Jiaming Shen
Qiasheng Zou

CS446 Project Proposal

Background:
   With the advancement of modern technology in the past decade, computer-driven trading algorithms have become a dominant force in global financial markets. The presence of High-frequency trading (HFT) has accounted for a majority of the trading volume across US.[1] In the past a few years, ideas of applying machine learning algorithms to making trading decisions have gained some popularity. Machine learning models like neural networks have more expressive power and therefore, in some cases, they might perform better than the traditional time series models.

Project idea:
   For this project, we want to utilize the pattern recognition advantage of neural networks to conduct technical analysis[2] on the stocks listed in the S&P500 index. Furthermore, we would like to build an automated-trading system that gives out buying, selling and holding signals based on the output of our neural networks. Users of the system can then choose to follow or ignore the signals generated by the trading system. If the users choose to follow the signals and take on a position in the market, the system will then determine the stop and exit signals based on users' risk preferences.

Machine Learning Component:

Neural Network One:
   MACD divergence pattern recognition.

---
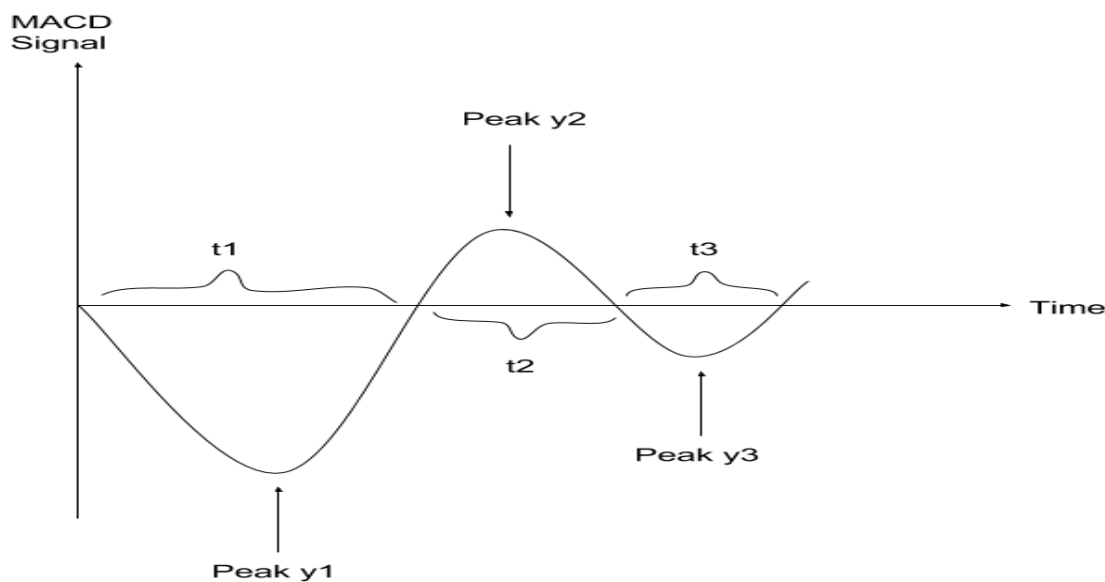
[1] https://en.wikipedia.org/wiki/High-frequency_trading

[2] In finance, technical analysis is a security analysis methodology for forecasting the direction of prices through the study of past market data, primarily price and volume.

Possible Features	
Feature 1	$\lvert y1\rvert/\lvert y2\rvert$
Feature 2	$\lvert y2\rvert/\lvert y3\rvert$
Feature 3	$\lvert y1\rvert/\lvert y3\rvert$
Feature 4	t1
Feature 5	t2
Feature 6	t3

Labels	
Label 1	neutral 0
Label 2	bullish divergence 1
Label 3	bearish divergence 2

Dataset:
1. Raw data: we will download original stock price data for stocks listed in S&P500 from Yahoo Finance or Quandl (an online financial database).
2. Training data: we will process the raw data to generate a series of MACD signal data for each stock. The data will be divided into many small samples and each sample be labeled as one of the three labels in the table shown above.

Neural Network Two:
Neural network classifier to generate buy/sell signals

Possible Features	
Feature 0	Output from MACD Pattern Recognition Classifier
Feature 1	Ratio of stock prices S1/S2
Feature 2	12-day stock price exponential moving average
Feature 3	26-day stock price exponential moving average
Feature 4	200-day long-term stock price simple moving average
Feature 5	Volume spike
Feature 6	Volume moving average
Feature 7	12-week stock price exponential moving average
Feature 8	26-week stock price exponential moving average
Feature 9	200-week long-term stock price simple moving average

Labels	
Label 1	Hold 0
Label 2	Long 1
Label 3	Short 2

Dataset:
1. Raw data: the same raw data we collected for neural network 1.
2. Training data: each training sample will have the features above, which can all be calculated from the raw data set. When we label the data, we will look at historical price patterns where MACD divergence has occurred and use the price movements afterwards as our labels. (For example, if the stock price moved upward after the MACD divergence pattern, we label it as a long; if the stock price moved downward we label it as a short)

System Design:

```
┌─────────────────────────┐        ┌──────────────┐        ┌──────────────┐
│   Online Data Source     │───────▶│  Clean and   │───────▶│   Database   │
│  (Yahoo Finance, etc)    │        │  Preprocess  │        │              │
│                          │        │     data     │        │              │
└─────────────────────────┘        └──────────────┘        └──────┬───────┘
                                                                   │
                                                                   ▼
                                                              ◇  input  ◇
                                                                   │
                                                                   ▼
                                        ┌────────────────────────────────┐   ┌──────────────┐
                                        │  MACD Divergence Pattern        │   │   Neural     │
                                        │  Recognition Classifier         │◀──│  Network 1   │
                                        └────────────────┬───────────────┘   └──────────────┘
                                                         │
                                                         ▼
                                                    ◇  input  ◇
                                                         │
                                                         ▼
┌──────────────────────────┐    ◇       ┌────────────────────────────────┐   ┌──────────────┐
│ Price levels, volumes,   │    input   │  Neural Network Classifier for  │   │   Neural     │
│ price moving averages,   │───────────▶│       trade entry signals       │◀──│  Network 2   │
│ etc..                    │    ◇       └────────────────┬───────────────┘   └──────────────┘
└──────────────────────────┘                            │
                                                         ▼
                                        ┌────────────────────────────────┐
                                        │   Human Intervention,           │
                                        │   manually take on trades       │
                                        └────────────────┬───────────────┘
                                                         │
                                                         ▼
                                        ┌────────────────────────────────┐
                                        │  Risk management system         │
                                        │  determines exit signal, stop   │
                                        │  signal.                        │
                                        └────────────────────────────────┘
```