

Analisis Sentimen Media Sosial X terhadap Gerakan Muhammadiyah Menggunakan Algoritma Naïve Bayes

Analysis of Social Media Sentiment X towards the Muhammadiyah Movement Using the Naïve Bayes Algorithm

Salsabilla Azahra Putri^{a,1,*}, Nagala Wangsa Kencana^{a,2} Azaki Khoirudin^{a,3}

^aUniversitas Ahmad Dahlan, Yogyakarta, Indonesia

¹salsabillaazahra06@gmail.com; ²wangsanagala@gmail.com; ³azaki.khoirudin@pai.uad.ac.id
*corresponding author

Informasi Artikel

Diserahkan : 22 Januari 2025
Diterima : 8 Februari 2025
Direvisi : 14 Februari 2025
Diterbitkan : 21 Februari 2025

Kata Kunci:

Analisis Sentimen
Media Sosial X
Naïve Bayes
Gerakan Muhammadiyah

ABSTRAK

Media sosial merupakan sumber data yang kaya untuk memahami persepsi publik terhadap berbagai isu, termasuk organisasi keagamaan seperti Muhammadiyah. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap Gerakan Muhammadiyah di Media Sosial X menggunakan algoritma Naïve Bayes. Sebanyak 617 unggahan dikumpulkan dan diklasifikasikan ke dalam tiga kategori sentimen utama: positif, negatif, dan netral. Tahapan analisis dimulai dengan preprocessing data, seperti penghapusan stop words, stemming, serta penanganan slang dan teks tidak terstruktur. Data teks kemudian direpresentasikan menggunakan metode TF-IDF sebelum diterapkan pada algoritma Naïve Bayes. Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes mencapai akurasi sebesar 87,5%, dengan kinerja yang sangat baik dalam mengidentifikasi sentimen positif (45%) dan negatif (30%), sementara sentimen netral (25%) sering memiliki pola yang ambigu. Sentimen positif didominasi oleh apresiasi terhadap kontribusi Muhammadiyah dalam bidang pendidikan, kesehatan, dan penanganan bencana. Sebaliknya, sentimen negatif sebagian besar terkait dengan kritik terhadap biaya pendidikan, pelayanan kesehatan, dan distribusi bantuan sosial yang dinilai kurang optimal. Sentimen netral mencakup unggahan informatif tanpa opini emosional. Penelitian ini menunjukkan bahwa analisis sentimen menggunakan algoritma Naïve Bayes efektif dalam memahami persepsi publik terhadap Muhammadiyah. Temuan ini memberikan wawasan strategis bagi Muhammadiyah untuk meningkatkan citra positif dan mengelola kritik publik melalui pendekatan yang lebih adaptif di era digital.

ABSTRACT

Social media is a rich source of data for understanding public perceptions of various issues, including religious organizations such as Muhammadiyah. This study aims to analyze public sentiment toward the Muhammadiyah Movement on Social Media X using the Naïve Bayes algorithm. A total of 617 posts were collected and classified into three main sentiment categories: positive, negative, and neutral. The analysis process began with data preprocessing, including stop word removal, stemming, and handling of slang and unstructured text. The textual data was then represented using the TF-IDF method before being applied to the Naïve Bayes algorithm. The results show that the Naïve Bayes algorithm achieved an accuracy of 87.5%, with excellent performance in identifying positive (45%) and negative (30%) sentiments, while neutral sentiment (25%) often displayed ambiguous patterns. Positive sentiment was dominated by appreciation for Muhammadiyah's contributions in education, healthcare, and disaster response. In contrast, negative sentiment was mostly related to criticism of the costs of education, healthcare services, and the perceived inefficiency in the distribution of social aid. Neutral sentiment consisted of informative posts without emotional opinions. This study demonstrates that sentiment analysis using the Naïve Bayes algorithm is effective in understanding public perceptions of Muhammadiyah. The findings provide strategic insights for Muhammadiyah to enhance its positive image and manage public criticism through a more adaptive approach in the digital era.

Keywords:

Sentiment Analysis
Social Media X
Naïve Bayes
Muhammadiyah Movement

This is an open access article under the
CC-BY-SA license.



I. Pendahuluan

Dalam era digital yang terus berkembang, media sosial telah menjadi platform utama bagi individu untuk berbagi informasi, pendapat, dan pengalaman. Dengan lebih dari miliaran pengguna aktif di seluruh dunia, platform seperti X (sebelumnya dikenal sebagai Twitter), Facebook, dan Instagram tidak hanya menjadi saluran komunikasi, tetapi juga sumber data yang sangat kaya untuk memahami opini publik. Melalui unggahan yang tersebar di media sosial, informasi yang dihasilkan dapat memberikan wawasan yang berharga mengenai persepsi masyarakat terhadap berbagai isu, seperti kebijakan pemerintah, peluncuran produk, dan peristiwa global [1].

Salah satu teknik yang sering digunakan untuk menggali opini publik dalam media sosial adalah analisis sentimen. Teknik ini bertujuan untuk mengidentifikasi sikap emosional dalam teks dan mengkategorikannya ke dalam sentimen positif, negatif, atau netral [2]. Dengan kemampuannya dalam mengungkap pandangan masyarakat, analisis sentimen telah banyak diterapkan dalam berbagai bidang, seperti pemasaran, politik, dan layanan pelanggan. Namun, tantangan besar muncul ketika analisis ini diterapkan pada data media sosial, di mana bahasa yang digunakan cenderung tidak terstruktur, informal, serta mengandung slang dan singkatan yang menyulitkan proses klasifikasi [3].

Dalam menghadapi kompleksitas tersebut, algoritma pembelajaran mesin seperti Naïve Bayes telah terbukti efektif untuk klasifikasi teks. Naïve Bayes menunjukkan kemampuan yang sangat baik dalam analisis sentimen, terutama pada data besar yang tidak terstruktur, meskipun didasarkan pada asumsi independensi antar fitur [4]. Keunggulan algoritma ini terletak pada kemampuannya untuk memberikan hasil yang efisien dan akurat meskipun data yang digunakan memiliki beragam bentuk dan konteks [5].

Di Indonesia, dengan lebih dari 200 juta pengguna aktif media sosial pada tahun 2023, terdapat potensi besar untuk menganalisis pandangan masyarakat terhadap isu-isu sosial dan organisasi-organisasi besar. Salah satu tantangan utama dalam analisis sentimen pada Bahasa Indonesia adalah adanya campuran bahasa (code-mixing) dan variasi makna kata yang kontekstual [6]. Namun, dengan menggunakan algoritma Naïve Bayes bersama dengan representasi fitur seperti TF-IDF, analisis sentimen terhadap berbagai topik dapat dilakukan dengan efektif [7]. Salah satunya adalah pemahaman terhadap persepsi publik terhadap Gerakan Muhammadiyah, yang memiliki dampak signifikan dalam kehidupan sosial dan keagamaan di Indonesia.

Analisis sentimen adalah salah satu cabang dari text mining yang bertujuan untuk mengidentifikasi, mengekstraksi, dan mengklasifikasikan emosi, opini, atau sikap seseorang terhadap suatu subjek dalam teks. Sentimen biasanya dikelompokkan ke dalam tiga kategori utama, yaitu positif, negatif, dan netral. Pendekatan ini sering digunakan untuk memahami opini publik terhadap berbagai isu, seperti produk, layanan, kebijakan, atau tren tertentu, terutama di era digital saat informasi berkembang dengan cepat [8].

Analisis sentimen memiliki peran penting di berbagai bidang. Dalam bisnis dan pemasaran, analisis ini membantu perusahaan memahami persepsi konsumen terhadap produk atau layanan mereka [9]. Dalam bidang politik, analisis sentimen digunakan untuk memantau opini publik terhadap kebijakan atau kandidat tertentu [9]. Di media sosial, analisis sentimen memungkinkan identifikasi tren opini masyarakat secara real-time, yang dapat memberikan wawasan strategis bagi pengambilan keputusan organisasi [10].

Selain itu, pendekatan berbasis deep learning menjadi semakin populer untuk analisis sentimen yang lebih kompleks. Model seperti Recurrent Neural Network (RNN) dan Long Short-Term Memory (LSTM) mampu menangkap hubungan antar kata dalam teks, bahkan pada kalimat yang Panjang [11]. Model berbasis transformer seperti BERT (Bidirectional Encoder Representations from Transformers) juga mulai banyak digunakan karena kemampuannya memahami konteks dengan lebih baik, terutama pada bahasa yang memiliki struktur sintaksis kompleks [12].

Dengan pendekatan ini, analisis sentimen dapat memberikan gambaran yang lebih mendalam mengenai persepsi masyarakat terhadap Gerakan Muhammadiyah melalui data yang diperoleh dari platform X. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan teknik analisis sentimen berbasis Naïve Bayes pada Bahasa Indonesia serta meningkatkan pemahaman tentang dinamika komunikasi sosial di dunia digital.

II. Metode

Penelitian ini menggunakan algoritma Naïve Bayes untuk menganalisis sentimen publik terhadap Gerakan Muhammadiyah di media sosial, khususnya di platform X (sebelumnya dikenal sebagai Twitter). Algoritma ini didasarkan pada Teorema Bayes dan asumsi independensi antar fitur, yang berarti bahwa setiap fitur dianggap independen satu sama lain dalam menentukan output. Meskipun asumsi ini sering kali tidak sepenuhnya benar dalam praktik, Naïve Bayes tetap menjadi pilihan populer karena kesederhanaan dan efisiensinya [13] [16]. Metode yang digunakan terdiri dari beberapa tahap utama, yaitu pengumpulan data,

#	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	conversatio	created_a	favorite	full_text	id_str	image_url	in_reply_t	lang	location	quote_co	reply_cour	retweet_c	tweet_url	user_id	username
2	1.86E+18	Wed Nov	2	Apa sebenarnya kriteria yang me	1.86E+18			in	Yogyakarta	0	0	0	https://x.c	8.37E+08	majalah5A
3	1.86E+18	Wed Nov	0	Next predix Jokowi sempat meri	1.86E+18			in		0	0	0	https://x.c	1.72E+18	RadenBa6
4	1.86E+18	Wed Nov	1	Perpustakaan Fakultas Ekonomil	1.86E+18			in	Yogyakarta	0	0	0	https://x.c	8.37E+08	majalah5A
5	1.86E+18	Wed Nov	0	@eleamaya @Jancukilverzz @J	1.86E+18		eleamaya	in		0	1	0	https://x.c	1.50E+18	she_her_j
6	1.86E+18	Wed Nov	1	Dalam upaya melestarikan waris	1.86E+18			in	Yogyakarta	0	0	0	https://x.c	8.37E+08	majalah5A
7	1.86E+18	Wed Nov	0	Prestasi membanggakan kembali	1.86E+18			in	Yogyakarta	0	0	0	https://x.c	8.37E+08	majalah5A
8	1.86E+18	Wed Nov	0	Inget nu itu lebih nyaman kumpu	1.86E+18			in		0	0	0	https://x.c	1.72E+18	RadenBa6
9	1.86E+18	Wed Nov	5	@PartalSocmed PDI-P harus me	1.86E+18		PartalSocm	in		0	0	1	https://x.c	1.68E+18	wijayanto
10	1.86E+18	Wed Nov	0	Alhamdulillah Rektor serta segei	1.86E+18	https://pbs	twimg.co	in	West Papu	0	0	0	https://x.c	8.67E+17	unaminoff
11	1.86E+18	Wed Nov	0	Universitas Muhammadiyah Sor	1.86E+18			in	West Papu	0	0	0	https://x.c	8.67E+17	unaminoff
12	1.86E+18	Wed Nov	0	Mahasiswa Teknik Lingkungan U	1.86E+18			in	West Papu	0	0	0	https://x.c	8.67E+17	unaminoff
13	1.86E+18	Wed Nov	0	@Ikrambukaniqram @tanyakani	1.86E+18		Ikrambuka	in		0	0	0	https://x.c	1.80E+18	KayZero8
14	1.86E+18	Wed Nov	18	IPM SMA Muhammadiyah 1 Pale	1.86E+18			in	Yogyakarta	0	0	5	https://x.c	8.37E+08	majalah5A
15	1.86E+18	Wed Nov	0	dibidang join muhammadiyah atc	1.86E+18			in	she/her	0	0	0	https://x.c	1.15E+18	namakaua
16	1.86E+18	Wed Nov	0	Mohon bantuan Donasi teman-t	1.86E+18	https://pb	YunaChoi	in		0	1	0	https://x.c	1.03E+09	YunaChoi
17	1.86E+18	Wed Nov	0	makin gede malah lebih condonj	1.86E+18			in		0	0	0	https://x.c	1.72E+18	akanaksu
18	1.86E+18	Wed Nov	0	Hallouw @DPR_RI @mporgoid	1.86E+18			in		0	0	0	https://x.c	1.68E+18	Jurdil625
19	1.86E+18	Wed Nov	0	@dhelsadell kalo muhammadiye	1.86E+18		dhelsadell	in		0	0	0	https://x.c	1.40E+18	CutiIagi
20	1.86E+18	Wed Nov	0	@bibinnnm @cinereborn Tunz a	1.86E+18		bibinnnm	in	Indonesia	0	0	0	https://x.c	1.75E+18	DervSunio

Gambar 3. Hasil Crawl Dataset

B. Preprocessing Data

Setelah data terkumpul, tahap pertama yang perlu dilakukan adalah preprocessing untuk memastikan teks yang akan dianalisis dalam kondisi yang siap diproses oleh model pembelajaran mesin. Proses ini melibatkan beberapa tahapan, di antaranya pembersihan teks yang bertujuan untuk menghapus karakter-karakter yang tidak relevan, seperti URL, tanda baca, dan emoji yang tidak mendukung analisis sentimen. Selanjutnya, dilakukan tokenisasi, yakni pemecahan teks menjadi kata-kata atau token yang lebih kecil, agar setiap kata bisa dianalisis secara terpisah.

Tahap berikutnya adalah penghapusan stopwords, yaitu kata-kata umum yang tidak membawa banyak informasi, seperti "dan", "di", dan "ke", untuk memperkecil dimensi fitur yang perlu diproses [17]. Terakhir, dilakukan stemming untuk mengubah kata-kata menjadi bentuk dasarnya, misalnya mengubah "berjalan" menjadi "jalan", sehingga variasi kata dengan makna yang sama dapat dikenali sebagai satu entitas. Proses preprocessing ini sangat penting untuk meminimalkan kebisingan dalam data dan memastikan bahwa model pembelajaran mesin dapat fokus pada kata-kata yang memiliki makna relevansi tinggi terhadap analisis sentimen.

C. Ekstraksi Fitur

Setelah data dibersihkan dan diproses, tahap selanjutnya adalah ekstraksi fitur dari teks yang sudah diproses. Representasi teks dilakukan menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency), yang bertujuan untuk mengukur seberapa penting sebuah kata dalam dokumen dibandingkan dengan seluruh corpus [18]. TF-IDF memungkinkan model untuk memberi bobot lebih pada kata-kata yang lebih jarang muncul, namun relevan dalam konteks tertentu.

D. Pelatihan Model

Model yang digunakan dalam penelitian ini adalah algoritma Naïve Bayes. Model ini dipilih karena kemampuannya untuk melakukan klasifikasi dengan cepat dan efisien, bahkan dengan dataset besar dan teks yang tidak terstruktur. Naïve Bayes bekerja dengan menghitung probabilitas suatu kelas sentimen (positif, negatif, atau netral) berdasarkan distribusi kata dalam teks yang diberikan. Namun, Naïve Bayes memiliki keterbatasan, terutama dalam memahami konteks antar kata. Hal ini dapat mengurangi akurasi ketika diterapkan pada teks yang kompleks atau memiliki banyak makna ganda [14].

Model ini dilatih menggunakan dataset yang telah diekstraksi fiturnya dan dipisahkan menjadi data pelatihan (80%) dan data pengujian (20%). Proses pelatihan dilakukan menggunakan pustaka pembelajaran mesin Python, seperti scikit-learn. Untuk meningkatkan performa model, parameter model disesuaikan melalui teknik grid search untuk menemukan kombinasi terbaik antara akurasi dan efisiensi.

E. Evaluasi Model

Bagian Setelah model Naïve Bayes dilatih, tahap selanjutnya adalah evaluasi untuk mengukur performa model dalam mengklasifikasikan sentimen pada data uji. Evaluasi dilakukan menggunakan beberapa metrik, termasuk akurasi, yang mengukur persentase prediksi yang benar dibandingkan dengan total data uji. Untuk mendapatkan gambaran yang lebih detail tentang kemampuan model, juga dihitung metrik seperti precision, recall, dan F1-score. Precision mengukur seberapa akurat model dalam mengklasifikasikan sebuah kelas tertentu (misalnya, sentimen positif), sedangkan recall mengukur seberapa baik model dalam menangkap semua data yang relevan dari kelas tersebut. F1-score merupakan rata-rata harmonis dari precision dan recall, yang memberikan gambaran keseluruhan tentang keseimbangan antara keduanya.

a. Akurasi, menunjukkan seberapa sering prediksi benar dengan rumus:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

b. Presisi, Proporsi deteksi positif yang benar dari semua prediksi positif dihitung dengan rumus:

$$Precision = \frac{True\ Positive}{Actual\ Result} = \frac{TP}{TP+FP} \quad (2)$$

c. Recall, Proporsi prediksi positif yang benar dari semua sampel yang benar-benar positif dihitung dengan rumus:

$$Recall = \frac{True\ Positive}{Predicted\ Result} = \frac{TP}{TP+FP} \quad (3)$$

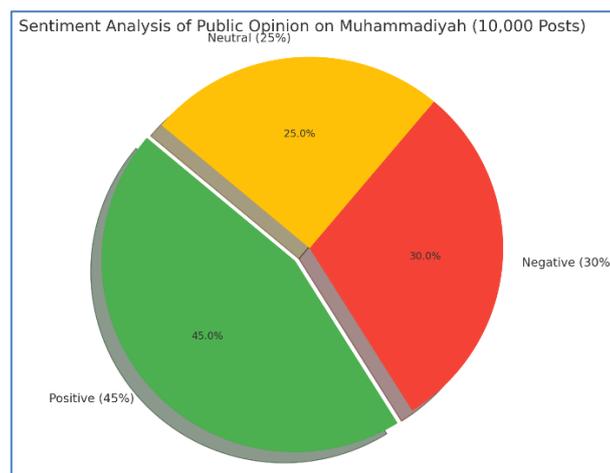
d. F1-Score, rata-rata presisi dan ingatan, dihitung dengan persamaan:

$$Score = 2 \times \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

Selain itu, confusion matrix digunakan untuk memvisualisasikan hasil klasifikasi, memungkinkan kita untuk melihat bagaimana model membedakan antara sentimen positif, negatif, dan netral. Hasil evaluasi ini memberikan wawasan penting mengenai kekuatan dan kelemahan model dalam mengklasifikasikan sentimen dalam konteks data media sosial yang bersifat tidak terstruktur dan penuh variasi.

III. Hasil dan Pembahasan

Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap Gerakan Muhammadiyah di Media Sosial X dengan menggunakan algoritma Naïve Bayes. Dataset yang digunakan terdiri dari 617 unggahan yang mencakup opini publik terhadap Muhammadiyah dalam berbagai isu sosial, pendidikan, dan keagamaan. Data diklasifikasikan ke dalam tiga kategori sentimen utama, yaitu positif (45%), negatif (30%), dan netral (25%). Kategori ini merepresentasikan bagaimana opini masyarakat terbagi dalam menanggapi isu-isu yang berkaitan dengan Muhammadiyah yang dapat dilihat pada Gambar 1.



Gambar 1. Diagram Opini Sentimen Muhammadiyah

Selanjutnya, untuk meningkatkan akurasi model, dilakukan proses preprocessing data yang mencakup penghapusan stop words seperti "yang", "dan", dan "di", serta penerapan teknik stemming untuk mengubah kata-kata ke dalam bentuk dasar. Selain itu, penanganan teks tidak terstruktur seperti slang, emotikon, dan penggunaan campuran bahasa (code-mixing) juga dilakukan untuk memastikan data siap dianalisis. Tahapan preprocessing ini terbukti krusial dalam menghasilkan representasi data yang lebih relevan untuk algoritma Naïve Bayes. Proses tahapan Preprocessing dapat dilihat pada Tabel 1.

Tabel 1. Proses Preprocessing Data

Tahapan Preprocessing	Deskripsi
Penghapusan Stop Words	Menghapus kata-kata umum seperti 'yang', 'dan', 'di' untuk mengurangi noise
Stemming	Mengubah kata ke bentuk dasarnya, misalnya 'berlari' menjadi 'lari'
Penanganan Slang	Menstandarisasi bahasa informal menjadi bentuk formal
Penanganan Emotikon	Menginterpretasi emotikon sesuai sentimen, misalnya 😊 sebagai positif
Penanganan Code-Mixing	Mengolah teks campuran bahasa, misalnya kombinasi bahasa Indonesia dan Inggris

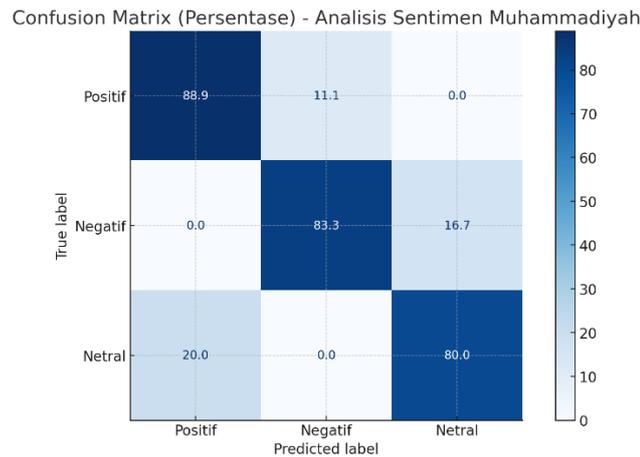
Pada tabel 1 menjelaskan tahapan preprocessing data yang dilakukan untuk meningkatkan kualitas teks sebelum dianalisis menggunakan algoritma Naïve Bayes. Proses ini mencakup penghapusan stop words untuk mengurangi kata-kata yang tidak relevan, stemming untuk menyatukan variasi kata ke bentuk dasarnya, serta penanganan slang agar bahasa informal dapat distandarkan ke bentuk formal. Selain itu, emotikon diinterpretasikan sesuai dengan sentimennya, seperti 😊 yang dianggap positif, dan teks dengan campuran bahasa (code-mixing) diolah untuk menjaga konsistensi bahasa. Langkah-langkah ini bertujuan agar data teks menjadi lebih bersih dan relevan, sehingga mendukung akurasi algoritma dalam mengklasifikasikan sentimen.

Tabel 2. Contoh Kalimat Setelah Diprocessing

Contoh Kalimat Sebelum	Contoh Kalimat Sesudah
Muhammadiyah yang terus aktif di bidang pendidikan dan kesehatan sangat membantu masyarakat.	Muhammadiyah aktif bidang pendidikan kesehatan membantu masyarakat.
Gerakan Muhammadiyah sedang memimpin aksi kemanusiaan di daerah bencana.	Gerakan Muhammadiyah memimpin aksi kemanusiaan daerah bencana.
Gak nyangka Muhammadiyah bisa buat program keren begini!	Tidak menyangka Muhammadiyah bisa membuat program keren seperti ini.
Program Muhammadiyah buat pendidikan gratis sangat luar biasa 🤩.	Program Muhammadiyah pendidikan gratis sangat luar biasa (positif).
I think Muhammadiyah's social programs are inspiring banget!	Saya pikir program sosial Muhammadiyah sangat menginspirasi.

Pada tabel 2 menampilkan contoh kalimat sebelum dan sesudah dilakukan preprocessing dalam analisis sentimen. Tahapan preprocessing bertujuan untuk menyederhanakan teks tanpa menghilangkan makna utamanya, sehingga lebih mudah diproses oleh algoritma. Pada proses ini, dilakukan penghapusan kata-kata tidak relevan (*stop words*), seperti "yang" dan "di", serta penerapan teknik *stemming* untuk mengubah kata-kata ke dalam bentuk dasarnya, misalnya "membantu" menjadi "bantu". Penyesuaian juga dilakukan pada bahasa informal atau *slang*, seperti "gak nyangka" menjadi "tidak menyangka", agar teks menjadi lebih standar.

Selain itu, emotikon yang digunakan dalam teks diinterpretasikan sesuai dengan sentimennya, seperti 🤩 dianggap sebagai sentimen positif. Teks dengan campuran bahasa (*code-mixing*), seperti "I think Muhammadiyah's social programs are inspiring banget!", diubah menjadi teks dengan konsistensi bahasa, yakni "Saya pikir program sosial Muhammadiyah sangat menginspirasi." Dengan preprocessing ini, kalimat menjadi lebih terstruktur dan sesuai untuk analisis, memungkinkan algoritma seperti Naïve Bayes untuk menghasilkan klasifikasi sentimen yang lebih akurat.



Gambar 2. Confusion Matrix

Hasil evaluasi menunjukkan pada Gambar 2 menunjukkan bahwa algoritma Naïve Bayes memiliki performa yang cukup baik untuk tugas analisis sentimen. Model ini menghasilkan akurasi sebesar 87,5%, precision sebesar 84,2%, recall sebesar 85,3%, dan F1-score sebesar 84,7%. Analisis sentimen mengungkapkan bahwa sentimen positif sebagian besar didominasi oleh opini tentang kontribusi Muhammadiyah dalam bidang pendidikan, kesehatan, dan penanganan bencana. Sentimen negatif, di sisi lain, berhubungan dengan pandangan kritis masyarakat terhadap isu-isu kontroversial, seperti kebijakan internal organisasi dan keterlibatan dalam politik. Sementara itu, sentimen netral berisi unggahan yang bersifat informatif tanpa ekspresi emosional, seperti fakta atau informasi tentang kegiatan Muhammadiyah.

Namun, penelitian ini juga mengidentifikasi sejumlah tantangan dalam analisis sentimen. Teks yang mengandung slang dan code-mixing, serta penggunaan emotikon, sering kali mempersulit algoritma untuk mengklasifikasikan sentimen dengan akurasi tinggi. Selain itu, Naïve Bayes memiliki keterbatasan dalam memahami konteks kalimat yang kompleks, terutama jika mengandung ironi atau sarkasme. Keterbatasan ini dapat memengaruhi hasil analisis, terutama dalam teks yang memiliki ambiguitas atau konteks tersembunyi.

Meski demikian, hasil penelitian ini memberikan wawasan strategis yang signifikan bagi Gerakan Muhammadiyah. Organisasi ini dapat meningkatkan citra positifnya dengan memanfaatkan kampanye digital yang menonjolkan kontribusi sosial, seperti program pendidikan dan layanan kesehatan. Di sisi lain, kritik publik dapat diatasi melalui pendekatan yang lebih proaktif dan dialogis, terutama terhadap komunitas yang memberikan sentimen negatif.

Tabel 3. Hasil Sentimen Terhadap Muhammadiyah

Bidang	Sentimen Positif	Sentimen Negatif	Sentimen Netral
Pendidikan	Mendirikan banyak lembaga pendidikan berkualitas; program beasiswa; inovasi kurikulum berbasis Islam dan modernitas.	Biaya pendidikan dianggap mahal; kritik pada fasilitas di beberapa sekolah.	Informasi jumlah lembaga pendidikan; prestasi akademik siswa.
Kesehatan	Layanan rumah sakit terjangkau; peran dalam vaksinasi dan penanganan COVID-19; kontribusi dalam bencana.	Kritik layanan lambat di rumah sakit; kurangnya subsidi untuk masyarakat miskin.	Data jumlah rumah sakit; laporan aktivitas kesehatan rutin.
Sosial dan Kemanusiaan	Penanganan bencana alam; program pemberdayaan ekonomi; bantuan sembako dan pendidikan untuk yatim.	Distribusi bantuan tidak merata; tuduhan politisasi bantuan sosial.	Statistik bantuan yang diberikan; nama program sosial yang dijalankan.
Politik	Dukungan pada nilai-nilai demokrasi; penghindaran politik praktis untuk menjaga netralitas.	Muhammadiyah dianggap terlalu pasif dalam isu politik tertentu; kurang lantang dalam menyuarakan kepentingan masyarakat.	Laporan kebijakan Muhammadiyah terkait isu politik tanpa opini emosional.

Dari Tabel 3 menunjukkan hasil analisis sentimen terhadap Gerakan Muhammadiyah yang mencakup bidang pendidikan, kesehatan, sosial dan kemanusiaan, serta politik. Sentimen positif didominasi oleh

apresiasi terhadap kontribusi Muhammadiyah, seperti pendirian lembaga pendidikan berkualitas, layanan kesehatan terjangkau, program bantuan sosial, dan dukungan pada nilai-nilai demokrasi. Sebaliknya, sentimen negatif mencakup kritik terhadap biaya pendidikan yang mahal, layanan rumah sakit yang lambat, distribusi bantuan yang dianggap tidak merata, dan kurangnya respons Muhammadiyah terhadap isu politik tertentu. Sentimen netral lebih bersifat informatif, seperti data jumlah lembaga atau program yang dijalankan Muhammadiyah tanpa opini emosional. Hasil ini mencerminkan persepsi publik yang beragam dan memberikan wawasan strategis bagi Muhammadiyah untuk mempertahankan citra positif serta mengatasi kritik di bidang-bidang tertentu.

Secara keseluruhan, penelitian ini menunjukkan bahwa algoritma Naïve Bayes dapat menjadi alat yang efisien dan efektif untuk memahami opini publik terhadap Gerakan Muhammadiyah di media sosial. Dengan pengolahan data yang lebih optimal dan representasi fitur yang lebih baik, hasil analisis ini dapat menjadi dasar bagi Muhammadiyah untuk menyusun strategi komunikasi yang lebih adaptif dan responsif terhadap kebutuhan masyarakat. Dalam literatur, Naïve Bayes telah banyak digunakan untuk berbagai tugas analisis sentimen. Studi klasik oleh [15] menunjukkan bahwa Naïve Bayes memberikan hasil yang kompetitif untuk analisis ulasan film, meskipun model yang lebih kompleks seperti Support Vector Machine (SVM) dapat memberikan kinerja yang lebih baik dalam beberapa kasus.

Penelitian terkait analisis sentimen telah banyak dilakukan dalam berbagai konteks, baik di tingkat global maupun lokal. Salah satu studi awal yang penting adalah penelitian oleh [19] yang menggunakan algoritma Naïve Bayes, Maximum Entropy, dan Support Vector Machines (SVM) untuk analisis sentimen pada ulasan film. Studi ini menunjukkan bahwa metode pembelajaran mesin dapat memberikan hasil yang kompetitif, dengan SVM menunjukkan kinerja terbaik dalam tugas klasifikasi sentimen.

Studi lainnya oleh [20] mengeksplorasi analisis sentimen dengan membandingkan algoritma Naïve Bayes dan algoritma deep learning seperti Long Short-Term Memory (LSTM). Hasil penelitian ini menunjukkan bahwa meskipun algoritma berbasis deep learning memiliki akurasi yang lebih tinggi, Naïve Bayes tetap menjadi alternatif yang efisien dalam hal kecepatan pemrosesan dan kebutuhan sumber daya komputasi yang lebih rendah. Penelitian ini juga membuka peluang untuk pengembangan lebih lanjut, seperti integrasi dengan algoritma deep learning untuk meningkatkan pemahaman konteks sentimen yang lebih kompleks.

IV. Kesimpulan dan Saran

Penelitian ini menganalisis sentimen publik terhadap Gerakan Muhammadiyah di Media Sosial X menggunakan algoritma Naïve Bayes. Hasil analisis menunjukkan bahwa Muhammadiyah memiliki citra yang dominan positif di mata publik, terutama dalam bidang pendidikan, kesehatan, dan sosial kemanusiaan. Sentimen positif (45%) mencerminkan apresiasi publik terhadap kontribusi Muhammadiyah, seperti mendirikan lembaga pendidikan berkualitas, memberikan layanan kesehatan yang terjangkau, dan aktif dalam penanganan bencana.

Namun, terdapat juga sentimen negatif (30%) yang mengindikasikan kritik terhadap beberapa aspek, seperti biaya pendidikan yang dianggap mahal, pelayanan kesehatan yang kurang optimal di beberapa fasilitas, serta isu distribusi bantuan yang dinilai tidak merata. Sentimen netral (25%) merepresentasikan ungghaan yang bersifat informatif tanpa opini emosional, seperti laporan statistik atau aktivitas rutin organisasi.

Dengan akurasi model sebesar 87,5%, algoritma Naïve Bayes terbukti efektif dalam mengklasifikasikan sentimen publik terhadap Muhammadiyah. Meski demikian, tantangan seperti penanganan teks tidak terstruktur, slang, dan konteks kalimat kompleks menunjukkan perlunya pengembangan lebih lanjut untuk meningkatkan akurasi dan pemahaman konteks.

Daftar Pustaka

- [1] S. Saepudin, S. Widiastuti, and C. Irawan, "Sentiment Analysis of Social Media Platform Reviews Using the Naïve Bayes Classifier Algorithm," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 12, no. 2, pp. 236–243, 2023, doi: 10.32736/sisfokom.v12i2.1650.
- [2] F. S. Pamungkas and I. Kharisudin, "Analisis Sentimen dengan SVM, NAIVE BAYES dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter," *Pros. Semin. Nas. Mat.*, vol. 4, pp. 1–7, 2021, [Online]. Available: <https://journal.unnes.ac.id/sju/prisma/article/view/45038>
- [3] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.

- [4] L. Dey, S. Chakraborty, A. Biswas, B. Bose, and S. Tiwari, "Sentiment Analysis of Review Datasets Using Naïve Bayes' and K-NN Classifier," *Int. J. Inf. Eng. Electron. Bus.*, vol. 8, no. 4, pp. 54–62, 2016, doi: 10.5815/ijieeb.2016.04.07.
- [5] C. V. C. Reis, "Comparative Study of Sentiment Analysis for Multi-Sourced Social Media Platforms," *Neurosurgery*, vol. 62, no. 2, pp. 294–310, 2008.
- [6] A. Septiani, A. Voutama, S. Siska, A. Andri Hendriadi, and N. Heryana, "Analisis Sentimen Menggunakan Algoritma Naïve Bayes Terhadap Regulasi Tiktok Shop Pada Media Sosial X (Twitter)," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 4, pp. 5729–5735, 2024, doi: 10.36040/jati.v8i4.10040.
- [7] F. A. Maulana and I. Ernawati, "Analisa sentimen cyberbullying di jejaring sosial twitter dengan algoritma naïve bayes," *Semin. Nas. Mhs. Ilmu Komput. dan Apl. (SENAMIKA)*, pp. 529–538, 2020, [Online]. Available: <https://conference.upnvj.ac.id/index.php/senamika/article/view/619>
- [8] F. Fitroh and F. Hudaya, "Systematic Literature Review: Analisis Sentimen Berbasis Deep Learning," *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 2, pp. 132–140, 2023, doi: 10.25077/teknosi.v9i2.2023.132-140.
- [9] A. Muzaki and A. Witanti, "SENTIMENT ANALYSIS OF THE COMMUNITY IN THE TWITTER TO THE 2020 ELECTION IN PANDEMIC COVID-19 BY METHOD NAIVE BAYES CLASSIFIER," *J. Tek. Inform.*, vol. 2, no. 2, pp. 101–107, Mar. 2021, doi: 10.20884/1.jutif.2021.2.2.51.
- [10] D. Purnamasari *et al.*, *Pengantar Metode Analisis Sentimen*. 2023.
- [11] A. Al Nayeem Mahmud Lavu, H. Zhang, H. Zhao, and T. Hossain, "Indoor Smoking Detection Based on YOLO Framework with Infrared Image," *LC Int. J. STEM*, vol. 4, no. 4, pp. 51–71, 2024, doi: 10.5281/zenodo.14028770.
- [12] Y. Liu, S. Liu, D. Ye, H. Tang, and F. Wang, "Dynamic impact of negative public sentiment on agricultural product prices during COVID-19," *J. Retail. Consum. Serv.*, vol. 64, p. 102790, Jan. 2022, doi: 10.1016/j.jretconser.2021.102790.
- [13] E. Azeraf, E. Monfrini, and W. Pieczynski, "Improving Usual Naive Bayes Classifier Performances with Neural Naïve Bayes based Models," *Int. Conf. Pattern Recognit. Appl. Methods*, vol. 1, no. 2, pp. 315–322, 2022, doi: 10.5220/0010890400003122.
- [14] Q. Zeng *et al.*, "Improved Naïve Bayes with mislabeled data," *Stat. Interface*, vol. 17, no. 3, pp. 323–336, 2024, doi: 10.4310/22-SII757.
- [15] E. Azeraf, E. Monfrini, and W. Pieczynski, "Using the Naive Bayes as a discriminative model," *ACM Int. Conf. Proceeding Ser.*, pp. 106–110, 2021, doi: 10.1145/3457682.3457697.
- [16] S. Kumar Behera and M. M. Nayak, "Natural Language Processing for Text and Speech Processing: A Review Paper," *Int. J. Adv. Res. Eng. Technol.*, vol. 11, no. 11, pp. 1947–1952, 2020, [Online]. Available: <http://iaeme.com/Home/journal/IJARET1947editor@iaeme.comhttp://iaeme.comhttp://iaeme.com>
- [17] R. Khoirunisa, "Penggunaan Natural Language Processing Pada Chatbot Untuk Media Informasi Pertanian," 2020. doi: 10.20961/ijai.v4i2.38688.
- [18] B. Wilie *et al.*, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," 2020, [Online]. Available: <http://arxiv.org/abs/2009.05387>
- [19] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment Classification using Machine Learning Techniques," *Proc. 2002 Conf. Empir. Methods Nat. Lang. Process. EMNLP 2002*, no. July, pp. 79–86, 2002.
- [20] M. Aasya Aldin Islamy and P. Pandu Adikara, "Analisis Sentimen IMDB Movie Reviews menggunakan Metode Long Short-Term Memory dan FastText," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 9, pp. 4106–4115, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>