#Importation de la table

Challenge <- read.csv(file="/Users/diao028606/Documents/2022/UC/Challenge QoE/Base/V2/challenge\_qoe\_day4.csv", sep = ",", header = TRUE, dec=";")

#suppression des variables unitiles

Challenge$qoe\_challenge\_table\_for\_model\_v2.year <- NULL

Challenge$qoe\_challenge\_table\_for\_model\_v2.month <- NULL

Challenge$qoe\_challenge\_table\_for\_model\_v2.imsi <- NULL

#mise à jour des colonnes

Challenge$msisdn <- Challenge$qoe\_challenge\_table\_for\_model\_v2.msisdn

Challenge$qoe\_challenge\_table\_for\_model\_v2.msisdn <- NULL

Challenge$coupure <- Challenge$qoe\_challenge\_table\_for\_model\_v2.coupure\_appel\_nb\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.coupure\_appel\_nb\_daily <- NULL

Challenge$nb\_hour\_total\_coupure <- Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_hour\_coupure\_appel\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_hour\_coupure\_appel\_daily <- NULL

Challenge$nb\_echec <- Challenge$qoe\_challenge\_table\_for\_model\_v2.echec\_appel\_nb\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.echec\_appel\_nb\_daily <- NULL

Challenge$nb\_hour\_echec <- Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_hour\_echec\_appel\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_hour\_echec\_appel\_daily <- NULL

Challenge$paging\_without\_response\_nb <- Challenge$qoe\_challenge\_table\_for\_model\_v2.paging\_without\_response\_nb\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.paging\_without\_response\_nb\_daily <- NULL

Challenge$nb\_hour\_paging\_without\_response <- Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_hour\_paging\_without\_response\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_hour\_paging\_without\_response\_daily <- NULL

Challenge$nb\_call\_est <- Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_call\_est\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_call\_est\_daily <- NULL

Challenge$reiter\_call\_interval\_nb <- Challenge$qoe\_challenge\_table\_for\_model\_v2.reiter\_call\_interval\_nb\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.reiter\_call\_interval\_nb\_daily <- NULL

Challenge$nb\_hour\_reiter\_call <- Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_hour\_reiter\_call\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_hour\_reiter\_call\_daily <- NULL

Challenge$call\_est\_bad\_setupduration\_nb <- Challenge$qoe\_challenge\_table\_for\_model\_v2.call\_est\_bad\_setupduration\_nb\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.call\_est\_bad\_setupduration\_nb\_daily <- NULL

Challenge$nb\_hour\_call\_est\_bad\_setupduration <- Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_hour\_call\_est\_bad\_setupduration\_daily

Challenge$qoe\_challenge\_table\_for\_model\_v2.nb\_hour\_call\_est\_bad\_setupduration\_daily <- NULL

Challenge$qoe\_challenge\_table\_for\_model\_v2.day <- NULL

#suppression des ligne avec un numéro vide

Challenge$test <- ifelse(is.na(Challenge$msisdn),0,Challenge$msisdn)

Challenge <- Challenge[!(Challenge$test == 0),]

#préparation de la matrice de travail pour faire le modèle non supervisé

write.table(Challenge\_ech, file = "/Users/diao028606/Documents/2022/UC/Challenge QoE/Base/V2/table\_matrice\_challengeQOE.csv",row.names=FALSE,col.names=TRUE, sep=";")

#importation des résultats avec les classes

Challenge\_kmeans <- read.csv(file="/Users/diao028606/Documents/2022/UC/Challenge QoE/Base/result\_kmeans\_Challenge\_QoEV2.csv", sep = ",", header = TRUE, dec=";")

# selection des classes extrêmes suite résultats Kmeans

Cluster\_normal <- Challenge\_kmeans[(Challenge\_kmeans$Cluster == 0),]

Cluster\_probleme <- Challenge\_kmeans[(Challenge\_kmeans$Cluster == 2),]

Base\_model2 <- rbind(Cluster\_normal,Cluster\_probleme)

Base\_model2$Cible <- ifelse(Base\_model2$Cluster==2 ,1,0)

#importation de la base pour effectuer le scoring

write.table(Base\_model2, file = "/Users/diao028606/Documents/2022/UC/Challenge QoE/Base/V2/table\_matrice\_challengeQOEV2.csv",row.names=FALSE,col.names=TRUE, sep=";")

# construction de la base d'apprentissage

#résultats from python

0.6439 \* coupure

+ 1.5524 \* nb\_hour\_total\_coupure

+ 0.5181 \* nb\_echec

+ 1.1697 \* nb\_hour\_echec

+ 0.7126 \* paging\_without\_response\_nb

+ 0.7246 \* nb\_hour\_paging\_without\_response

+ 0.2972 \* reiter\_call\_interval\_nb

+ 0.5476 \* nb\_hour\_reiter\_call

+ 1.4551 \* call\_est\_bad\_setupduration\_nb

+ 2.1038 \* nb\_hour\_call\_est\_bad\_setupduration

+ [-33.50265418] (intercept)

#impélementation du modèle

Challenge\_kmeans$pred <- 0.6439\*Challenge\_kmeans$coupure +

1.5524\*Challenge\_kmeans$nb\_hour\_total\_coupure + 0.5181\*Challenge\_kmeans$nb\_echec +

1.1697 \* Challenge\_kmeans$nb\_hour\_echec +

0.7126 \* Challenge\_kmeans$paging\_without\_response\_nb +

0.7246 \* Challenge\_kmeans$nb\_hour\_paging\_without\_response +

0.2972 \* Challenge\_kmeans$reiter\_call\_interval\_nb +

0.5476 \* Challenge\_kmeans$nb\_hour\_reiter\_call +

1.4551 \* Challenge\_kmeans$call\_est\_bad\_setupduration\_nb +

2.1038 \* Challenge\_kmeans$nb\_hour\_call\_est\_bad\_setupduration -33.50265418

Challenge\_kmeans$pred1 <- exp(Challenge\_kmeans$pred)

Challenge\_kmeans$predf <- 100-round((Challenge\_kmeans$pred1 / (1+ Challenge\_kmeans$pred1))\*100,2)

Challenge\_kmeans$test <- 1

test1 <-aggregate(data.frame(client = Challenge\_kmeans$test,

nb\_cause = as.numeric(as.character(Challenge\_kmeans$nb\_jour\_total\_cause)),

nb\_paging = as.numeric(as.character(Challenge\_kmeans$nb\_jour\_paging\_without\_response)),

nb\_reiter\_call = as.numeric(as.character(Challenge\_kmeans$nb\_jour\_reiter\_call)),

nb\_setupduration = as.numeric(as.character(Challenge\_kmeans$nb\_jour\_call\_est\_bad\_setupduration)),

top\_cause = as.numeric(as.character(Challenge\_kmeans$top\_cause)),

top\_paging = as.numeric(as.character(Challenge\_kmeans$top\_paging)),

top\_reiter = as.numeric(as.character(Challenge\_kmeans$top\_reiter)),

top\_setupduration = as.numeric(as.character(Challenge\_kmeans$top\_setupduration))),

by = list(Cluster =Challenge\_kmeans$Cluster), mean)

# implémentation du résultats Kmeans R

Challenge$d0 <- sqrt((Challenge$nb\_jour\_total\_cause-11.761297)^2 + (Challenge$nb\_jour\_paging\_without\_response-5.6070895)^2 +

(Challenge$nb\_jour\_reiter\_call-3.353767)^2 + (Challenge$nb\_jour\_call\_est\_bad\_setupduration-0.7952208)^2 +

(Challenge$top\_cause-1.997322)^2 + (Challenge$top\_paging-3.309034)^2 +

(Challenge$top\_reiter-1.727491)^2 + (Challenge$top\_setupduration-1.441301)^2 )

Challenge$d1 <- sqrt((Challenge$nb\_jour\_total\_cause-3.535535)^2 + (Challenge$nb\_jour\_paging\_without\_response-0.9285452)^2 +

(Challenge$nb\_jour\_reiter\_call-2.200010)^2 + (Challenge$nb\_jour\_call\_est\_bad\_setupduration-0.1625146)^2 +

(Challenge$top\_cause-1.170420)^2 + (Challenge$top\_paging-1.357630)^2 +

(Challenge$top\_reiter-1.512893)^2 + (Challenge$top\_setupduration-1.116349)^2 )

Challenge$d2 <- sqrt((Challenge$nb\_jour\_total\_cause-23.883199)^2 + (Challenge$nb\_jour\_paging\_without\_response-15.3096060)^2 +

(Challenge$nb\_jour\_reiter\_call-18.438962)^2 + (Challenge$nb\_jour\_call\_est\_bad\_setupduration-14.9097152)^2 +

(Challenge$top\_cause-7.506314)^2 + (Challenge$top\_paging-6.109987)^2 +

(Challenge$top\_reiter-5.928971)^2 + (Challenge$top\_setupduration-3.972418)^2 )

Challenge$d3 <- sqrt((Challenge$nb\_jour\_total\_cause-20.684505)^2 + (Challenge$nb\_jour\_paging\_without\_response-5.5657382)^2 +

(Challenge$nb\_jour\_reiter\_call-9.204865)^2 + (Challenge$nb\_jour\_call\_est\_bad\_setupduration-7.2919759)^2 +

(Challenge$top\_cause-4.695176)^2 + (Challenge$top\_paging-3.017182)^2 +

(Challenge$top\_reiter-3.183324)^2 + (Challenge$top\_setupduration-3.583580)^2 )

Challenge$d4 <- sqrt((Challenge$nb\_jour\_total\_cause-20.558639)^2 + (Challenge$nb\_jour\_paging\_without\_response-14.2365649)^2 +

(Challenge$nb\_jour\_reiter\_call-8.454938)^2 + (Challenge$nb\_jour\_call\_est\_bad\_setupduration-3.4120222)^2 +

(Challenge$top\_cause-3.840741)^2 + (Challenge$top\_paging-6.011990)^2 +

(Challenge$top\_reiter-3.173464)^2 + (Challenge$top\_setupduration-2.406955)^2 )

Challenge$d5 <- sqrt((Challenge$nb\_jour\_total\_cause-11.449623)^2 + (Challenge$nb\_jour\_paging\_without\_response-3.1355932)^2 +

(Challenge$nb\_jour\_reiter\_call-13.798767)^2 + (Challenge$nb\_jour\_call\_est\_bad\_setupduration-1.1787899)^2 +

(Challenge$top\_cause-2.671001)^2 + (Challenge$top\_paging-2.436287)^2 +

(Challenge$top\_reiter-5.423573)^2 + (Challenge$top\_setupduration-1.656887)^2 )

#Challenge$d <- min(Challenge$d0,Challenge$d1,Challenge$d2,Challenge$d3,Challenge$d4,Challenge$d5)

names(Challenge)

library(dplyr)

#Leader\_S\_2021 %>% rowwise() %>% mutate(d = min(Leader\_S\_2021$d0, Leader\_S\_2021$d1))

df <- Challenge[c(20,21,22,23,24,25)]

df$test <- apply(df, 1, FUN=min)

Challenge$d <- df$test

Challenge$Cluster <- ifelse(Challenge$d ==Challenge$d0,0,

ifelse(Challenge$d ==Challenge$d1,1,

ifelse(Challenge$d ==Challenge$d2,2,

ifelse(Challenge$d ==Challenge$d3,3,

ifelse(Challenge$d ==Challenge$d4,4,

ifelse(Challenge$d ==Challenge$d5,5,8))))))

Challenge$d0 <- NULL

Challenge$d1 <- NULL

Challenge$d2 <- NULL

Challenge$d3 <- NULL

Challenge$d4 <- NULL

Challenge$d5 <- NULL

Challenge$d <- NULL

Challenge$top\_cause <- NULL

Challenge$top\_paging <- NULL

Challenge$top\_reiter <- NULL

Challenge$top\_setupduration <- NULL

Challenge$decile\_cause <- NULL

Challenge$decile\_paging <- NULL

Challenge$decile\_reiter <- NULL

Challenge$decile\_setupduration <- NULL

Challenge$test <- 1

test2 <-aggregate(data.frame(nb\_client = Challenge$test,

cause = Challenge$total\_cause,

paging = Challenge$paging\_without\_response\_nb\_monthly,

reiter = Challenge$reiter\_call\_interval\_nb\_monthly,

setupduration = Challenge$call\_est\_bad\_setupduration\_nb\_monthly),

by = list(Cluster =Challenge$Cluster), sum)

test3 <-aggregate(data.frame(nb\_client = Challenge$test,

cause = Challenge$nb\_jour\_total\_cause,

paging = Challenge$nb\_jour\_paging\_without\_response,

reiter = Challenge$nb\_jour\_reiter\_call,

setupduration = Challenge$nb\_jour\_call\_est\_bad\_setupduration),

by = list(Cluster =Challenge$Cluster), mean)