###############################################################################################

Code to create dataset with weigthing Datasets (Binary, TFIDF etc) and Feature Selection (Chi, oneR Gainratio )

Office Computer

#################################################################################################

Dataset.file<- "D:/Experimentjournal/FinalFeatureList.csv"

Dataset <- read.csv(Dataset.file,header=TRUE)

Binary <- Dataset[Dataset$Datset == 'Binary' & Dataset$Collection == 'Secur\_Spars',]

TFIDF <- Dataset[Dataset$Datset == 'TFIDF' & Dataset$Collection == 'Secur\_Spars',]

TFC <- Dataset[Dataset$Datset == 'TFC' & Dataset$Collection == 'Secur\_Spars',]

LTC <- Dataset[Dataset$Datset == 'LTC' & Dataset$Collection == 'Secur\_Spars',]

Entropy <- Dataset[Dataset$Datset == 'Entropy' & Dataset$Collection == 'Secur\_Spars',]

Binary <- Binary[,-c(1,2,3)]

TFIDF <- TFIDF[,-c(1,2,3)]

TFC <- TFC [,-c(1,2,3)]

LTC <- LTC [,-c(1,2,3)]

Entropy <- Entropy[,-c(1,2,3)]

label1 <- c("Binary", "Entropy", "LTC", "TFC", "TFIDF")

label2 <- c("Chi","oneR", "InfoG", "GR", "RF")

setwd("D:/Experimentjournal/catalogs/SecurityFS/80Sparse")

temp = list.files(pattern="\*.csv")

for (i in 1:length(temp)) assign(temp[i], read.csv(temp[i]))

totalfiles<- length(temp)

Dataset<- read.csv(temp[5],header=TRUE)

Totalweigtht <- nrow(TFIDF)

for(k in 1:Totalweigtht)

{

Sel\_Features <- TFIDF [k,]

All\_Features <- names(Dataset)

N <- length(All\_Features)

M <- length(Sel\_Features)

List1<- c()

for(i in 1:M)

{

for(j in 1:N)

{

if(strcmp(All\_Features[j], Sel\_Features[i]))

{List1 <-append(List1, j) }

}

}

Filename <- paste("Secur\_Spars\_",label1[5],"\_",label2[k], ".csv", sep="")

ds<- Dataset[,List1]

write.csv(ds, Filename)

}