

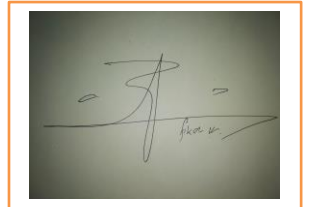


Laboratorium Komputer Teknik Informatika
Fakultas Teknologi Industri
Universitas Ahmad Dahlan

Jalan Ring Road Selatan, Tamanan, Banguntapan, Bantul Yogyakarta 55166 Email: laboratorium@tif.uad.ac.id Web: labs.tif.uad.ac.id

RESPONSI PRAKTIKUM
PEMROSESAN BAHASA ALAMI

NAMA : ERVIN FIKOT M
NIM : 1700018127
SLOT : SELASA, 09.00 WIB
KELAS : B
TANGGAL : 13 JULI 2020



Jawaban :

Dari Proyek Pengolahan Dokument terdapat beberapa Pengembangan antara lain :

1. Pengembangan Scopus Bahasa Rejang

Data Kosa kata dan artinya didapatkan melalui wawancara dengan tokoh adat budaya rejang dan mendapatkan sekitar 147 kosa kata yang biasanya digunakan oleh masyarakat.

2. Pemenggalan Kalimat menjadi kata (Tokenisasi)

Menggunakan Library Natural Language Tool kit untuk melakukan pemotongan kalimat mejadi kata sehingga nantinya dapat di proses oleh system untuk mencari kecocokan dalam directory.

3. Menghilangkan tanda baca (Filterasi)

Fungsi utama dalam menghilangkan tanda baca yaitu agar system dapat bekerja secara maksimal dan menghindari bug, selain itu filterisasi bertujuan untuk menghilangkan imbuhan yang tidak diperlukan.

4. Membakukan kata dengan Bahasa Indonesia yang benar (Stemmer)

Dikarenakan dalam direktori hanya ada Bahasa Indonesia yang baku, maka plaintext harus di bakukan agar dapat sesuai dan terdeteksi.

5. Pencocokan Text Bahasa Indonesia dengan kosa kata scopus Bahasa rejang Bengkulu

Pencocokan String bertujuan untuk Mendeteksi makna dar kata yang di maksud kemudia seluruh terjemahan kata akan digabungkan menjadi sebuah kalimat kembali, contoh output dapat dilihat di bawah :

```
In [42]: print("Kalimat Berbahasa Indonesia : ", plaintext)
         print("Kalimat Berbahasa Rejang Bengkulu : ", terjemah)

Kalimat Berbahasa Indonesia : dari tadi pagi saya menunggu kamu di rumah, tapi kamu tidak datang
Kalimat Berbahasa Rejang Bengkulu : kuni nano pueng uku tunggu ko nak umeak tapi ko coa t ko
```