

## ***Part-Of-Speech Tag Studi Kasus Kelas Kata dalam Al-Quran***

**Muhammad Amal Dzaky<sup>1</sup>, Moch. Arif Bijaksana<sup>2</sup>, Arif Fatchul Huda<sup>3</sup>**

<sup>1,2,3</sup>Fakultas Informatika, Universitas Telkom, Bandung

<sup>1</sup>mhdamaldzaky@students.telkomuniversity.ac.id,

<sup>2</sup>arifbijaksana@telkomuniversity.ac.id, <sup>3</sup>afh@uinsgd.ac.id

---

### **Abstrak**

Indonesia adalah negara dengan populasi terbesar ke empat dan menjadi negara terbesar untuk populasi muslim dunia. Sebagaimana seharusnya, Al-Qur'an menjadi pedoman untuk hidup para muslimin. Al-Qur'an sudah ditulis 1400 tahun yang lalu dan hingga sekarang umat islam mempelajari isi dari Al-Qur'an, bahkan dimulai dari kecil. Al-Qur'an memiliki 114 Surat, 6000 lebih ayat, dan 128219 kata. Dikarenakan Al-Qur'an menggunakan bahasa Arab, yang dimana bahasa Arab terbilang rumit secara morfologi. Dikarenakan Al-Qur'an memiliki banyak kata, alangkah lebih baik jika proses pembelajaran Al-Qur'an dibantu dengan memanfaatkan *Natural language Procесing*(NLP). Pada penelitian kali ini akan dilakukan klasifikasi *Part-Of-Speech Tagging* pada Al-Qur'an menggunakan *Term Frequency Inverse Document Frequency* (TF-IDF). Hasil dari klasifikasi memiliki nilai akurasi *Naive Bayes* sebesar 24.7725% untuk *corpus* berbahasa Arab, 54.9836% untuk tulisan Arab yang sudah di normalisasi, dan 64.4283% untuk *corpus* notasi *Buckwalter*, untuk *Random Forest* sebesar 38.077% untuk *corpus* berbahasa Arab, 69.6771% untuk tulisan Arab yang sudah di normalisasi, dan 81.685% untuk notasi *Buckwalter*.

**Kata kunci :** *Part-Of-Speech, Random Forest, Naive Bayes, Term Frequency Inverse Document Frequency, Natural language Procесing, Al-Qur'an.*