

Abstrak

Al-Hadits merupakan kumpulan dari sabda, perbuatan, ketetapan, dan persetujuan Rasulullah *Shallallahu 'Alaihi wa Salam* yang merupakan sumber hukum Islam kedua setelah Al-Qur'an. Sebagai dasar agama Islam, Muslim wajib mempelajari, menghafalkan, dan mengamalkan Al-Quran dan Al-Hadits. Satu dari imam besar sekaligus orang yang meriwayatkan Al-Hadits adalah imam Bukhari. Beliau menghabiskan waktu selama 16 tahun dalam meriwayatkan Al-Hadits yang jumlahnya sebanyak 2602 Hadits tanpa perulangan dan lebih dari 7000 jika dengan perulangan[9].

Kategorisasi teks otomatis merupakan sebuah kegiatan membangun perangkat lunak yang mampu mengklasifikasikan teks dokumen atau *Hypertext* ke dalam kategori atau kode subjek yang sudah ditentukan sebelumnya. Algoritma yang akan digunakan adalah *Random Forest* yang merupakan perkembangan dari *Decision Tree*.

Dalam penelitian tugas akhir ini, penulis memutuskan untuk membuat sebuah sistem yang mampu mengkategorisasikan teks dokumen yang memuat Hadits yang diriwayatkan oleh imam Bukhari berdasarkan kategori anjuran, larangan, dan informasi. Adapun dalam metode evaluasinya, perhitungan *K-Fold Cross Validation* dengan *F1-Score* yang didapat sebesar 90%.

Kata Kunci: Kategorisasi Teks Dokumen, Hadits Sahih Al-Bukhari, *Random Forest*, *K-fold cross validation*, *micro average f1-score*.