

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Lazada merupakan salah satu platform *e-commerce* yang memiliki konsep *E-Marketplace* [1]. Namun ketika melakukan belanja secara *online* terdapat beberapa hal yang perlu diperhatikan, di antaranya mungkin tidak terdapat ketidaksesuaian produk yang sampai ke tangan pembeli dengan gambar yang ada pada situs. Kemudian barang juga mungkin diterima dalam kondisi yang rusak maupun terjadinya kesalahan pengiriman atau pengemasan sehingga terjadinya perbedaan produk yang sampai maupun produk tidak sampai. Terakhir, hilangnya barang pada saat pengiriman dan juga terjadi penipuan [2].

Karena hal tersebut, pengguna melakukan pertimbangan untuk melakukan pembelian di *e-commerce* dengan melihat ulasan yang telah diberikan pembeli sebelumnya terhadap produk yang ingin dibeli. Ulasan yang diberikan oleh pembeli sebelumnya menjadi salah satu pertimbangan utama dalam mengambil keputusan pembelian di suatu platform *e-commerce*. Jika suatu produk memiliki ulasan positif, kebanyakan calon pembeli akan merasa lebih yakin untuk membeli produk tersebut dan sebaliknya [3].

Analisis sebuah ulasan produk memang dapat secara mudah dilakukan dengan melihat jumlah bintang yang diberikan oleh pembeli, tetapi tidak selalu jumlah bintang tersebut mewakili keseluruhan isi ulasan. Sehingga diperlukan untuk melihat keseluruhan ulasan untuk mengetahui maksud dari sebuah ulasan. Sebuah ulasan dapat memiliki lebih dari satu penilaian terhadap sebuah produk, sementara jumlah bintang tidak dapat mewakili seluruh ulasan yang diberikan. Untuk menganalisa sebuah teks ulasan memang mungkin untuk dilihat satu persatu, namun jika jumlah ulasannya sudah menjadi terlalu banyak, akan lebih mudah jika menggunakan sebuah sistem untuk melakukan analisis sentimen secara otomatis [4].

Text mining merupakan proses pengambilan informasi berkualitas yang berasal dari teks. Teknik ini menghasilkan informasi berupa pola dan kecenderungan melalui pembelajaran pola statistik dan lainnya. Tujuan dari *text mining* adalah untuk mengolah data yang tidak terstruktur dari sebuah teks, dan menentukan sebuah pola dan keputusan. Sehingga, banyaknya jumlah teks akan

mempengaruhi tingkat kesulitan dalam menentukan pola dan membuat keputusan. Analisis sentimen merupakan salah satu contoh dari *text mining* [5, 6].

Analisis sentimen merupakan salah satu cara untuk mengetahui pendapat seseorang atau sekelompok orang mengenai suatu topik, produk, layanan, dan sebagainya. Analisis sentimen memiliki tujuan untuk menentukan polaritas dari suatu sentimen positif, negatif, dan netral [7]. Analisis sentimen dapat dilakukan dengan *Natural Language Processing* (NLP) yaitu melakukan teknik komputasi untuk mengotomatisasi ekstraksi atau klasifikasi sentimen dari sebuah ulasan sentimen. Analisis sentimen sudah digunakan di banyak bidang seperti informasi konsumen, pemasaran, aplikasi, web, dan media sosial. Tujuan utama dari analisis sentimen adalah untuk melakukan analisa dari sebuah ulasan dan menghitung skor dari sentimennya. Analisa sentimen dibagi ke dalam tingkat dokumen, tingkat kalimat, tingkat kata, dan tingkat aspek. Pada proses analisa, proses evaluasi sentimen merupakan hal yang penting karena terdapat evaluasi untuk menentukan tinggi nilai polaritas sentimen [8].

XLM-R merupakan salah satu model yang dapat digunakan dalam melakukan klasifikasi sentimen *multilingual*. XLM-R sendiri merupakan pengembangan dari XLM dan BERT. XLM-R dilatih menggunakan 100 bahasa, sehingga memiliki kemampuan untuk melakukan klasifikasi teks *multilingual* dan mendapatkan tingkat akurasi yang lebih tinggi dibandingkan dengan model klasifikasi *monolingual* [9].

Pada penelitian yang dilakukan, XLM-R digunakan untuk melakukan klasifikasi sentimen *multilingual* dengan *dataset* evaluasi yaitu sekumpulan ulasan Bahasa Indonesia pada kategori produk yang berbeda. Sayangnya masih banyak teks pada ulasan produk dalam bentuk bahasa yang informal. Selain bahasa informal, terdapat juga kalimat yang tidak terstruktur dengan baik [10]. Selain itu, Bahasa Indonesia yang merupakan turunan bahasa dari bahasa Melayu yang termasuk ke dalam *low-resource languages* [11]. *Low-resource languages* merupakan sebuah bahasa yang ketersediaan datanya sedikit untuk melakukan pelatihan *natural language processing* (NLP) [12]. Karena Bahasa Indonesia yang termasuk ke dalam *low-resource language*, maka model XLM-R digunakan karena sudah dilatih dengan 100 bahasa dan mendapatkan SOTA pada berbagai macam tugas *multilingual* pada NLP seperti XNLI, *Named Entity Recognition*, *Cross-lingual question answering*, dan *GLUE benchmark* dan dapat dengan baik melakukan klasifikasi bahasa dengan sumber daya yang rendah [9].

Data latih yang *multilingual* digunakan pada penelitian ini, karena model

multilingual yang dilatih juga dengan bahasa yang *multilingual* akan menghasilkan performa yang lebih baik dari pada model *monolingual* terlebih lagi jika data targetnya terbatas. Selain itu model *multilingual* juga dapat digunakan untuk melakukan *transfer learning* sehingga lebih mudah untuk melakukan pelatihan untuk analisa sentimen dengan *low-resource language* dibandingkan dengan melatih model dari awal [12].

Beberapa penelitian sebelumnya oleh Keenan, menunjukkan bahwa XLM-R dapat melakukan klasifikasi kategori teks berita *multilingual* dengan akurasi yang sangat baik, dengan nilai *accuracy*, *precision*, *recall*, dan *F1* sebesar 0.902 dengan model yang dilatih menggunakan *Indonesian News Dataset* berukuran besar (70.304) [13]. Terdapat penelitian lain yang membandingkan model *multilingual* dan *monolingual* untuk melakukan klasifikasi dan melakukan pelabelan teks Bahasa Indonesia menggunakan IndoBERT yang merupakan pengembangan dari model BERT [14]. Ada pun penelitian yang menggunakan model XLM-R untuk melakukan klasifikasi sentimen pada data dengan sumber daya yang rendah yaitu Bahasa Hindi dengan menggunakan konsep *zero-shot transfer learning*, dan menghasilkan nilai *recall* sebesar 73.5 dan *accuracy* sebesar 71.8 dengan data latih Bahasa Inggris [15]. Pada penelitian lain, dilakukan perbandingan performa antara model XLM-R dengan M-BERT, dan menyimpulkan bahwa tingkat kebenaran dari model yang digunakan sangat tergantung kepada pembagian *batch* pada *dataset*, dengan hasil *F1-Score* rata-rata 0.90 pada XLM-R dan *F1-Score* rata-rata 0.83 pada M-BERT [16]. Terakhir, terdapat penelitian yang membandingkan model XLM-R dengan M-BERT dan menggunakan skenario *fine tuning* dan *zero-shot transfer learning* dan mendapatkan hasil *error percentage* yang lebih besar pada model mBert dibandingkan dengan XLM-R pada skenario *fine tuning* dan pada skenario *zero-shot transfer learning*, XLM-R juga memiliki hasil evaluasi yang lebih unggul dibandingkan dengan mBert. Pada penelitian disimpulkan bahwa penggunaan XLM-R dengan skenario *zero-shot transfer learning* mendapatkan hasil yang cukup baik dalam melakukan *cross-lingual transfer learning* mengingat bahwa model tidak pernah melihat bahasa dari *dataset* evaluasi [17].

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah dalam penelitian “Implementasi Klasifikasi Sentimen Ulasan Produk E-Commerce Lazada Bahasa Indonesia Menggunakan XLM-R” adalah sebagai berikut.

1. Bagaimana cara mengimplementasikan model XLM-R untuk melakukan klasifikasi sentimen pada ulasan produk *e-commerce* Lazada?
2. Bagaimana performa akurasi dari model XLM-R untuk melakukan klasifikasi sentimen pada ulasan produk *e-commerce* Lazada?

1.3 Batasan Permasalahan

Terdapat beberapa batasan masalah dalam penelitian yang dilakukan. Adapun batasan-batasan masalah dari penelitian adalah sebagai berikut.

1. *Dataset* yang digunakan merupakan *dataset* "Lazada Indonesian Review" yang di dapat dari Kaggle dan juga *dataset* Amazon *multi review* yang diambil dari huggingface.co. Kedua *dataset* diambil pada tanggal 10 Maret 2023
2. Label yang akan digunakan pada klasifikasi sentimen pada penelitian ini akan dibagi menjadi 3 yaitu positif, netral, dan negatif?

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Melakukan implementasi model XLM-R untuk melakukan klasifikasi sentimen ulasan produk.
2. Menguji performa akurasi model XLM-R dalam melakukan klasifikasi sentimen pada ulasan produk *e-commerce* Lazada

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Menambah pengetahuan mengenai *Natural Language Processing* (NLP) di bidang *multilingual*.
2. Melakukan pengujian performa dari model XLM-R untuk melakukan klasifikasi multibahasa.

1.6 Sistematika Penulisan

Dalam penyusunan laporan skripsi ini, sistematika penulisan yang digunakan sebagai berikut:

1. Bab 1 PENDAHULUAN

Pada bab ini, dijelaskan latar belakang masalah, batasan masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan laporan skripsi.

2. Bab 2 LANDASAN TEORI

Pada bab dua, dijelaskan mengenai teori-teori yang digunakan dalam penelitian yang dilaksanakan. Teori yang dijelaskan dapat berupa konsep dan juga metode yang digunakan untuk mencapai hasil yang diinginkan.

3. Bab 3 METODOLOGI PENELITIAN

Pada bab ini, metode penelitian yang digunakan dijabarkan secara berurutan beserta rancangan sistem dalam bentuk diagram alir atau *flowchart*

4. Bab 4 HASIL DAN DISKUSI

Bab ini membahas spesifikasi perangkat yang digunakan selama penelitian berlangsung dan hasil pengujian dari implementasi model yang digunakan.

5. Bab 5 KESIMPULAN DAN SARAN

Pada bab terakhir, dijabarkan kesimpulan dari penelitian yang telah dilakukan. Dari kesimpulan yang telah dibuat, diberikan juga saran kepada pengembang penelitian selanjutnya yang akan meneliti topik terkait lebih lanjut.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A