

Abstrak

Saat ini semakin banyak para pedagang yang menawarkan produknya melalui internet. Mereka biasanya meminta pelanggan untuk memberikan *review* atas produk yang dibelinya. Hal tersebut bertujuan untuk mempermudah calon pembeli lain menentukan produk apa yang akan dibelinya. Banyaknya *review* yang diberikan membuat para calon pembeli kesulitan dalam membaca satu persatu *review* yang ada. Oleh sebab itu, pada Tugas Akhir ini dibangun sebuah sistem yang dapat melakukan suatu *mining* dalam *review* yang ada tersebut.

Proses *mining* yang dilakukan disebut *opinion mining* atau analisis sentimen. Dimana hasil yang diharapkan adalah mengetahui apakah suatu kalimat tersebut merupakan kalimat opini positif atau negatif. Tugas Akhir ini berorientasi pada *feature-based opinion* dimana orientasi opini ditentukan, sebelum terlebih dahulu mengekstrak entitas produk yang dikomentari. Proses pertama yang dilakukan adalah mengekstrak fitur beserta opini yang mengikutinya, kemudian mengekstrak kalimat opini, dan terakhir menentukan orientasi opini (positif atau negatif). Metode yang digunakan dalam sistem ini adalah *Lexicalized Hidden Markov Model (L-HMM)* yang mengintegrasikan fitur linguistik, seperti *part-of-speech*, menjadi sebuah mesin pembelajaran otomatis. Sistem ini diharapkan mampu mengenali kosakata baru yang belum muncul pada data latih, yang akan diterapkan bersama dengan teknik *smoothing*. Dalam memperoleh data train beserta fitur, opini dan orientasinya dalam jumlah banyak, dilakukan teknik bootstrapping.

Hasil pengujian yang diberikan menunjukkan bahwa teknik bootstrapping dapat digunakan untuk memperkaya data latih yang digunakan. Selain itu, penggunaan L-HMM mampu mengekstraksi dengan baik fitur beserta orientasi opini dalam suatu kalimat.

Kata Kunci : *data mining, sentiment analysis, feature-based opinion, part-of-speech, lexicalized hidden markov models (L-HMM)*