

Abstrak

Hadis merupakan pedoman hidup yang berisi perbuatan ataupun perkataan Nabi Muhammad SAW, sehingga banyak umat muslim yang ingin mempelajarinya. Permasalahan dalam mempelajari hadis adalah jumlah hadis yang sangat banyak dan memiliki kategori untuk setiap hadis. Sehingga dibutuhkan sistem yang mampu mengelompokkan *corpus* pada hadis, supaya hadis mudah dipelajari sesuai dengan kelasnya. Oleh karena itu, dibangun yang mampu sistem mengklasifikasikan hadis ke dalam kelas anjuran, larangan, dan informasi dengan menggunakan kombinasi metode *preprocessing*; pembobotan kata menggunakan TF- IDF; seleksi fitur *Information Gain*; klasifikasi *Multinomial Naïve Bayes*; dan evaluasi *Confusion Matrix*. Hasil penelitian terhadap sistem menunjukkan kinerja terbaik didapatkan menggunakan *preprocessing* tanpa teknik *stopword removal* dan tanpa pengaplikasian seleksi fitur *Information Gain* dengan akurasi 88%. Kombinasi tanpa menggunakan *stopword removal* menghasilkan akurasi yang lebih tinggi karena model dapat memiliki lebih banyak fitur yang dapat diklasifikasikan.

Kata kunci : hadis, *information gain*, *multinomial*, *multi label*