**Business Problem:**

Suppose we are interested in the factors that influence whether a political candidate wins an election. The outcome (response) variable is binary (0/1); win or lose. The predictor variables of interest are the amount of money spent on the campaign, the amount of time spent campaigning negatively and whether or not the candidate is an incumbent.

**Data:**

Data in the form of continuous and discrete (numeric) format.

**Pre-processing Data:**

All the data in the form of continuous and discrete (numeric) format there was no outlier but there is an observation which has only NA’s in the data. So deleted the complete observation. Delete the unused feature like ‘Election id’ from the processing.

**Model Building:**

The dataset contains the 3 features. Our goal is find out whether the political candidate win or not. End result is discrete. So I’m proceeding with Binomial Regression. Initially I’, trying with all features and building the model.

model <- glm(Result~.,data=claimants2,family = "binomial") =affairs1,family = "binomial")

**Summary of the model1:**

Null deviance: 1.3460e+01 on 9 degrees of freedom

Residual deviance: 6.5897e-10 on 6 degrees of freedom

AIC: 8

Null deviance and Residual deviance is almost zero and AIC 8.

Hence model3 is final model and we go for further calculation.

**Reference**

**Prediction**  **0 1**

**0** 4 0

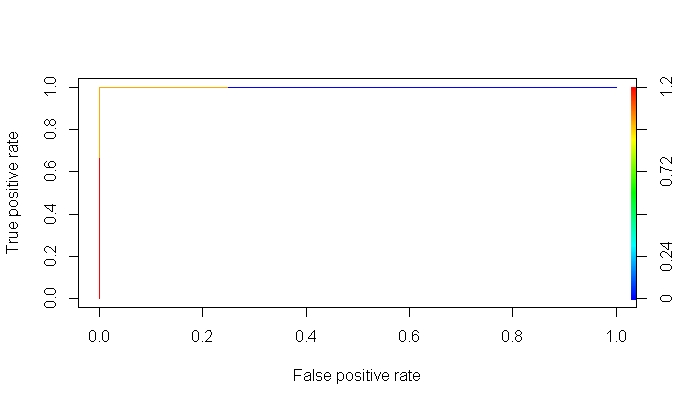
**1**  0 6

|  |  |
| --- | --- |
| True Positive | 6 |
| True Negative | 4 |
| False Positive | 0 |
| False negative | 0 |
| Accuracy | 100% |
| Sensitivity | 100% |
| Specificity | 100% |

**Conclusion:**

Political candidate who are having good popularity and little old, they are winning. Some times less popular persons also winning if they spent lot of money.

**Cut-off Value using ROC curve:**



As per the ROC Curve, more under the true positive rate and less false positive rate is good. So am per my understanding 90% is good.