A Comparative Analysis on Stock Price Prediction Model using DEEP LEARNING Technology

Rupashi Jaiswal¹
Department of Computer Science
University of Engineering & Management, Kolkata
Kolkata, India
rupashijaiswal@gmail.com

Kunal Mahato¹
Department of Computer Science
University of Engineering & Management, Kolkata
Kolkata, India
kunalmahato34@gmail.com

Pankaj Kapoor¹
Department of Computer Science
University of Engineering & Management, Kolkata
Kolkata, India
pankajkapoor580@gmail.com

Dr. Sudipta Basu Pal²
Department of CST & CSIT
University of Engineering & Management, Kolkata
Kolkata, India
sudipta_basu 68@yahoo.com

Abstract—

In today's world, Artificial Intelligence and Deep Learning are getting popular regularly. The various applications areas of artificial intelligence are related to human activity. One of the general application areas of neural networks and artificial intelligence is prediction analysis. In this paper, the authors also have performed one comparative study based on artificial intelligence. Authors have performed stock market predictions using different models. In reality, stock markets are entirely volatile, so there is very much a requirement of good prediction analysis for judging the stocks prices and their ups and downs with time. The stock prices can easily be predicted using machine learning algorithms on data available in financial news, as this data can also change investors' interests. However, traditional prediction methods have become obsolete and do not provide accurate predictions over non-stationary time series data. This paper proposes a stock price prediction method that gives accurate results with the advancements in deep learning technologies.

Keywords
Artificial Intelligence, ARIMA Model, LSTM Model, Prediction Analysis

I. INTRODUCTION

Stock is a financial product characterized by high risk, high return and flexible trading, which many favours. Investors can get great returns by accurately estimating stock price trends. However, the stock price is influenced by many factors such as macroeconomic situation, market condition, significant social and economic events, investors' preferences and companies' managerial decisions. Therefore, prediction of the stock price has always been the focus and challenging research topic. Statistical and econometric models are generally used in traditional stock price prediction, but these methods cannot deal with the dynamic and complex environment of the stock market. Since 1970, with the rapid development of computer technology, researchers have begun using machine learning to predict stock prices and fluctuations, helping investors determine investment strategies to reduce risk and increase returns.

The stock market is a highly complex time series scenario and has typical dynamic characteristics. There will be much active stock trading after the market's opening, and the stock price will change accordingly. Moreover, the stock price is affected by many unpredicted factors, which results in a typical non-stationary stock price time-series data. Therefore, stock price prediction is one of the most
challenging problems in all kinds of prediction research. In the past decades, scholars have studied stock price prediction from many perspectives. The improvement of prediction models and the selection of model features are the two most essential directions among them. Most of the early studies used econometric models, such as autoregressive integrated moving average (ARIMA) and autoregressive conditional heteroskedastic-autoregressive integrated moving average (ARCH-ARIMA) (Booth et al., 1994; Engle, 2001), to predict stock price. However, it is difficult for econometric models to consider the impact of other factors on stock price fluctuations. They have strong assumptions about the data, which are often difficult to meet (Le and Xie, 2018). Therefore, machine learning has been widely used in stock price prediction in recent years and many more suitable models for stock prediction have been proposed. Many studies have shown that deep learning has superior efficiency than other models (Marmer, 2008) and neural network models excel regression and discriminant models (Refenes et al., 1994).

Previous literature extensively investigates the stock price prediction methods and many advanced prediction models are proposed. However, many approaches on stock price prediction models have used RNNs because of their ability to retain memory of past records of sequential data. However, this approach has one main limitation. RNNs are trained by backpropagation. During backpropagation, RNNs suffer from gradient vanishing problem. The gradient is the value used to update Neural Networks’ weight. The gradient vanishing problem is when gradient shrink as it back propagates through time. Therefore, layers that get a small gradient do not learn and they cause the network to have short-term memory. Thus, this method is not suitable for the stock price prediction.

To fill the research gap discussed above, this paper uses LSTM to predict the stock prices. We have given opening, closing, high and low stock prices as input features, and LSTM is adopted to predict the stock price.

The rest of this paper is structured as follows. We review the literature on stock price prediction in Section 2 and introduce our method in Section 3. We explain the research data and experimental process in Section 4. Finally, we conclude the paper with a summary and possible future research directions in Section 5.

II. Deep Learning Models

In this paper authors predict the stock price based on deep learning models. The complete research work is mainly about the prediction model and feature selection of the prediction model. In the next sections the methodology depicted clearly.

A. A NEW STOCK PRICE PREDICTION METHOD

We propose a new stock price prediction model based on deep learning technology, which integrates LSTM model. It uses stock financial features to predict future stock price. The model mainly includes various complex tasks. We have developed this project into two parts –

- First, we will learn how to predict stock price using LSTM neural network
- Then we will build a dashboard using Plotly dash for stock analysis.

A. Feature selection

Our model uses financial features as input features. Financial features are descriptions of the daily trading data of stocks, which can reflect basic information of stock prices. Text features of social media refer to the text feature vector containing effective information in social media, which includes investors’ comments in financial social media and the news published by companies.

Financial features are described by daily transaction data and financial technical indicators (Bao et al., 2017) and include 21 features. Daily transaction data include open/close price, low/high price, trading volume, change amount and change rate, which are daily first-hand trading information in the stock market. These characteristics can directly reflect the historical trading situation of stocks. Financial technical indicators refer to indicators calculated based on stock trading data, including CCI (commodity channel index), ATR (average true range), SMI (stochastic momentum index), etc. These
indicators can often reflect some regular characteristics of stock movements. Table 1 gives the list of financial features.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily transaction data</td>
<td></td>
</tr>
<tr>
<td>Open/close price</td>
<td>Nominal daily open/close price</td>
</tr>
<tr>
<td>High/low price</td>
<td>Nominal daily highest/lowest price</td>
</tr>
<tr>
<td>Volume</td>
<td>Daily trading volume</td>
</tr>
<tr>
<td>Price_change</td>
<td>Change volume</td>
</tr>
<tr>
<td>P_change</td>
<td>Change rate</td>
</tr>
<tr>
<td>Technical indicators</td>
<td></td>
</tr>
<tr>
<td>MACD</td>
<td>Moving average convergence divergence: displays trends following characteristics and momentum characteristics</td>
</tr>
<tr>
<td>CCI</td>
<td>Commodity channel index: helps to find the start and end of trend</td>
</tr>
</tbody>
</table>
series modelling. Therefore, we choose it as the final prediction model.

LSTM is composed of multiple neurons. In each neuron, data first enters forget gate. The forget gate determines which input information will be forgotten so it will not affect the update of the next neuron. In the second step, the input gate decides which information is allowed to be added. The output of the previous neuron and input of the local neuron are processed by the sigmoid function and tanh function to generate two results. And then which information needs to be updated is decided based on these two results. The results will be saved for the output gate. Finally, the output gate determines which result obtained in the input gate can be generated. The results from the output gate of one neuron will be inputted to the next neuron, etc.

The input data of LSTM is a three-dimension array, representing time dimension, sample dimension and feature dimension, respectively. The time dimension represents the sliding time window, the sample dimension represents the sample size of training and testing and the feature dimension represents the number of input features. We choose 7 days as a time window to predict the close price on day 8. The input features are 21 dimensions, which are financial features.

A. EXPERIMENTAL Methodology

Data Acquisition

We chose to predict the stock price of Tata Beverages from Tata Global Beverages Limited, National Stock Exchange of India, Apple Inc, Microsoft Corporation, Facebook, Inc. and Tesla, Inc. That is because these are renowned companies which always make the headlines and are followed by the masses.

We also collected daily transaction data of these companies from various financial databases available, including open, close, high, low, trading volume, change volume and change rate. In addition, we calculated financial technical indicators based on daily transaction data (Bao et al., 2017).

B. Metrics

We used MAE, RMSE and MSE as measures to evaluate the performance of the prediction methods. MAE measures error without considering the directions of the predicted values. RMSE measures the average magnitude of estimation error in predicted values. MAE and RMSE are measures of closeness which evaluates the accuracy of the predicted value to the actual price. The MSE is the second moment (about the origin) of the error, and thus incorporates both the variance of the estimator (how widely spread the estimates are from one data sample to another) and its bias (how far off the average estimated value is from the true value). For an unbiased estimator, the MSE is the variance of the estimator. We hope the model has low MAE and RMSE and an MSE value near to zero (zero means the model is perfect, which is not possible). The three metrics are defined as follows:

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} | \hat{x}_i - x_i | \\
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \\
\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2
\]

where \(x, y\) represents the predicted value of the target variable of sample \(i\), and \(x, \hat{y}\) represents the predicted value of the target variable of sample \(i\), represents the mean of the target variable's true value of all samples and \(n\) represents the total number of samples.

C. Experiment

We train our LSTM model on stock data of Tata Beverages from Tata Global Beverages Limited. We use this trained model to predict the stocks of other companies that we have selected. We chronologically divided the text data of "Tata Beverages" into three parts, i.e. the first 80% data is train set, the middle 10% data is validation set and the last 10% data is test set. Normalization is a necessary step in many machine learning algorithms as the input data is expected to have different scaling. Our dataset is
uniform and clean, so we move forward to building and training the actual LSTM model that is going to perform the predictions. For this, we use the Adam Optimizer over others due to its ability to converge towards maxima easily and efficiently.

In the proposed prediction methods, LSTM is chosen as the prediction model, which also has some parameters to be identified. The most commonly used parameters of LSTM include hidden layers, dropout, number of neurons, optimizer, batch sizes and epochs. The hidden layer is set to 1–3 layers according to experience and the computing power of the machine. Dropout is usually between 0.2 and 0.5. According to Kolmogorov's theorem, the number of neurons in the hidden layer is set as double of the input dimensions plus one (Greff et al., 2017). Finally, we adjusted the parameters of LSTM where the optimal parameter combination was 2 hidden layers, 7-time windows, 85 neurons, 0.5 dropout, Adam optimizer, 1 batch size and 1 epoch. Now that our entire model is ready, we test the model on our test dataset. We have plotted the original dataset and the predicted stock market values together for years that fall under the 20% of the test dataset, which in this case was after the year 2014, that the model was not trained on.

IV. Result Analysis

The results shown in Fig 1, show that the predicted values and original values almost overlap each other. This indicates the high accuracy of the model that we have built. The metrics of the LSTM model (MAE = 0.0222, MSE = 9.2324E-04, RMSE = 0.03038) are better than the other baseline models. Figure 1 intuitively observes that the LSTM model fits the curve better, proving that the proposed model can effectively predict the fluctuation of stock prices. Table 2 summarizes these metrics and gives us an intuitive understanding on the accuracy of this model.
### Table 1 – Stock financial features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Daily transaction data</strong></td>
<td></td>
</tr>
<tr>
<td>Open/close price</td>
<td>Nominal daily open/close price</td>
</tr>
<tr>
<td>High/low price</td>
<td>Nominal daily highest/lowest price</td>
</tr>
<tr>
<td>Volume</td>
<td>Daily trading volume</td>
</tr>
<tr>
<td>Price_change</td>
<td>Change volume</td>
</tr>
<tr>
<td>P_change</td>
<td>Change rate</td>
</tr>
<tr>
<td><strong>Technical indicators</strong></td>
<td></td>
</tr>
<tr>
<td>MACD</td>
<td>Moving average convergence divergence: displays trends following characteristics and momentum characteristics</td>
</tr>
<tr>
<td>CCI</td>
<td>Commodity channel index: helps to find the start and end of trend</td>
</tr>
<tr>
<td>ATR</td>
<td>Average true range: measures the volatility of price</td>
</tr>
<tr>
<td>BOLL</td>
<td>Bollinger band: provides a relative definition of high and low, which aids in rigorous pattern recognition</td>
</tr>
<tr>
<td>EMA20</td>
<td>20 days exponential moving average</td>
</tr>
<tr>
<td>MA5/MA10</td>
<td>5/10 days moving average</td>
</tr>
<tr>
<td>V_MA5/V_MA10</td>
<td>5/10 days trading volume average</td>
</tr>
<tr>
<td>MTM6/MTM12</td>
<td>6/12 months momentum: helps pinpoint the end of a decline or advance</td>
</tr>
<tr>
<td>ROC</td>
<td>Price rate of change: shows the speed at which stock’s price rate is changing</td>
</tr>
<tr>
<td>SMI</td>
<td>Stochastic momentum index: shows where the close price is relative to the midpoint of the same range</td>
</tr>
<tr>
<td>WVAD</td>
<td>William’s variable accumulation/distribution: measures the buying and selling</td>
</tr>
</tbody>
</table>
Table 2 - Testing result

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE</th>
<th>RMSE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSTM</td>
<td>0.0222</td>
<td>0.03038</td>
<td>92324E-04</td>
</tr>
</tbody>
</table>

V. CONCLUSIONS

In this paper authors used two DL architectures for the stock price prediction. Here authors have trained mainly LSTM networks with the stock price of Tata Beverages from Tata Global Beverages Limited., it is clear that the models are capable of identifying the patterns existing in both the stock markets. In this paper authors have tried to find out why the LSTM model is best option for the stock price predictions. It is clearly visible from the above graph that LSTM model, is capable of capturing the abrupt changes in the system since a particular window is used for predicting the next instant.

REFERENCES


