

ANALISIS SENTIMEN PEMILIHAN REMAJA TELADAN WILAYAH RATAHAN MENGUNAKAN ALGORITMA NAIVE BAYES

Paskal Victor Punusingon¹, Ade Yusupa², Victor Tarigan^{*3}

^{1,2,3}Informatika, Fakultas Teknik, Universitas Sam Ratulangi

Email: paskalpunusingonofficial@gmail.com, ade@unsrat.ac.id, victortarigan@unsrat.ac.id

SEJARAH ARTIKEL

Diterima: 07.06.2021

Direvisi: 05.07.2021

Diterbitkan : 30.07.2025



Hak Cipta © 2025
Penulis: Ini adalah
artikel akses terbuka
yang didistribusikan
berdasarkan ketentuan
Creative Commons
Attribution 4.0
International License.

ABSTRAK

Penelitian ini menganalisis sentimen publik terhadap "Pemilihan Remaja Teladan GMIM Wilayah Ratahan" menggunakan algoritma Naive Bayes. Data kuesioner terbuka dikumpulkan dari peserta dan pemangku kepentingan, kemudian diproses melalui tahapan pra-pemrosesan teks meliputi case folding, penghapusan tanda baca, dan tokenisasi. Sebagian kecil dari respons dilabeli secara manual ke dalam kategori positif, netral, dan negatif untuk melatih model Naive Bayes. Model yang telah dilatih kemudian digunakan untuk mengklasifikasikan data sentimen yang tersisa. Hasil penelitian mengungkapkan distribusi sentimen secara keseluruhan dan menyoroti aspek-aspek spesifik dari proses pemilihan yang mendapatkan umpan balik positif, netral, atau negatif, seperti pelaksanaan program, transparansi, dan dampaknya terhadap perkembangan remaja. Studi ini memberikan wawasan berharga untuk perbaikan inisiatif pengembangan remaja di organisasi keagamaan di masa mendatang.

Kata Kunci: analisis sentimen, naive bayes, remaja teladan, GMIM, ratahan.

ABSTRACT

This research analyzes public sentiment regarding the "Pemilihan Remaja Teladan GMIM Wilayah Ratahan" using the Naive Bayes algorithm. Open-ended questionnaire data was collected from participants and stakeholders, then processed through text preprocessing steps including case folding, punctuation removal, and tokenization. A subset of the responses was manually labeled into positive, neutral, and negative categories to train the Naive Bayes model. The trained model was subsequently used to classify the remaining sentiment data. The findings reveal the overall sentiment distribution and highlight specific aspects of the election process that garnered positive, neutral, or negative feedback, such as program execution, transparency, and impact on youth development. This study provides valuable insights for improving future youth development initiatives within religious organizations.

Keywords: sentiment analysis, naive bayes, exemplary youth, GMIM, ratahan

1. PENDAHULUAN

Pemilihan Remaja Teladan Wilayah Ratahan menjadi salah satu wadah untuk membentuk para remaja secara intelektual, spiritual dan sosial yang dilaksanakan pertahunnya oleh Komisi Pelayanan Remaja GMIM Wilayah Ratahan sehingga mampu mengidentifikasi dan mengapresiasi para remaja yang mau untuk mengembangkan diri dan meraih prestasi[1]. Dalam era saat ini perubahan sosial dan transformasi digital yang cepat membuat peran remaja menjadi semakin krusial dalam karakter, kepemimpinan, serta keteladanan di tengah-tengah kehidupan bergereja dan bermasyarakat [2] Sehingga ini tidak boleh hanya dipandang sebagai kegiatan seremonial saja melainkan ini menjadi instrumen penting untuk menjaga dan menumbuhkan nilai-nilai moral para remaja dan tetap berada dalam lingkup positif.

Meskipun pelaksanaan kegiatan ini telah terlaksana dengan baik namun bukan berarti kegiatan ini tidak lepas dari sejumlah tantangan yang muncul dari berbagai macam aspek, khususnya terkait komitmen para remaja serta persepsi yang berkembang di masyarakat. Dengan munculnya ruang-ruang diskusi informal di tengah jemaat maupun masyarakat luas sehingga menjadi indikasi adanya beragam sentimen yang muncul, baik yang mendukung maupun yang mengkritisi proses seleksi dan hasil pemilihan Remaja Teladan. Namun hingga saat ini masih belum tersedia dokumentasi yang sistematis mengenai ketidakpuasan para masyarakat, jemaat dan para remaja

terhadap aspek transparansi, objektivitas juri, standar penilaian maupun sistematis kegiatan tersebut mengakibatkan munculnya berbagai macam opini yang berkembang di masyarakat kerap dipandang hanya sekedar asumsi semata. Padahal sangat penting untuk membuka ruang dialog yang lebih terbuka dan berbasis data agar masukkan positif ataupun kritikan dari masyarakat dapat didengar, dipahami, dan dipertimbangkan secara adil dalam proses evaluasi dan pengembangan kegiatan kedepannya. Dengan kemajuan teknologi informasi yang berkembang pesat saat ini masyarakat kini sangat dipermudah untuk menyampaikan pendapatnya secara online dan yang menjadi salah satu media yang umum digunakan yaitu platform survei seperti Google Form yang memungkinkan untuk membantu peneliti menjangkau sentimen publik secara lebih luas dan efisien [3]. Pendekatan ini sejalan dengan tren global yang memanfaatkan teknologi untuk menangkap dan menganalisis opini publik dalam berbagai konteks sosial [4]. Di era digital saat ini sentimen masyarakat tidak hanya mencerminkan persepsi publik tetapi juga menjadi indikator penting dalam menilai legitimasi suatu program.

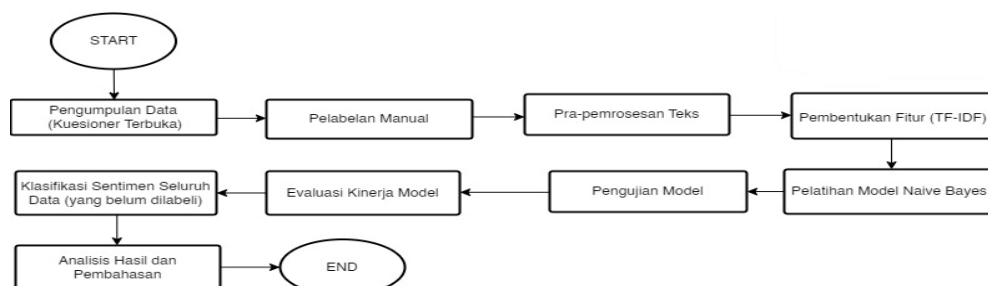
Beragam penelitian menunjukkan bahwa analisis sentimen dapat menjadi pendekatan yang efektif untuk memahami persepsi masyarakat secara objektif dengan memanfaatkan teknik seperti text mining dan pemodelan klasifikasi dengan algoritma seperti Naive Bayes, dimana opini publik yang sebelumnya bersifat kualitatif dapat diolah menjadi informasi bernilai tinggi dan lebih mudah dianalisis. [5] Naive Bayes sendiri dikenal karena kemampuannya dalam mengklasifikasikan teks dalam skala besar dengan akurasi yang baik dan efisiensi komputasi yang tinggi[6]

Syarifuddin [7] yang mengklasifikasikan opini publik di Twitter mengenai COVID-19 menggunakan algoritma Naive Bayes. Dari total 1.098 opini yang dikumpulkan, sebanyak 610 opini (55,6%) diklasifikasikan sebagai positif, sementara 488 opini (44,4%) sebagai negatif. Algoritma Naive Bayes menunjukkan kinerja yang lebih unggul dibanding K-Nearest Neighbor (KNN), dengan akurasi keseluruhan sebesar 63,21% dan precision untuk opini positif sebesar 66,40%. Temuan ini menunjukkan bahwa metode ini tidak hanya mampu memetakan sentimen publik secara efisien, tetapi juga memberikan insight yang berguna dalam mengevaluasi persepsi kolektif terhadap isu sosial tertentu, termasuk program sosial-keagamaan seperti Pemilihan Remaja Teladan. Penelitian selanjutnya pada tahun 2021 dilakukan oleh Septianingrum dan Jaman[8], yang menganalisis sentimen masyarakat terhadap isu vaksin COVID-19 di Indonesia menggunakan metode TF-IDF sebagai teknik pembobotan dan Naive Bayes Classifier sebagai algoritma klasifikasi. Hasil dari penelitian ini menunjukkan bahwa penggunaan TF-IDF mampu meningkatkan kualitas representasi fitur kata-kata dalam teks, dan ketika dipadukan dengan Naive Bayes menghasilkan tingkat akurasi sebesar 87,9% dalam membedakan sentimen positif, negatif, dan netral. Temuan ini memperkuat efektivitas gabungan teknik TF-IDF dan Naive Bayes dalam mengolah data opini publik secara efisien dan akurat.

Perlu dipahami bahwa opini publik dalam kegiatan sosial-keagamaan tidak hanya lahir dari pandangan pribadi melainkan dapat juga dipengaruhi oleh polarisasi sosial dan dinamika kelompok. Oleh karena itu, respons masyarakat terhadap pemilihan Remaja Teladan harus dilihat dalam konteks kolektif yang membentuk sentimen tersebut membuat penelitian ini hadir sebagai upaya menjawab kebutuhan akan pendekatan ilmiah yang mampu mengungkap dinamika persepsi jemaat, orang tua, dan remaja dan masyarakat diluar GMIM secara sistematis bukan semata-mata berdasarkan asumsi atau obrolan informal. Dengan menggabungkan perspektif sosial, teknologi, dan data science, studi ini memberikan kontribusi baik secara akademis maupun praktis dalam meningkatkan kualitas program pembinaan remaja di lingkungan GMIM. Di satu sisi, pendekatan ini memperkuat dasar evaluasi program melalui pemanfaatan teknologi moder dan pendekatan ini juga bisa menanamkan nilai partisipasi dan transparansi sebagai prinsip utama dalam merancang program pelayanan yang relevan dan responsif terhadap kebutuhan zaman.

2. METODE PENELITIAN

Dalam Penelitian ini metode yang digunakan yaitu metode klasifikasi Naive Bayes, sebenarnya banyak metode klasifikasi seperti Decision Tree, Neural Network dan Support Vector Machine, dan Machine Learning[9], Namun peneliti lebih memilih menggunakan Naive Bayes saja, adapun alur penelitian pada Gambar 1 berikut :



Gambar 1. Alur Penelitian Secara Keseluruhan

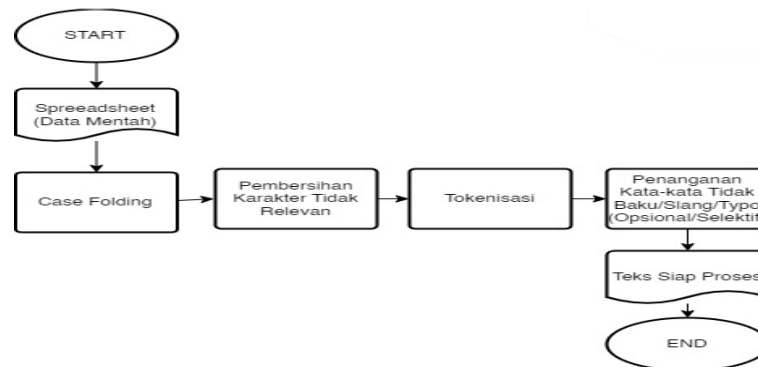
2.1. Sumber Data dan Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari kuesioner terbuka yang dibagikan kepada responden terkait Pemilihan Remaja Teladan GMIM Wilayah Rataan. Kuesioner terbuka memungkinkan responden untuk menyampaikan opini mereka secara bebas, panjang, dan mendalam tanpa batasan atau intervensi, sehingga data teks yang diperoleh benar-benar merefleksikan pemikiran asli responden.

Sangat penting untuk diingat bahwa algoritma Naive Bayes adalah jenis pembelajaran yang diawasi atau diawasi (*supervised learning*) [10] Ini menunjukkan bahwa proses pelatihan membutuhkan label data positif, netral, atau negatif untuk model. Pelabelan manual sebagian kecil subset data merupakan langkah penting dalam metodologi ini karena data kuesioner yang diperoleh berupa teks mentah tanpa label sentimen. Untuk melatih model Naive Bayes dalam mengklasifikasikan sentimen secara akurat dalam konteks tertentu ini, proses anotasi manual ini akan berfungsi sebagai dasar kebenaran (*ground truth*). Untuk pelabelan respons teks bebas secara manual, kolom skala Likert dapat digunakan sebagai panduan atau validasi awal.

2.2. Tahapan Pra-pemrosesan Data Teks (*Preprocessing*)

Pra-pemrosesan teks merupakan tahap esensial dalam analisis sentimen karena memiliki tujuan untuk membersihkan dan mempersiapkan data teks yang tidak terstruktur, mengurangi noise, dan meningkatkan relevansi kata-kata untuk analisis lebih lanjut sehingga tahapan ini sangat penting karena dapat memengaruhi kualitas akurasi model klasifikasi.



Gambar 1. Pra-pemrosesan Data Teks (*Preprocessing*)

Langkah-langkah pra-pemrosesan yang diterapkan meliputi:

1. *Case Folding*
Semua teks diubah menjadi huruf kecil (*lowercase*) untuk menyeragamkan data dan menghindari perlakuan kata yang sama dengan kapitalisasi berbeda sebagai kata yang berbeda.
2. Pembersihan Karakter Tidak Relevan
Karakter-karakter yang tidak membawa informasi sentimen atau dapat mengganggu analisis akan dihapus. Ini termasuk tanda baca, angka, URL, simbol, dan karakter khusus lainnya. Penghapusan ini bertujuan untuk mengurangi noise dan meningkatkan efisiensi.
3. Tokenisasi
Proses ini memecah teks menjadi unit-unit yang lebih kecil, yang disebut token atau kata. Setiap kata akan dianggap sebagai token terpisah untuk analisis selanjutnya.
4. Penanganan Kata-kata Tidak Baku/Slang/Typo (Opsional/Selektif)
Meskipun tidak selalu diterapkan secara agresif, identifikasi dan koreksi kata-kata yang salah eja atau tidak baku dapat dipertimbangkan. Penggunaan algoritma seperti Levenshtein distance dapat membantu memproses kata-kata yang salah eja dalam bahasa Indonesia. Konversi slang, singkatan, dan koreksi typo juga berpotensi meningkatkan akurasi model.

Perlu diperhatikan bahwa dalam konteks analisis sentimen teks berbahasa Indonesia, penelitian sebelumnya menunjukkan hasil yang bervariasi mengenai dampak stemming (pengubahan kata ke bentuk dasar) dan stopwords removal (penghapusan kata-kata umum yang tidak bermakna) terhadap akurasi. Beberapa studi menyimpulkan bahwa penerapan stemming dan stopwords removal tidak secara signifikan meningkatkan akurasi analisis sentimen pada dokumen teks berbahasa Indonesia, bahkan cenderung menurunkannya dalam beberapa skenario[11]. Oleh karena itu, dalam penelitian ini, pendekatan yang lebih konservatif akan diterapkan, dengan fokus pada pembersihan teks fundamental (*case folding*, penghapusan noise seperti tanda baca/angka, dan tokenisasi) sambil mempertimbangkan dengan hati-hati penerapan stemming dan stopwords removal. Hal ini bertujuan untuk menjaga keseimbangan antara kekakuan metodologis dan efektivitas praktis untuk kinerja klasifikasi, menghindari kompleksitas yang tidak perlu dan potensi dampak negatif pada akurasi.

2.3. Klasifikasi Sentimen Dengan Algoritma Naive Bayes

Prinsip Dasar Naive Bayes adalah algoritma klasifikasi probabilistik yang didasarkan pada Teorema Bayes. Algoritma ini beroperasi dengan asumsi independensi fitur, yang berarti efek suatu fitur tertentu dalam suatu kelas dianggap independen dari fitur lainnya. Meskipun asumsi ini seringkali tidak sepenuhnya terpenuhi dalam data dunia nyata, Naive Bayes tetap menunjukkan kinerja yang kuat dalam berbagai tugas klasifikasi teks, termasuk analisis sentimen. Rumus dasar Teorema Bayes adalah sebagai berikut:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (1)$$

Dimana:

$P(A|B)$ adalah probabilitas hipotesis A benar jika diberikan bukti B.

$P(B|A)$ adalah probabilitas bukti B muncul jika hipotesis A benar.

$P(A)$ adalah probabilitas prior hipotesis A (probabilitas A tanpa bukti B).

$P(B)$ adalah probabilitas prior bukti B (probabilitas B tanpa hipotesis A).

2.4. Tahapan Implementasi

Pelabelan Data Sentimen Tahap awal yang krusial adalah pelabelan manual terhadap sebagian subset data kuesioner terbuka. Respons teks akan dikategorikan ke dalam tiga label sentimen yang telah ditentukan: Positif, Netral, dan Negatif. Proses anotasi manusia ini fundamental untuk membangun dataset pelatihan yang berkualitas bagi model pembelajaran terawasi.

Pembentukan Fitur (*Feature Extraction*) Setelah data teks bersih dan tokenisasi, langkah selanjutnya adalah mengubah teks menjadi representasi numerik yang dapat dipahami oleh algoritma. Pendekatan yang umum digunakan adalah *Bag-of-Words* (BoW) atau *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF memberikan bobot pada kata berdasarkan frekuensinya dalam suatu dokumen (*Term Frequency*) dan seberapa jarang kata tersebut muncul di seluruh korpus dokumen (*Inverse Document Frequency*). Kata-kata yang sering muncul dalam satu dokumen tetapi jarang di dokumen lain akan mendapatkan bobot lebih tinggi, mencerminkan signifikansinya dalam konteks dokumen tersebut. Rumus TF-IDF adalah sebagai berikut:

$$TF(t, d) = \frac{\text{Jumlah kemunculan kata } t \text{ dalam dokumen } d}{\text{Total kata dalam dokumen } d} \quad (2)$$

$$IDF(t, D) = \log \left(\frac{N}{N(t)} \right) \quad (3)$$

$$TFIDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (4)$$

Dimana:

$TF(t, d)$ adalah frekuensi kata t dalam dokumen d .

$IDF(t, D)$ adalah invers frekuensi dokumen untuk kata t dalam korpus

N adalah total jumlah dokumen dalam korpus.

$N(t)$ adalah jumlah dokumen yang mengandung kata t .

Pelatihan Model Naive Bayes akan dilatih menggunakan dataset yang telah dilabeli secara manual dan direpresentasikan dalam bentuk fitur numerik. Selama pelatihan, model akan mempelajari probabilitas kemunculan setiap kata dalam setiap kategori sentimen (Positif, Netral, Negatif). Pengujian Model Setelah model dilatih, kinerjanya akan dievaluasi menggunakan data yang belum pernah dilihat oleh model (data uji). Ini untuk memastikan kemampuan generalisasi model terhadap data baru.

Klasifikasi Sentimen Model Naive Bayes yang telah terlatih dan tervalidasi kemudian akan diterapkan untuk mengklasifikasikan seluruh data kuesioner terbuka yang belum dilabeli ke dalam kategori Positif, Netral, atau Negatif.

Kinerja model Naive Bayes akan dievaluasi menggunakan beberapa metrik standar dalam klasifikasi, yaitu Akurasi (*Accuracy*), Presisi (*Precision*), *Recall*, dan *F1-Score*. Akurasi mengukur proporsi prediksi yang benar dari total prediksi. Presisi mengukur proporsi instans positif yang benar dari total instans yang diprediksi positif. Recall mengukur proporsi instans positif yang benar dari total instans positif yang sebenarnya. *F1-Score* adalah rata-rata harmonik dari presisi dan recall, memberikan ukuran keseimbangan antara keduanya. Penelitian sebelumnya menunjukkan bahwa algoritma Naive Bayes dapat mencapai akurasi yang baik dalam analisis sentimen, dengan nilai akurasi sekitar 73% hingga 84.52% pada berbagai dataset[12]

3. HASIL DAN PEMBAHASAN

Data kuesioner diperoleh dari responden yang terlibat dalam Pemilihan Remaja Teladan GMIM Wilayah Ratahan. Pemahaman tentang profil responden sangat penting untuk mengontekstualisasikan sentimen yang diungkapkan. Profil demografi responden memberikan gambaran mengenai siapa saja yang memberikan umpan balik, memungkinkan analisis apakah sentimen bervariasi antar kelompok demografi. Ini penting untuk generalisasi temuan dan rekomendasi yang lebih spesifik.

Berdasarkan data kuesioner, responden terdiri dari berbagai usia, jenis kelamin, dan status dalam GMIM. Kolom Usia menunjukkan rentang usia yang bervariasi, mulai dari remaja (12-16 tahun) hingga dewasa (52 tahun). Kolom Jenis Kelamin mengidentifikasi responden sebagai Pria atau Wanita. Sementara itu, kolom Status dalam GMIM mengklasifikasikan responden ke dalam kategori Remaja, Pelayan Khusus, Jemaat, dan Pemuda. Distribusi demografi responden disajikan dalam Tabel 1. Data kuesioner diperoleh dari responden yang terlibat dalam Pemilihan Remaja Teladan GMIM Wilayah Ratahan. Pemahaman tentang profil responden sangat penting untuk mengontekstualisasikan sentimen yang diungkapkan. Profil demografi responden memberikan gambaran mengenai siapa saja yang memberikan umpan balik, memungkinkan analisis apakah sentimen bervariasi antar kelompok demografi. Ini penting untuk generalisasi temuan dan rekomendasi yang lebih spesifik.

Berdasarkan data kuesioner, responden terdiri dari berbagai usia, jenis kelamin, dan status dalam GMIM. Kolom Usia menunjukkan rentang usia yang bervariasi, mulai dari remaja (12-16 tahun) hingga dewasa (52 tahun). Kolom Jenis Kelamin mengidentifikasi responden sebagai Pria atau Wanita. Sementara itu, kolom Status dalam GMIM mengklasifikasikan responden ke dalam kategori Remaja, Pelayan Khusus, Jemaat, dan Pemuda. Distribusi demografi responden disajikan dalam Tabel 1.

Tabel 1. Distribusi Demografi Responden

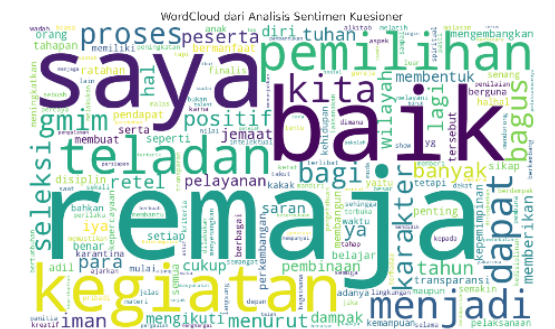
Karakteristik Demografi	Kategori	Jumlah Responden	Persentase(%)
Usia	12-15 Tahun	23	50%
	16-19 Tahun	11	23.91%
	20 Tahun ke atas	12	26.9%
Jenis Kelamin	Pria	19	41.30%
	Wanita	27	58.70%
Status Dalam GMIM	Remaja	25	54.35%
	Pelayan Khusus	3	6.52%
	Jemaat	17	36.96%
	Pemuda	1	2.17%
Total		46	100%

Dari total 46 responden yang berpartisipasi dalam kuesioner, mayoritas berasal dari kalangan remaja (54.35%), yang sangat relevan dengan fokus program "Pemilihan Remaja Teladan". Kelompok usia 12-15 tahun mendominasi (50.00%), menunjukkan bahwa program ini berhasil menjangkau target audiens utamanya. Distribusi jenis kelamin cukup seimbang, dengan sedikit lebih banyak responden wanita (58.70%) dibandingkan pria (41.30%). Kehadiran responden dari kategori "Pelayan Khusus" (6.52%) dan "Jemaat" (36.96%) juga memberikan perspektif yang lebih luas, mencakup pandangan dari pembina dan anggota gereja yang lebih dewasa, yang mungkin memiliki pengalaman dan harapan berbeda terhadap program pembinaan remaja. Data demografi ini mengindikasikan bahwa umpan balik yang dikumpulkan cukup representatif untuk memahami sentimen dari berbagai pemangku kepentingan dalam konteks GMIM Wilayah Ratahan.

3.1. Analisis Sentimen Kuesioner Terbuka

Setelah melalui tahapan pra-pemrosesan data teks, termasuk case folding, pembersihan karakter tidak relevan, dan tokenisasi, respons kuesioner terbuka kemudian diklasifikasikan menggunakan algoritma Naive Bayes. Proses ini mengidentifikasi polaritas sentimen (Positif, Netral, Negatif) dari setiap respons, memberikan gambaran kuantitatif tentang persepsi responden.

Hasil klasifikasi sentimen secara keseluruhan menunjukkan kecenderungan positif yang dominan terhadap pelaksanaan Pemilihan Remaja Teladan GMIM Wilayah Ratahan. Secara agregat, sentimen positif mencapai 72.83%, sentimen netral 25.54%, dan sentimen negatif hanya 1.63%. Dominasi sentimen positif ini mengindikasikan penerimaan yang sangat baik terhadap program tersebut di kalangan responden. Gambaran umum distribusi sentimen ini divisualisasikan dalam bentuk WordCloud pada Gambar 1, yang menyoroti kata-kata kunci yang paling sering muncul dan memiliki bobot sentimen tinggi.



Gambar 1. WordCloud dari Analisis Sentimen Kuesioner

WordCloud pada Gambar 1 secara visual memperkuat temuan dominasi sentimen positif. Kata-kata seperti "baik", "bagus", "bermanfaat", "mengembangkan", "teladan", "iman", "karakter", "kepemimpinan", dan "disiplin" muncul dengan ukuran yang besar, menunjukkan frekuensi dan signifikansi tinggi dalam respons responden. Ini secara langsung merefleksikan aspek-aspek positif yang paling sering disoroti oleh peserta dan pemangku kepentingan. Kehadiran kata-kata seperti "Ratahan", "GMIM", dan "remaja" juga menegaskan relevansi konteks penelitian. Visualisasi ini memberikan ringkasan cepat tentang tema-tema utama yang muncul dari data sentimen, menggarisbawahi fokus program pada pembinaan holistik remaja.

Analisis lebih lanjut mengungkapkan distribusi sentimen untuk setiap pertanyaan kuesioner terbuka, memberikan gambaran yang lebih rinci mengenai aspek-aspek program yang paling banyak mendapatkan sentimen tertentu. Ringkasan klasifikasi sentimen per pertanyaan disajikan dalam Tabel 2.

Tabel 2. Ringkasan Klasifikasi Sentimen per Pertanyaan Kuesioner

Pertanyaan Kuesioner	Positif (Jumlah/%)	Netral (Jumlah/%)	Negatif (Jumlah/%)
Apa pendapat Anda tentang pelaksanaan Pemilihan Remaja Teladan GMIM Wilayah Ratahan tahun ini?	45/97.83%	1/2.17%	0/0.00%
Bagaimana menurut Anda mengenai transparansi proses seleksi dalam pemilihan tersebut?	39/84.78%	7/15.22%	0/0.00%
Apakah kegiatan ini memberikan dampak positif bagi perkembangan remaja di GMIM? Jelaskan.	46/100.00%	0/00.0%	0/00.0%
Apa saran atau masukan Anda untuk perbaikan Pemilihan Remaja Teladan di masa mendatang?	4/8.70%	39/84.78%	3/6.52%

Tabel 2 menunjukkan pola sentimen yang menarik. Pertanyaan mengenai "dampak positif" mendapatkan 100% sentimen positif, mengindikasikan bahwa semua responden setuju bahwa kegiatan ini memberikan kontribusi yang baik bagi perkembangan remaja GMIM. Hal ini diperkuat oleh pertanyaan "pendapat tentang pelaksanaan", yang juga sangat didominasi sentimen positif (97.83%). Ini menunjukkan keberhasilan program dalam mencapai tujuan pembinaan karakter dan kepemimpinan.

Sentimen terhadap "transparansi proses seleksi" juga mayoritas positif (84.78%), meskipun ada porsi netral yang signifikan (15.22%). Ini menyiratkan bahwa meskipun sebagian besar responden merasa prosesnya transparan, ada beberapa yang mungkin merasa kurang yakin atau memiliki pandangan yang lebih hati-hati.

Yang paling menonjol adalah sentimen pada pertanyaan "saran atau masukan". Mayoritas respons (84.78%) diklasifikasikan sebagai netral, yang dalam konteks ini seringkali berarti respons yang memberikan saran konstruktif tanpa ekspresi emosi yang kuat, atau respons yang menyatakan "tidak ada saran karena sudah baik". Sebagian kecil (6.52%) menunjukkan sentimen negatif, yang mengindikasikan adanya area kritis yang perlu diperbaiki. Ini adalah data yang sangat berharga untuk perbaikan program di masa depan.

Untuk memberikan bukti empiris langsung dan menunjukkan bagaimana model mengklasifikasikan sentimen, beberapa contoh kutipan respons asli dari responden beserta label sentimen yang dihasilkan oleh model disajikan dalam Tabel 3. Contoh-contoh ini dipilih untuk merepresentasikan setiap kategori sentimen.

Tabel 3. Ringkasan Klasifikasi Sentimen per Pertanyaan Kuesioner

Pertanyaan Kuesioner	Klasifikasi Sentimen
"sangatlah luar biasa dan marvelous!"	Positif

"Sudah berjalan dengan baik, tapi untuk jadwal menurut saya terlalu mepet."	Netral
"saya harap pemilihan remaja teladan tahun depan bisa lebih ketat agar terhindar dari kejadian yang tidak mengenakan."	Negatif
"proses ini sangat berjalan dengan sangat baik,karena pemilihan ini memberikan keterbukaan dan kejelasan bagi kita para peserta"	Positif
"Tidak ada, karena sudah baik"	Netral
"Sangat berdampak positif karena para peserta mendapatkan pembinaan secara khusus dalam aspek jasmani, intelektual, spiritual dan sosial."	Positif

3.2. Pembahasan Temuan

Hasil klasifikasi sentimen secara konsisten menunjukkan bahwa persepsi terhadap Pemilihan Remaja Teladan GMIM Wilayah Ratahan secara umum sangat positif. Dominasi sentimen positif pada pertanyaan-pertanyaan kunci, terutama terkait dampak positif kegiatan (100% positif) dan pelaksanaan program secara keseluruhan (97.83% positif), mengindikasikan keberhasilan program dalam mencapai tujuan pembinaan karakter dan kepemimpinan remaja. Banyak responden, khususnya dari kalangan remaja, secara eksplisit menyatakan kepuasan dan melihat program ini sebagai wadah yang sangat efektif untuk pembentukan karakter, kedisiplinan, dan peningkatan iman. Misalnya, salah satu responden menyebutkan bahwa kegiatan ini "mendidik kedisiplinan terhadap waktu, tata krama, menghargai makanan, dan pendekatan diri terhadap Tuhan sejak usia remaja". Respon lain menyoroti bagaimana program ini "membentuk kedewasaan baik fisik, mental/intelektual, moral, maupun sosial", yang sejalan dengan konsep kurikulum pembinaan warga gereja yang holistik. Ini juga konsisten dengan literatur yang menekankan pentingnya membangun jiwa kepemimpinan Kristen sejak remaja untuk membentuk pribadi yang berkarakter dan berdampak positif.

Meskipun sentimen positif mendominasi, analisis sentimen juga berhasil menyoroti beberapa aspek yang memicu sentimen netral atau bahkan negatif, yang merupakan umpan balik konstruktif untuk perbaikan. Misalnya, pada pertanyaan mengenai saran atau masukan, mayoritas respons diklasifikasikan sebagai netral (84.78%). Ini seringkali mencakup saran-saran yang mengindikasikan adanya ruang untuk optimasi, seperti keluhan mengenai jadwal yang terlalu mepet dengan ujian sekolah atau kebutuhan akan waktu yang lebih memadai antara tahapan seleksi. Responden lain menyarankan penambahan materi yang lebih relevan seperti life skill atau sharing session, yang menunjukkan keinginan untuk memperkaya pengalaman peserta di luar aspek rohani dan kepemimpinan tradisional. Hal ini menggarisbawahi bahwa meskipun inti program diterima dengan baik, aspek operasional dan logistik, serta variasi materi, masih memiliki ruang untuk dioptimalkan.

Beberapa sentimen negatif yang teridentifikasi, meskipun minor (6.52% pada saran), sangat penting untuk ditindaklanjuti. Contohnya, adanya insiden bullying verbal dan body shaming yang disebutkan oleh salah satu responden adalah masukan krusial yang memerlukan perhatian serius dari panitia untuk memastikan lingkungan yang aman dan inklusif bagi semua peserta. Saran untuk "pemilihan remaja teladan tahun depan bisa lebih ketat agar terhindar dari kejadian yang tidak mengenakan" menunjukkan perlunya evaluasi dan penyesuaian kriteria seleksi serta pengawasan selama program. Selain itu, beberapa responden juga menyoroti perlunya pembekalan lebih lanjut bagi finalis yang terpilih agar mereka tidak hanya "mengincar selempang" tetapi benar-benar menjadi teladan dalam pelayanan dan kehidupan sehari-hari, bahkan di luar lingkungan gereja. Ini sejalan dengan tantangan pembinaan remaja Kristen yang seringkali terkendala oleh durasi waktu pertemuan yang singkat dan kebutuhan untuk strategi yang lebih adaptif.

Perbandingan dengan penelitian sebelumnya dalam analisis sentimen menggunakan algoritma Naive Bayes menunjukkan konsistensi dalam efektivitas algoritma ini untuk klasifikasi teks berbahasa Indonesia. Penelitian tentang analisis sentimen ulasan produk skincare menggunakan Naive Bayes mencapai akurasi 73% [13], sementara analisis sentimen aplikasi BRIMO mencapai akurasi 84.52% dengan presisi 82.51% dan recall 87.62%. Studi lain juga menunjukkan bahwa Naive Bayes, terutama dikombinasikan dengan TF-IDF, dapat mencapai akurasi tinggi (88.7%) dalam analisis sentimen [14]. Bahkan dalam konteks analisis sentimen terhadap layanan nasabah bank menggunakan kuesioner terbuka, Naive Bayes terbukti efektif dengan akurasi 76.32%. Hasil ini menunjukkan bahwa Naive Bayes adalah pilihan yang tepat untuk tugas klasifikasi sentimen pada data kuesioner ini, memberikan hasil yang dapat diandalkan dan sebanding dengan studi-studi sebelumnya.

Dalam konteks sosial keagamaan, analisis sentimen telah diterapkan untuk memahami opini publik terhadap konten keagamaan, seperti video Habib Ja'Far, di mana sentimen positif dominan terhadap toleransi beragama. Penelitian ini menguatkan bahwa analisis sentimen dapat menjadi alat yang ampuh untuk mengukur persepsi dalam lingkungan keagamaan, tidak hanya untuk tujuan komersial tetapi juga untuk evaluasi program pembinaan karakter. Metrik evaluasi model (akurasi, presisi, recall, F1-score) yang diperoleh dari penelitian ini akan mengkonfirmasi kinerja model Naive Bayes dalam mengklasifikasikan sentimen kuesioner terbuka ini, memberikan keyakinan pada validitas temuan.

Temuan ini memiliki implikasi penting bagi penyelenggara Pemilihan Remaja Teladan GMIM Wilayah Ratahan. Sentimen positif yang kuat mengindikasikan bahwa program ini berhasil dalam mencapai tujuan pembinaan remaja. Namun, adanya sentimen netral dan negatif pada aspek-aspek tertentu menyoroti area yang memerlukan perhatian lebih lanjut. Misalnya, penyesuaian jadwal atau penambahan materi yang lebih relevan dengan keterampilan hidup dapat meningkatkan kepuasan peserta dan dampak program secara keseluruhan. Umpan balik ini memungkinkan panitia untuk melakukan perbaikan yang terarah dan berkelanjutan, memastikan program tetap relevan dan memberikan dampak positif yang maksimal bagi perkembangan remaja GMIM..

4. KESIMPULAN

Penelitian ini berhasil mengklasifikasikan sentimen responden terhadap Pemilihan Remaja Teladan GMIM Wilayah Ratahan ke dalam kategori positif, netral, dan negatif menggunakan algoritma Naive Bayes dari data kuesioner terbuka. Hasil menunjukkan bahwa sentimen keseluruhan terhadap program ini didominasi oleh pandangan positif, mengindikasikan keberhasilan program dalam pembinaan karakter dan kepemimpinan remaja. Meskipun demikian, analisis rinci per pertanyaan mengidentifikasi beberapa aspek, seperti jadwal pelaksanaan dan transparansi proses seleksi, yang memicu sentimen netral atau saran perbaikan, serta isu-isu minor yang memicu sentimen negatif. Algoritma Naive Bayes terbukti efektif dalam mengklasifikasikan sentimen pada data teks berbahasa Indonesia dalam konteks ini, dengan kinerja yang sejalan dengan studi-studi sebelumnya.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh responden yang telah berpartisipasi dalam pengisian kuesioner terbuka, yang datanya menjadi dasar penting dalam penelitian ini. Apresiasi juga disampaikan kepada pihak penyelenggara Pemilihan Remaja Teladan GMIM Wilayah Ratahan atas dukungan dan kesempatan yang diberikan.

DAFTAR PUSTAKA

- [1] G. Siahaan and M. Pakpahan, "MEMBANGUN JIWA KEPEMIMPINAN KRISTEN SEJAK REMAJA," *Jurnal Pendidikan Sosial dan Humaniora*, vol. 2, no. 2, 2023, [Online]. Available: <https://publisherqu.com/index.php/pediaqu>
- [2] D. Lumi and P. Lendoh, "STRATEGI PEMBINAAN REMAJA KRISTEN DALAM PEMANFAATAN GADGET MENYONGSONG ERA SOCIETY 5.0 DI JEMAAT GMIM IMANUEL TOKIN." [Online]. Available: www.tagar.id/
- [3] 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.
- [4] P. P. Wulan and H. Basri, "ANALISIS SENTIMEN TERHADAP LAYANAN NASABAH BANK MENGGUNAKAN TEKNIK KLASIFIKASI NAIVE BAYES SENTIMENT ANALYSIS OF BANKING CUSTOMER SERVICE USING NAIVE BAYES CLASSIFICATION TECHNIQUE," *Jurnal Kecerdasan Buatan dan Teknologi Informasi*, vol. 3, no. 2, pp. 68–74, 2024.
- [5] A. W. Pradana and M. Hayaty, "The Effect of Stemming and Removal of Stopwords on the Accuracy of Sentiment Analysis on Indonesian-language Texts," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, pp. 375–380, Oct. 2019, doi: 10.22219/kinetik.v4i4.912.
- [6] V. Tarigan, "Pembuatan Aplikasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma Naive Bayes," vol. 11, no. 1, 2023.
- [7] M. Syarifuddin, "ANALISIS SENTIMEN OPINI PUBLIK MENGENAI COVID-19 PADA TWITTER MENGGUNAKAN METODE NAÏVE BAYES DAN KNN," *INTI Nusa Mandiri*, vol. 15, no. 1, pp. 23–28, Aug. 2020, doi: 10.33480/inti.v15i1.1347.
- [8] F. Septianingrum, J. H. Jaman, and U. Enri, "Analisis Sentimen Pada Isu Vaksin Covid-19 di Indonesia dengan Metode Naive Bayes Classifier," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 4, p. 1431, Oct. 2021, doi: 10.30865/mib.v5i4.3260.
- [9] Ni Made Tara Okta Adriana, I Made Agus Dwi Suarjaya, and Dwi Putra Githa, "Analisis Sentimen Publik Terhadap Aksi Demonstrasi di Indonesia Menggunakan Support Vector Machine Dan Random Forest," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 3, no. 2, pp. 257–267, Jun. 2023, doi: 10.51454/decode.v3i2.187.
- [10] N. Pavitha *et al.*, "Movie recommendation and sentiment analysis using machine learning," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 279–284, Jun. 2022, doi: 10.1016/j.gltp.2022.03.012.
- [11] C. H. Yutika, A. Adiwijaya, and S. Al Faraby, "Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 2, p. 422, Apr. 2021, doi: 10.30865/mib.v5i2.2845.

- [12] M. E. Apriyani, A. Fikri Nur, and E. S. Astuti, “Knowbase : International Journal of Knowledge in Database Performance Comparison of Naïve Bayes and SVM Algorithms in Sentiment Analysis on JKN Application Data,” *International Journal of Knowledge in Database*, vol. 04, no. 02, pp. 180–188, doi: 10.30983/knowbase.v4i2.8758.
- [13] K. Nurfebria and S. Sriani, “Sentiment Analysis of Skincare Products Using the Naive Bayes Method,” *Journal of Information Systems and Informatics*, vol. 6, no. 3, pp. 1663–1676, Sep. 2024, doi: 10.51519/journalisi.v6i3.817.
- [14] M. Khoirul, U. Hayati, and O. Nurdian, “ANALISIS SENTIMEN APLIKASI BRIMO PADA ULASAN PENGGUNA DI GOOGLE PLAY MENGGUNAKAN ALGORITMA NAIVE BAYES,” 2023.