

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Selama sepuluh tahun terakhir, kemajuan teknologi digital telah mempercepat pertumbuhan industri game global menjadi salah satu sektor hiburan paling dominan. Pasar *game* PC diprediksi akan terus tumbuh signifikan, didukung oleh akses internet yang lebih terjangkau serta inovasi perangkat keras yang memberikan pengalaman visual lebih mendalam.

Salah satu ekosistem utama yang menopang ekosistem *game* PC adalah Steam. Steam adalah platform distribusi digital yang dikembangkan oleh Valve Corporation. Dalam platform ini, pengguna dapat membeli dan mengunduh *game*, memainkan *game* bersama teman, serta mengakses berbagai fitur seperti forum diskusi, profil pengguna, koleksi *game*, hingga layanan kirim hadiah. Saat ini, Steam merupakan salah satu platform distribusi *game* terbesar di dunia. Berdasarkan laporan tahunan Valve Corporation (2024), platform ini mencatat puncak aktivitas pengguna yang terus meningkat, dengan jutaan pengguna aktif harian yang mengakses ribuan judul permainan dari berbagai genre.

Dalam lingkungan yang kompetitif ini, ulasan pengguna (*user review*) bukan lagi sekadar komentar, melainkan aset strategis bagi pengembang untuk memahami tingkat kepuasan pelanggan dan kualitas produk secara lebih mendalam. Salah satu *game* yang menonjol di platform ini adalah *Sons of the Forest* karya Endnight Games. Sejak rilis Februari 2023, *game* ini sukses besar dengan menjual sekitar 2 juta salinan hanya dalam 24 jam pertama dan menjadi salah satu judul yang paling dinantikan di Steam (Bankhurst, 2023). Pada 15 Desember 2025, volume ulasan untuk *game* ini mencapai lebih dari 280.000 ulasan, di mana sekitar 120.000 di antaranya ditulis dalam bahasa Inggris (Valve Corporation, n.d.). Jumlah ini memberikan peluang besar untuk dianalisis, tetapi pada saat yang sama juga menghadirkan tantangan karena sulit ditangani secara manual.

Genre *survival horror* yang diusung *Sons of the Forest* menambah kompleksitas analisis. Dalam konteks ini, istilah yang secara umum bernada negatif

seperti “takut” atau “mengerikan” sering kali justru digunakan untuk memuji kualitas atmosfer permainan. Misalnya, salah satu pemain menulis ulasan, “*The only time I’ve been chased by naked women. I was so scared of them that I close the game. I recommend it 10/10.*” Ungkapan “*so scared*” yang menggambarkan rasa sangat takut di sini tetap diakhiri dengan rekomendasi positif (*recommend it 10/10*). Contoh lain adalah ulasan, “*After 6 hours my conclusion is it’s beautiful, it’s terrifying and I will 100% die here again.*” yang menggunakan kata “*terrifying*” dalam konteks pengalaman yang justru ingin diulang. Fenomena ini menunjukkan bahwa kata-kata berkonotasi takut tidak selalu merepresentasikan sentimen negatif secara keseluruhan pada genre *survival horror*.

Selain itu, satu ulasan sering kali bersifat campuran (*multipolar*), misalnya dalam satu teks pemain memuji visual dan atmosfer hutan yang imersif, namun sekaligus mengeluhkan aspek lain seperti *bug*, performa yang tidak stabil, atau alur cerita yang terasa belum selesai. Contoh yang muncul pada ulasan adalah: “*Great atmosphere, solid survival mechanics and co-op is a blast. Still buggy and the story feels unfinished.*” Dari ulasan tersebut, sudah tampak adanya kombinasi apresiasi positif terhadap atmosfer, mekanik *survival*, dan mode kooperatif, sekaligus keluhan negatif terhadap stabilitas teknis (*buggy*) dan kualitas cerita (*story feels unfinished*).

Analisis sentimen konvensional yang hanya membagi ulasan menjadi “positif” atau “negatif” secara umum pada level dokumen menjadi kurang memadai untuk menangkap nuansa seperti ini. Pengembang membutuhkan wawasan yang lebih detail untuk menentukan prioritas perbaikan berdasarkan bagian-bagian tertentu dalam permainan. Oleh karena itu, diperlukan pendekatan *Aspect-Based Sentiment Analysis* (ABSA) untuk mengekstraksi dan membedah opini target secara lebih spesifik (Zhang et al., 2022). Dalam penelitian ini, analisis sentimen berbasis aspek diterapkan pada ulasan permainan video dengan mengelompokkan ulasan ke dalam lima aspek utama, yaitu *gameplay*, *graphics*, *performance*, *story*, dan *atmosphere*. Dengan ABSA, satu ulasan dapat diurai menjadi beberapa pasangan aspek sentimen, sehingga pengembang dapat mengetahui secara eksplisit aspek mana yang mendapat apresiasi dan aspek mana yang menjadi sumber keluhan pemain.

Dalam implementasinya, klasifikasi teks pada ulasan game menghadapi tantangan dimensi data yang tinggi akibat representasi fitur teks serta penggunaan bahasa tidak baku, slang, humor, dan hiperbola. Pada ulasan *Sons of the Forest* misalnya, muncul ungkapan seperti “*Kelvin and I be tearing it up!!*”, “*This game has consumed every waking hour of my life*” hingga kalimat hiperbolik bercampur humor seperti “*I will 100% die here again.*” Bentuk bahasa yang informal, penuh hiperbola, dan kadang mengandung kata kasar seperti ini menambah tantangan bagi sistem NLP dalam mengekstraksi fitur sentimen secara akurat.

Algoritma *Support Vector Machine* (SVM) dipilih dalam penelitian ini karena keunggulannya dalam menangani data berdimensi tinggi. Cortes et al. (1995) menyatakan bahwa SVM bekerja berdasarkan prinsip *Structural Risk Minimization* (SRM) yang bertujuan meminimalkan batas atas kesalahan generalisasi, sehingga model memiliki risiko *overfitting* yang lebih rendah dibandingkan metode konvensional. Studi komparatif terbaru terhadap SVM, KNN, dan Naïve Bayes pada klasifikasi sentimen ulasan aplikasi menunjukkan bahwa SVM dengan kernel *linear* dapat memberikan akurasi tertinggi di antara ketiga algoritma tersebut pada data teks berbahasa Indonesia (Aprilianti et al., 2025). Hal ini sejalan dengan karakteristik data ulasan Steam yang bersifat *sparse* dan berdimensi tinggi.

Permasalahan lain yang sering muncul dalam analisis sentimen adalah ketidakseimbangan kelas data (*imbalanced data*), di mana jumlah data pada satu kelas sentimen jauh lebih dominan dibandingkan kelas lainnya. Ketidakseimbangan ini dapat menyebabkan model *machine learning*, seperti *Support Vector Machine* (SVM), menjadi bias dan cenderung memprediksi kelas mayoritas, sementara performa pada kelas minoritas menjadi rendah. Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan teknik *Text Augmentation*. *Text Augmentation* bekerja dengan cara memodifikasi data teks asli untuk menghasilkan variasi data baru tanpa mengubah makna semantik labelnya. Menurut Wei & Zou (2019), teknik ini terbukti efektif meningkatkan performa klasifikasi teks, terutama pada dataset dengan jumlah sampel yang terbatas atau tidak seimbang, dengan tetap mempertahankan karakteristik linguistik dari ulasan pengguna. Penelitian terdahulu menunjukkan bahwa *text augmentation* dapat meningkatkan *F1-Score* kelas

minoritas secara signifikan, khususnya ketika menggunakan teknik *back translation* (Rahma & Suadaa, 2023).

Melalui pendekatan ini, diharapkan celah penelitian mengenai penggunaan empat kernel SVM pada domain *survival horror* dengan karakteristik bahasa yang unik dapat terisi.

1.2 Identifikasi Masalah

Berdasarkan latar belakang tersebut, masalah yang dapat diidentifikasi meliputi:

1. Volume data ulasan pengguna yang sangat masif dan terus bertambah menyebabkan metode analisis manual menjadi tidak efisien dan tidak efektif dalam mengekstraksi wawasan strategis secara menyeluruh.
2. Keterbatasan analisis sentimen pada tingkat dokumen dalam mengidentifikasi opini campuran yang terdapat dalam satu ulasan.
3. Risiko bias model akibat dominasi ulasan positif di platform Steam yang menyebabkan data tidak seimbang.
4. Belum terdapat referensi komprehensif yang berfokus pada evaluasi empat kernel SVM, dengan KNN digunakan sebagai *baseline*, serta efektivitas *text augmentation* pada domain *game horror*.

1.3 Pembatasan Masalah

Penelitian ini dibatasi pada hal-hal berikut untuk menjaga fokus analisis:

1. Data yang digunakan merupakan data sekunder dari repositori publik Kaggle, yang berisi ulasan pengguna platform Steam terhadap game *Sons of the Forest* secara spesifik pada masa peluncuran awal (*Early Access*), dengan rentang waktu pengambilan data dari Februari hingga Maret 2023.
2. Menggunakan ulasan berbahasa Inggris.
3. Fokus pada lima aspek: *Gameplay*, *Graphics*, *Performance*, *Story*, dan *Atmosphere*.

4. Algoritma utama yang digunakan adalah *Support Vector Machine* (SVM), dengan algoritma *K-Nearest Neighbor* (KNN) digunakan sebagai metode *baseline*.
5. Variasi fungsi kernel pada *Support Vector Machine* (SVM) dalam penelitian ini dibatasi pada empat jenis kernel bawaan yang tersedia pada pustaka *scikit-learn*, yaitu *Linear*, *Polynomial*, *Radial Basis Function* (RBF), dan *Sigmoid*.
6. Penanganan masalah ketidakseimbangan data dilakukan menggunakan teknik *Text Augmentation* dengan metode *Synonym Replacement* dan *Random Insertion*.
7. Klasifikasi sentimen dibagi menjadi tiga kategori: Positif, Negatif, dan Netral.

1.4 Rumusan Masalah

Berdasarkan penjelasan pada latar belakang, permasalahan yang dapat dirumuskan dalam penelitian ini yaitu, “Bagaimana hasil implementasi algoritma *Support Vector Machine* (SVM) dengan variasi kernel yang dipadukan dengan teknik *text augmentation* dalam melakukan analisis sentimen berbasis aspek pada ulasan game *Sons of the Forest*?”

1.5 Tujuan Penelitian

Tujuan penelitian ini adalah mengimplementasikan algoritma *Support Vector Machine* (SVM) yang dipadukan dengan teknik *text augmentation* untuk melakukan analisis sentimen berbasis aspek pada ulasan game *Sons of the Forest*, sehingga diperoleh model klasifikasi yang mampu memetakan sentimen positif, negatif, dan netral terhadap aspek *gameplay*, *graphics*, *performance*, *story*, dan *atmosphere*.

1.6 Manfaat Penelitian

Penelitian ini memiliki dua jenis manfaat utama, yaitu manfaat teoritis dan manfaat praktis. Kedua manfaat ini diharapkan dapat memberikan kontribusi baik bagi pengembangan ilmu pengetahuan maupun bagi pihak-pihak yang berkepentingan secara langsung.

1.6.1 Manfaat Teoritis:

1. Memberikan kontribusi pada literatur ABSA dengan studi kasus mendalam pada genre *survival horror*.
2. Menyediakan bukti empiris mengenai penggunaan *text augmentation* pada model SVM untuk klasifikasi sentimen ulasan *game* berbahasa Inggris.
3. Menjadi referensi kinerja komparatif antara berbagai kernel SVM dan algoritma KNN sebagai *baseline* pada tugas klasifikasi teks berdimensi tinggi.

1.6.2 Manfaat Praktis:

1. Bagi Pengembang Game: Membantu dalam menentukan prioritas perbaikan berdasarkan pemetaan sentimen per aspek pada ulasan pemain, sehingga pengembangan konten dan *patch* dapat lebih terarah.
2. Bagi Industri Game: Menawarkan kerangka kerja analisis otomatis yang dapat digunakan untuk memantau respons pasar terhadap produk *game*.
3. Bagi Peneliti dan Mahasiswa: Menyediakan contoh alur kerja pemrosesan data teks, implementasi SVM dengan berbagai kernel, serta penerapan teknik penanganan *class imbalance* pada domain ulasan game yang dapat dijadikan rujukan untuk penelitian selanjutnya.

Intelligentia - Dignitas