

## Modul 1 Praktikum Natural Language Processing Pengenalannya Pengolahan Data Teks menggunakan Python

### Deskripsi Singkat

Deskripsi

### Tujuan

1. Tujuan
2. Tujuan
3. Tujuan

### Materi 1 - Pengumpulan Data

#### Studi Kasus

Anda ingin mengumpulkan data menggunakan API Twitter

#### Langkah-langkah

1. Menginstall library tweepy

```
pip install tweepy
```

2. Membuat script python untuk mengumpulkan data

```
1 import tweepy
2 import numpy as np
3
4
```

### Materi 2 - Membaca data dari PDF

Seringkali data kita disimpan dalam format PDF, dan untuk keperluan analisis lebih lanjut dalam pemrosesan bahasa alami, kita perlu mengekstrak teks dari file tersebut.

Teks yang diambil dari file PDF ini dapat digunakan untuk berbagai aplikasi dalam NLP, seperti analisis sentimen, ekstraksi informasi, atau pengenalan entitas. Dengan menggunakan pustaka Python seperti PyPDF2, kita bisa dengan mudah mengakses dan menyiapkan teks dari dokumen PDF untuk proses analisis lebih lanjut.

```
1 import tweepy
2 import numpy as np
3
4
```

**Materi 3 - Membaca data dari File Word (Microsoft Word)**

Isi materi 3

**Materi 4 - Membaca data dari JSON**

Isi materi 4

**Materi 5 - Membaca data dari HTML**

Isi materi 3

## **Tugas**

1. Kerjakan semua materi 1 - 5 diatas