PUDUCHERRY TECHNOLOGICAL UNIVERSITY

**STOCK** **PRICE** **PREDICTION**

GUIDED BY DR. **Dr.** **Salini** **P** DEPARTMENT-CSE

PUDUCHERRY TECHNOLOGICAL UNIVERSITY

*Submitted* *by* SAURABH KISHOR (19CS1100)

SHUBHAM CHOUDHURY(19CS1103)

UTKARSH(19CS1113)

**INTRODUCTION**

• In this project we attempt to implement machine learning approach to predict stock prices. Machine learning is effectively implemented in forecasting stock prices. The objective is to predict the stock prices in order to make more informed and accurate investment decisions. We propose a stock price prediction system that integrates mathematical functions, machine learning, and other external factors for the purpose of achieving better stock prediction accuracy and issuing profitable trades. In this project, LSTM model is used to predict the stock price.

**LIST OF ABBREVATIONS**



**LSTM** : Long Short-Term Memory **ATS** : Automated Trading System **GRU** : Gated Recurrent Unit **ML** : Machine Learning

**SVM** : Support Vector Machine **EMH** : Efficient Market hypothesis **AI** : Artificial Intelligence **NN** : Neural Networks

**ARMA** : Autoregressive Moving Average **DRL** : Deep Reinforcement Learning **LMS** : Least Mean Square

**UML** : Unified modelling Language **MSE** : Mean Squared Error

**RMSE** : Root Mean Squared Error

• **Long** **short-term** **memory** **network:**

Long short-term memory network (LSTM) is a particular form of recurrent neural network (RNN).

• **Working** **of** **LSTM:**

• LSTM is a special network structure with three “gate” structures. Three gates are placed in an LSTM unit, called input gate, forgetting gate and output gate. While information enters the LSTM’s network, it can be selected by rules. Only the information conforms to the algorithm will be left, and the information that does not conform will be forgotten through the forgetting gate.

**Problem Definition**

The stock market allows investors to own shares of public companies through trading either by exchange or over the counter markets.

This market has given investors the chance of gaining money and having a prosperous life through investing small initial amounts of money, low risk compared to the risk of opening new business or the need of high salary career.

Stock markets are affected by many factors causing the uncertainty and high volatility in the market.

Stock market is a typical area that presents time-series data.

Stock prices are volatile in nature and price depends on various factors. The main aim of this project is to predict stock prices using Long short term memory (LSTM).

**OBJECTIVE**

Our objective is to predict the future price and calculate the future growth of the company in the different time span.

Then we analyze the prediction error for each company of different sector. Based on that we conclude which time span is best for future prediction of

that particular sector.

We first predict the future closing price of 5 different companies from some predecided sectors with the help of LSTM(Long Short-Term

Memory). This prediction will be done on historical data & the future prediction will be done f or 3-month, 6-month, 1 year & 3 years. In these four different time spans (3 & 6 months, 1 & 3 years), we c alculate the growth of those companies. Then by analyzing the deviations of closing price for each time span, we took the resultant time span which has maximum growth, i.e. less error for the particular sector

**LITERATURE** **SURVEY**

• We apply LSTM recurrent neural networks (RNN) in predicting the stock price correlation coefficient of two individual stocks. RNN’s are competent in understanding temporal dependencies.

• The use of LSTM cells further enhances its long-term predictive properties.

• To encompass both linearity and nonlinearity in the model, we adopt the ARIMA model as well.

• The ARIMA model filters linear tendencies in the data and passes on the residual value to the LSTM model.

**Random** **forests** **and** **LSTM** **networks** (more precisely CuDNNLSTM) as training methodologies to analyse their effectiveness in forecasting out- of-sample directional movements of constituent stocks of the S&P 500 from January 1993 till December 2018 for intraday trading

We will introduce a multi-feature setting consisting not only of the returns with respect to the closing prices, but also with respect to the opening prices and intraday returns.

As trading strategy, we use and, on each trading day, buy the 10 stocks with the highest probability and sell short the 10 stocks with the lowest probability to outperform the market in terms of intraday returns – all with equal monetary weight

**EXISTING** **METHODS**

Stock market prediction is an act of trying to determine the future value of a stock other financial instrument traded on a financial exchange.

. The technical and fundamental or the time series analysis is used by the most of the stockbrokers while making the stock predictions. The programming language is used to predict the stock market using machine learning is Python.

**Using Technique**

Automated Stock Price Prediction Using Machine Learning

Traditionally and in order to predict market movement, investors used to analyse the stock prices and stock indicators in addition to the news related to these stocks. Hence, the importance of news on the stock price movement.

In this work, **we** **propose** **an** **automated** **trading** **system** **that** **integrates** **mathematical** **functions,** **machine** **learning,** **and** **other** **external** **factors** **such** **as** **news’** **sentiments** **for** **the** **purpose** **of** **achieving** **better** **stock** **prediction** **accuracy** **and** **issuing** **profitable** **trades**.

Particularly, **we** **aim** **to** **determine** **the** **price** **or** **the** **trend** **of** **a** **certain** **stock** **for** **the** **coming** **end-of-day** **considering** **the** **first** **several** **trading** **hours** **of** **the** **day.** **To** **achieve** **this** **goal,** **we** **trained** **traditional** **machine** **learning** **algorithms** **and** **created/trained** **multiple** **deep** **learning** **models** **taking** **into** **consideration** **the** **importance** **of** **the** **relevant** **news**

• **OVERALL** **LIMITATION**

• Stock market is all about prediction and rapid decision making about investment, which cannot be done without thorough analysis of the market.

• If we can predict the stock market by analysing historical data properly, we can avoid the consequences of serious market collapse and to be able to take necessary steps to make market immune to such situations.

automated trading systems (ATS) that are operated by the implementation of computer programs can perform better and with higher momentum in submitting orders than any human.

**PROPOSED** **WORK**

The prediction methods can be roughly divided into two categories, statistical methods and artificial intelligence methods.

**Statistical** **methods** **:** It include logistic regression model, ARCH model, etc.

**Artificial** **intelligence** **methods:** It include multi-layer perceptron, convolutional neural network, naive Bayes network, back propagation network, single-layer LSTM, support vector machine, recurrent neural network, etc. They used Long short-term memory network (LSTM).

It is needed to find an optimization algorithm that requires less resources and has faster convergence speed.

• **Used** **Long** **Short-term** **Memory** **(LSTM)** **with** **embedded** **layer** **and** **the** **LSTM** **neural** **network** **with** **automatic** **encoder.**

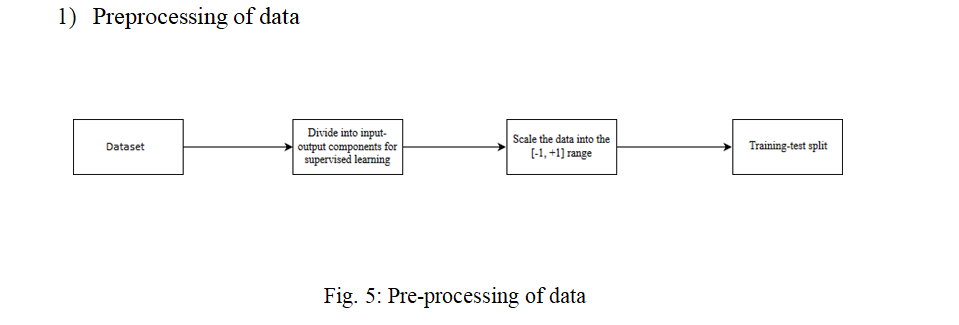
• **LSTM** **is** **used** **instead** **of** **RNN** **to** **avoid** **exploding** **and** **vanishing** **gradients.** • **In** **this** **project** **python** **is** **used** **to** **train** **the** **model.** **MySQL** **is** **used** **as** **a**

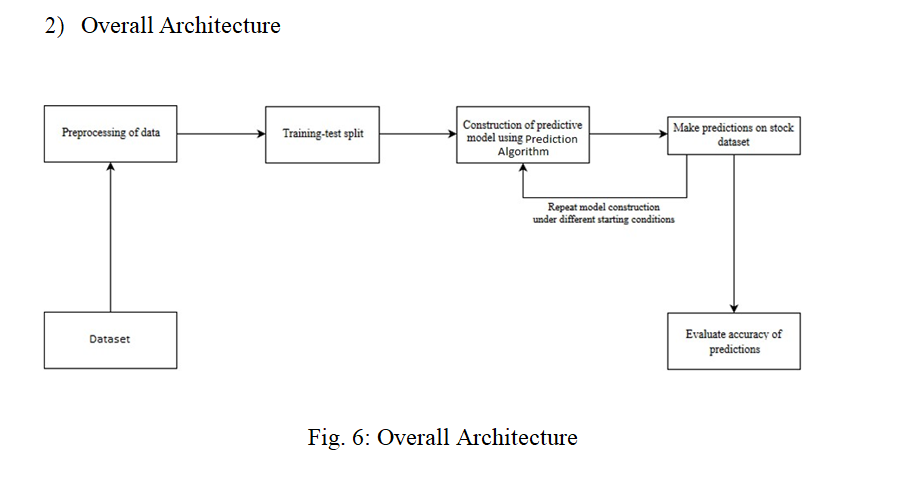
**dataset** **to** **store** **and** **retrieve** **data.**

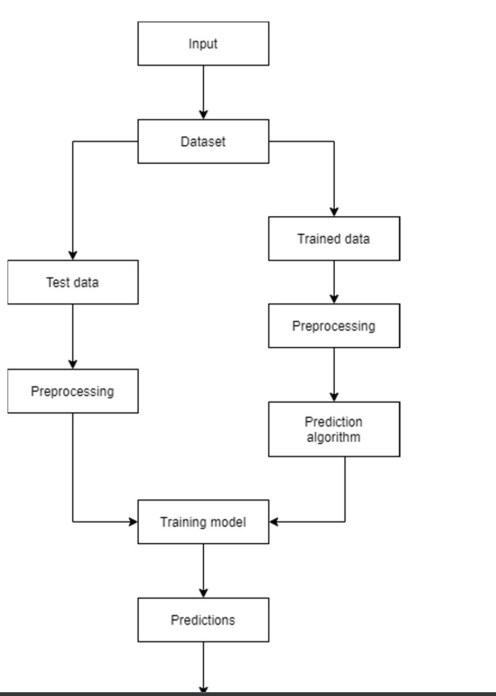
• **The** **historical** **stock** **data** **table** **contains** **the** **information** **of** **opening** **price,** **the** **highest** **price,** **lowest** **price,** **closing** **price,** **transaction** **date,** **volume** **and** **so** **on.**

• **The** **accuracy** **of** **this** **LSTM** **model** **used** **in** **this** **project** **is** **57%.**

**Proposed Work**

• **SYSTEM** **ARCHITECTURE**



Training and prediction

. **PLAN OF ACTION**

***Step1:*** ***Raw*** ***Stock*** ***Price*** ***Dataset:*** Day-wise past stock prices of selected companies are collected fromthe BSE (Bombay Stock Exchange) official website.

***Step*2:** ***Pre-processing***: This step incorporates the following:

a) Data discretization: Part of data reduction but with particular importance, especially for numerical data b) Data transformation: Normalization.

c) Data cleaning: Fill in missing values.

d) Data integration: Integration of data files. After the dataset is transformed into a clean dataset, the dataset is divided into training and testing sets so as to evaluate. Creating a data structure with 60 timesteps and 1 output.

***Step3:*** ***Feature*** ***Selection:*** In this step, data attributes are chosen that are going to be fed to the neural network. In this study Date & Close Price are chosen as selected features.

***Step*** ***4:*** ***Train*** ***the*** ***NN*** ***model***: The NN model is trained by feeding the training dataset. The model s initiated using random weights and biases. Proposed LSTM model consists of a sequential input layer

followed by 3 LSTM layers and then a dense layer with activation. The output layer again consists ofa dense layer with a linear activation function.

***Step5:*** ***Output*** ***Generation:*** The RNN generated output is compared with the target values and errordifference is calculated. The Backpropagation algorithm is used to minimize the error difference byadjusting the biases and weights of the neural network.

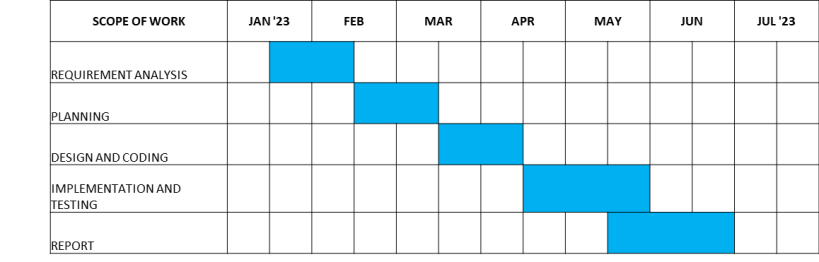
***Step*** ***6:*** ***Test*** ***Dataset*** ***Update:*** Step 2 is repeated for the test data set.

***Step*** ***7:*** ***Error*** ***and*** ***companies’*** ***net*** ***growth*** ***calculation:*** By calculating deviation we check thepercentage of error of our prediction with respect to actual price.

***Step*** ***8****:* ***Visualization***: Using Keras[21] and their function APIs the prediction is visualized.

***Step*** ***9***: Investigate different time interval: We repeated this process to predict the price at different time intervals. For our case, we took 2-month dataset as training to predict 3-month, 6-month, 1 year & 3 years of close price of the share. In this different time span, we calculate the percentage of error in the future prediction. This would be different for different sectors. So, this will help to find a framefor the particular sector to predict

future companies’ net growth.

**PLAN** **OF** **ACTION**

**DATA** **SET**

• historical data downloaded from the Internet

• we downloaded live datasets namely google, nifty, reliance, etc. from the Yahoo Finance website (https://finance.yahoo.com/)

TOOL/PLATFORM USED

Hardware Requirements: • RAM: 4 GB

• Storage: 500 GB

• CPU: 2 GHz or faster

• Architecture: 32-bit or 64-bit Software Requirements:

• Python 3.5 in Google Colab is used for data pre-processing, model training and

prediction.

• Operating System: windows 7 and above or Linux based OS or MAC OS

**Conclusion**

In this project, we are predicting closing stock price of any given organization, we developed a web application for predicting close stock price using LMS and LSTM algorithms for prediction. We have applied datasets belonging to Google, Nifty50, TCS, Infosys and Reliance Stocks and achieved above 95% accuracy for these datasets

**REFERENCES**

Datasets: Nifty 50 data , tw\_spydata\_raw(trading data) Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose1, Giridhar Maji, Narayan C. Debnath, Soumya Sen S. Selvin, R. Vinayakumar, E. A.

Gopalkrishnan, V. K. Menon and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in International Conference on Advances in Computing, Communications and Informatics, 2017

Mariam Moukalled Wassim El-Hajj Mohamad Jaber Computer Science Department American University of Beirut - Automated Stock Price Prediction Using Machine Learning. F. a. o. Eugene, "Efficient capital markets: a review of theory and empirical work," Journal of finance, vol. 25, no. 2, pp. 383-417, 1970.

J. Chen and N. S. Chaudhari, "Segmented-memory recurrent neural networks," IEEE transactionson neural networks, vol. 20, no. 8, pp. 1267-1280, 2009. F. A.

Gers and J. Schmidhuber, "Recurrent nets that time and count," in Proceedings of the IEEE- INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, IEEE, 2000, pp. 189- 194.

