

BAB II

TINJAUAN PUSTAKA

2.1 Studi Sebelumnya

Menurut penelitian terdahulu ada beberapa penelitian serupa, yang mana para peneliti tersebut melakukan evaluasi terhadap Analisis Sentimen pada aplikasi dengan metode yang sama. Berikut lima studi yang berkaitan:

Penelitian yang dilakukan oleh Idris, Mustofa, dan Salihi (2023) menganalisis sentimen ulasan pengguna aplikasi *Shopee* menggunakan algoritma *Support Vector Machine* (SVM). Data ulasan dikumpulkan melalui *scraping* dan dilabeli berdasarkan *rating*, kemudian diproses melalui tahapan normalisasi, *case folding*, penghapusan *stopword*, dan *cleansing* sebelum diklasifikasikan. Hasil penelitian menunjukkan bahwa SVM mampu mengklasifikasikan sentimen positif dan negatif dengan sangat baik, dibuktikan dengan akurasi mencapai 98% dan F1-score sebesar 0,98. Temuan ini memperkuat bukti bahwa SVM merupakan salah satu algoritma yang unggul dalam tugas klasifikasi sentimen pada data teks, khususnya dalam konteks ulasan aplikasi *e-commerce* di Indonesia [4].

Berdasarkan penelitian yang dilakukan oleh Wahyudi dan Kusumawardhana (2021) dalam jurnal "Analisis Sentimen pada *Review* Aplikasi Grab di Google Play Store Menggunakan *Support Vector Machine*", analisis sentimen terhadap ulasan pengguna aplikasi Grab di Google Play Store dilakukan dengan metode *Support Vector Machine* (SVM) untuk mengklasifikasikan *review* menjadi sentimen positif dan negatif. Data yang digunakan berasal dari lebih dari 1.000 *review* pengguna yang dikumpulkan melalui teknik web *scraping*, kemudian diproses melalui tahapan *preprocessing* seperti *case folding*, *tokenizing*, *filtering*, *spelling correction*, dan pembentukan *n-gram*. Selanjutnya, data dibagi menjadi data *training* dan testing dengan perbandingan 9:1, dan proses klasifikasi dilakukan menggunakan algoritma SVM yang diimplementasikan pada *RapidMiner*. Validasi model dilakukan dengan *k-fold cross validation* (fold 2–10), dan hasil terbaik menunjukkan akurasi sebesar

85,54%. Selain itu, penelitian ini menemukan bahwa kata "ovo" sering muncul pada *review* positif, sedangkan kata "driver" dominan pada *review* negatif. Hasil penelitian ini menunjukkan bahwa SVM efektif dalam mengklasifikasikan sentimen ulasan aplikasi dan dapat menjadi acuan bagi pengembang aplikasi untuk meningkatkan kualitas layanan berdasarkan opini pengguna[5].

Berdasarkan penelitian yang dilakukan oleh Fitriyana et al. (2023) dalam jurnal "Analisis Sentimen Ulasan Aplikasi Jamsostek Mobile Menggunakan Metode *Support Vector Machine*", analisis sentimen diterapkan pada ulasan pengguna aplikasi Jamsostek Mobile (JMO) untuk mengklasifikasikan opini ke dalam kategori positif, negatif, dan netral menggunakan metode *Support Vector Machine* (SVM) dengan kernel linear. Penelitian ini menggunakan tahapan *text mining* seperti *scrapping* data ulasan dari Google Play Store, *pre-processing* (*case folding*, *tokenizing*, *filtering*, *stemming*), pembobotan kata dengan TF-IDF, serta evaluasi model menggunakan *confusion matrix* untuk memperoleh nilai akurasi, presisi, *recall*, dan f1-score. Hasil penelitian menunjukkan bahwa metode SVM dengan kernel linear mampu menghasilkan akurasi klasifikasi sebesar 96%, *precision* 92%, *recall* 96%, dan f1-score 94%, dengan mayoritas ulasan pengguna termasuk dalam kategori positif. Temuan ini menegaskan bahwa SVM sangat efektif dalam melakukan analisis sentimen pada data teks ulasan aplikasi, sehingga dapat menjadi alat evaluasi yang bermanfaat bagi pengembang aplikasi maupun calon pengguna dalam menilai kualitas layanan JMO[6].

Berdasarkan penelitian yang dilakukan oleh Apriani dan Gustian (2019) dalam jurnal "Analisis Sentimen dengan *Naïve Bayes* terhadap Komentar Aplikasi Tokopedia", analisis sentimen dilakukan terhadap komentar pengguna aplikasi Tokopedia di Google Play Store dengan menggunakan metode *Naïve Bayes* untuk mengklasifikasikan komentar ke dalam kategori positif dan negatif. Penelitian ini melibatkan proses pengumpulan data melalui web *scraping*, tahapan *preprocessing* seperti *case folding*, pembersihan karakter, dan *filtering*, serta penggunaan *Rapidminer* untuk proses pelatihan dan pengujian model. Hasil penelitian menunjukkan bahwa metode *Naïve Bayes* mampu memberikan performa yang

sangat baik dengan tingkat akurasi sebesar 97,13%, *precision* 1, *recall* 95,49% (untuk kelas negatif), dan nilai AUC 0,980. Dari 1.500 komentar yang dianalisis, ditemukan bahwa 63% komentar bersentimen negatif dan 36% bersentimen positif, yang menunjukkan perlunya peningkatan kepercayaan dan kualitas layanan pada aplikasi Tokopedia. Temuan ini menegaskan bahwa metode *Naive Bayes* efektif dalam menganalisis sentimen pada ulasan aplikasi dan dapat menjadi masukan penting bagi pengembang untuk meningkatkan mutu layanan dan kepuasan pengguna[7].

Berdasarkan penelitian yang dilakukan oleh Jeremi Azero Putra, Alexander Dharmawan, dan Jutono Gondohanindijo (2024) dalam jurnal “Sentimen Analisis Aplikasi *Digitalent Mobile* Menggunakan *Naïve Bayes* dan SVM dengan Ekstraksi Fitur TF-IDF”, analisis sentimen dilakukan terhadap ulasan pengguna aplikasi *Digitalent Mobile* yang dikumpulkan dari Google Play Store melalui teknik *scraping*. Penelitian ini menerapkan tahapan *preprocessing* seperti *cleaning*, *case folding*, *stopword removal*, *tokenization*, dan *stemming*, serta menggunakan metode TF-IDF untuk ekstraksi fitur sebelum proses klasifikasi sentimen dilakukan dengan algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)*. Evaluasi model dilakukan menggunakan metrik akurasi, presisi, *recall*, dan F1-score, dengan hasil menunjukkan bahwa *Naïve Bayes* mampu memberikan performa akurasi sebesar 85,71% dan SVM sebesar 82,14%. Temuan ini menegaskan bahwa pemilihan algoritma dan teknik ekstraksi fitur sangat berpengaruh dalam analisis sentimen pada ulasan aplikasi, serta dapat memberikan informasi penting bagi pengembang untuk meningkatkan kualitas layanan dan kepuasan pengguna berdasarkan opini yang teridentifikasi [8].

Setiap tabel dan gambar harus digunakan atau dirujuk. Contoh simbol-simbol dari *flowchart* yang biasa digunakan akan dijelaskan pada Tabel 2.1.

Tabel 2. 1 Studi Sebelumnya

No.	Peneliti	Tahun	Tujuan	Metode	Hasil
1.	Idris, Mustofa, dan Salihi	2023	Menganalisis sentimen ulasan pengguna aplikasi Shopee	SVM; data dari scraping; preprocessing: normalisasi, case folding, stopword removal, cleansing	Akurasi 98%, F1-score 0,98; SVM efektif untuk klasifikasi sentimen pada ulasan aplikasi e-commerce di Indonesia
x2.	Wahyudi dan Kusumawardhana	2021	Mengklasifikasikan review pengguna Grab menjadi sentimen positif dan negatif	SVM; >1.000 review dari scraping; preprocessing: case folding, tokenizing, filtering, spelling correction, n-gram; Rapid Miner	Akurasi 85,54%; “ovo” dominan pada review positif, “driver” pada review negatif; SVM efektif dalam klasifikasi ulasan aplikasi
3.	Fitriyana et al.	2023	Mengklasifikasikan ulasan Jamsostek Mobile ke	SVM dengan kernel linear; scraping; preprocessing: case	Akurasi 96%, Precision 92%, Recall 96%, F1-score 94%;

No.	Peneliti	Tahun	Tujuan	Metode	Hasil
			dalam sentimen positif, negatif, dan netral	folding, tokenizing, filtering, stemming; TF-IDF	mayoritas ulasan bersentimen positif
4.	Apriani dan Gustian	2019	Mengklasifikasikan komentar Tokopedia ke dalam kategori positif dan negatif	Naïve Bayes; scraping; preprocessing: case folding, pembersihan karakter, filtering; RapidMiner	Akurasi 97,13%, Precision 1, Recall 95,49% (kelas negatif), AUC 0,980; 63% komentar negatif, 36% positif; perlunya peningkatan kualitas layanan
5.	Jeremi Azero Putra, Alexander Dharmawan, Jutono Gondohanindijo	2024	Menganalisis sentimen ulasan aplikasi Digitalent Mobile	Naïve Bayes & SVM; scraping; preprocessing: cleaning, case folding, stopword removal, tokenization, stemming; TF-IDF	Naïve Bayes akurasi 85,71%, SVM akurasi 82,14%; performa dipengaruhi oleh pemilihan algoritma dan fitur; dapat bantu pengembang meningkatkan kualitas layanan aplikasi

2.2 Dasar Teori

1. Analisis Sentimen

Analisis sentimen adalah proses otomatis untuk memahami, mengekstrak, dan mengolah data teks guna mengidentifikasi sentimen yang terkandung dalam sebuah opini, baik berupa evaluasi, ekspresi emosional, maupun efek emosional yang ingin ditimbulkan terhadap pembaca. Tujuan utama analisis sentimen adalah menentukan pandangan atau sikap seseorang terhadap suatu topik serta memahami polaritas keseluruhan dalam sebuah dokumen, yang dapat berupa sentimen positif, negatif, atau netral[9].

Proses analisis sentimen melibatkan beberapa tahapan penting, mulai dari pengumpulan data ulasan atau komentar pengguna, *preprocessing* data yang mencakup *case folding*, pembersihan karakter, stopword removal, tokenisasi, dan *stemming*, hingga ekstraksi fitur menggunakan metode seperti TF-IDF. Term Frequency (TF) memiliki proporsi kepentingan sesuai total kemunculannya dalam teks atau dokumen. Inverse Document Frequency (IDF) merupakan metode pembobotan token yang berfungsi untuk memonitor kemunculan token dalam himpunan teks[9]. Setelah itu, data hasil ekstraksi fitur diklasifikasikan menggunakan algoritma pembelajaran mesin seperti *Naïve Bayes*, *Support Vector Machine* (SVM), atau model berbasis deep learning seperti BERT untuk memperoleh hasil klasifikasi sentimen yang akurat [10].

Evaluasi model klasifikasi sentimen dilakukan dengan menggunakan metrik seperti akurasi, presisi, recall, dan F1-score yang dihasilkan dari *confusion matrix*. Penggunaan aplikasi restoran menyoroti pentingnya teknologi informasi dalam meningkatkan efisiensi operasional dan mutu layanan di bidang kuliner. Sistem informasi manajemen restoran berbasis web maupun *mobile* dikembangkan untuk menyelesaikan berbagai tantangan operasional, seperti pencatatan pesanan secara manual yang rawan kesalahan, rekapitulasi data yang memakan waktu, serta pelayanan yang kurang optimal. Dengan bantuan aplikasi ini, pemesanan makanan dan minuman bisa dilakukan secara digital, di mana pesanan langsung tersambung ke dapur dan sistem kasir, serta laporan penjualan dapat dihasilkan secara otomatis.

Implementasi ini membantu mempercepat pelayanan, mengurangi antrean, dan meningkatkan kepuasan pelanggan[11].

2. Aplikasi Restoran

Aplikasi Restoran merupakan *platform* digital yang dirancang untuk memfasilitasi proses pemesanan, pembayaran, dan pelayanan di industri kuliner secara lebih cepat dan efisien. Aplikasi *mobile* pada restoran dapat dimanfaatkan oleh siapa saja untuk menyederhanakan dan mempercepat berbagai proses bisnis serta kegiatan sehari-hari, termasuk pemesanan makanan dan minuman di restoran. Restoran sebagai salah satu jenis usaha yang sangat bergantung pada kecepatan dan efisiensi pelayanan dapat memanfaatkan teknologi untuk meningkatkan performa operasionalnya. Dengan demikian, teknologi berperan penting dalam mengoptimalkan proses pemesanan di restoran agar lebih efisien dan efektif[12]. Salah satu contoh restoran yang memiliki aplikasi untuk proses penjualan merupakan McDonald's.

Aplikasi *McDonald's* merupakan salah satu bentuk penerapan teknologi digital dalam industri restoran cepat saji yang bertujuan untuk meningkatkan kualitas layanan dan pengalaman pelanggan. Melalui aplikasi ini, pengguna dapat melakukan pemesanan makanan dan minuman secara *online*, memilih metode pembayaran digital, serta menentukan opsi pengambilan pesanan seperti *dine-in* atau *take away*. Selain itu, aplikasi *McDonald's* juga menyediakan berbagai fitur pendukung, seperti promo dan penawaran khusus, program loyalitas, serta informasi menu yang lengkap. Kehadiran aplikasi ini tidak hanya memudahkan pelanggan dalam melakukan transaksi, tetapi juga membantu pihak restoran dalam meningkatkan efisiensi operasional, mempercepat proses pelayanan, dan mengelola data pelanggan secara lebih optimal. Dengan demikian, aplikasi *McDonald's* menjadi salah satu contoh pemanfaatan aplikasi restoran yang efektif dalam mendukung kegiatan bisnis dan kepuasan pelanggan.

3. Support Vector Machine (SVM)

Support Vector Machine (SVM) diperkenalkan oleh Vapnik pada tahun 1992 sebagai salah satu metode dalam *machine learning* yang didasarkan pada prinsip *Structural Risk Minimization* (SRM). SVM merupakan teknik pembelajaran mesin yang digunakan untuk mengklasifikasikan opini, ulasan, atau komentar ke dalam kategori sentimen seperti positif, negatif, maupun netral. Metode ini bekerja dengan cara memisahkan data teks berdasarkan fitur yang telah diekstraksi, misalnya melalui pendekatan TF-IDF, sehingga mampu membedakan kelas sentimen secara optimal. Tujuan utamanya adalah menemukan *hyperplane* terbaik yang dapat memisahkan dua kelas berbeda dalam ruang *input* secara efektif [13].

Dibandingkan dengan metode lain seperti *Naive Bayes*, SVM memiliki keunggulan tersendiri. *Support Vector Machine* (SVM) dan *Naive Bayes* adalah dua algoritma yang banyak digunakan dalam tugas klasifikasi sentimen, namun keduanya memiliki ciri khas serta kelebihan masing-masing. Secara umum, SVM memberikan hasil yang lebih baik dari segi akurasi dan kinerja, terutama saat digunakan pada data dengan dimensi yang tinggi serta distribusi kelas yang tidak seimbang, kondisi yang kerap ditemukan pada data ulasan aplikasi maupun media sosial[14]. Penelitian terbaru menunjukkan bahwa SVM mampu mencapai akurasi yang lebih tinggi, berkisar antara 79% hingga 89%, sementara *Naive Bayes* biasanya berada di rentang 61% hingga 72%[15]. Walaupun *Naive Bayes* memiliki keunggulan dalam kecepatan, kemudahan implementasi, serta penggunaan sumber daya komputasi yang lebih efisien, tingkat kompleksitas yang lebih tinggi pada SVM melalui proses pemilihan kernel dan penyesuaian parameter justru mampu menghasilkan model yang lebih akurat dan tangguh, sehingga sangat sesuai digunakan dalam analisis sentimen yang menuntut presisi tinggi[16].

4. Bahasa Pemrograman Python

Python adalah bahasa pemrograman tingkat tinggi yang mudah dipelajari, fleksibel, dan serbaguna. *Python* dapat diterapkan dalam berbagai bidang, seperti pengembangan web menggunakan *framework Django* atau *Flask*, analisis data dan

data science dengan library *Pandas* serta *NumPy*, hingga *machine learning* dan kecerdasan buatan menggunakan *TensorFlow* atau *Scikit-learn* [17].

5. SMOTE

Synthetic Minority Over-sampling Technique (SMOTE) merupakan metode *oversampling* yang menghasilkan sampel sintetis pada kelas minoritas. Teknik ini berpotensi memberikan kinerja yang lebih baik dibandingkan *oversampling* sederhana dan telah digunakan secara luas. Metode ini biasa digunakan untuk mengatasi permasalahan ketidakseimbangan kelas pada data klasifikasi. Ketidakseimbangan kelas terjadi ketika jumlah data pada kelas mayoritas jauh lebih banyak dibandingkan kelas minoritas, sehingga dapat menyebabkan model cenderung bias terhadap kelas mayoritas. SMOTE bekerja dengan cara menghasilkan data sintetis pada kelas minoritas berdasarkan kedekatan antar sampel, bukan dengan menduplikasi data yang sudah ada. Dengan demikian, distribusi data menjadi lebih seimbang dan model pembelajaran mesin dapat mempelajari karakteristik kelas minoritas secara lebih optimal. Penerapan SMOTE diharapkan mampu meningkatkan kinerja model klasifikasi, khususnya dalam hal akurasi dan kemampuan prediksi pada kelas minoritas[18].