

Makalah Penelitian

Analisis Sentimen terhadap Kurikulum Merdeka di MTS Al Ishlah Medan Menggunakan Algoritma Naïve Bayes

Wirabuana¹, Afif Badawi²

^{1,2} Fakultas Sains dan Teknologi, Program Studi Sistem Komputer, Universitas Pembangunan Panca Budi, Medan, Indonesia
¹wirabuana113344@gmail.com, ²afifbadawi@pancabudi.ac.id

Corresponding Author: Wirabuana¹, Afif Badawi²

ABSTRACT

The Independent Curriculum is one of the latest educational innovations in Indonesia that aims to increase students' learning freedom and teachers' autonomy in designing learning. However, its implementation has generated various responses from education practitioners. This study aims to analyze student and teacher sentiments towards the implementation of the Independent Curriculum at MTS Al Ishlah Medan. Using the Naïve Bayes algorithm, comment data is classified into positive and negative sentiments. Data were obtained through questionnaires, interviews, and observations, then processed using text preprocessing techniques such as tokenization, normalization, stopword removal, TF-IDF, and labeling using the InSetLexicon dictionary. The results of the analysis of 60 comments showed that 75% had positive sentiments and 25% were negative. The model evaluation showed an accuracy value of 83% or 0.83, precision and recall respectively and an F1-score of 91% or 0.91 for the negative class. These findings indicate that the majority of respondents responded positively to the Independent Curriculum. This study provides useful input for educational institutions and policy makers in evaluating the effectiveness of this new curriculum.

Keywords: Sentiment Analysis; Independent Curriculum; Naïve Bayes Algorithm; InSetLexicon; TF-IDF

ABSTRAK

Kurikulum Merdeka merupakan salah satu inovasi pendidikan terbaru di Indonesia yang bertujuan meningkatkan kebebasan belajar peserta didik dan otonomi guru dalam menyusun pembelajaran. Namun, implementasinya menimbulkan beragam tanggapan dari para pelaku pendidikan. Penelitian ini bertujuan untuk menganalisis sentimen siswa dan guru terhadap penerapan Kurikulum Merdeka di MTS Al Ishlah Medan. Dengan menggunakan algoritma Naïve Bayes, data komentar diklasifikasikan menjadi sentimen positif dan negatif. Data diperoleh melalui kuesioner, wawancara, dan observasi, kemudian diolah menggunakan teknik preprocessing teks seperti tokenisasi, normalisasi, stopword removal, TF-IDF, dan pelabelan menggunakan kamus InSetLexicon. Hasil analisis terhadap 60 komentar menunjukkan bahwa 75% memiliki sentimen positif dan 25% negatif. Evaluasi model menunjukkan nilai akurasi sebesar 83% atau 0.83, precision dan recall masing-masing serta F1-score 91% atau 0.91 untuk kelas negatif. Temuan ini menunjukkan bahwa mayoritas responden merespon Kurikulum Merdeka secara positif. Penelitian ini memberikan masukan yang berguna bagi institusi pendidikan dan pembuat kebijakan dalam mengevaluasi efektivitas kurikulum baru ini.

Kata Kunci: Analisis Sentimen; Kurikulum Merdeka; Algoritma Naïve Bayes; InSetLexicon; TF-IDF

1. Pendahuluan

Di Indonesia, program pendidikan telah mencapai tahap pengembangan Kurikulum Merdeka, yang merupakan bentuk kurikulum darurat yang dibuat sebagai tanggapan terhadap dampak pandemi Covid-19 [1]. Kurikulum baru ini mengutamakan pembelajaran yang berpusat sepenuhnya pada siswa, mengangkat konsep “belajar bebas”. Istilah tersebut dimaksudkan sebagai cara yang memungkinkan siswa memilih pelajaran yang mereka anggap menarik. Sekolah berhak dan bertanggung jawab untuk membuat kurikulum yang sesuai dengan kebutuhan dan karakteristik setiap sekolah. Kebijakan pemilihan kurikulum yang diharapkan dapat mempercepat reformasi kurikulum nasional. Kebijakan yang memungkinkan sekolah memilih program studi adalah salah satu cara untuk mengelola perubahan [2].

Tujuan utama pendidikan adalah untuk mencetak generasi yang cerdas dan memiliki karakter yang sesuai dengan nilai-nilai bangsa Indonesia. Pendidikan juga diharapkan dapat mendorong siswa menuju perubahan yang lebih



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

baik dari generasi sebelumnya. Dengan demikian, pendidikan diharapkan dapat melahirkan siswa yang kreatif, inovatif, dan membawa perubahan bagi bangsa Indonesia [3].

Kurikulum Merdeka Belajar, yang diluncurkan baru-baru ini oleh Kementerian Pendidikan dan Kebudayaan Republik Indonesia (Kemendikbud RI), bertujuan untuk meningkatkan kemampuan pedagogis guru dengan menerapkan 5M: Memanusiakan Hubungan, Memahami Konsep, Membangun Keberlanjutan, Memilih Tantangan, dan Memberdayakan Konteks. Tujuan dari kurikulum ini adalah untuk memberi siswa kesempatan untuk belajar secara mandiri [4].

Guru selalu berusaha memberikan yang terbaik kepada peserta didik saat menerapkan kurikulum merdeka. Namun, pada awal pelaksanaannya terdapat beberapa tantangan, seperti perlunya kolaborasi antara guru dan peserta didik, administrasi dalam perencanaan kegiatan pembelajaran, karena perangkat pembelajaran harus banyak diubah, dan guru juga harus belajar mengukur penilaian karena kurikulum merdeka melibatkan penilaian [5].

Analisis sentimen adalah teknik yang secara otomatis mengidentifikasi, mengekstraksi, dan memproses informasi tekstual untuk menemukan informasi emosional dalam ekspresi pikiran [6]. Klasifikasi tersebut akan mengetahui apakah data tersebut bersifat positif atau negatif [7]. Pada penelitian kali ini analisis sentimen menggunakan algoritma Naive Bayes untuk menentukan apakah sebuah data bersifat positif atau negatif. Berdasarkan teorema Bayes bahwa semua kegiatan memberikan kontribusi yang sama penting atau saling bebas pada pemilihan kelas tertentu, Algoritma Naive Bayes adalah salah satu algoritma klasifikasi yang paling banyak digunakan dalam jenis data dan teks [8].

Penelitian ini adalah text mining yang merupakan salah satu contoh dari machine learning atau biasa disebut analisis teks [9]. Dalam penelitian sebelumnya, yang dilakukan oleh Refandah Puspitasari dan Aries Dwi Indriyanti (2024) menggunakan metode Naive Bayes untuk melakukan analisis sentimen opini publik tentang skripsi kebijakan terbaru di media sosial Twitter. Hasil penelitian menunjukkan akurasi sebesar 76% pada percobaan pertama dan 77% pada percobaan kedua, dengan rasio perbandingan antara data pelatihan dan data uji sebesar 90%:10% dan 80%:20%. Penelitian lain oleh Elly Indrayuni (2019) yaitu Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes, menemukan bahwa hasilnya memiliki akurasi 90,50% [10].

Kemudian penelitian terdahulu berikutnya yang dilakukan oleh Veny Amilia Fitri, Rachmadita Andreswari dan Muhammad Azani Hasibuan (2019) menggunakan metode Naive Bayes, Decision Tree, dan Random Forest dalam menganalisis studi kasus kampanye anti LGBT, dari penelitian tersebut menghasilkan bahwa naive bayes memperoleh akurasi tertinggi yaitu 83,43%, dari metode decision tree dan random forest yang memiliki tingkat akurasi yang cukup rendah. Penelitian terdahulu menunjukkan bahwa Algoritma Naive Bayes memberikan tingkat akurasi yang baik dalam analisis sentimen [11].

Penelitian ini bertujuan untuk menganalisis sentimen atau tanggapan dari siswa dan guru di MTS Al Ishlah Medan terhadap kurikulum merdeka pada tahun 2025 melalui penerapan Algoritma Naive Bayes dalam menganalisis sentimen positif dan negatif. Oleh karena itu, penulis menggunakan algoritma ini untuk menganalisis sentimen terhadap Kurikulum Merdeka di MTS Al Ishlah Medan. Hasil analisis ini diharapkan dapat memberikan wawasan bagi pemerintah, satuan pendidikan, guru, dan masyarakat mengenai pandangan dari mereka yang terlibat langsung dalam penerapan Kurikulum Merdeka, apakah itu bernada positif, negatif, atau netral.

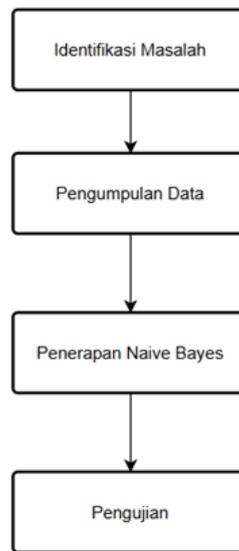
2. Metode Penelitian

Dalam studi ini, peneliti mengadopsi pendekatan kuantitatif, yang memang dirancang untuk mengungkap fenomena secara lebih holistik dan kontekstual. Dengan menggunakan pendekatan analisis masalah, peneliti melakukan penelitian yang berfokus pada serangkaian studi lapangan, situasi, atau aktivitas individu dan kolektif yang berkembang dari waktu ke waktu dalam konteks tertentu [12]. Metodologi ini dipilih karena mampu memberikan informasi yang lebih terukur dan objektif, terutama saat menganalisis pendapat dan perasaan orang tentang Kurikulum Merdeka di MTS Al Ishlah Medan. Berikut tahapan penelitian yang ditunjukkan pada gambar 1 dibawah [13].



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.



Gambar 1. Tahapan Penelitian

Pada gambar 1. Terdapat beberapa tahapan penelitian yang dilakukan yaitu dengan identifikasi masalah, pengumpulan data dan penerapan algoritma naive bayes serta melakukan pengujian dengan menghitung nilai akurasi, presisi, recall dan f-1 score.

2.1 Identifikasi Masalah

Kurikulum Merdeka adalah kebijakan baru dalam pendidikan di Indonesia yang memberi keleluasaan kepada sekolah dan guru untuk menyusun kurikulum dengan cara yang lebih fleksibel. Namun, penerapan kurikulum ini masih melibatkan pendapat dari guru dan siswa. Ada yang mendukung karena dianggap lebih inovatif, sementara yang lain merasa masih ada kendala dalam pelaksanaannya. Jadi, penting untuk melakukan analisis sentimen agar kita bisa mengetahui bagaimana guru dan siswa merespons Kurikulum Merdeka ini [14].

2.2 Pengumpulan Data

Dalam penelitian ini, sentimen data dikumpulkan dari berbagai sumber, seperti kuesioner dan ulasan dari guru serta siswa mengenai penerapan Kurikulum Merdeka di MTS Al Ishlah Medan. Data yang diperoleh akan diolah menggunakan teknik text preprocessing, seperti tokenisasi, stopword removal, stemming, dan TF-IDF untuk mendapatkan fitur-fitur penting yang akan digunakan dalam klasifikasi sentimen.

2.3 Penerapan Naive Bayes

Pemrosesan Bahasa Alami (Natural Language Processing/NLP) merupakan cabang dari kecerdasan buatan yang berfokus pada bagaimana komputer dapat memahami, menafsirkan, dan memproses bahasa manusia secara efektif. Dalam konteks penelitian ini algoritma Naïve Bayes dimanfaatkan sebagai metode klasifikasi dalam NLP untuk mengelompokkan komentar siswa dan guru menjadi sentimen positif atau negatif [15]. Sebelum melakukan klasifikasi sentimen, dilakukan proses TF-IDF (Term Frekuensi - Inverse Document Frekuensi) untuk mengekstraksi fitur dari teks. TF-IDF merupakan metode yang mengukur seberapa penting suatu kata dalam dokumen dibandingkan dengan keseluruhan dataset.

$$tf - idf = tf \times \log \frac{N}{df} \quad (1)$$

Dimana :

tf : Term Frequency

idf : Inverse Document Frequency

n : Jumlah Total Dokumen

df : Jumlah Dokumen yang Mengandung term sebuah kata

Setelah itu dilakukan klasifikasi data dengan menggunakan metode Naïve Bayes Classifier, Dapat dirumuskan sebagai berikut:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (2)$$



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

Dimana :

B : Data dengan class yang belum diketahui.

A : Hipotesis data A yang merupakan suatu data class spesifik.

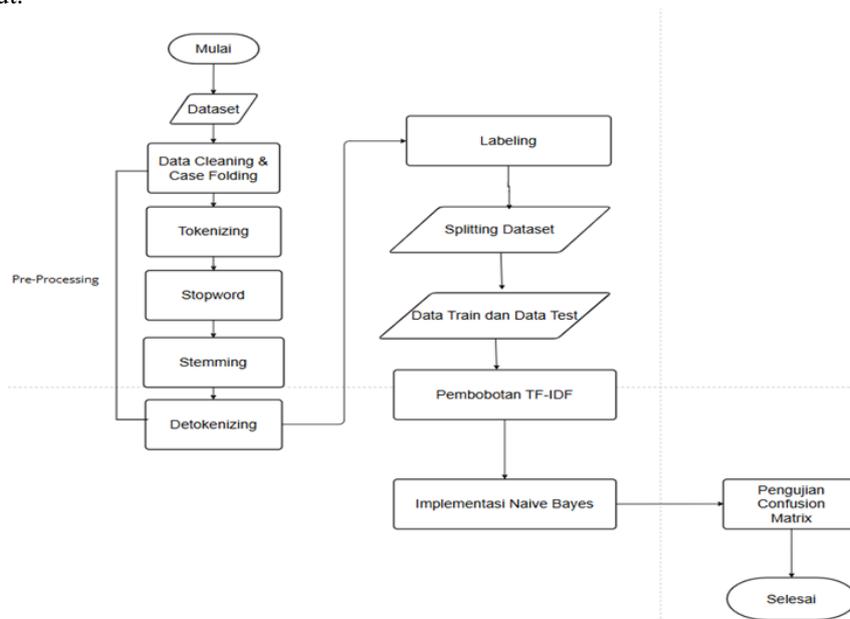
$P(A|B)$: Peluang kejadian A apabila B terjadi (posterior probabily).

$P(B|A)$: Peluang kejadian B apabila A terjadi (likelihood).

$P(B)$: Probabilitas dari B (prior probability).

$P(A)$: Probabilitas hipotesis A.

Algoritma Naive Bayes dapat dipahami dengan menggunakan rumus conditional probabilities (probabilitas bersyarat). Flowchart dapat mempermudah dalam menemukan dan menyelesaikan suatu masalah khususnya pada masalah yang perlu dievaluasi dan diperelajari lebih dalam. Untuk alur prosesnya dapat dilihat pada gambar flowchart berikut.



Gambar 3. Flowchart Algoritma Naive Bayes

Pada gambar 3. Menunjukkan alur bagan atau flowchart dalam penelitian dimana menjelaskan tentang tahap-tahap yang dilakukan dalam mengklasifikasi kelas sentimen kurikulum merdeka di Mts Al Ishlah Medan.

2.4 Pengujian

Tahap pengujian pada penelitian analisis sentimen ini menggunakan metode Confusion Matrix yang nantinya diperoleh nilai accuracy, precision, recall dan f-1 score dari klasifikasi yang telah dilakukan.

3. Hasil

Pada penelitian ini penulis melakukan proses pengumpulan data diambil dari Kuisioner yang diberikan langsung maupun secara online kepada responden menggunakan media google form yang diberikan pada guru dan kertas yang akan diisi oleh murid MTS Al Ishlah Medan. Data komentar diambil dengan menggunakan proses wawancara dan observasi. Data yang diambil dengan proses wawancara dan observasi tersebut akan disimpan kedalam file berekstensi .csv, total data yang diambil berjumlah 100 data komentar, dimana komentar membahas tentang sentimen Kurikulum Merdeka di Mts Al Ishlah Medan. Komentar nantinya akan diberikan label positif dan negatif dan netral menggunakan InsetLexikon dan Algoritma Naive Bayes dalam mencari accuracy, precision, recall, dan F1-Score untuk analisis hasil sentimenzx.

Tabel 1. Sampel Data Siswa dan Guru

No.	Full text
1	Saya cukup aktif dalam kurikulum merdeka tetapi kurang kebebasan dan terlalu banyak tugas
...	...
100	Kurangnya sosialisasi dan pelatihan mengajar model kurikulum merdeka



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

Langkah berikutnya dalam penelitian ini adalah menyaring data komentar untuk menghilangkan entri yang duplikat. Langkah ini bertujuan untuk memastikan setiap komentar yang dianalisis bersifat unik, sehingga hasil analisis menjadi lebih akurat dan representatif. Seperti pada tabel berikut yang awalnya data 100 menjadi 60 komentar.

Tabel 2. Sampel Data Akhir

No.	Full text
1	Saya cukup aktif dalam kurikulum merdeka tetapi kurang kebebasan dan terlalu banyak tugas
...	...
60	Kurangnya sosialisasi dan pelatihan mengajar model kurikulum merdeka

3.1 Pre-Processing Data

a. Text Cleaning, Lower Text, and Case Folding

Pada tahap cleaning, komentar dibersihkan dari karakter atau elemen yang tidak relevan untuk analisis sentimen. Proses ini meliputi penghapusan tanda baca, angka, emoji, URL, karakter khusus, serta spasi berlebih yang dapat mengganggu analisis teks. Setelah itu, dilakukan case folding, yaitu mengubah semua huruf menjadi huruf kecil (lowercase) untuk memastikan konsistensi dalam pemrosesan teks.

Tabel 3. Proses *Cleaning, Lower Text* Dan *Case Folding* Pada Text

No.	Full text	Clean
1	Saya cukup aktif dalam kurikulum merdeka tetapi kurang kebebasan dan terlalu banyak tugas	saya cukup aktif dalam kurikulum merdeka tetapi kurang kebebasan dan terlalu banyak tugas
...
60	Kurangnya sosialisasi dan pelatihan mengajar model kurikulum merdeka	kurangnya sosialisasi dan pelatihan mengajar model kurikulum merdeka

b. Tokenize

Pada tahap tokenizing, setiap komentar yang telah melewati proses cleaning dan case folding dipecah menjadi unit-unit kata yang lebih kecil, yang disebut token.

Tabel 4. Tokenize

No.	Clean	Tokenize
1	saya cukup aktif dalam kurikulum merdeka tetapi kurang kebebasan dan terlalu banyak tugas	['saya', 'cukup', 'aktif', 'dalam', 'kurikulum', 'merdeka', 'tetapi', 'kurang', 'kebebasan', 'dan', 'terlalu', 'banyak', 'tugas']
...
60	kurangnya sosialisasi dan pelatihan mengajar model kurikulum merdeka	['kurangnya', 'sosialisasi', 'dan', 'pelatihan', 'mengajar', 'model', 'kurikulum', 'merdeka']

c. Text Normalization (Slang Words)

Pada tahap normalisasi, kata-kata dalam komentar yang telah melewati proses tokenisasi disesuaikan dengan bentuk baku atau standar untuk memastikan konsistensi dalam analisis teks. Proses normalisasi ini menggantikan kata-kata gaul, singkatan, atau ejaan tidak standar dengan versi yang lebih umum digunakan dalam bahasa Indonesia. Misalnya, kata "gpp" diubah menjadi "tidak apa-apa", "btw" menjadi "ngomong-ngomong", dan "wkwk" dihapus karena tidak memiliki makna sentimen.

Tabel 5. Penerapan *Slang Words*

No.	Tokenize	Normalisasi
1	['saya', 'cukup', 'aktif', 'dalam', 'kurikulum', 'merdeka', 'tetapi', 'kurang', 'kebebasan', 'dan', 'terlalu', 'banyak', 'tugas']	['saya', 'cukup', 'aktif', 'dalam', 'kurikulum', 'merdeka', 'tetapi', 'kurang', 'kebebasan', 'dan', 'terlalu', 'banyak', 'tugas']
...
60	['kurangnya', 'sosialisasi', 'dan', 'pelatihan', 'mengajar', 'model', 'kurikulum', 'merdeka']	['kurangnya', 'sosialisasi', 'dan', 'pelatihan', 'mengajar', 'model', 'kurikulum', 'merdeka']

d. Stopwords

Pada tahap stopwords removal, kata-kata yang tidak memiliki makna signifikan dalam analisis sentimen dihapus dari teks. Stopwords adalah kata-kata umum seperti "yang", "dan", "di", "ke", "dari", yang tidak



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

memberikan kontribusi langsung terhadap sentimen suatu komentar. Dalam penelitian ini, stopwords dihapus menggunakan daftar kata dari NLTK atau sumber lainnya yang relevan dengan bahasa Indonesia.

Tabel 6. Stopwords

No.	Normalisasi	Stopwords
1	['saya', 'cukup', 'aktif', 'dalam', 'kurikulum', 'merdeka', 'tetapi', 'kurang', 'kebebasan', 'dan', 'terlalu', 'banyak', 'tugas']	['aktif', 'kurikulum', 'merdeka', 'kebebasan', 'tugas']
...
60	['kurangnya', 'sosialisasi', 'dan', 'pelatihan', 'mengajar', 'model', 'kurikulum', 'merdeka']	['kurangnya', 'sosialisasi', 'pelatihan', 'mengajar', 'model', 'kurikulum', 'merdeka']

e. Steaming

Pada tahap stemming, setiap kata dalam komentar dikonversi ke bentuk dasarnya (root word) untuk mengurangi variasi kata dengan makna yang sama. Proses ini dilakukan menggunakan algoritma Sastrawi, yang merupakan metode stemming populer untuk bahasa Indonesia. Misalnya, kata "membeli" diubah menjadi "beli", "berjalan" menjadi "jalan", dan "terbaik" menjadi "baik".

Tabel 7. Steaming

No.	Stopwords	Steaming
1	['aktif', 'kurikulum', 'merdeka', 'kebebasan', 'tugas']	['aktif', 'kurikulum', 'merdeka', 'bebas', 'tugas']
...
60	['kurangnya', 'sosialisasi', 'pelatihan', 'mengajar', 'model', 'kurikulum', 'merdeka']	['kurang', 'sosialisasi', 'latih', 'ajar', 'model', 'kurikulum', 'merdeka']

f. Detokenized Text

Detokenized text merupakan proses yang penggabungan kembali beberapa kata-kata menjadi bentuk kalimat lengkap dan dapat dibaca.

Tabel 8. Detokenized Text

No.	Steaming	Detokenized Text
1	['aktif', 'kurikulum', 'merdeka', 'bebas', 'tugas']	aktif kurikulum merdeka bebas tugas
60	['kurang', 'sosialisasi', 'latih', 'ajar', 'model', 'kurikulum', 'merdeka']	kurang sosialisasi latih ajar model kurikulum merdeka

3.2 Pemberian Label Menggunakan InsetLexicon

Tahap selanjutnya dalam analisis sentimen adalah pelabelan data secara otomatis menggunakan kamus InSetLexicon, yang berisi kata-kata dengan nilai polaritas dan bobot tertentu. Sistem menghitung total bobot dari kata-kata dalam teks ulasan untuk menentukan skor polaritasnya. Jika skor tersebut lebih dari nol, ulasan diklasifikasikan sebagai sentimen positif, sedangkan skor di bawah nol menunjukkan sentimen negatif. Metode ini memungkinkan identifikasi sentimen secara efisien tanpa intervensi manual.

Tabel 9. Nilai Polaritas Positive Corpus

No.	Word	Weight
1	Hai	3
2	Tetap	3
3	Detail	2
4	Bagus	2
...
3610	Orisinal	3

Tabel 10. Nilai Polaritas Negatif Corpus

No.	Word	Weight
1	Isak	-5
2	Sakit	-5



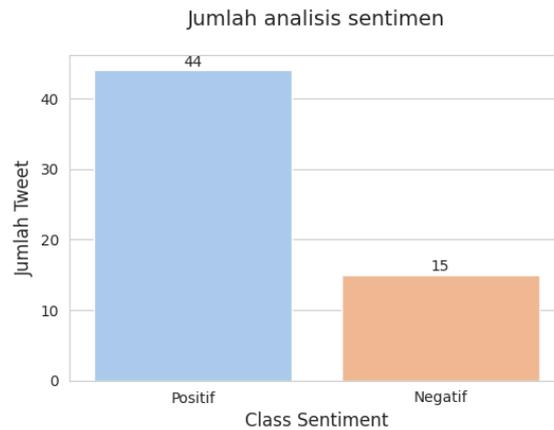
3	Mulu	-1
4	Gamau	-4
...
6610	Mencoreng	-4

Proses perhitungan dalam pelabelan sentimen menggunakan InsetLexicon, dimana jika hasil polaritas menunjukkan nilai positif maka sentimen dikategorikan menjadi positif, sebaliknya jika nilai polaritas menunjukkan nilai negatif, maka hasil pelabelan pada sentimen dikategorikan menjadi negatif.

Tabel 11. Hasil Pelabelan

Ulasan 1 : materi kurikulum mudah dimengerti							
	Materi	Kurikulum	Mudah	dimengerti			Polaritas
Bobot Kata	-	-	4	2			6
							Label
							Positif
Ulasan 2 : tidak menyenangkan karena pelajarannya sangat sulit							
	Tidak	Menyenangkan	Karena	Pelajarannya	Sangat	Sulit	Polaritas
Bobot Kata	-5	4	-3	-	3	-5	-6
							Label
							Nrgatif

Gambar 5 menunjukkan hasil analisis sentimen terhadap Kurikulum Merdeka di MTS Al Ishlah Medan menggunakan algoritma Naive Bayes dengan pelabelan sentimen yang dilakukan secara otomatis menggunakan kamus InSetLexicon. Dari total tweet yang dianalisis, sebanyak 44 tweet (sekitar 75%) dikategorikan sebagai sentimen positif, sementara 15 tweet (sekitar 25%) tergolong sentimen negatif. Hasil ini menunjukkan bahwa mayoritas tanggapan masyarakat terhadap penerapan Kurikulum Merdeka di MTS Al Ishlah Medan bersifat positif.

**Gambar 5.** Klasifikasi Kelas Sentimen

3.3 Pembobotan TF-IDF

Tahap awal yang dilakukan adalah menghitung nilai TF dari semua kata yang terdapat dalam teks (menghitung seberapa sering kata muncul). Menggunakan 3 kalimat yaitu :

- harap kurikulum merdeka maju dorong semangat ajar anak bangsa
- aktif kurikulum merdeka bebas tugas

Tabel 12. Term-Frequency (TF)

Kata	Frekuensi		DF (Total)	TF (Normalized)
	Dokumen 1	Dokumen 2		
harap	1	0	1	1/9=0.111
kurikulum	1	0	1	1/9=0.111
merdeka	1	0	1	1/9=0.111
maju	1	0	1	1/9=0.111
dorong	1	0	1	1/9=0.111



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

semangat	1	0	1	1/9=0.111
ajar	1	0	1	1/9=0.111
anak	1	0	1	1/9=0.111
bangsa	1	0	1	1/9=0.111
aktif	0	1	1	1/5=0.2
kurikulum	0	1	1	1/5=0.2
merdeka	0	1	1	1/5=0.2
bebas	0	1	1	1/5=0.2
tugas	0	1	1	1/5=0.2

Kemudian langkah selanjutnya menghitung nilai IDF menggunakan persamaan :

$idf_{(t,D)} = \log \frac{\text{Total Jumlah Dokumen}}{\text{Jumlah Dokumen dari TF}}$, sebagai berikut :

- $idf_{harap} = \log \frac{3}{1} = 0.477$
- $idf_{kurikulum} = \log \frac{3}{1} = 0.477$
- $idf_{merdeka} = \log \frac{3}{1} = 0.477$

Tabel 13. Perhitungan IDF

Kata	DF	IDF
harap	1	0.477
kurikulum	1	0.477
merdeka	1	0.477
maju	1	0.477
dorong	1	0.477
semangat	1	0.477
ajar	1	0.477
anak	1	0.477
bangsa	1	0.477
aktif	1	0.477
kurikulum	1	0.477
merdeka	1	0.477
bebas	1	0.477
tugas	1	0.477

Dan langkah berikutnya adalah perhitungan nilai TF dan IDF yang dihitung untuk masing-masing kata dalam komentar siswa dan guru di MTS Al-Ishlah Medan. Langkah perhitungan TF dan IDF untuk masing-masing kata ditunjukkan di sebagai berikut :

$tf - idf_{t,d} = tf_{td} \times idf_t$

- $tf - idf_{(harap,D1)} = tf_{harap,D1} \times idf_{harap} = 0.111 \times 0.477 = 0.053$
- $tf - idf_{(kurikulum,D1)} = tf_{kurikulum,D1} \times idf_{kurikulum} = 0.111 \times 0.477 = 0.053$
- $tf - idf_{(merdeka,D1)} = tf_{merdeka,D1} \times idf_{merdeka} = 0.111 \times 0.477 = 0.053$
- $tf - idf_{(aktif,D2)} = tf_{aktif,D2} \times idf_{aktif} = 0.2 \times 0.477 = 0.0954$
- $tf - idf_{(bebas,D2)} = tf_{bebas,D2} \times idf_{bebas} = 0.2 \times 0.477 = 0.0954$
- $tf - idf_{(tugas,D2)} = tf_{tugas,D2} \times idf_{tugas} = 0.2 \times 0.477 = 0.0954$

Tabel 14. Perhitungan TF-IDF

Kata	Dokumen	TF-IDF
harap	1	0.053
kurikulum	1	0.053
merdeka	1	0.053
maju	1	0.053
dorong	1	0.053
semangat	1	0.053
ajar	1	0.053
anak	1	0.053
bangsa	1	0.053
aktif	2	0.0954
kurikulum	2	0.0954
merdeka	2	0.0954
bebas	2	0.0954



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

Kata	Dokumen	TF-IDF
tugas	2	0.0954

3.4 Pengujian Algoritma Naïve Bayes

Jumlah data awal yang dianalisis adalah sebanyak 100 data dimana setelah dilakukan proses drop data sentimen yang memiliki kesamaan menjadi 60 data. Data ini dibagi menjadi dua bagian dengan perbandingan 80% untuk data latih dan 20% untuk data uji, yaitu dengan total 48 data latih dan 12 data uji. Proses pelatihan menggunakan metode pembobotan TF-IDF untuk menentukan bobot setiap kata berdasarkan frekuensinya dalam masing-masing kelas sentimen. Sebagai ilustrasi, digunakan 2 data ulasan sebagai data latih dan satu data sebagai data uji untuk menunjukkan alur proses klasifikasi sentimen menggunakan algoritma Naïve Bayes.

Tabel 15. Sampel Data Latih

Latih Komentar	Label
suka kurikulum merdeka	Positif
kurikulum merdeka sangat ribet	Negatif

Penentuan sentimen pada komentar dilakukan dengan menghitung probabilitas dokumen data uji berdasarkan distribusi kata yang telah diperoleh dari data latih. Klasifikasi ini dilakukan secara otomatis menggunakan algoritma Naïve Bayes, khususnya dengan memanfaatkan fungsi MultinomialNB. Dalam proses ini, bobot kata-kata dalam data uji dibandingkan dengan bobot kata-kata yang telah dihitung dari data latih. Dengan demikian, setiap dokumen uji akan diklasifikasikan ke dalam kelas positif dan negatif berdasarkan proporsi probabilitas tertinggi yang dihasilkan dari model yang telah dilatih sebelumnya.

Tabel 16. Pembobotan Kata pada Data Latih

No.	Kosa Kata	Tf(positif)	Tf(Negatif)
1.	suka	1	0
2.	kurikulum	1	0
3.	merdeka	1	0
4.	kurikulum	0	1
5.	merdeka	0	1
6.	sangat	0	1
7.	ribet	0	1
Jumlah Term		3	4

Dari hasil perhitungan, diketahui bahwa nilai term frequency (TF) untuk kelas positif adalah 3 dan untuk kelas negatif adalah 4, dengan total 7 kata. Proses klasifikasi dimulai dengan menghitung probabilitas awal (prior probability), diikuti oleh probabilitas bersyarat dari setiap kata terhadap kelasnya, dan diakhiri dengan perhitungan probabilitas posterior untuk menentukan kelas sentimen akhir. Langkah-langkah ini merupakan inti dari penerapan algoritma Naïve Bayes dalam proses pelatihan data latih.

1. Perhitungan nilai probabilitas

$$P(\text{Kelas} | \text{Komentar}) = \frac{\text{Jumlah Kelas } X}{\text{Jumlah Komentar}}$$

Dengan menggunakan persamaan di atas, kita akan memperoleh probabilitas setiap kelas dalam sentimen.

$$P(\text{Positif} | \text{Komentar}) = \frac{1}{2} = 0.5$$

$$P(\text{Negatif} | \text{Komentar}) = \frac{1}{2} = 0.5$$

2. Perhitungan Nilai Probabilitas Bersyarat :

$$P(\text{Term} | \text{Kelas}) = \frac{\text{Total bobot TF - IDF Term pada Kelas} + 1}{\text{Bobot TF - IDF Kelas} + \text{Total Bobot TF - IDF}}$$

Dengan menggunakan persamaan di atas, kita akan memperoleh probabilitas suku-suku di setiap kelas sentimen.

Probabilitas kata "suka"

$$P(\text{suka} | \text{Positif}) = \frac{1+1}{3+7} = \frac{2}{10} = 0.2$$

$$P(\text{suka} | \text{Negatif}) = \frac{0+1}{4+7} = \frac{1}{11} = 0.09$$

Probabilitas kata "kurikulum"

$$P(\text{kurikulum} | \text{Positif}) = \frac{1+1}{3+7} = \frac{2}{10} = 0.2$$

$$P(\text{kurikulum} | \text{Negatif}) = \frac{1+1}{4+7} = \frac{2}{11} = 0.19$$

Probabilitas kata "merdeka"

$$P(\text{merdeka} | \text{Positif}) = \frac{1+1}{3+7} = \frac{2}{10} = 0.2$$



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

$$P(\text{merdeka} | \text{Negatif}) = \frac{1+1}{4+7} = \frac{2}{11} = 0.19$$

Probabilitas kata “ribet”

$$P(\text{ribet} | \text{Positif}) = \frac{0+1}{3+7} = \frac{1}{10} = 0.1$$

$$P(\text{ribet} | \text{Negatif}) = \frac{1+1}{4+7} = \frac{1}{11} = 0.09$$

Selanjutnya adalah mengambil data uji yaitu dengan klasifikasi data uji dengan mengalikan semua peluang. Nilai yang lebih tinggi merupakan kelas baru dari data tersebut.

Tabel 17. Sampel Data Uji

Test Sentimen	Label
["kurikulum", "merdeka", "mudah", "dimengerti"]	?

Pada data uji “kurikulum merdeka mudah dimengerti” yang termasuk ke dalam data *training* adalah kata “kurikulum” dan “merdeka”.

3. Perhitungan nilai probabilitas posterior

$$P(\text{Uji} | \text{Positif}) = P(\text{Komentar} | \text{Kelas}) = P_{Term 1} \times \dots \times P_{Term n} \times P(\text{Kelas} | \text{Komentar})$$

$$= P(\text{positif}) \times P(\text{kurikulum} | \text{positif}) \times P(\text{merdeka} | \text{positif})$$

$$= 0.5 \times 0.2 \times 0.2$$

$$= 0.02$$

$$P(\text{Uji} | \text{Negatif}) = P(\text{positif}) \times P(\text{kurikulum} | \text{negatif}) \times P(\text{merdeka} | \text{negatif})$$

$$= 0.5 \times 0.19 \times 0.19$$

$$= 0.01805$$

Kesimpulan yaitu Nilai probabilitas tertinggi yaitu sebesar 0.08 pada P (Uji | Positif) sehingga komentar tersebut diklasifikasikan ke dalam kelas “Positif”.



Gambar 6. Wordcloud Sentimen Positif



Gambar 7. Wordcloud Sentimen Negatif

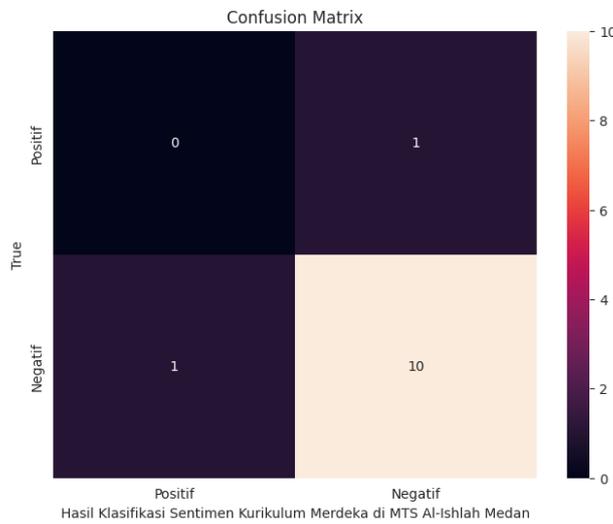
Setelah pengujian sentimen dilakukan menggunakan algoritma Naïve Bayes, sistem menghasilkan label sentimen untuk setiap data uji. Label ini kemudian dibandingkan dengan label asli untuk mengevaluasi kinerja model.



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

Pengukuran dilakukan melalui metrik evaluasi seperti akurasi, presisi, recall, dan F1-score pada dataset sentimen terkait topik kurikulum merdeka.



Gambar 9. Confusion Matrix

Hasil klasifikasi untuk mencari *accuracy*, *precision*, *recall* dan *F1-score* mengikuti rumus berikut:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{0 + 10}{0 + 10 + 1 + 1} = 0.83$$

metrik tersebut juga untuk kelas **Negatif**, berikut hasilnya:

$$precision = \frac{TN}{TN + FN} = \frac{10}{10 + 1} = 0.91$$

$$recall = \frac{TN}{TN + FP} = \frac{10}{10 + 1} = 0.91$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2 \times 0.91 \times 0.91}{0.91 + 0.91} = 0.91$$

```

Accuracy: 0.8333333333333334
      precision  recall  f1-score  support
Negatif      0.00      0.00      0.00         1
Positif      0.91      0.91      0.91        11

accuracy                0.83         12
macro avg      0.45      0.45      0.45         12
weighted avg   0.83      0.83      0.83         12
    
```

Gambar 10. Hasil Confusion Matrix



4. Pembahasan

Pada penelitian ini terhadap sentimen “Kurikulum Merdeka” yang diperoleh melalui kuisisioner, wawancara, dan observasi dari siswa serta guru di MTS Al Ishlah Medan menunjukkan bahwa hasil klasifikasi menggunakan metode Naïve Bayes menghasilkan nilai akurasi sebesar 83% atau 0.83, precision dan recall sebesar 91% atau 0.91, serta F1-score sebesar 91% atau 0.91 untuk kelas negatif. Berdasarkan data dan hasil analisis dapat disimpulkan bahwa mayoritas responden memberikan tanggapan positif terhadap Kurikulum Merdeka, dengan jumlah sentimen positif sebanyak 44 komentar (75%) dan sentimen negatif sebanyak 15 komentar (25%). Temuan ini menunjukkan bahwa Kurikulum Merdeka mendapat respons yang cukup baik dari warga sekolah. Hasil ini berbeda dengan penelitian Refandah Puspitasari dan Aries Dwi Indriyanti (2024) yang menganalisis sentimen opini publik terhadap kebijakan skripsi di Twitter, di mana akurasi tertinggi hanya mencapai 77% dengan dominasi sentimen negatif. Pada penelitian ini, algoritma Naïve Bayes diimplementasikan dengan pelabelan otomatis menggunakan kamus InSetLexicon dan pembobotan TF-IDF, serta pengujian dilakukan menggunakan metode Confusion Matrix. Akurasi tertinggi diperoleh pada rasio pembagian data latih dan uji sebesar 80:20. Berdasarkan hasil tersebut, dapat disimpulkan bahwa algoritma Naïve Bayes cukup efektif dalam mengklasifikasikan sentimen pendidikan khususnya terhadap Kurikulum Merdeka di lingkungan MTS Al Ishlah Medan.

5. Kesimpulan

Analisis sentimen dengan algoritma Naïve Bayes berhasil melakukan klasifikasi terhadap sentimen siswa dan guru di MTS Al Ishlah Medan mengenai penerapan Kurikulum Merdeka. Dari proses pengumpulan data melalui kuisisioner, wawancara, dan observasi, diperoleh sebanyak 100 data komentar, yang kemudian disaring menjadi 60 data unik untuk dianalisis. Setelah dilakukan proses klasifikasi, diketahui bahwa 44 data tergolong ke dalam sentimen positif dan 15 data tergolong sentimen negatif. Hasil klasifikasi menggunakan metode Naïve Bayes Classifier memperoleh nilai akurasi sebesar 83% atau 0.83, precision, recall dan F1-score sebesar 91% atau 0.91 untuk kelas negatif. Berdasarkan data dan hasil analisis, dapat disimpulkan bahwa mayoritas warga sekolah memberikan tanggapan positif terhadap penerapan Kurikulum Merdeka. Hal ini menunjukkan bahwa kebijakan pendidikan ini diterima dengan baik oleh guru dan siswa di lingkungan MTS Al Ishlah Medan.

6. Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam penyusunan jurnal berjudul "Analisis Sentimen terhadap Kurikulum Merdeka di MTS Al Ishlah Medan Menggunakan Algoritma Naïve Bayes." Ucapan terima kasih disampaikan kepada redaksi jurnal JNASEK yang telah menerima dan mempublikasikan artikel ini, serta kepada kedua orang tua yang selalu memberikan doa dan semangat. Penulis juga menyampaikan apresiasi sebesar-besarnya kepada dosen pembimbing yang telah membimbing dan mengarahkan selama proses penelitian berlangsung, serta kepada guru dan siswa MTS Al Ishlah



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

Medan yang telah berpartisipasi dalam pengumpulan data. Tak lupa, ucapan terima kasih ditujukan kepada teman-teman seperjuangan yang telah memberikan dukungan moral dan motivasi selama penyusunan jurnal ini.

REFERENSI

- [1] Rudi Hartono, I Wayan Suastra, and I Wayan Lasmawan, “Implementasi Kurikulum Merdeka Dalam Melestarikan Budaya Nusantara,” *EDUKASIA: Jurnal Pendidikan dan Pembelajaran*, vol. 4, no. 2, pp. 823–828, 2023, doi: 10.62775/edukasia.v4i2.356.
- [2] N. A. Fatirul, “Metode Penelitian Pengembangan Bidang Pembelajaran (Edisi Khusus Mahasiswa Pendidikan dan Pendidik),” *Sanskara Pendidikan dan Pengajaran*, vol. 1, no. 02, pp. 56–67, 2022, doi: 10.58812/spp.v1.i02.
- [3] D. A. Ramdani, A. Hasanah, and B. S. Arifin, “Core Ethical Values Pendidikan Karakter Berbasis Nilai-Nilai Islam,” *Jiip - Jurnal Ilmiah Ilmu Pendidikan*, vol. 6, no. 10, pp. 7891–7899, 2023, doi: 10.54371/jiip.v6i10.3010.
- [4] Q. A’yun, M. B. A. Pamungkas, I. S. D. Agustin, I. Zahroh, R. G. Afandi, and Z. Zulkarnaen, “Penerapan nilai iman, takwa dan akhlak mulia profil pelajar Pancasila dalam membentuk karakter islami siswa,” *Tawazun: Jurnal Pendidikan Islam*, vol. 16, no. 1, pp. 9–20, 2023, doi: 10.32832/tawazun.v16i1.8650.
- [5] F. Juwita, M. Masudi, and S. Zulaiha, “Pelaksanaan Proyek Penguatan Profil Pelajar Pancasila Pada Kurikulum Merdeka Dalam Menanamkan Karakter Siswa Di Sdit Cahaya Rabbani Kepahiang,” 2024.
- [6] Alfandi Safira and F. N. Hasan, “Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier,” *ZONAsi: Jurnal Sistem Informasi*, vol. 5, no. 1, pp. 59–70, 2023, doi: 10.31849/zn.v5i1.12856.
- [7] N. Mayasari, C. Rizal, and F. Wulandari, “Penerapan Metode Klasifikasi Karakteristik Kepribadian Berbasis Website,” vol. 06, no. 02, pp. 27–30, 2024.
- [8] N. Khoirunnisaa, K. N. N. Kesuma, S. Setiawan, and A. Y. P. Yusuf, “Klasifikasi Teks Ulasan Aplikasi Netflix Pada Google Play Store Menggunakan Algoritma Naive Bayes dan SVM,” *SKANIKA: Sistem Komputer Dan Teknik Informatika*, vol. 7, no. 1, pp. 64–73, 2024.
- [9] W. Erika, “Application of Data Mining to Predict Sales of the Best-Selling Dolls at Chudalla Store Using the K-Nearest Neighbor (K-NN) Method,” vol. 3, no. 1, 2023.
- [10] R. Puspitasari, “KEBIJAKAN BARU SKRIPSI PADA MEDIA SOSIAL,” vol. 05, no. 03, pp. 37–42, 2024.
- [11] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, “Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm,” *Procedia Comput Sci*, vol. 161, pp. 765–772, 2019, doi: 10.1016/j.procs.2019.11.181.
- [12] S. Batubara, T. Wahyono, A. Khaliq, Y. Ennisa Lubis, and D. Julia Sari, “Website-Based Online Job Training Application Design Using the Unified Modeling Language,” *Sciences Development and Technology*, vol. 2, no. 1, 2022.



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

- [13] Z. Syahputra, M. I. Sarif, R. S. Ritonga, A. Husniyyah, U. Pembangunan, and P. Budi, “E-KEDE APPLICATION DEVELOPMENT IN INCREASING UMKM SALES BASED ON SDLC,” vol. 4307, no. 4, pp. 1605–1610, 2024.
- [14] I. Ramadhan, “Kurikulum Merdeka : Proses Adaptasi dan Pembelajaran di Sekolah Menengah Atas,” *Journal of Education Research*, vol. 4, no. 4, pp. 1846–1853, 2023.
- [15] A. Badawi, “the Effectiveness of Natural Language Processing (Nlp) As a Processing Solution and Semantic Improvement,” *International Journal of Economic, Technology and Social Sciences (Injects)*, vol. 2, no. 1, pp. 36–44, 2021, doi: 10.53695/injects.v2i1.194.



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.