



International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Stock Prediction Using Machine Learning

Prof.Dhanashree Kangane, Mohammed Tabish, Naval Chawan, Rakib Sayed

B.E Computer Engineering, Rizvi College of Engineering, Mumbai, India.

mohammedtabish98@eng.rizvi.edu.in, navalchawan@eng.rizvi.edu.in, rakibsayed@eng.rizvi.edu.in, dhanashree.kangane@eng.rizvi.edu.in

ABSTRACT

Stock prediction using Machine Learning works to predict upcoming worth of stock for an organization. The Prediction is based on the values of previous and current stocks.

In addition, the proposed has a great feature of predicting stocks and giving specific output to the user based on two means, social statistics and gathering earlier data with respect to the stocks. This project focuses on several methods to predict the worth of stocks which aims to predict the future movement of the stock's value of a financial exchange.

Keywords: Stock, Shares, Prediction, ML, Machine Learning, Data Science, Analysis

1. Introduction

Stock market is trading platform where different investors sale and purchase shares according to stock availability. Stock market ups and downs effects the profit of stakeholders. If market prices going up with available stock then stakeholders get profit with their purchased stocks. In other case, if market going down with available stock prices then stakeholders have to face losses. Buyers buy stocks with low prices and sell stocks at high prices and try to get huge profit. Similarly, sellers sell their products at high prices for profit purpose. Stock market (SM) work as trusty platform among sellers and buyers.

Advances in Artificial Intelligence (AI) supporting a lot in each field of life with its intelligent features. Several algorithms present in AI that performing their role in future predictions. Machine learning (ML) is a field of artificial intelligence (AI) that can be considered as we train machines with data and analysis future with test data. Machines can be trained on the basis of some standard that are called algorithms. Stock market predictions can be great beneficial to businessman. SMP provide future trend of stock prices on the basis of previous history.

If stakeholders get future predictions then investment can lead him toward profit. Predictions can be 50% correct and 50% wrong as it is risk of business. Risks facing capability in business filed can lead toward success. In any field of life, we take risks for success. Similarly, we rely on ML predictions about future prices of stock. In this chapter we would like to explain these ML algorithms with the help of their working methodologies and examples. Before working on actual problem SMP, complete understanding of ML algorithms role in prediction is also necessary. That's why in this chapter we explained complete working scenario and problem definition. Several Machine learning algorithms can be used for stock market prediction but in this research we used few algorithms like Long Short-Term Memory (LSTM), Three Months Moving Average (3MMA) and Exponential Smoothing and if we further consider many other algorithms can also be used for Stock Market Prediction (SMP).

2. Literature Review

i. According to Researcher's Mehar Vijh, Deeksha Chandola, Vinay Anand Tikkiwal, Arun Kumar (2019) .

They have done research on Stock Closing Price Prediction using Machine Learning Techniques. This research work has utilized artificial neural

* Corresponding author. Mob No : +917900071666

E-mail address: navalchawan@eng.rizvi.edu.in

network and random forest technique for predicting stock prices for 5 different companies of different sectors of operations. The financial data which they have used in their studies includes open, high, low and stock close prices. They have evaluated the models using standard strategic indicators: RMSE and MAPE. It has been observed that the values of these 2 indicators are low which shows that the models which are proposed by them are efficient in predicting stock closing price. According to Research conducted by Lufuno Ronald, Marwala (2010-09-13) on topic Forecasting the Stock Market Index Using Artificial Intelligence Techniques The frail type of Efficient Market hypothesis (EMH) states that it is difficult to figure what's to come cost of a resource dependent on the data contained in the authentic costs of a resource. This implies that the market carries on as a random walk and as a result makes prediction impossible. Furthermore, financial forecasting is a tough task due to the basic difficulty of the monetary structure. The objective of this work was to utilize artificial intelligence (AI) techniques to model and forecast the upcoming value of a stock market record.

ii. Stock Market Prediction Using Ensemble Of Graph Theory

Stock Market is the marketplace for security where coordinated issuance and exchanging of Stocks happen either through trade or over the counter in electronic or actual structure. It assumes a significant part in canalizing capital from the financial backers to the business houses, which thus prompts the accessibility of assets for business extension. Prediction of the Bombay Stock Exchange (BSE) Market Returns Using Artificial Neural Network and Genetic Algorithm, Research By Yusuf Perwej, Asif Perwej (May 2012) have investigated to predict the daily excess returns of Bombay Stock Exchange (BSE) indices over the individual Treasury bill rate returns. Initially, they prove that the extra return time series do not vary randomly. They have applied the prediction models of Autoregressive feed forward Artificial Neural Networks (ANN) to predict the excess return time series using lagged value. For the Artificial Neural Networks model using a Genetic Algorithm is constructed to choose the optimal topology. This paper analyses the attainability of the expectation task and gives proof that the business sectors are not fluctuating arbitrarily lastly, to apply the most appropriate forecast model and measure their effectiveness.

iii. Efficient Machine Learning Techniques for Stock Market Prediction.

The computerized PC programs utilizing information mining and prescient innovations do a charge measure of exchanges the business sectors. Data mining is all around established on theory that the historic data holds the essential memory anticipating the future bearing. This innovation is intended to assist financial backers with finding concealed examples from the noteworthy information that have plausible prescient ability in their speculation choices. The prediction of securities exchanges is viewed as a difficult assignment of monetary time arrangement prediction. Data analysis is one method of foreseeing if future stocks costs will increment or decline. Five strategies for breaking down stocks were joined to anticipate if the day's end cost would increase or decrease. These methods were Typical Price (TP), Bollinger Bands, Relative Strength Index (RSI), CMI and Moving Average (MA). Financial Stock Market Forecast using Data Mining Techniques Research By K. Senthamarai Kannan, P. Sailapathi Sekar, M. Mohamed Sathik and P. Arumugam (March 17-19 2010) talked about different procedures which can foresee with future shutting stock cost will increment or lessening better than level of importance. Additionally, they examined different worldwide events and their issues predicting on stock markets

3. Scope

In Stock Market Prediction using Machine Learning we planned the usage of information collected from various organizations of financial markets with lots of research regarding past and present movements in the stock market. Data is collected regarding stock market ups and downs, the entire history of previous 5 years. The project gives a precise output compared to the targets. Algorithm such as Support Vector Machine (SVM), Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF) are used to predict the stock worth.

SVM works with bulk dataset value which is gathered from worldwide economic markets. Problem of over fitting never occurs in SVM. One of the suitable algorithms for prediction.

A Radial Basis Function is the real-valued function which depends on the value of the distance from the origin. RBF is a popular algorithm. The RBF Kernel acts as the former which selects the smooth solutions.

Feedforward neural networks qualified with back propagation algorithm are Multilayer Perceptron's (MLPs). Multilayer Perceptron's comes under the supervised networks hence these MLPs need a chosen outcome to learn. MLPs have 3 layers, which consists of the input layer, followed by the hidden layer and the output layer. MLPs are used to give approximate presentation of optimal statistical classifiers in tough problem.

4. Requirements and Analysis

4.1. Requirements Specification

- The requirement is to create an algorithm that could predict the closing price of a particular quote which will help the client to develop different strategies considering the predicted rate.
- The requirement is to create an algorithm that could predict the closing price of a particular quote which will help the client to develop different strategies considering the predicted rate.
- The predicted rate should be accurate or close to the predicted rate.
- The selection of the prediction algorithm should depend on the market conditions and stock instrument to be analyzed.

- Factors like instrument volatility, market sentiments, quantities being trade also come into picture.
- A slight mistake in strategy selection could have an adverse effect in P\L calculations.
- Besides strategy selection other factors like accurate and up to date market data, historical data selection, accurate back-testing are all other factors to be considered.
- Current machine learning algorithms provide varied results and also costs us a lot of computation problems.
- High power consumption, incorrect results, no reliable data are all the other factors providing a clumsy response.

4.2. Software and Hardware Requirements

Hardware Requirement

- Windows 7 or Higher
- Mac OS X 10.11 or higher , 64-bit
- Linux : RHEL 6/7, 64-bit (almost all libraries also work in Ubuntu)
- Ram: 4 GB
- Disk Space requirement: 100 MB
- Graphic Card: NVIDIA GEFORCE GTX 1050ti

Software Requirement

- Compiler: Python compiler 3.8.5
- Software used: visual studio code

LINKS TO THE SOFTWARE:

-Python compiler: <https://www.python.org/downloads/release/python-390/>

-VS CODE: <https://code.visualstudio.com/>

5. System Design

5.1. Basic Modules

i. Index module:

- It contains a title, label with the text login module 1.0 providing the description of the preliminary screen.
- Two buttons Login and Register directing the user to the login and register page.

ii. Login module:

- It contains a label stating the client to enter his details.
- Two input boxes with a label providing details of actions to be performed by the client.
- Client is supposed to enter his username in the first input box and password in the second input box.
- If the information provided in the input boxes is valid the client is directed to the next page.

iii. Registration module:

- It contains a label stating the client to enter his details.
- Two input boxes with a label providing details of actions to be performed by the client.
- Client is supposed to enter his username in the first input box and password in the second input box.
- The user is expected to remember his credentials.

iv. Back-Testing module:

- The first input tab is for the instrument name.
- If the instrument is not available an alert box is displayed stating so.
- A list of instruments would be provided by default.
- The second entry tab is for number of epochs to be performed on the data.
- Once the number is entered epochs are calculated and a graph is uploaded on the screen.

v. Trading module:

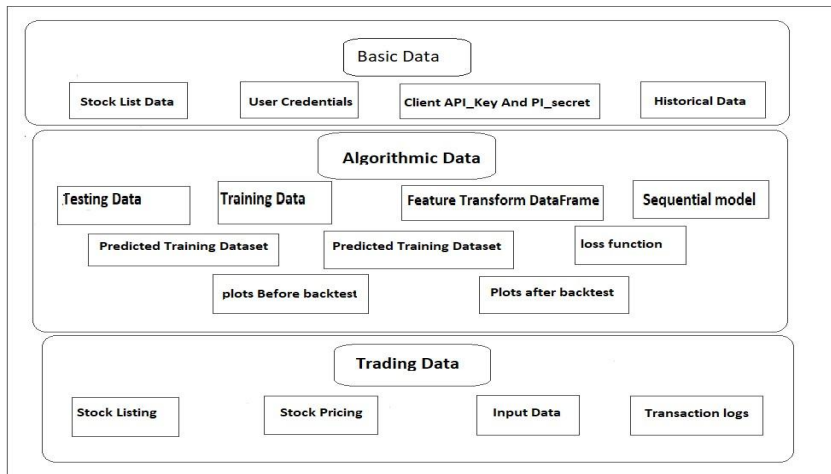
- Label stating the Description of the module.
- User should provide its Account's API_ID and API_SECRET.
- Input tab for user input. • Quantity tab for entering amount of quantity.
- A dropdown list which contains all the available contracts.
- Buy button which buys the selected contract with the selected quantity.
- Sell button which buys the selected contract with the selected quantity.

5.2 Data Design

i. Managing the data

- All client credentials are stored inside a secure(password) protected repository.
- Client credentials have been encrypted.
- The instrument's historical data is stored in .csv format.
- All transactions made by the user are stored in the form of log files.
- User's net positions, historical trades are all stored in .csv format

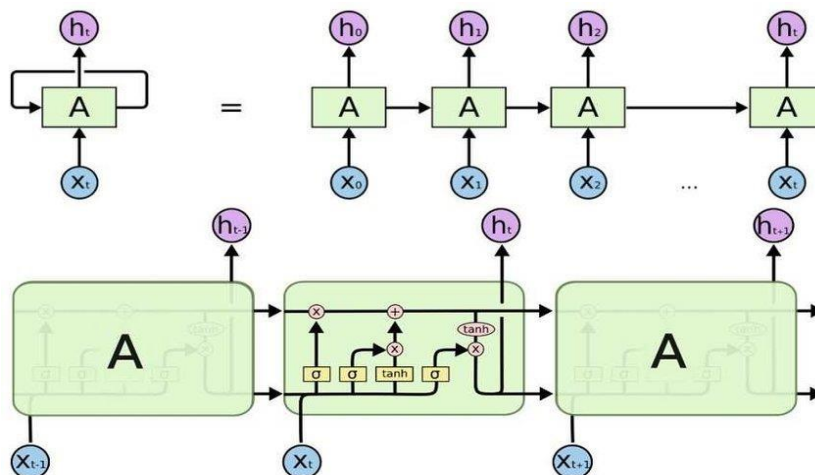
ii. Schema Design



5.3.Procedural design

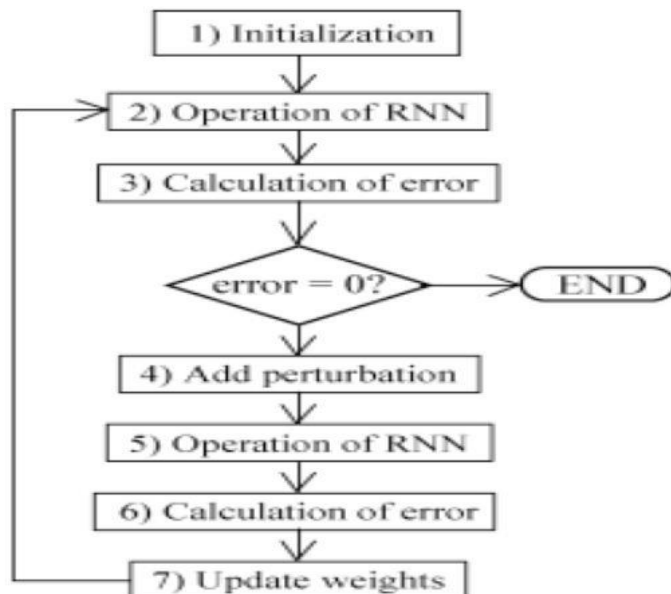
i. Internal LSTM mechanism

- Long Short-Term Memory (LSTM) networks are type of recurrent neural network capable of learning order dependence in sequence prediction problem.
- It contains sigmoid activation functions used for long-short memory manipulation.
- Increase in number of hidden layers increases the complexity and accuracy of the LSTM model and hence is called Stacked LSTM.



- Long Short-Term Memory (LSTM) networks are a type of recurrent neural network capable of learning order dependence in sequence prediction problem.
- It contains sigmoid activation functions used for long-short memory manipulation.
- Increase in number of hidden layers increases the complexity and accuracy of the LSTM model and hence is called Stacked LSTM.

ii. Control Flow Chart



- RNN (Recurrent Neural Network) manages the control flow and memory of the previous data in order to design patterns and thereby analyze the curve and make prediction.

6. Implementation and Experiments Performed

6.1. Implementation

Accurate prediction of stock market returns is a very challenging task due to volatile and non-linear nature of the financial stock markets. With the introduction of Machine Learning and increased computational capabilities, programmed methods of prediction have proved to be more efficient in predicting stock prices. For Stock Prediction Market using Machine Learning we have used python as a programming language with all the libraries which is required for the project. The version used is 3.9.4 which supports libraries like tkinter, pandas, numpy, sklearn, matplotlib.

1. Tkinter

- Using tkinter we have created a login and registration page in order to provide a UI for users to authenticate themselves and use the software.
- Alternatives to tkinter are Kiwi, PyQt, PySide.
- Tkinter v3.9.4 provides an easy to use UI building and handling properties helping a technical person focus on the principle requirement.
- Lower versions of tkinter have a lagging issue, also they lack attractive UI properties.

2. Pandas Datareader

- Library provides a variety of methods to retrieve data from various sources.
- Get_data_tingo method is recent addition to pandas_datareader for receiving stock market historical data from various sources.
- A tiingo account api key is a prerequisite for using this method.
- Other sources like nasdaq_trader, moex, bankofcanada are some of its other methods to retrieve historical data.
- Google Finance have deprecated the use of pandas_datareader discontinuing the free forex international market data.

3. Pandas

- Pandas libraries used for all the dataframe calculations required for cleaning the data and data engineering processes.
- Dask, Vaex, PySpark, Modin are some of the pandas alternatives.
- Any pandas version would be suitable for the project, latest one recommended.

4. Matplotlib

- Extremely helpful library for visualization purposes. Graphs are the most basic amenities provided.

5. Numpy

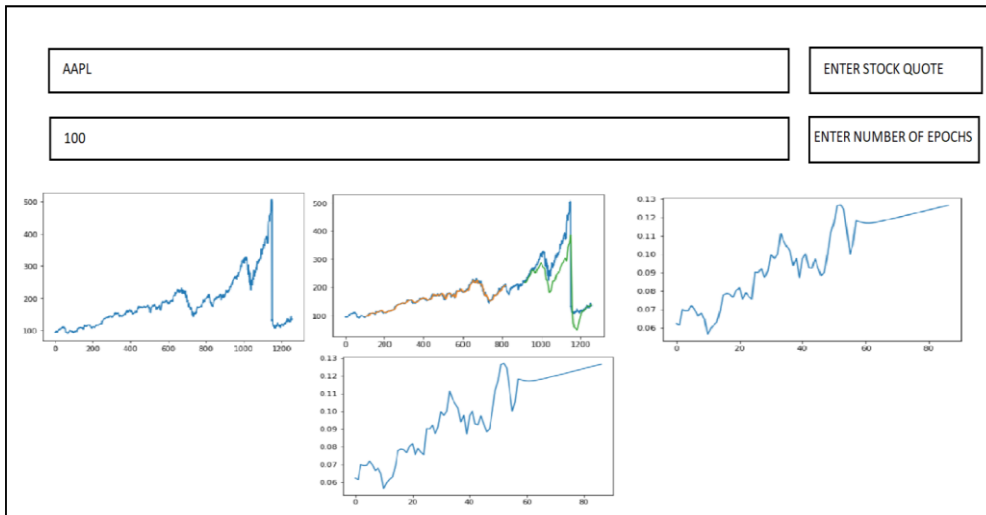
- Numerical data pre processing library used for converting the close price to a appropriate pre-processing format.
- Here our training and testing datasets are converting into a dataset matrix for the building layers of neural network.

6. LSTM

- The actual data back testing is done using the stacked LSTM model.
- Number of hidden layers, loss function and optimizer are to be set for our purpose of preprocessing.
- Tensorflow models are used for setting a network of neurons(nodes) interconnecting each other.

7. SKLEARN

- Sklearn metrics are used for measuring the mean squared error optimizer(loss function) in order to mitigate the errors and making the model as overfitting as possible depending on the epochs provided by the user.
- Once the model is developed a prediction algorithm is used to predict the pattern and provide output for the next 30 days.



➤ OUTPUT

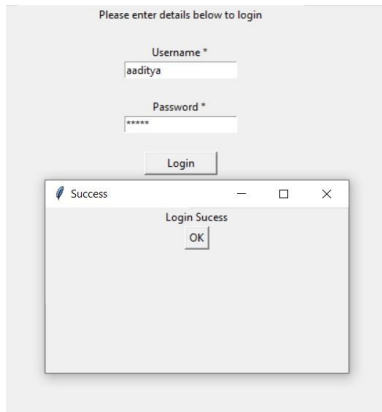
- The output are in the form of graphs providing insights about the trading pattern and prediction for the next 30 days for a particular stock in this case Apple.
- The first graph is the actual market data fluctuation of the last 5 year Apple close price.
- Second graph is the insightful fitting of the fluctuation our model is trying to mimic.
- Third graph is the separated prediction based on the data provided to the model.
- Fourth graph is the joint and complete prediction graph the trader can grab insights from

6.2. Experiments

1. Authentication Page

- Only the registered account could access the software.

- A success message would be generated if the registration is successful.



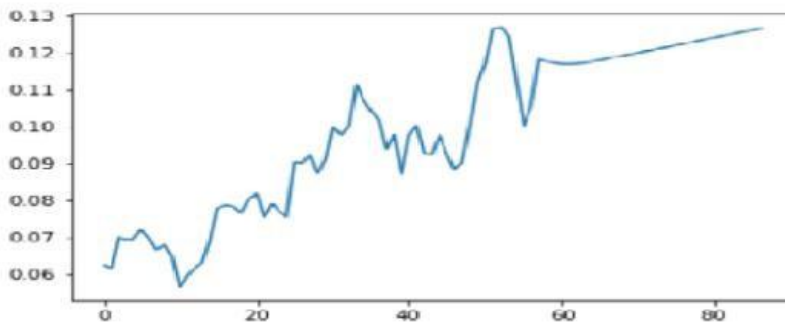
- successful login credentials log's you into the software.



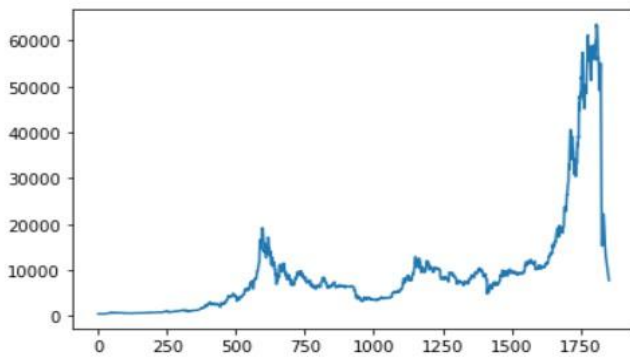
- Unsuccessful login credentials prohibits you from login in.

2. LSTM Experiments

- LSTM models can be used for long term investment opportunities for a product having steady growth patterns and less day to day volatility.
- For example: For a Stock like Apple or Microsoft with steady growth LSTM works best in order make predictions.



- a steady smooth prediction can be expected.
- Were as if we consider a predictor for bitcoin with extremely high volatility the prediction looks irregular and inaccurate



7. Conclusion

Information on stock developments by a small part of a second can prompt high benefits financial backers can make which makes securities exchange reads a significant inspiration for a specialist. The incredible advances and achievement of characteristic language interaction and assessment investigation of online news dependent on AI and profound learning have acquired enormous ubiquity as of late in the monetary area particularly in market forecast models. This overview has talked about the new current examinations on market forecast frameworks dependent on text mining strategies with extensive explaining of the model's fundamental impediments and future improvement techniques. The review was embraced on many significant segments, for example, text preprocessing, AI Algorithms, assessment systems, discoveries, and limits related with point-by-point conversation and clarification of the best utilized procedures. In addition, this audit gives a genuine endeavor to address the issue of market expectation dependent on the latest content mining techniques and give a reasonable perspective on the future examination course. As of late, more broad perceptions into the monetary business sectors are needed in the current unique world, since its shortfall can detrimentally affect the speculations all throughout the planet. It is hence fundamental to embrace forecast models dependent on text mining research as a useful arrangement that can prompt a lot more prominent level of trust in the comprehension of market developments and make important ventures. With the significant measure of textbased information accessible on the web, the need to assemble particular content digging frameworks slowly develops for each field of market examination.

In the wake of going through this capstone project it is clear that utilizing LSTM networks in blend with stacked autoencoder and notion examination yields results acceptable enough to use in live exchanging. Future works for this task will incorporate and require a more in profundity research with an end goal to adjust the best outcome yielding model, the hybrid model to the Moroccan market. One fundamental issue we may experience for this situation would be the sheer volume of accessible information which isn't almost enough for an ideal and productive deep learning model.

This could be tackled by applying some information expansion methods to the all-around existing informational indexes in requests to develop it in size and make it reasonable for a deep learning project.

REFERENCES

- i. Stock Price Prediction Using Machine Learning 2020 <https://thecleverprogrammer.com/2020/11/14/stock-price-prediction-using-machine-learning>
- ii. Stock Price Prediction Using Machine Learning | Deep Learning 2021 <https://www.analyticsvidhya.com/blog/2018/10/predicting-stock-price-machine-learningnddeep-learning-techniques-python/>
- iii. Stock Price Prediction with Machine Learning 2020 <https://thecleverprogrammer.com/2020/05/11/stock-price-prediction-with-machine-learning>
- iv. A Machine Learning Model For Stock Market Prediction <https://www.researchgate.net/publication/259240183>
- v. The Easy Way To Predict Stock Prices Using Machine Learning <https://pedrolealdino.medium.com/the-stupidly-easy-way-to-predict-stock-prices-using-machine-learning-dbb65873cac8>
- vi. Stock Price Prediction - Machine Learning Project in Python 2021 <https://data-flair.training/blogs/stock-price-prediction-machine-learning-project-in-python/>
- vii. Stock price prediction using LSTM (Long Short-Term Memory) <https://www.datasciencecentral.com/profiles/blogs/stock-price-prediction-using-lstm-longshort-term-memory>
- viii. Stock Price Prediction Using Python & Machine Learning 2019 <https://www.youtube.com/watch?v=QIUxPv5PJOY>
- ix. Predicting Stock Prices Using Reinforcement Learning <https://www.analyticsvidhya.com/blog/2020/10/reinforcement-learning-stock-priceprediction>
- x. Stock Closing Price Prediction using Machine Learning SVM Model https://www.academia.edu/44538799/Stock_Closing_Price_Prediction_using_Machine_Learning_SVM_Model