

Sentiment Analysis Of Public Reviews Of Public Services At The Dpmptsp Of Bengkulu Province Using The Naive Bayes Method

Khairullah ¹⁾; Muhammad Husni Rifqo ²⁾; Ujang Juhardi ³⁾; Na'im Mukarrom Nawwar ⁴⁾

^{1,2,3,4)} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Bengkulu

Email: ¹⁾ khairullah@umb.ac.id

How to Cite :

Khairullah, Rifqo, M., H., Juhardi, U., Nawwar, N., M. (2026). Sentiment Analysis of Public Reviews of Public Services at The Dpmptsp of Bengkulu Province Using The Naive Bayes Method. Jurnal Media Computer Science, 5(4).

ARTICLE HISTORY

Received [10 Agustus 2026]

Revised [25 Oktober 2026]

Accepted [27 Oktober 2026]

KEYWORDS

Sentiment Analysis, Public Services, Naive Bayes, TF-IDF, NLP.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentimen ulasan masyarakat terhadap pelayanan publik di DPMPTSP Provinsi Bengkulu menggunakan metode Naive Bayes. Data penelitian diperoleh dari ulasan masyarakat pada periode 2025 - 2026 yang kemudian diproses menggunakan teknik *Natural Language Processing* (NLP), meliputi *case folding*, *cleaning*, *tokenizing*, *stopword removal*, dan *stemming*. Selanjutnya data direpresentasikan menggunakan TF-IDF dan diklasifikasikan ke dalam sentimen positif, negatif, dan netral menggunakan algoritma Naive Bayes. Hasil penelitian menunjukkan bahwa metode Naive Bayes mampu mengklasifikasikan sentimen ulasan masyarakat dengan baik serta memberikan gambaran mengenai tingkat kepuasan dan persepsi masyarakat terhadap pelayanan yang diberikan. Hasil analisis ini diharapkan dapat menjadi bahan evaluasi bagi DPMPTSP Provinsi Bengkulu dalam upaya meningkatkan kualitas pelayanan publik.

ABSTRACT

This study aims to analyze public sentiment toward public services at DPMPTSP of Bengkulu Province using the Naive Bayes method. The research data were obtained from public reviews collected during the 2025–2026 period and processed using *Natural Language Processing* (NLP) techniques, including *case folding*, *cleaning*, *tokenizing*, *stopword removal*, and *stemming*. Furthermore, the data were represented using the TF-IDF method and classified into positive, negative, and neutral sentiments using the Naive Bayes algorithm. The results indicate that the Naive Bayes method is capable of classifying public review sentiments effectively and providing an overview of public satisfaction and perceptions regarding the services provided. The findings of this study are expected to serve as evaluation material for DPMPTSP of Bengkulu Province in its efforts to improve the quality of public services.

PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi pada era digital saat ini telah membawa perubahan yang sangat signifikan dalam berbagai aspek kehidupan masyarakat. Teknologi informasi tidak hanya dimanfaatkan dalam dunia bisnis dan pendidikan, tetapi juga telah menjadi bagian penting dalam penyelenggaraan pemerintahan, khususnya dalam meningkatkan kualitas pelayanan publik. Menurut Indrajit (2020), teknologi informasi merupakan kombinasi antara teknologi komputer dan teknologi komunikasi yang digunakan untuk mengolah, menyimpan, serta menyampaikan informasi secara cepat dan akurat. Pemanfaatan teknologi informasi memungkinkan instansi pemerintah untuk meningkatkan efektivitas, efisiensi, transparansi, dan akuntabilitas dalam memberikan pelayanan kepada masyarakat. Pelayanan publik merupakan salah satu fungsi utama pemerintah yang secara langsung berhubungan dengan kebutuhan masyarakat.

Kualitas pelayanan publik menjadi salah satu indikator keberhasilan suatu instansi pemerintah dalam menjalankan tugas dan tanggung jawabnya. Semakin baik kualitas pelayanan yang diberikan, maka semakin tinggi pula tingkat kepuasan masyarakat terhadap instansi tersebut. Sebaliknya, pelayanan yang kurang optimal dapat menimbulkan berbagai keluhan dan penilaian negatif dari masyarakat. Oleh karena itu, evaluasi terhadap kualitas pelayanan publik perlu dilakukan secara berkelanjutan agar instansi pemerintah dapat mengetahui kekurangan dan melakukan perbaikan yang diperlukan.

Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu (DPMPTSP) Provinsi Bengkulu merupakan salah satu instansi pemerintah yang memiliki peran penting dalam memberikan pelayanan kepada masyarakat, khususnya dalam bidang perizinan, investasi, dan pelayanan administrasi lainnya. Sebagai instansi yang berinteraksi langsung dengan masyarakat, DPMPTSP dituntut untuk mampu memberikan pelayanan yang cepat, mudah, transparan, dan profesional. Tingkat kepuasan masyarakat terhadap pelayanan yang diberikan menjadi salah satu ukuran keberhasilan instansi dalam menjalankan tugas dan fungsinya. Seiring dengan perkembangan teknologi digital, masyarakat kini tidak hanya menjadi penerima layanan, tetapi juga aktif memberikan tanggapan, kritik, saran, dan penilaian terhadap pelayanan yang mereka terima. Berbagai platform digital seperti Google Review, media sosial, dan website resmi instansi menjadi sarana bagi masyarakat untuk menyampaikan pengalaman mereka terkait pelayanan publik. Ulasan yang diberikan masyarakat dapat berupa opini positif, negatif, maupun netral yang mencerminkan tingkat kepuasan terhadap layanan yang diterima. Informasi yang terkandung dalam ulasan tersebut sangat berharga karena dapat digunakan sebagai bahan evaluasi dan dasar pengambilan keputusan dalam upaya meningkatkan kualitas pelayanan. Namun, jumlah ulasan masyarakat yang terus bertambah setiap waktu menyebabkan proses analisis secara manual menjadi kurang efektif. Selain membutuhkan waktu yang lama, analisis manual juga rentan terhadap subjektivitas dan kesalahan dalam penafsiran data. Data ulasan yang berbentuk teks tidak terstruktur memerlukan metode khusus agar dapat diolah menjadi informasi yang mudah dipahami dan digunakan dalam proses evaluasi. Oleh karena itu, diperlukan suatu teknik yang mampu mengolah data teks secara otomatis, cepat, dan sistematis. Salah satu teknik yang dapat digunakan untuk mengolah data teks adalah analisis sentimen (*sentiment analysis*). Menurut Asrumi, Suharijadi, Setiari, dan Wulanda (2023), analisis sentimen merupakan bidang studi yang menganalisis opini, sentimen, evaluasi, sikap, dan emosi masyarakat terhadap suatu objek, layanan, organisasi, maupun peristiwa tertentu. Analisis sentimen bertujuan untuk mengetahui kecenderungan opini masyarakat yang dapat dikategorikan ke dalam sentimen positif, negatif, dan netral. Melalui analisis sentimen, instansi pemerintah dapat memahami persepsi masyarakat secara lebih objektif berdasarkan data yang tersedia.

Dalam implementasinya, analisis sentimen memanfaatkan teknik *Natural Language Processing (NLP)* untuk mengolah data teks sehingga dapat diproses oleh algoritma klasifikasi. Salah satu algoritma yang banyak digunakan dalam analisis sentimen adalah Naive Bayes. Metode Naive Bayes merupakan metode klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes untuk menentukan kategori suatu data. Metode ini memiliki beberapa keunggulan, seperti proses perhitungan yang cepat, mudah diterapkan, mampu menangani data dalam jumlah besar, serta memiliki tingkat akurasi yang cukup baik dalam klasifikasi teks.

LANDASAN TEORI

Artificial Intelligence (AI)

Artificial Intelligence (AI) atau kecerdasan buatan merupakan salah satu cabang ilmu komputer yang berfokus pada pengembangan sistem atau mesin yang mampu meniru kemampuan berpikir manusia dalam menyelesaikan berbagai permasalahan. AI dirancang agar komputer dapat melakukan aktivitas yang umumnya membutuhkan kecerdasan manusia, seperti belajar, memahami bahasa, mengenali pola, mengambil keputusan, dan memecahkan masalah.

Natural Language Processing (NLP)

Natural Language Processing (NLP) adalah cabang dari ilmu komputer yang berfokus pada interaksi antara komputer dan bahasa manusia. Tujuannya adalah memungkinkan mesin untuk "memahami", menganalisis, dan menghasilkan bahasa yang digunakan oleh manusia secara alami. Dalam konteks analisis sentimen, NLP berperan sebagai fondasi utama dalam mengekstraksi makna dari teks menurut Hidayah, A. K., Sunardi, D., & Sahputra, E. (2025).

Analisis Sentimen

Menurut Asrumi, Suharijadi, D., Setiari, A. D., & Wulanda, D. P (2023) Analisis sentimen adalah bidang studi yang menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan emosi masyarakat terhadap entitas seperti produk, layanan, organisasi, individu, isu, peristiwa, topik, dan atributnya. Analisis sentimen biasanya dilakukan dalam konteks data teks, seperti ulasan produk, *tweet*, posting media sosial, artikel berita, dan lainnya. Tujuan analisis sentimen adalah untuk memahami bagaimana orang merasakan atau berpikir tentang suatu topik, produk, layanan, merek, atau peristiwa.

Aplikasi

Aplikasi adalah suatu perangkat lunak yang dikembangkan untuk menjalankan fungsi atau tugas tertentu pada berbagai jenis perangkat elektronik, seperti computer, smartphone, maupun tablet. Aplikasi berfungsi sebagai alat bantu pengguna dalam melakukan aktivitas digital secara lebih efisien dan terarah sesuai tujuan pembuatannya. (Sura & Fauzi, 2025).

METODE PENELITIAN

Metode Penelitian

Penelitian ini menggunakan jenis penelitian kuantitatif, pendekatan data mining, metodologi CRISP-DM, serta algoritma Naïve Bayes sebagai metode klasifikasi untuk menganalisis sentimen ulasan masyarakat terhadap pelayanan DPMPSTSP Provinsi Bengkulu. Instrumen yang digunakan dalam penelitian ini terdiri dari perangkat keras (*hardware*) dan perangkat lunak (*software*). Pada penelitian ini metode pengumpulan data dilakukan melalui beberapa teknik, yaitu observasi, studi pustaka, dan dokumentasi. Data penelitian diperoleh dari ulasan masyarakat pada periode 2025 - 2026 yang kemudian diproses menggunakan teknik *Natural Language Processing (NLP)*, meliputi *case folding*, *cleaning*, *tokenizing*, *stopword removal*, dan *stemming*. Selanjutnya data direpresentasikan menggunakan TF-IDF dan diklasifikasikan ke dalam sentimen positif, negatif, dan netral menggunakan algoritma Naive Bayes.

HASIL DAN PEMBAHASAN

Hasil

Hasil penelitian menunjukkan bahwa metode Naive Bayes mampu mengklasifikasikan sentimen ulasan masyarakat dengan baik serta memberikan gambaran mengenai tingkat kepuasan dan persepsi masyarakat terhadap pelayanan yang diberikan. Hasil analisis ini diharapkan dapat menjadi bahan evaluasi bagi DPMPSTSP Provinsi Bengkulu dalam upaya meningkatkan kualitas pelayanan publik.

Evaluasi Model

Pada klasifikasi yang digunakan dalam penelitian ini, data dibagi ke dalam dua kategori sentimen, yaitu Netral dan Positif. Fokus pembahasan meliputi evaluasi performa model menggunakan *confusion matrix*, *precision*, *recall*, dan *f1-score*. Skenario ini digunakan untuk melihat kemampuan algoritma Naive Bayes dalam membedakan ulasan yang bersifat informatif (Netral) dengan ulasan yang mengandung penilaian positif terhadap pelayanan publik DPMPSTSP Provinsi Bengkulu.

Gambar 1 Evaluasi Model

...	precision	recall	f1-score	support
Netral	1.00	0.22	0.36	23
Positif	0.81	1.00	0.89	75
accuracy			0.82	98
macro avg	0.90	0.61	0.62	98
weighted avg	0.85	0.82	0.77	98

Setelah proses klasifikasi sentimen menggunakan algoritma Naïve Bayes selesai dilakukan, tahap berikutnya adalah mengevaluasi performa model menggunakan *Classification Report*. Evaluasi ini bertujuan untuk mengetahui tingkat kemampuan model dalam mengklasifikasikan data sentimen berdasarkan beberapa metrik evaluasi, yaitu precision, recall, F1-score, dan accuracy. Seluruh proses evaluasi dilakukan menggunakan bahasa pemrograman Python pada platform Google Colab dengan memanfaatkan pustaka (library) Scikit-learn.

Pada Gambar 4.11 ditampilkan hasil *Classification Report* yang diperoleh dari proses pengujian model Naïve Bayes terhadap data uji (*testing data*). Laporan tersebut menyajikan nilai *precision*, *recall*, *F1-score*, dan *support* untuk masing-masing kelas sentimen, yaitu Netral dan Positif, serta nilai *accuracy*, *macro average*, dan *weighted average* sebagai ukuran performa model secara keseluruhan.

Berdasarkan hasil pada Gambar 4.11, kelas Netral memperoleh nilai *precision* sebesar 1,00, *recall* sebesar 0,22, dan *F1-score* sebesar 0,36 dengan jumlah data (*support*) sebanyak 23. Nilai *precision* yang mencapai 1,00 menunjukkan bahwa seluruh data yang diprediksi sebagai sentimen netral benar-benar merupakan data netral. Akan tetapi, nilai *recall* yang hanya sebesar 0,22 menunjukkan bahwa model masih belum mampu mengenali seluruh data netral yang terdapat pada data uji. Hal ini mengakibatkan nilai *F1-score* untuk kelas netral menjadi relatif rendah, yaitu sebesar 0,36.

Sementara itu, untuk kelas Positif, model memperoleh nilai *precision* sebesar 0,81, *recall* sebesar 1,00, dan *F1-score* sebesar 0,89 dengan jumlah data sebanyak 75. Nilai *recall* sebesar 1,00 menunjukkan bahwa seluruh data yang sebenarnya berlabel positif berhasil dikenali oleh model. Selain itu, nilai *F1-score* yang mencapai 0,89 menunjukkan bahwa model memiliki performa yang sangat baik dalam mengklasifikasikan sentimen positif. Hasil tersebut juga menunjukkan bahwa sebagian besar data pada penelitian ini berhasil diprediksi dengan benar sebagai sentimen positif.

Secara keseluruhan, model Naïve Bayes memperoleh nilai *accuracy* sebesar 0,82 atau 82%, yang menunjukkan bahwa model mampu mengklasifikasikan 82% dari seluruh data uji dengan benar. Nilai *macro average* menghasilkan *precision* sebesar 0,90, *recall* sebesar 0,61, dan *F1-score* sebesar 0,62, sedangkan *weighted average* menghasilkan *precision* sebesar 0,85, *recall* sebesar 0,82, dan *F1-score* sebesar 0,77. Nilai *weighted average* memberikan gambaran performa model dengan mempertimbangkan jumlah data pada masing-masing kelas sehingga lebih representatif terhadap kondisi dataset yang tidak seimbang.

Berdasarkan hasil evaluasi tersebut dapat disimpulkan bahwa algoritma Naïve Bayes memiliki performa yang baik dalam mengklasifikasikan sentimen masyarakat terhadap pelayanan publik DPMPSTSP Provinsi Bengkulu, terutama pada kelas Positif yang merupakan kelas dengan jumlah data terbanyak. Namun demikian, kemampuan model dalam mengenali sentimen Netral masih relatif rendah. Hal ini disebabkan oleh jumlah data netral yang lebih sedikit dibandingkan data positif sehingga model cenderung lebih mudah mengenali pola pada kelas positif. Meskipun demikian, tingkat akurasi sebesar 82% menunjukkan bahwa model Naïve Bayes mampu memberikan hasil klasifikasi yang cukup baik dan layak digunakan dalam analisis sentimen terhadap ulasan masyarakat mengenai pelayanan publik DPMPSTSP Provinsi Bengkulu.

Penghitungan Manual Evaluasi Model :

Netral: Precision = 1,00, Recall = 0,22, F1-Score = 0,36, Support = 23

Positif: Precision = 0,81, Recall = 1,00, F1-Score = 0,89, Support = 75

Accuracy = 0,82 (82%)

Perhitungan Accuracy

Rumus

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Substitusi nilai

$$Accuracy = \frac{5 + 75}{5 + 75 + 0 + 18}$$

$$= \frac{80}{98}$$

$$= 0,8163$$

$$= 81,63\% \approx 82\%$$

Keterangan:

- TP (Netral) = 5
- FN (Netral) = 18
- FP (Netral) = 0
- TN (Netral) = 75

Jumlah data: $5 + 18 + 0 + 75 = 98$

Perhitungan Precision Kelas Netral

Rumus

$$Precision = \frac{TP}{TP + FP}$$

$$= \frac{5}{5 + 0}$$

$$= 1,00$$

Artinya seluruh data yang diprediksi Netral memang benar Netral.

Perhitungan Recall Kelas Netral

Rumus

$$Recall = \frac{TP}{TP + FN}$$

$$= \frac{5}{5 + 18}$$

$$= \frac{5}{23}$$

$$= 0,2174$$

$$\approx 0,22$$

Perhitungan F1-Score Netral

Rumus

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

$$= 2 \times \frac{1,00 \times 0,22}{1,00 + 0,22}$$

$$= \frac{0,44}{1,22}$$

$$= 0,36$$

Perhitungan Precision Positif

Untuk kelas Positif

TP = 75

FP = 18

$$\begin{aligned} \text{Precision} &= \frac{75}{75 + 18} \\ &= \frac{75}{93} \\ &= 0,806 \\ &\approx 0,81 \end{aligned}$$

6. Perhitungan Recall Positif

$$\begin{aligned} \text{Recall} &= \frac{75}{75 + 0} \\ &= 1,00 \end{aligned}$$

7. Perhitungan F1-Score Positif

$$\begin{aligned} F1 &= 2 \times \frac{0,81 \times 1,00}{0,81 + 1,00} \\ &= \frac{1,62}{1,81} \\ &= 0,895 \\ &\approx 0,89 \end{aligned}$$

Analisis Wordcloud

Pada penelitian ini, klasifikasi sentimen dilakukan dengan mengelompokkan data ulasan masyarakat menjadi dua kategori, yaitu Positif dan Netral. Untuk mengetahui karakteristik setiap kategori sentimen, dilakukan visualisasi *WordCloud* berdasarkan frekuensi kemunculan kata pada hasil stemming. Visualisasi ini bertujuan untuk memberikan gambaran mengenai kata-kata yang paling sering muncul pada setiap kelas sentimen sehingga dapat diketahui pola bahasa yang digunakan oleh masyarakat dalam memberikan ulasan terhadap pelayanan publik Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu (DPMPTSP) Provinsi Bengkulu.

Selain digunakan untuk melihat karakteristik data, visualisasi *WordCloud* juga melengkapi hasil evaluasi model yang diperoleh melalui *Confusion Matrix*, *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Dengan demikian, penelitian ini tidak hanya menilai tingkat performa algoritma Naïve Bayes dalam mengklasifikasikan sentimen, tetapi juga memberikan gambaran mengenai kata-kata yang paling dominan pada setiap kategori sentimen.

Pada Gambar 4.12 ditampilkan hasil visualisasi *WordCloud* untuk kategori Sentimen Positif. Berdasarkan visualisasi tersebut, terlihat bahwa kata "baik" merupakan kata yang memiliki ukuran paling besar, yang menunjukkan bahwa kata tersebut memiliki frekuensi kemunculan tertinggi pada ulasan masyarakat. Selain itu, terdapat beberapa kata lain yang juga sering muncul, seperti "sangat", "layan", "amanah", "bagus", "cepat", "ramah", "tahan", "sopan", "terima kasih", "pegawai", "informasi", "prosedur", "surat izin", "muas", "tingkat", "moga", dan "semangat". Semakin besar ukuran suatu kata pada *WordCloud*, semakin tinggi pula frekuensi kemunculan kata tersebut pada data ulasan.

Dominasi kata "baik", "sangat", "layan", "ramah", "cepat", dan "amanah" menunjukkan bahwa sebagian besar masyarakat memberikan penilaian yang positif terhadap kualitas pelayanan yang diberikan oleh DPMPTSP Provinsi Bengkulu. Kata-kata tersebut menggambarkan bahwa masyarakat merasa pelayanan yang diberikan telah memenuhi harapan, baik dari segi keramahan petugas, kecepatan pelayanan, maupun kualitas pelayanan secara keseluruhan. Selain itu, kemunculan kata "terima kasih", "bagus", "muas", dan "semangat" menunjukkan adanya bentuk apresiasi dan kepuasan masyarakat terhadap pelayanan yang diterima.

Beberapa kata lain, seperti "informasi", "prosedur", "surat izin", "pegawai", dan "digital", menunjukkan bahwa ulasan masyarakat juga banyak membahas aspek pelayanan administratif yang diberikan oleh DPMPTSP Provinsi Bengkulu. Hal ini menunjukkan bahwa masyarakat tidak hanya memberikan penilaian terhadap sikap petugas, tetapi juga terhadap proses pelayanan, penyampaian informasi, kemudahan pengurusan perizinan, serta inovasi pelayanan yang diberikan.

Secara keseluruhan, hasil visualisasi WordCloud menunjukkan bahwa ulasan masyarakat didominasi oleh kata-kata yang memiliki makna positif. Dominasi kata-kata tersebut memperlihatkan bahwa sebagian besar masyarakat merasa puas terhadap pelayanan yang diberikan oleh DPMPTSP Provinsi Bengkulu. Hasil ini juga sejalan dengan proses pelabelan sentimen yang menunjukkan bahwa mayoritas data termasuk ke dalam kategori Positif. Dengan demikian, visualisasi WordCloud mampu memberikan gambaran yang jelas mengenai karakteristik sentimen positif sekaligus memperkuat hasil klasifikasi yang diperoleh menggunakan algoritma Naïve Bayes.

Gambar 2 Wordcloud Positif



Pada Gambar 2 ditampilkan hasil visualisasi WordCloud untuk kategori Sentimen Netral. Visualisasi ini menampilkan kumpulan kata yang paling sering muncul pada ulasan masyarakat yang diklasifikasikan ke dalam kelas netral. Ukuran setiap kata pada WordCloud menunjukkan tingkat frekuensi kemunculan kata tersebut, sehingga semakin besar ukuran kata maka semakin sering kata tersebut muncul pada data ulasan.

Berdasarkan hasil visualisasi tersebut, terlihat bahwa kata "tidak" merupakan kata yang memiliki ukuran paling besar, yang menunjukkan bahwa kata tersebut memiliki frekuensi kemunculan tertinggi pada kelompok sentimen netral. Selain itu, terdapat beberapa kata lain yang juga sering muncul, seperti "responsif", "belum", "comen", "camkoha", "mantapo", "baikk", "bagusss", "hebat", "sangan", "peleybaananan", "kerenn", "sarai", dan "komentar". Sebagian kata tersebut merupakan hasil penulisan masyarakat yang tidak baku atau mengandung variasi ejaan, namun tetap dipertahankan karena tidak seluruhnya dapat dinormalisasi oleh kamus normalisasi yang digunakan dalam penelitian.

Dominasi kata "tidak" dan "belum" pada kategori sentimen netral tidak selalu menunjukkan adanya sentimen negatif. Pada banyak ulasan, kedua kata tersebut digunakan dalam kalimat yang bersifat informatif, misalnya untuk menyampaikan kondisi tertentu, memberikan keterangan, atau menjawab pertanyaan tanpa disertai penilaian yang menunjukkan kepuasan maupun

ketidakpuasan terhadap pelayanan. Oleh karena itu, ulasan tersebut tetap dikategorikan sebagai sentimen netral berdasarkan hasil pelabelan menggunakan pendekatan *lexicon-based*.

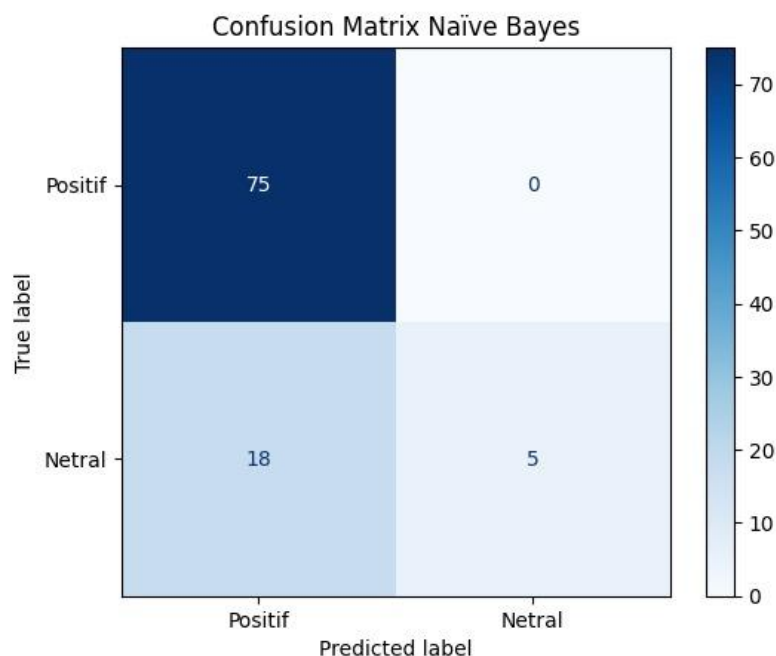
Selain itu, kemunculan kata seperti "camkoha", "comen", "mantapo", "baikkk", "bagusss", dan "kerenn" menunjukkan bahwa masih terdapat variasi penulisan kata yang digunakan oleh masyarakat dalam memberikan ulasan. Beberapa kata tersebut merupakan bentuk penulisan tidak baku, singkatan, atau kata yang mengalami pengulangan huruf sehingga tidak seluruhnya dapat diubah menjadi bentuk baku pada proses normalisasi. Meskipun demikian, kata-kata tersebut tetap menjadi bagian dari fitur yang digunakan dalam proses klasifikasi karena muncul pada data penelitian.

Secara umum, hasil visualisasi WordCloud menunjukkan bahwa karakteristik sentimen netral berbeda dengan sentimen positif. Pada kategori netral, kata-kata yang muncul cenderung bersifat informatif, berupa tanggapan singkat, komentar sederhana, atau pernyataan yang tidak menunjukkan kecenderungan emosi yang kuat terhadap kualitas pelayanan. Berbeda dengan sentimen positif yang didominasi oleh kata-kata seperti "baik", "ramah", "cepat", dan "terima kasih", kategori netral lebih banyak berisi kata-kata yang tidak secara langsung menggambarkan kepuasan ataupun ketidakpuasan masyarakat terhadap pelayanan yang diberikan.

Hasil visualisasi WordCloud ini mendukung proses klasifikasi yang dilakukan menggunakan algoritma Naïve Bayes. Perbedaan pola kemunculan kata antara sentimen positif dan sentimen netral memberikan informasi yang penting bagi model dalam membedakan kedua kategori sentimen. Dengan demikian, visualisasi WordCloud tidak hanya memberikan gambaran mengenai karakteristik data, tetapi juga memperkuat hasil analisis sentimen terhadap ulasan masyarakat mengenai pelayanan publik DPMPTSP Provinsi Bengkulu.

Analisis Confusion Matrix

Gambar 3 Confusion Matrix



Setelah proses klasifikasi sentimen menggunakan algoritma Naïve Bayes selesai dilakukan, tahap berikutnya adalah mengevaluasi performa model menggunakan Confusion Matrix. Confusion Matrix merupakan salah satu metode evaluasi yang digunakan untuk membandingkan hasil prediksi model dengan label aktual sehingga dapat diketahui jumlah prediksi yang benar maupun yang salah

pada setiap kelas sentimen. Melalui Confusion Matrix, performa model dapat dianalisis secara lebih rinci sebelum dihitung nilai accuracy, precision, recall, dan F1-score. Pada kelas Positif, seluruh data yang diprediksi sebagai Netral tidak ditemukan, sehingga menunjukkan bahwa model mampu mengenali karakteristik sentimen positif secara konsisten. Dominasi kata-kata seperti "baik", "sangat", "layan", "cepat", "ramah", dan "terimakasih" pada *wordcloud* turut memperkuat kemampuan model dalam membedakan sentimen positif dari kategori lainnya.

Secara keseluruhan, dari 98 data uji yang digunakan pada proses evaluasi, model berhasil mengklasifikasikan 80 data dengan benar, yaitu terdiri atas 75 data Positif dan 5 data Netral, sedangkan 18 data mengalami kesalahan klasifikasi. Hasil tersebut menunjukkan bahwa algoritma Naïve Bayes memiliki kemampuan yang baik dalam mengidentifikasi sentimen positif, namun masih memiliki keterbatasan dalam mengenali sentimen netral. Hal ini terlihat dari banyaknya data netral yang diprediksi sebagai sentimen positif.

Berdasarkan hasil Confusion Matrix tersebut dapat disimpulkan bahwa performa model Naïve Bayes pada penelitian ini cukup baik dalam mengklasifikasikan sentimen masyarakat terhadap pelayanan publik DPMPTSP Provinsi Bengkulu. Kemampuan model dalam mengenali kelas Positif sangat tinggi, sedangkan kemampuan dalam mengenali kelas Netral masih perlu ditingkatkan. Perbedaan performa tersebut dipengaruhi oleh distribusi data yang tidak seimbang, di mana jumlah data sentimen positif lebih banyak dibandingkan data sentimen netral.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian mengenai analisis sentimen terhadap pelayanan publik Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu (DPMPTSP) Provinsi Bengkulu menggunakan algoritma Naïve Bayes, dapat disimpulkan bahwa metode yang diterapkan mampu mengklasifikasikan sentimen masyarakat dengan baik. Penelitian ini menggunakan sebanyak 490 data ulasan yang diperoleh dari DPMPTSP Provinsi Bengkulu. Sebelum dilakukan proses klasifikasi, seluruh data terlebih dahulu melalui tahapan *preprocessing* yang meliputi *cleaning*, *case folding*, normalisasi, *tokenizing*, *stopword removal*, dan *stemming*. Tahapan tersebut berhasil menghasilkan data yang lebih bersih, seragam, dan siap digunakan pada proses pembobotan kata menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)*.

Proses pelabelan sentimen menggunakan pendekatan *lexicon-based* berhasil mengelompokkan data ulasan ke dalam dua kategori, yaitu Positif dan Netral. Hasil visualisasi *WordCloud* menunjukkan bahwa sentimen positif didominasi oleh kata-kata seperti *baik*, *layan*, *ramah*, *cepat*, dan *terima kasih* yang mencerminkan kepuasan masyarakat terhadap kualitas pelayanan yang diberikan. Sementara itu, sentimen netral didominasi oleh kata-kata yang bersifat informatif dan tidak menunjukkan kecenderungan emosi yang kuat terhadap pelayanan.

Berdasarkan hasil pengujian menggunakan *Confusion Matrix* dan *Classification Report*, algoritma Naïve Bayes memperoleh nilai *accuracy* sebesar 82%, yang menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengklasifikasikan sentimen masyarakat terhadap pelayanan publik DPMPTSP Provinsi Bengkulu. Model memiliki performa yang sangat baik dalam mengenali sentimen positif, namun masih mengalami keterbatasan dalam mengidentifikasi sentimen netral karena jumlah data pada kelas tersebut lebih sedikit dibandingkan kelas positif. Secara keseluruhan, hasil penelitian menunjukkan bahwa algoritma Naïve Bayes dapat digunakan sebagai metode klasifikasi sentimen masyarakat dan mampu memberikan informasi yang bermanfaat sebagai bahan evaluasi dalam upaya meningkatkan kualitas pelayanan publik di DPMPTSP Provinsi Bengkulu.

Saran

Berdasarkan hasil penelitian yang telah dilakukan, penulis memberikan beberapa saran yang diharapkan dapat menjadi masukan bagi penelitian selanjutnya maupun bagi DPMPTSP Provinsi Bengkulu. Penelitian selanjutnya disarankan menggunakan jumlah data ulasan yang lebih banyak dan lebih beragam agar model dapat mempelajari pola sentimen secara lebih optimal serta meningkatkan performa klasifikasi. Selain itu, penelitian berikutnya dapat membandingkan algoritma Naïve Bayes dengan metode klasifikasi lainnya, seperti *Support Vector Machine (SVM)*, *Random Forest*, atau metode berbasis *Deep Learning*, sehingga dapat diketahui algoritma yang memberikan hasil terbaik pada analisis sentimen pelayanan publik.

Di samping itu, proses pelabelan sentimen dapat dikembangkan dengan menggunakan kamus sentimen yang lebih lengkap atau melalui pelabelan secara manual oleh pakar agar hasil klasifikasi menjadi lebih akurat. Penelitian selanjutnya juga disarankan untuk mengatasi ketidakseimbangan jumlah data pada setiap kelas sentimen sehingga kemampuan model dalam mengenali seluruh kategori sentimen dapat ditingkatkan. Bagi DPMPTSP Provinsi Bengkulu, hasil penelitian ini diharapkan dapat dimanfaatkan sebagai bahan evaluasi untuk mempertahankan kualitas pelayanan yang telah dinilai baik oleh masyarakat serta melakukan perbaikan terhadap aspek pelayanan yang masih perlu ditingkatkan sehingga kualitas pelayanan publik dapat terus berkembang sesuai dengan harapan masyarakat.

DAFTAR PUSTAKA

- Asrumi, Suharijadi, D., Setiari, A. D., & Wulanda, D. P. (2023). *Analisis Sentimen dan Penggalian Opini*. Eureka Media Aksara.
- Hardiany, R., & Saputra, A. (2022). Penerapan Algoritma Naive Bayes untuk Klasifikasi Sentimen Masyarakat terhadap Kualitas Layanan Pemerintah. *Jurnal Informatika dan Teknologi Sistem Informasi*.
- Hidayat, T. (2025). *Pedoman Teknis Pengolahan Data Kuantitatif untuk Skripsi Teknik Informatika*. Andi Offset.
- Hidayah, A. K., Sunardi, D., & Sahputra, E. (2025). *Analisis Sentimen di Era Digital: Konsep, Algoritma, dan Studi Kasus Media Sosial*. PT Literasi Nusantara Abadi Grup.
- IBM Education. (2023). *What is Artificial Intelligence (AI)?* Tersedia di: <https://www.ibm.com/topics/artificial-intelligence>.
- Indrajit, R. E. (2020). *Manajemen Sistem Informasi dan Teknologi Informasi*. Prehallindo.
- Jurafsky, D., & Martin, J. H. (2020). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Stanford University.
- Liu, B. (2020). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press.
- Pratama, R. (2021). *Evaluasi Pelayanan Publik di Era Digital: Tantangan dan Peluang*. Universitas Muhammadiyah Bengkulu Press.
- Sari, D. P., & Kurniawan, D. (2024). Analisis Kepuasan Masyarakat Menggunakan Data Media Sosial dengan Pendekatan Machine Learning. *Jurnal Sistem Informasi dan Bisnis*.
- Wicaksono, A. (2023). *Natural Language Processing untuk Bahasa Indonesia: Teori dan Implementasi Preprocessing*. Informatika.
- Zhang, H., Li, J., & Liu, Y. (2020). Exploring the Performance of Naive Bayes in Text Classification. *Journal of Computer Science and Technology*.