International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

# STOCK MARKET PRICE PREDICTION USING ANN-DT

S.MARIS SEENIVAS
Department Of Computer Science And Engineering
PSN College Of Engineering And Technology, ,Tirunelveli

Dr.S.Radhakrishnan,
Associate Professor,
Department Of Computer Science And Engineering.
PSN College Of Engineering And Technology, TIrunelveli.

*Abstract*— **The majority of people turn to the performance of a country's stock market as thebest indicator of how well that economy is doing. Stock markets cover all industries across all sectors of the economy. This means they serve as a barometerof what cycle the economy is in and the hopes and fears of the population who generate growth and wealth. In this project, the stock market predictions are doneby using an ANN-DT method. By using ANN-DT we predict the closing price ofa stock at the end of the day. If the closing price is predicted the stock buyers canable to buy a stock based on the predicted price. ANN-DT is a hybrid machine learning technique so that the machine can learn details by utilising both the Artificial Neural Network and decision tree algorithm.**

## I. INTRODUCTION

The majority of people turn to the performance of a country'sstock marketas the best indicator of how well that economy is doing. Stock markets cover allindustries across all sectors of theeconomy. This means they serve as a barometerof what cycle the economy is in and the hopes and fears of the population who generate growth and wealth Stock markets have existed for centuries and will no doubt go on being the mainpublic, regulated marketplaces where people can buy and sell shares of differentcompanies

investing in stocks typically far outweigh those of holdingmoney in lower-return assets like cash.

- Diversification

Trading a variety of stocks can help it spread a risk across different asset classes, economic sectors, and geographicallocations.

- Income

Some stocks provide income as regular dividends, even if the stock has lost value That is income it can keep or reinvest. It can also register votes in company activities

- Control

Of course, today's markets are very different from share trading in the Dutch East India Company back in 1602 but stocks still remain the most popular investment choice thanks to their potential for returns and their opportunity to invest directly in individual companies.

**Essential of stock markets**

Stock markets enable companies to be traded publicly and raise capital The transfer of capital and ownership is traded in a regulated, secure environment. Stock markets promote investment markets promote investment. The raising of capital allows companies to grow their businesses, expand operations and create jobs in the economy. This investment is a key driver for economic trade, growth and prosperity investors, stock markets provide a way to invest money in order to potentially earn a share of the company's profits (knowing that the risk oflosses exists too). Active investors and traders can easily buy and sell theirsecurities due to the abundant liquidity in most major stock markets.

**Stocks trade:**

There are numerous reasons why companies, banks,funds, investors and traders buy and sell company stocks:

- Investment Gains

Stock ownership may help it money grow. Over the long-term,the benefits of

Stocks trade by the millions every day so it can easily trade, buy and sell stocks and shares when it want. This flexibility also means it decide which company toinvest in and when.

Stock market prediction and analysis are some of the most difficult jobs to complete. There are numerous causes for this, including market.

**Problem statement for Stock Market Prediction:**

Let us see the data on which we will be working before we begin implementing the software to anticipate stock market values. In this section, we will examine the stock price of Microsoft Corporation (MSFT) as reported by the National Association of Securities Dealers Automated Quotations

(NASDAQ). The stock price data will be supplied as a Comma Separated File (.csv), that may be opened and analyzed in Excel or a Spreadsheet.

MSFT's stocks are listed on NASDAQ and their value is updated every working day of the stock market. It should be noted that the market does not allow trading on Saturdays and Sundays, therefore there is a gap between the two dates. The Opening Value of the stock, the Highest and Lowest values of that stock on the same days, as well as the Closing Value at the end of the day, are all indicated for each date.

The Adjusted Close Value reflects the stock's value after dividends have been declared (too technical!). Furthermore, the total volume of the stocks in the market is provided, With this information, it is up to the job of a Machine Learning/Data Scientist to look at the data and develop different algorithms that may extract patterns from the historical data of the Microsoft Corporation stock.

However, patterns that allow the prediction of some movements can be found. Stock market analysis deals with the study of these patterns. It uses different techniques and strategies, mostly automatic that trigger buying and selling orders depending on different decision making algorithms. It can be considered as an intelligent treatment of past and present financial data in order to predict the stock market future behavior. Therefore it can be viewed as an artificial intelligence problem in the data mining field. This paper aims to study,construct and evaluate these investment strategies in order to predict future stock exchanges.

**DATA MINING**:

Data mining is the process of sorting through large data sets to identify patterns and relationships that can help solve business problems through data analysis. Data mining techniques and tools enable enterprises to predict future trends and make more-informed business decisions.

Data mining is a key part of data analytics overall and one of the core disciplines in data science, which uses advanced analytics techniques to find useful information in data sets. At a more granular level, data mining is a step inthe knowledge discovery in databases (KDD) process, a data science methodology for gathering, processing and analyzing data. Data mining and KDD are sometimes referred to interchangeably, but they're more commonly seen as distinct things

**Important of data mining:**

Data mining is a crucial component of successful analytics initiatives in organizations. The information it generates can be used in business intelligence (BI) and advanced analytics applications that involve analysis of historical data, as well as real-time analytics applications that examine streaming data as it's created or collected.

Effective data mining aids in various aspects of planning business strategies and managing operations. That includes customer-facing functions such as marketing, advertising, sales and customer support, plus manufacturing, supply chain management, finance.

**Data mining works:**

Effective data mining aids in various aspects of planning business strategies and managing operations. That includes customer-facing functions such as marketing, advertising, sales and customer support, plus manufacturing, supply chain management, finance.

Its core elements include machine learning and statistical analysis, along with data management tasks done to prepare data for analysis. The use of machine learning algorithms and artificial intelligence (AI) tools has automated more of the process and made it easier to mine massive data sets, such as customer databases, transaction records and log files from web servers, mobile apps and sensors.

1. **Data gathering**. Relevant data for an analytics application is identified and assembled. The data may be located in different source systems, a data warehouse or a data lake, an increasingly common repository in big data environments that contain a mix of structured and unstructured data. External data sources may also be used. Wherever the data comes from, a data scientist often moves it to a data lake for the remaining steps in the process.

2. **Data preparation**. This stage includes a set of steps to get the data ready to be mined. It starts with data exploration, profiling and pre-processing, followed by data cleansing work to fix errors and other data quality issues. Data transformation is also done to make data sets consistent, unless a data scientist is looking to analyze unfiltered raw data for a particular application.

3. **Mining the data.** Once the data is prepared, a data scientist chooses the appropriate data mining technique and then implements one or more algorithms to do the mining. In machine learning applications, the algorithms typically must be trained on sample data sets to look for the information being sought before they're run against the full set of data.

**Types of data mining techniques**

Various techniques can be used to mine data for different data science applications. Pattern recognition is a common data mining use case that's enabled by multiple techniques, as is anomaly detection, which aims to identify outlier values in data sets. Popular data mining techniques include the following types:

• Association rule mining. In data mining, association rules are if-then statements that identify relationships between data elements. Support and confidence criteria are used to assess the

relationships -- support measures how frequently the related elements appear in a data set, while confidence reflects the number of times an if-then statement is accurate.

• Classification. This approach assigns the elements in data sets to different categories defined as part of the data mining process. Decision trees, Naive Bayes classifiers, k-nearest neighbor and logistic regression are some examples of classification methods.

• Neural networks. A neural network is a set of algorithms that simulates the activity of the human brain. Neural networks are particularly useful in complex pattern recognition applications involving deep learning, a more advanced offshoot of machine learning.

## II. LITERATURE SURVEY

**Short-term prediction for opening price of stock market basedon self- adapting variant PSO-Elman neural network**

Zhang et al took Elman network to predict the opening price of stock market. Considering that Elman network is limited, this paper adopts self- adapting variant PSO algorithm to optimize the weights and thresholds of network. Afterwards, the optimized data, regarded as initial weight and threshold value, is given to Elman network for training, accordingly the prediction model for opening price of stock market based on self-adapting variant PSO- Elman network is formed. Finally, this paper verifies that model by some stock prices, and compares with BP network and Elman network, so as to draw the result that shows the precision and stability of this predication model both are superior to the traditional neural network.

**Survey of stock market prediction using machine learning approach**

Sharma et al minimized their risks. Prediction plays a very important role in stock market business which is very complicated and challenging process. Employing traditional methods like fundamental and technical analysis may not ensure the reliability of the prediction. To make predictions regression analysis is used mostly. In this paper we survey of well-known efficient regression approach to predict the stock market price from stock market data based. In future the results of multiple regression approach could be improved using more number of variables.

**Combining of random forest estimates using LSboost for stock market index prediction**

Sharma et al explored to combine the predictions/estimates of the ensemble of trees in a Random Forest using LSboost (i.e. LS-RF). The prediction performance of the proposed model is compared with that of well-known Support Vector Regression. Technical indicators are selected as inputs to each of the prediction models. The closing value of the stock price is the predicted variable. Results show that the proposed scheme outperforms Support Vector Regression and can be applied successfully for building predictive models for stock prices prediction.

**Improving Traditional Stock Market Prediction Algorithms using Covid-19 Analysis**

Jindal et al resulted in the poor performance of various traditional trend prediction algorithms because these algorithms do not account for the impact of the pandemic on the stock market trends. The proposed work aims to enhance the stock market prediction ability of various common prediction models by taking into account the factors related to COVID-19. The forecasting techniques analysed are Decision Tree Regressor, Random Forest Regressor and Support Vector Regressor (SVR).

Currently the most affected countries by COVID-19 are the United States of America, India and Russia.

**Stock Market Behaviour Prediction using Stacked LSTM Networks**

analysed an entity such as the stock market that appears to lack synchronicity yet seemingly influenced by historic events remains an open issue in research. This study adopts a stacked Long Short Term Memory network model for predicting stock market behabiour. The data used is composed of historic stockmarket data from the American Stock Exchange, NASDAQ Composite (IXIC). Results obtained show that by making use of a stacked Long Stort Term Memory network model, future stock market behavior can be predicted.

**Efficacy of News Sentiment for Stock Market Prediction**

Kalra et al focused to observe fluctuations in stock prices with respect to the relevant news articles of a company. In this paper, a daily prediction model is proposed using historical data and news articles to predict the Indian stock market movements. Classifier Naïve Bayes is used to categorize the news text having negative or positive sentiment. The count of the positive and negative sentiment of news articles for each day and variance of adjacent days close price along with historical data is used for prediction purpose and an accuracy ranging from 65.30 to 91.2

% achieved with various machine learning techniques. **Prediction of Stock Market Using Artificial Intelligence** Shah et al presented an Artificial intelligence is used to identifythe unknown samples and analyze them. These values are takenas input they are used to predict the stock market. Predictingstock market is a very ambiguous tasks it has many methodswhich make it possible to predict such artificial neural network(ANN), adaptive neuro fuzzy inference (ANFIS), SwarmIntelligence etc. There are future methods like LevenbergMarquardt which is a part of neural which help not only to predictaccurately but also being more efficient than ANFIS in manyways such as time taken, memory allocation, accuracy etc.

**Analysis of various machine learning algorithm and hybrid model for stock market prediction using python**

Vazirani et al performed an extensive comparative analysis of already implemented models for stock trade market. On the basis of results obtained, new linear regression models are proposed which provide much significant error reduction. It was observed that appending two linear regression models where output of first block is fed to the input of second linear regression model gives the most efficient prediction model.

**Literature review on Artificial Neural Networks Techniques Application for Stock Market Prediction and as Decision Support Tools**

Firdaus et al aimed to explore the use Artificial Neural Network (ANN) techniques in the field of stock market prediction. Design: Content analysis research technique. Data sources: Information retrieved from ProQuest electronic databases. Review methods: Utilizing key terms and phrases associated with Artificial Neural Network Stock Market Prediction from 2013-2018. Out of the 129 scholarly journal reviewed, there are 4 stock market studies met the inclusion criteria. The analysis and the evaluation includes 6 ANN derivatives techniques used to predict. Results: Findings from the reviewed studies revealed that all studies shows consistency that the accuracy rate of ANN stock market prediction is high. 2 Studies shows accuracy above 90%, 2 studies shows accuracy above 50%. Conclusion: This study reveals that the ability of ANN shows consistency of an accuracy rate of stock market prediction. Four method in predicting stock market had an accuracy above 95%. The highest accuracy achieved by using Signal Processing/Gaussian Zero-Phase Filter (GZ-Filter) with 98.7% prediction accuracy.

**Stock Market Prediction based on Social Sentiments using Machine Learning**

Mankar et al added to their annual income, nowadays, people have started looking at stock investments as a lucrative option. With expert guidance and intelligent planning, we can almost double our annual revenue through stock returns. That said, stock investment still remains a risky proposition for the uninitiated. Exorbitant wages of the investment experts coupled

with a general ignorance pertaining to the financial matters among the public, deters many from trading in stocks. The fear of losses also acts as a deterrent to many. These facts propelled us to harness the power of machine learning to predict the movement of stocks. Using sentiment analysis on the tweets collected using the Twitter API and also the closing values of various stocks, we seek to build a system that forecasts the stock price movement of various companies. Such a prediction would greatly help a potential stock investor in taking informed decisions which would directly contributeto his profits.

**Stock Market Prediction with Deep Learning Using Public Disclosure Platform Data**

Aslan et al studied stock value prediction has been made for the three sport clubs Galatasaray (GSRAY), Fenerbahçe (FENER) and Beşiktaş (BJKAS) that have stock values in Borsa Istanbul. Sentiment analysis results that have been obtained using the Public Disclosure Platform notification data about GSRAY, FENER and BJKAS have been used in addition to financial data that belong to these companies for stock market prediction. Long-short term memory, which is one of the deep learning methods, has been used for prediction. According to experimental results, when Public Disclosure Platform notification data about these companies have been used in prediction, the prediction results are more accurate.

**Predicting the Effects of News Sentiments on the Stock Market**

Shah et al retrieved, extracted, and analyzed the effects of news sentiments on the stock market. Our main contributions include the development of a sentiment analysis dictionary for the financial sector, the development of a dictionary-based sentiment analysis model, and the evaluation of the model for gauging the effects of news sentiments on stocks for the pharmaceutical market. Using only news sentiments, we achieved a directional accuracy of 70.59% in predicting the trends in short-term stock price movement.

**Developing a Prediction Model for Stock Analysis**

Nivetha et al worked a various prediction algorithms are analyzed to build a prediction model. The prediction model will be based on monthly prediction and daily prediction to forecast the next day market price. This model estimates the open value of the next day in the market. Sentiment Analysis needs to identify and extract sentiments from each individual in the social media. The correlation between the sentiments and the stock value is to be determined. A comparative study of these three algorithms: Multiple Linear Regression, Support Vector Machine and Artificial Neural Network are

done. The stock price is predicted by sentiment analysis with the best forecasting algorithm.

**Analysis of Investor Sentiment and Stock Market Volatility Trend Based on Big Data Strategy**

Peng et al studied the specific mechanism of investor sentiment affecting stock market volatility. With the help of Pollet and Wilson's theory of volatility decomposition, it performs a comparative analysis based on big data strategy and sources. This paper collects the data of web news emotion index, web search volume, social network emotion index, social network heat index, and establishes corresponding analysis index. After correlation analysis and Granger causality tests, it extracts the indicators which have significant correlation with the financial market and brings them into forecasting analysis. The model constructs market volatility index and analyzes the correlation between investor sentiment and stock price changes. In empirical study, the deviation between stock price and value is introduced as an explanatory variable, and the logarithmic return of stock is used to measurethe volatility of stock price.

**Indices prediction of Bangladeshi stock by using time series forecasting and performance analysis**

Majumder et al predicted the indices of Bangladeshi stock by using various stock prediction algorithms such as Feed-Forward Neural Network (FFNN), Auto Regressive Integrated Moving Average (ARIMA), Linear model & Holt- Winter approaches and analyzed the performance of these algorithms over 35 Bangladeshi stocks. Time series analysis is considered to this work performance of algorithms is computed by calculating percentage accurate prediction. From analysis, it is found that, ARIMA (1,0,0) gives maximum prediction accuracy (82.1%) in average among all and FFNN shows best algorithm in forecasting stocks index. FFNN gives maximum accuracy in 14 out of 35 stocks.

**Research on Stock Price Prediction Method Based on Convolutional Neural Network**

Sayavong et al developed a stock price prediction model based on convolution neural network, which has obvious self-adaptability and self- learning ability. Combining the characteristics of CNN (Convolution Neural Network) and Thai stock market, the data set is trained and tested after pretreatment. On this basis, three stocks (BBL, CAPLL&PTT) listed on the Thai Stock Exchange are tested and compared with the actual stock price. The results show that the model based on CNN can effectively identify the changing trend of stock price and predict it which can provide valuable reference for stock price forecast. The prediction accuracy is high, and it is worth further promotion in the financial field.

**Classification Study**

Misra et al performed by mainly two means, one by using previous data available against the stock and the other by analysing the social media information. Predictions based on previous data lack accuracy due to changing patterns in the stock market, some fields might have been missed due to their insignificance in some stocks or unavailability of data. For example, some models may require `return rate' as a parameter for stock prediction, but the available data might not have it. On the other hand, a model predicting onlyon the basis of the return rate may find opening and closing priceto be insignificant parameters. The data has to be cleansed beforeit can be used for predictions. This paper focuses on categorising various methods used for predictive analytics indifferent domains to date, their

shortcomings.

**Optimized Prediction Model for Stock Market TrendAnalysis**

Soni et al added to the academic understanding of stock market analysis using some well defined algorithms and machine learning techniques. Stock price forecasting is a popular and important topic in financial studies and at academic levels. Share Market is not a neat place for

analyzing since there are no significant rules to estimate or predict the price of share in the share market. Many a method like technical analysis, fundamental analysis, time series analysis and statistical analysis, etc. have been used in an attempt to analyze the share trends in the market but none of these methods have so far proved to be a universal approach for acceptance as a prediction tool.

**Analysis of Stock Market using Streaming data Framework**

Umadevi et al tried to design a stock market prediction model which is considers different parameters of a particular stock. Analysis is performed after obtaining the stock scores. This analysis involves visualization of stock scores in the form of various plots and prediction of the scores using a time series model known as ARIMA (auto regressive moving average). Theresults shows that the time series model performed a descent prediction of the market scores with considerably high accuracy.Each factor was studied independently to find out its associationwith market performance. Furthermore the result suggests that behavior of market can be predicted using machine learning techniques.

## III. EXISTING SYSTEM

**Stock Price Prediction Based on Morphological Similarity Clustering and Hierarchical Temporal Memory**

WANG et al presented a clustering method for mining similar stocks, which combines morphological similarity distance (MSD) and kmeans clustering. Subsequently, Hierarchical Temporal Memory (HTM), an online learning model, is used to learn patterns from similar stocks and make predictions at last, denoted as C-HTM. The experiments on the price prediction show that 1) compared with HTM which has not learned similar stock patterns, C- HTM has better prediction accuracy, 2) in terms of short-term prediction, the performance of C-HTM is better than all baseline models.

In this section, we introduce the method used in this paper, including KMSD clustering, HTM, baseline models and evaluation measures for clustering and prediction.

KMEANS WITH MORPHOLOGICAL SIMILARITY DISTANCE

Kmeans [19] is a common clustering algorithm in machine learning tasks. Given n samples of time series (x1, x,
. . . , xn), where each sample is a d- dimensional real vector, kmeans clustering aims to partition the n samples into k(≤ N) sets S = (S1, S2, . . . , Si). In other words, its goal is
to find the cluster Si that satisfies the following formula:

$$\arg\min_{S} \sum_{i=1}^{k} \sum_{x \in S_i} f(x, \mu_i)$$

where μi is the mean value of all series in Si , the function f is the similarity between sample x and μi . We use MSD as the similarity measure, so f can be expressed as follows:

$$f(x, \mu_i) = ED \times (2 - \frac{ASD}{SAD})$$
$$= \sqrt{\sum_{j=1}^{d} (x_j - \mu_{ij})^2} \times (2 - \frac{|\sum_{j=1}^{d} (x - \mu_i)|}{\sum_{j=1}^{d} |x - \mu_i|})$$

where ED is Euclidean distance, ASD is the absolute sum of the difference, and SAD is Manhattan distance.

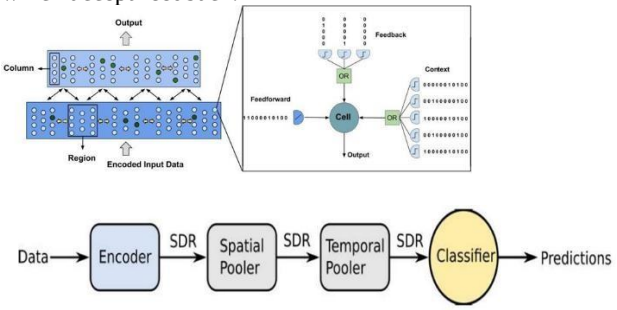
**HIERARCHICAL TEMPORAL MEMORY**

In this section, we only make a brief introduction to HTM. For readers who are not familiar with HTM.

**HTM STRUCTURE**

HTM is an unsupervised machine learning technology that simulates the working principle of the neocortex of the mammalian (mainly human) brain, proposed by originally. The structure of an abstract HTM model with two level is shown in Figure. A typical HTM model is a tree-like hierarchical structure. Each level is composed of smaller elements called regions, while single level in a hierarchical structure may contain multiple regions. Generally, the higher a level is in the model, the less regions it contains. Furthermore, each region iscomposed of columns of multiple neurons.

The design of neuron is like the pyramidal cells in the brain. It is a multipolar neuron that includes three types of dendrites: proximal dendrites, which receive feedforward input information; terminal dendrites, which receive contextual information; apical dendrites, which accept feedback.




**HTM WORKFLOW**

As shown in Figure, the HTM workflow mainly includes four parts: encoding, spatial pooling, temporal pooling and classifier prediction. Firstly, the encoder transforms the input into sparse distributed representations (SDRs), data structure composed of binary, in which the number of 1 is much smaller than 0. SDRs is the basis for the robustness of the model. Then, the spatial pooling algorithm and the temporal pooling algorithm integrate similar patterns and divide time groups, respectively. Finally, the prediction result of the model is given by the classifier.

**SPARSE DISTRIBUTED REPRESENTATION**

As far as nerual network is concerned, sparsity terms has been proven to improve prediction accuracy. As sparse data representation in HTM, SDRs ensures the robustness of

noise and the sensitivity to input. An SDR consists of thousands of binary bits in which 1 represents a relatively active neuron and a 0 represents a relatively inactive neuron. However, the number of 1 is much smaller than 0, which is why we think SDRs is sparse. In general, SDR is considered to be the core concept of HTM.

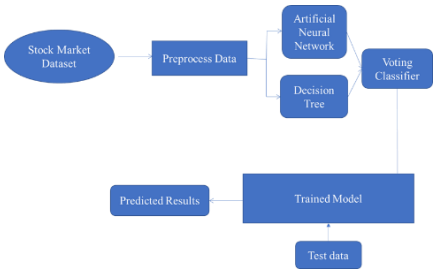
**BASELINE MODELS**

In this model, we use RNN, LSTM and GRU as baseline models. As a class of neural network, RNN is widely used in processing variable-length time series because of its internal memory. In addition, LSTM and GRU are variants of RNN, which avoid the gradient vanishing problem and prove to be more suitable for processing long time series.

**PROPOSED SYSTEM**

The chapter aims is to implement supervised machine learning concept by using datasets regarding stock market collected that are used for detection and prediction. These datasets are send as the input which is trained by artificial neural network algorithm and apply decision tree classification learning algorithm to perform classification which results in recognizing and also possibly predict the disease based on the nature of stock price and classify accordingly. Artificial Neural network algorithm seems to avoid pruning problem and has higher efficiency and accuracy in training the datasets. Also the using of Decision tree is because they are easy to interpret, understand and also possess non-linear characteristics between values. This holds well the performance of the tree constructed which gives better outputs.

**Proposed ANN-DT algorithm**




At first the datas are getted from the input dataset by using pandas library. Then we preprocess the data by droping null values, then we make feature selection by selecting input features for feeding it in the Hybrid MACHINE Learning module, we design a Hybrid ensemble machine learning model by using ANN and Decision tree with voting classifier which can able to give high accuracy, the extracted features are inserted in to the Hybrid model and the machine gets trained. After training we predict the Stock price.

As is shown in Figure, the ANN-DT algorithm generates a univariate decision tree from a trained neural network. A given decision node in the tree returns “true” if the node’s single specified variable exceeds a certain

threshold and “false” otherwise. For each node the algorithm decides on which variable to partition the set of data, after which the threshold of that variable has to be determined. More specifically, the univariate decision tree is constructed by use of the algorithm to examine the responses of the neural network in the feature space and to conduct a sensitivity or significance analysis of the different attributes or explanatory variables pertaining to these responses.

**Training of the Artificial Neural Networks**

The ANN-DT algorithm is depicted schematically in Fig. 2. In the first stage of the procedure, a neural-network model capable of generalizing underlying trends in the data has to be constructed. In this particular investigation both multilayer perceptron and radial basis function neural networks were used. The former was trained by a gradient descent method, in which the error was propagated backward through the network . Training the radial basis function network entails finding a suitable set of basis nodes and weights to the output layer. Since the connecting weights of the output layer of the radial basis function neural network can be found by solution of a least squares problem, the most difficult task is finding a good set of basis functions.

**Induction of Rules from Sampled Points in the Feature Space**

A systematic search through the input or feature space, as is accomplished by the methods of Fu, Gallant, Thrun and Saito and Nakano, can lead to an overwhelming number of rules for many real-world problems. Even when the feature space is properly bounded, decision boundaries generated by the neural network which are not parallel to the boundaries of the feature space can lead to an intractably large number of rules required to represent the behavior of the network to some arbitrary degree of accuracy. For example, Thrun has investigated the movement of a robotic arm that has required more than 8000 rules to describe. Consequently restrictions to the depths of these kinds of searches have to be imposed. Even when the depths of these searches are restricted, the number of rules can still be large, as was suggested by a study of Saito and Nakano[14].

The ANN-DT algorithm uses sampling, which can similarly lead to numerous rules depending on the number of sample points that are generated. However, it is necessary to ensure that the sampled points are restricted to those regions of the (often high- dimensional) search space on which the neural- network model is based. This is accomplished by taking the joint distributions of points in the feature space into account, not only during sampling of the network, but also when the resulting branches of the trees require pruning.

**SOFTWARE SPECIFICATION**

**SOFTWARE REQUIRED:**

- IDLE 1.7
- PYTHON1.7.6

**HARDWARE REQUIRED:**

- System : Windows Xp Professional Service Pack 2
- Processor :Up to 1.5 GHz
- Memory :Up to 512 MB RAM

**PYTHON**

The Python language had a humble beginning in the late 1980s when a Dutchman Guido Von Rossum started working

on a fun project, which would be a successor to ABC language with better exception handling and capability to interface with OS Amoeba at Centrum Wiskunde and Informatica. It first appeared in 1991. Python 2.0 was released in the year 2000 and Python 3.0 was released in the year 2008. The language was named Python after the famous British television comedy show Monty Python's Flying Circus, which was one of Guido's favorite television programmes. Here we will see why Python has

**Why Python?**

Now you might be suddenly bogged with the question, why Python? According to Institute of Electrical and Electronics Engineers (IEEE) 2016 ranking Python ranked third after C and Java. As per Indeed.com's data of 2016, the Python job market search ranked fifth. Clearly, all the data points to the ever rising demand in the job market for Python. It's a cool language if you want to learn just for fun or if you want to build your career around Python, you will adore the language. At school level, many schools have started including Python programming for kids. With new technologies taking the market by surprise Python has been playing a dominant role. Whether it is cloud platform, mobile app development, Big Data, IoT with Raspberry Pi, or the new Block chain technology, Python is being seen as a niche language platform to develop anddeliver a scalable and robust applications.

Some key features of the language are:

- Python programs can run on any platform, you can carry code created in Windows machine and run it on Macor Linux
- Python has inbuilt large library with prebuilt and portable functionality, also known as the standard library
- Python is an expressive language
- Python is free and open source
- Python code is about one third of the size of equivalent C++ and Java code
- Python can be both dynamically and strongly typed--dynamically typed means it is a type of variable

that is interpreted at runtime, which means, in Python,there is no need to define the type (int or float) of the variable

**Python applications**

One of the most famous platforms where Python is extensively used is YouTube. The other places where you will find Python being extensively used are the special effects in Hollywood movies, drug evolution and discovery, traffic control systems, ERP systems, cloud hosting, e-commerce platform, CRM systems, and whatever field you can think of.

In order to effectively classify the emotions, This paper [19] make use of SVM classification algorithm for both DEAP and SEED dataset. For DEAP dataset events are separated before the data is pre-processed. Then statistical, frequency and other features are extracted. Then the obtained data is given to the classifier. Both the datasets provided best results.

**Versions**

At the time of writing this book, two main versions of the Python programming language were available in the market, which are Python 2.x and Python 3.x. The stable release as of writing the book were Python 2.7.13 and Python 3.6.0.

Implementations of Python

Major implementations include CPython, Jython, IronPython, MicroPython, and PyPy.

**Installation**

Here we will look forward to the installation of Python on three different OS platforms, namely, Windows, Linux,and Mac OS. Let's begin with the Windows platform.

## IV. SYSTEM IMPLEMENTATION

**USECASE**

In software and systems engineering, a use case is a list of actions or event steps, typically defining the interactions between a role (known in the Unified Modeling Language as an actor) and a system, to achieve a goal. The actor can be a human or other external system. In systems engineering, use cases are used at a higher level than within software engineering, often representing missions or stakeholder goals. The detailed requirements may then be captured in the Systems Modeling Language (SysML) or as contractual statements. Use case analysis is an important and valuable requirement analysis technique that has been widely used in modern software engineering since its formal introduction by Ivar Jacobson in 1992. Use case driven development is a key characteristic of many process models and frameworks such as ICONIX, the Unified Process (UP), the IBM Rational Unified Process (RUP), and the Oracle Unified Method (OUM). With its inherent iterative, incremental and evolutionary nature, use case also fits well for agile development.

**SEQUENCE DIAGRAM**

A sequence diagram is an interaction diagram that shows how objects operate with one another and in what order. It is a construct of a message sequence chart. A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios. A sequence diagram shows, as parallel vertical lines, different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them, in the order in which they occur. This allows the specification of simple runtime scenarios in a graphical manner.

**ACTIVITY DIAGRAM**

Activity diagram is another important diagram in UML to describe dynamic aspects of the system. Activity diagram is basically a flow chart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. So the control flow is drawn

from one operation to another.

**COLLABORATIVE DIAGRAM**

A collaboration diagram resembles a flowchart that portrays the roles, functionality and behavior of individual objects as well as the overall operation of the system in real time. Objects are shown as rectangles with naming labels inside. These labels are preceded by colons and may be underlined. The relationships between the objects are shown as lines connecting the rectangles. The messages between objects are shown as arrows connecting the relevant rectangles along with labels that define the message sequencing.

**CLASS DIAGRAM**

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure

diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among objects.

## CONCLUSION

In this project, the stock market predictions are done by using a ANN-DT method. This method has been developed to extract decision trees from trained feedforward neural networks, regardless of the structures of these networks. It was found that in some cases these rules were significantly more representative of the behavior of the neural network than rules extracted from the training data only. Alternatively, the algorithm can be used as a method to extract rules from data sets. These rules appear to be of similar accuracy as those obtained with CART. In fact, in some cases a significant improvement could be obtained with the ANN-DT algorithm. In one particular case it was demonstrated that the significance analysis of the ANN- DT(s) could correctly identify the most important attributes and build valid sets of rules. In contrast to this, a greedy error driven procedure, such as used in CART and ANN-DT(e), failed to identify the most important attributes. As a result, rules derived with CART and ANN- DT(e) were comparatively inaccurate, while the ANN- DT(e) algorithm could only find more accurate rules by using many more sample points than ANN- DT(s).

## REFERENCES

K. S. Umadevi, A. Gaonka, R. Kulkarni and R. J. Kannan, "Analysis of Stock Market using Streaming data Framework," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018, pp. 1388-1390, doi: 10.1109/ICACCI.2018.8554561.

D. Soni, S. Agarwal, T. Agarwel, P. Arora and K. Gupta, "Optimised Prediction Model for Stock Market Trend Analysis," 2018 Eleventh International Conference on Contemporary Computing (IC3), 2018, pp. 1-3, doi: 10.1109/IC3.2018.8530457.

T. Tantisripreecha and N. Soonthomphisaj, "Stock Market Movement Prediction using LDA-Online Learning Model," 2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2018, pp. 135-139, doi: 10.1109/SNPD.2018.8441038.

M. Misra, A. P. Yadav and H. Kaur, "Stock Market Prediction using Machine Learning Algorithms: A Classification Study," 2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE), 2018, pp. 2475-2478, doi: 10.1109/ICRIEECE44171.2018.9009178.

L. Sayavong, Z. Wu and S. Chalita, "Research on Stock Price Prediction Method Based on Convolutional Neural Network," 2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS), 2019, pp. 173-176, doi: 10.1109/ICVRIS.2019.00050.

M. M. R. Majumder, M. I. Hossain and M. K. Hasan, "Indices prediction of Bangladeshi stock by using time series forecasting and performance analysis, 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1-5, doi: 10.1109/ECACE.2019.8679480.

D. Peng, "Analysis of Investor Sentiment and Stock Market Volatility Trend Based on Big Data Strategy," 2019 International Conference on Robots & Intelligent System (ICRIS), 2019, pp. 269-272, doi: 10.1109/ICRIS.2019.00077.

R. Y. Nivetha and C. Dhaya, "Developing a Prediction Model for Stock Analysis," 2017 International Conference on Technical Advancements in Computers and Communications (ICTACC), 2017, pp. 1-3, doi: 10.1109/ICTACC.2017.11.

D. Shah, H. Isah and F. Zulkernine, "Predicting the Effects of News Sentiments on the Stock Market," 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 4705-4708, doi: 10.1109/BigData.2018.8621884.

B. Aslan and R. C. Erdur, "Stock Market Prediction with Deep Learning Using Public Disclosure Platform Data," 2020 Innovations in Intelligent Systems and Applications Conference (ASYU), 2020, pp. 1-5, doi: 10.1109/ASYU50717.2020.9259836.

T. Mankar, T. Hotchandani, M. Madhwani, A. Chidrawar and C. S. Lifna, "Stock Market Prediction based on Social Sentiments using Machine Learning," 2018 International Conference on Smart City and Emerging Technology (ICSCET), 2018, pp. 1-3, doi: 10.1109/ICSCET.2018.8537242.

M. Firdaus, S. E. Pratiwi, D. Kowanda and A. Kowanda, "Literature review on Artificial Neural Networks Techniques Application for Stock Market Prediction and as Decision Support Tools," 2018 Third International Conference on Informatics and Computing (ICIC), 2018, pp. 1-4, doi: 10.1109/IAC.2018.8780437