Financial Time-series Data Analysis using Deep Convolutional Neural Networks

Jou-Fan Chen, Wei-Lun Chen, Chun-Ping Huang
Institute of Information Management
National Chiao Tung University
HsinChu, Taiwan, R.O.C.

Szu-Hao Huang, An-Pin Chen
Department of Information Management and Finance
National Chiao Tung University
HsinChu, Taiwan, R.O.C.

Abstract*—A novel financial time-series analysis method based on deep learning technique is proposed in this paper. In recent years, the explosive growth of deep learning researches have led to several successful applications in various artificial intelligence and multimedia fields, such as visual recognition, robot vision, and natural language processing. In this paper, we focus on the time-series data processing and prediction in financial markets. Traditional feature extraction approaches in intelligent trading decision support system are used to applying several technical indicators and expert rules to extract numerical features. The major contribution of this paper is to improve the algorithmic trading framework with the proposed planar feature representation methods and deep convolutional neural networks (CNN). The proposed system is implemented and benchmarked in the historical datasets of Taiwan Stock Index Futures. The experimental results show that the deep learning technique is effective in our trading simulation application, and may have greater potentialities to model the noisy financial data and complex social science problems. In the future, we expected that the proposed methods and deep learning framework could be applied to more innovative applications in the next financial technology (FinTech) generation.

***Keywords*—Deep learning, data visualization, trend prediction, convolutional neural networks, machine learning**

I. Introduction

With the wide popularity of internet and mobile platforms like smartphones, wearable devices, and IOT, the total amount of collected data had also been grown explosively. The big data analysis focuses on the information discovering and knowledge extraction from the huge and growing data. A variety of artificial intelligence methods originated from various research fields, such as pattern recognition, data mining, and machine learning, are applied to solve the data-driven models and to develop practical information systems. Data analysis in the research field of Economics and Finance had also been developed vigorously. There are many machine learning methods that have been successfully applied in various interesting financial problems, such as trend approximation, bankruptcy prediction, investment targets, credit evaluation, and portfolio selection. In this paper, we developed a deep learning framework based on convolutional neural networks to analyze the trading time-series data. The experimental results show that the proposed learning system can understand more complex and useful information directly from the raw price and volume sequence than traditional rule-based trading system which is defined by the expert experience.

In 2012, the Google team used the machine learning method to learn the classification rules from ten million unlabeled pictures, and was successful in image classification. In comparison with the previous classifier, this machine learning method significantly improved the accuracy. It's Deep Learning! In 2015, the Google team has a new initiative: AlphaGo. AlphaGo is composed of the convolutional neural network (CNN) model from deep learning concept and became the first Go program to beat the professional Go player in 2015. In addition, AlphaGo dueled the world Go champion Lee Se-dol five times and won four games. Since then, deep learning becomes much more popular in the world.

CNN has obtained good effect in various fields. However, CNN needs to be fed two-dimensional data as input, leading to seldom apply in the analysis of financial time series data. Therefore, this paper tried to develop 2D mapping method of time series data, and used CNN model to predict the future trend of Taiwan index futures. We will evaluate the feasibility of deep learning application in financial prediction.

The main contribution of this paper are (1) we develop the mean average mapping method(MAM) and the double moving average mapping method(DMAM) to transform the time series data into 2D images. The transformed images are not allowed to lose any information and can be recognized by CNN for training. (2) Using these transformed images, CNN can capture useful features and successfully classify the price trend. It means that CNN won't be limited to 2D visualized data and has the potentiality of finance prediction or other time series analysis.

We believe that this work is a good example of novel Financial Technology (FinTech) applications which tried to apply advanced technology to solve financial problems and develop innovative applications. Artificial Intelligence and big data analysis also play important roles in this emerging research field to extract more hidden knowledge and contribute to the service automation. It also can be utilized in the intelligent trading system, such as high-frequency trading and algorithmic trading, or various customized, personalized, and unmanned service design.

978-1-5090-3555-7/16 $31.00 © 2016 IEEE
DOI 10.1109/CCBD.2016.51

In outline, we review machine learning in finance prediction, deep learning and convolutional neural network in Section II. Section III describes our proposed methods, including training data, our CNN model design and the time series data transformation method. Next, we show some experiment to decide the value of parameter and CNN's classification result in section IV Finally, we give the conclusions and future directions in section V.

II. LITERATURE REVIEW

A. Machine learning in financial prediction

Machine learning methods already had been used to predict financial time-series applications [31] for several years. SVM's classification ability is often used as a financial market predictor. [1] applied SVM in the prediction of financial time series data, and used five different groups of futures data. It was verification that SVM in financial time series forecasting could be an ideal tool besides BPNN. In addition, [2] used SVM to classify the short term price movements from high frequency data of limit orders. [3] [4],[5],[6] researched financial textual data with SVM.

Moreover, neural network also takes a place in the financial market forecast. [7], [8], [9] used the traditional neural network to predict the price for stock index. [10] proved that forward neural network had an outstanding effect in prediction on S&P500. [11] proposed a hybrid model, the combination of genetic algorithm for technical indicators and neural network for other data, as a new time series prediction method. [12] also used a hybrid model, the combination of recurrent neural network and genetic algorithm, to predict for NASDAQ, other 36 companies stock prices and textual feature data, and got more favorable results than traditional financial method. [13] confirmed that the nonlinear neural network has a significant effect on the processing of textual data, and can be used to predict the market.

B. Deep Learning

The concept of deep learning is similar with further extracting feature from the former feature. Through extracting many times, deep learning will obtain higher dimensional feature for classification. So far, deep learning has an outstanding result in image classification, audio or natural language processing applications.

In the beginning, the concept of deep learning was proposed by Lecun, Y in 1989. Lecun also completed [14] and [15] research. However, deep learning was hard to apply because it spent much time to establish the model. In 2006, Hinton, G.E's research ,[16] 、[17], put an emphasis on deep learning again in the academia, and explained that deep learning used high dimensions feature, which could obtain the more classification accuracy. Deep learning is the combination of supervised learning and unsupervised learning to extract feature from deeper structure.

Deep learning researches had been published can be divided into two part, theory and application. In theory part, the main target placed in how to speed up the learning time and optimize classification [16]-[23]. In application part, deep learning applied in images, computer visions, text, and other characteristics of more complex data to learn[24]-[28][32][33].

However, deep learning is rarely used in the field of finance. [29]- [30] used finance newspaper textual data to predict the price movement under deep learning. [31] applied deep learning in bankruptcy prediction.

C. Convolutional Neural Networks

In the machine learning field, not only deep learning becomes more and more popular, but also convolutional neural network (CNN) is attracted attention recently. CNN is mainly applied in image recognition, shown in Figure 1[14]. To feed two-dimensional data as input, and process with multi-layer of convolution. Each layer of convolution will receive the output of the previous layer of convolution as the input. CNN keeps pooling and convolution to extract useful feature for each layer and finally obtain high-quality feature, which can get a favorable ability of image classification.

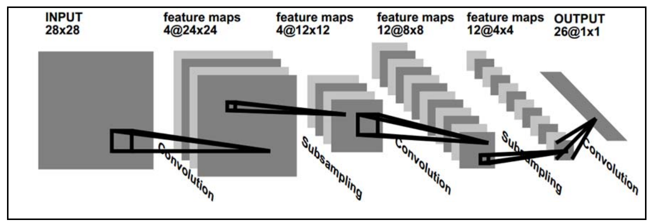

Fig.1. Convolutional networks illustration in the previous works [14].

III. METHODOLOGY

This paper attempted to analysis financial market price data and transformed financial time series data into two-dimensional images as input for classifying the trend in future. Trend categories are rise, fair, fall. We used three kinds of visualization methods as the experimental group and the candlestick chart as the control group for performance evaluation.

A. Financial Data and Computational Evironemnt

This paper included approximately 1,054,959 historical intraday one-minute data of Taiwan index futures, which covered from January 2, 2001 to April 24, 2015. Each one-minute data contains the opening price, the closing price, the highest price and the lowest price. For example, the plot of the closing price is shown in Figure 2.

Fig.2. TXF closing price plot.

In our experiments, we used the computing platform with Intel® core™ I7-4790 CPU @ 3.60 GHz and 8.00GB RAM. We referenced a framework of convolutional neural network designed by [29] and the programming environment is MATLAB. This CNN model contains 2 layers of convolution, 2 layers of pooling, and a layer of output.

B. Parameters

The main variables are as following:

1) Two-dimensional data size (D)：To decide the image size. We preset D = 20 to avoid the lack of memory problems. Each 2-D data size is 20*20, it means that the width and the height are 20 pixels for one image. Applied to our problem, every image included 400 financial data.

2) Training times (N)：To decide the training times of the model. Since the training data set is enormous, we preset N = 20 to avoid the lack of memory problems.

3) Prediction future time (F)：To decide the time interval of asset holdings after the on-line training sequence collected. Due to short F is easily affected by the short-term fluctuations, this paper attempts to predict trend stably. We preset F = 7.

4) Threshold：Through sigmoid function, CNN will give each category a corresponding score based on input data. The range of score is between 0 and 1. If the score of a category is higher, it means the output would be more similar with the category. So the result of the prediction would be the category with the highest score. However, some result of the highest score is not that high enough. We consider this is not a valid classification. To prevent the influence of uncertain classification, we checked the highest score to be greater than a threshold. If it met the threshold, we adopted the classification. After a lot of experiments, we preset Threshold = 0.65.

5) Training period (TP): To decide the length of training data. Through a lot of experiments, the number of training data needs to be sufficient in order to have a certain stability, and the more training data may bring a higher accuracy. However, we don't use longer TP because of the lack of memory problem. We preset 5-years data as a set, which is composed of 4-years data for training and 1-year data for testing. There're 9 sets by scrolling year by year. We finally compute training and testing accuracy on average. About the reason for each group of five years, there are pretest we will show in part (D)-1.

6) Time frame (TF)：To decide time frame of each time series data. As mentioned in III-A, the time frame of raw data is one-minute that included the opening price, highest price, lowest price and the closing price, we called that one-minute Candlestick (TF=1). However, we thought that the more diversity of the data, the more information we got. So we turned the raw data to five-minute Candlestick (TF=5) as another input. Shown in Figure 3, the time frame in five-minute Candlestick (TF=5) is 5 times as much as one-minute Candlestick (TF=1). We preset TF = 1 and 5 which most of ordinary investors use in investment. About the parameters we preset, there are pretest we will show in part (D)-2.

Fig.3. Time frame covered illustration.

C. Data visualization method

CNN is mainly used for features capture and classification of two-dimensional image data, nevertheless, the time series data of financial market is in one dimension. So we needed to convert the data into 2D image through data pre-processing, thereby feeding into CNN for identification and classification as inputs. There are three methods as the experimental group we implemented in this paper.

1) Gramian Angular Field (GAF)

In [30], it analyzed time series data by CNN model, which adopted Gramian Angular Field (GAF) to convert the data to 2D. The application of Gramian Matrix used in computing the nonlinear data is the conception of GAF. The one-dimensional data is as a vector, then it turns to a symmetric matrix derived from an inner product. We called this symmetric matrix as GAF.

$$G = \begin{bmatrix} <\tilde{x}_1, \tilde{x}_1> & \cdots & <\tilde{x}_1, \tilde{x}_n> \\ \vdots & \ddots & \vdots \\ <\tilde{x}_n, \tilde{x}_1> & \cdots & <\tilde{x}_n, \tilde{x}_n> \end{bmatrix} \quad (1)$$

$$\text{where } <\mathrm{x}, \mathrm{y}> = \mathrm{x} \cdot \mathrm{y} - \sqrt{1-x^2} \cdot \sqrt{1-y^2}$$

The concept and steps of GAF are as following:

Normalize the time series data $\mathrm{X} = \{x_1, x_2, \dots, x_n\}$ to fall within the interval $[-1,1]$ or$[0,1]$, to generate$\tilde{X} = \{\widetilde{x_1}, \widetilde{\mathrm{x}_2}, \dots, \widetilde{\mathrm{x}_\mathrm{n}}\}$. This paper use interval$[0,1]$.

According to [30], there are several advantages of GAF representation method: the bottom right element in the matrix contained the last information in the time series. In other words, the old information may be placed on the upper and left parts, and the timestamp of each data element in this matrix will be increased with rightward and downward axes. This attribute means that the generated visualized graph can retain the temporal dependency, making the transformed 2D images without losing the property of the original data. Moreover, GAF also contains the temporal correlations. Because the time interval between the neighbor signals of the input 1D sequence may not be constant, more hidden information can be modeled in the G(i,j) value, such as the relationship of the time delay between the absolution timestamps of data point i and j.

2) Moving Average Mapping (MAM)

Moving Average (MA) is one of the widely known and accepted technical analysis indicators. Some people may have regarded MA as a characteristic in the investment decision making, such as the so-called ''golden cross'', ''death cross'' that were staggered by the different timeframe of MA lines. Compared to the original data, MA has a smoother curve, so that it has more temporal dependency.

For each input data in particular historical time, various moving average values are calculated with different timeframe settings. The output is two-dimensional data. The formula is as following:

$$G = \begin{bmatrix} MA_{t-D+1,1} & \cdots & MA_{t,1} \\ \vdots & \ddots & \vdots \\ MA_{t-D+1,D} & \cdots & MA_{t,D} \end{bmatrix} \quad (2)$$

where $MA_{i,j}$ is the average of the closing price from the time period i to j.

3) Double Moving Average Mapping (DMAM)

We extended to depict another visualization method, named as "double MAM". The visualization method was very similar to the MAM, which had been mentioned above. The difference between these two methods is the input data, as MAM used the closing price, DMAM used the mean value of the opening price and closing price. From doing this, we expected DMAM covered more information than MAM.

In addition, no matter which method is used, all of the 2D data will be standardization, so that the input is scaling to the interval[0,1]. The standardization function is as the formula (3).

$$X_{new} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3)$$

D. Experimental procedure

At first, we pretested some parameters that will affect the process of our experiment. Then, shown in Figure 4, we mapped time series sequence data (shown in red frame) into single 2D plane, and wanted to predict the trend during the consecutive seven days (shown in blue arrow). The sliding window shifted by 1-time unit. Finally, we evaluated the accuracy rate under the different parameters and the 2D mapping methods.

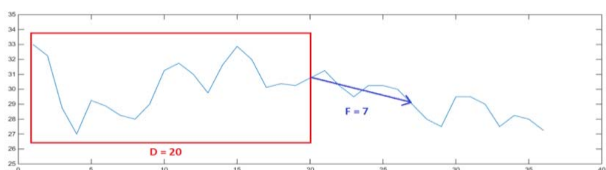

Fig.4. The instance of a sliding window

The detail of the experimental procedure is as following:

1) The setting of Training period (TP)

The amount of the data is over 100 million in the experiment, and the length of the time period is more than 14 years. However, the trading pattern and trend are caused by humans, of course the data. It probably changed along with time. For the reason, we assumed that under the different length of time to train, CNN may capture different characteristics, so as the classification results. Therefore, it's necessary to find suitable training period.

2) The setting of Time frame (TF)

We carried out the experiment of the four mapping methods and compare its results. Since then, we tried to observe the effects of TF, and find out the better TF value in order to be adopted in the later experiment.

3) Experimental group

Each method was described above generated two images, which matched the best two value of TF. These two visualization methods were fed into CNN model to train at the same time, and the depth of the convolution was two. The dimension of the input will be 20*20*2. After the training, we built a model with the historical data, to classify the categories of the trend in the future.

4) Control group

The basic way to present the price in the financial market is candlestick chart, therefore, we used candlestick chart as input to train the CNN model and became our control group. It was for the purpose of understanding whether CNN can also effectively acquired information from Candlestick chart or not, and evaluated the performance for comparing with the experimental group which people can't grasp from the graph.

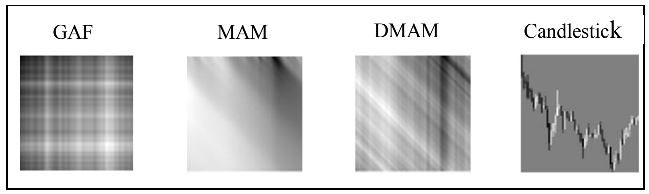

Fig.5. The example of the same data in the different graphic mapping

5) Performance evaluation

The performance evaluation of this paper is based on the evaluation of the accuracy. The paper distinguished the trend into three categories: rise, fair and fall. The formula to compute the accuracy is as formula (4).

$$\text{Accuracy} = \frac{\sum_{i=1}^{n} Correct_i}{N} \quad (4)$$

When the prediction of the i-th data is the same as the actual situation, $Correct_i = 1$, or 0 and N is total number of the test data.

Due to the division of three categories, the expected value of the prediction is about 33.33%.

IV. Experimental Results

In the beginning, we pretested the suitable parameters TP, TF for the experiment. Figure 6 displayed the accuracy of the different training time period for the four mapping methods. In Figure 6, we can roughly have observed that the performance tended to be stable when TP is greater than or equal to 4. However, the more TP was, the more training time was spent. It would lead to insufficient memory problem. So, the paper adopted the length of 4 for TP.

Figure 7 showed the accuracy of the different time frame (TF) for the four mapping methods. By observing Figure 7, we obviously found, the shorter the TF was, the higher the performance was. As a result, we adopted TF of 1, 5 to carry out the following experiment.

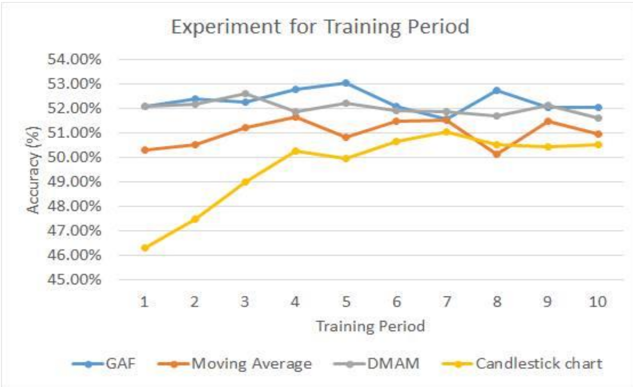

Fig.6. The accuracy of the different training time period

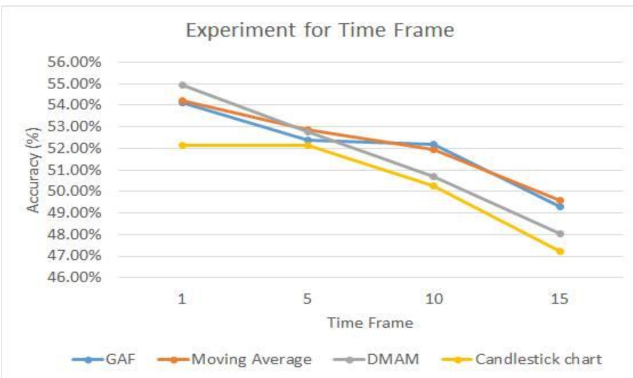

Fig.7. The accuracy of the different time frame

The paper respectively modeled the three graphics method into CNN according to the parameter setting, and compared with the control group, to verify the ability of the feature capture and classification of the designed 2D image.

Table 1 showed the accuracy results of all the methods in the different periods of the test data. Figure 6 showed the relatively relations between all the results.

TABLE I. ACCURACY COMPARISON

Testing Year	GAF	MAM	DMAM	Candlestick chart
2006	56.53%	57.28%	56.05%	55.04%
2007	55.97%	56.86%	57.06%	55.72%
2008	55.73%	55.62%	55.57%	53.87%
2009	53.84%	53.71%	53.81%	53.91%
2010	57.06%	56.42%	56.42%	56.20%
2011	**57.88%**	56.42%	57.10%	55.36%
2012	52.52%	52.44%	52.75%	53.60%
2013	51.28%	51.63%	52.07%	50.66%
2014	52.93%	52.47%	52.36%	49.50%
Average	54.86%	54.76%	54.80%	53.76%

In table 1, the highest accuracy can be achieved to 57.88%, that is, it had a better performance than expected rate (33.33%) for 24.55%. It was in the condition of GAF mapping and the testing data in 2011. It had been considered a very good results.

However, we can find something interesting in Figure 8, the accuracy seems to be lower after 2012. We guess that the frequency of the artificial operation in the market is increasing, or the more artificial intelligence is invested in the market. As a result, the pattern in the market rapidly changed, and the captured feature is hard to sustain for a long time. In spite of that, the overall performance got close to 53%, which improved around 20% than expected rate. The ranking from high to low is about: GAF >DMAM >MA > Candlestick chart. It's worth attention that the accuracy of all the experimental group is better than the control group.

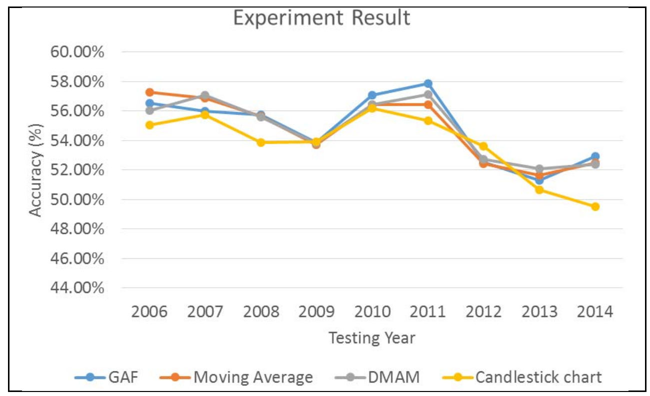

Fig.8. Broken-line graph of the accuracy

From the results, we discovered that the graph people understand was not necessarily valuable for CNN, instead, CNN can effectively classified the graph for some useful information that people can't grasp.

V. CONCLUSIONS

This paper transforms Taiwan index futures time series data into the two-dimensional image data as input for CNN through different visualization methods. We got good accuracy to the classification of future trends. It was proved that CNN could extract useful features and recognize the behaviors of financial markets based on the proposed two-dimensional time-series data representation methods. The trend of stock price actually has its characteristics and predictability.

CNN still has certain feature acquisition and classification ability regarding to the protean finance market. can try more data needed for prediction in the future. Even if original one or higher dimensional data, all you need to do is find appropriate method of conversion. CNN can still play an excellent ability of prediction, and should not be limited in two-dimensional image data. Here are some directions and suggestions for further researches:

1) More technical indicators as input：

The data formed by human knowledge is even valuable. Our experiment only used the simplest data to classify. If there is more valuable information, it could reduce the difficulty of classification.

2) Prediction of future trends of financial products in other countries：

Taiwan financial product's volume is limited, but there is large trading volume in the United States or other global financial products. These global financial products are affected by the great impact of much more traders. It's worthy to discovery their trading behaviors.

3) Investment portfolio analysis：

Investment portfolio is also a popular financial application. It's interesting to predict the future profitability of a variety of investment portfolio with a certain amount of data as input.

ACKNOWLEDGEMENT

The authors would like to thank the Ministry of Science and Technology of the Republic of China, Taiwan, for partially financially supporting this research under contract number MOST 104-2410-H-009-017-MY2 and MOST 105-2221-E-009-153. This work was also supported by the FinTech Innovation Research Center, National Chiao Tung University.

REFERENCES

[1] Cao, L. J., & Tay, F. E. (2003). Support vector machine with adaptive parameters in financial time series forecasting. Neural Networks, IEEE Transactions on, 14(6), 1506-1518.

[2] Kercheval, A. N., & Zhang, Y. (2013). Modeling high-frequency limit order book dynamics with support vector machines.

[3] Smailović, J., Grčar, M., Lavrač, N., & Žnidaršič, M. (2014). Stream-based active learning for sentiment analysis in the financial domain. Information Sciences, 285, 181-203.

[4] Kranjc, J., Smailović, J., Podpečan, V., Grčar, M., Žnidaršič, M., & Lavrač, N. (2015). Active learning for sentiment analysis on data streams: Methodology and workflow implementation in the ClowdFlows platform. Information Processing & Management, 51(2), 187-203.

[5] Hagenau, M., Liebmann, M., & Neumann, D. (2013). Automated news reading: Stock price prediction based on financial news using context-capturing features. Decision Support Systems, 55(3), 685-697.

[6] Li, X., Xie, H., Chen, L., Wang, J., & Deng, X. (2014). News impact on stock price return via sentiment analysis. Knowledge-Based Systems, 69, 14-23.

[7] Chapados, N., & Bengio, Y. (2001). Cost functions and model combination for VaR-based asset allocation using neural networks. Neural Networks, IEEE Transactions on, 12(4), 890-906.

[8] Sitte, R., & Sitte, J. (2000). Analysis of the predictive ability of time delay neural networks applied to the S&P 500 time series. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 30(4), 568-572.

[9] Zhang, G. P. (2007). A neural network ensemble method with jittered training data for time series forecasting. Information Sciences, 177(23), 5329-5346.

[10] Gençay, R., & Gibson, R. (2007). Model risk for European-style stock index options. Neural Networks, IEEE Transactions on, 18(1), 193-202.

[11] Armano, G., Marchesi, M., & Murru, A. (2005). A hybrid genetic-neural architecture for stock indexes forecasting. Information Sciences, 170(1), 3-33.

[12] Kwon, Y. K., & Moon, B. R. (2007). A hybrid neurogenetic approach for stock forecasting. Neural Networks, IEEE Transactions on, 18(3), 851-864.

[13] Geva, T., & Zahavi, J. (2014). Empirical evaluation of an automated intraday stock recommendation system incorporating both market data and textual news. Decision support systems, 57, 212-223.

[14] LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361(10), 1995.

[15] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, ''Gradient-based learning applied to document recognition,'' Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[16] Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. Neural computation, 18(7), 1527-1554.

[17] Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. Science, 313(5786), 504-507.

[18] Zhang, Z., Luo, P., Loy, C. C., & Tang, X. (2015). Learning deep representation for face alignment with auxiliary attributes.

[19] Kavukcuoglu, K., Ranzato, M. A., Fergus, R., & Le-Cun, Y. (2009, June). Learning invariant features through topographic filter maps. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 1605-1612). IEEE.

[20] Wong, W. K., & Sun, M. (2011). Deep learning regularized Fisher mappings.Neural Networks, IEEE Transactions on, 22(10), 1668-1675.

[21] Larochelle, H., Bengio, Y., Louradour, J., & Lamblin, P. (2009). Exploring strategies for training deep neural networks. The Journal of Machine Learning Research, 10, 1-40.

[22] Erhan, D., Bengio, Y., Courville, A., Manzagol, P. A., Vincent, P., & Bengio, S. (2010). Why does unsupervised pre-training help deep learning?. The Journal of Machine Learning Research, 11, 625-660.

[23] Montavon, G., Braun, M. L., & Müller, K. R. (2011). Kernel analysis of deep networks. The Journal of Machine Learning Research, 12, 2563-2581.

[24] Yuan, Y., Mou, L., & Lu, X. (2015). Scene recognition by manifold regularized deep learning architecture. IEEE transactions on neural networks and learning systems, 26(10), 2222-2233.

[25] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[26] Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., & Ng, A. Y. (2011). Multimodal deep learning. In Proceedings of the 28th international conference on machine learning (ICML-11) (pp. 689-696).

[27] Theis, L., Gerwinn, S., Sinz, F., & Bethge, M. (2011). In all likelihood, deep belief is not enough. The Journal of Machine Learning Research, 12, 3071-3096.

[28] Nair, V., & Hinton, G. E. (2009). 3D object recognition with deep belief nets. InAdvances in Neural Information Processing Systems (pp. 1339-1347).

[29] Palm, R. B. (2012). Prediction as a candidate for learning deep hierarchical models of data. Technical University of Denmark, Palm, 25.

[30] Wang, Z., & Oates, T. (2015). Imaging time-series to improve classification and imputation. arXiv preprint arXiv:1506.00327.

[31] Huang, S.H., Tai, S.H., & Lai, S.H. (2011) A Learning based Contrarian Trading Strategy via Dual Classifiers Model, ACM Transactions on Intelligent Systems and Technology, 2(3).

[32] Huang, S.H. & Pan, Y.C. (2015) Automated Visual Inspection in Semiconductor Industry: A Survey, Computers in Industry, 66, 1-10.

[33] Huang, S.H. & Pan, Y.C. (2014) Ergonomic job Rotation Strategy Based on An Automated RGB-D Anthropometric Measuring System. Journal of Manufacturing Systems, 33(4), 699-710.