

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Natural Language Processing (NLP)

Natural Language Processing merupakan bidang interdisipliner yang menggabungkan linguistik komputasional, ilmu komputer, kecerdasan buatan, dan ilmu kognitif untuk memungkinkan komputer memahami, menafsirkan, dan menghasilkan bahasa manusia secara bermakna dan bermanfaat (Jurafsky & Martin, 2023; Bird et al., 2009). NLP mencakup spektrum tugas yang luas, mulai dari pemrosesan sintaksis tingkat rendah (tokenisasi, penandaan kelas kata, penguraian sintaksis) hingga pemahaman semantik tingkat tinggi (pengenalan entitas bernama, ekstraksi relasi, tanya jawab, penerjemahan mesin) dan interpretasi pragmatik (analisis sentimen, deteksi sarkasme, manajemen dialog).

Tugas-tugas NLP yang relevan untuk analisis sentimen meliputi: (1) Tokenisasi: segmentasi aliran teks menjadi unit-unit diskrit (kata, subkata, karakter) yang berfungsi sebagai satuan pemrosesan dasar, dengan tantangan penanganan tanda baca, kontraksi, kata majemuk, dan konvensi ortografis bahasa (Manning et al., 2008). (2) Normalisasi: standarisasi representasi teks melalui pelipatan kasus, koreksi ejaan, perluasan singkatan, dan penanganan variasi bentuk (misalnya, 'color' vs 'colour'), yang sangat penting dalam konten buatan pengguna yang banyak menggunakan bahasa informal (Eisenstein, 2013). (3) Stemming dan

Lemmatisasi: mereduksi kata ke bentuk dasarnya untuk memungkinkan pengenalan kesetaraan semantik (misalnya, ‘berlari’, ‘berlarian’, ‘lari’ → ‘lari’), di mana stemming menggunakan penghapusan akhiran berbasis aturan (Porter, 1980), sementara lemmatisasi menggunakan analisis morfologi untuk mendapatkan bentuk dasar yang sesuai konteks (Lovins, 1968). (4) Penandaan Kelas Kata (POS Tagging): pemberian kategori gramatikal (kata benda, kata kerja, kata sifat, dan sebagainya) yang memungkinkan pemrosesan berbasis sintaksis, dicapai menggunakan model statistik (Hidden Markov Models, Maximum Entropy Markov Models) atau pendekatan neural (BiLSTM-CRF) (Toutanova et al., 2003). (5) Penghapusan Kata Henti (Stop Word Removal): penyaringan kata fungsi umum (‘the’, ‘a’, ‘is’) yang biasanya tidak memberikan informasi diskriminatif untuk tugas klasifikasi, meskipun memerlukan pertimbangan khusus dalam analisis sentimen di mana negasi (‘tidak’, ‘jangan’) sangat penting bagi sentimen (Saif et al., 2014).

Dalam konteks permainan digital secara khusus, NLP menghadapi tantangan unik yang memerlukan penanganan tersendiri, antara lain: banyaknya terminologi khusus domain (istilah mekanika permainan, nama karakter, nama lokasi), penggunaan singkatan dan akronim yang luas (‘GG’ untuk permainan yang bagus, ‘OP’ untuk terlalu kuat, ‘AFK’ untuk meninggalkan keyboard), variasi ejaan kreatif dan kesalahan ejaan yang disengaja (‘pwned’, ‘noob’), penggunaan emoji dan emotikon yang sering membawa informasi sentimen (‘😊’, ‘😡’), pencampuran bahasa dengan fenomena alih kode (Li, 2023), penggunaan ragam

bahasa yang sangat informal khas komunikasi daring santai, serta sarkasme dan ironi yang mempersulit interpretasi sentimen secara harfiah (Poria et al., 2016).

2.1.2 Algoritma Naïve Bayes: Fondasi Matematis

Naïve Bayes merupakan keluarga pengklasifikasi probabilistik yang berakar pada penerapan teorema Bayes dengan asumsi independensi kondisional yang kuat (naïve) antar fitur (Zhang, 2004). Meskipun asumsi independensi tersebut sering dilanggar dalam data dunia nyata, pengklasifikasi Naïve Bayes menunjukkan kinerja yang sangat tangguh di berbagai aplikasi, khususnya klasifikasi teks (Domingos & Pazzani, 1997).

Teorema Bayes menetapkan hubungan antara probabilitas kondisional:

$$P(C|X) = [P(X|C) \times P(C)] / P(X)$$

Dimana $P(C|X)$ adalah probabilitas posterior dari kelas C dengan diketahuinya fitur yang diamati X , $P(X|C)$ adalah kemungkinan (likelihood), $P(C)$ adalah probabilitas prior, dan $P(X)$ adalah probabilitas bukti.

Untuk klasifikasi teks dengan representasi kantong kata (bag-of-words), Multinomial Naïve Bayes menggunakan formula yang menguraikan kemungkinan (likelihood) sebagai hasil kali probabilitas fitur individual. Keputusan klasifikasi mengikuti prinsip Maximum A Posteriori (MAP), yaitu memilih kelas yang memaksimalkan probabilitas posterior. Keunggulan teoretis yang menjelaskan kinerja yang tangguh meliputi: efisiensi komputasional, efisiensi sampel (berkinerja baik dengan himpunan data latih yang relatif kecil), ketahanan terhadap fitur yang tidak relevan, keluaran probabilistik yang memberikan estimasi

ketidakpastian, dan justifikasi teoretis bahwa Naïve Bayes dapat mencapai klasifikasi optimal bahkan ketika asumsi independensi dilanggar, asalkan dependensi terdistribusi secara merata di seluruh kelas.

Analisis Sentimen, juga disebut opinion mining, merupakan studi komputasional tentang opini, sentimen, evaluasi, sikap, penilaian, dan emosi yang diekspresikan dalam bentuk tekstual (Liu, 2020). Tujuan utama melibatkan penentuan keadaan afektif atau orientasi subjektif (positif, negatif, netral) penulis terhadap topik, produk, layanan, organisasi, individu, atau peristiwa tertentu.

Analisis sentimen dapat didekati melalui berbagai paradigma metodologis: (1) Pendekatan Berbasis Leksikon yang bergantung pada kamus yang telah ditentukan sebelumnya yang memetakan kata atau frasa ke skor sentimen (VADER, SentiWordNet, AFINN), (2) Pendekatan Pembelajaran Mesin yang memperlakukan analisis sentimen sebagai masalah klasifikasi terawasi, melatih model pada contoh berlabel (Naïve Bayes, SVM, Decision Trees, Random Forests), dan (3) Pendekatan Pembelajaran Mendalam yang memanfaatkan arsitektur neural yang mampu pembelajaran fitur otomatis dari teks mentah (RNN, LSTM, GRU, CNN, Transformers seperti BERT).

Granularitas analisis sentimen bervariasi: tingkat dokumen (menentukan sentimen keseluruhan dari seluruh ulasan), tingkat kalimat (menganalisis sentimen dari kalimat individual), dan berbasis aspek (mengidentifikasi sentimen terhadap fitur atau aspek spesifik, yang paling dapat ditindaklanjuti untuk perbaikan yang ditargetkan).

2.1.3 Model IndoBERT sebagai Algoritma Pembanding

IndoBERT merupakan model bahasa pre-trained berbasis arsitektur Transformer yang dikembangkan khusus untuk memproses teks berbahasa Indonesia, mengikuti kerangka Bidirectional Encoder Representations from Transformers (BERT) yang diperkenalkan oleh Devlin et al. (2019). Berbeda dengan model word embedding statis, IndoBERT menghasilkan representasi kontekstual di mana makna sebuah kata dapat berubah bergantung pada konteks kalimat tempat kata tersebut muncul, sehingga mampu menangkap nuansa semantik, negasi, dan ambiguitas bahasa yang sulit ditangkap oleh pendekatan bag-of-words seperti Naïve Bayes (Koto et al., 2020). Model ini dilatih awal (pre-trained) pada korpus besar teks Bahasa Indonesia melalui skema masked language modeling dan next sentence prediction, kemudian dapat disesuaikan (fine-tuned) untuk tugas hilir tertentu, termasuk klasifikasi sentimen (Wilie et al., 2020).

Dalam penelitian ini, IndoBERT digunakan sebagai algoritma pembanding terhadap Multinomial Naïve Bayes untuk mengukur sejauh mana representasi kontekstual berbasis Transformer memberikan peningkatan akurasi klasifikasi sentimen ulasan game 'Fish It', sekaligus mengevaluasi trade-off dari sisi kompleksitas komputasional, kebutuhan sumber daya, dan waktu pelatihan. Penelitian terdahulu pada ulasan Roblox berbahasa Indonesia menunjukkan bahwa IndoBERT yang dikombinasikan dengan teknik penyeimbangan kelas SMOTE mampu meningkatkan akurasi klasifikasi secara signifikan dibandingkan baseline (Wahyuni & Romli, 2024), sehingga menjadikannya pembanding yang relevan bagi pendekatan Naïve Bayes yang lebih ringan secara komputasional.

2.1.4 Analisis Sentimen: Teori dan Pendekatan

Analisis Sentimen, juga disebut opinion mining, merupakan studi komputasional tentang opini, sentimen, evaluasi, sikap, penilaian, dan emosi yang diekspresikan dalam bentuk tekstual (Liu, 2020). Tujuan utama melibatkan penentuan keadaan afektif atau orientasi subjektif (positif, negatif, netral) penulis terhadap topik, produk, layanan, organisasi, individu, atau peristiwa tertentu.

Analisis sentimen dapat didekati melalui berbagai paradigma metodologis:

(1) Pendekatan Berbasis Leksikon yang bergantung pada kamus yang telah ditentukan sebelumnya yang memetakan kata atau frasa ke skor sentimen (VADER, SentiWordNet, AFINN), (2) Pendekatan Pembelajaran Mesin yang memperlakukan analisis sentimen sebagai masalah klasifikasi terawasi, melatih model pada contoh berlabel (Naïve Bayes, SVM, Decision Trees, Random Forests), dan (3) Pendekatan Pembelajaran Mendalam yang memanfaatkan arsitektur neural yang mampu pembelajaran fitur otomatis dari teks mentah (RNN, LSTM, GRU, CNN, Transformers seperti BERT).

Granularitas analisis sentimen bervariasi: tingkat dokumen (menentukan sentimen keseluruhan dari seluruh ulasan), tingkat kalimat (menganalisis sentimen dari kalimat individual), dan berbasis aspek (mengidentifikasi sentimen terhadap fitur atau aspek spesifik, yang paling dapat ditindaklanjuti untuk perbaikan yang ditargetkan).

2.1.5 Game ‘Fish It’ Karakteristik dan Mekanika sebagai Objek Penelitian

Fish It merupakan game bergenre simulasi memancing (fishing simulation) sekaligus role-playing game (RPG) dunia terbuka yang dikembangkan oleh

kelompok pengembang independen Fish Atelier di platform Roblox sejak Oktober 2024. Game ini terinspirasi oleh judul-judul sejenis yang telah lebih dahulu populer, yaitu Fisch dan Fishing Simulator, namun mengembangkan sistem permainan yang lebih kompleks melalui penambahan elemen eksplorasi pulau, sistem cuaca dinamis, dan variasi spesies ikan yang sangat beragam. Dalam waktu relatif singkat, Fish It berhasil menempati posisi sebagai salah satu game dengan jumlah pemain aktif tertinggi di platform Roblox, sehingga menghasilkan volume ulasan dan komentar pengguna yang sangat besar di berbagai media sosial, termasuk TikTok, yang menjadikannya objek yang relevan bagi penelitian analisis sentimen berskala besar.

Mekanika inti permainan Fish It berpusat pada aktivitas memancing yang melibatkan beberapa komponen saling terkait. Pertama, sistem Fishing Rod (joran pancing) menyediakan berbagai tingkatan alat dengan statistik berbeda yang dapat diperoleh melalui pembelian dalam game maupun hadiah event. Kedua, sistem Enchanting memungkinkan pemain memberikan mantra tambahan pada joran untuk meningkatkan peluang mendapatkan ikan langka atau ikan bermutasi. Ketiga, sistem Bait/Bobber (umpan) memengaruhi statistik Luck pemain sehingga memengaruhi kualitas hasil tangkapan. Keempat, permainan menghadirkan sistem eksplorasi berbasis Islands (pulau-pulau) yang dapat dijangkau menggunakan Boats (kapal), dengan masing-masing pulau memiliki karakteristik ekosistem ikan yang berbeda. Kelima, permainan menerapkan sistem cuaca dinamis (dynamic weather system) yang memengaruhi kemunculan spesies ikan tertentu, misalnya beberapa spesies hanya muncul pada kondisi badai atau malam hari. Kompleksitas mekanika

ini menjadi sumber terminologi khusus (in-game jargon) yang banyak muncul dalam ulasan pengguna, seperti ‘rod’, ‘bobber’, ‘mutation’, ‘enchant’, dan nama-nama pulau atau ikan tertentu, sehingga relevan dengan tantangan NLP domain-spesifik.

Salah satu elemen yang menjadi daya tarik utama Fish It adalah sistem rarity (tingkat kelangkaan) dan mutation (mutasi) pada ikan hasil tangkapan. Setiap ikan diklasifikasikan ke dalam beberapa tingkat kelangkaan, mulai dari common hingga legendary, dengan lebih dari 300 spesies unik yang tersebar di berbagai lokasi. Selain tingkat kelangkaan spesies, setiap ikan juga berpeluang muncul dengan mutasi kosmetik acak (misalnya Gold, Albino, Shiny, hingga Galaxy) yang memengaruhi nilai jual dan tampilan visual ikan, meskipun tidak mengubah mekanika permainannya. Kombinasi faktor keberuntungan (random number generator/RNG), strategi peralatan, dan kondisi lingkungan inilah yang mendorong pemain untuk aktif berbagi pengalaman, kekecewaan, maupun kegembiraan mereka dalam bentuk ulasan dan komentar di media sosial, yang menjadi sumber data utama penelitian ini.

Dari sisi sosial dan ekonomi dalam game, Fish It menyediakan fitur trading yang memungkinkan antar-pemain saling menukar ikan maupun peralatan, serta fitur monetisasi berupa game pass dan pembelian menggunakan mata uang virtual Robux untuk mempercepat progres permainan. Karakteristik ini menjadikan ulasan pengguna Fish It kerap menyentuh isu keseimbangan permainan (game balance), kewajaran sistem monetisasi (pay-to-win), maupun dinamika interaksi sosial antarpemain, yang secara langsung berkorelasi dengan dimensi Sistem

Perkembangan & Penghargaan serta Monetisasi & Persepsi Nilai pada kerangka evaluasi UX lima dimensi.

Karakteristik-karakteristik tersebut menjadikan Fish It sebagai objek penelitian yang tepat bagi penerapan analisis sentimen berbasis Multinomial Naïve Bayes dan IndoBERT. Ulasan pengguna Fish It di TikTok umumnya ditulis dalam bahasa Indonesia informal yang bercampur dengan istilah gaul, singkatan, dan terminologi khas permainan, sehingga menuntut tahapan prapemrosesan data yang cermat, sekaligus menyediakan konteks yang kaya untuk menghubungkan hasil klasifikasi sentimen dengan dimensi-dimensi User Experience secara konkret pada BAB IV.

2.1.6 Platform Roblox: Karakteristik dan Ekosistem

Roblox merepresentasikan platform hibrida unik yang menggabungkan mesin game, alat pengembangan, pasar distribusi, dan jejaring sosial, memungkinkan pengguna untuk membuat dan berbagi pengalaman 3D yang imersif (Craig & Pepler, 2023). Platform ini menggunakan bahasa pemrograman Lua dan lingkungan pengembangan Roblox Studio yang bersifat eksklusif, menawarkan hambatan masuk yang relatif rendah dibandingkan dengan mesin game tradisional sekaligus memberikan kemampuan yang andal bagi pengembang berpengalaman.

Statistik platform (per Q4 2024) mendemonstrasikan skala besar: 70,2 juta pengguna aktif harian (DAU), 320+ juta pengguna aktif bulanan (MAU), rata-rata 2,4 jam keterlibatan harian per pengguna, 15 juta+ pengalaman buatan pengguna

tersedia, \$2,8 miliar didistribusikan kepada pengembang/kreator pada 2023, pengguna yang mencakup 180+ negara dengan jangkauan global.

2.1.7 Pengalaman Pengguna (User Experience) dalam Konteks Desain Game

Pengalaman Pengguna (User Experience/UX) mencakup semua aspek interaksi pengguna akhir dengan perusahaan, layanannya, dan produknya, didefinisikan oleh Organisasi Internasional untuk Standardisasi (ISO 9241-210) sebagai persepsi dan respons seseorang yang dihasilkan dari penggunaan atau antisipasi penggunaan produk, sistem, atau layanan (ISO, 2019). Dalam konteks permainan digital, UX melampaui kegunaan tradisional untuk mencakup kualitas hedonis termasuk kesenangan, penghayatan, tantangan, konektivitas sosial, keterlibatan emosional, dan kondisi aliran (flow).

Kerangka UX spesifik permainan mengidentifikasi berbagai dimensi: Game Experience Questionnaire (GEQ), model Player Experience of Need Satisfaction (PENS), dan model GameFlow. Berdasarkan sintesis literatur desain game, penelitian ini mengadopsi kerangka UX lima dimensi: (1) Mekanika Permainan (sistem interaksi inti termasuk respons kontrol, keseimbangan permainan, perkembangan tingkat kesulitan), (2) Desain Visual & Audio (kualitas estetika termasuk kualitas grafis, gaya artistik, kejernihan visual, musik), (3) Fitur Sosial & Komunitas (fungsi multipemain, alat komunikasi, dinamika kompetitif versus kolaboratif), (4) Sistem Perkembangan & Penghargaan (mekanika peningkatan level, struktur pencapaian, rasa keberhasilan), dan (5) Monetisasi & Persepsi Nilai (model penetapan harga, pembelian dalam game, kekhawatiran bayar-untuk-menang, nilai yang dipersepsikan atas biaya yang dikeluarkan)