

Analisis Sentimen Kritik Dan Saran Peserta Pelatihan Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine

Ade Rudi Masa'id, Aldi Rosyid*

Teknologi Informasi, Magister Ilmu Komputer, Universitas Budi Luhur, Jakarta, Indonesia

Email: ¹aderudimasaid1@gmail.com, ^{2*}aldirosyid@gmail.com

Email Penulis Korespondensi: aldirosyid@gmail.com

Submitted 15-04-2026; Accepted 27-04-2026; Published 30-04-2026

Abstrak

Penelitian ini bertujuan untuk mengklasifikasikan sentimen kritik dan saran peserta pelatihan ke dalam kategori positif, negatif, dan netral menggunakan pendekatan machine learning, serta membandingkan kinerja algoritma Naïve Bayes dan Support Vector Machine (SVM). Selain itu, penelitian ini juga menganalisis pengaruh optimalisasi hyperparameter terhadap performa model analisis sentimen. Metode penelitian meliputi prapemrosesan teks, ekstraksi fitur menggunakan Term Frequency Inverse Document Frequency (TF-IDF), pemodelan klasifikasi, serta evaluasi performa menggunakan metrik akurasi, precision, recall, dan F1-score. Optimalisasi model dilakukan menggunakan metode Grid Search dan Random Search. Hasil penelitian menunjukkan bahwa dari 487 data komentar peserta pelatihan, distribusi sentimen didominasi oleh sentimen positif. Pengujian model menunjukkan bahwa algoritma Support Vector Machine secara konsisten menghasilkan akurasi yang lebih tinggi dibandingkan Naïve Bayes, dengan akurasi tertinggi sebesar 79,0%, sedangkan Naïve Bayes mencapai akurasi tertinggi sebesar 65,3%. Selain itu, penerapan optimalisasi hyperparameter terbukti meningkatkan performa kedua algoritma, khususnya pada algoritma Naïve Bayes. Implikasi dari penelitian ini dapat digunakan sebagai alat pendukung evaluasi awal untuk membantu memahami kecenderungan umum persepsi peserta berdasarkan data teks, tanpa menggantikan metode survei atau evaluasi manajemen pelatihan yang telah ada.

Kata Kunci: Analisis Sentimen; Machine Learning; Naïve Bayes; Support Vector Machine; Klasifikasi Teks

Abstract

This study aims to classify the sentiment of training participants' feedback into positive, negative, and neutral categories using a machine learning approach, as well as to compare the performance of the Naïve Bayes and Support Vector Machine (SVM) algorithms. In addition, this study examines the effect of hyperparameter optimization on the performance of sentiment analysis models. The research methodology includes text preprocessing, feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF), sentiment classification modeling, and performance evaluation using accuracy, precision, recall, and F1-score metrics. Model optimization is conducted through hyperparameter tuning using Grid Search and Random Search methods. The results show that, out of 487 participant feedback comments, the sentiment distribution is dominated by positive sentiment. Model evaluation indicates that the Support Vector Machine algorithm consistently achieves higher accuracy than Naïve Bayes, with the highest accuracy reaching 79.0%, while Naïve Bayes achieves a maximum accuracy of 65.3%. Furthermore, hyperparameter optimization is shown to improve the performance of both algorithms, particularly for Naïve Bayes. The findings of this study can serve as a tool to support preliminary evaluations, helping to understand general trends in participants' perceptions based on textual data, without replacing existing survey methods or training management evaluation processes.

Keywords: Sentiment Analysis; Machine Learning; Naïve Bayes; Support Vector Machine; Text Classification

1. PENDAHULUAN

Pelaksanaan evaluasi dalam program pelatihan merupakan aspek penting untuk menjamin kualitas dan efektivitas proses pembelajaran, di mana umpan balik peserta berfungsi sebagai sumber data utama bagi penyelenggara. Berbagai kerangka evaluasi pelatihan menekankan bahwa reaksi peserta terhadap pelatihan harus diukur secara sistematis karena dapat menjadi dasar bagi perbaikan desain, metode, maupun penyampaian materi [1]. Dalam konteks ini, PT Brainmatics Indonesia Cendekia sebagai penyelenggara pelatihan di bidang teknologi informasi telah menyelenggarakan lebih dari 2.000 sesi pelatihan bagi instansi pemerintah, swasta, maupun individu, dan setiap pelaksanaan selalu disertai evaluasi melalui kuesioner yang mencakup penilaian kuantitatif serta komentar terbuka berupa kritik dan saran dari peserta. Namun, realitas di lapangan menunjukkan bahwa data kualitatif berupa komentar peserta tersebut masih jarang dimanfaatkan secara optimal karena sifatnya yang tak terstruktur (unstructured data) dan sulit dianalisis secara manual, padahal potensi informasinya sangat besar untuk memberikan wawasan lebih mendalam mengenai pengalaman belajar peserta [2].

Sebagai upaya menemukan informasi penting dari data yang tidak terstruktur seperti komentar terbuka, analisis sentimen dapat menjadi pendekatan yang terstandar dalam mengekstraksi polaritas (positif/negatif/netral) dan sikap dari teks, sehingga institusi dapat memetakan persepsi peserta secara agregat maupun spesifik topik. Bidang ini telah matang sebagai salah satu area riset paling aktif dalam NLP dan text mining, dengan fondasi teoretis dan praktik yang kuat di berbagai domain, termasuk pendidikan dan layanan [3]. Literatur menunjukkan bahwa analisis sentimen efektif untuk memproses umpan balik mahasiswa/peserta belajar guna memetakan apresiasi, masalah, dan emosi yang muncul selama proses pembelajaran. Di antara algoritma pembelajaran mesin untuk klasifikasi teks, Naïve Bayes (NB) dan Support Vector Machine (SVM) menempati posisi klasik karena kesederhanaan, efisiensi, dan kinerja yang kompetitif. NB berangkat dari asumsi independensi bersyarat antar fitur dan telah lama digunakan untuk pemodelan dokumen dengan varian event model yang terdokumentasi baik dalam literatur awal text classification. Sementara itu, SVM

memaksimalkan margin pemisah pada ruang fitur berdimensi tinggi, sebuah sifat yang secara empiris sangat sesuai untuk representasi sparse pada teks [4].

Bertolak dari kebutuhan institusi pelatihan untuk memperoleh bukti berbasis data dari komentar peserta, penelitian ini diarahkan untuk menjawab dua permasalahan utama. Pertama, bagaimana sentimen kritik dan saran peserta pelatihan di PT Brainmatics Indonesia Cendekia, apakah cenderung positif, negatif, atau netral, sehingga dapat memberikan gambaran umum mengenai persepsi peserta terhadap kualitas pelatihan yang telah dilaksanakan. Kedua, bagaimana kinerja algoritma Naïve Bayes dan Support Vector Machine dalam mengklasifikasikan sentimen kritik dan saran tersebut, sehingga dapat diketahui algoritma mana yang lebih sesuai dan akurat untuk diterapkan pada konteks data evaluasi pelatihan. Dengan begitu, penelitian ini dapat memberikan kontribusi teoritis dalam kajian analisis sentimen sekaligus memberikan kontribusi praktis untuk pengelolaan data komentar peserta pelatihan PT Brainmatics Indonesia Cendekia.

Sejumlah penelitian terdahulu telah mengkaji penerapan analisis sentimen dalam konteks pendidikan maupun layanan berbasis umpan balik pengguna. Penelitian oleh [5] menunjukkan bahwa analisis sentimen mampu mengidentifikasi kecenderungan persepsi peserta pelatihan secara efektif, dengan tingkat akurasi yang cukup tinggi menggunakan algoritma klasifikasi tradisional. Sementara itu, [6] menemukan bahwa metode Support Vector Machine (SVM) cenderung memberikan performa yang lebih baik dibandingkan Naïve Bayes (NB) dalam klasifikasi teks berbahasa Indonesia, khususnya pada data yang memiliki kompleksitas tinggi dan distribusi fitur yang tidak seimbang.

Di sisi lain, beberapa penelitian lain melaporkan hasil yang berbeda, di mana Naïve Bayes justru menunjukkan kinerja yang kompetitif bahkan unggul dalam kondisi dataset tertentu, terutama ketika jumlah data relatif kecil dan fitur yang digunakan sederhana. Perbedaan hasil ini menunjukkan bahwa performa algoritma sangat bergantung pada karakteristik data, seperti ukuran dataset, distribusi kelas, serta teknik preprocessing yang digunakan [7].

Meskipun demikian, sebagian besar penelitian sebelumnya masih berfokus pada domain umum seperti ulasan produk, media sosial, atau evaluasi pembelajaran formal di institusi pendidikan. Penelitian yang secara khusus mengkaji analisis sentimen terhadap komentar evaluasi pelatihan profesional di bidang teknologi informasi, khususnya dalam konteks lembaga pelatihan di Indonesia, masih relatif terbatas.

Berdasarkan hal tersebut, terdapat kesenjangan penelitian (research gap) dalam pemanfaatan analisis sentimen untuk mengolah data evaluasi pelatihan yang bersifat tidak terstruktur, serta dalam perbandingan kinerja algoritma Naïve Bayes dan Support Vector Machine pada konteks data pelatihan profesional. Oleh karena itu, penelitian ini tidak hanya berfokus pada klasifikasi sentimen, tetapi juga memberikan evaluasi komparatif terhadap kedua algoritma tersebut untuk menentukan metode yang paling sesuai dengan karakteristik data evaluasi pelatihan di PT Brainmatics Indonesia Cendekia.

2. METODOLOGI PENELITIAN

2.1 Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen komparatif untuk mengevaluasi performa dua algoritma machine learning, yaitu Naïve Bayes (NB) dan Support Vector Machine (SVM), dalam melakukan analisis sentimen terhadap kritik dan saran peserta pelatihan di PT Brainmatics Indonesia Cendekia. Pemilihan kedua algoritma ini didasarkan pada literatur yang menunjukkan efektivitas keduanya dalam klasifikasi teks, khususnya pada domain umpan balik pendidikan maupun ulasan produk [8].

2.2 Sampling/Metode Pemilihan Sampel

Metode pemilihan sampel dalam penelitian ini menggunakan total sampling, yaitu teknik pengambilan sampel dengan cara menggunakan seluruh anggota populasi sebagai sampel penelitian (Sugiyono, 2019). Metode ini dipilih karena populasi penelitian ini memiliki batasan yang jelas serta jumlah data yang tersedia masih memungkinkan untuk dianalisis secara menyeluruh. Selain itu, penggunaan total sampling bertujuan untuk memperoleh gambaran sentimen peserta pelatihan secara komprehensif serta meminimalkan potensi bias dalam pemilihan sampel [9]. Adapun populasi dalam penelitian ini adalah seluruh komentar kritik dan saran peserta pelatihan yang diperoleh dari hasil evaluasi pelatihan di PT Brainmatics Indonesia Cendekia pada periode 3 Januari 2025 sampai dengan 20 November 2025 dengan total komentar kritik dan saran sebanyak 518. Dari populasi tersebut, seluruh data komentar digunakan sebagai sampel penelitian tanpa dilakukan proses pengambilan sebagian data tertentu.

2.3 Metode Pengumpulan Data

Metode pengumpulan data dalam penelitian ini dilakukan dengan menggunakan data sekunder yang diperoleh dari hasil evaluasi peserta pelatihan di PT Brainmatics Indonesia Cendekia. Data tersebut berasal dari instrumen evaluasi pelatihan berupa kuesioner yang disebarluaskan kepada peserta setelah mengikuti kegiatan pelatihan. Kuesioner evaluasi terdiri atas dua bagian utama, yaitu penilaian kuantitatif dan kolom komentar terbuka yang berisi kritik serta saran dari peserta pelatihan. Penelitian ini memfokuskan pada data komentar terbuka yang bersifat kualitatif atau unstructured data. Data komentar tersebut merepresentasikan opini, pengalaman, serta persepsi peserta terhadap pelaksanaan pelatihan yang telah diikuti, sehingga relevan untuk dianalisis menggunakan pendekatan analisis sentimen berbasis text mining.

2.4 Teknik Analisis Data

Teknik analisis data dalam penelitian ini diawali dengan proses pengumpulan data berupa kritik dan saran peserta pelatihan. Data yang diperoleh merupakan data teks bebas (unstructured text), sehingga memerlukan tahapan pembersihan dan transformasi sebelum dapat digunakan dalam analisis sentiment [10]. Tahapan analisis diawali dengan proses text preprocessing, yang meliputi beberapa langkah berikut:

a. Cleansing

Menghapus karakter atau simbol yang tidak relevan, seperti angka, tanda baca, emoji, atau karakter khusus lainnya. Langkah ini bertujuan agar dataset dapat diolah dan mengurangi data noise yang terdapat didalamnya.

b. Case Folding

Menyeragamkan seluruh huruf dalam teks menjadi huruf kecil (lowercase) untuk memastikan konsistensi bentuk kata, seperti kata “Pelayanan” diubah menjadi “pelayanan” [11].

c. Normalization

Tahap ini untuk memperbaiki kata singkatan atau kata yang penulisannya tidak tepat/typo, tahap ini juga merubah kata yang tidak baku menjadi bentuk kata dasar. Pada kalimat yang tidak terstruktur biasanya banyak kata yang disingkat dan/atau tidak baku, serta kesalahan saat pengetikan (typo), text transformation adalah tahap memperbaiki kata tersebut [12].

d. Stopword Removal (Filtering)

Menghapus kata-kata umum yang tidak memberikan kontribusi signifikan terhadap klasifikasi sentimen, seperti “yang”, “dan”, “di”, atau “untuk”. Proses ini dilakukan menggunakan daftar stopwords Bahasa Indonesia dari pustaka NLTK, ditambah daftar stopwords tambahan yang disusun secara manual .

e. Tokenizing

Memisahkan teks, baik dalam bentuk kalimat maupun paragraf, menjadi unit kata (token) agar dapat dianalisis secara kata per kata.

f. Stemming

Mengembalikan kata ke bentuk dasarnya dengan bantuan pustaka Sastrawi. Misalnya, kata “diajarkan”, “mengajari”, dan “mengajarkan” akan diubah menjadi bentuk dasar “ajar”.

Setelah tahap berikutnya adalah pelabelan sentimen dengan menggunakan metode lexicon. Pada tahap ini, setiap data ulasan diberikan label sentimen positif atau negatif. Selanjutnya, teks yang telah diberi label diubah menjadi representasi numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk menghitung bobot kepentingan suatu kata dalam dokumen, dengan langkah-langkah sebagai berikut:

1. Term Frequency (TF): Mengukur frekuensi kemunculan kata dalam suatu ulasan.
2. Inverse Document Frequency (IDF): Mengukur tingkat kelangkaan kata di seluruh ulasan, sehingga kata yang umum akan memiliki bobot lebih rendah.
3. Skor TF-IDF: Hasil perkalian antara nilai TF dan IDF, yang menghasilkan representasi numerik kata sebagai fitur masukan bagi algoritma klasifikasi.

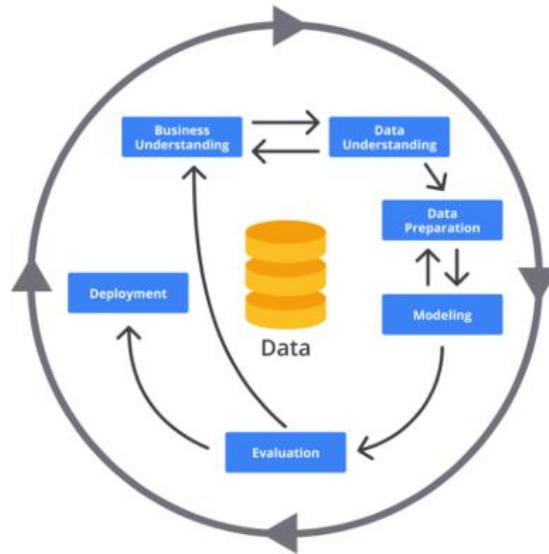
Hasil dari proses TF-IDF berupa matriks fitur yang dapat digunakan oleh model machine learning untuk melakukan klasifikasi sentimen.

2.5 Langkah-Langkah Penelitian

Penelitian ini menggunakan kerangka kerja Cross Industry Standard Process for Data Mining (CRISP-DM) sebagai acuan dalam pelaksanaan analisis sentimen. Kerangka kerja ini dipilih karena menyediakan tahapan yang sistematis dan terstruktur dalam proses pengolahan dan analisis data teks [13]. Tahapan penelitian dimulai dari Business Understanding, yaitu penentuan permasalahan dan tujuan penelitian terkait pemanfaatan kritik dan saran peserta pelatihan sebagai bahan evaluasi kualitas pelatihan di PT Brainmatics Indonesia Cendekia. Selanjutnya, pada tahap Data Understanding, dilakukan pengumpulan dan pemahaman awal terhadap data komentar evaluasi peserta pelatihan yang bersifat tidak terstruktur.

Tahap Data Preparation mencakup proses pembersihan dan penyiapan data melalui tahapan text preprocessing serta pelabelan sentimen. Data yang telah disiapkan kemudian diubah menjadi fitur numerik menggunakan metode TF-IDF. Pada tahap Modeling, dilakukan pembangunan model klasifikasi sentimen menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM). Tahap Evaluation dilakukan untuk menilai kinerja model klasifikasi menggunakan metrik evaluasi berbasis confusion matrix. Tahap terakhir adalah Deployment, yaitu pemanfaatan hasil analisis sentimen sebagai dasar pendukung evaluasi dan perbaikan pelaksanaan pelatihan.

Alur tahapan penelitian berdasarkan kerangka kerja CRISP-DM disajikan pada gambar berikut.



Gambar 2. Cross Industry Standard Process for Data Mining (CRISP-DM)

3. HASIL DAN PEMBAHASAN

3.1 Business Understanding

Pada tahap business understanding, penelitian ini difokuskan pada pemahaman permasalahan yang dihadapi oleh PT Brainmatics Indonesia Cendekia dalam mengevaluasi kualitas pelaksanaan pelatihan. Berdasarkan hasil observasi terhadap data evaluasi pelatihan, khususnya pada kolom kritik dan saran peserta, ditemukan bahwa informasi kualitatif tersebut belum dianalisis secara sistematis, meskipun mengandung masukan penting terkait pengalaman dan persepsi peserta pelatihan. Seiring dengan meningkatnya jumlah dan variasi program pelatihan, perusahaan membutuhkan metode evaluasi yang mampu mengolah data dalam jumlah besar secara objektif dan efisien. Evaluasi berbasis skor kuantitatif saja dinilai belum cukup untuk menggambarkan kepuasan dan ketidakpuasan peserta secara menyeluruh [14]. Oleh karena itu, diperlukan pendekatan analisis yang mampu mengungkap kecenderungan sentimen peserta pelatihan, baik positif, netral, maupun negatif, dari data komentar yang bersifat tidak terstruktur.

3.2 Data Understanding

Pada tahapan data understanding, peneliti melakukan pengumpulan dan pemahaman awal terhadap data yang digunakan dalam penelitian ini. Data diperoleh dari fungsi training di PT Brainmatics Indonesia Cendekia, berupa komentar kritik dan saran yang diberikan oleh peserta setelah mengikuti dan menyelesaikan program pelatihan. Pengisian komentar dilakukan melalui formulir evaluasi pelatihan yang disediakan oleh pihak penyelenggara sebagai bagian dari proses penilaian pascapelatihan.

No	Name Training	Nama Perusahaan	Tanggal Pelaksanaan		Komentar dan Saran
			Mulai	Selesai	
1	Data Mining Fundamentals	PT Geo Dipa Energi	20 January 2025	23 Januari 2025	Ahmad Ryan Suryansyah : Banyak istilah asing yg belum pernah dipahami... pengenalan stelah membutuhkan waktu yang lebih banyak Amalia : sebetulnya dari contoh kasus yang diberikan oleh perusahaan, penyelenggara training memastikan tujuan apa saja yang ingin dicapai, kelengkapan datanya disesuaikan, sehingga saat pelaksanaan training semua tujuan dapat terjawab Charunnisa Burhanuddin : Mengingat di awal kebutuhan dari perusahaan trainee sehingga lebih dapat memperkaya studi kasus Genry Muzlani Yudawinata : Pertahankan kualitas pengajar dan sarana training Mohamad Nida Dharma : Bagus Retno Andriyana : Semoga after training dapat memberikan asistensi sistem Paradeo : Terima kasih untuk kesempatan belajar... Data mining ini hal baru bagi kami, utk materi yg banyak perlu Waktu Indri Hanjaya : Sudah baik dan diperhatikan Weni Kusumaningrum : Pertahankan pelayanannya
2	Data Analytics for Surveillance	PT Bank Negara Indonesia Tbk	21 January 2025	23 Januari 2025	Amri Fadel : Sangat menambah wawasan Jheni Fernando : Training sangat bagus dan detail serta ramah untuk pengajarnya erdas Yusuf Priandana : Lebih banyak materinya lagi Agung Stimoraningrum : Materinya sangat bagus dan relevan dengan kebutuhan pekerjaan Dinar Nur Khalifah : Sudah baik Rika Angel Karini Er Saragih : Pengajar telatua cepat dalam menjelaskan materi sehingga saya sering kelupaan. Ada beberapa materi yang tidak dipelajari secara rinci. Dwi Andini Hidayat : Agar menambahkan study kasus yang berkaitan dengan jobsdes peserta Lisa Maghdalena : Sudah baik Prima Sholi Ramadhani : menggunakan data real yang mendekati dengan data yang digunakan sehari-hari. Rifqi Tyas S : Waktu disesuaikan dengan materi agar tidak telatua cepat. Dwi Retnoningrum : Materi dan penyampaian sangat menarik dan mudah dimengerti. Adelia Yansuanti Andria : menambah soal latihan dan di tesion 1:1 maka dari itu codingan Venda Putri Aprilyanta : Penyampaian materi dapat diterima dengan baik, dapat dipahami, dan fasilitator serta anggota tim dapat merespon dengan baik permasalahan yang terjadi ketika proses pelatihan Ada materi & contoh soal latihan sudah cukup, peluang perbaikan selanjutnya bisa menggunakan ini data dari pengguna (ohi, peserta pelatihan) untuk memberikan pemahaman lebih baik bagi pengguna khususnya untuk oan & visualisasi ini data yang selama ini terdenda oleh format data yang berbeda dan data kotor Rima Ardanti : Lebih banyak praktik Nuhfa Anggia : semoga bisa lebih fokus saja pada 1 alat analisa data, agar peserta bisa fokus dan bisa perbanyak latihan dengan waktu terbatas dan materi yang bagus ini, tidak cukup waktunya Aryni Fregina Rahmadhani : Metode pembelajaran menggunakan full workshop dan beberapa usecase (dibagi kelompok) sehingga setiap orang mendapatkan ilmu yang berbeda dan di anggap dengan sharing sesama peserta (dita metode presentasi) Ila Filia : Sharing soal-soal baru yang bisa kami gunakan untuk practice lebih sering setiap 2 minggu sekali dan kami diberikan kemudahan untuk menghubungi Pak Riki untuk konsultasi lebih lanjut Dwi Sucati Apriliane : Instruktur dan tim sudah sangat banyak membantu namun bisa lebih ditanyakan untuk perbanyak latihan soal dan praktik yang berhubungan dengan organisasi Ricky Andan : Pemaparan materi agar bisa lebih pelan dan setiap teori yang diajarkan dipaparkan terlebih dahulu oleh peserta

Gambar 3. Contoh Dataset Penelitian

Pada gambar di atas terlihat bahwa dalam formular evaluasi terdapat beberapa atribut data, antara lain nama program pelatihan yang diikuti, nama perusahaan peserta, tanggal mulai dan tanggal selesai pelatihan, nama peserta, serta komentar kritik dan saran yang diberikan. Namun demikian, penelitian ini secara khusus memfokuskan analisis hanya pada data komentar kritik dan saran peserta pelatihan. Atribut lain yang bersifat identitas maupun administratif tidak dilibatkan dalam proses analisis, karena tidak berkontribusi secara langsung terhadap tujuan penelitian, yaitu mengidentifikasi sentimen peserta pelatihan berdasarkan opini yang disampaikan

3.3 Data Preparation

Data ulasan yang diperoleh dan telah melalui eksplorasi awal masih mengandung berbagai elemen yang tidak berkaitan dengan analisis sentimen, seperti simbol khusus, tanda baca, angka, serta istilah yang tidak memberikan kontribusi signifikan terhadap proses analisis. Oleh karena itu, tahap persiapan data menjadi langkah penting untuk memastikan kualitas data yang digunakan dalam proses pemodelan berada pada kondisi optimal.

3.3.1 Data Cleansing

Pada tahap ini, dilakukan penghapusan berbagai karakter dan elemen yang berpotensi menimbulkan noise dalam proses analisis, seperti emotikon, tag HTML, hashtag, tautan (URL), username, serta tanda baca dan simbol khusus yang tidak relevan, antara lain (., “?!@%&* :+ - _ =). Keberadaan elemen-elemen tersebut dapat mengganggu proses pemodelan karena tidak memiliki makna sentimen yang jelas serta dapat memengaruhi representasi fitur teks[15]

Tabel 1. Contoh Hasil Data Cleansing pada Data Kritik Saran Peserta Pelatihan

Sebelum	Sesudah
Sebaiknya satu hari untuk pemaparan satu hari lagi untuk langsung mengerjakan Pengerjaan latihan sebaiknya menggunakan PowerPoiwnt pada GWS atau Canva sehingga lebih kreatif dan waktu pengerjaan lebih cepat	sebaiknya satu hari untuk pemaparan satu hari lagi untuk langsung mengerjakan. pengerjaan latihan sebaiknya menggunakan powerpoiwnt pada gws atau canva, sehingga lebih kreatif dan waktu pengerjaan lebih cepat

3.3.3 Normalisasi Data

Proses normalisasi pada penelitian ini bertujuan untuk menyeragamkan bentuk kata dalam teks kritik dan saran peserta pelatihan agar sesuai dengan format bahasa Indonesia yang baku dan konsisten. Normalisasi dilakukan untuk mengatasi variasi penulisan kata yang sering muncul dalam komentar peserta, seperti penggunaan singkatan, bahasa tidak formal, kata serapan, maupun kesalahan pengetikan. Variasi tersebut, jika tidak ditangani, dapat menyebabkan perbedaan representasi kata yang sebenarnya memiliki makna yang sama dalam proses analisis sentimen [16].

Tabel 2. Contoh Hasil Normalisasi Data pada Data Kritik Saran Peserta Pelatihan

Sebelum	Sesudah
sebaiknya satu hari untuk pemaparan satu hari lagi untuk langsung mengerjakan. pengerjaan latihan sebaiknya menggunakan powerpoiwnt pada gws atau canva, sehingga lebih kreatif dan waktu pengerjaan lebih cepat	sebaiknya satu hari untuk pemaparan satu hari lagi untuk langsung mengerjakan. pengerjaan latihan sebaiknya menggunakan powerpoint pada gws atau canva, sehingga lebih kreatif dan waktu pengerjaan lebih cepat

3.3.4 Stopwords

Pada tahap ini, dilakukan proses stopwords removal untuk menyaring kata-kata umum yang sering muncul dalam komentar kritik dan saran peserta pelatihan, namun tidak memiliki kontribusi yang signifikan terhadap penentuan sentimen. Kata-kata seperti “yang”, “dan”, “ke”, “dari”, “pada”, dan “dengan” kerap muncul dalam teks, tetapi tidak secara langsung mencerminkan opini atau penilaian peserta terhadap pelatihan yang diikuti. Penghapusan stopwords bertujuan untuk memfokuskan analisis hanya pada kata-kata yang bermakna dan relevan, sehingga informasi sentimen yang terkandung dalam komentar dapat ditangkap secara lebih akurat [17].

Tabel 3. Contoh Hasil Stopwords Removal pada Data Kritik Saran Peserta Pelatihan

Sebelum	Sesudah
sebaiknya satu hari untuk pemaparan satu hari lagi untuk langsung mengerjakan. pengerjaan latihan sebaiknya menggunakan powerpoint pada gws atau canva, sehingga lebih kreatif dan waktu pengerjaan lebih cepat	pemaparan langsung pengerjaan latihan powerpoint gws canva kreatif pengerjaan cepat

3.3.5 Tokenisasi

Proses tokenisasi ini dilakukan menggunakan library pemrosesan bahasa alami pada Python, seperti NLTK (Natural Language Toolkit), yang menyediakan fungsi `word_tokenize`, yang secara otomatis memisahkan teks berdasarkan spasi dan tanda baca yang telah dibersihkan sebelumnya. Hasil tokenisasi berupa daftar kata (token) kemudian digunakan sebagai input pada tahapan selanjutnya [18].

Tabel 4. Contoh Hasil Tokenisasi pada Data Kritik Saran Peserta Pelatihan

Sebelum	Sesudah
pemaparan langsung pengerjaan latihan powerpoint gws canva kreatif pengerjaan cepat	['pemaparan', 'langsung', 'pengerjaan', 'latihan', 'powerpoint', 'gws', 'canva', 'kreatif', 'pengerjaan', 'cepat']

3.3.6 Stemming

Proses stemming pada penelitian ini dilakukan untuk mengembalikan kata-kata dalam komentar kritik dan saran peserta pelatihan ke bentuk dasarnya dengan menghilangkan imbuhan, baik awalan, sisipan, maupun akhiran. Tahap ini bertujuan untuk menyatukan berbagai bentuk kata yang memiliki makna dasar yang sama agar dapat diperlakukan sebagai satu entitas oleh algoritma klasifikasi. Dalam implementasinya, proses stemming dilakukan menggunakan modul StemmerFactory dari library Sastrawi, yang dirancang khusus untuk Bahasa Indonesia. Melalui modul ini, variasi kata seperti pelatihan, melatih, dan dilatih disederhanakan menjadi bentuk dasar latih, sehingga kata-kata dengan akar yang sama dapat diidentifikasi secara konsisten [19].

Tabel 5. Contoh Hasil Stemming pada Data Kritik Saran Peserta Pelatihan

Sebelum	Sesudah
pemaparan langsung pengerjaan latihan powerpoint gws canva kreatif pengerjaan cepat	apar langsung kerja latih powerpoint gws canva kreatif kerja cepat

3.3.7 Proses Data Labeling

Pada tahap pelabelan data, penelitian ini menggunakan pendekatan lexicon-based yang diimplementasikan menggunakan bahasa pemrograman Python. Pelabelan sentimen dilakukan secara otomatis dengan memanfaatkan kamus sentimen InSet (Indonesia Sentiment Lexicon) sebagai acuan dalam menentukan polaritas sentimen setiap komentar kritik dan saran peserta pelatihan. Dataset teks yang telah melalui seluruh tahapan preprocessing kemudian diproses menggunakan Python untuk mencocokkan setiap kata dalam komentar dengan daftar kata positif dan negatif yang terdapat pada kamus InSet. Proses ini dilakukan secara komputasional dengan menghitung skor sentimen berdasarkan selisih antara jumlah kemunculan kata positif dan kata negatif dalam setiap teks komentar. Tabel di bawah ini menunjukkan contoh hasil proses data labeling berdasarkan kamus sentiment InSet (Indonesia Sentiment Lexicon) [20].

Tabel 6. Contoh Hasil Data Labeling dengan Kamus Sentimen InSet

Text	Score	Label
sebaiknya satu hari untuk pemaparan satu hari lagi untuk langsung mengerjakan pengerjaan latihan sebaiknya menggunakan powerpoint pada gws atau canva sehingga lebih kreatif dan waktu pengerjaan lebih cepat	3	positif
metode pembelajaran full workshop usecase dibagi kelompok orang ilmu berbeda dilengkapi sharing peserta metode presentasi	-4	negatif
materi disampaikan dengan study case dilapangan	0	netral

Pada tabel 6 menunjukkan penentuan label sentimen dilakukan berdasarkan nilai skor yang dihasilkan. Apabila skor sentimen bernilai lebih besar dari nol, maka komentar diberi label positif. Sebaliknya, apabila skor sentimen bernilai kurang dari nol, maka komentar diklasifikasikan sebagai negatif. Sementara itu, komentar dengan skor sentimen sama dengan nol dikategorikan sebagai netral. Seluruh proses pelabelan ini dilakukan secara otomatis menggunakan skrip Python, sehingga dapat mengurangi subjektivitas dan mempercepat proses penyiapan dataset.

Tabel 7. Distribusi Sentimen

No.	Pelabelan	Jumlah Data
1	Positif	283
2	Negatif	154
3	Netral	50
Grand Total		487

Berdasarkan tabel 7 di atas proses data labeling yang telah dilakukan menggunakan pendekatan lexicon-based dengan kamus sentimen InSet, diketahui bahwa sentimen yang paling dominan pada data komentar dan saran peserta pelatihan adalah sentimen positif, yakni sebanyak 283 komentar. Dominasi sentimen positif ini menunjukkan bahwa secara umum peserta memberikan tanggapan yang baik terhadap pelatihan yang diikuti. Meskipun demikian, masih terdapat komentar dengan sentimen netral dan negatif dalam jumlah yang cukup signifikan. Sentimen netral mencerminkan adanya komentar yang bersifat informatif atau deskriptif tanpa kecenderungan emosi tertentu, sedangkan sentimen negatif menunjukkan adanya kritik dan saran dari peserta yang perlu mendapat perhatian lebih lanjut. Keberadaan variasi sentimen ini menunjukkan bahwa dataset memiliki keragaman opini, sehingga layak dan relevan untuk digunakan pada tahap pemodelan analisis sentimen menggunakan algoritma klasifikasi pada penelitian ini.

3.3.8 Proses Feature Extraction

Pada tahap ekstraksi fitur, teks hasil preprocessing ditransformasikan ke dalam bentuk numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk merepresentasikan tingkat kepentingan suatu kata dalam sebuah dokumen relatif terhadap keseluruhan korpus, sehingga data teks dapat diproses lebih lanjut oleh algoritma machine learning [21].

	text	label	baik	materi	lebih	dapat	dan	di	training	sudah	waktu	untuk	yang
0	banyak istilah asing yang belum pernah dipahami pengenalan istilah membutuhkan	positif	0	0	0.3951	0	0	0	0	0	0.5642	0	0.725
1	sebaiknya dari contoh kasus yang diberikan oleh perusahaan penyelenggara training	negatif	0	0	0	0.4407	0	0	0.7212	0	0	0	0.5345
2	menggal di awal kebutuhan dari perusahaan trainee sehingga lebih dapat memperk	positif	0	0	0.4592	0.6946	0	0.5538	0	0	0	0	0
3	pertahankan kualitas pengajar dan sarana training	positif	0	0	0	0	0.6009	0	0.7993	0	0	0	0
4	bagus	positif	0	0	0	0	0	0	0	0	0	0	0
5	semoga after training dapat memberikan asistensi sistem	netral	0	0	0	0.7739	0	0	0.6333	0	0	0	0
6	terima kasih untuk kesempatan belajar data mining ini hal baru bagi kami untuk mat	positif	0	0.3671	0	0	0	0	0	0	0.4805	0.7342	0.3088
7	sudah baik dan dipertahankan	positif	0.666	0	0	0	0.4781	0	0	0.5727	0	0	0
8	pertahankan pelayanannya	positif	0	0	0	0	0	0	0	0	0	0	0
9	sangat menambah wawasan	positif	0	0	0	0	0	0	0	0	0	0	0

Gambar 4. Hasil Ekstraksi Fitur Term Frequency - Inverse Document Frequency (TF-IDF)

3.4 Modeling

Pada penelitian ini digunakan dua algoritma klasifikasi, yaitu Naïve Bayes dan Support Vector Machine (SVM), untuk mengklasifikasikan sentimen kritik dan saran peserta pelatihan ke dalam kelas positif, negatif, dan netral. Proses pemodelan dilakukan dengan menerapkan variasi rasio pembagian data latih dan data uji sebesar 90:10, 80:20, 70:30, dan 60:40. Variasi rasio ini bertujuan untuk menganalisis ketahanan serta performa masing-masing algoritma terhadap perubahan komposisi data yang digunakan dalam proses pelatihan dan pengujian model.

Tabel 8. Perbandingan Akurasi Algoritma Naive Bayes dan Support Vector Machine

Pembagian Data Training & Testing	Akurasi	
	Naïve Bayes	Support Vector Machine (SVM)
90:10	61.2%	73.5%
80:20	64.3%	76.5%
70:30	65.3%	75.5%
60:40	64.1%	79.0%

Tabel 8 menunjukkan perbandingan akurasi algoritma Naïve Bayes dan Support Vector Machine (SVM) pada berbagai skenario pembagian data latih dan data uji. Hasil pengujian menunjukkan bahwa akurasi Naïve Bayes relatif stabil pada kisaran 61,2%–65,3%, sedangkan akurasi SVM cenderung lebih tinggi pada seluruh variasi rasio data. Akurasi tertinggi Naïve Bayes diperoleh pada pembagian data 70:30 sebesar 65,3%, sementara SVM mencapai akurasi tertinggi pada rasio 60:40 sebesar 79,0%. Secara umum, SVM menunjukkan performa yang lebih unggul dibandingkan Naïve Bayes, yang dikarenakan kemampuannya bekerja efektif pada data berdimensi tinggi hasil transformasi TF-IDF serta dalam menentukan decision boundary yang optimal. Sebaliknya, performa Naïve Bayes yang lebih rendah kemungkinan dipengaruhi oleh asumsi independensi antar kata yang tidak sepenuhnya terpenuhi pada data teks kritik dan saran.

3.5 Analisis Hasil Pengujian

Berdasarkan hasil pengujian perbandingan berbagai komposisi antara data training dan data testing sebagaimana ditampilkan pada Tabel 4.26, diketahui bahwa algoritma Support Vector Machine memberikan nilai akurasi lebih tinggi jika dibandingkan dengan algoritma Naive Bayes. Berikut adalah ringkasan hasil pengujian dari algoritma Naive Bayes dan Support Vector Machine pada berbagai komposisi pembagian data latih dan data uji.

Tabel 9. Hasil Pengujian Terbaik

Pembagian Data	Naïve Bayes			SVM		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Data 90:10	0.66	0.61	0.48	0.73	0.73	0.73
Data 80:20	0.67	0.64	0.54	0.76	0.76	0.75
Data 70:30	0.68	0.65	0.56	0.74	0.75	0.74
Data 60:40	0.64	0.64	0.54	0.78	0.79	0.77

Pada tabel 9 di atas, konsisten menunjukkan performa yang lebih baik dibandingkan Naïve Bayes pada seluruh skenario pembagian data. Pada rasio data 90:10, SVM menghasilkan nilai precision, recall, dan F1-score yang seimbang sebesar 0,73, lebih tinggi dibandingkan Naïve Bayes yang memiliki F1-score relatif rendah (0,48). Pada pembagian 80:20

dan 70:30, kinerja SVM tetap unggul dengan F1-score masing-masing sebesar 0,75 dan 0,74, sementara Naïve Bayes menunjukkan peningkatan yang lebih terbatas. Performa terbaik SVM terlihat pada pembagian data 60:40, dengan precision 0,78, recall 0,79, dan F1-score 0,77. Secara keseluruhan, hasil ini menunjukkan bahwa SVM lebih stabil dan andal dalam mengklasifikasikan sentimen pada berbagai variasi pembagian data, sedangkan Naïve Bayes cenderung memiliki performa yang lebih rendah, terutama pada keseimbangan antara precision dan recall.

3.6 Optimalisasi Model

3.6.1 Optimalisasi Model Naive Bayes

Optimalisasi model Naïve Bayes dilakukan untuk meningkatkan performa klasifikasi sentimen melalui penyesuaian nilai hiperparameter yang digunakan. Proses ini bertujuan untuk memperoleh konfigurasi model yang lebih optimal dibandingkan dengan model baseline dalam mengklasifikasikan kritik dan saran peserta pelatihan.

Tabel 10. Parameter Baseline Naive Bayes

No.	Parameter	Nilai Baseline
1	alpha	1.0
2	class_prior	None
3	fit_prior	True
4	force_alpha	True

Tabel 10 di atas menampilkan parameter baseline yang digunakan pada algoritma Naïve Bayes sebelum proses optimalisasi. Nilai alpha sebesar 1.0 diterapkan untuk smoothing, sedangkan pengaturan class_prior bernilai None dan fit_prior bernilai True memungkinkan model menghitung probabilitas kelas secara otomatis. Parameter force_alpha bernilai True digunakan untuk memastikan penerapan nilai alpha selama proses pelatihan.

a. Optimalisasi Model Naive Bayes dengan Grid Search

Pada tahap ini dilakukan optimalisasi model Naïve Bayes menggunakan metode Grid Search dengan menguji berbagai kombinasi parameter secara sistematis. Pendekatan ini bertujuan untuk memperoleh konfigurasi parameter terbaik yang mampu meningkatkan kinerja model dibandingkan kondisi baseline.

Tabel 11. Hasil Perbandingan Performa Model Naive Bayes Sebelum dan Sesudah Optimalisasi dengan Grid Search

Matrix	Baseline	Sesudah Tuning	Peningkatan
Akurasi	0.6310	0.7347	0.0937
F1-Score	0.5454	0.7156	0.1702

Tabel 11 di atas menunjukkan perbandingan performa algoritma Naïve Bayes sebelum dan sesudah dilakukan optimalisasi parameter menggunakan Grid Search. Hasil pengujian menunjukkan peningkatan kinerja yang signifikan, di mana nilai akurasi meningkat dari 0,6310 menjadi 0,7347 dan nilai F1-score meningkat dari 0,5454 menjadi 0,7156. Peningkatan ini mengindikasikan bahwa proses tuning parameter mampu memperbaiki kemampuan model dalam mengklasifikasikan sentimen secara lebih seimbang.

Tabel 12. Confusion Matrix pada Optimalisasi Akurasi Naive Bayes dengan Grid Search

	Pred. Negatif	Pred. Netral	Pred. Positif
True Negatif	20	0	11
True Netral	4	2	4
True Positif	6	1	50

Tabel 12 di atas menampilkan confusion matrix hasil optimalisasi algoritma Naïve Bayes menggunakan Grid Search dengan fokus pada peningkatan akurasi. Dari tabel tersebut terlihat bahwa model mampu mengklasifikasikan sentimen positif dengan cukup baik, sementara masih terdapat kesalahan klasifikasi pada kelas sentimen negatif dan netral. Hasil ini menunjukkan adanya perbaikan kinerja model, meskipun tantangan dalam membedakan kelas minoritas masih tetap ada.

Tabel 13. Hasil Perbandingan Performa Model Naive Bayes Sebelum dan Sesudah Optimalisasi f1_weighted dengan Grid Search

Matrix	Baseline	Sesudah Tuning	Peningkatan
Akurasi	0.6410	0.6837	0.0426
F1-Score	0.5454	0.6812	0.1358

Tabel 13 di atas menunjukkan perbandingan performa algoritma Naïve Bayes sebelum dan sesudah optimalisasi parameter menggunakan Grid Search dengan metrik F1-weighted. Hasil pengujian menunjukkan adanya peningkatan kinerja, di mana nilai akurasi meningkat dari 0,6410 menjadi 0,6837 dan nilai F1-score meningkat dari 0,5454 menjadi 0,6812. Peningkatan ini menunjukkan bahwa optimalisasi berbasis F1-weighted mampu memperbaiki keseimbangan kinerja model antar kelas sentimen.

Tabel 14. Confusion Matrix pada Optimalisasi fl_weighted Naïve Bayes dengan Grid Search

	Pred. Negatif	Pred. Netral	Pred. Positif
True Negatif	20	1	10
True Netral	6	3	1
True Positif	12	1	44

Tabel 14 menampilkan confusion matrix hasil optimalisasi algoritma Naïve Bayes menggunakan Grid Search dengan metrik F1-weighted. Hasil klasifikasi menunjukkan adanya peningkatan kemampuan model dalam mengenali kelas sentimen netral dan negatif dibandingkan kondisi sebelum optimalisasi, meskipun masih terdapat kesalahan klasifikasi pada beberapa data. Secara keseluruhan, optimalisasi berbasis F1-weighted membantu meningkatkan keseimbangan kinerja model antar kelas sentimen.

Secara keseluruhan, berdasarkan hasil optimalisasi yang dilakukan, algoritma Naïve Bayes mengalami peningkatan performa setelah diterapkan *hyperparameter tuning* menggunakan metode *Grid Search*. Pengujian variasi nilai parameter *alpha* pada rentang 0,10 hingga 1,00 berhasil meningkatkan akurasi model dari kondisi *baseline* sebesar 64,10% menjadi 73,47%, atau mengalami peningkatan sebesar 9,37 poin persentase. Peningkatan kinerja juga terlihat pada nilai *F1-score* yang meningkat dari 54,54% menjadi 71,56%, menunjukkan perbaikan keseimbangan klasifikasi antar kelas sentimen. Analisis *confusion matrix* mengindikasikan bahwa model menjadi lebih baik dalam mengenali kelas sentimen negatif, sehingga kesalahan klasifikasi dapat dikurangi secara signifikan.

b. Optimalisasi Model Naive Bayes dengan Random Search

Pada tahap ini dilakukan optimalisasi model Naïve Bayes menggunakan metode Random Search untuk mengeksplorasi kombinasi parameter secara acak. Pendekatan ini bertujuan untuk mengevaluasi potensi peningkatan kinerja model dibandingkan konfigurasi baseline dan hasil optimalisasi sebelumnya.

Tabel 15. Hasil Perbandingan Performa Model Naive Bayes Sebelum dan Sesudah Optimalisasi dengan Random Search

Matrix	Baseline	Sesudah Tuning	Peningkatan
Akurasi	0.6410	0.7245	0.0835
F1-Score	0.5454	0.7111	0.1657

Tabel 15 di atas menunjukkan perbandingan performa algoritma Naïve Bayes sebelum dan sesudah optimalisasi parameter menggunakan metode Random Search. Hasil pengujian menunjukkan peningkatan kinerja, di mana nilai akurasi meningkat dari 0,6410 menjadi 0,7245 dan nilai F1-score meningkat dari 0,5454 menjadi 0,7111. Peningkatan ini mengindikasikan bahwa Random Search juga efektif dalam meningkatkan performa model, meskipun hasilnya sedikit lebih rendah dibandingkan optimalisasi menggunakan Grid Search.

Tabel 16. Confusion Matrix pada Optimalisasi Naïve Bayes dengan Random Search

	Pred. Negatif	Pred. Netral	Pred. Positif
True Negatif	20	1	10
True Netral	6	2	2
True Positif	6	2	49

Pada tabel 16 di atas, hasil klasifikasi menunjukkan bahwa model mampu mengidentifikasi sentimen positif dengan baik, sementara peningkatan pada kelas sentimen negatif dan netral masih terbatas. Secara keseluruhan, optimalisasi dengan Random Search memberikan perbaikan kinerja model, meskipun belum sepenuhnya mengatasi kesalahan klasifikasi pada kelas minoritas. Secara keseluruhan, hasil optimalisasi menggunakan metode Random Search menunjukkan adanya peningkatan performa model Naïve Bayes dibandingkan kondisi baseline, baik pada metrik akurasi maupun F1-score. Namun demikian, peningkatan yang diperoleh tidak menunjukkan perbedaan yang signifikan apabila dibandingkan dengan hasil optimalisasi menggunakan Grid Search. Hal ini mengindikasikan bahwa ruang parameter yang dieksplorasi relatif sederhana dan telah tercakup secara optimal melalui pendekatan Grid Search yang sistematis, sehingga Random Search tidak memberikan keuntungan tambahan yang berarti pada kasus penelitian ini.

3.6.2 Optimalisasi Model Support Vector Machine

Optimalisasi model Support Vector Machine dilakukan untuk memperoleh kombinasi hiperparameter yang paling optimal dalam meningkatkan kinerja klasifikasi sentimen. Proses ini bertujuan untuk mengevaluasi sejauh mana penyesuaian parameter mampu memperbaiki performa model dibandingkan dengan konfigurasi baseline.

Tabel 17. Parameter Baseline SVM

No.	Parameter	Nilai Baseline
1	C	1.0
2	class_weight	balanced
3	dual	auto
4	fit_intercept	True
5	intercept_scaling	1
6	loss	Square_hinge
7	max_iter	5000
8	multi_class	ovr
9	penalty	l2

Tabel 17 di atas menunjukkan konfigurasi baseline algoritma Support Vector Machine (SVM) yang digunakan sebelum proses optimalisasi. Model menggunakan nilai C = 1.0 dengan class_weight balanced untuk mengakomodasi ketidakseimbangan kelas, serta fungsi loss square_hinge dan penalty l2 yang umum digunakan pada klasifikasi linier. Selain itu, pengaturan multi_class one-vs-rest (ovr) dan batas iterasi maksimum sebesar 5000 bertujuan memastikan proses pelatihan berjalan stabil dan konvergen.

a. Optimalisasi Model Support Vector Machine dengan Grid Search

Optimalisasi model Support Vector Machine dengan Grid Search dilakukan untuk mencari kombinasi hiperparameter terbaik secara sistematis berdasarkan nilai yang telah ditentukan. Pendekatan ini bertujuan untuk meningkatkan kinerja model sekaligus mengevaluasi stabilitas performa SVM pada klasifikasi sentimen kritik dan saran.

Tabel 18. Hasil Perbandingan Performa Model SVM Sebelum dan Sesudah Optimalisasi Akurasi, f1_weighted, recall_weighted dengan Grid

Matrix	Baseline	Sesudah Tuning	Peningkatan
Akurasi	0.7897	0.7551	-0.0346
F1-Score	0.7746	0.7522	-0.0224

Tabel 18 di atas menunjukkan bahwa proses optimalisasi model SVM menggunakan Grid Search dengan fokus pada metrik akurasi, f1_weighted, dan recall_weighted tidak menghasilkan peningkatan performa dibandingkan konfigurasi baseline. Nilai akurasi dan F1-Score justru mengalami penurunan masing-masing sebesar 3,46% dan 2,24%. Hal ini mengindikasikan bahwa parameter baseline SVM yang digunakan sebelumnya sudah berada pada konfigurasi yang relatif optimal untuk data penelitian ini.

Tabel 19. Confusion Matrix Optimalisasi Akurasi, f1_weighted, recall_weighted pada Model SVM dengan Grid

	Pred. Negatif	Pred. Netral	Pred. Positif
True Negatif	23	1	7
True Netral	4	4	2
True Positif	10	0	47

Berdasarkan Tabel 19, hasil confusion matrix menunjukkan bahwa model SVM hasil optimalisasi masih mampu mengklasifikasikan kelas positif dengan cukup baik, terlihat dari 47 data positif yang berhasil diprediksi secara benar. Namun, kesalahan klasifikasi masih terjadi pada kelas negatif dan netral, di mana sebagian data negatif dan netral diprediksi sebagai kelas positif. Kondisi ini sejalan dengan penurunan nilai akurasi dan F1-Score, yang mengindikasikan bahwa proses optimalisasi belum mampu meningkatkan keseimbangan prediksi antar kelas.

Tabel 20. Hasil Perbandingan Performa Model SVM Sebelum dan Sesudah Optimalisasi precision_weighted dengan Grid

Matrix	Baseline	Sesudah Tuning	Peningkatan
Akurasi	0.7897	0.5306	-0.2591
F1-Score	0.7746	0.5020	-0.2726

Berdasarkan Tabel 20, optimalisasi model SVM dengan fokus pada metrik precision_weighted justru menyebabkan penurunan performa yang signifikan. Nilai akurasi turun dari 78,97% menjadi 53,06% dan F1-Score menurun dari 77,46% menjadi 50,20%, yang menunjukkan bahwa konfigurasi parameter hasil tuning tidak mampu mempertahankan keseimbangan prediksi antar kelas. Hasil ini mengindikasikan bahwa precision_weighted kurang sesuai digunakan sebagai tujuan optimalisasi pada kasus analisis sentimen dalam penelitian ini.

Tabel 21. Confusion Matrix Optimalisasi precision_weighted pada Model SVM dengan Grid

	Pred. Negatif	Pred. Netral	Pred. Positif
True Negatif	31	0	0
True Netral	6	4	0

True Positif	37	3	17
---------------------	----	---	----

Berdasarkan Tabel 21 di atas, confusion matrix hasil optimalisasi precision_weighted pada model SVM menunjukkan kecenderungan model untuk memprediksi kelas negatif secara dominan. Seluruh data negatif berhasil diklasifikasikan dengan benar, namun sebagian besar data positif justru salah diklasifikasikan sebagai negatif, terlihat dari tingginya nilai false negative. Kondisi ini menyebabkan penurunan kinerja model secara keseluruhan karena ketidakseimbangan prediksi antar kelas sentimen.

Secara keseluruhan, proses optimalisasi pada model Support Vector Machine menunjukkan hasil yang bervariasi dan tidak selalu memberikan peningkatan performa dibandingkan model baseline. Meskipun Grid Search dengan variasi parameter seperti nilai C, gamma, dan kernel mampu mengeksplorasi kombinasi parameter yang lebih luas, hasil pengujian menunjukkan bahwa performa terbaik SVM justru telah dicapai pada konfigurasi baseline. Hal ini mengindikasikan bahwa model SVM pada penelitian ini sudah cukup optimal sejak awal, sehingga proses tuning lanjutan tidak memberikan peningkatan yang signifikan dan bahkan berpotensi menurunkan kinerja model.

b. Optimalisasi Model Support Vector Machine dengan Random Search

Optimalisasi model Support Vector Machine dengan Random Search dilakukan untuk mengeksplorasi ruang hiperparameter secara acak dengan tujuan menemukan konfigurasi yang berpotensi meningkatkan performa model. Pendekatan ini digunakan sebagai pembandingan terhadap Grid Search dalam menilai efektivitas strategi pencarian parameter pada tugas klasifikasi sentimen.

Tabel 22. Hasil Perbandingan Performa Model SVM Sebelum dan Sesudah Optimalisasi Akurasi, f1_weighted, recall_weighted dengan Random

Matrix	Baseline	Sesudah Tuning	Peningkatan
Akurasi	0.7897	0.7551	-0.0346
F1-Score	0.7746	0.7522	-0.0224

Berdasarkan Tabel 22 di atas, optimalisasi model Support Vector Machine menggunakan Random Search dengan metrik akurasi, f1_weighted, dan recall_weighted tidak menghasilkan peningkatan performa dibandingkan model baseline. Nilai akurasi menurun dari 78,97% menjadi 75,51%, sementara F1-Score juga mengalami penurunan dari 77,46% menjadi 75,22%. Hasil ini menunjukkan bahwa konfigurasi parameter hasil Random Search belum mampu melampaui performa baseline, sehingga model awal SVM dapat dikatakan sudah berada pada kondisi yang relatif optimal.

Tabel 23. Confusion Matrix Optimalisasi Akurasi, f1_weighted, recall_weighted pada Model SVM dengan Random

	Pred. Negatif	Pred. Netral	Pred. Positif
True Negatif	23	1	7
True Netral	4	4	2
True Positif	10	0	47

Berdasarkan Tabel 23, confusion matrix hasil optimalisasi model SVM dengan Random Search menunjukkan bahwa model masih mampu mengklasifikasikan kelas positif dengan cukup baik, terlihat dari jumlah prediksi benar yang relatif tinggi pada kelas tersebut. Namun, kesalahan klasifikasi masih cukup sering terjadi pada kelas negatif dan netral yang cenderung diprediksi sebagai kelas positif. Hal ini mengindikasikan bahwa proses optimalisasi belum memberikan perbaikan signifikan dalam membedakan kelas minoritas secara lebih akurat.

Tabel 24. Hasil Perbandingan Performa Model SVM Sebelum dan Sesudah Optimalisasi precision_weighted dengan Random

Matrix	Baseline	Sesudah Tuning	Peningkatan
Akurasi	0.7897	0.5306	-0.2591
F1-Score	0.7746	0.5020	-0.2726

Berdasarkan Tabel 24, optimalisasi model SVM menggunakan Random Search dengan metrik precision_weighted justru menyebabkan penurunan performa yang signifikan dibandingkan model baseline. Nilai akurasi menurun dari 78,97% menjadi 53,06%, sedangkan F1-Score turun dari 77,46% menjadi 50,20%. Hasil ini menunjukkan bahwa konfigurasi parameter hasil Random Search belum mampu mempertahankan keseimbangan kinerja model dalam mengklasifikasikan seluruh kelas sentimen secara optimal.

Tabel 25. Confusion Matrix Optimalisasi precision_weighted pada Model SVM dengan Random

	Pred. Negatif	Pred. Netral	Pred. Positif
True Negatif	31	0	0
True Netral	6	4	0
True Positif	37	3	17

Berdasarkan Tabel 26, confusion matrix hasil optimalisasi precision_weighted pada model SVM dengan Random Search menunjukkan kecenderungan model dalam memprediksi kelas negatif secara dominan dan mengabaikan sebagian

variasi kelas lainnya. Seluruh data negatif berhasil diprediksi dengan benar, namun banyak data positif dan netral yang salah diklasifikasikan sebagai negatif atau tidak terdistribusi secara seimbang. Kondisi ini menjelaskan penurunan akurasi dan F1-Score, karena meskipun presisi pada kelas tertentu meningkat, kemampuan generalisasi model terhadap seluruh kelas sentimen menjadi menurun.

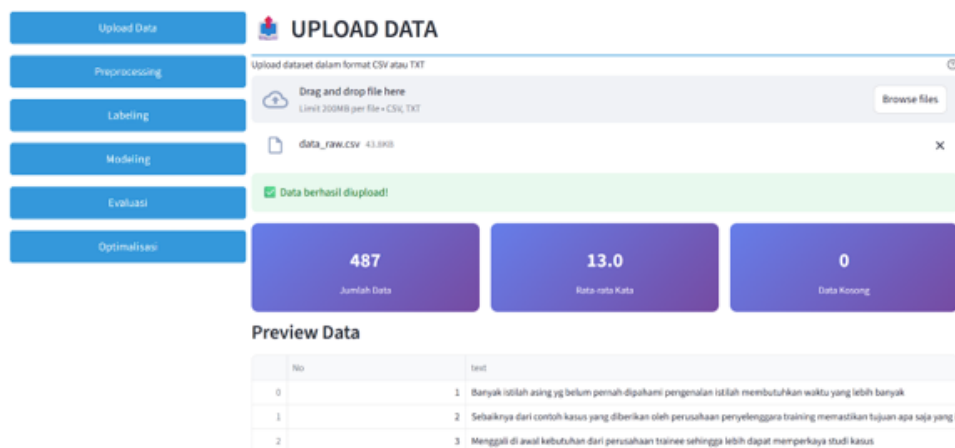
Secara keseluruhan, hasil optimalisasi model Support Vector Machine menggunakan Random Search menunjukkan penurunan performa dibandingkan model baseline. Pada evaluasi berbasis accuracy, nilai akurasi menurun dari 78,97% menjadi 75,51% (penurunan 3,46 poin persentase), yang mengindikasikan bahwa pencarian parameter secara acak belum mampu menemukan kombinasi hiperparameter yang optimal. Penurunan yang lebih signifikan terlihat pada optimalisasi dengan metrik precision_weighted, di mana akurasi turun drastis menjadi 53,06%, menunjukkan bahwa model menjadi sangat sensitif terhadap pemilihan metrik evaluasi dan cenderung menghasilkan konfigurasi yang tidak seimbang dalam mengklasifikasikan seluruh kelas sentimen.

Berdasarkan hasil eksperimen, Grid Search terbukti lebih efektif daripada Random Search untuk kasus spesifik analisis sentimen ini. Pada model Naïve Bayes, Grid Search berhasil meningkatkan F1-Score sebesar 17.02%, sementara pada model SVM meskipun terdapat variasi hasil, pendekatan sistematis Grid Search menunjukkan konsistensi yang lebih baik. Temuan ini kontras dengan teori umum yang menyatakan Random Search lebih efisien untuk ruang parameter besar, namun dalam konteks penelitian ini dengan ruang parameter yang relatif terbatas (alpha: 10 nilai, C: 4 nilai, gamma: 4 nilai, kernel: 2 nilai), Grid Search mampu melakukan eksplorasi komprehensif tanpa beban komputasi berlebihan. Total kombinasi parameter untuk SVM adalah 32 kombinasi (4×4×2), yang masih dalam rentang feasible untuk dieksplorasi secara lengkap.

Pemilihan metrik evaluasi terbukti memiliki pengaruh signifikan terhadap hasil optimalisasi. Metrik accuracy dan f1_weighted menunjukkan pola peningkatan yang konsisten pada model Naïve Bayes, sementara precision_weighted dan recall_weighted menghasilkan peningkatan yang lebih moderat. Pada model SVM, sensitivitas terhadap metrik evaluasi lebih terlihat jelas, dimana tuning dengan fokus pada precision_weighted justru menghasilkan penurunan performa yang signifikan. Hal ini mengindikasikan bahwa optimasi untuk precision tertentu dapat mengorbankan balance keseluruhan model. Hasil terbaik secara konsisten diperoleh dengan menggunakan metrik f1_weighted, yang menyeimbangkan precision dan recall, terutama penting mengingat karakteristik data yang memiliki distribusi kelas tidak seimbang.

3.7 Deployment

Deployment sistem dilakukan melalui pengembangan prototipe berbasis web menggunakan bahasa pemrograman Python dan framework Streamlit. Pendekatan ini memungkinkan sistem analisis sentimen dapat diakses dan digunakan secara interaktif oleh pengguna tanpa memerlukan kemampuan pemrograman lanjutan. Prototipe dirancang dengan konsep end-to-end pipeline yang mencakup seluruh tahapan analisis sentimen, mulai dari input data kritik dan saran peserta pelatihan hingga evaluasi hasil klasifikasi menggunakan algoritma Naïve Bayes dan Support Vector Machine. Untuk meningkatkan kemudahan penggunaan, aplikasi dilengkapi dengan menu navigasi pada sidebar yang mengorganisasi setiap tahapan proses ke dalam fitur-fitur yang terstruktur dan sistematis.



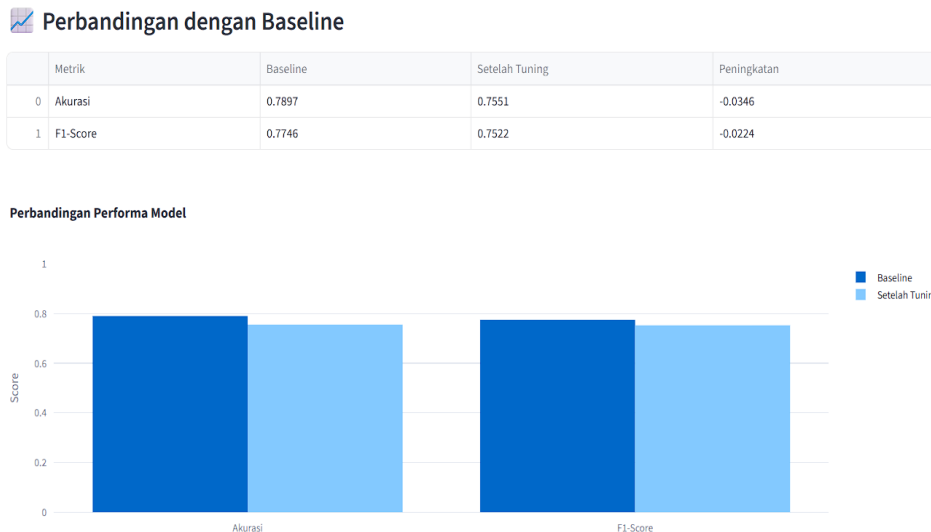
Gambar 5. Tampilan Menu Upload Data

Pada gambar 5 menampilkan menu upload data yang berfungsi sebagai pintu masuk data ke dalam sistem analisis sentimen. Melalui menu ini, pengguna dapat mengunggah dataset dalam format CSV maupun memasukkan data teks secara manual langsung melalui antarmuka aplikasi. Fitur ini memberikan fleksibilitas dalam pengujian sistem, baik menggunakan data historis yang telah tersedia maupun data baru secara langsung, sehingga mendukung proses analisis dan evaluasi kinerja algoritma Naïve Bayes dan Support Vector Machine.

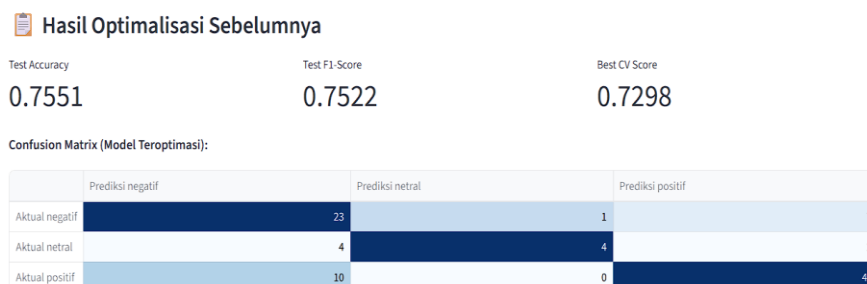


Gambar 6. Tampilan Parameter Baseline Naïve Bayes

Sebagai hasil akhir, menu Optimalisasi menyediakan antarmuka interaktif yang memungkinkan pengguna melakukan eksperimen parameter untuk menemukan konfigurasi optimal dari model Naïve Bayes dan Support Vector Machine (SVM) sebagaimana terlihat pada Gambar 6. Implementasi ini bertujuan mengatasi keterbatasan model baseline dengan menyesuaikan parameter algoritma terhadap karakteristik spesifik data sentimen dalam domain pendidikan.



Gambar 7. Tampilan Perbandingan Performa Model Sebelum dan Setelah Optimalisasi



Gambar 8. Tampilan Confusion Matrix Setelah Optimalisasi

Pada gambar 20 terdapat komponen pertama, grouped bar chart, menampilkan perbandingan side-by-side antara akurasi dan F1-score model baseline dengan model teroptimasi, dimana peningkatan performa diwakili oleh perbedaan tinggi batang yang secara visual langsung teridentifikasi. Komponen kedua, box plot distribusi skor CV, mengungkapkan variabilitas internal dari proses tuning melalui display kuartil, median, dan outliers dari semua kombinasi parameter yang dievaluasi, memberikan insight mengenai robustness konfigurasi yang ditemukan. Komponen ketiga, tabel ranking 10 kombinasi terbaik, menyajikan data terstruktur dengan kolom mean score, standard deviation, dan representasi string parameter, memungkinkan analisis trade-off antara performa rata-rata dan konsistensi. Komponen keempat, heatmap-style confusion matrix, menggunakan gradien warna biru untuk mengindikasikan frekuensi prediksi, dengan intensitas warna yang proporsional terhadap jumlah instance, memudahkan identifikasi pola kesalahan klasifikasi antar kelas.

4. KESIMPULAN

Analisis sentimen yang dilakukan telah berhasil mengintegrasikan seluruh tahapan pemrosesan teks mulai dari preprocessing, lexicon-based labeling, feature extraction menggunakan TF-IDF, modeling dengan dua algoritma berbeda, hingga optimalisasi hyperparameter secara sistematis. Implementasi berbasis Streamlit ini terbukti memberikan antarmuka yang interaktif dan user-friendly, memungkinkan pengguna dengan berbagai tingkat keahlian teknis untuk melakukan eksperimen dan analisis tanpa kompleksitas implementasi yang berlebihan. Proses optimalisasi hyperparameter melalui Grid Search dan Random Search berhasil meningkatkan performa model secara signifikan. Pada algoritma Naïve Bayes, implementasi Grid Search dengan konfigurasi alpha 0.10 hingga 1.00 berhasil meningkatkan akurasi dari 64.10% menjadi 73.47% dan F1-Score dari 54.54% menjadi 71.56%. Meskipun peningkatan ini cukup substansial, model SVM menunjukkan performa baseline yang lebih tinggi namun dengan hasil optimalisasi yang lebih variabel tergantung pada metode tuning dan metrik evaluasi yang digunakan. Perbandingan efektivitas metode tuning menunjukkan bahwa Grid Search lebih konsisten dan efektif dibandingkan Random Search untuk kasus spesifik analisis sentimen pendidikan ini. Hal ini terutama disebabkan oleh ruang parameter yang relatif terbatas sehingga memungkinkan eksplorasi komprehensif tanpa beban komputasi yang berlebihan. Temuan ini memberikan wawasan penting bahwa pemilihan metode tuning harus mempertimbangkan karakteristik spesifik dataset dan kompleksitas model, bukan hanya mengikuti praktik umum.

REFERENCES

- [1] A. A. Adesina, T. V. Iyelolu, and P. O. Paul, "Optimizing Business Processes with Advanced Analytics: Techniques for Efficiency and Productivity Improvement," *World Journal of Advanced Research and Reviews*, vol. 22, no. 3, pp. 1917–1926, Jun. 2024, doi: 10.30574/wjarr.2024.22.3.1960.
- [2] J. O. Leandro and M. I. Fianty, "Evaluation of Sentiment Analysis Methods for Social Media Applications: A Comparison of Support Vector Machines and Naïve Bayes," *JOIV: International Journal on Informatics Visualization*, vol. 2, no. 9, pp. 796–807, Mar. 2025. [Online]. Available: www.joiv.org/index.php/joiv
- [3] I. Ali Kandhro, M. Ameen Chhajro, K. Kumar, H. N. Lashari, and U. Khan, "Student Feedback Sentiment Analysis Model Using Various Machine Learning Schemes A Review," *Indian J. Sci. Technol.*, vol. 14, no. 12, pp. 1–9, Apr. 2019, doi: 10.17485/ijst/2019/v12i14/143243.
- [4] F. Pozzi, E. Fersini, E. Messina, and B. Liu, *Sentiment analysis in social networks*. Elsevier, 2017.
- [5] A. A. Hisyam and A. T. Ayunda, "Analisis Sentimen Persepsi Publik Terhadap Bank DKI pada Twitter Menggunakan Metode Naïve Bayes Classifier," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 14, no. 1, Jan. 2026, doi: 10.23960/jitet.v14i1.8842.
- [6] A. S. Lubis and R. A. Putri, "Analisis Sentimen Mahasiswa Terhadap Penggunaan E-Learning dengan Algoritma Support Vector Machine (SVM)," *Decode: Jurnal Pendidikan Teknologi Informatika*, vol. 5, no. 2, pp. 778–788, Jul. 2025, doi: 10.51454/decode.v5i2.1247.
- [7] F. A. Dara and I. Pratama, "Sentiment Analysis of the TikTok Tokopedia Seller Center Application Using Support Vector Machine (SVM) and Naive Bayes Algorithms," *International Journal Software Engineering and Computer Science (IJSECS)*, vol. 5, no. 1, pp. 177–189, Apr. 2025, doi: 10.35870/ijsecs.v5i1.3463.
- [8] M. A. Ullah, "Sentiment Analysis of Students Feedback: A Study towards Optimal Tools," in *IWCI 2016 - 2016 International Workshop on Computational Intelligence*, IEEE, Dec. 2016.
- [9] B. Gupta, I. M. Negi, K. Vishwakarma, G. Rawat, P. Badhani, and B. Tech, "Study of Twitter Sentiment Analysis using Machine Learning Algorithms on Python," *Int. J. Comput. Appl.*, vol. 165, no. 9, pp. 975–8887, 2017.
- [10] W. Kaur, V. Balakrishnan, and B. Singh, "Improving teaching and learning experience in engineering education using sentiment analysis techniques," in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, Jun. 2020, doi: 10.1088/1757-899X/834/1/012026.
- [11] J. Han, M. Kamber, and J. Pei, *Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)*. Elsevier, 2022.
- [12] N. Z. B. Jannah and K. Kusnawi, "Comparison of Naïve Bayes and SVM in Sentiment Analysis of Product Reviews on Marketplaces," *Sinkron: Jurnal dan Penelitian Teknik Informatika*, vol. 8, no. 2, pp. 727–733, Mar. 2024, doi: 10.33395/sinkron.v8i2.13559.
- [13] N. Sharma and V. Jain, "Evaluation and summarization of student feedback using sentiment analysis," in *Advances in Intelligent Systems and Computing*, Springer, 2021, pp. 385–396. doi: 10.1007/978-981-15-3383-9_35.
- [14] M. Umar, N. B. Ahmad, and A. Zainal, "Sentiment Analysis of Student's Opinion on Programming Assessment: Evaluation of Naïve Bayes over Support Vector Machines," *International Journal of Innovative Computing*, vol. 10, no. 2, Nov. 2020, doi: 10.11113/ijic.v10n2.278.
- [15] M. Fejfarová and J. Fejfar, "EMPLOYEE TRAINING AND DEVELOPMENT AND COMPETENCYBASED APPROACH: ANY RELATIONSHIP?," *Journal on Efficiency and Responsibility in Education and Science*, vol. 15, no. 4, pp. 242–250, 2022, doi: 10.7160/eriesj.2022.150405.
- [16] M. Umair, A. Hakim, A. Hussain, and S. Naseem, "Sentiment Analysis of Students' Feedback before and after COVID-19 Pandemic," *International Journal on Emerging Technologies*, vol. 12, no. 2, pp. 177–182, 2021, [Online]. Available: www.researchtrend.net
- [17] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, Dec. 2014, doi: 10.1016/j.asej.2014.04.011.
- [18] U. Lestari, T. Romadhani, and E. Fatkhiyah, "SENTIMENT ANALYSIS FOR EXTRACTING STUDENT OPINION DATA ON HIGHER EDUCATION SERVICES USING THE NAIVE BAYES CLASSIFIER AND SUPPORT VECTOR MACHINE METHODS (CASE STUDY AKPRIND INSTITUTE OF SCIENCE AND TECHNOLOGY YOGYAKARTA)," *Technology*

- Acceptance Model*) *Jurnal TAM*, vol. 13, no. 1, 2022, [Online]. Available: <https://ojs.stmikpringsewu.ac.id/index.php/JurnalTam/index>
- [19] A. Chourasiya, A. Khan, K. Bajaj, M. Tomar, T. Kohli, and D. Chauhan, “A Review of Sentiment Analysis and Emotion Detection from Text using Different Models,” 2025.
- [20] Marjo and E. Sumantri, “IMPLEMENTASI DATA MINING DALAM PREDIKSI TARGET PRODUKSI PADA PROSES KERJA MESIN MOLDING MENGGUNAKAN ALGORITMA LINEAR REGRESSION (STUDI KASUS : PT. AIM KARAWANG),” *Jurnal Indonesia : Manajemen Informatika dan Komunikasi*, vol. 4, no. 3, pp. 1694–1703, Sep. 2023, doi: 10.35870/jimik.v4i3.397.
- [21] I. Lazrig and S. L. Humpherys, “Using Machine Learning Sentiment Analysis to Evaluate Learning Impact,” 2022. [Online]. Available: <https://isedj.org/>; <https://iscap.info>