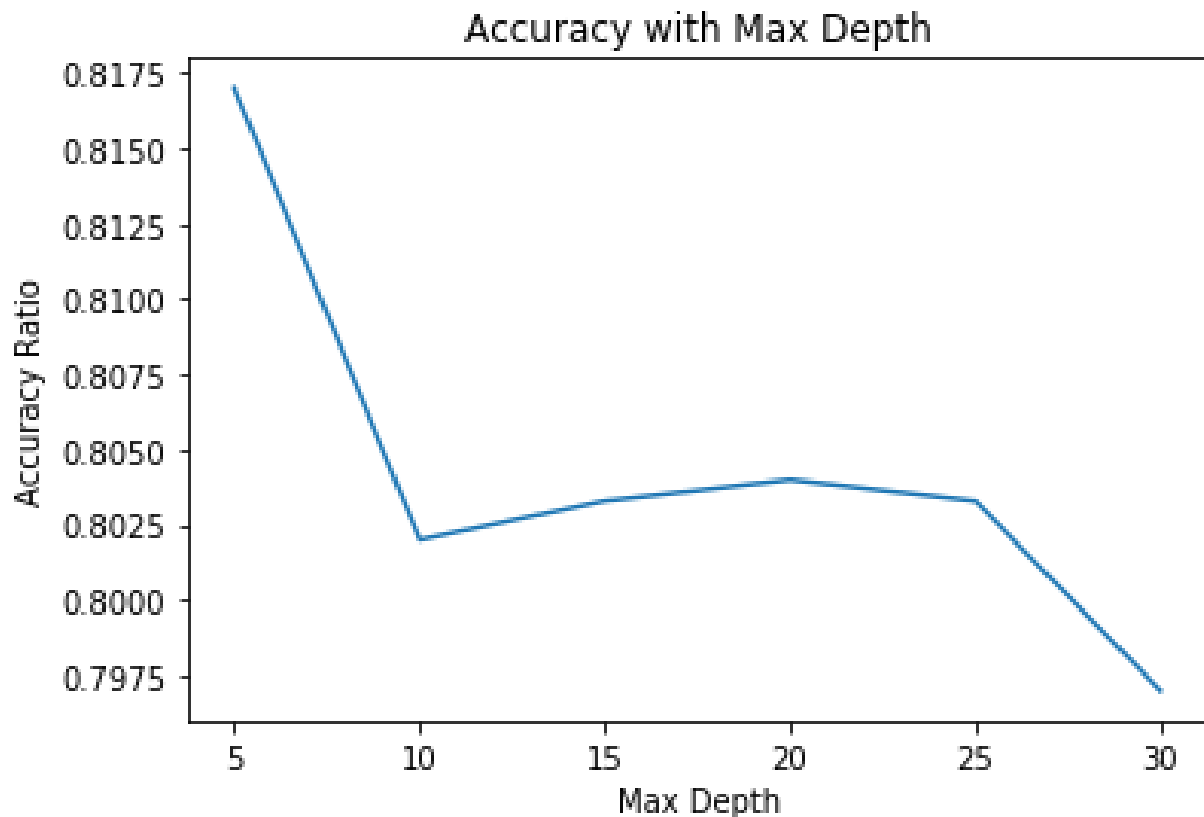# Decision Tree (CART Model)

For the purpose of this assignment, I first analyzed the dataset and found out that most of the columns were object field therefore for the purpose of decision tree, I decided to do a multi split for object fields and a binary split based on median value for integer field. I created a simple function to calculate the gini index at each split and then a split function that actually decides the best split. I also made a node class to actually store the nodes of the tree and predict from the given input data. I used multiple stopping criterion for terminal nodes such as, the number of examples being less than 5, ratio of true to false or false to true being more than 99 or the maximum depth being 5. I varied the maximum depth criterion and then tested for different max depth values before deciding a suitable one.

The model can further be improved by using some better pruning algorithm.