

ABSTRAK

Nama : Muhammad Hafidz Naufal Hilal

NIM : 1167010044

Judul Skripsi : Analisis *Part Of Speech Tagging* Dengan Menggunakan *Hidden Markov Model* Pada Data Al-Qur'an

Part of Speech (POS) tagging merupakan bagian dari Natural Language Processing untuk menentukan dengan benar label kata pada suatu kalimat dari input kata yang diberikan secara otomatis. Pengembangan teknik POS tagging untuk bahasa Arab masih sedikit dilakukan. Penelitian ini membahas POS tagging menggunakan metode Hidden Markov Model pada data teks Al-Qur'an. Dalam penelitian ini dataset terdiri dari tiga kategori data yang berasal dari quran corpus terdiri dari 150 kalimat sempurna sederhana, 50 kalimat dengan S/P/O/K lebih dari satu dan 50 ayat Al-Qur'an pilihan. Percobaan data dilakukan dengan menggunakan teknik K-Fold Cross Validation. Dataset dibagi menjadi dua, yaitu data latih dan data uji. Data latih digunakan untuk mencari probabilitas emisi dan transisi, sedangkan pengujian data menggunakan algoritma Viterbi untuk menentukan tag terbaik pada setiap kata. Hasil percobaan didapat akurasi rata-rata sebesar 89,44% untuk kategori pertama yaitu dataset kalimat sederhana, 74,18% untuk kategori kedua yaitu dataset ada anak kalimat, dan 69,04% untuk kategori ketiga yaitu dataset ayat lengkap.

Kata Kunci : *Part of Speech Tagging, Hidden Markov Model, Algoritma Viterbi, K-Fold Cross Validation*

