Project Title: Stock Market Analysis and Forecasting Using Deep Learning
Student: Emily Liang
Course: COE 379L

Introduction and Problem Statement
I always felt that the stock market is a complex and dynamic system influenced by numerous factors, making accurate predictions difficult. Traditional methods often cannot capture the intricate patterns and non-linear relationships that are inherent in financial data. Therefore, I believe that with deep learning, it can be leveraged with these advanced algorithms to analyze historical stock data and be able to accurately predict future trends better.

This project aims to utilize deep learning techniques, specifically Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, to analyze and predict stock prices. By focusing on historical data from major companies, the goal is to develop models that can accurately predict what the stock will do.

Data Source

I will be using historical stock data from the following companies:
- Google: 2006–2018
- Microsoft: 2006–2018
- IBM: 2006–2018
- Amazon: 2006–2018

These will be the base datasets, I will try to do as many companies as I can. However, those are the ones I've found good datasets on. Other companies I would like to add:
- Facebook (Meta)
- Apple
- Netflix
- Tesla

Each dataset includes the data with the following attributes:
- Open Price
- High Price
- Low Price
- Close Price
- Volume

These datasets are publicly available and can be obtained from financial data providers. I found two good ones on Kaggle. I'll need to look into each one more since they are very large datasets.

Methods, Techniques, and Technologies

- **Data Preprocessing and Analysis:**
    - Handling missing values and outliers
    - Normalization or standardization of data
    - Exploratory Data Analysis (EDA) to understand trends and patterns
    - Visualization of stock price movements and volume
- **Model Development:**
    - Implementation of CNN/ RNN and LSTM models using Keras (I want to learn more about LSTM, since it seems like the best for time series forecasting.)
    - Training models on historical stock data
    - Evaluate the prediction

And just for fun, using the prediction and comparing it with the market of that day.

- **Technologies and Tools:**
    - Programming Language: Python
    - Libraries: Pandas, NumPy, Matplotlib, Keras (Maybe PyTorch, I see a lot of resources online)
- **Development Environment**: Jupyter Notebook

Project Deliverables

- **Final Report (PDF):**
    - Comprehensive documentation of the project, including methodology, results, and conclusions
    - Discussion on the effectiveness of deep learning models in stock price prediction
- **Code:**
    - Notebooks containing all the code used in the project
    - Instructions for replicating the analysis and models
- **Visualizations:**
    - Graphs and plots illustrating stock price trends and model predictions
- **Presentation:**
    - Slides
    - Video presentation

References:
https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/
https://www.kaggle.com/datasets/jacksoncrow/stock-market-dataset
https://www.kaggle.com/datasets/paultimothymooney/stock-market-data