Using Long Short-Term Memory to Predict Cash Dividend

Chu-Hui Lee*
Department of Information Management, Chaoyang University of Technology, Taiwan

Chia-Ling Hsu
Department of Information Management, Chaoyang University of Technology, Taiwan

ABSTRACT

Investing in stocks has been very popular in recent years. Investors hope to make a profit by investing in stocks. However, stock prices are highly volatile. Many investors judge whether to invest in stocks based on historical stock prices, technical indicators and market conditions. There are many researches about stock price forecasting in the research field but seldom for cash dividend. The profitable methods of investing in stocks include earning the stock spread and the cash dividends that the company distributes to shareholders. However, cash dividends are major profit for the long-term investor. In this research, we proposed a Stacked Long Short-Term Memory model to predict the value of the company's annual distribution of cash dividends that provided a basis for managers and investors to make investment decisions. The experiments showed the accuracy of prediction for Cash Dividend of the Formosa Plastics Corporation is excellent that the MAPE is only 2.95%.

CCS CONCEPTS

• **Computing methodologies**; • **Machine learning**; • **Machine learning approaches**; • **Neural networks**;

KEYWORDS

Time Series Prediction, Predict stocks, Cash dividend, Long Short-Term Memory

ACM Reference Format:

Chu-Hui Lee* and Chia-Ling Hsu. 2021. Using Long Short-Term Memory to Predict Cash Dividend. In *2021 5th International Conference on E-Education, E-Business and E-Technology (ICEBT 2021), June 18–20, 2021, Beijing, China.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3474880.3474898

1 INTRODUCTION

According to the Taiwan Stock Exchange Corporation Accumulated Investor Accounts statistics table, the number of new investor accounts is increasing yearly in the Taiwan stock market. In the past five years, an average increase of four hundred thousand accounts. In 2019, a single year increase of six hundred thousand accounts and the cumulative total number of accounts is 19 million [1]. It shows that many investors use stocks as financial management tools. There are two profit methods for investing in stocks. The first is to earn the short-term spread when buying and selling stocks, which is the selling price minus the buying price. The second is to earn dividends per share that the company distributes its annual earnings to shareholders. Dividend distribution methods are divided into cash dividends and stock dividends. When investors judge whether to buy or sell stocks, they are susceptible to the new or revised policies of government announcements, international economy trade, the company's operating conditions and profitability will also affect stock prices.

With the rapid development of information technology, investors can quickly view real-time information on global stock markets through online platforms. Many researchers use online platforms to collect daily stock price data in the stock market for stock technical indicator analysis and price prediction. In the analysis of stock technical indicators, daily stock prices are used to calculate technical indicators, such as Moving Average Convergence Divergence, Relative Strength Index, Bias Ratio, etc. In computer science, many re-searchers use daily stock prices and stock technical indicators to analyze and predict stock prices. P. F. Pai et al. [2] proposed Autoregressive Integrated Moving Average and Support Vector Machines mixed model, which solved the problem of nonlinear regression in time series and the problem of predicting stock prices, effectively. J. K. Bae [3] combined a decision tree, Multi-Layer Perceptron and Support Vector Machine to establish a model for predicting dividend policy and compare accuracy and performance. This model provided that the company should increase or decrease the dividend distribution depending on increasing the company's value and profits.

In neural network models, the Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) models are popular for prediction. These models can solve long-term dependent problems and are widely used to predict nonlinear time series. In RNN, with the increase of data, the problem of gradient disappearance will become more and more serious. And then LSTM was evolved. It's suitable for dealing with long-term sequence problems. It can preserve and process the previous information and solve the problem of gradient disappearances in RNN. K. A. Althelaya et al. [4] used a Bidirectional and Stacked LSTM model to predict the Standard & Poor's 500 index close price. The re-search results show that the stacked LSTM model has better performance and convergence in prediction.

In the stock investment market, many researchers have applied machine learning and neural network models to predict stock prices. These researches used different input features or combine various training models to improve the accuracy of the prediction of stock prices. In addition, many researches use the application of information technology to forecast dividend policy. It mainly discusses the impact of dividend policy on stock prices and the company's future earnings or predicts that the company should increase or de-crease the distribution of dividends but seldom to predict cash dividends that are a major profit for a long-term investor.

ICEBT 2021, June 18–20, 2021, Beijing, China

ACM ISBN 978-1-4503-8960-0/21/06…$15.00
https://doi.org/10.1145/3474880.3474898

To sum up, this research selects Formosa Plastics Co., Ltd. as a research example from the Taiwan stock market to predict the trend of the company's annual distribution of cash dividends. Formosa Plastics Co., Ltd. is a Taiwanese traditional industry company engaged in the production and processing of plastic materials. Its business has expanded into various industries, such as petrochemical, oil refining, textile, fiber, medical, education and biotechnology, etc. From 2000 to 2020, Taiwan experienced three global crises, the Internet bubble in 2000, SARS in 2003 and the financial crisis in 2008. Its revenue is still the first and second prize among Taiwanese companies. Therefore, this research proposed to use the Stacked LSTM model to predict the trend of the company's annual distribution of cash dividends. A reference basis for managers and investors was provided to make investment decisions.

This paper is organized as follows. In Section II, related technologies and research on Time Series Analysis and LSTM is discussed. Understanding the methods, advantages, disadvantages and applications are used by researchers in solving time series problems are presented. In section III, the research LSTM structure and research process used in this research is described. In section IV, experimental results are demonstrated. Section V is the conclusion of this research.

2 RELATED WORKS

2.1 Time series analysist

From the past to the present, different data are generated due to different time, and the data is sorted according to time. The time unit of the time series can be divided into seconds, minutes, hours, days, months, years, etc. A collection of multiple different time series is called time series data. Using time series data to analyze the linear changes of features in the sequence and finding important information from past historical data can be used to predict future information.

Time series analysis is widely used in various fields. For example, R. C. Chuentawat et al. [5] proposed a method based on a support vector machine and genetic algorithm for univariate and multivariate time series forecasting PM2.5. This research used hourly temperature, pressure, and cumulative wind speed time series data for analysis. The results of the study have lower errors in univariate time series than the multivariate time series model. D. Liu et al. [6] proposed a time series analysis model based on support vector machines to predict energy consumption of buildings. This model used hourly energy consumption data to establish a predictive model applied to buildings with different consumables. Experimental results show that machine learning methods are better than traditional time series forecasting models. Y. Zhang et al. [7] proposed a risk management framework based on time series analysis. The framework includes risk assessment, dynamic risk management and risk prediction. Risk prediction used time series models to study changes in risk and predict future risks.

Stock forecasting is the category of time series analysis. There are many different disciplines to research methods of predicting stocks such as Finance, Statistics and Computer Science. In computer science, it can be divided into traditional models, machine learning models and neural network models.

In traditional models, Autoregressive Integrated Moving Average (ARIMA) and Generalized-Autoregressive Conditional Heteroskedasticity (GARCH) assume that a linear correlation structure exists in time series values. P. W. Tsai et al. [8] used GARCH to analyze the impact of signing international agreements on Taiwan stock market. The impacts of Taiwan's signing of international agreements with other countries on Taiwan's stock prices were explored. The time series is divided into the period before and after the signing of the agreement to analyze the stock market price. S. M. Idrees et al. [9] used ARIMA to predict the trend of the India stock market. This research collects five years of historical data on stocks including daily opening price, lowest price, closing price, highest price, adjusted closing price, and transaction volume of two Indian stock exchange index-es. The average percentage error between the predicted value and the actual value of the training result is 5%.

In machine learning, W. Huang et al. [10] used economic indicators such as the term structure of interest rates, short-term interest rate, long-term interest rate, Consumer Price Index (CPI), Gross National Product (GNP), Gross Domestic Product (GDP) as the features of the training model. This research proposed that support vector machines combined with classification methods to predict the weekly fluctuation direction of the Nikkei index can be more accurate than support vector machines. J. Mager et al. [11] proposed support vector machine combined with a dynamic time warping algorithm to predict the closing price of the financial stock market.

In deep learning, J. Wang et al. [12] proposed a case study of crude oil price fluctuations. Recurrent Neural Network (RNN) was used to predict energy market index. This research sets different model learning rates and training iterations for different datasets and uses linear regression analysis to compare predicted values with actual prices. S. Das et al. [13] collects real-time sentiment from Twitter financial news and inputs it to a RNN model to predict the stock prices of Google, Apple, and Microsoft. Z. Berradi et al. [14] combined Principal Components Analysis (PCA) and RNN to predict the stock price of Casablanca Stock Exchange. This research used PCA to reduce the dimensionality of the data. The original number of eight features is reduced to six features and input to the RNN model to predict the closing price.

2.2 Application of Dividend Policy

The dividend policy is determined by the company's board of directors that determines the method of dividend payment, the number of dividends to be paid and the date of dividend payment. Dividends are the surplus earned by the company from investment and operations, which are distributed to investors. Investors can understand the company operating conditions through the announced dividend policy.

To sum up, the dividend policy issued by the company will affect the amount of profit and investment confidence of investors. Therefore, there have been researches on predicting dividend policies in the stock investment market. It enables investors to understand the impact of dividend payment on stock prices or financial indicators and predicts that the company will increase or decrease the distribution of dividends in the future. C. Won et al. (2012) [15] proposed Genetic Algorithm Knowledge Refinement (GAKR) to predict the dividend policy. The input features are current dividend and past and current stock prices. Target feature is transformed dividend policy into increase or decrease. This research combined the traditional rule technology and genetic algorithms to compare the accuracy of CHAID, CART, QUEST and C5.0 classifiers. The results of the GAKR are better than other classifiers models in predicting the dividend policy. D. Nissim et al. (2001) [16] researched the relationship between dividend volatility and the company earnings capacity. The results of the research found that the dividend changes are positive related to the company earnings changes in the future. The total dividend changes are positively related to earnings anomalies in the future. After controlling the book value, the dividend changes can reflect the future profitability of the company.

2.3 Long Short-Term Memory

S. Hochreiter and J. Schmidhuber proposed a Long Short-Term Memory (LSTM) model in 1997 [17]. The LSTM model is composed of input gates, forget gates, output gates and cell state. The input gate determines which values to update to the cell state after calculating the input characteristic value. The forget gate determines how much previous information should be forgotten for the cell status. The output gate determines which part of the cell state should be output after calculation. The cell status will remember the previous information and be updated with the new information. LSTM is suitable for dealing with long-term sequence problems. It can save and process previous information and solve the problem that the gradient disappears due to the excessive amount of data in RNN. Y. Liu et al. [18] proposed a model combining Gated Recurrent Unit (GRU) and LSTM to predict the closing price of stocks. This model is composed of two GRU layers, an LSTM layer and a fully connected layer which is called the normalized GRU-LSTM model. Regularization is a regression which can reduce the complexity and instability of the model in the learning process. A. Moghar et al. [19] used four LSTM layers and two Dropout layers models to predict the opening prices of GOOGL and NKE stocks in the New York stock market. This research compares the loss values of different training epochs. Dropout procedure can be used to discard some neurons when the model is overfitting in training. S. Jain et al. [20] used Convolutional Neural Network (CNN) for data preprocessing to extract important features. The results with LSTM were combined to predict the stock price of two listed companies on the National Stock Exchange of India.

Stock forecasting is a type of time series analysis. Future information was predicted through sequential information in the past. Its analysis methods include traditional models, machine learning and deep learning. The studies listed above have successfully applied these methods to predict stock prices. In this research, we use the LSTM model to predict the trend of annual dividend quotas for listed companies in Taiwan.

3 PROPOSED METHOD

This research focuses on the company's annual cash dividends when investing in stocks. This paper proposed a prediction model that used Long Short-Term Memory models to predict the company's distribution of cash dividends. Company stock price and monthly financial statement data was collected. A long and short-term memory model was established to predict cash dividends after data pre-processing.

3.1 Data collection

The source of data collection is various securities information provided by the Taiwan Stock Exchange Corporation (TWSE). TWSE is the operating organization of the Taiwan Securities Exchange Market [21]. Its purpose is to maintain market order and transaction security, strengthen risk management operations for securities firms, provide investors with real-time trading information, operate overview of listed companies and financial statements, etc.

3.2 Data calculation

The company allocates cash dividends once a year most often. However, this re-search uses monthly stock prices, technical indicators and financial statement information to train the model. When approaching the ex-dividend trading day, the predicted cash dividend should be closer to the actual cash dividend distributed. Therefore, the conversion of annual cash dividends into monthly cash dividends is expressed linearly.

Input features include stock prices, stock technical indicators and financial statement information are shown in Table 1. The monthly stock price characteristics were calculated as the opening price on the first day of the month, the closing price on the last day of the month, the highest price of the month, and the lowest price of the month. Stock technical indicators use monthly Earnings Per Share (EPS) and Price-to-Earnings Ratio (PER). EPS is to understand the company's profitability. That is how much you can earn by holding one share. PER determines whether the stock price is expensive or cheap. When the company's profits decrease, after-tax net profit and earnings per share decrease, the P/E ratio will increase. It means that investors need to spend more time to make a profit.

The output feature of this research is cash dividends. The company distributes cash dividends from operating profit to shareholders. Cash dividends are allotted once a year shown in Table 2. The annual cash dividend distribution is based on the Ex-Dividend Date, and the annual cash dividend is converted into monthly cash dividend shown in Figure 1. The increasing or decreasing rate was this year's dividend to subtract last year's dividend divided by twelve months. Cash dividends are transferred to linear instead of step function that makes the predicted information more practical and more accurate.

3.3 Data normalization

This research uses the Min-Max Normalization method to compress the value of each feature in the data set to the interval of 0 to 1. The calculation method is the original value to subtract the minimum value divided by the maximum value to subtract the minimum value, as in (1). It can use the MinMaxScaler in the Scikit-learn package for data normalization.

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \in [0, 1] \tag{1}$$

Table 1: Input features description

Feature	Description
Open price	Opening price on the first day of each month.
Close price	Closing price on the last day of each month.
Max price	Highest monthly price.
Min price	Lowest monthly price.
Earnings Per Share (EPS)	Formula: Net Income / Outstanding Shares. Understand the company's profitability. That is how much you can earn by holding one share.
Price-to-Earnings Ratio (PER)	Formula: Price / EPS. Determine whether the stock price is expensive or cheap.
Revenue	Revenue from the company providing products or services.

Table 2: Input features description

Year	Cash dividend (dollar)
2008	1.8
2009	4
2010	6.8
2011	4
2012	1.2
2013	1.9
2014	1.7
2015	3.6
2016	4.6
2017	5.7
2018	5.8
2019	4.4

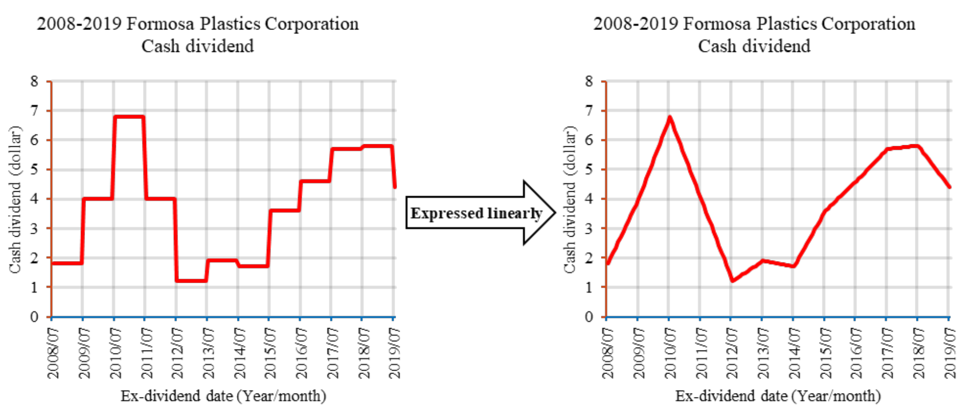

Figure 1: Cash dividend is expressed linearly

3.4 Construction of Sliding Window

The purpose of constructing the sliding window is to learn the characteristics of each time series in historical data when the model is training. This research uses a sliding window to divide the data into three time-step sequence data and establishes an overlap-ping time series using a sliding window to predict cash dividends for the next month.

3.5 Construction of Long Short-Term Memory

The LSTM model loads data for three time steps each time during training and establishes a cell state to learn information from the past to the current mission. This research proposes a Stacked-LSTM model. Our Stacked-LSTM model is composed of the input layer, two LSTM layers with sigmoid and tanh activation functions and dense layer with linear activation function as shown in Figure 2 [22]. There are three types of information for each layer that are

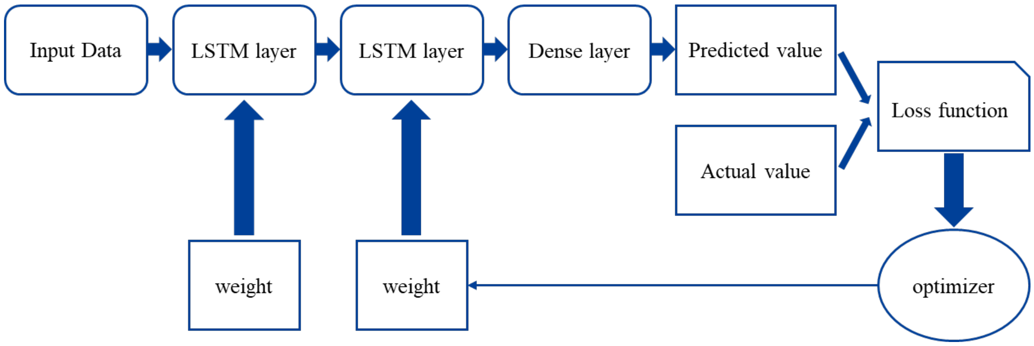

Figure 2: Structure of the proposed Stacked-LSTM model

Table 3: Parameter of the model

Parameter	Value
Learning rate	0.001
Batch size	32
Epochs	350
Loss function	MSE
Optimizer	Adam

input value, output value and cell state. After the input value is multiplied by the activation functions and weights, the output value is sent to the next state. In addition to the model being affected by the number of model layers, the parameter setting is important shown in Table 3. The weight is calculated with the input data to find the correct weight value. The loss function evaluates the error between the predicted value and the actual value and measures the output quality of the neural network. The optimizer is to fine-tune the weights of each layer to reduce the loss score of each learning.

3.6 Evaluation Standard

The predictive performance of the LSTM model is evaluated. This study uses Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) to measure the error between actual and predicted values. MAE is to calculate the absolute error between the predicted value and the actual value. The smaller error of these evaluation functions is better as in (2). MAPE is the absolute error between the predicted value and the actual value divided by the actual value. It measures the ratio of error and actual value. If the MAPE is closer to 0%, the prediction accuracy of the model is better as in (3). RMSE calculates the deviation between the predicted value and the actual value. It can be used to measure the precision of the prediction results as in (4). The formulas are used for evaluation standards, where $\hat{y}_i$ is prediction value and y_i is actual value.

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|\hat{y}_i - y_i| \tag{2}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{3}$$

$$RMSE = \sqrt{\sum_{t=1}^{n}\frac{(\hat{y}_i - y_i)^2}{n}} \tag{4}$$

4 EXPERIMENTS AND RESULTS

This research used Python language to build predictive models. The Scikit-learn package was used to normalize the data. The NumPy library was used to divide the data set into training set and validation set and performs dimensional array operations on the data. TensorFlow library and Keras neural network library was used to build a Stacked-LSTM model in Python 3.6.10 environment.

For the dataset, we collected daily stock prices, technical indicators and monthly financial statements of Formosa Plastics Corporation from July 2008 to June 2019 at TWSE. The monthly maximum, minimum or average values were calculated as the features of the training and testing model with the sliding window size of 3 months. We collected 132 samples of data for 11 years. From the dataset 108 samples are used for the training set and 24 samples for the validation set. This dataset contains 7 features which belong to 3 different types to predict cash dividends and compare their error values. The first feature type is stock prices that have four features. It includes open price, close price, max price and min price. The second feature type is technical indicators that have two features. It includes EPS and PER. EPS is used to understand the profitability of the company. PER is based on the company's stock market price to determine the amount of investment that will take several years to pay back. Both of these indicators are related to whether the company makes a profit. The third feature type is financial feature that is revenue. It means the income from the company providing products or services. When the company's revenue becomes higher, it means that the company's profit is growing. The company's annual cash dividend may increase. Hence the level of profit is related

Table 4: Comparative result using different features with Stock prices, Technical indicators and Financial features

Parameter	Error Comparison		
	MAE	MAPE	RMSE
Stock prices	0.4636	8.6761%	0.5423
Stock prices and Technical indicators	0.2185	4.0869%	0.2730
Stock prices, Technical indicators and Financial	0.1623	2.9548%	0.2057

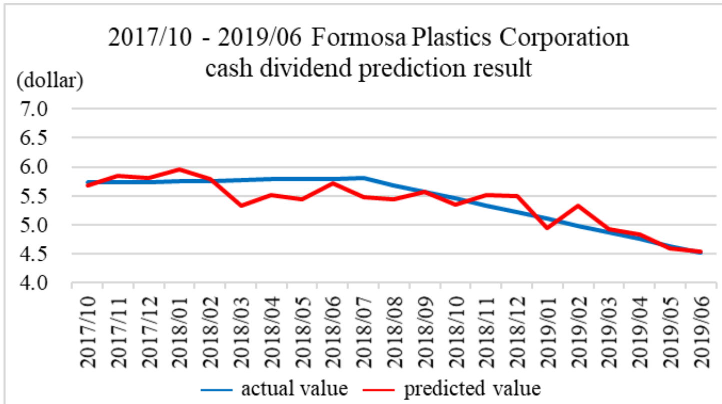

Figure 3: Cash dividend prediction result

to the level of the company's cash dividend. After collecting feature data, each vector of the sequence is normalized and converted into three-dimensional data which are samples, timesteps, and features. This research model has 105 samples with 3 timesteps and 7 features. The training data into the Stacked-LSTM model was inputted. The number of hidden layer nodes and activation function of the model was set. This research used the Stacked-LSTM model. Therefore, the return-sequences parameter needs to be set to true. The two-dimensional data output from the first LSTM layer is converted into three-dimensional that is the input for the second LSTM layer. In the training process, if the data are overfitting, a dropout parameter is set to discard some of the node data and the activation function of the dense layer is set. The MSE loss function is used to calculate the weight and the Adam optimizer function to update the weight of the LSTM model. The predicted results use three different types of evaluation functions including MAE, MAPE and RMSE. The errors of using different features to predict cash dividends were compared.

In Table 4, the research experiment compared the prediction results of different feature combinations. The first feature combination used stock prices. This result is MAE of 0.4636, MAPE of 8.6761, and RMSE of 0.5423. The second feature combination used stock prices and Technical indicators. This result is MAE of 0.2185, MAPE of 4.0869, and RMSE of 0.2730. The third feature combination used Stock prices, Technical indicators and Financial. This feature combination achieves the best results with a MAE of 0.1623, MAPE of 2.9548, and RMSE of 0.2057.

In Figure 3, it shows the prediction results of the Formosa Plastics Corporation for October 2017 to June 2019. The blue curve is the actual price and the red curve is the predicted price. When the blue curve is closer to the red curve, it means that the predicted price is closer to the actual value. It can be seen from Figure 3 that the error between the actual value and the predicted value is less than 0.5dollar cash dividend for each time. The average error is 0.16dollar cash dividend.

In Table 5, the experiment uses the best combination that includes all stock prices, technical indicators and financial features to compare the different size of sliding window and different LSTM layer. The error results of the sliding window with size of 3 months are lower than the sliding window with size of 1 month. Every 3 months is used as a time step to predict the cash dividend for the next month. Constantly observing the data of the previous 3 months reduces the error in predicting cash dividends. In the experiment, multiple recurrent layers are added and the numbers of units are increased in the layer can in-crease the capacity and performance of the LSTM. The error results of the Stacked-LSTM are lower than the single LSTM. The best model for this experiment used Stacked-LSTM with two layers and sliding window with size of 3 months.

5 CONCLUSIONS AND FUTURE WORK

Stock prices are highly volatile in the stock market. Therefore, the prediction of stock prices is challenging. There are many researches on predicting stock prices in the fields of finance, statistics and computer science. Researchers research the application of new techniques to predict the stock prices but seldom cash dividends. In recent years, when investors invest in the stocks market, they will consider whether the company's profits are stable. Because investors want to earn cash dividends issued by the company as one of the long-term investments and financial management methods.

Table 5: Comparative result using different Sliding Window and LSTM layer

Parameter	Sliding window size	Error Comparison		
		MAE	MAPE	RMSE
Single-LSTM	1	0.3105	5.6705%	0.3505
	3	0.2100	3.8359%	0.2517
Stacked-LSTM	1	0.3010	5.5755%	0.3574
	3	0.1623	2.9548%	0.2057

Therefore, this research collected stock prices, technical indicators and revenue in financial statements as the features of the training model. We used the Stacked Long Short-Term Memory model to predict the trend of the company's cash dividends. We used all the features to predict cash dividends to get the best prediction results and the predict-ed value is closer to the actual value. It can provide a reference basis for managers and investors to make investment decisions. The experiments show the prediction of our model is excellent. The MAPE of prediction for cash dividend of the Formosa Plastics Corporation is only 2.95%. In future research, we can increase the types of input features. For example, extracting sentiment features from financial news and other technical indicators. It enables the forecasting model to forecast cash dividends for companies in the same stock class.

ACKNOWLEDGMENTS

This research was supported by the Ministry of Science and Technology of Taiwan under Grant MOST-109-2221-E-324-023.

REFERENCES

[1] TWSE 2021. Number of Accounts Opened at Securities Companies Annual Statistics. Retrieved March 05, 2021 from https://www.twse.com.tw/zh/statistics/index/07

[2] P. F. Pai and C. S. Lin. 2005. A hybrid ARIMA and support vector machines model in stock price forecasting. Omega. 33, 6 (December 2005), 497-505. https://doi.org/10.1016/j.omega.2004.07.024

[3] J. K. Bae. 2010. Forecasting Decisions on Dividend Policy of South Korea Companies Listed in the Korea Exchange Market Based on Support Vector Machines. Convergence Information Technology. 5, 8 (Oct. 2010), 186-194. https://doi.org/10.4156/jcit.vol5.issue8.20

[4] K. A.Althelaya, E. El-Alfy and S. Mohammed. 2018. Evaluation of bidirectional LSTM for short-and long-term stock market prediction. In Proceedings of International Conference on Information and Communication Systems (ICICS). Irbid, Jordan, 151-156. 10.1109/IACS.2018.8355458

[5] R. Chuentawat and K. N. Yosporn. 2018. The Comparison of PM2.5 forecasting methods in the form of multivariate and univariate time series based on Support Vector Machine and Genetic Algorithm. In Proceedings of the 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). Chiang Rai, Thailand, 572-575. 10.1109/ECTICon.2018.8619867

[6] D. Liu, Q. Chen and K. Mori. 2015. Time series forecasting method of building energy consumption using support vector regression. In Proceedings of the IEEE International Conference on Information and Automation (ICIA). Lijiang, China. 1628-1632. 10.1109/ICInfA.2015.7279546

[7] Y. Zhang, X. Tan, H. Xi and X. Zhao. 2008. Real-time risk management based on time series analysis. In Proceedings of the 7th World Congress on Intelligent Control and Automation (WCICA). Chongqing, China. 2518-2523. 10.1109/WCICA.2008.4593320

[8] P. W. Tsai, Y. X. Lin, J. Zhang, J. F. Chang, J. S. Chen and J. Zhang. 2016. Time-Series Analysis on the Impact Level upon Taiwan Stock Market with Events of Signing International Agreements. In Proceedings of the 3rd International Conference on Computing Measurement Control and Sensor Network (CMCSN). Matsue, Japan. 126-129. 10.1109/CMCSN.2016.51

[9] S. Idrees, M. Alam and P. Agarwal. 2019. A Prediction Approach for Stock Market Volatility Based on Time Series Data. IEEE Access. 7 (January 2019), 17287-17298. 10.1109/ACCESS.2019.2895252

[10] W. Huang, Y. Nakamori and S. Y. Wang. 2005. Forecasting stock market movement direction with support vector machine. Computers & Operations Research. 32, 10 (October 2005), 2513-2522. 10.1016/j.cor.2004.03.016

[11] J. Mager, U. Paasche and B. Sick. 2008. Forecasting financial time series with support vector machines based on dynamic kernels. In Proceedings of the IEEE Conference on Soft Computing in Industrial Applications (SMCia). Muroran, Japan. 252-257. 10.1109/SMCIA.2008.5045969

[12] Jie. Wang and Jun. Wang. 2016. Forecasting energy market indices with recurrent neural networks: Case study of crude oil price fluctuations. Energy. 102 (May 2016), 365-374. 10.1016/j.energy.2016.02.098

[13] S. Das, R. K. Behera, M. Kumar and S. K. Rath. 2018. Real-Time Sentiment Analysis of Twitter Streaming data for Stock Prediction. Procedia Computer Science. 132 (June 2018), 956-964. 10.1016/j.procs.2018.05.111

[14] Z. Berradi and M. Lazaar. 2019. Integration of Principal Component Analysis and Recurrent Neural Network to Forecast the Stock Price of Casablanca Stock Exchange. Procedia Computer Science. 148 (February 2019), 55-61. 10.1016/j.procs.2019.01.008

[15] C. Wona, J. Kima and J. K. Bae. 2012. Using genetic algorithm based knowledge refinement model for dividend policy forecasting. Expert Systems with Applications. 39, 18 (December 2012), 13472-13479. 10.1016/j.eswa.2012.06.001

[16] D. Nissim and A. Ziv. 2002. Dividend Changes and Future Profitability. The Journal of Finance. 56, 6 (December 2002), 2111-2133. 10.1111/0022-1082.00400

[17] S. Hochreiter and J. Schmidhuber. 1997. Long Short-Term Memory. Neural Computation. 9, 8 (November 1997), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

[18] Y. Liu, Z. Wang and B. Zheng. 2019. Application of Regularized GRU-LSTM Model in Stock Price Prediction. In Proceedings of the 5th International Conference on Computer and Communications (ICCC). Chengdu, China. 1886-1890. 10.1109/ICCC47050.2019.9064035.

[19] A. Moghar and M. Hamiche. 2020. Stock Market Prediction Using LSTM Recurrent Neural Network. Procedia Computer Science. 170 (April 2020), 1168-1173. 10.1016/j.procs.2020.03.049

[20] S. Jain, R. Gupta and A. A. Moghe. 2018. Stock Price Prediction on Daily Stock Data using Deep Neural Networks. In Proceedings of the International Conference on Advanced Computation and Telecommunication (ICACAT). Bhopal, India. 1-13. 10.1109/ICACAT.2018.8933791

[21] TWSE 2021. Taiwan Stock Exchange Corporation Service. Retrieved March 05, 2021 from https://www.twse.com.tw/zh/page/about/company/service.html

[22] François Chollet. 2018. Deep Learning with Python. Manning Publications, USA.