CSE 635, Spring 2021, Homework 7

Sima Shafaei

Process “cardiotocographic.txt“ to develop a multinomial model using training subset to predict the dependent variable NSP as a function of the independent variables: {LB, AC, FM, UC}

Use the instruction set.seed(777) to split the data set into training and testing subsets with probabilities = 0.6, 0.4 respectively.

Code:

```
ctg=read.table("CTG_Dataset.txt",header = TRUE)
ctg=ctg[,c("LB","AC","FM","UC","NSP")]
set.seed(777)
ind=sample(2,nrow(ctg), replace=TRUE, prob = c(0.6,0.4))
training= ctg[ind==1,]
testing= ctg[ind==2,]
nrow(ctg)
nrow(training)
nrow(testing)

head(training)
head(testing)
```

Results:

nrow(ctg)	**2126**
nrow(training)	**1314**
nrow(testing)	**812**

Develop a multinom model m using the training subset. List the coefficients of m. Use the instruction predict(m) to predict NSP. Then compute the confusion matrix and the accuracy ACC

Code:

```
library(nnet)
m=multinom(training$NSP~training$LB+training$AC+training$FM+training$UC)
s=summary(m)
c=s$coefficients
c
fitted.values(m)
p=predict(m)  #actual prediction

t=table(training$NSP,p)
t
acc=sum(diag(t))/sum(t)
acc
```

Results:

Coefficients of model:

	(Intercept)	LB	AC	FM	UC
2	-15.4971450	0.11212469	-1.1527784	0.005414034	-0.2129547
3	0.5754737	-0.01386717	-0.8768855	0.008730868	-0.0084898

Confusion matrix:

	1	2	3
1	978	38	0
2	89	95	1
3	88	18	7

Accuracy: 0.8219178

Follow Lecture 10 to construct matrix X for the testing data subset. Then use coefficients of model m to compute the responses of the model. Compute the probabilities and predictions using matrix X. Finally find the confusion matrix and accuracy ACC

Code:

```
X=cbind(rep(1,nrow(testing)),testing[,-5])
head(X)
class(X)
X=as.matrix(X)
class(X)
Y=X%*%t(c)
head(Y)
Y=exp(Y)
head(Y)

pb=cbind(1/(1+Y[,1]+Y[,2]),Y[,1]/(1+Y[,1]+Y[,2]),Y[,2]/(1+Y[,1]+Y[,2]))
head(pb)
p=which.max(pb[1,])
for (i in 2:nrow(testing)) {p=c(p, which.max(pb[i,]))}
p
t=table(testing$NSP,p)
t
sum(diag(t))/sum(t)
```

Results:

Confusion matrix:

	1	2	3
1	611	28	0
2	55	53	2
3	49	12	2

Accuracy: 0.82019