

BAB III

METODOLOGI PENELITIAN

3.1 Gambaran Umum Objek Penelitian

Objek penelitian yang digunakan dalam penelitian ini adalah Kumpulan teks dari halaman ulasan *game* Marvel Rivals di website Steam. Data yang digunakan dalam penelitian ini mencakup komentar yang diberikan oleh pengguna mengenai aspek-aspek permainan, termasuk mekanik, grafik, karakter, dan pengalaman bermain secara keseluruhan. Dengan mengolah dan menganalisis data ini, penelitian ini bertujuan untuk mengidentifikasi pola sentimen yang muncul serta faktor-faktor utama yang mempengaruhi persepsi pemain terhadap game Marvel Rivals. Data dikumpulkan melalui website Apify yang memberikan layanan untuk *web scraping* dari berbagai sumber dari internet.

3.2 Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan membandingkan performa berbagai algoritma klasifikasi dalam menganalisis sentimen pemain terhadap game Marvel Rivals. Model yang diuji mencakup algoritma *machine learning* seperti Random Forest, Multinomial Naïve Bayes, dan XGBoost, serta algoritma *deep learning* yaitu Bi-LSTM (Bidirectional Long Short-Term Memory). Untuk mengarahkan proses analisis data secara sistematis dan terstruktur, penelitian ini mengacu pada kerangka kerja CRISP-DM (Cross-Industry Standard Process for Data Mining). Dalam proses *data mining*, terdapat beberapa framework yang umum digunakan, yaitu KDD (Knowledge Discovery in Databases), SEMMA (Sample, Explore, Modify, Model, Assess), dan CRISP-DM. Masing-masing memiliki kelebihan dan kekurangan yang dirangkum dalam Tabel 3.1.

Tabel 3. 1 Perbandingan Metode Data Mining

Metode	Kelebihan	Kekurangan	Tahapan
CRISP-DM	Mencakup siklus hidup proyek data mining yang lengkap	Fase deployment bisa menjadi sangat kompleks dan	<i>Business Understanding,</i> <i>Data Understanding,</i> <i>Data Preparation,</i>

Metode	Kelebihan	Kekurangan	Tahapan
	dari konsepsi hingga implementasi.	menantang dalam implementasi praktisnya di lingkungan bisnis nyata.	<i>Modeling, Evaluation, Deployment</i>
KDD	Memberikan pemahaman konseptual yang baik tentang proses penemuan pengetahuan dari data secara keseluruhan.	Tidak memiliki penekanan yang kuat pada implementasi.	<i>Selection, Preprocessing, Transformation, Data Mining, Interpretation/Evaluation</i>
SEMMA	Dikembangkan oleh SAS, sehingga sangat relevan dan praktis bagi pengguna ekosistem SAS	Tidak menekankan fase pemahaman bisnis dan interpretasi yang mendalam.	<i>Sample, Explore, Modify, Model, Assess</i>

Kerangka CRISP-DM dipilih dalam penelitian ini karena menawarkan pendekatan yang komprehensif, fleksibel, dan terstruktur, yang sesuai untuk studi analisis sentimen berbasis *machine learning* dan *deep learning*. CRISP-DM mencakup aspek pemahaman bisnis yang mendefinisikan tujuan dan kebutuhan proyek dari perspektif masalah yang ingin diselesaikan, yakni memahami sentimen pemain terhadap *game* Marvel Rivals untuk memberikan rekomendasi

kepada pengembang. Adanya tahap *deployment* memastikan bahwa solusi teknis yang dikembangkan relevan dan memberikan nilai. Selain itu, framework ini bersifat iteratif, memungkinkan penyesuaian pada tiap tahap berdasarkan temuan sebelumnya, yang sangat penting dalam proses eksploratif seperti analisis sentimen terhadap opini pemain *game*.

Keempat model yang dipilih—Random Forest (RF), Multinomial Naïve Bayes (MNB), XGBoost, dan Bidirectional Long Short-Term Memory (Bi-LSTM) memiliki pendekatan yang berbeda dalam pemrosesan data dan penentuan prediksi. Model-model dalam penelitian ini mencakup metode probabilistik (MNB), teknik ensemble yang kuat (RF dan XGBoost), dan arsitektur deep learning yang memproses data secara sekuensial (Bi-LSTM). Tabel 3.2 menguraikan cara kerja umum, serta kelebihan dan kekurangan dari masing-masing algoritma secara umum.

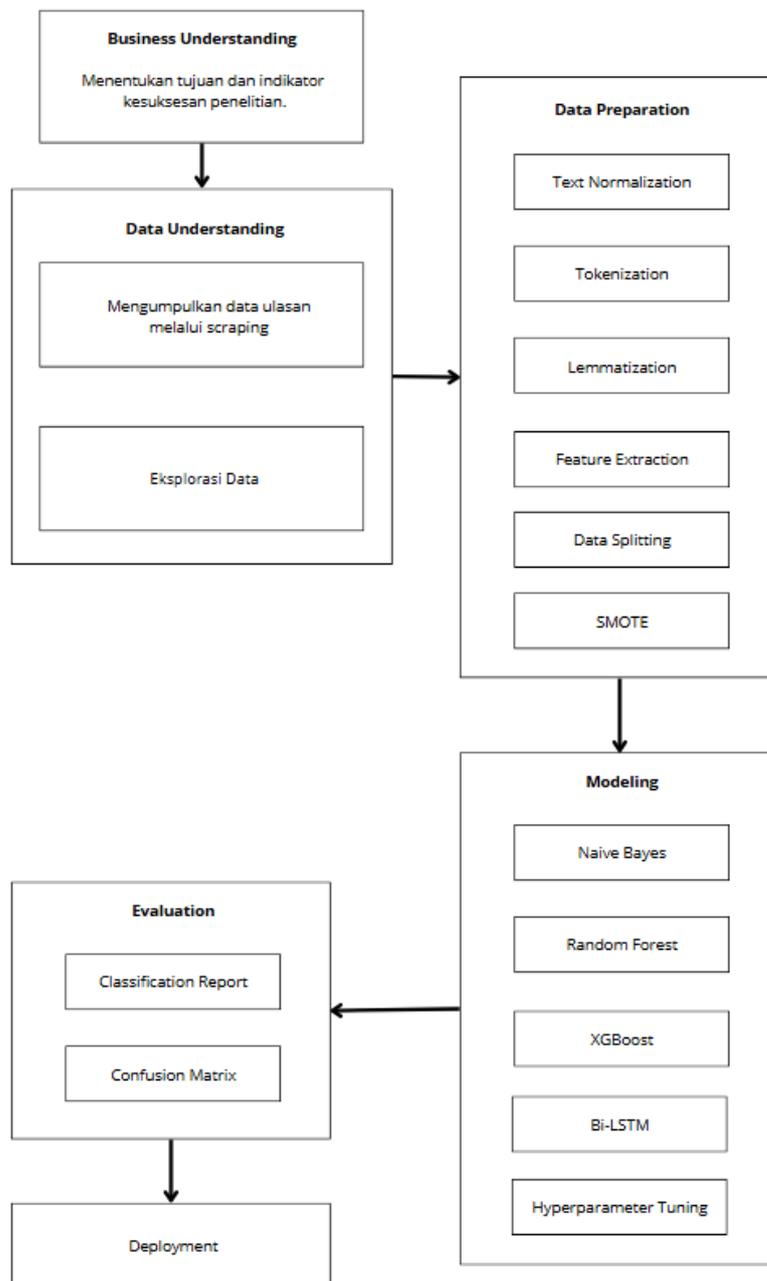
Tabel 3. 2 Perbandingan Algoritma

Algoritma	Cara Kerja	Kekurangan	Kelebihan
Multinomial Naïve Bayes (MNB)	Menghitung probabilitas sebuah data masuk ke kelas tertentu berdasarkan frekuensi fitur (kata), dengan asumsi fitur bersifat independen.	Asumsi bahwa setiap fitur (kata) bersifat independen tidak sesuai dengan bahasa natural.	Sangat cepat dan efisien secara komputasi
Random Forest (RF)	Membangun banyak pohon keputusan secara independen dan menggunakan	Tidak menangkap konteks kata.	Kuat terhadap overfitting. Fleksibel untuk berbagai jenis data.

Algoritma	Cara Kerja	Kekurangan	Kelebihan
	suara mayoritas untuk prediksi akhir.		
Extreme Gradient Boosting (XGBoost)	Menggunakan boosting dengan menambahkan pohon secara bertahap; setiap pohon baru memperbaiki kesalahan pohon sebelumnya berdasarkan gradien dari loss.	Rentan overfitting jika tidak dituning dengan benar.	Memiliki regularisasi internal untuk mencegah overfitting.
Bidirectional Long Short-Term Memory (Bi-LSTM)	Memproses data sekuensial (teks) dari dua arah (maju dan mundur) untuk memahami konteks penuh dari setiap kata..	Intensif secara komputasi	Unggul dalam memahami konteks dan urutan kata.

3.3 Alur Penelitian

Alur kerja penelitian mengikuti kerangka kerja CRISP-DM dan digambarkan secara visual dalam Gambar 3.1. Berikut adalah penjelasan tahapan-tahapannya:



Gambar 3. 1 Alur Kerja Penelitian

Pada tahap *Business Understanding* , ditentukan tujuan dan indikator kesuksesan penelitian. Ini melibatkan pendefinisian tujuan dan kebutuhan proyek dari sudut pandang bisnis atau masalah yang ingin diselesaikan, serta kriteria keberhasilan. Dalam konteks penelitian ini, ini berarti memahami mengapa analisis sentimen ulasan Marvel Rivals penting bagi pengembang dan apa yang diharapkan dari hasil penelitian.

Tahap *Data Understanding* berfokus pada pengumpulan data awal dan eksplorasi untuk memahami karakteristik data tersebut. Data ulasan pemain Marvel Rivals dikumpulkan dari platform Steam menggunakan teknik web scraping dengan bantuan website Apify. Setelah data terkumpul, dilakukan eksplorasi awal untuk memahami kualitas, kuantitas, dan karakteristik umum dari data ulasan tersebut. Ini termasuk melihat jumlah data, contoh ulasan, serta distribusi sentimen.

Tahap *Data Preparation* mencakup semua aktivitas untuk membersihkan dan mengubah data mentah menjadi dataset yang siap untuk tahap pemodelan. Langkah-langkah yang dilakukan meliputi:

- 1) Text Normalization: Membersihkan teks ulasan dari karakter yang tidak diperlukan (seperti URL, mentions, tanda baca, angka) dan mengubah semua teks menjadi huruf kecil.
- 2) Tokenization: Memecah kalimat dalam ulasan menjadi unit-unit kata individual (token).
- 3) Lemmatization: Mengubah setiap kata ke bentuk dasarnya (lemma) untuk mengurangi variasi kata dan dimensi fitur. (Sebelumnya juga ada stopwords removal setelah tokenisasi).
- 4) Feature Extraction: Mengubah teks yang telah diproses menjadi fitur numerik menggunakan TF-IDF (Term Frequency-Inverse Document Frequency).
- 5) Data Splitting: Membagi dataset menjadi data latih (*training set*) dan data uji (*test set*).
- 6) SMOTE (Synthetic Minority Over-sampling Technique): Menerapkan teknik SMOTE pada data latih untuk menangani ketidakseimbangan kelas antara sentimen positif dan negatif dengan membuat data sintesis pada kelas minoritas.

Pada tahap *Modeling*, empat algoritma diterapkan untuk memprediksi sentimen ulasan. Algoritma yang digunakan adalah Multinomial Naive Bayes, Random Forest, Extreme Gradient Boosting, dan Bidirectional Long Short-Term Memory. Proses pemodelan ini juga melibatkan *hyperparameter tuning*, di mana parameter dari masing-masing algoritma disesuaikan untuk menemukan kombinasi yang menghasilkan performa terbaik. Model-model ini dilatih dan diuji dalam beberapa

skenario, yaitu tanpa SMOTE dan tanpa *hyperparameter tuning*, dengan SMOTE dan tanpa *hyperparameter tuning*, tanpa SMOTE dan dengan *hyperparameter tuning*, serta dengan SMOTE dan dengan *hyperparameter tuning*.

Pada tahap *Evaluation*, keempat model dari setiap skenario dievaluasi menggunakan *classification report* dan *confusion matrix*. *Classification report* Menyajikan nilai precision, recall, dan F1-score untuk setiap kelas sentimen (positif dan negatif), serta akurasi keseluruhan model. Sedangkan *confusion matrix* digunakan untuk melihat secara detail jumlah prediksi yang benar dan salah untuk setiap kelas. Evaluasi ini bertujuan untuk membandingkan kinerja antar model dan skenario, serta menentukan model terbaik yang paling andal dalam mengklasifikasikan sentimen.

Tahap *Deployment* bertujuan untuk mengaplikasikan model dan pengetahuan yang didapat ke dalam penggunaan praktis. *Deployment* diwujudkan dengan mengembangkan aplikasi berbasis web interaktif menggunakan Streamlit. Aplikasi ini mampu melakukan web scraping ulasan terbaru, memproses data, melakukan prediksi sentimen menggunakan model terbaik yang terpilih, dan memberikan penjelasan sentimen menggunakan LLM Llama 3 dengan metode RAG (Retrieval-Augmented Generation).

3.4 Teknik Analisis Data

Dalam penelitian ini, Bahasa pemrograman Python digunakan dalam *software* Jupyter Notebook. Penelitian ini menggunakan analisis kuantitatif oleh karena penerapan *machine learning* dan *deep learning* yang menerapkan konsep matematika untuk melakukan *sentiment analysis*. Data yang didapatkan akan diklasifikasi menjadi dua kategori, yakni negatif dan positif, Tabel 3.2 berisi daftar *library* yang digunakan dalam analisis.

Tabel 3. 3 Daftar Library Python Penelitian

Nama Library	Fungsi
pandas	Manipulasi dan analisis data dalam bentuk DataFrame
numpy	Komputasi numerik, termasuk operasi matriks dan array
re	Memanipulasi <i>string</i> menggunakan <i>regular expression</i>

Nama Library	Fungsi
string	Manipulasi string, seperti penghapusan tanda baca
nltk	Pustaka untuk tugas pemrosesan bahasa alami
nltk.corpus.stopwords	Menghapus kata-kata umum (stopwords)
nltk.stem.WordNetLemmatizer	Mengubah kata ke bentuk dasarnya
nltk.tokenize.word_tokenize	Memisahkan teks menjadi kata-kata
sklearn.model_selection.train_test_split	Membagi dataset menjadi data latih dan uji
sklearn.model_selection.GridSearchCV	Mencari parameter terbaik dalam model dengan cross-validation
sklearn.feature_extraction.text.TfidfVectorizer	Mengubah teks menjadi vektor numerik berbasis TF-IDF
imblearn.over_sampling.SMOTE	Menangani ketidakseimbangan data dengan oversampling
sklearn.metrics.classification_report	Menampilkan laporan evaluasi model klasifikasi
sklearn.metrics.accuracy_score	Menghitung akurasi model
sklearn.metrics.confusion_matrix	Membuat confusion matrix
sklearn.metrics.roc_curve, auc	Menghitung dan memvisualisasikan kurva ROC
sklearn.ensemble.RandomForestClassifier	Algoritma Random Forest untuk klasifikasi
sklearn.naive_bayes.MultinomialNB	Algoritma Naive Bayes
sklearn.preprocessing.LabelEncoder	Mengubah label kategori menjadi angka
xgboost	Model XGBoost
matplotlib.pyplot	Membuat grafik dan visualisasi data
seaborn	Membuat grafik dan visualisasi data
wordcloud.WordCloud	Membuat visualisasi word cloud dari teks
joblib	Menyimpan dan memuat objek model
tensorflow	framework <i>machine learning</i> open-source untuk membangun dan menjalankan model AI
tensorflow.keras.models.Sequential	Membuat model neural network secara berurutan layer per layer.
tensorflow.keras.layers.Dense	Lapisan Dense
tensorflow.keras.layers.Embedding	Lapisan Embedding

Nama Library	Fungsi
tensorflow.keras.layers.LSTM	Lapisan LSTM
tensorflow.keras.layers.Bidirectional	Menjalankan Bi-LSTM
tensorflow.keras.preprocessing.text.Tokenizer	Mengubah teks menjadi token angka yang bisa diproses oleh model Bi-LSTM
keras.preprocessing.sequence.pad_sequences	Menyamakan panjang urutan (sequence) teks sebelum masuk ke model.
tensorflow.keras.optimizers.Adam	Algoritma optimisasi
keras_tuner	Mencari hyperparameter terbaik secara otomatis saat melatih model.
tensorflow.keras.models.load_model	Memuat model yang sudah dilatih sebelumnya.
random	Menghasilkan angka acak

3.5 Teknik Pengujian atau Validasi Sistem

3.5.1 *Confusion Matrix*

Confusion Matrix adalah metode yang digunakan untuk menilai sejauh mana model klasifikasi berkinerja dengan membandingkan label yang diprediksi oleh model dengan label sebenarnya dalam dataset. *Confusion Matrix* memberikan ringkasan tentang hasil klasifikasi, termasuk jumlah prediksi yang benar, prediksi yang salah, serta sejauh mana model mampu membedakan antara kelas positif dan negatif [96].

Dalam kasus klasifikasi biner, *confusion matrix* adalah sebuah matriks 2x2 yang terdiri dari empat sel: *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN) [96]. Sel TP adalah jumlah kasus positif yang diprediksi dengan benar, sel FP adalah jumlah kasus negatif yang diprediksi sebagai kasus positif, sel TN adalah jumlah kasus negatif yang diprediksi dengan benar, dan sel FN adalah jumlah kasus positif yang diprediksi sebagai kasus negatif [96]. Visualisasi *confusion matrix* dapat dilihat pada Gambar 2.3.

		Predicted Value	
		Positive	Negative
Actual Value	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

Gambar 3. 2 Confusion Matrix

3.5.2 Accuracy

Accuracy adalah jumlah TP dan TN dibagi dengan jumlah total dataset, dengan nilai terbaik adalah 1.0 dan terburuk 0.0 [96], rumus *accuracy* yakni:

$$\frac{TP + TN}{P + N}$$

Rumus 3. 1 Accuracy

3.5.3 Precision

Precision adalah jumlah TP dibagi dengan jumlah TP dan FP, dengan nilai terbaik adalah 1.0 dan terburuk 0.0 [96], rumus *precision* yakni:

$$\frac{TP}{TP + FP}$$

Rumus 3. 2 Precision

3.5.4 Recall

Recall adalah jumlah TP dibagi dengan jumlah semua positif, dengan nilai terbaik adalah 1.0 dan terburuk 0.0 [96], rumus *recall* yakni:

$$\frac{TP}{TP + FN}$$

Rumus 3. 3 Recall

3.5.5 F1-Score

F1-Score adalah suatu nilai tunggal yang mewakili *harmonic mean* dari *precision* dan *recall* [97]. Rumus *F1-Score* yakni:

$$2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Rumus 3. 4 F1-Score