



NUS

National University
of Singapore

STOCK MARKET PREDICTION

MACHINE REASONING-REASONING SYSTEMS PROJECT

GROUP MEMBERS

Akshay Sachdeva
Aniket Mohan Arasanipalai
Siddhant Naveria
Ma Weizhong

STUDENT ID

A0198532W
A0198481N
A0198485H
A0198426R

Contents

1. BUSINESS PROBLEM BACKGROUND	3
2. OBJECTIVES AND SUCCESS MEASUREMENTS	4
2.1 OBJECTIVES	4
2.2 SUCCESS MEASUREMENTS	4
3. SOLUTION	5
3.1 TARGET AUDIENCE	5
3.2 DATA QUALITY AND KNOWLEDGE BASE	5
3.3 PROJECT SCOPE	5
3.3 PROCESS FLOW	6
4. IMPLEMENTATION	7
5. PERFORMANCE AND VALIDATION	12
6. CONCLUSION AND FUTURE STEPS	17
7. REFERENCES	18

EXECUTIVE SUMMARY

Stock market prediction is the act of trying to determine the future value of a company stock or other financial instrument traded on a financial exchange. The successful prediction of a stock's future price will maximize investor's gains.

Most adults nowadays are very much interested in trying to understand the functioning of stocks and when a particular time would be right to invest in it. These sentiments can change on a day to day basis. Based on this, our team has come up with a lightweight yet easily scalable solution to predict the *status of a stock* related to a specific company on an everyday basis.

Based on machine reasoning concepts and how logical rules and knowledge can be inferred with a reasoning based system, our group of four members, brainstormed and decided to build a system with hybrid architecture (Rules Engine, Genetic Algorithms and Web MVC technology) and automate, as much as we can, the business or functions based on the knowledge gleaned from the stock market industry

We began by getting a good understanding of the overall key processes of the proposed solution by mapping them out with data acquisition modelling. Next, we determined the areas which needed targeting that we agreed should be implemented into business processes. We also made sure we analysed the processes with our improvements and see how these improvements were handled with test cases to validate our assumptions.

Our primary objective is to enable a user to understand the historical stock prices and provide him the visibility into the movement of the stock prices for the following day. For this project, we have considered companies from the NASDAQ 100 Stock List.

1. BUSINESS PROBLEM BACKGROUND

One of the more interesting areas over recent time has been in the area of stocks. In Stock Market Prediction, the aim is to predict the future value of the financial stocks of a company.

A correct prediction of stocks can lead to huge profits for the seller and the broker. Frequently, it is brought out that prediction is chaotic rather than random, which means it can be predicted by carefully analysing the history of respective stock market.

Stock market prediction seems a complex problem because there are many factors that have yet to be addressed and it doesn't seem statistical at first. But by proper use of techniques, one can relate previous data to the current data and train the machine to learn from it and make appropriate assumptions.

The recent trend in stock market prediction technologies is the use of genetic algorithms which makes predictions based on the values of current stock market indices by training on their previous values. Genetic algorithms itself employs different iterations to make prediction easier and authentic. There are various factors considered such as open, close, low, high and volume that go a long way to help predicting the outcome.

2. OBJECTIVES AND SUCCESS MEASUREMENTS

2.1 OBJECTIVES

The objective of this project is to create an intelligent framework to recommend and enable a user to understand the stock prices for a specific company or a set of companies.

With the vast majority of population, looking to invest in stocks, and with so many factors and categories to keep track of, it becomes easy if they are able to identify the stock which are performing well in the market, in order to minimize the losses.

2.2 SUCCESS MEASUREMENTS

It is vital to understand the success of the system based on a few measures.

There are a few key measures of our system:

1. Whether the system was able to understand the user
2. Whether the system was able to respond to the specific query being asked
3. Whether the system was able to present the related information
4. Whether the system was able to provide alternatives in case information related to user's query was not present.

3. SOLUTION

3.1 TARGET AUDIENCE

The major target audience would be the users/stock traders who want information on the stock prices so that they can invest in the same based on the status of the price (up or down).

.

3.2 DATA QUALITY AND KNOWLEDGE BASE

All information used to build the ‘Stock Market Prediction’ system is from the stock information and indicators provided online. It is assumed that the data present on these websites are correct.

3.3 PROJECT SCOPE

In this project, a stock prediction system is created, using the front end – back end framework, with the help of various stock market indicators and rules that are generally used in everyday stock market price calculation. The stocks associated with the system are considered from the NASDAQ100 list.

The system will respond to user preferences and provide the answer on a web application.

3.3 PROCESS FLOW

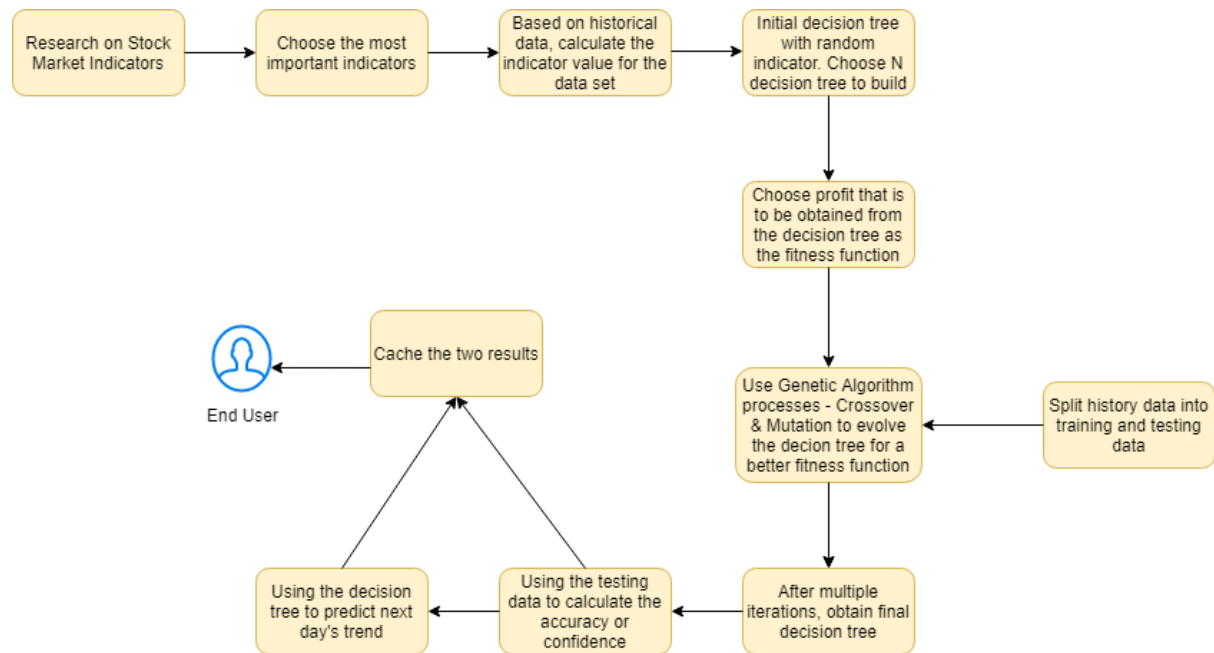


Fig 3.3 Process Flow

The above diagram displays the process flow of the project.

We initially choose the most important stock market indicators that are used as part of the stock market industry. Once we obtain these, we calculate the indicator value for the data set. We use a decision tree Algorithm to build a decision trees with random indicators.

The fitness function will enable us to obtain the profit value, while the usage of Genetic Algorithm will help to evolve the decision tree for a better fitness function. Furthermore, we split the historical data into training and testing data. After multiple iterations, we obtain the final decision tree.

On caching the results obtained from predicting the trend as well as the confidence value, the user can see the output on the UI.

4. IMPLEMENTATION

4.1 SYSTEM DEVELOPMENT

System Development is the process of defining, designing, testing and implementing a new software application or a program. It can include the internal development of customized systems, the creation of database systems or the acquisition of third party developed software.

4.2 SYSTEM ARCHITECTURE

The system architecture is based on three main principles which are intelligence and robustness, scalability and modularity. The JDK Engine running Java is intelligent and robust, as it can perform the required machine reasoning based on the specified statements and provide consistent and accurate outputs.

The scalability and modularity of the system architecture allows easy deployment to docker, Kubernetes and other cloud platforms for rapid expansion of operations as well as integration with other backend systems to provide more functionalities.

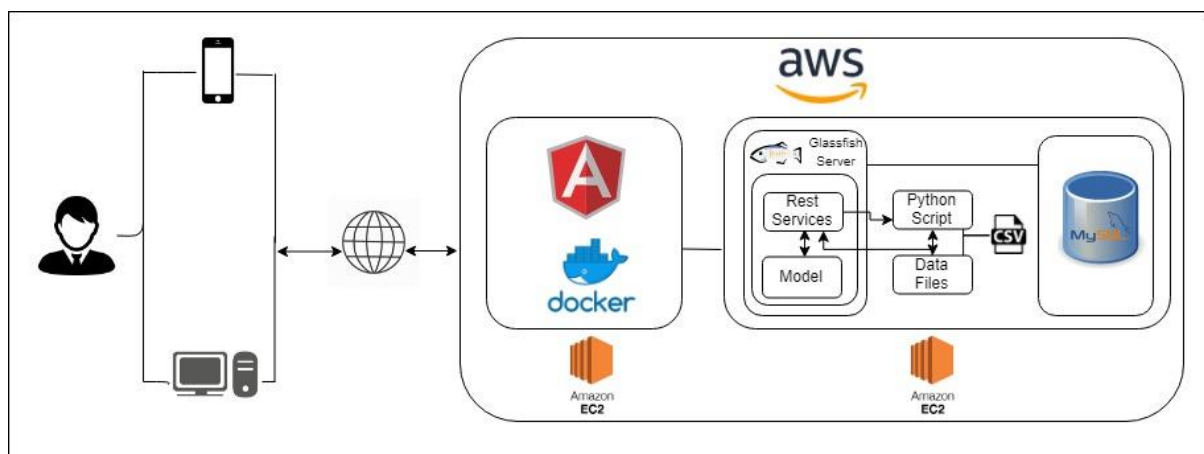


Fig 4.2 System Architecture

Fig 4.2 shows the system architecture of our Stock Market Prediction. The system has three main components.

Front End Framework:

The front end framework has been designed using Angular JS. The major benefits of using Angular JS are –

- **Simple architecture**
AngularJS development is considered as the simpler design architecture which is used nowadays and easy to catch up for any developers and enhance the features to any extent which the client requires. It works extremely well when it comes to managing the heavy web applications that contain a number of components and complex requirements
- **Improved Design Architecture**
Some large application contains a bulk of components numbers more than 60. AngularJS makes it easy to manage them even after a new programmer joins the project in the middle of the process. The architecture is built in a way which helps the programmer to locate and develop the code without any difficulty.
- **Lesser code and improved reusability**
With lesser coding required, the focus can be put towards improving the efficiency of the system. Moreover, the same chunk of code can be reused.

Back End Framework:

The backend framework is wholly built in Java and consists of various machine intelligence algorithms such as decision trees and genetic algorithms.

- **Genetic Algorithms -**
Genetic algorithm is an optimization technique, which tries to find out such values of input so that we get the best output values or results.
The working of the Genetic Algorithm is as below –

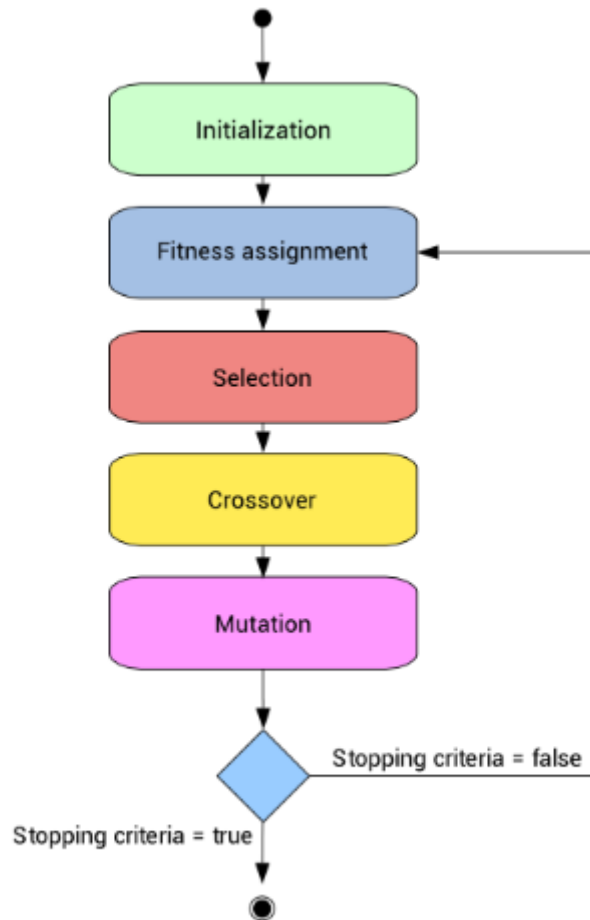


Fig 4.2 Process Flow of Genetic Algorithm

- Initialization**
 To solve this problem using genetic algorithm, our first step would be defining our population. So, our population will contain individuals, each having their own set of chromosomes, in our case, this would be random population assigned based on the output of the decision tree (which uses the market indicators).
- Fitness Assignment**
 The fitness function in Stock Market Prediction is the profit function which helps us understand the status of the stock price
- Selection**
 Once the fitness function is assigned, we can select fit chromosomes and allow them to produce off-springs. To avoid having less diversity in the chromosomes, we have used the *Roulette Wheel Selection Method*
- Crossover**
 As we have selected parent chromosomes that will produce off springs in the previous step, cross over is directly related to reproduction. This is done to produce better offspring's

- Mutation

This is generally done, to initiate a random tweak in the chromosome. In our project, we have considered mutation after every 2000 generations to improve the accuracy of our output.

So the entire process is summarised as shown in the figure.

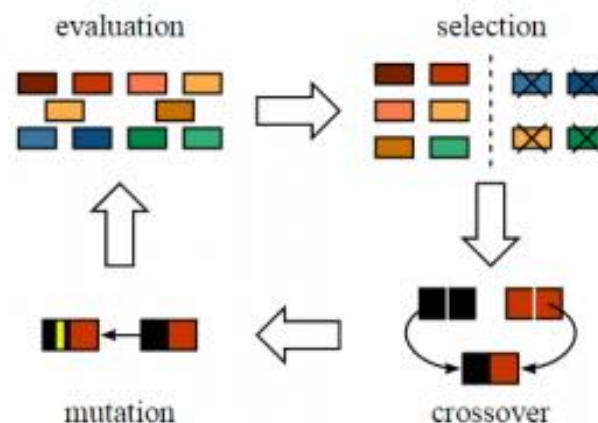


Fig 4.2 Genetic Algorithm Summary

Once we have reached our best possible solution, we terminate the algorithm.

- Decision Tree

Decision trees are commonly used in machine learning, help identify a strategy most likely to reach a goal.

Decision Trees have been used to help reach the goal which in our project is identifying the stock price status. There are certain stock market indicators like moving averages, relative strength index, on balance volume (OBV) etc. that have been considered while creating the nodes of the decision tree.

The choice of this structure

- **Business Agility:** Separating the Java components from the main application allow business analysts to work on building the business rules and business process, without the need to care about any application development. This greatly speeds up the production deployment of any enhancement to business rules or processes
- **High Availability and Scalability:** The JDK powered Java Application supports the deployment to OpenShift and docker container.

Deployment:

The full project has been deployed on AWS Educate which provides easy accessibility, Flexibility and ease of use. The application works as cloud-app in which cloud-based components work together on remote servers processing logic that is accessed through a web browser with a continual internet connection.

Few Benefits of Deploying the application on AWS cloud platform are:

- **Fast Response:** As the application is deployed on cloud, there is no need for the local setup of the software and databases required to operate the application.
- **API use:** Cloud application can be kept smaller by using APIs to hand data to applications or API-based back-end services for processing or analytics computations.

5. PERFORMANCE AND VALIDATION

We have performed performance on 2 different platforms to ensure that the project provides the correct expected output: locally on the Desktop and on the Cloud AWS.

5.1 LOGIN PAGE:

While testing the login page if the user enters the correct login details which are: username: “admin” and the password:” admin”. The details get authenticated from the user list which is stored in MYSQL Database. IF the details entered are authenticated and correct, Home Page is opened.

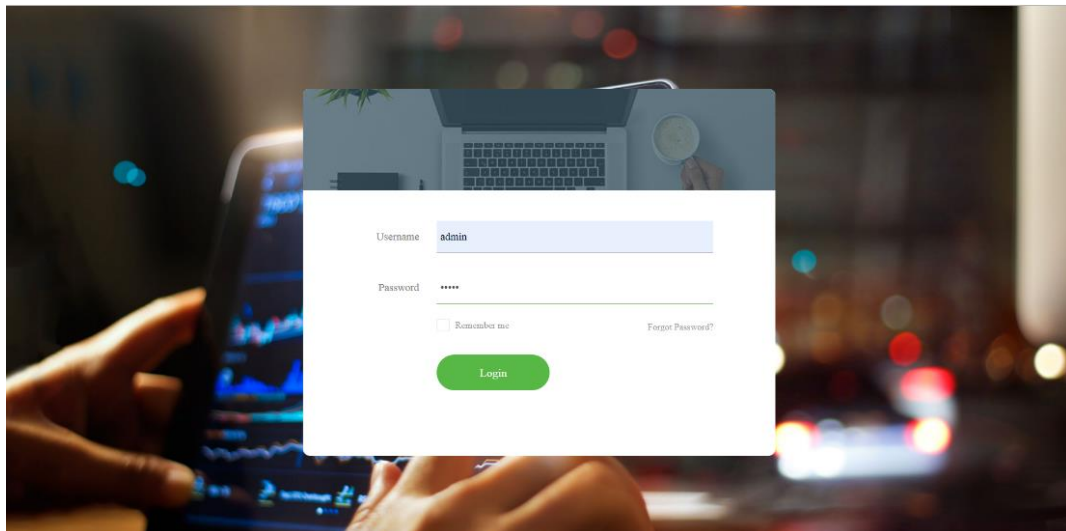


Fig. 5.1 Login Page

5.2 HOME PAGE:

Loading the Home Page provides with the list of companies that are fetched from the Database. The Database contains the mapping of the company’s name and symbol, when we select the company name and click on Submit button the name is matched and mapped to the symbol.

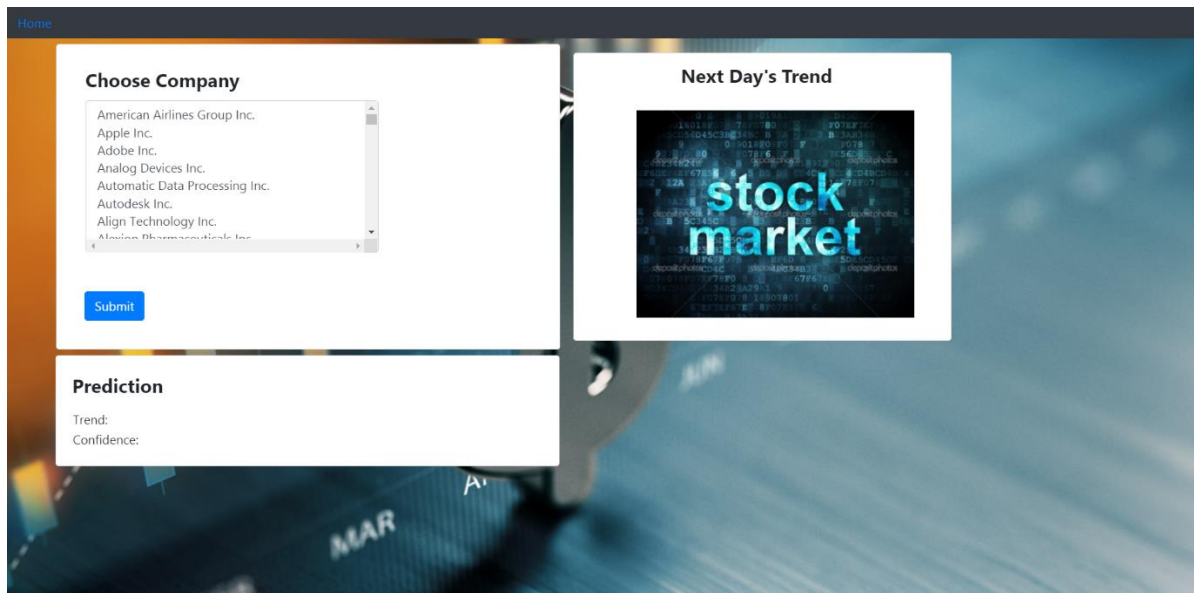


Fig. 5 Home Page

5.3 CREATION OF CSV FILES

Once the user selects the name of the company, it creates the list of all the Nasdaq 100 companies and performs the analysis on those. To make the processing fast and easier for the end user, it fetches all the list of the companies in one go, applies the indicator on those and generates the csv file. This process is done by the python script and later used by Genetic Algorithm to predict the stocks

nasdaq	18-09-2019 14:30	File folder	
AAL.csv	18-09-2019 18:03	Microsoft Excel Co...	45 KB
AAPL.csv	18-09-2019 18:03	Microsoft Excel Co...	45 KB
ADBE.csv	18-09-2019 18:03	Microsoft Excel Co...	44 KB
ADI.csv	18-09-2019 18:03	Microsoft Excel Co...	44 KB
ADP.csv	18-09-2019 18:04	Microsoft Excel Co...	45 KB
ADSK.csv	18-09-2019 18:03	Microsoft Excel Co...	44 KB
ALGN.csv	18-09-2019 18:03	Microsoft Excel Co...	44 KB
ALXN.csv	18-09-2019 18:03	Microsoft Excel Co...	45 KB
AMAT.csv	18-09-2019 18:03	Microsoft Excel Co...	44 KB
AMD.csv	18-09-2019 18:03	Microsoft Excel Co...	45 KB
AMGN.csv	18-09-2019 18:03	Microsoft Excel Co...	45 KB
AMZN.csv	18-09-2019 18:03	Microsoft Excel Co...	42 KB
ASML.csv	18-09-2019 18:03	Microsoft Excel Co...	45 KB
ATVI.csv	18-09-2019 18:03	Microsoft Excel Co...	45 KB
AVGO.csv	18-09-2019 15:32	Microsoft Excel Co...	44 KB
BIDU.csv	18-09-2019 15:32	Microsoft Excel Co...	44 KB
BIIB.csv	18-09-2019 15:32	Microsoft Excel Co...	44 KB
BKNG.csv	18-09-2019 15:32	Microsoft Excel Co...	43 KB
BMRN.csv	18-09-2019 15:32	Microsoft Excel Co...	44 KB
CDNS.csv	18-09-2019 15:32	Microsoft Excel Co...	43 KB
CELG.csv	18-09-2019 15:32	Microsoft Excel Co...	43 KB
CERN.csv	18-09-2019 15:33	Microsoft Excel Co...	44 KB
CHKP.csv	18-09-2019 15:33	Microsoft Excel Co...	44 KB
CHTR.csv	18-09-2019 15:33	Microsoft Excel Co...	43 KB
CMCSA.csv	18-09-2019 15:33	Microsoft Excel Co...	44 KB
COST.csv	18-09-2019 15:33	Microsoft Excel Co...	44 KB
CSCO.csv	18-09-2019 15:33	Microsoft Excel Co...	45 KB

Fig. 5.3.1 List of CSV's

Each CSV file must generate the data having multiple indicators providing the binary output which are in turn used to train the model. Each file creates the data with one year of values from the data on which the system is being executed. So for every time, the system is executed these files are generated.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
1	Date	open	high	low	close	volume	SP-SMAS	SP-SMA7	SP-SMA14	SP-SMA21	SP-SMA28	SP-SMA35	SP-SMA42	SP-SMA49	SP-SMA56	SP-SMA63	SP-SMA70	SP-SMA77	SP-SMA84	SP-SMA91	SP-SMA98	SP-SMA105	SP-SMA112	SP-SMA119	SP-SMA126	SP-SMA133	SP-SMA140
2	15-10-2018	31	32.24	30.82	31.78	1.1E+07	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
3	16-10-2018	32.06	33.44	31.98	33.28	1.1E+07	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
4	17-10-2018	34.45	35.4	33.53	33.57	1.5E+07	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
5	18-10-2018	33.58	33.76	31.87	32.06	1.3E+07	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
6	19-10-2018	32.09	32.65	31.81	32.04	1E+07	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
7	22-10-2018	32.27	32.52	31.82	32.16	7572100	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
8	23-10-2018	31.53	32.44	30.6	32.38	1.2E+07	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
9	24-10-2018	32.48	32.75	30.24	30.34	1.5E+07	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
10	25-10-2018	31.19	33.33	31.12	32.37	1.8E+07	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
11	26-10-2018	31.4	33.01	31.3	32.46	1.1E+07	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
12	29-10-2018	33.14	33.99	31.95	32.6	9405500	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
13	30-10-2018	32.44	34.78	32.33	34.66	1.5E+07	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
14	31-10-2018	35.16	35.9	34.81	35.08	1.1E+07	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
15	01-11-2018	35.3	36.47	35.07	36.37	9643200	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
16	02-11-2018	36.59	37.46	35.64	36.33	1E+07	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
17	05-11-2018	36.35	36.52	35.13	35.72	1.1E+07	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
18	06-11-2018	35.6	35.96	34.84	35.17	1.1E+07	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
19	07-11-2018	35.57	37.39	35.48	36.97	1.1E+07	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
20	08-11-2018	36.77	37.05	35.97	36.86	6884800	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
21	09-11-2018	36.7	37.26	36.03	36.22	6794400	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
22	12-11-2018	36.31	37.3	35.78	36.86	9360800	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
23	13-11-2018	37.15	38.42	37.1	37.78	9694100	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
24	14-11-2018	38	38.59	37.45	38.11	7288400	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
25	15-11-2018	37.87	38.16	36.31	37.82	8296700	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
26	16-11-2018	37.4	37.53	36.5	36.75	1.1E+07	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
27	19-11-2018	36.76	37.08	35.31	35.82	9421000	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
28	20-11-2018	35.05	36.39	34.86	35.59	7059600	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
29	21-11-2018	36.09	37.51	35.97	36.31	8701500	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
30	23-11-2018	36.66	38.27	36.66	37.95	6121300	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
31	26-11-2018	38.3	38.62	37.7	38.1	9203100	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
32	27-11-2018	38.21	38.77	37.81	38.29	8143000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
33	28-11-2018	38.59	38.97	37.78	38.94	6673800	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
34	29-11-2018	38.8	39.34	38.4	38.42	5178100	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
35	30-11-2018	38.42	40.2	38.42	40.16	1E+07	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

Fig. 5.3.2 Data Generation in CSV

5.4 ANALYSING THE TREND AND CONFIDENCE VALUES:

The Trends and Confidence Values can be analysed from the User-Interface and from the console of eclipse. The Trends Provides us the information about the stocks predicted for next day. If the stocks are down that means, that the person should not buy the stocks of a company as they are going to get lower prices on the next day. The confidence value provides the accuracy of the trend that it's true.

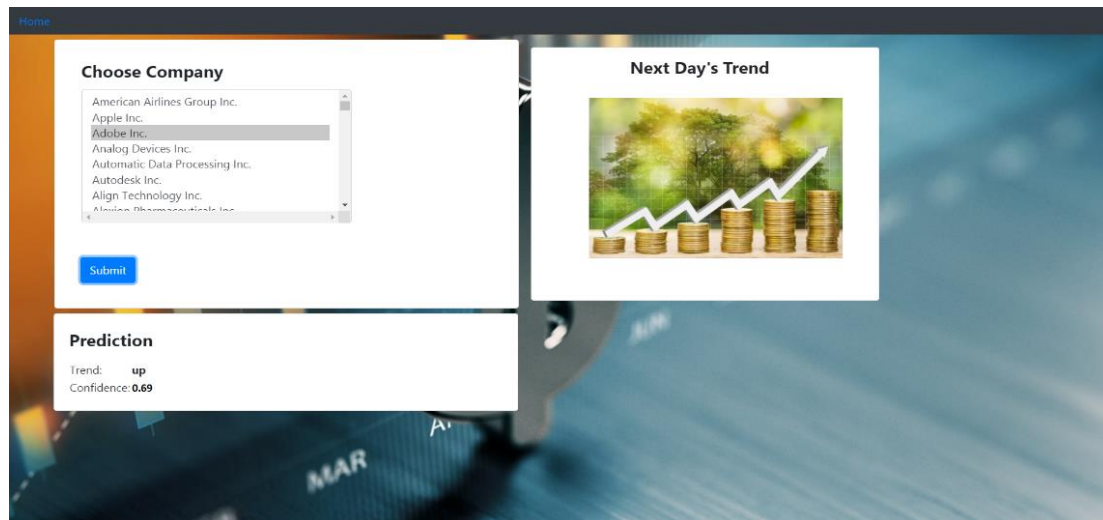


Fig. 5.4.1 Up-Trend

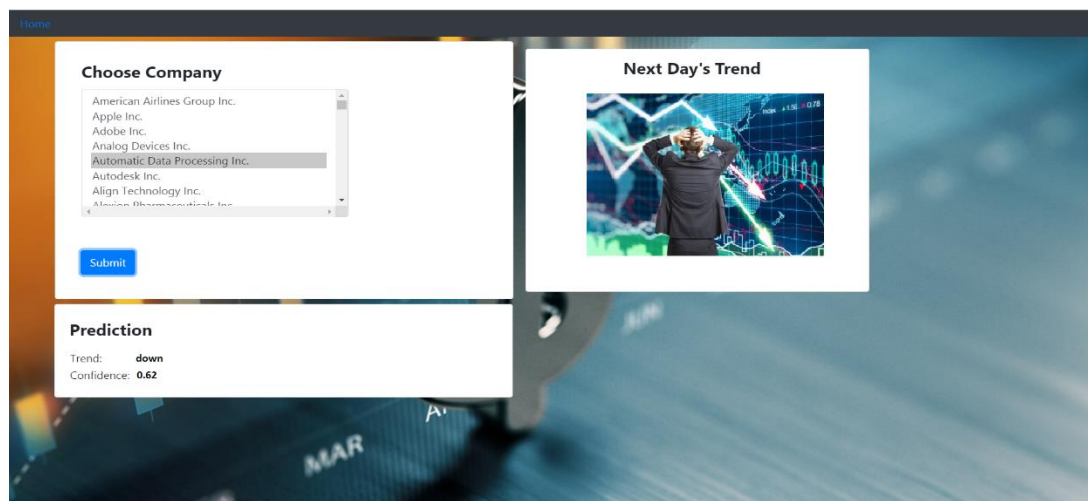
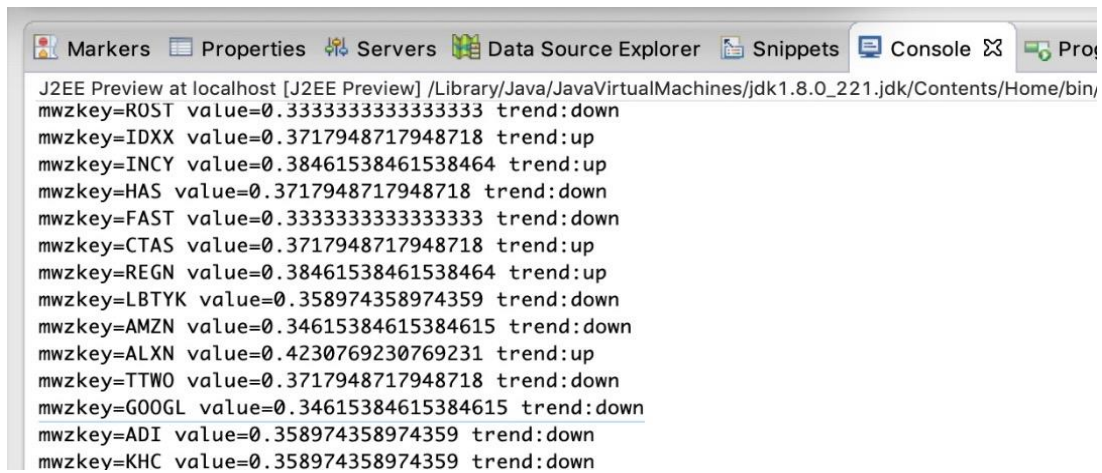


Fig. 5.4.2 Down-Trend

The output is also generated on the console when the system is executed locally. This provides the insight of the trend and the confidence value of the output generated.



```
J2EE Preview at localhost [J2EE Preview] /Library/Java/JavaVirtualMachines/jdk1.8.0_221.jdk/Contents/Home/bin/
mwzkey=ROST value=0.3333333333333333 trend:down
mwzkey=IDXX value=0.3717948717948718 trend:up
mwzkey=INCY value=0.38461538461538464 trend:up
mwzkey=HAS value=0.3717948717948718 trend:down
mwzkey=FAST value=0.3333333333333333 trend:down
mwzkey=CTAS value=0.3717948717948718 trend:up
mwzkey=REGN value=0.38461538461538464 trend:up
mwzkey=LBTYK value=0.358974358974359 trend:down
mwzkey=AMZN value=0.34615384615384615 trend:down
mwzkey=ALXN value=0.4230769230769231 trend:up
mwzkey=TTWO value=0.3717948717948718 trend:down
mwzkey=GOOGL value=0.34615384615384615 trend:down
mwzkey=ADI value=0.358974358974359 trend:down
mwzkey=KHC value=0.358974358974359 trend:down
```

Fig. 5.4.3 Console Page on Execution of Code

6. CONCLUSION AND FUTURE STEPS

In this project, a Stock Market prediction system is designed, developed and deployed via Amazon AWS to provide the ease to predict the stock prices of the next day, which in turn provides the Trend and the confidence for NASDAQ 100 companies.

Firstly, the system is executed using Angular App deployed on AWS which authenticates the user through the REST API and displays the list of over 100 companies, on which the prediction can be made. When the Home page appears, python script which fetches the data from the yahoo-finance library is executed. After the text is retrieved, indicators are applied, and the csv files are generated. These .csv files are used to train the Genetic Algorithm model which provides the Trend and confidence for every stock rate of the next day. The result is pulled from the backed and displayed to the User Interface. These predictions can help the user to invest in stock market accordingly.

Therefore, the stock prediction system can serve the user to predict the next day stock prices and provide the insight of the investment to be made.

The inherent limitation of the system is that it can predict the stocks only for the next day and provides model accuracy of about 67.52%. Currently, the model only predicts for NASDAQ 100 companies which can be modified to predict the data for other companies too.

Future enhancements of the project could involve, increasing the accuracy of the model and to predict the stocks for other companies.

7. REFERENCES

“Information for Building the Generic Algorithm for Stock Prediction”

- <https://ieeexplore.ieee.org/document/8212724>
- <https://www.investfly.com/builder/documentation>
- <https://www.sciencedirect.com/science/article/pii/S1877042812023555>

“For Indicators Selection”

- <https://economictimes.indiatimes.com/wealth/invest/7-stock-market-technical-indicators-that-can-help-you-invest/articleshow/63342749.cms?from=mdr>
- <https://www.investopedia.com/articles/active-trading/041814/four-most-commonlyused-indicators-trend-trading.asp>

“For Deploying on AWS”

- <https://docs.travis-ci.com/user/docker/>
- <https://www.cloudmantra.net/blog/glassfish-installation-on-ec2-amazon-linux/>
- <https://www.cloudjojo.com/how-to-install-mysql-server-in-amazon-ec2-instance/>