BAB II
LANDASAN TEORI

2.1 Penelitian Terkait

Penelitian terkait dilakukan untuk meninjau berbagai studi sebelumnya yang membahas analisis sentimen menggunakan algoritma *Naive Bayes*. Kajian ini bertujuan untuk memahami pendekatan yang telah digunakan, tahapan pengolahan data, serta performa model yang diperoleh dalam penelitian terdahulu.

Penelitian yang dilakukan oleh (Rahmatullah *et al*., 2025) menyebutkan bahwa pemilihan algoritma *Naive Bayes* dikarenakan kesederhanaannya efisiensinya dalam memproses data besar serta kemampuannya dalam memberikan hasil yang cukup akurat dalam klasifikasi teks. Penelitian ini menghasilkan akurasi yang cukup tinggi sebesar 87,35%. Hasil tersebut menunjukkan bahwa *Naive Bayes* memiliki performa klasifikasi yang cukup tinggi pada data sentimen media sosial.

Penelitian yang dilakukan oleh (Fatah *et al*., n.d.) menerapkan algoritma *Naive Bayes* guna mengklasifikasikan pandangan publik terhadap kebijakan penyitaan tanah oleh pemerintah. Penelitian tersebut melalui proses prapemrosesan dan pembobotan fitur berbasis TF-IDF. Analisis yang dilakukan menghasilkan tingkat akurasi yang cukup tinggi sebesar 90%. Hasil *precision* dan *recall* yang cukup tinggi mengindikasikan bahwa model dapat mengenali sentimen dengan baik. Berdasarkan hasil tersebut menunjukkan bahwa *Naive Bayes* memiliki performa klasifikasi yang cukup tinggi.

Penelitian yang dilakukan oleh (Lestari *et al*., 2026) menerapkan algoritma *Naive Bayes* guna menganalisis sentimen mengenai keterlibatan anak dalam

aktivitas militer terhadap komentar YouTube. Penelitian ini dilakukan secara mendalam meliputi *cleaning*, *tokenizing*, *stopword removal* dan *stemming*. Penelitian yang dilakukan menghasilkan tingkat akurasi yang cukup tinggi sebesar 84,26%. Penelitian ini menunjukkan bahwa penerapan *Naive Bayes* menghasilkan capaian akurasi yang cukup memadai dalam proses klasifikasi sentimen.

Penelitian yang dilakukan oleh (Aprianti *et al*., 2025) menerapkan algoritma *Naive Bayes* dalam mengklasifikasikan pandangan publik terhadap Program Makan Bergizi Gratis berdasarkan sentimen komentar dari platform X. Pembagian data dengan rasio 80% untuk data latih dan 20% untuk data uji dipilih karena pada pengujian sebelumnya terbukti menghasilkan performa yang baik dibandingkan skema lain. Penelitian ini juga menghasilkan tingkat akurasi yang cukup tinggi yaitu diatas 86%. Hasil ini mengindikasikan bahwa porsi data latih yang besar (80%) membuat model lebih mampu mengenali pola data, sehingga performanya menjadi optimal. Kinerja yang seimbang antara *precision, recall* dan *F1-score* juga menunjukkan bahwa model ini tidak hanya tepat dalam memprediksi, tetapi juga mampu menangkap sebagian besar data yang relevan secara akurat.

Penelitian yang dilakukan oleh (Pratama *et al*., 2023) juga menerapkan algoritma *Naive Bayes* dalam mengklasifikasi analisis sentimen Chat GPT sebagai masa depan pekerja pada platform YouTube. Hasil penelitian menunjukkan bahwa model menghasilkan akurasi tinggi sebesar 95,88%.

Penelitian yang dilakukan oleh (Kartika *et al*., 2023) menerapakan pelabelan IndoBert *(pseudo-labelling)* dan teknik *Easy Data Augmentation* (EDA)

dengan strategi *Random Word Swap* untuk mengatasi ketidakseimbangan distribusi kelas berhasil mencapai 89,7% dengan kombinasi metode ini.

2.2 Text Mining

Text mining merupakan proses pengolahan data teks yang bersifat tidak terstruktur menjadi informasi yang dapat dianalisis secara sistematis. Data teks seperti komentar media sosial, ulasan pengguna atau opini publik umumnya tidak memiliki format yang terstruktur sehingga memerlukan tahapan pengolahan khusus sebelum dapat digunakan dalam proses analisis.

Dalam konteks penelitian analisis sentimen berbasis media sosial, *text mining* umumnya melibatkan beberapa tahapan utama, yaitu prapemrosesan, ekstraksi fitur dan klasifikasi (Rahmatullah *et al*., 2025). Tahapan tersebut bertujuan untuk meningkatkan kualitas data serta mentransformasikan teks menjadi vektor numerik yang selanjutnya digunakan dalam tahap klasifikasi dengan *Naive Bayes*.

Sebagaimana diterapkan dalam penelitian (Aprianti *et al*., 2025), *text mining* menjadi langkah penting dalam mempersiapkan data agar siap digunakan dalam proses klasifikasi menggunakan algoritma *Naive Bayes*.

2.3 Analisis Sentimen

Analisis sentimen merupakan teknik dalam *text mining* yang digunakan untuk mengkaji dan mengklasifikasikan opini ke dalam kategori tertentu, seperti positif, negatif dan netral. Melalui pendekatan ini, kecenderungan pandangan masyarakat terhadap suatu isu dapat diidentifikasi secara lebih sistematis dan terukur (Fatah *et al*., n.d.; Rahmatullah *et al*., 2025).

Analisis sentimen atau *sentiment analysis*, yang juga dikenal sebagai *opinion mining*, merupakan salah satu bidang penelitian dalam *Natural Language Processing* (NLP) yang berfokus pada proses identifikasi, pengolahan, dan analisis opini, emosi, penilaian, serta sikap seseorang terhadap suatu entitas tertentu. Entitas yang dimaksud dapat berupa produk, layanan, organisasi, individu, peristiwa, isu sosial, maupun atribut tertentu yang berkaitan dengan objek pembahasan. Tujuan utama dari analisis sentimen adalah untuk memperoleh informasi subjektif dari data berbentuk teks sehingga dapat diketahui kecenderungan opini yang terkandung di dalamnya, apakah bersifat positif, negatif atau netral (Agustina & Herliana, 2025).

Dalam konteks media sosial, analisis sentimen dipandang sebagai metode relevan karena pengguna secara aktif menyampaikan pendapat, kritik dan dukungan terhadap berbagai kebijakan secara terbuka. Data tersebut dapat dimanfaatkan untuk menggambarkan opini publik secara sistematis berdasarkan kategori sentimen yang telah ditentukan.

Dalam perkembangannya, analisis sentimen menjadi salah satu metode yang sangat penting di era digital karena meningkatnya jumlah data teks yang dihasilkan melalui media sosial, forum diskusi, situs ulasan, blog, dan platform daring lainnya. Melalui analisis sentimen, perusahaan maupun organisasi dapat memahami persepsi masyarakat terhadap produk atau layanan yang mereka tawarkan. Selain itu, metode ini juga sering dimanfaatkan dalam berbagai bidang, seperti pemasaran, politik, pendidikan, kesehatan, hingga analisis opini publik terhadap suatu kebijakan atau peristiwa tertentu (Lutfina *et al*., 2024).

Secara umum, analisis sentimen dapat dibedakan berdasarkan tingkat analisisnya menjadi tiga kategori utama, yaitu tingkat dokumen *(document – level)*, tingkat kalimat *(sentence – level)* dan tingkat aspek *(aspect – level)*. Pada tingkat dokumen, keseluruhan isi dokumen dianalisis untuk menentukan apakah dokumen tersebut memiliki sentimen positif, negatif atau netral secara keseluruhan. Pendekatan ini biasanya digunakan ketika sebuah dokumen hanya membahas satu topik utama sehingga klasifikasi sentimen dapat dilakukan secara menyeluruh terhadap isi teks (Sarang Narendra Shirole, 2023).

Dalam implementasinya, analisis sentimen memanfaatkan berbagai teknik, mulai dari metode berbasis kamus *(lexicon – based approach)* hingga metode pembelajaran mesin *(machine learning)* dan *deep learning*. Berbagai algoritma seperti *Naive Bayes*, *Support Vector Machine* (SVM), dan jaringan saraf tiruan sering digunakan untuk meningkatkan akurasi klasifikasi sentimen. Dengan perkembangan teknologi kecerdasan buatan, analisis sentimen kini mampu memberikan hasil yang lebih akurat dan efisien dalam memahami opini masyarakat dari data teks dalam jumlah besar (Ratnaswari *et al*., 2025).

2.4 Text Preprocessing

Text preprocessing merupakan tahap awal dalam pengolahan data teks yang bertujuan untuk membersihkan dan menyiapkan data sebelum dilakukan proses klasifikasi. Data dari media sosial umumnya mengandung *noise* seperti URL, simbol, tanda baca, singkatan, serta variasi penulisan kata yang dapat memengaruhi hasil analisis.

Tahapan *preprocessing* penelitian ini mencakup proses *cleaning*, *case*

folding, *tokenizing*, *normalizing, stopword removal* dan *stemming*. Tahapan ini sejalan dengan penelitia yang dilakukan oleh (Fatah *et al*., n.d.; Rahmatullah *et al*., 2025) sebelum dilakukan ekstraksi fitur. Proses tersebut berperan dalam membersihkan data dari elemen yang mengganggu guna mendukung performa model klasifikasi.

2.5 Term Frequency-Inverse Document Frequency (TF-IDF)

TF-IDF merupakan metode pembobotan kata yang digunakan guna mengonversi teks menjadi representasi numerik. Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen serta tingkat keunikannya dalam keseluruhan kumpulan dokumen.

Pengunaan TF-IDF sebelum proses klasifikasi menggunakan *Naive Bayes* diterapkan dalam penelitian (Fatah *et al*., n.d.). Pembobotan fitur memungkinkan model menilai tingkat kepentingan setiap kata sehingga proses penentuan kategori sentimen menjadi lebih baik dan akurat.

2.6 Algoritma Naive Bayes

Naive Bayes merupakan algoritma pembelajaran mesin yang digunakan untuk melakukan klasifikasi berbasis probabilitas. Algoritma ini bekerja dengan menggunakan konsep Teorema Bayes dengan asumsi bahwa setiap fitur (kata) dalam teks dianggap independen satu sama lain *(conditional independence assumption)*. Meskipun asumsi ini jarang terpenuhi sepenuhnya pada data teks nyata, *Naive Bayes* tetap menunjukkan performa yang kompetitif dalam tugas klasifikasi teks, termasuk analisis sentimen. Algoritma ini menghitung probabilitas

posterior setiap kelas sentimen berdasarkan distribusi kata yang muncul dalam teks, kemudian memilih kelas dengan probabilitas tertinggi sebagai hasil prediksi.

Rumus dasar *Naive Bayes* adalah :

$$P(C \mid X) = \frac{P(X \mid C) \times P(C)}{P(X)}$$

Penjelasan :

$P(C \mid X)$ adalah probabilitas suatu data termasuk dalam kategori tertentu, $P(X \mid C)$ adalah probabilitas munculnya data pada kategori tersebut, $P(C)$ adalah probabilitas awal dari kategori, $P(X)$ adalah probabilitas keseluruhan data.

Keunggulan algoritma *Naive Bayes* meliputi proses komputasi yang sederhana dan cepat, kemampuan untuk bekerja dengan baik pada data berdimensi tinggi seperti matriks TF-IDF, serta kebutuhan data pelatihan yang relatif kecil dibandingkan algoritma yang lebih kompleks. Berbagai penelitian terdahulu telah membuktikan kemampuan *Naive Bayes* dalam mencapai tingkat akurasi yang tinggi pada analisis sentimen media sosial berbahasa Indonesia (Fatah *et al*., n.d.; Rahmatullah *et al*., 2025; Wahyuni *et al*., 2025). Varian *Multinomial Naive Bayes* digunakan dalam penelitian ini karena lebih sesuai untuk data teks yang direpresentasikan dalam bentuk frekuensi kata atau bobot TF-IDF.

2.7 Pseudo-labelling dengan IndoBERT

Pseudo-labelling merupakan teknik *semi-supervised learning* di mana model yang telah dilatih sebelumnya *(pre-trained model)* digunakan untuk memberikan label secara otomatis pada data yang belum berlabel. Dalam konteks penelitian ini, model IndoBERT (Indonesian RoBERTa) yang tersedia di Hugging Face dengan

nama w11wo/Indonesian-Roberta-base-sentiment-classifier digunakan sebagai pelabel otomatis. Model ini merupakan varian arsitektur RoBERTa *(Robustly Optimized BERT Pretraining Approach)* yang telah di-fine-tuning secara khusus untuk tugas klasifikasi sentimen Bahasa Indonesia, dengan kemampuan mengenali tiga kelas output : positif, negatif dan netral.

IndoBERT menggunakan mekanisme *attention* berbasis arsitektur transformer yang memungkinkannya memahami konteks kalimat secara lebih mendalam dibandingkan pendekatan berbasis leksikon atau aturan. Penggunaan IndoBERT sebagai *pseudo-labeler* memberikan beberapa keuntungan, antara lain : mengurangi subjektivitas yang muncul pada pelabelan manual, meningkatkan konsistensi penentuan label, serta mempercepat proses pelabelan pada dataset berukuran besar. Model ini juga mampu menangani variasi bahasa informal yang umum ditemukan pada komentar media sosial, menjadikannya pilihan yang tepat untuk konteks penelitian ini (Aprilah *et al*., 2025; Khomsah *et al*., 2023).

2.8 Easy Data Augmentation (EDA)

Easy Data Augmentation (EDA) merupakan teknik augmentasi teks yang diperkenalkan oleh (Wei & Zou, 2019) untuk meningkatkan performa model klasifikasi teks, khususnya pada kondisi keterbatasan data. Teknik ini bekerja dengan cara menciptakan variasi baru dari data teks yang sudah ada melalui empat operasi sederhana, yaitu : *Synonym Replacement* (mengganti kata dengan sinonimnya), *Random Insertion* (memasukkan kata acak ke dalam kalimat), *Random Swap* (menukar posisi dua kata secara acak) dan *Random Deletion*

(menghapus kata secara acak). Keempat operasi ini dirancang agar tidak mengubah makna semantik dari teks asli secara signifikan.

Dalam penelitian ini, diterapkan strategi *Random Swap* dengan parameter aug_p = 0,2 yang berarti 20% kata dalam kalimat dipertukangkan posisinya secara acak. Teknik ini dipilih karena relatif sederhana namun efektif dalam menciptakan variasi data yang cukup berbeda dari teks asli. EDA digunakan untuk mengatasi permasalahan *class imbalance*, yaitu kondisi di mana distribusi data antar kelas sentimen tidak merata, dengan cara menambahkan data sintetis pada kelas minoritas (netral dan positif) hingga jumlahnya seimbang dengan kelas mayoritas (negatif). Pendekatan ini terbukti mampu meningkatkan generalisasi model tanpa memerlukan pengumpulan data tambahan dari lapangan (Nur Azizah *et al*., 2023).

2.9 Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan untuk mengukur tingkat keberhasilan algoritma dalam mengklasifikasikan data sentimen. Metrik evaluasi yang umum digunakan meliputi *accuracy*, *precision*, *recall* dan *F1-Score* (Fatah *et al*., n.d.; Rahmatullah *et al*., 2025; Wahyuni *et al*., 2025). *Accuracy* mengukur proporsi prediksi yang benar dari keseluruhan data uji. *Precision* mengukur ketepatan model dalam memprediksi kelas tertentu, yaitu rasio data yang benar-benar termasuk kelas tersebut dari seluruh data yang diprediksi sebagai kelas tersebut. *Recall* mengukur kelengkapan model dalam menemukan seluruh data yang seharusnya termasuk ke dalam kelas tertentu. *F1-Score* merupakan *harmonic mean* dari *precision* dan *recall* yang memberikan gambaran keseimbangan antara kedua metrik tersebut. Penggunaan metrik tersebut memungkinkan penilaian performa model secara

menyeluruh serta membandingkan hasil pada setiap skenario klasifikasi yang diterapkan dalam penelitian ini. Melalui evaluasi tersebut, dapat diketahui sejauh mana algoritma *Naive Bayes* mampu mengelompokkan data sentimen secara efektif dan konsisten.

2.10 Platform YouTube sebagai Sumber Data

YouTube merupakan platform media sosial berbasis video yang memungkinkan pengguna untuk berinteraksi secara langsung melalui fitur komentar. Penelitian ini memanfaatkan YouTube sebagai platform media sosial dalam analisis sentimen karena menyediakan data teks dalam jumlah besar yang dapat diolah menggunakan pendekatan *text mining*. (Fatah *et al*., n.d.; Rahmatullah *et al*., 2025; Wahyuni *et al*., 2025)