

REPORT

Preprocessing: As we go through analysis of data of any form we first need to process the data. Preprocessing of data helps around 60-70% of data analysis.

In this project we have been successfully in this task. We have categorized the training and the test sets, which prove to be helpful in model fitting.

Model Fitting: I have implemented poisson, negative binomial regression models and weighted regression for the models fitting. Amongst all I choose to go with cTree which closely relates to the value of prediction made from Google's Prediction API.

Future Work: We can go through the poisson and negative binomial model but first we need to perform subset selection methods to get a good result out of the predictable variable and use the appropriate regression models.

Conclusion: According to my prediction model I tend to believe that the predictions are good and related to Google's Prediction API. We can work to reduce the errors and build a better model by implanting the methods described in the .R script.