**PROBLEM DESCRIPTION**

Predicting stock market prices has been an age old problem in financial data analysis. Though there are several parameters that can be predicted for each share, for the purpose of this competition, we limit ourselves to predicting the open and close prices of fifty randomly chosen shares from the National Stock Exchange. The objective of the challenge is to utilize the data provided and come up with predictions of open and close prices for one week in advance.

This competition is being organized as a pre-runner to the Deloitte USI Analytics Summit, 2017 and will be active from 2nd May 2017. Top finishers on the leaderboard will be invited for a live data hackathon on the hackathon D-Day in one of the Deloitte campuses. The winners will be decided based on the results of the hackathon D-Day and final results announced during the Summit.

**DATA**

One consolidated file is available for download. This file contains:

|  |  |
| --- | --- |
| Symbol | Share name |
| Date | The trade date |
| Prev Close | Previous day’s closing rate of the share |
| Open Price | opening rate of the share for the trade date |
| High Price | Highest traded price for the trade date |
| Low Price | Lowest traded price for the trade date |
| Last Price | Last traded price for the trade date |
| Close Price | Closing price for the trade date |
| Average Price | Average traded price for the trade date |
| Total Traded Quantity | Total quantity of shares traded on the trade date |
| Turnover in Lacs | Turnover in Lacs (of rupees) on the trade date |
| Deliverable Qty | Number of shares delivered on the trade date |
| % Dly Qt to Traded Qty | Percentage of shares delivered as a percentage of total traded shares on the trade date |
|  |  |

1. The consolidated file contains information provided in the above table for 50 shares.
2. sample\_submission.csv: This is a sample submission file. Any submissions made should exactly be the same as this file. Some points to note:
   1. The file should be comma-separated ‘csv’ file
   2. The header should be ‘Date’, ‘Share’, ‘Open’, ‘Close’
   3. The column ‘share’ should be sorted in ascending order, from 1 to 50 (as seen in the sample submission file).
   4. The columns ‘Open’ and ‘Close’ can be any real number
   5. There should be a total of 101 rows (including the headers).

If any of these points are not followed, it will result in a submission error.

**PREDICTIONS**

Participants are to predict the opening and close prices of the given stocks for the 18th and 19th of May 2017. No data apart from the data provided can be used for the purpose of this challenge. Make special note of the fact that the data provided is till the 28th of April and the modelling has to be done only on data till the 28th of April. Participants cannot use data post the 28th of April to tune parameters of the model. It is compulsory to submit codes used for making the leaderboard submissions. Any hardcoding of parameters that look suspicious will invite disqualifications of the team. If at any point of time, the participating teams are found to be using data after the 28th of April, the organizers reserve all rights to disqualify these teams in spirit of the competition.

## EVALUATION CRITERIA

The models are being assessed on their performance on real-life data. The participants are allowed to use the provided dataset for modelling till the 17th of May. The submission portal would be open from the 2nd of May and closes on the 17th of May. Participants can make only one submission. No request for re-submissions will be entertained. On the 20th of May, a leaderboard would be active, which will show the rankings of the participating teams, calculated based on the Mean Absolute Percentage Error (MAPE) values of the submission, found from the real-life stock open and close prices for the 18th and 19th of May, 2017.

**SHORTLISTING**

Based on the leaderboard positions on the 20th of May, 2017, top teams will be invited for a live Hackathon D-Day to be held on the 22nd of May, 2017. Final results of the data hackathon will be announced during the awards ceremony of the analytics summit.