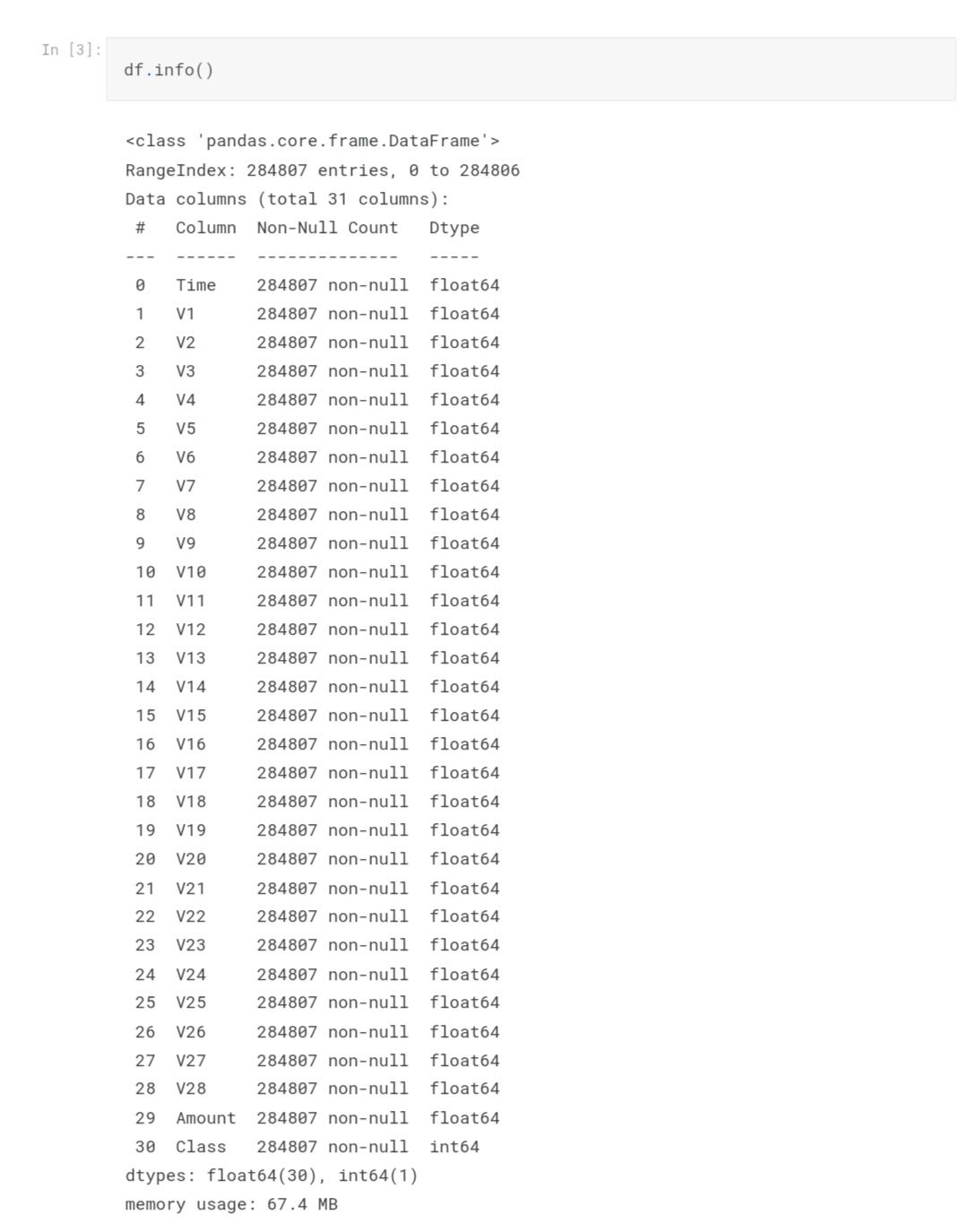
- Retrain the model periodically to adapt to evolving fraud patterns.

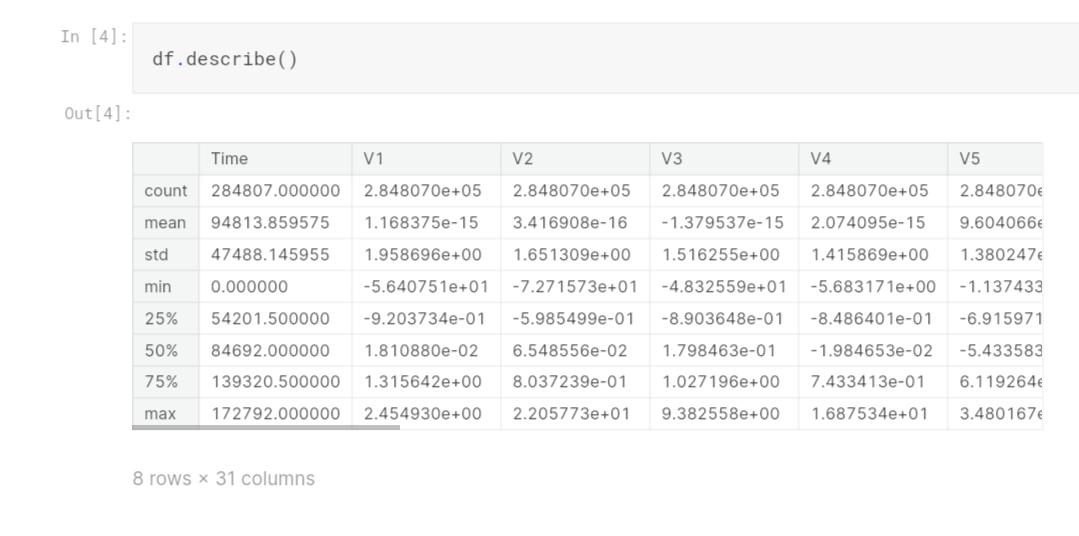
Libraries:

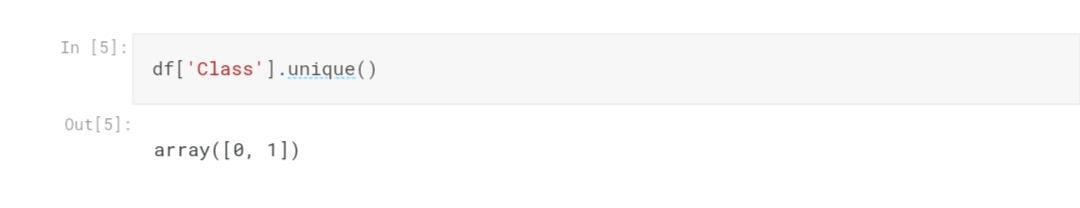


Pre-processing:

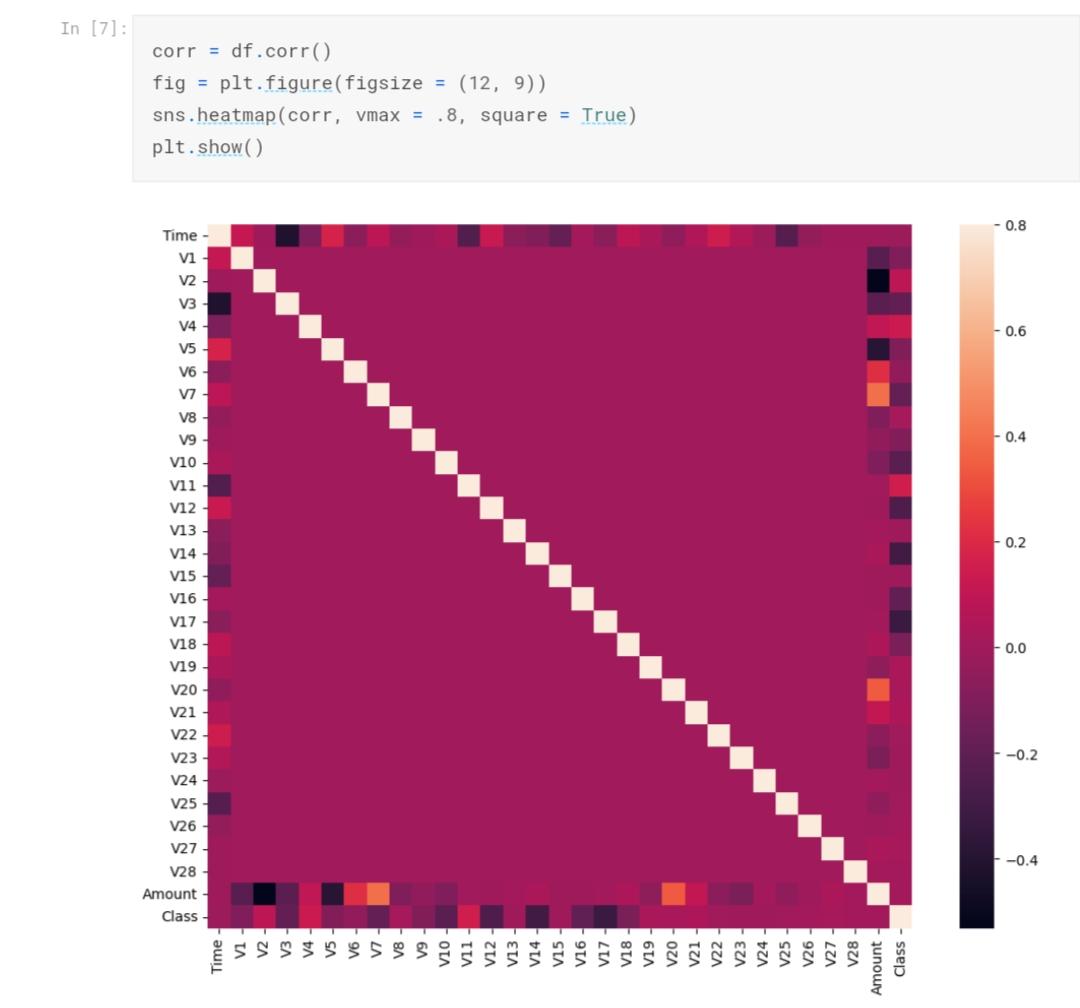


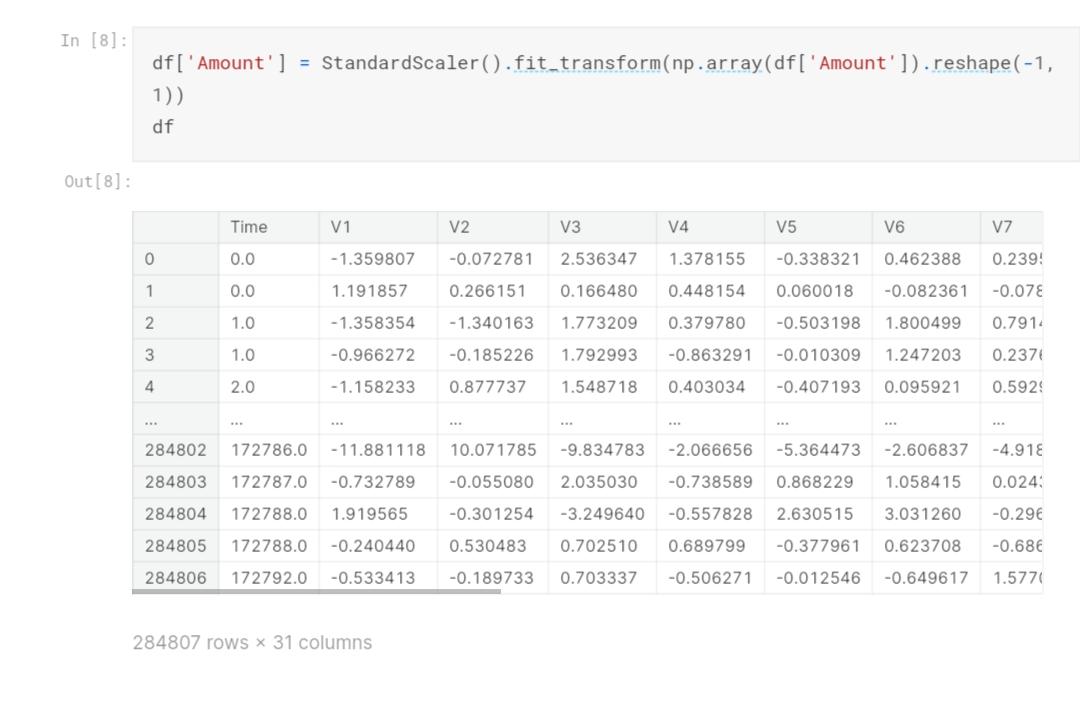








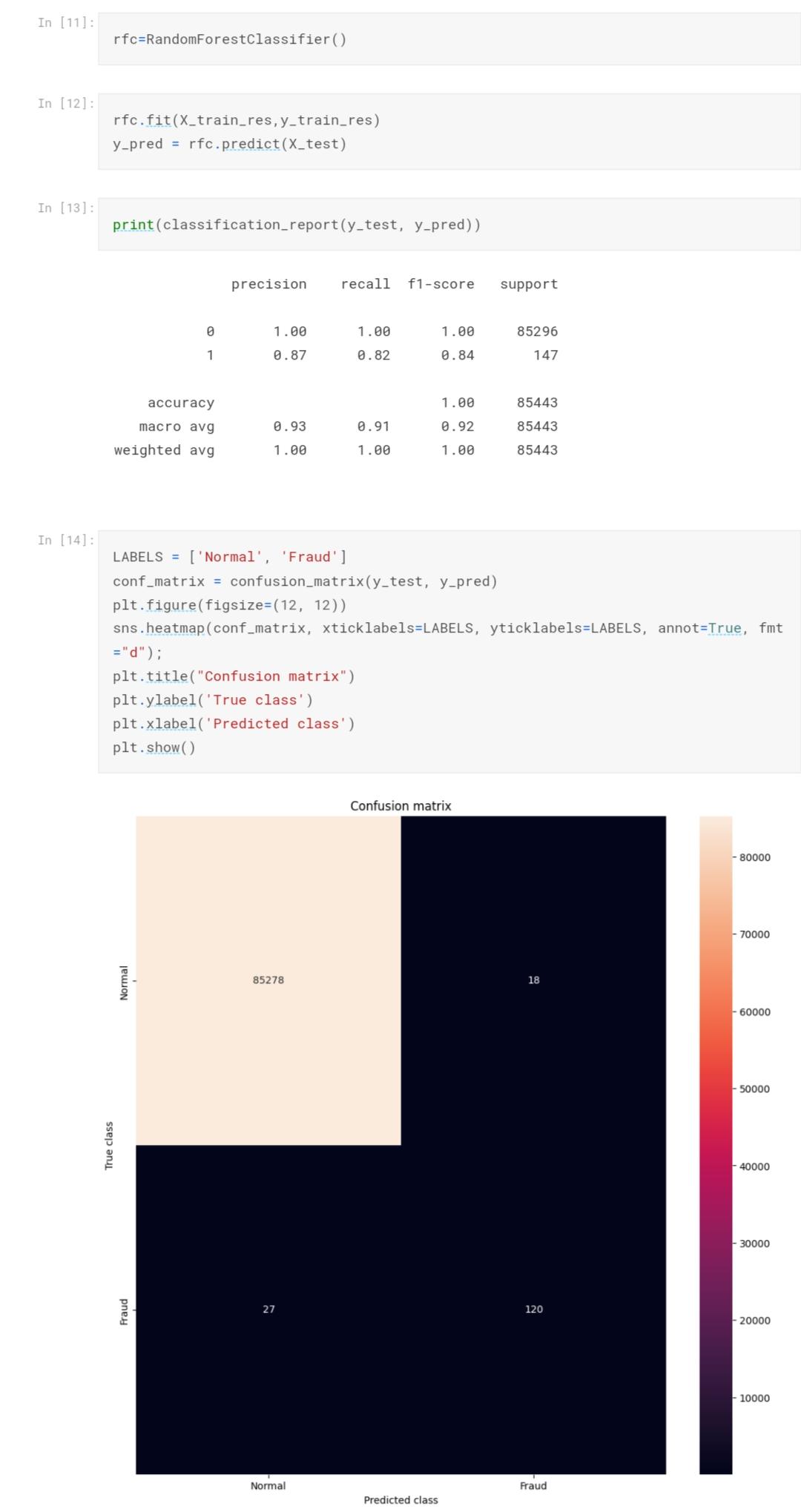






SMOTE:





\*\*Conclusion:\*\* Credit card fraud detection using data science is a critical application that can help financial institutions safeguard their customers' transactions. The success of such a project relies on robust data preprocessing, effective model selection, and ongoing monitoring to stay ahead of emerging fraud threats.