| **Nama :**  **Nim :** | **Praktikum Penambangan Teks** | **MODUL 4**    **Nama Dosen :**  **Syandra Sari, S.Kom, M.Kom** |
| --- | --- | --- |
| **Hari dan Tanggal :** | **Nama Asisten Laboratorium :**   1. **Ridho Rachmat Giffary (064001800006)** 2. **Tasya Aulia (065001800022)** |

**Praktikum 4 :**

**Tokenizing dan Stemming**

**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 instalasi Library Stemming | Mampu melakukan instalasi Library Katadasar untuk Stemming | 30 menit | 2 |
| 2 | Melakukan proses Tokenizing dan Stemming | Mampu melakukan proses Tokenizing dan Stemming | 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 Stemming dna Tokenizing 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 instalasi library katadasar

**Kompetensi Dasar:**

Mampu melakukan proses instalasi library katadasar untuk proses Stemming

**Latihan 1:**

**Penjelasan Singkat:**

Pada latihan ini anda diminta untuk melakukan proses instalasi library katadasar untuk stemming

R

| library(githubinstall)  gh\_install\_packages("nurandi/katadasaR")  library(katadasaR)  katadasar("memberikan") |
| --- |

**ELEMEN KOMPETENSI 2:**

**Deskripsi:**

Melakukan proses Tokenizing dan Stemming

**Kompetensi Dasar:**

Mampu melakukan prosesTokenizing dan Stemming

**Latihan 2:**

**Penjelasan Singkat:**

Pada latihan ini anda diminta untuk melakukan proses Tokenizing dan Stemming

| Nama\_Kelompok = read.csv("namadatahasilcleansing.csv")  stem <- as.character(Nama\_Kelompok$x)  stem[20:25]  library(katadasaR)  stemming <- function(x){  paste(lapply(x,katadasar),collapse = " ")}  library(tokenizers)  tweets <- lapply(tokenize\_words(stem[]), stemming)  tweets[1:5]  Nama\_Kelompok2<- unlist(tweets)  View(Nama\_Kelompok2)  write.csv(Nama\_Kelompok2, file = 'C:/Users/ibrahim/Downloads/latihan TA/namadatakelompokhasilstemming.csv', row.names = F) |
| --- |



| **No** | **Elemen Kompetensi** | **Penyelesaian**  **Cek List** | |
| --- | --- | --- | --- |
| **Selesai** | **Tidak** |
| **1** | **Melakukan proses instalasi library KatadasaR** |  |  |
| **2.** | **Melakukan proses Tokenizing dan Stemming** |  |  |

1. **Form Umpan Balik**

| **Elemen Kompetensi** | **Waktu Pengerjaan** | **Kriteria** |
| --- | --- | --- |
| **Melakukan proses Instalasi Library KatadasaR** |  |  |
| **Melakukan proses Tokenizing dan Stemming** |  |  |

Kriteria

1.Sangat Menarik

2.Cukup Menarik

3.Kurang Menarik

4.Sangat Kurang Menarik