**Binary logistic regression on Band Dataset**

step1-> check for na values in dataset

Step2->changed the below columns into numeric from categorical

factor columns

default

housing

loan

y

Setp3-> changed the below column into numeric from categorical

dummy columns

poutcome

contact

marital

job

Step4-> changed month column into numerical values like "jan" = 1,"feb"= 2

Step5-> applied glm model into dataset

model <- glm(bank\_filtered$y~.,data = bank\_filtered, family = 'binomial')

Step6-> eliminated non contributing column and again applied glm model

bank\_filtered$job.job.unknown <- NULL

bank\_filtered$marital.marital.single <- NULL

bank\_filtered$education.education.unknown <- NULL

bank\_filtered$contact.contact.unknown <- NULL

bank\_filtered$poutcome.poutcome.unknown <- NULL

got below results

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 32631 on 45210 degrees of freedom

Residual deviance: 22592 on 45178 degrees of freedom

AIC: 22658

Number of Fisher Scoring iterations: 6

accuray -> 0.9002455

Step7-> Applied PCA concept on data frame and consider those pc values in model making.

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 3.2631e+04 on 45210 degrees of freedom

Residual deviance: 2.6392e-07 on 45182 degrees of freedom

AIC: 58

Here you can find the Null deviance and AIC value is dramatically changed and lesser the value of AIC better the model is.

Accuracy-> 1

So the Second module is the best module.