

KOMPARASI METODE ALGORITMA NAIVE BAYES DAN K-NEAREST NEIGHBOR PADA ANALISIS SENTIMEN APLIKASI KOPI KENANGAN

Achmad farhan Surachman ¹, Suhardjono ², Dwi Irawati ³

^{1,2,3} Program Studi Sistem Informasi, Fakultas Teknik dan Informatika
Universitas Bina Sarana Informatika
achmadfarhan376@gmail.com

ABSTRAK

Banyaknya ulasan pengguna aplikasi Kopi Kenangan pada Google Play Store menghasilkan data teks dalam jumlah besar sehingga menyulitkan analisis manual, padahal informasi sentimen pengguna penting untuk meningkatkan kualitas layanan dan daya saing perusahaan. Penelitian ini bertujuan membandingkan kinerja algoritma *Naive Bayes* dan *K-Nearest Neighbor* atau KNN dalam mengklasifikasikan sentimen ulasan tersebut berdasarkan akurasi, *precision*, *recall*, dan *ROC-AUC*. Data berupa 1.000 ulasan dikumpulkan dari Google Play Store menggunakan Python dengan *library google-play-scraper*, kemudian dianalisis dengan kedua algoritma melalui tahapan *pre-processing* teks dan dievaluasi menggunakan *confusion matrix* serta kurva *ROC*. Hasil menunjukkan KNN memiliki akurasi dan kemampuan diskriminasi kelas lebih tinggi, yaitu akurasi sebesar 85,63% dan *ROC-AUC* sebesar 0,821 yang termasuk kategori “Baik”, sedangkan *Naive Bayes* memperoleh akurasi sebesar 69,68% dan *ROC-AUC* sebesar 0,612 yang termasuk kategori “Kurang Baik”. Meskipun demikian, *Naive Bayes* unggul pada beberapa metrik berbasis kelas, yaitu *precision* kelas positif sebesar 88,93% dibandingkan KNN 85,51%, serta *recall* kelas negatif sebesar 68,52% dibandingkan KNN sebesar 39,81%. Dengan demikian, KNN lebih baik berdasarkan akurasi keseluruhan dan *ROC-AUC*, sedangkan *Naive Bayes* unggul pada kemampuan tertentu dalam mengklasifikasikan kelas positif dan negatif. Hasil ini menunjukkan bahwa penilaian kinerja algoritma perlu mempertimbangkan karakteristik setiap metrik evaluasi, tidak hanya satu ukuran performa.

Keyword : Analisis Sentimen, *Naive Bayes*, *K-Nearest Neighbor*, *Kopi Kenangan*, *Google Play Store*

1. PENDAHULUAN

Industri kopi di Indonesia memiliki peran yang cukup besar dalam sektor perkebunan dan konsumsi domestik. Berdasarkan data Badan Pusat Statistik, produksi kopi Indonesia pada tahun 2023 mencapai sekitar 758,73 ribu ton, dengan sebagian besar produksi berasal dari perkebunan rakyat. Data tersebut menunjukkan bahwa Indonesia memiliki kapasitas produksi kopi yang besar, meskipun produksi pada periode 2021–2023 mengalami penurunan. Di sisi konsumsi, permintaan kopi domestik juga tetap tinggi seiring berkembangnya kebiasaan konsumsi kopi di masyarakat. Kondisi tersebut memperlihatkan bahwa industri kopi Indonesia tetap memiliki pasar yang luas, baik dari sisi produksi maupun konsumsi. [1].

Pertumbuhan ini turut didorong oleh transformasi digital yang mengubah pola konsumsi dan perilaku bisnis di sektor kopi. Generasi muda menjadikan kopi sebagai bagian dari gaya hidup, sementara kemajuan teknologi digital mendorong lahirnya aplikasi mobile yang mempermudah interaksi konsumen dengan merek serta mempercepat proses transaksi [2]. Salah satu pelaku industri yang memanfaatkan tren ini secara masif adalah Kopi Kenangan, yang sejak tahun 2017 mengandalkan teknologi digital untuk memperkuat posisinya sebagai pemain dominan dalam industri kopi lokal. Ekspansi jaringan gerai Kopi Kenangan tercatat mencapai 868 outlet pada tahun 2023, dengan target lebih dari 1.100 gerai pada akhir tahun

2025 [3]. Strategi digitalisasi ini juga diwujudkan melalui *Kenangan App*, aplikasi pemesanan yang mengintegrasikan berbagai metode pembayaran digital dan promosi berbasis media sosial. Tingginya penerimaan pengguna terhadap aplikasi ini tercermin dari rating 4,9 dari sekitar 160.000 ulasan serta jumlah unduhan yang telah melampaui 1 juta unduhan di *Google Play Store* [3].

Di balik capaian tersebut, volume ulasan yang besar justru menimbulkan tantangan baru dalam pengelolaan dan pemahaman data, mengingat data ulasan berbentuk teks tidak terstruktur dalam jumlah besar sulit dianalisis secara manual, padahal pemahaman terhadap sentimen pengguna sangat penting bagi pengembangan layanan dan daya saing perusahaan. Analisis sentimen hadir sebagai pendekatan untuk memproses teks yang tidak terstruktur dengan tujuan menemukan pola sentimen positif atau negatif, dan telah banyak diterapkan pada berbagai bidang seperti sosial, politik, dan ekonomi, termasuk dalam menganalisis emosi masyarakat di media sosial [4].

Dalam penerapannya, pemilihan algoritma klasifikasi menjadi faktor penentu akurasi hasil analisis sentimen. Dua algoritma yang sering dipakai adalah *Naive Bayes*, yang beroperasi secara probabilistik dan efisien dalam pengolahan data teks, dan *K-Nearest Neighbor* (KNN), yang mengklasifikasikan data berdasarkan kedekatan jarak antar data [5]. Beberapa penelitian terdahulu menunjukkan hasil yang bervariasi ketika

membandingkan performa kedua algoritma ini. Penelitian terhadap ulasan aplikasi e-wallet menemukan bahwa KNN unggul dengan akurasi 89,44% dibandingkan *Naive Bayes* yang hanya mencapai 73,03% [5]. Temuan serupa juga dilaporkan pada analisis ulasan aplikasi Rosalia Indah Transport, di mana KNN mencatatkan akurasi 81,54% dan *Naive Bayes* 64,94% [6]. Penelitian terdahulu menunjukkan bahwa kinerja *Naive Bayes* dan KNN dipengaruhi oleh karakteristik dataset dan rancangan pengujian yang digunakan. Maharani et al. [5] membandingkan kedua algoritma pada data sentimen penggunaan e-wallet dan memperoleh akurasi KNN sebesar 89,44%, lebih tinggi dibandingkan *Naive Bayes* sebesar 73,03%. Sementara itu, Satrio dan Pratama [6] menggunakan ulasan aplikasi Rosalia Indah Transport dan memperoleh akurasi KNN sebesar 81,54%, sedangkan *Naive Bayes* sebesar 64,94%. Perbedaan sumber data, distribusi kelas, tahapan *pre-processing*, representasi fitur, konfigurasi parameter, strategi validasi, dan metrik evaluasi dapat memengaruhi hasil klasifikasi. Informasi mengenai konfigurasi parameter pada beberapa penelitian terdahulu juga belum selalu dilaporkan secara rinci sehingga hasil antardataset tidak dapat langsung digeneralisasikan. Kondisi tersebut menunjukkan perlunya pengujian pada dataset dengan karakteristik berbeda dan evaluasi menggunakan lebih dari satu metrik.

Berdasarkan kesenjangan tersebut, penelitian ini membandingkan *Naive Bayes* dan KNN pada ulasan pengguna aplikasi Kopi Kenangan di *Google Play Store*. Kebaruan penelitian ini bersifat empiris dan evaluatif, bukan pengembangan algoritma baru, yaitu menguji perilaku kedua algoritma pada dataset ulasan Kopi Kenangan yang memiliki distribusi kelas tidak seimbang, terdiri atas 780 ulasan positif dan 216 ulasan negatif dari 996 data setelah normalisasi. Perbandingan dilakukan menggunakan *accuracy*, *precision*, dan *recall* pada masing-masing kelas serta ROC-AUC sehingga dapat menunjukkan *trade-off* performa yang tidak terlihat apabila evaluasi hanya didasarkan pada akurasi keseluruhan. Dengan demikian, penelitian ini memberikan informasi yang lebih spesifik mengenai kemampuan kedua algoritma dalam mengenali sentimen positif dan negatif pada karakteristik data ulasan aplikasi dengan distribusi kelas tidak seimbang.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen merupakan cara untuk menganalisis nada emosional dalam teks digital guna mengetahui apakah itu positif, negatif, atau netral. Analisis sentimen merupakan proses menilai pandangan, perasaan, penilaian, emosi, sikap, atau opini mengenai suatu produk, tokoh, institusi, isu, layanan, serta kejadian. Di samping itu, analisis sentimen selalu berkaitan dengan masyarakat karena

sumber datanya berasal dari media sosial, di mana masyarakat menggunakannya [7].

2.2. Text Mining

Pengolahan teks, juga dikenal sebagai text mining, adalah proses penggalian data berbasis teks yang relevan dengan tujuan pengguna untuk menemukan hubungan atau korelasi yang signifikan dengan memanfaatkan alat analisis yang merupakan bagian dari proses data mining. Proses pengolahan teks terdiri dari beberapa tahap *pre-processing* atau pembersihan data teks sebelum data diklasifikasikan ke tahap selanjutnya [8].

2.3. Naive Bayes

Naive Bayes merupakan algoritma klasifikasi probabilistik yang menggunakan probabilitas bersyarat berdasarkan *Teorema Bayes* untuk menentukan kemungkinan suatu data termasuk ke dalam kelas tertentu. Pada penelitian ini, algoritma *Naive Bayes* diimplementasikan menggunakan operator *Naive Bayes* pada *RapidMiner* dengan konfigurasi parameter default tanpa penyesuaian lebih lanjut. Model digunakan untuk mengklasifikasikan ulasan pengguna ke dalam kelas sentimen positif dan negatif. Rincian varian dan konfigurasi parameter model perlu disesuaikan dengan pengaturan operator yang digunakan pada *RapidMiner* [9].

Berikut bentuk secara umum *Teorema Bayes* berbentuk:

$$P(H|X) = \frac{(P(X|H) \times P(H))}{P(X)} \quad (1)$$

Pada persamaan (1), X adalah sampel data yang kelas atau labelnya belum diketahui, sedangkan H adalah hipotesis bahwa X termasuk dalam class tertentu. $P(H|X)$ menunjukkan kemungkinan hipotesis H berdasarkan kondisi X , $P(X|H)$ menunjukkan kemungkinan munculnya X dengan asumsi H benar, dan $P(X)$ menunjukkan kemungkinan umum munculnya data X .

2.4. K-Nearest Neighbor (KNN)

K-Nearest Neighbor (*K-Nearest Neighbor* (KNN)) merupakan algoritma klasifikasi yang menentukan kelas suatu data berdasarkan kelas mayoritas dari sejumlah k tetangga terdekat pada data latih. Kedekatan antardata ditentukan berdasarkan ukuran jarak tertentu, sedangkan nilai k menentukan jumlah tetangga yang digunakan dalam proses klasifikasi. Pada penelitian ini, algoritma KNN diimplementasikan menggunakan *RapidMiner* dengan konfigurasi parameter default tanpa modifikasi lebih lanjut. Konfigurasi model mencakup nilai k , ukuran jarak, dan strategi pembobotan yang digunakan dalam proses klasifikasi. Seluruh konfigurasi tersebut diterapkan secara konsisten pada proses *cross validation*. Pada ruang ini, sebuah titik diidentifikasi dengan kelas tertentu, yaitu label kelas yang paling sering muncul di antara k buah tetangga terdekat pada titik itu, yang

mana jarak antar tetangga umumnya dihitung dengan menggunakan jarak *Euclidean* [10]. Rumus *K-Nearest Neighbor* berikut dapat digunakan untuk menghitung jarak *Euclidean*:

$$d(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2} \quad (2)$$

Pada persamaan (2), d menyatakan jarak antara dua titik data yang dibandingkan, yaitu titik x dan titik y . Variabel n menunjukkan jumlah kriteria atau atribut yang digunakan dalam perhitungan, sedangkan x_i dan y_i masing-masing menyatakan nilai kriteria ke- i dari data pertama dan data kedua yang dibandingkan.

2.5. RapidMiner

Rapidminer adalah aplikasi atau perangkat lunak untuk pengetahuan data mining yang dikembangkan oleh perusahaan yang berfokus pada semua aplikasi menangani data besar dalam sektor perdagangan, penelitian, pendidikan, pelatihan, dan pembelajaran. *Rapidminer* adalah software independen yang dapat diintegrasikan dengan berbagai bahasa pemrograman dan memiliki kemampuan untuk menganalisis dan menambang data [11].

2.6. Confusion Matrix

Confusion Matrix menyajikan hasil prediksi secara keseluruhan dari tingkat akurasi dan untuk mengevaluasi kinerja pengklasifikasi, yaitu seberapa sering kasus kelas X yang benar diidentifikasi sebagai kelas X atau terjadi kesalahan klasifikasi kelas yang lain. Pada situasi di mana dataset hanya terdiri dari dua kelas, yaitu kelas positif dan kelas negatif. Jumlah catatan positif dalam dataset yang diklasifikasikan sebagai positif disebut *True Positif* (TP) dan *False Positif* (FP) [12].

Tabel 1. *Confusion Matrix*

Prediction	Actual	
	Positive	Negative
Positive	True Positive	False Negative
Negative	False Positive	True Negative

Perhitungan nilai *accuracy*, *precision*, dan *recall* dalam studi ini didasarkan pada *confusion matrix* seperti ditunjukkan pada Tabel 1, yang terdiri atas empat komponen utama. *True Positive* (TP) adalah data positif yang berhasil diidentifikasi secara akurat oleh model, sedangkan *False Positive* (FP) merupakan data negatif yang keliru dikategorikan sebagai positif. Sebaliknya, *True Negative* (TN) merupakan data negatif yang berhasil diidentifikasi secara akurat, sementara *False Negative* (FN) merupakan data positif yang keliru dikategorikan sebagai negatif.

Berdasarkan Tabel 1, dapat dihitung tiga metrik evaluasi, yaitu *accuracy*, *precision*, dan *recall*. Ketiga metrik tersebut dihitung menggunakan rumus berikut.

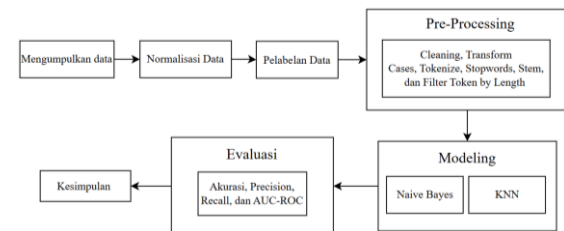
$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (5)$$

3. METODE PENELITIAN

Penelitian ini dilaksanakan melalui berbagai tahap seperti yang terlihat pada Gambar 1. Tahap pertama merupakan pengumpulan data, dilanjutkan dengan pelabelan data. Selanjutnya, data melalui tahap *pre-processing* yang meliputi *cleaning*, *transform cases*, *tokenize*, *stopwords*, *stemming*, dan *filter token by length*. Data yang telah melalui tahap *pre-processing* selanjutnya digunakan sebagai masukan dalam proses pemodelan menggunakan algoritma *Naive Bayes* dan KNN pada RapidMiner. Representasi fitur numerik mengikuti konfigurasi pada proses pengolahan teks yang digunakan dalam RapidMiner. Kedua model kemudian dievaluasi melalui *cross validation* menggunakan metrik *accuracy*, *precision*, *recall*, dan ROC-AUC untuk membandingkan performa klasifikasi sentimen.



Gambar 1. Tahapan penelitian

3.1. Pengumpulan data

Penelitian ini menggunakan data ulasan aplikasi Kopi Kenangan pada *Google Play Store* dengan populasi sebanyak 203.000 ulasan. Teknik pengambilan sampel yang digunakan adalah *non-probability sampling* dengan metode *purposive sampling*. Data dikumpulkan berdasarkan kriteria ulasan berbahasa Indonesia (*lang='id'*), berasal dari pengguna Indonesia (*country='id'*), mencakup seluruh rentang rating 1–5 (*filter_score_with=None*), dan diurutkan berdasarkan tingkat relevansi (*Sort.MOST_RELEVANT*) menggunakan *library google-play-scraper*. Berdasarkan kriteria tersebut, ditetapkan sebanyak 1.000 ulasan sebagai dataset awal penelitian. Data dikumpulkan melalui *Visual Studio Code* menggunakan bahasa pemrograman Python, kemudian disimpan dalam format CSV. Dari atribut yang diperoleh, penelitian ini menggunakan dua atribut utama, yaitu *content* sebagai teks ulasan dan *score* sebagai dasar pelabelan sentimen. Setelah proses normalisasi untuk menghapus data duplikat serta data kosong pada atribut *content* dan *score*, jumlah data yang digunakan dalam analisis menjadi 996 ulasan. Pemilihan 1.000 ulasan dilakukan untuk memperoleh jumlah data yang memadai dalam perbandingan dua algoritma klasifikasi dengan tetap

mengikuti kriteria pengambilan data yang telah ditetapkan.

3.2. Normalisasi Data

Pada tahapan normalisasi berguna untuk data lebih efisien dengan dilakukan menggunakan filter *exampel* sebagai menghilangkan data yang tidak lengkap, *remove duplicate* digunakan menghapus data duplikat, dan *nominal to text* digunakan untuk mengonversi format atribut content menjadi format teks.

3.3. Pelabelan Data

Pelabelan sentimen dilakukan berdasarkan rating bintang sebagai *proxy label*, dengan rating 4–5 dikategorikan positif dan 1–3 dikategorikan negatif. Pendekatan ini dipilih untuk memungkinkan pelabelan data secara konsisten pada dataset berukuran besar. Meskipun demikian, rating numerik tidak selalu sepenuhnya merepresentasikan sentimen teks, khususnya pada ulasan dengan rating 3 yang dapat mengandung opini campuran. Oleh karena itu, keterbatasan validitas label berbasis rating perlu dipertimbangkan dalam interpretasi hasil, dan penelitian selanjutnya disarankan melakukan validasi manual atau anotasi independen.

3.4. Pre-Processing

Pre-processing adalah langkah untuk mengonversi data tidak terstruktur menjadi data terstruktur sesuai dengan format yang diperlukan, dengan tujuan agar data tersebut mudah untuk dibaca [13].

a. Cleaning

Cleaning digunakan untuk menghilangkan istilah yang tidak penting dari dokumen. Tujuan langkah ini adalah untuk membuat dokumen yang sedang dikerjakan lebih sesuai. Sebagai contoh, menghilangkan simbol seperti: `[-!"#$%&'()*+,-./:;<=>?@[\\]\_`{|}~]` [14].

b. Transform Cases

Proses *Transform Cases* mengubah mengubah seluruh huruf dalam teks menjadi satu bentuk huruf, biasanya huruf kecil (*lowercase*). Proses ini digunakan dengan tujuan membuat teks lebih konsisten dan lebih mudah untuk dipahami [4].

c. Tokenize

Proses ini memecah data teks menjadi elemen-elemen kecil yang disebut token, seperti kata atau frasa, sehingga komputer lebih mudah memahami dan menganalisis teks [4].

d. Stopwords

Stopword diubah menjadi kalimat yang istilah yang sering muncul dan dianggap tidak relevan, seperti kata penghubung atau penunjuk waktu, yang perlu dihapus [4].

e. Stemming

Tahap *stemming* dilakukan untuk mengubah kata berimbuhan menjadi bentuk dasar agar variasi kata yang memiliki akar sama dapat

diperlakukan sebagai satu istilah. Pada penelitian ini, proses *stemming* dilakukan menggunakan operator *Stem (Dictionary)* pada RapidMiner dengan memanfaatkan kamus kata sebagai acuan perubahan bentuk kata. Proses ini tidak mencakup koreksi ejaan secara khusus, sehingga fungsi *stemming* dibatasi pada normalisasi kata ke bentuk dasarnya. Hasil proses ini dapat dilihat pada perubahan beberapa kata, seperti “memudahkan” menjadi “mudah” dan “penjelasan” menjadi “jelas”. [4].

f. Filter Tokens by Length

Tahap *Filter Tokens by Length* digunakan untuk menyaring token berdasarkan panjang karakter. Pada penelitian ini, token dengan panjang kurang dari 4 karakter dan lebih dari 25 karakter dihapus sebelum tahap pemodelan. Batas tersebut mengikuti konfigurasi *pre-processing* yang digunakan dalam penelitian dan bertujuan untuk mengurangi token yang sangat pendek atau sangat panjang yang dinilai kurang informatif bagi proses klasifikasi. Contoh token yang terhapus pada tahap ini antara lain “yg”, “jg”, “aja”, dan token pendek lainnya. Meskipun demikian, penerapan ambang 4–25 karakter berpotensi menghilangkan token pendek yang masih memiliki makna sentimen atau konteks tertentu. Oleh karena itu, pengaruh pemilihan ambang panjang token perlu dipertimbangkan sebagai keterbatasan penelitian, dan penelitian selanjutnya dapat melakukan analisis sensitivitas terhadap beberapa variasi batas panjang token untuk mengetahui pengaruhnya terhadap performa klasifikasi. [6].

3.5. Modeling

Tahapan modeling dilakukan dengan perangkat lunak RapidMiner sebagai alat untuk diterapkan dalam analisis data dipilih, yaitu *Naive Bayes* dan *K-Nearest Neighbor*. Tujuan pemilihan algoritma ini, untuk memastikan bahwa metode yang dipilih adalah yang paling efektif untuk mencapai hasil analisis data yang diinginkan.

3.6. Evaluasi

Evaluasi dilaksanakan untuk menilai performa dua teknik klasifikasi, yakni *K-Nearest Neighbors* (KNN) dan *Naive Bayes*, dengan memanfaatkan *confusion matrix*. Indikator evaluasi yang diterapkan mencakup akurasi, *precision*, *recall*, dan kurva ROC-AUC. Teknik *cross-validation* digunakan untuk mengevaluasi kemampuan generalisasi model terhadap data yang tidak digunakan secara langsung pada proses pelatihan. Pada penelitian ini, evaluasi dilakukan menggunakan 10-fold *cross-validation* dengan *stratified sampling*, sehingga proporsi kelas tetap terjaga pada setiap *fold* pengujian. Pendekatan ini digunakan untuk memperoleh estimasi performa model yang lebih stabil dan mengurangi

ketergantungan hasil pada satu pembagian data latihan dan data uji.

4. HASIL DAN PEMBAHASAN

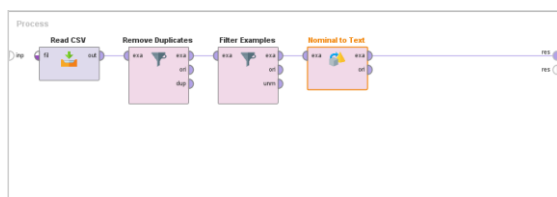
4.1. Pengumpulan Data

Data diambil dari *review* aplikasi "Kopi Kenangan" yang ada di *Google Play Store* dengan menggunakan bahasa pemrograman *python* serta *library google-play-scraper* untuk memperoleh data berdasarkan relevansi. Informasi yang didapat dari ulasan 1000 dataset mencakup beberapa atribut, seperti *field reviewId*, *userName*, *userImage*, *content*, *score*, *thumbsUpCount*, *reviewCreatedVersion*, *at*, *replyContent*, *repliedAt*, dan *appVersion*. Penulis hanya menggunakan dua atribut, yaitu "*content*" dan "*score*" dari seluruh atribut yang diperoleh.

Gambar 2. Hasil Pengumpulan Data

4.2. Normalisasi Data

Untuk meningkatkan efisiensi pemrosesan data, hasil pengumpulan data dinormalisasi. Proses ini bertujuan untuk menghilangkan data duplikat serta data kosong atau tidak ada dari atribut "*content*" dan "*score*" untuk meningkatkan alur kerja serta tidak menghambat dalam pemrosesan. Dengan demikian, tidak hanya tingkat efisiensi yang meningkat, tetapi juga kualitas dan keandalan data yang diperoleh selama proses analisis meningkat.



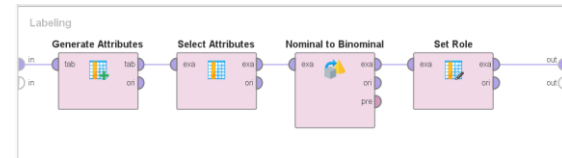
Gambar 3. Operator Normalisasi Data

Hasil normalisasi menghasilkan 996 data yang berubah dari 1000 data awal. Hasil menunjukkan data berulang, atribut *content* dan *score* yang kosong atau nihil

4.3. Pelabelan Data

Hasil normalisasi data akan dibagi menjadi dua kelompok, yaitu kelompok positif dan kelompok negatif sebagai label. Kelompok ini ditetapkan berdasarkan bintang atau "*score*", ulasan bintang 4 dan 5 menunjukkan kelompok positif, sedangkan bintang 3, 2 dan 1 menunjukkan kelompok negatif. Pelabelan data dilakukan dengan *software*

Rapidminer dengan menggunakan beberapa operator seperti Gambar 4.



Gambar 4. Operator Pelabelan Data

Hasil proses pelabelan menghasilkan 996 data yang terdiri atas 780 ulasan berkategori positif dan 216 ulasan berkategori negatif. Distribusi tersebut menunjukkan adanya ketidakseimbangan kelas, dengan proporsi ulasan positif yang lebih besar dibandingkan ulasan negatif. Kondisi ini perlu diperhatikan dalam interpretasi hasil klasifikasi karena penggunaan nilai akurasi saja berpotensi memberikan gambaran performa yang kurang representatif terhadap kelas minoritas. Oleh karena itu, evaluasi model tidak hanya didasarkan pada *accuracy*, tetapi juga mempertimbangkan *precision* dan *recall* pada masing-masing kelas, ROC-AUC, serta metrik yang lebih sensitif terhadap ketidakseimbangan kelas seperti *macro-averaged F1-score* dan *balanced accuracy*. Berikut hasil pelabelan kalimat ulasan aplikasi Kopi Kenangan pada Tabel 2.

Tabel 2. Hasil Proses *Labeling*

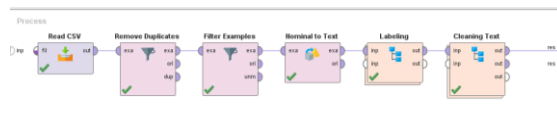
Content	Label
rasa kopi seperti kopi murah beda dgn dulu bahan kopinya sekarang kualitas rendah rasa pahit kurang pandai meracikanya roti kempes harga mahal roti coklatnya dicampur sagu rasa jd kurang enak harap perbaiki harga kopi mahal pdhl rasa tak jauh dari kopi kaki lima jarang promo dan kalau promo juga tidak sesuai harganya tetap mahal	Positif
Aplikasi apa ini? memudahkan tapi Customer service balasnya lama sekali seperti balas artis papan atas, saya ada masalah di aplikasi ini, pihak kopi kenangan tidak menemukan driver, sedangkan uang saya sudah masuk di rekening kopi kenangan, minta penjelasan, tidak ada kejelasannya sekali, balasan customer servicenya lama, seperti artis atau selebgram, lama sekali, padahal pelanggan sedang kebingungan, malah tidak segera membuka chat customer	Negatif

Berdasarkan metrik yang mempertimbangkan ketidakseimbangan kelas, *Naive Bayes* memperoleh *macro F1-score* sebesar 63,92% dan *balanced accuracy* sebesar 69,26%, sedangkan KNN memperoleh *macro F1-score* sebesar 73,04% dan *balanced accuracy* sebesar 69,07%. Hasil ini menunjukkan bahwa keunggulan KNN pada akurasi keseluruhan tidak sepenuhnya tercermin pada *balanced accuracy*, karena kemampuan kedua

algoritma dalam mengenali masing-masing kelas relatif lebih berimbang ketika sensitivitas terhadap distribusi kelas diperhitungkan.

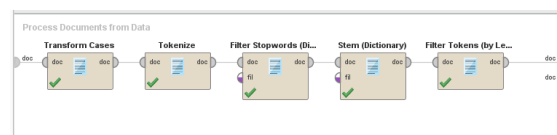
4.4. Pre-Processing

Perangkat lunak *RapidMiner* melalui proses blank sebagai titik awal pengerjaan. Proses membaca file dengan tipe csv, *Remove Duplicates* untuk menghapus ulasan yang bersifat duplikat, *Filter Examples* untuk menyaring kalimat yang kurang relevan, *Cleaning* untuk mengeliminasi karakter *noise* seperti simbol dan tanda baca, serta *Nominal to Text* untuk mengkonversi tipe atribut dari *nominal* menjadi teks, sebagaimana ditunjukkan pada Gambar dibawah ini.



Gambar 5. Operator *Cleaning Data*

Lima *sub operator* dalam *operator Proses Documents from Data* tersusun secara sekuensial, yakni *Tokenize* yang memecah teks ulasan menjadi unit kata individual, *Transform Cases* untuk menyeragamkan seluruh karakter menjadi huruf kecil, *Filter Stopwords Dictionary* untuk menghapus kata tidak bermakna berdasarkan file kamus *stopwords* bahasa Indonesia, *Stemming Dictionary* untuk mengubah kata-kata menjadi bentuk asalnya, serta *Filter Token by Length* yang menyaring token yang berada di luar rentang 4 hingga 25 karakter, sebagaimana diperlihatkan pada Gambar 6. *Process Documents from Data*.



Gambar 6. Operator *Pre-Processing*

4.5. Cleaning

Pada tahap ini, proses ini bertujuan untuk menghilangkan simbol, karakter khusus, dan angka. Berikut adalah hasil dari *cleaning* pada kalimat ulasan aplikasi Kopi Kenangan.

Tabel 3. Hasil Proses *Cleaning*

Sebelum <i>Cleaning</i>	Sesudah <i>Cleaning</i>
Dari sekian byk coffe shop aq psti lbh pilih kopi kenangan. pertama kali cobain kopi kenangan mantan lgsung jatuh cinta â?â? krn beneran berasa kopinya dan ga cuma rasa susu yg dominan. krn aq kurang suka susu. jdi cocok banget buat aq	Dari sekian byk coffe shop aq psti lbh pilih kopi kenangan pertama kali cobain kopi kenangan mantan lgsung jatuh cinta krn beneran berasa kopinya dan ga cuma rasa susu yg dominan krn aq kurang suka susu jdi cocok banget buat aq penikmat

Sebelum <i>Cleaning</i>	Sesudah <i>Cleaning</i>
penikmat kopi ringan dan yg ga susu banget.	kopi ringan dan yg ga susu banget

4.6. Transform Cases

Tahapan ini, setiap karakter dalam teks diubah menjadi huruf kecil. Tujuannya ialah untuk membandingkan kata-kata yang seharusnya identik namun memiliki huruf yang berbeda.

Tabel 4. Hasil Proses *Transform Cases*

Sebelum <i>Transform Cases</i>	Sesudah <i>Transform Cases</i>
Dari sekian byk coffe shop aq psti lbh pilih kopi kenangan pertama kali cobain kopi kenangan mantan lgsung jatuh cinta krn beneran berasa kopinya dan ga cuma rasa susu yg dominan krn aq kurang suka susu jdi cocok banget buat aq penikmat kopi ringan dan yg ga susu banget	dari sekian byk coffe shop aq psti lbh pilih kopi kenangan pertama kali cobain kopi kenangan mantan lgsung jatuh cinta krn beneran berasa kopinya dan ga cuma rasa susu yg dominan krn aq kurang suka susu jdi cocok banget buat aq penikmat kopi ringan dan yg ga susu banget

4.7. Tokenize

Tokenize berguna untuk memisahkan teks menjadi elemen-elemen kecil yang disebut token dalam data ulasan. Setelah menggunakan *tokenize* pada kalimat ulasan aplikasi Kopi Kenangan, hasilnya adalah sebagai berikut.

Tabel 5. Hasil Proses *Tokenize*

Sebelum <i>Tokenize</i>	Sesudah <i>Tokenize</i>
dari sekian byk coffe shop aq psti lbh pilih kopi kenangan pertama kali cobain kopi kenangan mantan lgsung jatuh cinta krn beneran berasa kopinya dan ga cuma rasa susu yg dominan krn aq kurang suka susu jdi cocok banget buat aq penikmat kopi ringan dan yg ga susu banget	[dari, sekian, banyak, coffee, shop, aku, pasti, lebih, pilih, kopi, kenangan, pertama, kali, cobain, kopi, kenangan, mantan, langsung, jatuh, cinta, karena, beneran, berasa, kopinya, dan, tidak, cuma, rasa, susu, yang, dominan, karena, aku, kurang, suka, susu, jadi, cocok, banget, buat, aku, penikmat, kopi, ringan, dan, yang, tidak, susu, banget]

4.8. Stopwords

Tujuan dari proses ini untuk menghapus kata-kata yang tidak memengaruhi makna kalimat, seperti "ada", "atas", "belum", "dia", "itu", "siapa", dll.

Tabel 6. Hasil Proses *Stopwords*

Sebelum <i>Stopwords</i>	Sesudah <i>Stopwords</i>
[dari, sekian, banyak, coffee, shop, aku, pasti, lebih, pilih, kopi, kenangan, pertama, kali, cobain, kopi, kenangan, mantan, langsung, jatuh, cinta,	[sekian, coffee, shop, pilih, kopi, kenangan, kali, cobain, kopi, kenangan, mantan, langsung, jatuh, cinta,

Sebelum Stopwords	Sesudah Stopwords
cinta, karena, beneran, berasa, kopinya, dan, tidak, cuma, rasa, susu, yang, dominan, karena, aku, kurang, suka, susu, jadi, cocok, banget, buat, aku, penikmat, kopi, ringan, dan, yang, tidak, susu, banget]	beneran, berasa, kopinya, susu, dominan, suka, susu, cocok, banget, penikmat, kopi, ringan, susu, banget]

4.9. Stem

Tujuan dari proses ini untuk mengembalikan kata-kata ke bentuk asalnya. Ini berarti menghapus semua Imbuhan yang terdapat pada kata itu.

Tabel 7. Hasil Proses Stem

Sebelum Stem	Sesudah Stem
[sekian, coffee, shop, pilih, kopi, kenangan, kali, cobain, kopi, kenangan, mantan, langsung, jatuh, cinta, beneran, berasa, kopinya, susu, dominan, suka, susu, cocok, banget, penikmat, kopi, ringan, susu, banget]	[sekian, coffee, shop, pilih, kopi, kenangan, kali, cobain, kopi, kenangan, mantan, langsung, jatuh, cinta, benar, berasa, kopinya, susu, dominan, suka, susu, cocok, banget, penikmat, kopi, ringan, susu, banget]

4.10. Filter Tokens by Length

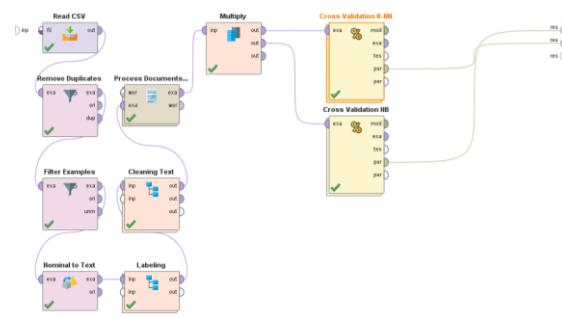
tahap *filter token by length* adalah untuk membatasi jumlah karakter yang termasuk dalam setiap kata yang panjangnya kurang dari 4 atau lebih dari 25 karakter akan dibuang karena tidak memiliki arti.

Tabel 8. Hasil Proses Filter Token by Length

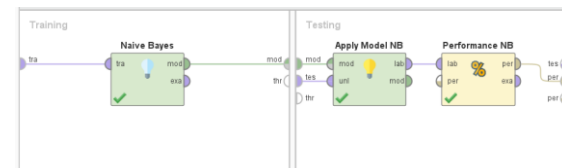
Sebelum Filter Token by Length	Sesudah Filter Token by Length
[sekian, coffee, shop, pilih, kopi, kenangan, kali, cobain, kopi, kenangan, mantan, langsung, jatuh, cinta, benar, berasa, kopinya, susu, dominan, suka, susu, cocok, banget, penikmat, kopi, ringan, susu, banget]	[sekian, coffee, shop, pilih, kopi, kenangan, kali, cobain, kopi, kenangan, mantan, langsung, jatuh, cinta, benar, berasa, kopinya, susu, dominan, suka, susu, cocok, banget, penikmat, kopi, ringan, susu, banget]

4.11. Modeling

Dalam tahapan modeling, algoritma klasifikasi text mining *Naive Bayes* dan *K-Nearest Neighbor* akan diterapkan. Metode *cross validation* akan digunakan pada fase pengujian, dan *confusion matrix* akan diterapkan untuk mengevaluasi kinerja model secara lebih tepat dalam mengklasifikasikan data ulasan aplikasi Kopi Kenangan ke dalam positif atau negatif berdasarkan sentimen yang terkandung ulasan tersebut.

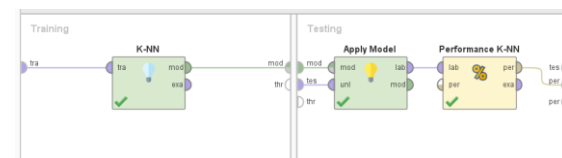


Gambar 7. Proses Modeling Algoritma Naive Bayes dan K-Nearest Neighbor



Gambar 8. Proses Cross Validation Naive Bayes

Proses *Cross validation* data dilakukan dengan memanfaatkan algoritma klasifikasi *Naive Bayes*, yang meliputi proses *training* dan *testing* dengan penerapan *apply model* serta *performance*. Kemudian, algoritma klasifikasi *Naive Bayes* digunakan untuk memodelkan untuk memperoleh nilai *accuracy*, *precision*, dan *recall* yang didasarkan pada *performance vector*.



Gambar 9. Proses Cross Validation K-Nearest Neighbor

Proses *Cross validation* data dilakukan dengan menggunakan algoritma klasifikasi *K-Nearest Neighbor*, yang meliputi proses *training* dan *testing* dengan penerapan *apply model* serta *performance*. Setelahnya, pemodelan algoritma klasifikasi *K-Nearest Neighbor* dilaksanakan untuk mendapatkan nilai *accuracy*, *precision*, dan *recall* yang didasarkan pada *performance vector*.

4.12. Evaluasi

Proses evaluasi dilaksanakan untuk menilai hasil model yang telah dibuat melalui uji *cross validation* dan menghasilkan *confusion matrix*. Proses evaluasi dilakukan untuk menilai kinerja model melalui *10-fold cross-validation* yang menghasilkan *confusion matrix* pada setiap proses pengujian. Pada penelitian ini digunakan stratified sampling agar proporsi kelas positif dan negatif tetap terjaga pada setiap *fold*. Proses pembagian data dilakukan secara acak dengan local random seed sebesar 1992, sehingga pembentukan *fold* dapat direproduksi secara konsisten. Penggunaan *10-fold cross-validation* bertujuan untuk memperoleh

estimasi performa model yang lebih stabil dibandingkan hanya menggunakan satu pembagian data latih dan data uji.

Hasil yang dilaporkan dalam penelitian ini merupakan hasil agregasi evaluasi *cross-validation*. Nilai rata-rata dan standar deviasi performa pada masing-masing *fold* belum dilaporkan secara terpisah, sehingga hal tersebut menjadi keterbatasan dalam penyajian stabilitas performa model.

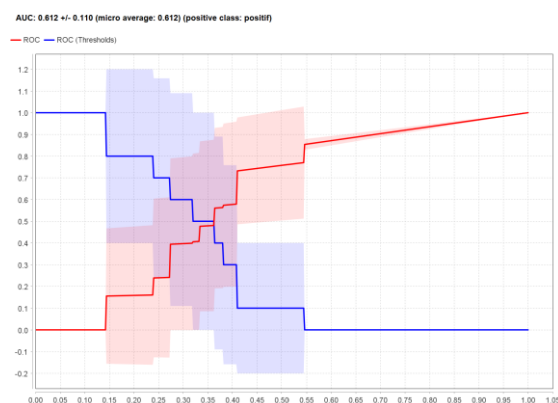
4.13. Evaluasi Naïve Bayes

accuracy: 69.68% +/- 3.63% (micro average: 69.68%)

	true negatif	true positif	class precision
pred. negatif	148	234	38.74%
pred. positif	68	546	88.93%
class recall	68.52%	70.00%	

Gambar 10. Hasil Akurasi dan *Confusion Matrix Naive Bayes*

Gambar 10 menunjukkan hasil evaluasi algoritma *Naive Bayes* menggunakan *cross-validation* pada 996 data ulasan aplikasi Kopi Kenangan. Model memperoleh akurasi sebesar 69,68%. Pada kelas positif, nilai *precision* sebesar 88,93% dan *recall* sebesar 70,00%, sehingga diperoleh *F1-score* sebesar 78,34%. Pada kelas negatif, nilai *precision* sebesar 38,74% dan *recall* sebesar 68,52%, sehingga diperoleh *F1-score* sebesar 49,50%. Berdasarkan kedua kelas tersebut, nilai *macro-averaged F1-score* mencapai 63,92%. Hasil ini menunjukkan adanya perbedaan performa yang cukup besar antara kelas positif dan negatif, sehingga penilaian kinerja model tidak hanya didasarkan pada akurasi keseluruhan, tetapi juga mempertimbangkan kemampuan model dalam mengklasifikasikan masing-masing kelas.



Gambar 11. Hasil Kurva AUC ROC *Naive Bayes*

Kurva ROC pada Gambar 11 menunjukkan hubungan antara *True Positive Rate* (TPR) dan *False Positive Rate* (FPR) pada berbagai nilai *threshold*. Nilai AUC sebesar 0,612 menunjukkan kemampuan diskriminasi model *Naive Bayes* yang masih terbatas. Area bayangan pada visualisasi tidak diinterpretasikan sebagai bukti ketidakstabilan antar-*fold*, karena penelitian ini tidak melaporkan ukuran variasi kuantitatif seperti rata-rata dan

standar deviasi AUC pada setiap *fold*. Evaluasi model karena itu didasarkan pada AUC agregat serta metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

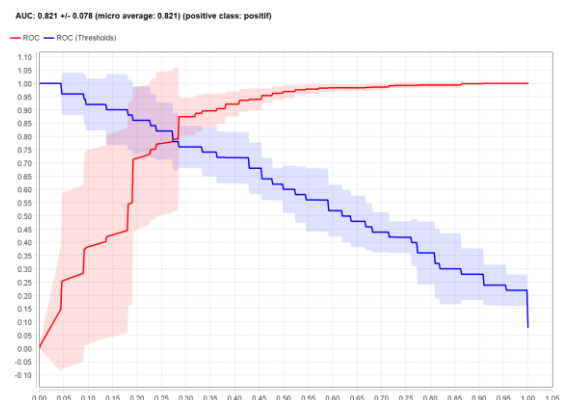
4.14. Evaluasi K-Nearest Neighbor

accuracy: 85.63% +/- 3.09% (micro average: 85.64%)

	true negatif	true positif	class precision
pred. negatif	86	13	86.87%
pred. positif	130	767	85.51%
class recall	39.81%	98.33%	

Gambar 12. Hasil Akurasi dan *Confusion Matrix K-Nearest Neighbor*

Hasil evaluasi algoritma *K-Nearest Neighbor* (KNN) pada Gambar 12 diperoleh melalui proses *cross-validation* terhadap 996 data ulasan aplikasi Kopi Kenangan. Berdasarkan hasil pengujian, model menghasilkan akurasi sebesar 85,63%. Untuk kelas positif, model memperoleh *precision* sebesar 85,51% dan *recall* sebesar 98,33%, dengan *F1-score* sebesar 91,75%. Sementara itu, pada kelas negatif diperoleh *precision* sebesar 86,87% dan *recall* sebesar 39,81%, sehingga menghasilkan *F1-score* sebesar 54,62%. Nilai *macro-averaged F1-score* sebesar 73,19% menunjukkan rata-rata kinerja model pada kedua kelas. Hasil tersebut memperlihatkan bahwa KNN memiliki kemampuan yang lebih tinggi dalam mengenali ulasan positif, sedangkan kemampuan dalam mendeteksi ulasan negatif masih lebih rendah berdasarkan nilai *recall* yang diperoleh



Gambar 13. Hasil Kurva AUC ROC *K-Nearest Neighbor*

Gambar di atas menunjukkan kurva ROC hasil klasifikasi sentimen menggunakan algoritma *K-Nearest Neighbor* (KNN), dengan garis ROC merepresentasikan hubungan antara *True Positive Rate* (TPR) dan *False Positive Rate* (FPR) pada berbagai nilai *threshold*. Model memperoleh nilai AUC sebesar 0,821 yang menunjukkan kemampuan diskriminasi yang baik secara keseluruhan. Meskipun demikian, nilai ROC-AUC tersebut perlu diinterpretasikan bersama dengan performa pada masing-masing kelas. KNN memperoleh *precision* sebesar 85,51% dan *recall* sebesar 98,33% pada kelas positif, sedangkan pada kelas negatif diperoleh *precision* sebesar 86,87% dan *recall* sebesar

39,81%. Rendahnya *recall* kelas negatif menunjukkan bahwa model masih mengalami keterbatasan dalam mengenali sebagian ulasan negatif. Dengan demikian, nilai AUC yang relatif tinggi tidak menunjukkan bahwa kemampuan klasifikasi KNN sama kuat pada kedua kelas, tetapi menggambarkan kemampuan diskriminasi model secara keseluruhan yang tetap perlu dibaca bersama metrik kelas spesifik.

Kurva yang naik cepat dari kiri bawah ke kanan atas mengindikasikan tingkat deteksi positif yang tinggi meskipun pada nilai FPR yang rendah, sementara area bayangan merah yang lebih kecil dibandingkan *Naive Bayes* mencerminkan hasil prediksi yang lebih stabil dan konsisten antar-fold pengujian. Adapun garis biru yang merepresentasikan perubahan nilai *threshold* tidak dijadikan tolak ukur utama dalam penilaian performa model.

5. KESIMPULAN DAN SARAN

Temuan penelitian menunjukkan bahwa algoritma *Naive Bayes* memperoleh akurasi sebesar 69,68%, dengan *recall* kelas negatif sebesar 68,52% dan kelas positif sebesar 70,00%, serta *precision* kelas negatif sebesar 38,74% dan kelas positif sebesar 88,93%. Nilai AUC sebesar 0,612 menunjukkan bahwa kemampuan diskriminasi model masih terbatas. Sementara itu, algoritma *K-Nearest Neighbor* (KNN) memperoleh akurasi sebesar 85,63%, dengan *recall* kelas negatif sebesar 39,81% dan kelas positif sebesar 98,33%, serta *precision* kelas negatif sebesar 86,87% dan kelas positif sebesar 85,51%. Nilai AUC KNN sebesar 0,821 menunjukkan kemampuan diskriminasi yang lebih baik secara keseluruhan. Berdasarkan hasil tersebut, KNN menunjukkan kinerja yang lebih tinggi pada akurasi keseluruhan, ROC-AUC, *precision* kelas negatif, dan *recall* kelas positif. Sebaliknya, *Naive Bayes* menunjukkan kinerja yang lebih tinggi pada *precision* kelas positif dan *recall* kelas negatif. Dengan demikian, hasil penelitian tidak menunjukkan keunggulan satu algoritma pada seluruh dimensi performa, melainkan memperlihatkan adanya perbedaan karakteristik kinerja kedua algoritma pada masing-masing metrik evaluasi. Untuk penelitian selanjutnya, disarankan memperluas jumlah dan variasi dataset, mempertimbangkan ketidakseimbangan kelas, serta menggunakan metrik tambahan seperti *macro F1-score* dan *balanced accuracy*. Penggunaan algoritma lain juga dapat dipertimbangkan sebagai pembandingan untuk memperoleh gambaran performa klasifikasi yang lebih komprehensif.

DAFTAR PUSTAKA

[1] A. S. Wahyusesa, "PEMBUATAN APLIKASI PEMESANAN PADA COFFEE SEDUH SPACE CIKARANG BERBASIS ANDROID MENGGUNAKAN METODE

WATERFALL," vol. 13, no. 3, 2023, doi: 10.23960/jitet.v13i3S1.7601.

- [2] D. Afryla and Asnawati, "Pengaruh Price Discount dan Usability serta Customer Experience Terhadap Keputusan Pembelian di Aplikasi Kopi Kenangan," vol. 1, no. 4, pp. 2265–2277, 2025, doi: 10.63822/t7y4yt82.
- [3] S. Humairah and S. A. Maldin, "Pengaruh Kemudahan Penggunaan Aplikasi Terhadap Minat Beli Pelanggan Di Kopi Kenangan Kota Batam," vol. 5, no. 3, pp. 36390–36400, 2025.
- [4] M. A. Fadli, E. Novalia, and A. L. Hananto, "KLASIFIKASI ULASAN APLIKASI KOPI KENANGAN PADA GOOGLE PLAYSTORE MENGGUNAKAN ALGORITMA NAIVE BAYES," vol. 6, no. 2, pp. 593–605, 2025, doi: 10.46576/djtechno.v6i2.7037.
- [5] F. M. D. Maharani, A. L. Hananto, S. S. Hilabi, F. N. Apriani, A. Hananto, and B. Huda, "Perbandingan Metode Klasifikasi Sentimen Analisis Penggunaan E-Wallet Menggunakan Algoritma Naïve Bayes dan K-Nearest Neighbor," vol. 6, no. 2, 2022, doi: 10.47002/metik.v6i2.372.
- [6] H. Satrio and E. K. Pratama, "Perbandingan Algoritma Naive Bayes Dan K-Nearest Neighbor Terhadap Analisis Sentimen Aplikasi Rosalia Indah Transport," vol. 13, no. 1, pp. 7–18, 2025, doi: 10.33592/jutis.v13i1.5327.
- [7] L. Nursinggah, Ruuhwan, and T. Mufizar, "ANALISIS SENTIMEN PENGGUNA APLIKASI X TERHADAP PROGRAM MAKAN SIANG GRATIS," vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4336 ANALISIS.
- [8] A. N. Indraini and I. Ernawati, "Analisis Sentimen Terhadap Pembelajaran Daring Di Indonesia Menggunakan Support Vector Machine (SVM)," vol. XIV, no. 1, pp. 68–80, 2022, doi: 10.22441/fifo.2022.v14i1.007.
- [9] Fathoni, A. P. Maretta, A. N. Kusuma, R. M. Sasmita, A. F. Rizkyllah, and A. Ibrahim, "PERBANDINGAN METODE NAÏVE BAYES DAN K-NEAREST NEIGHBOR TERHADAP ANALISIS SENTIMEN ULASAN PROGRAM MAKAN SIANG GRATIS DI INDONESIA," vol. 9, no. 4, 2025, doi: 10.36040/jati.v9i4.14084.
- [10] S. I. Maiyanti *et al.*, "Perbandingan Klasifikasi Penyakit Kanker Paru-paru menggunakan Support Vector Machine dan K-Nearest Neighbor," vol. 18, no. 1, pp. 54–62, 2023, doi: 10.33998/processor.2023.18.1.700.
- [11] S. Rahmah and A. B. Fawait, "IMPLEMENTASI DATA MINING DENGAN ALGORITMA REGRESI LINEAR SEDERHANA UNTUK MEMPREDIKSI NILAI EKSPOR DI KALIMANTAN TIMUR

- DENGAN APLIKASI RAPIDMINER,” vol. 5, no. 1, pp. 12–21, 2024, doi: 10.21093/dijitac.v5i1.8952.
- [12] R. Maulana, P. Hertaryawan, M. Raihan, and I. Santoso, “KOMPARASI ALGORITMA NAIVE BAYES DAN K- NEAREST NEIGHBOR PADA ANALISIS SENTIMEN,” vol. 17, no. 2, pp. 177–189, 2023, [Online]. Available: <https://pdfs.semanticscholar.org/0aeb/508ffb3d29f804375e405c72eac6ba923680.pdf>
- [13] P. K. Sari and R. R. Suryono, “KOMPARASI ALGORITMA SUPPORT VECTOR MACHINE DAN RANDOM FOREST UNTUK ANALISIS SENTIMEN METAVERSE,” vol. 7, no. 1, pp. 31–39, 2024, doi: 10.36040/mnemonic.v7i1.8977.
- [14] M. R. Firdaus, N. Rahaningsih, and R. D. Dana, “Analisis Sentimen Aplikasi Shopee di Goole Play Store Menggunakan Klasifikasi Algoritma Naïve Bayes,” vol. 6, no. 1, 2024, doi: 10.36499/jinrpl.v6i1.10302.