

BAB I

PENDAHULUAN

Bab pendahuluan ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan yang digunakan dalam dokumen skripsi ini.

1.1 Latar Belakang

Game Online adalah hiburan yang disenangi oleh banyak masyarakat di dunia. *Game Online* dapat dimainkan melalui berbagai *platform* seperti komputer, perangkat mobile, ataupun konsol *game*. Beragam jenis *game* seperti strategi, *role playing game* (RPG), aksi, dan lainnya yang mendominasi pasar *game online*. Salah satu *game* populer sampai saat ini adalah Genshin Impact, sebuah *game online* yang memadukan petualangan dan eksplorasi dengan grafis berkualitas tinggi dan mekanisme permainan yang kompleks. Sejak dirilis pada tahun 2020 oleh miHoYo, Genshin Impact masih menjadi *game* dengan pemain aktif sebanyak 60 juta dan menjadi salah satu *game* RPG dengan pendapatan tertinggi. Kesuksesan *game* Genshin Impact tidak terlepas dari royalti dan loyalitas para pemainnya yang kerap memberikan ulasan mengenai *game* Genshin Impact. Pemain secara aktif memberikan opini terhadap pembaruan konten yang dirilis secara berkala oleh pengembang. Pemain tidak hanya memberikan apresiasi tetapi juga menyampaikan kritik terhadap berbagai aspek permainan, meliputi mekanisme *gameplay*, pengembangan alur cerita, serta implementasi sistem gacha yang diterapkan dalam permainan (Safrudin dkk., 2024). Ulasan-ulasan yang diberikan para pemain memberikan wawasan penting bagi pengembang mengenai kepuasan, harapan, serta kritik dari para pemain.

Opini para pemain dapat dipahami melalui teknik khusus yang disebut analisis sentimen. Analisis sentimen adalah studi yang menganalisis opini, sentimen, evaluasi, dan emosi seseorang kepada entitas seperti produk, organisasi, dan individual yang diekspresikan dalam bentuk teks (Liu, 2012). Analisis sentimen mengklasifikasikan pendapat pemain dalam golongan positif, negatif, ataupun netral (Wang dkk., 2018). Oleh karena itu, analisis sentimen penting untuk dilakukan sebagai tolok ukur bagi pengembang Genshin Impact agar pengembang dapat mengetahui aspek positif yang layak dipertahankan dan aspek negatif yang perlu diperbaiki.

Penelitian terdahulu telah melakukan penelitian terhadap ulasan Genshin Impact, seperti pada penelitian yang dilakukan oleh Safrudin, dkk. (2024) menggunakan metode *Naïve Bayes* dan SVM. Akurasi yang didapatkan pada penelitian yang dilakukan oleh Safrudin, dkk. (2024) menggunakan metode *Naïve Bayes* pada penelitian analisis sentimen ulasan Genshin Impact sebesar 75,5%. Sedangkan akurasi yang didapatkan menggunakan metode SVM sebesar 83%. Hal tersebut membuktikan bahwa SVM lebih unggul dibandingkan *Naïve Bayes* untuk melakukan penelitian analisis sentimen. Penelitian lainnya, yaitu oleh Kusnadi dkk. (2021) menggunakan metode *fine-tuning BERT* pada ulasan Genshin Impact. Metode *BERT* dipilih oleh peneliti tersebut dikarenakan kemampuannya dalam memahami konteks kalimat secara mendalam melalui representasi *bidirectional*, yang menunjukkan model ini menangkap informasi semantik dari konteks keseluruhan teks. Hasil penelitian yang dilakukan oleh Kusnadi dkk. (2021) dalam analisis sentimen ulasan Genshin Impact menggunakan metode *BERT* mendapatkan akurasi sebesar 74%. Metode *BERT* memiliki kekurangan dikarenakan metode tersebut membutuhkan sumber daya komputasi yang tinggi, baik dalam hal waktu pelatihan maupun kebutuhan memori. Tantangan umum yang belum dibahas dalam penelitian-penelitian tersebut adalah masalah ketidakseimbangan data yang sering terjadi dalam *dataset* ulasan, yang ditunjukkan oleh distribusi sentimen positif dan negatif yang tidak merata. Ketidakseimbangan ini menyebabkan metode pembelajaran cenderung bias terhadap kelas mayoritas, sehingga kinerja model pada kelas minoritas seringkali menurun secara signifikan (Krawczyk, 2016).

Keterbatasan dari penelitian sebelumnya dan permasalahan ketidakseimbangan kelas mendorong peneliti untuk mengusulkan penggunaan metode *Support Vector Machine* yang dikombinasikan dengan teknik *random oversampling*. *Support Vector Machine* memiliki kelebihan dibandingkan dengan kedua metode sebelumnya karena fleksibilitasnya dalam penggunaan berbagai kernel untuk menangani pola-pola data non-linier yang sering terjadi dalam analisis sentimen (Joachims, 1998). Selain itu, *Support Vector Machine* terbukti optimal dalam menangani *dataset* yang telah diseimbangkan karena kemampuannya dalam menemukan *hyperplane* optimal yang memaksimalkan margin antara kelas-kelas yang ada (Vapnik, 2013). Penelitian yang dilakukan Pambudi dkk. (2024) menunjukkan bahwa kombinasi metode *Support Vector Machine* dengan teknik *random oversampling* dapat meningkatkan akurasi klasifikasi dari 94,14% menjadi 97,92%.

Metode *Support Vector Machine* digunakan dengan pelabelan menggunakan kamus *Indonesian Sentiment lexicon*. Metode *Support Vector Machine* merupakan metode pembelajaran mesin yang dapat digunakan untuk klasifikasi dan regresi (Agustina & Putri, 2024). Metode *Support Vector Machine* menjadi salah satu metode terbaik karena kapabilitasnya dalam pengenalan pola yang dicapai dengan melakukan transformasi dalam ruang input ke ruang dimensi yang lebih tinggi (ruang fitur) dan melakukan pengoptimalan dalam ruang vektor yang baru (Sukma Utama dkk., 2017). Sedangkan kamus *Indonesian sentiment lexicon* adalah metode umum untuk menentukan apakah sebuah kalimat, teks, atau komentar memiliki nilai sentimen positif, negatif, atau netral (Nahar dkk., 2020). Nilai tersebut akan digunakan sebagai data input pada metode *Support Vector Machine*.

Penelitian ini bertujuan untuk menganalisis dan membandingkan performa berbagai kernel yang ada pada metode *Support Vector Machine* (Linear, RBF, dan *Polynomial*) dalam mengklasifikasikan sentimen ulasan aplikasi Genshin Impact dengan menerapkan teknik *oversampling* untuk mengatasi masalah ketidakseimbangan data. Evaluasi performa akan dilakukan menggunakan berbagai metrik seperti akurasi, *precision*, *recall*, dan *f1-score* untuk memberikan pemahaman yang komprehensif tentang kinerja masing-masing kernel dalam menangani karakteristik data yang berbeda. Hasil dari penelitian ini diharapkan dapat memberikan wawasan tentang efektivitas kombinasi metode *Support Vector Machine* dan teknik *oversampling* dalam analisis sentimen, serta memberikan rekomendasi kernel yang optimal untuk kasus serupa.

1.2 Rumusan Masalah

Rumusan masalah berdasarkan identifikasi masalah penelitian ini adalah sebagai berikut:

1. Bagaimana performa masing-masing kernel metode SVM dalam mengklasifikasikan sentimen positif dan negatif pada ulasan Genshin Impact terhadap data tidak seimbang dan seimbang?
2. Bagaimana komparasi performa metode SVM pada data tidak seimbang dan data seimbang?

1.3 Tujuan dan Manfaat Penelitian

Pada penelitian ini yang menjadi tujuan utama adalah sebagai berikut:

1. Membangun model analisis sentimen menggunakan metode SVM pada data ulasan Genshin Impact yang diterapkan pada data tidak seimbang maupun seimbang.
2. Mencari kernel terbaik untuk mendapatkan model analisis sentimen menggunakan *hyperparameter tuning*.

Adapun manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Memberikan kontribusi pengetahuan dalam penerapan metode SVM dengan berbagai kernel untuk klasifikasi sentimen, terkhusus dalam menangani ketidakseimbangan data pada ulasan *game*.
2. Memberikan referensi bagi peneliti lain mengenai perbandingan performa berbagai kernel SVM dalam kondisi data tidak seimbang dan data seimbang,

1.4 Ruang Lingkup

Untuk memastikan skripsi ini lebih terarah dan mencapai tujuan yang diinginkan, diperlukan penetapan ruang lingkup. Ruang lingkup dalam skripsi ini meliputi hal-hal berikut:

1. Data yang digunakan sebagai dasar dalam penelitian ini adalah data ulasan pada aplikasi Google Play Store dengan parameter utama yang diambil yaitu parameter ulasan, skor, dan tanggal yang diambil menggunakan metode *scraping* pada rentang Oktober 2023 – Oktober 2024.
2. Analisis sentimen pada data dilakukan dengan klasifikasi sentimen secara biner ke dalam 2 kelas yaitu positif dan negatif. Data dilabelisasi menggunakan kamus *Indonesian sentiment lexicon*.
3. Pengolahan data analisis sentimen menggunakan metode *Support Vector Machine*, bahasa pemrograman Python, dan perangkat lunak Google Colab.

1.5 Sistematika Penulisan

Sistematika penulisan bertujuan untuk memberikan gambaran yang terstruktur dan jelas mengenai isi laporan skripsi ini. Berikut adalah beberapa topik utama yang dibahas dalam skripsi ini:

BAB I PENDAHULUAN

Bab ini berisi mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat penelitian, ruang lingkup, serta sistematika penulisan dalam skripsi Perbandingan Performa Penanganan Data *Imbalance* dan Kernel pada Analisis Sentimen Ulasan *Game Genshin Impact* Berbasis *Support Vector Machine*.

BAB II LANDASAN TEORI

Bab ini berisi studi pustaka mengenai penelitian terdahulu (*state-of-the-art*) serta dasar teori yang mencakup *Genshin Impact*, analisis sentimen, *data preprocessing*, kamus *Indonesian sentiment lexicon*, *oversampling*, *Term Frequency-Inverse Document Frequency*, *grid search cross validation*, *Support Vector Machine*, dan evaluasi model dalam skripsi Perbandingan Performa Penanganan Data *Imbalance* dan Kernel pada Analisis Sentimen Ulasan *Game Genshin Impact* Berbasis *Support Vector Machine*.

BAB III METODOLOGI PENELITIAN

Bab ini berisi mengenai metodologi yang digunakan dalam skripsi Perbandingan Performa Penanganan Data *Imbalance* dan Kernel pada Analisis Sentimen Ulasan *Game Genshin Impact* Berbasis *Support Vector Machine*. Bab ini membahas tentang alur penelitian, pengumpulan data, labelisasi data, prapemrosesan data, pembagian data, dan klasifikasi dari perhitungan bobot.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi mengenai hasil dan analisis Perbandingan Performa Penanganan Data *Imbalance* dan Kernel pada Analisis Sentimen Ulasan *Game Genshin Impact* Berbasis *Support Vector Machine*. Bab ini membahas tentang implementasi pelatihan

dan skenario pengujian menggunakan 3 jenis kernel yaitu Linear, RBF, dan *Polynomial* pada kondisi dua data yang berbeda.

BAB V KESIMPULAN

Bab ini berisi mengenai kesimpulan dari bab-bab yang telah dijelaskan dan saran untuk pengembangan selanjutnya.