Abstrak

Hadis merupakan sumber hukum bagi umat muslim setelah al-qur'an yang di dalamnya terdapat petunjuk baik berupa perkataan, perbuatan, sikap, dan lain-lain. Hadis wajib dipelajari dan diamalkan oleh umat muslim, lalu dijadikan sebagai pedoman hidup setelah al-qur'an. Melakukan klasifikasi pada hadis merupakan cara untuk memudahkan umat muslim dalam mempelajari hadis dengan cara melihat pola teks pada terjemahan hadis bukhari berdasarkan tiga kelas atau kategori berdasarkan anjuran, larangan dan informasi. Klasifikasi yang dilakukan adalah klasifikasi *multi-label*. Pada proses klasifikasi menggunakan *N-gram* dan *TF-IDF* sebagai ekstraksi fitur, lalu menggunakan *CART* dan *bagging* sebagai metode klasifikasi dan *hamming loss* sebagai metode evaluasi. *Bagging* digunakan untuk menutupi kekurangan dari *CART* yaitu model *CART* kurang stabil yang mana jika terdapat sedikit perubahan pada data latih akan berpengaruh besar pada model pembelajaran yang dihasilkan. Pada penelitian ini dilakukan beberapa metode pengujian untuk mendapatkan nilai *hamming loss* terbaik. Berdasarkan beberapa pengujian yang telah dilakukan didapatkan nilai hamming loss terbaik sebesar 0,1914 atau sebesar 80,86%. Hasil tersebut menunjukan bahwa penggunaan *bagging* dapat membantu meningkatkan akurasi sebesar 5%.

Kata kunci : Klasifikasi, hadis bukhari, preprocessing, ekstraksi fitur, CART, Bagging