A Comparative Study on Trend Forecasting Approach for Stock Price Time Series

Wei Li[1*], Jian Liao[2]
[1]China Tobacco ZheJiang Industrial CO. LTD, Hangzhou 310000, China
[2]School of Computer Science Fudan University, Shanghai 200433, China
lwei@zjtobacco.com; jliao@fudan.edu.cn

***Abstract*—Trend forecasting is considered a difficult task, especially for China stock market due to its highly uncertainty. The study compares six forecasting models, i.e., Support Vector Machine (SVM), Naive Bayes, Decision Tree, Multilayer perceptron (MLP), Recurrent neural network (RNN), and Long Short-Term Memory (LSTM). 9 features combinations are selected based on 23 technical indicators which are commonly used in stock market analysis, and trainsets of 12 different records numbers are chosen to compare the performance of the models under different scenarios. Evaluation is carried out on 8 years of historical data from 2008 to 2015 of the listed company (000592) in China stock market. Experimental results show that the performance of deep learning models MLP, RNN, LSTM is better than other models with respect to the index of accuracy. MLP is 20.75% higher than Decision Tree, Decision Tree is better than others under f-measure, and Decision Tree is 40.02% higher than Naive Bayes. Experimental results also show that in the imbalanced stock market data, the performance of models RNN and Decision Tree is better than others.**

Keywords-Stock market; Trend forecasting; Deep learning

I. Introduction

Due to the complex, evolutionary and nonlinear dynamic financial market, predicting stock price trend and stock price indices is difficult. There are many complicated factors that affect the stock market, such as the business cycle, interest rates, regulatory policies, political situation and so on.

We find that stock time series prediction can be divided into two types: one is the value forecast; the other is the trend forecast, that the up and down of the stock price movement, the latter considered as a classification task, in which the classes can be divided into two (up and down) or three (up, down and flat) classes. Classification models, including Support Vector Machine (SVM), Naive Bayes, Decision Tree, Neural Network and so on, can be selected for the prediction[1][2][3].

A Decision Tree is a decision supporting tool that uses a tree-like graph or a model of decisions and their possible outcomes, and decision tree algorithm is also used for stock trend prediction[4][5][6]. Nair used a Decision Tree for stock market trend prediction. Their work designed a method of evaluating the performance of a hybrid Decision Tree rough set based system for predicting the next-day trend. The result showed that the performance of the system was better than that of the ANN and Naive Bayes[4]. Wang and Chan used a two-layer bias Decision Tree to reduce the classification error and to improve purchasing accuracy[5].

Artificial Neural Networks(ANN) are inspired by functioning of biological neural networks, and they are a dense network of inter-connected neurons which get activated based on inputs. Since the 1990s, ANN was trying to forecast stock price and stock trend[7][8]. The study of Nair also designed a novel method of predicting the next day's closing price of a stock market, they built a system based on adaptive ANN to finish this task. The system used the genetic algorithms to adapt itself dynamically, which tuned the parameters of the neural network at the end of each trading session[9].

Recently, deep learning techniques has received widespread attention in the field of machine learning and pattern recognition. Deep learning technology is based on the neural network, and deep learning models include Multilayer perceptron (MLP), Deep Belief Network (DBN), Recurrent neural network (RNN), Long Short-Term Memory (LSTM), etc….

A new method based on mixture of MLP experts was presented by Ebrahimpour. Three neural network combining methods and an Adaptive Network-Based Fuzzy Inference System are applied to trend forecasting[10]. A study had shown that four predictive models (MLP, ANN, Logistic Regression (LR), and Bagging of Logistic Regression (BLR)) had been developed for classification task in predicting the direction of movement in ISEN data, which showed that MLP was significantly better than the other classifiers[10]. Hsieh presented an integrated system, which combined recurrent neural network based on artificial bee colony algorithm and wavelet transforms for stock price forecasting[11]. A study by Yoshihara proposed an approach to market trend prediction based on a recurrent deep neural network to model temporal effects of past events[12]. A paper used LSTM to predict China stock returns by Chen transformed the historical data of China stock market into 30-day-long sequences with 10 learning features and 3-day earning rate labeling. The study demonstrated the power of LSTM in stock market prediction in China, it also showed that China stock market was unpredictable[14].

The conventional method is to analyze the stock market via technical indicators, while it has been shown that news events influence the trends of stock price movements. Duan had explored a prediction approach based on the behavior of analysts' recommendations[15]. Some researchers showed that predicting the event-driven stock market trend better using deep learning models[16][17].

Although, the deep learning shows excellent results in a lot of areas, there is little research on deep learning for trend forecasting of stock price. This study pays attention to compare prediction performance of SVM, Naive Bayes, Decision Tree,

* Corresponding author

978-1-5386-0533-2/17/$31.00 ©2017 IEEE

MLP, RNN, LSTM algorithms for the trend forecasting of stock time series. Our purpose is to compare the performance of shallow machine learning (SVM, Naive Bayes and Decision Tree) and deep learning (MLP, RNN, LSTM). We design nine features combinations, which are used as the input to these models. Our trend prediction categories fall into three classes, namely up, down and flat. The data we used came from the listed company (code 000592) in China stock market. The experiments are carried out using eight years of historical data, which consists of a large number of different price moving patterns.

The remainder of the paper is organized as follows: Section 2 describes research data, the pre-processing of the data, the design of nine features combinations, and the categories that we want to forecast; Section 3 describes the models that we use in this study; Section 4 shows the experimental results; and we discuss the results and conclude the study in Section 5.

II. DATASET

This study uses the eight years of data that come from the listed company (code 000592) from May 2008 to Dec 2015. The data contains 23 technical indicators, such as opening price, rate of return, etc., which are presented in TABLE I. (1-23). We normalize the 23 technical indicators using (1) to reduce the scale difference between the dimensions before the data are organized into a dataset. The entire dataset is shown in TABLE I. , including 1401 cases. TABLE II. shows the proportions of up, down and flat in the entire dataset. The categories of up, down and flat can be calculated from the technical indicators of opening price and closing price. The opening price is shown in Fig. 1, which has small spreads with closing price.

$$x = \frac{x - min}{max - min} \quad (1)$$

The categories we used include up, down and flat, which are derived from the opening price and closing price via (2). OP_i is the opening price of the i-th day, and CP_i is the closing price of the i-th day. If $c > 0$, the value of class is up; $c = 0$, the value is flat; $c < 0$, the value is down. In addition, we map the values of three classes to numbers, that is, (up, down, flat) are mapped into (0, 1, 2).

This study uses 8000 trainsets. In order to evaluate the influence of time period we choose on the prediction performance of the models, we experiment with 12 different ti-

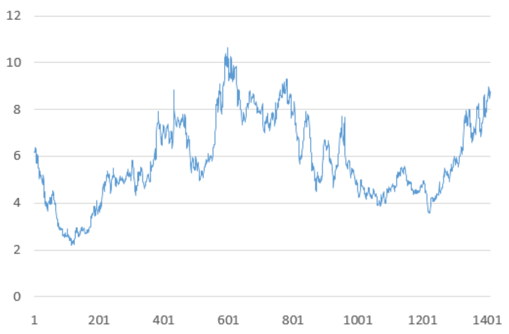

Figure 1. The trend of opening price

TABLE I. THE ENTIRE DATASET, 1-23 ARE TECHNICAL INDICATORS, 24 IS CATEGORIES. MA IS THE MOVING AVERAGE.

SN	Technical Indicator
1	5MA-10MA
2	10MA-20MA
3	5MA-20MA
4	10MA-20MA
5	The ratio of the lower shallow to the K line
6	The ratio of the rod to the K line
7	The ratio of the upper shallow to the K line
8	Closing price – 5MA
9	Minimum price – 5MA
10	Maximum price – 5MA
11	Opening price – 5MA
12	Rate of return – Opening price
13	Rate of return – Maximum price
14	Rate of return – Minimum price
15	Rate of return – Closing price
16	$(turnover_{t+1}-turnover_t)/ turnover_t$
17	$(5MA_{t+1}-5MA_t)/ 5MA_t$
18	$(10MA_{t+1}-10MA_t)/ 10MA_t$
19	$(20MA_{t+1}-20MA_t)/ 20MA_t$
20	Opening price
21	Maximum price
22	Minimum price
23	Closing price
24	Categories(up, down, flat)

TABLE II. THE NUMBER OF UP, DOWN AND FLAT CASES PROPORTION IN THE ENTIRE DATASET

Up	%	Down	%	Flat	%
465	33.2	733	52.3	203	14.5

$$c = \frac{OP_{i+1} - CP_i}{CP_i} \quad (2)$$

TABLE III. THE 12 TIME PERIOD, SN IS THE SERIAL NUMBER, THE PERIOD IS THE NUMBER OF THE 12 TIME PERIOD

SN	1	2	3	4	5	6
Period	30	50	100	200	300	400
SN	7	8	9	10	11	12
Period	500	600	700	800	900	1000

me periods shown in TABLE III. for each of which we repeatedly apply the forecasting method for 400 times on a moving-window basis. In addition, we also compare the relationship between different features combinations and the performance of models. We designed 9 features combinations, and the specific features show in TABLE IV.

The time series trainsets include a lot of patterns so the result of the comparative study is effective. Fig. 2 shows some examples of different time series patterns, and TABLE V. shows the proportions among the three categories. The patterns are the trend of opening price time series, and they are different. For example, the trend in Fig. 2(a) is that it is a decline for the beginning and then a rise, the trend in Fig. 2(b) is always a decline. There are four simple patterns in the example, and more complex patterns exist in the dataset.

We find that the values of categories are imbalanced, and in trainsets this phenomenon also exists which we will discuss this issue later. In a nutshell, the technical indicators are normalized to reduce the scale difference of dimensions. Then we designed difference features combinations as the input to the models SVM, Naive Bayes, Decision Tree, MLP, RNN, and LSTM. The categories in the entire dataset and the trainsets are imbalanced.

III. PREDICTION METHOD AND MODELS

A. Prediction method

We use six models to predict the trend of stock price, and the process of forecasting is shown in Fig. 3. Under a certain condition, we build training set and test set as input of the model, training set for model training and test set for model testing, and then we get the test result from the output of the model, which is used to calculate the value of indicators to evaluate the performance of the model.

The input of the models are the combinations of periods (TABLE III.) and 9 features combinations (TABLE IV.). We evaluate the influence of this two parameters on prediction of performance in the experiment. First parameter period is the number of records of the training set, and with the days of data we predict the next day's trend. Second parameter features combinations is formed in combination with different technical indicators. The training sets are dynamically constructed based on these two parameters.

In other words, the trainsets are constructed through the interval $[t-trl, t]$ in the features combinations, trl representing the number of records of the trainset, which is also the time period of data we used. We predict the trend of day $t+1$, which means the test set is one-day data, that is, the number of records of test set is 1.

TABLE IV. THE SPECIFIC FEATURES, THE TECHNOCAL INDECATORS IN FEATURES COMBINATIONS

SN	Features combinations
1	15
2	15, 5, 7
3	15, 1, 2, 3, 4
4	15, 8, 9, 10, 11
5	15, 16
6	15, 17, 18, 19
7	15, 5, 7, 1, 2, 3, 4
8	15, 5, 7, 16
9	15, 5, 7, 8, 9, 10, 11

TABLE V. THE PERPORTION OF THE THREE CATEGORIES

SN	Up	%	Down	%	Flat	%
a	137	34.3	202	50.4	61	15.3
b	57	28.5	114	57.0	29	14.5
c	30	30.0	61	61.0	9	9.0
d	125	41.7	126	42.0	49	16.3

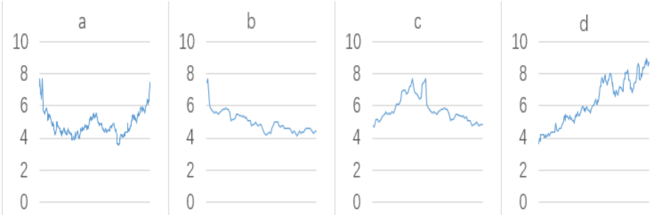

Figure 2. Different time series patterns

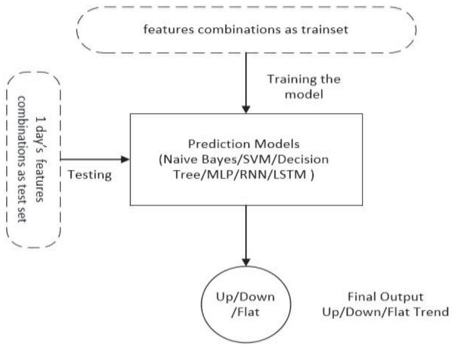

Figure 3. The process of trend forecasting

In the experiment, we constructed 8000 training sets. Out of the 1401 records, we choose the former 1000 to build the training sets, and the remaining to build the test set except the record 1401, which means that we performed 400 test experiments in each group. We use a moving-window basis to build the input. In a certain parameter, we build the input to forecast the trend of day $t+1$, and then we rebuild the input to predict the trend of day $t+2$, continuing the above until record 1400 as the test set.

B. Prediction models

1) Naive-Bayes classifier

Naive Bayes classifier is based on applying Bayes' theorem with strong independence assumptions among the features. Naive Bayes classifier assumes the attribute conditional independence; for a given category, assume that all attributes are independent of each other. Bayesian classifier classifies the test data into a class with the highest probability.

2) SVM

SVM is a learning method that uses high dimensional feature space. If the original space is finite, then there must be a high-dimensional feature space so that the sample space can be divided. The purpose of SVM find the hyperplane with maximum margin.

In general, regularization parameter c controls the trade-off between margin and misclassification error, and in our experiment c is 1.0. The kernel function we use is radial basis function (rbf). We have to deal with a three classification problem, and SVM in dealing with multi-classification problems, the commonly used methods are one-versus-rest and one-versus-one. In our models, we use one-versus-one SVM.

3) Decision tree

Decision tree is a common type of machine learning method. A Decision Tree consists of a root node, a number of internal nodes, and leaf nodes. Decision tree has three learning algorithms, ID3, C4.5 and CART. CART (Classification and Regression Tree) is a well-known Decision Tree learning algorithm, available for classification and regression tasks. CART uses Gini index to select the partition attribute.

4) MLP

Multilayer perceptron is a feedforward artificial neural network model, which maps sets of input data onto a set of appropriate outputs. MLP is composed of multiple nodes in the directed graph, and the layers are fully connected. Expect the input nodes, each node is a neuron with a nonlinear function, such as relu, sigmod. It is difficult to determine the parameters of the MLP model, the number of layers, the number of nodes per layer, the size of batch size, and the number of iterations, the activation function, the optimizer and so on. The model of we build, parameters show in the TABLE VI. .

5) RNN

Recurrent neural network (RNN) is a class of artificial neural network where connections between units form a directed cycle. This creates an internal state of the network which allows it to exhibit dynamic temporal behavior. The RNN model is embodied in the form that the network will remember the previous information and apply it to the current output calculation, that is, the nodes between the hidden layers are connected, and the hidden layer's input includes not only the output of the input layer but also the output of the hidden layer at the last time. We guess that there is a kind of relationship about time in our features. The parameters of the RNN model are batch size and epoch mainly. The parameters of RNN are shown in the TABLE VI. .

6) LSTM

Long short-term memory (LSTM) is a recurrent neural network (RNN) architecture. Unlike traditional RNN, LSTM is good at learning from experience to classify, predict time series, because LSTM unit is a recurrent network unit that excel at remembering values for either long or short durations of time. However, RNN does not have this function. This unit is the single memory block, which contains memory cells with self-connections storing the temporal state of the network in addition to special multiplicative units called gates to control the flow of information. LSTM block contain three gates, the input gate controls the flow of input activations into the memory cell. The output gate controls the output flow of cell activations into the rest of the network. Later, the forget gate was added to the memory block[18].The parameters of LSTM are the same with RNN's.

TABLE VI. THE PARAMETERS OF MLP AND RNN

MLP		RNN	
layers	5	batch size	32
nodes	15-10-5-4-3	epoch	20
batch	32	Dropout_W	0.2
epoch	20	dropout_U	0.2
activation	relu	optimizer	Adam
dropout	0.2	classifier	softmax
optimizer	RMSProp		
classifier	softmax		

IV. EXPERIMENTAL RESULTS

The models SVM, Naive Bayes, Decision Tree we used are from scikit-learn, and the models MLP, RNN, LSTM are built by Keras. We test the performance of the models using parameters according to Section 3, which are set with reference to Chen's paper. And the Accuracy and F-measure are used to evaluate the performance of the six models.

First we compare the experimental results under different combinations of features. We select the trainsets with time period 50. The values of the two performance indexes under each of the features combinations are shown in TABLE VII. . The features combinations 4,7,9 are better than others according to TABLE VII. The performance of models MLP and Decision Tree are better than the others. The accuracy of MLP is 20.75%, 12.28%, 13.77% higher than Decision Tree in features combinations 4,7,9. Decision Tree is better under f-measure, which is 32.19%, 44.02%, 37.71% higher than Naive Bayes, which Naive Bayes is the worst under f-measure. Also, it can be seen from the results that the technical indicator 15 is useful, which can be obtained from the features combination 1. The accuracy and f-measure of the features combinations 1 have little difference with the others.

Second, we compare the effects of training sets of different time period. We select the features combination 9 as the best features combination, because the average value of accuracy of the six models are 44.8%, 44.8%, 45.8% in features combinations 4,7,9. And we present the performance in the test of the time period 50, 100, 200, 300 in TABLE VIII. . The performance of other time periods are very similar to the time period 300, because the accuracy of model Naive Bayes, SVM MLP, LSTM is 45.8% which is a bad result we will explain later. The time period 50,100,200 is better than others. The performance of models MLP and LSTM are better under accuracy, and the Decision Tree is better under f-measure. Accuracy of MLP is 18.6% higher than Decision Tree in time period 200, and LSTM is 28.6% higher than Decision Tree in

TABLE VII. THE SIX MODELSPPERFORMANCE ON FEATURES COMBINATIONS

	DET	NB	SVM	MLP	RNN	LSTM
the feature conbinations 1						
Acc	0.395	0.450	0.450	0.460	0.470	0.455
Fmea	0.323	0.264	0.287	0.323	0.301	0.298
the feature conbinations 4						
Acc	0.398	0.453	0.438	0.480	0.458	0.460
Fmea	0.342	0.259	0.282	0.346	0.316	0.307
the feature conbinations 7						
Acc	0.428	0.465	0.455	0.480	0.435	0.425
Fmea	0.384	0.267	0.293	0.325	0.324	0.283
the feature conbinations 9						
Acc	0.418	0.465	0.458	0.475	0.460	0.473
Fmea	0.371	0.269	0.295	0.334	0.336	0.321

TABLE VIII. THE SIX MODELS PERFORMANCE ON DIFFERENT TIME PERIOD

	DET	NB	SVM	MLP	RNN	LSTM
the length:50						
Acc	0.408	0.450	0.458	0.458	0.415	0.473
Fmea	0.368	0.289	0.295	0.329	0.307	0.321
the length:100						
Acc	0.368	0.465	0.460	0.450	0.465	0.473
Fmea	0.321	0.269	0.257	0.282	0.335	0.312
the length:200						
Acc	0.390	0.460	0.465	0.463	0.455	0.448
Fmea	0.332	0.236	0.234	0.258	0.311	0.238
the length:300						
Acc	0.430	0.458	0.458	0.458	0.450	0.458
Fmea	0.378	0.209	0.209	0.209	0.281	0.213

TABLE IX. THE RESULT BASED ON TIME PERIOD 50 AND FEATURES COMBINATIONS 9

Models	Acc	F-mea			F-mea-avg
		Up	*Down*	*Flat*	
DET	0.408	0.385	0.495	0.225	0.368
NB	0.450	0.283	0.584	0.000	0.289
SVM	0.458	0.292	0.592	0.000	0.295
MLP	0.458	0.359	0.573	0.055	0.329
RNN	0.415	0.342	0.528	0.050	0.307
LSTM	0.473	0.372	0.592	0.000	0.321

time period 100. F-measure of Decision Tree is 41.8% higher than SVM in time period 200. Against our wish, as the time period of trainset increases, the models MLP, Naive Bayes, SVM, and LSTM classifies all test samples into the category down. According to statistics of the trainsets, the category flat is fewer as the time period increases in the imbalanced trainsets.

TABLE IX. shows the experimental results in the case of time period 50 and features combinations 9 in detail. It also shows that the three categories up, down, flat of f-measure, which we see f-measure of Naive Bayes, SVM, LSTM models for category flat is 0 , and both the precision and recall values which constitute f-measure in the models are 0 in TABLE IX. . And the performance of RNN and Decision Tree is better than others in the imbalanced train set, which are these two models that can classify test cast into category flat.

As a matter of fact, the values of accuracy and f-measure as a whole are low. The reason, we think, is that the features combinations contain too similar values with different categories to make the models hard to finish the classification task. And with few records of trainset to predict the trend better, few records of dataset is more likely to show data trend possibly. Although our entire study shows that the accuracy and f-measure of all models is relatively low for stock price trend prediction in China, it also shows that the China stock market is much more unpredictable.

V. CONCLUSIONS

In this paper, we design 9 features combinations, and design 12 time period of trainsets. We compare the performance of models SVM, Naive Bayes, Decision Tree, MLP, RNN, and LSTM. Experimental results show that the performance of deep learning modes MLP, RNN, LSTM is better than other models under accuracy, and in the imbalanced stock market data, Decision Tree is better than others under f-measure. The results also show that in the imbalanced stock market data, the performance of models RNN and Decision Tree is better than others.

REFERENCES

[1] W.~Huang, Y.~Nakamori, S.-Y. "Wang, Forecasting stock market movement direction with support vector machine," Computers& Operations Research 32~(10), pp. 2513-2522, 2015.

[2] M.~R. Hassan, B.~Nath, M.~Kirley, "A fusion model of hmm, ann and ga for stock market forecasting," Expert Systems with Applications 33~(1) , pp. 171-180, 2007.

[3] J.~Patel, S.~Shah, P.~Thakkar, K.~Kotecha, "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques," Expert Systems with Applications 42~(1), pp. 259-268, 2015.

[4] B.~B. Nair, V.~P. Mohandas, N.~R. Sakthivel, "A decision tree- rough set hybrid system for stock market trend prediction," International Journal of Computer Applications 6~(9), pp. 1-6, 2010.

[5] J.~L. Wang, S.~H. Chan, "Stock market trading rule discovery using two-layer bias decision tree," Expert Systems with Applications 30~(4) pp. 605-611, 2006.

[6] R.~K. Lai, C.~Y. Fan, W.~H. Huang, P.~C. Chang, "Evolving and clustering fuzzy decision tree for financial time series data forecasting," Expert Systems with Applications An International Journal 36~(2), pp. 3761-3773, 2009.

[7] J.~H. Wang, J.~Y. Leu, "Stock market trend prediction using arima-based neural networks," IEEE International Conference on Neural Networks, pp. 2160—2165,1996.

[8] E.~W. Saad, D.~V. Prokhorov, "Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks," IEEE Transactions on Neural Networks 9~(6), pp. 1456-1470, 1998.

[9] B.~B. Nair, S.~G. Sai, A.~N. Naveen, A.~Lakshmi, G.~S. Venkatesh, V.~P. Mohandas, "A GA-Artificial Neural Network Hybrid System for Financial Time Series Forecasting," Springer Berlin Heidelberg, 2011.

[10] R.~Ebrahimpour, H.~Nikoo, S.~Masoudnia, M.~R. Yousefi, M.~S. Ghaemi, "Mixture of mlp-experts for trend forecasting of time series: A case study of the tehran stock exchange," International Journal of Forecasting 27~(3), pp. 804-816, 2011.

[11] P.~Tüfekci, "Classification-based prediction models for stock price index movement," Intelligent Data Analysis 20~(2), pp. 357-376, 2016.

[12] T.~J. Hsieh, H.~F. Hsiao, W.~C. Yeh, "Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm," Applied Soft Computing 11~(2), pp. 2510-2525, 2011.

[13] A.~Yoshihara, K.~Fujikawa, K.~Seki, K.~Uehara, "Predicting Stock Market Trends by Recurrent Deep Neural Networks", Springer International Publishing, 2014.

[14] K.~Chen, Y.~Zhou, F.~Dai, "A lstm-based method for stock returns prediction: A case study of china stock market," IEEE International Conference on Big Data, pp. 2823-2824, 2015.

[15] J.~Duan, H.~Liu, J.~Zeng, "Posterior probability model for stock return prediction based on analyst's recommendation behavior," Knowledge-Based Systems 50, pp. 151-158, 2013.

[16] X.~Ding, Y.~Zhang, T.~Liu, J.~Duan, "Deep learning for event-driven stock prediction," International Conference on Artificial Intelligence, pp. 2327-2333, 2015.

[17] R.~Akita, A.~Yoshihara, T.~Matsubara, K.~Uehara, "Deep learning for stock prediction using numerical and textual information," Ieee/acis International Conference on Computer and Information Science, pp. 1-6, 2016.

[18] H.~Sak, A.~Senior, F.~Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," Computer Science, pp. 338-342, 2014.