Imbalanced dataset, as we don’t have the capabilities to handle imbalanced datasets as of right now, we decided to use different metrics to represent the results better.

We need to use precision, recall, f1score

Implement a function to save the values into a file to make life easier

Before outlier removal, the categories of the outliers should be checked to ensure that categories with low sample number don’t get heavily affected by the removal of a few datasets.