

## ABSTRAK

**Nama : Fauziah**

**NIM : 1167010023**

**Judul : Analisis *Part Of Speech* (POS) *Tagging* Bahasa Arab Menggunakan Algoritma *Bee Colony Optimization* (BCO) pada Data Al Quran**

*Part Of Speech* (POS) *tagging* merupakan proses otomatis untuk menentukan label tata bahasa yang tepat atau kategori sintaksis dari sebuah kata tergantung pada konteksnya. POS *tagging* merupakan salah satu proses penting pada aplikasi *Natural Language Processing* (NLP) seperti *summarization text*, *Speech Recognition* (SR), *Question Answering* (QA) dan *Information Retrieval* (IR). POS *tagging* otomatis dibutuhkan karena POS *tagging* secara manual membutuhkan waktu yang lama dan biaya mahal karena memerlukan ahli bahasa. Masalah utama dalam POS *tagging* secara otomatis adalah kata yang memiliki sifat berbeda jika ditempatkan pada konteks yang berbeda (ambigu) dan kata yang ada dalam *corpus* uji namun tidak ada dalam *corpus* latih *Out Of Vocabulary* (OOV). Dalam penelitian ini akan dibahas sebuah pendekatan POS *tagging* yang efisien untuk teks bahasa Arab menggunakan algoritma *Bee Colony Optimization* (BCO). Masalah POS *tagging* direpresentasikan sebagai graf dan teknik pembobotan baru diusulkan untuk menetapkan nilai transisi pada setiap label kelas kata yang mungkin tidak secara probabilitas, kemudian lebah mencari jalur solusi terbaik. *Dataset* yang digunakan dalam penelitian ini berasal dari *Quranic corpus* yang sudah ditransliterasi terdiri dari 150 kalimat sempurna sederhana sebagai kategori *dataset* mudah, 50 kalimat dengan S/P/O/K lebih dari satu sebagai kategori *dataset* sedang, dan 50 ayat Al Quran pilihan sebagai kategori *dataset* sulit. Pendekatan yang diusulkan dievaluasi menggunakan teknik validasi silang yaitu *k-fold cross validation*. Hasil penelitian menunjukkan akurasi rata-rata sebesar 100% untuk kategori *dataset* mudah, 98.96% untuk kategori *dataset* sedang, dan 94.96% untuk kategori *dataset* sulit.

**Kata Kunci: *Part of Speech Tagging*, *Bee Colony Optimization* (BCO)**

## **ABSTRACT**

**Name** : Fauziah

**NIM** : 1167010023

**Title** : *Arabic Part Of Speech (POS) Tagging Analysis Using Bee Colony Optimization (BCO) Algorithm on Quran Corpus*

*Part-of-Speech (POS) tagging is an automatic process for determining the proper grammatical tag or syntactic category of a word depending on the context. POS tagging is one of essential processes in many Natural Language Processing (NLP) applications such as text summarization, Speech Recognition (SR), Question Answering (QA) and Information Retrieval (IR). Automatic POS tagging is needed because manually POS tagging takes a long time and is expensive because it requires a linguist. The main problem of automatic POS tagging is words that have different properties if they are placed in different contexts (ambiguous) and words that are in the test corpus but are not in the training corpus Out-of-Vocabulary (OOV). This research will discuss an efficient tagging approach for the Arabic language using Bee Colony Optimization (BCO) algorithm. The POS tagging problem is represented as a graph and a new weighting technique is proposed to assign a transition value to each word tag not probability, then bee looking for the best solution path. The dataset used in this study comes from the transliterated Quranic Corpus consists of 150 simple perfect sentences as the easy dataset category, 50 sentences with more than one S/P/O/K as the medium dataset category, and 50 selected verses of the Quran as difficult dataset categories. The proposed approach was evaluated using a cross validation technique, k-fold cross validation. The results showed an average accuracy of 100% for the easy category, 98.96% for the medium category, and 94.96% for the difficult category.*

**Keyword:** *Part of Speech Tagging, Bee Colony Optimization (BCO)*