| **Nama :**  **Nim :** | **Praktikum Penambangan Teks** | **MODUL 6**    **Nama Dosen :**  **Syandra Sari, S.Kom, M.Kom** |
| --- | --- | --- |
| **Hari dan Tanggal :** | **Nama Asisten Laboratorium :**   1. **Ridho Rachmat Giffary (064001800006)** 2. **Tasya Aulia (065001800022)** |

**Praktikum 6 : Naive Bayes**

**KODE POKOK BAHASAN:**

**DESKRIPSI POKOK BAHASAN:**

Mengenalkan dan mengajarkan cara melakukan salah satu proses pada tahapan Analisis Sentimen yaitu menganalisis data menggunakan metode Machine Learnning salah satunya adalah Naive Bayes

| No | Elemen Kompetensi | Indikator Kerja | Jml Jam | Hal |
| --- | --- | --- | --- | --- |
| 1 | Melakukan Analisis menggunakan Naive Bayes | Mampu melakukan Analisis menggunakan Naive Bayes | 60 menit | 2 |
| 2 | Melakukan penjelasan terhadap setiap elemen Script | Mampu melakukan penjelasan terhadap setiap elemen Script | 30 menit |  |

**TUGAS PENDAHULUAN:**

Hal yang harus dilakukan dan acuan yang harus dibaca sebelum praktikum:

1. Mengetahui pemahaman dasar tentang Kegunaan penarikan data.
2. Mempersiapkan Aplikasi R atau Jupyter Notebook yang sudah terinstall di Laptop masing - masing.

**TEORI SINGKAT:**

| Tuliskan Teori singkat terkait Naive Bayes minimal 2 Paragraf |
| --- |

**LAB SETUP:**

Hal yang harus disiapkan dan dilakukan oleh praktikan untuk menjalankan praktikum modul ini.

**Alat dan Bahan:** Laptop

**Software** : R atau Jupyter Notebook

**ELEMEN KOMPETENSI I:**

**Deskripsi:**

Melakukan Analisis menggunakan Naive Bayes

**Kompetensi Dasar:**

Mampu melakukan Analisis menggunakan Naive Bayes

**Latihan 1:**

**Penjelasan Singkat:**

Pada latihan ini anda diminta untuk melakukan Analisis menggunakan Naive Bayes

| library(caret)  library(RTextTools)  library(dplyr)  library(tm)  df <- read.csv("directorydata/data.csv", stringsAsFactors = FALSE)  glimpse(df)  table(df$klasifikasi)  df <- df[sample(nrow(df)), ]  glimpse(df)  df$klasifikasi<- as.factor(df$klasifikasi)  corpus <- Corpus(VectorSource(df$text))  corpus  inspect(corpus[1:3])  dtm <- DocumentTermMatrix(corpus)  inspect(dtm)  findFreqTerms(dtm, 5) #mencari kata yang mempunyai frekuensi muncul lebih dari 5 kali  tdm <- TermDocumentMatrix(corpus)  inspect(tdm)  ## SPLITTING DATA  library(e1071) #package for classification  n <- nrow(df)  df.train <- df[1:round(.9 \* n),] #percentage split for 90:10, do to another composition 80:20, 70:30, 60:40, and 50:50  df.test <- df[(round(.9 \* n)+1):n,]  nn <- length(corpus)  corpus.train <- corpus[1:round(.9 \* nn)]  corpus.test <- corpus[(round(.9 \* nn)+1):nn]  nnn <- nrow(dtm)  dtm.train <- dtm[1:round(.9 \* nnn),]  dtm.test <- dtm[(round(.9 \* nnn)+1):nnn,]  dim(dtm.train)  fivefreq <- findFreqTerms(dtm.train, 10)  length((fivefreq))  fivefreq  dtm.train.nb <- DocumentTermMatrix(corpus.train, control=list(dictionary = fivefreq))  dim(dtm.train.nb)  dtm.test.nb <- DocumentTermMatrix(corpus.test, control=list(dictionary = fivefreq))  dim(dtm.test.nb)  convert\_count <- function(x) {  y <- ifelse(x > 0, 1,0)  y <- factor(y, levels=c(0,1), labels=c("Yes", "No"))  y  }  # Apply the convert\_count function to get final training and testing DTMs  trainNB <- apply(dtm.train.nb, 2, convert\_count)  testNB <- apply(dtm.test.nb, 2, convert\_count)  # Train the classifier  system.time( classifier <- naiveBayes(trainNB, df.train$klasifikasi, laplace = 1) )  # Use the NB classifier we built to make predictions on the test set.  system.time( pred <- predict(classifier, newdata=testNB) )  # Create a truth table by tabulating the predicted class labels with the actual class labels  library(gmodels)  conf.mat <- confusionMatrix(pred, df.test$klasifikasi)  conf.mat |
| --- |

**ELEMEN KOMPETENSI 2:**

**Deskripsi:**

Melakukan penjelasan terhadap setiap elemen Script

**Kompetensi Dasar:**

Mampu melakukan penjelasan terhadap setiap elemen Script

**Latihan 2:**

**Penjelasan Singkat:**

Pada latihan ini anda diminta untuk melakukancpenjelasan terhadap setiap elemen Script

| Letakan Code kalian disini dan jelaskan setiap baris pada code tersebut! |
| --- |



| **No** | **Elemen Kompetensi** | **Penyelesaian**  **Cek List** | |
| --- | --- | --- | --- |
| **Selesai** | **Tidak** |
| **1** | **Melakukan Pelabelan berdasarkan Kamus positif dan Kamus negatif**  **Melakukan proses Pembobotan** |  |  |
| **2.** | **Melakukan penjelasan terhadap setiap elemen Script** |  |  |

1. **Form Umpan Balik**

| **Elemen Kompetensi** | **Waktu Pengerjaan** | **Kriteria** |
| --- | --- | --- |
| **Melakukan Analisis menggunakan Naive Bayes** |  |  |
| **Melakukan Pelabelan berdasarkan Kamus positif dan Kamus negatif**  **Melakukan proses Pembobotan** |  |  |

Kriteria

1.Sangat Menarik

2.Cukup Menarik

3.Kurang Menarik

4.Sangat Kurang Menarik