Stock Prediction based on Bayesian-LSTM

Biao Huang, Qiao Ding, Guozi Sun, Huakang Li*
Jiangsu Key Lab of Big Data and Security and Intelligent Processing
Nanjing University of Posts and Telecommunications, Nanjing, 210023, China
Collaborative Innovation Center for Economics crime investigation and prevention technology, Jiangxi, China
State Key Laboratory of Mathematical Engineering and Advanced Computing, Wuxi, 214125, China
huakanglee@163.com

ABSTRACT
Fluctuations in stock market represent the changes in national economic objectively. Some machine learning algorithms, such as linear fitting and sequence mining, are employed to predict the stock market. However, linear fitting faces issue of over-fitting and black relationships with historical data, while sequence mining is short in efficiency and lack dynamic adaptations. This paper builds a modified Bayesian-LSTM (B-LSTM) model for stock prediction. Six indicators of the Chinese stock market in every day are the basic input for LSTM. In order to represent the economic wave, we defined a data set unit by week which means the basic unit in LSTM is data in one week. Furthermore, a Bayesian optimization model is proposed to estimate the unit number dynamically in different economic cycle. The experimental results demonstrated that the B-LSTM increase over 25% prediction than the conventional LSTM.

CCS Concepts
• Computing methodologies→Neural networks.

Keywords
Stock market; B-LSTM; data set unit; Stock prediction; Bayesian optimization

1. INTRODUCTION
Since stock market is known to be of both high risk and profit, it has seized great amount of attention of investors. There are many stock data generated worldwide. Some investors and investment institutions make analysis and prediction for more profits, mainly through history stock data. It is proved that the stock trend is predictable in a short period [15]. Some machine learning algorithms, such as structural learning [24] and neural network [19], have been employed to predict stock market automatically. Although they satisfy some investment goals to some extends, due to the complexity and fuzziness in mechanism, the influence of conventional machine learning algorithms is limited, which cannot reflect the dynamic changes in stock market.

There are two characteristics that stock market presents [21][16], the changes in both long and short time and the variance in periods. If research simply applies neural network algorithm on history stock data, the effectiveness is limited. Such issue can be categorized into 4 parts. First, the parameters in training period is all static. The parameters throughout the whole process keep unchanged, which cannot make adjustments to cater to the dynamic changed in real stock market. Second, the principle in splitting training and testing dataset is out-of-date. Some researches simply spilt raw dataset with regard to percentage, which neglects the variances and periods in stock market. Third, using single-dim input for training. There are many standards in stock data and they are interacted. For instance, for close prediction, one cannot singly use history close data for input. Some standards, such as open, volume should also be included. Last, a large range of companies should be covered. There maybe relationships between companies. Also, a single company cannot reflect the whole market environment.

In this paper, the upper four issues are deeply discussed and successfully solved. Specifically, issues about dynamic parameters are emphasized. This paper introduces a novel B-LSTM (Bayesian-LSTM) model, with LSTM to be main part and Bayesian Optimization for dynamic parameter selection. B-LSTM involves discussion in influences that changes in window size *U* brings about towards prediction in a period *P*. The optimized number of units, U_{opt}, in *P* is determined by Bayesian Optimization. Finally, by training on the time sequence $U_{opt1}, U_{opt2}, \ldots, U_{optn}$, future optimized window size is dynamically selected, and can be taken into prediction model. The split of raw dataset is determined by window size *U*. For training set, *U* units are to be input. The $U+1^{th}$ unit are to be output. The evaluation criterion is accuracy(Acc). In this paper, number of input dim is set to be 6, including open, close, max, min, volume and value. Output dim is close only.

According to related socioeconomic concepts, including psychology of stock market and economic reflection in stock market, the trend presented in $\mathrm{U_{opt}}$ sequence is suited to characteristics introduced in socioeconomic fields. What's more, contrary is revealed both between established methods and B-LSTM and between different industry sectors. With regard to comparison results, the significant effectiveness of B-LSTM is pointed out.

Section 2 presents related works. Conventional LSTM is introduced in section 3. In section 4, methodology of B-LSTM is presented in detail. Experiment and results are demonstrated in section 5. Section 6 draws conclusion.

ICMLC 2018, February 26–28, 2018, Macau, China.

ACM ISBN 978-1-4503-6353-2/18/02...$15.00.
DOI: https://doi.org/10.1145/3195106.3195170

2. RELATED WORK

2.1 Machine Learning in Stock Prediction

Due to the noise and volatile features in stock market, the prediction to stock market is considered to be the most challenged time sequence problem [20]. In 1996, Yaser S. Bbu-Mostafa [1] make a brand new introduction to finance problems through research on commodity feature and foreign exchange. Yaser claims that an effective way to reduce noise is to enlarge size of dataset. However, owing to financial data is always non-stationary, comparing to old data, new data can be totally different. Two aspects are discussed.

2.1.1 Linear Algorithms

For linear models in business fields, the most popular methods are linear fitting and linear regression. Sanjiban Sekhar Roy and Dishant Mittal discussed linear regression in stock market in 2015 [6]. In their research, they build a LASSO linear regression model, which is proved better than ridge linear regression model. What's more, YE Cakra and BD Trisedya put forward linear regression model based on sentiment analysis and finally achieves better performance.

Adaline, short for adaptive linear element, one of neural networks in linear expression, is introduced in 1950s. E. Schöneburg, proposed its application in stock market in 1990 [8]. In his research, both adaline and madaline are added into comparison for real stock dataset. In all, although linear regression a fitting achieves effectiveness to some extends, the performance is still limited. They neglect the interference of history data, so the results are insulating in time.

2.1.2 Sequence Mining Algorithms

There are generally four sequence mining algorithms, AprioriAll, GSP, Freespan and Prefixspan. AprioriAll and GSP are Apriori algorithms while Freesapan and Frefixspan are pattern growth algorithms. For business applications, YP Wu and KP Wu analyzed performance of AprioriAll in stock market and verifies its long-term effects [23]. J Han takes a deep look at Freespan model and proves that the efficiency of Freespan is better than Apriori [15].

In the field of recurrent neural network, LSTM is special RNN with feedback links connecting different dims. LSTM can solve vanishing gradient problem through memory units which retain information of an arbitrary amount of time [18]. Some researches [3][11] have been used to verify LSTM is more effective than conventional recurrent network. It is an algorithm, that is widely accepted, for filtering single-dimensional issues [13][22][12]. It is used to reduce the noise in financial data and later feed back into deep learning framework.

2.2 Multiple Index Analysis

For simple prediction model, one input dimension is enough. However, in most cases, only one input dim cannot satisfy the demanding accuracy. In 2015, Kai Chen and Yi Zhou [2] have taken a close look at the effects of number of inputs on accuracy in stock prediction. From their work, one could safely find that the more input dimensions, the more accuracy will achieve. With all 5 principles, open, close, max, min and value, the accuracy can reach the peak. Also, the effects of input dims are presented by Ryo Akita [7], when using numerical and textual information for stock prediction. The result is similar, the more input dims, the more factors will be taken into consideration, which will finally lead to better outcomes. Therefore, in this paper, 6 dimensions are used for prediction.

2.3 Bayesian Optimization

The Bayesian optimization algorithm is a complete global optimization algorithm for noisy and expensive black-box functions [17][5]. Bayesian optimization algorithm mainly depends on probilistic model which defines distribution in objective function from input space to objective of interest.

The recent contribution to Bayesian optimization function includes elegant theoretical results [9], multitask and transfer optimization [4] and application to diverse tasks such as sensor set selection [10]. The performance of Bayesian optimization algorithm in parameter selection is outstanding.

3. B-LSTM MODEL

For LSTM, there are little parameter could be changed. Therefore, the reflection of dynamics could be determined by window size selection in a fixed time period. For instance, in a month, the number of days to be input to predict the next day. In an economic cycle, there are *n* months. Thus, a time sequence containing *n*-elements is established. A line chart reflecting monthly variation in an economic period is presented by plotting the time sequence. Such kind of dynamic changes in number of units in a fixed period can represent variation of stock market in a fixed period.

In this paper, seven attributes, time, open, max, min, close, volume and value, are covered in raw data. Time step is one day. According to socioeconomic concepts, the history data are split into n parts. Each part contains data from time period t_o to t_f. Take single part for study. In the targeted part, window w_d, which includes two parts of data, one from t_i to t_p, the other from t_p to t_j, is determined. Data length from t_p to t_j is fixed. The study lies on the determination of data length U from t_i to t_p to optimize the effectiveness when predicting data from t_p to t_j using data from t_i to t_p. Finally, through plotting optimized *U* sequence from all parts, we can make analysis and prediction to stock market trend.

3.1 Data Split

For more detailed processes, first, we split history stock data into *n* parts, $T = \{t_1, t_2, \dots, t_n\}$. Then, in each part, daily data are divided into units. Units are stored in an array $A = \{a_1, a_2, \dots, a_n\}$. In every unit, the daily data contains six attributes, o(open), max, min, c(close), v(volume) and V(value). When it comes to LSTM, the input dim is set to be 6(o, max, min, c, v and V). The input are data in *U* units. The output dim is 1, and the output is close in the next 1 unit.

There are five layers in LSTM, one for input layer, one for output layer and three for hidden layer. Because the data in single training is modicus, full-batch training is preferred for reduction in error. Training is repeated 10000 times. Every 1000 times, models and parameters are saved once. There are *n* units in one part. Set the first *m* units for training and the rest *(n-m)* units for testing. To the beginning, set current size to be *U*. Thus, the first *m* units are categorized into Set $U = \{a_1, a_2, \dots, a_u,\ a_{u+1};\ a_2, a_3, \dots, a_{u+!}, a_{u+2}; \dots; a_{m-1-u}, a_{m-u}, \dots, a_{m-1}, a_m\}$. There are *U+1* data in one category, where *U* data for input and 1 for output. The same in categorization for testing. The evaluation criterion is acc with predicted data and real data.

4. FINDING OPTIMIZED PARAMETER --- BAYESIAN OPTIMIZATION

The most important part in B-LSTM is to determine U_{opt} precisely and quickly. It is a waste of time to find optimized *U* with brute force. We cannot ensure the function form that *U* emerges, thus the monotonicity is uncertain. Bayesian optimization is one valid candidate solution for finding optimized *U* effectively.

Bayesian optimization is designed for functions without expressions. Despite the absense of expressions, for every input *x*, the function generates a output *y*. When a certain amount of data *(x, y)* is provided, through Bayesian optimiztion, the function form can be estimatied, so that the extremum is predicted. There are two goals of Bayesian optimization, one to learn the function form, the other to find the extremum. Bayesian can weigh pros and cons when facing with two conflicting targets and make adjustments automatically, in order to find the extremum. The next sampling point is chosen on the basis of forlulation as presented:

$$x_t = \underset{x \in D}{\operatorname{argmax}}\, \mu_{t-1}(x) + \beta_t^{\frac{1}{2}} \sigma_{t-1}(x)$$

In this paper, the main goal is to find the minimum value. The first step is to build regression model based on Gaussian process by using existing sample points. With Gaussian process, the model can predict the mean value $\mu_{t-1}(x)$ and standard deviation $\sigma_{t-1}(x)$ in unknown position. The next sampling point is determined by the position of the maximum sum of mean value $\mu_{t-1}(x)$ and standard deviation $\sigma_{t-1}(x)$. The sum function is called acquisition function with a wright in the function, $\beta_t^{1/2}$. For acquisition function, if the standard deviation is larger, it means the model knows little about the point. It can better understand the function form through multiple sampling. If the mean value is larger, it means that the point is supposed to be the minimum. We can locate the minimum faster by samples with large mean value. To the beginning, the model will choose points with largest standard deviation. When sampling reaches a certain amount, the standard deviation will be greatly reduced. Thus, sampling process will be influenced more by mean value and the sampling point will be closer to the real minimum value.

4.1 General Algorithm Design

Bayesian optimization are employed to determine the optimized *U* value. Through taking samples in different *U* value, Bayesian optimization can predict functions in shape and finally find out the value of *U* when it reaches the minimum. The whole model is presented in Fig 1:

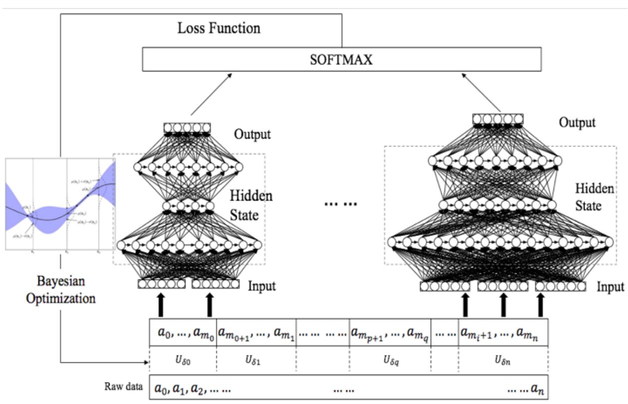

Figure 1. Model Implementation.

And the frame design is shown as follows:

Algorithm B-LSTM

---

input data S

output $U_{opt1}, U_{opt2}, \dots, U_{optn}$

step 1: Split S into *T* parts, $p_1, p_2, \dots, p_T$

step 2: Split p_i into *n* units, $a_1, a_2, \dots, a_u$

step 3: Bind *U+1* units as a sequence

step 4: Function Bayesian Optimization(acc)

step 5: Run LSTM, output acc sent to step 4

step 6: If there are parts left, run Step 2

End B-LSTM

In Bayesian optimization, with regard to statistical characteristics of selected samples, acquisition function is used to determine the optimized *U* dynamically. The optimized *U* is marked as U_{opt}. Since there are *n* parts, totally *n* iterations are required. In every iteration, U_{opt} is emerged. Once plotting the line chart of U_{opt} sequence from *n* parts, economical trend in stock market is presented intuitively. Furthermore, the trend is discussed by introduction to related socioeconomically concepts so that the reality and precision of B-LSTM is verified.

5. EXPERIMENT & RESULTS

5.1 Data

Shanghai Stock Exchange composite index and Shenzhen composite index from 1990 to 2016. Sector data including Coal, Security, Realty, Travel and Electrical components from 2007 to 2016. All data is abstracted as *S*.

5.2 Evaluation Criterion

The accuracy(Acc) is resented as $\mathrm{acc}(A, B) = \frac{|A \cap B|}{|A|}$.

5.3 Results

5.3.1 Prediction to Optimized U

According to calculations and comparisons of acc, B-LSTM figures out U_{opt} for every year based on selection from Bayesian optimization. For Shanghai Stock Exchange composite index(SHCI), the optimized number of week input U_{opt} sequence with time is presented as follows:

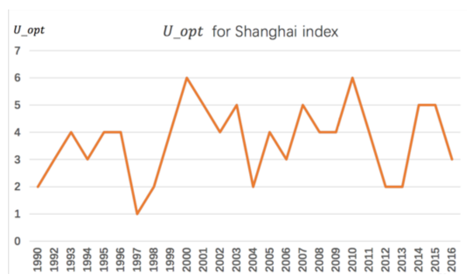

Figure 2. U_{opt} for SHCI.

For Shenzhen composite index (SZCI), the U_{opt} sequence can be plotted in Fig 3:

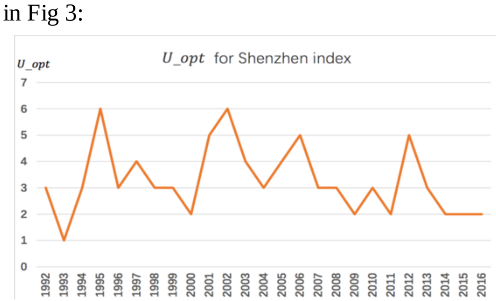

Figure 3. U_{opt} for SZCI.

Take Shanghai Stock Exchange composite index for example. As is shown in figure, the U_{opt} is a time sequence with a year as an interval. We can input them into conventional LSTM to predict U_{opt} for the future year, 2017. After training, the output of LSTM is 3, in other words, $U_{opt-2017} = 3$.

According to related socioeconomic concepts, within a given period, the stock market is constituted of a series of macro trends. As is show in figure above, from 1997 to 2012 is a sample macro trend. In a specific macro trend, each micro trends represent the psychological changes of investors. Take a period from 2004 to 2005 for instance. The stock market comes through a recovery period. However, with a desire to stability, some of investors choose to sell shares, which finally leads to a slip in a temporary time. Nonetheless, the macro trend is ascending, which reflects the economic growth of the whole nation. Analysis towards B-LSTM in stock market verifies the effectiveness in socioeconomic field. Same as Shenzhen composite index, the trend fits characteristics in socioeconomic theories.

5.3.2 Comparisons with Conventional LSTM

For conventional LSTM, time step U_c is set by average value in B-LSTM. After calculating average value of U_{opt} series in 26 years and rounding, U_c is set to be 4. Take $U_c = 4$, *input dim*=6(o, max, min, c, v, V) as parameter, run LSTM to predict in single year. For Shanghai Stock Exchange composite index (SHCI), the acc comparison is demonstrated as follows:

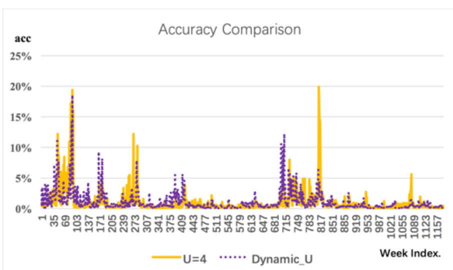

Figure 4. Accuracy Comparison.

Take week as unit to make prediction to close value. Since the data of weeks in 26 years is huge in size, we choose data from the first week in 1992 to the first week in 1997 for demonstration:

Figure 5. Comparison in Close of SHCI.

For Shenzhen composite index (SZCI), comparison between real data, B-LSTM and conventional LSTM is presented in figure 6:

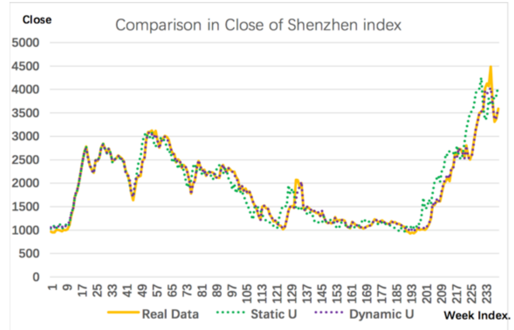

Figure 6. Comparison in Close of SZCI

Take Shanghai Stock Exchange composite index for study. From the graph, we can intuitively find that the results predicted by conventional LSTM, which is based on the fixed U, have disparities with those predicted by B-LSTM. For conventional LSTM, during prediction, it generates over-fitting results, thus the deviation to real data is significant. Therefore, conventional LSTM cannot fit in acute changes in dynamic stock market. When it comes to B-LSTM, it can quickly make adjustments to changes in market, so that difficulty in predicting policy oriented market is reduced.

5.3.3 Comparison of Different Methods in Prediction

What's more, static machine learning algorithm, linear regression, is also used for series prediction. Comparison between LSTM, B-LSTM and linear regression is demonstrated in the following table:

Table 1. Comparison of Different Methods in Prediction

	B-LSTM	**LSTM**	**Linear Regression**
	Dynamic U	Static U	
Shanghai	1.01%	1.24%	8.97%
Shenzhen	0.97%	1.44%	9.88%

The evaluation criterion is acc in percentage. In the table above, we select 10-year data from 2007 to 2016 for comparison in methodology. From table above, LSTM is much effective than linear regression. But for B-LSTM, as discussed in upper section, is more adaptable and precise than conventional LSTM.

5.3.4 Comparison in Different Section Data

Furthermore, industry sector data can represent changes in certain fields. Thus, we pick 5 different sectors which stand out in their fields for prediction. Both LSTM and B-LSTM is used for comparison. In Table 2, Stat stands for conventional LSTM, Dyn stands for B-LSTM and EI stands for Electronic Component.

Table 2. Comparison in Predicting Different Section Data

	2007		**2008**		**2009**		**2010**	
Sector (%)	**Stat**	**Dyn**	**Stat**	**Dyn**	**Stat**	**Dyn**	**Stat**	**Dyn**
Coal	1.21	1.17	1.25	1.11	1.88	1.88	0.97	0.70
Security	2.10	1.00	2.31	0.97	2.07	1.09	1.33	1.03
Realty	2.67	1.44	2.34	1.23	1.66	1.66	1.24	0.67
Travel	4.32	1.37	5.67	1.49	4.34	1.00	10.23	1.76
EI	2.58	0.33	3.21	0.88	2.65	1.04	7.44	1.16
	2011		**2012**		**2013**		**2014**	
Sector	**Sta**	**Dyn**	**Sta**	**Dyn**	**Sta**	**Dyn**	**Sta**	**Dyn**

	t		t		t		t	
Coal	2.12	1.77	1.33	1.09	1.55	1.55	2.21	1.56
Security	0.91	0.91	1.55	0.48	1.47	0.76	0.55	0.55
Realty	1.34	0.88	0.67	0.67	0.78	0.75	1.43	0.99
Travel	8.25	1.02	3.77	0.99	6.12	1.53	4.32	0.56
EI	2.01	0.34	4.01	1.34	7.12	1.98	3.99	0.86
	2015				**2016**			
Sector (%)	**Stat**		**Dyn**		**Stat**		**Dyn**	
Coal	1.45		0.66		1.89		0.77	
Security	0.98		0.77		0.67		0.63	
Realty	2.03		1.22		1.09		1.00	
Travel	2.45		1.24		3.00		0.76	
EI	3.00		1.01		4.01		1.00	

The evaluation criterion is acc in percentage. We select 10-year data from 2007 to 2016 in all five sectors. In experiment, it is interesting to find out that in both Travel and Electronic Components sector, the U_{opt} sequence changes more frequently, in the range [1, 3]. Such finding fit in the characteristics in the two industries, which is always known to sell fast and change rapidly. From table above, B-LSTM is more effective than conventional LSTM in Travel, Electronic Components and Security. However, such advantage appears obscure when predicting Coal and Realty. Such deviation is determined by characteristics of industries. Both Coal and Realty are conventional industries, there is little fluctuations. U_{opt} in B-LSTM is close to U_c . Therefore, in industries like Coal and Realty, the difference is vague. In conclusion, B-LSTM stands out.

6. CONCLUSION

This paper successfully builds Bayesian-LSTM for choosing adaptive cycle dynamically, and B-LSTM is proved to well suited economical concepts. In order to present the periodicity and relative concepts in socioeconomic fields, Bayesian optimization are able to find the optimized adaptive cycle, U_{opt}, in a certain economic wave. The optimized adaptive cycle in the future year is predicted through analyzing time sequence of U_{opt}.

This paper takes 26-year history data of Shanghai and Shenzhen composite index for experiment. First, we separate history data by year. In each year, data in one week are regarded as a unit. B-LSTM are applied to find and plot the time sequence of $\mathrm{U_{opt}}$. We find that the trend suits relative socioeconomic concepts and well represents the growth and decline in national stock market. Finally, we make comparisons between different algorithms and dataset, which reveals the effectiveness and strong adaptability.

For future work, despite the improvements comparing with conventional algorithms, B-LSTM still needs future modification for policy-oriented market. Some information in related fields, namely, textual information in newspaper, websites as well as blogs, should be added to make adjustments to LSTM. Therefore, evaluations and parameters should be added into neural network, or in more innovative way, add more interacting layers.

7. ACKNOWLEDGEMENT

This work was supported by the NSFC (No.61502247, 11501302, 61502243, 91646116), China Postdoctoral Science Foundation (No.2016M600434), Natural Science Foundation of Jiangsu Province (BK20140895, BK20150862), Scientific and Technological Support Project (Society) of Jiangsu Province (No.BE2016776), and Postdoctoral Science Foundation of Jiangsu Province(1601128B), Opening Project of Collaborative Innovation Center for Economics crime investigation and prevention technology(JXJZXTCX-015), and Open Project Program of the State Key Laboratory of Mathematical Engineering and Advanced Computing(2017A10).

8. REFERENCES

[1] Abu-Mostafa, Y. S., & Atiya, A. F. (1996). Introduction to financial forecasting. *Applied Intelligence*, 6(3), 205-213.

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

[3] Baccouche, M., Mamalet, F., Wolf, C., Garcia, C., & Baskurt, A. (2010). Action Classification in Soccer Videos with Long Short-Term Memory Recurrent Neural Networks. *Artificial Neural Networks* – ICANN 2010. Springer Berlin Heidelberg.

[4] Bardenet, R., Brendel, M., Kégl, B., & Sebag, M. (2013). Collaborative hyperparameter tuning. *International Conference on International Conference on Machine Learning* (Vol.28, pp.II-199). JMLR.org.

[5] Brochu, E., Brochu, T., & Freitas, N. D. (2010). A Bayesian interactive optimization approach to procedural animation design. *Eurographics/acm SIGGRAPH Symposium on Computer Animation, SCA 2010*, Madrid, Spain (pp.103-112). DBLP.

[6] Cakra, Y. E., & Trisedya, B. D. (2016). Stock price prediction using linear regression based on sentiment analysis. *International Conference on Advanced Computer Science and Information Systems* (pp.147-154). IEEE.

[7] Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. *IEEE International Conference on Big Data* (pp.2823-2824). IEEE.

[8] E. Schöneburg. (1990). Stock price prediction using neural networks: a project report. *Neurocomputing*, 2(1), 17-27.

[9] Freitas, N. D., Smola, A., & Zoghi, M. (2012). Exponential regret bounds for gaussian process bandits with deterministic observations. Icml.

[10] Garnett, R., Osborne, M. A., & Roberts, S. J. (2010). Bayesian optimization for sensor set selection. *ACM/IEEE International Conference on Information Processing in Sensor Networks* (pp.209-219). ACM

[11] Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: continual prediction with lstm. *Neural Computation*, 12(10), 2451.

[12] Graves, A. (2012). Supervised Sequence Labelling with Recurrent Neural Networks. Springer Berlin Heidelberg.

[13] Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., & Schmidhuber, J. (2015). Lstm: a search space odyssey. *IEEE Transactions on Neural Networks & Learning Systems*, PP(99), 1-11.

[14] Han, J., Pei, J., Hsu, M. C., Hsu, M. C., Hsu, M. C., & Hsu, M. C. (2000). FreeSpan: frequent pattern-projected sequential pattern mining. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp.355-359). ACM.

[15] Kohara K, Ishikawa T, Fukuhara Y, et al. Stock Price Prediction Using Prior Knowledge and Neural Networks. *Intelligent Systems in Accounting Finance & Management*, 2015, 6(1):11-22.

[16] Levine, Ross, Zervos, & Sara. (2013). Stock market development and long-run growth. Policy Research Working Paper, 10(2), 323-339.

[17] Lizotte, D. J. (2008). Practical bayesian optimization. University of Alberta.

[18] Osborne, M., & Briscoe, T. (1997). Learning stochastic categorial grammars. Conll, 80--87.

[19] Tkáč, M., & Verner, R. (2016). Artificial neural networks in business: two decades of research. *Applied Soft Computing*, 38, 788-804.

[20] Wang B, Huang H, Wang X. A novel text mining approach to financial time series forecasting. *Neurocomputing*. 2012; 83(6):136–45.

[21] Wei, Y., & Huang, D. S. (2003). A study on the multifractal feature of chinese stock market. Systems Engineering.

[22] Werbos, P. J. (1990). *Backpropagation through time: what it does and how to do it. Proceedings of the IEEE*, 78(10), 1550-1560.

[23] Wu, Y. P., Wu, K. P., & Lee, H. M. (2013). Stock Trend Prediction by Sequential Chart Pattern via K-Means and AprioriAll Algorithm. *Technologies and Applications of Artificial Intelligence* (Vol.30, pp.176-181). IEEE.

[24] Zhang, Y. P., Zhang, L., Wu, T., Xu, F., Zhang, M., & Wang, L. W. (2004). Structural learning algorithm based on covering algorithm and its application in stock forecasting. *Journal of Computer Research & Development.*