**Project name**

Predicting which stocks to buy based on financial fundamental data

**1.Business problem we are trying to solve**

**Motivation Behind Choosing this project**

There is too much information in the financial statements of the company and seeing each company's financial data (balance sheet, cash flow, income statements & financial ratios) before making investment decisions is a challenging task .

**How is this problem being currently solved?**

Big investor/ rating agencies have their extensive own research team that study each company’s financial statements in detail and figure out the best companies to invest in

Investors who can’t afford such detailed analysis , generally rely on few important financial metrics like Price/Earning ratio , EBITDA etc as a rule of thumb before making investment decisions. But such methods are only heuristics and not guaranteed to work.

**How are we trying to solve this problem effectively?**

We are trying to build a model which takes around 200+ financial indicators of a company & predicts if we as an investor should buy the stock or not. This would automate the whole investing process which currently is very research intensive or rely on guesswork as mentioned above

**Business Impact of this project**

Investors could use this tool to make informed investment decisions with ease. This would specially empower the small investor which can’t afford extensive research.

**2. Detailed Project Description**

It is well acknowledged that the buy & hold strategy is able to outperform many of the algorithmic strategies at least in the long run

Taking this insight forward, Based on the financial data of the previous year, we will try to predict by how much percent the price will change in the next full year

In the process of doing so, We would be able to find sector wise factors which are detrimental for the success of a company. This would further help the investors in making informed decisions.

**The problem statement consist of these three parts:**

**Regression Problem**

* 200+ financial indicators for the year 2015 (for eg);
* Target variable is the percent price variation (PRICE VAR [%] ) for the year 2016 (meaning from the first trading day on Jan 2016 to the last trading day on Dec 2016

**Classification Problem**

* for each stock, if the PRICE VAR [%] value is positive, class = 1. From a trading perspective, the 1 identifies those stocks that an hypothetical trader should BUY at the start of the year and sell at the end of the year for a profit.
* for each stock, if the PRICE VAR [%] value is negative, class = 0. From a trading perspective, the 0 identifies those stocks that an hypothetical trader should NOT BUY, since their value will decrease, meaning a loss of capital.

**Clustering problem**

The third-to-last column, Sector, lists the sector of each stock. We will perform per-sector analyses and comparisons