BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 4, June 2025 | Hal 813-820
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i4.530

Implementasi Algoritma Random Forest untuk Analisis Sentimen Ulasan Pengguna Aplikasi Merdeka Mengajar

Yuwan Jumaryadi*, Ruci Meiyanti, Riri Fajriah, Athiyyah Nisrina Mahsyar, Puspita Sari Anggraeni

Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Mercu Buana, Jakarta, Indonesia
Email: [1,*]yuwan.jumaryadi@mercubuana.ac.id, [2]ruci@mercubuana.ac.id, [3]riri.fajriah@mercubuana.ac.id,
[4]nisrinathiyyah@gmail.com, [5]puspitasarianggraeni9c@gmail.com
Email Penulis Korespondensi: yuwan.jumaryadi@mercubuana.ac.id

Abstrak−Pendidikan mempunyai peran besar dalam menentukan kualitas sumber daya manusia. Peranan guru amatlah penting sebagai pendidik yang memberikan bimbingan dan pembelajaran. Sebagai upaya dalam memudahkan para guru untuk melakukan tugas dan tanggung jawabnya terutama pada kurikulum merdeka, Kemendikbud mengembangkan aplikasi bernama Merdeka Mengajar. Namun belum adanya metode untuk mengklasifikasikan sentiment atau opini dari data komentar pada survei kepuasan pengguna aplikasi merdeka mengajar di google playstore, guna mengetahui sejauh mana kepuasan pengguna terhadap aplikasi merdeka mengajar. Penelitian ini bertujuan untuk mengamati analisis sentimen mengenai opini pemakai terhadap aplikasi merdeka mengajar di google playstore menggunakan algoritma Random Forest, SVM dan Naïve Bayes dengan menggunakan pembobotan TF-IDF untuk proses klasifikasi. Pengkajian ini memakai data sekunder yang berasal melalui ulasan pengguna aplikasi merdeka mengajar dan diklasifikasi dengan metode Random Forest, SVM, dan Naïve Bayes. Hasil dari pengklasifikasian tersebut menunjukkan bahwa, algoritma Random Forest merupakan algoritma terbaik dalam meprediksi ulasan pengguna apliksi merdeka mengajar dibandingkan dengan Naive Bayes dan SVM.

Kata Kunci: Analisis Sentimen; Merdeka Mengajar; Naive Bayes; SVM; Random Forest

Abstract−Education plays a major role in determining the quality of human resources. The role of teachers is very important as educators who provide guidance and learning. As an effort to facilitate teachers to carry out their duties and responsibilities, especially in the Merdeka Mengajar curriculum, the Ministry of Education and Culture has developed an application called Merdeka Mengajar. However, there is no method to classify sentiment or opinions from comment data on the Merdeka Mengajar application user satisfaction survey on the Google Playstore, in order to determine the extent of user satisfaction with the Merdeka Mengajar application. This study aims to observe sentiment analysis regarding user opinions on the Merdeka Mengajar application on the Google Playstore using the Random Forest, SVM and Naïve Bayes algorithms using TF-IDF weighting for the classification process. This study uses secondary data derived from user reviews of the Merdeka Mengajar application and is classified using the Random Forest, SVM, and Naïve Bayes methods. The results of the classification show that the Random Forest algorithm is the best algorithm in predicting Merdeka Mengajar application user reviews compared to Naive Bayes and SVM.

Keywords: Sentiment Analysis; Merdeka Mengajar; Naive Bayes; SVM; Random Forest

1. PENDAHULUAN

Berdasarkan penelitian PISA dari tahun 2000 hingga 2018, bahwa pendidikan di Indonesia masih jauh dari sempurna [1]. Akan tetapi pesatnya kemajuan teknologi membuat penggunaan teknologi dalam pendidikan menjadi semakin penting untuk meningkatkan kompetensi guru dan siswa. Kurikulum merdeka yang diterapkan oleh Kemendikbud pada tahun 2022 bertujuan untuk meningkatkan kualitas tenaga didik dan siswa, sejalan dengan peran pendidikan dalam menentukan kualitas sumber daya manusia. Guru memiliki peran krusial dalam pelaksanaan pendidikan, dengan kemampuan pedagogik yang mencakup pengelolaan pembelajaran dan pemanfaatan teknologi untuk meningkatkan kualitas pembelajaran [2].

Peran utama pemberian rating aplikasi Merdeka Mengajar di Google Playstore sangat penting dalam memberikan layanan kepada pengguna. Dalam memastikan bahwa kebutuhan pengguna terpenuhi, maka aplikasi tersebut terus berkembang sesuai dengan perkembangan teknologi informasi. Saat ini, ada tren yang meningkat dalam literatur ilmiah terkait prediksi tingkat kepuasan pengguna terhadap aplikasi menggunakan teknik Data Mining [3]. Dalam era Big Data, minat terhadap Data Science dan Machine Learning semakin berkembang [4], dan analisa sentimen ulasan pengguna menjadi semakin relevan dan penting [5].

Dalam konteks aplikasi Merdeka Mengajar di Google Playstore, analisis sentimen ulasan pengguna menjadi penting untuk memahami tanggapan dan persepsi pengguna terhadap layanan yang disediakan [6]. Kepuasan pengguna merupakan indikator utama dalam menilai kualitas layanan, sedangkan keluhan pengguna dapat menjadi sinyal adanya permasalahan dalam penyelenggaraan layanan. Melalui penerapan algoritma Naïve Bayes untuk analisis sentimen, dapat diidentifikasi pola-pola dalam ulasan pengguna yang mencerminkan tingkat kepuasan atau ketidakpuasan terhadap aplikasi Merdeka Mengajar. Penelitian ini akan membandingkan kinerja algoritma Random Forest dengan algoritma Naïve Bayes dan SVM.

Hasil survei pengguna Aplikasi Merdeka Mengajar tahun 2024 di Google Playstore menunjukkan bahwa aplikasi ini sudah mencapai lebih dari 1 juta unduhan secara jumlah ulasan meraih 168 ribu. Meskipun banyak komentar yang berisi kritik dan saran konstruktif, tidak sedikit pula ulasan negatif yang menyoroti kurangnya kualitas yang memenuhi harapan pengguna. Meski demikian, hasil survei tersebut belum mampu memberikan informasi yang cukup spesifik bagi pemilik layanan untuk memantau sentimen pengguna secara menyeluruh terhadap Aplikasi Merdeka Mengajar.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 4, June 2025 | Hal 813-820
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i4.530

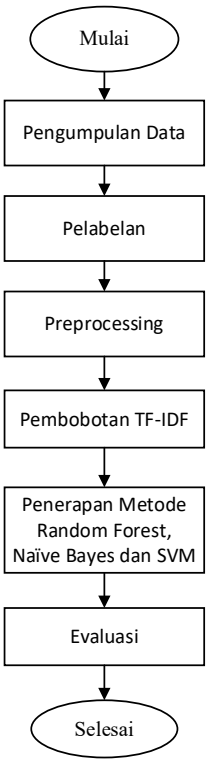
Komentar-komentar negatif dan kritik yang diberikan oleh pengguna memberikan gambaran bahwa masih terdapat beberapa aspek yang perlu ditingkatkan dalam aplikasi ini. Mungkin terdapat masalah terkait fungsionalitas, tampilan antarmuka, atau kualitas konten yang disajikan. Dengan demikian, upaya lebih lanjut diperlukan untuk menganalisis ulasan-ulasan tersebut secara lebih mendalam guna mendapatkan wawasan yang lebih detail mengenai kebutuhan dan harapan pengguna terhadap Aplikasi Merdeka Mengajar.

Ulasan yang dibagikan pemakai Aplikasi Merdeka Mengajar memiliki dampak yang signifikan terhadap keputusan penggunaan aplikasi ini oleh pegawai. Ulasan juga mempengaruhi reputasi aplikasi, sehingga ulasan dan rating yang buruk dapat berdampak negatif bagi penyedia layanan. Karena pengguna aplikasi memiliki ekspektasi yang tinggi terhadap layanan yang diberikan, data ulasan dan komentar dari hasil survei menjadi penting untuk membantu pemilik Aplikasi Merdeka Mengajar dalam menyusun dan merencanakan peningkatan layanan. Studi empiris menunjukkan bahwa ulasan pengguna yang panjang dan mudah dipahami, terutama dengan ulasan yang sedikit negatif, dapat meningkatkan kemungkinan perbaikan dan pembaharuan layanan Aplikasi Merdeka Mengajar di masa mendatang.

Pemanfaatan metode Naive Bayes Classifier diharapkan dapat memberikan kontribusi dalam menganalisis komentar pada hasil survei kepuasan pemakai Aplikasi Merdeka Mengajar di Google Play Store. Saat ini, komentar melalui pemakai aplikasinya belum dikaitkan dengan hasil survei yang telah dilakukan sebelumnya. Sebelumnya, tingkat kepuasan pengguna dievaluasi secara kuantitatif menggunakan skala likert. Dengan menggunakan metode analisis sentimen seperti Naive Bayes Classifier dan Random Forest, diharapkan dapat memberikan analisa lebih dalam mengenai persepsi dan sentimen pengguna terhadap Aplikasi Merdeka Mengajar.

2. METODOLOGI PENELITIAN

Dalam penerapan metode ini dilakukan langkah-langkah untuk menganalisis opini masyarakat tentang Aplikasi Merdeka Mengajar. Berikut penerapan metode seperti pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Pada Langkah ini dilakukan scraping data dengan menggunakan kata kunci “Merdeka Mengajar”. Adapun data yang didapatkan sebanyak 4993 data.

2.2 Pelabelan

Tahapan pelabelan data merupakan tahapan dimana ditentukannya suatu keterangan sentimen positif dan negatif pada setiap tweet [7]. Proses pelabelan dilakukan dengan library TextBlob. TextBlob digunakan untuk melakukan pemprosesan terhadap data yang sifatnya tekstual yang salah satunya digunakan untuk kasus analisis sentimen [8].

TextBlob menentukan sentimen sebuah teks dengan menggunakan analisis leksikal berbasis kamus (lexicon-based), bukan machine learning. TextBlob memanfaatkan library Pattern untuk menganalisis kata-kata dalam teks dan

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 4, June 2025 | Hal 813-820
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i4.530

menghitung Polaritas dan Subjektivitas. Pada Tabel 1 merupakan rumus untuk menentukan klasifikasi sentimen berdasarkan nilai polaritas.

Tabel 1. Klasifikasi Polaritas

Rentang Polaritas	Klasifikasi Sentimen
> 0	Positif
< 0	Negatif
= 0	Netral

2.3 Preprocessing

Preprocessing merupakan tahap awal dalam pengolahan data teks yang bertujuan untuk mengubah data yang belum terstruktur menjadi bentuk yang lebih terorganisir sehingga siap untuk dianalisis. Dalam proses text processing, terdapat beberapa langkah penting sebagai berikut:

a. *Case Folding*
   Case folding adalah langkah awal dalam pengolahan teks yang meskipun sederhana dan efisien, sering diabaikan. Proses ini mengubah semua huruf dalam teks menjadi huruf kecil, sehingga karakter huruf kapital seperti "A" hingga "Z" diubah menjadi "a" hingga "z" [9].
b. *Cleansing*
   Cleansing merupakan tahap pembersihan data dengan mengevaluasi dan memperbaiki kualitas teks. Ini dilakukan dengan menghapus tanda baca, karakter non-huruf, mention, email, dan URL, sehingga menghasilkan data yang bersih dan berkualitas tinggi [10].
c. *Tokenizing*
   Tokenizing merupakan proses di mana sekumpulan kata yang tersusun dalam sebuah kalimat akan dipisahkan menjadi bagian-bagian kata tunggal atau dalam bentuk token [11].
d. *Slang word*
   Slang word merupakan proses penyesuaian kata-kata tidak baku atau singkatan dalam teks. Kata-kata tersebut dikumpulkan ke dalam kamus slang word, kemudian diganti dengan padanan yang sesuai dengan kaidah Bahasa Indonesia [12].
e. *Stop word*
   Stop word mengacu pada kata-kata yang kurang memberikan kontribusi makna dalam suatu dokumen [13]. Proses ini bertujuan untuk menghapus kata-kata yang tidak relevan atau tidak penting dalam kalimat. Kata-kata ini biasanya sudah terdaftar dalam pustaka Sastrawi [14].
f. *Stemming*
   Stemming adalah proses untuk mengubah kata berimbuhan menjadi bentuk dasarnya dengan cara menghapus awalan maupun akhiran pada kata tersebut. Tahap ini menggunakan library Sastrawi sebagai alat bantu [15][16].

2.3 TF-IDF

TF-IDF (*Term Frequency-Inverse Document Frequency*) merupakan sebuah metode yang digunakan untuk menentukan seberapa sering sebuah kata muncul dalam suatu dokumen, sekaligus mengukur seberapa jarang kata tersebut ditemukan di seluruh kumpulan dokumen [17]. Teknik ini berfungsi untuk menilai tingkat relevansi suatu kata terhadap isi dokumen. TF-IDF merupakan algoritma yang umum digunakan dalam analisis big data. Algoritma ini memberikan nilai bobot pada setiap kata kunci dalam tiap kategori guna mengidentifikasi kesesuaian antara kata kunci dan kategori yang tersedia. TF-IDF sangat penting dalam proses klasifikasi teks karena berfungsi sebagai teknik ekstraksi fitur, yang mengubah dokumen teks menjadi representasi numerik yang bisa diproses oleh algoritma machine learning.

2.4 Random Forest

Random Forest adalah salah satu algoritma ensemble learning yang banyak digunakan dalam berbagai aplikasi data mining dan machine learning karena keandalannya dalam menghasilkan prediksi yang akurat dan kemampuannya menangani data dengan dimensi tinggi serta variabel yang saling berinteraksi [18]. Salah satu keunggulan utama Random Forest adalah kemampuannya dalam mengurangi overfitting yang sering terjadi pada decision tree tunggal, berkat penggunaan teknik bootstrap aggregating (bagging) dan pemilihan fitur secara acak pada setiap pemisahan node [19]. Selain itu, Random Forest juga menyediakan informasi penting mengenai pentingnya variabel, yang sangat berguna dalam proses seleksi fitur. Berbagai penelitian menunjukkan bahwa Random Forest secara konsisten memberikan performa yang kompetitif atau bahkan lebih baik dibandingkan algoritma lain dalam banyak kasus, termasuk dalam bidang bioinformatika, keuangan, dan sistem rekomendasi. Namun, meskipun unggul dalam akurasi, interpretabilitas model ini cenderung lebih rendah dibandingkan dengan model linear atau pohon tunggal, yang menjadi salah satu keterbatasan utamanya.

2.5 Evaluasi

Penelitian ini melakukan pengujian dengan menggunakan metode Random Forest, SVM, dan Naive Bayes untuk menilai tingkat akurasi, presisi, dan recall dari masing-masing metode. Evaluasi dilakukan dengan menggunakan confusion

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 4, June 2025 | Hal 813-820
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i4.530

matrix, yaitu sebuah tabel yang umum digunakan dalam machine learning untuk menilai kinerja model klasifikasi. Confusion matrix membandingkan nilai aktual dengan hasil prediksi model dan mencakup sembilan kemungkinan hasil klasifikasi, yaitu TP, FN1, FNt1, FP1, TN, FNt1, FNt2, FP2, FN2, dan TNt. Penjelasan mengenai *confusion matrix* yang digunakan dalam penelitian ini disajikan pada Tabel 2.

Tabel 2. *Confusion Matrix*

		Actual Values		
		Positive	Negative	Neutral
Predicted Values	Positive	TP (True Positive)	FN1 (False Negative1)	FNt1 (False Netral1)
	Negative	FP1 (False Positive1)	TN (True Negative)	FNt2 (False Netral2)
	Neutral	FP2 (False Positive2)	FN2 (False Negative2)	TNt (True Netral)

Setelah jumlah data yang diperoleh dari tabel *confusion matrix*, dari itu dapat dihitung nilai akurasi, presisi dan *recall* sebagai berikut:

2.2.1. Akurasi

Akurasi merupakan ukuran yang menunjukkan seberapa tepat model dalam mengklasifikasikan data, dihitung dengan membagi jumlah prediksi benar untuk kelas positif dan negatif dengan total keseluruhan data dalam dataset. Rumus perhitungan akurasi dapat dilihat pada Persamaan (1).

$$Akurasi = \frac{TP+TN}{TP+FP+TN+FN} \quad x\ 100\% \tag{1}$$

2.2.2. Presisi

Presisi mengukur proporsi data yang diprediksi sebagai positif yang benar-benar termasuk dalam kategori positif. Nilai ini diperoleh dari rasio antara jumlah data positif yang terklasifikasi dengan benar terhadap seluruh data yang diprediksi sebagai positif [20]. Rumus presisi dapat dilihat pada Persamaan (2).

$$Presisi = \frac{TP}{TP+FP} \quad x\ 100\% \tag{2}$$

2.2.3. Recall

Recall adalah menggambarkan persentase data kategori positif yang terklasifikasi dengan benar oleh sistem [20]. Adapun rumus perhitungan dari *recall* dapat dilihat pada persamaan (3).

$$Presisi = \frac{TP}{TP+FN} \quad x\ 100\% \tag{3}$$

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Proses penghimpunan informasi dan data guna sebuah pengkajian disebut teknik pengumpulan data. Targetnya guna menghimpun data yang dibutukan guna meraih target pengkajian. Peneliti harus mahir dalam prosedur pengumpulan data karena mereka bertujuan untuk mengatur data yang dapat diandalkan untuk penelitian. Oleh karena itu, penting untuk memahami metode pengumpulan data yang akan diterapkan dalam suatu penelitian.

Teknik pengumpulan data mencakup 2 golongan berupa kuantitatif & kualitatif. Data kuantitatif adalah informasi yang diperoleh dari temuan penelitian yang terorganisir yang dikumpulkan sebagai data numerik atau angka, yang akan dianalisis oleh perangkat lunak analitis. Data kualitatif, di sisi lain adalah informasi deskriptif dan tidak terstruktur yang diberikan secara verbal, bukan numerik, dan dikumpulkan melalui observasi objek atau wawancara. Pendekatan kuantitatif merupakan strategi pengumpulan data untuk penelitian ini, dan datanya bersumber dari Google Playstore. Google adalah perusahaan yang menjalankan dan mengembangkan platform distribusi digital Google Playstore. Selain memberikan pengguna akses ke media digital yang tersedia bagi pengguna Android, Google Playstore menawarkan fitur yang memungkinkan pengguna mengevaluasi dan meninjau aplikasi yang mereka gunakan. Pengujian model digunakan untuk mengetahui model kinerja terhadap model klasifikasi dengan metode tertentu. Terdapat dataset yang digunakan untuk pelatihan dan pengujian dalam melakukan pengujian model. Semakin banyak data yang digunakan maka akan menjadi lebih baik juga model tersebut. Data training digunakan sebagai pelatihan model sedangkan untuk data testing digunakan sebagai pengujian model [20]. Adapun data ulasan pengguna aplikasi merdeka mengajar pada tahun 2024 sebanyak 6783 ulasan.

3.2 Preprocessing

Tahapan preprocessing digunakan untuk pembersihan data dan juga penyederhanaan data sehingga menjadi data yang terstruktur. Preprocessing dilakukan melalui beberapa proses tahapan diantaranya cleaning data, case folding, tokenization, normalization, stopword, dan stemming. Pada tahap preprocessing ada 503 data ulasan yang dihapus,

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 4, June 2025 | Hal 813-820
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i4.530

sehingga hanya 6281 data ulasan yang akan digunakan untuk proses klasifikasi. Adapun contoh hasil dari tahapan preprocessing dapat dilihat pada Tabel 3.

Tabel 3. Hasil Preprocessing Data

Tahapan Preprocessing	Data Tweet
Data awal	Bagus banget, informatif sekali, aplikasi sangat bagus
Cleaning data	Bagus banget informatif sekali aplikasi sangat bagus
Case folding	bagus banget informatif sekali aplikasi sangat bagus
Tokenization	['bagus', 'banget', 'informatif', 'sekali', 'aplikasi', 'sangat', 'bagus']
Stopword removal	['bagus', 'banget', 'informatif', 'aplikasi', 'bagus']
Normalization	['bagus', 'banget', 'informatif', 'aplikasi', 'bagus']
Stemming	bagus banget informatif sekali aplikasi sangat bagus

3.3 Pelabelan

Setelah Preprocessing, maka dilakukan pelabelan terhadap data tweet yang dihasilkan untuk memudahkan penilaian skor positif, negatif, dan netral. Pada Gambar 2 adalah data yang diperoleh serta sudah diberikan label.

	text	score	at
14224	mantap	1	2024-01-01 05:03:07
9922	sangat membantu	5	2024-01-01 05:39:38
5340	Sangat membantu tugas saya selaku guru	5	2024-01-01 06:23:51
2412	sangat membantu dalam proses kbm belajar jadi ...	5	2024-01-01 10:39:47
14235	Solusitif	5	2024-01-01 11:28:08
...	...	...	...
18941	apk bagus	1	2024-12-30 10:26:36
123	Aplikasi PMM terbaru malah tidak Compatible di...	5	2024-12-30 12:04:56
10716	semangat belajar	5	2024-12-30 13:36:51
5011	sangat membantu	5	2024-12-30 14:18:41
12178	sangat terbantu	5	2024-12-30 23:51:07

6281 rows × 3 columns

Gambar 2. Hasil Pelabelan Data

Pada Tabel 3 merupakan dataset ulasan pengguna aplikasi Merdeka Mengajar setelah proses preprocessing, dimana ada 503 data ulasan yang dihapus sehingga data yang digunakan adalah 6281 data ulasan. Berdasarkan data ulasan pada aplikasi merdeka mengajar, ada 5765 data ulasan dengan label positif, 154 Netral, dan 362 Negatif. Perbedaan jumlah dataset label positif dengan Negatif dan Netral yang tidak seimbang dapat menyebabkan Performa model menurun pada kelas minoritas (kelas yang jarang muncul). Sehingga untuk menangani permasalahan tersebut, maka pada penelitian ini akan menggunakan teknik *oversampling* dengan SMOTE. Selain itu, penelitian ini akan menggunakan *precision*, *recall*, dan *f-measure* untuk mengukur kinerja model pada kelas-kelas tersebut.

Tabel 3. Dataset Ulasan Pengguna Aplikasi Merdeka Mengajar

Actual Values		
Positive	Netral	Negatif
5765	154	362

3.4 Uji Model

Pengujian model bertujuan untuk mengevaluasi kinerja model klasifikasi berdasarkan metode tertentu. Proses ini melibatkan penggunaan dataset yang dibagi untuk pelatihan dan pengujian. Semakin besar jumlah data yang digunakan, maka kualitas model yang dihasilkan cenderung semakin baik. Data training digunakan untuk membangun model, sementara data testing digunakan untuk menilai performa model tersebut [20]. Adapun Pembagian data dilakukan dengan rasio 80% untuk data training dan 20% untuk data testing.

3.5 Evaluasi Model

Evaluasi model dilakukan setelah proses klasifikasi dan pengujian selesai dilaksanakan [21]. Tujuan dari evaluasi ini adalah untuk mengukur kinerja metode yang digunakan, dalam hal ini menggunakan algoritma Naïve Bayes Classifier. Hasil pengukuran performa, seperti nilai Accuracy, Precision, Recall, dan F-measure dari algoritma Naïve Bayes, Random Forest, dan SVM, disajikan pada tabel 4, 5, dan 6.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 5, No 4, June 2025 | Hal 813-820
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v5i4.530

Tabel 4. Nilai Accuracy, Precision, Recall, F-measure Naive Bayes

Klasifikasi	*Accuracy*	*Precision*	*Recall*	*F-measure*
Negative	*0.75*	*0.91*	*0.72*	*0.81*
Netral	*0.75*	*0.87*	*0.68*	*0.77*
Positif	*0.75*	*0.59*	*0.85*	*0.70*

Tabel 5 Nilai Accuracy, Precision, Recall, F-measure Random Forest

Klasifikasi	*Accuracy*	*Precision*	*Recall*	*F-measure*
Negative	*0.90*	*0.96*	*0.88*	*0.92*
Netral	*0.90*	*0.99*	*0.86*	*0.92*
Positif	*0.90*	*0.79*	*0.97*	*0.87*

Tabel 6. Nilai Accuracy, Precision, Recall, F-measure SVM

Klasifikasi	*Accuracy*	*Precision*	*Recall*	*F-measure*
Negative	0.89	0.98	0.87	0.92
Netral	0.89	0.98	0.84	0.90
Positif	0.89	0.77	0.98	0.86

Proses pembagian data menghasilkan 5025 data training dan 1256 data testing. Pada tahapan pengujian dengan data Testing 20%, dan data training 80%, secara naratif kinerja ketiga algoritma berdasarkan akurasi keseluruhan. Misalnya, "Dari ketiga algoritma yang diuji, Random Forest menunjukkan akurasi tertinggi sebesar 90 %, diikuti oleh SVM (89%), dan Naïve Bayes (75%). Hal ini mengindikasikan bahwa Random Forest memiliki kemampuan generalisasi yang lebih baik dalam mengklasifikasikan sentimen ulasan Aplikasi Merdeka Mengajar.

Adapun untuk precision tertinggi pada Random Forest untuk mengklasifikasikan kelas Netral untuk mengklasifikasi kelas Netral dengan nlai 0,99, dan terendah pada Naive Bayes untuk mengklasifikasikan kelas positif dengan nilai 0,59. Sehingga dalam data ulasan aplikasi merdeka mengajar, maka precision tidak mampu memprediksi kelas positif dengan baik. Sedangkan untuk recall, SVM dapat mengklasifikasikan kelas positif dengan tertinggi yaitu 0,98, dan terendah yaitu Naive Bayes dalam mengklasfikasikan label Netral dengan nilai 0,68. Sehingga dapat disimpulkan bahwa model SVM berhasil mengklasifikasi hampir data positif dengan baik. Adapun untuk *F-Measure*, Rendom Forest dapat mengklasifikasikan kelas Negative dan Netral dengan nilai 0,92, dan SVM juga dapat mengklasifikasikan kelas Negatif dengan nilai 0,92. Adapun terendah didapatkan naive Bayes dengan nilai 0,79 untuk mengklasifikasikan kelas positif.

Berdasarkan penelitian yang dilakukan, dapat diketahui bahwa algoritma Naive Bayes tidak mampu mengklasifikasikan ulasan aplikasi merdeka belajar dengan baik. Hal ini dapat dikarenakan nilai accuracy dan F-measure algoritma Naive Bayes lebih rendah dari SVM dan Random Forest.

3.6 Visualisasi

Setelah tahapan selesai semua dilakukan, selanjutnya adalah melakukan visualisasi berupa representasi wordcloud. Pada tampilan wordcloud merupakan bentuk gambaran yang mempresentasikan kata yang sering muncul pada data [22]. Pada gambar 3 merupakan WordCloud untuk sentimen bernilai positif.

Gambar 3. Tampilan WordCloud Sentimen Positif

4. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa mayoritas ulasan pengguna terhadap aplikasi Merdeka Mengajar bersifat positif, dengan 5765 dari 6281 ulasan tergolong dalam kategori positif. Hal ini mencerminkan bahwa secara umum, aplikasi telah diterima dengan baik oleh penggunanya. Meskipun demikian, masih terdapat sejumlah ulasan negatif (362) dan netral (154) yang perlu menjadi perhatian, karena ulasan tersebut dapat mencerminkan adanya aspek-aspek layanan yang belum sepenuhnya memuaskan.Dari sisi teknis, algoritma Random Forest menunjukkan kinerja klasifikasi sentimen yang paling tinggi dengan tingkat akurasi 90%, mengungguli SVM (88%) dan Naïve Bayes (75%). Implikasi dari temuan ini adalah bahwa Random Forest dapat direkomendasikan sebagai algoritma untuk memantau sentimen pengguna secara lebih akurat. Model seperti Naïve Bayes sering kali kurang akurat dalam menangani ketidakseimbangan ini karena asumsi distribusi probabilistik yang kaku. Random Forest, dengan mekanisme voting dari banyak decision tree, lebih adaptif terhadap dominasi kelas mayoritas, sehingga lebih mungkin memberikan prediksi yang lebih tepat pada kelas positif. Selain itu juga, Support Vector Machine (SVM) kurang optimal dalam menangani data tidak seimbang, seperti dataset ulasan aplikasi Merdeka Mengajar yang didominasi oleh ulasan positif. SVM cenderung bias terhadap kelas mayoritas, sehingga kurang akurat dalam mengklasifikasikan kelas minoritas seperti sentimen negatif. Dalam konteks data teks berdimensi tinggi, SVM juga membutuhkan sumber daya komputasi yang lebih besar dan waktu pelatihan yang lebih lama dibandingkan Random Forest. Sementara itu, Random Forest lebih fleksibel dan efisien dalam menangani fitur dalam jumlah besar serta memberikan interpretasi yang lebih baik melalui informasi pentingnya fitur. Oleh karena itu, Random Forest menunjukkan performa klasifikasi yang lebih unggul dan stabil dalam analisis sentimen ini.

REFERENCES

[1] L. Hewi and M. Shaleh, "Refleksi Hasil PISA (The Programme For International Student Assesment): Upaya Perbaikan Bertumpu Pada Pendidikan Anak Usia Dini)," *J. Golden Age*, vol. 4, no. 01, pp. 30–41, 2020, doi: 10.29408/jga.v4i01.2018.

[2] A. Sudianto, H. Bahtiar, M. F. Wajdi, and M. Mahpuz, "Penerapan Aplikasi Merdeka Belajar Kampus Merdeka (MBKM) Sebagai Upaya Peningkatan Kualitas Perguruan Tinggi," *Infotek J. Inform. dan Teknol.*, vol. 6, no. 2, pp. 421–430, 2023, doi: 10.29408/jit.v6i2.17183.

[3] D. Fatmawati, W. Trisnawati, Y. Jumaryadi, and G. Triyono, "Klasifikasi Tingkat Kepuasan Penggunaan Layanan Teknologi Informasi Menggunakan Decision Tree," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 3, no. 6, pp. 1056–1062, 2023, doi: 10.30865/klik.v3i6.803.

[4] N. Irzavika *et al.*, *Big Data*. Yogyakarta: Penamuda Media, 2024.

[5] A. K. Bitto, M. H. I. Bijoy, M. S. Arman, I. Mahmud, A. Das, and J. Majumder, "Sentiment analysis from Bangladeshi food delivery startup based on user reviews using machine learning and deep learning," *Bull. Electr. Eng. Informatics*, vol. 12, no. 4, pp. 2282–2291, 2023, doi: 10.11591/eei.v12i4.4135.

[6] R. S. Aryanti, W. Yudiana, and R. A. Sulistiobudi, "Aplikasi Kurikulum Merdeka Belajar-Kampus Merdeka (MBKM) pada Perguruan Tinggi Terhadap Karier Mahasiswa," *J. Paedagogy*, vol. 10, no. 1, p. 74, 2023, doi: 10.33394/jp.v10i1.6307.

[7] D. Moldovan, "A majority voting framework for reliable sentiment analysis of product reviews," *PeerJ Comput. Sci.*, vol. 11, pp. 1–31, 2025, doi: 10.7717/peerj-cs.2738.

[8] J. Asian, M. Dholah Rosita, and T. Mantoro, "Sentiment Analysis for the Brazilian Anesthesiologist Using Multi-Layer Perceptron Classifier and Random Forest Methods," *J. Online Inform.*, vol. 7, no. 1, p. 132, 2022, doi: 10.15575/join.v7i1.900.

[9] M. Isnan, G. N. Elwirehardja, and B. Pardamean, "Sentiment Analysis for TikTok Review Using VADER Sentiment and SVM Model," *Procedia Comput. Sci.*, vol. 227, pp. 168–175, 2023, doi: 10.1016/j.procs.2023.10.514.

[10] G. A. Buntoro, R. Arifin, G. N. Syaifuddiin, A. Selamat, O. Krejcar, and H. Fujita, "Implementation of a Machine Learning Algorithm for Sentiment Analysis of Indonesia's 2019 Presidential Election," *IIUM Eng. J.*, vol. 22, no. 1, pp. 78–92, 2021, doi: 10.31436/IIUMEJ.V22I1.1532.

[11] D. Hermawan, A. Aditya, and F. E. Purwiantono, "Sentiment Analysis of Vaccine Perceptions in Indonesia: A Comparative Study Using the Naïve Bayes Multinomial Text Algorithm on Twitter Data," in *2024 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, Mataram: IEEE, 20245. doi: 10.1109/ISITIA63062.2024.10667933.

[12] A. Amalia, D. Gunawan, and K. Nasution, "Sentiment analysis of GO-JEK services quality using Multi-Label Classification," *J. Phys. Conf. Ser.*, vol. 1830, no. 1, 2021, doi: 10.1088/1742-6596/1830/1/012003.

[13] P. Songram, I. Ritthisit, C. Jareanpon, and N. Muangnak, "A Comparative Study of Machine Learning and Deep Learning Models for Sentiment Analysis on Thai Restaurant Reviews," *ICIC Express Lett. Part B Appl.*, vol. 16, no. 4, pp. 379–386, 2025, doi: 10.24507/icicelb.16.04.379.

[14] D. Fimoza, A. Amalia, and T. H. F. Harumy, "Sentiment Analysis for Movie Review in Bahasa Indonesia Using BERT," in *2021 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA)*, Medan: IEEE, 2021. doi: 10.1109/DATABIA53375.2021.9650096.

[15] B. A. Abdelfattah, S. M. Darwish, and S. M. Elkaffas, "Enhancing the Prediction of Stock Market Movement Using Neutrosophic-Logic-Based Sentiment Analysis," *J. Theor. Appl. Electron. Commer. Res.*, vol. 19, no. 1, pp. 116–134, 2024, doi: 10.3390/jtaer19010007.

[16] D. Alfiyanti and Indra, "Penerapan Algoritma Multinomial Naïve Bayes dan Support Vector Machine (SVM) untuk Analisis Sentimen terhadap Ulasan Google Maps di Taman Mini Indonesia," *Simetris*, vol. 15, no. 1, pp. 85–102, 2024.

[17] Y. Yunitasari, A. Musdholifah, and A. K. Sari, "Sarcasm Detection For Sentiment Analysis in Indonesian Tweets," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 13, no. 1, p. 53, 2019, doi: 10.22146/ijccs.41136.

[18] S. F. Hilal Z and Rushendra, "The Efficiency of Machine Learning Techniques for DDoS Attack Detection: Random Forest, Logistic Regression, and Neural Networks," *Sink. J. dan Penelit. Tek. Inform.*, vol. 9, no. 1, pp. 522–530, 2025.

[19] J. Novakovic and S. Markovic, "Classifier Ensembles for Credit Card Fraud Detection," *2020 24th Int. Conf. Inf. Technol. IT 2020*, no. February, 2020, doi: 10.1109/IT48810.2020.9070534.

[20] Z. Wei and S. Zhang, "A structured sentiment analysis dataset based on public comments from various domains," *Data Br.*, vol. 53, p. 110232, 2024, doi: 10.1016/j.dib.2024.110232.
[21] H. Ahmadian, T. F. Abidin, H. Riza, and K. Muchtar, "Hybrid Models for Emotion Classification and Sentiment Analysis in Indonesian Language," *Appl. Comput. Intell. Soft Comput.*, vol. 2024, 2024, doi: 10.1155/2024/2826773.
[22] A. Quilez-Robres, M. Acero-Ferrero, D. Delgado-Bujedo, R. Lozano-Blasco, and M. Aiger-Valles, "Social Networks in Military Powers: Network and Sentiment Analysis during the COVID-19 Pandemic," *Computation*, vol. 11, no. 6, 2023, doi: 10.3390/computation11060117.