Summary

In this assignment i have first done the eda and removed some of the useless columns and then found the outliers and removed them made the boxplot after that and after that i have used pca and mutual info and random forest importance to find the best 25 columns and after than i have used randomforest and logistic regression of which logistic worked best with pca columns