

ABSTRAK

Hadis adalah sumber hukum fundamental Islam kedua setelah Al-Qur'an. Hadis perlu dikategorikan agar mudah memahami apa tujuan dari hadis itu. Kami menggunakan teks Hadis Sahih Al-Bukhari yang diterjemahkan bahasa Indonesia. Hadist memiliki karakteristik teks yang unik karena hadits memiliki dua bahasa (Indonesia dan Arab) yang ada dalam satu dokumen. Dalam penelitian ini, klasifikasi teks digunakan untuk mengkategorikan hadits terjemahan bahasa Indonesia. Seperti yang sudah dikatakan sebelumnya dalam satu hadits mungkin memiliki dua bahasa. Masalah tersebut perlu diselesaikan dengan pemrosesan teks khusus untuk mengetahui makna setiap katanya. Berdasarkan masalah tersebut, penelitian ini dilakukan dalam dua fase utama untuk klasifikasi teks. Fase pertama adalah preprocessing yang melibatkan penyisipan kata dengan Skip-Gram Negative-Sample (SGNS) untuk mendapatkan konteks setiap kata. Fase kedua adalah klasifikasi dengan membandingkan Backpropagation (BP) dan Support Vector Machine (SVM). Berdasarkan hasil evaluasi, diketahui bahwa sistem mampu melakukan klasifikasi teks terbaik menggunakan BP dengan F1-Score tertinggi pada 82,05%, sedangkan SVM hanya mendapatkan F1-Score tertinggi pada 76.06%.