**Fraudulent transaction prediction**

**Business objectives:**

Objective of this project is to use python and machine learning to detect and prevent fraudulent transactions in online platforms. This will help reduce the financial losses, enhance the customer trust, and improve the security of online transactions.

**Problem definition:**

Develop a system that can automatically detect and prevent fraudulent transactions using machine learning techniques. The problem is to use python and transaction data to build and evaluate predictive models that can classify transactions as fraudulent or not.

**Implementation process:**

* **Import and explore the dataset:** here we will import our various machine learning libraries, load our data and explore statistical pattern in our data
* **Data Cleaning and Data Preparation:** here we will remove irrelevant features, handling missing values, and dealing with outliers
* **Exploratory Data Analysis**: here we Dive deep into the transactions by **analyzing spending patterns, pinpoint outliers, check correlations between features.**
* **Feature Engineering:** involves creating new features from existing ones to improve model performance.
* **Model Building evaluation:**  various machine learning algorithms are trained on the preprocessed dataset. The models’ performances are then evaluated using appropriate metrics

**Reference**

**Problem statement:** Devtown

**Data:** https://drive.google.com/uc?export=download&confirm=6gh6&id=1VNpyNkGxHdskfdTNRSjjyNa5qC9u0JyV

**GitHub:** **https: https://github.com/armandjucelin/fraudulent-transaction-prediction**

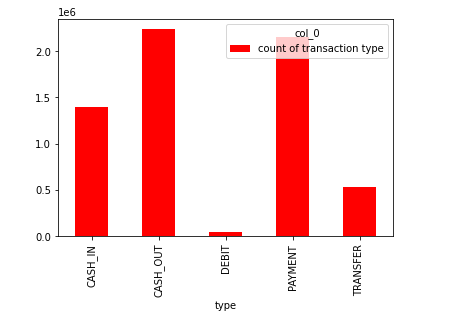
Tech Stack Used:

* Python,
* Machine Learning

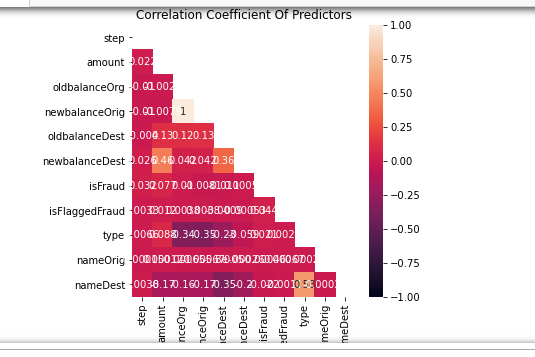
**Results**

**Correlation matrix:**

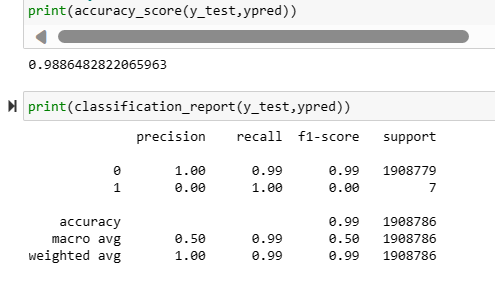
* CASH\_OUT and PAYMENT have the most transaction type made while DEBIT comes with the least transaction made



* the oldbalanceOrg and newbalanceOrg are highly related
* Namedest and type are related
* key factors that predict fraudulent customer are isFraud and amount



**Evaluation:**

* The accuracy of the isolation forest algorithm was approximately 0.98, showing that this model is good. 

**Application**

* **Banking and Finance**: Fraudulent transaction detection is crucial in the banking sector to identify and prevent credit card fraud, identity theft, and money laundering. It helps banks save millions of dollars and maintain customer trust.
* **E-commerce**: Online retailers can use it to detect and prevent fraudulent purchases, reducing losses and improving customer experience. It can also help in identifying return frauds and fake reviews.
* **Insurance**: It can be used to detect fraudulent claims, which are a significant issue in the insurance industry. This leads to reduced operational costs and premiums for customers.

**Strategy for future action**

* **Enhance Data Collection**: Collect more diverse and comprehensive data to improve the model’s ability to generalize. This could include transaction details, user behavior data, and historical fraud data.
* **Improve Feature Engineering**: Continually refine and create new features that can better capture the patterns associated with fraudulent transactions. This could involve using advanced techniques like deep learning for automatic feature extraction.
* **Real-time Fraud Detection**: Develop a system for real-time fraud detection. This would allow for immediate action when a fraudulent transaction is detected, further minimizing potential losses.