

Article

# Deep Learning-Based Corporate Performance Prediction Model Considering Technical Capability

Joonhyuck Lee <sup>1</sup>, Dongsik Jang <sup>1</sup> and Sangsung Park <sup>2,\*</sup>

<sup>1</sup> Department of Industrial Management Engineering, Korea University, Seoul 02841, Korea; iguana751@korea.ac.kr (J.L.); jang@korea.ac.kr (D.J.)

<sup>2</sup> Graduate School of Management of Technology, Korea University, Seoul 02841, Korea

\* Correspondence: hanyul@korea.ac.kr; Tel.: +82-10-8700-0469; Fax: +82-2-3290-3900

Academic Editor: Giuseppe Ioppolo

Received: 15 May 2017; Accepted: 24 May 2017; Published: 26 May 2017

**Abstract:** Many studies have predicted the future performance of companies for the purpose of making investment decisions. Most of these are based on the qualitative judgments of experts in related industries, who consider various financial and firm performance information. With recent developments in data processing technology, studies have started to use machine learning techniques to predict corporate performance. For example, deep neural network-based prediction models are again attracting attention, and are now widely used in constructing prediction and classification models. In this study, we propose a deep neural network-based corporate performance prediction model that uses a company's financial and patent indicators as predictors. The proposed model includes an unsupervised learning phase and a fine-tuning phase. The learning phase uses a restricted Boltzmann machine. The fine-tuning phase uses a backpropagation algorithm and a relatively up-to-date training data set that reflects the latest trends in the relationship between predictors and corporate performance.

**Keywords:** prediction model; corporate performance prediction; deep learning; deep belief network; technical indicator

## 1. Introduction

Many studies have predicted the future performance of companies for the purpose of making investment decisions [1–7]. Most of these are based on the qualitative judgment of experts in related industries, who consider various financial and firm performance information [8,9]. However, qualitative judgments are highly subjective, and limited in the sense that conclusions come at a significant cost in terms of time and money. With recent developments in technology, researchers have begun using machine learning techniques to predict corporate performance [1–7]. For example, artificial neural networks, which have relatively good predictive ability in various fields, are widely used in such studies [4]. However, models based on artificial neural networks often suffer from the problem of overfitting on training data. Moreover, training a deep neural network takes a long time, and the propagation of errors, based on a backpropagation algorithm, back to the input layer can be difficult.

The results of many studies have shown that prediction models constructed using a support vector machine (SVM), as suggested by Vapnik, have good predictive performance and a fast learning speed [10]. As a result, many researchers have investigated using a SVM to predict corporate performance and stock prices [1,3,5,7].

Recently, artificial neural network-based prediction models have again been attracting attention owing to the development of parallel processing technology, as well as algorithms that overcome the limitations of deep neural networks [11–13]. A typical algorithm used to train a deep neural network

is the deep belief network (DBN). A DBN performs pre-training through unsupervised learning using a restricted Boltzmann machine (RBM), and then fine-tunes the network via supervised learning on training data. In addition, convolutional neural networks, widely used in image processing and voice recognition, demonstrate good performance and are widely used in constructing classification models in various fields [14,15].

In order to construct a corporate performance prediction model, predictors are needed to predict the performance of companies. Most corporate performance prediction models use a company's financial performance data and financial indicators as predictors. However, there has been a recent increase in the proportion of technology-intensive firms whose technological capability significantly influences their corporate performance. Thus, in order to predict corporate performance more accurately, it is necessary to use both a company's financial information and its technical information as predictors. As a result, many recent studies have proposed indicators that show the technological competitiveness of a company [16–18]. Many of these studies apply patent data, because they are easy to use in quantitative analyses, have an internationally uniform structure, and contain citation information [19–22].

Among the studies that predict a company's corporate performance, there are few that construct a prediction model using patent data and a deep learning algorithm. In this study, we propose a deep neural network-based corporate performance prediction model that uses a company's financial and technical indicators as predictors.

The proposed model includes an unsupervised learning phase, using an RBM, in which training uses the entire training data set. Then, there is a fine-tuning phase, which uses a backpropagation algorithm and a relatively up-to-date training data set. These data reflect the latest trends in the correlation between predictors and corporate performance in forecasting in order to improve the prediction accuracy of the network. In general, managerial environments change over time [23–27]. Accordingly, a prediction model that cannot reflect recent trends will find it difficult to achieve sustainable prediction performance. The proposed model is expected to maintain sustainable prediction performance in a volatile business environment by fine-tuning the pre-trained model using the up-to-date data set.

## 2. Related Studies

With the development of technology, many researchers have attempted to use machine learning for investments and managerial decision-making. Here, artificial neural networks and SVMs show relatively good performance [1–7]. The studies reviewed in this section predict a company's performance and stock price using artificial neural networks.

Yoon and Swales [28] proposed an artificial neural network-based model that predicts stock prices using economic and financial indices as predictors. Furthermore, they compared the prediction performance of their proposed model to that of a multiple discriminant analysis-based prediction model [28]. Ahn et al. (2000) conducted a study to predict the bankruptcy of firms using a neural network algorithm [4]. They used companies' past financial performance indicators as predictors in the model. In addition, they improved the prediction accuracy of the model by adopting a preprocessing procedure, using a rough set approach to obtain a reduced information table. Finally, Lam (2004) proposed a model to predict the financial performance of firms using an artificial neural network with backpropagation algorithms [29].

However, prediction models based on artificial neural network algorithms may degrade in prediction performance because of overfitting on training data. When training a multi-layer neural network, nodes in the lower layer of the network find it difficult to achieve meaningful learning because of the vanishing gradient problem [30]. Therefore, studies often use a SVM to develop prediction models, because a SVM increases the prediction performance of a model by applying a maximum margin to reduce the possibility of overfitting [1,3,5,7].

Many studies predict increases and decreases in stock prices, as well as corporate default, using SVMs. Li and Sun (2009) suggested a SVM-based corporate default prediction model using the K-nearest neighbor algorithm [31]. Huang et al. (2005) proposed a model to predict the fluctuation of a stock market using a SVM [7]. The authors verified the proposed model empirically using NIKKEI stock average data. Their proposed SVM-based model showed superior prediction performance when compared with artificial neural network-based models.

Other studies have developed prediction models using a support vector regression (SVR), which is a modified SVM applied to a regression problem [3,32–36]. Hsu et al. (2009) predict the stock price of a firm using a SVR and a self-organizing map [3]. Lee et al. (2016) proposed a SVR-based corporate prediction model that searches for optimal SVR parameters using a genetic algorithm [37]. Hinton et al. (2006) proposed a DBN that can train deep neural networks effectively using a RBM and the wake—sleep algorithm [11].

The development of parallel processing technology has greatly reduced the learning time of neural networks. Thus, many studies have begun using deep neural networks again in various fields, such as speech recognition, image processing, and prediction models [11–15]. Deep neural networks are also used to predict firms' performance. Ribeiro and Lopes (2011) proposed a model that uses a DBN to predict the defaults of French companies [12]. The authors also compared the prediction performance of their model to that of a SVM-based prediction model. Their results showed that the performance of the DBN-based prediction model surpassed that of the SVM-based prediction model.

Shen et al. (2015) used an improved DBN to predict exchange rate movements [13]. They applied continuous restricted Boltzmann machines (CRBM) and a conjugated gradient method in the training process. They conducted a comparative study to verify the prediction performance of the proposed model empirically by using the same time series data of exchange rates. Their results showed that the proposed model outperforms a feedforward neural network-based model.

In addition, DBNs are widely used for constructing prediction models in various fields. Huang et al. (2014) proposed a deep neural network model for traffic flow prediction using a DBN. They adopted a grouping method based on the weights of the top layer to train the proposed model more effectively [38].

Kuremoto et al. (2014) proposed a modified DBN-based model for time series forecasting [39]. They adopted a particle swarm optimization algorithm during the pre-training process using a RBM. Furthermore, they added a preprocessing process that removes the seasonal factors of the original data ahead of the model training step to obtain smoothed training data [39]. They compared the prediction performance of the modified DBN model they proposed to that of a multi-layer perceptron-based model. It was shown that the proposed model had the lowest approximation error.

### 3. Deep Belief Networks

In this study, we propose a corporate performance prediction model using a RBM as the main component of a DBN. After pre-training using the RBM, the proposed model is fine-tuned using a backpropagation algorithm.

A DBN is a generative model with many layers [11]. The training process of a DBN can be divided into two phases. The first phase is the pre-training step, using unsupervised learning and a RBM. The second phase fine-tunes the pre-trained neural network for predictive purposes on a training data set. Figure 1 shows a DBN with three hidden layers.

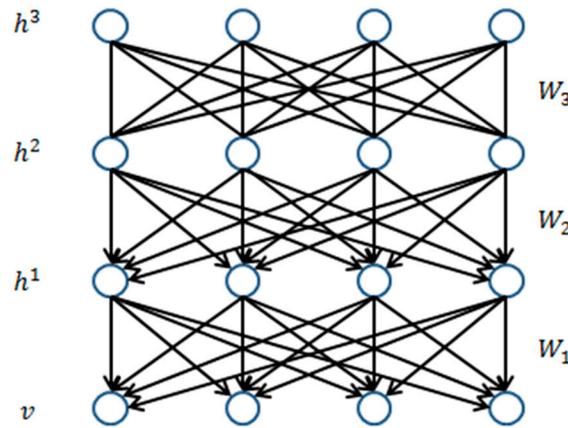


Figure 1. Architecture of a deep belief network (DBN).

As shown in Figure 1, the connection between the H3 and H2 layers is undirected, while the connections between the lower layers are directed downward. Here, the top two layers of the DBN consist of a RBM. This acts as an associative memory for the top-level features in the DBN [13]. The lower layers consist of a sigmoid belief network (SBN).

For example, the top two layers of a DBN, with N hidden layers, contain a RBM composed of  $h^N$ ,  $h^{N-1}$ ,  $W_N$ . The probability distribution of  $h^{N-1}$  is expressed as follows [40]:

$$p(h^{N-1} | W_N) = P_{RBM}(v = h^{N-1} | W_N). \tag{1}$$

The probability distribution of  $h^{N-2}$ , located in the next layer, follows the probability distribution of an SBN, which is expressed as follows:

$$\begin{aligned} p(h^{N-2} | W_{N-1}, W_N) &= \sum_{h^{N-1}} p(h^{N-2}, h^{N-1} | W_{N-1}, W_N) \\ &= \sum_{h^{N-1}} P_{Sig}(v = h^{N-2} | h^{N-1}, W_{N-1}) P_{RBM}(v = h^{N-1} | W_N). \end{aligned} \tag{2}$$

The probability distribution of the  $i$ -th hidden layer  $h^i$  is expressed as follows:

$$\begin{aligned} p(h^i | W_{i+1}, \dots, W_{N-1}, W_N) &= \sum_{h^{i+1}} p(h^i, h^{i+1} | W_{i+1}, \dots, W_{N-1}, W_N) \\ &= \sum_{h^{i+1}} P_{Sig}(v = h^i | h^{i+1}, W_{i+1}) p(h^{i+1} | W_{i+2}, \dots, W_{N-1}, W_N). \end{aligned} \tag{3}$$

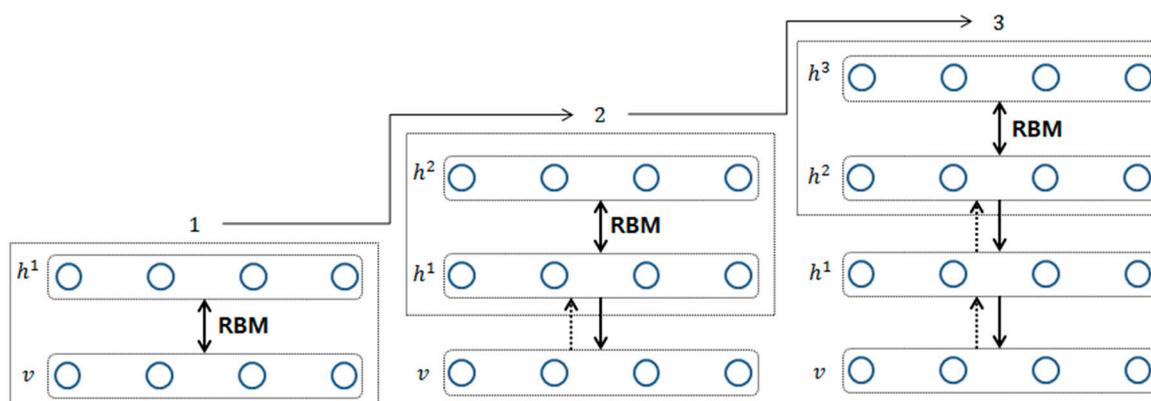
Lastly, the probability distribution of the visible layer  $v$  at the bottom of the network is expressed as follows:

$$p(v | W_1, \dots, W_{N-1}, W_N) = \sum_{h^1} P_{Sig}(v | h^1, W_1) p(h^1 | W_2, \dots, W_{N-1}, W_N). \tag{4}$$

The training process of the DBN is performed by modifying its parameters so that the relationship between “data” and “feature” is expressed well, as described above [11]. For a DBN with N hidden layers, the parameters  $W_1, W_2, \dots, W_N$  are trained by searching  $W_1, W_2, \dots, W_N$ , which contain the maximum log-likelihood of training data  $v$  for  $W_1, W_2, \dots, W_N$ . The log-likelihood of  $v$  for the DBN parameters is expressed as follows [40]:

$$\ln p(v | W_1, \dots, W_{N-1}, W_N) = \ln \sum_{h^1} P_{Sig}(v | h^1, W_1) p(h^1 | W_2, \dots, W_{N-1}, W_N). \tag{5}$$

As with the SBN, it is difficult to obtain the log-likelihood values directly if the hidden nodes of one layer are associated with a different layer's hidden nodes. Therefore, a DBN is trained by identifying a parameter that maximizes the lower bound ( $L_p$ ) of the log-likelihood [40]. When the upper parameters  $W_2, \dots, W_{N-1}, W_N$  and the lower parameter  $W_1$  are bound, the processes of maximizing the lower bound of the log-likelihood and training an RBM using  $W_1$  as the parameter are the same. Using this, Hinton et al. (2006) proposed a method of training a DBN by training one layer of the network at a time, using multiple RBMs [11]. The proposed training process is shown in Figure 2.



**Figure 2.** Training Process of a DBN using multiple restricted Boltzmann machines (RBMs).

As shown in Figure 2,  $W_1$  is trained using a RBM with given training data  $v$ . Then, the trained  $W_1$  is fixed and samples  $h^{1(k)}$  by  $P_{RBM}(h^1|v, W_1)$ . To train the second layer,  $W_2$  is trained using a RBM with the sampled  $h^{1(k)}$ . Then,  $h^{2(k)}$  for  $h^{1(k)}$  is sampled by  $P_{RBM}(h^2|h^1, W_2)$ . By repeating this step and stacking the hidden layers incrementally, it is possible to construct a directional neural network similar to a DBN [40]. As previously described, the pre-trained neural network must be fine-tuned for a specific purpose. For fine-tuning, common methods include the wake–sleep algorithm, and adding a final layer to the upper layer of the pre-trained network, then fine-tuning using a backpropagation algorithm [41].

In this study, we propose a deep neural network-based corporate performance prediction model trained in the same way as a DBN. In general, the impact of each corporate performance predictor changes with the market and business environment. Therefore, more recent trends should have priority when predicting corporate performance. In this study, the distribution of a company's performance predictors is pre-trained using unsupervised learning, and applying a RBM with the full training data set (see Figure 3). Then, we modify the pre-trained neural network using a RBM-to-general feedforward neural network (FNN) structure by stacking one output layer at the top of the pre-trained network. The initial values of the neural network parameters from the second layer to the visible layer of the modified neural network are replaced by the pre-trained parameters, and the uppermost parameters are set to the random initial values. Then, the network parameters are fine-tuned by applying a backpropagation algorithm, using relatively recent training data. Using recent data in the fine-tuning process means recent trends are given priority in predictions, thus improving the accuracy of predictions of time-series data (e.g., corporate performance) [37,42]. In addition, the model is expected to have sustainable prediction performance if new data are periodically added to the training data used in the fine-tuning phase.

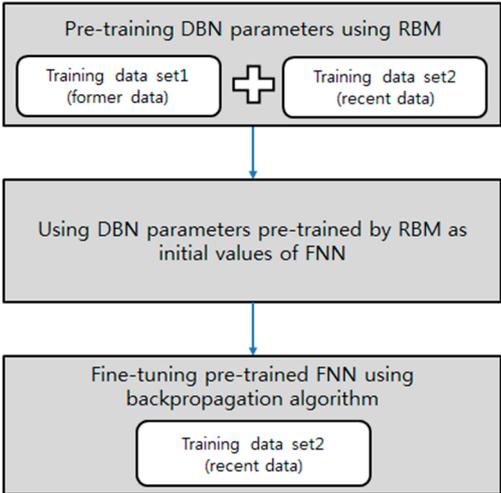


Figure 3. Summary of the proposed model’s pre-training and fine-tuning.

4. Experiments and Results

In this study, we propose a prediction model that predicts the revenue, operating profit, and net profit of a company, based on a RBM and a backpropagation algorithm. Figure 4 summarizes the proposed model.

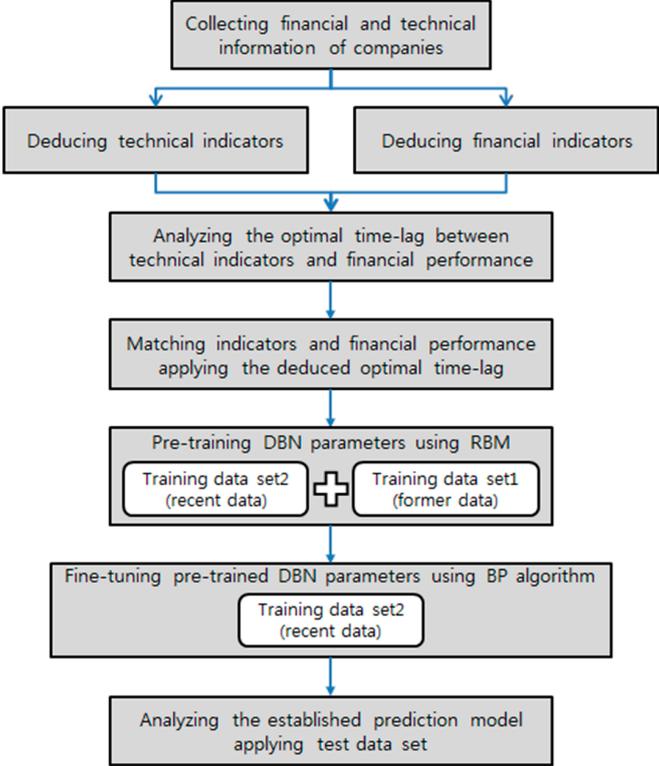


Figure 4. Summary of the proposed prediction model.

In this study, we analyzed the prediction performance of the proposed DBN-based prediction model using the financial, technological, and patent data of 22 pharmaceutical companies listed on the US stock market from 2000 to 2015. Using the verification data, we compared the predictability of the proposed model to that of models constructed using a SVR and a FNN.

#### 4.1. Predictors of the Proposed Prediction Model

We used 11 financial indicators and four patent indicators as predictors in the proposed model to predict corporate performance. The descriptive statistics of the 11 financial indicators are shown in Table 1.

**Table 1.** Descriptive statistics of the derived financial indicators.

| Financial Indicator                  | Minimum    | Maximum    | Average   | Standard Deviation |
|--------------------------------------|------------|------------|-----------|--------------------|
| Revenue (US\$ Million)               | 0.20       | 74,331.00  | 18,027.55 | 18,132.86          |
| Revenue Growth (\$ Million)          | −19,160.00 | 40,360.00  | 1498.12   | 4679.32            |
| Operating Profit (\$ Million)        | −3,016.00  | 22,193.00  | 4226.28   | 4676.32            |
| Operating Profit Growth (\$ Million) | −11,314.00 | 13,986.00  | 365.74    | 2354.49            |
| Net Profit (\$ Million)              | −4,886.00  | 22,003.00  | 3407.01   | 4016.59            |
| Net Profit Growth (\$ Million)       | −12,867.00 | 12,773.00  | 341.74    | 2601.51            |
| Gross Profit (\$ Million)            | −49.16     | 52,340.00  | 13,044.44 | 13,327.01          |
| Operating Expense (\$ Million)       | 9.51       | 53,833.00  | 13,772.44 | 14,054.36          |
| Shareholder's Equity (\$ Million)    | −72.13     | 90,450.00  | 17,401.59 | 21,827.93          |
| Total Assets (\$ Million)            | 8.73       | 212,950.00 | 36,185.63 | 42,421.32          |
| Capital Adequacy Ratio (%)           | −1.54      | 0.88       | 0.48      | 0.21               |

In recent years, there have been studies using technical indicators that represent the technological capability of a company as predictors to predict corporate performance [16–18,37]. Many of these studies apply patent data to derive the technical indicators [16–22,37]. We collected 59,740 patent data items on 22 pharmaceutical companies in order to derive the technical indicators of companies' technological capability. Using these data, four patent indicators (the number of patent applications, patent share, the number of Patent Cooperation Treaty (PCT) patent applications, and PCT patent share) were identified and used as predictors in the proposed model. The number of patent applications represents the total number of patent applications filed in a year by a particular company. The patent share denotes a company's share of all patents in the same industry. Unlike a general patent application, a patent registered under the PCT can obtain exclusive rights in many countries at the same time. Because PCT patents are relatively costly in terms of application and retaining rights, they tend to be used for technically significant and valuable inventions. Therefore, we also include the number of PCT patent applications and the PCT patent share as predictors in the proposed model. The descriptive statistics of the four patent indicators are shown in Table 2.

**Table 2.** Descriptive statistics of the derived patent indicators.

| Patent Indicator  | Minimum | Maximum | Average | Standard Deviation |
|---|---------|---------|---------|--------------------|
| The number of patent applications                                 | 0.000   | 893.000 | 169.716 | 174.817            |
| Patent share  | 0.000   | 0.202   | 0.045   | 0.045              |
| The number of Patent Cooperation Treaty (PCT) patent applications | 0.000   | 444.000 | 96.580  | 102.631            |
| PCT patent share  | 0.000   | 0.261   | 0.045   | 0.049              |

In many cases, it takes a number of years to commercialize a technology after applying for a patent and, thus, improve corporate performance [37]. Therefore, we conduct a correlation analysis to determine the time-lag that maximizes the correlation coefficient between technical indicators and corporate performance. The results show that the average correlation coefficient is highest (0.5015)

for a time-lag of three years, suggesting that the four technical indicators best explain and predict corporate performance after three years.

#### 4.2. Model Design

The purpose of this study is to construct a model that predicts a company's future revenue, operating profit, and net profit using past financial and patent data. Changes in revenue, operating profit, and net profit are more stable than changes in a stock price. Therefore, if the model is designed to predict the corporate performance for the next quarter, a comparison with the performance of models developed using other algorithms, such as a SVR or a FNN, may be unclear. Therefore, we designed the proposed model to predict corporate performance using the financial indicators of the previous year, and the patent indicators from three years previously. To do so, we matched the financial indicators with the corporate performance one year later, and matched the technical indicators with the corporate performance three years later, as shown in Figure 5.

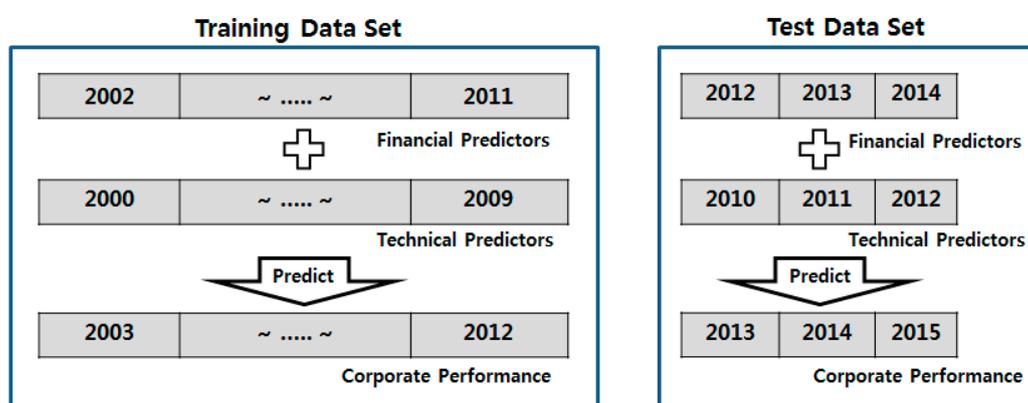


Figure 5. Summary of the training and test data sets.

In order to verify and analyze the prediction performance of the proposed model empirically, the model uses the financial indicators of 2002–2011, the technical indicators of 2000–2009, and the corporate performance of 2003–2012 as the training data set. The neural network architecture of the proposed model has six layers, consisting of an input layer with 15 visible nodes, four hidden layers each with 200 hidden nodes, and an output layer.

In the pre-training phase with a RBM, the financial indicators of 2002–2011 and technical indicators of 2000–2009 are used to train the model. In the RBM training parameters, the learning rate is set to 0.85 and the number of iterations (epochs) to 150,000.

For fine-tuning, the pre-trained weight and bias parameters are set as the initial values of the neural network. Then, the output layer is placed on top of the pre-trained network to make a basic FNN architecture. In order to better reflect recent relations between the predictors and corporate performance, we fine-tune the parameters of the neural network using a backpropagation algorithm. Thus, the model predicts the corporate performance of 2009–2012 using the financial indicators of 2008–2011 and the technical indicators of 2006–2009. In the training parameters of the FNN with a backpropagation algorithm (BP-FNN), the learning rate was set to 0.85, and the number of iterations (epochs) to 250,000.

#### 4.3. Experiment Results

In order to verify and analyze the prediction performance of the proposed model empirically, we predicted the corporate performance (revenue, operating profit, and net profit) of 2013–2015 using the trained model and the test data set (i.e., the financial indicators of 2012–2014 and the technical indicators of 2010–2012), as shown in Figure 5.

The root mean square error (RMSE) value between the actual and predicted revenue for 2013–2015 is 3146.44M. The proposed model predicts a company's revenue accurately.

The RMSE value between the actual and predicted operating profit for 2013–2015 is \$2876.36M. The proposed model is relatively accurate in predicting operating profit compared to other machine learning algorithm-based models. However, because operating profit is, in general, more volatile than revenue, the model predictions are not as accurate as in the case of revenue.

The RMSE value between the actual and predicted net profit for 2013–2015 is \$2783.44M. If, compared to the previous year, the net profit improves or falls sharply, the predictability of the proposed model decreases. However, the proposed model predicts net profit more accurately than other machine learning-based models do.

In order to compare the predictability of the proposed model to that of existing algorithm-based models, we used the DBN, FNN, and support vector regression (SVR) learning algorithms to construct corporate performance prediction models using the same training data. Then, we estimated the prediction performance of the four models using the RMSE, as in related studies [43]. The RMSE values between the actual and predicted corporate performance values for the models, using the same test data, are shown in Table 3.

**Table 3.** The root mean square error (RMSE) values of established prediction models.

|                  | The Proposed Model | DBN     | FNN     | SVR     |
|------------------|--------------------|---------|---------|---------|
| Revenue          | 3146.44            | 3841.88 | 5077.01 | 4798.97 |
| Operating Profit | 2876.36            | 3880.24 | 4288.95 | 3707.85 |
| Net Profit       | 2783.44            | 3681.07 | 4015.48 | 3514.35 |

In general, the prediction performance of SVM-based models is superior to that of neural network-based models in forecasting data with high volatility such as net profit and operating profit [3,7,32,33,35,37]. However, as shown in Table 3, the prediction performance of the proposed DBN-based model is the most accurate. Its performance is 1.3–1.5 times better than that of the SVR-based model, which has shown good prediction performance in many studies [3,7,32–37]. Furthermore, a comparison of the RMSE of the proposed model and that of the general DBN-based model shows that fine-tuning the pre-trained model using relatively recent training data leads to better predictions of time-series data.

## 5. Results

Many recent studies have attempted to predict the performance of companies quantitatively using machine learning algorithms. Many of these predict the bankruptcy of companies, or increases and decreases in their stock prices. In this study, we proposed a model that predicts a company's performance in terms of revenue, operating profit, and net profit. We used a RBM for pre-training, and a backpropagation algorithm for fine-tuning, similarly to the way in which a DBN is trained.

We constructed the proposed model using financial and patent data on 22 bio-pharmaceutical companies listed on the US stock market, and then verified the prediction performance of the model empirically. The results suggest that the proposed model shows good prediction performance. We also constructed corporate performance prediction models using various existing algorithms with the same training data set to compare the predictability between them. The results of the comparison show that the prediction performance of the proposed model is superior to that of the models constructed using the DBN, SVR, and FNN algorithms. Furthermore, the prediction accuracy of time-series data can be improved by using relatively recent data when fine-tuning the pre-trained network with a RBM.

The proposed model shows a 1.3–1.5 times reduction in the prediction error compared to that of the SVR-based model, which many studies have shown to have excellent prediction performance because of its high generalization performance. In particular, the proposed model predicts the performance of firms that record earnings surprises or shocks better than SVR-based models do. This shows that the

proposed model has sustainable predictability while the general prediction models show degradation of prediction performance over time.

In future research, we expect that the accuracy of the proposed model will be higher if it includes R&D investment data, patent citation data, and patent valuations as predictors. In addition, it is necessary to construct a model and analyze its performance using long-term–short-term memory and a recurrent neural network, which show excellent performance in predicting time-series data.

Although using a deep neural network such as a DBN shows high prediction performance, it takes much longer to train than do models based on shallow learning algorithms, such as a SVR. Therefore, algorithms that can train a deep neural network in less time are necessary. It is also necessary to develop algorithms that can determine the optimal learning parameters, such as the learning rate and the number of iterations, within a required prediction accuracy.

**Acknowledgments:** This research was supported by the Basic Science Research Program through the National Research Foundation of Korea, funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1D1A1A01059742). This research was also supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT & Future Planning (NRF-2017R1A2B1010208). Lastly, this research was supported by the BK 21 Plus (Big Data in Manufacturing and Logistics Systems, Korea University).

**Author Contributions:** Joonhyuck Lee and Sangsung Park conceived and designed the experiments; Dongsik Jang analyzed the data to show the validity of this study; Joonhyuck Lee wrote the paper and performed the entire research steps. In addition, all authors have cooperated with each other in revising the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ince, H.; Trafalis, T.B. Short-term forecasting with support vector machines and application to stock price prediction. *Int. J. Gen. Syst.* **2008**, *37*, 677–687. [[CrossRef](#)]
2. Estrella, A.; Mishkin, F.S. Predicting US recessions: Financial variables as leading indicators. *Rev. Econ. Stat.* **2006**, *80*, 45–61. [[CrossRef](#)]
3. Hsu, S.; Hsieh, J.; Chih, T.; Hsu, K. A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. *Expert Syst. Appl.* **2009**, *36*, 7947–7951. [[CrossRef](#)]
4. Ahn, B.; Cho, S.; Kim, C. The integrated methodology of rough set theory and artificial neural network for business failure prediction. *Expert Syst. Appl.* **2000**, *18*, 65–74. [[CrossRef](#)]
5. Yeh, C.; Chi, D.; Hsu, M. A hybrid approach of DEA, rough set and support vector machines for business failure prediction. *Expert Syst. Appl.* **2010**, *37*, 1535–1541. [[CrossRef](#)]
6. Huang, S.; Tsai, C.; Yen, D.C.; Cheng, Y. A hybrid financial analysis model for business failure prediction. *Expert Syst. Appl.* **2008**, *35*, 1034–1040. [[CrossRef](#)]
7. Huang, W.; Nakamori, Y.; Wang, S. Forecasting stock market movement direction with support vector machine. *Comput. Oper. Res.* **2005**, *32*, 2513–2522. [[CrossRef](#)]
8. Ang, J.S.; Chua, J.H.; Sellers, R. Generating cash flow estimates: An actual study using the Delphi Technique. *Financ. Manag.* **1979**, *8*, 64–67. [[CrossRef](#)]
9. Doyle, P.; Fenwick, I. Sales forecasting—Using a combination of approaches. *Long Range Plan.* **1976**, *9*, 60–64. [[CrossRef](#)]
10. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, Pittsburgh, PA, USA, 27–29 July 1992; pp. 144–152.
11. Hinton, G.E.; Osindero, S.; Teh, Y.W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [[CrossRef](#)] [[PubMed](#)]
12. Ribeiro, B.; Lopes, N. Deep belief networks for financial prediction. In *Neural Information Processing, Proceedings of the International Conference on Neural Information Processing, Shanghai, China, 13–17 November 2011*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 766–773.
13. Shen, F.; Chao, J.; Zhao, J. Forecasting exchange rate using deep belief networks and conjugate gradient method. *Neurocomputing* **2015**, *167*, 243–253. [[CrossRef](#)]

14. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
15. Karpathy, A.; Toderici, G.; Shetty, S.; Leung, T.; Sukthankar, R.; Fei-Fei, L. Large-scale video classification with convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1725–1732.
16. Shane, H.; Klock, M. The relation between patent citations and Tobin’s Q in the semiconductor industry. *Rev. Quant. Financ. Account.* **1997**, *9*, 131–146. [[CrossRef](#)]
17. Hall, B.H.; Jaffe, A.; Trajtenberg, M. Market value and patent citations. *Rand J. Econ.* **2005**, *36*, 16–38.
18. Lanjouw, J.O.; Schankerman, M. Patent quality and research productivity: Measuring innovation with multiple indicators. *Econ. J.* **2004**, *114*, 441–465. [[CrossRef](#)]
19. Jun, S.; Park, S.; Jang, D. Technology forecasting using matrix map and patent clustering. *Ind. Manag. Data Syst.* **2012**, *112*, 786–807. [[CrossRef](#)]
20. Thoma, G. Composite value index of patent indicators: Factor analysis combining bibliographic and survey data sets. *World Pat. Inf.* **2014**, *38*, 19–26. [[CrossRef](#)]
21. Martino, J. *Technology Forecasting for Decision Making*, 3rd ed.; McGraw-Hill: New York, NY, USA, 1993; pp. 93–96.
22. Moge, M. Using patent data for technology analysis and planning. *Res. Tech. Manag.* **1991**, *34*, 43–49.
23. Urbano, D.; Toledano, N.; Ribeiro-Soriano, D. Socio-cultural factors and transnational entrepreneurship: A multiple case study in Spain. *Int. Small Bus. J.* **2011**, *29*, 119–134. [[CrossRef](#)]
24. Soriano, D.R.; Huarng, K.H. Innovation and entrepreneurship in knowledge industries. *J. Bus. Res.* **2013**, *66*, 1964–1969. [[CrossRef](#)]
25. Mas-Tur, A.; Soriano, D.R. The level of innovation among young innovative companies: The impacts of knowledge-intensive services use, firm characteristics and the entrepreneur attributes. *Serv. Bus.* **2014**, *8*, 51–63. [[CrossRef](#)]
26. Del Mar Benavides-Espinosa, M.; Ribeiro-Soriano, D. Cooperative learning in creating and managing joint ventures. *J. Bus. Res.* **2014**, *67*, 648–655. [[CrossRef](#)]
27. Parellada, F.S.; Soriano, D.R.; Huarng, K.H. An overview of the service industries’ future (priorities: Linking past and future). *Ser. Ind. J.* **2011**, *31*, 1–6. [[CrossRef](#)]
28. Yoon, Y.; Swales, G. Predicting stock price performance: A neural network approach. In Proceedings of the IEEE Twenty-Fourth Annual Hawaii International Conference on System Sciences, Kauai, HI, USA, 8–11 January 1991; pp. 156–162.
29. Lam, M. Neural network techniques for financial performance prediction: Integrating fundamental and technical analysis. *Decis. Support Syst.* **2004**, *37*, 567–581. [[CrossRef](#)]
30. Pascanu, R.; Mikolov, T.; Bengio, Y. On the difficulty of training recurrent neural networks. *ICML* **2013**, *3*, 1310–1318.
31. Li, H.; Sun, J. Predicting business failure using multiple case-based reasoning combined with support vector machine. *Expert Syst. Appl.* **2009**, *36*, 10085–10096. [[CrossRef](#)]
32. Yuan, F. Parameters optimization using genetic algorithms in support vector regression for sales volume forecasting. *Appl. Math.* **2012**, *3*, 1480–1486. [[CrossRef](#)]
33. Jin, X.; Zhang, Y.; Yao, D. Simultaneous prediction of network traffic flow based on PCA-SVR. *Adv. Neural Netw.* **2007**, *4492*, 1022–1031.
34. Son, H.; Kim, C.; Kim, C. Hybrid principal component analysis and support vector machine model for predicting the cost performance of commercial building projects using pre-project planning variables. *Automat. Constr.* **2012**, *27*, 60–66. [[CrossRef](#)]
35. Chen, K.; Wang, C. Support vector regression with genetic algorithms in forecasting tourism demand. *Tour. Manag.* **2007**, *28*, 215–226. [[CrossRef](#)]
36. Chen, K. Forecasting systems reliability based on support vector regression with genetic algorithms. *Reliab. Eng. Syst. Saf.* **2007**, *92*, 423–432. [[CrossRef](#)]
37. Lee, J.; Kim, G.; Park, S.; Jang, D. Hybrid corporate performance prediction model considering technical capability. *Sustainability* **2016**, *8*, 640. [[CrossRef](#)]
38. Huang, W.; Song, G.; Hong, H.; Xie, K. Deep architecture for traffic flow prediction: Deep belief networks with multitask learning. *IEEE Trans. Intell. Transp. Syst.* **2014**, *15*, 2191–2201. [[CrossRef](#)]

39. Kuremoto, T.; Kimura, S.; Kobayashi, K.; Obayashi, M. Time series forecasting using a deep belief network with restricted Boltzmann machines. *Neurocomputing* **2014**, *137*, 47–56. [[CrossRef](#)]
40. Bengio, Y. Learning deep architectures for AI. *Found. Trends Mach. Learn.* **2009**, *2*, 1–127. [[CrossRef](#)]
41. Mohamed, A.R.; Dahl, G.; Hinton, G. Deep belief networks for phone recognition. In Proceedings of the NIPS Workshop on Deep Learning for Speech Recognition and Related Applications, Whistler, BC, Canada, 12 December 2009; p. 39.
42. Troiano, L.; Kriplani, P. Predicting trend in the next-day market by Hierarchical Hidden Markov Model. In Proceedings of the International Conference on Computer Information Systems and Industrial Management Applications, CISIM, Krakow, Poland, 8–10 October 2010; pp. 199–204.
43. Yu, J. Forecasting volatility in the New Zealand stock market. *Appl. Financ. Econ.* **2002**, *12*, 193–202. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).