**DATA SCIENCE PROJECT 2 (MAST90107)   
Semester 2 2021**

**GROUP 17**

Meeting No. 3

**Date & Time:** 24th August 2021, 6:00PM - 6:PM

**Agenda:** Catch up with Supervisor

**Attendees:** Kartika Waluyo, Vrinda Rajanahally, Ziad Al Bkhetan

**Discussion:**

1. Understanding the Input, Output, Parameters for autoencoders
2. Mathematical rigour and what it could mean for our project

**Next steps:**

1. Try implementing this code for XGB:

train <- as.data.frame(allMol\_data)

test <- as.data.frame(allMol\_data)

nrow(train)

ncol(train)

models <- list()

gene\_id <- 1

for (gene\_id in 1:3)

{

  y <- train[, gene\_id]

  x <- train[, setdiff(1:length(train), c(gene\_id))]

  models[[gene\_id]] <- randomForest(y = y, x = x)

  my\_pred <- predict(models[[gene\_id]],

                     newdata = test[, setdiff(1:length(train), c(gene\_id))])

  print(cor(my\_pred, test[, gene\_id]))

}

1. Send Ziad the Friday meeting email