> file<-"E:\\Ph.D\\3-JOURNALS\\Obj\_4-AllergyApplication\\Allergy - APP\\Allergy Data\\allergy\_data\_ubtrain.csv"

> aapp\_ubtrain <- read.csv(file, header = FALSE)

> colnames(aapp\_ubtrain) <- c("housedust","cottondust","aspergilus","pollen", "parthenium", "cockroach","catdander","dogfur","roaddust","oldpdust","PSdust","MilkP","MilkC","curd","coffee","tea","beef","chicken","mutton","egg","fishA","fishB","crab","prawns","shark","avaraikai","banana","beans","beetroot","brinjal","cabbage","capsicum","chillie","cauliflower","carrot","chowchow","corn","cucumber","drumstick","greens","gourds","kovaikai","kothavarai","Lfinger","malli","mango","mushroom","nuckol","onion","peas","potroot","paneer","potato","pumkin","pudina","radish","tomato","tondaikai","vazpooThandu","yams","gram","channa","dhal","maida","oats","ragi","rice","wheat","coconut","oil","garlic","ginger","pepper","tamarind","aginomoto","spices","coco","horlicks","boost","sdrink","hdrink","sfruit","wfruit","lime","nuts","runningnose","sneeze","cough","wheezeBlocks","headache","itching","swelling","redrashes","FHistory","gender","Class")

> col\_names <- names(aapp\_ubtrain)

> aapp\_ubtrain[,col\_names] <- lapply(aapp\_ubtrain[,col\_names] , factor)

> library("UBL")

> aapp\_ideal<-SmoteClassif(Class ~ ., aapp\_ubtrain, C.perc ="balance",dist="HEOM")

> table(aapp\_ideal$Class)

**N OT RH RH\_OT RH\_UT UT UT\_OT**

**203 203 203 203 203 203 203**

> write.csv(aapp\_ideal,"E:\\Ph.D\\3-JOURNALS\\Obj\_4-AllergyApplication\\Allergy - APP\\Allergy Data\\allergy\_data\_ideal.csv")

> aapp\_single<-SmoteClassif(Class ~ ., aapp\_ubtrain, C.perc =list(N=1.4,OT=1.25,RH=1.01,RH\_O=1,RH\_UT=1,UT=0.85,UT\_O=1),dist="HEOM")

> dim(aapp\_single)

[1] 1423 96

> table(aapp\_single$Class)

**N OT RH RH\_OT RH\_UT UT UT\_OT**

**78 282 390 90 72 451 60**

> write.csv(aapp\_single,"E:\\Ph.D\\3-JOURNALS\\Obj\_4-AllergyApplication\\Allergy - APP\\Allergy Data\\allergy\_data\_single.csv")

aapp\_complete<-SmoteClassif(Class ~ ., aapp\_ubtrain, C.perc =list(N=1.4,OT=0.95,RH=0.9,RH\_O=1,RH\_UT=1.6,UT=0.8,UT\_O=1),dist="HEOM")

> table(aapp\_complete$Class)

**N OT RH RH\_OT RH\_UT UT UT\_OT**

**78 214 348 90 115 424 60**

> dim(aapp\_complete)

[1] 1329 96

> write.csv(aapp\_complete,"E:\\Ph.D\\3-JOURNALS\\Obj\_4-AllergyApplication\\Allergy - APP\\Allergy Data\\allergy\_data\_complete.csv")

Added 50 RH\_OT samples, 40 UT\_OT samples, 5 RH\_UT samples manually

> table(aapp\_complete$Class)

**N OT RH RH\_OT RH\_UT UT UT\_OT**

**78 214 348 140 120 424 100**

> dim(aapp\_complete)

[1] 1424 96