

Name: Hari Yadav
Class: Into to Stat
Date : 02/22/2020

#lab3

#2.a

```
library(ISLR)
summary(Weekly)
cor(Weekly[, -9])
attach(Weekly)
plot(Volume)
```

#2.b

```
fit.glm <- glm(Direction ~ Lag1 + Lag2 + Lag3 + Lag4 + Lag5 + Volume,
data = Weekly, family = binomial)
summary(fit.glm)
```

#2.c

```
probs <- predict(fit.glm, type = "response")
pred.glm <- rep("Down", length(probs))
pred.glm[probs > 0.5] <- "Up"
table(pred.glm, Direction)
```

#2.d

```
train <- (Year < 2009)
Weekly.20092010 <- Weekly[!train, ]
Direction.20092010 <- Direction[!train]
fit.glm2 <- glm(Direction ~ Lag2, data = Weekly, family = binomial,
subset = train)
summary(fit.glm2)
```

```
probs2 <- predict(fit.glm2, Weekly.20092010, type = "response")
pred.glm2 <- rep("Down", length(probs2))
pred.glm2[probs2 > 0.5] <- "Up"
table(pred.glm2, Direction.20092010)
```

#2.e

```
library(MASS)
fit.lda <- lda(Direction ~ Lag2, data = Weekly, subset = train)
fit.lda
```

```
pred.lda <- predict(fit.lda, Weekly.20092010)
table(pred.lda$class, Direction.20092010)
```

#2.f

```
fit.qda <- qda(Direction ~ Lag2, data = Weekly, subset = train)
fit.qda
pred.qda <- predict(fit.qda, Weekly.20092010)
table(pred.qda$class, Direction.20092010)
```

#f.g

```
library(class)
train.X <- as.matrix(Lag2[train])
```

```

test.X <- as.matrix(Lag2[!train])
train.Direction <- Direction[train]
set.seed(1)
pred.knn <- knn(train.X, test.X, train.Direction, k = 1)
table(pred.knn, Direction.20092010)

#2.h.i
# Logistic regression with Lag2:Lag1
fit.glm3 <- glm(Direction ~ Lag2:Lag1, data = Weekly, family = binomial,
subset = train)
probs3 <- predict(fit.glm3, Weekly.20092010, type = "response")
pred.glm3 <- rep("Down", length(probs3))
pred.glm3[probs3 > 0.5] = "Up"
table(pred.glm3, Direction.20092010)

mean(pred.glm3 == Direction.20092010)

# LDA with Lag2 interaction with Lag1
fit.lda2 <- lda(Direction ~ Lag2:Lag1, data = Weekly, subset = train)
pred.lda2 <- predict(fit.lda2, Weekly.20092010)
mean(pred.lda2$class == Direction.20092010)

# QDA with sqrt(abs(Lag2))
fit.qda2 <- qda(Direction ~ Lag2 + sqrt(abs(Lag2)), data = Weekly, subset
= train)
pred.qda2 <- predict(fit.qda2, Weekly.20092010)
table(pred.qda2$class, Direction.20092010)

mean(pred.qda2$class == Direction.20092010)

# KNN k =10
pred.knn2 <- knn(train.X, test.X, train.Direction, k = 10)
table(pred.knn2, Direction.20092010)

mean(pred.knn2 == Direction.20092010)

# KNN k = 100
pred.knn3 <- knn(train.X, test.X, train.Direction, k = 100)
table(pred.knn3, Direction.20092010)

mean(pred.knn3 == Direction.20092010)

#3.a
attach(Auto)
mpg01 <- rep(0, length(mpg))
mpg01[mpg > median(mpg)] <- 1
Auto <- data.frame(Auto, mpg01)

#3.b
cor(Auto[, -9])

pairs(Auto)
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.c

```

train <- (year %% 2 == 0)
Auto.train <- Auto[train, ]
Auto.test <- Auto[!train, ]
mpg01.test <- mpg01[!train]

```

#3.d

```

fit.lda <- lda(mpg01 ~ cylinders + weight + displacement + horsepower,
data = Auto, subset = train)
fit.lda
pred.lda <- predict(fit.lda, Auto.test)
table(pred.lda$class, mpg01.test)
mean(pred.lda$class != mpg01.test)

```

#3.e

```

fit.qda <- qda(mpg01 ~ cylinders + weight + displacement + horsepower,
data = Auto, subset = train)
fit.qda
pred.qda <- predict(fit.qda, Auto.test)
table(pred.qda$class, mpg01.test)
mean(pred.qda$class != mpg01.test)

```

#3.f

```

fit.glm <- glm(mpg01 ~ cylinders + weight + displacement + horsepower,
data = Auto, family = binomial, subset = train)
summary(fit.glm)
probs <- predict(fit.glm, Auto.test, type = "response")
pred.glm <- rep(0, length(probs))
pred.glm[probs > 0.5] <- 1
table(pred.glm, mpg01.test)
mean(pred.glm != mpg01.test)

```

#3.g

```

train.X <- cbind(cylinders, weight, displacement, horsepower)[train, ]
test.X <- cbind(cylinders, weight, displacement, horsepower)[!train, ]
train.mpg01 <- mpg01[train]
set.seed(1)
pred.knn <- knn(train.X, test.X, train.mpg01, k = 1)
table(pred.knn, mpg01.test)
mean(pred.knn != mpg01.test)
pred.knn <- knn(train.X, test.X, train.mpg01, k = 10)
table(pred.knn, mpg01.test)
mean(pred.knn != mpg01.test)
pred.knn <- knn(train.X, test.X, train.mpg01, k = 100)
table(pred.knn, mpg01.test)
mean(pred.knn != mpg01.test)

```