Deep Learning with Stock Indicators and Two-Dimensional Principal Component Analysis for Closing Price Prediction System

Tingwei Gao[1,2,3],Xiu Li[1,2],Yueting Chai[1,3],Youhua Tang[1,2]
1 Department of Automation, Tsinghua University, Beijing 100084,China
2 Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055,China
3 National Engineering Laboratory for E-Commerce Technologies,Beijing,100084,China

Abstract—**The stock market is an important component in the current economic market. And stock price prediction has recently garnered significant interest among investment brokers , individual investors and researchers. In general, stock market is very complex nonlinear dynamic system. Accordingly, accurate prediction of stock market is a very challenging task,owing to the inherent noisy environment and high volatility related to outside factors .In this paper , we focus on deep learning method to achieve high precision in stock market forecast. And a deep belief networks(DBNs) ,which is a kind of deep learning algorithm model,coupled with stock technical indicators(STIs) and two-dimensional principal component analysis($(2D)^2$PCA) is introduced as a novel approach to predict the closing price of stock market. A comparison experiment is also performed to evaluate this model.**

Keywords-Stock Prediction, Deep Learning, Technical Indicators, 2D Principal Component Analysis

I. INTRODUCTION

As the rapid development of economy and technology, the stock market has become an important part of our daily life. Accordingly, the stock trend predict has greatly been one of the focuses of public topic. And the stock market forecast is the act of trying to learn the future value and trend of stock market.

In recent years, a large number of research applying machine learning algorithms for forecasting the stock market has been developed. It contains many artificial neural networks (ANNs) approaches. And ANNs often combine with other approaches to predict the stock market.Two neural network architectures and generalized regression neural networks was proposed to predict the stock price (Mostafa 2010)[1]. A model combined wavelet transform and ANNs was applied to forecast the stock price(S. Kumar Chandar, M. Sumathi and S. N. Sivanandam 2016)[2].Particle swarm optimization of ensemble neural networks with fuzzy aggregation for prediction of the stock exchange(Martha Pulido, Patricia Melin, Oscar Castillo 2014)[3]. ANNs and support vector machines(SVMs) were used to predict stock movement(Kara, Boyacioglu, and Baykan 2011)[4].The ANNs integrated with an improved bacterial chemotaxis optimization have been used to predict the stock indices (Zhang &Wu, 2009)[5].A hybrid neural networks and multiple regression model was designed for time series forecasting(Alexey Averkin, Sergey Yarushev, Igor Dolgy and Andrey Sukhanov2016)[6].

In this paper, we propose a novel prediction method based on DBNs, STIs and $(2D)^2$PCA.This model performs very well in practical prediction by experiment. In summary, this system encompasses the following characteristics:There are strong correlation between the inputs and output;The deep learning algorithm has been proved to have excellent performance for prediction and analysis;The dimensionality reduction and feature extract approach has good performance.

The rest of this paper is structured as follows. Section II presents the related work and background knowledge.In Section III, we introduce our deep learning architecture for closing price prediction. And section IV gives the comparison experimental results and the analysis of results..

II. RELATED WORK

A. Stock Prediction Methodologies

In general, there are two main stock market prediction method. One is fundamental analysis(FA) and the other is technical analysis(TA).FA is concerned more with the company rather than the stock price and volume data. And the analysts and investors predict the stock trend mainly based on the performance of the company, political events and stock market news.In contrast to FA,TA focuses on the price data. TA often applies machine learning to achieve stock trend forecast.

B. Technical Indicators

STIs have been proved to have abundant and latent information about stock market. And a lot of stock indicators have good performance for analysis the stock.For example, the relative strength index(RSI) is a well-known indicator and it is used to signal overbought and oversold conditions in a security. And the moving average convergence divergence(MACD) is often used to reveal changes in the strength direction and duration of a trend.

C. Deep Learning Algorithms

Deep learning is a branch of machine learning. It attempts

978-1-4673-9904-3/16/$31.00 ©2016 IEEE

to model high-level abstractions in data by multiple neural layers.And there are various deep learning models such as DBNs, convolutional deep neural networks and recurrent neural networks .A large number of deep learning models have been applied to many fields and produce more and more state-of-the-art results.The key idea of deep learning algorithm comprises: unsupervised learning mode for pre-train; train the layer one by one ,and the train result will be the input of next layer; adjust all layers by supervised mode.

D. *Dimensionality Reduction*

In both machine learning and statistics, dimensionality reduction is critical step for data pre-process and analysis.And it is the process of reducing the number of input variables. In general, this process can be divided into feature selection and feature extraction. By this process, we can map the data from a high dimensional space to a lower dimensional space. Some well-known dimensionality reduction algorithms comprise principal component analysis(PCA),linear discriminant analysis(LDA),locally linear embedding(LLE) and Laplacian Eigenmaps.

III. PREDICTION MODEL

A. *DBNs Architecture and Training*

DBNs is a generative graphical model, or a type of deep neural networks, composed of multiple layers of hidden units, with connections between the layers but not between units within each layer.DBNs can be viewed as a composition of several restricted boltzmann machines(RBMs).

RBM is a generative stochastic artificial neural network. It consist of two layers.One is visible layer and the other one is hidden layer. Units in each layer have no connections between themselves. While they are connected to all units of other layer. Connections between the neurons are bidirectional. A graphical depiction of an RBM is shown in Fig. 1.

Let V denotes visible layer and H denotes hidden layer. Then we can define the energy function E(V,H) as below, where b is the value of bias,w is the value of weight.

$$E(V,H)=-\sum_i\sum_j\sum_k W_{ij}^k h_j v_i^k-\sum_i\sum_k v_i^k b_i^k-\sum_j h_j b_j \tag{1}$$

The joint over the visible and hidden variables is given by the Gibbs Distribution.

$$P(V,H)=\prod_u P(V_u,H_u)=\prod_u\frac{\exp(-E(V_u,H_u))}{\sum_u \exp((-E(V_u,H_u))} \tag{2}$$

The conditional $P^{\theta}(v_i|H)$ take the form of softmax function,where K is the maximum rating value.

$$p(v_i^k=1|h)=\frac{\exp(b_i^k+\sum_j h_j W_{ij}^k)}{\sum_l \exp(b_i^l+\sum_j h_j W_{ij}^l)} \tag{3}$$

Given the observed ratings V, we can make a prediction of

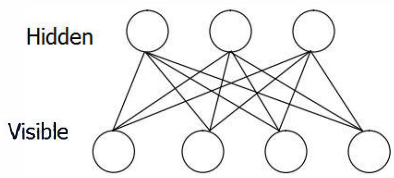

Fig. 1 Single RBM structure

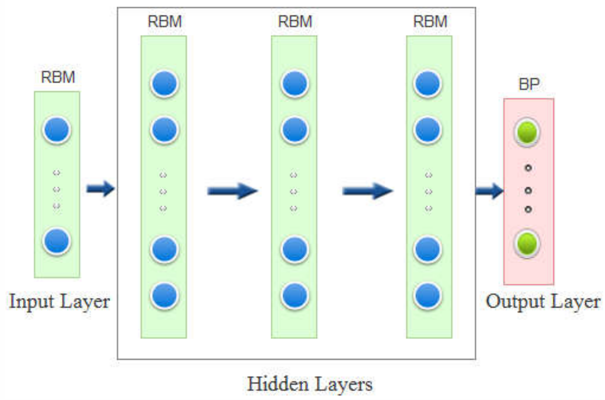

Fig. 2 DBNs architecture

rating for a new query movie q in time linear in the number of hidden units:

$$p(v_q^k=1|V)=\exp(v_q^k b_q^k)\prod_j(1+\exp(\sum_{il} v_i^l W_{ij}^l+v_q^k W_{qj}^k+b_j)) \tag{4}$$

In our model,DBNs consists of four RBMs layers and a back propagation (BP) layer.The framework of the DBNs is illustrate as Fig. 2. BP comprises two processes—propagation and weight update. And the propagation involves two steps—forward propagation generates output activations, and backward propagation generates the deltas. The weight update gradient descent.

And the training consists of two steps, one is pre-training and another is fine-tuning. The pre-training step is unsupervised layer-wise training, which aims to minimize the cost of each layer.Fine-tuning is to learn and tune the parameters by minimizing the cost by back-propagation. And fine-tuning is a process of supervised learning,

B. *Input Variables Selection*

Here,the inputs of the networks comprise two sets—basic inputs and indicators inputs.

There is open price(Op),high price(Hp),low price (Lp),close price(Cp),adjusted price(Ad) and volume(Vo) in basic inputs.

The indicators inputs consist of some well-known stock market indicators. Technical indicators are effective tools to characterize the real market situation. A total of ten ones are selected considered as input variables. And there is Chaikin volatility(CV), Williams%R(W%R), typical price(TP), Stochastic oscillator(SO), weighted close(WC), price rate of change(PROC), bolume rate of change(VROC), volume- Price

TABLE 1: Selected technical indicators

Indicator Name	formula
MACD	EMA(close，12）-EMA（close，26）
CV	(EMA [H-L] - EMA [H-L_{t-10}]) / EMA [H- L_{t-10}] * 100
W%R	(Hi_n-Cl_t)/(Hi_n-Lo_n)*100
TP	(Hi+Lo+Cl)/3
SO	(Cl-Lo_{min})/(Hi_{max}-Lo_{min})*100
ADO	((Cl_{max}-Lo_{max})-(Hi_{max}-Cl_{max}))/(Hi_{max}-Lo_{max})*Sum(Vo)
PROC	(Cl-Cl_{t-12})/Cl_{t-12}
VROC	(Vo-Vo_{t-12})/Vo_{t-12}
VPT	VPT_{t-1}+Vo*(Cl-Cl_{t-1})/Cl_{t-1}
WC	(Hi+Lo+Cl*2)/4
RSI	100-100/(1+Sum(Up)/Sum(Down))
OSCP	MA_5-MA_{10}/MA_5

trend(PVT), relative strength index(RSI),accumulation distribution oscillator(ADO),price oscillator(OSCP),\and moving average convergence divergence(MACD).

To explain the technical indicators ,we introduce some basic concepts and formulas-- exponential moving average (EMA), Moving Average(MA).

$$MA(Cl_N) = (Cl_1 + Cl_2 + ... + Cl_N)/N \tag{5}$$

$$EMA_t = \alpha Cl_t + (1-\alpha)EMA_{t-1} \tag{6}$$

C. $(2D)^2$ PCA

Let X denote an n-dimensional unitary column vector to project an m × n matrix A. And Y is the projected feature vector of A.

$$Y = AX \tag{7}$$

We can define the total scatter of the projected samples..

$$J(X) = tr(X^T E((A-EA)^T (A-EA))X) \tag{8}$$

The covariance matrix S_t can be denoted as below.

$$S_t = \frac{1}{M}\sum_{i=1}^{M}(A_i - \overline{A})(A_i - \overline{A})^T \tag{9}$$

And we can get the eigenvalue λ_i of A_i and eigenvector v_i of S_t. X_{opt} is composedby the orthonormal eigenvectors $X_1,\ldots,X_d$ of S_t corresponding to the d largest eigenvalue.

We can let Ai to be denoted by an n-dimensional column vector.

$$A_i = [A_i^{(1)}, A_i^{(2)}, \ldots, A_i^{(m)}] \tag{10}$$

$$S_t = \frac{1}{M}\sum_{i=1}^{m}\sum_{j=1}^{n}(A_i^{(j)} - \overline{A^{(j)}})(A_i^{(j)} - \overline{A^{(j)}})^T \tag{11}$$

Z is an m×q matrix with orthonormal columns.

$$J(Z) = tr(Z^T E((A-EA)^T (A-EA))Z) \tag{12}$$

TABLE 2: Stock data for experiment

	Op	Hi	Lo	Cl	Vo	Ad
2004-01-02	1111.92	1118.85	1105.08	1108.48	1153200000	1108.48
2004-01-05	1108.48	1122.22	1108.48	1122.22	1578200000	1122.22
2004-01-06	1122.22	1124.46	1118.44	1123.67	1494500000	1123.67
…	…	…	…	…	…	…
2016-04-13	2065.92	2083.18	2065.92	2082.42	4191830000	2082.42
2016-04-14	2082.89	2087.84	2078.13	2082.78	3765870000	2082.78
2016-04-15	2083.10	2083.22	2076.31	2080.73	3701450000	2080.73

Fig. 3 Chart of stock data for experiment

The optimal projection matrix Z_{opt} can be obtained by computing the eigenvectors Z1,…,Zq corresponding to the q largest eigenvalues.

Now we can get the projection matrices X and Z, , yielding a q× d matrix C, which is our target martrix .

$$C = Z^T AX. \tag{13}$$

IV. EXPERIMENTAL RESULTS

A. Data Description

Experiments were performed with the real-word data of Standard & Poor's (S&P500) to validate and compare the effectiveness of the model.And we get the data from the Yahoo Finance by R.The total number of samples is 3,093 trading days, from January 2004 to April 2016.Each sample consisted of daily information including low price, high price, opening price, closing price, adjusted price and trading volume.And the dataset is divided into two parts: training and testing.Table 2 and Fig 3 show the dataset.

B. Comparison Experiments and Performance Evaluation

We compare the predict performance of three models.

Model A:The inputs just comprise the basic variables.And the networks is the common BP neural networks(BPNNs).

Model B: The inputs comprises both basic variables and STIs.And the netwoks is the DBNs.

Model C: DBNs with basic inputs ,STIs and $(2D)^2$PCA.

Fig. 4 Experiment result of Model A

Fig. 5 Experiment result of Model B

Fig. 6 Experiment result of Model C

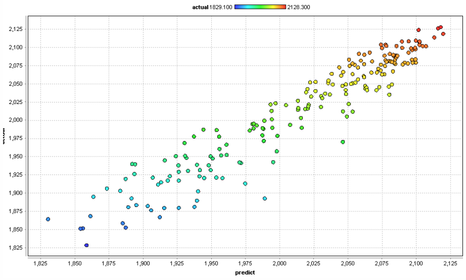

Fig. 7 Experiment result of Model A

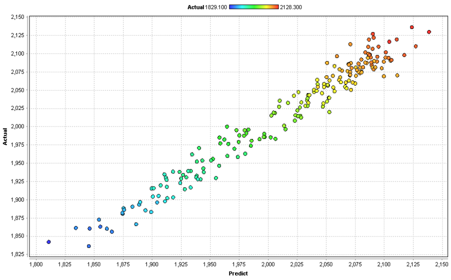

Fig. 8 Experiment result of Model B

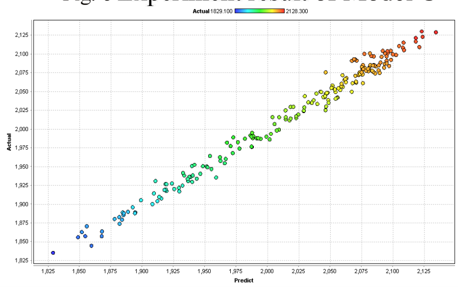

Fig. 9 Experiment result of Model

TABLE 3: Experiment Results

	MAE	RMSE
Model A	17;934	22.867
Model B	12.046	15.548
Model C	10.192	12.918

The test data comprise stock data of 200 days.To evaluate the performance,we conducted a series of experiments and the results were measured by mean absolute error (MAE) and root mean square error(RMSE).Also we plot the charts of results,Please refer to Fig 4-Fig 9.

In statistics, MAE is given by:

$$MAE = \frac{\sum_{i=1}^{N} |p_i - q_i|}{N} \tag{14}$$

And RMSE is computed by:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - q_i)^2} \tag{15}$$

By experiment, we can find that the prediction accuracy of the DBNs is better than the prediction errors of BPNNs.And the (2D)2PCA has improved the performance of system.

REFERENCES

[1] Mostafa M M. Forecasting stock exchange movements using neural networks: Empirical evidence from Kuwait[J]. Expert Systems with Applications, 2010, 37(9): 6302-6309.

[2] Chandar S K, Sumathi M, Sivanandam S N. Prediction of Stock Market Price using Hybrid of Wavelet Transform and Artificial Neural Network[J]. Indian Journal of Science and Technology, 2016, 9(8).

[3] Pulido M, Melin P, Castillo O. Particle swarm optimization of ensemble neural networks with fuzzy aggregation for time series prediction of the Mexican Stock Exchange[J]. Information Sciences, 2014, 280: 188-204.

[4] Kara Y, Boyacioglu M A, Baykan Ö K. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange[J].Expert systems with Applications, 2011, 38(5): 5311-5319.

[5] Zhang Y, Wu L. Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network[J].Expert systems with applications, 2009, 36(5): 8849-8854.

[6] Averkin A, Yarushev S, Dolgy I, et al. Time Series Forecasting Based on Hybrid Neural Networks and Multiple Regression[C].Proceedings of the First International Scientific Conference "Intelligent Information Technologies for Industry"(IITI'16). Springer International Publishing, 2016: 111-121.

[7] Zhang D, Zhou Z H. (2D) 2PCA: Two-directional two-dimensional PCA for efficient face representation and recognition[J].Neurocomputing, 2005, 69(1): 224-231.hj

[8] Hinton G E. Deep belief networks[J]. Scholarpedia, 2009, 4(5): 5947.

[9] Huang W, Song G, Hong H, et al. Deep architecture for traffic flow prediction: deep belief networks with multitask learning[J].Intelligent Transportation Systems, IEEE Transactions on, 2014,15(5): 2191-2201.

[10] Jabbarzadeh A, Shavvalpour S, Khanjarpanah H, et al. A Multiple-Criteria Approach for Forecasting Stock Price Direction: Nonlinear Probability Models with Application in S&P 500 Index[J] .International Journal of Applied Engineering Research, 2016, 11(6)

[11] Guo Z, Wang H, Yang J, et al. A Stock Market Forecasting Model Combining Two-Directional Two-Dimensional Principal Component Analysis and Radial Basis Function Neural Network[J]. PloS one, 2015, 10(4): e0122385.

[12] Qiu M, Song Y. Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model[J].PloS one, 2016, 11(5): e0155133.

[13] Kong H, Wang L, Teoh E K, et al. Generalized 2D principal component analysis for face image representation and recognition[J].Neural Networks, 2005, 18(5): 585-594

[14] Qiu M, Song Y, Akagi F. Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market[J]. Chaos, Solitons & Fractals, 2016, 85: 1-7

[15] Hrasko R, Pacheco A G C, Krohling R A. Time Series Prediction Using Restricted Boltzmann Machines and Backpropagation[J]. Procedia Computer Science, 2015, 55: 990-999.