

Abstrak—Klasifikasi multi-label adalah tugas penting dalam analisis teks, terutama untuk dataset kompleks seperti ayat-ayat Al-Qur'an yang sering mengandung beberapa label tematik. Studi ini meneliti penggunaan metode ensemble dengan mengombinasikan model pembelajaran mesin tradisional, seperti Support Vector Machine (SVM) dan Naïve Bayes, dengan model berbasis transformer, yaitu BERT. Penelitian ini mengevaluasi kinerja model individu dan ensemble di bawah berbagai kondisi prapemrosesan serta menggunakan Hamming Loss sebagai metrik evaluasi utama. SVM muncul sebagai model tunggal paling efektif dengan Hamming Loss terendah sebesar 0,0881, sementara ensemble SVM dan Naïve Bayes menunjukkan hasil yang kompetitif dengan Hamming Loss sebesar 0,0891. Menariknya, prapemrosesan minimal menghasilkan kinerja lebih baik dibandingkan dengan transformasi teks yang ekstensif, menekankan pentingnya mempertahankan kekayaan semantik dalam analisis ayat-ayat Al-Qur'an. Meskipun penyertaan BERT dalam ensemble menjanjikan, model ini sering menunjukkan performa yang kurang optimal karena sensitivitasnya terhadap dataset kecil dan ketergantungan kontekstual. Penelitian ini menyoroti potensi integrasi antara model pembelajaran mesin tradisional dan model berbasis transformer untuk klasifikasi teks Al-Qur'an serta memberikan wawasan berharga dalam mengoptimalkan strategi klasifikasi multi-label dalam konteks serupa.

Kata kunci—Klasifikasi multi-label, metode ensemble, BERT, Al-Qur'an, SVM, Naïve Bayes