Anomaly Detection in banking dataset comprising a series of mobile money transactions.

1.About FinHack:

- FinHack is UTD's JSOM Lab's 2-day annual flagship event.
- During this event, participants tackle a problem statement by leveraging their expertise in Finance and Analytics to devise solutions.

2.Objectives:

- To perform exploratory analysis on dataset comprising of mobile money transactions to detect anomalies.
- To develop models using Machine learning algorithms to predict the anomalies in the dataset.

3.Data:

- Data consists of 101613 records of 11 variables.

4.Data Description:

The dataset contains 11 variables as shown below:

Column	Description
Step	Unit of time
type	CASH-IN, CASH-OUT, DEBIT, PAYMENT, and TRANSFER.
amount	Transaction amount in local currency
nameOrig	Transaction originator
oldbalanceOrg	Initial balance (before transaction)
newbalanceOrig	New balance (after transaction)
nameDest	Transaction recipient
oldbalanceDest	Initial balance before transaction
newbalanceDest	New balance after transaction
isFraud	Fraud agent takes control of customers' accounts and attempts to empty it by transferring to another account and then cashing out
isFlaggedFraud	An illegal attempt to transfer massive amount of money in a single transaction

5.Data Analysis:

- Utilized Python's pandas to remove missing values and duplicates from the dataset.
- Created a correlation matrix plot to identify relationships among variables.

6.Classification Models:

- Partitioned the data into 75%-25%
- Developed models using KNN Classifier, Logistic Regression, and Decision Trees algorithms.

7. Results:

- From correlation matrix, we can say that variables newbalanceDest and oldbalanceDest are positively correlated.
- The models achieved an AUC of 0.94 and accuracy of 99%.