Gankhuyag, Khulan & Alghanim, Khatoon F. – Contribution

Decision Tree:

install.packages("rpart")

install.packages("rpart.plot")

install.packages("RWeka")

install.packages("C50")

install.packages("caret")

install.packages("e1071")

library(readr)

WineMag <- read.csv('C:\\Users\\gank7343\\Desktop\\WineMag.csv')

head(WineMag)

library(rpart)

library(rpart.plot)

library(RWeka)

library(C50)

install.packages("rpart")

install.packages("rpart.plot")

install.packages("RWeka")

install.packages("C50")

install.packages("caret")

install.packages("e1071")

library(readr)

WineMag <- read.csv('C:\\Users\\gank7343\\Desktop\\WineMag.csv')

head(WineMag)

library(rpart)

library(rpart.plot)

library(RWeka)

library(C50)

library(caret)

library(e1071)

install.packages("lattice")

install.packages("ggplot2")

str(WineMag)

S<- sample(97851,80000)

trainIndex = sample(1:97851, size = round(0.7\*97851), replace=FALSE)

train = WineMag[trainIndex ,]

test = WineMag[-trainIndex ,]

## remove Description and designation ###

df = subset(train, select = -c(Description, Designation))

## To convert the price into binary ###

train$Price[train$Price < 50] <- 0

train$Price[train$Price > 50] <- 1

train$Price[train$Price == 50] <- 1

##to change the coloumn position in DF file

setcolorder(df, c("Price", 1: (length(Df)-1) ) )

df$Price1<-df$Price

df2 = subset(df, select = -c(Price))

#### perform the test ##

X <- df2[ ,1:7]

y <- df2[,8]

trainX <- X[1:68496,]

trainy <- y[1:68496]

model <- C50::C5.0( trainX, trainy )

summary( model )