
VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

# PROJECT REPORT

Topic: **Stock Market prediction using Bio - Inspired Computing (Ant Colony Optimisation)**

## Slot: B2+TB2

**Dr. Rajeshkannan R**

## Members:

Avik Mitra : 17BCE0208
Achyut Tripathi : 17BCE0954
Priyanshu Mittal : 17BCE0027

## ABSTRACT:

Nature is an immense source of inspiration for solving hard and complex problems in computer science. It exhibits extremely diverse, dynamic, robust, complex and fascinating phenomenon. It always finds the optimal solution to solve its problem and maintains perfect balance among its components. This has been the boost behind bio-inspired computing. Nature inspired algorithms are meta heuristics that mimics the nature for solving optimization problems opening a new era in computation.

In our project we plan to create an artificial intelligence system that can predict closing prices of stock markets, when trained by the Swarm based Optimisation algorithm: Ant Colony Optimisation.

## INTRODUCTION:

Our project aims to utilise the ant colony optimisation algorithm to predict stock prices to a number of days into the future. The ant colony method involves several ants that take different pathways and leave a chemical trail of pheromones which is later picked up by several ants following it. This creates the most useful or optimised path for the ants. In our context it helps us determine the future pathway the ants are going to take given a certain period of time.

## METHOD:

The project involves an optimisation procedure towards a basic stock database. The best route path of ants (representing stocks) and cities (representing dates) is charted through the bio inspired ant colony optimisation algorithm and then the qualities are passed to the prediction model that is created within the ACO. Once applied to the stocks through the stock access module stocker, the best path as well as the price on the given day of the prediction is calculated and displayed along with the confidence upper limit and lower limit, the area within which is shaded. The time is also a key factor in determining the best route for the problem as it helps set a trend and therefore train

through the available data. It will then be able to decide on its current path which route to take (predict) based on its past decisions in the dataset as accessed by the stocker module into the program.

## DATASET SPECIFICATION:

The datasets used are accessed through a module called stocker. It is a python module that has a repository of major stocks and needs access to the internet to attract a specific stock when called upon.

The stocker repository contains historical data exceeding 20 years (approx.) for all major companies/players in the stock market. We can make intelligent decisions based on it. Huge datasets give us the chance to fine tune our prediction model and serve a better estimation of the future.

Sample:

For Apple Stocks (Code 'AAPL'):

```
AAPL Stocker Initialized. Data covers 1980-12-12 to 2018-03-27.
Maximum Adj. Close = 181.72 on 2018-03-12.
Minimum Adj. Close = 0.16 on 1982-07-08.
Current Adj. Close = 168.34 on 2018-03-27.
```

Plotted:

**TABULATED DATASET:**

| Date | Open | High | Low | Close* | Adj. close** | Volume |
|---|---|---|---|---|---|---|
| 23-Mar-2018 | 1,539.01 | 1,549.02 | 1,495.36 | 1,495.56 | 1,495.56 | 80,06,000 |
| 22-Mar-2018 | 1,565.47 | 1,573.85 | 1,542.40 | 1,544.92 | 1,544.92 | 63,20,600 |
| 21-Mar-2018 | 1,586.45 | 1,590.00 | 1,563.17 | 1,581.86 | 1,581.86 | 47,50,800 |
| 20-Mar-2018 | 1,550.34 | 1,587.00 | 1,545.41 | 1,586.51 | 1,586.51 | 45,81,600 |
| 19-Mar-2018 | 1,554.53 | 1,561.66 | 1,525.35 | 1,544.93 | 1,544.93 | 65,80,800 |
| 16-Mar-2018 | 1,583.45 | 1,589.44 | 1,567.50 | 1,571.68 | 1,571.68 | 54,25,000 |
| 15-Mar-2018 | 1,595.00 | 1,596.91 | 1,578.11 | 1,582.32 | 1,582.32 | 40,69,700 |
| 14-Mar-2018 | 1,597.00 | 1,606.44 | 1,590.89 | 1,591.00 | 1,591.00 | 42,59,400 |
| 13-Mar-2018 | 1,615.96 | 1,617.54 | 1,578.01 | 1,588.18 | 1,588.18 | 65,31,900 |
| 12-Mar-2018 | 1,592.60 | 1,605.33 | 1,586.70 | 1,598.39 | 1,598.39 | 51,74,200 |
| 09-Mar-2018 | 1,563.50 | 1,578.94 | 1,559.08 | 1,578.89 | 1,578.89 | 45,34,300 |
| 08-Mar-2018 | 1,550.00 | 1,554.88 | 1,545.25 | 1,551.86 | 1,551.86 | 38,62,600 |
| 07-Mar-2018 | 1,526.52 | 1,545.90 | 1,522.51 | 1,545.00 | 1,545.00 | 41,74,100 |
| 06-Mar-2018 | 1,533.20 | 1,542.13 | 1,528.00 | 1,537.64 | 1,537.64 | 45,61,700 |
| 05-Mar-2018 | 1,494.24 | 1,525.38 | 1,481.00 | 1,523.61 | 1,523.61 | 52,33,900 |
| 02-Mar-2018 | 1,469.10 | 1,501.05 | 1,455.01 | 1,500.25 | 1,500.25 | 65,87,600 |
| 01-Mar-2018 | 1,513.60 | 1,518.49 | 1,465.00 | 1,493.45 | 1,493.45 | 68,35,200 |

## Attributes of the dataset:

Date: Mentions the date of the stock fluctuation

Open: Price of the stock when market opens

High: Highest price per stock on the given date

Low: Lowest price per stock on the given date

Close: Closing price, that is, the price of the stock when the market closed on that day

Adj.close: Closing price that is adjusted in case of error or anomalies.

Volume: Total volume or number of stocks traded (bought as well as sold)

## RESULTS:

```
Predicted Price on 2019-03-27 = $2589.65
```

## ALGORITHM TRACEOUT:

**Begin**

Initialize

**While** stopping criterion not satisfied **do**

Position each ant in a starting node

**Repeat**

**For each** ant **do**

Choose next node by applying the state transition rule

Apply step by step pheromone update

**End for**

**Until** every ant has built a solution

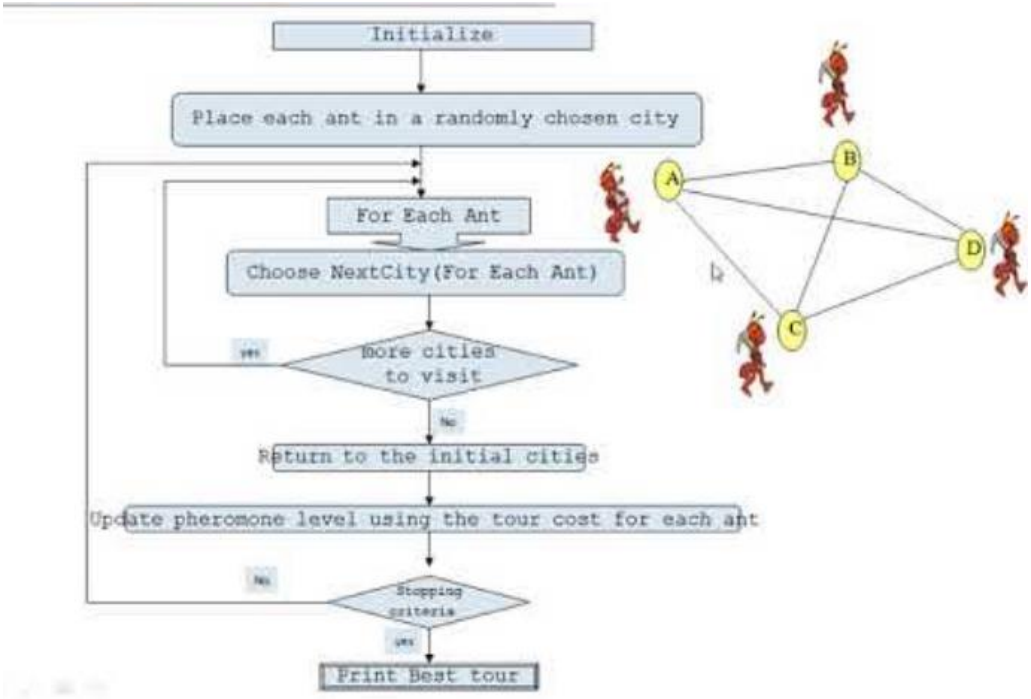
Update best solution

Apply offline pheromone update

**End While**

**End**

## Algorithm Flowchart:



## FORMULA USED:

$$\Delta\tau_{i,j}^{k} = \begin{cases} \frac{1}{L_k} \\ 0 \end{cases}$$ **kth ant travels on the edge i to j**

**Amt of Pheromone deposited.**

$$\tau_{i,j}^{k} = \sum_{k=1}^{m} \Delta\tau_{i,j}^{k}$$ Without vaporization

$$\tau_{i,j}^{k} = (1-\rho)\,\tau_{i,j} + \sum_{k=1}^{m} \Delta\tau_{i,j}^{k}$$ With vaporization

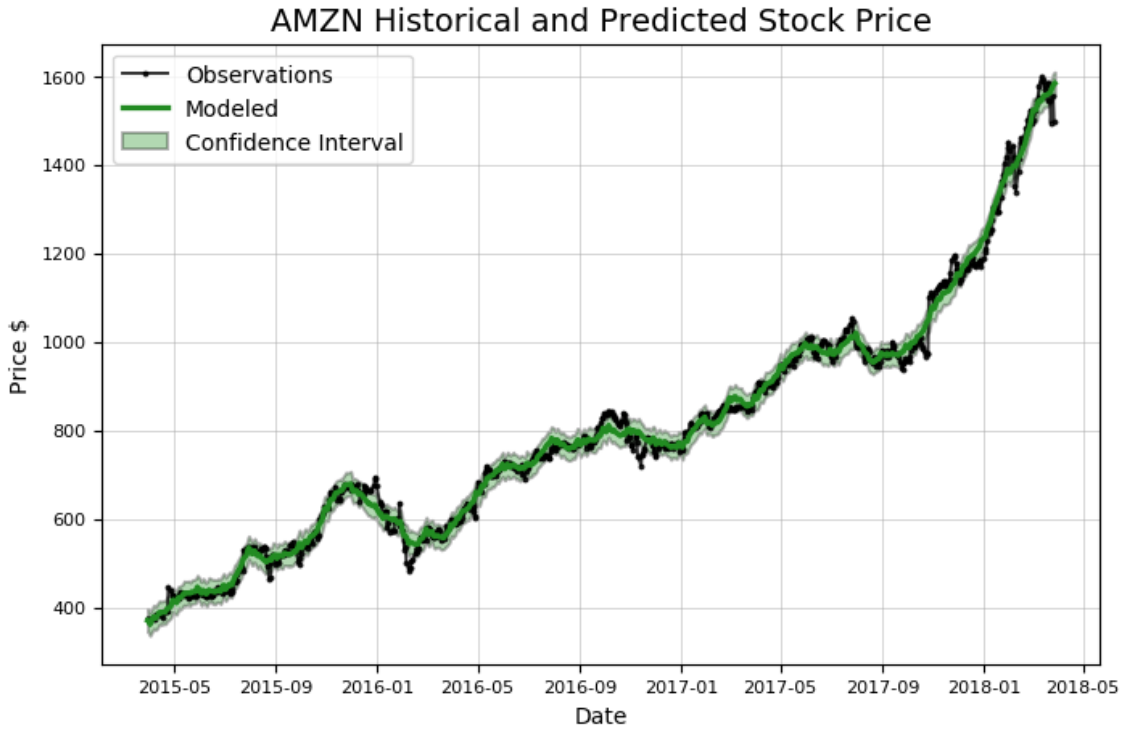
$$P_{i,j} = \frac{(\tau_{i,j})^{\alpha}(\eta_{i,j})^{\beta}}{\sum\left((\tau_{i,j})^{\alpha}(\eta_{i,j})^{\beta}\right)}$$ **Path Probability.**

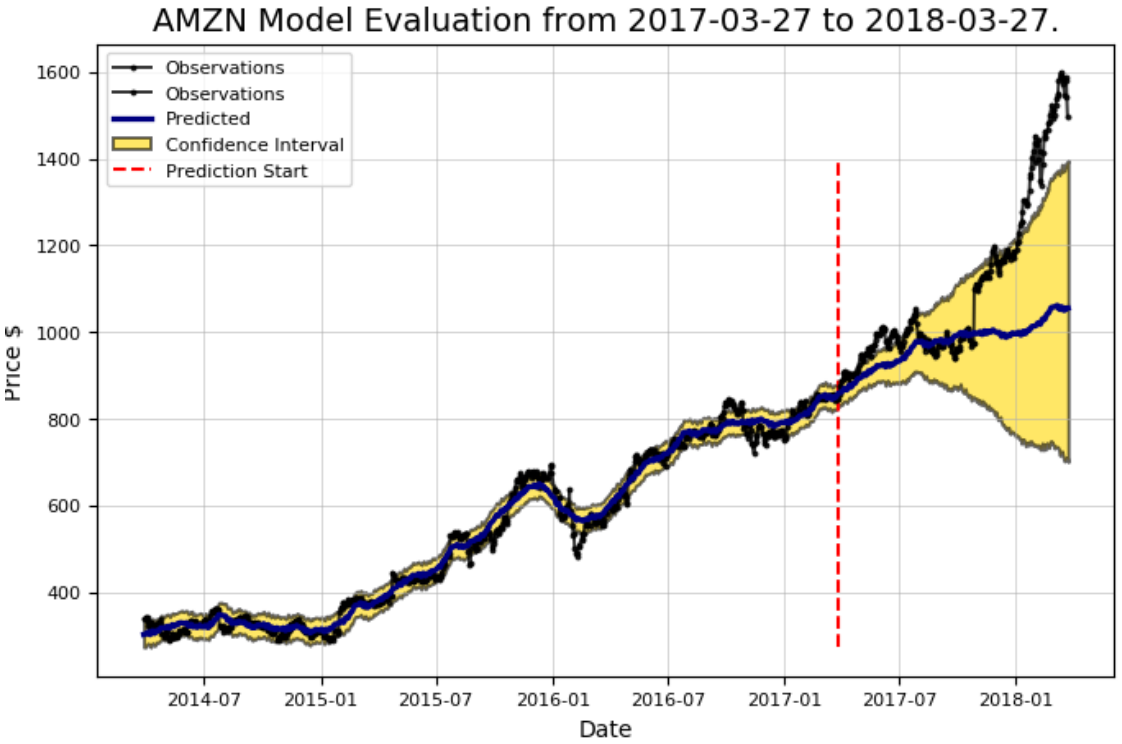
$\tau_{i,j}$ **is the pheromone level of the path**
**alpha here is the pheromone factor**
**beta is the visibility factor**

$(\tau_{i,j})^{\alpha}$ **is the pheromone feature**

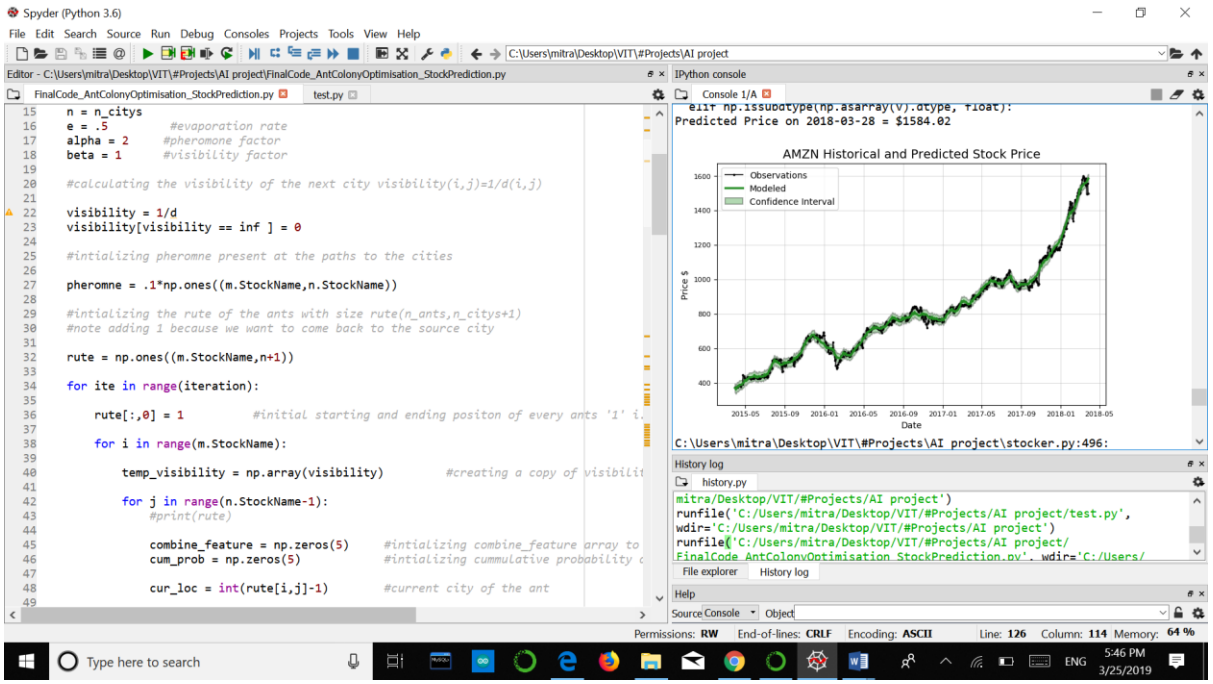
$(\eta_{i,j})^{\beta}$ **is the visibility feature (visibility = 1/path cost)**

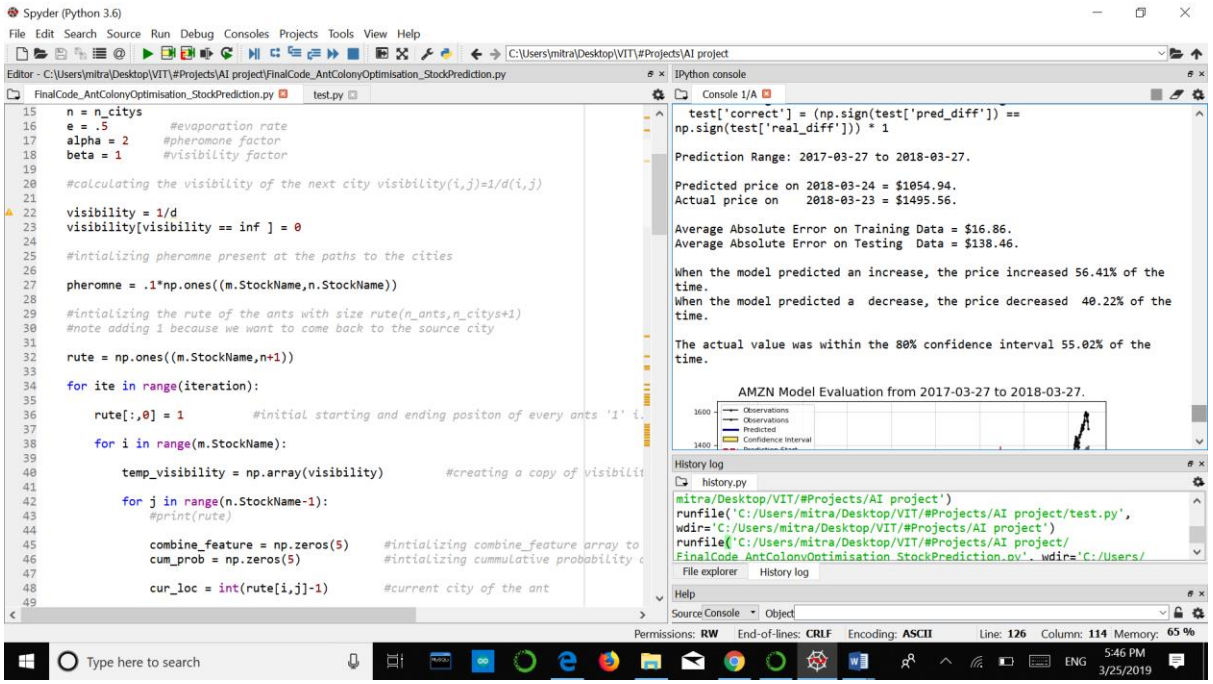
## GRAPHS:

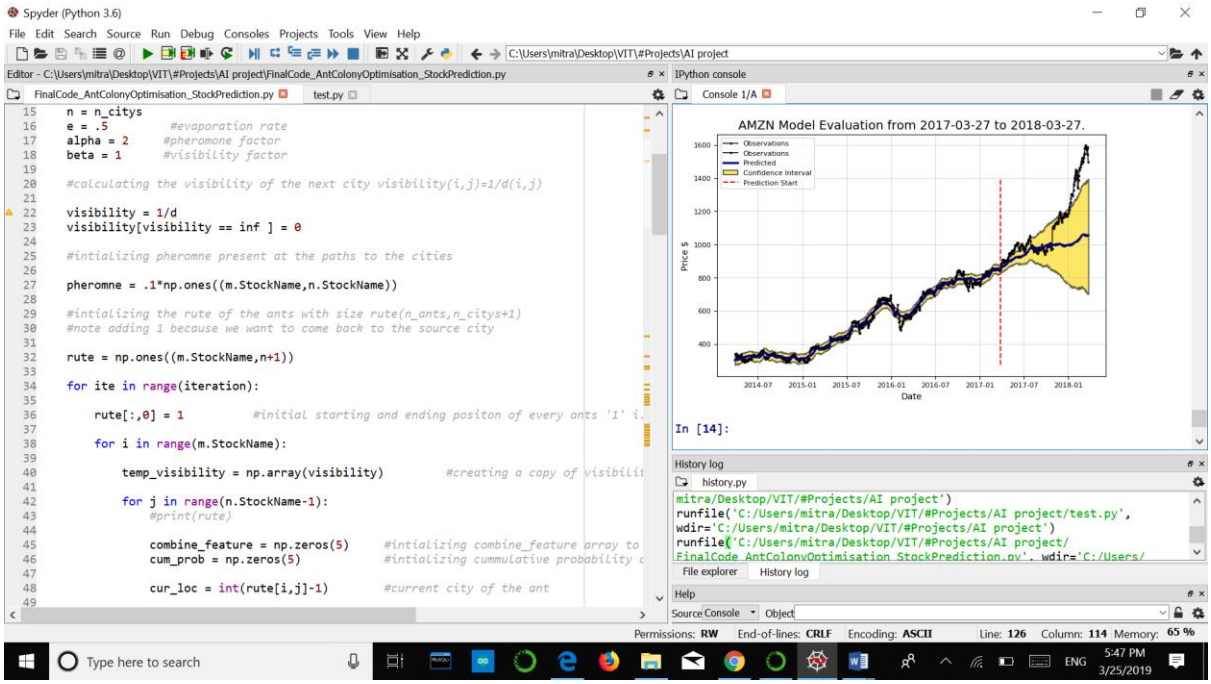

AMZN Historical and Predicted Stock Price
Observations
Modeled
Confidence Interval
Price $
1600
1400
1200
1000
800
600
400
2015-05
2015-09
2016-01
2016-05
2016-09
2017-01
2017-05
2017-09
2018-01
2018-05
Date


AMZN Model Evaluation from 2017-03-27 to 2018-03-27.
Observations
Observations
Predicted
Confidence Interval
Prediction Start
Price $
1600
1400
1200
1000
800
600
400
2014-07
2015-01
2015-07
2016-01
2016-07
2017-01
2017-07
2018-01
Date

# PROGRAM SCREENSHOTS:






## ACCURACY MEASURE:

Here we use a stocker module based confusion matrix with arguments placed as the start and end date from the existing database. This method gives us the confidence level of the data and the accuracy in the prediction of this said data.

```
Prediction Range: 2017-03-27 to 2018-03-27.

Predicted price on 2018-03-24 = $1054.94.
Actual price on    2018-03-23 = $1495.56.

Average Absolute Error on Training Data = $16.86.
Average Absolute Error on Testing  Data = $138.46.

When the model predicted an increase, the price increased 56.41% of the
time.
When the model predicted a  decrease, the price decreased  40.22% of the
time.

The actual value was within the 80% confidence interval 56.63% of the
time.
```

## CONCLUSION:

As we have seen in this project, the analysis of stocks using bio inspired algorithms like ACO has helped us achieve a quick solution to stock market prediction methodology. Further studies can be done to determine other impactful utilization of other similar bio inspired procedures or algorithms such as bacterial foraging algorithm or bee hive algorithm.

Other adjustments to the ACO algorithm may also give out better results to the user in whichever field it is used in. adjustments may come in the form of fine tuning the pheromone factors. Other changes in the system or algorithm may result in better and more optimised outcomes of future applications.