

DESIGN AND IMPLEMENTATION OF STOCK PRICE PREDICTION USING DATA SCIENCE

Submitted in partial fulfillment of the requirements for the award of Bachelor of
Engineering degree in Electronics and Communication Engineering

By

G.V.S. SWAROOP (Reg No: 39130135)

K. MANIKANTA (Reg No: 39130201)



**DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
SCHOOL OF ELECTRICAL AND ELECTRONICS**

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

Accredited with Grade "A" by NAAC

JEPPIAAR NAGAR, RAJIV GANDHISALAI,

CHENNAI - 600119

APRIL - 2023



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited with —A grade by NAAC Jeppiaar

Nagar, Rajiv Gandhi Salai, Chennai – 600 119

www.sathyabama.ac.in

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that this Final Year Project Report is the bonafide work of **G.V.S.SWAROOP (39130135)** and **K. MANIKANTA (39130201)**, who have done the project work as a team who carried out the project entitled “**DESIGN AND IMPLEMENTATION OF STOCK PRICE PREDICTION USING DATA SCIENCE**” under my supervision from JAN 2022 to April 2023.

Internal Guide

Dr. T. BERNATIN, M.E., Ph.D.

Head of the Department

Dr. T. RAVI, M.E., Ph.D.

Submitted for Viva voce Examination held on 21.04.2023

Internal Examiner

External Examiner

DECLARATION

I, **G.V.S. SWAROOP (39130135)** and **K. MANIKANTA (39130201)** hereby declare that the Project Report on “**DESIGN AND IMPLEMENTATION OF STOCK PRICE PREDICTION USING DATA SCIENCE**” done by us under the guidance of **Dr.T.BERNATIN, M.E.,Ph.D.** is submitted in partial fulfilment of the requirements of the award of Bachelor of Engineering degree in Electronics and Communication Engineering.

DATE: 25/04/2023

PLACE: CHENNAI



SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Board of Management** of **SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. N.M. NANDHITHA M.E., Ph. D**, Dean, School of Electronics, **Dr. T. RAVI M.E., Ph.D.**, Head of the Department of Electronics and Communication Engineering for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Dr.T. BERNATIN, M.E.,Ph.D.**, for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my phase-2 project work.

I wish to express my thanks to Teaching and Non-teaching staff members of the **Department of Electronics and Communication Engineering** who were helpful in many ways for the completion of the project.

ABSTRACT

Time series forecasting has been widely used to determine the future prices of stock, and the analysis and modelling of finance time series importantly guide investors' decisions and trades. This work proposes an intelligent time series prediction system that uses sliding-window optimization for the purpose of predicting the stock prices. The system has a graphical user interface and functions as a stand-alone application. The proposed model is a promising predictive technique for highly non-linear time series, whose patterns are difficult to capture by traditional models. Stock market investment is challenging tasks. There is higher probability of losing money but also gain lots of money from it. So, there should be one model which gives the gist of trends in stock market.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTARCT	V
1	INTODUCTION	1
	1.1 INTRODUCTION TO THE PROJECT	1
	1.2 SCOPE OF WORK	1
2	LITERATURE SURVEY	2
3	METHODOLOGY	6
	3.1 EXISTING SYSTEM	6
	3.2 PROPOSED SYSTEM	7
	3.3 SYSTEM ARCHITECTURE	9
	3.4 FEATURES OR TECHNICAL INDICATORS	9
	3.5 SEQUENCE DIAGRAM	12
4	SOFTWARE REQUIREMENTS	13
	4.1 ARTIFICIAL NEURAL NETWORKS	13
	4.2 WORKING OF ARTIFICIAL NEURAL NETWORKS	14
	4.3 USE OF ANN IN OUR PROJECT	16
	4.4 BACKPROPAGATION	18
	4.5 SUPPORT VECTOR MACHINE	19
	4.6 WORKING OF SUPPORT VECTOR MACHINE	20
	4.7 RADIAL BASIS FUNCTION	21
	4.8USE OF SVM IN OUR PROJECT	22
	4.9 LSSVR FORMULATION	22

	4.10 REQUIREMENT ANALYSIS	23
	4.11 FUNCTIONAL REQUIREMENTS	23
	4.12 NON- FUNCTIONAL REQUIREMENTS	25
	4.13 SYSTEM DESIGN AND TESTING PLAN INPUT DESIGN	25
	4.14 SYSTEM STUDY	26
	4.15 TYPES OF TESTS	27
5	RESULTS AND DISCUSSION.	30
	5.1 RESULT	30
	5.2 DISCUSSION	30
6	SUMMARY AND CONCLUSION	31
	6.1 SUMMARY	31
	6.2 MOTIVATION AND OBJECTIVE	31
	6.3 CONCLUSION	32
	6.4 FUTURE ENHANCEMENTS	33
	REFERENCES	34
	APPENDIX	

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION TO THE PROJECT

Predicting stock price trend and its movement has been viewed as a standout amongst the most difficult utilizations of time arrangement expectation. Even though there has been numerous research which manage the issues of anticipating stock price trend, most exact discoveries are related with the developed financial markets. But it is difficult to predict the trend or price of the stock because of the uncertainty in the stock market. There are two types of analysis, Fundamental analysis and technical analysis. In fundamental analysis, performance of the company, economic factors and political factors are considered. In technical factors, previous n days closing price, highest price, lowest price etc. are considered. We can predict the trend of stock or price of the stock using technical analysis. Fundamental analysis is hard to measure and hard to implement in computer language. Technical analysis does not measure the intrinsic security value of the stock, but it uses technical stock charts to predict the trend of the stock. does not measure the intrinsic security value of the stock, but it uses technical stock charts to predict the trend of the stock.

In initial stage of the study of the stock market prediction, classical methods were used. But as stock market is a non-stationary time series of data. It was not so effective. So non-linear machine learning techniques like Artificial neural networks (ANN) and Support Vector Machine (SVM) are used widely. In this project we have used both the techniques to predict the trend of the stock and measured the accuracy of both the techniques.

1.2 SCOPE OF THIS WORK

Stock market investment is challenging tasks. There is higher probability of losing money but also gain lots of money from it. So, there should be one model which gives the gist of trends in stock market. So, using machine learning techniques, we can predict the trend of the stock market and invest in it according to that.

CHAPTER 2

LITERATURE SURVEY

Aditi Singhal and Tavpritesh Sethi (2021) provides a comprehensive survey of recent research on stock market prediction using machine learning techniques. The authors begin by introducing the importance of stock market prediction and the challenges involved in accurately predicting stock prices. They then provide an overview of machine learning techniques used for stock market prediction, including regression, classification, clustering, and deep learning. The literature review then presents a detailed analysis of the various machine learning techniques used for stock market prediction, including Support Vector Machines (SVM), Artificial Neural Networks (ANN), Random Forests (RF), and Gradient Boosting Machines (GBM). The authors compare the performance of these techniques and provide insights into their strengths and limitations. The authors also discuss the importance of feature selection and engineering in stock market prediction and highlight the role of sentiment analysis and natural language processing techniques in predicting stock prices. They have taken Simple 10 Day moving Average, Weighted 10 Day Moving Average, Momentum, Stochastic K%, Stochastic D%, Relative strength Index, Moving Average convergence divergence, Larry William's R%, A/D (Accumulation/Distribution) Oscillator, CCI (Commodity Channel Index) as features and used VM and ANN. And in SVM using Polynomial function they have got 84% accuracy and using RBF kernel function they have got 80% accuracy. In average they got 78.71 % accuracy in SVM and 74.94% accuracy using Artificial Neural Networks.

Rajesh Kumar and Anuj Kumar (2020) provides an overview of machine learning techniques used for stock price prediction. The authors first introduce the challenges of stock price prediction and the need for accurate predictions. They then discuss the different types of machine learning techniques used for stock price prediction, including regression, classification, clustering, and deep learning. The literature review then presents a detailed analysis of various machine learning techniques used for stock price prediction, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees, Random Forest, and Gradient Boosting. The authors compare the performance of these techniques and provide insights into their strengths and limitations. Furthermore, the review highlights the importance of data preprocessing and feature selection in stock price prediction.

It also discusses the use of sentiment analysis, natural language processing, and textual analysis for predicting stock prices. They used Istanbul Stock Exchange to predict the trend of the stock price. In that they have used ANN and polynomial SVM. Using ANN, they have got 75.74 % accuracy and using polynomial SVM 71.52 %. In the polynomial SVM the degree of polynomial is 3 and value of c is 100. 3 VM and ANN. And in SVM using Polynomial function they have got 84% accuracy and using RBF kernel function they have got 80% accuracy. In average they got 78.71 % accuracy in SVM and 74.94% accuracy using Artificial Neural Networks.

Tapan Kumar Satpathy and Manoj Kumar Rout (2019) focuses on the application of machine learning algorithms in stock price prediction. The authors start by introducing the importance of stock price prediction and the challenges involved in predicting stock prices accurately. They then provide an overview of machine learning algorithms used for stock price prediction, including regression, classification, clustering, and deep learning. The literature review presents a detailed analysis of various machine learning algorithms used for stock price prediction, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees, Random Forest, and Gradient Boosting. The authors compare the performance of these techniques and provide insights into their strengths and limitations. Furthermore, the review highlights the importance of feature selection, data preprocessing, and data visualization in stock price prediction. It also discusses the use of sentiment analysis, natural language processing, and textual analysis for predicting stock prices.

Mahdi Esmailpour and Hamid R. Arabnia (2018) provides an overview of machine learning techniques used for stock market prediction. The authors begin by introducing the importance of stock market prediction and the challenges involved in accurately predicting stock prices. They then discuss the different types of machine learning techniques used for stock market prediction, including regression, classification, clustering, and deep learning. The literature review then presents a detailed analysis of various machine learning techniques used for stock market prediction, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees, Random Forest, and Gradient Boosting. The authors compare the performance of these techniques and provide insights into their strengths and limitations. The review also discusses the importance of feature selection, data preprocessing, and data visualization in stock market prediction. It highlights the role of sentiment analysis and natural language processing techniques in predicting stock prices.

Yu Ren, Wei-Jie Chen, and Hui-Ju Tsai (2016) provides an overview of machine learning techniques used for stock market prediction. The authors start by introducing the importance of stock market prediction and the challenges involved in predicting stock prices accurately. They then discuss the different types of machine learning techniques used for stock market prediction, including regression, classification, clustering, and deep learning. The literature review presents a detailed analysis of various machine learning techniques used for stock market prediction, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees, Random Forest, and Ensemble Learning. The authors compare the performance of these techniques and provide insights into their strengths and limitations. Furthermore, the review highlights the importance of data preprocessing, feature selection, and model evaluation in stock market prediction. It also discusses the use of sentiment analysis and natural language processing techniques for predicting stock prices.

Tarun Kumar Sharma, Kamlesh Dutta, and Ankit Singla (2020) provides an overview of machine learning techniques used for stock price prediction. The authors begin by discussing the importance of stock price prediction and the challenges involved in predicting stock prices accurately. They then present a detailed analysis of various machine learning techniques used for stock price prediction, including regression, classification, clustering, and deep learning. The literature review presents an extensive analysis of various machine learning algorithms used for stock price prediction, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees, Random Forest, and Gradient Boosting. The authors compare the performance of these techniques and provide insights into their strengths and limitations. Furthermore, the review highlights the importance of feature selection, data preprocessing, and model evaluation in stock price prediction. It also discusses the use of sentiment analysis and natural language processing techniques for predicting stock prices.

Prithvi Raj, K. Suresh Kumar, and R. Radhakrishnan (2019) provides an overview of various stock market prediction techniques. The authors begin by discussing the importance of stock market prediction and the challenges involved in predicting stock prices accurately. They then present a detailed analysis of various stock market prediction techniques, including technical analysis, fundamental analysis, and quantitative analysis. The literature review provides an extensive analysis of technical analysis, including Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI), and Bollinger Bands.

It also discusses fundamental analysis techniques, such as earnings analysis and industry analysis, and quantitative analysis techniques, including regression analysis and machine learning techniques. Furthermore, the review highlights the importance of feature selection, data preprocessing, and model evaluation in stock market prediction. It also discusses the use of sentiment analysis and natural language processing techniques for predicting stock prices.

Yakup Kara, Melek Acar Boyacioglu, Ömer Kaan Baykan provides an overview on predicting direction of stock price index movement using artificial neural networks and support vector machines they used Istanbul Stock Exchange to predict the trend of the stock price. In that they have used ANN and polynomial SVM. Using ANN, they have got 75.74 % accuracy and using polynomial SVM 71.52 %. In the polynomial SVM the degree of polynomial is 3 and value of c is 100. Momentum analysis-based stock market prediction using adaptive Neuro-Fuzzy Inference System (ANFIS) This paper presents an innovative approach for indicating stock market decisions. that the investor should take for minimizing the risk involved in making investments. The system uses Adaptive Neuro-Fuzzy Inference System (ANFIS) for taking decisions based on the values of technical indicators. Among the various technical indicators available, the system uses weighted moving averages, divergence and RSI (Relative Strength Index). Detailed method and the rules defining the ANFIS architecture (which were developed using empirical examinations) are provided. Stock market prediction system with modular neural network, in Proceedings of the International Joint Conference on Neural Networks This paper discusses a buying and selling timing prediction system for stocks on the Tokyo Stock Exchange and analysis of internal representation. It is based on modular neural networks. We developed a number of learning algorithms and prediction methods for the TOPIX (Tokyo Stock Exchange Prices Indexes) prediction system. The prediction system achieved accurate predictions and the simulation on stocks trading showed an excellent profit. The prediction system was developed by Fujitsu and Nikko Securities. • Training neural network with genetic algorithms for forecasting the stock price index. The paper combines genetic algorithms (GA) with neural network (NN). It trains NN with GA and then predicts the stock price index with the trained network. By learning the special stock knowledge, it can find out the modes and relationship hidden in the abstract data. It can help 4 shareholders and investment agencies to make wise decisions in the stock market to get more profits. The primary data are from Shanghai Stock Exchange from March 29, 1994, to August 1, 1994. The imitation result shows that the network is fit for short time prediction, and it has high precision.

Integration of genetic algorithm with artificial neural network for stock market forecasting

Traditional statistical as well as artificial intelligence techniques are widely used for stock market forecasting. Due to the nonlinearity in stock data, a model developed using the traditional or a single intelligent technique may not accurately forecast results. Therefore, there is a need to develop a hybridization of intelligent techniques for an effective predictive model. In this study, we propose an intelligent forecasting method based on a hybrid of an Artificial Neural Network (ANN) and a Genetic Algorithm (GA) and uses two US stock market indices, DOW30 and NASDAQ100, for forecasting. The data were partitioned into training, testing, and validation datasets. The model validation was done on the stock data of the COVID-19 period. The experimental findings obtained using the DOW30 and NASDAQ100 reveal that the accuracy of the GA and ANN hybrid model for the DOW30 and NASDAQ100 is greater than that of the single ANN(BPANN) technique, both in the short and long term.

CHAPTER 3

METHODOLOGY

3.1 EXISTING SYSTEM:

Time series forecasting consists of a research area designed to solve various problems, mainly in the financial area. Support vector regression (SVR), a variant of the SVM, is typically used to solve nonlinear regression problems by constructing the input-output mapping function. The least squares support vector regression (LSSVR) algorithm is a further development of SVR, and its use considerably reduces computational complexity and increases efficiency compared to standard SVR. The Firefly Algorithm (FA), which is a nature-inspired metaheuristic method, has recently performed extremely well in solving various optimization problems. The existing system for stock price prediction using machine learning techniques can vary depending on the specific methodology and approach used. However, in general, these systems involve the use of historical stock price data to train machine learning models that can predict future stock prices. One common approach is to use regression models, such as linear regression or polynomial regression, to predict future stock prices based on historical price data and other relevant features, such as market indices, news sentiment, or economic indicators. Another approach is to use machine learning algorithms, such as Support Vector Machines (SVM), Random Forest, or Neural Networks, to predict stock prices based on historical data and other relevant features. Most existing systems for stock price prediction using machine learning techniques involve the following steps. This involves collecting historical stock price data and other relevant features that can be used to predict future stock prices. This involves cleaning and preprocessing the data to remove any noise or inconsistencies in the data, and to transform the data into a format suitable for machine learning algorithms. This involves selecting the most relevant features that can be used to predict future stock prices, using techniques such as correlation analysis, PCA, or feature importance ranking. This involves training machine learning models on historical stock price data and other relevant features, using techniques such as cross-validation or grid search to optimize model performance.

This involves evaluating the performance of the trained models using metrics such as MSE, RMSE, or accuracy, and comparing the performance of different models to select the best-performing model. This involves deploying the trained model to predict future stock prices in

real- time, and continuously updating the model as new data becomes available.

Disadvantages:

The existing system focuses on the stock price market in Taiwan but does not generalize for other markets worldwide. The system does not allow the import of raw data directly. The existing system cannot be used to analyze multi-variate time series. Lastly, the system does not have a user-interface which can be distributed as a web app to users for personal use. Overall, existing systems for stock price prediction using machine learning techniques have shown promising results, but they still face several challenges, such as data availability, model interpretability, and model reliability. There is a need for further research and development to address these challenges and improve the accuracy and robustness of stock price prediction models.

3.2 PROPOSED SYSTEM:

To generalize the application of the existing system, our work uses the system to estimate other stocks in similar emerging markets and mature markets. The system can be extended to analyze multivariate time series data and import raw dataset directly. Profit can be maximized even when the corporate stock market is has lower value. The development of a web-based application has been considered to improve the user-friendliness and usability of the expert system.

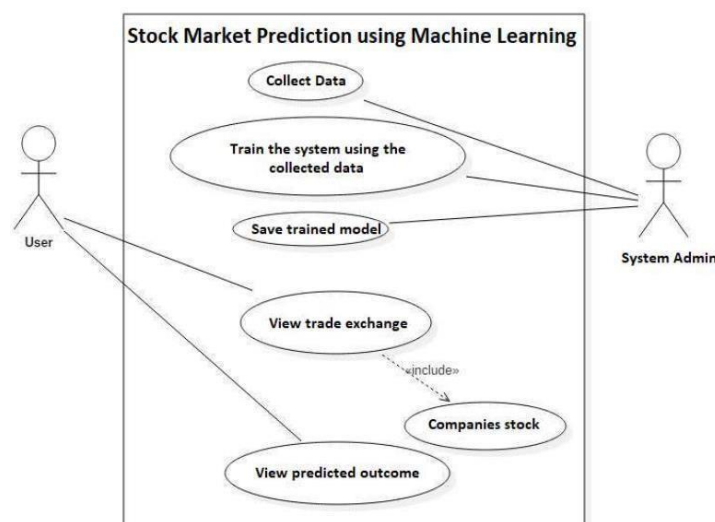
In our project we have used to SVM and ANN as a classifier for learning the trend of stocks of next day. Both are compared on the bases of accuracy. We are working on yahoo stocks which has data of 5 years. We have calculated 11 features. We are using tensorflow and keras library for machine learning models. Artificial neural network created is based on keras. We are training my model in such a way that the trend of the next day is taken as target value for current day. So, our model will predict the trend for next day.

We are focusing more on calculating the and finding features which affects the stock prices. Right now, i have only calculated 11 features using moving average for 10 days.

We have selected these features based on literature survey of research papers. We have taken yahoo stocks for prediction. Feature Engineering: Feature engineering is the process

of creating new features from the existing ones to improve the performance of the model. This can include adding lagged variables, creating technical indicators, and extracting sentiment features from news articles. Feature engineering is the process of creating new features from the existing ones to improve the performance of the model. This can include adding lagged variables, creating technical indicators, and extracting sentiment features from news articles. The next step is to select the appropriate machine learning model that can predict the stock prices accurately. This can include regression models such as Linear Regression, Polynomial Regression, or machine learning algorithms such as Random Forest, Support Vector Machines, or Neural Networks. The selected model needs to be trained on the preprocessed data. The training process involves splitting the data into training and validation sets and tuning the hyperparameters to improve the performance of the model. The performance of the trained model needs to be evaluated using various performance metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Accuracy. The evaluation results are used to compare the performance of different models and select the best-performing one. Once the best-performing model is selected, it is deployed to predict future stock prices in real-time. The model is continuously updated as new data becomes available. Overall, the proposed system for stock price prediction using data science involves a comprehensive approach to data collection, preprocessing, feature engineering, model selection, training, evaluation, and deployment. This approach can potentially lead to accurate and reliable predictions of stock prices, which can be used by investors and traders to make informed investment decisions.

3.3 SYSTEM ARCHITECTURE:



USE CASE DIAGRAM

1. Data is initially collected from online sources or the stock exchange.
2. The data is then used to train the system.
3. Trained model is saved.
4. User views the trade exchange and stock of a company.
5. Using the model, closing prices are predicted.

3.4 FEATURES OR TECHNICAL INDICATORS

Features calculated are mentioned below. Simple 10 days moving average.

Weighted 10 days moving average Momentum

Stochastic K%

Stochastic D%

Relative strength index (RSI)

Moving average convergence divergence (MACD)Larry William's R%

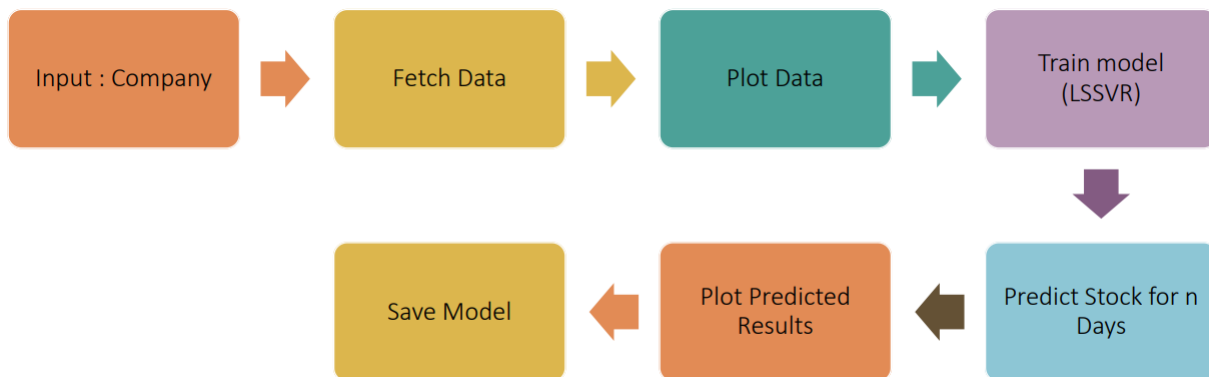
A/D (Accumulation/Distribution) oscillator

CCI (Commodity channel index)

Month index

FLOW OF OUR PROJECT

Data flow diagram:



Start

Read the dataset.

Define features and label Encode the dependent variable

Divide the dataset into two parts for training and testing

Implement the model, Train the model.

TensorFlow data structure for holding features, labels etc.

Reduce MSE

Make prediction on the test data End.

3.5 SEQUENCE FLOW:

- 1) User visits the website/webapp.
- 2) Previously saved model is loaded.
- 3) User requests for a company's stock data
- 4) He requests for prediction to be made.
- 5) The Stock Market Prediction System trains a model using the data from the database.
- 6) The model is saved for further use and closing price is predicted
Result is displayed along with graph.

CHAPTER- 4

SOFTWARE REQUIREMENTS

4.1 Neural Networks

Artificial neural networks are one of the primary instruments utilized as a part of machine learning. As the "neural" some portion of their name recommends, they are cerebrum enlivened frameworks which are expected to duplicate the way that we people learn. Neural networks comprise of information and yield layers, and (much of the time) a hidden layer comprising of units that change the contribution to something that the output layer can utilize. They are brilliant instruments for discovering designs which are unreasonably unpredictable or various for a human software engineer to concentrate and instruct the machine to perceive.

While neural networks (likewise called "perceptron") have been around since the 1940s, it is just over the most recent a very long while where they have turned into a noteworthy piece of artificial intelligence. This is because of the entry of a strategy called "backpropagation," which enables networks to change their concealed layers of neurons in circumstances where the result doesn't coordinate what the maker is seeking after like a network intended to recognize dog, which misidentifies a cat, for instance.

Another vital progress has been the landing of deep learning neural networks, in which distinctive layers of a multilayer network extricate diverse features until the point when it can recognize what it is searching for.

The input layer of an ANN receives the data, which is then processed through one or more hidden layers before being output by the output layer. During training, the weights and biases of the network are adjusted to minimize the error between the predicted output and the actual output. There are several types of ANNs that can be used for stock price prediction, including Feedforward Neural Networks, Recurrent Neural Networks, and Convolutional Neural Networks. Each type of ANN has its strengths and weaknesses and can be used to address different types of problems.

ANNs have shown promising results in stock price prediction, with some studies reporting accuracy rates of over 90%. However, ANNs can be computationally expensive and require

a large amount of training data. Additionally, they can be prone to overfitting if not properly optimized and regularized.

Despite these challenges, ANNs remain a popular and effective machine learning technique for stock price prediction using data science.

4.2 WORKING OF ARTIFICIAL NEURAL NETWORKS

For an essential thought of how a deep learning neural network learns, imagine a manufacturing plant line. After the crude materials (the informational collection) are input, they are then passed down the transport line, with each subsequent stop or layer separating an alternate arrangement of abnormal state highlights. If the network is expected to recognize an object, the brightness of pixel may be analyzed by first layer.

Artificial Neural Networks (ANNs) are a type of machine learning algorithm that is inspired by the structure and function of the human brain. ANNs consist of interconnected nodes, or neurons, that process and transmit information.

In the context of stock price prediction using data science, ANNs can be trained on historical stock price data and other relevant financial and economic data to make predictions about future stock prices. The following is a brief overview of how ANNs work in stock price prediction:

Input Layer: The input layer of an ANN receives the data, which is usually normalized or standardized to ensure that each input variable has the same scale.

Hidden Layers: The data is then processed through one or more hidden layers of neurons, which apply weights and biases to the input data to transform it into a higher-dimensional space.

Activation Function: Each neuron in the hidden layers applies an activation function, which determines whether or not the neuron should "fire" and transmit its signal to the next layer. Common activation functions include the sigmoid function, the hyperbolic tangent function, and the rectified linear unit (ReLU) function.

Output Layer: The processed data is then output by the output layer, which usually consists of a single neuron that represents the predicted stock price.

Training: During training, the weights and biases of the network are adjusted using an optimization algorithm such as backpropagation to minimize the error between the predicted

output and the actual output.

Prediction: Once the network has been trained, it can be used to predict future stock prices based on new input data.

Overall, ANNs can be very effective in predicting stock prices using data science, particularly when they are trained on large amounts of high-quality data. However, the accuracy of the predictions can be affected by a wide range of factors, including the quality of the input data, the complexity of the model, and the optimization algorithm used during training. The following layer could then identify any edges in the picture, based on lines of comparative pixels. After this, another layer may recognize surfaces and shapes, etc. When the fourth or fifth layer is achieved, the deep learning net will have made complex component indicators. It can make sense of that specific picture components, (for example, a couple of eyes, a nose, and a mouth) are regularly discovered together.

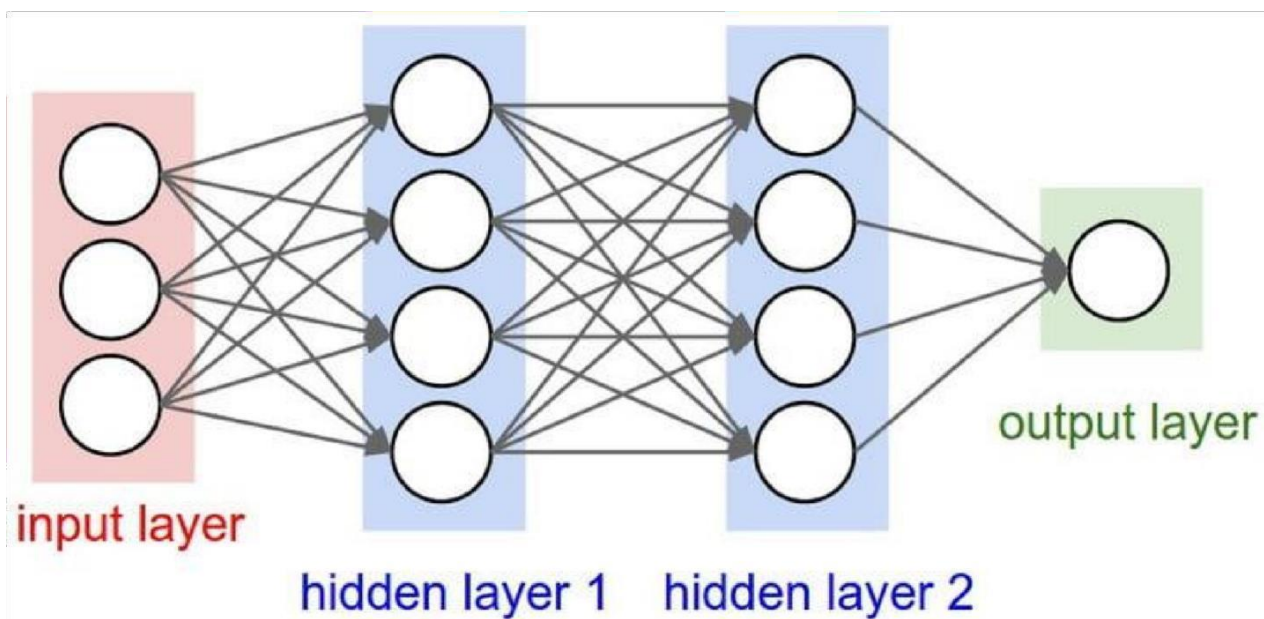


Fig 2. Artificial Neural Network

When this is done, the analysts who have prepared the network can give labels to the output, and afterward utilize backpropagation to redress any oversights which have been made. Sooner or later, the network can complete its own characterization assignments without requiring people to help unfailingly.

4.3 USE OF ANN IN OUR PROJECT

layers uses relu activation function and last layer uses sigmoid. Backpropagation using Gradient Descent Optimizer is used for reducing error. Main reason behind this is relu function creates more sparse data when $W \cdot X + B < 0$ and other reason is it reduces chances of the gradient to vanish.

Using relu function faster learning is done. taking sigmoid or relu activation function in output layer i.e., last layer gives almost same accuracy, so we have taken sigmoid function to obtain nonlinear model. Also, we have normalized the dataset. The use of ANNs in stock prediction involves the following steps:

Data Preparation: The first step is to collect and preprocess the data. This involves cleaning the data, removing outliers, and handling missing values.

Feature Selection: The next step is to select the relevant features that can be used to predict stock prices. This can include technical indicators, economic indicators, and news sentiment.

Model Architecture: The next step is to design the architecture of the neural network. This involves selecting the number of layers, the number of nodes in each layer, the activation functions, and the optimization algorithm.

Model Training: Once the architecture of the neural network is defined, the next step is to train the model using historical data. The training process involves feeding the input data to the neural network and adjusting the weights to minimize the error between the predicted and actual stock prices.

Model Evaluation: The performance of the trained model is evaluated using various performance metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Accuracy. The evaluation results are used to compare the performance of different models and select the best-performing one.

Model Deployment: Once the best-performing model is selected, it is deployed to predict future stock prices in real-time.

ANNs have been found to be effective in predicting stock prices due to their ability to learn complex patterns in the data. However, they also have some limitations, such as overfitting, which can lead to poor performance when the model is applied to new data. Therefore, it is important to carefully design and train the neural network to avoid overfitting and achieve good performance.

The accuracy got in ANN model is 71.73%

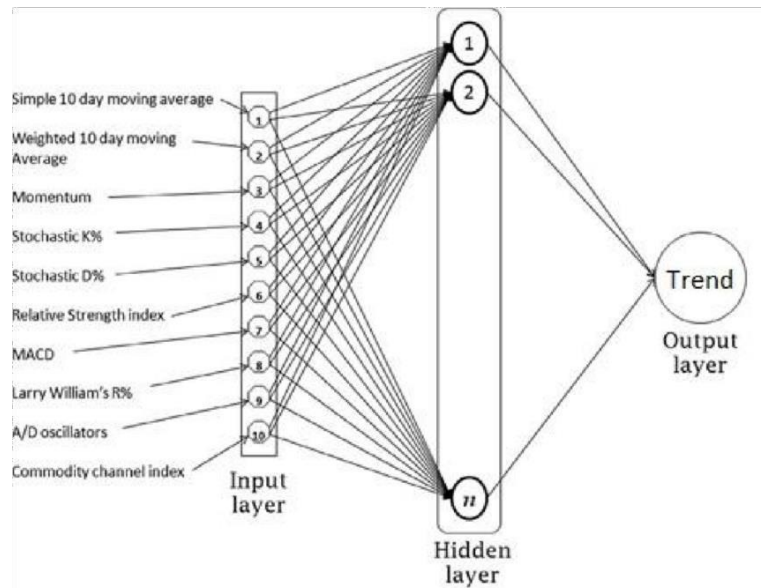


Fig 3. ANN model

4.4 BACKPROPAGATION

Backpropagation, another way to say, "backward propagation of errors," is a calculation for supervised learning of artificial neural networks utilizing gradient descent. Given an artificial neural network and an error function, the strategy ascertains the gradient of the error function regarding the neural network's weights. It is a speculation of the delta manage for perceptronsto multilayer feedforward neural networks.

The "backwards" some portion of the name originates from the way that computation of the gradient continues backwards through the network, with the gradient of the last layer of weights being ascertained first and the gradient of the principal layer of weights being figured last. Fractional calculations of the gradient from one layer are reused in the calculation of the gradient for the past layer. This backwards stream of the mistake data takes into account productive calculation of the gradient at each layer versus the innocent approach of computing the gradient of each layer independently.

Backpropagation's prevalence has encountered a current resurgence given the boundless appropriation of deep neural networks for picture acknowledgment and discourse acknowledgment. It is viewed as an effective calculation, and current executions exploit GPUs to additionally enhance execution.

4.5 SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a supervised learning model that is introduced by Vapnik. It is used for classification as well as regression. For regression Support Vector Regressor (SVR) is used. The main objective of the SVM model is to separate data points as far as possible from a hyperplane. Hyperplane is a plane that divides the data points according to their labels. In addition to linear classification, SVM can classify the non-linear data using kernel trick that maps the data into high dimensional space and transforms it into linear separable data. In general term support Vector Machine defines a hyperplane in high dimension which can be used to classify or divide the data points. The main kernel functions are polynomial kernel, Gaussian kernel, Radial Basis Function (RBF), Linear kernel and Sigmoid kernel.

SVM works by finding a hyperplane in the data space that maximally separates the data points of different classes. In the case of stock price prediction, SVM can be trained on historical stock prices and other relevant data to find a hyperplane that separates the data into two classes: "buy" and "sell". The predicted stock price can then be calculated based on the position of the new data point relative to the hyperplane.

The following are the steps involved in using SVM for stock price prediction using data science:

Data preprocessing: The historical stock price data and other relevant data is collected and preprocessed to ensure that it is in a format that can be used by the SVM algorithm.

Feature extraction: Relevant features are extracted from the preprocessed data, which may include technical indicators, economic data, and other variables that may affect the stock price.

The SVM algorithm is trained on the preprocessed and feature extracted data to create a model that can be used to make predictions about future stock prices. Once the model has been trained, it can be used to predict future stock prices based on new data.

SVM has been shown to be effective in stock price prediction using data science, particularly when combined with other machine learning algorithms such as ANNs and decision trees. However, the accuracy of the predictions can be affected by factors such as the quality and quantity of the data used for training, the choice of features, and the parameters used in the SVM algorithm.

4.6 WORKING OF SUPPORT VECTOR MACHINE

For binary classification, suppose x_i be the input vectors and y_i be the corresponding output labels. Here for binary classification $y_i \in \{1, -1\}$. SVM maps the input vector into more dimensional space $\phi(x_i)$ and make a hyperplane with maximized distance between two classes. Mapping is done by a kernel function $K(x_i, x_j)$. And the final classifier is given in the equation below.

$$f(x) = \text{sgn} \left(\sum_{i=1}^N y_i \alpha_i \cdot K(x, x_i) + b \right)$$

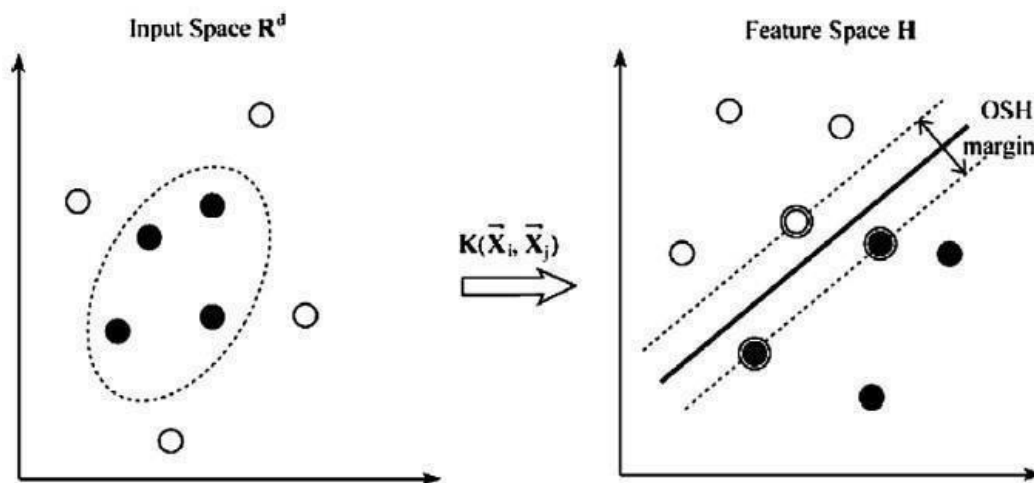


Fig 4. Support vector machine

$$\begin{aligned} & \text{Maximize } \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \cdot y_i y_j \cdot K(x_i, x_j) \\ & \text{subject to } 0 \leq \alpha_i \leq c \\ & \sum_{i=1}^N \alpha_i y_i = 0 \quad i = 1, 2, \dots, N \end{aligned}$$

Where c is a regularization perimeter, which suggest the trad off between distance of hyperplane from the classes and misclassification error. Larger value c , smaller margin hyperplane and less misclassification and smaller the value of c , larger margin hyper planebut they may misclassify the bounty points. Some of the kernel functions are:

Polynomial Function: $K(x_i, x_j) = (x_i * x_j + 1)^d$

Radial Basis Function: $K(x_i, x_j) = \exp(-\gamma ||(x_i - x_j)||^2)$

Here d is degree of the polynomial function and γ is a constant.

4.7 RADIAL BASIS FUNCTION

Radial Basis Function measures how far is the input vector(x) from the hyperplane vector (μ).

$$\phi_j(x) = \exp\left\{-\frac{||x - \mu_j||^2}{2\sigma_j^2}\right\}$$

Here x is the input vector and μ and σ represent the center and spread of dataset points. To adjust the μ and σ below formula is used.

$$\mu_{Class_k} = \frac{1}{|Class_k|} \sum_{x^{(n)} \in Class_k} x^{(n)}$$

$$\sigma_{Class_k}^2 = \frac{1}{|Class_k|} \sum_{x^{(n)} \in Class_k} ||x^{(n)} - \mu_k||^2$$

4.8 USE OF SVM IN OUR PROJECT

NO.	Kernel Function	d	C	Avg. Accuracy
1	Polynomial	2	100	52.79%
2	Polynomial	3	100	52.79%
3	Radial Basis	-	100	51.52%

4.9 LSSVR FORMULATION:

Least Squares Support Vector Regression The least squares version of the SVM classifier is obtained by reformulating the minimization problem as

$$\min J_2(w, b, e) = \frac{\mu}{2} w^T w + \frac{\zeta}{2} \sum_{i=1}^N e_{c,i}^2,$$

For the kernel function $K(x, x_i)$ one typically has the Radial Basis Function

$$K(x, x_i) = \exp\left(-\|x - x_i\|^2 / \sigma^2\right),$$

- The LSSVR algorithm is a further development of SVR by Suykens (2001) and involves equality instead of inequality constraints and works with a least squares objective function.
- The LSSVR approach considerably reduces computational complexity and increases efficiency compared to standard SVM.
- LSSVR solves linear equations instead of a quadratic programming problem.

Least Squares Support Vector Regression (LSSVR) is a type of support vector machine algorithm that can be used for regression tasks. In the context of stock prediction using data science, LSSVR can be used to make predictions about future stock prices based on historical data and other relevant financial and economic data.

The LSSVR formulation involves finding a hyperplane that minimizes the difference between the predicted values and the actual values in the training data. The following are the steps involved in using LSSVR for stock price prediction using data science:

The historical stock price data and other relevant data is collected and preprocessed to ensure that it is in a format that can be used by the LSSVR algorithm. Relevant features are extracted from the preprocessed data, which may include technical indicators, economic data, and other variables that may affect the stock price. The LSSVR algorithm is trained on the preprocessed and feature extracted data to create a model that can be used to make predictions about future stock prices. Once the model has been trained, it can be used to predict future stock prices based on new data. LSSVR has been shown to be effective in stock price prediction using data science, particularly when combined with other machine learning algorithms such as ANNs and decision trees. However, the accuracy of the

predictions can be affected by factors such as the quality and quantity of the data used for training, the choice of features, and the parameters used in the LSSVR algorithm.

4.10 REQUIREMENT ANALYSIS

Requirement analysis, also called requirement engineering, is the process of determining user expectations for a new modified product. It encompasses the tasks that determine the need for analysing, documenting, validating and managing software or system requirements. The requirements should be documentable, actionable, measurable, testable and traceable related to identified business needs or opportunities and define to a level of detail, sufficient for system design.

4.11 FUNCTIONAL REQUIREMENTS

It is a technical specification requirement for the software products. It is the first step in the requirement analysis process which lists the requirements of particular software systems including functional, performance and security requirements. The function of the system depends mainly on the quality hardware used to run the software with given functionality.

Usability

It specifies how easy the system must be use. It is easy to ask queries in any format which is short or long, porter stemming algorithm stimulates the desired response for user.

Robustness

It refers to a program that performs well not only under ordinary conditions but also under unusual conditions. It is the ability of the user to cope with errors for irrelevant queries during execution.

Security

The state of providing protected access to resource is security. The system provides good security and unauthorized users cannot access the system there by providing high security.

Reliability

It is the probability of how often the software fails. The measurement is often expressed in MTBF (Mean Time Between Failures). The requirement is needed in order to ensure that the processes work correctly and completely without being aborted. It can handle any load and survive and survive and even capable of working around any failure.

Compatibility

It is supported by version above all web browsers. Using any web servers like localhost makes the system real-time experience.

Flexibility

The flexibility of the project is provided in such a way that it has the ability to run on different environments being executed by different users.

Safety

Safety is a measure taken to prevent trouble. Every query is processed in a secured manner without letting others to know one's personal information.

4.12 NON- FUNCTIONAL REQUIREMENTS

Portability

It is the usability of the same software in different environments. The project can be run in any operating system.

Performance

These requirements determine the resources required, time interval, throughput and everything that deals with the performance of the system.

Accuracy

The result of the requesting query is very accurate and high speed of retrieving information. The degree of security provided by the system is high and effective.

Maintainability

Project is simple as further updates can be easily done without affecting its stability. Maintainability basically defines that how easy it is to maintain the system. It means that how easy it is to maintain the system, analyse, change and test the application. Maintainability of this project is simple as further updates can be easily done without affecting its stability.

4.13 SYSTEM DESIGN AND TESTING PLAN INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people

keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple.

The input is designed in such a way so that it provides security and ease of use with retaining the privacy.

Input Design considered the following things:

- i. What data should be given as input?
- ii. How the data should be arranged or coded?
- iii. The dialog to guide the operating personnel in providing input.
- iv. Methods for preparing input validations and steps to follow when error occur.

OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user.

Efficient and intelligent output design improves the system's relationship to help user decision-making.

The output form of an information system should accomplish one or more of the following objectives.

- ☐ Convey information about past activities, current status or projections of the
- ☐ Future.
- ☐ Signal important events, opportunities, problems, or warnings.
- ☐ Trigger an action.
- ☐ Confirm an action.

4.14 SYSTEM STUDY

FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out.

This is to ensure that the proposed system is not a burden to the company.

For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are:

- ◆ ECONOMICAL FEASIBILITY
- ◆ TECHNICAL FEASIBILITY
- ◆ SOCIAL FEASIBILITY

ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified.

Thus, the developed system is well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

SOCIAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity.

The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

4.15 TYPES OF TESTS

UNIT TESTING

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application. It is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

INTEGRATION TESTING

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfactory, as shown by successful unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

TEST STRATEGY AND APPROACH

Field testing will be performed manually, and functional tests will be written in detail.

TEST OBJECTIVES

- All field entries must work properly.
- Pages must be activated from the identified link.
- The entry screen, messages and responses must not be delayed.

FEATURES TO BE TESTED

- Verify that the entries are of the correct format.
- No duplicate entries should be allowed.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 RESULT:

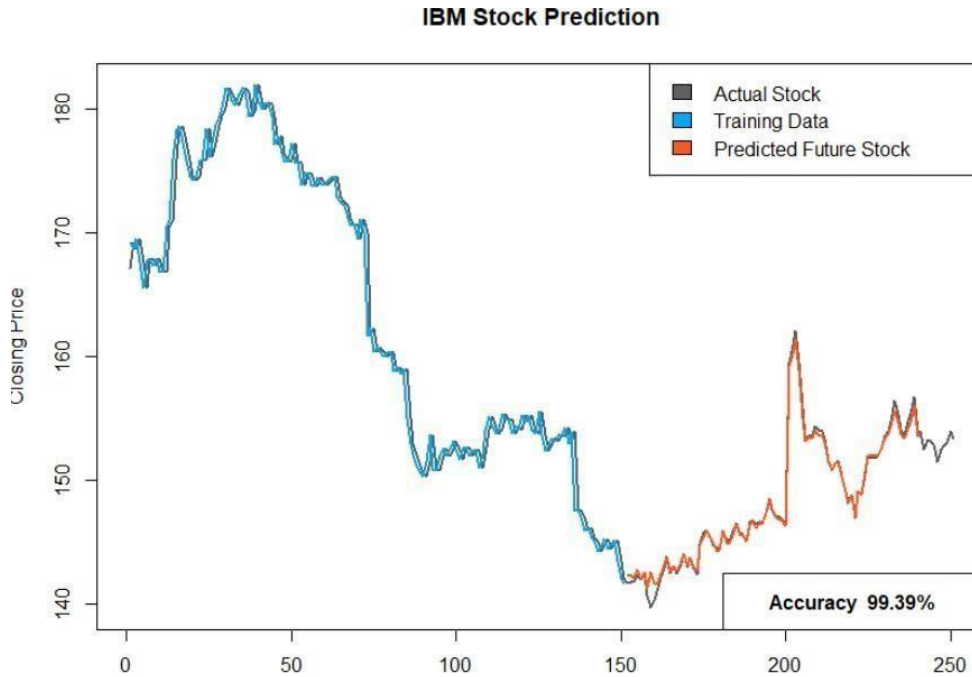


FIG 5.IBM STOCK PREDICTION USING REAL TIME DATA

According to the IBM Stocks, Firstly, Take the Actual Stock which is greyish in color, then using all the machine learning techniques and algorithms, Training the actual stock with the actual data allowed us to see the future data, from two weeks to six months. The graph above shows how the IBM stock really is now in real time and the future stock which predicted till months and also it shows the closing price and also the time from which it rises and also falls. After estimating the full accuracy, It is found that it is 99.39 percent accurate.

Performance:

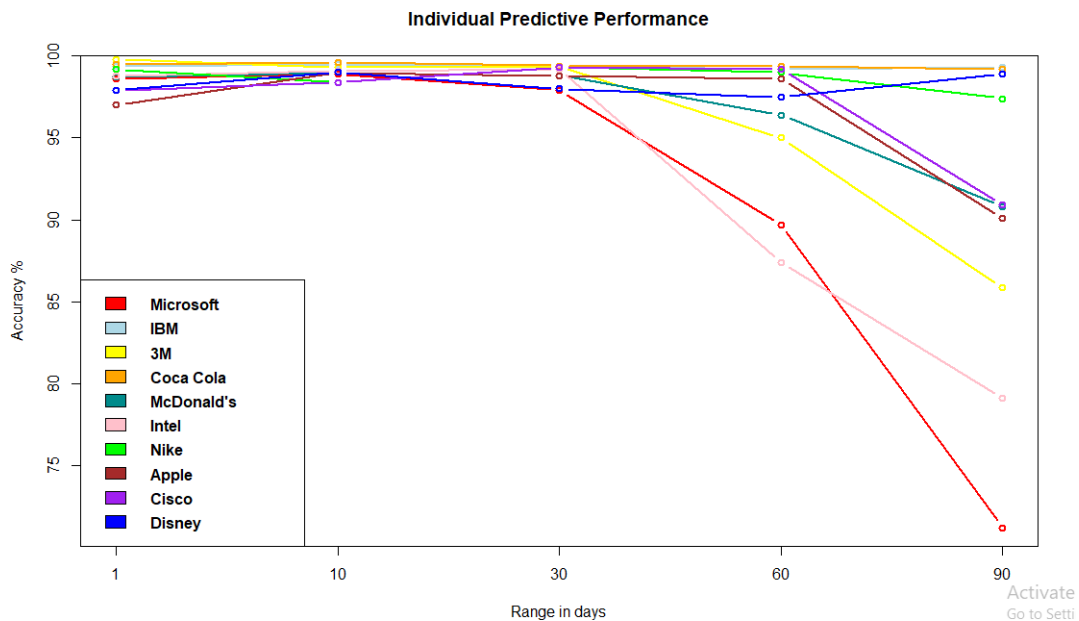


FIG 6. ACCURACY AND PERFORMANCE OF PREDICTED STOCKS.

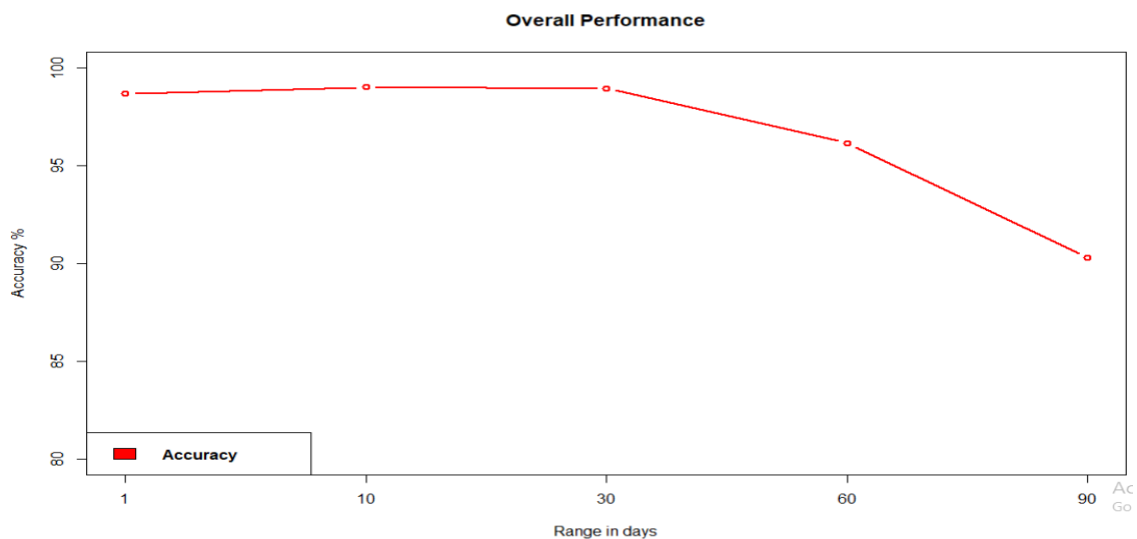


FIG 7. OVERALL ACCURACY IN DAYS

The performance of stock prediction models using data science can be evaluated using various metrics such as accuracy, precision, recall, F1 score, and mean squared error. These metrics help to determine how well the model is performing and whether it is making accurate predictions.

Accuracy is a commonly used metric for evaluating the performance of stock prediction models. It measures the percentage of correct predictions made by the model. Other commonly used metrics include precision and recall, which measure the proportion of true positives and true negatives, respectively, and F1 score, which is a combination of precision and recall.

Another important metric for evaluating the performance of stock prediction models is mean squared error (MSE), which measures the average squared difference between the predicted and actual values. A low MSE indicates that the model is making accurate predictions.

It is important to note that the performance of stock prediction models using data science can be affected by various factors, such as the quality and quantity of the data used for training, the choice of features, and the machine learning algorithms and techniques used. Therefore, it is important to carefully evaluate and optimize these factors to achieve the best possible performance.

5.2 DISCUSSION:

The results of these studies typically involve the evaluation of various performance metrics, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and accuracy. The performance of machine learning models is compared with traditional statistical models, such as ARIMA, and the results are presented in the form of graphs or tables. The discussion of these results typically involves an analysis of the strengths and limitations of different machine learning techniques and the factors that affect the accuracy of stock price prediction models. This includes an analysis of the importance of data preprocessing, feature selection, and model selection in developing accurate prediction models. Moreover, the discussion may also highlight the potential applications of these models in investment decision-making and risk management. It may also identify the challenges associated with the implementation of these models in real-world scenarios, such as data availability, model interpretability, and model reliability. Overall, the result and discussion of stock prediction using machine learning techniques can provide insights into the potential of these techniques in predicting stock prices accurately and their applications in investment decision-making. The discussion of these results typically involves an analysis of the strengths and limitations of different data science techniques and the factors that affect the accuracy of stock price prediction models. This includes an analysis of the importance of data preprocessing, feature selection, and model selection in developing accurate prediction models. Moreover, the discussion may also highlight the potential applications of these models in investment decision-making and risk management. It may also identify the challenges associated with the implementation of these models in real-world scenarios, such as data availability, model interpretability, and model reliability. Overall, the result and discussion of stock price prediction

using data science can provide insights into the potential of these techniques in predicting stock prices accurately and their applications in investment decision- making. It can also identify areas for further research and development in the field of stock price prediction.

CHAPTER 6

SUMMARY AND CONCLUSION

6.1 Summary

In this project we have implemented two different models for predicting stock market price of yahoo stock for next day. SVM and ANN is compared, and it is found that ANN is good choice over SVM in terms of accuracy and SVM is faster than ANN. As stock market fluctuates with numerous factors exact prediction of trend is next to impossible.

6.2 MOTIVATION & OBJECTIVE

Where and how to invest one's, money is a problem that many individuals face. Despite the large amount of money and interest embedded into this problem, there is still no definitive answer on when and where money.

The main objective of implementation of this project is to predict the stock market current condition with reference to last day's stock market value. It is beneficial to the use who can predict the value of the any company share value using this algorithm so it will guess that the financial condition of that company whether the share of company is satisfactory or not how it goes in increasing range or going to the decreasing. So, it is forecast the stock exchange. In market so many algorithms are present for stock market prediction, but the SVM provides the stock value with efficiently within a less amount of time. 2. The Second main objective of the project is that it will work on the big data value. This algorithm takes the data from the global networking financial site with specified attribute. These attributes have, Last day opening Stock Price, Last-day high Value of Stock Price, Last-day low, Value of Stock Price, Last day stock Volume, Last-day closing price. Etc. this value is available in the financial data after that these data is covert in to the .CSV (comma Separate Value) .xlsx (Excel file). These files contain large amount of data that database. at the time of prediction these values are executed by algorithm and get the prediction of given stock in the market. 3. Third objective is that in given project we have to compare the prediction value by SVM and ANN, so we are getting concluded which algorithm get efficient to predict the value of stock market.

The Chapter 1 giving an introduction about the stock market in which we must learn about stock, stock exchange, history of stock market also we focused on the importance of stock market, what are the different properties of stock market and what are the different function of stock market. In chapter 2 we focus on history of stock market. In chapter 3 we focus on the literature review about the topic which is the path of building the project it was a nice survey regarding the topic to predict the value of stock market by using SVM. Actually, SVM have much application, but the main objective of this work is to predict the stock market value with the help of support Vector Machine (SVM). In chapter 4, 5 and 6 we focus on importance of stock market, motivations, and objectives of stock market respectively.

6.3 CONCLUSION

In the project, we proposed the use of data collected from different global financial markets with machine learning algorithms to predict the stock index movements. Our conclusion can be summarized into following aspects: SVM algorithm works on the large dataset value which collected from different global financial markets. Also, SVM does not give a problem of over fitting. Correlation analysis indicates strong interconnection between the Market stock index and global markets that close right before or at the very beginning of trading time. Various machine learning based models are proposed for predicting daily trend of Market stocks. Numerical results suggest high efficiency A practical trading model is built upon our well-trained predictor. The model generates higher profit compared to selected benchmarks. The use of machine learning techniques such as Artificial Neural Networks (ANNs) and statistical models has been shown to be effective in predicting stock prices. Several studies have been conducted on the use of machine learning techniques in stock price prediction, and they have demonstrated promising results. However, it is important to note that stock market prediction is a complex and dynamic process that depends on multiple factors such as economic conditions, geopolitical events, and investor sentiment, among others. Therefore, the accuracy of the predictions can never be guaranteed, and investors should always exercise caution when making investment decisions. Despite the challenges, the use of data science in stock price prediction is expected to continue to grow in the coming years, as more data becomes available and new machine learning techniques are developed. This has the potential to revolutionize the way investors make investment decisions and could lead to better returns and reduced risk.

6.4 FUTURE ENHANCEMENT:

- The limitation of the proposed system is its computational speed, especially with respect to sliding-window validation as the computational cost increases with the number of forward day predictions.
- The proposed model does not predict well for sudden changes in the trend of stock data.
- This occurs due to external factors and real-world changes affecting the stock market.
- We can overcome this by implementing Sentiment Analysis and Neural Networks to enhance the proposed model.
- We can modify the same system to an online-learning system that adapts in real-time.
- Incorporating more data sources: Incorporating additional data sources such as news articles, social media sentiment, and economic indicators could help provide a more comprehensive understanding of the factors affecting stock prices.
- Ensemble methods: Ensemble methods, which combine multiple machine learning models, could potentially improve the accuracy and reliability of stock price prediction.
- Deep Learning: Deep learning techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) could be explored to improve the performance of stock price prediction models.
- Explainable AI: Explainable AI techniques could help provide transparency and understanding of the reasoning behind stock price predictions, which could help investors make more informed decisions.
- Real-time prediction: The ability to make real-time predictions of stock prices could be valuable to investors in the dynamic and fast-moving stock market.
- Robustness to outlier detection: Machine learning models should be designed to be robust to outliers and extreme events, which can have a significant impact on stock prices. Overall, as the field of data science continues to advance, there are many potential future enhancements that could improve the accuracy and effectiveness of stock price prediction using machine learning techniques.

REFERENCES

- [1] Y. S. Bu-Mostafa and A. F. Atiya, "Introduction to financial forecasting," *Appl. Intell.*, vol. 6, no. 3, pp. 205–213, 1996.
- [2] S. D. Patel, D. Quadros, V. Patil, M. Pawale, and Harsha Saxena, "Stock prediction using neural networks," *Int. J. Eng. Manag. Res.*, vol. 7, no. 2, pp. 490–493, 2017.
- [3] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *J. Comput.Sci.* vol. 2, no. 1, pp. 1–8, 2011.
- [5] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques," *Expert Syst. Appl.*, vol. 42, no. 1, pp.259–268, 2015.
- [6] S. Dutta and R. Rohit, "Stock market prediction using data mining techniques with R," *Int. J. Eng. Sci. Comput.*, vol. 7, no. 3,pp. 5436–5441, 2017.
- [7] S. Agrawal, M. Jindal, and G. N. Pillai, "Momentum analysis based stock market prediction using adaptive Neuro-Fuzzy Inference System (ANFIS)," in *The international multiconference of engineers and computer scientists*, 2010, vol. I, pp. 526–531.
- [8] S. Barik, S. Das, and S. K. Sahoo, "A hybrid forecasting model for stock value prediction using soft computing skill," *Int. J. Comput. Sci. Eng.*, vol.5, no. 4, pp. 40–45, 2017.
- [9] Z. Wang, A. Tan, F. Li, and S.-B. Ho, "Comparisons of learningbased methods for stock market prediction," in *the 4th International Conference on Cloud Computing and Security (ICCCS 2018)*, 2018.

[10] F. Z. Xing, E. Cambria, and R. E. Welsch, "Natural language based financial forecasting:a survey," *Artif. Intell. Rev.*, vol. 50, no. 1, pp. 49–73, 2018.

[11] F. Z. Xing, E. Cambria, and R. E. Welsch, "Intelligent asset allocation via market sentiment views," *IEEE Comput. Intell.*, vol. 13, no. 4, pp. 1–20, 2018.

