

Abstrak

Kebebasan konsumen dalam mengemukakan pendapat di media menyebabkan ulasan menjadi menumpuk dan sulit menyaring fitur dari pembahasan produk. Oleh karena itu berdasarkan permasalahan tersebut pada tugas akhir ini topik yang diangkat adalah *sentiment analysis* namun fokusnya pada ekstraksi fitur dari ulasan produk. Pada penelitian sebelumnya sudah dilakukan ekstraksi fitur produk terhadap ulasan konsumen dalam bahasa Inggris menggunakan metode *frequent itemset mining* pada algoritma apriori dan *pruning*. Performansi yang didapat berkisar 20%-40%, karena masih terdapat fitur produk yang belum terekstrak. Oleh karena itu pada tugas akhir ini dilakukan penelitian dengan menambahkan kombinasi *coreference resolution* dan *iob-chunking* dalam *preprocessing*. Ekstraksi fitur yang dilakukan adalah *frequent itemset mining* pada *association mining* menggunakan algoritma apriori, kemudian ekstraksi *pruning* dari hasil ekstraksi *frequent itemset mining*, selanjutnya ekstraksi fitur-fitur yang teridentifikasi memiliki opini dari hasil *pruning*, dan terakhir ekstraksi taksonomi fitur dengan memanfaatkan *lexicon wordnet 2.1*. Pada penelitian ini dari penambahan kombinasi *preprocessing* yang dilakukan menghasilkan kenaikan performansi pada metode *frequent itemset mining* dan *pruning* sebesar 20%-57%. Ekstraksi fitur opini yang dilakukan memberikan performansi sebesar 6%-19%. Hal ini karena dipengaruhi oleh hasil fitur *pruning* yang berulang-sera ada fitur yang tidak memiliki opini. Pada ekstraksi fitur taksonomi dari hasil fitur yang memiliki opini performansi yang didapat sebesar 1%-11.98% dan ekstraksi fitur taksonomi menggunakan *feature expert judgement* hasilnya sebesar 3%-6%, hasil ekstraksi taksonomi dengan menggunakan fitur yang mempunyai opini lebih besar dibanding menggunakan *feature expert judgement*, hal ini disebabkan karena ekstraksi taksonomi ini sudah menggunakan seluruh fitur asli pada dataset sehingga fitur yang dicari taksonominya lebih bervariasi.

Kata kunci: *Frequent Itemset Mining, Coreference Resolution, IOB Chunking, ekstraksi fitur yang memiliki opini, Lexicon Taxonomi-Based.*