

1. Pendahuluan

Bagian ini berisi penjelasan latar belakang dipilihnya permasalahan penelitian, perumusan masalah, tujuan, metodologi, dan sistematika penulisan tugas akhir.

Latar Belakang

Part of Speech Tagging (POS-Tagging) adalah suatu proses yang memberikan label kelas kata secara otomatis pada suatu kata dalam kalimat[1]. Hasil dari *Part of Speech Tagging* (POS) ini sangat berpengaruh terhadap keluaran dari proses *Parsing*[2]. Masalah yang muncul adalah bagaimana cara mendapatkan pelabelan kelas kata yang tepat dalam konteks kalimat. Dengan adanya sistem pelabelan, pengguna akan lebih mudah untuk melakukan pelabelan kalimat, sehingga pengguna tidak perlu repot melakukan pencarian kalimat yang diinginkan. Sistem pelabelan sendiri harus dapat menganalisis sekian banyak data tentang kalimat yang tersedia, dapat juga didukung dengan data *rating* agar hasil yang diberikan lebih akurat.

Ada banyak metode yang dapat digunakan untuk membangun *POS Tagging* kalimat. Contohnya menggunakan metode *Flair NLP*, *Maximum Entropy Models*, *Hidden Markov Model*, dan masih banyak lagi. Salah satu metode yang paling menonjol yaitu dengan menggunakan implementasi Metode *Hidden Markov Model (HMM)*. *Hidden Markov Model* dapat melakukan pengamatan pada suatu yang berhubungan (misalnya variabel yang berkorelasi) dengan *state* sebelumnya[3]. Karakteristik yang demikian membuat HMM menjadi metode yang tepat dalam merekomendasi kalimat *real-time*, karena HMM mengamati kalimat sebelumnya yang dimasukkan dan memperkirakan kalimat yang direkomendasikan berdasarkan kondisi sebelumnya. Dari karakteristik yang demikian, membuat HMM menjadi metode yang tepat untuk membuat prediksi pada suatu kalimat.

Sistem *POS Tagging* ini diharapkan dapat membantu orang dalam menentukan tata bahasa Indonesia yang benar khususnya yang baru mengenal dalam Tagging Bahasa Indonesia. Diharapkan untuk kedepannya para peneliti di bidang pemrosesan bahasa alami untuk bahasa Indonesia tidak lagi harus membuat sistem rekomendasi kalimat secara manual dari awal.

Topik dan Batasannya

Berdasarkan dari latar belakang yang telah dijelaskan, pertanyaan penelitian yang diangkat dalam tugas akhir ini sebagai berikut:

1. Bagaimana performansi metode *Hidden Markov Model* dalam pembuatan *POS Tagging* bahasa Indonesia.
2. Bagaimana sistem yang dibangun akan menampilkan *output* berupa tag yang sudah dihitung berdasarkan probabilitas tag yang tertinggi.

Adapun batasan dari permasalahan yang ada pada tugas akhir ini sebagai berikut:

1. Kalimat hanya bisa menggunakan alfabet bahasa Indonesia, angka, dan simbol tertentu.
2. Tag yang bisa diidentifikasi oleh sistem sesuai dengan Tagset Bahasa Indonesia.
3. Kata akan di set tag menjadi "NN" apabila kata yang dimasukkan tidak ada di *dataset*.

Tujuan

Tujuan dari Tugas Akhir ini adalah sebagai berikut:

1. Membuat *POS tagging* Bahasa Indonesia menggunakan *Hidden Markov Model*.
2. Mengetahui performansi *Hidden Markov Model* dalam sistem *POS Tagging* Bahasa Indonesia.