

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berdasarkan hasil asesmen PISA (*Programme for International Student Assesment*) pada tahun 2022, Indonesia menduduki peringkat ke-71 dari 81 negara yang berpartisipasi [1]. Meskipun peringkat Indonesia mengalami kenaikan, nilai skor rata-rata yang diperoleh justru menurun dibandingkan tahun-tahun sebelumnya. Pada tahun 2018, skor rata-rata bidang literasi mencapai 371, sedangkan pada tahun 2022 turun menjadi 359 [2]. Hal ini menunjukkan bahwa kemampuan literasi siswa di Indonesia masih tergolong rendah.

Berbagai upaya telah dilakukan dan terus ditingkatkan untuk mengatasi permasalahan tersebut dengan melibatkan berbagai pihak. Salah satu upaya untuk meningkatkan literasi di Indonesia adalah mengadakan taman baca yang menyediakan berbagai bacaan. Bagi sebagian orang, untuk memasuki dunia literasi, dapat diawali dengan membaca sebuah novel. Karakteristik yang dimiliki oleh novel, seperti memiliki cerita yang menarik dan mudah dimengerti, serta genre yang bermacam-macam, mampu meningkatkan minat membaca. Sehingga, novel menjadi salah satu jenis buku yang banyak digemari oleh hampir seluruh kalangan usia. Ketersediaan berbagai jenis novel dibanyaknya toko buku dan perpustakaan, membuat novel mudah untuk didapatkan.

Di Indonesia, banyak penulis yang memiliki novel-novel menarik, hingga karya nya disebut sebagai novel *best seller*. Untuk menilai sebuah novel, tidak bisa dilandaskan hanya dari kepopuleran dan judul buku saja, karena judul buku tidak bisa menggambarkan keseluruhan cerita dari novel tersebut. Selain itu, sinopsis novel tidak sepenuhnya mencerminkan isi cerita. Maka untuk menentukan sebuah novel yang akan dibaca atau dibeli, biasanya seseorang akan membaca ulasan dari orang-orang yang telah membaca novel tersebut terlebih dahulu. Melalui ulasan atau *review* tersebut, dapat menjadi penentu apakah buku tersebut bagus dan layak dibaca atau tidak. Hal ini didukung dengan hasil penelitian yang menunjukkan

bahwa sebanyak 82% konsumen akan membaca ulasan suatu produk sebelum membuat keputusan pembelian [3].

Perkembangan teknologi informasi telah mendorong kemunculan berbagai *platform* digital yang menyediakan akses terhadap informasi serta ulasan buku. Tidak hanya *review* atau ulasannya saja, bahkan berbagai macam buku bacaan termasuk novel pun kini sudah tersedia secara *online*. Goodreads menjadi salah satu situs yang menyediakan informasi terkait buku secara *online*, yang di dalamnya membahas mengenai informasi buku, informasi penulis, serta berbagi rekomendasi buku bacaan kepada sesama anggota [4].

Ulasan yang ditulis oleh banyaknya pengguna terkadang memberikan kesulitan untuk menemukan informasi yang sesuai dengan keinginan pengguna. Hal ini disebabkan karena banyaknya sudut pandang dan cara penulisan yang berbeda, sehingga pengguna harus memilah satu per satu ulasan yang ada. Selain itu, tidak jarang pula ditemukan sebuah ulasan yang ditulis dengan begitu panjang dan bertele-tele, sehingga pengguna kesulitan dalam menyimpulkannya poin utama dari ulasan novel tersebut secara cepat dan tepat. Untuk memecahkan permasalahan ini, maka dibutuhkan penerapan dua teknik yang dapat saling melengkapi, yaitu analisis sentimen dan *text summarization*.

Analisis sentimen merupakan salah satu teknik *text mining*, dimana bertujuan untuk melakukan proses ekstraksi, mengolah, serta memahami data secara otomatis untuk memperoleh informasi sentimen dalam data dokumen teks yang berupa pendapat atau opini yang bersifat positif, negatif, atau netral [5]. Sedangkan *text summarization* merupakan sebuah proses untuk meringkas teks secara otomatis dengan tetap mempertahankan poin penting dari teks aslinya, sehingga akan menciptakan sebuah rangkuman teks yang lebih singkat. Metode yang dapat digunakan dalam *text summarization* terbagi menjadi dua, yaitu *extractive* yang berarti tetap mempertahankan teks aslinya karena akan memilih kalimat yang penting untuk dijadikan sebagai ringkasan dan metode *abstractive* yang berarti menciptakan ringkasan dengan kata dan struktur baru dan berbeda dari teks asli [6].

Model *Bidirectional Encoder Representations from Transformers* (BERT) dapat diimplementasikan untuk tugas analisis sentimen, karena merupakan sebuah model yang dapat menyelesaikan berbagai tugas *Natural Language Processing* (NLP). BERT tersedia di dalam berbagai bahasa, termasuk bahasa Indonesia yang dikenal sebagai IndoBERT. Model IndoBERT telah dilatih menggunakan 220 juta kata dalam Bahasa Indonesia [7]. Sehingga IndoBERT memiliki pemahaman lebih dalam mengenai Bahasa Indonesia dan efektif untuk tugas NLP dengan dataset berbahasa Indonesia.

Keunggulan IndoBERT juga telah dibuktikan melalui beberapa penelitian terdahulu. Penelitian pertama mengenai ulasan aplikasi BRImo menunjukkan bahwa IndoBERT mampu menganalisis sentimen dengan cukup akurat dan memperoleh akurasi sebesar 90% [8]. Penelitian kedua pada ulasan pengunjung objek wisata Baturraden menunjukkan bahwa IndoBERT mampu menganalisis sentimen dengan efektif, dan memperoleh akurasi sebesar 94.61% [9]. Penelitian ketiga membandingkan performa IndoBERT dengan SVM dan *Naïve Bayes*, dimana performa IndoBERT lebih unggul dibandingkan metode SVM dan *Naïve Bayes*, dengan memperoleh akurasi sebesar 83.5% [10].

Kemudian, salah satu algoritma yang banyak digunakan untuk *text summarization* adalah TextRank. TextRank merupakan algoritma berbasis graf yang memberi peringkat pada kata atau kalimat dengan melihat hubungan antar unit teks, sehingga menghasilkan ringkasan teks bersifat *extractive* [11]. Penelitian sebelumnya menunjukkan bahwa TextRank dapat meringkas teks berita berbahasa Indonesia dengan memperoleh nilai *precision*, *recall*, dan *f-measure* sebesar 0.552 pada *compression rate* 30% [12]. Penelitian lain pada teks putusan pengadilan berbahasa Indonesia, menunjukkan bahwa TextRank mampu memperoleh nilai *precision* sebesar 0.62 pada *compression rate* 75%, serta hasil evaluasi pada ringkasan yang dihasilkan oleh algoritma TextRank lebih seimbang dibandingkan dengan algoritma LSA [13]. Oleh karena itu, hasil temuan-temuan tersebut mendukung pemilihan IndoBERT sebagai metode untuk analisis sentimen dan TextRank sebagai algoritma untuk meringkas teks ulasan.

Penelitian ini menawarkan sebuah solusi untuk permasalahan di atas dengan mengembangkan dua pendekatan secara terpisah, yaitu analisis sentimen berbasis IndoBERT dan model *text summarization* berbasis TextRank. Kedua pendekatan ini digunakan untuk mendukung pembaca dalam memahami opini, isi ulasan, serta respon terhadap buku Laut Bercerita karya Leila S. Chudori. Melalui kombinasi tersebut, penelitian ini bertujuan untuk meringkas ulasan dan mengidentifikasi sentimen pembaca, sehingga dapat memudahkan dalam memahami isi dan persepsi yang berkembang terhadap sebuah novel secara efisien. Selain itu, penelitian ini berusaha untuk mengevaluasi kemampuan kedua pendekatan tersebut dalam menyajikan isi dan opini ulasan novel secara relevan.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah pada penelitian ini adalah sebagai berikut.

1. Bagaimana menerapkan model analisis sentimen dengan IndoBERT dan model *text summarization* dengan algoritma TextRank pada ulasan novel berbahasa Indonesia?
2. Bagaimana performa model algoritma TextRank dalam meringkas teks ulasan novel dan model IndoBERT dalam menganalisis sentimen terhadap data ulasan novel asli maupun data ulasan novel hasil ringkasan?

1.3 Batasan Masalah

Agar penelitian dapat mengarah kepada permasalahan, maka dibutuhkan pembatasan masalah sehingga hasil dari penelitian lebih fokus dan terarah.

1. Data ulasan yang akan digunakan diambil melalui situs goodreads.com.
2. Ulasan novel yang digunakan yaitu ulasan pada novel karya Leila S. Chudori yang berjudul Laut Bercerita.

1.4 Tujuan

Berdasarkan permasalahan yang ada, maka penelitian ini memiliki tujuan sebagai berikut.

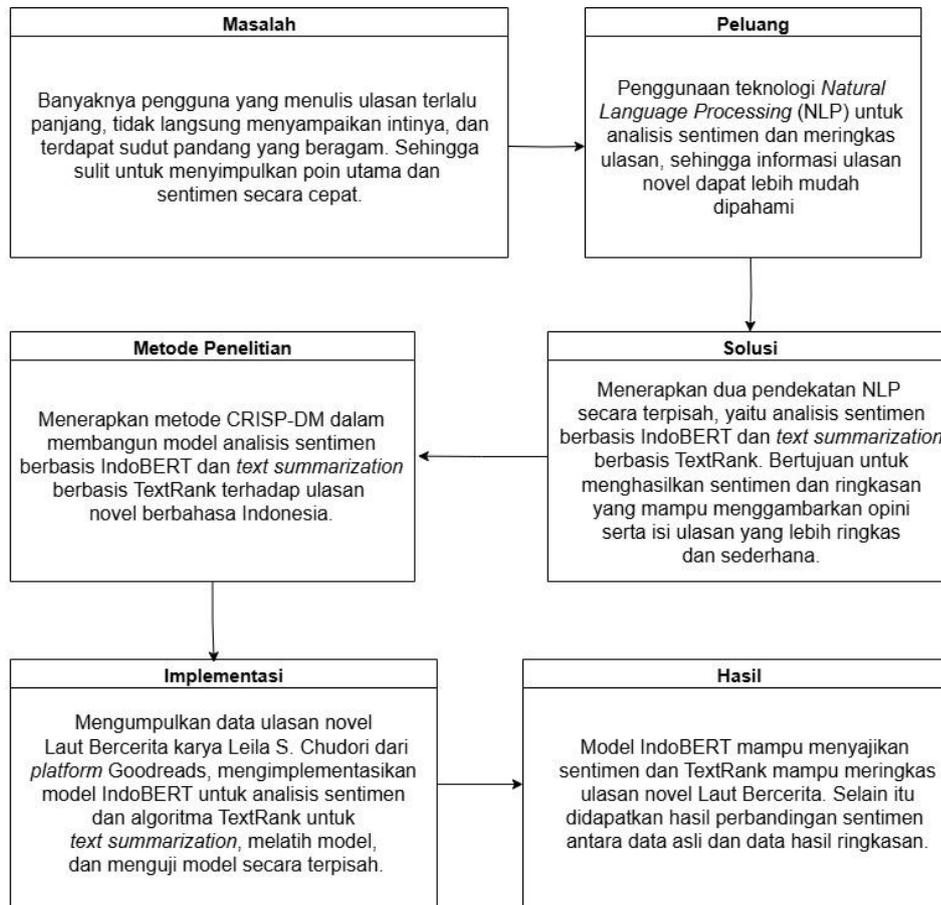
1. Menerapkan model IndoBERT dan algoritma TextRank untuk melakukan analisis sentimen dan *text summarization* pada ulasan novel Laut Bercerita karya Leila S. Chudori.
2. Mengevaluasi performa model IndoBERT dan algoritma TextRank dalam menghasilkan analisis sentimen dan ringkasan yang merepresentasikan isi dan opini, serta mengidentifikasi perubahan sentimen pada ulasan novel Laut Bercerita karya Leila S. Chudori.

1.5 Manfaat

1. Hasil penelitian ini diharapkan dapat membantu calon pembaca dalam memahami pandangan umum melalui isi dan opini ulasan pada ringkasan dan analisis sentimen terhadap novel Laut Bercerita karya Leila S. Chudori.
2. Hasil penelitian ini diharapkan dapat menjadi gambaran dan informasi tambahan mengenai dua pendekatan NLP, yaitu analisis sentimen berbasis IndoBERT dan *text summarization* berbasis TextRank untuk mengolah ulasan novel berbahasa Indonesia. Sehingga dapat dikembangkan oleh penelitian selanjutnya dalam lingkup yang sama di masa yang akan datang.
3. Hasil penelitian ini diharapkan mampu memperluas pengetahuan, pengalaman, serta wawasan mengenai penerapan teknologi *Natural Language Processing* (NLP) dalam konteks sastra, yang mengimplementasikan dua pendekatan berbeda dan terpisah untuk mengolah teks ulasan berbahasa Indonesia.

1.6 Kerangka Pemikiran

Berdasarkan permasalahan tersebut, maka kerangka pemikiran dari penelitian ini ditampilkan pada Gambar 1.1.



Gambar 1. 1 Kerangka Pemikiran

1.7 Sistematika Penulisan

Sistematika penulisan akan memaparkan struktur atau urutan penulisan yang menggambarkan alur penelitian serta hubungan pada setiap bab atau pembahasan yang saling berkaitan dari awal hingga akhir di dalam laporan hasil penelitian. Sistematika penulisan pada penelitian ini adalah sebagai berikut.

BAB I PENDAHULUAN

Pada bab ini menyajikan pembahasan terkait latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, kerangka pemikiran, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini, menyajikan pembahasan mengenai hasil penelitian terdahulu dan landasan teori yang relevan dan mendukung penelitian.

BAB III METODOLOGI PENELITIAN

Pada bab ini menyajikan pembahasan terkait metode penelitian, tahapan, dan teknik-teknik yang digunakan di dalam tugas akhir ini.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini menyajikan pembahasan terkait hasil penelitian dan analisis terhadap temuan-temuan yang diperoleh.

BAB V SIMPULAN DAN SARAN

Pada bab ini menyajikan pembahasan terkait kesimpulan dari hasil penelitian dengan saran-saran yang dapat diberikan untuk penelitian selanjutnya.

