

## Penerapan *Cross Validation* sebagai Analisis Sentimen Pelayanan Publik Kereta Api Lokal Daop 8 Menggunakan Metode *Multinomial Naïve Bayes*

Risnaldy Novendra Irawan<sup>1✉</sup>, Kartika Maulida Hindrayani<sup>2</sup>, Mohammad Idhom<sup>3</sup>

<sup>1,2,3</sup> Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur, Indonesia

### Informasi Artikel

#### Riwayat Artikel

**Diserahkan** : 14-03-2024

**Direvisi** : 16-03-2024

**Diterima** : 21-03-2024

#### Kata Kunci:

Analisis Sentimen;  
Multinomial Naïve Bayes;  
K-Fold Cross Validation;  
Kereta Api Subsidi.

#### Keywords :

Sentiment Analysis;  
Multinomial Naïve bayes; K-  
Fold Cross Validation;  
Subsidized Railway  
Transportation.

#### Corresponding Author :

Risnaldy Novendra Irawan  
Sains Data, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur  
Jl. Rungkut Madya No.1, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur 60294  
Email: [risnaldy19@gmail.com](mailto:risnaldy19@gmail.com)

### ABSTRAK

Dalam menyediakan layanan yang efisien dan berkualitas untuk Masyarakat pada sektor transportasi kereta api wilayah daerah operasional 8 Surabaya, perlu dilakukan langkah untuk memenuhi harapan pengguna kereta api subsidi. Penelitian ini bertujuan untuk mengetahui pengaruh *cross validation* terhadap metode *Multinomial Naïve Bayes* dalam analisis sentimen menggunakan metode *10-fold cross validation*. Langkah-langkah *preprocessing* data dilakukan sehingga didapatkan data sebanyak 1123 komentar dengan dua kelas yaitu kelas positif sebanyak 778 komentar dan negatif sebanyak 345 komentar. Analisis dilakukan sebanyak 2 kali dengan proses validasi *10-fold cross validation* dan pengujian *Multinomial Naïve Bayes*. Berdasarkan hasil pengujian *Multinomial Naïve Bayes* menggunakan data uji sebanyak 225 data, didapatkan 14 data positif dan 160 data negatif yang ditinjau dari performa pengujian terbaik pada parameter nilai validasi fold = 1. Hasil akhir didapatkan nilai akurasi 77%, presisi 81%, recall 77%, dan f1 score 71%, yang mengungkapkan bahwa model efektif dalam mengklasifikasi komentar negatif dari keseluruhan data uji.

### ABSTRACT

To provide efficient and quality services for community in the railway transportation sector of operational region 8 Surabaya, steps need to be taken to meet the expectations of subsidized train users.. This research aims to determine the influence of cross-validation on the Multinomial Naïve Bayes method in sentiment analysis using 10-fold cross-validation method. The data preprocessing steps were carried out to obtain a total of 1123 comments with two classes: 778 positive comments and 345 negative comments. The analysis was conducted twice with 10-fold cross-validation and Multinomial Naïve Bayes testing processes. Based on the Multinomial Naïve Bayes testing results using 225 test data, 14 positive data and 160 negative data were obtained. The best testing performance was observed the parameter validation fold value of 1. Final result yielded accuracy of 77%, precision of 81%, recall of 77%, and f1 score of 71% that reveals model effective in classifying negative comments.

## PENDAHULUAN

Pemenuhan fasilitas yang memadai di lingkungan Masyarakat merupakan peran pemerintah dalam memberikan kemudahan transportasi yang aman dan nyaman sehingga memudahkan dalam mobilisasi penduduk. Salah satu bentuk layanan yang memiliki peranan vital dalam mendukung mobilitas masyarakat adalah transportasi kereta api bersubsidi. Menurut peraturan menteri perhubungan (RI, 2010) tentang penyelenggaraan kewajiban pelayanan publik angkutan orang dengan kereta api kelas ekonomi tahun anggaran 2010, Program ini merupakan bagian dari *Public Service Obligation* (PSO) yang ditetapkan pemerintah untuk meningkatkan kualitas layanan dengan membayar selisih tarif antara penetapan pemerintah dan tarif yang dihitung oleh PT KAI. Hal ini seiring dengan pemberlakuan grafik perjalanan kereta api terbaru yang mulai dilaksanakan pada 1 Juni 2023 menjadikan banyak problematika di sarana transportasi kereta api khususnya kereta api lokal Daop 8 Surabaya.

Dalam menyelesaikan permasalahan ini, perlu dilakukan sebuah analisis sentimen pengguna terhadap layanan kereta api subsidi dalam aturan terbaru yang sudah ditetapkan. Analisis sentimen menjadi relevan karena dapat memberikan wawasan yang berharga mengenai bagaimana pengguna sebenarnya merasakan dan menanggapi kualitas layanan yang diberikan (Muhaimin et al., 2022). Penelitian ini akan fokus pada evaluasi kualitas layanan Transportasi Kereta Api Subsidi melalui pengumpulan data yang melibatkan tanggapan dan pandangan dari pengguna. Faktor-faktor kritis seperti kebersihan fasilitas, ketepatan waktu kedatangan dan keberangkatan, kenyamanan selama perjalanan, serta interaksi dengan layanan pelanggan akan menjadi pokok analisis. Pengumpulan data akan mencakup aspek-aspek positif dan negatif dalam merespons pengalaman pengguna.

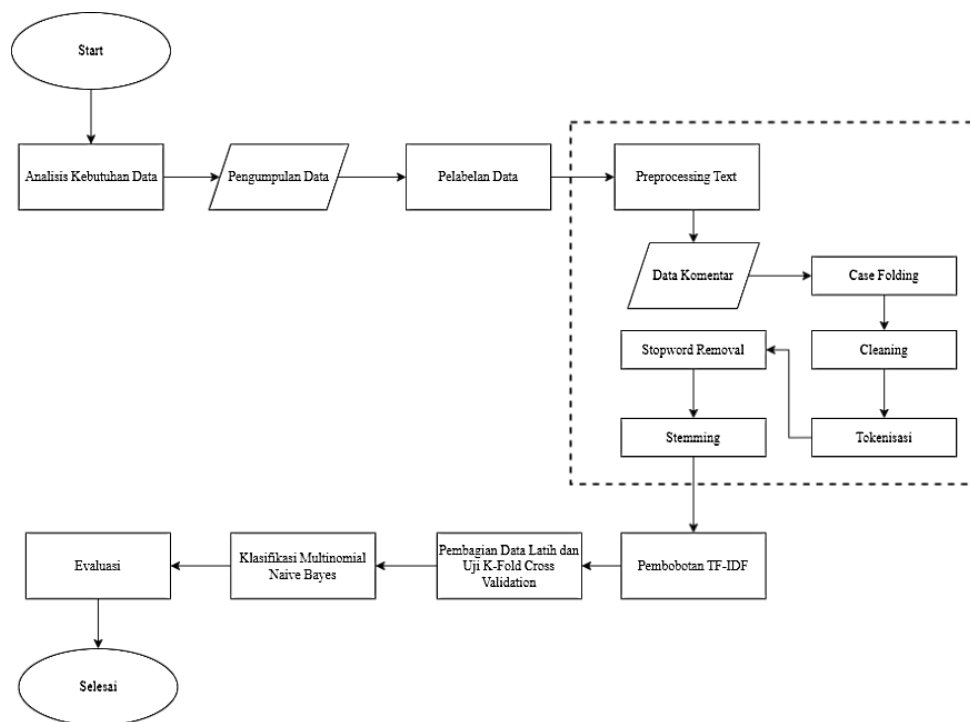
Analisis sentimen adalah suatu disiplin ilmu yang berfokus pada teknik analisis pendapat, pandangan, dan penilaian individu terhadap suatu subjek (Husada & Paramita, 2021). Oleh karena itu, diperlukan suatu pendekatan yang efisien dan metode yang terstruktur untuk mengelola data tersebut untuk dapat membedakan antara ulasan positif, negatif, dan netral (Estika et al., 2021). Dalam proses analisis sentimen, teknik Pemrosesan Bahasa Alami (NLP) digunakan untuk mengidentifikasi dan mengukur emosi dalam teks yang dilakukan oleh mesin untuk memprediksi hasil analisis (Muktafin et al., 2020). Analisis sentimen banyak diterapkan dalam mengolah teks atau komentar, seperti dalam evaluasi moda transportasi di Indonesia (Dananjaya & Irawan, 2023). Ulasan dari pelanggan menjadi indikator utama dalam menilai kualitas suatu layanan umum, terutama dalam konteks moda transportasi seperti kereta api. Ulasan juga dapat memberikan gambaran tentang pengalaman pengguna sebelumnya dengan moda transportasi tersebut. Studi yang dilakukan oleh (Chen, 2020) menunjukkan bahwa ulasan online berpengaruh signifikan terhadap persepsi dan keputusan konsumen dalam pembelian produk. Salah satu metode yang digunakan dalam analisis sentimen adalah klasifikasi Naïve Bayes, dengan salah satu tekniknya yaitu Multinomial Naïve Bayes.

Proses yang dilakukan dalam analisis sentimen untuk mendapatkan hasil terbaik memiliki langkah dan metode yang bervariasi, salah satu metode optimasi yang dapat digunakan untuk menunjang algoritma multinomial naïve bayes dalam studi kasus analisis sentimen yaitu *cross validation* (Oktaviani Putri & Cahya Wihandika, 2020). *Cross validation* atau biasa disebut *K-fold cross validation* merupakan sebuah teknik validasi model untuk menilai hasil analisis statistik yang digunakan untuk melakukan prediksi model dan memperkirakan seberapa akurat sebuah model prediktif ketika dijalankan dalam praktiknya sehingga mengurangi bias pada suatu data (Tempola et al., 2018). Pada penelitian yang dilakukan oleh (Zhafira et al., 2021) tentang analisis sentimen menggunakan naïve bayes dengan penerapan pembobotan TF-IDF, dan *10-fold cross validation* terhadap kebijakan kampus merdeka berdasarkan komentar pada Youtube menghasilkan akurasi lebih tinggi yakni sebesar 97% dibandingkan naïve bayes dengan pembobotan TF-IDF tanpa menggunakan optimasi *cross validation*. Yang berarti *cross validation* berpengaruh terhadap hasil akurasi yang didapatkan

Penelitian ini bertujuan untuk menerapkan *cross validation* untuk optimasi dari metode *Multinomial Naïve Bayes* dengan membagi datanya ke setiap fold. Metode ini dilakukan karena efektif dalam mengelola data teks dan mampu mempertimbangkan frekuensi kemunculan kata. Tahapan dilakukan dengan analisa kebutuhan data, persiapan data, analisis sentimen, evaluasi sistem, dan terakhir skenario pengujian untuk pembagian data latih dan data uji. Untuk dapat memahami hasil digunakan evaluasi confusion matrix yang menampilkan nilai akurasi, presisi, recal, dan f1 score.

## METODE PENELITIAN

Pada penelitian ini terdapat tahapan – tahapan penelitian yang dilakukan, tahapan penelitian tersebut digambarkan dalam sebuah diagram alir penelitian yang berguna untuk menggambarkan secara jelas penerapan *cross validation* terhadap *Multinomial Naïve Bayes* dalam melakukan analisis sentimen untuk topik yang dibahas, gambar 1 berikut merupakan tahapan penelitian yang dilakukan:



**Gambar 1. Alur Penelitian Proses Analisis Sentimen Menggunakan Cross Validation dan Multinomial Naïve Bayes**

### Analisis Kebutuhan Data

Analisis kebutuhan data merupakan tahap penentuan dalam pengambilan data. Hal ini menjadi dasar untuk mengkaji layanan kereta api subsidi Daop 8 Surabaya. Pada penelitian yang dilakukan, penulis menggali informasi dari PPID Daop 8 Surabaya terkait komentar pengguna yang menggunakan moda transportasi kereta api lokal subsidi.

### Pengumpulan Data

Pengumpulan data merupakan tahapan yang perlu dilakukan dalam melakukan analisis sentimen sebagai langkah awal melakukan penelitian dengan mengetahui struktur data dan juga model data (Jon, 2022). Pengumpulan data sentimen dapat dilakukan dengan dua metode, yakni metode primer dan sekunder. Pengumpulan data metode primer yakni dilakukan secara langsung dengan melakukan pengumpulan informasi menggunakan kuesioner, survei, wawancara, atau observasi, sedangkan pengumpulan data metode sekunder adalah mendapatkan data dari pihak yang memiliki kewenangan atau dari sumber yang dapat diakui kebenaran sumber data nya. Pada

penelitian ini, data yang dikumpulkan adalah komentar atau ulasan dari pelanggan yang menggunakan moda transportasi kereta api subsidi yang berisikan tentang nama pelanggan, nilai kepuasan, dan juga ulasan terhadap pelayanan yang diberikan oleh PT. KAI dalam bentuk csv.

## Pelabelan Data

Pelabelan data merupakan proses untuk menentukan label atau kategori tertentu pada data teks atau konten untuk mengindikasikan sentimen atau perasaan yang terkandung dalam suatu teks dengan tujuan untuk pembelajaran pada model klasifikasi yang akan dibangun (Telnoni, 2021). Pelabelan dilakukan menggunakan library *Lexicon* yang dilakukan secara otomatis oleh python. Sistem kerja *Lexicon* yakni menganalisis tiap kata sehingga menghasilkan *compound* dengan luaran yang berupa sentimen positif, negatif, atau netral. Detail nilai dalam tiap kategori data yang digunakan untuk analisis sentimen disajikan dalam tabel 1

**Tabel 1. Penjelasan Nilai Compound terhadap Sentimen**

| <i>Nilai Compound</i> | <i>Kategori</i> |
|-----------------------|-----------------|
| <0                    | <i>Negatif</i>  |
| 0                     | <i>Netral</i>   |
| >0                    | <i>Positif</i>  |

## Preprocessing Text

*Preprocessing text* bertujuan untuk pembersihan, persiapan dan pengorganisasian data mentah dalam memastikan data yang dianalisis telah tepat, relevan, dan siap untuk digunakan. (Febrisa Sidabutar et al., 2023). Tahapan dalam *preprocessing* diantaranya adalah case folding, untuk menjadikan huruf kecil untuk konsistensi, selanjutnya adalah cleansing untuk menghapus karakter, simbol, atau noise yang dianggap tidak penting, lalu dilakukan tokenisasi untuk memisahkan karakter dalam sebuah kalimat menjadi unit kata, proses selanjutnya stopword removal untuk menghapus kata yang kurang memberikan informasi dengan bantuan library Sastrawi, dan proses terakhir stemming di mana kata-kata yang memiliki imbuhan diubah menjadi kata dasar yang dipilih setelah melakukan stopword removal.

## Pembobotan TF-IDF

Pembobotan kata atau *word weighting* merupakan metode yang dilakukan untuk memberikan bobot pada setiap kata dalam teks sehingga dapat meningkatkan kemampuan analisis sentimen (Hanif Razka et al., 2023). Metode yang digunakan pada pembobotan adalah *Term Frequency Inverse Document Frequency* (TF-IDF). Dalam melakukan perhitungan bobot kata, perlu diketahui terlebih dahulu nilai  $tf_{t,d}$  sebagai bobot dari term  $t$  dan nilai  $idf$  yang merupakan banyak dokumen yang berisi istilah yang dicari. Apabila kedua nilai telah ditemukan maka untuk melakukan perhitungan adalah dengan cara mengalikan kedua nilai tersebut untuk mendapatkan bobot ( $\omega$ ). Penjabaran rumus adalah sebagai berikut:

$$tf - idf_{t,d} = tf_{t,d} * idf_t$$

## Pembagian Data Latih dan Uji K-Fold Cross Validation

*K-fold cross validation* merupakan teknik *resampling* untuk menyesuaikan parameter data saat membangun model. (Oktaviani Putri & Cahya Wihandika, 2020). ‘fold’ pada *k-fold cross validation* yang berarti lipatan merupakan bentuk dari jumlah himpunan pelatihan yang akan dihasilkan. Model-model dilatih menggunakan subset pembelajaran dan divalidasi menggunakan subset validasi (Farisi et al., 2019). Proses ini diterapkan pada dataset yang telah melalui tahap *preprocessing* dengan membagi dalam dua subset, yakni dataset pelatihan dan dataset pengujian. Pembagian ini menggunakan metode alokasi 20% dari total dataset sebagai dataset pengujian, sementara 80% sisanya digunakan sebagai dataset pelatihan. Proses berikutnya adalah pembentukan model pada 80% data pelatihan, di mana fokus utama adalah membuat model pembagian data eksperimental. Eksperimen akan diulang sebanyak  $k$  kali, dengan bergantian antara data pelatihan dan data validasi pada setiap iterasi. Pemilihan nilai  $k$  dapat berdasarkan kebutuhan riset. Secara umum, penelitian menggunakan  $k = 10$  untuk mendapatkan akurasi

optimal. Nilai akurasi dari Cross Validation dapat dilakukan menggunakan persamaan sebagai berikut (Sari, 2019):

$$\text{Akurasi} = \frac{\text{jumlah klasifikasi benar}}{\text{jumlah data uji}} \times 100\%$$

Persamaan diatas merupakan nilai dari jumlah klasifikasi benar yang dibuat oleh model selama proses cross-validation, kemudian jumlah data uji merupakan total jumlah data yang diuji selama seluruh proses cross-validation atau banyaknya K, dan untuk mencari rasio, digunakan jumlah klasifikasi benar dengan 100%

### Klasifikasi Multinomial Naïve Bayes

*Multinomial Naïve Bayes* merupakan algoritma berbasis *supervised learning* yang menggunakan probabilitas teorema *Bayes*, yang mana fokus klasifikasi terdapat pada suatu teks (Farisi et al., 2019). *Multinomial Naïve Bayes* menggunakan pendekatan *term frequency* untuk menghitung seberapa sering kata tertentu muncul dalam sebuah dokumen. Model ini mempertimbangkan dua aspek, yaitu keberadaan kata dalam dokumen dan frekuensi kemunculannya dalam dokumen. *Multinomial Naïve Bayes* merupakan model yang cocok dalam pengklasifikasian teks atau dokumen yang diimplementasikan dengan persamaan berikut:

$$V_{MAP} = \underset{v_j \in V}{\operatorname{argmax}} P(v_j) \prod_i P(a_i | v_j)$$

### Evaluasi

Evaluasi dilakukan untuk menilai kinerja analisis yang telah dilakukan dalam proses validasi dan pengujian menggunakan metode Multinomial Naïve Bayes dengan dua pendekatan, yaitu *convusion matrix* dan klasifikasi hasil pengujian secara manual untuk mengkategorikan nilai akurasi berdasarkan kelas. Dalam *convusion matrix*, parameter yang umum digunakan adalah Akurasi, Presisi, Recall, dan F1-Score. Dalam penelitian ini, terdapat dua kelas yang dianalisis dalam *convusion matrix*, yaitu kelas positif dan negatif. Detail kelas dari *convusion matrix* disajikan dalam Tabel 2.

**Tabel 2. Convusion Matrix Evaluasi Output Data**

|                        | <i>Predicted Positive</i> | <i>Predicted Negative</i> |
|------------------------|---------------------------|---------------------------|
| <i>Actual Negative</i> | <i>True Positive</i>      | <i>False Negative</i>     |
| <i>Actual Positive</i> | <i>False Positive</i>     | <i>True Negative</i>      |

Evaluasi kedua dilakukan untuk melakukan peninjauan terhadap seluruh pengujian yang sudah dilakukan. Pengklasifikasian hasil uji yang dilakukan pada penelitian ini dibagi kedalam tiga kategori, yakni Akurat, Kurang Akurat, dan Tidak Akurat. Tahapan dalam melakukan klasifikasi adalah dengan mengetahui nilai akurasi tertinggi dari proses validasi, dan juga dilakukan peninjauan untuk nilai performansi yang terdiri dari presisi, recall, dan F1 Score. Dari hasil yang diketahui kemudian dievaluasi secara manual untuk dapat mengetahui akurasi terbaik dari salah satu fold. Dan kemudian diproses pada langkah pengujian untuk dapat menentukan hasil. Tabel klasifikasi hasil uji disajikan pada tabel 3

**Tabel 3. Klasifikasi Hasil Uji Analisis Model**

| <i>Range Nilai Akurasi</i> | <i>Kelompok</i>      |
|----------------------------|----------------------|
| 81-100%                    | <i>Akurat</i>        |
| 70-80%                     | <i>Kurang Akurat</i> |
| <70%                       | <i>Tidak Akurat</i>  |



## HASIL DAN PEMBAHASAN

### Pengumpulan Data

Data yang digunakan dalam penelitian terdiri dari sentimen atau ulasan yang diberikan oleh pengguna mengenai layanan kereta api subsidi yang dioperasikan oleh PT KAI Daop 8 Surabaya. Data diperoleh secara sekunder dari Bagian Humas PPID Daop 8 Surabaya dengan rentang waktu bulan Juli sampai September 2023, dan seluruh komentar disampaikan dalam bahasa Indonesia. Data yang didapatkan sebanyak 1125 baris yang mencakup nama pengguna, penilaian yang diberikan, serta komentar yang disertakan mengenai pengalaman pengguna dengan layanan yang diberikan. Semua data disimpan dalam format file csv untuk mempermudah tahap *preprocessing*.

### Preprocessing Data

Proses *preprocessing* untuk memperbaiki struktur data dilakukan sebanyak 5 tahap. Dengan urutan yaitu case folding, cleaning, tokenisasi, stopword removal, dan stemming (Fahrudin et al., 2022). Tujuan utama dari proses ini adalah mempersiapkan data untuk dapat dilakukan proses analisis. Kelima proses diatas dilakukan menggunakan library umum yang ada pada python, yakni nltk, pandas, dan sastrawi sebagai pemrosesan Bahasa alami dalam bahasa Indonesia. Setelah melakukan preprocessing untuk memperbaiki struktur data, dilakukan proses lanjutan yaitu pelabelan dengan menggunakan dua library pada python diantaranya yaitu *vader lexicon* dan *SentimentIntensityAnalyzer*. *SentimentIntensityAnalyzer* digunakan untuk menghitung nilai polaritas seluruh data pada kolom ulasan hingga menghasilkan nilai *compound* yang mana nilai *compound* ini digunakan untuk menentukan label dari dataset. Apabila nilai *compound* kurang dari 0 maka data tersebut akan diberikan label negatif, sebaliknya apabila nilai *compound* bernilai lebih dari 0 maka data akan diberikan label positif. Lalu, jika nilai *compound* bernilai sama dengan 0 maka data akan diberikan label netral. Tetapi, dalam penelitian ini data dengan label netral tidak akan digunakan karena batasan sentimen hanya menggunakan dua label yaitu positif dan negatif sehingga data dengan label netral akan di-drop dari tabel. Untuk melakukan *vader lexicon*, diperlukan *translate* komentar kedalam Bahasa Inggris terlebih dahulu menggunakan *library translator*, yang kemudian analisis dapat dilakukan dan kemudian data dinormalisasi kedalam Bahasa Indonesia kembali untuk menyesuaikan kebutuhan analisis lanjutan. Seluruh proses telah dilakukan dan dapat ditunjukkan contoh hasil *preprocessing* data yang dilakukan terlihat seperti pada Tabel 4

Tabel 4. Hasil Preprocessing Data

| Nama             | Penilaian | Ulasan (Sebelum)                                                                      | Ulasan (Sesudah)                                            | Compound | Sentimen |
|------------------|-----------|---------------------------------------------------------------------------------------|-------------------------------------------------------------|----------|----------|
| Yudi Guntara     | 4         | Mendingan ga terlalu tegak sandarannya<br>Karena jarak tpt duduk berhadapan terlalu   | [lebih, baik, tidak tegak, sandar]                          | 0.4404   | Positif  |
| Ricard Staq      | 3.5       | dekat, maka saat penumpang penuh, maka terjadilah adu dengkul                         | [jarak, tempat, duduk, hadap, penumpang, penuh, adu, lutut] | -0.4019  | Negatif  |
| Rizki            | 4         | Kebersihan fasilitas toilet di kereta selalu terjaga dengan baik. toilet selalu sulit | [bersih, fasilitas, toilet, kereta, jaga]                   | 0.2732   | Positif  |
| Jeffery Wahyudi  | 3         | disiram, apalagi aliran airnya yang nggak begitu cepat                                | [toilet, sulit, siram, alir, air, tidak cepat]              | -0.3612  | Negatif  |
| Mohammad Farizal | 3         | pelayanan masih sama bahkan pola operasi ka tersebut lebih buruk dari sebelumnya      | [layan, pola, operasi, kereta, buruk]                       | -0.4767  | Negatif  |

Positif

Negatif



Pada Gambar 2 Hasil dari proses wordcloud terhadap data komentar menunjukkan penumpang yang merasa senang dengan layanan termasuk dalam label positif sering mengucapkan “kereta”, dan “Nyaman”, sedangkan label negatif, penumpang sering mengucapkan “tidaknyaman”, “tiket”, dan “kursi”. Hal ini diidentifikasi dari kata-kata yang paling sering muncul dalam data, yang ditandai dengan ukuran font yang lebih besar.

Tahap pembobotan TF – IDF merupakan pendefinisian fungsi dari seluruh kata yang ada pada data menjadi bentuk token atau numerik. Dalam penelitian yang dilakukan, library python yang digunakan yakni sklearn. Library ini akan mendefinisikan tiap kata yang ada menjadi bentuk numerik pada salah satu kolom yang dipilih menggunakan fungsi CountVectorizer dan TfidfVectorizer. Hasil dari pembobotan ini akan menjadi sebuah dataframe baru yang berisi representasi TF-IDF dari setiap kata yang ada. Nama-nama kolom dalam dataframe tersebut sesuai dengan kata-kata yang muncul dalam seluruh dataset. Hasil dari proses pembobotan TF-IDF dapat dilihat pada gambar 3

|     | ac       | aces | acuan | acuh | acung | ada | adem | adik | adil | adu      | ... | waspada | wc       | wifi | wilayah | wisata | wkwk | wlingi | wonokromo | ya  | zaman |
|-----|----------|------|-------|------|-------|-----|------|------|------|----------|-----|---------|----------|------|---------|--------|------|--------|-----------|-----|-------|
| 0   | 0.000000 | 0.0  | 0.0   | 0.0  | 0.0   | 0.0 | 0.0  | 0.0  | 0.0  | 0.000000 | ... | 0.0     | 0.000000 | 0.0  | 0.0     | 0.0    | 0.0  | 0.0    | 0.0       | 0.0 | 0.0   |
| 1   | 0.000000 | 0.0  | 0.0   | 0.0  | 0.0   | 0.0 | 0.0  | 0.0  | 0.0  | 0.450872 | ... | 0.0     | 0.000000 | 0.0  | 0.0     | 0.0    | 0.0  | 0.0    | 0.0       | 0.0 | 0.0   |
| 2   | 0.000000 | 0.0  | 0.0   | 0.0  | 0.0   | 0.0 | 0.0  | 0.0  | 0.0  | 0.000000 | ... | 0.0     | 0.000000 | 0.0  | 0.0     | 0.0    | 0.0  | 0.0    | 0.0       | 0.0 | 0.0   |
| 3   | 0.239981 | 0.0  | 0.0   | 0.0  | 0.0   | 0.0 | 0.0  | 0.0  | 0.0  | 0.000000 | ... | 0.0     | 0.000000 | 0.0  | 0.0     | 0.0    | 0.0  | 0.0    | 0.0       | 0.0 | 0.0   |
| 4   | 0.208395 | 0.0  | 0.0   | 0.0  | 0.0   | 0.0 | 0.0  | 0.0  | 0.0  | 0.000000 | ... | 0.0     | 0.344359 | 0.0  | 0.0     | 0.0    | 0.0  | 0.0    | 0.0       | 0.0 | 0.0   |
| ... | ...      | ...  | ...   | ...  | ...   | ... | ...  | ...  | ...  | ...      | ... | ...     | ...      | ...  | ...     | ...    | ...  | ...    | ...       | ... | ...   |
| 95  | 0.000000 | 0.0  | 0.0   | 0.0  | 0.0   | 0.0 | 0.0  | 0.0  | 0.0  | 0.000000 | ... | 0.0     | 0.000000 | 0.0  | 0.0     | 0.0    | 0.0  | 0.0    | 0.0       | 0.0 | 0.0   |
| 96  | 0.319058 | 0.0  | 0.0   | 0.0  | 0.0   | 0.0 | 0.0  | 0.0  | 0.0  | 0.000000 | ... | 0.0     | 0.000000 | 0.0  | 0.0     | 0.0    | 0.0  | 0.0    | 0.0       | 0.0 | 0.0   |
| 97  | 0.357792 | 0.0  | 0.0   | 0.0  | 0.0   | 0.0 | 0.0  | 0.0  | 0.0  | 0.000000 | ... | 0.0     | 0.000000 | 0.0  | 0.0     | 0.0    | 0.0  | 0.0    | 0.0       | 0.0 | 0.0   |
| 98  | 0.000000 | 0.0  | 0.0   | 0.0  | 0.0   | 0.0 | 0.0  | 0.0  | 0.0  | 0.000000 | ... | 0.0     | 0.000000 | 0.0  | 0.0     | 0.0    | 0.0  | 0.0    | 0.0       | 0.0 | 0.0   |
| 99  | 0.000000 | 0.0  | 0.0   | 0.0  | 0.0   | 0.0 | 0.0  | 0.0  | 0.0  | 0.000000 | ... | 0.0     | 0.000000 | 0.0  | 0.0     | 0.0    | 0.0  | 0.0    | 0.0       | 0.0 | 0.0   |

100 rows x 1299 columns

## Klasifikasi Multinomial Naïve Bayes

Klasifikasi yang dilakukan pada penelitian ini terdiri dari dua metode percobaan, metode pertama adalah validasi menggunakan 10 K-Fold Cross validation yang menggunakan 80% data uji untuk mengetahui nilai terbaik pada lipatan percobaan yang dilakukan oleh model, lalu metode

kedua adalah proses pengujian yang dilakukan dari nilai yang sudah didapatkan dari fold terbaik hasil validasi. Hasil validasi metode pertama terlihat pada Tabel 5

**Tabel 5. Tabel Validasi 10 K-Fold Cross Validation Multinomial Naïve Bayes**

| K  | Hasil Klasifikasi |    |    |    | Performansi |        |          |          |
|----|-------------------|----|----|----|-------------|--------|----------|----------|
|    | TP                | TN | FN | FP | Precision   | Recall | F1 Score | Accuracy |
| 1  | 23                | 65 | 10 | 15 | 0.61        | 0.7    | 0.65     | 0.78     |
| 2  | 23                | 63 | 9  | 18 | 0.56        | 0.72   | 0.63     | 0.76     |
| 3  | 26                | 52 | 7  | 28 | 0.48        | 0.79   | 0.6      | 0.69     |
| 4  | 26                | 64 | 7  | 15 | 0.63        | 0.79   | 0.7      | 0.8      |
| 5  | 18                | 66 | 12 | 16 | 0.53        | 0.6    | 0.56     | 0.75     |
| 6  | 18                | 64 | 15 | 15 | 0.55        | 0.55   | 0.55     | 0.73     |
| 7  | 30                | 56 | 11 | 15 | 0.67        | 0.73   | 0.7      | 0.77     |
| 8  | 31                | 54 | 11 | 16 | 0.66        | 0.74   | 0.7      | 0.76     |
| 9  | 20                | 60 | 10 | 22 | 0.48        | 0.67   | 0.56     | 0.71     |
| 10 | 28                | 54 | 10 | 20 | 0.58        | 0.74   | 0.65     | 0.73     |

Tabel diatas menunjukkan bahwa sebanyak 80% data yang digunakan untuk pelatihan sebanyak k=10 kali percobaan mendapatkan beberapa hasil klasifikasi dan nilai performansi yang berbeda. Dari validasi yang dilakukan, di dapatkan nilai optimal untuk pengujian yaitu pada fold atau percobaan ke-1. Dengan nilai performansi accuracy sebesar 78%, precision sebesar 61%, recall sebesar 70%, dan F1 Score sebesar 65% yang selanjutnya nilai optimal pada hasil klasifikasi yang ada pada percobaan ke 1, akan digunakan sebagai pengujian pada 20% dari sisa dataset. untuk pengujian dan didapatkan nilai optimal pada fold = 1 akan diterapkan sebagai uji coba pada sisa 20%. Hasil pengujian terhadap proses validasi terlihat pada tabel 6

**Tabel 6. Tabel Hasil Uji Validasi Menggunakan Multinomial Naïve Bayes**

| Fold | TP | TN  | FN | FP | Precision | Recall | F1 Score | Accuracy |
|------|----|-----|----|----|-----------|--------|----------|----------|
| 1    | 14 | 160 | 50 | 1  | 0.81      | 0.77   | 0.71     | 0.77     |

Dari tabel diatas, hasil yang didapatkan dari pengujian model Multinomial Naïve Bayes memiliki nilai prediksi positif sebesar 14 data sedangkan nilai prediksi negatif sebesar 160 data.

## Evaluasi

Evaluasi pada penelitian ini dilakukan untuk dapat melakukan peninjauan secara teoritis terkait hasil yang didapatkan setelah pengujian model. Pada Gambar 4 Berikut adalah hasil classification report keseluruhan untuk skenario pengujian yang dilakukan

| Classification Report: |           |        |          |         |      |           |          |          |          |
|------------------------|-----------|--------|----------|---------|------|-----------|----------|----------|----------|
|                        | precision | recall | f1-score | support |      |           |          |          |          |
| Negatif                | 0.93      | 0.22   | 0.35     | 64      |      |           |          |          |          |
| Positif                | 0.76      | 0.99   | 0.86     | 161     |      |           |          |          |          |
| accuracy               |           |        | 0.77     | 225     |      |           |          |          |          |
| macro avg              | 0.85      | 0.61   | 0.61     | 225     |      |           |          |          |          |
| weighted avg           | 0.81      | 0.77   | 0.72     | 225     |      |           |          |          |          |
| Hasil Pengujian:       |           |        |          |         |      |           |          |          |          |
| TP                     | TN        | FN     | FP       | Total   | Test | Precision | Recall   | F1Score  | Accuracy |
| 0                      | 14        | 160    | 50       | 1       | 225  | 0.810667  | 0.773333 | 0.718007 | 0.773333 |

**Gambar 4. Classification Report Terhadap Skenario Pengujian Model Multinomial Naïve bayes**

Hasil dari penerapan MultinomialNB pada 20% sisa dari data total diperoleh akurasi sebesar 77% sedangkan nilai Precision yang diperoleh sebesar 93% untuk sentimen negatif dan



76% untuk sentimen positif, yang artinya dari total 64 data yang diprediksi mendapatkan sentimen positif, terdapat 14 data yang benar benar terprediksi positif dan 50 data salah terprediksi. Lalu, berdasarkan 161 data yang diprediksi mendapatkan sentimen negatif, terdapat 160 data yang benar benar diprediksi negatif dan 1 data salah diprediksi. Kemudian untuk nilai Recall mendapatkan hasil sebesar 22% untuk sentimen negatif dan 99% untuk sentimen positif. Artinya, berdasarkan 15 data yang seharusnya sentimen positif, terdapat 14 data yang berhasil diketahui oleh model dan 1 data yang tidak diketahui oleh model. Lalu, berdasarkan 210 data yang seharusnya sentimen negatif, terdapat 160 data yang berhasil diketahui model dan 50 data yang tidak diketahui oleh model. Setelah didapatkan nilai Precision dan Recall, didapatkan pula nilai F1- Score yang merupakan nilai rata-rata dari precision dan recall. Hasil dari F1-Score yaitu sebesar 35% untuk sentimen negatif dan 86% untuk sentimen positif. Dan dapat disimpulkan bahwa hasil percobaan penelitian ini memiliki akurasi yang rendah atau kurang akurat, karena model dapat menganalisis sangat baik untuk kelas negatif, namun pada kelas positif masih terjadi ketidakstabilan hasil uji.

## KESIMPULAN DAN SARAN

### Kesimpulan

Penelitian tentang model Multinomial Naïve Bayes yang dilakukan pengujian model dengan perbandingan data latih dan data uji sebesar 80%:20% dengan hasil akurasi yang didapatkan sebesar 77% dan nilai F1 Score sebesar 71%, nilai ini merupakan hasil awal yang memiliki Tingkat keakuratan rendah yang mana dibuktikan dari presicion dan recall hasil dari True Positif dan True Negatif. Model multinomial naïve bayes memprediksi sebanyak 14 data benar dan 50 salah dari 64 total keseluruhan untuk presisi sentimen positif, namun untuk presisi sentimen negatif, model dapat menganalisis sangat baik dengan 1 nilai eror dari 161 total keseluruhan data. Hal ini yang menyebabkan akurasi dari percobaan pertama masih memiliki akurasi yang rendah atau kurang akurat

### Saran

Berdasarkan penelitian yang sudah dilakukan, saran yang dapat diberikan oleh penulis untuk menunjang penelitian ini lebih baik lagi di masa mendatang adalah dengan melakukan perbandingan dengan model lain sebagai analisis tambahan untuk mengetahui kinerja awal model, selain itu dapat dilakukan metode yang lebih komprehensif terkait seleksi fitur yang dapat dilakukan dengan model Multinomial Naive Bayes sehingga hasil penelitian dapat lebih akurat dan bervariasi

## UCAPAN TERIMA KASIH

Terima kasih kepada pihak pihak yang terlibat dalam penelitian ini, yaitu PT KAI Daop 8 Surabaya yang menunjang dalam pemberian data sehingga penelitian dapat dilakukan dengan sebaik mungkin oleh penulis

## REFERENSI

- Chen, M. (2020). *expressions of reviewers ' avatars*. <https://doi.org/10.1108/OIR-08-2018-0251>
- Dananjaya, A. B. T., & Irawan, M. I. (2023). Analisis Sentimen pada Komentar terhadap Kebijakan Perjalanan Domestik yang Dikelompokkan Menggunakan Metode Self-Organizing Maps. *Jurnal Sains Dan Seni ITS*, 12(1). <https://doi.org/10.12962/j23373520.v12i1.107484>
- Estika, I. Di, Irfan, D., & Pratiwi, O. N. (2021). Analisis Sentimen Ulasan Pengguna Untuk Peningkatan Layanan Menggunakan Algoritma Naive Bayes (Studi kasus: Bukalapak). *E-Proceeding of Engineering*, 8(2), 2735–2745.
- Fahrudin, T. M., Sari, A. R. F., Iffadah, A. S., Windyadari, C. C., & ... (2022). Pemodelan Teks

- Tweet pada Isu Pelecehan Seksual Berbasis Analisis Sentimen dan Leksikon Emosi. *Senada*, 2022(Senada), 12–23. <https://senada.upnjatim.ac.id/index.php/senada/article/view/33%0Ahttps://senada.upnjatim.ac.id/index.php/senada/article/download/33/19>
- Farisi, A. A., Sibaroni, Y., & Faraby, S. Al. (2019). Sentiment analysis on hotel reviews using Multinomial Naïve Bayes classifier. *Journal of Physics: Conference Series*, 1192(1). <https://doi.org/10.1088/1742-6596/1192/1/012024>
- Febrisa Sidabutar, A., Habibi, R., & Isti Rahayu, W. (2023). Perbandingan Metode Klasifikasi Untuk Pengelompokan Risiko Magang Mahasiswa. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(3), 2071–2076. <https://doi.org/10.36040/jati.v7i3.7026>
- Hanif Razka, M., Theresiawati, T., & Chamidah, N. (2023). Analisis Sentimen terhadap Aplikasi PeduliLindungi pada Jejaring Sosial Twitter Menggunakan Algoritma Naïve Bayes dan Seleksi Fitur Particle Swarm Optimization. *Informatik : Jurnal Ilmu Komputer*, 19(1), 68–80. <https://doi.org/10.52958/iftk.v19i1.4688>
- Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Teknika*, 10(1), 18–26. <https://doi.org/10.34148/teknika.v10i1.311>
- Jon, A. M. (2022). *Analisis Sentimen Pada Media Sosial Instagram Klub Persija Jakarta Menggunakan Metode Naive Bayes*. 958.
- Muhaimin, A., Rafka, M., & Alamsyah, R. B. (2022). Citayam Fashion Week dalam Twitter. *Senada*, 2022(Senada), 61–65. <https://senada.upnjatim.ac.id/index.php/senada/article/view/47%0Ahttps://senada.upnjatim.ac.id/index.php/senada/article/download/47/25>
- Muktafin, E. H., Kusri, K., & Luthfi, E. T. (2020). Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing. *Jurnal Eksplora Informatika*, 10(1), 32–42. <https://doi.org/10.30864/eksplora.v10i1.390>
- Oktaviani Putri, F., & Cahya Wihandika, R. (2020). Analisis Sentimen pada Ulasan Pengguna MRT Jakarta Menggunakan Metode Neighbor-Weighted K-Nearest Neighbor dengan Seleksi Fitur Information Gain. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 4(7), 2195–2203. <http://j-ptiik.ub.ac.id>
- RI, K. (2010). *Peraturan Menteri Perhubungan Nomor : Km. 29 Tahun 2010 Tentang Penyelenggaraan Kewajiban Pelayanan Publik Angkutan Orang Dengan Kereta Api Kelas Ekonomi Tahun Anggaran 2010*.
- Sari, C. Y. (2019). *Implementasi Metode K-Nearest Neighbor Dalam Memprediksi Bau, Rasa Dan Warna Air Sumur Di Kelurahan Selat Baru Bengkalis*. <https://repository.uir.ac.id/8862/%0Ahttps://repository.uir.ac.id/8862/1/153510266.pdf>
- Telnoni, P. A. (2021). Pelabelan Data Dengan Latent Dirichlet Allocation dan K-Means Clustering pada Data Twitter Menggunakan Bahasa Indonesia. *Jurnal Elektro Dan Telekomunikasi Terapan*, 7(2), 885. <http://journals.telkomuniversity.ac.id/jett/article/view/3442>
- Tempola, F., Muhammad, M., & Khairan, A. (2018). Perbandingan Klasifikasi Antara KNN dan Naive Bayes pada Penentuan Status Gunung Berapi dengan K-Fold Cross Validation. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 5(5), 577. <https://doi.org/10.25126/jtiik.201855983>
- Zhafira, D. F., Rahayudi, B., & Indriati, I. (2021). Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar pada Youtube. *Jurnal Sistem Informasi, Teknologi Informasi, Dan Edukasi Sistem Informasi*, 2(1), 55–63. <https://doi.org/10.25126/justsi.v2i1.24>