STOCK MARKET PREDICTION AND ANALYSIS USING LSTM NEURAL NETWORK

Kothapally srujani[1], Bhavagna Uppala[2], and K Jaya Laxmi[3].

[1] , Student, Department of Information technology, BV Raju institute of technology, Telangana, India

[2] , Student, Department of Information technology, BV Raju institute of technology, Telangana, India

[3] , Assistant Professor, Department of Information technology, BV Raju institute of technology, Telangana, India

ABSTRACT

The prediction of stock value is a complex task which needs a robust algorithm background in order to compute the long-term share prices. Stock prices are correlated within the nature of market; hence it will be difficult to predict the costs. Prior studies concentrated on the factors that can affect investors emotions. The researchers completed studies based on social media, the period of the stock market, and the use of various models to extract the feature of stocks. To accurately anticipate the stock, they initially used NLP and GBDT, which primarily focus on emotion and select information from the news (which didn't give the accurate predictions).

The proposed algorithm using the market data to predict the share price using machine learning techniques like recurrent neural network named as Long Short-Term Memory (LSTM), in that process weights are corrected for each data points using stochastic gradient descent. This system will provide accurate outcomes in comparison to currently available stock price predictor algorithms. The network is trained and evaluated with various sizes of input data to urge the graphical outcomes.

Keyword: *LSTM Neural Network, Stock, Prediction, Gradient Descent, Market, Shares,Accurate.*

1.INTRODUCTION:

1.1 Background Information:

In order to forecast the future value of firm stock or other financial instruments traded on an exchange, one must do stock market prediction and analysis. The country's stock market, which is a significant component of the economy, is crucial to the development of the nation's industry and commerce, which ultimately has an impact on the economy. A company's principal source of capital for corporate expansion is the stock market.

Current business study interests have been focused on areas of future stock price movement predictions, which makes it difficult and demanding. Scholars, corporate groups, and interested consumers are eager to find the stock price forecast of movements in because they believe that future occurrences depend on current and historical data. Financial data, on the other hand, is seen as being complex data to anticipate or predict.

In addition to buying and selling stocks and shares on stock exchanges, each stock is distinguished by a few other factors, such as the closing price, which is the most crucial factor for determining the next day's price for a particular stock. All the factors that affect stock movements over time have a link and exhibit distinct behavior. In the stock market, several variables, including news events, economic indices, and business financial reports, can affect stock prices. While analyzing historical data, LSTM neural networks can be utilized to spot patterns and trends that can be used to forecast future stock values.

We first need to gather historical data for the companies we wish to investigate in order to deploy LSTM neural networks for stock market prediction and analysis. The information should contain historical stock price, volume, and other pertinent metrics. The values are then normalized, and the data is divided into training and testing sets as part of the pre-processing step.

The LSTM neural network can then be trained on the training set using a series of historical data to predict future values. The network gains the ability to spot data linkages and patterns that can be utilized to forecast outcomes.

After the model has been trained, we may assess how well it performs on the test set. To gauge the precision of the predictions, we can use metrics like mean squared error (MSE) and root mean squared error (RMSE).

Many real-world uses exist for stock market forecasting and analysis utilizing LSTM neural networks, including portfolio management, risk analysis, and trading tactics. But it's crucial to remember that stock market prediction is fundamentally speculative and dependent on a variety of outside variables. Instead of relying solely on these models, it's critical to use them as tools to aid in decision-making.

1.2 Existing System:

- The KNN algorithm is a machine learning algorithm that has a high computing cost and is regarded as a slow learning algorithm.
- While utilizing the KNN technique, the data is split into two groups: the training data, which the system uses to create predictions, and the test data. The weighted Euclidean distance, which is essentially the separation between the two components, is the most widely used approach.
- The accuracy of the KNN algorithm is approximately 95%.

1.3 Disadvantages of Existing System:

- Accuracy is less in existing system.
- Time consumed for execution is high compared to LSTM model.
- With large data prediction might be slow.

1.4 Proposed System:

- Long short-term memory (LSTM) is one of the many types of recurrent neural network RNN, it's also capable of catching data from past stages and using it for future predictions.
- Its architecture comprises the cell, input gate, output gate, and forget gate.
- Firstly, the dataset is downloaded from the yahoo finance Dataset. The data collected will be correlated. Then the data is split into train and test the number of train samples (Data Preprocessing). Lastly, run the LSTM algorithm.

1.5 Advantages of Proposed System:

- There are many other regression algorithms like the current proposed system, but the current proposed system is faster and quicker in execution aswell as in giving the user outputs.
- Not all existing system's supports huge volume of data, as we know stock analysis needs huge data, hence the proposed system easily tackles this situationby handling large volume of data without any hesitation.
- The other main goal of this system is, it is mainly designed to analyze andpredict the data accurately which cannot be found in other systems.
- Not only with large datasets, but it also works well with small and medium data sets as well. The proposed system is Highly Flexible when compared to the existing system.

1.6 Requirement Specifications:

1.6.1 Software Requirements:
- Language and Technology used : Python
- Operating Systems Supported : Windows
- Packages : Scikit-Learn ,numpy ,pandas.
- Tools : Python IDLE

1.6.2 Hardware Requirements:
- Processor : Pentium IV or higher
- RAM : 2GB

2. IMPLEMENTATION:

Stock market forecasting is a difficult task, and no forecast can be made with absolute certainty. Yet employing LSTM neural networks can make our predictions based on historical data more precise. Here is an example of how to use LSTM neural networks for stock market prediction and analysis:

1. Data gathering and preparation: Get historical stock market information for the particular stock you want to forecast. The data is preprocessed by removing any blanks or gaps, standardizing the data, and dividing it into training and test sets.
2. constructing the LSTM model: Provide the LSTM model's structural details, such as the number of LSTM layers, neurons, and activation methods. Choose the proper optimizer and loss function before compiling the model. Use the fit method to train the model on the training set.

3. Model Evaluation: Use the evaluate method to gauge the model's performance on the testing set. Plot or chart the model's performance to see it in action.
4. Analysis and forecast: To forecast stock market prices for the upcoming day or week, use the model.Review the forecasts and contrast them with the current stock market pricing.
   If necessary, change the model's parameters and retrain it to increase forecast accuracy.

2.1 Flow Chart:

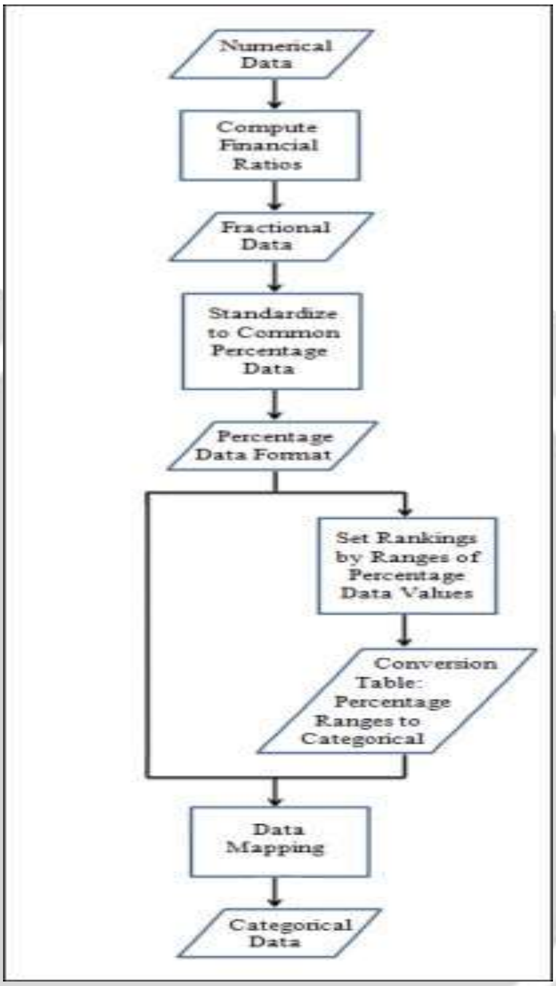


Fig-1: Flowchart for data transformation

2.2 Algorithm used:

Recurrent neural networks include long short-term memory. The output from the previous phase is sent into the current step of an RNN as input. Hochreiter & Schmidhuber created LSTM. It addressed the issue of long-term RNN dependency, wherein the RNN can predict words from current data but cannot predict words kept in long-term memory. As the gap length grows, RNN's performance becomes less effective. By default, LSTM may store information for a long time. It uses time series data for processing, forecasting, and classification.

Steps for algorithm:-

Input: Historic stock data

Output: Prediction of stock price using price variation

STEP 1**:** Start

STEP 2: Data Preprocessing after getting the historic data from the market for aparticular share.

STEP 3: Import the dataset to the data structure and read the open price.

STEP 4: Do a feature scaling on the data so that the data values will vary from 0 to 1

STEP 5: Creating a data structure with 60 timestamps and 1 output.

STEP 6: Building the RNN for step 5 data set and initialize the RNN by usingsequential processor.

STEP 7: Adding the first LSTM layer and some Dropout regularization forremoving unwanted values.

STEP 8: Adding the output layer.

STEP 9: Compiling the RNN by adding Adam optimization and the loss as mean squared error.

STEP 10: Making the prediction and visualizing the results using plottingtechniques.

2.3 Description of Modules:

Modules present in this Stock market prediction and analysis using LSTMproject are:

User Module
Prediction Module

- **User module:** Takes the input from the users thefeatures/attributes considered for predicting the stock.
- **Prediction module:** Responsible for building the classification model that is used for prediction of stock.

2.4 Data Sets:

The terminology used to name the aspects in the data sets, such as debt/equity, asset turnover, cash flow, and return on equity, closely reflect those used to name financial ratios. In data set 1, the initial source of the information includes numerical values that were later transformed into categorical values.

The independent variables are debt equity, asset turnover, cash flow, and return on equity. The class variable Price Trend's value is predicted using the values of the independent variables. As different stocks have different prices, the variable named Price in data set 1 comprises numerical values that do not provide semantic meanings of the stocks. In contrast, the data set 2 variable titled Price Trend comprises standardized categorical values that denote a positive or negative price trend.

Feature	Data Type
Debt_Equity	Numeric
Asset_Turnover	Numeric
Cash_Flow	Numeric
Return_On_Equity	Numeric
Price	Numeric

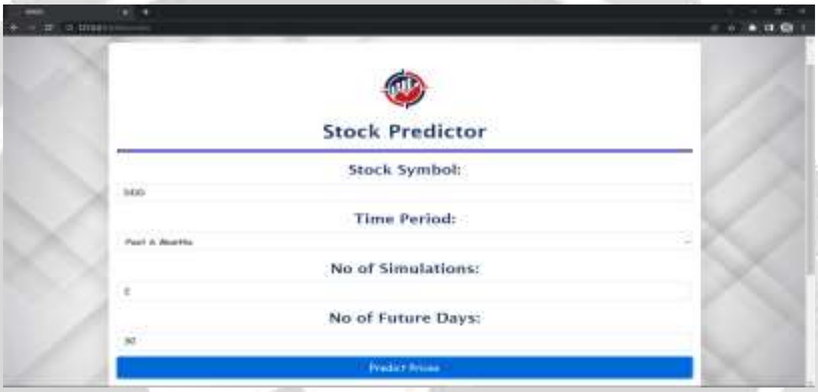
Table-1: Dataset 1

Feature	Data Type
Debt_Equity	Categorical
Asset_Turnover	Categorical
Cash_Flow	Categorical
Return_On_Equity	Categorical
Price_Trend	Categorical

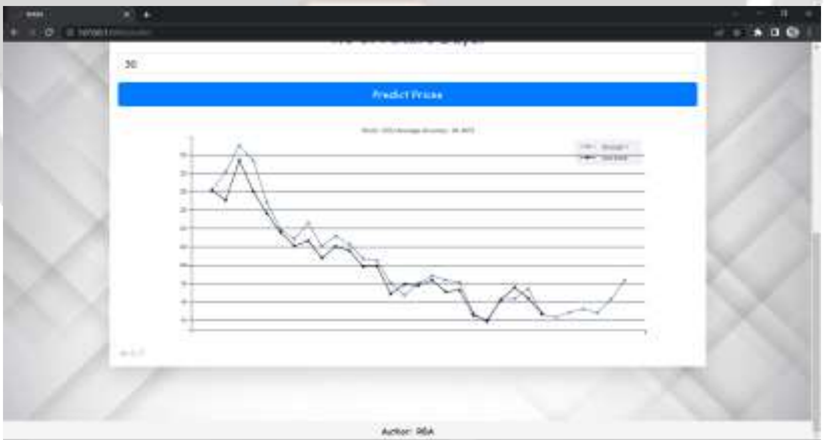
Table-2: Dataset 2

3.SAMPLE OUTPUT/INPUT:

3.1 Input:



3.2 Output:



3.3 Results:

The proposed model was tested and compared with four other standard algorithms, including KNN, Naïve Bayes, OneR and ZeroR. The test examined how accurate the tested algorithms predict the stock price trends, and evaluated the MAE and RMSE. Table 5 presents the test results. The hybrid KNN-Probabilistic model has allowed us to achieve an estimated accuracy of 89.1725%, exceeding the stand alone KNN reported accuracy of 86.6667% and the Naive Bayes accuracy of 76.1194%. The accuracy rates for OneR and ZeroR classifiers were 71.6418% and 64.1791% respectively. KNN-Probabilistic model has MAE rate of 0.0667% and RMSE rate of 0.2582% which are much lower than the other classifiers.

Classifier	Accuracy (%)	MAE	RMSE
KNN-Probabilistic	93.3333	0.0667	0.2582
KNN	86.6667	0.1333	0.3651
Naive Bayes	76.1194	0.1726	0.2824
OneR	71.6418	0.5325	0.6139
ZeroR	64.1791	0.4619	0.4805

Table-3:prediction results of classifier

Overall, KNN-Probabilistic model has better accuracy rate and error rates than the other classifiers used for comparisons. The test demonstrated that the hybrid mechanism of KNN and probabilistic method produced significantly improved results, compared with each of the KNN and Naïve Bayes classifiers.

4. CONCLUSION:

In essence, instead of utilizing a machine learning model, we used a neural network (deep learning) model to construct a stock market price prediction application. The project's greater accuracy in comparison to other existing models is its key benefit. In terms of functionality, this program has been enhanced with more special features that are simple to use and evaluate. These features, which are extremely flexible in nature, can be a major benefit to users who track stocks and occasionally attempt to buy in them. They also perform well with small datasets. The project's user-friendly interface can be accessible in a single step. The scope of this initiative is very broad because a growing number of individuals are investing in stocks, which will eventually lead to economic growth and a host of other significant developments.

5. REFERENCES:

[1]. Benjamin Graham, Jason Zweig, and Warren E. Buffett, The Intelligent Investor, Publisher: Harper Collins Publishers Inc, 2003.
[2]. Charles D. Kirkpatrick II and Julie R. Dahlquist, Technical Analysis: The Complete Resource for Financial Market Technicians (3rd Edition), Pearson Education, Inc., 2015.
[3]. Bruce Vanstone and Clarence Tan, A Survey of the Application of Soft Computing to Page | 60 Investment and Financial Trading, Proceedings of the Australian and New Zealand Intelligent Information Systems Conference, Vol. 1, Issue 1, http://epublications.bond.edu.au/infotech_pubs/ 13/, 2003, pp. 211–216.
[4]. Monica Tirea and Viorel Negru, Intelligent Stock Market Analysis System - A Fundamental and Macro-economical Analysis Approach, IEEE, 2014.
[5]. Kian-Ping Lim, Chee-Wooi Hooy, Kwok-Boon Chang, and Robert Brooks, Foreign investors and stock price efficiency: Thresholds, underlying channels and investor heterogeneity, The North American Journal of Economics and Finance. Vol. 36, http://linkinghub.elsevier.com/retrieve/pii/S106 2940815001230, 2016, pp. 1–28.
[6]. Lamartine Almeida Teixeira and Adriano Lorena Inácio de Oliveira, A method for automatic stock trading combining technical analysis and nearest neighbor classification, Expert Systems with Applications, http://linkinghub.elsevier.com/retrieve/pii/S095 7417410002149, 2010, pp. 6885–6890.