
Abstrak

Hadits memiliki beberapa tingkat keaslian, di antaranya adalah hadits yang lemah (dhaif), dan hadits palsu (maudhu) yang mungkin tidak berasal dari Nabi Muhammad SAW, dan dengan demikian tidak boleh dianggap dalam menyimpulkan hukum Islam (syariah). Namun, banyak hadits semacam itu yang pada umumnya dikacaukan sebagai hadits otentik di kalangan Muslim biasa. Untuk dengan mudah membedakan hadits-hadits semacam itu, makalah ini mengusulkan suatu metode untuk memeriksa keaslian hadits dengan membandingkannya dengan koleksi hadits-hadits palsu dalam bahasa Indonesia. Metode yang diusulkan menggunakan model ruang vektor dan juga melakukan koreksi ejaan menggunakan symspell untuk memeriksa apakah penggunaan pemeriksaan ejaan dapat meningkatkan keakuratan pengambilan hadits, karena belum pernah dilakukan dalam karya sebelumnya dan kesalahan ketik umum pada hadits yang diterjemahkan dalam bahasa Indonesia pada teks mentah web dan media sosial. Hasil percobaan menunjukkan bahwa penggunaan pemeriksaan ejaan meningkatkan rata-rata presisi dan recall menjadi masing-masing 81% (dari 73%) dan 89% (dari 80%). Oleh karena itu, peningkatan keakuratan dengan menerapkan koreksi ejaan membuat sistem pengambilan hadits lebih layak dan didorong untuk diimplementasikan dalam karya-karya ke depan karena dapat memperbaiki kesalahan ketik yang umum dalam teks mentah di Internet.

Keywords: hadits, model ruang vector, symspell, perbaikan ejaan, pengambilan dokumen
