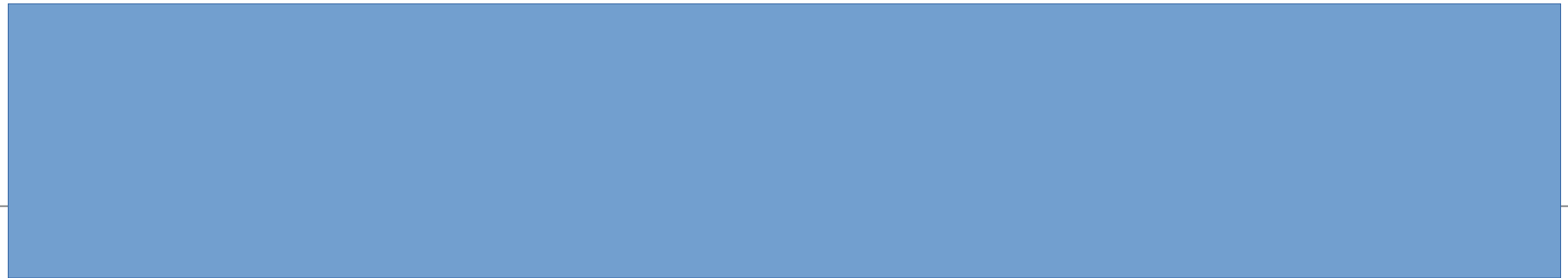
Machine Learning Documentation
Fun with Forecast

Hackerrank Handle:yodude
Institute:Indian Institute of Technology Guwahati
Name: Akul Agrawal
Roll No: 160101085

Attempt 1

Supervised Learning was used as described in the following points:

- The problem was changed into a classification problem. This was done by including the number of last entries to be used to predict as features to train and predict Usage. For example, for domain A, last 100 entries have to be used to predict next 20 entries. So, last 100 entries were used as features to predict next 20 entries.
- Note that number of components having A and B, and C[i] and C[ii] are different.
- Xgboost and GBM models were tried.



Time Series Analysis

- The data and trends were analyzed and different plots were made.
- Data for each domain: A, B, C[i], C[ii] were seperated and analyzed.
- AR, MA and ARIMA models were tried to be applied. For the same, the data was refined, for eg, the data was not stationary and it had to be made stationary to apply ARIMA model on it.

Strategy to decide final list

ARIMA Model

This strategy doesn’t make “extensive” use of the training data but analyzes trends in the test data as described in the problem statement (like last 100 entries to predict 20 entries in domain A).

- ARIMA Model is one of the models used to predict stock prices since it is also a time series problem.
- Data was processed and made stationary for this
- ARIMA model was trained on last entries to be used to predict entries of a domain and the entries were predicted from the trained model.