**PGP BDML CAPSTONE PROJECT – PROPOSAL (GROUP 4)**

1. **Domain and Context :**

Nowadays, as the connections between worldwide economies are tightened by globalization, external perturbations to the financial markets are no longer domestic. With evolving capital markets, more and more data is being created daily by various sources and machines.

The intrinsic value of a company’s stock is the value determined by estimating the expected future cash flows of which is determined by the company’s stock price. This market value of a stock can deviate from the intrinsic value due to reasons unrelated to the company’s fundamental operations, such as market sentiment.

The fluctuation of stock market is violent and there are many complicated financial indicators. Only few people with extensive experience and knowledge can understand the meaning of the indicators and use them to make good prediction to get fortune. Most people must rely solely on luck to earn money from stock trading. However, the advancement in technology and machine learning provides an opportunity to gain steady fortune from stock market and can help experts to find out the most informative indicators to make better prediction and trends. The prediction of the market value is of paramount importance to help in maximizing the profit of stock option purchase while keeping the risk as low as possible.

1. **Problem Statement:**

The aim of the project is to examine several different forecasting/regression techniques to predict future stock returns based on past returns and numerical news indicators to construct a portfolio of multiple stocks to diversify the risk of stock. We do this by applying supervised learning methods for stock price forecasting by interpreting the seemingly chaotic market data.

1. **Dataset(s):**

The data set is used from Kaggle <https://www.kaggle.com/ramamet4/nse-stocks-database#nifty50.csv> ( banknifty.csv, nifty50.csv)

|  |  |
| --- | --- |
| Field description | |
| Context | The NIFTY 50 index is National Stock Exchange of India's benchmark stock market index for Indian equity market. It is a well diversified 50 stock index accounting for 22 sectors of the economy. It is used for a variety of purposes such as bench-marking fund portfolios, index based derivatives and index funds. (nifty50.csv) |
| Bank Nifty represents the 12 most liquid and large capitalized stocks from the banking sector which trade on the National Stock Exchange (NSE). It provides investors and market intermediaries a benchmark that captures the capital market performance of Indian banking sector.( banknifty.csv) |
| Content | A data frame with 8 variables: index, date, time, open, high, low, close and id. For each year from 2013 to 2016, the number of trading data of each minute of given each date. The currency of the price is Indian Rupee (INR).   * index : market id * date: numerical value (Ex. 20121203- to be converted to 2012/12/03) * time: factor (Ex. 09:16) * open: numeric (opening price) * high: numeric (high price) * low: numeric (low price) * close: numeric (closing price) |