

ABSTRAK

Abstrak - Hadis merupakan pedoman dan sumber ajaran Islam yang kedua setelah Al-Qur'an. Salah satu kitab hadis yang paling shahih adalah kitab Shahih Bukhori. Hadis Shahih Bukhari memiliki rantai perawi, nomor hadis dan isi kandungan yang berbeda. Hadis ini juga terdapat ilmu yang membahas sejarah para perawi hadis yang disebut dengan Ilmu Rijalul Hadis. Pada hadis Shahih Bukhori terdapat nama-nama periwayat hadis yang mempunyai kesamaan nama, sehingga menimbulkan ambiguitas antar nama. Hal itu membuat banyak orang awam sulit mengerti akan nama-nama yang ambigu tersebut, karena belum diketahui apakah kedua nama yang sama tersebut benar orang yang sama atau tidak. Sehingga, menimbulkan permasalahan suatu ambiguitas nama bagi orang-orang awam yang tidak dapat membedakan apakah nama perawi tersebut merupakan orang yang sama atau tidak. Untuk memecahkan permasalahan tersebut dibangun solusi yaitu disambiguasi nama guna menghilangkan ambiguitas nama dengan mengecek nama, nomor hadis, rantai perawi, isi, topik, kalangan, negeri dan sahabat Rasulullah yang dilihat dari 3 nama terakhir sebelum Rasulullah berdasarkan rantai perawi. Selain itu, solusi tersebut dibantu dengan menggunakan metode Mining Evidence dengan beberapa pendekatan lainnya yaitu document label association, label word association, context similarity, cosine similarity, dan word2vec untuk memperoleh seluruh nilai similarity antar entitas nama. Setelah nilai similarity didapat, data tersebut dikelompokkan menggunakan algoritma Clustering. Sistem ini dibangun diharapkan dapat menghasilkan performansi sistem yang baik dengan confusion matrix berdasarkan nilai precision, recall dan accuracy.

Keywords: Disambiguasi, Entitas Nama, Mining Evidence, Sahih Bukhari, Similarity