**Project Proposal**

Project team: Wen-Chih Li, Jianqiao Liu

1. **Introduction**

In the real world, there are plenty of researchers working through predicting the value of each company in the stock market. However, it is still an undefined topic, my team would like to use our knowledge that both learning in class and do some research outside the class to challenge this topic.

1. **Method**

**2.1 Data Visualization**

For data visualization, we will first management the data which means that we will reorder our data such that make the visualization clearer and more reasonable. For instance, we will use four datasets (Google, Apple, Amazon, Microsoft) to find out their relationship. We can plot the data by finding their correlation. Based on the correlation, we can make a theory, like the curve of "Apple" closing price has the same trend as "Google" closing price so that we can make a prediction of either "Google" closing price or “Apple" closing price by combining these two datasets and takes average. Also, we can compare the daily percentage return of two stocks to check how correlated. For example, we'll use "joinplot" to compare the daily returns of "Google" and "Microsoft". If two stocks are perfectly (and positively) correlated with each other a linear relationship between their daily return values should occur.

**2.2 LSTM**

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture [1] used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can process not only single data points (such as images), but also entire sequences of data (such as speech or video). For example, LSTM is applicable to tasks such as unsegmented, connected handwriting recognition,[2] speech recognition [3][4], and anomaly detection in network traffic or IDSs (intrusion detection systems). In short word, a LSTM model is one that you can use to make predictions on new data. That is, given new examples of input data, you want to use the model to predict the expected output. This may be a classification (assign a label) or a regression (a real value). Also, LSTM, short for Long Short-term Memory, is an extremely powerful algorithm for time series. It can capture historical trend patterns and predict future values with high accuracy. [5]

1. **Data**

Data source: 1. <https://finance.yahoo.com/quote/GOOG/history/>

2. <https://www.kaggle.com/tarunpaparaju/apple-aapl-historical-stock-data>

3. <https://www.kaggle.com/varpit94/amazon-stock-data?select=AMZN.csv>

4. <https://www.kaggle.com/vijayvvenkitesh/microsoft-stock-time-series-analysis>

We will use the dataset gain from “Kaggle.com” and “finance.yahoo.com/” which is Amazon, Apple, Google, and Microsoft historical trading data for each day.

1. **Relevant background work**
2. **Tentative plan**

**References:**

[1] Sepp Hochreiter; Jürgen Schmidhuber (1997). "Long short-term memory". Neural Computation. 9 (8): 1735–1780. doi:10.1162/neco.1997.9.8.1735. PMID 9377276. S2CID 1915014.

[2] Graves, A.; Liwicki, M.; Fernandez, S.; Bertolami, R.; Bunke, H.; Schmidhuber, J. (2009). "A Novel Connectionist System for Improved Unconstrained Handwriting Recognition" (PDF). IEEE Transactions on Pattern Analysis and Machine Intelligence. 31 (5): 855–868. CiteSeerX 10.1.1.139.4502. doi:10.1109/tpami.2008.137. PMID 19299860. S2CID 14635907.

[3] Sak, Hasim; Senior, Andrew; Beaufays, Francoise (2014). "Long Short-Term Memory recurrent neural network architectures for large scale acoustic modeling" (PDF). Archived from the original (PDF) on 2018-04-24.

[4] Li, Xiangang; Wu, Xihong (2014-10-15). "Constructing Long Short-Term Memory based Deep Recurrent Neural Networks for Large Vocabulary Speech Recognition". arXiv:1410.4281 [cs.CL].

[5] Li, K. (yi). (2021, July 29). Predicting stock prices using machine learning. Neptune.Ai. <https://neptune.ai/blog/predicting-stock-prices-using-machine-learning>

Resource:

<https://neptune.ai/blog/predicting-stock-prices-using-machine-learning>

https://www.kaggle.com/faressayah/stock-market-analysis-prediction-using-lstm

<https://machinelearningmastery.com/make-predictions-long-short-term-memory-models-keras/>

https://keras.io/api/layers/recurrent\_layers/lstm/