

1. Pendahuluan

1.1. Latar Belakang

Dengan adanya *review* produk, informasi mengenai kelebihan dan kekurangan terhadap sebuah produk akan lebih bervariasi. Informasi mengenai kelebihan dan kekurangan terhadap sebuah produk yang bervariasi pada *review* produk, merupakan bahan analisis yang baik bagi sebuah perusahaan terkait untuk melakukan pengembangan terhadap produk tersebut. Tetapi semakin banyaknya informasi dari berbagai macam *review* produk yang dikumpulkan juga menyebabkan masalah bagi perusahaan. Dikarenakan hal tersebut akan menyulitkan perusahaan terkait untuk menganalisis informasi dari suatu produk yang diulas dalam banyak *review* produk.

Pada penelitian sebelumnya telah ditemukan solusi untuk menyelesaikan permasalahan tersebut, yaitu dikembangkan sebuah sistem yang mampu memberikan sebuah ringkasan dari sebuah *review* produk dengan menggunakan *opinion summarization*. Secara garis besar, *opinion summarization* melakukan identifikasi dan ekstraksi fitur produk serta identifikasi polaritas opini yang berhubungan dengan fitur produk yang berhasil diekstrak dari sebuah *review* produk. Perlu diketahui pada penelitian sebelumnya proses ekstraksi fitur dari sebuah *review* produk dilakukan menggunakan metode *Association Rule Mining* dengan algoritma apriori dan *pruning*. Metode tersebut mampu mengidentifikasi *frequent itemset* yang cenderung merupakan sebuah fitur produk dengan memperhatikan frekuensi kemunculan *itemset* tersebut dalam sebuah *database* serta mengeliminasi kandidat fitur produk yang berhasil terekstrak karena dianggap kurang relevan [1].

Akan tetapi proses ekstraksi fitur yang dilakukan pada penelitian sebelumnya masih ditemukan beberapa fitur yang belum terekstrak dikarenakan beberapa faktor seperti keterbatasan ruang lingkup metode *preprocessing* data yang digunakan dan belum diterapkannya proses identifikasi yang mampu menangani masalah ungkapan kata ganti untuk entitas yang sama pada kalimat dalam *dataset*. Oleh karena itu perlu dilakukan pengembangan terhadap proses ekstraksi fitur

sebelumnya guna memperoleh hasil analisis apakah faktor-faktor tersebut berpengaruh terhadap hasil ekstraksi fitur yang dilakukan.

Berdasarkan permasalahan di atas, pada tugas akhir ini akan dilakukan penambahan identifikasi *noun phrase* pada proses *preprocessing* data menggunakan *noun phrase parser* serta penerapan *coreference resolution* untuk menangani identifikasi kata ganti yang mengacu pada entitas yang sama dalam sebuah kalimat. Analisis juga akan dilakukan hingga tahap identifikasi orientasi opini fitur dengan menggunakan *opinion lexicon* yang berbeda guna mengetahui pengaruh *opinion lexicon* terhadap identifikasi orientasi fitur opini yang diperoleh. Diharapkan dengan dilakukannya hal tersebut dapat menangani permasalahan pada analisis yang dilakukan sebelumnya.

1.2. Perumusan Masalah

Tugas akhir ini memiliki beberapa batasan masalah seperti; berfokus untuk melakukan analisis pada proses ekstraksi fitur pada tugas akhir Adina Kusuma [1]. Penggunaan *dataset* yang diperoleh dari penelitian yang dilakukan oleh Hu Mingqing dan Bing Liu [2]. Selain itu dalam tugas akhir ini perbandingan *opinion lexicon* tidak dilakukan hingga tahap klasifikasi melainkan hanya sampai pada tahap identifikasi kata opini yang merupakan pasangan fitur opini yang diperoleh pada tahap identifikasi orientasi fitur opini. Berdasarkan latar belakang yang telah dipaparkan sebelumnya terdapat susunan rumusan masalah sebagai berikut:

- a. Bagaimana pengaruh penerapan *coreference resolution* dan *noun phrase parser* yang ditambahkan pada tahap *preprocessing* data terhadap hasil ekstraksi fitur menggunakan pendekatan *association rule mining* ?
- b. Bagaimana pengaruh perbandingan *opinion lexicon* yang digunakan pada tahap identifikasi orientasi fitur opini ?

1.3. Tujuan

Berdasarkan susunan rumusan masalah diatas dapat diperoleh beberapa tujuan dalam tugas akhir ini, diantaranya adalah:

- a. Melakukan analisis terhadap pengaruh penerapan *coreference resolution* dan *noun phrase parser* yang ditambahkan pada tahap *preprocessing* data terhadap hasil ekstraksi fitur menggunakan pendekatan *association rule mining*.

- b. Melakukan analisis terhadap pengaruh perbandingan opinion lexicon yang digunakan terhadap proses identifikasi orientasi opini fitur.

1.4. Metodologi Penyelesaian Masalah

Berikut ini adalah tahap-tahap yang dilakukan dalam metodologi penyelesaian masalah pada tugas akhir:

- a. Tahap Studi Literatur

Pada tahap ini dilakukan kegiatan studi yang terkait dengan materi *opinion extraction*, *feature extraction* dan *association rule mining* dari berbagai macam referensi yang ada.

- b. Analisis dan Perancangan Sistem

Pada tahap ini dilakukan kegiatan analisis dan perancangan sistem berdasarkan referensi yang telah diperoleh pada tahap sebelumnya.

- c. Implementasi Metode

Pada tahap ini dilakukan implementasi metode yang digunakan dalam melakukan ekstraksi fitur produk. Metode tersebut adalah *Association Rule Mining*.

- d. Pengujian

Pada tahap ini dilakukan pengujian sistem terhadap *dataset* yang digunakan. Pengukuran yang dilakukan pada pengujian ini akan menggunakan evaluasi *precision* dan *recall* sesuai dengan yang dilakukan pada penelitian sebelumnya.

- e. Analisis Hasil Pengujian dan Kesimpulan

Pada tahap ini dilakukan analisis terhadap hasil penelitian dan membuat kesimpulan berdasarkan hasil analisis tersebut.

- f. Pembuatan Laporan

Pada tahap ini dilakukan pembuatan dokumen yang berisi hasil penelitian yang dilakukan selama mengerjakan tugas akhir.

1.5. Sistematika Penulisan

Berikut ini adalah gambaran singkat mengenai informasi yang ada pada setiap bab dalam laporan tugas akhir ini.

a. BAB 1 PENDAHULUAN

Bab ini berisi gambaran ruang lingkup penelitian yang dilakukan dalam tugas akhir yang dituangkan kedalam beberapa sub bab yaitu latar belakang, perumusan masalah, tujuan, metodologi penyelesaian masalah dan sistematika penulisan.

b. BAB 2 KAJIAN PUSTAKA

Bab ini berisi sumber, acuan atau teori pendukung yang digunakan dalam penelitian pada tugas akhir ini.

c. BAB 3 PERANCANGAN SISTEM

Bab ini berisi penjelasan mengenai ruang lingkup pengerjaan sistem yang dilengkapi dengan alur proses pada sistem secara detail.

d. BAB 4 PENGUJIAN DAN ANALISIS

Bab ini berisi pengujian dan hasil analisis yang dilakukan setelah melakukan penelitian.

e. BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dan saran penulis terhadap penelitian yang telah dilakukan pada tugas akhir ini.