

ABSTRAK

Bahasa Arab memiliki sistem morfologi yang kompleks, khususnya pada pembentukan *fi'il* (kata kerja) yang dalam ilmu *sharaf* diklasifikasikan menjadi *fi'il mujarrad* dan *fi'il mazid*. *Fi'il mujarrad* merupakan kata kerja dasar tanpa tambahan huruf, sedangkan *fi'il mazid* adalah kata kerja yang mengalami penambahan huruf sehingga membentuk pola atau *wazan* tertentu. Permasalahan dalam penelitian ini adalah adanya kemiripan pola morfologi antara *fi'il* dan non-*fi'il* yang memengaruhi ketepatan klasifikasi. Penelitian ini bertujuan untuk mengidentifikasi *fi'il* dalam Surah Al-Baqarah serta mengklasifikasikan jenisnya menjadi *mujarrad* dan *mazid* beserta pola *wazan*-nya secara otomatis menggunakan pendekatan *rule-based pattern matching*. Metode yang digunakan menerapkan aturan berbasis pola morfologi Arab, seperti awalan *fi'il mudhari'*, panjang kata, serta keberadaan huruf tambahan sesuai kaidah *sharaf*. Data penelitian berupa teks Al-Qur'an yang diproses melalui tahapan *preprocessing* meliputi *load data* dan tokenisasi kata. Sistem mendeteksi sebanyak 916 kandidat *fi'il*, kemudian 468 data digunakan sebagai data evaluasi yang dibandingkan dengan anotasi manual menggunakan *confusion matrix*. Hasil penelitian menunjukkan bahwa sistem memperoleh akurasi klasifikasi jenis *fi'il* sebesar 75% dengan *precision*, *recall*, dan *F1-score* masing-masing sebesar 0,77, 0,75, dan 0,75. Pada klasifikasi pola *wazan*, sistem memperoleh akurasi sebesar 69,23% dengan *weighted average precision* sebesar 0,66, *recall* sebesar 0,69, dan *F1-score* sebesar 0,65. Hasil tersebut menunjukkan bahwa pendekatan *rule-based* cukup mampu mendeteksi *fi'il mujarrad* dan *mazid*, meskipun performa beberapa pola *wazan* masih rendah akibat kemiripan struktur. Keterbatasan tersebut berdampak pada menurunnya ketepatan sistem dalam membedakan beberapa pola yang memiliki karakteristik serupa. Oleh karena itu, diperlukan pengembangan aturan yang lebih spesifik serta integrasi metode pembelajaran mesin untuk meningkatkan akurasi sistem.

Kata kunci: *fi'il*, *wazan*, bahasa Arab, *rule-based*, klasifikasi, Al-Qur'an.

ABSTRACT

Arabic has a complex morphological system, particularly in the formation of fi'il (verbs), which in sharaf are classified into fi'il mujarrad and fi'il mazid. Fi'il mujarrad refers to basic verbs without additional letters, while fi'il mazid involves added letters that form specific wazan (patterns). The main problem addressed in this study is the similarity of morphological patterns between fi'il and non-fi'il, which affects classification accuracy. This study aims to identify fi'il in Surah Al-Baqarah and classify them into mujarrad and mazid, along with their wazan patterns, automatically using a rule-based pattern matching approach. The method applies rules based on Arabic morphological patterns, such as fi'il mudhari' prefixes, word length, and the presence of additional letters in accordance with sharaf principles. The data consist of Qur'anic text processed through preprocessing stages, including load data and tokenization. The system detected 916 candidate fi'il, of which 468 data points were used for evaluation by comparison with manual annotations using a confusion matrix. The results show that the system achieved an accuracy of 75% for fi'il type classification, with precision, recall, and F1-score of 0.77, 0.75, and 0.75, respectively. For wazan classification, the system achieved an accuracy of 69.23%, with weighted average precision of 0.66, recall of 0.69, and F1-score of 0.65. These findings indicate that the rule-based approach is sufficiently effective in detecting fi'il mujarrad and mazid, although performance for certain wazan patterns remains limited due to structural similarities. These limitations result in reduced system accuracy in distinguishing between certain patterns that share similar characteristics. Therefore, further development of more specific rules and integration with machine learning methods are recommended to improve system accuracy.

Keywords: verb (fi'il), wazan, Arabic language, rule-based, classification, Quran