Mendeteksi Otomatis Aspek Produk

dari Ulasan Pelanggan

Fakry Adi Permana (1301164034) Anggi Yuniar Putri (1301154492)  
Program Studi Teknik Informatika Telkom University

*Abstract*—Makalah ini mengusulkan pendekatan baru untuk meendeteksi aspek secara otomatis. Tujuan dari algoritma yang kami gunakan untuk memperoleh ringkasan aspek dari produk tertentu, yang berasal dari ulasan pelanggan. Pendekatan kami dimulai dengan mencocokkan *dependency path* buatan dalam kalimat untuk menemukan aspek kandidat pada pendapat yang diungkapkan. Pendekatan kami tidak membutuhkan sepatah katapun … atau pengetahuan khusus domain, karena hanya mempekerjakan leksikon sentiment. Yang kami bahas hanya hasil dari akurasi yang telah dilakukan berdasarkan data dari ulasan pelanggan.

Keywords—aspek, dependency path, lexicon, sentiment.

# Pendahuluan

Pada saat ini minat seseorang untuk menganalsis telah bergeser ke arah analisis yang lebih kualitatif, yang bertujuan untuk mendeteksi aspek-aspek yang berbeda dari suatu topik ke arah mana suatu pendapat tersebut diungkapkan. Misalnya, dalam menganalisis ulasan sebuah produk. Dalam pengulasannya terhadap aspek-aspek seperti kondisi barang, kinerja produk, dan kualitas produk tersebut.

Dalam tulisan ini kami memperkenalkan pendekatan baru yaitu Aspektaktor, sebuah algoritma baru mendeteksi dan menilai aspek produk dari ulasan pelanggan. Namun saat ini kami batasi pengerjaan hanya sampai mendeteksi aspek. Aspektaktor bisa menemukan aspek kandidat dengan mencocokkan sintaks *dependency path*, sedangkan pendekatan lain membutuhkan inputan kata-kata awal dan menggunakan sintaks *dependency* atau untuk menemukan kata baru dan hubungannya. Akibatnya, sistem yang diusulkan dapat mendeteksi dan menilai aspek produk dalam domain apapun. Secara konkret Aspektaktor menggabungkan aspek kandidat yang telah diesktraksi dari kalimat individu melalui sintaks *dependency path*, dengan mencari akurasi, presisi dan recall dari hasil aspek-aspek yang telah dikeluarkan sistem berdasarkan data dari ulasan pelanggan.

# Tinjauan Pustaka

Pada tahun 2004 Hu dan Liu memperkenalkan masalah yang lebih menarik dari analisis sentiment berbasis aspek, dimana polaritas tidak ditugaskan untuk kalimat atau dokumen tetapi untuk aspek tunggal yang dibahas dalamnya. Dalam pendekatannya diberikan sejumlah ulasan besar pada produk tertentu, pertama-tama melakukan pengidentifikasian aspek-aspek produk yang menarik, menggunakan asosiasi penambangan, kemudian melampirkan nilai sentiment untuk setiap aspek yang telah dieskploitasi dari sekumpulan kata-kata opini, bersamaan dengan kata sinonim dan antonim. Tahap selanjutnya mendeteksi kata pendapat yang kemudian di ekstraksi menjadi beberapa aspek-aspek.

Pekerjaan yang telah dibahas sejauh ini mengusulkan solusi dari domain independent yang diposisikan atau digunakan untuk menganalisis sentiment berbasis aspek. Dalam pengklasifikasiannya ulasan produk, membutuhkan inputan awal untuk membaca algoritma khusus mengenai *value* produk tersebut seperti kondisi produk, kualitas produk dan kinerja produk. Yang di ikuti dengan kata opini khusus sebagai domain untuk algorima mereka. Seperti yang diharapkan domain menghasilkan klasifikasi sentiment yang lebih akurat. Karena semua aspek domain yang relevan diberikan sebagai inputan, secara eksklusif pendeteksian mengenai opini masih berfokus tehadap aspek yang diberikan (1) Menangkap kata-kata opni baru melalui *dependency sintaks,* dan (2) Menilai aspek produk berdasarkan eksternal leksikon sentiment dan beberapa kata opini khusus yang telah diberikan domain.

Demikian pula,menurut Zhu, menggunakan aspek produk dan beberapa istilah terkait aspek sebagai input untuk algoritma mereka, tetapi kemudian berusaha untuk menemukan istilah terkait aspek baru dengan menerapkan algoritma bootstrap berdasarkan pada kejadian antara istilah kata utama dan istilah kandidat baru. Skor sentimen kembali diperoleh dengan mengakses leksikon sentimen eksternal. Sementara pendekatan kami mempertahankan dari karya-karya ini penggunaan leksikon eksternal, itu tidak memerlukan contoh berlabel atau pengetahuan khusus domain, sehingga memiliki penerapan yang lebih luas.

## Mendeteksi Aspek Produk

Langkah pertama adalah untuk mengekstrak dari ulasan pelanggan menyebutkan aspek produk dan kata-kata yang mengekspresikan pendapat penulis terhadap aspek tersebut. Misalnya, diberi kalimat:

“i'd say that this camera is perfect.”

Aspektaktor akan mengekstrak kalimat menjadi berikut:

< Camera >

Aspect

Mention

Aspectator mengekstraksi pasangan pendapat dengan menggunakan enam dependency path sederhana buatan tangan secara manual, dalam tiga langkah:

1. Untuk setiap kalimat, aspektator mengekstrak sintak dependency tree menggunakan stanford depedency parser
2. Jika ada dependency tree, aspektator mencoba untuk mengeluarkan pasangan pendapat dasar yang dibentuk oleh pengubah sentimen kata tunggal
3. Jika ada pasangan pendapat yang cocok, aspektator mencoba untuk memperluas kecocokan dengan kata terdekat dengan menerapkan depedency path tambahan

Teknik kami untuk mengekstraksi aspek menyebutkan dan pengubah sentimen menghasilkan recall tinggi, ketepatannya rendah, karena beberapa urutan kata yang tidak relevan ditangkap. Namun demikian, langkah-langkah berikut memungkinkan sistem kami untuk memberikan kepercayaan yang lebih rendah untuk ekstraksi yang salah, sehingga pada akhirnya menghasilkan ekstraksi peringkat atas yang akurat.

# Metodologi

Keluaran dari aspektator dievaluasi secara manual pada sebagian dari dua set data public dari domain yang berbeda oleh dua anotasi, fimana hanya satu yang merupakan penulis kedua. Dataset pertama adalah kumpulan ulasan produk yang merupakan penjualan sebuah kamera