**Brief Report**

Lue XIONG, Xinhui Gu, Tianyi Peng, Qingyao Sun

**1. Evaluation Metric**

The stock price data we used is MSTF price from 2014/3/27 to 2017/9/12. We split the data into three sets: training set from 2014/3/27 to 2016/8/29 with 611 observations; validation set from 2016/8/30 to 2017/5/9 with 173 observations; test set from 2017/5/10 to 2017/9/12 with 86 observations.

We would compare our model performance on the test set with a benchmark model and a naive model as detailed in section 2 and section 3 respectively. We would also evaluate our model performance with a p-value style metric on the test set. Concretely, we would count the number of correct stock price movement predictions produced by our model, and then calculate the probability that we observe this number or more correct predictions assuming the number of correct predictions follows a binomial distribution with n=86 and p=0.5.

**3. Naive Model**

In our training set, there are 310 up moves out of 611 observations. Therefore, under this naive model, we would randomly predict a up move on each date with probability 310/611=50.73% on our test set and then we would compare our prediction with the actual labels of the test set. It turns out that our naive model made 45 correct predictions on test set, with probability of correct prediction being 52.32%.