

Perbandingan Algoritma *Naïve Bayes* Dan *Decision Tree* Untuk Analisis Sentimen Komentar *Youtube* Terhadap Tragedi Robohnya Musala Al-Khoziny

Comparison Of Naïve Bayes And Decision Tree Algorithms For YouTube Comment Sentiment Analysis On The Tragedy Of The Collapse Of The Musala Al-Khoziny

Nauradhia Shofarianti Putri*¹, Zaehol Fatah², Irma Yunita³

^{1,3}Teknologi Informasi, ²Sistem Informasi
Universitas Ibrahimy, Situbondo Indonesia
Email: ¹nauradhiasp23@gmail.com

Article Info:	Received 17 Mei 2026	Revised 17 Mei 2026	Accepted 08 Juni 2026	Published: 08 Juni 2026
----------------------	-------------------------	------------------------	--------------------------	----------------------------

Abstrak

Keterbatasan dalam penanganan ketidakseimbangan data serta minimnya perbandingan algoritma pada analisis sentimen berbasis kasus lokal menjadi tantangan dalam menghasilkan model yang akurat. Pada peristiwa robohnya musala Pondok Pesantren Al-Khoziny Sidoarjo, beragam opini publik di YouTube belum banyak dianalisis dengan pendekatan yang mempertimbangkan kondisi tersebut. Penelitian ini bertujuan untuk mengidentifikasi sentimen publik serta membandingkan kinerja algoritma *Naïve Bayes* dan *Decision Tree* pada data yang tidak seimbang. Data yang digunakan berupa kurang lebih 2.200 komentar yang dikumpulkan pada Januari–Maret 2026 melalui web scraping. Tahapan analisis meliputi preprocessing teks, pelabelan sentimen ke dalam kategori positif, negatif, dan netral, ekstraksi fitur menggunakan TF-IDF, serta penerapan SMOTE pada data latih untuk mengatasi ketidakseimbangan kelas. Model kemudian dilatih dan diuji menggunakan kedua algoritma tersebut. Hasil menunjukkan bahwa *Decision Tree* unggul dengan accuracy sebesar 0,99 dan F-1 Score sebesar 0,99, sedangkan *Naïve Bayes* memperoleh accuracy sebesar 0,68 dan F-1 Score sebesar 0,74. Temuan ini menunjukkan bahwa penanganan data tidak seimbang dan pemilihan algoritma berpengaruh signifikan terhadap hasil analisis sentimen. Temuan ini dapat menjadi acuan bagi pengembang sistem analisis sentimen yang bekerja dengan dataset berbahasa Indonesia, khususnya dalam konteks peristiwa sosial-keagamaan dengan distribusi kelas yang tidak seimbang.

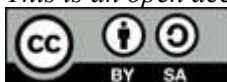
Kata Kunci: Analisis sentimen, YouTube, *Naïve Bayes*, *Decision Tree*, SMOTE, TF-IDF.

Abstract

Limitations in handling imbalanced data and the lack of algorithm comparisons in sentiment analysis based on local cases remain challenges in developing accurate models. In the case of the collapse of Musala at Pondok Pesantren Al-Khoziny Sidoarjo, public opinions expressed on YouTube have widely analyzed using approaches that consider these conditions. This study aims to identify public sentiment and compare the performance of the Naïve Bayes and Decision Tree algorithms on imbalanced data. The dataset consists of approximately 2,200 comments collected from January to March 2026 through web scraping. The analysis stages include text preprocessing, sentiment classification into positive, negative, and neutral categories, feature extraction using TF-IDF, and the application of SMOTE on the training data to address class imbalance. The models were then trained and tested on both algorithms. The results show that Decision Tree outperforms Naïve Bayes, achieving an accuracy of 0.99 and an F-1 Score of 0.99, while Naïve Bayes obtains an accuracy of 0.68 and an F-1 Score of 0.74. These findings indicate that handling imbalanced data and selecting appropriate algorithms significantly influence sentiment analysis results. The findings can serve as a reference for developing sentiment analysis systems using Indonesian-language datasets, particularly in the context of social media religious events with imbalanced class distributions.

Keywords: Sentiment Analysis, YouTube, Naïve Bayes, Decision Tree, SMOTE, TF-IDF.

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



1. PENDAHULUAN

Perkembangan teknologi digital mendorong meningkatnya penggunaan media sosial sebagai sarana penyebaran informasi dan opini publik. *YouTube* menjadi salah satu platform yang berpengaruh karena tidak hanya berfungsi sebagai media hiburan, tetapi juga sebagai ruang interaksi melalui fitur komentar, sehingga berpotensi menjadi sumber data dalam analisis opini masyarakat [1][2]. Tingginya intensitas interaksi serta banyaknya komentar yang muncul pada konten terkait peristiwa tertentu menjadikan platform ini relevan untuk mengkaji dinamika opini publik secara lebih mendalam. Tragedi robohnya Musala Pondok Pesantren Al-Khoziny di Sidoarjo memunculkan beragam tanggapan publik yang mencerminkan sentimen seperti simpati, kritik, dan opini netral, sehingga penting untuk dianalisis secara sistematis.

Dalam konteks peristiwa sosial-keagamaan, opini publik yang tersebar di media sosial mencerminkan persepsi, sikap, dan respons emosional masyarakat yang beragam. Pemahaman terhadap opini tersebut menjadi penting untuk melihat kecenderungan pandangan publik serta potensi dampaknya terhadap kehidupan sosial. Analisis sentimen sebagai bagian dari *Natural Language Processing (NLP)* digunakan untuk mengklasifikasikan opini ke dalam kategori tertentu, seperti positif, negatif, dan netral, dengan memanfaatkan algoritma machine learning seperti *Naïve Bayes* dan *Decision Tree* serta teknik ekstraksi fitur *TF-IDF*[3][4][5].

Sejumlah penelitian terdahulu menunjukkan bahwa algoritma *Naïve Bayes* dan *Decision Tree* memiliki kinerja yang cukup baik dalam analisis sentimen, namun performanya bergantung pada karakteristik data dan metode yang digunakan. Pada data Twitter, *Decision Tree* terbukti lebih unggul dengan akurasi 71,96% dibandingkan *Naïve Bayes* sebesar 66,11%[6][7]. Namun, penelitian tersebut belum mempertimbangkan penanganan ketidakseimbangan kelas yang dapat memengaruhi hasil klasifikasi. Sementara itu, pada data komentar YouTube, kedua algoritma mampu mengklasifikasikan sentimen secara efektif, meskipun *Naïve Bayes* cenderung lebih sederhana dan sensitif terhadap variasi bahasa, sedangkan *Decision Tree* lebih baik dalam menangkap pola data yang kompleks[9][10]. Akan tetapi, kajian tersebut masih terbatas pada kondisi data tertentu dan belum mengeksplorasi pengaruh distribusi kelas yang tidak seimbang

secara mendalam. Selain itu, penerapan teknik *SMOTE* terbukti dapat meningkatkan performa klasifikasi dengan mengatasi ketidakseimbangan data[11] namun implementasinya belum banyak dikaji pada konteks data komentar berbahasa Indonesia. Sementara itu, pada dataset berbahasa Indonesia, tantangan utama terletak pada penggunaan bahasa informal yang memengaruhi hasil analisis[12]. Pertama, sebagian penelitian belum secara khusus menangani permasalahan ketidakseimbangan kelas yang berpotensi menurunkan akurasi model. Kedua, kajian yang ada masih terbatas pada kondisi dan karakteristik data tertentu, sehingga belum mencerminkan kompleksitas distribusi data di dunia nyata. Ketiga, penelitian pada data berbahasa Indonesia masih menghadapi kendala bahasa informal dan belum banyak mengintegrasikan teknik penyeimbangan data seperti *SMOTE* secara optimal.

Berdasarkan gap tersebut, penelitian ini bertujuan untuk mengidentifikasi sentimen publik pada komentar *YouTube* terkait peristiwa sosial-keagamaan lokal, menerapkan teknik *SMOTE* untuk mengatasi ketidakseimbangan kelas pada data latih, serta membandingkan kinerja algoritma *Naïve Bayes* dan *Decision Tree* secara sistematis dalam kondisi data yang tidak seimbang. Penelitian ini juga memberikan kontribusi melalui pemanfaatan dataset berbahasa Indonesia yang merepresentasikan opini masyarakat secara kontekstual. Hasil penelitian ini diharapkan dapat memberikan gambaran yang lebih akurat mengenai kecenderungan opini publik serta menjadi dasar dalam pengembangan sistem analisis sentimen yang lebih adaptif terhadap karakteristik data lokal.

Penelitian ini memberikan kontribusi dalam tiga aspek utama, yaitu pemanfaatan dataset komentar *YouTube* berbahasa Indonesia yang berkaitan dengan peristiwa sosial-keagamaan lokal, penerapan teknik *SMOTE* untuk mengatasi ketidakseimbangan kelas pada data latih, serta perbandingan kinerja algoritma *Naïve Bayes* dan *Decision Tree* secara sistematis dalam kondisi data yang tidak seimbang.

Meskipun penelitian sebelumnya masih terbatas pada penggunaan dataset umum, minim konteks peristiwa lokal, serta belum optimal dalam menggabungkan penyeimbangan data dan perbandingan algoritma pada kondisi tidak seimbang, penelitian ini menghadirkan kebaruan melalui pemanfaatan dataset komentar *YouTube* pada peristiwa sosial-keagamaan, penerapan metode *SMOTE*, serta analisis perbandingan kinerja *Naïve Bayes* dan *Decision Tree* dalam konteks yang lebih spesifik[13]. Tujuan penelitian ini adalah menganalisis sentimen komentar *YouTube* terkait tragedi robohnya Musala Pondok Pesantren Al-Khoziny serta membandingkan kinerja kedua algoritma dalam mengklasifikasikan sentimen. Hasil penelitian diharapkan dapat memberikan gambaran opini publik sekaligus menjadi referensi bagi pengembangan metode analisis sentimen berbasis data media sosial.

2. METODE PENELITIAN

Penelitian ini mengadopsi pendekatan kuantitatif dan komparatif karena data yang dianalisis berupa metrik kinerja model seperti akurasi, presisi, dan recall, serta seluruh proses dilakukan secara terstruktur dan berbasis angka[14]. Metode penelitian mencakup langkah sistematis mulai dari pengumpulan data melalui scraping, pemrosesan teks, pelatihan model, hingga evaluasi kinerja, yang sejalan dengan pendekatan kuantitatif deskriptif dan inferensial dalam analisis sentimen berbasis data[15].

2.1 Teknik Pengumpulan Data

Data penelitian diperoleh dari sekitar 2.200 komentar *YouTube* terkait peristiwa robohnya Musala Pondok Pesantren Al-Khoziny Sidoarjo yang dikumpulkan pada Januari–Maret 2026 dari beberapa video relevan. Pengambilan data dilakukan menggunakan teknik scraping berbasis *Python* melalui *YouTube API* dengan bantuan library *google-api-python-client* dan *pandas*, kemudian disimpan dalam bentuk dataset terstruktur berbahasa Indonesia untuk tahap analisis selanjutnya[16].

2.1.1 Web Scraping YouTube Comments

Proses pengambilan data dilakukan dengan *Python* menggunakan *YouTube API* melalui *library google-api-python-client*. Data kemudian diolah dengan *pandas* dalam format terstruktur untuk memudahkan analisis. Teknik ini memungkinkan pengumpulan komentar secara otomatis, sehingga lebih efisien dan sistematis dibanding metode manual. Selain itu, penggunaan *API YouTube* mendukung ketersediaan data yang lebih lengkap dan konsisten, termasuk metadata yang diperlukan dalam analisis sentimen[17][18][19].

Data yang dikumpulkan berasal dari beberapa video *YouTube* yang relevan dengan peristiwa yang diteliti. Pemilihan video didasarkan pada tingkat keterkaitan dengan topik, jumlah komentar, serta tingkat interaksi pengguna. Selanjutnya, dilakukan proses penyaringan untuk memastikan bahwa komentar yang digunakan sesuai dengan kebutuhan analisis. Tabel berikut menunjukkan kriteria sumber data dan proses penyaringan komentar:

Tabel 1. Kriteria Sumber Data dan Penyaringan Komentar

ASPEK	KRITERIA
Jumlah Video	10 video
Kriteria Pemilihan Video	Relevan dengan peristiwa, memiliki jumlah komentar dan interaksi tinggi
Periode Data	Januari – Maret 2026
Bahasa Komentar	Bahasa Indonesia
Filter Komentar	Menghapus duplikasi, komentar kosong, spam, dan karakter tidak relevan
Relevansi	Komentar yang berkaitan langsung dengan peristiwa

2.2 Metode Pengembangan Sistem dan Alur Penelitian

Penelitian ini mengacu pada tahapan *Knowledge Discovery in Databases (KDD)* sebagai kerangka dalam mengolah dan menganalisis data[20][21]. Tahapan *KDD* tersebut diimplementasikan secara langsung dalam alur penelitian yang dilakukan. Pada tahap pemilihan data, dilakukan pengumpulan komentar *YouTube* yang relevan dengan peristiwa yang dikaji. Selanjutnya, tahap pembersihan diwujudkan melalui proses seleksi dan penghilangan noise untuk meningkatkan kualitas data.

Tahap transformasi data mencakup *preprocessing* teks, pelabelan sentimen ke dalam kategori positif, negatif, dan netral[22], serta ekstraksi fitur menggunakan *TF-IDF*[23]. *Preprocessing* dilakukan melalui beberapa tahapan, yaitu *case folding* untuk menyeragamkan penulisan huruf, *tokenizing* menggunakan *library NLTK* untuk memecah teks menjadi unit kata, *stopword removal* menggunakan daftar *stopword* Bahasa Indonesia untuk menghilangkan kata yang tidak memiliki makna penting, serta *stemming* menggunakan *Sastrawi* karena sesuai untuk memproses kata berbahasa Indonesia.

Tabel 2. Tahapan Preprocessing

TAHAP	HASIL
<i>Case Folding</i>	“Buat musala saja dah ambruk bagai mana pula mau buat mesjid?? Pokoknya proyeklah, orang mati ngak masalah yg penting proyek ku lancar guys??”
<i>Case Folding</i>	“buat musala saja dah ambruk bagai mana pula mau buat mesjid?? pokoknya proyeklah, orang mati ngak masalah yg penting proyek ku lancar guys??”
<i>Cleaning</i>	“buat musala saja dah ambruk bagai mana pula mau buat masjid br pokoknya proyeklah orang mati ngak masalah yg penting proyek ku lancar guys”
<i>Tokenizing</i>	["buat", "musala", "saja", "dah", "ambruk", "bagai", "mana", "pula", "mau", "buat", "masjid", "br", "pokoknya", "proyeklah", "orang", "mati", "ngak", "masalah", "yg", "penting", "proyek", "ku", "lancar", "guys"]

<i>Normalisasi</i>	["buat", "musala", "saja", "sudah", "ambruk", "bagaimana", "pula", "mau", "buat", "masjid", "biar", "pokoknya", "proyeklah", "orang", "mati", "tidak", "masalah", "yang", "penting", "proyek", "ku", "lancar", "teman"]
<i>Stopword Removal</i>	["musala", "ambruk", "masjid", "pokoknya", "proyeklah", "orang", "mati", "masalah", "penting", "proyek", "lancar"]
<i>Stemming</i>	["musala", "ambruk", "masjid", "pokok", "proyek", "orang", "mati", "masalah", "penting", "proyek", "lancar"]

TF-IDF digunakan untuk mengubah data teks menjadi representasi numerik berdasarkan bobot kata. Pada penelitian ini, *TF-IDF* diterapkan menggunakan parameter default dari *library scikit-learn*, sehingga setiap kata direpresentasikan berdasarkan frekuensi kemunculan dan tingkat kepentingannya dalam dokumen. Penggunaan parameter *default* dipilih untuk menjaga kesederhanaan model serta menghindari *overfitting* pada data yang relatif terbatas. Dataset kemudian dibagi menjadi data latih dan data uji dengan perbandingan 80:20 menggunakan teknik *stratified sampling* untuk menjaga proporsi kelas tetap seimbang. Untuk mengatasi ketidakseimbangan kelas, diterapkan metode *SMOTE* pada data latih setelah proses pembagian data. Tahap evaluasi dalam *KDD* diimplementasikan melalui proses pelatihan model *Naïve Bayes* dan *Decision Tree* menggunakan data latih, serta pengujian menggunakan data uji[24]. Hasil klasifikasi selanjutnya dievaluasi menggunakan *confusion matrix* dan metrik evaluasi, serta dibandingkan untuk menentukan model dengan performa terbaik[25].

Tabel 3. Distribusi Data Hasil Splitting

KELAS	TRAINING	TESTING	TOTAL
Negatif (0)	141	36	177
Positif (1)	45	11	56
Netral (2)	1536	384	1920
TOTAL	1722	431	2153

Alur penelitian ini divisualisasikan dalam bentuk diagram yang menunjukkan keterkaitan antar tahapan, mulai dari pengumpulan data, pembersihan, preprocessing dan pelabelan, transformasi data, penyeimbangan kelas, hingga tahap pelatihan, pengujian, dan evaluasi model. Dengan demikian, setiap tahapan *KDD* tidak hanya menjadi kerangka konseptual, tetapi juga diimplementasikan secara sistematis dalam proses penelitian yang dilakukan.



Gambar 1 Diagram Alur Penelitian

2.3 Pelabelan Sentimen

Dilakukan pelabelan data ke dalam tiga kategori sentimen, yaitu negatif (0), positif (1), dan netral (2) berdasarkan konteks dan kecenderungan opini pada setiap komentar. Proses pelabelan dilakukan secara manual dengan mempertimbangkan makna kalimat agar hasil klasifikasi lebih representatif. Untuk menjaga konsistensi, dilakukan pengecekan ulang pada sebagian data guna meminimalkan kesalahan dalam pemberian label. Hasil pelabelan ini kemudian digunakan sebagai dasar dalam proses pelatihan dan pengujian model. Distribusi hasil pelabelan menunjukkan adanya ketidakseimbangan kelas, di mana kelas tertentu memiliki jumlah data yang lebih dominan. Oleh karena itu, pada tahap selanjutnya diterapkan metode *SMOTE* untuk menyeimbangkan distribusi data sebelum dilakukan proses klasifikasi.

Tabel 4. Tahapan Labeling

No	Komentar	Label	Jumlah
1	buat musala saja dah ambruk bagai mana pula mau buat masjid br pokoknya proyeklah orang mati ngak masalah yg penting proyek ku lancar guys	0	177
2	sungguh luar biasa mukjizat tuhan	1	56
3	kontraktor pasti kena sangsi pidana perdata pidana	2	1920

3. HASIL DAN PEMBAHASAN

3.1 Transformasi

Tahap *transformasi* fitur mengubah data teks menjadi representasi numerik dengan metode *TF-IDF*, yaitu teknik pembobotan yang mempertimbangkan frekuensi kata dalam dokumen serta kepentingannya dalam dataset. Kata yang relevan terhadap konteks sentimen akan memiliki bobot lebih tinggi. Secara matematis, nilai *TF-IDF* dapat dihitung menggunakan rumus sebagai berikut:

$$TF-IDF(t, d) = TF(t, d) \times \log \left(\frac{N}{df(t)} \right) \quad 1$$

dengan t sebagai *term*, d dokumen, N jumlah dokumen, dan $df(t)$ jumlah dokumen yang memuat *term* tersebut. Hasilnya berupa vektor numerik sebagai input algoritma klasifikasi.

Tabel 5. Kata dengan Bobot *TF-IDF* Tertinggi

No	Kata	Bobot
1	bangun	0.0398
2	allah	0.0212
3	nya	0.0176
4	ponpes	0.0148
5	takdir	0.0144

Tabel 5 menunjukkan kata-kata dengan bobot *TF-IDF* tertinggi yang merepresentasikan topik dominan dalam komentar. Kata “bangun” dan “ponpes” mengindikasikan adanya pembahasan terkait proses pembangunan dan lokasi peristiwa, sedangkan kata “allah” dan “takdir” mencerminkan respons masyarakat yang bernuansa religius, seperti ungkapan keimanan, doa, maupun penerimaan terhadap kejadian yang terjadi. Sementara itu, kemunculan kata “nya” menunjukkan karakteristik bahasa informal dalam komentar yang masih sering digunakan oleh pengguna.

Secara keseluruhan, kata-kata dengan bobot tinggi tersebut menunjukkan bahwa opini publik tidak hanya berfokus pada aspek teknis peristiwa, tetapi juga dipengaruhi oleh perspektif sosial dan keagamaan yang kuat. Hal ini sesuai dengan konteks peristiwa yang terjadi di lingkungan pesantren, sehingga respons yang muncul cenderung mengandung nilai religius dan emosional. Untuk memperkaya analisis, visualisasi tambahan seperti *wordcloud* dapat digunakan untuk menggambarkan distribusi frekuensi kata secara lebih intuitif, sehingga pola kata yang dominan dalam dataset dapat lebih mudah diidentifikasi.

3.2 Penanganan Data Tidak Seimbang (*SMOTE*)

Dataset dalam penelitian ini memiliki distribusi kelas yang tidak seimbang, sehingga berpotensi menyebabkan model lebih dominan mengenali kelas mayoritas. Untuk mengatasi hal tersebut, diterapkan metode *SMOTE* yang menghasilkan data sintesis pada kelas minoritas berdasarkan pendekatan *k-nearest neighbors*. Implementasi dilakukan menggunakan *library imbalanced-learn (imblearn)* dengan parameter $k=5$. Melalui proses ini, distribusi data menjadi lebih seimbang sehingga model dapat mempelajari pola setiap kelas secara lebih optimal dan mengurangi bias terhadap kelas mayoritas.

Tabel 6. Distribusi Data Sebelum *SMOTE*

Kelas	Jumlah Data
Negatif (0)	177
Positif (1)	56
Netral (2)	1920
Total	2153

Tabel 7. Distribusi Data Setelah *SMOTE* (Data Latih)

Kelas	Jumlah Data
Negatif (0)	1536
Positif (1)	1536
Netral (2)	1536
Total	4608

3.3 Pembangunan Model Klasifikasi

3.3.1 Algoritma Naïve Bayes

Tabel 8 menampilkan hasil evaluasi *Naïve Bayes* dalam klasifikasi sentimen. Kelas negatif memperoleh *F-1 Score* 0,35, kelas positif hanya 0,16, sedangkan kelas netral lebih tinggi dengan 0,79. Secara keseluruhan, akurasi model berada di angka 0,68, menunjukkan performa yang kurang optimal terutama pada kelas minoritas, meskipun precision untuk kelas netral cukup tinggi.

Tabel 8. Hasil Naive Bayes

Kelas	Precision	Recall	F-1 Score
Negatif	0,23	0,75	0,35
Positif	0,10	0,36	0,16
Netral	0,96	0,68	0,79
Accuracy	0,68		

3.3.2 Algoritma Decision Tree

Tabel 9 menunjukkan performa *Decision Tree* dalam klasifikasi sentimen. Kelas negatif meraih *F-1 Score* 0,94, kelas positif 0,91, dan kelas netral hampir sempurna dengan 1,00. Secara keseluruhan, akurasi mencapai 0,99, menegaskan konsistensi dan efektivitas *Decision Tree* dalam mengenali semua kategori sentimen.

Tabel 9. Hasil Decision Tree

Kelas	Precision	Recall	F-1 Score
Negatif	0,97	0,92	0,94
Positif	0,91	0,91	0,91
Netral	0,99	1,00	1,00
Accuracy	0,99		

3.7 Evaluasi

Tabel 10 menunjukkan perbandingan performa *Naïve Bayes* dan *Decision Tree*. *Naïve Bayes* hanya mencapai akurasi 0,68 dengan *F-1 Score* 0,74, menandakan keterbatasan dalam mengenali kelas minoritas. Sebaliknya, *Decision Tree* unggul dengan akurasi, precision, recall, dan *F-1 Score* konsisten di 0,99, membuktikan kemampuannya menangani data tidak seimbang setelah transformasi *TF-IDF* dan penyeimbangan dengan *SMOTE*. Dengan demikian, tabel ini menegaskan bahwa *Decision Tree* lebih efektif dan sesuai digunakan dibandingkan *Naïve Bayes*.

Tabel 10. Hasil Evaluasi Model

Model	Accuracy	Precision	Recall	F-1 Score
Naïve Bayes	0,68	0,87	0,68	0,74
<i>Decision Tree</i>	0,99	0,99	0,99	0,99

3.8 Visualisasi Wordcloud

Visualisasi *wordcloud* menampilkan kata-kata paling sering muncul pada tiap kategori sentimen, dengan ukuran kata menunjukkan frekuensi komentar. Pemisahan *wordcloud* positif, negatif, dan netral memperjelas pola kata serta karakteristik opini pengguna. Pada sentimen negatif, kata dominan seperti *roboh*, *proyek*, dan *tanggung jawab* menunjukkan kritik, terbagi menjadi kritik struktural terhadap sistem/pihak terkait serta kritik deskriptif tentang kondisi fisik peristiwa. Pada sentimen positif, kata seperti *allah*, *moga*, dan *selamat* mencerminkan empati dan dukungan, berupa ungkapan religius/doa serta simpatik yang menunjukkan harapan dan kepedulian. Sementara itu, sentimen netral menampilkan kata seperti *bangun*, *lantai*, dan *ponpes* yang bersifat informatif tanpa emosi kuat, berupa istilah faktual dan deskriptif tentang kondisi tanpa penilaian emosional.



Gambar 2 Visualisasi *Wordcloud*

3.9 Analisis Hasil

Hasil penelitian menunjukkan bahwa *Decision Tree* jauh lebih unggul dibandingkan *Naïve Bayes* dalam klasifikasi sentimen komentar YouTube. *Decision Tree* mencapai akurasi 0,99 dengan precision, recall, dan *F-1 Score* yang konsisten tinggi, sedangkan *Naïve Bayes* hanya mencapai akurasi 0,68 dan menunjukkan kelemahan pada kelas minoritas. Perbedaan performa ini dipengaruhi oleh karakteristik data yang digunakan, yaitu didominasi oleh bahasa informal serta mengandung konteks sosial-keagamaan yang beragam. *Naïve Bayes* yang mengasumsikan independensi antar fitur cenderung kurang mampu menangkap hubungan kompleks antar kata dalam teks, terutama pada data dengan variasi bahasa yang tinggi, sedangkan *Decision Tree* lebih adaptif karena mampu memodelkan hubungan antar fitur dan menangkap pola non-linear dalam data. Selain itu, penerapan *TF-IDF* membantu merepresentasikan bobot kata berdasarkan tingkat kepentingannya dalam dokumen, sehingga kata-kata yang relevan terhadap sentimen memiliki kontribusi lebih besar dalam proses klasifikasi. Dukungan metode *SMOTE* juga berperan dalam meningkatkan performa model dengan menyeimbangkan distribusi data, sehingga model tidak bias terhadap kelas mayoritas. Meskipun demikian, nilai akurasi yang sangat tinggi pada *Decision Tree* berpotensi mengindikasikan terjadinya overfitting apabila model terlalu menyesuaikan diri dengan data latih. Untuk menjaga validitas hasil, penelitian ini menggunakan teknik stratified sampling dalam pembagian data serta hanya menerapkan *SMOTE* pada data latih setelah proses splitting, sehingga tidak terjadi data leakage antara data latih dan data uji dan hasil evaluasi tetap mencerminkan kemampuan generalisasi model terhadap data baru. Visualisasi *wordcloud* turut mendukung hasil analisis dengan menunjukkan perbedaan karakteristik kata pada setiap kategori sentimen, sehingga memperkuat bahwa *Decision Tree* lebih efektif dalam mengenali pola linguistik yang kompleks pada data komentar. Secara keseluruhan, hasil penelitian menunjukkan bahwa keberhasilan model tidak hanya dipengaruhi oleh algoritma yang digunakan, tetapi juga oleh karakteristik data serta teknik pengolahan data yang diterapkan.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa sentimen publik terhadap robohnya Musala Pondok Pesantren Al-Khoziny Sidoarjo didominasi oleh komentar netral, diikuti negatif, dan paling sedikit positif. *Decision Tree* terbukti lebih unggul dibandingkan *Naïve Bayes*, dengan akurasi 0,99 serta precision, recall, dan *F-1 Score* yang konsisten tinggi, sementara *Naïve Bayes* hanya mencapai akurasi 0,68 dan menunjukkan kelemahan pada kelas minoritas. Keunggulan *Decision*

Tree dipengaruhi oleh kemampuannya dalam membentuk aturan keputusan yang bersifat nonlinier serta menangkap hubungan kompleks antar fitur dalam data teks. Hal ini membuat *Decision Tree* lebih adaptif terhadap karakteristik data yang mengandung variasi bahasa informal dan konteks sosial-keagamaan. Sementara itu, *Naïve Bayes* yang mengasumsikan independensi antar fitur cenderung kurang mampu menangkap keterkaitan antar kata dalam teks. Selain itu, penerapan metode *SMOTE* berkontribusi dalam meningkatkan performa model dengan menyeimbangkan distribusi data, sehingga model tidak bias terhadap kelas mayoritas.

Penelitian ini juga memastikan validitas hasil dengan menerapkan pembagian data menggunakan stratified sampling serta hanya menerapkan *SMOTE* pada data latih setelah proses splitting untuk menghindari data leakage. Kontribusi penelitian ini terletak pada penggunaan dataset komentar YouTube berbahasa Indonesia dalam konteks sosial-keagamaan lokal serta penerapan teknik penyeimbangan data yang relevan untuk meningkatkan performa klasifikasi. Validasi lanjutan dengan k-fold cross-validation disarankan untuk memastikan kemampuan generalisasi model. Adapun keterbatasan penelitian meliputi jumlah data yang relatif terbatas, proses pelabelan manual, serta penggunaan algoritma yang masih sederhana. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar, metode yang lebih kompleks seperti deep learning, serta kategori sentimen yang lebih beragam.

KONFLIK KEPENTINGAN

Para penulis menyatakan bahwa tidak terdapat konflik kepentingan dalam penelitian ini, baik antara para penulis maupun dengan objek penelitian yang diteliti.

DAFTAR PUSTAKA

- [1] D. S. Rahmawati and N. Pratiwie, "Analyzing YouTube Comments on Predatory Pricing: A Digital Citizenship on Netnography Study," *Mediator: Jurnal Komunikasi*, vol. 18, no. 1, pp. 12–26, Jun. 2025, doi: 10.29313/mediator.v18i1.3347.
- [2] Y. A. Prasetyo, E. Utami, and A. Yaqin, "Pengaruh Komposisi Split Data Terhadap Performa Akurasi Analisis Sentimen Algoritma *Naïve Bayes* dan SVM," *Journal homepage: Journal of Electrical Engineering and Computer (JEECOM)*, vol. 6, no. 2, pp. 1–9, Oct. 2024, doi: 10.33650/jeeecom.v4i2.
- [3] T. H. Rochadiani, "Sentiment Analysis of YouTube Comments Toward Chat GPT," *Jurnal Transformatika*, vol. 21, no. 1, Aug. 2023, doi: 10.26623/transformatika.v21i2.7033.
- [4] Ardiansyah and Kurniawan, "Optimasi Metode *Naïve Bayes* Classifier Menggunakan Pendekatan Term Frequency-Inverse Document Frequency (*TF-IDF*) Pada Analisis Sentimen," *JSAI: Journal Scientific and Applied Informatics*, vol. 7, no. 3, pp. 1–7, Nov. 2024, doi: 10.36085.
- [5] B. Ariyadi Jasno, A. Ariful Fathoni, D. Dharma Putra, M. Zidane Hasan, F. Amsury, and H. Supendar, "ANALISIS SENTIMEN KOMENTAR BERPOTENSI TOXIC PADA MEDIA SOSIAL TIKTOK MENGGUNAKAN METODE *DECISION TREE*," vol. 16, p. 26207427, doi: 10.52972/hoaq.voll6no2.
- [6] R. Hidayat, A. B. Hakim, and R. Nugraha, "Perbandingan Metode *Naïve Bayes* Dan *Decision Tree* C4.5 untuk Analisis Sentimen Produk Es Teh Indonesia di Media Sosial Twitter," Feb. 2024.
- [7] D. Elreedy, A. F. Atiya, and F. Kamalov, "A theoretical distribution analysis of synthetic minority oversampling technique (*SMOTE*) for imbalanced learning," *Mach. Learn.*, vol. 113, no. 7, pp. 4903–4923, Jul. 2024, doi: 10.1007/s10994-022-06296-4.
- [8] N. Bayes, O. : Sulistia, M. Harahap, and R. Kurniawan, "Analisis Sentimen Komentar Youtube terhadap Food Vlogger dengan Menggunakan Metode Analisis Sentimen Komentar Youtube terhadap Food Vlogger dengan Menggunakan Metode *Naïve Bayes*," vol. 9, no. 1, [Online]. Available: http://ejournal.ust.ac.id/index.php/Jurnal_Means/

- [9] A. Wahid and G. Saputri, "Analisis Sentimen Komentar Youtube Tentang Relawan Patwal Ambulance Menggunakan Algoritma *Naïve Bayes* dan *Decision Tree*," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 2, p. 319, Dec. 2022, doi: 10.30865/json.v4i2.4941.
- [10] B. Franko, N. Wilyanto, and H. Irsyad, "Analisis Sentimen terhadap Naturalisasi Pemain pada Youtube Menggunakan *Decision Tree* dan *Naive bayes*," *Software Development, Digital Business Intelligence, and Computer Engineering*, vol. 3, no. 1, pp. 8–16, Sep. 2024, doi: 10.57203/session.v3i1.2024.8-16.
- [11] A. Milka Martin and A. Nur Rohman, "PENERAPAN ALGORITMA *NAÏVE BAYES* DENGAN TEKNIK *SMOTE* UNTUK KLASIFIKASI SENTIMEN KURSUS ONLINE SKILL ACADEMY," *Rabit : Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 11, no. 1, pp. 1922–1933, Feb. 2026, doi: 10.36341/rabit.v11i1.7545.
- [12] Y. Fauziah, B. Yuwono, and A. S. Aribowo, "Lexicon Based Sentiment Analysis in Indonesia Languages : A Systematic Literature Review," *RSF Conference Series: Engineering and Technology*, vol. 1, no. 1, pp. 363–367, Dec. 2021, doi: 10.31098/cset.v1i1.397.
- [13] D. Lado Kaka, G. Kopong Pati, and K. Wulla Rato, "Analisis Sentimen Komentar SIAKAD Menggunakan Metode *Naive Bayes Classifier*." [Online]. Available: www.tripadvisor.com.
- [14] S. Helmiyah and R. Pramestiawan, "Analisis Komparatif Algoritma Machine Learning dengan Metrik Akurasi, Presisi, Recall, dan *F-1 Score* pada Dataset Kacang Kering," *Jurnal Ilmu Komputer dan Teknologi*, vol. 6, no. 3, pp. 152–159, Oct. 2025, doi: 10.35960/ikomti.v6i3.2031.
- [15] M. A. Palomino and F. Aider, "Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis," *Applied Sciences (Switzerland)*, vol. 12, no. 17, Sep. 2022, doi: 10.3390/app12178765.
- [16] H. Santani Mulyono and U. Saprudin, "Efektivitas Logistic Regression dalam Analisis Sentimen Berbahasa Indonesia pada Komentar YouTube tentang Isu Ketenagakerjaan," Sep. 2025. doi: 10.63447.
- [17] D. J. Anggraini, "Analisis Sentimen Komentar Audiens terhadap Tayangan 'LIVE-Anies dan Efek Kejut di Pilkada Jakarta | ROSI' pada Kanal YouTube KOMPASTV: Studi tentang Persepsi Publik di Media Digital," *CARAKA: Indonesia Journal of Communication*, vol. 5, no. 2, pp. 130–141, doi: 10.25008/caraka.
- [18] M. Mubarak, L. Tanti, and R. Rosnelly, "Perbandingan Algoritma *Decision Tree* dan *Naive Bayes* Pada Analisis Sentimen Masyarakat Terhadap Pejabat Pertamina Pasca Kasus Pertamina Oplosan," *Jurnal Minfo Polgan*, vol. 15, no. 1, pp. 179–188, Mar. 2026, doi: 10.33395/jmp.v15i1.15971.
- [19] L. Rofiqi and M. Akbar, "Analisis Sentimen Terkait RUU Perampasan Aset dengan Support Vector Machine," *JEKIN - Jurnal Teknik Informatika*, vol. 4, no. 3, pp. 529–538, Aug. 2024, doi: 10.58794/jekin.v4i3.824.
- [20] W. Gata, "Perbandingan Analisis Sentimen Game Honor of Kings dengan VADER, TextBlob, Rating dan IndoBERT 1*," 2025.
- [21] B. F. S. Supriyanto and S. Rosalin, "Analisis Sentimen Program Merdeka Belajar dengan Text Analysis *Wordcloud* & Word Frequency," *Jurnal Minfo Polgan*, vol. 12, no. 1, pp. 25–32, Mar. 2023, doi: 10.33395/jmp.v12i1.12312.
- [22] Z. Fatah and L. Syarifah, "Analisis Sentimen Komentar YouTube pada Video Terkait Insiden Pengemudi Ojek Online dan Anggota Brimob Menggunakan Algoritma *Naive Bayes*."
- [23] M. Z. Naeem, F. Rustam, A. Mehmood, Mui-zzud-din, I. Ashraf, and G. S. Choi, "Classification of movie reviews using term frequency-inverse document frequency and

- optimized machine learning algorithms,” *PeerJ Comput. Sci.*, vol. 8, 2022, doi: 10.7717/PEERJ-CS.914.
- [24] L. D. Cahya, A. Luthfiarta, J. I. T. Krisna, S. Winarno, and A. Nugraha, “Improving Multi-label Classification Performance on Imbalanced Datasets Through *SMOTE* Technique and Data Augmentation Using IndoBERT Model,” *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 9, no. 3, pp. 290–298, Jan. 2024, doi: 10.25077/teknosi.v9i3.2023.290-298.
- [25] P. P. Armaeni, I. K. A. G. Wiguna, and W. G. S. Parwita, “Sentiment Analysis of YouTube Comments on the Closure of TikTok Shop Using *Naïve Bayes* and *Decision Tree* Method Comparison,” *Jurnal Galaksi*, vol. 1, no. 2, pp. 70–80, Aug. 2024, doi: 10.70103/galaksi.v1i2.15.