

Abstrak

Pemanfaatan informasi dalam kalimat opini di surat kabar bukanlah hal yang mudah dilakukan karena banyak kasus anaphora kalimat dalam kalimat di surat kabar. Dalam tugas akhir ini akan dilakukan identifikasi terhadap satu jenis anafora disebut *definite noun phrase anaphora* dengan menggunakan *Support Vector Machine* dengan *feature-Based Radial Basis Function* (RBF). Sistem ini menggunakan *feature set* yang digunakan untuk membentuk vektor fitur sebagai masukan kepada *Support Vector Machine* (SVM). Sistem ini dibagi menjadi beberapa proses, yaitu proses *preprocessing*, proses ekstraksi fitur, proses pelatihan dan proses pengujian.

Dalam proses pengujian, akan dilakukan 4 jenis pengujian, pertama adalah *cross validation* untuk menemukam model data yang paling baik[16]. Pengujian yang kedua adalah pengujian parameter. Pada pengujian parameter digunakan data hasil *cross validation* untuk menemukan yang kombinasi parameter terbaik untuk menemukan akurasi tertinggi. Dalam pengujian nilai C ditemukan bahwa semakin besar nilai C maka semakin tinggi nilai akurasi. Sedangkan untuk pengujian parameter gamma ditemukan nilai gamma dapat menyebabkan akurasi menjadi stabil atau konsisten pada nilai tertentu sesuai dengan nilai C. Pada pengujian selanjutnya ditunjukkan bahwa persebaran data yang digunakan pada tugas akhir ini tidak mempengaruhi akurasi. sedangkan pada pengujian terakhir menunjukkan bahwa akurasi sistem mencapai 100% berdasarkan pengujian *white-box*.

Kata Kunci: *definite noun phrase anaphora*, RBF, SVM, *feature set*