



Penggunaan Metode Naïve Bayes untuk Klasifikasi Topik Tweet Bidang dan Non-Bidang Rektorat Telkom University pada Akun Telyufess

Mukhammad Hafiz Bima Ibrahim*, Anisa Dzulkarnain, Alqis Rausanfit

Program Studi Sistem Informasi, Universitas Telkom, Surabaya, Indonesia

Email: ^{1,*}hafizbimaaa@student.telkomuniversity.ac.id, ²anisadzulkarnain@telkomuniversity.ac.id, ³alqisfita@telkomuniversity.ac.id

Email Penulis Korespondensi: hafizbimaaa@student.telkomuniversity.ac.id

Abstrak—Keluhan mahasiswa Telkom University yang disampaikan melalui akun Telyufess di media sosial X belum dimanfaatkan secara optimal sebagai bahan evaluasi layanan kampus. Penelitian ini bertujuan untuk mengklasifikasikan tweet pada akun Telyufess ke dalam dua kategori, yaitu bidang (berkaitan dengan unit-unit resmi rektorat seperti akademik, keuangan, dan layanan kampus) dan non-bidang (keluhan umum yang tidak terhubung langsung dengan unit tertentu). Permasalahan utama yang diangkat adalah pentingnya pemetaan keluhan secara otomatis untuk mendukung evaluasi layanan kampus. Metode klasifikasi yang digunakan adalah Naïve Bayes, dengan tahapan pelabelan manual oleh peneliti dan asisten annotator (disertai validasi antar-rater), praproses teks (normalisasi menggunakan kamus baku dan library Sastrawi, penghapusan karakter khusus, penggunaan stop word berbasis Bahasa Indonesia yang dilengkapi kata unik “telyu!”), tokenisasi, stemming, pembobotan TF-IDF, serta pembagian data ke dalam rasio 65:35, 70:30, 80:20, dan 90:10. Sebanyak 1.090 tweet dikumpulkan dalam rentang 1 Januari 2023 hingga 1 Januari 2025 menggunakan Tweet Harvest API, dengan kriteria tweet berupa keluhan, opini, dan saran (retweet tidak disertakan). Hasil eksperimen menunjukkan bahwa model memperoleh akurasi tertinggi sebesar 87,27% pada rasio 90:10, diikuti oleh 81,74% (80:20), 78,35% (70:30), dan 76,96% (65:35). Meskipun terdapat indikasi overfitting pada data pelatihan (akurasi >99%), metode Naïve Bayes terbukti efektif dalam klasifikasi otomatis tweet, serta memberikan kontribusi terhadap pemanfaatan media sosial sebagai sumber data evaluasi layanan kampus.

Kata Kunci: Klasifikasi; Naïve Bayes; Bidang Rektorat; Telyufess

Abstract—Student complaints submitted through the Telyufess account on the social media platform X have not been optimally utilized as input for evaluating campus services at Telkom University. This study aims to classify tweets from the Telyufess account into two categories: domain-related (linked to official university units such as academics, finance, and campus services) and non-domain (general complaints unrelated to specific units). The main issue addressed is the need for an automated mapping system of student complaints to support campus service evaluations. The classification method used is Naïve Bayes, involving manual labeling by the researcher and assistant annotators (with inter-rater validation), text preprocessing (normalization using a standard dictionary and the Sastrawi library, removal of special characters, stop word filtering based on Indonesian language lists augmented with the unique term “telyu!”), tokenization, stemming, TF-IDF weighting, and dataset splitting in ratios of 65:35, 70:30, 80:20, and 90:10. A total of 1,090 tweets were collected between January 1, 2023 and January 1, 2025 using the Tweet Harvest API, based on criteria including complaints, opinions, and suggestions (retweets were excluded). The highest accuracy was achieved at 87.27% with a 90:10 split, followed by 81.74% (80:20), 78.35% (70:30), and 76.96% (65:35). Although the model showed signs of overfitting on training data (accuracy >99%), the results demonstrate that Naïve Bayes is effective for automated tweet classification and contributes to the use of social media as a data source for evaluating campus services.

Keywords: Classification; Naïve Bayes; Rectorate Department; Telyufess

1. PENDAHULUAN

Teknologi Informasi (TI) saat ini menjadi aspek fundamental dalam mendukung operasional berbagai organisasi bisnis maupun institusi berskala besar. Banyak entitas melihat TI sebagai sarana yang mendukung implementasi strategi demi tercapainya visi, misi, dan tujuan yang telah dirancang [1]. Di lingkungan Telkom University, berbagai sistem telah diimplementasikan untuk menunjang kegiatan perkuliahan, baik melalui pengembangan sistem baru maupun sebagai pengganti sistem yang telah ada sebelumnya. Namun demikian, meskipun pemanfaatan sistem telah cukup luas, perlu dilakukan evaluasi untuk menilai sejauh mana sistem-sistem tersebut mampu memenuhi kebutuhan mahasiswa secara optimal [2]. Meskipun Telkom University telah mengembangkan berbagai sistem, masih terdapat beberapa kelemahan yang dirasakan oleh komunitas akademik. Salah satu keluhan yang paling sering muncul adalah terkait gangguan atau kesalahan pada server. Keluhan-keluhan ini umumnya disampaikan oleh civitas akademika melalui media sosial X, terutama melalui akun bot bernama Telyufess yang berperan sebagai kanal aspirasi mahasiswa dalam menyampaikan opini, masukan, maupun kritik terhadap berbagai aspek di lingkungan kampus.

Penelitian ini bertujuan untuk mengelompokkan *tweet* menjadi dua kategori utama, yaitu *tweet* yang berkaitan langsung dengan bidang-bidang di bawah naungan rektorat Telkom University dan *tweet* yang tidak berkaitan. Kategori pertama disebut sebagai *tweet* bidang, yakni *tweet* yang secara eksplisit mengarah pada unit-unit kerja atau layanan yang berada di bawah koordinasi rektorat. Sementara itu, *tweet* non bidang merupakan *tweet* yang tidak mengindikasikan keterkaitan langsung dengan bidang-bidang tersebut. Pemisahan ini dimaksudkan agar informasi yang terkandung dalam *tweet* bidang dapat dimanfaatkan sebagai bahan evaluasi dan pengembangan layanan di berbagai unit kerja Telkom University. Penggunaan akun Telyufess sebagai sumber data keluhan dinilai relevan, karena akun ini menyediakan saluran komunikasi yang bersifat real-time bagi mahasiswa untuk menyampaikan aspirasi mereka. Berbeda dengan metode survei konvensional yang dilakukan secara berkala dan terbatas waktunya, Telyufess memungkinkan mahasiswa menyampaikan keluhan kapan saja tanpa dibatasi oleh periode tertentu.



Penelitian yang dilakukan oleh Alvina dan rekan-rekannya dilatarbelakangi oleh cepatnya penyebaran virus corona, yang mendorong WHO menetapkan COVID-19 sebagai pandemi pada 11 Maret 2020. Perubahan status tersebut berlangsung secara tiba-tiba dan menyebabkan hampir seluruh negara terdampak oleh penyebaran virus ini. Oleh karena itu, penelitian tersebut dimanfaatkan untuk memprediksi tingkat penyebaran COVID-19. Studi ini memfokuskan analisis pada prediksi jumlah kasus COVID-19 dengan menggunakan algoritma klasifikasi Naïve Bayes [3].

Alita dan Shodiqin (2023) mengemukakan bahwa pelaksanaan vaksinasi COVID-19 di Indonesia menimbulkan berbagai tanggapan, baik yang mendukung maupun yang menolak. Pemerintah mendorong masyarakat untuk mengikuti program vaksinasi guna mencapai kekebalan kelompok (*herd immunity*). Penelitian yang mereka lakukan bertujuan untuk mengevaluasi kualitas layanan yang diberikan oleh tenaga kesehatan kepada masyarakat selama proses vaksinasi berlangsung. Studi ini juga menggunakan metode klasifikasi Naïve Bayes untuk memprediksi jumlah kasus COVID-19 [4].

Studi yang dilakukan Safitri dan kawan-kawan (2023) menerapkan algoritma Support Vector Machine (SVM) untuk mengklasifikasi sentimen pengguna Twitter berkaitan dengan grup idola Korea, BTS. Tujuannya adalah mengidentifikasi faktor-faktor apa saja yang dijadikan pertimbangan konsumen dalam menentukan figur publik yang layak menjabat sebagai brand ambassador suatu produk. Dengan demikian, penelitian ini berfokus pada eksplorasi sentimen netizen terhadap BTS sebagai bagian dari evaluasi kriteria pemilihan duta merek [5].

Penelitian lain yang sejenis dilakukan oleh Prasetyo dan rekan-rekannya pada tahun 2023, yang membahas analisis sentimen terkait isu relokasi Ibu Kota Negara (IKN) dengan menerapkan algoritma Naïve Bayes dan K-Nearest Neighbor (KNN). Topik utama dalam penelitian ini adalah sensitivitas perbincangan publik mengenai pemindahan IKN [6]. Perbedaan utama antara penelitian tersebut dengan penelitian ini terletak pada pendekatan metode yang digunakan. Penelitian ini hanya menggunakan algoritma Naïve Bayes, sementara studi Prasetyo dkk. menggabungkan dua metode, yaitu Naïve Bayes dan KNN.

Merawati dan rekan-rekannya dalam penelitiannya memanfaatkan algoritma Naïve Bayes dan Latent Dirichlet Allocation (LDA) untuk menganalisis sentimen wisatawan yang disampaikan melalui akun media sosial pribadi. Penelitian ini berfokus pada berbagai opini dan komentar terkait pariwisata Lombok yang diunggah oleh wisatawan, khususnya di platform Twitter. Banyaknya topik yang dibicarakan membuat proses identifikasi sentimen dan tema pembahasan secara manual menjadi tidak efisien, sehingga diperlukan pendekatan otomatis. Opini wisatawan mengenai destinasi Lombok menjadi perhatian utama dalam penelitian ini karena dianggap memiliki potensi sebagai sumber informasi yang bernilai untuk pengembangan pariwisata daerah tersebut [7]. Perbedaan antara penelitian ini dan penelitian yang akan dilakukan terletak pada metode yang digunakan. Penelitian tersebut menerapkan dua metode analisis, sedangkan penelitian yang akan dilakukan hanya menggunakan satu metode, yaitu algoritma Naïve Bayes.

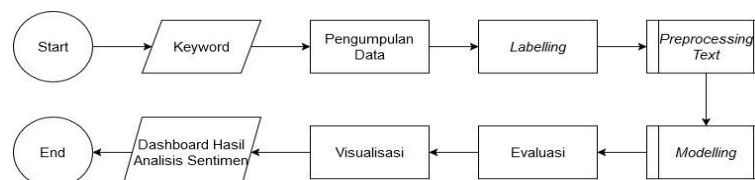
Teknologi informasi (TI) saat ini menjadi fondasi penting dalam mendukung layanan pendidikan di lingkungan perguruan tinggi, termasuk di Telkom University. Berbagai sistem telah diimplementasikan untuk menunjang proses perkuliahan serta layanan administrasi. Namun, mahasiswa masih kerap menyampaikan keluhan, khususnya terkait permasalahan layanan maupun gangguan sistem. Salah satu media yang digunakan untuk menyampaikan keluhan tersebut adalah akun bot Telyufess di platform media sosial X, yang memungkinkan mahasiswa menyampaikan aduan secara langsung dan bersifat real-time. Kendati demikian, data keluhan yang tersedia melalui akun tersebut belum dimanfaatkan secara optimal sebagai sumber umpan balik resmi untuk pengembangan kinerja di masing-masing unit rektorat.

Berbagai penelitian sebelumnya telah mengaplikasikan algoritma Naïve Bayes dalam pengolahan data dari media sosial, antara lain untuk memprediksi penyebaran COVID-19 maupun mengkaji sentimen masyarakat. Namun, sebagian besar studi tersebut berfokus pada isu-isu kesehatan dan opini publik, bukan pada aspek pengelolaan keluhan institusional. *Research gap* yang diidentifikasi adalah masih terbatasnya kajian yang memanfaatkan data keluhan mahasiswa dari media sosial sebagai dasar evaluasi kinerja unit layanan di lingkungan kampus. Penelitian ini memberikan kontribusi dengan mengimplementasikan algoritma Naïve Bayes pada data Telyufess untuk mengklasifikasikan keluhan mahasiswa, sehingga hasilnya dapat dijadikan bahan pertimbangan bagi pihak rektorat Telkom University dalam proses pengambilan keputusan untuk perbaikan dan peningkatan kualitas layanan.

Di samping itu, penelitian ini juga memiliki tujuan untuk melakukan klasifikasi *tweet* pada akun Telyufess ke dalam dua kategori, yaitu bidang dan non-bidang, dengan menerapkan metode Naïve Bayes. Hasil dari proses klasifikasi ini diharapkan dapat dimanfaatkan sebagai dasar dalam merancang strategi pengembangan serta peningkatan kinerja pada unit-unit rektorat Telkom University.

2. METODOLOGI PENELITIAN

Gambar 1 berikut merupakan tahapan penelitian yang dilakukan.



Gambar 1. Tahapan Penelitian



2.1 Pengumpulan Data

Seperti ditunjukkan pada Gambar 1, tahap ini melibatkan proses pengumpulan data dari akun Twitter @telyufess, yang memuat berbagai cuitan dari mahasiswa. Data dikumpulkan menggunakan teknik crawling yang mencakup seluruh tweet yang dipublikasikan oleh akun tersebut. Sebanyak 1.090 tweet berhasil dihimpun dalam rentang waktu 1 Januari 2023 hingga 1 Januari 2025 dengan memanfaatkan *Tweet Harvest* API. Hanya tweet yang memuat keluhan, opini, dan saran dari mahasiswa yang diambil, sementara retweet serta konten promosi dikecualikan untuk menjaga relevansi dataset terhadap tujuan analisis. Proses *crawling* dilakukan melalui *Application Programming Interface* (API) guna memastikan data yang diperoleh bersifat lengkap dan valid. Untuk mendukung proses pengambilan data ini, diperlukan alat bantu khusus yang dirancang untuk melakukan crawling pada situs target, salah satunya adalah *Tweet Harvest* [8]. *Tweet Harvest* merupakan sebuah *tools* yang digunakan untuk melakukan proses *crawling* data pada platform media sosial X dengan memanfaatkan *Application Programming Interface* (API). [9].

2.2 Labelling

Proses *labelling* dilakukan dengan membagi tweet ke dalam dua kategori utama, yaitu bidang dan non-bidang. Kegiatan ini dilakukan secara manual oleh peneliti bersama dengan asisten annotator untuk mempercepat proses klasifikasi data. Dengan tujuan memastikan konsistensi hasil pelabelan, tingkat kesepakatan antar-annotator diukur menggunakan nilai Cohen's Kappa. Nilai tersebut digunakan untuk menilai tingkat reliabilitas proses pelabelan sebelum data digunakan dalam tahap analisis berikutnya. Metode manual dipilih karena ukuran dataset masih dalam skala kecil hingga menengah, sehingga peninjauan langsung dianggap memberikan hasil yang lebih akurat [10].

2.3 Normalisasi

Proses awal pengolahan data pada file CSV hasil *crawling* diawali dengan eliminasi karakter non-alfabet, seperti (#, !, dan \$), guna menghindari gangguan dalam tahap analisis. Selanjutnya, kata-kata yang tidak sesuai dengan ejaan baku disesuaikan menggunakan kamus resmi bahasa Indonesia serta pemanfaatan *library* sastra berbasis *rule-based*. Tahap ini bertujuan memastikan seluruh kata tidak baku dikonversi secara otomatis sebelum memasuki proses analisis lanjutan [11].

2.4 Lowercasing

Setelah karakter-karakter khusus dihapus, tahap selanjutnya adalah proses *lowercasing* untuk menyeragamkan seluruh huruf menjadi bentuk kecil. Penyamaan ini dilakukan guna menghindari perbedaan makna akibat penggunaan huruf kapital. Namun, pada kasus tertentu, seperti pada singkatan-singkatan (misalnya BUMN, RI, dan sebagainya), proses *lowercasing* diterapkan secara selektif agar tidak mengubah arti dari singkatan tersebut. [12].

2.5 Penghapusan Stop Word

Tahap berikutnya adalah penghapusan *stop word*, yaitu kata-kata yang memiliki frekuensi kemunculan tinggi namun tidak memberikan kontribusi signifikan terhadap makna teks. Sebelum proses ini dilakukan, terlebih dahulu disusun daftar kata yang dikategorikan sebagai *stop word* untuk dijadikan acuan dalam proses eliminasi [13]. Tahap ini bertujuan untuk mengeliminasi kata-kata yang sering muncul namun tidak memiliki nilai informasi yang penting dalam konteks analisis, seperti kata hubung atau kata seruan. Daftar *stop word* yang digunakan mengacu pada daftar standar bahasa Indonesia yang umum dipakai dalam pemrosesan bahasa alami (NLP), kemudian disempurnakan dengan menambahkan kata-kata kontekstual tertentu seperti "telyu!" yang kerap muncul dalam dataset tetapi tidak relevan untuk proses klasifikasi. Penyesuaian tersebut dilakukan agar penghapusan *stop word* menjadi lebih optimal dan kontekstual.

2.6 Tokenisasi

Setelah proses penghapusan stop word, tahap selanjutnya adalah tokenisasi. Pada tahap ini, digunakan *library* sastra yang dirancang untuk pemrosesan bahasa alami (*Natural Language Processing*) dalam bahasa Indonesia, mengingat mayoritas tweet yang dianalisis menggunakan bahasa tersebut. Tokenisasi menghasilkan pemisahan kalimat menjadi unit-unit kata (token), sehingga setiap kata dapat dianalisis secara individual pada tahapan pemrosesan berikutnya [14].

2.7 Stemming

Setelah melewati sejumlah tahapan sebelumnya, proses selanjutnya adalah *stemming*, yaitu mengubah setiap kata dalam kalimat ke bentuk dasarnya. Penggunaan kata dasar ini penting dalam proses klasifikasi tweet agar sistem dapat mengenali dan memproses kata-kata secara lebih akurat. Untuk tahap *stemming* ini, digunakan *library* sastra yang memang dirancang untuk bahasa Indonesia [15].

2.8 Perhitungan TF-IDF

Setelah melalui berbagai tahap penyesuaian format teks, langkah berikutnya adalah proses vektorisasi, yakni mengonversi data berbasis teks menjadi representasi numerik melalui pembobotan menggunakan metode TF-IDF. Pada tahap ini, perhitungan *Term Frequency* (TF) dilakukan terlebih dahulu dengan menghitung seberapa sering suatu kata muncul dalam satu dokumen [16]. Tahap berikutnya adalah menghitung *Document Frequency* (DF), yaitu jumlah dokumen yang mengandung kata tertentu. Setelah nilai DF diperoleh, dilanjutkan dengan perhitungan *Inverse Document Frequency*



(IDF). Hasil dari perhitungan TF dan IDF kemudian dikombinasikan untuk menghasilkan nilai TF-IDF, yang membentuk *dataset* dalam bentuk pembobotan numerik [17].

2.9 Modelling

Setelah melalui *preprocessing text*, data kemudian memasuki tahap *modelling*. Proses ini terdiri dari dua bagian utama, yaitu *training* dan *testing* data. Sebelum dilakukan pemodelan, data dibagi (*splitting*) ke dalam empat skenario rasio, yakni 65:35, 70:30, 80:20, dan 90:10, dengan proporsi data yang lebih besar dialokasikan untuk training. Tujuan dari training data adalah untuk melatih model dalam mengenali pola-pola yang membedakan tweet kategori bidang dan non-bidang. Sementara itu, testing data digunakan untuk mengevaluasi sejauh mana model mampu melakukan klasifikasi secara akurat [18].

2.10 Evaluasi

Setelah data melewati berbagai tahap sebelumnya, langkah selanjutnya adalah evaluasi untuk menilai kemampuan model dalam mengklasifikasikan *tweet* ke dalam kategori bidang dan non-bidang secara akurat. Evaluasi dalam penelitian ini mencakup beberapa proses, antara lain penyusunan *confusion matrix*, perhitungan akurasi (*accuracy*), ketepatan (*precision*), daya jangkau (*recall*), serta nilai *F1-score* sebagai metrik gabungan dari *precision* dan *recall*. [19].

2.11 Visualisasi

Hasil evaluasi dari tahap sebelumnya akan disajikan dalam bentuk dashboard. Pembuatan dashboard menggunakan bahasa pemrograman *JavaScript* karena dianggap lebih fleksibel dalam pengelolaan tampilan antarmuka pengguna (*User Interface*). Dashboard ini dirancang untuk menampilkan kata-kata yang paling sering muncul dalam kategori bidang dan non-bidang, serta menampilkan *tweet-tweet* dari kedua kategori yang paling sering dikeluhkan [20].

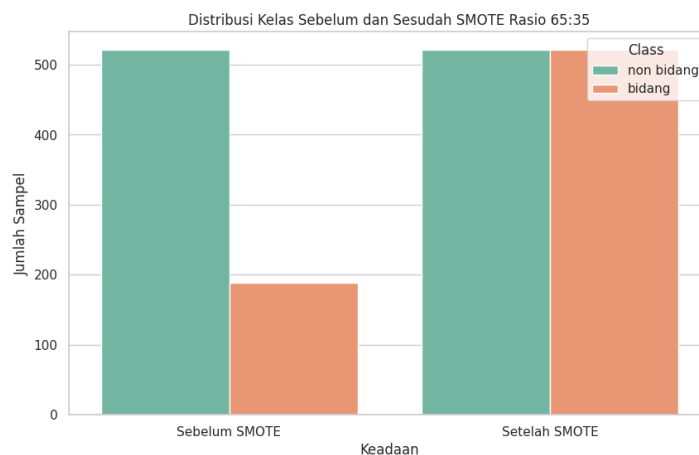
3. HASIL DAN PEMBAHASAN

3.1 Hasil Imbalance Data Handling

Metode SMOTE (*Synthetic Minority Oversampling Technique*) digunakan untuk menyeimbangkan distribusi data antara kelas “bidang” yang memiliki jumlah lebih sedikit dengan kelas “non-bidang”. Teknik ini bekerja dengan menghasilkan data sintetis pada kelas minoritas guna mengurangi potensi bias model terhadap kelas mayoritas. Namun, perlu dicermati bahwa penambahan data sintetis juga berisiko menimbulkan noise, sehingga hasil evaluasi model harus dianalisis secara hati-hati.

3.1.1 Rasio 65:35

Distribusi data sebelum dan sesudah proses penyeimbangan kelas dengan metode SMOTE pada rasio 65:35 ditampilkan pada Gambar 2. Visualisasi tersebut menunjukkan perbandingan jumlah data pada kelas “bidang” dan “non-bidang” sebelum serta setelah penambahan data sintetis dilakukan.



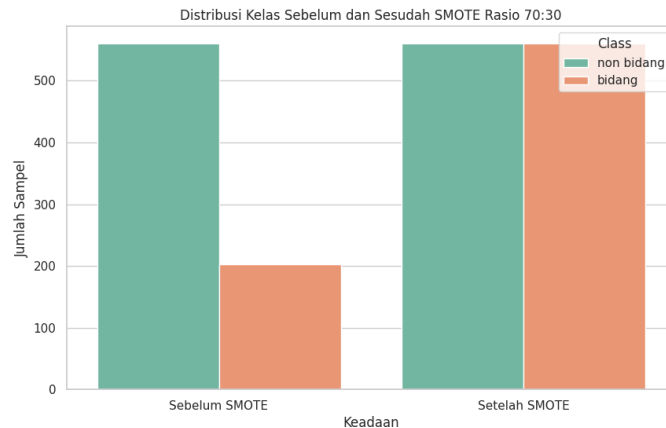
Gambar 2. Distribusi Kelas Sebelum dan Sesudah SMOTE Rasio 65:35

Pada Gambar 2, jumlah data pada kelas “non-bidang” dan “bidang” masing-masing tercatat sebanyak 521 dan 188, yang berasal dari total 709 data training. Ketimpangan ini menunjukkan adanya ketidakseimbangan distribusi antar kelas (*class imbalance*). Untuk mengatasi permasalahan tersebut, diterapkan teknik *imbalance data handling*, salah satunya adalah metode SMOTE. Teknik ini bertujuan menyeimbangkan jumlah data dengan menambahkan data sintetis pada kelas minoritas. Dalam konteks ini, SMOTE digunakan untuk menambah jumlah data pada kelas “bidang” hingga setara dengan kelas “non-bidang”, yakni masing-masing menjadi 521 data



3.1.2 Rasio 70:30

Distribusi data untuk kelas “bidang” dan “non-bidang” sebelum dan sesudah penyeimbangan dengan metode SMOTE pada rasio 70:30 ditampilkan pada Gambar 3. Gambar tersebut menggambarkan selisih jumlah data antara kedua kelas, serta menunjukkan hasil akhir setelah penambahan data sintetis guna mencapai distribusi yang seimbang.

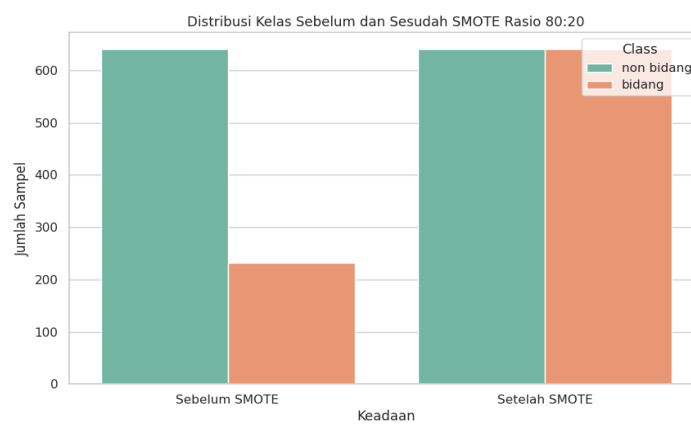


Gambar 3. Distribusi Kelas Sebelum dan Sesudah SMOTE Rasio 70:30

Pada Gambar 3, jumlah data untuk kelas “non-bidang” tercatat sebanyak 560, sedangkan kelas “bidang” sebanyak 203, yang seluruhnya berasal dari 763 data pelatihan. Perbedaan jumlah yang signifikan antara kedua kelas ini mencerminkan adanya ketidakseimbangan data (*imbalance*). Untuk menangani kondisi tersebut, digunakan pendekatan khusus yang dikenal sebagai *imbalance data handling*. Salah satu teknik yang umum diterapkan adalah SMOTE, yang berfungsi menambahkan data sintetis pada kelas minoritas. Dalam konteks ini, SMOTE digunakan untuk memperbanyak data pada kelas “bidang” dari 203 menjadi 560, agar distribusinya setara dengan kelas “non-bidang”.

3.1.3 Rasio 80:20

Distribusi jumlah data pada kelas “bidang” dan “non-bidang” sebelum dan sesudah dilakukan penyeimbangan dengan metode SMOTE pada rasio 80:20 ditampilkan pada Gambar 4. Visualisasi tersebut menunjukkan kondisi awal ketidakseimbangan antar kelas serta hasil akhir setelah penambahan data sintetis pada kelas minoritas untuk mencapai distribusi yang seimbang.

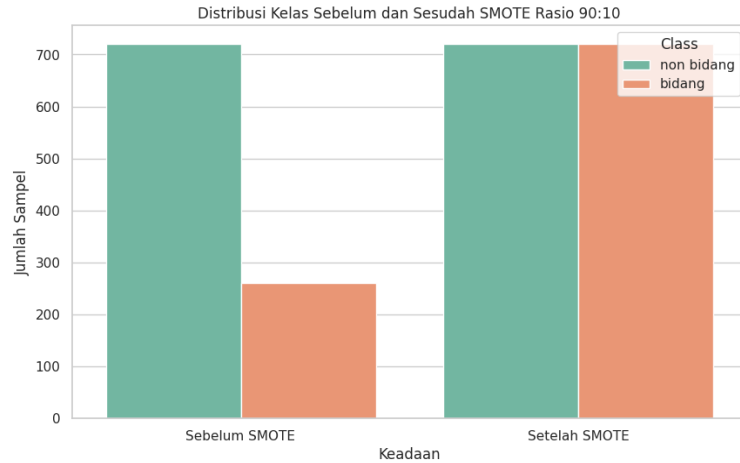


Gambar 4. Distribusi Kelas Sebelum dan Sesudah SMOTE Rasio 80:20

Gambar 4 memperlihatkan bahwa jumlah data pada kelas “non-bidang” sebanyak 640 dan kelas “bidang” sebanyak 232, dari total 872 data pelatihan. Ketimpangan ini menunjukkan adanya kondisi ketidakseimbangan data (*imbalanced data*). Untuk mengatasi hal tersebut, diterapkan strategi penanganan khusus yang dikenal sebagai *imbalance data handling*. Salah satu metode yang umum digunakan adalah SMOTE, yang berfungsi untuk menghasilkan data sintetis pada kelas minoritas. Dalam hal ini, SMOTE digunakan untuk menambah jumlah data pada kelas “bidang” dari 232 menjadi 640, sehingga distribusinya menjadi seimbang dengan kelas “non-bidang”.

3.1.4 Rasio 90:10

Distribusi data pada kelas “bidang” dan “non-bidang” sebelum maupun sesudah dilakukan penyeimbangan dengan metode SMOTE pada rasio 90:10 disajikan pada Gambar 5. Gambar tersebut menampilkan perbedaan jumlah data antara kedua kelas serta hasil akhir penyeimbangan setelah penambahan data sintetis pada kelas minoritas.



Gambar 5. Distribusi Kelas Sebelum dan Sesudah SMOTE Rasio 80:20

Pada Gambar 5, jumlah data pada kelas “non-bidang” tercatat sebanyak 720, sedangkan kelas “bidang” berjumlah 261 dari total 981 data pelatihan. Perbedaan jumlah yang cukup besar ini mencerminkan adanya ketidakseimbangan kelas (*class imbalance*). Untuk menangani kondisi tersebut, diperlukan pendekatan khusus yang dikenal sebagai *imbalance data handling*. Salah satu metode yang banyak digunakan adalah SMOTE, yaitu teknik yang menghasilkan data sintesis untuk menambah jumlah data pada kelas minoritas. Dalam konteks ini, SMOTE digunakan untuk meningkatkan jumlah data pada kelas “bidang” dari 261 menjadi 720, sehingga distribusinya setara dengan kelas “non-bidang”.

3.2 Confusion Matrix

3.2.1 Rasio 65:35

Hasil *confusion matrix* untuk pembagian data dengan rasio 65:35 dapat dilihat pada Tabel 1. Tabel ini menunjukkan distribusi prediksi model pada kelas “bidang” dan “non-bidang”.

Tabel 1. Confusion matrix rasio 65:35

Confusion Matrix		Prediksi	
		Bidang	Non Bidang
Aktual	Bidang	89	13
	Non Bidang	75	205

Tabel 1 menyajikan hasil confusion matrix berdasarkan rasio data pelatihan dan pengujian sebesar 65:35. Berdasarkan tabel tersebut, nilai *True Positive* (TP) tercatat sebesar 89, yang mengindikasikan bahwa sebanyak 89 data pada kelas “bidang” berhasil diklasifikasikan dengan benar. Nilai *True Negative* (TN) sebesar 205 menunjukkan bahwa 205 data pada kelas “non-bidang” juga diklasifikasikan secara akurat. Sementara itu, terdapat 75 data kelas “non-bidang” yang salah terklasifikasi sebagai “bidang”, yang ditunjukkan oleh nilai *False Positive* (FP). Adapun nilai *False Negative* (FN) sebesar 13, menandakan bahwa 13 data kelas “bidang” secara keliru diklasifikasikan sebagai “non-bidang”.

3.2.2 Rasio 70:30

Hasil *confusion matrix* untuk rasio 70:30 dapat dilihat pada Tabel 2.

Tabel 2. Confusion matrix Rasio 70:30

Confusion Matrix		Prediksi	
		Bidang	Non Bidang
Aktual	Bidang	74	4
	Non Bidang	58	183

Tabel 2 memperlihatkan hasil confusion matrix dari pembagian data dengan rasio 70:30. Berdasarkan tabel tersebut, nilai *True Positive* (TP) sebesar 74 menunjukkan bahwa sebanyak 74 data pada kelas “bidang” berhasil diklasifikasikan dengan benar. Sementara itu, *True Negative* (TN) tercatat sebesar 183, menandakan bahwa 183 data pada kelas “non-bidang” juga berhasil diklasifikasikan secara akurat. Di sisi lain, nilai *False Positive* (FP) sebesar 58 mengindikasikan adanya 58 data kelas “non-bidang” yang keliru diklasifikasikan sebagai “bidang”. Sedangkan nilai *False Negative* (FN) sebesar 13 menunjukkan bahwa 13 data pada kelas “bidang” salah diklasifikasikan sebagai “non-bidang”.

3.2.3 Rasio 80:20

Hasil *confusion matrix* untuk rasio 80:20 dapat dilihat pada Tabel 3.

**Tabel 3.** *Confusion matrix* rasio 80:20

<i>Confusion Matrix</i>		Prediksi	
		Bidang	Non Bidang
Aktual	Bidang	54	13
	Non Bidang	36	125

Tabel 3 menampilkan hasil confusion matrix dari pembagian data dengan rasio 80:20. Berdasarkan hasil tersebut, nilai *True Positive* (TP) sebesar 54 menunjukkan bahwa 54 data pada kelas “bidang” berhasil terklasifikasi dengan benar. Sementara itu, *True Negative* (TN) sebesar 125 mengindikasikan bahwa 125 data pada kelas “non-bidang” juga diklasifikasikan secara akurat. Di sisi lain, terdapat 36 data kelas “non-bidang” yang salah terklasifikasi sebagai “bidang”, sebagaimana ditunjukkan oleh nilai *False Positive* (FP). Adapun nilai *False Negative* (FN) sebesar 4 menunjukkan bahwa 4 data pada kelas “bidang” keliru diklasifikasikan sebagai “non-bidang”..

3.2.4 Rasio 90:10

Hasil *confusion matrix* untuk rasio 90:10 dapat dilihat pada Tabel 4.

Tabel 4. *Confusion matrix* rasio 90:10

<i>Confusion Matrix</i>		Prediksi	
		Bidang	Non Bidang
Aktual	Bidang	28	1
	Non Bidang	13	68

Tabel 4 dan Gambar 5.8 menampilkan hasil klasifikasi dengan nilai *True Positive* (TP) sebesar 28, yang menunjukkan bahwa 28 data pada kelas “bidang” berhasil diklasifikasikan secara tepat. Sementara itu, nilai *True Negative* (TN) sebesar 68 mengindikasikan bahwa 68 data pada kelas “non-bidang” juga berhasil dikenali dengan benar. Selain itu, terdapat 13 data *False Positive* (FP), yakni data dari kelas “non-bidang” yang secara keliru diklasifikasikan sebagai “bidang”, serta 1 data *False Negative* (FN), yaitu data kelas “bidang” yang salah diklasifikasikan sebagai “non-bidang”.

3.3 Evaluasi Performansi

Hasil evaluasi performansi berdasarkan metrik akurasi, presisi, recall, dan F1-score untuk setiap rasio pembagian data dapat dilihat pada Tabel 5.

Tabel 5. Hasil evaluasi performansi

Rasio	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
65:35	0.7696	0.7415	0.8023	0.7462
70:30	0.7835	0.7471	0.8050	0.7567
80:20	0.8174	0.7845	0.8537	0.7959
90:10	0.8727	0.8342	0.9025	0.8533

Berdasarkan Tabel 5, dapat diamati bahwa peningkatan proporsi data pelatihan cenderung sejalan dengan peningkatan akurasi model. Rasio 90:10 menghasilkan tingkat akurasi tertinggi, yaitu sebesar 87,27%. Namun, rasio ini juga menunjukkan indikasi adanya potensi *overfitting*, ditandai dengan akurasi data latih yang melebihi 99%. Oleh karena itu, disarankan agar penelitian selanjutnya mempertimbangkan penerapan metode seperti *cross-validation* atau teknik *regularisasi* untuk mengurangi risiko *overfitting*. Di samping itu, keberadaan *False Positive* (FP) dan *False Negative* (FN) memiliki implikasi penting. FP yakni *tweet* “non-bidang” yang salah diklasifikasikan sebagai “bidang” dapat menyebabkan alokasi sumber daya rektorat terhadap keluhan yang tidak relevan. Sementara itu, FN *tweet* “bidang” yang tidak teridentifikasi secara benar berisiko lebih besar karena berpotensi mengabaikan permasalahan penting yang berkaitan langsung dengan unit rektorat

Confusion matrix yang telah diperoleh digunakan sebagai dasar untuk mengevaluasi performa model. Dalam proses evaluasi ini, *confusion matrix* dimanfaatkan untuk menghitung metrik-metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Nilai-nilai tersebut disajikan pada Tabel 5.5. Berdasarkan hasil evaluasi performa yang dihitung melalui *confusion matrix*, dapat disimpulkan bahwa semakin besar jumlah data yang digunakan untuk pelatihan, maka semakin tinggi pula nilai evaluasi yang diperoleh.

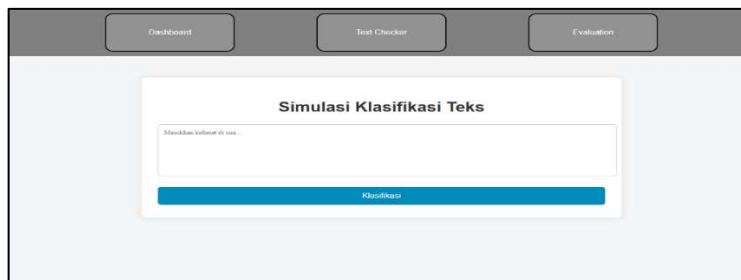
3.4 Visualisasi

Hasil visualisasi data dan antarmuka sistem dapat dilihat pada Gambar 6, Gambar 7, dan Gambar 8.



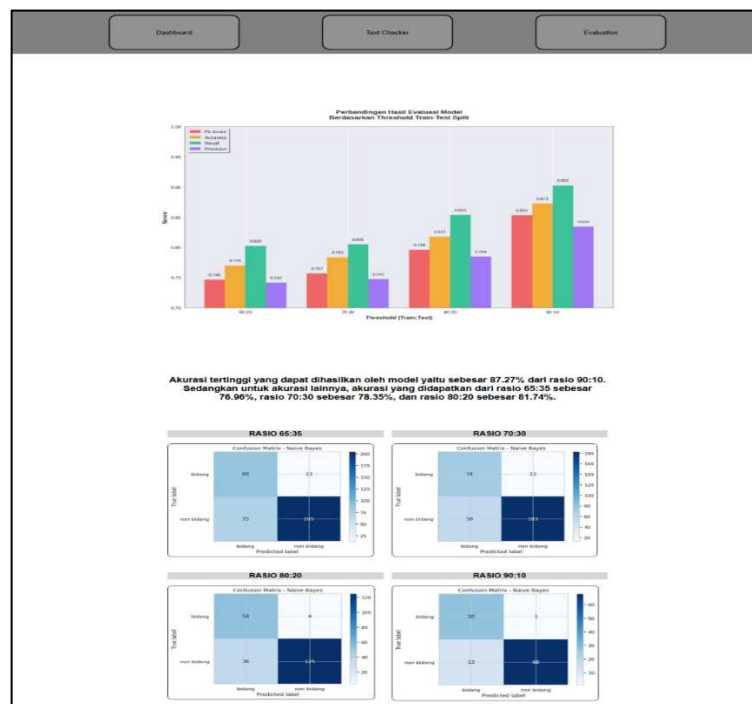
Gambar 6. Visualisasi dashboard

Halaman dashboard ini menyajikan kata-kata yang paling sering muncul pada masing-masing kelas. Berdasarkan Gambar 6, terlihat bahwa kata “semester” merupakan salah satu kata yang umum muncul di kedua kelas. Namun, terdapat perbedaan yang mencolok dalam konteks penggunaannya: kata-kata yang dominan pada kelas “bidang” umumnya berkaitan dengan lingkungan kampus, sedangkan kata-kata yang sering muncul pada kelas “non-bidang” lebih banyak mencerminkan aspek kehidupan mahasiswa secara umum.



Gambar 7. Visualisasi text checker

Halaman yang ditampilkan pada Gambar 7 memperlihatkan simulasi pemrosesan teks yang menyerupai tampilan aplikasi X. Pada halaman ini terdapat beberapa elemen, yaitu kolom input untuk memasukkan teks, tombol untuk memulai proses analisis, serta output yang menampilkan hasil dari pemrosesan teks tersebut.



Gambar 8. Visualisasi hasil evaluasi

Gambar 8 menyajikan visualisasi dari hasil evaluasi model yang telah dilakukan sebelumnya. Visualisasi ini mencakup hasil evaluasi dari keempat rasio pembagian data, yaitu 65:35, 70:30, 80:20, dan 90:10. Selain metrik evaluasi, gambar tersebut juga menampilkan confusion matrix untuk masing-masing rasio tersebut.



Penerapan metode SMOTE berhasil menyamakan distribusi data antara kelas “bidang” (sebagai kelas minoritas) dan “non-bidang”, sehingga model tidak bias terhadap kelas mayoritas. Penambahan data sintetis dari proses *oversampling* turut berkontribusi dalam peningkatan akurasi, meskipun berisiko menimbulkan noise karena tidak sepenuhnya mencerminkan pola data yang sebenarnya. Oleh karena itu, pada penelitian selanjutnya disarankan untuk menggunakan teknik validasi silang (*cross-validation*) atau metode regularisasi guna meminimalkan kemungkinan *overfitting* akibat penggunaan data sintetis.

Analisis terhadap nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) memberikan wawasan yang lebih komprehensif mengenai performa model. *False Positive* (FP), yaitu ketika *tweet* “non-bidang” keliru diklasifikasikan sebagai “bidang”, dapat menyebabkan pihak rektorat memfokuskan perhatian pada keluhan yang tidak relevan dengan layanan kampus. Sebaliknya, *False Negative* (FN), yakni *tweet* “bidang” yang salah terklasifikasi sebagai “non-bidang”, berdampak lebih serius karena dapat menyebabkan keluhan penting dari mahasiswa yang berkaitan dengan layanan rektorat terlewatkan. Oleh sebab itu, strategi pengoptimalan model di masa mendatang sebaiknya diarahkan untuk meminimalkan nilai FN agar keluhan yang krusial tetap dapat terdeteksi dengan baik.

Hasil penelitian menunjukkan bahwa rasio 90:10 memberikan akurasi tertinggi sebesar 87,27%, disertai dengan metrik evaluasi lainnya seperti *precision*, *recall*, dan *F1-score* yang relatif stabil. Capaian ini dipengaruhi oleh besarnya jumlah data pelatihan, yang memungkinkan model untuk mengenali lebih banyak pola. Namun, kondisi ini juga memunculkan *trade-off* berupa potensi *overfitting* (ditandai dengan akurasi data latih yang melebihi 99%) serta waktu pelatihan yang lebih lama. Oleh karena itu, disarankan untuk menerapkan metode seperti *cross-validation* guna memastikan performa generalisasi model tetap optimal.

Berdasarkan visualisasi yang ditampilkan pada dashboard (Gambar 5), kata “ya” merupakan kata yang paling sering muncul karena berfungsi sebagai transisi atau respons dalam interaksi, bukan sebagai kata yang memiliki makna substantif. Pada kelas “bidang”, kata-kata yang dominan umumnya berkaitan dengan layanan akademik, keuangan, serta fasilitas kampus, mencerminkan fokus keluhan yang melibatkan unit-unit di bawah rektorat. Sebaliknya, kata-kata yang sering muncul pada kelas “non-bidang” lebih merepresentasikan aspek sosial dalam kehidupan mahasiswa. Temuan ini dapat dimanfaatkan oleh Telkom University untuk menetapkan prioritas peningkatan layanan, khususnya dalam hal akademik dan fasilitas kampus pada kelas “bidang”, serta untuk memahami dinamika sosial mahasiswa melalui analisis kelas “non-bidang”.

4. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan *tweet* dari akun Telyufess ke dalam kategori “bidang” dan “non-bidang” dengan tingkat akurasi yang bervariasi, bergantung pada rasio pembagian data. Akurasi tertinggi dicapai pada rasio 90:10 dengan nilai 87,27%, diikuti oleh 81,74% (80:20), 78,35% (70:30), dan 76,96% (65:35). Temuan ini mengindikasikan bahwa semakin besar proporsi data pelatihan, maka kinerja model pada data uji cenderung semakin baik. Namun demikian, penelitian juga menemukan adanya indikasi *overfitting*, ditunjukkan oleh performa model pada data latih yang rata-rata melebihi 99%, sementara performa pada data uji mengalami penurunan yang cukup signifikan. Hal ini mengisyaratkan bahwa model terlalu menyesuaikan diri dengan data pelatihan dan berisiko kurang efektif saat diterapkan pada data baru. Untuk mengatasi permasalahan tersebut, disarankan agar penelitian selanjutnya menerapkan teknik regularisasi, *cross-validation*, atau pendekatan *ensemble learning* seperti *Random Forest* dan *XGBoost*. Selain itu, pengujian dengan *dataset* tambahan dari periode waktu yang berbeda juga direkomendasikan untuk mengukur kemampuan generalisasi model di luar data pelatihan.

REFERENCES

- [1] Q. E. Muftikhali, Y. C. Kurniawan, and D. W. Rahma, “Impact of enterprise systems on user performance using the IGRACIAS V.1 application at Telkom University Jakarta,” *J. Adv. Inf. Ind. Technol.*, vol. 6, no. 1, pp. 41–50, 2024, doi: 10.52435/jaiit.v6i1.540.
- [2] D. Rahmawati, N. P. Istyanto, and M. Nasrullah, “Adoption of knowledge management system: A study on how IGracias should be adopted for Telkom Campus,” *J. KomtekInfo*, vol. 7, no. 3, pp. 218–227, 2020, doi: 10.35134/komtekinfo.v7i3.82.
- [3] Rayuwati, H. Gemasih, and I. Nizar, “Implementasi algoritma Naïve Bayes untuk memprediksi tingkat penyebaran COVID,” *J. Ris. Rumpun Ilmu Tek.*, vol. 1, no. 1, pp. 38–46, 2022, doi: 10.55606/jurritek.v1i1.127.
- [4] D. Alita and R. A. Shodiqin, “Sentimen analisis vaksin Covid-19 menggunakan Naive Bayes dan *Support Vector Machine*,” *J. Artif. Intell. Technol. Inf.*, vol. 1, no. 1, pp. 1–12, 2023, doi: 10.58602/jaiti.v1i1.20.
- [5] T. Safitri, Y. Umaidah, and I. Maulana, “Analisis sentimen pengguna Twitter terhadap BTS menggunakan algoritma *Support Vector Machine*,” *J. Appl. Informatics Comput.*, vol. 7, no. 1, p. 34, 2023, doi: 10.30871/jaic.v7i1.5039.
- [6] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, “Analisis sentimen lokasi Ibukota Nusantara menggunakan algoritma Naïve Bayes dan KNN,” *J. KomtekInfo*, vol. 10, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [7] N. L. P. Merawati, A. Z. Amrullah, and Ismarmiaty, “Analisis sentimen dan pemodelan topik pariwisata Lombok menggunakan algoritma Naïve Bayes dan *Latent Dirichlet Allocation*,” *J. RESTI*, vol. 5, no. 1, pp. 123–131, 2021, doi: 10.29207/resti.v5i1.2587.
- [8] A. Z. Putri, S. Defit, and W. G. Nurcahyo, “Analisis sentimen publik terhadap program penurunan angka prevalensi *stunting* Indonesia menggunakan data Twitter dengan menggunakan metode Naïve Bayes,” *J. Soc. Sci. Res.*, vol. 4, no. 5, pp. 1978–1989, 2024, doi: 10.31004/innovative.v4i5.15180.
- [9] S. A. Putra and A. Wijaya, “Analisis sentimen artificial intelligence (AI) pada media sosial Twitter menggunakan metode *Lexicon*



- Based,” *JuSiTik J. Sist. dan Teknol. Inf. Komun.*, vol. 7, no. 1, pp. 21–28, 2023, doi: 10.32524/jusitik.v7i1.1042.
- [10] R. Puspita and A. Widodo, “Perbandingan metode KNN, *decision tree*, dan Naïve Bayes terhadap analisis sentimen pengguna layanan BPJS,” *J. Inform. Univ. Pamulang*, vol. 5, no. 4, p. 646, 2021, doi: 10.32493/informatika.v5i4.7622.
- [11] I. Habib Kusuma and N. Cahyono, “Analisis sentimen masyarakat terhadap penggunaan *E-Commerce* menggunakan algoritma *K-Nearest Neighbor*,” *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, 2023, doi: 10.30591/jpit.v8i3.5734.
- [12] E. W. Sholeha, S. Yunita, R. Hammad, V. C. Hardita, and K. Kaharuddin, “Analisis sentimen pada agen perjalanan *online* menggunakan Naïve Bayes dan *K-Nearest Neighbor*,” *JTIM J. Teknol. Inf. dan Multimed.*, vol. 3, no. 4, pp. 203–208, 2022, doi: 10.35746/jtim.v3i4.178.
- [13] K. Kevin, M. Enjeli, and A. Wijaya, “Analisis sentimen penggunaan aplikasi Kinemaster menggunakan metode Naive Bayes,” *J. Ilm. Comput. Sci.*, vol. 2, no. 2, pp. 89–98, 2024, doi: 10.58602/jics.v2i2.24.
- [14] Muhammad Daffa Al Fahreza, Ardytha Luthfiarta, Muhammad Rafid, and Michael Indrawan, “Analisis sentimen: pengaruh jam kerja terhadap kesehatan mental generasi Z,” *J. Appl. Comput. Sci. Technol.*, vol. 5, no. 1, pp. 16–25, 2024, doi: 10.52158/jacost.v5i1.715.
- [15] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, “Pengaruh *text preprocessing* terhadap analisis sentimen komentar masyarakat pada media sosial Twitter (studi kasus pandemi COVID-19),” *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 406, 2021, doi: 10.30865/mib.v5i2.2835.
- [16] M. H. Mahendra, D. T. Murdiansyah, and K. M. Lhaksmana, “Analisis sentimen *tweet* COVID-19 menggunakan *K-Nearest Neighbors* dengan TF-IDF dan ekstraksi fitur *CountVectorizer*,” *DIKE J. Ilmu Multidisiplin*, vol. 1, no. 2, pp. 37–43, 2023, doi: 10.69688/dike.v1i2.35.
- [17] O. I. Gifari, M. Adha, F. Freddy, and F. F. S. Durrand, “Film review sentiment analysis using TF-IDF and Support Vector Machine,” *J. Inf. Technol.*, vol. 2, no. 1, pp. 36–40, 2022, doi: 10.46229/jifotech.v2i1.330.
- [18] R. Merdiansah, S. Siska, and A. Ali Ridha, “Analisis sentimen pengguna X Indonesia terkait kendaraan listrik menggunakan IndoBERT,” *J. Ilmu Komput. dan Sist. Inf.*, vol. 7, no. 1, pp. 221–228, 2024, doi: 10.55338/jikomsi.v7i1.2895.
- [19] R. Nurhidayat and K. E. Dewi, “Penerapan algoritma *K-Nearest Neighbor* dan fitur ekstraksi N-Gram dalam analisis sentimen berbasis aspek,” *Komputa J. Ilm. Komput. dan Inform.*, vol. 12, no. 1, pp. 91–100, 2023, doi: 10.34010/komputa.v12i1.9458.
- [20] I. G. B. Premana Putra, M. Sudarma, and I. B. G. Manuaba, “Penerapan metode *extreme programming* pada rancang bangun sistem analisis sentimen portal berita,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 6, pp. 1369–1378, 2023, doi: 10.25126/jtiik.2023106904.