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

**GROUP 17**

Meeting No. 2

**Date & Time:** Tuesday 10th August 2021, 6pm - 6.30pm

**Agenda:** Discuss the error from autokeras

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

**Discussion:**

1. XGBoost is not a random forest, it’s a decision tree
2. Use this for sampling for training and test set

library(caret)

letters <- read.table("letter-recognition.data.txt", sep=",")[1:2000,]

head(letters)

indxTrain <- createDataPartition(y = letters$V1, p = 0.75)

lettersTrain <- letters[indxTrain$Resample1,]

lettersTest <- letters[-indxTrain$Resample1,]

**Next steps:**

1. Ziad will take a look at the code and try to help figure out what causes the error
2. Ziad will read our semester 1 report and give feedback
3. Send the Rdata as csv to Ziad (kartika)
4. Send Friday’s meeting link to Ziad