|  |  |  |
| --- | --- | --- |
| **Text Mining** | | |
| Praktikan | **Aslab** | |
| Nama: xxxx  Nim: xxxx | Annur Hangga Prihadi | 065001800028 |
| Faiz Kumara | 065001800003 |

**PRAKTIKUM 5**

**DATA SAINS DAN ANALITIK**

Topik pertemuan praktikum ke-lima adalah mengolah teks untuk kegiatan analisa bahan data berupa teks.

**Source Code:** https://github.com/hanggaa/PrakDSDA/blob/main/Prak%205%20Text%20Processing.ipynb

**Latihan 1**

1. **Memasang library yang dibutuhkan**

|  |
| --- |
|  |

1. **Membaca file**

|  |
| --- |
|  |

1. **Mengubah huruf menjadi huruf kecil**

|  |
| --- |
|  |

1. **Menghapus angka pada paragraf**

|  |
| --- |
|  |

1. **Menghapus tanda baca**

|  |
| --- |
|  |

1. **Menghapus karakter kosong**

|  |
| --- |
|  |

1. **Proses tokenizing**

|  |
| --- |
|  |

1. **Memasang library Sastrawi**

|  |
| --- |
|  |

1. **Proses filter menggunakan Sastrawi**

|  |
| --- |
|  |

1. **Proses Stemming Sastrawi**

|  |
| --- |
|  |

1. **Frekuensi huruf yang muncul menggunakan Sastrawi**

|  |
| --- |
|  |

1. **Kata yang sering muncul menggunakan Sastrawi**

|  |
| --- |
|  |

1. **Keseluruhan kata yang muncul**

|  |
| --- |
|  |

1. **Plot data kata yang sering muncul**

|  |
| --- |
|  |

1. **Memasang library Porter**

|  |
| --- |
|  |

1. **Proses filter menggunakan Porter**

|  |
| --- |
|  |

1. **Proses Stemming menggunakan Porter**

|  |
| --- |
|  |

1. **Frekuensi kata yang muncul menggunakan Porter**

|  |
| --- |
|  |

1. **Kata yang sering muncul menggunakan Porter**

|  |
| --- |
|  |

1. **Keseluruhan kata yang muncul**

|  |
| --- |
|  |

1. **Plot data yang sering muncul**

|  |
| --- |
|  |

**Latihan 2**

1. Diberikan 2 bahan berita (news2 dan news3) silahkan lakukan olah teks mining seperti diatas (Silahkan pilih salah satu bahan berita)

**Lampiran Screenshot hasil**

|  |
| --- |
| **Input screenshot disini** |