

Abstrak

Pesatnya penggunaan Internet mengakibatkan pertumbuhan dan pertukaran informasi yang sangat cepat. Informasi yang terdapat pada Internet sangat berlimpah dan beragam, sehingga mengakibatkan jumlah informasi terus meningkat secara eksponensial. Perkembangan informasi ini memungkinkan informasi dapat diakses dengan mudah oleh pengguna. Namun, dengan jumlah informasi yang terus bertambah setiap harinya menimbulkan masalah dan tantangan yang cukup besar. Oleh karena itu, diperlukan suatu pengkategorisasian terhadap artikel berita yang memudahkan pengguna untuk mencari artikel yang diinginkan. Salah satu cara yang dapat mengkategorikan dokumen adalah dengan menggunakan teknik kategorisasi dalam data mining.

Akan tetapi jumlah dimensi yang besar membuat performansi classifier kurang baik. Untuk mengatasinya digunakan teknik *feature selection*. Pada tugas akhir ini, digunakan pendekatan feature selection dengan *wrapper feature selection*. Sedangkan metoda pencarian subset untuk *wrapper* adalah *hill-climbing search* dan *best first search* dengan menggunakan teknik klasifikasi *Naive Bayes* dari tools WEKA 3.5. Pencarian feature subset dilakukan dengan menghitung nilai *macro average F-measure* dari setiap node dan akan dihasilkan feature terbaik.

Kata kunci : *feature selection, wrapper feature selection, best first search, hill climbing search, macro average F-measure*