

1. Pendahuluan

Latar Belakang

Pada saat ini, perkembangan teknologi terjadi secara pesat sehingga membuat informasi tentang jual beli menjadi lebih mudah didapatkan. Salah satunya adalah e-commerce yang menyediakan informasi tentang produk yang dijual ataupun dibeli oleh penggunanya. E-commerce menjadi pilihan oleh banyak orang untuk kegiatan jual beli secara daring karena, selain menyediakan informasi, e-commerce juga menjadi perantara pembayaran sehingga membuat penggunanya merasa aman dari penipuan. Pada e-commerce juga terdapat informasi tentang ulasan produk dari toko daring oleh pembeli yang telah melakukan transaksi sebelumnya. Ini tentunya akan membuat pengguna e-commerce lainnya menjadi lebih percaya untuk membeli produk dari toko daring yang mempunyai penilaian tinggi.

Beberapa e-commerce yang terkenal adalah Shopee, Tokopedia, Bukalapak, OLX dan lainnya. Pada tahun 2019, jumlah pengguna e-commerce di Indonesia diproyeksikan akan mencapai 168,3 juta pengguna dan terus meningkat mencapai 212,2 juta pada 2023 [1]. Semua ulasan produk pada e-commerce ini mempunyai pola yang sama, yaitu mengulas tentang produk ataupun toko dengan kalimat yang mempunyai sentimen positif atau negatif. Informasi ini dapat dikumpulkan menjadi data untuk menentukan toko daring mana yang mempunyai kualitas yang bagus. Analisis sentimen adalah teknik yang bisa dilakukan untuk menentukannya. Pada penelitian ini, data yang diambil berasal dari situs e-commerce Shopee sebagai sampel dari e-commerce.

Pada analisis sentimen, terdapat dua pendekatan algoritma yang bisa digunakan untuk menyelesaikan permasalahan ini. Pendekatan algoritma tersebut adalah pendekatan lexicon dan pendekatan learning. Pendekatan menggunakan lexicon bergantung pada kamus sentimen yang berupa data kata yang berisi kumpulan kalimat yang mempunyai nilai sentimen [2]. Sedangkan pendekatan menggunakan learning pendekatan yang menerapkan algoritma machine learning dengan menggunakan fitur linguistik yang ada [2]. Kedua pendekatan ini mempunyai kelemahan, maka untuk mengatasi kelemahan tersebut dibuatlah sistem dengan kombinasi dua pendekatan ini.

Topik dan Batasannya

Pendekatan lexicon bisa dipakai untuk memberi label pada data agar pelabelan bisa dilakukan secara otomatis atau untuk mengklasifikasi langsung suatu kalimat yang berisi sentimen. Tetapi, pendekatan lexicon terkadang tidak bisa mendeteksi skor dari beberapa kata opini karena kata tersebut tidak terdapat di dalam kamus opini yang dipakai. Skor positif dan skor negatif yang dihasilkan oleh pendekatan lexicon juga bisa sama besarnya yang membuat kebingungan dalam pemberian label. Maka dari itu, untuk menutupi kelemahan tersebut digunakan juga pendekatan learning yang sifatnya bisa melabeli data tanpa perlu kamus opini atau sistem skor.

Data yang digunakan pada penelitian ini berasal dari ulasan pada situs e-commerce Shopee yang berbahasa inggris. Data tersebut berupa kalimat tentang ulasan suatu produk yang dibeli dari suatu toko. Data akan diberi label positif (1) dan negatif (0) sesuai dengan sentimennya secara manual. Sistem analisis sentimen yang dibuat untuk mendeteksi data tersebut tidak bersifat real-time online sehingga jika ada data baru maka model dengan pendekatan learning harus dilatih ulang.

Tujuan

Untuk menganalisa apakah metode kombinasi yang digunakan lebih baik daripada hanya menggunakan pendekatan lexicon. Mempelajari proses metode kombinasi yang dipilih agar dapat

meningkatkan kinerja lexicon dalam permasalahan analisis sentimen. Melihat dampak penggunaan metode kombinasi dibanding hanya menggunakan pendekatan lexicon.

Organisasi Tulisan

Pada bagian 1 dijelaskan mengenai latar belakang masalah, perumusan masalah, dan tujuan dari penelitian. Pada bagian 2 dijelaskan landasan teori yang terkait dengan penelitian ini. Pada bagian 3 dijelaskan gambaran sistem yang di bangun, pada bagian 4 dijelaskan mengenai hasil dan analisis dari metode tersebut. Pada bagian 5 merupakan kesimpulan yang menjawab permasalahan pada penelitian ini serta saran dari penulis untuk penelitian selanjutnya.