**Solution Sheet**

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

Model is CatBoostRegressor with only *NAV* and *P/E Ratio* columns.

Why only 2 features?

*Stock Index* is unique in each row, and the rest of

the Categorical Columns *Index and Industry* have

one-to-one association with Numerical Columns

*General Index* and *Covid Impact (Beta)* respectively.

So, all 3 of them are ignored.

Correlation shows except *NAV* and *P/E Ratio*,

all Numerical Columns are insignificant.

[Feature Importance also reinforces the same.]

With more Domain understanding,

there is Future Scope of Feature Engineering.

Evaluation Metric is Mean Absolute Error as it is preffered for Stock Price Prediction. PyCaret and Random Forest with RandomizedSearchCV is also used for experimentation.

Output file is *output file 01.csv*.

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

Stock Price for 16-AUG can not be predicted.

Put-Call Ratio Time-Series can not be predicted as it has no trend or pattern. It has linear relation with the past date as follows:

11-AUG: 1.20 X 10-AUG - 0.30

12-AUG: 1.14 X 11-AUG + 0.02 = 1.368 X 10-AUG - 0.322

13-AUG: 1.26 X 12-AUG + 0.01 = 1.72368 X 10-AUG - 0.39572

14-AUG: 1.12 X 13-AUG - 0.20 = 1.9305216 X 10-AUG - 0.6432064

15-AUG: 1.13 X 14-AUG - 0.12 = 2.181489408 X 10-AUG - 0.846823232

Furthermore, even if Put-Call Ratio had a pattern and we were able to predict it for the following day, 16-AUG, Correlation Matrix shows Put-Call Ratio does not influences Stock Price (at least not significantly enough to be used for prediction.) And, the two biggest predictors of Stock Price are Net Asset Value and Price-Earnings Ratio which are also time dependent, i.e., they change daily. So, to predict Stock Price for 16-AUG, we need NAV and P/E Ratio of the day.