Flipr Competition.

This project we have done as a part of Flipr Machine Learning Hackethon Series.

Aim

The project has two objective, first one was to predict the stock price based on given features (which was just a regression problem), and the second was to do a time series prediction based on past data of put-call ratios and further based on this predicted value we had to predict the future stock price.

Things we did

As this was a competition, so I was provided with two data set, the first one was data of features responsible for stock price and second was put-call ratio data of 6 days and we need to predict the put-call ratio of 7^{th} day.

For the first part, using the co-relation matrix we filtered the features that had high co-relation value with the label and removed the features, which are not that much significant. Next we identified the features which had missing values, in which I removed the features having missing value more than 50 percent, rest features are imputed using imputer function from scikit learn. Going further with the pre-processing step, we implemented label encoder to those columns which had categorical variable but contains strings. We also used some visualization tools to analyse features.

Coming to the model, we have used Random Forest Regressor which we have used it with Grid SearchCrossValidation to identify the optimal hypermeters. We then trained the model and predicted the stock price of test dataset.

Moving on to the second part, we need to perform time series analysis. In this with the data given of 6 days put-call ratio, we applied ARIMA time series model to predict the value of 7^{th} day. We need to then use this predicted put-call ratio to predict the future stock price.

Evaluation

Evaluation part was done by Flipr itselt, based on the evaluation we got an assessment report, in which our model was given A+ for prediction Algorithm, A for Time-series Algorithm and A+ for accuracy.

Application

The time series prediction holds lots of applications in industry today, for example PSUs can use this for inventory management, groceries can use it product stock regulation, and financial and stock companies can use it for investment portfolio management.