1. **Deskripsi Masalah**

Membuat POSTagger sederhana berbasis 3 metode :

* Metode baseline
* Metode klasifikasi konvensional non-sekuensial (Algoritma Decision Tree)
* Metode HMM-Viterbi

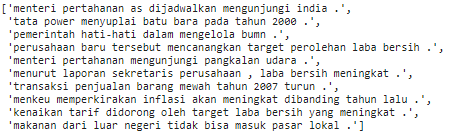
Terdapat 50 kalimat untuk training, dan 10 kalimat untuk testing.

1. **Perancangan Sistem & Analisis**

* Read File TSV



* Kalimat Uji



* Hasil Tagging pada metode Baseline





















* Hasil Tagging pada metode Klasifikasi





















* Hasil Tagging pada metode HMM-Viterbi





















* Akurasi pada metode Baseline



* Akurasi pada metode Klasifikasi



* Akurasi pada metode HMM-Viterbi



1. **Kesimpulan**

Kesimpulan yang didapat dari 3 metode yang dijalankan, akurasi tertinggi didapatkan dari metode baseline, dan yang paling rendah didapatkan dari metode HMM-Viterbi. Hal ini dapat terjadi karena pada metode baseline, kata-kata dari kalimat uji akan dicari dari corpus, jika terdapat kata yang sama, maka tag kata tersebut di assign tag yang paling sering muncul dari kata tersebut, jika tidak ada kata yang sama, maka tag kata tersebut akan di assign tag yang paling sering muncul dari corpus.

Sedangkan pada metode klasifikasi dengan algoritma decision tree mendapat akurasi tertinggi kedua setelah baseline. Hal ini dapat terjadi dalam proses pemodelan terdapat aturan yang ambigu, sehingga terdapat beberapa kesalahan.

Kemudian pada metode HMM-Viterbi mendapatkan akurasi paling rendah dibandingkan 2 metode lainnya. Hal ini dapat terjadi pada saat proses decoding Viterbi dan backtracking, terdapat nilai 0, sehingga tag yang didapat adalah ‘start’.