Question 1

1. *R code*:

data <- read.csv("Weekly.csv")

data$Direction <- as.factor(data$Direction)

logit\_fit <- glm(data = data, Direction ~ Lag1 + Lag2 + Lag3 + Lag4 + Lag5 + Volume, family = binomial)

summary(logit\_fit)

*Output*:

Text

Description automatically generated with medium confidence

1. Based on the above results, we can only reject the null for for Lag2, with Lag2 positively correlating with likelihood of Direction being “Up”. (Note, the intercept is also statistically significant and nonzero, and non-negative. That is, without any predictors, Direction is more likely to be “Up” than “Down”.)
2. *R Code:*

Direction\_hat\_logit <- rep("Down", nrow(data))

Direction\_hat\_logit\_prob <- predict(logit\_fit, type = "response")

Direction\_hat\_logit[which(Direction\_hat\_logit\_prob >= 0.5)] <- "Up"

Direction\_hat\_logit <- as.factor(Direction\_hat\_logit)

table(Direction\_hat\_logit,data$Direction)

mean(Direction\_hat\_logit == data$Direction)

*Output*:

Text

Description automatically generated

1. *R Code:*

lda\_fit <- MASS::lda(data = data, Direction ~ Lag1 + Lag2 + Lag3 + Lag4 + Lag5 + Volume)

lda\_pred <- predict(lda\_fit, data)

Direction\_hat\_lda <- lda\_pred$class

table(Direction\_hat\_lda,data$Direction)

mean(Direction\_hat\_lda == data$Direction)

*Output:*

Text

Description automatically generated with medium confidence

Both LDA and Logit perform equally in terms of overall accuracy. Though Logit is *slightly* more sensitive, while LDA is *slightly* more specific.

Can be confirm by outputs of the following code:

install.packages("caret")

library(caret)

confusionMatrix(data = as.factor(Direction\_hat\_logit), reference = data$Direction)

confusionMatrix(data = as.factor(Direction\_hat\_lda), reference = data$Direction)

Question 2

Text, letter

Description automatically generated

A piece of paper with writing on it

Description automatically generated

1. *R Code:*

#clear environment

rm(list=ls())

#read data, set column names, code outcomes as factors

logit\_data <- read.csv("logit\_data.csv", header = FALSE)

colnames(logit\_data) <- c("A","study\_hours")

#assignment of initial variables

y <- logit\_data$A

x <- logit\_data$study\_hours

alpha <- 0.001

b\_last <- matrix(ncol=2,nrow=1,data=0)

colnames(b\_last) <- c("b0","b1")

b <- matrix(ncol=2,nrow=1,data=0)

colnames(b\_last) <- c("b0","b1")

err <- 100

eps <- 10^-4

b\_history <- matrix(ncol=2, data = 0)

colnames(b\_history) <- c("b0","b1")

grad <- matrix(ncol=1,nrow=2)

#calculation of grad descent

while(err > eps) {

grad[1,] <- sum(y-exp(b\_last[,1]+b\_last[,2]\*x)/(1+exp(b\_last[,1]+b\_last[,2]\*x)))

grad[2,] <- sum(x\*y-x\*exp(b\_last[,1]+b\_last[,2]\*x)/(1+exp(b\_last[,1]+b\_last[,2]\*x)))

b[,1] = b\_last[,1] + alpha\*grad[1,]

b[,2] = b\_last[,2] + alpha\*grad[2,]

err\_temp <- abs(b - b\_last)

err <- sqrt(err\_temp[,1]^2+err\_temp[,2]^2)

b\_last <- b

b\_history <- rbind(b\_history, b\_last)

}

#print final betas

cat("b\_0 =",b\_last[,1])

cat("b\_1 =",b\_last[,2])

#plot betas

plot(b\_history[,1], xlab = "Iterations", ylab = "beta\_0")

plot(b\_history[,2], xlab = "Iterations", ylab = "beta\_1")

*Output:*

Text

Description automatically generated

A picture containing shape

Description automatically generated

Chart

Description automatically generated

1. *R Code:*

b\_0 <- b\_last[,1]

b\_1 <- b\_last[,2]

x\_input <- 5

prob\_a <- exp(b\_0+b\_1\*x\_input)/(1+exp(b\_0+b\_1\*x\_input))

prob\_not\_a <- 1-prob\_a

prob\_not\_a

*Output:*

The probability of not getting an A with 5 study hours per week is 0.3131283.