

BAB I PENDAHULUAN

1.1 Latar Belakang

Di era *modern* saat ini banyak platform belanja online yang memanfaatkan teknologi dalam memasarkan barangnya, salah satu *marketplace* tersebut yaitu shopee. Aplikasi shopee merupakan *platform* belanja *online* terdepan di Asia Tenggara dan Taiwan. Aplikasi tersebut membantu para penjual dan pembeli untuk melakukan transaksi penjualan dan pembelian barang maupun jasa. *Platform* tersebut menyediakan banyak data teks singkat mengenai ulasan berbagai macam produk yang dijual. Kelas sentimen positif dan negatif yang terdapat pada ulasan *online* jelas mencerminkan persepsi pengguna tentang berbagai jenis layanan yang ada. Selain itu, pelanggan berpotensi yang membaca ulasan *online* dapat secara signifikan terpengaruh oleh polaritas sentimen dari ulasan yang tertera pada kolom ulasan. Hal ini menandakan ulasan yang bersentimen positif ataupun negatif yang ditinggalkan oleh pengguna, sangat mempengaruhi pengguna lainnya dalam memilih layanan ataupun produk yang dicari [1].

Analisis sentimen atau *opinion mining* merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini [2]. Analisis sentimen dapat digunakan untuk melakukan klasifikasi ulasan ke dalam sentimen tertentu dengan memperhatikan faktor pada objek sentimennya. Dalam analisis sentimen, ada sebuah metode yang lebih maju yaitu menggunakan *word embedding*, representasi kata dalam vektor, banyak peneliti yang telah menggunakan metode ini dalam penelitiannya. Penelitian terdahulu yang telah menggunakan *word embedding* menyatakan bahwa pada penggunaan *word embedding* akan mendapatkan hasil kualitas yang lebih baik dalam tugas pemrosesan teks tingkat lanjut, termasuk analisis sentimen [3].

Beberapa peneliti telah melakukan analisis sentimen menggunakan *word embedding* pada ulasan produk. Pada penelitian menggunakan metode *word*

embedding ini akan menghasilkan ruang vektor kata yang substruktur yang bermakna proses ini menghasilkan fitur dimensi tinggi space yang memrepresentasikan teks dalam dokumen dengan sekumpulan kata, dengan representasi vektornya dapat dibentuk menjadi semantik model distribusi dengan menangkap kata yang ada [4] Terdapat beberapa penelitian yang menyediakan vektor *embedding* yang sudah melalui proses pembelajaran (*pre-trained*) yaitu *Word2Vec* dan *Glove* [5]. Kedua metode ini dipilih karena dapat menangkap makna semantik, sintaksis, dan urutan bahkan konteks disekitar kata jika dibandingkan dengan *feature engineering* tradisional seperti *bag of words* [6]. Semakin populernya penggunaan *word embedding* digunakan dalam NLP (*Natural Language Processing*) menjadi motivasi untuk melakukan observasi kinerja dari setiap model *word embedding* yang ada. Oleh karena itu penulis mengajukan penelitian terkait *word embedding* yang berjudul "Perbandingan Metode *Word Embedding* untuk Analisis Sentimen pada Data Ulasan *Marketplace*".

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, dapat dirumuskan permasalahannya yaitu keberhasilan kinerja classifier terbaik dengan metode *word embedding Glove* dan *Word2vec* dalam representasi kata untuk menentukan sentimen suatu kalimat dalam ulasan *marketplace*?

1.3 Batasan Masalah

Pada penelitian ini akan meneliti pengelompokan sentimen dari ulasan *marketplace* berbahasa indonesia menggunakan *word embedding* dengan metode *Word2Vec* dan metode *Global Vector (GloVe)*.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, tujuan dilakukannya penelitian ini yaitu untuk mengetahui akurasi akhir dan *vocabulary* yang dihasilkan oleh *word embedding* menggunakan metode *Word2vec* lebih baik atau tidak dibandingkan dengan penggunaan metode *Glove* serta di klasifikasikan menggunakan *LSTM* dalam menganalisis sentimen ulasan *shopee* dalam bahasa Indonesia.

1.5 Manfaat Penelitian

Adapun manfaat penelitian adalah membantu melakukan generalisasi polaritas opini pada *user feedback* suatu aplikasi, mempercepat proses klasifikasi *user feedback* yang dilakukan secara otomatis menggunakan ekstraksi fitur khususnya *word embedding Word2Vec* dan *GloVe* serta klasifikasi *LSTM*.

1.6 Sistematika Penulisan

Untuk mengatahui uraian singkat yang memuat gambaran singkat secara keseluruhan isi masing-masing bab, maka dibuat sistematika sebagai berikut :

| | |
|----------------|--|
| BAB I | PENDAHULUAN Bab ini membahas tentang uraian latar belakang, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, manfaat penelitian, metode penelitian serta sistematika penulisan. |
| BAB II | LANDASAN TEORI Bab ini berisi tentang teori-teori yang digunakan serta konsep dasar yang menjadi landasan dari penelitian ini. |
| BAB III | METODOLOGI PENELITIAN Dalam bab ini berisi alur tahapan penelitian, alat berupa metode serta perangkat dan bahan berupa data yang akan digunakan dalam penelitian. |
| BAB IV | HASIL DAN PEMBAHASAN Bab ini membahas tentang hasil yang didapatkan dari penelitian yang sudah dilakukan oleh penulis. |
| BAB V | PENUTUP Berisi kesimpulan dari penelitian yang dibuat serta saran untuk memperbaiki kekurangan yang ada pada penelitian. |