ANALISIS SENTIMEN UMPAN BALIK PESERT PELATIHAN PROXSIS ACADEMY MENGGUNAKAN ALGORITMA NAÏVE BAYES

Louqman El Hakim[1*], Agus Junaidi[2]

[1,2]*Universitas Bina Sarana Informatika*
E-mail : louqmanelhakim@gmail.com

ABSTRACT

Manual evaluation of participant feedback at Proxsis Academy is inefficient and prone to bias. This study aims to automate the process through Sentiment Analysis based on Natural Language Processing (NLP) using the Multinomial Naïve Bayes algorithm. From 1,884 initial data, cleansing and preprocessing yielded 1259 valid text rows, which were then extracted using Term Frequency-Inverse Document Frequency (TF-IDF). Testing with a 70:30 ratio resulted in an overall Accuracy of 99.21%, weighted Precision of 98%, Recall of 99%, and F1-Score of 99%. Despite the high accuracy, evaluation revealed a bias due to imbalanced data, where the model perfectly predicted the Positive class (100% Recall) but failed on the Negative class (0% Recall). Qualitatively, WordCloud showed positive sentiment was dominated by appreciation for instructors, while negative sentiment highlighted technical constraints. In conclusion, the model effectively reduces evaluation inefficiency and provides decision-supporting insights for management. Future research is recommended to apply the Synthetic Minority Over-sampling Technique (SMOTE) to address minority class bias.

Keywords : *Sentiment Analysis, TF-IDF, Multinomial Naïve Bayes, Feedback, Imbalanced*

ABSTRAK

Evaluasi umpan balik peserta di Proxsis Academy secara manual dinilai tidak efisien dan rentan bias. Penelitian ini bertujuan mengotomatisasi proses tersebut melalui Analisis Sentimen berbasis *Natural Language Processing* (NLP) menggunakan algoritma Multinomial Naïve Bayes. Dari 1.884 data awal, proses *cleansing* dan *preprocessing* (*case folding, tokenization, stopword removal, stemming*) menghasilkan 1.259 data teks valid yang kemudian diekstraksi menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Pengujian dengan rasio 70:30 menghasilkan Akurasi keseluruhan 99,21%, rata-rata tertimbang Presisi 98%, *Recall* 99%, dan *F1-Score* 99%. Meskipun akurasi tinggi, evaluasi mengungkap bias akibat ketidakseimbangan data (*imbalanced data*), di mana model memprediksi kelas Positif secara sempurna (*Recall* 100%) namun gagal pada kelas Negatif (*Recall* 0%). Secara kualitatif, *WordCloud* menunjukkan sentimen positif didominasi apresiasi terhadap instruktur, sementara sentimen negatif menyoroti kendala teknis operasional. Kesimpulannya, model ini efektif mereduksi inefisiensi evaluasi manual dan memberikan wawasan pendukung keputusan bagi manajemen. Untuk penelitian mendatang, disarankan penerapan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) guna mengatasi bias pada kelas minoritas.

Kata Kunci: Analisis Sentimen, TF-IDF, Multinomial Naïve Bayes, Umpan Balik, *Imbalanced Data.*

PENDAHULUAN

Perkembangan industri dan pelatihan profesional di Indonesia saat ini mengalami pertumbuhan yang sangat signifikan seiring dengan meningkatnya kebutuhan organisasi terhadap pengayaan kompetensi sumber daya manusia (SDM) yang berkelanjutan. Dalam ekosistem bisnis yang semakin kompetitif, lembaga pelatihan dituntut untuk tidak hanya mampu menyelenggarakan program yang relevan dengan tren dan kebutuhan industri, tetapi juga harus secara konsisten menjaga, mengevaluasi, dan meningkatkan kualitas layanan yang diberikan kepada para peserta (Al Lutfani et al., 2025). Salah satu mekanisme manajerial yang paling umum dan fundamental digunakan untuk mengukur kualitas pelatihan adalah pengumpulan umpan balik (feedback) dari peserta melalui instrumen evaluasi daring (online survey) (Fahima & Sumathipala, 2025; Rahman et al., 2025)

Proxsis Academy merupakan salah satu lembaga pelatihan profesional terkemuka yang aktif menyelenggarakan berbagai program krusial, mulai dari bidang manajemen risiko, sistem manajemen mutu, keamanan informasi, *Score Environmental, Social, and Governance* (ESG), hingga pengembangan kompetensi Teknologi Informasi (TI) lainnya. Dalam setiap kegiatan yang diselenggarakan, Proxsis Academy secara rutin mengumpulkan umpan balik dari peserta melalui formulir interaktif. Data umpan balik yang dikumpulkan mencakup penilaian kuantitatif terhadap berbagai dimensi kualitas pelatihan menggunakan parameter skala penilaian, serta menyediakan ruang bagi opini kualitatif melalui kolom teks bebas berupa "Kesan & Pesan". Berdasarkan pangkalan data operasional awal yang tersedia, telah terkumpul himpunan dataset ekstensif yang berisi sebanyak 1.884 ulasan peserta yang mengikuti beragam topik pelatihan di bawah bimbingan puluhan pemateri profesional.

Meskipun mekanisme pengumpulan data umpan balik teks telah berjalan secara rutin dan terstruktur, pemanfaatan data tersebut untuk mendukung pengambilan keputusan strategis oleh manajemen masih menghadapi kendala teknis. Selama ini, analisis terhadap ulasan teks bebas dilakukan secara manual atau hanya dibaca sepintas. Pendekatan observasi manual ini sangat tidak efisien, membutuhkan waktu ekstraksi yang lama, rentan terhadap bias dan subjektivitas penilai, serta tidak mampu mengekstraksi pola kecenderungan opini secara komprehensif dari volume data yang terus bertambah besar (Zhafira et al., 2021). Akibatnya, informasi krusial mengenai keluhan teknis, sentimen kepuasan peserta, maupun masukan konstruktif bagi para pemateri sering kali terlambat diidentifikasi, sehingga memperlambat langkah perbaikan mutu layanan.

Untuk mengatasi batasan operasional tersebut, dibutuhkan penerapan teknik *Data Mining*, khususnya pendekatan *Text Mining* dan *Sentiment Analysis* (Analisis Sentimen). Analisis sentimen memungkinkan sistem komputasi untuk mengenali, membedah, dan mengelompokkan ulasan teks secara otomatis ke dalam kategori polaritas opini positif atau negatif (Hidayat et al., 2025). Mengingat dimensi data teks yang tidak terstruktur (*unstructured data*) memiliki kompleksitas ruang fitur yang sangat tinggi, diperlukan sebuah algoritma klasifikasi yang berdaya komputasi tinggi namun tetap efisien. Salah satu algoritma klasifikasi *supervised learning* yang terbukti sangat efektif dalam pengolahan teks berdimensi tinggi adalah *Naïve Bayes Classifier* (NBC) yang dipadukan dengan pendekatan pembobotan matriks fitur *Term Frequency - Inverse Document Frequency* (Ariany et al., 2026; Kirana & Oktadini, 2025; Kusuma et al., 2026) . Algoritma Naïve Bayes bekerja berbasis teorema probabilitas bersyarat Bayes dan terbukti mampu menghasilkan metrik akurasi yang tangguh dengan beban komputasi yang relatif ringan dan terukur (Al Lutfani et al., 2025; Sudarsono et al., 2021).

Sejumlah penelitian terdahulu telah menerapkan analisis sentimen menggunakan Naïve Bayes dan metode sejenis pada ulasan publik di berbagai sektor, seperti opini publik di media sosial X terkait isu Pilkada Hidayat et al., (2025) dan kebijakan pemerintah Pramudita et al., (2024) komentar YouTube terkait kebijakan pendidikan Zhafira et al., (2021), ulasan aplikasi permainan digital di Play Store)Hisyam & Ayunda, (2026); Insan et al., (2023); Rahman et al., (2025), ulasan pelanggan pada platform(Al Lutfani et al., 2025) Anggraini & Wahyuni, (2025)e-commerce (Al Lutfani et al., 2025; Hasibuan & Heriyanto, 2022) hingga ulasan pengguna aplikasi media sosial X (Agustina & Herliana, 2025; Kirana & Oktadini, 2025; Tundo & Rachmawati, 2024). Namun demikian, tinjauan literatur menunjukkan bahwa sebagian besar penelitian tersebut sebatas mengukur performa komputasi algoritma tanpa memberikan kontrol penuh terhadap manipulasi *Natural Language Processing* (NLP) secara mendalam, serta belum melanjutkannya hingga tahap perumusan rekomendasi tindakan operasional yang konkrit. Terdapat *research gap* yang signifikan dalam mentransformasikan hasil label klasifikasi sentimen ulasan peserta menjadi kerangka strategi peningkatan kualitas layanan yang transparan, dapat direproduksi (*reproducible*), dan tervalidasi secara komputasional (Kirana & Oktadini, 2025; Putra & Ratnawati, 2025).

Oleh karena itu, penelitian ini dirancang untuk menerapkan algoritma Naïve Bayes Classifier dan pembobotan TF-IDF pada data umpan balik peserta Proxsis Academy dengan melakukan migrasi konseptual dan teknikal menuju ekosistem bahasa pemrograman Python di lingkungan Jupyter Notebook / Google Colab. Pemilihan Python didasarkan pada kebutuhan akademik akan ekosistem pustaka komputasi sains data yang komprehensif, seperti Pandas untuk manipulasi struktur data, NLTK dan Sastrawi untuk pemrosesan teks bahasa Indonesia yang akurat, serta Scikit-Learn untuk ekstraksi fitur TF-IDF dan inferensi *machine learning* (Kirana & Oktadini, 2025). Penelitian ini bertujuan untuk mengukur tingkat akurasi model klasifikasi serta menganalisis ulasan bersentimen guna menyusun rekomendasi peningkatan kualitas pelatihan yang konkrit dan berbasis data (*data-driven*) bagi manajemen Proxsis Academy.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen berbasis *Natural Language Processing* (NLP) untuk mengklasifikasikan umpan balik peserta pelatihan Proxsis Academy menggunakan algoritma Multinomial Naïve Bayes. Data penelitian diperoleh dari *Google Form* dan berjumlah 1.884 data dengan 22 atribut, kemudian difokuskan pada kolom "Kesan & Pesan" sebagai data teks utama. Tahapan pengolahan data meliputi seleksi data, *data cleansing*, pelabelan sentimen, *text preprocessing*, pembentukan fitur, pembagian data, pemodelan, serta evaluasi. Proses *data cleansing* dilakukan dengan menghapus data kosong, karakter atau simbol yang tidak diperlukan, angka, spasi berlebih, dan data duplikat. Pelabelan sentimen ditentukan berdasarkan rata-rata 12 kolom penilaian dengan skala 1–5, yaitu nilai ≥3 dikategorikan positif dan nilai <3 dikategorikan negatif. Selanjutnya, *text preprocessing* dilakukan melalui *case folding*, *tokenization*, penyaringan panjang token, *stopword removal*, dan *stemming* menggunakan pustaka Sastrawi, sehingga diperoleh 1.259 data valid. Data teks kemudian dikonversi menjadi fitur numerik menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) dengan konfigurasi *unigram*, *min_df*=2, dan *max_df*=0,90. Dataset dibagi menjadi 70% data pelatihan dan 30% data pengujian dengan *stratified random sampling* dan *random state* 42. Pemodelan dilakukan menggunakan algoritma Multinomial Naïve Bayes dengan

Laplace smoothing (*alpha*=1,0), sedangkan kinerja model dievaluasi menggunakan *confusion matrix* berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil klasifikasi selanjutnya divisualisasikan menggunakan *WordCloud* untuk mengidentifikasi kata-kata yang paling dominan dalam umpan balik peserta.

HASIL PENELITIAN DAN PEMBAHASAN

Hasil Pengumpulan dan Seleksi Data

Tahap awal penelitian ini adalah memuat dataset primer dari file Dataset_Proxsis_Academy.xlsx (sheet "Dataset Gabungan"). Berdasarkan hasil ekstraksi awal, dataset ini memiliki dimensi yang cukup besar, yakni 1.884 baris observasi dengan 22 kolom atribut. Karena fokus penelitian ini adalah analisis sentimen berbasis teks (*Text Mining*), proses seleksi fitur dilakukan dengan hanya mengekstraksi kolom "Kesan & Pesan". Pada tahap pengecekan nilai kosong (*missing values*), ditemukan 1 baris data yang tidak memiliki teks ulasan. Data kosong tersebut dieliminasi, sehingga hasil akhir seleksi data menyusut menjadi 1.883 baris.



Sumber: Penulis (2026)

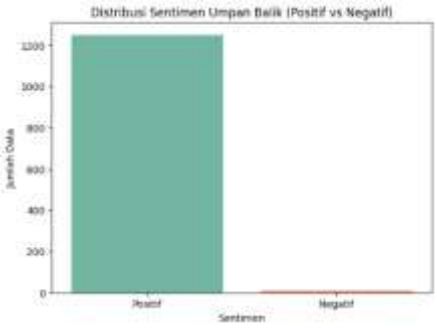
Gambar 1. Gambar *Screenshoot* file Data Mentah

Gambar I menunjukkan tampilan data mentah hasil pengumpulan umpan balik peserta pelatihan yang belum melalui proses pengolahan. Data terdiri dari beberapa atribut, seperti Timestamp, nama, perusahaan, jabatan, email, judul pelatihan, nama pemateri, serta beberapa nilai penilaian peserta terhadap pelaksanaan pelatihan. Data mentah ini selanjutnya akan melalui tahap seleksi dan data *cleansing* untuk memperoleh data yang relevan sebagai bahan analisis sentimen.

Hasil Data *Cleansing* dan Pelabelan Sentimen

Umpan balik yang diberikan oleh peserta pelatihan merupakan data teks tidak terstruktur yang sarat dengan noise. Oleh karena itu, data *cleansing* diimplementasikan menggunakan fungsi Regular *Expressions* (Regex) untuk membersihkan karakter non-ASCII (seperti emoji), tanda baca, simbol, angka, dan spasi berlebih. Setelah format teks dibersihkan, sistem mendeteksi adanya 327 baris data duplikat yang berpotensi menyebabkan bias pada saat pelatihan model. Setelah duplikasi dihapus, jumlah dataset bersih yang tersisa adalah 1.310 baris. Data bersih tersebut kemudian dipetakan dengan file referensi Selesi_Text_Sentimen.xlsx untuk mendapatkan

ground truth atau label aktual (Positif dan Negatif). Dari hasil pemetaan ini, diperoleh 1.309 baris data yang berhasil dilabel



Sumber: Penulis (2026)

Gambar 2 Gambar Distribusi Label Sentimen

Gambar 2 Gambar Distribusi Label Sentimen menunjukkan distribusi sentimen pada data umpan balik peserta pelatihan yang terdiri dari sentimen positif dan negatif. Berdasarkan grafik, sentimen positif memiliki jumlah data yang jauh lebih besar, yaitu 1.249 data, sedangkan sentimen negatif berjumlah 10 data. Hal ini menunjukkan bahwa sebagian besar peserta memberikan umpan balik yang bersifat positif terhadap pelaksanaan pelatihan. Perbedaan jumlah tersebut menunjukkan adanya ketidakseimbangan kelas (*class imbalance*) yang cukup ekstrem pada dataset, sehingga perlu diperhatikan pada tahap pemodelan agar hasil klasifikasi sentimen tidak terlalu didominasi oleh kelas positif.

Hasil *Preprocessing* Teks

Data berlabel (1.309 baris) diproses lebih lanjut melalui fungsi preprocess_text() guna menstandardisasi struktur linguistiknya. Tahapan ini menghasilkan teks yang seragam melalui proses Case Folding, Tokenization, penyaringan token berdasarkan panjang karakter (4 hingga 25 karakter), *Stopword Removal* menggunakan 757 daftar kata tugas bahasa Indonesia, serta Stemming menggunakan pustaka Sastrawi untuk mengembalikan kata ke bentuk dasarnya. Pemotongan imbuhan dan penghapusan kata hubung menyebabkan beberapa ulasan kehilangan seluruh elemen kata pembentuknya, menghasilkan teks kosong. Setelah dilakukan pembersihan data kosong (dropna), sebanyak 50 baris data terhapus. Hasil akhir dari tahapan ini adalah 1.259 baris data yang valid dan siap diekstraksi menjadi fitur.

	Text_Asli	Text_Clean	Sentimen	Text_Preprocessed
0	Sangat bagus, saran dilanjut dengan kunjungan ...	Sangat bagus saran dilanjut dengan kunjungan k...	Positif	bagus saran lanjut kunjung lapang lengsung ala...
1	Alhamdulilah penyampaian sudah baik	Alhamdulilah penyampaian sudah baik	Positif	alhamdulilah sampai
2	- Beberapa pertanyaan dari peserta tidak terja...	Beberapa pertanyaan dari peserta tidak terjawa...	Positif	serta jawab latsana lath
3	Terima kasih atas kesempatan yang diberikan \n	Terima kasih atas kesempatan yang diberikan	Positif	terima kasih sempat
4	Sangat bermanfaat	Sangat bermanfaat	Positif	manfaat
8	Materi cocoknya untuk enjiner instrumen , tida...	Materi cocoknya untuk enjiner instrumen tidak ...	Positif	materi cocok enjiner instrumen relevan kerja c...
11	Mohon untuk bach selanjutnya di sediakan lapto...	Mohon untuk bach selanjutnya di sediakan lapto...	Positif	mohon bach sedia laptop acces lancar gial lati...
12	Pemateri komunikatif, asik dan kompeten dibida...	Pemateri komunikatif asik dan kompeten dibidan...	Positif	materi komunikatif asik kompeten bidang lath ...
14	mantap luar biasa ilmunya sangat bermanfaat, a...	mantap luar biasa ilmunya sangat bermanfaat sa...	Positif	mantap ilmu manfaat moga sehat sukses
16	Terima Kasih	Terima Kasih	Positif	terima kasih

Sumber: Penulis (2026)

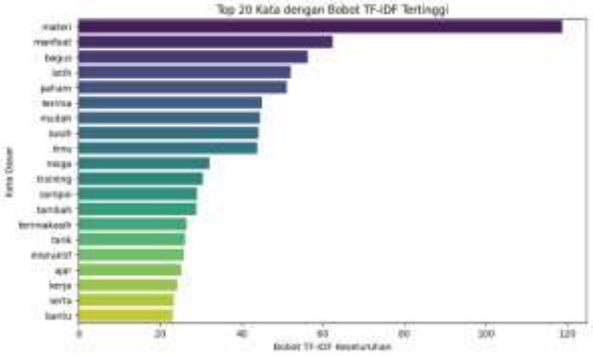
Gambar 3 Gambar *Screenshot* Hasil *Preprocessing* Text

Gambar 3 Gambar *Screenshot* Hasil *Preprocessing* Text menunjukkan hasil teks menjadi lebih seragam dan ringkas, seperti terlihat pada kolom Text_Preprocessed, sehingga kata-kata yang tidak relevan dapat dihilangkan dan kata dikembalikan ke bentuk dasarnya. Setelah

pembersihan data kosong, diperoleh 1.259 baris data valid yang siap digunakan pada tahap pembentukan fitur TF-IDF.

Hasil Ekstraksi Fitur (TF-IDF)

Representasi teks ke dalam ruang vektor numerik diimplementasikan menggunakan algoritma *Term Frequency-Inverse Document Frequency* (TF-IDF). Parameter yang diterapkan adalah Unigram (ngram_range=(1,1)), dengan pemangkasan fitur ekstrem melalui min_df=2 (mengabaikan kata yang muncul pada kurang dari 2 dokumen) dan max_df=0.90 (mengabaikan kata yang muncul pada lebih dari 90% dokumen). Penalaan hyperparameter ini terbukti krusial dalam menyaring istilah yang tidak relevan secara prediktif. Transformasi ini menghasilkan matriks fitur berdimensi (1.259, 562), yang menunjukkan bahwa 1.259 dokumen ulasan peserta berhasil direpresentasikan menjadi 562 fitur kata unik.



Sumber: Penulis (2026)

Gambar 4. Gambar Bobot TF-IDF Tertinggi

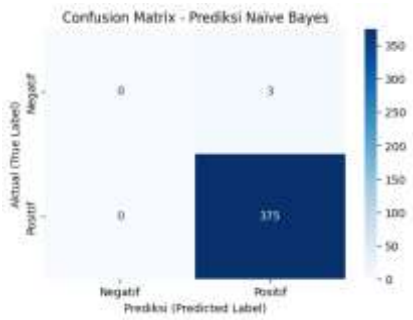
Gambar 4 Gambar Bobot TF-IDF Tertinggi pada data umpan balik peserta pelatihan. Kata "materi" memiliki bobot tertinggi, diikuti oleh "manfaat", "latih", "bagus", "paham", dan "ilmu", yang menunjukkan bahwa kata-kata tersebut memiliki tingkat kepentingan yang tinggi dalam merepresentasikan isi ulasan. Hasil ini menunjukkan bahwa fitur TF-IDF mampu mengidentifikasi kata-kata yang paling informatif dalam dataset dan selanjutnya digunakan sebagai fitur masukan dalam proses klasifikasi menggunakan algoritma Multinomial Naïve Bayes.

Hasil Pembagian Data dan Pemodelan

Dalam upaya mengevaluasi kemampuan generalisasi algoritma terhadap data yang belum pernah dilihat (unseen data), hasil dataset (1.259 baris) dipartisi menjadi data pelatihan dan pengujian dengan rasio 70:30. Penggunaan metode stratified random sampling (stratify=y) memastikan distribusi kelas Positif dan Negatif pada data uji mewakili proporsi yang sama dengan data latih, mencegah bias akibat ketidakseimbangan kelas. Hasil pembagian ini mengalokasikan 881 data untuk pelatihan (training) dan 378 data untuk pengujian (testing). Model klasifikasi diinduksi menggunakan algoritma Multinomial Naïve Bayes (MultinomialNB) dengan nilai smoothing Laplace (alpha=1.0). Hasil dari model ini adalah kalkulasi probabilitas frekuensi kemunculan fitur pada setiap kelas pelatihan, yang memberikan keunggulan berupa efisiensi komputasi yang tinggi pada matriks data teks.

Hasil Evaluasi Kinerja Model

Kinerja model MultinomialNB yang telah dilatih kemudian dievaluasi menggunakan data pengujian (378 baris). Alat ukur utama yang digunakan adalah Confusion Matrix, yang menghasilkan nilai metrik evaluasi berupa Akurasi, Presisi, Recall, dan F1-Score.



Sumber: Penulis (2026)

Gambar 5 Gambar Confusion Matrix

Gambar 5 Gambar Confusion Matrix menunjukkan hasil evaluasi model Naïve Bayes menggunakan confusion matrix. Dari hasil tersebut, sebanyak 375 data positif berhasil diprediksi dengan benar sebagai positif, sedangkan 3 data negatif diprediksi sebagai positif dan tidak terdapat data yang diprediksi sebagai negatif. Hal ini menunjukkan bahwa model mampu mengenali kelas positif dengan baik, tetapi masih mengalami kesulitan dalam mengidentifikasi sentimen negatif. Kondisi tersebut berkaitan dengan ketidakseimbangan jumlah data antara sentimen positif dan negatif, sehingga prediksi model cenderung lebih dominan pada kelas positif.

```
•••  Akurasi Keseluruhan (Accuracy) : 99.21%

    --- CLASSIFICATION REPORT ---
                  precision    recall  f1-score   support

         Negatif       0.00      0.00      0.00         3
         Positif       0.99      1.00      1.00       375

        accuracy                           0.99       378
       macro avg       0.50      0.50      0.50       378
    weighted avg       0.98      0.99      0.99       378
```

Sumber: Penulis (2026)

Gambar 6 Gambar Hasil Classification Report

Gambar 6 Hasil *classification report* menunjukkan bahwa model Naïve Bayes memperoleh akurasi sebesar 99,21% dari 378 data pengujian. Model memiliki performa sangat baik pada kelas positif dengan *precision* 0,90, recall 1,00, dan *F1-Score* 1,00, sedangkan kelas negatif memiliki nilai *precision, recall*, dan *F1-Score* sebesar 0,00 karena seluruh data negatif tidak berhasil diklasifikasikan dengan benar. Nilai macro average sebesar 0,57 menunjukkan bahwa performa model belum seimbang pada kedua kelas, sementara weighted average sebesar 0,99 menunjukkan performa keseluruhan yang cukup baik karena dipengaruhi oleh dominasi data positif. Hasil ini memperlihatkan bahwa meskipun akurasi model tinggi, model masih perlu ditingkatkan dalam mengenali sentimen negatif.

Hasil Visualisasi Sentimen (*WordCloud*)

Untuk memfasilitasi interpretasi kualitatif terhadap aspek apa saja yang paling diapresiasi maupun dikeluhkan oleh peserta pelatihan Proxsis Academy, hasil ekstraksi kata-kata kunci divisualisasikan melalui metode *WordCloud*. Ukuran tipografi pada visualisasi ini

merepresentasikan frekuensi kemunculan kata; semakin besar ukuran sebuah kata, semakin tinggi intensitas kemunculannya pada kelas sentimen tertentu.



Sumber: Penulis (2026)

Gambar 6. Gambar Hasil *WordCloud*

Gambar 6 Gambar Hasil *WordCloud* menunjukkan *WordCloud* dari umpan balik peserta Proxsis Academy, dengan kata "materi", "paham", "manfaat", "training", dan "terima kasih" sebagai kata yang paling dominan. Hal ini menunjukkan bahwa sebagian besar umpan balik peserta berfokus pada materi pelatihan, pemahaman peserta, manfaat yang diperoleh, serta apresiasi terhadap pelaksanaan training. Dominasi kata-kata tersebut memberikan gambaran bahwa umpan balik peserta cenderung positif dan menekankan manfaat serta pemahaman terhadap materi yang diberikan.

Pembahasan Hasil Penelitian

Berdasarkan hasil penelitian, implementasi *Natural Language Processing* (NLP) dengan kombinasi pembobotan TF-IDF dan algoritma Multinomial Naïve Bayes menunjukkan bahwa pendekatan tersebut mampu digunakan untuk mengolah dan mengklasifikasikan umpan balik teks peserta pelatihan Proxsis Academy. Dari data awal sebanyak 1.884 baris, setelah melalui tahapan seleksi, data cleansing, pelabelan sentimen, dan preprocessing, diperoleh 1.259 data valid yang kemudian direpresentasikan menjadi 562 fitur menggunakan TF-IDF. Hasil ini menunjukkan bahwa tahapan pengolahan data memiliki peranan penting dalam menghasilkan representasi teks yang lebih terstruktur dan mengurangi noise sebelum data digunakan dalam proses pemodelan.

Hasil pembentukan fitur TF-IDF juga menunjukkan bahwa kata-kata seperti materi, manfaat, latih, bagus, paham, dan ilmu memiliki bobot yang relatif tinggi, yang menggambarkan topik utama dalam umpan balik peserta. Temuan tersebut diperkuat oleh visualisasi *WordCloud* yang memperlihatkan dominasi kata seperti materi, paham, manfaat, training, dan terima kasih. Dominasi kata-kata tersebut menunjukkan bahwa sebagian besar umpan balik peserta berorientasi pada pengalaman positif, terutama berkaitan dengan materi pelatihan, pemahaman peserta, manfaat yang diperoleh, serta apresiasi terhadap pemateri dan pelaksanaan pelatihan.

Pada tahap evaluasi, model Multinomial Naïve Bayes menghasilkan akurasi sebesar 99,21% pada 378 data pengujian. Model mampu mengklasifikasikan 375 data positif dengan benar, tetapi seluruh 3 data negatif diklasifikasikan sebagai positif. Kondisi tersebut menyebabkan nilai precision, recall, dan *F1-Score* untuk kelas negatif masing-masing sebesar 0,00, sedangkan kelas positif memperoleh *precision* sebesar 0,99, recall sebesar 1,00, dan *F1-Score* sebesar 1,00. Dengan demikian, tingginya akurasi belum sepenuhnya menunjukkan bahwa model memiliki performa yang seimbang karena terdapat ketimpangan distribusi kelas antara sentimen positif dan

negatif. Model cenderung mengikuti pola kelas mayoritas sehingga kurang mampu mengenali karakteristik kelas negatif.

Ketidakseimbangan tersebut menjadi salah satu temuan penting dalam penelitian karena dapat menyebabkan bias dalam hasil klasifikasi. Oleh karena itu, evaluasi model tidak cukup hanya didasarkan pada nilai akurasi, tetapi perlu mempertimbangkan precision, recall, f1-score, dan confusion matrix untuk setiap kelas. Selain faktor distribusi kelas, keterbatasan Multinomial Naïve Bayes dalam merepresentasikan hubungan antarkata juga dapat memengaruhi kemampuan model dalam memahami konteks, ambiguitas, atau makna tertentu pada teks umpan balik. Meskipun demikian, penggunaan TF-IDF dan Multinomial Naïve Bayes tetap memberikan keunggulan dari sisi kesederhanaan dan efisiensi komputasi untuk proses klasifikasi teks.

Secara praktis, hasil penelitian menunjukkan bahwa analisis sentimen dapat membantu Proxsis Academy mengolah umpan balik peserta secara lebih sistematis dibandingkan dengan rekapitulasi secara manual. Hasil klasifikasi dan visualisasi kata dapat digunakan sebagai informasi pendukung untuk mengetahui aspek pelatihan yang paling sering dibicarakan peserta serta mengidentifikasi area yang memerlukan perhatian. Namun, agar model dapat memberikan hasil yang lebih representatif terhadap seluruh kelas sentimen, penelitian selanjutnya dapat mempertimbangkan teknik penyeimbangan data, seperti SMOTE, serta pengembangan metode *preprocessing* dan pemodelan yang mampu menangkap karakteristik teks secara lebih baik. Dengan demikian, hasil penelitian tidak hanya menunjukkan kemampuan TF-IDF dan Multinomial Naïve Bayes dalam melakukan klasifikasi, tetapi juga mengidentifikasi aspek penting yang perlu diperbaiki untuk meningkatkan keandalan analisis sentimen pada data umpan balik peserta pelatihan.

KESIMPULAN

Berdasarkan keseluruhan tahapan penelitian mulai dari pengumpulan data, prapemrosesan, hingga evaluasi model terkait Analisis Sentimen Umpan Balik Peserta Pelatihan Proxsis Academy menggunakan algoritma Multinomial Naïve Bayes berbasis Python, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Efektivitas Prapemrosesan Data (*Text Preprocessing*): Penerapan teknik pembersihan teks dan *preprocessing* termasuk *case folding*, *tokenization*, penghapusan karakter/simbol, *stopword removal*, dan *stemming* terbukti sangat krusial dalam menyaring *noise*. Dari total 1.884 baris data mentah observasi awal, proses ini berhasil menyaring dan menghasilkan 1.259 baris data teks bersih dan terstruktur yang tervalidasi serta siap diolah oleh algoritma *machine learning*.
2. Kinerja Algoritma Multinomial Naïve Bayes: Berdasarkan pengujian pada rasio data 70:30, model berhasil mencapai tingkat Akurasi keseluruhan sebesar 99,21%, dengan rata-rata tertimbang (*weighted average*) Presisi 98%, *Recall* 99%, dan *F1-Score* 99%. Namun, tingginya akurasi ini dipengaruhi oleh ketidakseimbangan data (*imbalanced data*) yang cukup ekstrem. Model terbukti sangat luar biasa dalam memprediksi kelas Positif (*Recall* 100%), namun gagal mengidentifikasi kelas minoritas (Negatif) dengan nilai *Recall* 0%. Hal ini mengonfirmasi bahwa meskipun Naïve Bayes sangat efisien untuk klasifikasi teks, algoritma ini sangat sensitif terhadap dominasi kelas mayoritas.
3. Wawasan Kualitatif (*Business Insight*): Visualisasi menggunakan *WordCloud* berhasil memetakan pola opini peserta secara intuitif. Pada kelas sentimen positif, dominasi kata seperti ["materi", "latih", "paham"] menunjukkan keberhasilan Proxsis Academy dalam aspek

penyampaian edukasi dan kompetensi pengajar. Sebaliknya, pada kelas sentimen negatif, kemunculan kata seperti ["waktu", "koneksi", "kurang"] merefleksikan bahwa keluhan peserta lebih banyak bersumber pada kendala operasional dan teknis selama pelatihan berlangsung.

4. Otomatisasi Evaluasi: Migrasi dan implementasi metode klasifikasi menggunakan bahasa pemrograman Python di lingkungan Google Colab berhasil mereduksi inefisiensi rekapitulasi data umpan balik yang sebelumnya dilakukan secara manual. Sistem ini memungkinkan pihak manajemen untuk melakukan evaluasi yang lebih cepat, objektif, dan terukur.

DAFTAR PUSTAKA

Agustina, V., & Herliana, A. (2025). *Analisis Sentimen Publik atas Kebijakan Efisiensi Anggaran 2025 dengan Text Mining dan Natural Language Processing*. https://doi.org/https://doi.org/10.55338/jumin.v6i3.6301

Al Lutfani, T. K., Astuti, R., Prihartono, W., & Hamonangan, R. (2025). Penerapan Naive Bayes Untuk Analisis Sentimen Pada Ulasan Pelanggan Di Lazada: Studi Kasus Toko Mawar Collection. *Jurnal Informatika Dan Teknik Elektro Terapan*, *13*(2). https://doi.org/10.23960/jitet.v13i2.6391

Anggraini, D., & Wahyuni, E. I. (2025). Analisis Sentimen Pengguna Terhadap Aplikasi Glints Di Google Play Store Menggunakan Metode Naive Bayes. *JMA)*, *3*(5), 3031–5220. https://doi.org/https://doi.org/10.62281/v3i5.1903

Apridiansyah, Y., Veronika, N. D. M., & Erwin, D. P. (2021). Prediksi Kelulusan Mahasiswa Fakultas Teknik Informatika Universitas Muhammadiyah Bengkulu Menggunakan Metode Naive Bayes. *JSAI : Journal Scientific and Applied Informatics*, *4*(2). https://doi.org/https://doi.org/10.36085/jsai.v4i2.1701

Ariany, F., Tri Prastowo, A., Ahmad, I., & Jafar Adrian, Q. (2026). *Algoritma Naïve Bayes dalam Analisis Komperatif Sentimen Dompet Digital* (Vol. 20, Number 1). https://doi.org/https://doi.org/10.33365/jtk.v20i1.1454

Fahima, M. M. F., & Sumathipala, P. (2025). Sentiment Analysis and its Application in the Education Sector: A Comprehensive Review. *Elementaria: Journal of Educational Research*, *3*(1), 105. https://doi.org/https://doi.org/10.61166/elm.v3i1.89

Firda, H., Putra, P., Oktadini, N. R., & Sevtiyuni, P. E. (2025). *Perbandingan Pelabelan Rating-based dan Inset Lexicon-based dalam Analisis Sentimen Menggunakan SVM (Studi Kasus: Ulasan Aplikasi GoBiz di Google Play Store)*. https://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/viewFile/4795/908

Hasibuan, E., & Heriyanto, E. A. (2022). Analisis Sentimen Pada Ulasan Aplikasi Amazon Shopping Di Google Play Store Menggunakan Naive Bayes Classifier. *JTS*, *1*(3). https://doi.org/https://doi.org/10.56127/jts.v1i3.434

Hidayat, T., Siddiq, M. J., Jayasri, S., Suhendi, A., & Rizky, R. (2025). Analisis Sentimen Opini Masyarakat Terhadap Pilkada 2024 Di Media Sosial Twitter Menggunakan Algoritma Naïve Bayes. *Jurnal Informatika Dan Teknik Elektro Terapan*, *13*(2). https://doi.org/10.23960/jitet.v13i2.6280

Hisyam, A. A., & Ayunda, A. T. (2026). Analisis Sentimen Persepsi Publik Terhadap Bank Dki Pada Twitter Menggunakan Metode Naive Bayes Classifier. *Jurnal Informatika Dan Teknik Elektro Terapan*, *14*(1). https://doi.org/http://dx.doi.org/10.23960/jitet.v14i1.8842

Insan, M. K., Hayati, U., & Nurdiawan, O. (2023). Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Number 1). https://doi.org/https://doi.org/10.36040/jati.v7i1.6373

Kirana, C. C., & Oktadini, N. R. (2025). Analisis Sentimen Naïve Bayes dengan TF-IDF dan 10-Fold pada Ulasan Aplikasi X. *Jurnal Sistem Komputer Dan Informatika (JSON)*, *7*(2), 523–535. https://doi.org/10.30865/json.v7i2.9007

Kusuma, W. T., Kurniawan, L. C., Rikatsih, N., & Wardoyo, P. (2026). Analisis Sentimen Ulasan Pengguna Aplikasi Akupunktur di Play Store Menggunakan Naïve Bayes. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *6*(2), 632–639. https://doi.org/https://doi.org/10.57152/malcom.v6i2.2620

Listiawan, D. A. B., Wicaksono, W. A., & Wiyanti. (2025). Analisis Sentimen Publik Pada Komentar Instagram Postingan Akun Tempodotco Terhadap Penerapan Kurikulum Ai Di Indonesia Menggunakan Metode Naïve Bayes. In *Jurnal Mahasiswa Teknik Informatika)* (Vol. 9, Number 5). https://doi.org/https://doi.org/10.36040/jati.v9i5.14934

Maulana, R., Voutama, A., & Ridwan, T. (2022). Analisis Sentimen Ulasan Aplikasi Mypertamina Pada Google Play Store Menggunakan Algoritma NBC. *Jurnal Teknologi Terpadu*. https://doi.org/https://doi.org/10.54914/jtt.v9i1.609

Pramudita, D., Akbar, Y., & Wahyudi, T. (2024). Analisis Sentimen Terhadap Program Kartu Indonesia Pintar Kuliah pada Media Sosial X Menggunakan Algoritma Naive Bayes. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*(4), 1420–1430. https://doi.org/https://doi.org/10.57152/malcom.v4i4.1565

Putra, A. R., & Ratnawati, D. E. (2025). Analisis Sentimen Berbasis Aspek pada Aplikasi Mobile Menggunakan Naïve Bayes berdasarkan Ulasan Pengguna Playstore (Studi Kasus : Jconnect Mobile). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *12*(2), 293–300. https://doi.org/https://doi.org/10.25126/jtiik.2025127556

Rahman, R. N., Rahim, A., & Pranoto, W. J. (2025). *Analisis Sentimen Ulasan Game eFootball 2024 Pada Playstore menggunakan Algoritma Naïve Bayes*. https://doi.org/https://doi.org/10.33884/jif.v13i01.9913

Sani, R. R., Pratiwi, A. Y., Winarno, S., Udayanti, E. D., & Zami, F. Al. (2022). *Analisis Perbandingan Algoritma Naive Bayes Classifier dan Support Vector Machine untuk Klasifikasi Hoax pada Berita Online Indonesia* (Vol. 13, Number 2). https://doi.org/https://doi.org/10.14710/jmasif.13.2.47983

Saputra, R. A., & Fatimah, T. (2025). *Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Buku Populer Berbasis Web* (Vol. 4, Number 2). https://senafti.budiluhur.ac.id/senafti/article/view/1834/911

Sejati, V. R. A., Painem, Ferdiansyah, & Pramusinto, W. (2023). Analisis Sentimen Ulasan Pengguna Aplikasi Ipusnas Pada Google Play Store Dengan Multinomial Naïve Bayes. *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, *2*(1). https://senafti.budiluhur.ac.id/senafti/article/view/534

Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). ANALISIS DATA MINING DATA NETFLIX MENGGUNAKAN APLIKASI RAPID MINER. *JBASE - Journal of Business and Audit Information Systems*, *4*(1). https://doi.org/10.30813/jbase.v4i1.2729

Thomas, V. W. D., & Rumaisa, F. (2022). Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, *6*(3), 1767. https://doi.org/10.30865/mib.v6i3.4218

Tundo, & Rachmawati, D. N. (2024). Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Terhadap Program Makan Siang Gratis. In *Jurnal Indonesia : Manajemen Informatika dan Komunikasi (JIMIK)* (Vol. 5, Number 3). https://doi.org/https://doi.org/10.35870/jimik.v5i3.978

Zhafira, D. F., Rahayudi, B., & Indriati. (2021). *Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes Dan Pembobotan Tf-Idf Berdasarkan Komentar Pada Youtube* (Vol. 2, Number 1). https://doi.org/https://doi.org/10.25126/justsi.v2i1.24