Project 4 (Team 3)

Title: Stock Price Prediction using Machine Learning/Neural Networks

Members: *Salvador Alvarenga, Mohammad Lu6or, Mohamed Akl*

Steps:

1- **Selecting a dataset (Data Collection)**: Picking 3 Stocks (e.g, AAPL, MSFT, AMZON)
2- **Data Cleaning/Processing**: Cleaning data and prepping it for analysis using various tools
3- **Setting up back-end:** SQL Database, executing Machine learning model
4- **Front-end:** Deploying our tool/website

Assumptions:

- Stock data, crypto data, sentiment data for 2 years (2023 & 2024)
- Assume data is close to 100% accurate
- Target Variable being predicted:
    - Adjus. Close Price
    - Overall behavior/patterns/trends for the stock

Research Questions:

1- What features are needed to predict "historical" stock price with "high accuracy"?
2- How close our Machine Learning model to actual behavior of the stock (i.e., KPI: accuracy, confusion matrix, mean square error (MSE))

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

3- What is the Visualization of our prediction's vs actual performance of a stock?
4- How accurate is the model to predict the "future price" of the stock (e.g., Feb 2024)?

Final Production (for now):

- A Jupyter notebook with our analysis and code, and graphs
- If possible, deploying our Machine learning model on a local host website (maybe create an API with our model on databricks to the website)

Tools/Libraries:

- Python Pandas, Python Matplotlib, SQL Database, scikit learn, Long Short-Term Memory (LSTM) neural network model, Keras, Tensor flow, Python API, FLASK, Numpy