

Abstrak

Sentimen pada sebuah kalimat merupakan nilai emosi yang terkandung didalam kalimat dapat berisi emosi positif ataupun negatif dan biasanya sebuah kalimat yang tidak memiliki nilai sentiment dianggap sebagai kalimat yang bersentimen netral. Sentiment ini bisa di berbagai lingkungan seperti *review* pada sebuah produk, status seseorang di media social dan pada kalimat sitasi didalam karya ilmiah. Umumnya dalam karya ilmiah pembuatan kalimat sitasi cenderung dibuat dengan orientasi netral, sehingga dapat menghindari kritik secara langsung pada karya ilmiah yang diacu. Tetapi tidak menutup kemungkinan untuk pembuatan kalimat sitasi dalam kalimat bersentimen positif maupun negatif. Nilai sentimen didalam kalimat sitasi dapat digunakan untuk mengetahui kualitas dari sebuah karya ilmiah, mendeteksi relasi antar karya ilmiah, summarisasi dan lainnya. Banyak metode yang dapat diimplementasikan untuk melakukan analisis sentimen seperti untuk menentukan *classifier* terbaik dan juga fitur terbaik. Dikarenakan banyaknya metode yang dapat dipakai maka muncul permasalahan tentang menemukan kombinasi *classifier* dan fitur yang digunakan untuk menghasilkan performansi yang tinggi. Di dalam penelitian sebelumnya sudah didapatkan kombinasi yang terbaik dalam mengklasifikasi sentiment pada kalimat sitasi. Akan tetapi ada fitur yang potensial yang tidak diujikan. Dalam penelitian ini berfokus pada analisis performansi dalam penggunaan kombinasi *classifier* dengan kombinasi fitur serta melengkapi kombinasi fitur yang belum diuji pada penelitian sebelumnya. Hasil dari penelitian ini didapatkan bahwa kombinasi fitur *1-3 Grams +POSTag* dengan *classifier Support Vector Machine* dan kombinasi *1-3Grams + Dparser* dengan *classifier Logistic Regression* sebagai kombinasi fitur dan *classifier* yang terbaik. Hasil kombinasi tersebut menjadi yang terbaik pada saat pengujian untuk *5-fold validation* dimana untuk setiap *fold* dibentuk secara manual untuk dapat menyerupai kejadian riilnya. Hasil dari ke dua kombinasi tersebut memiliki *micro-f* lebih dari 99%.

Kata Kunci: sentimen, performansi, sitasi, *classifier*, fitur