A Project Review-2 Report

on

**STOCK PREDICTION AND
STOCK RECOMMENDATION
USING ML**

*Submitted in partial fulfillment of the
requirement for the award of the degree of*

B.Tech CSE

GALGOTIAS UNIVERSITY

(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

Under The Supervision of
Rajkamal Kishor Gupta
Asst. Professor

Submitted By

KESHAV GAUR (19SCSE1010660)
ALOK KUMAR YADAV (19SCSE1010677)

**SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND
ENGINEERING GALGOTIAS UNIVERSITY, GREATER
NOIDA
INDIA
DECEMBER,
2021**

ABSTRACT

Stock market is one of the most important activities going on in today's financial world. It has been a source of income for many from a long time (around 200 years) and due to the increasing demands for stocks it will continue to do so. Now with benefits, volatility also steps in. Most favorable intensions of many financial analysts is predicting the rise and fall of the stock market to be precise but there is not much success till now. With so much improvement in technology, processing power, better algorithms we can make it more possible to do so. The new solution based on the machine learning approach includes the stocks recommendation based on the investment and also the risk associated and expected value increment or decrement with respect to his or her own investment, it will use the previous stock dataset for learning and understanding, will use tweets and fresh news for sentiment analysis and risk management according to the current scenario to overcome the existing barrier and making the best predictions of stock market available for all.

.

Keywords – stocks market; machine learning; predictions; support vector machine; feature engineering; data mining.

LIST OF FIGURES:

TABLE OF CONTENTS

CHAPTER - 1

INTRODUCTON

1.1 INTRODUCTION

The stock market is one of the major areas in which investors are committed, so stock market pricing remains a hot topic for researchers from both financial and technical backgrounds. In this study, our aim is to build a high-quality speculation model for price forecasting. The stock market fluctuates with the share of participants profits. When market prices increase with availability stock then participants benefit from their purchased shares. Alternatively, if the market decreases with availability stock prices where participants have to face losses.

Basically, most traders with a lot of money from stock markets buy stocks derivatives, equities and commodities at a cheap price and later selling them at a higher price. When we talk about stock prediction and their short-term recommendation, the trend is not a new thing and this issue is kept discussed by many organizations. Almost every investor performs two types of analysis before selecting a particular stock, first is the fundamental analysis, in this investor look at the foundation, performance of industry, economy, intrinsic value of stocks. On the other hand, technical analysis is an evolution of stock by studying the statistics generated by market activity, such as previous trends, volumes.

Stock market price prediction for short time appears to be random but over a long period of time the stock price movement develops a linear curve. The uncertainty in the stock market refrain people from investing in stocks. Thus, there is a need to accurately predict the stock market which can be implemented in real life.

The prediction target for the probable stock market can be volatility of the prices, future prices or say the market trend. There can be two types in this prediction like dummy and real time prediction which is used in stock market prediction system. In dummy prediction we can predict the share prices by calculating the average price and in real time prediction

we see current prices of shares using internet.

Over the past few years, trading stocks has become a center of attraction for technological advances. Investors look for efficient tools and techniques that can reduce the risk and hence increase their profit. Stock Market Prediction is not an easy task due to its unreliable, dynamic and non–linear nature.

The advancement in computation led to introduction of machine learning for making the predictive system in financial market. In this paper we are using machine learning i.e., we will use python for programming and support vector machine (SVM) to predict the stock market.

1.2 Formulation of problem

Everyone want to be rich in his life with great advantages and less efforts. Hearing all the big profits investors taking on a daily basis driving more and more people to invest in this financial world. Investors are familiar with the saying, 'buy low and sell high" but this does not provide enough context to make the right move. To make proper investment decision, an investor needs to be aware how the market behaves before buying a stock.

It is very important to select right stock at a right time, failing in this can have disastrous results. Investors today are facing a really big problem as they do not properly understand which stock to buy or sell to get maximum profit. Long term prediction can be comparatively easy than prediction on day-to-day basis as we experience more volatility in short periods.

Stock market prediction is generally done using quarterly financial ratio using the dataset but a single dataset is not much reliable and less sufficient for the prediction and can give inaccurate results.

Predicting how a stock will perform is quite difficult and estimation will remain a problem if a better algorithm is not provided. For accurate

prediction, we need a lot of factors to keep in mind like the sentiments of thousands of investors, recent news based on a particular stock, their events, upgrades happened or going to happen in a short term, a statement by a political leader. It can also be an international event like big movements in currencies or commodities etc. All these events affect the market and it is beyond the scope of almost all investors to correctly and consistently predict these parameters. Once we collect the right data, then it can be used to train the machine more effectively and to generate a predictive result.

To tackle this problem the best and the most effective method could be using supervised machine learning as this field has already proven its capability in predicting markets with usage of proper algorithms as well as using proper training and testing dataset

1.2.1 Tools and technologies

Tools Used:

Hardware Requirements:
- Processor: Pentium or later processors (2.4 GHz).
- Hard Disk: 10 GB.
- Ram: 4GB or more
- Monitor
- Keyboard & Mouse

Software Requirements:
- Operating system: Windows 7 or later versions .
- Coding Language: Python
- IDE: Python ide (Google Collaboratory).

Technologies Used:

- Supervised ML
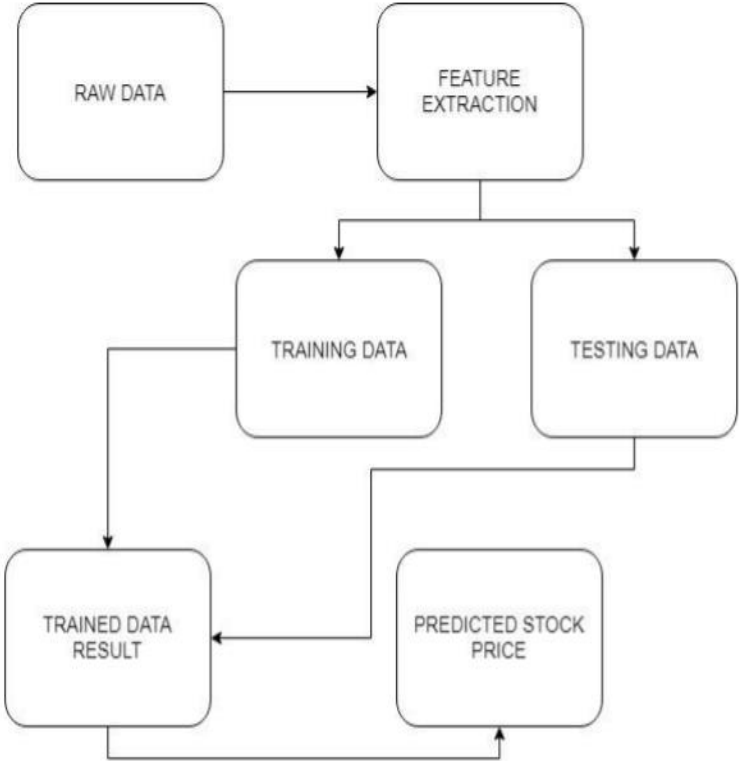- Natural Language Processing

1.2.2   Workflow Diagram

Fig 1 Workflow

1.3 Classical approaches for SMP

There exist two main traditional approaches for the analysis of the stock market : (1) Fundamental Analysis and (2) Technical Analysis.

1.3.1 Fundamental Analysis

Fundamental Analysis calculates a genuine value of a company and arbitrate the value that one share of that company should cost. A presumption is made that, if sufficient time is given to a company, the company will move to a cost agreeing with the prediction made or very close to it. This analysis is based on the fact that if the company/sector is overvalued, then its market price should fall and conversely if a company/sector is undervalued, then the market value should rise with time. This analysis is performed considering various factors of that company/sector, such as a future prospectus, balance sheets, reports and company's work environment. Fundamental analysis can be used to find mostly the undervalued stocks to perform really well in the coming future and also can be used for the deliberation of financial ratios to extricate poor stocks from quality stocks. The two common metrics used to predict price movements yearly for fundamental analysis are (a) price by book ratio (P/B) and (b) price by Earning ratio (P/E). The companies with a lower P/E ratio yield higher returns than the companies with a high P/E ratio. The P/B ratio compares the value of a company specified by the market. If P/B ratio is high, the company might be overvalued and it can fall with time. Fundamental analysis, firstly, lacks sufficient knowledge of the rules governing the working of system and secondly, there is non-linearity in system.

1.3.2 Technical Analysis

Technical analysis is basically the study of price action in stock, by following the previous patterns, charts and the future movement based on these factors to make profit, or to make better decisions on your investments. Technical analysis anticipate the direction of the future

price movements of stocks based on their historical data, and helps to analyze financial time series data using technical indicators to forecast stock prices. Meanwhile, it is assumed that the price moves in a trend and has momentum . Technical analysis uses price charts and certain formulae, and studies patterns to predict future stock prices; it is mainly used by short-term investors. The price would be considered high, low or open, or the closing price of the stock, where the time points would be daily, weekly, monthly, or yearly. Dow theory puts forward the main principles for technical analysis, which are that the market price discounts everything, prices move in trends, and historic trends usually repeat the same patterns . There are several technical indicators, such as the Moving Average (MA), Moving Average Convergence/Divergence (MACD), the Aroon indicator, and the money flow index, etc. The evident flaws of technical analysis as per are that expert's opinions define rules in technical analysis, which are fixed and are reluctant to change. Various parameters that affect stock prices are ignored.

The prerequisite is to overcome the deficiencies of fundamental and technical analysis, and the evident advancement in the modelling techniques has motivated various researchers to study new methods for stock price prediction.

1.4 Modern approaches for SMP

There are some modern approaches that can be functional and fruitful for SMP that would enhance prediction accuracies. In this review, we will highlight some modern functional approaches.

1.4.1 Machine Learning approach

Because of global digitization, SMP has entered a technological era. Machine learning in stock price prediction is used to discover patterns in data [22]. Usually, a tremendous amount of structured and unstructured heterogeneous data is generated from stock markets. Using machine learning algorithms, it is possible to quickly analyze more complex heterogeneous data and generate more accurate results. Various machine learning methods have been used for SMP [23]. The

machine learning approaches are mainly categorized into supervised and unsupervised approaches. In the supervised learning approach, named input data and the desired output are given to the learning algorithms. Meanwhile, in the unsupervised learning approach, unlabeled input data is provided to the learning algorithm, and the algorithm identifies the patterns and generates the output accordingly. Furthermore, different algorithmic approaches have been used in SMP, such as the Support Vector Machine (SVM), k Nearest Neighbors (kNN), Artificial Neural Networks (ANN), Decision Trees, Fuzzy Time-Series, and Evolutionary Algorithms. The SVM is a supervised machine learning technique that limits error and augments geometric margins, and is a pattern classification algorithm [24]. In terms of accuracy, the SVM is an important machine learning algorithm compared to the other classifiers [25]. In the kNN, stock prediction is mapped into a classification based on closeness. Using Euclidean distance, the kNN classifies the "k" nearest neighbors in the training set. The ANN is a nonlinear computational structure for various machine learning algorithms to analyze and process complex input data together. The FIS (Fuzzy Inference Systems) apply rules to fuzzy sets and then apply de-fuzzification to give crisp outputs for decision making [26]. The evolutionary algorithms include gene-inspired neuro-fuzzy and neuro-genetic algorithms, mimic the natural selection theory of species, and can give an optimal output.

1.4.2. Sentiment Analysis Approach

One of the phenomena of current times that is changing the world is the global availability of the internet. The most-used platforms on the internet are social media. It is estimated that social media users all over the world will number around 3.07 billion. There is a high association between stock prices and events related to stocks on the web. The event information is extracted from the internet to predict stock prices; such an approach is known as event-driven stock prediction. Through social networks, people generate tremendous amounts of data that is filled with emotions. Much of this data is related to user perceptions and concerns. Sentiment analysis is a field

of study that deals with the people's concerns, beliefs, emotions, perceptions, and sentiments towards some entity. It is the process of analyzing text corpora, e.g., news feeds or stock market specific tweets, for stock trend prediction. The Stock Twits, Twitter, Yahoo Finance, and so on are well-known platforms used for the extraction of sentiments. There is a significant importance of using sentimental data for enhancing the prediction of volatility in the stock market. The 'Wisdom of Crowds' and sentiment analysis generate more insights that can be used to increase the performance in various fields, such as box office sales, election outcomes, SMP, and so on. This suggests that a good decision can be made by taking the opinions and insights of large groups of people with varied types of information. The information generated through social media allows us to explore vast and diverse opinions. Exploring sentiments from social media in addition to numeric time-series stock data would enhance the accuracy of the prediction. Using time-series data as well as social media data would intensify the prediction accuracy. Different approaches and techniques have been proposed over time to anticipate stock prices through numerous methodologies, thanks to the dynamic and challenging panorama of stock markets.

2. Research Methodology

This section explains the overall process of the literature collection on SMP using machine learning. Initially, the phrase "stock market prediction using machine learning" was keyed to various search engines, digital libraries and databases, including 'google scholar', 'research gate', 'ACM digital library', 'IEEE Explore', 'Scopus', and so on. During the process of literature collection, various phrases like "stock market prediction methods", "impact of sentiments on stock market prediction", and "machine learning-based approach for stock market prediction" were keyed. The OR and AND operators were used for the keyword searches in single and multiple classes, respectively. As a result, some of the fundamental papers in the field of stock market prediction were retrieved. By the careful analysis of a few basic papers, a primary insight into the domain was obtained.

The search criteria were further modified to collect the literature of the last decade, in order to enhance and improve the domain. In addition, the literature selected was screened by applying quality criteria, where metrics such as indexing, quartiles, impact factors and publishers were observed. Figure 1 presents the steps followed in the literature collection.
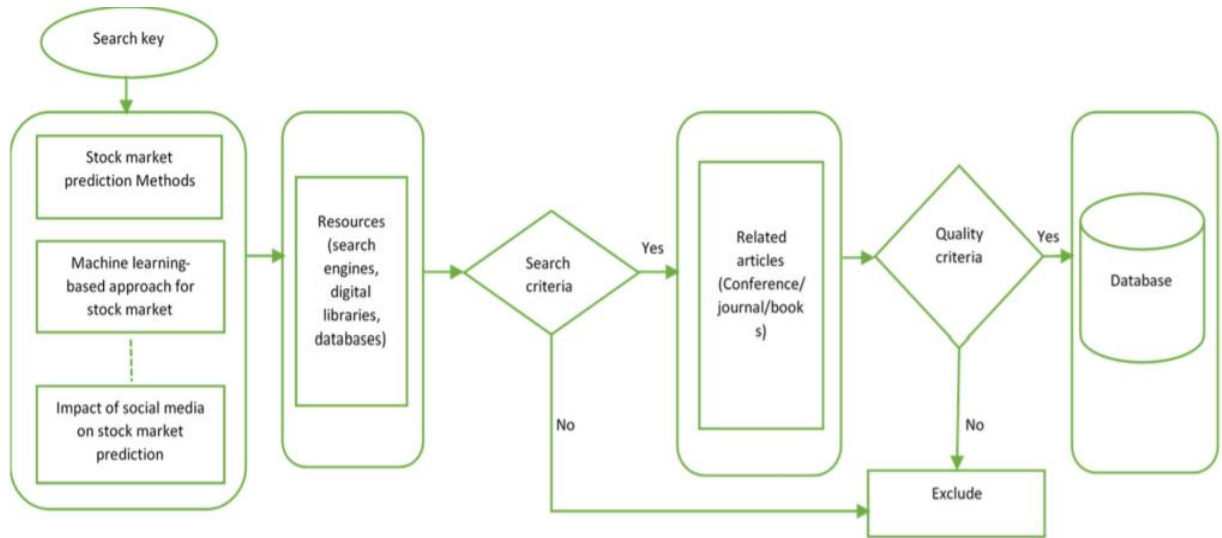
Figure 2. literature collection process

3. **Generic Scheme for SMP**

Figure 2 describes the generic process involved in SMP. The process starts with the collection of the data, and then pre-processing that data so that it can be fed to a machine learning model. The prediction models generally use two types of data: market and textual data. The literature of both types is discussed in the following section. The next section classifies the previous studies based on the type of data used. Furthermore, the next section surveys the previous studies based on the various data-preprocessing approaches applied. Moreover, the literature is further surveyed based on the machine learning algorithms used by different systems.

Figure 3. Generic Scheme for SMP(Stock Market Prediction)

CHAPTER - 2
LITERATURE SURVEY

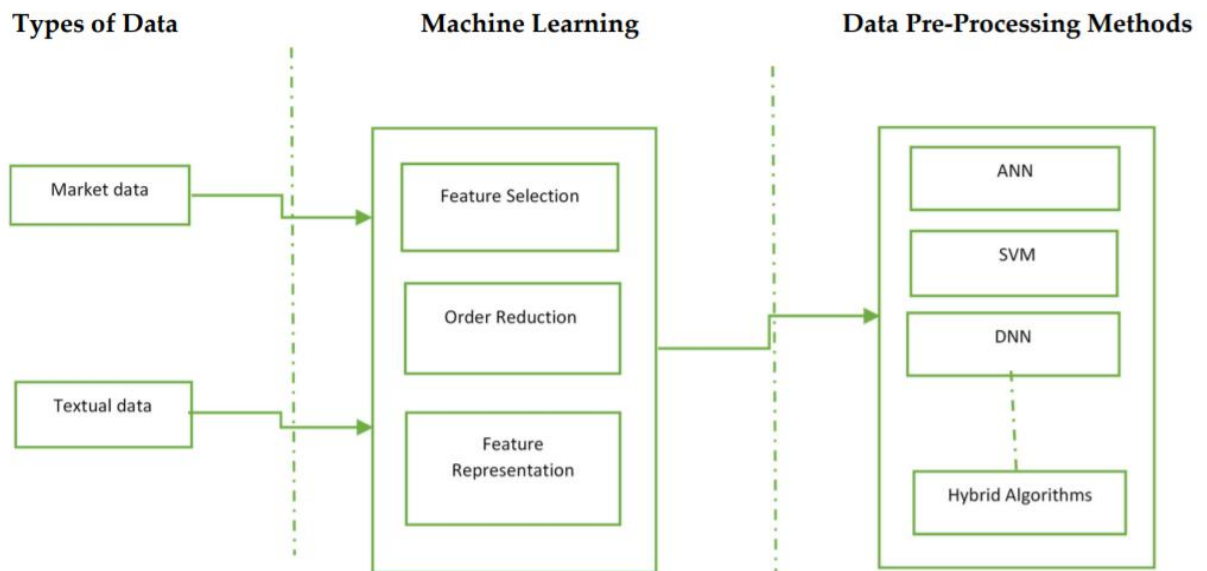**[1]** 2 July 2019, International Journal of Recent Technology and Engineering (IJRTE). They inspected the efficiency of applying technical analysis to the stock prices. We analyzed whether investors could manage more profits than suggested by the recent research of Pang X. et al. (2018).

In future the research work can be extended by applying more STIs (Stock Technical Indicators) and can be evaluated against several ML and deep learning approaches. The proposed model can be further evaluated and optimized for stock indices. The proposed deep learning algorithm can also be further enhanced to optimize the performance.

[2] January 2020, International Conference on Computational Intelligence and Data Science (ICCIDS). The comparative analysis indicates, that various companies, ANN (Artificial Neural Network) proves to be a better technique, giving better RMSE (Root Mean Square Error) and MAPE (Mean Absolute Percentage Error) values.

For future work, deep learning models could be developed which consider financial news articles along with financial parameters such as a closing price, traded volume, profit and loss statements etc., for possibly better results.

[3] April 2019, International Journal of Engineering and Advanced Technology (IJEAT). Other data like previous year stock prices are also considered. The relationship between various data points is considered, and a prediction is made on these data points.

Future scope of this project will involve adding more parameters and factors like the financial ratios, multiple instances, etc. The more the parameters are taken into account more will be the accuracy.

[4] May 2021, Social Science Research Network (SSRN). By estimating the parameters of accuracy, it was found that the most appropriate method for stock market prediction is to use multiple algorithms such as ANN, Random Forest, SVM, LSTM and combine their results for predictions, and

also considering all the factors affecting the stock market price such as news sentiment, currency, commodity prices and other international stock exchange market data.

Future work contains providing predictions and recommendations for individual according to its risk and reward appetite, and also providing a portfolio management to mitigate the risk associated with investing in stock market. Process of placing orders and making trades can be automated which work according to suggestions given by AI and ML algorithms, which can remove human errors and factors such as fear and greed of human being which can result in faster and more profitable trading.

[5] The stock market is highly volatile and complex in nature. Technical analysts in general use Technical Analysis (TA) for historical pricing data, which is also a complete work may produce incorrect predictions. Machine learning is combined with basic and / or Technical Analysis also yields satisfactory results for the stock market forecast. In this make an effort to predict pricing and stock price trends using fairness Short-Term Memory (O-LSTM) in depth learning and stock market technology Indicators (STIs). We also examined the buy-sell decision model at the end of day. To improve in-depth learning functionality, we have used the concept of Correlation-Tensor built-in have appropriate STIs. An accurate measurement of the predictions achieved using the proposed model by 59.25%, above the number of shares, which is much higher than the comparative methods.

[6] By using the introduction of practical skills and increased numeracy skills, planned forecasting methods have been shown to be very effective in stock forecasting prices. In this work, the Artificial Neural Network and random strategies Forestry have been used to predict closing rates the next day in five different companies

functional areas. Financial data: Open, high, low and closed stock prices used to create new variables used as model inputs. Such models tested using standard strategic indicators: RMSE (Root Mean Square Error) and MAPE (Means the Total Percentage Error). The low values of these two indicators indicate that models that work well in predicting stock closing prices.

[7] By considering the various strategies and variables to be considered, we found that techniques such as random forest, vector support mechanism were not used fully. In this paper we will introduce and review the most probable method of forecasting stock movement with high accuracy. The first thing we considered a stock market price index from last year. The database is pre-processed and prepared for actual analysis. Therefore, our paper will also focus on data processing for green database. Second, after processing the data, we will review the random usage in the forest, vector support database and output results. In addition, the proposed paper evaluates the use of the forecast system in real-world settings as well problems related to the accuracy of the total values provided. The paper also highlights a machine learning model to predict long-term stock market competitiveness. The stock market is a place where people buy and sell shares of publicly listed companies. Every buyer and seller try to predict the price of the stock market in order to maximize profits profit and small loss. Using state-of-the-art technology such as AI can improve stock price forecast. In the process of considering the strategies and flexibility that should be in hindsight, we found ML algorithmics such as random forest, LSTM, SVM, ANN fully utilized. In this model we will introduce and review further the possible forecast method stock movement with high accuracy. The first thing we considered in the previous data annual stock market prices, financial history prices and the stock market as well historical news articles. The databases are pre-processed and configured for real analysis.

Existing Survey Limit

1. The existing system fails when there are unusual results or predictions, such as an algorithm based on a sample bootstrap.

2. Previous results show that stock prices are not predictable if they are traditional a classifier is used.

3. The presence system reported the most predictable values, with appropriate selections their testing time to get the most speculative points.

4. The existing system does not work properly if there is a change in performance nature.

5. It does not focus on external events in the area, such as news or social events the media.

6. It uses only one data source, thus being highly biased.

7. The existing system requires some form of interpretation, thus requiring scale.

8. Does not use pre-data processing techniques to remove inconsistencies once data imperfections.

PROPOSED SYSTEM

Introduction

The program is divided into the following modules:

• Data Collection - Data is collected from a variety of sources, such as yahoo finance and google finance, in the form of .csv format. Because The news data collection system uses the NSE India api to obtain real-time data.

• Data Transformation and Visual Analysis - Collected data is purified and processed in order to be suitable for use with incorrect algorithms and models. and this data visualization is also performed.

• Build Model - Data refined and processed for construction, construction and train various machine learning algorithms that can be used in predictions.

• Guess the Results - After the model has been successfully constructed, the next thing to make predicting the effect pattern of a particular stock as well check the accuracy of the predictions

• Predicting to integrate the results of all algorithms - After the construction of the models the system integrates its own results for better prediction and higher accuracy and predict output based on real-time data provided in it.

In this proposed program, we aim to predict future prices using various instruments learning methods. In the proposed program, we were able to train and evaluate ML algorithms from different data points from the past to create future sharing value ratio. We have reviewed EOD (End OF the Day) data for the past years to update train ML models. We have used

various ML libraries and frameworks to achieve the goal The system uses Numpy, Pandas, data and visual effects, Sklearn, Kears, Tensorflow, for machine learning models and mathematical performance data. Another website was historical articles used to analyze emotions. Main The library was a NumPy library, often used for cheating and cleaning and data sets, converting it into a form that can be used directly on ML models. One was sklearn a library, used for real statistics, estimates and predictions. Data we used they were historical stock markets, and they were collected from different communities and opened resources are available online, and 85 percent or more of the data is used for the training of ML models depending on the need and some are used for verification and testing purpose.

CHAPTER – 3
WORKING OF PROJECT

3.1 Types Of Data

SMP systems can be categorized according to the type of data they use as input. Most studies used market data in their analysis. Recent research has been considered text data from online sources as well. At this stage, subjects are categorized based the type of data they use for predictive purposes. At the end of this section, Table 1 points exclude data sources comparisons, input type and forecast duration used in lessons to date.

Market Data-

Market data is temporary data for price history related to financial markets. Analysts and traders use data to analyze historical trends and the latest stock prices in the market. They represent the information needed to understand market behavior. Market data is usually free, and can be downloaded directly from market websites. Various researchers have used this data to predict price movements using machine learning algorithms. Previous research has focused on two types of predictions. Some studies have used stock index predictions such as Dow Jones Industrial Average (DJIA) [35], Nifty [36], Standard and Poor's (S&P) 500 [37], National Association of Securities Dealers Automated quotes (NASDAQ) [38], Deutcher
Aktien Index (DAX) [39], as well as multiple indicators [40,41]. Some studies have used individual stock predictions based on specific companies such as Apple [42], Google [43], or corporate groups [12,44].

Textual Data-

Text data is used to analyze the effect of emotions on the stock market. Public sentiments have been shown to have a profound effect on the market. The most challenging part is converting text information into numerical values to provide a predictive model. In addition, extracting text data is a

challenging task. Text data has many sources, such as financial news websites, general news, and social media platforms [52]. Most studies are done on text data trying to predict whether the feeling directed at a particular cell is right or wrong. Previous research has examined several sources of SMP literature, such as the Wall Street Journal [53], Bloomberg [22], CNBC and Reuters [54], Google Finance [55], and Yahoo Finance [56]. Released stories may be general news or specific financial news, but most researchers use financial news, as it is considered not easy to be noisy [57]. Some researchers have used informal text data, such as message boards.

3.2 Data Pre-Processing

After the data is available, it needs be some pre-processing so that it can be fed to a machine learning model. Significance of the output depends on the pre-processing of the data.

Comparison of the data sources, type of input and prediction duration.

References	Data	Type of Input	Prediction Duration
[37]	S&P 500	Market data	Few days ahead
[38]	NASDAQ index	Market data	Few days ahead
[39]	DAX 30	broker house newsletters, RSS market feeds, and stock exchange data	Intraday
[56]	Yahoo Finance	Financial News	Intraday
[44]	DGAP, Euro-Adhoc	Corporate announcements financial new	Daily

[58]	Yahoo finance (18 Stock Companies data)	Market data, yahoo finance message board data	Daily
[35]	DJIA	Market data and Twitter	Daily
[32]	BSE and NSE stocks	Market data, technical indicators, Twitter data	Intraday
[36]	Nifty and Sensex	Market data and news	Intraday

| References	Feature Selection	Order Reduction	Feature Representation
[58]	Bag of words, LDA, JST, Aspect Based	-	TF-IDF
[47]	Correlation	Lemmatization	Boolean
[69]	Bag of Words	Chi2, Information Gain, Document Frequency, Occurrence	TF-IDF
[77]	Bag-of-word, Word2vec		TF-IDF
[78]	GA	PCA, FA, FO	-
[79]	N-grams	SVM based Recursive Feature Elimination, PCA, KPCA, and XGB	-
[73]	Bag-of-words	Occurrence	TF-IDF
[80]	GA, Feature Ranking	PCA-SVM, DA-RNN	-

Architecture / Framework
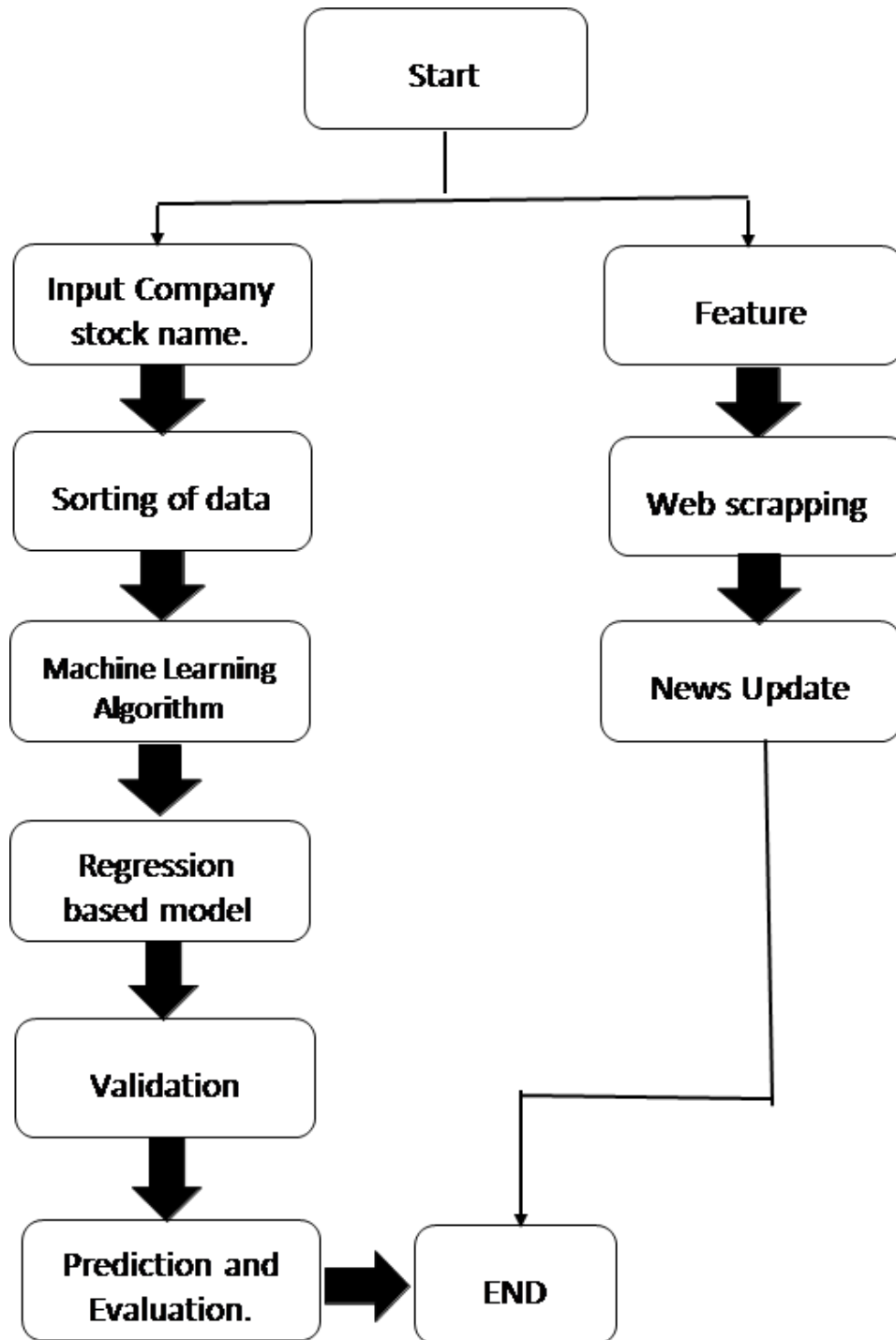
Fig 4. FrameWork

3.3 Comparative Analysis

The distribution of the number of papers published in recent years. The number of publications increased since 2009, and was very high in 2019, but during the last two years, the number of publications was low. The distribution of machine learning algorithms used for SMP, where SVM was the most popular method used. However, ANN and DNN have attracted the attention of the research community over the past few years. Traditional methods of neural network may not produce as accurate SMPs as before; the severity of randomly selected problems may suffer for good localization, and result in incorrect predictions. In-depth learning methods are used to analyze complex patterns in stock data, and provide very fast results. Moreover, there is no one such method that can promise the best results.

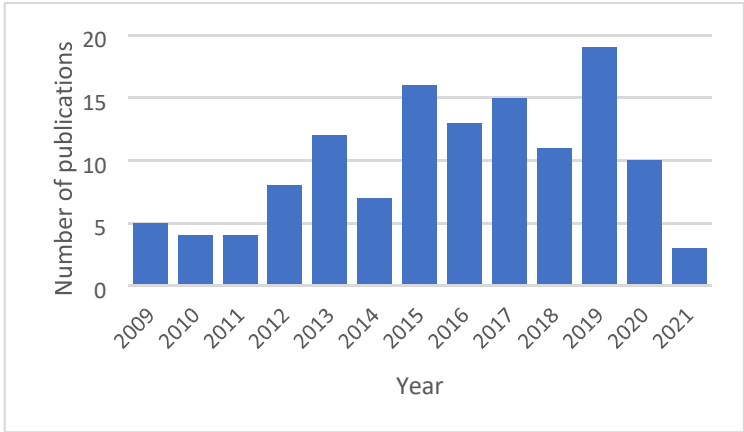Number of Publication per year-

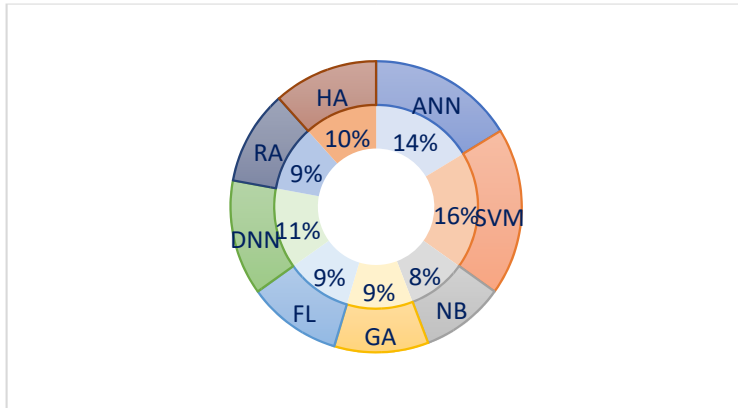**Fig 5.** publication per year

Distribution of the SMP techniques

Fig 6. Distribution of SMP techniques

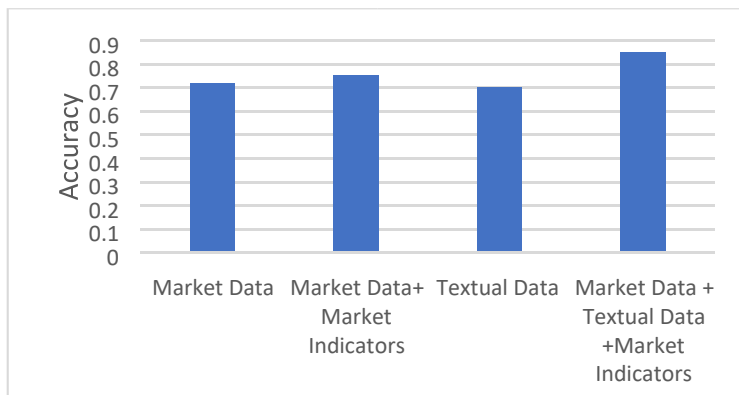**Comparison of the accuracies with different types of data**

Fig 7. Comparing different types of data

3.4 Predictability/Prediction Strategies

The latest methods of predicting the stock market have been presented and provided comparisons
analysis [10] of all of these Strategies.
Great prediction techniques such as data mining, machine learning and depth
learning strategies used to estimate future stock prices based on these strategies
and discuss their pros and cons. They are,
1. Holt-Winters
2. Artificial Neural Network
3. Hidden Markov Model
4. ARIMA Model
5. Time Series Linear Model
6. Recurrent Neural Networks.

3.5 Methodology-
In this project the prediction of stock market is done by the Support Vector Machine (SVM) and Radial Basis Function (RBF).

3.5.1 Support Vector Machine:
Support Vector Machine (SVM) discriminates against them the officially defined category is the dividing hyperplane. In other words, the training data provided by the label (supervised learning), the algorithm produces the appropriate hyperplane separating new models. In the two-dimensional
Space is the hyperplane line that divides the plane into two parts where each class lies side by side.
Support Vector Machine (SVM) is considered as one of the most relevant algorithms available at the time series predictions.

In machine learning, kernel radial base function, or RBF kernel, is a popular kernel function used in various kernelized learning algorithms. Mostly, of course widely used in vector support machine

separation. Radial base function is a function with its true value the number depends only on the distance from the source, so that; or otherwise in the distance from another point, called the center, too. Any work that satisfies the structure radial activity.

$$RBF = Local\ Response\ Task$$

The RBF Kernel is nothing more than a low-band pass filter, best known for Processing Cells as a tool for this smooth images. The RBF Kernel acts as a preferential option smooth solution.

3.5.2 Stock Market Prediction Steps:

Step 1: This step is important for the download data from the net. We are predicting the financial market value of any stock. So that the share value up to the closing date are download from the site.

Step 2: In the next step the data value of any stock that can be converted into the CSV file (Comma Separate Value) so that it will easily load into the algorithm.

Step 3: In the next step when the GUI is turned on and on click the SVM button will show the window from which it is generated we select the stock database database file.

Step 4: After selecting the stock database database file in its folder will show the stock before the map and the stock after to make a map.

Step 5: Next step algorithm calculates log2c once log2g value reduction error. Therefore, it will predict a graph with data rate successfully.

Step 6: In the final step algorithm display the predicted value selected stock graph showing actual value once the predicted amount of stock.
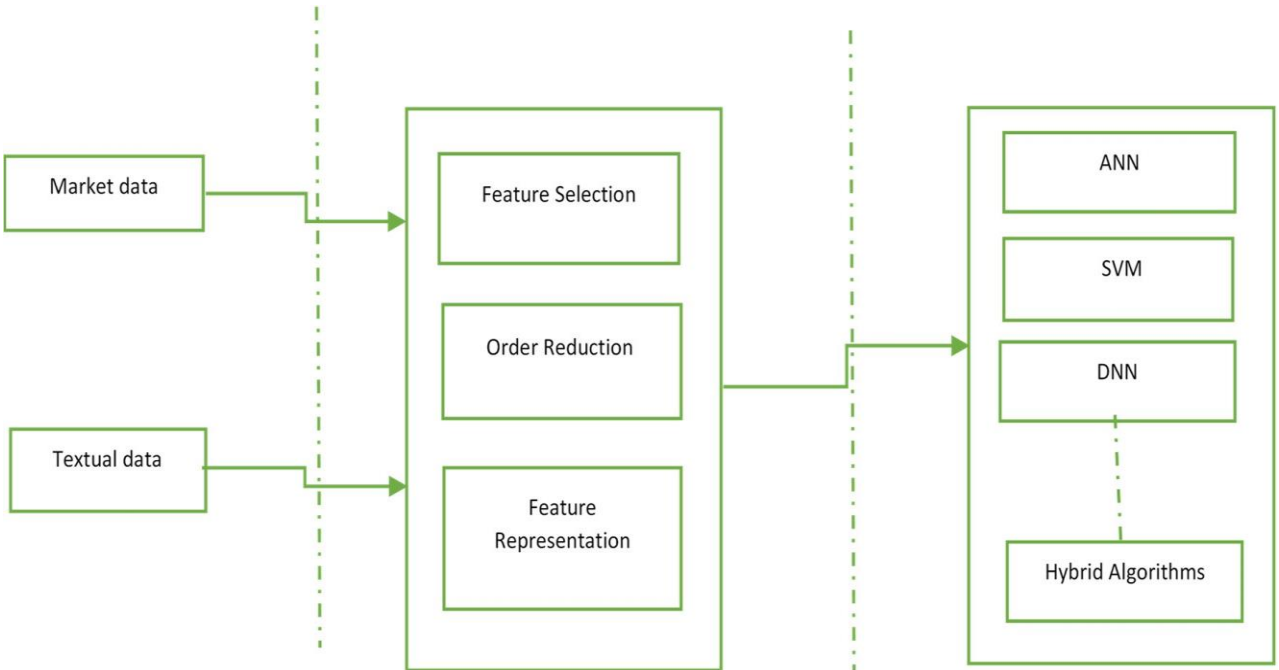
Fig 8. Prediction model

3.5.3 Step by Step implementation:

Step 1: Import the libraries

```
Python3

# Machine learning
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score

# For data manipulation
import pandas as pd
import numpy as np

# To plot
import matplotlib.pyplot as plt
plt.style.use('seaborn-darkgrid')

# To ignore warnings
import warnings
warnings.filterwarnings("ignore")
```

Fig 9.

Step 2: Read stock data

We will read the Stock Data Downloaded from yahoo Finance Website. The data is stored in OHLC(Open, High, Low, Close) format in a CSV file. To read a CSV file, you can use the read_csv() method of pandas.

```
Python3

# Read the csv file using read_csv
# method of pandas
df = pd.read_csv('RELIANCE.csv')
df
```

Fig.10

Output:

	Date	Open	High	Low	Close	Adj Close	Volume
0	2020-09-16	2320.000000	2369.350098	2310.550049	2324.550049	2317.096191	15668979.0
1	2020-09-17	2320.000000	2333.699951	2291.850098	2298.750000	2291.378906	11919927.0
2	2020-09-18	2314.250000	2319.449951	2276.550049	2305.699951	2298.306396	15264068.0
3	2020-09-21	2300.000000	2336.000000	2247.350098	2255.850098	2248.616455	15519031.0
4	2020-09-22	2277.000000	2277.000000	2201.550049	2211.149902	2204.059570	16056620.0
...	...	...	...	...	...	...	...
244	2021-09-09	2427.899902	2437.850098	2416.100098	2425.600098	2425.600098	4136538.0
245	2021-09-13	2433.000000	2433.000000	2368.050049	2371.550049	2371.550049	7527598.0
246	2021-09-14	2375.000000	2394.000000	2366.000000	2368.449951	2368.449951	4111205.0
247	2021-09-15	2368.500000	2395.750000	2368.500000	2378.300049	2378.300049	4186300.0
248	2021-09-16	2381.550049	2436.750000	2367.000000	2428.199951	2428.199951	6204799.0

249 rows × 7 columns

Fig 11.

Step 3: Define the explanatory variables
Explanatory or independent variables are used to predict the value response variable. The X is a dataset that holds the variables which are used for prediction. The X consists of variables such as 'open – close' and 'High – low'.

```Python3
# Create predictor variables
df['Open-Close'] = df.Open - df.Close
df['High-Low'] = df.High - df.Low

# Store all predictor variables in a variable X
X = df[['Open-Close', 'High-Low']]
X.head()
```

Fig 12.

Output:

Date	Open-Close	High-Low
2020-09-16	-4.550049	58.800049
2020-09-17	21.250000	41.849853
2020-09-18	8.550049	42.899902
2020-09-21	44.149902	88.649902
2020-09-22	65.850098	75.449951

Fig 13.

Step 4: Define the target variable
The target variable is the outcome which the machine learning model will predict based on the explanatory variables. Y is the target dataset storing the correct trading signal.

```Python3
# Target variables
y = np.where(df['Close'].shift(-1) > df['Close'], 1, 0)
y
```

Fig 14.

Step 5: Split the data into train and test

```Python3
split_percentage = 0.8
split = int(split_percentage*len(df))

# Train data set
X_train = X[:split]
y_train = y[:split]

# Test data set
X_test = X[split:]
y_test = y[split:]
```

Fig 15.

Step 6: Support vector classifiers(SVC)

```Python3
# Support vector classifier
cls = SVC().fit(X_train, y_train)
```

Fig 16.

Step 7: Strategy Implementation

We will predict the signal (buy or sell) using the cls.predict() function.

```Python3
df['Predicted_Signal'] = cls.predict(X)
```

Fig 17.

Calculate daily returns

```python
# Calculate daily returns
df['Return'] = df.Close.pct_change()
```

Fig 18.

Calculate strategy returns

```python
# Calculate strategy returns
df['Strategy_Return'] = df.Return *df.Predicted_Signal.shift(1)
```

Fig 19.

Plot strategy returns vs original returns

```python
import matplotlib.pyplot as plt
%matplotlib inline

plt.plot(Df['Cum_Ret'],color='red')
plt.plot(Df['Cum_Strategy'],color='blue')
```

Fig 20.

4. Result:

Out[59]: [<matplotlib.lines.Line2D at 0x7f6d741a5550>]

Fig 21. output from implementation

Open Challenges and problems-
Financial market analysis and forecasting continue to be an exciting and challenging issue. Today, access to data is simpler, but the difficulty increases with data acquisition and processing to extract important information and analyze their impact on stock prices. The removal of a feature from financial data is a challenging task, as it is important to look at the variability used to predict. Financial data sets are generally noisy. The data rate greatly affects SMPs.

Many books that predict stock about live test ensure that the previously proposed methods can be used in real time. However, these methods may work in controlled conditions. However, the biggest challenge will be the live prediction test. Live testing also comes with challenging features, such as price differences, sound, and unexpected events. One such example is Knight Capital Tragedy, where a loss of $ 440 million was tolerated by the company.

Market volatility is the extent to which the market price of an investment

fluctuates. The main reasons for volatility are uncertainty and inflation, and risk increases as the market fluctuates. The effects of mood swings are endless. Stock forecasting is a challenge when the market is volatile. One of the reasons for market instability is algorithmic trading. One such example was a flash crash, which cost $ 860 billion within 30 minutes of the US stock market. International politics also plays a major role in the evolution of the stock market.

Statement Of Data Availability: No new data were created or analyzed in this study.

Stock dataset for IBM Inc. in CSV file

	A	B	C	D	E	F
	VarName1	VarName2	VarName3	VarName4	VarName5	VarName6
	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER
251	192.00999	193	191.2	192.14999	3878600	187.30965
252	189.8	192.81	189.74001	192.27	5419700	187.42664
253	187.25	190.7	187.00999	190.00999	11255500	185.22356
254	198.05	198.71001	195	196.39999	8527300	191.45259
255	195.98	197.41	195.42	197.02	5351300	192.05698
256	196.24001	198.24001	195.88	197.77	5431100	192.78809
257	195.03999	197	194.27	195.19	4835000	190.27307
258	196.06	199.21001	195.66	195.67999	8417900	190.75072
259	193.88	197	193.60001	196.64	5480100	191.68654
260	193.89	194.17	192.69	193.28999	5125000	188.42093
261	191.72	195.31	191.57001	194.52	6740500	189.61995
262	193.12	193.97	191.28	191.77	6089900	186.93922
263	193.2	193.38	192.35001	192.69	4091100	187.83604
264	194.24001	194.5	192.49001	193.55	4924100	188.67439
265	193.12	195.13	192.78	194.5	5394100	189.60045

Fig.22 Stock dataset for IBM Inc.
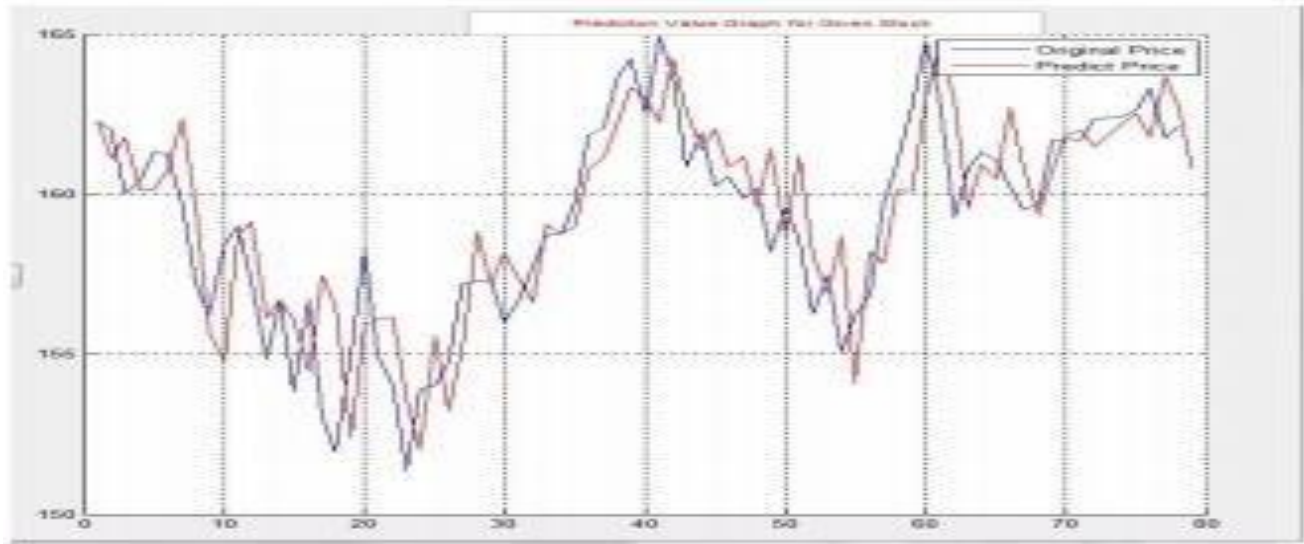
Predicted Output by SVM for IBM Inc.

Fig 23. Predicted Output for IBM Inc.

5. CONCLUSION:

In the project, we proposed the use of data collected from different global financial markets with machine learning algorithms to predict stock index movements. The SVM algorithm works with large amounts of data collected in different global financial markets. Also, SVM does not present the problem of overload. Various machine-based models have been proposed to predict the daily trends of Market stocks. The numerical results suggest higher efficiency. Effective trading models are built on our well-trained forecast. The model produces a higher profit compared to the selected benchmarks.

Financial markets offer an excellent platform for investors and traders, who can trade from any online gadget. In the last few years, people have done just that very attractive stock trading. Like any other lifestyle, the stock market has it also changed with the advent of technology. Now, people can invest their money grow. Online trading has only changed the way people buy and sell stocks. I budget markets have developed rapidly, and have built a interconnected world market place. This development paves the way for new opportunities. Unlike conventional frameworks, SMP is currently developed using a machine reading, big data analysis, and in-depth reading,

35

which provides the best decision to do. Stock markets, today, are exposed to the emotions of social media and cyber-attacks. Researchers can play an important role and thrive in these areas through development better and more secure trading structures.

REFERENCES

[1] Zhen Hu, Jibe Zhu, and Ken Tse "Stocks Market Predicting Using a Vector Support Machine", 6th
International Conference on Information Management, Innovation and Industrial Engineering, 2013.M.

[2] Wei Huang, Yoshiteru Nakamori, Shou-Yang Wang, "Predicting the direction of stock market movements with support vector mechanism", Computers & Operations Research, Volume 32, Issue 10, October 2005.

[3] N. Ancona, Vector Separation Properties Reversal equipment, technical report, RIIESI / CNR Nr. 02/99.

[4] K. jae Kim, "Financial forecast predictions are used support vector equipment," Neurocomputing, vol. 55, 2018.

[5] Debashish Das and Mohammad sharif Uddin data mining and neural

network strategies in the stock market forecasting: method review, international Journal of practical wisdom and applications, vol.4, no.1, January 2019.

[6] Krishna, V. ScienceDirect ScienceDirect NSE Stock Market Prediction Prediction Using Using Model Learning Models. Procedia Comput. Science. 2018.

[7] Market Fees for Listed Local Companies (US $ Current) Data. Available online: https://data.worldbank.org/ indicator / CM.MKT.LCAP.CD (accessed 19 May 2021).

[8] Upadhyay, A .; Bandyopadhyay, G. Predicting Stock Activity in the Indian Market using Multinomial Logistic Regression. J. Bus. Stud. Q. 2012.

[9] Tan, T.Z .; Quek, C .; Ng, G.S. Biological Brain-Inspired Genetic Complementary Learning for Stock Market and Bank Failure Prediction. Computer. Intell. 2007. [CrossRef]

[10] Ali Khan, J. Trending Predictions in Market Trading Using Machine Learning Phases. Science. Int. 2016.

[11] Gupta, R .; Garg, N .; Singh, S. Stock Market Prediction Accuracy Analysis Using Kappa Measure. Proceedings of the 2013 International Conference on Telecommunications Systems and Network Technology, Gwalior, India, 6–8 April 2013.

[12] Fama, E.F. It moves randomly with stock market prices. Finance. Indunu. J. 1995. [CrossRef]

[13] Bujari, A .; Furini, M .; Laina, N. By using castags to predict corporate stock trends. At the 14th Annual Conference of IEEE Consumer Communications and Communication (CCNC), Las Vegas, NV, USA, 8–11 January 2017.

[14] Inthachot, M .; Boonjing, V .; Intakosum, S. Artificial Neural Network and Genetic Algorithm Hybrid Intelligence Predicting Thai Stock Index Trends. Computer. Intell. Neurosci. 2016, 2016. [CrossRef]

[15] Park, C.-H .; Irwin, S.H. What Do We Know About the Benefits of Technology Analysis? J. Econ. Surv. 2007. [CrossRef]

[16] Venkatesh, C.K .; Tyagi, M. Basic Analysis As a Value for Value

Compared with Technical Analysis. Bangladesh Res. Published. J. 2011.

[17] Neir, B.B .; Mohandas, V. A wise stock exchange program. Intell. Decis. Technology. 2015. [CrossRef]

[18] Shiller, R. Balance expectations for investor and investor RJ Shiller. J. Psychology. Finance. Mark. 2000. [CrossRef]

[19] Aaron, D.Y .; Gavious, I .; Yosef, R. The effects of the stock market bubble on consolidation and acquisition. Q. Pastor Econ. Finance. 2010. [CrossRef]

[20] Kurach, R .; Słon´ski, T. PE Ratio and Predicted Growth — The Story of Poland. Folia Oecon. Stetin. 2015. [CrossRef]

[21] Deboeck, G. Edge Trading: Neural, Genetic, and Fuzzy Systems for Conflict Financial Markets; Wiley: New York, NY, USA, 1994.

[22] Dutta, A. Predicting Stock Acquisitions in Indian Commodity Market Through Depreciation. Learner job. J. Bus. Inf. 2012.