Mahardika Darmawan KW, M.Pd 74001019

REVISI FIX.docx

IRTA

Kampus 2

Perpustakaan

Document Details

Submission ID

trn:oid:::1:3651803875

Submission Date

Sep 17, 2026, 5:12 PM GMT+7

Download Date

Sep 17, 2026, 7:07 PM GMT+7

File Name

REVISI_FIX.docx

File Size

651.4 KB

7 Pages

3,317 Words

29,026 Characters

5% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography

Top Sources

5% Internet sources
2% Publications
1% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Top Sources

3% Internet sources
2% Publications
1% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1 Internet

archive.umsida.ac.id **3%**

2 Internet

jurnal.pustakagalerimandiri.co.id **<1%**

3 Internet

eprints.polbeng.ac.id **<1%**

4 Publication

Adinda Cahya Kamilla, Jadiaman Parhusip. "Pengukuran Usability Spotify Deskto... **<1%**

Design Of A Web-Based Patient Registration Management Information System At Dua Dental Care Clinic Using The Waterfall Method
[Perancangan Sistem Informasi Manajemen Pendaftaran Pasien pada Klinik Dua Dental Care Berbasis Web Menggunakan Metode Waterfall]

Muhammad Fikri Amin [1)], Novia Ariyanti [*2)]

[1)] Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia
[2)] Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia
*Email Penulis Korespondensi: yulianfindawati@umsida.ac.id

Abstrak. *Dental clinics that still register new patients manually often face long queues, recording errors, and disorganized patient data. This condition is also found at Klinik Dua Dental Care, where patients must come directly to the clinic, fill out forms manually, and wait without certainty regarding queue time. This research aims to design a web-based patient registration management information system for Klinik Dua Dental Care using the Waterfall method, which develops the system sequentially through problem identification, data collection, system design, testing, and evaluation/maintenance stages. The system is designed using Unified Modeling Language (UML) in the form of a Use Case Diagram and Data Flow Diagram (DFD), and is built using ViteJS for the front end, ExpressJS for the back end, and PostgreSQL as the database. The system provides three main features, namely online registration, real-time queue management based on the First In First Out (FIFO) principle for both online and walk-in patients, and structured patient data management. Functional testing using the Black-Box method on eleven test scenarios shows that the features that have been developed, including account registration, login, patient registration, and queue status monitoring, run according to expectations. It can be concluded that this system is expected to shorten the registration process, reduce queues, and improve the quality of service at Klinik Dua Dental Care. User satisfaction testing using a Likert scale involving 15 respondents (patients and administrative staff) across five assessment aspects yielded an overall score of 88.0%, which falls into the "Very Good" category, indicating that the system was well received by its users.*
Keywords: *Patient Registration Information System, Waterfall, UML, Black-Box Testing, Web*

Abstract. *Klinik gigi yang masih melakukan pendaftaran pasien baru secara manual kerap menghadapi antrean panjang, kesalahan pencatatan, serta data pasien yang tidak terorganisir dengan baik. Kondisi ini juga dialami oleh Klinik Dua Dental Care, di mana pasien harus datang langsung ke klinik, mengisi formulir secara manual, dan menunggu tanpa kepastian waktu antrean. Penelitian ini bertujuan untuk merancang sistem informasi manajemen pendaftaran pasien pada Klinik Dua Dental Care berbasis web menggunakan metode Waterfall, yang mengembangkan sistem secara berurutan melalui proses identifikasi masalah, pengumpulan data, perancangan sistem, pengujian, serta evaluasi dan pemeliharaan. Sistem dirancang menggunakan Unified Modeling Language (UML) berupa Use Case Diagram dan Data Flow Diagram (DFD), serta dibangun menggunakan ViteJS sebagai front end, ExpressJS sebagai back end, dan PostgreSQL sebagai basis data. Sistem ini menyediakan tiga fitur utama, yaitu pendaftaran online, manajemen antrean real-time berdasarkan prinsip First In First Out (FIFO) baik bagi pasien online maupun pasien yang datang langsung (walk-in), serta pengelolaan data pasien secara terstruktur. Pengujian fungsional menggunakan metode Black-Box pada sebelas skenario pengujian menunjukkan bahwa fitur-fitur yang telah dikembangkan, meliputi registrasi akun, login, pendaftaran pasien, hingga pemantauan status antrean, berjalan sesuai dengan yang diharapkan. Sistem ini dapat disimpulkan mampu mempersingkat proses pendaftaran, mengurangi antrean, serta meningkatkan kualitas pelayanan pada Klinik Dua Dental Care. Pengujian kepuasan pengguna dengan skala Likert yang melibatkan 15 responden (pasien dan petugas administrasi) pada lima aspek penilaian memperoleh persentase keseluruhan sebesar 88,0%, yang termasuk dalam kategori "Sangat Baik", sehingga menunjukkan bahwa sistem diterima dengan baik oleh penggunanya.*
Kata Kunci: *MBKM,Merdeka Belajar Kampus Merdeka,program MBKM,Kampus Merdeka,Merdeka Belajar*

I. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah mengubah cara masyarakat menyampaikan opini dan memperoleh informasi. Salah satu media yang banyak digunakan untuk berbagi pendapat adalah platform X (sebelumnya Twitter), yang memungkinkan pengguna menyampaikan tanggapan terhadap berbagai isu secara cepat dan terbuka. Data yang dihasilkan dari media sosial tersebut dapat dimanfaatkan untuk mengetahui persepsi masyarakat terhadap suatu kebijakan melalui teknik analisis sentimen, yaitu teknik dalam bidang pengolahan bahasa

alami yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini dalam teks ke dalam kategori tertentu seperti positif, netral, dan negatif [1], [2].

Program Merdeka Belajar Kampus Merdeka (MBKM) merupakan kebijakan yang diperkenalkan oleh Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia sebagai upaya meningkatkan kualitas pendidikan tinggi. Program ini memberikan kesempatan kepada mahasiswa untuk memperoleh pengalaman belajar di luar program studi melalui berbagai kegiatan seperti magang, pertukaran pelajar, proyek kemanusiaan, penelitian, kewirausahaan, dan asistensi mengajar. Implementasi MBKM memunculkan berbagai tanggapan dari mahasiswa, dosen, maupun masyarakat yang banyak disampaikan melalui media sosial.

Besarnya jumlah komentar yang dipublikasikan setiap hari menyebabkan proses analisis secara manual menjadi kurang efektif. Selain membutuhkan waktu yang lama, analisis manual juga rentan terhadap subjektivitas penilai. Oleh karena itu, diperlukan metode klasifikasi otomatis berbasis *text mining* dan *machine learning* yang mampu mengelompokkan opini ke dalam kategori positif, netral, dan negatif secara cepat dan konsisten [3].

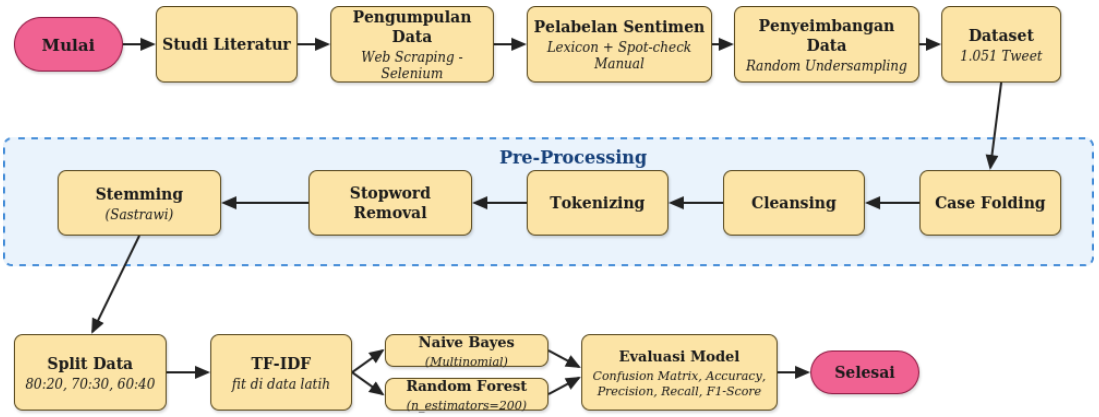
Beberapa penelitian sebelumnya menunjukkan bahwa algoritma *Naive Bayes* memiliki keunggulan dalam proses klasifikasi teks karena sederhana, memiliki waktu komputasi yang cepat, serta mampu memberikan performa yang baik pada data berdimensi tinggi [4]. Namun, algoritma ini memiliki asumsi bahwa setiap fitur bersifat independen sehingga pada beberapa kondisi performanya dapat menurun ketika hubungan antar kata cukup kompleks. Sebaliknya, algoritma *Random Forest* merupakan metode *ensemble learning* yang membangun sejumlah *decision tree* dan menggabungkan hasilnya untuk meningkatkan akurasi klasifikasi serta meminimalkan risiko *overfitting* [5].

Penelitian terdahulu menunjukkan hasil yang bervariasi terkait performa kedua algoritma tersebut pada kasus analisis sentimen. Zhafira dkk. menganalisis sentimen Kebijakan Kampus Merdeka berdasarkan komentar YouTube menggunakan *Naive Bayes* dan pembobotan TF-IDF, dan memperoleh akurasi terbaik sebesar 97% dengan rata-rata akurasi 91,8% [6]. Fikri dkk. memperoleh akurasi 73,65% menggunakan *Naive Bayes* pada analisis sentimen platform X [7], sementara Puspitasari dkk. melaporkan akurasi 92,00% pada analisis sentimen tweet pengguna e-commerce dengan metode yang sama [8]. Di sisi lain, beberapa penelitian menunjukkan *Random Forest* lebih unggul dibandingkan *Naive Bayes*: Admojo dkk. memperoleh akurasi *Random Forest* sebesar 88% berbanding 74% pada analisis sentimen aplikasi Webtoon [9], Nugroho dkk. melaporkan *Random Forest* unggul dengan akurasi hingga 93% dibandingkan *Naive Bayes* sebesar 85,86%–89% pada klasifikasi sentimen aplikasi Identitas Kependudukan Digital [10], dan Fitri dkk. melaporkan *Random Forest* sebagai algoritma terbaik dengan akurasi 97,16% dibandingkan *Naive Bayes* sebesar 94,16% dan Support Vector Machine sebesar 96,01% pada analisis sentimen aplikasi Ruangguru [11]. Sebaliknya, Ernamia dan Herliana [12] serta Prasetyo dkk. [13] menemukan hasil berlawanan, yaitu *Naive Bayes* justru lebih unggul pada topik penelitian masing-masing. Perbedaan hasil antar penelitian ini menunjukkan bahwa performa kedua algoritma sangat bergantung pada karakteristik data dan konteks penelitian, sehingga perbandingan *Naive Bayes* dan *Random Forest* saja belum cukup menjelaskan mengapa performa keduanya bervariasi antar studi. Sebagian besar penelitian terdahulu, seperti Zhafira dkk. [6] dan Puspitasari dkk. [8], menggunakan dataset dengan distribusi kelas yang cenderung tidak seimbang dan mengandalkan pelabelan manual penuh, sehingga akurasi yang dilaporkan berpotensi terinflasi oleh dominasi kelas mayoritas dan tidak mencerminkan kesulitan sesungguhnya dalam mengklasifikasikan kelas minoritas. Penelitian ini berbeda dari studi-studi tersebut dalam dua hal utama: pertama, dataset dikumpulkan secara bertahap melalui multi-sesi web scraping dan diseimbangkan lewat random undersampling sehingga distribusi kelas Negatif, Netral, dan Positif setara, memberikan gambaran performa model yang lebih adil pada ketiga kelas sekaligus; kedua, sebagian label data diperoleh melalui pendekatan lexicon dan bukan pelabelan manual sepenuhnya, sebuah karakteristik pengumpulan data yang jarang diungkap secara eksplisit pada penelitian sejenis padahal berpotensi memengaruhi kualitas ground truth dan performa klasifikasi. Dengan demikian, kontribusi penelitian ini tidak hanya terletak pada perbandingan akurasi *Naive Bayes* dan *Random Forest*, tetapi juga pada evaluasi bagaimana karakteristik dataset MBKM, yaitu keseimbangan kelas dan pendekatan pelabelan berbasis lexicon, memengaruhi performa kedua algoritma pada kasus nyata yang belum banyak dikaji pada penelitian sebelumnya.

Pada penelitian ini dilakukan pengumpulan data menggunakan teknik *web scraping* berbasis Selenium terhadap komentar yang mengandung kata kunci terkait Program MBKM. Pengumpulan data dilakukan secara bertahap dalam beberapa sesi hingga pertengahan tahun 2026, kemudian diseimbangkan melalui proses *random undersampling* sehingga diperoleh *dataset* akhir sebanyak 1.051 tweet dengan komposisi seimbang (351 Netral, 350 Positif, 350 Negatif). Dataset tersebut kemudian melalui tahapan *preprocessing*, yaitu *case folding*, *cleaning*, *tokenization*, *stopword removal*, dan *stemming*. Selanjutnya dilakukan proses ekstraksi fitur menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) sebelum dilakukan klasifikasi menggunakan algoritma *Naive Bayes* dan *Random Forest*. Evaluasi model dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

II. Metode Penelitian

Penelitian ini dilaksanakan melalui delapan tahapan utama secara berurutan, mulai dari pengumpulan data hingga evaluasi model, sebagaimana digambarkan pada diagram alir penelitian pada Gambar 1.



Gambar 1. Diagram Alir Penelitian

Pengumpulan Data

Data penelitian dikumpulkan dari platform X yang membahas Program MBKM menggunakan teknik *web scraping* berbasis Python, karena akses API resmi X memerlukan biaya berlangganan. Proses scraping memanfaatkan beberapa *library*, yaitu undetected-chromedriver untuk menjalankan *browser* Chrome dengan patch anti-deteksi otomatisasi, Selenium (WebDriver, By, WebDriverWait, expected_conditions) untuk otomasi interaksi dengan halaman web, pandas untuk menyusun data hasil scraping, python-dateutil (relativedelta) untuk memecah rentang tanggal pencarian menjadi periode bulanan, serta modul re, time, dan random untuk pemrosesan teks dan pengaturan jeda antar permintaan.

Pencarian dan pengumpulan data dilakukan menggunakan lima kata kunci yang berkaitan dengan Program MBKM, sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. Kata Kunci Pencarian Data

No	Kata Kunci
1	MBKM
2	Kampus Merdeka
3	Merdeka Belajar
4	Program MBKM
5	Merdeka Belajar Kampus Merdeka

Pengumpulan data dilakukan secara bertahap dalam beberapa sesi scraping. Sesi awal menghasilkan data dengan komposisi sentimen yang timpang (didominasi kelas Netral), sehingga sesi-sesi berikutnya diarahkan secara terkuota untuk menambah data pada kelas Positif dan Negatif yang masih kurang. Setiap tweet baru diklasifikasikan sentimennya secara otomatis menggunakan pendekatan *lexicon* (pencocokan kata kunci positif dan negatif berbahasa Indonesia) sebagai label awal, dan diperiksa duplikasinya terhadap data yang sudah terkumpul sebelumnya berdasarkan kemiripan teks yang dinormalisasi, sehingga tweet yang sama tidak diambil berulang kali. Pendekatan *lexicon* digunakan pada sekitar separuh dari total data yang terkumpul, sedangkan sisanya diperoleh dari proses pengumpulan bertahap lainnya. Untuk menilai validitas label *lexicon* tersebut, dilakukan pengecekan manual (*spot-check*) terhadap sebagian data berlabel *lexicon* dengan membandingkan label otomatis terhadap penilaian sentimen secara manual oleh peneliti. Hasil pengecekan menunjukkan tingkat kesesuaian sekitar 70-85%, dengan sebagian selisih ditemukan terutama pada teks yang bersifat ambigu antara sentimen Positif dan Netral, misalnya kalimat yang menyampaikan informasi atau apresiasi secara tidak eksplisit. Tingkat kesesuaian ini menunjukkan bahwa pendekatan *lexicon* cukup memadai untuk mempercepat proses pelabelan awal pada skala data yang besar, namun tetap berpotensi menimbulkan *noise* pada sekitar 15-30% label, sehingga kualitas *ground truth* pada sebagian data tidak sepenuhnya setara dengan pelabelan manual penuh dan perlu dipertimbangkan sebagai faktor yang membatasi performa klasifikasi.

Penyeimbangan Data (Random Undersampling)

Seluruh sesi pengumpulan data menghasilkan lebih dari 2.400 tweet unik dengan label sentimen, namun dengan distribusi kelas yang masih timpang karena kelas Netral jauh lebih dominan dibandingkan kelas Positif dan Negatif. Ketimpangan kelas (*class imbalance*) semacam ini perlu diatasi karena model klasifikasi yang dilatih pada data tidak seimbang cenderung bias terhadap kelas mayoritas, sehingga menghasilkan nilai *accuracy* yang tampak tinggi secara keseluruhan namun sebenarnya kurang mampu mengenali pola pada kelas minoritas.

Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan teknik *random undersampling*, yaitu salah satu pendekatan *resampling* yang menyeimbangkan distribusi kelas dengan cara mengurangi secara acak jumlah

data pada kelas mayoritas hingga jumlahnya setara dengan kelas minoritas, tanpa menambahkan data baru maupun data sintetis. Secara teknis, proses ini dilakukan melalui empat langkah, yaitu: (1) menghitung jumlah data pada masing-masing kelas sentimen (Negatif, Netral, dan Positif) dari *dataset* awal yang berjumlah lebih dari 2.400 tweet; (2) menetapkan jumlah data pada kelas dengan anggota paling sedikit, yaitu kelas Positif dan Negatif, sebagai jumlah target (*n_target*); (3) melakukan pengambilan sampel secara acak tanpa pengembalian (*random sampling without replacement*) pada kelas yang jumlah datanya melebihi *n_target*, terutama kelas Netral yang pada tahap pengumpulan awal jumlahnya jauh lebih dominan, menggunakan fungsi *resample* pada *library* scikit-learn dengan parameter *replace=False* dan *random_state* tetap agar hasil pengambilan sampel dapat direproduksi; dan (4) menggabungkan kembali data hasil pengambilan sampel dari ketiga kelas menjadi satu *dataset* akhir yang seimbang. Melalui proses tersebut, diperoleh *dataset* akhir sebanyak 1.051 tweet dengan komposisi seimbang, yaitu 351 tweet Netral, 350 tweet Positif, dan 350 tweet Negatif, sebagaimana ditunjukkan pada Gambar 2.

Random undersampling dipilih pada penelitian ini karena kesederhanaan implementasinya dan karena tidak menimbulkan risiko penambahan data sintetis yang berpotensi memunculkan pola kata yang tidak alami pada data teks, sebagaimana dapat terjadi pada teknik *oversampling* berbasis interpolasi seperti *Synthetic Minority Oversampling Technique* (SMOTE). Meskipun demikian, teknik ini memiliki konsekuensi berupa berkurangnya volume data yang tersedia, dari lebih dari 2.400 tweet pada tahap pengumpulan menjadi 1.051 tweet setelah penyeimbangan, karena sebagian data pada kelas mayoritas tidak diikutsertakan pada *dataset* akhir. Pengaruh berkurangnya volume data latih akibat proses *undersampling* ini terhadap performa kedua model dibahas lebih lanjut pada bagian pembahasan.

Preprocessing Data

Dataset tersebut selanjutnya melalui tahap *preprocessing* yang meliputi *case folding* untuk menyeragamkan huruf menjadi huruf kecil, *cleaning* untuk menghapus URL, *mention*, dan karakter khusus, *tokenization* untuk memecah teks menjadi unit kata, *stopword removal* untuk menghilangkan kata yang tidak bermakna, dan *stemming* untuk mengubah kata berimbuhan menjadi kata dasar menggunakan *library* Sastrawi.

Ekstraksi Fitur TF-IDF

Teks yang telah melalui *preprocessing* selanjutnya diubah menjadi representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF), yang memberikan bobot pada tiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen dan kelangkaannya pada keseluruhan korpus data [14].

Pembagian Data Latih dan Data Uji

Dataset yang telah diekstraksi fiturnya kemudian dibagi menjadi data latih dan data uji dengan proporsi 80% data latih (840 tweet) dan 20% data uji (210 tweet) menggunakan teknik *stratified split*, sehingga proporsi ketiga kelas sentimen tetap seimbang pada kedua subset data. Sebagai perbandingan, penelitian ini juga menguji dua skenario rasio pembagian data lain, yaitu 70% data latih : 30% data uji dan 60% data latih : 40% data uji, guna mengetahui pengaruh proporsi data latih terhadap performa kedua model (lihat Bagian 3.3).

Klasifikasi Data

Klasifikasi dilakukan menggunakan dua algoritma, yaitu Multinomial *Naive Bayes* dan *Random Forest*. Model *Random Forest* dilatih menggunakan 200 buah *decision tree* (n_estimators=200) tanpa batas kedalaman pohon maksimum, sedangkan model *Naive Bayes* dilatih menggunakan parameter bawaan (*default*) dari *library* scikit-learn [15].

Evaluasi Model

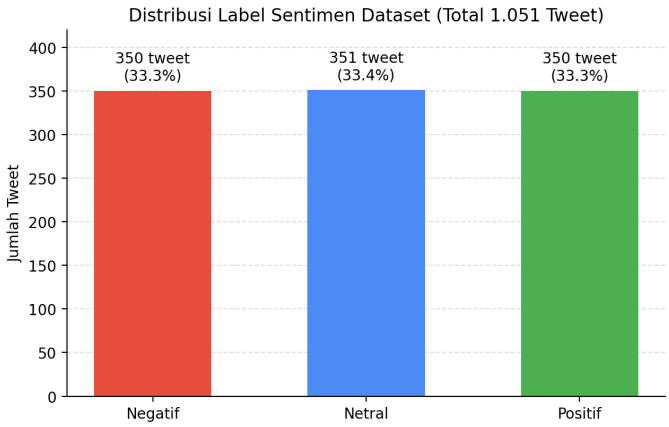
Evaluasi performa kedua model dilakukan menggunakan data uji dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, serta dianalisis lebih lanjut menggunakan confusion matrix untuk mengetahui kelas sentimen mana yang paling sering salah diklasifikasikan oleh masing-masing model.

III. HASIL DAN PEMBAHASAN

A. Hasil

Distribusi Dataset

Dataset akhir sebanyak 1.051 tweet dibagi menjadi 840 data latih dan 210 data uji. Kedua algoritma, yaitu *Naive Bayes* dan *Random Forest*, dilatih menggunakan data latih yang sama dan diuji pada data uji yang sama pula agar perbandingan performa kedua model dapat dilakukan secara adil. Distribusi label sentimen pada *dataset* akhir relatif seimbang antar ketiga kelas, sebagaimana ditunjukkan pada Gambar 2, sehingga model tidak bias terhadap kelas mayoritas tertentu. Hasil evaluasi kedua model disajikan pada Tabel 2.



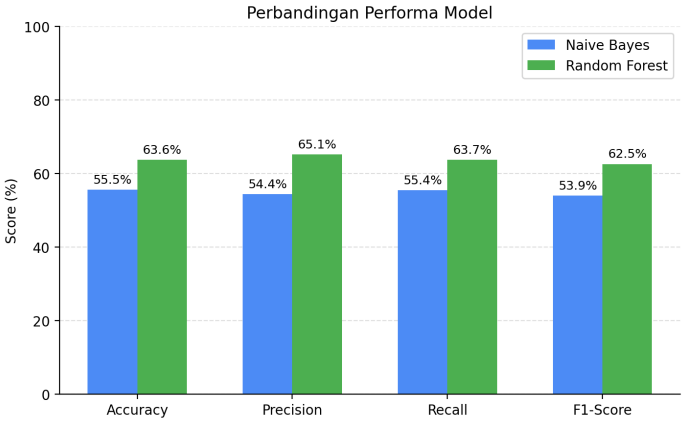
Gambar 2. Distribusi Label Sentimen Dataset (Total 1.051 Tweet)

Hasil Evaluasi Model

Tabel 2. Perbandingan Hasil Evaluasi *Naive Bayes* dan *Random Forest*

Model	**Accuracy**	**Precision**	**Recall**	***F1-score***
Naive Bayes	55,50%	54,35%	55,43%	53,93%
Random Forest	63,64%	65,14%	63,66%	62,55%

Tabel 2 menunjukkan bahwa *Random Forest* memperoleh performa lebih baik dibandingkan *Naïve Bayes* pada seluruh metrik evaluasi. *Random Forest* unggul sebesar 8,14 poin persentase pada *accuracy*, 10,79 poin pada *precision*, 8,23 poin pada *recall*, dan 8,62 poin pada *F1-score*. Keunggulan ini menunjukkan bahwa pendekatan *ensemble learning* pada *Random Forest*, yang menggabungkan hasil dari 200 *decision tree*, lebih mampu menangkap pola hubungan antar kata yang kompleks pada data tweet dibandingkan *Naïve Bayes* yang mengasumsikan independensi antar fitur. Perbandingan persentase keempat metrik evaluasi tersebut divisualisasikan pada Gambar 3.



Gambar 3. Perbandingan Persentase Accuracy, Precision, Recall, dan *F1-Score Naïve Bayes* dan *Random Forest*

Perbandingan Rasio Data Latih dan Data Uji

Tabel 3. Perbandingan Hasil Evaluasi pada Tiga Skenario Rasio Data Latih dan Data Uji

Rasio Split	Model	Data Latih	Data Uji	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
80:20	Naive Bayes	840	211	58,77%	58,76%	58,77%	57,22%
80:20	Random Forest	840	211	62,09%	63,59%	62,09%	61,67%
70:30	Naive Bayes	735	316	55,38%	54,66%	55,38%	53,56%
70:30	Random Forest	735	316	60,44%	61,79%	60,44%	60,35%
60:40	Naive Bayes	630	421	56,77%	56,92%	56,77%	55,54%
60:40	Random Forest	630	421	61,76%	64,38%	61,76%	61,39%

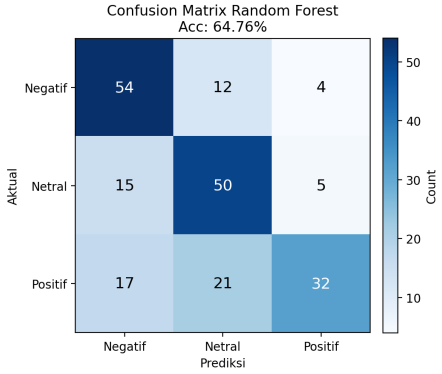
Tabel 3 menunjukkan hasil evaluasi Naïve Bayes dan Random Forest pada tiga 9cenario rasio pembagian data, yaitu 80:20, 70:30, dan 60:40. Random Forest secara konsisten memperoleh accuracy lebih tinggi dibandingkan Naïve Bayes pada ketiga 9cenario. Pada rasio 80:20, Random Forest memperoleh accuracy sebesar 62,09%, sedangkan Naïve Bayes sebesar 58,77%. Pada rasio 70:30, Random Forest memperoleh accuracy sebesar 60,44%, sedangkan Naïve Bayes sebesar 55,38%. Sementara itu, pada rasio 60:40, Random Forest memperoleh accuracy sebesar 61,76%, sedangkan Naïve Bayes sebesar 56,77%.

Confusion Matrix Model *Random Forest*

Tabel 4. Confusion Matrix Model *Random Forest*

Aktual \ Prediksi	Negatif	Netral	Positif
Negatif	54	12	4
Netral	15	50	5
Positif	17	21	32

Confusion matrix pada Tabel 4 menunjukkan bahwa *Random Forest* paling baik mengenali kelas Negatif (54 dari 70 data uji diklasifikasikan dengan benar) dan kelas Netral (50 dari 70 data), namun masih cukup banyak kesalahan pada kelas Positif, di mana hanya 32 dari 70 data yang diklasifikasikan dengan benar dan sisanya lebih sering keliru diklasifikasikan sebagai Negatif atau Netral. Kesalahan ini kemungkinan disebabkan oleh kemiripan kosakata antara tweet Positif dan Netral yang bersifat informatif, seperti penyampaian apresiasi yang disampaikan secara tidak eksplisit. Visualisasi confusion matrix tersebut disajikan pada Gambar 4.



Gambar 4. Confusion Matrix Model *Random Forest* pada Data Uji

B. Pembahasan

Pembahasan dan Perbandingan dengan Penelitian Terdahulu

Hasil penelitian ini sejalan dengan temuan Admojo dkk. [9], Nugroho dkk. [10], dan Fitri dkk. [11] yang juga melaporkan keunggulan *Random Forest* dibandingkan *Naive Bayes* pada analisis sentimen aplikasi Webtoon, aplikasi Identitas Kependudukan Digital, dan aplikasi Ruangguru. Namun demikian, hasil ini berbeda dengan temuan Zhafira dkk. [6], Fikri dkk. [7], dan Puspitasari dkk. [8] yang justru melaporkan performa baik dari *Naive Bayes* pada topik penelitian masing-masing. Perbedaan ini menegaskan bahwa performa kedua algoritma sangat bergantung pada karakteristik *dataset*, termasuk proses pelabelan, keseimbangan kelas, serta kompleksitas kosakata yang digunakan pada data penelitian masing-masing.

Akurasi *Random Forest* sebesar 63,64% pada penelitian ini tergolong lebih rendah dibandingkan sebagian besar penelitian terdahulu yang disebutkan sebelumnya, dan hal ini dapat dijelaskan oleh setidaknya tiga faktor. Pertama, *dataset* pada penelitian ini diseimbangkan melalui *random undersampling* sehingga ketiga kelas sentimen memiliki proporsi yang setara; kondisi ini membuat model tidak dapat "terbantu" oleh dominasi kelas mayoritas sebagaimana yang terjadi pada *dataset* tidak seimbang, misalnya pada penelitian Zhafira dkk. [6] dan Puspitasari dkk. [8] yang datasetnya didominasi kelas tertentu sehingga akurasi yang dilaporkan cenderung lebih tinggi namun kurang mencerminkan kemampuan model mengenali seluruh kelas secara merata. Kedua, proses *undersampling* turut mengurangi jumlah data latih yang tersedia (dari lebih dari 2.400 tweet awal menjadi 840 data latih), sehingga model *Random Forest* memiliki lebih sedikit variasi pola kata untuk dipelajari dibandingkan penelitian dengan volume data latih yang lebih besar. Ketiga, sebagaimana ditunjukkan pada Tabel 4, kesalahan klasifikasi paling banyak terjadi pada kelas Positif yang sering tertukar dengan kelas Netral, kemungkinan besar disebabkan oleh *noise* label dari pendekatan *lexicon* yang hanya memiliki tingkat kesesuaian sekitar 70-85% terhadap penilaian manual (lihat Bab II). Ketidaksesuaian label pada rentang 15-30% data tersebut berarti sebagian *ground truth* yang digunakan untuk melatih dan menguji model tidak sepenuhnya akurat, sehingga membatasi *accuracy* maksimum yang dapat dicapai oleh kedua algoritma, tidak hanya oleh *Random Forest* tetapi juga oleh *Naive Bayes* yang *accuracy*-nya juga tergolong rendah (55,50%). Ketiga faktor ini, yaitu keseimbangan kelas yang lebih ketat, volume data latih yang lebih kecil setelah *undersampling*, dan *noise* label dari pendekatan *lexicon*, secara bersama-sama menjelaskan mengapa performa kedua model pada penelitian ini berbeda dari penelitian-penelitian terdahulu yang

menggunakan *dataset* dengan karakteristik berbeda.

Keterbatasan Penelitian

Penelitian ini memiliki keterbatasan pada proses pelabelan sebagian data tambahan yang menggunakan pendekatan *lexicon* (pencocokan kata kunci) sebagai label awal, bukan pelabelan manual sepenuhnya. Meskipun pengecekan manual menunjukkan tingkat kesesuaian yang cukup baik (70-85%), sisa ketidaksesuaian tersebut berpotensi menimbulkan *noise* pada label sentimen yang dapat memengaruhi performa kedua model, khususnya pada kelas Positif yang menunjukkan tingkat kesalahan klasifikasi paling tinggi.

IV. SIMPULAN

Penelitian ini berhasil menganalisis sentimen masyarakat terhadap Program Merdeka Belajar Kampus Merdeka (MBKM) berdasarkan 1.051 tweet berbahasa Indonesia dari platform X dengan komposisi kelas yang seimbang (351 Netral, 350 Positif, 350 Negatif). Berdasarkan hasil pengujian, algoritma *Random Forest* memberikan performa klasifikasi yang lebih baik dibandingkan *Naive Bayes*, dengan nilai *accuracy* sebesar 63,64%, *precision* 65,14%, *recall* 63,66%, dan *F1-score* 62,55%, sedangkan *Naive Bayes* memperoleh *accuracy* 55,50%, *precision* 54,35%, *recall* 55,43%, dan *F1-score* 53,93%. Keunggulan *Random Forest* menunjukkan bahwa pendekatan *ensemble learning* lebih mampu menangkap pola hubungan antar kata yang kompleks pada data tweet dibandingkan *Naive Bayes* yang mengasumsikan independensi antar fitur. Penelitian selanjutnya disarankan untuk memperbanyak proporsi data berlabel manual guna mengurangi *noise* pada label sentimen, serta menguji algoritma klasifikasi lain seperti *Support Vector Machine* atau metode berbasis *deep learning* untuk memperoleh performa yang lebih optimal.

Hasil analisis sentimen menunjukkan opini masyarakat terhadap Program MBKM di platform X terbagi cukup merata antara positif, netral, dan negatif, yang mengindikasikan masih adanya pro-kontra di kalangan masyarakat terhadap kebijakan ini. Temuan ini dapat menjadi masukan bagi pengelola Program MBKM untuk lebih memperhatikan aspek-aspek yang sering dikeluhkan dalam komentar bersentimen negatif, serta memperkuat sosialisasi program guna mengurangi opini netral yang mengindikasikan kurangnya pemahaman masyarakat terhadap manfaat program