**Final Project: INFO 7390 Advances in Data Science**

**STOCK PRICE FORECASTING**



**TEAM 11**

Aditya Pawar

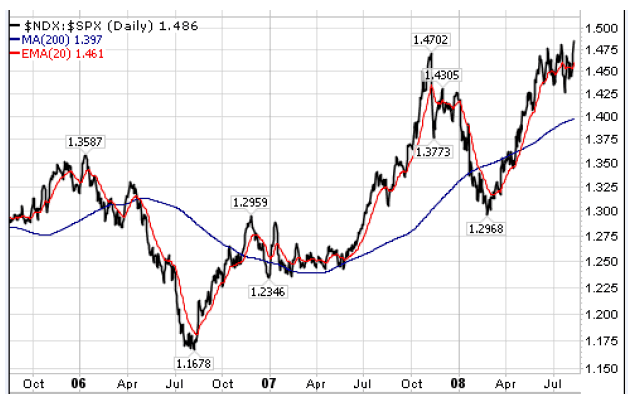
Dhairya Jaiswal

Ranga Chari

# OVERVIEW

Stock price forecasting is one of the most important area in today’s data science field as it helps in determining the future value of a company’s stock or other financial trade decisions. Depending on how well the prediction algorithms work it could yield a significant profit for the public/investors.

# 



**GOALS**

* To provide insightful decision-making trends for public/investors which would be profitable in the near future and long run.
* Approximate stock value for a particular company for a day.
* Factors affecting the high and low value of stock prices.
* Offering an interactive UI that lets the user read the description of a particular company’s data and also provides them future prediction values on a daily and weekly basis.

**USE CASES**

* Giving foresight to investors to make the right decisions and avoid making a loss.
* Knowing upfront highs and lows in stock trends.
* Up-trends, down-trends sideway moves of the market are some of the main features we plan to predict.

# 

# DATA

* **STOCK DATA** – Data source is attached and pushed to git hub repository.
* The variables (columns) of the dataset are as follows:

Date, Symbol, Open, Close, Low, high, Volume, etc.

**PROCESS OUTLINE**

* Data Preprocessing, Data Cleaning, handling missing values
* Exploratory Data Analysis
* Plot out the Normalized Close price and set last day close price as Y
* Study of time series approach and select the best models for prediction.
* Study of Unsupervised approaches like Rolling Linear regression, ARIMA(Auto Regressive Intensive Moving Average) .We will also try LSTM for prediction.
* Design of a pipeline and system to implement this approach and discussion on the system’s capabilities
* Deploy the Model on Azure/AWS or Google Cloud Computing Platform
* Build a web application to demonstrate the prediction results and detailed analysis of selected stock value.

## **DEPLOYMENT DETAILS:**

1. Language: Python
2. Pipeline: Airflow
3. Container: Docker
4. Cloud Tools/Platforms: Microsoft Azure Machine Learning Studio,AWS (Amazon WEb Services) EC2
5. Tools for Analysis: Microsoft Azure/ Visual Studio
6. Other Considerations: Google Cloud Platform

# Milestones

|  |  |
| --- | --- |
| **Timeframe** | **Delivery** |
|  |  |
| Day 1-2 | Data Preprocessing and Exploratory Data Analysis |
| Day 3-7 | Model Building ,Training ,Selection and selecting the best model |
| Day 8-9 | Deployment of model on cloud and building web application |
| Day 10-11 | System integration and documentation |