README: Blockchain Fraud Prediction using ML

CSCI E-118 Project by Kartik Srikumar

1) Installations:

   a. Pandas:
   Run the following in terminal: pip install pandas
   More info: https://pandas.pydata.org/pandas-docs/stable/getting_started/install.html

   b. Numpy:
   Run the following in terminal: pip install numpy
   More info: https://numpy.org/install/

   c. Sklearn:
   Run the following in terminal: pip install -U scikit-learn
   More info: https://scikit-learn.org/stable/install.html

   d. Seaborn:
   Run the following in terminal: pip install seaborn
   More info: https://seaborn.pydata.org/installing.html

   e. Matplotlib:
   Run the following in terminal: python -m pip install -U matplotlib
   More info: https://matplotlib.org/stable/users/installing.html

2) How to run code:

   a. Open the .ipynb file in Jupyter Lab
   b. Ensure the dataset (transaction_dataset.csv file) submitted as part of the project is in the current directory.
   c. Run all cells from the top

3) Demonstration:

   a. All the cells in the notebook have markdown cells preceding them, explaining the flow of the code, and navigating the reader through the pipeline.
   b. Additionally, there are comments within the code describing the logic and flow.

c. <u>NOTE</u>: The actual demonstration portion of this project will include a PowerPoint(.pptx) as well as a link to a video presentation on youtube.com.

4) References:

a. https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html?highlight=decision%20tree#sklearn.tree.DecisionTreeClassifier

b. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html?highlight=logistic%20regression#sklearn.linear_model.LogisticRegression

c. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html?highlight=random%20forest#sklearn.ensemble.RandomForestClassifier

d. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html?highlight=gradient%20boost#sklearn.ensemble.GradientBoostingClassifier

e. https://scikit-learn.org/stable/modules/model_evaluation.html#classification-metrics

f. https://www.kaggle.com/vagifa/ethereum-frauddetection-dataset