

BAB 1 PENDAHULUAN

1.1. Latar Belakang

Seiring dengan berkembangnya industri permainan digital, platform distribusi permainan digital seperti Steam telah menjadi tempat utama bagi para pengguna untuk membagikan pengalaman dan ulasan mereka [1]. Salah satu genre yang memiliki tingkat popularitas yang terus meningkat di Steam adalah Role-Playing Game (RPG) [2]. Ulasan pengguna tidak hanya mencerminkan kepuasan pemain tetapi juga menjadi indikator kualitas permainan digital dan memberikan informasi berharga bagi pengembang dan calon pembeli [3]. Namun, mengingat jumlah ulasan yang sangat banyak dan beragam, menjadi tantangan besar bagi pengembang permainan digital, pemain lain, maupun platform itu sendiri untuk menganalisis dan mengambil manfaat dari data tersebut secara efektif.

Salah satu metode yang kini berkembang untuk mengolah informasi dalam jumlah besar ini adalah analisis sentimen [4]. Analisis sentimen merupakan suatu teknik yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini atau perasaan yang terkandung dalam suatu teks, apakah positif, negatif, atau netral [5]. Dalam konteks ulasan permainan digital di Steam, analisis sentimen berfungsi untuk mengidentifikasi reaksi emosional pemain terhadap suatu permainan, baik yang bersifat positif maupun negatif.

Dalam beberapa tahun terakhir, kemajuan dalam hal kecerdasan buatan khususnya dalam bidang pemrosesan bahasa alami atau NLP (Natural Language Processing), telah membuka potensi besar bagi model-model berbasis Large Language Models (LLM) untuk melakukan analisis teks secara lebih mendalam dan akurat. Beberapa contoh model LLM yang populer yaitu BERT (Bidirectional Encoder Representations from Transformers), RoBERTa

(Robustly optimized BERT approach), dan DistilBERT (distilled version of BERT) [6][7][8]. Model berbasis LLM tersebut merupakan varian dari arsitektur transformer yang telah terbukti unggul dalam pemahaman konteks dan hubungan antar kata dalam sebuah kalimat, sehingga memungkinkan analisis sentimen yang lebih canggih dan berbasis konteks [9][10].

Beberapa penelitian terdahulu memberikan dasar penting bagi penelitian ini. Samir et al. (2021) menunjukkan bahwa BERT mampu memberikan tingkat akurasi sebesar 86.5% dalam tugas klasifikasi sentimen dalam *dataset* twitter [11], sementara peneliti lain oleh Chatzimina et al. (2023) menunjukkan bahwa RoBERTa mampu mengungguli GPT-2 dan XLNet dalam tugas klasifikasi sentimen pada percakapan klinis berbahasa Yunani, dengan tingkat akurasi mencapai 91,43%. Model ini terbukti efektif khususnya dalam mengenali nuansa sentimen negatif yang bersifat halus dan kontekstual [12]. Penelitian lain oleh Yejian et al. (2023) mengonfirmasi efektivitas BERT untuk proses klasifikasi data multi-label pada ulasan game di platform Steam, dengan tingkat akurasi sebesar 88,88% [13]. Selain itu, Bayat et al. (2023) mengevaluasi performa DistilBERT dan berbagai model pembelajaran mesin lainnya terhadap data ulasan produk di Amazon. Hasilnya menunjukkan bahwa DistilBERT mencapai akurasi tertinggi sebesar 96,10%, mengungguli model seperti Multi-Layer Perceptron (MLP), Naive Bayes, dan Decision Tree. Temuan ini menyoroti kemampuan DistilBERT dalam menangkap konteks sentimen secara efisien pada data ulasan dunia nyata [14].

Hingga saat ini, belum banyak penelitian yang secara khusus dan komprehensif membandingkan performa berbagai model berbasis *large language model* (LLM) untuk analisis sentimen dalam konteks ulasan permainan digital, khususnya pada genre role-playing game (RPG). Ulasan pengguna game di platform seperti Steam memiliki karakteristik linguistik yang sangat khas dengan gaya bahasa yang tidak baku, ekspresi subjektif yang kuat, serta penggunaan istilah domain-spesifik yang menghadirkan tantangan

tersendiri bagi model analisis sentimen. Hal ini menciptakan kebutuhan untuk mengevaluasi seberapa baik model LLM mampu memahami konteks semantik yang kompleks dan ambigu dalam teks-teks ulasan tersebut.

Untuk menjawab kebutuhan tersebut, penelitian ini dirancang guna membandingkan performa tiga model LLM populer, yaitu BERT, RoBERTa, dan DistilBERT, dalam tugas klasifikasi sentimen terhadap ulasan game RPG di Steam. Selain perbandingan performa berdasarkan metrik evaluasi, penelitian ini juga mengusulkan dan mengkaji RPGSentiLex, yaitu leksikon domain RPG yang digunakan untuk menganalisis distribusi sentimen istilah-istilah khas yang sering muncul dalam ulasan. Dengan demikian, penelitian ini memberikan kontribusi penting dalam pengembangan metode analisis sentimen yang lebih akurat, kontekstual, dan *domain-aware*, khususnya untuk platform ulasan digital dengan variasi data yang tinggi.

1.2. Rumusan Masalah

Atas dasar latar belakang yang telah disampaikan maka rumusan masalah dalam penelitian ini dapat dijabarkan sebagai berikut:

- Bagaimana membangun sistem klasifikasi sentimen yang efektif untuk ulasan game RPG pada platform Steam menggunakan model bahasa besar seperti BERT, RoBERTa, dan DistilBERT?
- Bagaimana pengaruh teknik praproses teks, seperti lemmatisasi dan penghapusan *stopword*, terhadap performa model-model tersebut dalam tugas klasifikasi sentimen?
- Bagaimana distribusi sentimen dari istilah-istilah domain-spesifik dalam ulasan RPG, dan sejauh mana konteks mempengaruhi interpretasi sentimen terhadap istilah-istilah tersebut?

1.3. Tujuan dan Manfaat

Tujuan dari penelitian ini adalah merancang dan mengevaluasi sistem klasifikasi sentimen terhadap ulasan pengguna game genre RPG di platform

Steam dengan memanfaatkan model LLM (*large language models*) berbasis arsitektur transformer. Sistem ini diharapkan mampu mengklasifikasikan opini pengguna ke dalam kategori sentimen positif atau negatif secara akurat dan efisien.

Secara khusus, penelitian ini bertujuan untuk membandingkan performa dari tiga model LLM, yaitu BERT, RoBERTa, dan DistilBERT, dalam tugas klasifikasi sentimen. Evaluasi dilakukan dengan mengukur performa masing-masing model menggunakan metrik akurasi, presisi, recall, dan *F1-score*, sehingga kondisi ideal yang ingin dicapai dapat dikuantifikasi secara jelas.

Selain itu, penelitian ini bertujuan untuk menganalisis pengaruh teknik praproses teks, seperti lemmatisasi dan penghapusan *stopword*, terhadap performa model, guna menilai apakah langkah-langkah tersebut meningkatkan atau justru menurunkan akurasi klasifikasi. Penelitian ini juga bertujuan untuk menyusun dan menganalisis leksikon domain-spesifik bernama RPGSentiLex untuk mengevaluasi distribusi sentimen terhadap istilah khas dalam ulasan game RPG, guna mengetahui sejauh mana konteks memengaruhi polaritas sentimen.

Melalui pencapaian tujuan-tujuan tersebut, sistem yang dikembangkan diharapkan dapat memberikan kontribusi nyata dalam peningkatan efisiensi proses analisis opini pengguna dalam domain game digital serta mendukung pengambilan keputusan yang lebih tepat oleh pengembang maupun konsumen.

Tabel 1.1 Keterkaitan antara Tujuan, Pengujian, dan Kesimpulan

No.	Tujuan	Pengujian	Kesimpulan
1	Mengevaluasi dan membandingkan performa model BERT, RoBERTa, dan DistilBERT dalam klasifikasi sentimen	Pengujian pada masing-masing model menggunakan metrik akurasi, presisi, recall, dan <i>F1-score</i>	RoBERTa menunjukkan performa terbaik dengan <i>F1-score</i> 97,50%, diikuti BERT (95,50%) dan DistilBERT (95,30%)

2	Mengetahui pengaruh teknik praproses (lemmatisasi dan penghapusan <i>stopword</i>) terhadap performa model	Pengujian dengan dan tanpa praproses pada tiap model	Praproses berlebihan justru menurunkan performa; model bekerja lebih baik dengan teks alami
3	Menganalisis distribusi sentimen istilah domain RPG untuk mengkaji konteksualitas makna	Pembuatan dan analisis RPGSentiLex terhadap lima istilah umum dalam ulasan	Mayoritas istilah yang berada di domain RPGSentiLex, seperti "boss", "map", dan "character" memiliki distribusi polaritas seimbang, tergantung konteks

1.4. Batasan Masalah

Penelitian ini memiliki beberapa batasan yang ditetapkan untuk menyederhanakan ruang lingkup pekerjaan agar dapat diselesaikan dalam waktu dan sumber daya yang tersedia selama satu semester. Batasan-batasan ini diperlukan untuk menjaga fokus penelitian serta memastikan kelayakan teknis dan operasional pengerjaan Tugas Akhir.

Pertama, data ulasan yang digunakan dibatasi hanya pada lima judul game RPG populer di platform Steam, yaitu Dark Souls 3, Dragon's Dogma 2, Elden Ring, The Elder Scrolls, dan Wukong. Pembatasan ini dilakukan karena proses pengumpulan dan anotasi manual ulasan membutuhkan waktu dan tenaga yang signifikan. Selain itu, keterbatasan perangkat komputasi yang tersedia juga menjadi pertimbangan dalam membatasi volume data, yaitu sebanyak 10.000 ulasan saja, meskipun secara umum penelitian analisis sentimen skala besar dapat menggunakan data dalam jumlah jauh lebih besar.

Kedua, proses anotasi data dilakukan secara manual oleh tiga annotator, tanpa memanfaatkan platform anotasi otomatis. Hal ini disebabkan belum tersedianya sistem anotasi otomatis yang memiliki akurasi

tinggi untuk domain game RPG, serta untuk memastikan validitas pelabelan dalam dataset. Pendekatan ini juga digunakan untuk menjaga konsistensi kualitas label sentimen, meskipun metode ini tidak skalabel untuk jumlah data yang lebih besar.

Ketiga, model yang dievaluasi terbatas pada tiga arsitektur LLM, yaitu BERT, RoBERTa, dan DistilBERT. Model lain yang lebih besar dan kompleks tidak disertakan karena keterbatasan perangkat keras (GPU) dan waktu pelatihan yang diperlukan. Oleh karena itu, fokus penelitian dibatasi pada model-model transformer yang umum digunakan dalam penelitian akademik dan memiliki dokumentasi pelatihan yang lebih ringkas.

Keempat, ruang lingkup eksperimen praproses dibatasi hanya pada teknik penghapusan *stopword* dan lemmatisasi, tanpa melibatkan metode praproses lanjutan seperti stemming berbasis statistik maupun augmentasi dengan substitusi sinonim. Pembatasan ini ditetapkan karena keterbatasan waktu pelaksanaan dan pertimbangan untuk tetap menjaga kejelasan interpretasi hasil terhadap respons model.

Kelima, analisis leksikon domain-spesifik (RPGSentiLex) dilakukan secara statistik tanpa integrasi ke dalam sistem klasifikasi sebagai fitur tambahan. Hal ini karena fokus penelitian ini adalah mengevaluasi performa model berbasis konteks penuh tanpa modifikasi arsitektur internal.

1.5. Metode Penelitian

Penelitian ini dilakukan dengan pendekatan kombinatif antara studi literatur, pengumpulan data, eksperimen, dan evaluasi. Pendekatan tersebut diterapkan secara bertahap untuk membangun sistem klasifikasi sentimen pada ulasan game bergenre RPG di platform Steam.

Tahapan pertama dimulai dengan studi literatur untuk mengkaji berbagai penelitian terdahulu yang menggunakan model bahasa besar berbasis arsitektur transformer seperti BERT, RoBERTa, dan DistilBERT dalam tugas klasifikasi sentimen. Studi ini mencakup pemahaman terhadap karakteristik

masing-masing model, strategi praproses, serta metrik evaluasi yang digunakan dalam domain serupa.

Selanjutnya dilakukan tahap pengumpulan data dengan metode web crawling terhadap ulasan pengguna dari lima judul game RPG yang populer. Data yang diperoleh kemudian dianotasi secara manual oleh tiga anotator independen untuk menentukan label sentimen positif atau negatif. Hasil label diperkuat melalui validasi dengan model pralatih sebagai acuan.

Setelah data terkumpul, dilakukan tahap praproses teks untuk membersihkan dan menstandarkan format input. Praproses yang digunakan meliputi penghapusan tanda baca, case folding, penghapusan *stopword*, dan lemmatisasi. Untuk menganalisis pengaruh praproses terhadap performa model, dilakukan eksperimen pembandingan antara data yang dipraproses dan yang tidak dipraproses.

Proses inti berupa pelatihan dan fine-tuning model klasifikasi dilakukan dengan mengimplementasikan ketiga arsitektur model LLM. Model dilatih pada dataset yang telah dibagi ke dalam set pelatihan, validasi, dan pengujian. Proses pelatihan disertai dengan tuning *hyperparameter*, khususnya learning rate, untuk menentukan konfigurasi optimal pada masing-masing model.

Evaluasi dilakukan dengan pendekatan analisis statistik menggunakan metrik kuantitatif: akurasi, presisi, recall, dan *F1-score*. Untuk mendukung interpretasi hasil, dibuat pula analisis leksikon domain-spesifik bernama RPGSentiLex yang digunakan untuk mengkaji distribusi sentimen terhadap istilah umum dalam game RPG, guna menunjukkan pentingnya pemodelan konteks.

Rangkaian metode ini secara keseluruhan dirancang untuk menjawab rumusan masalah dan mendukung pencapaian tujuan penelitian secara terukur dan sistematis.

1.6. Jadwal Pelaksanaan

Jadwal pelaksanaan Tugas Akhir dirancang untuk diselesaikan dalam jangka waktu enam bulan, dengan membagi pekerjaan ke dalam beberapa tahapan utama.

Berdasarkan Tabel 1.2, kegiatan dimulai dengan studi literatur yang dilaksanakan pada bulan pertama dan kedua. Tahapan ini menjadi fondasi penting dalam memahami teori, metodologi, serta studi terdahulu yang relevan dengan topik penelitian. Selanjutnya, pengumpulan data dilakukan pada bulan kedua, berfokus pada pengambilan dan kurasi data ulasan game RPG dari platform Steam yang akan digunakan sebagai dataset utama.

Setelah data terkumpul, tahap eksperimen dimulai pada bulan ketiga dan berlanjut hingga bulan keempat. Pada fase ini, proses *fine-tuning* dan evaluasi awal terhadap tiga model LLM (BERT, RoBERTa, dan DistilBERT) dilakukan secara sistematis berdasarkan skenario percobaan yang telah ditentukan. Setelah eksperimen selesai, proses evaluasi dan analisis hasil dilakukan secara mendalam pada bulan kelima, untuk menilai performa masing-masing model serta mengkaji kontribusi leksikon RPGSentiLex.

Terakhir, penyusunan laporan dan buku tugas akhir dilaksanakan secara intensif mulai bulan kelima hingga bulan keenam. Dokumen ini mencakup seluruh hasil eksperimen, analisis, dan pembahasan yang telah dilakukan selama proses penelitian. Jadwal ini juga mencakup milestone penting yang digunakan sebagai acuan dalam mengevaluasi kemajuan dan ketercapaian target di setiap tahapan pelaksanaan Tugas Akhir.

Tabel 1.2. Jadwal Pelaksanaan Tugas Akhir.

No.	Deskripsi Tahapan	Bulan 1	Bulan 2	Bulan 3	Bulan 4	Bulan 5	Bulan 6
1	Studi Literatur						
2	Pengumpulan Data						
3	Eksperimen						
4	Evaluasi						

5	Penyusunan Laporan/Buku TA						
---	-------------------------------	--	--	--	--	--	--