# Assignment 15

DataSet

a. Predict the no of comments in next H hrs

b. Use regression technique

c. Report the training accuracy and test accuracy

Sol:-

#a) & b)

#reading the dataset and viewing

slr <- read.csv(file.choose())

slr1<- slr

View(slr1)

#features

dim(slr1)

str(slr1)

library(psych)

describe(slr1)

summary(slr1)

#visualization

hist(slr1$Advt ,xlab = "advt", ylab = "Frequency",main="Histogram of advt",col="red")

hist(slr1$Sales ,xlab = "sales", ylab = "Frequency",main="Histogram of sales",col="blue")

plot(slr1$Advt,slr1$Sales)

#\*\*\*NOTE\*\*\*

#using linear regression model technique

#using slr1 dataset

#linear regression model

model<- lm(slr1$Advt~slr1$Sales)

model

#Important features

#multiple r squared value

#p value of slope test

#F stats

#predicting

Pred<- predict(lm(slr1$Sales~slr1$Advt))

Pred

pred<- predict(model,newdata= slr1Test,type = "response")

table(slr1$Advt,pred>= 0.5)

conf<- table(slr1$Advt,pred)

conf

predict(model)

Pred=predict(model)

slr1$predicted =NA

slr1$predicted =Pred

slr1$error =model$residuals

#verfify residuals

error<- residuals(lm(slr1$Sales~slr1$Advt))

error

summary(error)

#check and interpreting the summary

summary(model)

#\*\*NOTE\*\*

#Interpreting

#thus by multiple r squared value we see our model is good

#also our p value of slope test is <0.05 so good for our model

#adjusted r squared value is also good 0.891

#f stats value of 90.93 suggest our model is good and also its p value is <0.05

#our model accuracy is 0.9009 which is good

#result of all of our models

summary(model)

summary(model1)

summary(model2)

#model coefficients

model

model1

model2

slr1$coefficients<- NA

slr1$coefficients<- model$coefficients

slr1$coefficients

#c)

#test and training accuracy

#dataset slr1

set.seed(1)

split<- sample.split(slr1$Advt,SplitRatio = 0.70)

slr1Train <- subset(slr1,split == TRUE)

slr1Test<- subset(slr1, split == FALSE)

#train

model1<- lm(slr1Train$Advt~slr1Train$Sales)

model1

summary(model1)

#accuracy is 0.926

#test

model2<- lm(slr1Test$Advt~slr1Test$Sales)

model2

summary(model2)

#accuracy is 0.871