**Project Proposal**



**For**

**STOCK MARKET FORECASTING USING**

**TIME-SERIES ANALYSIS**

**Under guidance of**

**Professor. Varsha Nemade**

**MPSTME, NMIMS**

**Dipanshu Agarwal N201**

**Riya Airen N204**

**Saurabh Ajit N205**

**25th May, 2020**

**Synopsis**

Predicting the Stock Market has been the bane and goal of investors since its existence. Everyday billions of dollars are traded on the exchange, and behind each dollar is an investor hoping to profit in one way or another. Entire companies rise and fall daily based on the behaviour of the market. Should an investor be able to accurately predict market movements, it offers a tantalizing promises of wealth and influence. In the real world, the stock market predictions can be categorised in 2 parts, Fundamental Analysis and Technical Analysis.

In this undertaking, we will be creating a supervised machine learning model which will help us to somewhat predict the price value of stocks/security of a company i.e. **State Bank of India** to be specific. The Model will be using Time-Series Analysis, Time series is a set of observations or data points taken at specified time usually at equal intervals and it’s used to predict the **future**values based on the **previous**observed values.

**Goals**

The Chief goal of this project is to add to the academic understanding of stock market prediction. This project will focus exclusively on predicting the daily trend (price movement) of individual stocks. More so, the project will analyse the accuracies of these predictions.

**Methodology**

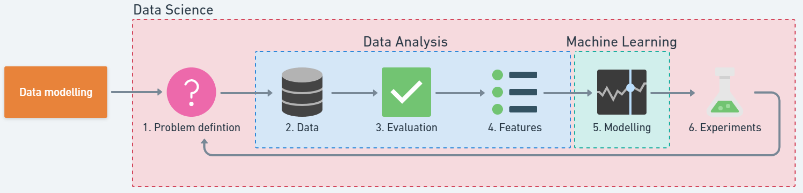


FIGURE 1: - STEPS IN A FULL MACHINE LEARNING PROJECT [1]

**Problem Definition: -** What business problem are we trying to solve? How can it be phrased as a machine learning problem?

🡪 Stock Market Price Prediction.

**Data: -** If machine learning is getting insights out of data, what data we have? Is our data structured or unstructured? Static or streaming?

🡪 Our Data will be structured in a CSV file, and it will be streaming (Ever-Changing) in nature. We will use 80% of the data collected to train our model and use the remaining 20% as Test Data.

**Evaluation: -** What defines success? Is a 95% accurate machine learning model good enough?

🡪 For this project to be successful, the model needs to be over 85% accurate at determining the price of the stock. That is, if the price of the stock is ₹ 100, the predicted price should not be less than ₹ 85 or more than ₹ 115.

**Features: -** What parts of our data are we going to use for our model?

🡪 We will be using the Opening Prices, Closing Prices, Daily High and Low Value of stock to train and test our model.

**Modelling: -** Which model should you choose? How do you compare it with other models?

🡪 We will be using ARIMA Model of the Time-Series analysis to build our project, The ARIMA model is a form of Regression analysis. An ARIMA model can be better understood by looking into its individual components: Auto regression (AR), Integrated (I) and Moving Averages (MA). In AR model, Partial Auto Correlation Function (PACF) graph is used to find P value and in MA model, Auto Correlation Function (ACF) graph to find q value. Integration Function is used to find the d value, i.e. the differentiation.

**Experiments: -** What else could we try? Does our deployed model do as we expected? How do the other steps change based on what we’ve found?

🡪 This step involves all the other steps. Because machine learning is a highly iterative process, we want to make sure our experiment is actionable.

**Schedule**

A time frame indicating steps that will be required and the expected date when they will be completed.

**04 May 2020 – 27 June 2020**

**Tasks: -**

1. Study of Basics of Time-Series Analysis and Fundamentals of Trading
2. Collecting of Dataset and Data-set Pre-Processing & Normalization
3. Study of Different Time-Series Models and Selection of Models
4. Study of the Mathematical concepts behind the model
5. Creation of Model using python and various python libraries as well as studying the various functions used in the model
6. Training, Testing and Validation of the Model and Calculation of Error Values
7. Final Project Report and Submission

**Reference**

1. A 6 Step Field Guide for Building Machine Learning Projects by Daniel Bourke
2. Introduction to Time Series Forecasting With Python by Jason Brownlee
3. Contreras, 1. Espinola, R.NogaJes, F1.and conejo,AJ.(2003) "ARIMA models to predict next day electricity prices", IFEE transactions on power system, vo1.18, noJ,pp: I 014-1 020.
4. Kumar; K Yadav;A.KSingh, M.P; Hassan and H.Jain,V.K(2004)"Forecasting Daily Maximum Surface Ozone".
5. Tsitsika,E.V;Maravelias,C.D& Haralatous,J. (2007)"Modelling and forecasting pelagic fish production using univariate and multivariate ARIMA models". Fisheries science volume 73,pp:979-988.
6. Datta K.(2011)"ARIMA forecasting of Inflation in the Bangladesh Economy",The IUP journal of bank management,voI.X,No.4,pp-7-15.
7. D. Banerjee, "Forecasting of Indian stock market using time-series ARIMA model," 2014 2nd International Conference on Business andInformation Management (ICBIM), Durgapur, 2014, pp. 131-135, doi: 10.1109/ICBIM.2014.6970973.
8. “Introduction to Time Series Analysis and Forecasting” by Douglas C. Montgomery, Cheryl L. Jennings, and Murat Kulahci
9. “Time Series Analysis: Forecasting and Control” by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung
10. Chris Chatfield, "The analysis of time series An introduction"