Assignment 3 MATH 208 (Question 1)

Prsa Yadollahi
260869949
MATH 208 - Assignment 3

(a)

```r
logistic_Regression = function(theta, x1, x2) {
  p = 1/(1+exp(-x1*theta[1] - x2*theta[2] - theta[3]))
  return(p)
}
theta = c(pi, pi/2, pi/2)
x1 = c(1,2,3)
x2 = c(4,5,6)
logistic_Regression(theta, x1, x2)
```

```
## [1] 0.9999832 0.9999998 1.0000000
```

(b)

```r
cross_entropy_loss = function(theta, x1, x2, y){
  p = logistic_Regression(theta, x1, x2)
  log_reg = 0
  # Negative sum from 1 to n
  for (i in seq_along(y)){
    single_pass =
        y[i]*log(p[i]) +
        (1-y[i])*log(1-p[i])
    log_reg = log_reg - single_pass
  }
  return(log_reg)
}
theta = c(pi, pi/2, pi/2)
x1 = c(1,2,3)
x2 = c(4,5,6)
y = c(3, 2, 1)
cross_entropy_loss(theta, x1, x2, y)
```

```
## [1] -37.69909
```

(c)

```r
loss_funct = function(col1,col2){
  x1_d = c(HTRU2[[col1]])
  x2_d = c(HTRU2[[col2]])
  y_d = c(HTRU2$Class)
  result = optim(
            par=c(0,0,0),
            fn=cross_entropy_loss,
            x1=x1_d,
            x2=x2_d,
            y=y_d
          )
}

result = loss_funct(1,5)
result
```

```
## $par
## [1] -0.10569326  0.01629013  7.28979911
##
## $value
## [1] 1991.015
##
## $counts
## function gradient
##      218       NA
##
## $convergence
## [1] 0
##
## $message
## NULL
```

```r
# Thetas = -0.10569326 0.01629013 7.28979911
# Value = 1991.015
```

(d)

```r
var_combs = combn(names(HTRU2[,-9]),2)
dim(var_combs)
```

```
## [1]  2 28
```

```r
result = NULL
for(i in seq_along(names(HTRU2[,-8]))){
  for(j in 1:8){
    if(j <= i){
      next
    }
    if(i == 8 & j == 8){
      break
    }
    cross_entropy = loss_funct(i,j)
    table = tibble(
      "column 1" = names(HTRU2)[i],
      "column 2" = names(HTRU2)[j],
      "cross entropy loss" = cross_entropy$value
    )
    result = bind_rows(table,result)
  }
}
kable(result[order(result$‘cross entropy loss‘),])
```

column 1	column 2	cross entropy loss
EK_IP	SD_DMSNR	1427.745
EK_IP	EK_DMSNR	1429.591
EK_IP	SKW_DMSNR	1434.257
EK_IP	SKW_IP	1450.829
Mean_IP	EK_IP	1483.505
SD_IP	EK_IP	1490.764
EK_IP	Mean_DMSNR	1502.008
Mean_IP	SD_DMSNR	1759.214
Mean_IP	EK_DMSNR	1763.425
Mean_IP	SKW_DMSNR	1790.573
SKW_IP	SD_DMSNR	1834.243
SKW_IP	EK_DMSNR	1839.221
SKW_IP	SKW_DMSNR	1875.364
Mean_IP	SKW_IP	1918.023
Mean_IP	Mean_DMSNR	1991.015
SKW_IP	Mean_DMSNR	2021.685
Mean_IP	SD_IP	2052.101
SD_IP	SKW_IP	2305.642
SD_IP	EK_DMSNR	2777.460
SD_IP	SD_DMSNR	2877.531
SD_IP	SKW_DMSNR	2953.056
SD_IP	Mean_DMSNR	3365.135
Mean_DMSNR	EK_DMSNR	3772.916
SD_DMSNR	SKW_DMSNR	3800.222
EK_DMSNR	SKW_DMSNR	3808.527
SD_DMSNR	EK_DMSNR	3809.508
Mean_DMSNR	SKW_DMSNR	3869.097
Mean_DMSNR	SD_DMSNR	3971.733

(e)

```
var_combs = combn(names(HTRU2[,-9]),2)

loss_funct2 = function(columns){
  x1_d = c(HTRU2[[toString(columns[[1]])]])
  x2_d = c(HTRU2[[toString(columns[[2]])]])
  y_d = c(HTRU2$Class)

  result = optim(
              par=c(0,0,0),
```

```r
              fn=cross_entropy_loss,
              x1 = x1_d,
              x2 = x2_d,
              y = y_d
            )

  table = tibble(
    "column 1" = columns[[1]],
    "column 2" = columns[[2]],
    "cross entropy loss" = result$value
  )
  return(table)
}
end_result = map_dfr(as.data.frame(var_combs), loss_funct2)
end_result = end_result[order(end_result$‘cross entropy loss‘),]
kable(end_result)
```

column 1	column 2	cross entropy loss
EK_IP	SD_DMSNR	1427.745
EK_IP	EK_DMSNR	1429.591
EK_IP	SKW_DMSNR	1434.257
EK_IP	SKW_IP	1450.829
Mean_IP	EK_IP	1483.505
SD_IP	EK_IP	1490.764
EK_IP	Mean_DMSNR	1502.008
Mean_IP	SD_DMSNR	1759.214
Mean_IP	EK_DMSNR	1763.425
Mean_IP	SKW_DMSNR	1790.573
SKW_IP	SD_DMSNR	1834.243
SKW_IP	EK_DMSNR	1839.221
SKW_IP	SKW_DMSNR	1875.364
Mean_IP	SKW_IP	1918.023
Mean_IP	Mean_DMSNR	1991.015
SKW_IP	Mean_DMSNR	2021.685
Mean_IP	SD_IP	2052.101
SD_IP	SKW_IP	2305.642
SD_IP	EK_DMSNR	2777.460
SD_IP	SD_DMSNR	2877.531
SD_IP	SKW_DMSNR	2953.056
SD_IP	Mean_DMSNR	3365.135
Mean_DMSNR	EK_DMSNR	3772.916
SD_DMSNR	SKW_DMSNR	3800.222
EK_DMSNR	SKW_DMSNR	3808.527
SD_DMSNR	EK_DMSNR	3809.508
Mean_DMSNR	SKW_DMSNR	3869.097
Mean_DMSNR	SD_DMSNR	3971.733