

BAB I

PENDAHULUAN

I.1 Latar Belakang

Memasuki era globalisasi, saat ini kemajuan pada bidang teknologi mengalami perkembangan yang sangat pesat dan menyesuaikan dengan perkembangan zaman. Kemajuan pada teknologi memiliki pengaruh yang besar bagi kehidupan manusia saat ini. Salah satu pengaruh dari kemajuan teknologi tersebut yaitu semakin canggihnya kemajuan teknologi informatika. Salah satunya yang merambah sampai ke pelosok yaitu internet. Banyak warung-warung internet atau warnet yang tersedia di pinggir jalan, yang menunjukkan banyaknya pengguna atau orang yang memanfaatkan sarana ini, diantara sarana hiburan yang sangat populer saat ini adalah *game online* yang sering dimanfaatkan oleh kalangan tua, muda bahkan anak-anak (Paremeswara & Lestari, 2021).

Game digital pada smartphone saat ini tidak hanya sebagai hiburan untuk mengisi waktu luang ataupun hobi dari penggunanya saja. Tingginya interaksi antar pengguna game untuk bermain secara kompetitif menghasilkan industri baru dalam olahraga yang dikenal dengan istilah E-Sport. E-Sport biasanya lebih mengarah kepada permainan game kompetitif yang melibatkan penonton (Freeman & Wohn, 2017). E-Sport merupakan singkatan dari Electronic Sport adalah cabang olahraga yang menggunakan media elektronik seperti komputer, konsol game, Smartphone dan lain sebagainya yang dimainkan dengan kompetitif dan dilakukan secara profesional sesuai dengan peraturan yang berlaku.

Perubahan perilaku pengguna dalam sebuah *game online* merupakan tantangan tersendiri bagi perusahaan agar dapat memenuhi kebutuhan dan keinginan penggunanya. Dengan bertambahnya jenis-jenis *game online* tentunya sangat mempengaruhi daya saing. Hal ini menyebabkan setiap pebisnis yang bergabung dengan perusahaan *game online* mempunyai cara tersendiri untuk meningkatkan daya tarik penggunanya pada *game online* tersebut. Hal-hal yang selalu diperhatikan pebisnis *game online* untuk meningkatkan daya tarik penggunanya yaitu memperhatikan kualitas pelayanan terhadap pengguna, menggunakan Newsletter, selalu mengerjakan upsell event, mengerjakan retargeting event. Namun walaupun demikian tidak jarang pengguna *game online* meragukan produk dari *game online* dikarenakan tidak mendapatkan product yang sesuai dengan event pada *game online*. Berdasarkan hasil analisis yang telah dilakukan sebagian besar setiap pengguna yang pernah memainkan *game online* ini memberikan review terhadap *game online*. Review maupun rating juga sangat penting untuk meningkatkan frekuensi pengguna *game online*, karena review dan rating dari pengguna dapat memberikan penilaian yang lebih bersifat akurat dan juga emosional karena diberikan oleh sesama pengguna *game online* sehingga memiliki nilai kepercayaan yang lebih tinggi. Oleh sebab itulah rating sebuah *game online* membutuhkan perhatian khusus dari pengelola *game online* untuk meningkatkan daya saingnya. (Farki, 2016)

Rating adalah bagian dari review yang menggunakan bentuk simbol bintang daripada bentuk teks dalam mengekspresikan pendapat dari pelanggan. Rating dapat diartikan sebagai penilaian dari pengguna pada preferensi suatu produk

terhadap pengalaman mereka mengacu pada keadaan psikologis dan emosional yang mereka jalani saat berinteraksi dengan produk virtual dalam lingkungan dimediasi. (Kim & Park, 2017) Namun kepercayaan pengguna *game online* terhadap rating dapat terhambat karena inkonsistensi jawaban seperti review yang memiliki teks negatif dengan nilai rating Positif. Adanya jawaban yang berbeda seperti rating bagus namun review buruk dan sebaliknya dapat membuat orang lain bingung. Untuk itu dibutuhkan penelitian untuk menemukan suatu cara yang akurat untuk mensinkronisasi antara review pengguna *game online* dengan rating yang diberikan karena pada dasarnya rating harus sesuai dengan review untuk memudahkan perusahaan *game online* agar meningkatkan kualitas pelayanan.(Amelia et al., 2021)

Topik dalam penelitian ini adalah bagaimana mengidentifikasi teks atau komentar-komentar pada rating *game online* yang akan di analisis dan divisualisasikan sehingga dapat dijadikan sebagai nilai rating dari *game online* tersebut. Proses mengidentifikasi teks pada sebuah database disebut dengan istilah *text mining*. Text mining adalah proses mengambil Informasi berkualitas tinggi dari teks informasi berkualitas tinggi. *Text mining* bertujuan untuk mendapatkan informasi yang berguna dari sebuah dokumen. (Ipmawati, 2016)

Pada fungsi *data mining*, dibutuhkan algoritma beserta metode yang sesuai dalam pengolahan input menjadi output yang diinginkan. Atas dasar inilah, terdapat banyak algoritma *data mining* untuk setiap fungsi. Misalnya, algoritma fungsi klasifikasi terdiri dari *Statistical-Based Algorithms*, *Distance-Based*

Algorithms, Decision Tree-Based Algorithms, Neural Network-Based Algorithms, dan Rule-Based Algorithms.(Rasheed et al., 2020).

Pada penelitian produk toko *online* berdasarkan hasil ulasan penggunanya, (Muzani et al., 2021) mengkombinasikan *text processing* berbasis TF-IDF dengan algoritma K-NN, dimana hasil terbaik diperoleh pada nilai K = 5 dengan nilai akurasi sebesar 88.89%. Dari penelitian ini, penulis mengambil nilai K = 5 sebagai referensi nilai yang digunakan dalam proses prediksi.

Penelitian yang dilakukan oleh (Imamah et al., 2020) tentang *text mining* pada ulasan turis terhadap lokasi wisata di Kabupaten Bangkalan menggunakan kombinasi antara analisis sentimen dan algoritma SVM, menyatakan bahwa algoritma SVM menghasilkan hasil akurasi sebesar 70% dalam proses klasifikasi ulasan bernilai positif dan negatif. Berdasarkan referensi penelitian ini, penulis menggunakan hasil analisis sentimen, setelah melalui tahapan *text processing*, sebagai target prediksi.

Kombinasi sentimen analisis berbasis TF-IDF dan algoritma Naïve Bayes yang dilakukan (Mandasari et al., 2022) menyatakan bahwa algoritma ini berhasil mengklasifikasikan sentimen positif, netral, dan negatif terhadap kualitas layanan transportasi *online*, dengan tingkat akurasi 86.57%. Respon sentimen dalam bentuk nilai positif, netral, dan negatif pada penelitian ini menginspirasi penulis untuk menggunakan kategori yang sama pada model prediksi di dalam penelitian ini.

Pada kasus penelitian analisis sentimen interaksi antara netizen Indonesia dengan pemerintah, (Anam et al., 2021) membandingkan performa algoritma K-NN dan Naïve Bayes, dengan hasil bahwa algoritma Naïve Bayes memiliki

performa yang lebih baik dibandingkan algoritma K-NN, dengan nilai akurasi 100% berbanding 98.25%. (Syahputra et al., 2022) pada penelitiannya mengenai analisis sentimen terhadap aplikasi Peduli Lindungi, melakukan komparasi terhadap algoritma SVM dan Naïve Bayes, dengan hasil bahwa algoritma SVM memiliki performa yang lebih baik dengan tingkat akurasi 86% dibandingkan algoritma Naïve Bayes dengan tingkat akurasi 85%. Komparasi antara algoritma SVM dan K-NN yang dilakukan oleh (Riefky & Pramesti, 2020) pada analisis sentimen terhadap penyelenggaraan SEA Games 2019 menyatakan bahwa SVM lebih unggul dibandingkan dengan K-NN, dengan perbandingan tingkat akurasi 92.96% dan 92.56%. Pada penelitian lain, (Utami et al., 2021) membandingkan ketiga algoritma ini pada penelitian analisis sentimen terhadap sistem pembelajaran *online* pada masa pandemi Covid-19, dengan hasil bahwa algoritma SVM menghasilkan performa terbaik dibandingkan kedua algoritma lainnya, dengan tingkat akurasi 87.67%. Melihat hasil penelitian-penelitian ini, penulis tertarik untuk membandingkan algoritma K-NN, SVM, dan Naïve pada model prediksi *rating game* PUBG yang dibangun.

Berdasarkan latar belakang tersebut, penelitian ini mencoba melakukan analisis pada komentar-komentar hasil review pada *game online* PUBG dengan membandingkan algoritma K-NN, SVM, dan Naïve Bayes untuk mendapatkan hasil prediksi berupa hasil *rating game* Bagus, Netral, dan Kurang Bagus. Adapun sumber data set adalah dari situs Kaggle, dengan menggunakan nilai akurasi, presisi dan *recall* yang dihasilkan dari evaluasi menggunakan 10-fold dan 20-fold *cross validation*. Hasil penelitian ini kemudian penulis tuangkan dalam bentuk sebuah

laporan tesis yang berjudul: “**Prediksi Rating *Game online* Pada PlayerUnknown Battle Ground Dengan K-NN, *Support Vector Machine* (SVM) dan *Naive Bayes*”**

I.2 Rumusan Masalah

Berdasarkan penjelasan dari latar belakang sebelumnya, rumusan masalah di dalam penelitian ini adalah bagaimana memprediksi *rating game* PUBG berdasarkan hasil ulasan pengguna, dengan membandingkan performa tiga algoritma: K-NN, SVM, dan Naïve Bayes, serta mengevaluasi hasilnya menggunakan nilai akurasi, presisi, dan *recall*.

I.3 Tujuan

Adapun tujuan yang ingin dicapai pada penelitian yang akan dilakukan ini di antaranya adalah:

1. Melakukan analisis sentimen terhadap ulasan pengguna *game* PUBG yang diperoleh dalam bentuk komentar teks, sehingga menghasilkan respon positif, netral, dan negatif, untuk digunakan sebagai target dalam proses prediksi.
2. Membangun model prediksi *rating game* PUBG, dengan menggunakan algoritma K-NN, SVM dan *Naïve Bayes*.
3. Menganalisis performa model dengan metode *cross validation*, untuk menghasilkan nilai akurasi, presisi, dan *recall* .

I.4 Manfaat

Manfaat yang diharapkan dari hasil penelitian yang akan dilakukan ini adalah sebagai berikut:

1. Menghasilkan model analisis sentimen yang memberikan output berupa respon positif, netral, dan negatif, berdasarkan dataset komentar teks dari ulasan pengguna *game* PUBG.
2. Menghasilkan analisis perbandingan performa algoritma K-NN, SVM, dan Naïve Bayes, yang diukur berdasarkan metode *cross validation*.
3. Menghasilkan model dengan performa terbaik dalam melakukan prediksi *rating game* PUBG berdasarkan nilai akurasi, presisi, dan recall.

I.5 Batasan Masalah

Adapun beberapa batasan masalah yang penulis tentukan di dalam penelitian prediksi *rating game* PUBG ini adalah sebagai berikut:

1. Data set yang digunakan adalah komentar yang berupa opini maupun review game online yang diambil dari situs Kaggle.
2. Analisis sentimen dilakukan dengan tahapan *text processing* untuk memperoleh token serta VADER untuk menghasilkan respon positif, netral dan negatif.
3. Model K-NN dibangun dengan menggunakan nilai $K = 5$, serta menggunakan dua variasi *distane metric*, yaitu Euclidean dan Manhattan.
4. Model SVM dibangun dengan menggunakan tiga variasi fungsi kernel, yaitu Polynomial, RBF, dan Sigmoid.

5. Evaluasi performa model dilakukan dengan metode *10-fold cross validation* dan *20-cross validation*, dengan nilai akurasi, presisi, serta recall sebagai bahan acuan.

I.6 Sistematika Penulisan

Sistematika proposal tesis yang akan dilakukan dalam penelitian ini terdiri dari 3 bab, yaitu:

BAB I PENDAHULUAN

Pendahuluan akan membahas tentang latar belakang dari masalah yang akan di angkat beserta rumusan masalah, tujuan penelitian, dan batasan masalah serta sistematika penulisan yang akan diuraikan dalam penyusunan laporan penelitian ini.

BAB II TINJAUAN PUSTAKA

Tinjauan pustaka menjelaskan tentang teori-teori yang akan diuraikan mengenai komentar, kritikan, saran pada aplikasi game online, beserta perangkat analisis sistem beserta metode analisis yang digunakan.

BAB III METODOLOGI PENELITIAN

Metodologi penelitian menjelaskan rangkaian tahapan dalam pembuatan sistem yang terdiri dari pengumpulan data, analisis data, perancangan sampai implementasi beserta analisis yang akan dilakukan.

BAB IV HASIL

Hasil penelitian yaitu berupa pemaparan hasil pelaksanaan metode/teknik penelitian dan menyajikan data yang mendukung hasil tersebut. Penyajian data dan penjelasannya dilakukan secara berurutan dan lengkap menggunakan teks dan ilustrasi lainnya (misalnya, tabel dan gambar) sebagai pendukung hasil.

BAB V PEMBAHASAN

Pembahasan penelitian yaitu menerjemahkan makna dari hasil yang diperoleh untuk menjawab pertanyaan atau masalah penelitian. Fungsi lainnya adalah untuk menjelaskan pemahaman baru yang didapatkan dari hasil penelitian, yang diharapkan berguna dalam pengembangan keilmuan. Dalam penelitian tingkat lanjut, hasil pembahasan diharapkan menunjukkan kontribusi penulis terhadap pengembangan keilmuan. Pada penelitian ini, hasil pembahasan bukan pengembangan keilmuan secara substansial, tetapi cukup terhadap pemahaman personal dalam implementasi konsep atau teori.

BAB VI KESIMPULAN DAN SARAN

Kesimpulan dan saran berisi hasil penelitian klasifikasi komentar maupun review yang dikelompokkan menjadi rating menggunakan metode SVM dan *Naive Bayes*, Visualisasi dan simulasi interaktif dengan menggunakan Phyton. Bagian kesimpulan berisi jawaban atas

rumusan masalah pada penelitian dan pada bagian saran berupa pengembangan penelitian dan kendala yang dihadapi dalam menyusun penelitian sehingga dapat memberikan masukan bagi penelitian selanjutnya.