| **Nama :**  **Nim :** | **Praktikum Penambangan Teks** | **MODUL 3**    **Nama Dosen :**  **Syandra Sari, S.Kom, M.Kom** |
| --- | --- | --- |
| **Hari dan Tanggal :** | **Nama Asisten Laboratorium :**   1. **Ridho Rachmat Giffary (064001800006)** 2. **Tasya Aulia (065001800022)** |

**Praktikum 3 :**

**Normalize & Stopword**

**KODE POKOK BAHASAN:**

**DESKRIPSI POKOK BAHASAN:**

Mengenalkan dan mengajarkan cara melakukan salah satu proses pada tahapan Pre-Processing yaitu Normalize dan Stopwors yang akan digunakan untuk membersihkan data yang telah ditarik dari platform social media twitter sebelumnya.

| No | Elemen Kompetensi | Indikator Kerja | Jml Jam | Hal |
| --- | --- | --- | --- | --- |
| 1 | Melakukan proses Normalize | Mampu melakukan proses Normalize terhadap data yang telah dihimpun | 30 menit | 2 |
| 2 | Melakukan proses Stopword | Mampu melakukan proses Stopword terhadap baris yang tidak diperlukan | 60 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 Normalize dan Stopword pada Text Mining 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 proses Normalize

**Kompetensi Dasar:**

Mampu melakukan proses Normalize terhadap data yang telah dihimpun

**Latihan 1:**

**Penjelasan Singkat:**

Pada latihan ini anda diminta untuk melakukan proses Normalize terhadap data yang telah dihimpun pada tugas sebelumnya.

Download data Normalize dan stopword pada link berikut ini : <https://drive.google.com/drive/folders/18QwLnE4VzG1zn0qeGTHF-0rtqinN8iys?usp=sharing>

R

| library(textclean)  Kelompok\_Praktikan <- read.csv('nama\_data.csv')  normalize <- read.csv("kamusnormalize.csv")  View(slangw)  Kelompok\_Praktikan <-replace\_internet\_slang(aya2, slang = paste0("\\b", normalize$sebelum, "\\b"),  replacement = normalize$setelah, ignore.case = TRUE)  View(Kelompok\_Praktikan)  write.csv(Kelompok\_Praktikan, file = directory/nama\_daya', row.names = F) |
| --- |

**ELEMEN KOMPETENSI 2:**

**Deskripsi:**

Melakukan proses Stopword

**Kompetensi Dasar:**

Mampu melakukan proses Stopword terhadap baris yang tidak diperlukan

**Latihan 2:**

**Penjelasan Singkat:**

Pada latihan ini anda diminta untuk melakukan proses Stopword terhadap baris yang tidak diperlukan.

| library(tm)  Kelompok\_Praktikan <- read.csv('cleanslangvaksinsinovac.csv')  Kelompok\_Praktikan <- iconv(Kelompok\_Praktikan$x, to="UTF-8")  Kelompok\_Praktikan <- Corpus(VectorSource(Kelompok\_Praktikan))  Kelompok\_Praktikan <- tm\_map(Kelompok\_Praktikan, tolower)  inspect(Kelompok\_Praktikan[20:25])  stopwordID <- "C:/Users/ibrahim/Downloads/latihan TA/stopword\_id.txt"  cStopwordID <- readLines(stopwordID)  Kelompok\_Praktikan <- tm\_map(Kelompok\_Praktikan, removeWords, cStopwordID)  Kelompok\_Praktikan <- tm\_map(Kelompok\_Praktikan, removeWords, c('jeni', 'samp', 'amp', 'jenisjenis', 'kanisu', 'teruk',  'elak...', 'jerung','u','ppp','kamukamukamukamu', 'ff', 'sinovac', 'vaksin', 'gitugitugitu'))  Kelompok\_Praktikan <- unlist(Kelompok\_Praktikan)  View(Kelompok\_Praktikan) |
| --- |



| **No** | **Elemen Kompetensi** | **Penyelesaian**  **Cek List** | |
| --- | --- | --- | --- |
| **Selesai** | **Tidak** |
| **1** | **Melakukan proses Normalize** |  |  |
| **2.** | **Melakukan proses Stopword** |  |  |

1. **Form Umpan Balik**

| **Elemen Kompetensi** | **Waktu Pengerjaan** | **Kriteria** |
| --- | --- | --- |
| **Melakukan proses Normalize** |  |  |
| **Melakukan proses Stopword** |  |  |

Kriteria

1.Sangat Menarik

2.Cukup Menarik

3.Kurang Menarik

4.Sangat Kurang Menarik