

Abstrak

Ulasan produk dari konsumen merupakan salah satu faktor yang penting dalam penjualan suatu produk. Menganalisis klasifikasi sentimen dan peringkasan suatu *review* produk memiliki tiga tahap yang harus dilakukan. Tahap pertama yaitu ekstraksi fitur menggunakan *frequent itemset mining* dengan algoritma apriori. Kemudian, dilakukan klasifikasi opini menggunakan SentiWordnet untuk penentuan polaritas kata opini. Tahap terakhir yaitu dilakukan peringkasan menggunakan semantic similarity scoring dan sentence clustering. Hasil dari penelitian ini didapat bahwa *filtering* kata yang sesuai juga mempengaruhi performansi dari ekstraksi pada penelitian ini. *Filtering* kata digunakan yaitu *Noun*, *Noun Phrase*, irisan serta gabungan keduanya, keempat *filtering* kata yang digunakan memiliki hasil yang cukup seimbang, gabungan dan irisan yang diharapkan dapat meningkatkan performansi juga masih didapat hasil yang tidak terlalu jauh dengan hanya *Noun* dan *Noun Phrase*. Hasil performansi ekstraksi pada penelitian ini adalah sekitar 20-40% pada dataset yang digunakan. Klasifikasi menggunakan SentiWordNet menunjukkan hasil performansi yang cukup baik namun pada beberapa dataset yang memiliki kompleksitas kalimat yang cukup tinggi juga terjadi penurunan walaupun tidak terlalu berbeda jauh dan masih pada sekitaran 40-90%. Peringkasan dokumen dapat dilakukan dengan baik pada dataset yang disediakan karena dataset memiliki jumlah kalimat ulasan produk yang memadai dan peringkasan dengan metode yang digunakan memperlihatkan beberapa representasi kalimat dari *clustering* dengan baik.

Kata kunci: analisis sentimen, ulasan produk, *frequent itemset mining*, *association mining*, *semantic smilarity scoring*, *sentence clustering*.