**TITANIC DATA SET PREDICTION**

**BY- HARSH RAJ**

**Roll no-J011**

**After analyzing the probability of survival with respect to different factors the Highest output was achieved when we take 3 major factors that are**

1. SIBSP-Spouses aboard the Titanic( Individual with spouses had higher chances as one of them took care of other to survive)
2. Sex-
3. Class-Ticket Class- Individual With first class ticket had better chance of survival then that of lower class

Algorithm-

split<-sample.split(bdata$Survived,SplitRatio = 0.85)

split

train<-subset(bdata,split==TRUE)

test<-subset(bdata,split==FALSE)

dim(train)

dim(test)

train.fit<-naiveBayes(Survived~SibSp+Sex+Pclass,data=train)

P.train<-predict(train.fit,train)

table(P.train,train$Survived)

P.test<-predict(train.fit,test)

table(P.test,test$Survived)

**Data set was split into 85-15**

**Training data set had Accuracy of – 80.07916%**

**Test Data Set Had an Accuracy of -74.43%**