**Tugas Besar Data Mining Kelompok 10**

videogamesales

**THEOFILUS.S (3311901070)**

**Muhammad Putra Caesar Zellya (3311901084)**

buat pengaturan lokasi directory yang foldernya dibuat adalah **data** dan nama file excel sesuai dengan nama nya

buatlah variable **lokasi\_kerja** sesuaikan nilainya dengan nama file excel

**lokasi\_kerja <- "D:/data"**

**getwd()**

perintah untuk pengaturan lokasi directory

**setwd(lokasi\_kerja)**

**getwd()**

perintah untuk Jalankan perintah untuk mengatur working directory

**dataset <- read.csv("videogamesales.csv", sep = ";")**

**head(dataset)**

Gunakan perintah library yang sudah didownload  
**library(arules)**

**library(arulesViz)**

**library(C50)**

**library(printr)**

Ketika belum ada maka menggunakan perintah **install.packages()** contoh **install.packages(arules)** dan seterusnya Ketika sudah maka tinggal di library

Import dataset yang sudah dibuat

**dataset <- read.transactions("videogamesales.csv", sep = ",")**

**inspect(dataset)**

Perintah rules

**rules<- apriori(dataset, parameter = list(support = 0.2, confidence = 0.6))**

**inspect(rules)**

pertama run dulu ulang perintah **dataset <- read.csv("videogamesales.csv", sep = ";")** dan **head(dataset)**

Dan gunakan perintah

**arulesViz::plotly\_arules(rules)**

Untuk melihat berbagai korelasi antara Na\_Sales dan Other\_sales di dataset videogame sales maka menggunakan

**cor(dataset$NA\_Sales, dataset$Other\_Sales)**

atau

**cor(dataset$NA\_Sales, dataset$Year)**

**cor(dataset$Other\_Sales, dataset$Year)**

selanjutnya ketik perintah

**class(dataset)**

**class(dataset$age)**

**class(dataset$buys\_computer)**

dan ketik

**dataset$NA\_Sales <- as.factor(dataset$NA\_Sales)**

**dataset$Other\_Sales <- as.factor(dataset$Other\_Sales)**

untuk merubah data set NA\_Sales dan Other\_Sales kedalam factor dan lanjut run pada perintah class dengan data set nya

modeling dalam alogaritma pada dataset NA\_Sales dan Other\_Sales yaitu

**model <- C5.0(NA\_Sales ~., data=dataset)**

dan untuk menampilkan model yaitu

**model**

dan juga bisa menggunkan data set satulagi

**model <- C5.0(Other\_Sales ~., data=dataset)**

**model**

dan mengetik

**summary(model)**

untuk menampilkan semua model dari alogaritma C5.0 dalam data set

untuk menampilkan plot pada data setnya yang sudah dirubah kedalam model maka mengetik

**plot(model)**

ketika menjadikan dataset, sebagai data testing di kolom 1,2,3,4 pada datanya maka menggunakan

**datatesting <- dataset[,1:4]**

untuk menampilkan prediksi model data testing pada datasetnya maka menggunakan

**predictions <- predict(model, datatesting)**

membandingkan hasil prediksi dengan dataset menggunakan

**table(predictions, dataset$NA\_Sales)**

dan selanjutnya dengan dataset yang lain seperti

**table(predictions, dataset$Other\_Sales)**

**table(predictions, dataset$Platform.)**

**table(predictions, dataset$Year)**