

Abstraksi

Nama perawi dalam Hadits merupakan salahsatu komponen terpenting dalam menentukan validitas suatu hadits, akan tetapi dengan banyaknya jumlah Hadits yang ada, menyebabkan proses penentuan validitas suatu Hadits secara manual berdasarkan nama perawi menjadi sulit, khususnya pada Hadits terjemahan bahasa Indonesia. Named Entity Recognition merupakan suatu metode yang bertujuan untuk menemukan entitas dalam suatu dokumen teks, dalam hal ini entitas meliputi nama orang, lokasi, organisasi, dsb. Paper ini akan membahas mengenai penerapan Named Entity Recognition pada koleksi Hadits terjemahan Bahasa Indonesia untuk menemukan nama perawi dari setiap Hadits. Pada paper ini sebanyak 200 Hadits yang berasal dari 9 kitab yang berbeda dan terdiri dari 2241 entitas nama perawi akan digunakan sebagai dataset. Dengan bentuk entitas yang bervariasi dan jumlah data yang digunakan maka pada Paper ini akan digunakan pendekatan supervised-learning, dan untuk memaksimalksimalkan performansi dari sistem Named Entity Recognition ini maka Support Vector Machine dipilih sebagai model classifier yang diketahui memiliki kemampuan generalisasi yang baik dalam mengklasifikasi data serta kemampuan dalam mengatasi data berdimensi tinggi. Beberapa kombinasi skenario pengujian pada model NER menunjukkan hasil F-1 tertinggi yaitu 0.9 dengan data training sebanyak 140 Hadits yang terdiri dari 1564 entitas dan testing sebanyak 60 Hadits yang terdiri dari 677 entitas. Nama perawi yang dihasilkan sistem NER ini kemudian akan digunakan sebagai indeks dari Hadits-hadits yang telah diriwayatkan oleh perawi tersebut dengan menggunakan metode Inverted Index.