

BAB I

PENDAHULUAN

1.1 Latar Belakang

Platform media sosial merupakan ruang penting bagi masyarakat untuk menyampaikan pendapat dan tanggapan terkait isu-isu terkini yang berkembang di Indonesia. Facebook, Instagram, dan YouTube sebagai *platform* dengan jumlah pengguna yang besar memungkinkan terjadinya interaksi dan diskusi publik secara cepat dan luas. Ketika isu nasional muncul, masyarakat cenderung langsung memberikan opini melalui komentar, unggahan, maupun video di media sosial. Fenomena ini menunjukkan bahwa Facebook, Instagram, dan YouTube dapat menjadi sumber informasi mengenai persepsi publik. Pendapat tersebut sering kali mencerminkan kepedulian masyarakat terhadap kebijakan pemerintah. Oleh sebab itu, pendapat publik di *platform* media sosial dapat dianalisis untuk memahami sikap masyarakat terhadap suatu kejadian. *Analisis* ini penting karena dapat memberikan informasi mengenai pandangan masyarakat yang dapat dijadikan dasar pertimbangan dalam proses pengambilan keputusan. Dengan demikian, media sosial menjadi *indikator dinamis* opini publik. (Pratiwi, 2024).

Korupsi merupakan permasalahan yang kerap terjadi di berbagai negara. Istilah kata “korupsi” diambil dari istilah Latin “*corruptio*,” yang diartikan “busuk,” rusak, atau goyah. Korupsi diartikan sebagai penyalahgunaan wewenang atau kekuasaan yang dimiliki seseorang dalam suatu jabatan untuk memperoleh keuntungan pribadi. Tindakan tersebut dilakukan dengan tujuan mendapatkan manfaat yang bukan menjadi hak atau kewajibannya, baik dalam bentuk uang maupun aset negara, sehingga dapat memperkaya diri sendiri. Praktik korupsi berpotensi melanda berbagai bidang, baik di perusahaan swasta ataupun lembaga pemerintah. Di Indonesia, tindak pidana korupsi diatur dalam Undang-Undang Nomor 31 Tahun 1999 tentang Pemberantasan Korupsi, sebagaimana telah diubah dengan Undang-Undang Nomor 20 Tahun 2001, yang

menjadi dasar hukum dalam penindakan terhadap pelaku korupsi, baik oleh individu maupun oleh organisasi. (Pebrianti & Karawang, 2025). Salah satu isu ramai diperbincangkan adalah dugaan kasus korupsi pengadaan Chromebook di Kemendikbudristek (Belarminus, 2025). Kasus ini muncul dalam konteks transformasi digital pada sektor pendidikan, yang seharusnya mendukung kegiatan belajar mengajar di sekolah. Namun, adanya dugaan penyimpangan anggaran membuat publik mempertanyakan integritas dan transparansi pemerintah dalam menjalankan program tersebut (Puspitasari, 2025). Berita terkait kasus ini dengan cepat menyebar dan memicu banyak respons negatif dari masyarakat, khususnya di Facebook, Instagram, dan YouTube. Masyarakat mengungkapkan rasa kecewa karena terjadinya dugaan korupsi pada sektor pendidikan yang sangat vital. Reaksi publik ini dapat menjadi sinyal bahwa kepercayaan masyarakat terhadap sistem pendidikan perlu dievaluasi. Maka dari itu, penting untuk mengkaji bagaimana opini publik terbentuk dan berkembang melalui komentar online. Hasil kajian tersebut dapat memberi gambaran tingkat kepercayaan publik terhadap pemerintah.

Analisis sentimen merupakan salah satu metode yang digunakan untuk mengidentifikasi atau mengumpulkan opini publik mengenai suatu isu tertentu, seperti layanan publik, isu yang sedang berkembang, kinerja pemerintahan, maupun topik lainnya (Suryono & Luthfi, n.d.). Analisis sentimen merupakan sebuah teknik yang bertujuan untuk mengenali sikap, opini, serta perasaan seseorang terkait dengan suatu produk, tokoh, isu, lembaga, ataupun layanan tertentu (Suryantari et al., 2025). Proses pelatihan model dalam analisis sentimen umumnya lebih kompleks dibandingkan dengan bidang *machine learning* lainnya karena data yang diolah berupa opini yang sifatnya subjektif serta tidak bisa dinilai secara langsung. (Sitanggang et al., 2024). Komentar yang muncul di Facebook, Instagram, dan Youtube dapat diolah sebagai data teks dan dianalisis pola emosinya terhadap kasus dugaan korupsi tersebut. Dengan memahami distribusi sentimen, peneliti dapat menilai seberapa besar dukungan atau kritik masyarakat terhadap pemerintah. Analisis sentimen juga dapat mengungkap isu-isu spesifik yang menjadi perhatian publik. Data mengenai pandangan masyarakat yang

diperoleh langsung dari sosial media dinilai lebih jujur karena disampaikan secara spontan. Oleh karena itu, analisis sentimen memiliki nilai penting dalam riset sosial modern. Hasilnya dapat menjadi referensi dalam menilai persepsi publik secara objektif (Pratiwi, 2024).

Dalam penelitian ini, metode Support Vector Machine (SVM) dan Multinomial Naive Bayes digunakan sebagai *algoritma* klasifikasi data komentar pengguna Facebook. SVM dikenal karena kinerjanya yang tinggi dalam mengelompokkan teks karena dapat memisahkan kelas data secara optimal. Sementara itu, Multinomial Naive Bayes merupakan metode klasifikasi berdasarkan pendekatan probabilitas sederhana tapi ampuh dalam memproses data teks. (Hartimar et al., 2025). Kedua *algoritma* ini sering digunakan dalam *Natural Language Processing* sebab mampu menghasilkan tingkat akurasi yang tinggi meskipun diterapkan pada dataset berukuran besar dan beragam. Melalui penerapan algoritma SVM dan MNB, hasil klasifikasi diperkirakan akan lebih akurat, *komprensif*, dan *representatif* melalui perbandingan *kinerja* kedua metode. Temuan dari penelitian ini diharapkan dapat menjadi bahan pertimbangan bagi pihak terkait dalam memperbaiki *regulasi* maupun tindakan *evaluatif*. Selain itu, hasil penelitian dapat memperkaya *kajian akademik* mengenai pengawasan publik berbasis media sosial. Pada akhirnya, penelitian ini diharapkan dapat berkontribusi dalam mendorong transparansi dan *akuntabilitas* pemerintah di sektor pendidikan. Masyarakat pun dapat lebih memahami bagaimana pendapat mereka turut memengaruhi kebijakan (Caroline et al., 2024).

Selain itu, berbagai studi terdahulu mengungkapkan bahwa algoritma Support Vector Machine (SVM) secara berulang kali menghasilkan performa paling unggul dalam menyelesaikan tugas klasifikasi sentimen berbasis teks, khususnya yang bersumber dari data media sosial. Misalnya, penelitian Hafika yang menganalisis sentimen masyarakat terhadap kebijakan efisiensi anggaran di Twitter. Hasil analisis menunjukkan bahwa SVM dengan kernel linier dapat mencapai akurasi sebesar 81,46%, yang lebih tinggi daripada akurasi kernel RBF sebesar 81,17%, yang

menandakan bahwa model linier menunjukkan kinerja yang lebih konsisten dalam mengklasifikasikan sentimen negatif dan netral (Hafika et al., 2025).

Penelitian Putri tentang sentimen publik terhadap kinerja Menteri Pendidikan Nadiem Makarim juga memperlihatkan bahwa SVM tetap unggul meskipun data yang digunakan tidak seimbang. Kombinasi TF-IDF dengan kernel linear menghasilkan F1-score 68%, dan tetap menjadi model paling stabil bahkan ketika dilakukan penyeimbangan kelas menggunakan SMOTE (Putri et al., 2025).

Hasil penelitian lainnya juga mendukung keunggulan SVM. Fatkhurrohman menunjukkan bahwa analisis sentimen terkait Program Makan Bergizi Gratis menghasilkan akurasi 80% serta weighted F1-score 0,76. model berhasil mengidentifikasi sentimen positif dengan sangat baik (recall 0,98), namun masih terdapat tantangan pada kategori netral dan negatif (Fatkhurrohman et al., 2025).

Hasil penelitian lainnya juga menerapkan *Term Frequency–Inverse Document Frequency* (TF-IDF) sebagai metode ekstraksi fitur untuk mengubah data teks ulasan menjadi representasi numerik. Selanjutnya, algoritma *Multinomial Naïve Bayes* digunakan untuk mengklasifikasikan sentimen ke dalam tiga kategori: positif, negatif, dan netral. Kumpulan data terdiri berdasarkan 1.000 ulasan aplikasi MyXL yang telah didapatkan dari *Kaggle*. Sebelum klasifikasi, data menjalani beberapa langkah prapemrosesan, termasuk pembersihan, *case folding*, normalisasi, tokenisasi, penghapusan *stopword*, dan *stemming*. Kumpulan data tersebut kemudian dibagi menjadi 80% data pelatihan dan 20% data uji. Kinerja model dievaluasi menggunakan matriks kebingungan (*confusion matrix*) serta metrik akurasi, presisi, *recall*, dan skor F1. Hasil evaluasi menunjukkan bahwa model mencapai akurasi sebesar 61,5%. Namun, model tersebut cenderung mengklasifikasikan lebih banyak data ke dalam kategori negatif karena jumlah titik data di kelas tersebut jauh lebih besar daripada di kategori lainnya. Akibatnya, kategori netral dan positif tidak dapat diidentifikasi secara optimal. Temuan ini menunjukkan bahwa meskipun algoritma *Naïve Bayes* cukup efektif untuk klasifikasi teks, kinerjanya dapat menurun ketika diterapkan pada *dataset* dengan distribusi kelas yang tidak seimbang (Hartimar et al., 2025).

Sementara itu, penelitian Rofiqi & Akbar mengenai sentimen terhadap RUU Perampasan Aset juga menunjukkan kinerja yang sangat baik dari SVM. Dengan kombinasi TF-IDF dan penyeimbangan data menggunakan SMOTE, model mencapai akurasi 79,8%, precision 80,58%, dan F1-score 79,96%, yang membuktikan stabilitas dan ketepatan SVM dalam memproses data komentar dari Twitter maupun YouTube (Rofiqi & Akbar, 2024).

Berdasarkan berbagai penelitian terdahulu, *Support Vector Machine* (SVM) telah terbukti memiliki performa yang unggul dalam analisis sentimen, terutama pada data teks yang kompleks, berdimensi tinggi, serta memiliki distribusi kelas yang tidak seimbang. Keunggulan utama SVM terletak pada kemampuannya membentuk *hyperplane* optimal yang berfungsi sebagai garis pemisah antar kelas, sehingga proses klasifikasi yang dihasilkan menjadi lebih konsisten dan akurat. Jika dibandingkan dengan metode lain, seperti *Multinomial Naïve Bayes* yang mengandalkan asumsi kebebasan antarfitur sehingga proses perhitungannya lebih sederhana, maupun *K-Nearest Neighbor* (KNN) yang cenderung kurang efisien ketika diterapkan pada data teks dalam jumlah besar, SVM umumnya memberikan kinerja yang lebih baik. Meskipun demikian, *Multinomial Naïve Bayes* tetap banyak digunakan karena merupakan metode klasifikasi probabilistik yang ringkas, cepat, serta efektif dalam mengolah data teks, oleh karena itu, sering digunakan sebagai acuan dalam penelitian analisis sentimen. Mengacu pada penelitian terdahulu yang telah menunjukkan tingkat akurasi tinggi serta kinerja yang konsisten, SVM dipilih sebagai metode utama pada penelitian ini, sedangkan *Multinomial Naïve Bayes* digunakan sebagai metode sebagai acuan perbandingan untuk mengevaluasi perbedaan kinerja klasifikasi. Penerapan kedua metode tersebut diharapkan mampu memberikan klasifikasi sentimen yang lebih akurat dalam menggambarkan tanggapan masyarakat terhadap dugaan korupsi pengadaan Chromebook berdasarkan komentar yang diperoleh dari Facebook, Instagram, dan YouTube.

1.2 Perumusan Masalah

Dari gambaran pada latar belakang diatas, dapat dirumuskan beberapa masalah diantaranya:

1. Seperti apa *persepsi* dan sentimen pengguna Facebook, Instagram, dan Youtube terhadap kasus dugaan korupsi pengadaan Chromebook di Kemendikbudristek?
2. Bagaimana proses pengolahan dan klasifikasi komentar pengguna Facebook, Instagram, dan Youtube dalam menerapkan metode *Support Vector Machine* (SVM) dan metode *Multinomial Naïve Bayes*?
3. Seberapa tinggi akurasi metode SVM dan *Multinomial Naïve Bayes* untuk mengklasifikasikan sentimen komentar pengguna Facebook, Instagram, dan Youtube terkait kasus dugaan korupsi chromebook tersebut?

1.3 Batasan Masalah

Dari rumusan masalah yang telah ditetapkan, berikut ini adalah batasan masalah untuk analisis sentimen pengguna facebook, Instagram, dan youtube terhadap kasus dugaan korupsi chromebook adalah:

1. Data yang digunakan dalam penelitian ini terdiri dari komentar pengguna Facebook, Instagram, dan Youtube terkait berita atau unggahan tentang dengan kata kunci dugaan korupsi Chromebook di Kemendikbudristek.
2. Sentimen yang dianalisis dibatasi pada tiga kategori, yaitu sentimen positif, netral, dan negatif.
3. Algoritma klasifikasi yang diterapkan dalam penelitian ini adalah *Support Vector Machine* (SVM) dan *Multinomial Naïve Bayes*.
4. Penelitian hanya fokus pada analisis sentimen, tidak membahas lebih lanjut mengenai penyelidikan hukum terkait kasus dugaan korupsi chromebook di Kemendikbudristek.

1.4 Tujuan dan Manfaat

Tujuan dan manfaat dari analisis pengguna facebook terhadap kasus dugaan korupsi chromebook adalah sebagai berikut:

a. Tujuan

1. Mengetahui persepsi dan kecenderungan sentimen publik di Facebook, Instagram, dan YouTube terhadap kasus dugaan korupsi Chromebook di Kemendikbudristek.
2. Mengolah dan mengklasifikasikan komentar pengguna Facebook, Instagram, dan YouTube dengan menerapkan algoritma *Support Vector Machine* (SVM) dan *Multinomial Naïve Bayes*.
3. Mengukur dan membandingkan tingkat akurasi metode SVM dan Multinomial Naïve Bayes dalam menentukan kecenderungan sentimen dalam komentar pengguna media sosial pada kasus dugaan korupsi Chromebook di Kemendikbudristek.

b. Manfaat

1. Membantu mengetahui bagaimana respon masyarakat Facebook, Instagram, dan YouTube dalam menanggapi kasus dugaan korupsi Chromebook, sehingga pihak terkait dapat memahami apa yang sebenarnya dirasakan dan dipikirkan oleh publik.
2. Memberikan contoh bagaimana algoritma SVM dan *Multinomial Naïve Bayes* mampu digunakan untuk mengolah dan mengelompokkan ulasan secara otomatis, sehingga proses membaca dan menganalisis opini Masyarakat menjadi lebih cepat dan efisien untuk dilakukan.
3. Menunjukkan dan membandingkan tingkat akurasi SVM dan Multinomial Naïve Bayes dalam menganalisis sentimen, sehingga hasilnya dapat dijadikan dasar atau referensi bagi penelitian lain yang ingin menggunakan metode serupa pada isu-isu berbeda di berbagai platform sosial media.

1.5 Metode Penelitian

Penelitian ini menerapkan pendekatan kuantitatif untuk mengukur dan mengklasifikasikan kecenderungan sentimen publik terhadap dugaan korupsi Chromebook di Kemendikbudristek melalui komentar pengguna Facebook, Instagram, dan Youtube. Seluruh proses dilakukan secara sistematis, meliputi pengumpulan data, analisis, perancangan sistem klasifikasi, implementasi model, pengujian performa, hingga dokumentasi akhir. Penelitian ini menggunakan pendekatan kuantitatif sebab berfokus pada proses pengolahan data dalam jumlah besar yang dianalisis secara komputasional menggunakan metode pembelajaran mesin.

Data penelitian dikumpulkan mulai 30 Mei 2025 hingga 18 Februari 2026 melalui berbagai unggahan Facebook dari Kompas.com, OfficialiNews, Tribunmedan.com, Melihatindonesia.id, Sindonews, Tribunsolo, Kumparan.com, INDONESIAKUDAMAI1, e100ss, serta video reel terkait. Unggahan Instagram dari medantalkviral, kompascom, gerpolrtv, ntvnews.id, selampung, theogidn, radar_politik, bumiclay, kompastv, liputan6, maoinfoindonesia, bisniscom, inilah_com, cnbcindonesia, zhaid.mushafi, jaksapedia, fyifact, wartaekonomi, fuadbakh23. Unggahan Youtube dari METROTV, tvOneNews, Liputan6, SINDOnews, Kompas.com, HarianKompas, Tribunnews. Total data yang dikumpulkan berjumlah 6.991 komentar, yang kemudian dijadikan dasar dalam analisis sentimen pada penelitian ini.

1. Pengumpulan Data

Tahap awal penelitian ini yaitu mengumpulkan data. Data diperoleh dari extension Chrome Esuit untuk mengekstraksi komentar Facebook terkait kasus dugaan korupsi Chromebook di Kemendikbudristek, dan extension Chrome IGCommentsExport untuk mengekstraksi komentar Instagram terkait kasus dugaan korupsi Chromebook di Kemendikbudristek, dan extension Chrome YoutubeCommentExporter untuk mengekstraksi komentar Youtube terkait kasus dugaan korupsi Chromebook di Kemendikbudristek, namun dengan tetap menjaga privasi pengguna. Data sekunder berupa literatur ilmiah tentang

analisis sentimen, NLP, serta metode klasifikasi seperti Support Vector Machine (SVM) dan Naïve Bayes digunakan sebagai dasar teoretis agar penelitian sesuai dengan standar akademik dan perkembangan ilmu terbaru.

2. Analisis Data

Setelah data terkumpul, dilakukan analisis awal untuk memastikan kelayakannya sebelum memasuki tahap pemodelan. Proses ini meliputi pembersihan data dari komentar duplikat, *spam*, komentar yang mengandung *emoji*, serta komentar yang tidak relevan. Komentar kemudian diberi label sentiment positif, negatif, atau netral dengan bantuan model machine learning. Untuk menjaga *konsistensi* dan akurasi, hasil pelabelan diverifikasi menggunakan label pembanding. Tahapan ini menjaga agar data yang dijadikan data latih memiliki mutu dan keandalan yang sesuai

3. Perancangan Sistem Klasifikasi

Pada tahap perancangan sistem, dilakukan penyusunan alur analisis sentimen yang terstruktur. Teks terlebih dahulu diproses melalui normalisasi, tokenisasi, stemming, dan penghapusan stopwords agar siap diolah. Selanjutnya, fitur diekstraksi dengan menerapkan TF-IDF untuk mengonversi data teks menjadi bentuk numerik. Selanjutnya, model klasifikasi dirancang dengan menggunakan algoritma *Support Vector Machine* (SVM) dan *Multinomial Naïve Bayes*. *Support Vector Machine* (SVM) dipilih sebab mampu mengolah data teks berdimensi tinggi dengan efektif serta mampu menghasilkan kinerja klasifikasi yang konsisten, sedangkan Multinomial Naïve Bayes digunakan sebagai metode pembanding yang sederhana dan berbasis probabilitas untuk melihat perbandingan kinerja klasifikasi.

4. Implementasi Sistem

Pada tahap implementasi, rancangan sistem diterjemahkan ke dalam program Python menggunakan *library* seperti *Scikit-learn*, *Pandas*, *NLTK*, dan *Sastrawi*. Komentar yang telah dipra-proses diubah menjadi *vektor* TF-IDF dan digunakan sebagai data pelatihan bagi model SVM dan Multinomial Naïve Bayes. Pada tahap ini juga dilakukan penyesuaian hyperparameter pada model

SVM guna meningkatkan akurasi dan kesesuaian prediksi sentimen terhadap pola data, sementara pada Multinomial Naïve Bayes dilakukan pengaturan parameter seperti jenis model (misalnya Multinomial Naïve Bayes) agar performa klasifikasi menjadi lebih optimal.

5. Pengujian Sistem

Setelah model dibangun, pengujian dilaksanakan dengan menggunakan dataset uji yang telah dipisahkan sebelumnya dengan rasio 80:20. Kinerja model dievaluasi berdasarkan metrik akurasi, presisi, *recall*, serta *F1-score* untuk mengukur kemampuan model dalam melakukan klasifikasi secara keseluruhan. Penggunaan *confusion matrix* bertujuan untuk melihat penyebaran hasil prediksi di tiap kategori sentimen. Tahapan ini bertujuan untuk memastikan bahwa model tidak sekadar mempelajari pola pada data latih, melainkan juga mampu mengklasifikasikan komentar baru secara efektif.

6. Dokumentasi Penelitian

Tahap akhir penelitian adalah penyusunan dokumentasi yang merangkum seluruh proses, mulai dari pengumpulan data hingga analisis hasil pengujian. Dokumentasi ini berfungsi sebagai catatan akademik yang memastikan transparansi metode serta memudahkan peneliti lain untuk mereplikasi atau mengembangkan penelitian di masa mendatang.

1.6 Sistematika Penulisan

Sistematika penulisan merupakan urutan atau struktur penyusunan isi dalam suatu karya tulis ilmiah yang bertujuan untuk membantu penulis dalam Menyusun isi karya tulis secara terstruktur, logis, dan runtut sehingga setiap bagian dapat disajikan dengan jelas dan tidak ada informasi penting yang terlewat.

BAB I PENDAHULUAN

Bab ini menyajikan gambaran umum mengenai penelitian yang dilakukan. Pada bagian awal dijelaskan latar belakang penelitian yang menguraikan fenomena perkembangan media sosial, khususnya Facebook, Instagram, dan YouTube sebagai media bagi masyarakat dalam menyampaikan pendapat terhadap berbagai isu publik,

termasuk kasus dugaan korupsi Chromebook di Kemendikbudristek. Selain itu, dijelaskan pula pentingnya analisis sentimen untuk mengetahui persepsi masyarakat terhadap isu tersebut serta alasan pemilihan metode Support Vector Machine (SVM) dan Multinomial Naïve Bayes dalam penelitian.

Bab ini juga memaparkan rumusan masalah yang menjadi fokus penelitian, batasan masalah yang ditetapkan untuk memperjelas batasan penelitian, serta tujuan dan manfaat penelitian. Metode penelitian secara umum, serta sistematika penulisan yang menjadi panduan dalam penyusunan tugas akhir ini.

BAB II TINJUAN PUSTAKA

Bab ini memaparkan landasan teori yang menjadi dasar pelaksanaan penelitian. Selain itu, bab ini mengulas berbagai penelitian terdahulu yang berkaitan dengan topik penelitian sebagai bahan referensi dan pembandingan untuk memperjelas posisi penelitian yang dilakukan.

Selain mengulas penelitian terdahulu, bab ini juga menguraikan berbagai teori yang mendukung pelaksanaan penelitian. Teori yang dibahas mencakup konsep media sosial, data, analisis sentimen, *text mining*, metode *Support Vector Machine* (SVM), *Multinomial Naïve Bayes*, *Google Colaboratory*, bahasa pemrograman Python, TF-IDF, *flowchart*, dan *extension* Google Chrome yang dimanfaatkan dalam proses pengumpulan data.

Teori-teori yang telah diuraikan pada bab ini menjadi landasan dalam perancangan sistem, implementasi model, dan analisis hasil penelitian. Dengan begitu, penelitian ini memiliki dasar teoretis dan akademik yang kuat sebagai acuan dalam setiap tahapan pelaksanaannya.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode dan tahapan penelitian yang dilakukan secara sistematis. Tahapan penelitian dimulai dari penyusunan kerangka penelitian dan proses

pengumpulan data komentar pengguna Facebook, Instagram, dan YouTube terkait kasus dugaan korupsi Chromebook di Kemendikbudristek.

Bab ini juga menjelaskan teknik pengumpulan data menggunakan extension Chrome seperti E-Suit Comments Exporter, IGCommentsExport, dan YoutubeCommentExporter, serta penentuan sampel data yang digunakan dalam penelitian. Selanjutnya dijelaskan tahap pra-pemrosesan data yang meliputi proses cleaning, normalisasi, case folding, filtering, stemming, dan tokenizing guna membersihkan sekaligus mempersiapkan data sebelum diolah lebih lanjut.

Selain itu, bab ini juga menguraikan tahapan penentuan polaritas sentimen, proses pelabelan data, ekstraksi fitur menggunakan *Term Frequency–Inverse Document Frequency* (TF-IDF), pembangunan model klasifikasi dengan *Support Vector Machine* (SVM) dan *Multinomial Naïve Bayes*, serta penilaian kinerja model yang diukur melalui metrik akurasi, presisi, *recall*, *F1-score*, dan *confusion matrix*.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil implementasi sistem beserta pembahasan terhadap langkah-langkah penelitian yang sudah dilakukan. Pembahasan diawali dengan penjelasan mengenai perangkat penelitian yang dipakai, terdiri atas perangkat keras (*hardware*) serta perangkat lunak (*software*). Selanjutnya, disajikan hasil pengumpulan data komentar yang berasal dari Facebook, Instagram, dan YouTube, diikuti dengan pembahasan mengenai tahap pra pemrosesan, yang meliputi *cleaning*, *case folding*, normalisasi, *filtering*, *stemming*, dan *tokenizing*.

Bab ini juga menguraikan proses pemberian label pada data dilakukan secara otomatis menggunakan pendekatan *lexicon-based*, ekstraksi fitur teks dengan *Term Frequency–Inverse Document Frequency* (TF-IDF), serta pembagian *dataset* menjadi data latih (*training*) dan data uji (*testing*) dengan perbandingan rasio 80:20.

Tahap selanjutnya adalah membangun dan melatih model dengan SVM dan MNB. Evaluasi dan perbandingan kinerja kedua model dilakukan berdasarkan metrik akurasi,

presisi, *recall*, dan *F1-score* untuk mengetahui metode mana yang memberikan kinerja terbaik dalam melakukan klasifikasi sentiment.

Selain itu, bab ini menyajikan visualisasi data menggunakan *WordCloud* untuk setiap kategori sentimen, yakni positif, negatif, dan netral, serta *confusion matrix* yang digunakan untuk mengevaluasi tingkat ketepatan klasifikasi model. Pada bagian akhir, dilakukan pengujian menggunakan beberapa sampel komentar guna menilai kemampuan model dalam mengklasifikasikan sentimen pada komentar baru.

BAB V KESIMPULAN DAN SARAN

Bab ini menyajikan kesimpulan yang dirumuskan berdasarkan hasil penelitian serta pembahasan yang telah dilakukan. Kesimpulan tersebut disusun untuk menjawab rumusan masalah mengenai pandangan masyarakat terhadap kasus dugaan korupsi pengadaan Chromebook di Kemendikbudristek, sekaligus membandingkan kinerja metode *Support Vector Machine* (SVM) dan *Multinomial Naïve Bayes* dalam analisis sentimen.

Selain memuat kesimpulan, bab ini juga menyampaikan sejumlah saran yang diharapkan dapat menjadi referensi bagi penelitian berikutnya. Saran tersebut mencakup pengembangan metode yang digunakan, penambahan jumlah dataset, penerapan teknik penyeimbangan data, serta penggunaan algoritma lain guna meningkatkan performa analisis sentimen.