```
---
title: "1P22CS020_ML_ASSIGNMENT_3"
author: "Nisha Mohanraj"
date: "2023-10-02"
output: html_document
---
```

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

1.
```{r}
library(ISLR)
summary(Auto)
```

```{r}
attach(Auto)
```


```{r}
mpg01 <- ifelse( mpg > median(mpg), yes = 1, no = 0)
Auto <- data.frame(Auto, mpg01)
Auto
```

2.
```{r}
cor(Auto[,-9])
```

```{r}
#library(corrplot)
```

scatter plot
```{r}
pairs(Auto)
```

boxplot
```{r}
par(mfrow=c(2,3))
boxplot(cylinders ~ mpg01, data = Auto, main = "Cylinders vs mpg01")
boxplot(displacement ~ mpg01, data = Auto, main = "Displacement vs mpg01")
boxplot(horsepower ~ mpg01, data = Auto, main = "Horsepower vs mpg01")
boxplot(weight ~ mpg01, data = Auto, main = "Weight vs mpg01")
boxplot(acceleration ~ mpg01, data = Auto, main = "Acceleration vs mpg01")
boxplot(year ~ mpg01, data = Auto, main = "Year vs mpg01")
```


3.
```{r}
Auto <- data.frame(mpg01, apply(cbind(cylinders, weight, displacement, horsepower,
acceleration),
                         2, scale), year)

```


```{r}
train <-  (year %% 2 == 0)
train
```
```{r}
test<- !train
```


```{r}
Auto.train <-  Auto[train,]
```

```
Auto.test <-  Auto[test,]
mpg01.test <-  mpg01[test]
```

```{r}
mpg01.test
```
4.
```{r}
library(MASS)
```

```{r}
lda.fit <-  lda(mpg01 ~ cylinders + weight + displacement + horsepower,
             data = Auto, subset = train)
lda.pred <-  predict(lda.fit, Auto.test)
mean(lda.pred$class != mpg01.test)
```
5.
```{r}
qda.fit <- qda(mpg01 ~ cylinders + horsepower + weight + acceleration, data=Auto, subset =
train)
qda.fit
```
```{r}
qda.pred<-predict(qda.fit, Auto.test)
mean(qda.pred$class != mpg01.test)

#$class
#table(qda.class, testing_data$mpg01)
```
6.
```{r}
glm.model <- glm(mpg01 ~ cylinders + weight + displacement + horsepower, data =
Auto,subset=train, family = binomial)
summary(glm.model)
```

```{r}
glm.fit <-  glm(mpg01 ~ cylinders + weight + displacement + horsepower,
             data = Auto,
             family = binomial,
             subset = train)
glm.probs <-  predict(glm.fit, Auto.test, type = "response")
glm.pred <-  rep(0, length(glm.probs))
glm.pred[glm.probs > 0.5] <- 1
mean(glm.pred != mpg01.test)
```
7.
```{r}
library(class)
train.X <-  cbind(cylinders, weight, displacement, horsepower)[train,]
test.X <-  cbind(cylinders, weight, displacement, horsepower)[test,]
train.mpg01 <-  mpg01[train]
set.seed(1)

# KNN (k=1)
knn.pred <-  knn(train.X, test.X, train.mpg01, k = 1)
mean(knn.pred != mpg01.test)
```

```{r}
knn.pred <-  knn(train.X, test.X, train.mpg01, k = 10)
mean(knn.pred != mpg01.test)
```

```{r}
```

```
knn.pred <-  knn(train.X, test.X, train.mpg01, k = 100)
mean(knn.pred != mpg01.test)
```