

1. Pendahuluan

1.1 Latar Belakang

Industri game dalam beberapa tahun terakhir sangat berkembang dengan banyaknya game yang dirilis setiap tahun. Pada saat yang sama, jumlah ulasan game di internet, terutama di platform seperti Steam, juga meningkat pesat. Ulasan yang diberikan oleh pengguna dapat memiliki dampak besar dalam menentukan pilihan konsumen dibandingkan dengan informasi yang diberikan oleh perusahaan [1].

Salah satu contoh yang signifikan adalah game Cyberpunk 2077 yang dirilis pada tahun 2020. Game yang sudah dinanti sejak 2018 tersebut mengecewakan banyak pengguna setelah dirilis. Oleh karena itu, analisis sentimen menjadi sangat penting bagi pengembang game untuk memahami perasaan dan opini pengguna terhadap game mereka. Analisis sentimen adalah teknologi yang memproses data ulasan untuk menunjukkan sikap pengguna. Ini adalah bidang studi yang menganalisa opini, sentimen, evaluasi, pertimbangan, dan emosi pengguna [2].

Penelitian sebelumnya, seperti yang dilakukan oleh Melisa Winda Pertiwi dalam analisis sentimen opini publik mengenai sarana dan transportasi mudik tahun 2019 di Twitter, menunjukkan bahwa metode K-Nearest Neighbors (KNN) dapat menghasilkan akurasi tinggi hingga 90,76% [7]. Penelitian yang menggunakan dataset yang sama, dilakukan oleh Muhammad Ro'id Akbar Aslami, dkk. dengan menggunakan metode Random Forest dan *feature extraction* TF-IDF, yang membuahkan 3 skenario. Dimana di skenario pertama mendapat F1-Score 59% dan 63%, di skenario kedua mendapat F1-Score 65% dan 40%, dan di skenario terakhir mendapat F1-Score 65% dan 37% [3].

Penelitian ini bertujuan untuk meningkatkan performansi dari penelitian sebelumnya. Analisis sentiment pada penelitian ini menggunakan data ulasan game Cyberpunk 2077 pada platform Steam dilakukan dengan metode K-Nearest Neighbors (KNN). Metode KNN dipilih karena telah terbukti menghasilkan akurasi tinggi, mencapai 96,8% saat digunakan bersama feature extraction Term Frequency Inverse Document Frequency (TF-IDF) [4]. Selain itu, kombinasi KNN dan TF-IDF juga menunjukkan akurasi sebesar 80,0% dibandingkan dengan metode ekstraksi fitur lainnya [5].

1.2 Topik dan Batasannya

Pada penelitian ini, penulis akan membuat model untuk analisis sentimen terhadap ulasan game. Penelitian ini memfokuskan pada proses *pre-processing*, *classification*, dan *feature extraction*. Pada proses *pre-processing* penulis akan membandingkan pengaruh penggunaan *lemmatization*. Proses *feature extraction* menggunakan TF-IDF dan *classification* menggunakan metode KNN. Dataset yang digunakan diambil dari penelitian sebelumnya yang diambil menggunakan *steam crawler*. Dataset yang diambil berjumlah 1000 ulasan yang sudah dilabeli menjadi positif dan negatif dengan semua ulasan berbahasa Inggris.

1.3 Tujuan

Tujuan penelitian ini untuk meningkatkan performansi dari penelitian sebelumnya tentang analisa terhadap sentimen yang ada didalam ulasan game yang diberikan oleh konsumen yang telah membeli game. Penelitian ini membandingkan pengaruh menggunakan metode K-Nearest Neighbors dengan metode Random Forest.

1.4 Organisasi Tulisan

Bagian selanjutnya pada penelitan ini adalah bagian 2 yang membahas studi yang terkait dengan penelitian yang dilakukan, bagian 3 membahas rancangan sistem yang dibangun, bagian 4 membahas evaluasi dari hasil pengujian, dan bagian 5 membahas kesimpulan dari penelitian ini dan saran untuk penelitian selanjutnya.