

## INTISARI

Al-Quran sebagai panduan hidup bagi umat Islam, memiliki kompleksitas yang sulit untuk diproses secara manual. Studi sebelumnya (Utomo et al., 2020) mengkaji penggunaan algoritma stemmer dalam pengembangan ontologi Al-Quran Indonesia. Dalam konteks klasifikasi teks, tantangan dimensi fitur yang tinggi sering dihadapi, yang mempengaruhi akurasi klasifikasi. Penelitian sebelumnya (Purwati et al., 2023) telah mengimplementasikan Chi-Square pada ontologi Al-Quran, namun hasilnya belum optimal. Oleh karena itu, penelitian ini bertujuan untuk menginvestigasi beberapa metode seleksi fitur seperti Chi-Square, SelectPercentile, Mutual Information, Information Gain, dan SelectKBest with `classif_f` dengan penggunaan Classifier SVM dan BPNN pada klasifikasi instansi ontologi Quran. Berdasarkan hasil eksperimen, dapat disimpulkan bahwa Seleksi fitur menggunakan metode SelectPercentile menghasilkan performa hasil kinerja tertinggi dengan memberikan nilai ukuran presisi tertinggi pada ukuran data uji 40% pada dataset Tafsir Quran dari Kementerian Agama Indonesia dengan menggunakan model BPNN = 0.7415 dan dengan SVM = 0.7113. Melalui eksperimen dan analisis yang dilakukan, dapat disimpulkan bahwa pemilihan fitur yang tepat dapat meningkatkan performa klasifikasi instance secara signifikan.

Kata kunci: Al-Quran, klasifikasi teks, seleksi fitur, Classification performance

## **ABSTRACT**

*The Quran, as a guiding light for the Islamic community, possesses a complexity that is challenging to process manually. Previous research (Utomo et al., 2020) examined the utilization of stemmer algorithms in the development of the Indonesian Quran ontology. In the context of text classification, the challenge of high feature dimensions is often encountered, which impacts classification accuracy. Earlier study (Purwati et al., 2023) implemented Chi-Square in the Quran ontology, yet the results were suboptimal. Therefore, this research aims to investigate various feature selection methods such as Chi-Square, SelectPercentile, Mutual Information, Information Gain, and SelectKBest with classif\_f using SVM and BPNN classifiers. Based on experimental results, it can be concluded that feature selection using the SelectPercentile method yields the highest performance results, providing the highest precision values at a 40% test data size on the Tafsir Quran dataset from the Indonesian Ministry of Religious Affairs, with BPNN achieving 0.7415 and SVM achieving 0.7113. Through the conducted experiments and analysis, it can be inferred that proper feature selection significantly enhances instance classification performance.*

*Keywords: Al-Quran, Text classification, Feature selection, Classification performance*