

# MAST90138 Assignment 3

Haonan Zhong

2022-09-18

## Question 1a

```
library(MASS)
suppressMessages(library(pls))
XGtrain <- read.table("XGtrainRain.txt", header = TRUE, sep = ",")
XGtest <- read.table("XGtestRain.txt", header = TRUE, sep = ",")

# Fitting the quadratic discriminant model
# qda_model <- qda(G~., data = XGtrain)
# Fitting the logistic regression model
logistic <- glm(G~., data = XGtrain, family = binomial(link = "logit"))

## Warning: glm.fit: algorithm did not converge

# summary(logistic)
```

When attempting to fit the quadratic discriminant model with all the  $p$  predictors in the training set, an error was shown that some groups are too small for fitting, and from the summary of the logistic regression fitted with all  $p$  predictors, the model is overfitted as it has zero degrees of freedom. And only 149 of the 365 explanatory variables were used to fit the model, mainly because we only have 150 instances in the training set, thus there are insufficient degrees of freedom to fit all  $p$  predictors. Therefore, it is not recommended to use these two classifier on the test set.