**Solution Sheet**

1. Which model have you used for stock price prediction? Explain your model.

As it is a supervised learning problem we used linear regression to predict the stock price.

First after importing packages and reading the Train\_dataset\_.csv file

We found the number of NaN values in each column, then we used box plot on each column to find outliers, then found the correlation between the attributes and also with target variable and the result was that stock price was independent of stock index. Following that the NaN values of each column are replaced by the mean of that column and the outliers in ‘Covid Impact(Beta)’ column were also replaced by the mean of the column.

Then the test data set was read. The outliers and NaN were removed by the same process used on the train data set. Then the model was trained and the predictions made were exported as ‘solution\_part1.csv’ file.

1. Which model have you used for Put-Call ratio Time series prediction? Explain your model.

We thought of applying a ML algorithm to find the put –call ratio time series prediction. But there were only 6 days of data was available and thought there could be huge variation in the result thus obtained. As there will huge variation in put call ratio in short time. We used the mean of previous 6 days of put call ratios and used it as the predicted put call ratio.

We trained the model to find the stock price in the same way as previous question. And replaced the Put-Call ratio column in the test set with the new values. The predictions made were exported as ‘solution\_part2.csv’ file.