Tradexa: Your Million Dollar Shortcut

Jash Mehta
Department of Information Technology
K. J. Somaiya College of Engineering
Mumbai, India
jash.jm@somaiya.edu

Jubin Kamdar
Department of Information Technology
K. J. Somaiya College of Engineering
Mumbai, India
jubin.kamdar@somaiya.edu

Mitanshu Gada
Department of Information Technology
K. J. Somaiya College of Engineering
Mumbai, India
mitanshu.g@somaiya.edu

Prof. Ravindra Divekar
Department of Information Technology
K. J. Somaiya College of Engineering
Mumbai, India
ravindradivekar@somaiya.edu

***Abstract*—The stock market is a place where buy and sell shares and any concerning individual would like to know if the stock which he/she has purchased will result in a profit to the user or not. The user base for this kind of project is already present but advanced systems that can accurately predict the future stock prices are still under development. Investing in shares is complicated and depends on many factors which a trained and experienced trader knows very well. To help the traders with investing into good shares with high returns we have incorporated the use of Machine learning techniques to help us build a model sufficient enough to cater to the needs of every individual. This study focuses on building a model that uses Recurrent Neural Networks (RNN) and specifically Long-Short Term Memory (LSTM) layers in the network for learning the stock trends and predicting the stock prices. The main objective behind this research is to predict & investigate the future closing prices of various shares and penny stocks that are often overlooked by high-end projects because they trade in one digit or at the price of a penny or those with the Most Low Market.**

Penny shares are generally regarded as stocks trading at a single-digit or penny price or those with a very Low Market value. Because of the very low price, such types of stocks can easily raise many times over when a strong buy interest rate hits the stock.

***Index Terms*—stock market, stock-prediction, statistics, prediction, deep-learning, recurrent neural network, long short term memory.**

I. Introduction

A handful of studies have been on the topic of using machine learning to learn and predict big money, predict the costs of managing and building the entire portfolio of assets, as well as, the process of investment, and numerous other activities that can be covered by machine learning algorithms. The term "machine learning" is generally used to refer to any method of computer algorithms that generate patterns in a database and does not use editorial guidelines. Predicting can be defined as the prognosis of future events or events by examining historical data.

Fundamental analysis is a type of investment analysis in which a company's share price is measured by testing its sales, profits, profits, and other economic factors. This method is most suitable for long-term forecasts.To predict the future price, Technical analysis uses the historical price of stocks . The algorithm commonly used in Technical Analysis is known as Central Navigation. It can be considered a weightless method of previous 'n' Data. This method is best fitted for temporary predictions. The third way is to analyze the time-series data.

There are many complex economic measures and the volatility of the stock market is very harsh. However, as technology evolves, the opportunity to gain solid wealth in the stock market grows and helps experts find the most instructive indicators to make better predictions. Market value forecast is very important to help increase the profitability of buying stock options and caring about the risk of loss. General neural networks (RNN) are one of the most powerful models for sequential data processing. Long-Term Memory is one of RNN's most prime structures. LSTM instigates a memory cell, a unit of calculation that replaces traditional neurons with a hidden network layer. With these memory cells, networks can effectively associate memories with remote input over time, which is why they are more suited to capturing data formation over time with higher prediction power.

II. Literature Review

Traditional methods include the use of accurate forecasting models that predict historical stock prices and Bowden et al [2] in his paper, suggested the use of an auto-regressive model called the ARIMA method, this method is used to create a default model to predict approximate values of time-series data. This method is mathematically effective, but the assumption that the data represents the distribution of statistics limits the ability to model data for an indirect time series.

There are many factors in play when we are talking about stock price prediction. With the increasing developments in the statistical methods in financial fields, many people have stored huge amounts of data for the sole purpose of performing analysis on it and studying it to find patterns. This huge dataset gives a platform for many machine learning models to find trends in them. Nair and his colleagues [1] in their paper used the method of decision trees and rough sets in order to integrate the pros of both the concepts, but this method was prone to overfitting as the data was huge and there was a lot of noise in the data which the model couldn't comprehend. Nottola et al. [9] have used neural networks for prediction and his work has proven to be more beneficial and accurate than the decision trees. However, there are some situations where the neural networks face problems such as the local optima problem and the vanishing gradients problem.

Cao et al. [3] in their paper displayed that Support vector machines can be used for stock prediction and that with the help of SVM the model is generalizing effectively. Deng [6] proved in his paper that the performance of random forests can be increased with parameter optimization and that the resultant model can perform better than SVM for prediction.

With the advent of different artificial intelligence technologies, deep learning has seen many advancements and everyone's attention is now towards deep learning because it has proven exceptional in language translation [5], sentiment recognition from audio [8], recognizing images using CNN [13], and other areas. Deep neural networks (DNN) when juxtaposed with the traditional model, can analyze the complex nonlinear relationships because of the layered architecture, and this also helps with a deeper understanding of the data where we have nonlinear patterns which are hard to predict [7].

Tsantekidis [12] used the convolutional neural network (CNN) as the model for prediction and used this to compare it with other models, but the main problem that got engendered was that CNN found it difficult to work efficiently because of the sequential dataset. In all previous approaches, there was no further study, but Recurrent Neural Networks and their way of applying output from previous iterations to the next iterations is where it outperforms in matters of sequential data, and Long Short Term Memory (LSTM) moves one step further with a gated approach to improve performance and terminate the drawbacks of conventional RNN network. Selvin [11] in his paper experimented on three stock prediction models comprising Convolutional Neural Networks (CNN), Recurrent neural networks (RNN), and LSTM deep learning networks. His analysis was done on the dataset that he had and used these three models to compare their performance on his sequential stock market data. Shile Chen and Changjun Zhou [4] in their paper have also proposed an LSTM-based approach but their architecture didn't perform well, also their model's performance was based on the Chinese stock market. The dataset they worked with was from the Chinese stock market and that market was influenced by other factors different from the Indian stock market. From their results, we concluded to create an LSTM-based neural network with the principal component analysis done on the dataset.

III. Related Work

A. Why LSTM?

Deep Neural Networks can be divided into three layers: a. Input layer b. output layer and c.hidden layers. There are hyperparameters like the number of hidden layers, the learning parameter (α), and other parameters, which can be tweaked to make the model more efficient. There is no memory concept in an artificial neural network and because of that we have Recurrent Neural Networks (RNNs). As we can see in the figure, with every iteration an output of the previous iteration is carried forward as memory and this is why RNNs are used in time-series datasets.

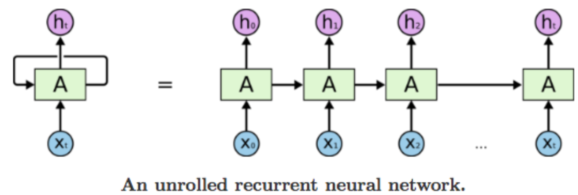


Fig. 1: Recurrent Neural Network

In the first iteration, the input given is X0 and the models output is h0; then for the next iteration the input x1 along with h0 is given to the model and the output is h1, which is then coupled with input x2 for the next iteration. [10]

$$h_t = f(h_{t-1}, x_t)$$

when applying the activation function the formula looks like this:

$$h_t = tanh(W_{hh}h_{t-1} + W_{xh}h_t)$$

where $tanh()$ is the tan hyperbolic activation function, W_{hh} is the weight for the previous hidden layer at timestamp $t-1$, W_{xh} is the weight for the current input state t.

The problem with RNNs arises when we have a large dataset and the RNN is unable to learn the long-term dependency. Another major issue that can arise is the problem of vanishing gradients when there are several hidden layers and during learning, the gradients become very small.

Long Short-Term Memory overcomes all the problems of traditional RNN architecture with the help of its gated architecture. Most importantly LSTM networks overcome the long-term dependency problem in RNN. Fig. 2 shows a typical LSTM node and we can see that each node can be separated by its functionality. 3 gates govern the flow of information through each unit and they are: a. Forget gate b. Input gate and c. output gate.

Input Gate: Determines which input should be used to adjust the condition of the cell. The sigmoid function allows certain values to pass and the tanh function gives weight to the input that determines its value.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$c_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

Forget the Gate: as the name suggests, this function is based on current entries and previous releases or do not omit current entries or delete them.

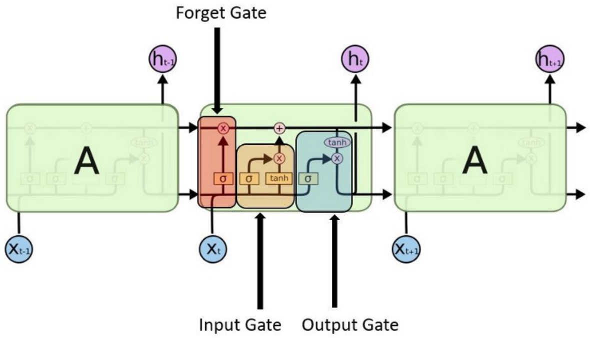


Fig. 2: Typical LSTM Cell

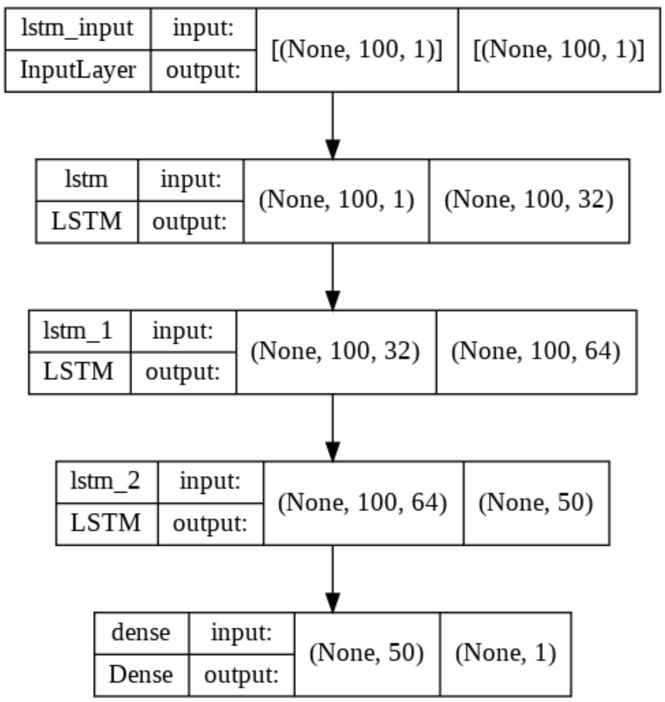


Fig. 3: Model 1

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Output gate: current state input and the previous cell input controls the output that is sent to the next recurrence.

$$O_t = \sigma(W_o[h_{t-1}, X_t] + b_o)$$

$$h_t = O_t * tanh(c_t)$$

IV. Implementation

In the current paper we are implementing two different LSTM neural network models. The stocks under consideration are Bajaj Finance, HDFC Bank, TCS and some penny stocks like JP Power and Suzlon.

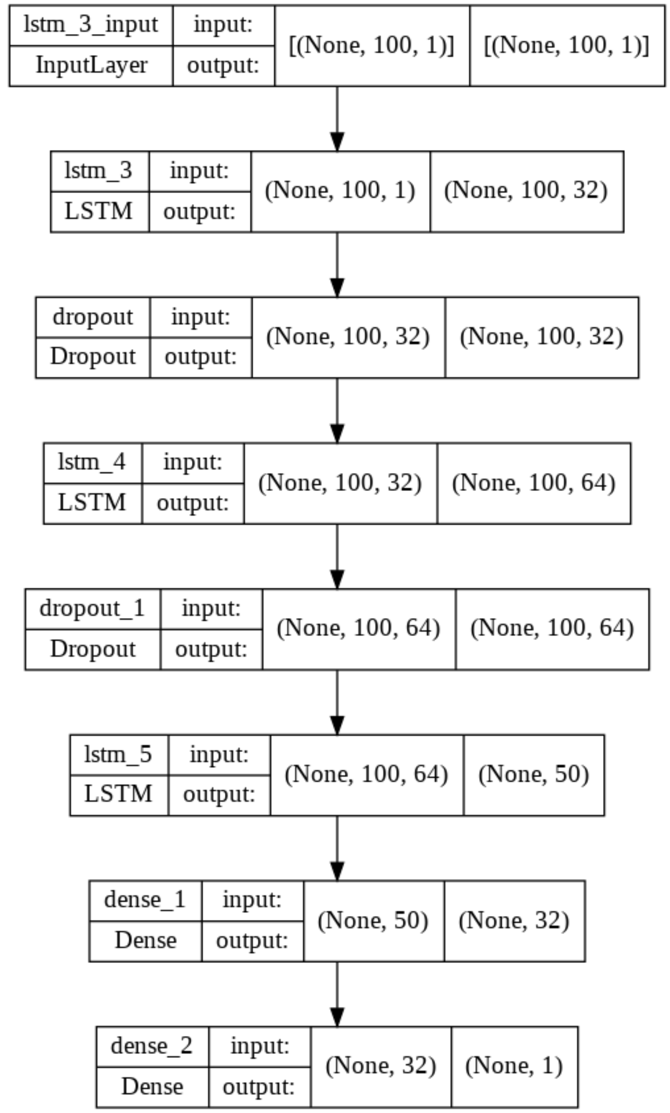
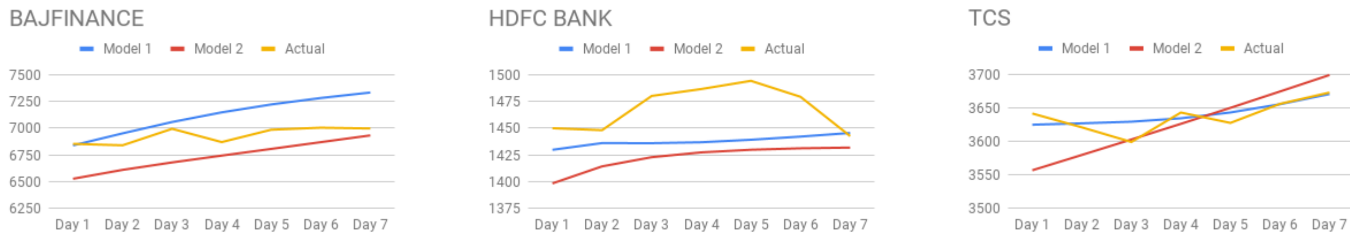


Fig. 4: Model 2

Fig. 3 represents the first LSTM network architecture that we have used for prediction. The first model consists of 3 LSTM layers, 1 input layer, and 1 dense output layer. The input given to the input layer is of size (100, 1) where 100 previous values are taken, and then again the next 100 values are taken so that the model can identify the pattern and what the next prediction will be. Next, 3 LSTM layers are performing the learning operation from the data and each layer has its unit size from 32, 64, and 50. Finally, as the model needs to predict the next value based on the input, we have a dense layer that combines the output from the previous layer and gives only 1 integer output. These models were trained on the past 5 years' worth of data where the train to test size ratio was 90:10 (i.e. 90% of the 1200 rows (approx) were considered for training and the rest 10% were considered as test data). The model was trained for 100 epochs on each stock and then the predictions were recorded for the next 7 days.



(a) Bajaj Finance (b) HDFC Bank (c) TCS

Fig. 5: Stock Predictions

(a) JP Power (b) Suzlon

Fig. 6: Penny-Stock Predictions

The second model which can be seen in Fig 4, displays the Network architecture which was devised from the results of model 1. Some layers are the same, but what has been added is the concept of dropout and can be justified by the results of model 1. Each dropout layer is added after the LSTM layers, and the probability is set to 0.25 i.e. 25%. Dropout is a regularization technique that is used when we have a model overfitting the training data. Overfitting means that the model trains very well on the train data, but is unable to perform on the test or validation data and as a result, we add a dropout layer which helps with this problem. Also, we have added an extra dense layer before the output layer to get the predictions as accurate as possible. In the next section, we'll carefully analyze the results of each model and why the second model was created, and how the results have then changed.

V. Results

In this research paper, we are trying to predict the future stock prices of stocks from the Indian stock market. The stocks under experimentation belong to companies Tata Consultancy Services (TCS), HDFC Bank, and Bajaj Finance. The future price prediction of these stocks is predicted by two different models, and results of model 2 are more accurate than model 1. From Fig. 5 we can see that the predictions made by model 1 are a little harsh and steep as compared to the predictions of model 2. The results also portray that the addition of the dropout layer to the model has produced promising results and that it has had a positive effect on the prediction overall. The actual stock price of Bajaj Finance for 7 days is increasing overall from the current price and the prediction drawn by model 2 is very accurate to it whereas the predictions of model 1 are a little off track. Also, it can be seen that the overfitting problem has arisen in model 1 and that a penalty method or a regularization technique is the next step forward. The stock prices of HDFC Bank predicted by model 1 and model 2 both are accurate overall while the actual price went through a rollercoaster, the graph carefully depicts the predicted prices of both the models and the actual price so anyone can understand the performance of the models. The predictions made by model 2 for TCS show us that error is low as well as the future price would increase which happened. From these predictions, we can see how well both models are performing and the importance of regularization in neural networks.

Predictions of penny stocks done correctly will have a massive impact on any trader or person's profits. Investing in penny stocks comes with high risks, but has high rewards in return and things can go awry at any second. So, we have tried to figure out if these models accurately predict the stock prices for these penny stocks. The results can be seen in Fig. 6. The penny stocks under experimentation are JP Power and Suzlon. These stocks have more than 5 years' worth of data for the model to train. From the JP Powers graph, we can see that model 2 is predicting more precisely than model 1, but the predictions for Suzlon say otherwise. Model 2 is not accurate enough in predicting stock prices for Suzlon as the price went down and the prediction was that the price is going up.

As we have discussed in this paper that the stock prices are dependent on various factors and from this paper, we can see that historic data can only help in predictions to a certain level, but to get truly accurate predictions some extra aspects can be considered and added to the prediction model.

VI. Conclusion

This paper proposes an LSTM Neural Network built to forecast future values for various stocks like Bajaj Finance, HDFC Bank, and TCS. Two models of different architectures are constructed for this experiment and each model's results for the next 7 days were recorded. From the results section, and the discussions done earlier, it is safe to say that LSTM has shown its worth in time-series forecasting, and with added other features like news analysis of a particular stock, and pre-defined weights of each stock can help a lot in predictions. The future work in this could be the addition of extra parameters like news analysis and how that affects the stock price, new architecture for the neural network with the inclusion of different layers like GRUs (Gated Recurrent Units) or LSTM or some other model which outperforms the current top of the table layers.

References

[1] Binoy.B.Nair, V.P Mohandas, and N. R. Sakthivel. "Article: A Decision tree- Rough set Hybrid System for Stock Market Trend Prediction". In: *International Journal of Computer Applications* 6.9 (Sept. 2010). Published By Foundation of Computer Science, pp. 1–6.

[2] Nicholas Bowden and James E. Payne. "Short term forecasting of electricity prices for MISO hubs: Evidence from ARIMA-EGARCH models". In: *Energy Economics* 30.6 (2008). Technological Change and the Environment, pp. 3186–3197. ISSN: 0140-9883. DOI: https://doi.org/10.1016/j.eneco.2008.06.003. URL: https://www.sciencedirect.com/science/article/pii/S0140988308000868.

[3] L.J. Cao and F.E.H. Tay. "Support vector machine with adaptive parameters in financial time series forecasting". In: *IEEE Transactions on Neural Networks* 14.6 (2003), pp. 1506–1518. DOI: 10.1109/TNN.2003.820556.

[4] Shile Chen and Changjun Zhou. "Stock Prediction Based on Genetic Algorithm Feature Selection and Long Short-Term Memory Neural Network". In: *IEEE Access* 9 (2021), pp. 9066–9072. DOI: 10.1109/ACCESS.2020.3047109.

[5] Marta Costa-jussa et al. "Introduction to the Special Issue on Deep Learning Approaches for Machine Translation". In: *Computer Speech Language* 46 (May 2017). DOI: 10.1016/j.csl.2017.03.001.

[6] J. Deng and L. Li. "School of mathematics, physics and statistics, Shanghai university of engineering Science; Application of parametric optimization random forests in stock forecasting". In: *Computer Engineering Software* 1 (2020), pp. 178–182.

[7] Lv Dongdong et al. "DNN models based on dimensionality reduction for stock trading". In: (2020). DOI: 10.3233/IDA-184403.

[8] Haytham M. Fayek, Margaret Lech, and Lawrence Cavedon. "Evaluating deep learning architectures for Speech Emotion Recognition". In: *Neural Networks* 92 (2017). Advances in Cognitive Engineering Using Neural Networks, pp. 60–68. ISSN: 0893-6080. DOI: https://doi.org/10.1016/j.neunet.2017.02.013. URL: https://www.sciencedirect.com/science/article/pii/S089360801730059X.

[9] Feng Li and Cheng Liu. "Application Study of BP Neural Network on Stock Market Prediction". In: *Proceedings of the 2009 Ninth International Conference on Hybrid Intelligent Systems - Volume 03*. HIS '09. USA: IEEE Computer Society, 2009, pp. 174–178. ISBN: 9780769537450. DOI: 10.1109/HIS.2009.248. URL: https://doi.org/10.1109/HIS.2009.248.

[10] Aditi Mittal. *Understanding RNN and LSTM*. https://aditi-mittal.medium.com/understanding-rnn-and-lstm-f7cdf6dfc14e. "[online]". 2019.

[11] Sreelekshmy Selvin et al. "Stock price prediction using LSTM, RNN and CNN-sliding window model". In: Sept. 2017, pp. 1643–1647. DOI: 10.1109/ICACCI.2017.8126078.

[12] Avraam Tsantekidis et al. "Forecasting Stock Prices from the Limit Order Book Using Convolutional Neural Networks". In: *2017 IEEE 19th Conference on Business Informatics (CBI)*. Vol. 01. 2017, pp. 7–12. DOI: 10.1109/CBI.2017.23.

[13] Junliang Xing et al. "Diagnosing deep learning models for high accuracy age estimation from a single image". In: *Pattern Recognition* 66 (2017), pp. 106–116. ISSN: 0031-3203. DOI: https://doi.org/10.1016/j.patcog.2017.01.005. URL: https://www.sciencedirect.com/science/article/pii/S0031320317300079.