

ABSTRAK

KLASIFIKASI *MULTI-LABEL* TEKS HADIST DENGAN MENGGUNAKAN K-NEAREST NEIGHBOR DAN DIINTEGRASIKAN LABEL POWERSSET

Oleh
Ian Aprilianto Said
16102196

Hadits merupakan salah satu dasar dari agama Islam selain Al-Qur'an, Hadits juga merupakan makna dari seluruh sikap, perkataan, dan perbuatan Rasulullah SAW dalam mengajarkan agama Islam. Hadits memiliki berbagai macam makna, Hadits yang berisi anjuran, informasi, serta larangan yang ditunjukkan untuk seluruh umat Islam. Pada hadits terdapat beberapa teks hadits yang berisi 3 makna dalam 1 hadits tersebut, atau bisa disebut juga dengan *multi-label*. Dengan jumlah teks hadits yang sangat banyak, saat ini sangat membutuhkan waktu untuk memberikan label pada setiap satu teks hadits, karena itu dibuat suatu model klasifikasi yang dapat mengelompokkan teks hadits dengan *multi-label*. Penelitian ini menggunakan *dataset* berupa hadits yang memiliki 1064 teks yang sudah dilakukan pelabelan yang sesuai. Pada penelitian sebelumnya telah dilakukan klasifikasi dengan menggunakan *dataset* yang sama berupa teks hadis shahih bukhari dengan hasil f1-score sebesar 0.8539 dengan pengujian menggunakan k-nearest neighbor dengan binary relevance dan hasil hamming loss sebesar 0.178403 dengan pengujian menggunakan convolutional neural network dengan label powerset. Pada dasarnya KNN merupakan algoritma yang berfokus pada kasus pengenalan pola, namun bisa juga digunakan untuk melakukan klasifikasi single-label, karena itu pada penelitian ini digunakan metode label powerset untuk mengklasifikasikan *multi-label* pada algoritma KNN dalam melakukan klasifikasi *multi-label* dari teks hadis shahih bukhari. Pada pengujian ini didapatkan hasil hamming loss terbaik dari penggunaan metode label powerset dengan nilai $k=4$ dan nilai rata-rata hamming loss 0.0795, dan hasil penggunaan metode binary relevance dengan nilai $k=5$ dan nilai rata-rata hamming loss 0.0821, sehingga dapat disimpulkan bahwa penggunaan metode binary relevance tidak lebih baik dibanding penggunaan label powerset pada penelitian ini.

Kata kunci : Hadis, KNN, *Label Powerset*, *Binary Relevance*, Klasifikasi Multi-label