Problem statement

You are the data scientist at a telecom company whose customers are churning out to its competitors. You have to analyse the data of the company and find insights and probably stop your customers from churning to other telecom companies.

How I went about the project/Normal data lifecycle:

1. I started with the data preprocessing and data exploration part where I tried to understand the structure of the dataset
2. I visualized the dataset and tried to understand whatever has happening underneath it.
3. After understanding and comprehending the data properly, I went ahead and built my model.
4. I found out my independent variables and dependent variable.
5. Then went ahead and divided the dataset into training and testing set
6. Built model on training set and then checked the model accuracy on test set.
7. First I built a linear regression model, then I built a logistic regression model, then a decision tree classifier, and finally a random forest classifier.
8. Finally I calculated the accuracy of each of the models. (Root mean squared error for regression, confusion matrix, ROC, for classification models)