

## COMPARISON OF LOGISTIC REGRESSION, MULTINOMIALNB, SVM, AND K-NN METHODS ON SENTIMENT ANALYSIS OF GOJEK APP REVIEWS ON THE GOOGLE PLAY STORE

Audenza Maulana<sup>\*1</sup>, Inayah Khasnaputri Afifah<sup>2</sup>, Asghafi Mubarrak<sup>3</sup>, Kiagus Rachmat Fauzan<sup>4</sup>, Ardhan Dwintara<sup>5</sup>, Bitu Parga Zen<sup>6</sup>

<sup>1</sup> Informatics Engineering, Faculty of Informatics, Institut Teknologi Telkom Purwokerto, Indonesia

<sup>2</sup> Informatics Engineering, Informatics Engineering, Faculty of Informatics, Universitas Alma Ata Yogyakarta, Indonesia

<sup>3</sup> Tourism Academy, Faculty of Tourism, STP AMPTA Yogyakarta, Indonesia

<sup>4</sup> Information Systems, Faculty of Computer Science, Universitas Amikom Purwokerto, Indonesia

<sup>5</sup> Informatics Engineering, Faculty of Engineering and Computer Science, Universitas Sains Al-Qur'an Wonosobo, Indonesia

<sup>6</sup> Institut Teknologi Telkom Purwokerto, Indonesia

Email: <sup>1</sup>[audenzamaul@gmail.com](mailto:audenzamaul@gmail.com), <sup>2</sup>[inayahkhasna@gmail.com](mailto:inayahkhasna@gmail.com), <sup>3</sup>[asghaf124@gmail.com](mailto:asghaf124@gmail.com),  
<sup>4</sup>[fauzanrachmat9@gmail.com](mailto:fauzanrachmat9@gmail.com), <sup>5</sup>[ardandwin@gmail.com](mailto:ardandwin@gmail.com), <sup>6</sup>[bitu@ittelkom-pwt.ac.id](mailto:bitu@ittelkom-pwt.ac.id)

(Article received: January 27, 2023; Revision: March 28, 2023; published: December 25, 2023)

### Abstract

*Today's community activities cannot be separated from the name of transportation because it makes it very easy for people's daily activities. With the existence of transportation, people can more quickly reach their destination. With the Gojek application on the Google Play Store, it will help people travel or deliver goods. To determine service quality, sentiment analysis can be used to classify reviews. The purpose of this study is to compare which method has the best accuracy, so that it can classify reviews into positive or negative sentiments. The methods tested in this study are Logistic Regression, MultinomialNB, SVM, and K-NN. Performance assessment methods include score accuracy, recall, and precision, classification reports, and confusion matrix to determine the appropriate method for classifying reviews into positive or negative categories. Of the four methods tested, the one with the highest performance is the Logistic Regression method. Accuracy, recall and precision scores of the Logistic Regression method were 82.45%, 82.49%, 82.45% and 82.43%, respectively. Classification report also shows good results. In the confusion matrix, there are 111 and 124 True positives and True negatives. There are only 22 and 28 False positive and False negative results respectively. The method that has the lowest score is K-NN, with score accuracy, recall, and precision respectively were 52.28%, 59.43%, 93, 52%, and 65.65%. Classification report shows quite bad results. In the connection matrix, it produces True positives and True negatives 130 and 19. There are only 127 and 9 False positive and False negative results respectively. The results of this study state that using the Logistic Regression method is suitable for use in classifying positive and negative reviews in the review dataset on the Gojek application on the Google Play Store.*

**Keywords:** k-nn, logistic regression, method comparison, multinomialnb, sentiment analysis, svm

## PERBANDINGAN METODE LOGISTIC REGRESSION, MULTINOMIALNB, SVM, DAN K-NN PADA ANALISIS SENTIMEN ULASAN APLIKASI GOJEK DI GOOGLE PLAY STORE

### Abstrak

Aktivitas masyarakat saat ini tidak lepas dari yang namanya transportasi karena sangat memudahkan aktivitas masyarakat setiap harinya. Dengan adanya transportasi, masyarakat dapat lebih cepat mencapai tempat tujuan mereka. Dengan adanya aplikasi Gojek di *Google Play Store* akan membantu masyarakat dalam bepergian atau pun mengantar barang. Untuk mengetahui kualitas pelayanan, analisis sentimen dapat digunakan untuk mengklasifikasikan ulasan. Tujuan penelitian ini untuk membandingkan metode mana yang memiliki akurasi terbaik, sehingga dapat mengklasifikasikan ulasan ke dalam sentiment positif atau negatif. Metode-metode yang diuji di dalam penelitian ini adalah *Logistic Regression*, *MultinomialNB*, *SVM*, dan *K-NN*. Penilaian performa metode berupa skor *accuracy*, *recall*, dan *precision*, *classification report*, serta *confusion matrix* agar dapat

menentukan metode yang cocok untuk mengklasifikasikan ulasan masuk ke kategori positif atau negatif. Dari keempat metode yang diuji, yang memiliki performa tertinggi adalah metode *Logistic Regression*. Skor *accuracy*, *recall*, dan *precision* dari metode *Logistic Regression* secara berturut-turut adalah 82,45%, 82,49%, 82,45%, dan 82,43%. *Classification report* juga menunjukkan hasil yang bagus. Pada *confusion matrix*, menghasilkan *True positive* dan *True negative* 111 dan 124. Hasil *False positive* dan *False negative* secara berturut-turut hanya terdapat 22 dan 28. Metode yang memiliki skor terendah adalah *K-NN*, dengan skor *accuracy*, *recall*, dan *precision* secara berturut-turut adalah 52,28%, 59,43%, 93, 52%, dan 65,65%. *Classification report* menunjukkan hasil yang cukup buruk. Pada *confusion matrix*, menghasilkan *True positive* dan *True negative* 130 dan 19. Hasil *False positive* dan *False negative* secara berturut-turut hanya terdapat 127 dan 9. Hasil penelitian ini menyatakan bahwa menggunakan metode *Logistic Regression* cocok untuk digunakan dalam mengklasifikasikan ulasan positif dan negatif pada *dataset* ulasan aplikasi Gojek di *Google Play Store*.

**Kata kunci:** analisis sentimen, k-nn, logistic regression, multinomialnb, perbandingan metode, svm

## 1. PENDAHULUAN

Perkembangan transportasi selalu meningkat setiap tahunnya, Kendaraan bermotor dimanfaatkan oleh masyarakat untuk membantu aktifitas harian. Bagi masyarakat yang tidak bisa mengendarai kendaraan bermotor akan terasa menyulitkan [1].

Perkembangan teknologi di Indonesia memunculkan banyaknya aktivitas manusia yang dapat didukung dengan sarana digital, salah satunya dalam bidang transportasi. Hal ini terlihat dari bermunculannya berbagai model transportasi berbasis online pada kota-kota besar di Indonesia. Salah satu transportasi online yang paling diminati masyarakat di Indonesia adalah Gojek. PT Aplikasi Karya Anak Bangsa (Gojek) merupakan sebuah perusahaan decacorn asal Indonesia yang didirikan pada tahun 2010 di Jakarta oleh Nadiem Makarim. Dalam menjalankan bisnisnya, PT Aplikasi Karya Anak Bangsa (Gojek) membuat suatu aplikasi yang didaftarkan pada berbagai platform yang terintegrasi dengan smartphone, salah satunya yaitu Google Play Store [2].

Meningkatnya penggunaan ponsel pintar atau biasa kita sebut dengan smartphone dan juga meningkatnya jumlah kendaraan yang ada, menjadikan semakin bertambahnya pengguna yang mengunduh aplikasi gojek, dan di dalam google play store, setiap pengguna dapat memberikan ulasan terkait aplikasi tersebut.

Ulasan pengguna sering kali digunakan sebagai alat yang efektif dan efisien dalam mencari informasi tentang suatu produk atau aplikasi. Bahwa studi terbaru menemukan bahwa hampir 50% pengguna internet mengandalkan rekomendasi dari mulut ke mulut (opini) sebelum menggunakan suatu produk, hal ini dikarenakan review dari pengguna lain dapat memberikan informasi terbaru mengenai produk tersebut berdasarkan perspektif pengguna lain yang telah menggunakan produk [3].

Analisis sentimen merupakan klasifikasi teks yang bertujuan untuk mengklasifikasikan teks (dokumen) yang berisi opini sebagai opini yang positif, negatif, atau netral [4].

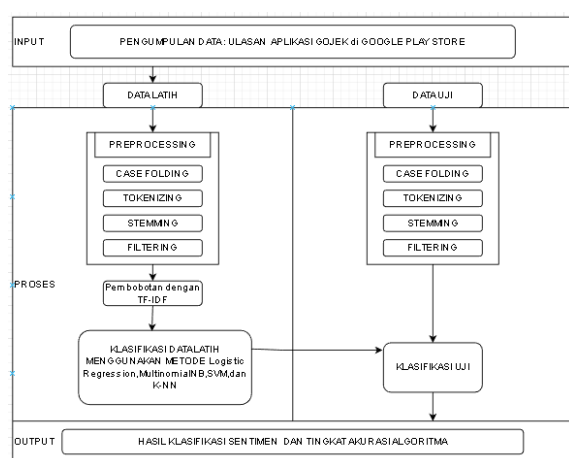
Metode yang cocok untuk mengumpulkan data informasi tersebut adalah penggunaan metode web scraping. Namun, review tersebut masih berupa teks tanpa makna yang konkrit. Oleh karena itu, diperlukan suatu analisis yang mampu mengklasifikasikan *review* pengguna tersebut menjadi sentimen positif dan negatif berdasarkan isi review yang diberikan dengan menggunakan text mining. Permasalahan yang sering muncul pada sebagian besar studi analisis sentimen, yaitu kebanyakan data review cenderung tidak seimbang (*imbalanced datasets*) dari segi jumlah tiap kelas individu, misalnya cenderung ke arah kelas positif atau sebaliknya. Secara umum, algoritma machine learning akan menghasilkan suatu model dengan tingkat sensitivitas yang rendah terhadap kelas minoritas ketika menerima set data yang tidak seimbang karena hal tersebut akan menyebabkan performa klasifikasi analisis sentimen yang dilakukan menjadi buruk. Setelah mendapatkan *dataset* sentimen positif dan negatif, kami menggabungkannya secara manual dan membaginya dengan cukup rata agar tidak terjadi ketidakseimbangan *dataset* (*imbalanced datasets*) [5].

Pada penelitian analisis sentimen yang berjudul “Analisis Sentimen Masyarakat Terhadap Kebijakan Vaksinasi Covid-19 pada Media Sosial Twitter Menggunakan Metode *Logistic Regression*” metode *Logistic Regression* mampu melakukan klasifikasi *tweet* dengan tingkat akurasi 67% [6]. Selanjutnya, pada penelitian yang berjudul “Implementasi Metode Multinomial Naïve Bayes Classifier untuk Analisis Sentimen” metode *MultinomialNB* menghasilkan rata-rata akurasi sebesar 79% dan maksimum akurasi sebesar 82%. [7]. Untuk metode SVM, pada penelitian dengan judul “Penerapan Algoritma SVM untuk Analisis Sentimen pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia”, diperoleh akurasi sebesar 82% untuk metode SVM [8]. Sedangkan untuk metode K-NN, pada penelitian dengan judul “Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia” diperoleh akurasi sebesar 100% [9].

Berdasarkan penjabaran tersebut, penelitian ini bertujuan melakukan analisis sentimen Ulasan Aplikasi Gojek di *Google Play Store*, dengan menggunakan metode klasifikasi supervised learning diantaranya metode *Logistic Regression*, *MultinomialNB*, *SVM*, dan *K-NN* yang mana metode klasifikasi tersebut kemudian dibandingkan kinerja performa klasifikasinya untuk mendapatkan metode terbaik berdasarkan *classification report* dan *confusion matrix*.

## 2. METODE PENELITIAN

Berikut ini adalah gambaran umum sistem pada penelitian ini.



Gambar 1. Gambaran umum tahapan penelitian

Gambaran umum pada gambar 1 merupakan bentuk kasaran kegiatan yang kami lakukan untuk melakukan uji coba perbandingan model yang kami teliti menggunakan *dataset* ulasan aplikasi Gojek.

### 2.1. Pengumpulan Data

Input yang digunakan yaitu review *Google Play Store* Gojek. Beberapa dari data ini akan menjadi data latih dan data uji. Data tersebut terlebih dahulu dilakukan *preprocessing* untuk membedakan antara komentar positif dan negatif. Tahap *preprocessing* data diperoleh dari data *review* Gojek di *Google Playstore* dan diproses melalui *proses case folding* dan *filtering*.

### 2.2. Tahap Preprocessing Data

Mentahan dari data review *Google Play Store* Gojek diolah melalui proses *case folding*, *tokenizing*, *stemming*, dan *filtering*. Klasifikasi data yang dilatih menggunakan *Logistic Regression*, *MultinomialNB*, *SVM*, dan *K-NN*. Berikut ini merupakan penjelasan dari tahapan *preprocessing*:

#### 1. Case folding

Pada tahap ini, dilakukan proses perubahan huruf besar menjadi huruf kecil dan menghilangkan seluruh tanda baca pada kalimat

#### 2. Tokenizing

Pada tahapan ini, setiap kata akan dipisahkan berdasarkan spasi yang ditemukan.

#### 3. Stemming

Pada tahapan *stemming*, yaitu perubahan kata berimbuhan menjadi kata dasar.

#### 4. Filtering

Pada tahapan ini, dilakukan proses pembuangan kata kata yang tidak penting dari hasil token [10].

### 2.3. TF-IDF

Untuk penelitian ini, kami menggunakan TF-IDF untuk pembobotan *dataset* kami. TF-IDF (*term frequency-inverse document frequency*) adalah metrik statistik untuk menentukan pentingnya kata atau kumpulan kata dalam sebuah *dataset*. Nantinya, kata yang sering muncul akan diberi nilai terendah. Kemudian, hasil dari pembobotan TF-IDF dapat diimplementasikan menggunakan algoritma yang telah kami persiapkan untuk dibandingkan performanya [11].

### 2.4. Tahap Klasifikasi Naïve Bayes

Model klasifikasi *Multinomial* dan *Multivariate Bernoulli* merupakan metode klasifikasi berbasis peluang yang paling sering digunakan. Model klasifikasi ini banyak digunakan karena mudah diaplikasikan dan prosesnya sederhana (Manning et al. 2008). Pada model *Multivariate Bernoulli* setiap dokumen memiliki atribut yang menunjukkan ada atau tidaknya kata-kata atau term dalam dokumen tersebut, tetapi jumlah kemunculan term dalam dokumen tidak ikut diperhitungkan. Pada model *Multinomial Naïve Bayes*, jumlah kemunculan term pada dokumen ikut diperhitungkan, setiap dokumen diwakili oleh kemunculan term dari dokumen. Penelitian ini menggunakan model *Multinomial Naïve Bayes*, dimana dokumen diwakili oleh serangkaian kemunculan dokumen term pada suatu dokumen, yaitu:

$$d = \langle t_1, \dots, t_k, \dots, t_{(n_d)} \rangle, t_k \in V \quad (1)$$

Dengan  $n_d$  adalah jumlah token unik pada dokumen dan  $V$  adalah himpunan kata unik yang terdapat dalam korpus data latih. Kata unik dalam hal ini disebut juga *vocabulary*. Pada model ini, setiap kata dalam data uji akan dihitung frekuensi katanya dengan persamaan.

$$P(t|c) = \frac{Tct}{\sum_{t' \in V} Tct'} \quad (2)$$

Selanjutnya jumlah kata pada setiap kelas akan dihitung, begitu juga dengan jumlah *vocabulary* ( $B$ ) yang ada. Setelah nilai frekuensi setiap kata dalam suatu dokumen, jumlah keseluruhan frekuensi kata dalam setiap kelas, dan jumlah *vocabulary* kata diketahui, maka setiap kata  $t$  dalam kelas  $c$  dapat dihitung nilai peluangnya  $t|c$ . Dari perhitungan peluang ini, dapat terjadi kemungkinan nilai peluang

tersebut nol jika frekuensi kata pada data uji di kelas  $c$  ( $Tct$ ) bernilai nol. Untuk itu digunakan *Laplace Smoothing*, dengan persamaan:

$$P(c|d) \propto P(c) \prod_{1 \leq k < n_d} P(t_k|c) \quad (3)$$

Parameter  $P(c)$  adalah peluang dokumen yang ada pada kelas  $c$ ,  $(tk|c)$  adalah peluang kata atau token  $tk$  yang muncul pada dokumen  $c$ , dan  $n_d$  adalah jumlah token unik pada dokumen. Untuk itu nilai peluang  $P(t|c)$  yang didapatkan dari *Laplace smoothing* sangat berpengaruh karena jika nilai peluangnya nol, maka hasil peluang  $P(c|d)$  data uji yang didapatkan juga akan nol [12].

### 2.5. Tahap Klasifikasi Logistic Regression

Tujuan melakukan analisis data kategori menggunakan *Logistic Regression* adalah mendapatkan model terbaik dan sederhana untuk menjelaskan hubungan antara keluaran dari variabel respons ( $y$ ) dengan variabel-variabel prediktornya ( $x$ ). Variabel respons dalam regresi logistik dapat berupa kategori atau kualitatif, sedangkan variabel prediktornya dapat berupa kualitatif dan kuantitatif. Jika variabel merupakan variabel biner atau dikotomi dalam artian variabel respons terdiri dari dua kategori yaitu “sukses” (= 1) atau “gagal” (= 0), maka variabel mengikuti sebaran *Bernoulli* yang memiliki fungsi densitas peluang.

Rumus *Logistic Regression*:

$$\sum_{i=1}^m \frac{y_{pred}^{(i)} = y_{true}^{(i)}}{m} \quad (4)$$

### 2.6. Tahap Klasifikasi SVM

*Support Vector Machine* (SVM) merupakan salah satu metode dalam *supervised learning* yang biasanya digunakan untuk klasifikasi (seperti *Support Vector Classification*) dan regresi (*Support Vector Regression*). Dalam pemodelan klasifikasi, SVM memiliki konsep yang lebih matang dan lebih jelas secara matematis dibandingkan dengan teknik-teknik klasifikasi lainnya. SVM juga dapat mengatasi masalah klasifikasi dan regresi dengan linier maupun nonlinier [13].

$$f(x_d) = \sum_{i=1}^{n_{sv}} a_i y_i (y x_i^T x + r)^p + b \quad (5)$$

### 2.7 Tahap Klasifikasi Naïve Bayes

Algoritma *K-Nearest Neighbor* merupakan salah satu metode untuk proses klasifikasi terhadap suatu objek berdasarkan data training yang memiliki jarak paling dekat dengan objek. Penentuan jarak dengan nilai terdekat dan terjauh dihitung berdasarkan jarak *Euclidean* (J. Nilson, 1996).

*Preprocessing*, langkah pertama adalah menyiapkan data *training* dan mendapatkan *tuple* himpunan  $DS = \{(d_i, c_j) \mid 1 \leq i < n, 0 \leq j < m\}$

dimana  $d_i$  adalah term *vector representation* teks dokumen dan  $c_j$  kategori label.

*Similarity Measure*, menggunakan *TF x IDF* untuk menghitung bobot setiap term dalam dokumen, sebagai variasi untuk meningkatkan akurasi yang signifikan. *Term Frequency* merupakan jumlah kemunculan sebuah term dalam sebuah dokumen. *TF* dapat dirumuskan pada persamaan berikut:

$$w_{t,d} = \log(tf_{t,d} + 1) \quad (6)$$

*Inverse Document Frequency* adalah log dari kebalikan probabilitas term yang ditemukan di dalam dokumen. *IDF* dapat dirumuskan pada persamaan berikut:

$$idf_t = \log\left(\frac{n}{n_t}\right) \quad (7)$$

*Cosine Similarity* adalah fungsi yang digunakan untuk menghitung kesamaan antara semua data latih dengan dokumen  $X$ . *Cosine similarity* dapat dirumuskan pada persamaan berikut (Suguna, 2010) [14]:

$$SIM(X, d_i) = \frac{\sum_{j=1}^m x_j \cdot d_{ij}}{\sqrt{(\sum_{j=1}^m x_j)^2} \sqrt{(\sum_{j=1}^m d_{ij})^2}} \quad (8)$$

## 3. HASIL DAN PEMBAHASAN

Tabel 1. Data sebelum dilakukan *preprocessing*

Parah ribet banget untuk cara cancel nya . driver pada cancel orderan dengan cara di abaikan supaya tercancel otomatis . Gara2 itu pihak resto jadi males buat orderan nya. Apalagi kalo lagi order malem2 gini pihak resto jadi sungkan untuk membuat orderan nya karna driver nya php. Balikin ajha lah cara cancel nya seperti dulu. Kalo kaya gini jadi nya ribet bikin pusing. Rencana pgn makan malah emosi Kesel gw lama2 sama gojek .
---

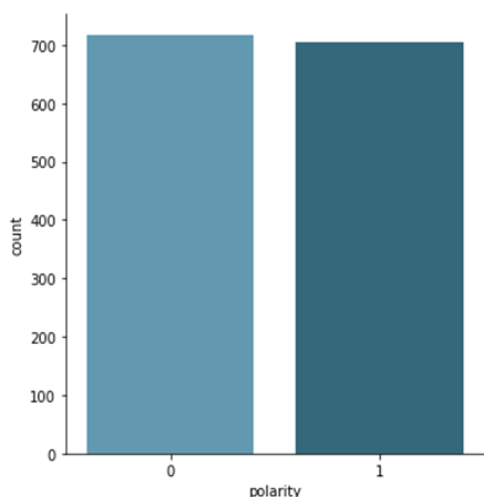
Tabel di atas merupakan data awal sebelum dilakukannya *preprocessing*. Di dalam tabel tersebut belum diklasifikasikan polaritasnya. Pengklasifikasian polaritas dilakukan menggunakan *labelling lexicon*. Data ulasan juga melalui proses *cleaning*, yang melibatkan *case folding*, *stemming*, *tokenizing*, dan *filtering*.

Tabel 2. Data setelah dilakukan *preprocessing*

<i>polarity</i>	<i>content_clean</i>
0	parah ribet banget untuk cara cancel nya . driver pada cancel orderan dengan cara di abaikan supaya

tercancel otomatis . gara2 itu pihak resto jadi males buat orderannya. apalagi kalo lagi order malem2 gini pihak resto jadi sungkan untuk membuatnya karna driver nya php. balikin ajha lah cara cancel nya seperti dulu. kalo kaya gini jadi nya ribet bikin pusing. rencana pgn makan malah emosi kesel gw lama2 sama gojek .

Untuk tabel 2, ulasan pada tabel 1 telah dilakukan *cleaning* dan *labeling lexicon*. Terdapat kolom polaritas pada tabel 2, yang berisikan 0, yang berarti memiliki polaritas negatif. Untuk polaritas positif, dilambangkan dengan angka 1.



Gambar 2. Grafik prediksi ulasan negatif (0) dan positif (1)

Dari 2000 sampel, diambil 717 data sentimen negatif dan 705 data sentimen positif untuk data latih, seperti yang terlihat pada gambar 2. Hasil sentimen yang bernilai netral tidak kami gunakan ke dalam perbandingan. Kemudian, diambil 191 data berupa review acak dari hasil sampel, yang akan digunakan sebagai data prediksi. Data prediksi ini nantinya akan digunakan untuk menentukan apakah Machine Learning dapat memprediksi ulasan sebagai ulasan positif atau negatif.

Untuk data latih, data dibagi menjadi 80% untuk data latih dan 20% data uji dengan *random\_state* sebesar 42. Pengujian yang akan menentukan mana metode yang terbaik adalah menggunakan *classification report* dan *confusion matrix*. *Classification report* adalah sebuah metrik evaluasi performa yang menunjukkan precision, recall, F1-score, dan metrik pendukung lainnya dari model yang telah dilatih [15]. *Confusion matrix*

adalah tabel dengan empat klasifikasi, yang berisi nilai prediksi dan nilai aktual . Empat klasifikasi tersebut adalah *True positive*, *True negative*, *False positive*, dan *False negative* [16].

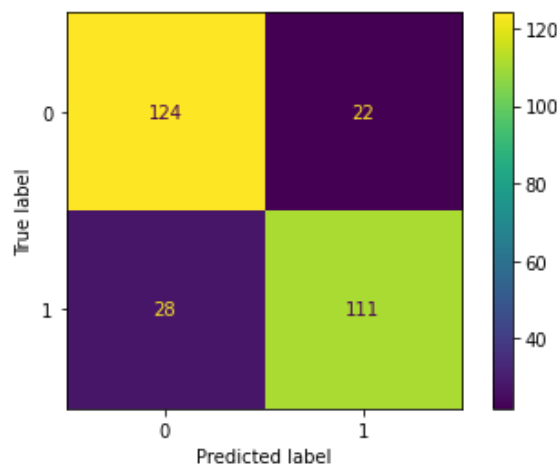
*Dataset* tersebut kemudian akan kami ujikan terhadap empat metode yang kami uji cobakan, yaitu *Logistic Regression*, *Multinomial Naïve Bayes*, *SVM*, dan *K-Nearest Neighbor (K-NN)*.

### 3.1. Logistic Regression

Metode pertama yang akan digunakan untuk menguji dan mengklasifikasikan data ulasan ini adalah *Logistic Regression*. Dengan menggunakan hyperparameter  $C = 3.7$ , *solver = liblinear*, dan  $max\_iter = 200$ , diperoleh nilai data latih sebesar 94,54%. Setelah menghitung nilai vektor latih, kami melakukan pengujian *classification report* dan *confusion matrix* pada data uji dengan hasil sebagai berikut.

Tabel 3. *Classification report* data uji *Logistic Regression*

	Precision	Recall	F1 Score	Support
0	0.82	0.85	0.83	146
1	0.83	0.80	0.82	139
accuracy			0.82	285
Macro avg	0.83	0.82	0.82	285
Weighted avg	0.82	0.82	0.82	285



Gambar 3. *Confusion matrix* data uji *Logistic Regression*

Hasil *classification report* dan *confusion matrix* pada tabel 1 dan gambar 3 akurat, di mana model memiliki skor yang cukup tinggi pada bagian precision, recall, f1-score, dan accuracy. Metode ini juga dapat memprediksi *True positive* dan *True negative* dengan cukup tepat. Hasil prediksi *False positive* dan *False negative* menggunakan metode ini cukup kecil, yaitu hanya terdapat 22 *False positive* dan 28 *False negative*.

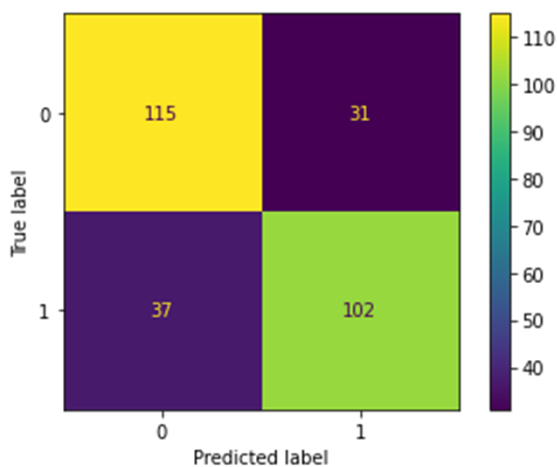
### 3.2. Multinomial Naïve Bayes

Metode yang diuji berikutnya adalah metode *Multinomial Naïve Bayes*. Hyperparameter yang digunakan *MultinomialNB* adalah  $alpha = 0.5$ .

Setelah menerapkan hyperparameter, kemudian vektor latih x dan y dilakukan fit. Menggunakan fungsi score, dapat diketahui score vektor latih x dan y, yaitu sebesar 83,90%. Setelah menghitung nilai vektor, dilakukan pengujian *classification report* dan *confusion matrix* dengan hasil sebagai berikut

Tabel 4. *Classification report* data uji MultinomialNB

	Precision	Recall	F1 Score	Support
0	0.76	0.79	0.77	146
1	0.77	0.73	0.75	139
accuracy			0.76	285
Macro avg	0.76	0.76	0.76	285
Weighted avg	0.76	0.76	0.76	285



Gambar 4. *Confusion matrix* data uji MultinomialNB

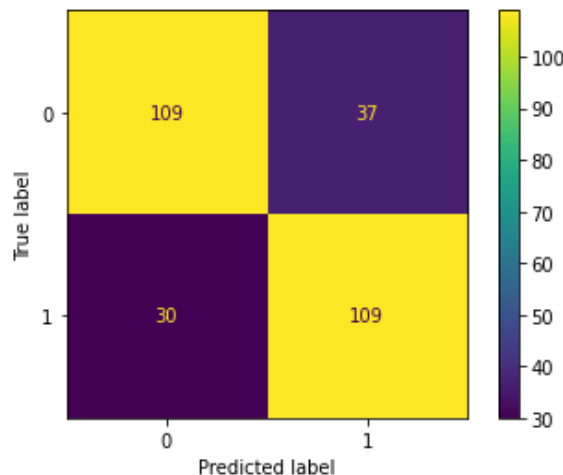
Berdasarkan tabel 2 dan gambar 4, hasil *classification report* dan *confusion matrix* terbilang cukup akurat. Skor precision, recall, f1-score, dan accuracy memiliki nilai yang cukup tinggi, yaitu di kisaran kurang dari (<) 80%, namun lebih dari (>) 70%. Hasil *confusion matrix*-nya pun cukup baik, walau hasilnya masih di bawah *Logistic Regression*. Hal tersebut dikarenakan angka *False positive* dan *False negative* yang diklasifikasikan oleh Multinomial Naïve Bayes berjumlah 31 *False positive* dan 37 *False negative*.

### 3.3. SVM

Berikutnya adalah metode SVM digunakan untuk menguji dan mengklasifikasikan data ulasan ini adalah SVM. Dengan menggunakan hyperparameter C = 3, kernel = rbf, dan max\_iter = 150, diperoleh nilai data latih sebesar 97,57%. Setelah menghitung nilai vektor latih, kami melakukan pengujian *classification report* dan *confusion matrix* pada data uji dengan hasil sebagai berikut.

Tabel 5. *Classification report* data uji SVM

	Precision	Recall	F1 Score	Support
0	0.78	0.75	0.76	146
1	0.75	0.78	0.76	139
accuracy			0.76	285
Macro avg	0.77	0.77	0.76	285
Weighted avg	0.77	0.76	0.76	285



Gambar 5. *Confusion matrix* data uji SVM

Berdasarkan tabel 3 dan gambar 5, hasil *classification report* dan *confusion matrix* terbilang cukup akurat, namun performanya masih lebih rendah dibandingkan *Logistic Regression* dan Multinomial NB. Skor precision, recall, f1-score, dan accuracy memiliki nilai yang cukup tinggi, yaitu di kisaran kurang dari (<) 80%, namun lebih dari (>) 70%. Hasil *confusion matrix*-nya pun cukup baik Hal tersebut dikarenakan angka *False positive* dan *False negative* yang diklasifikasikan oleh SVM berjumlah 37 *False positive* dan 30 *False negative*.

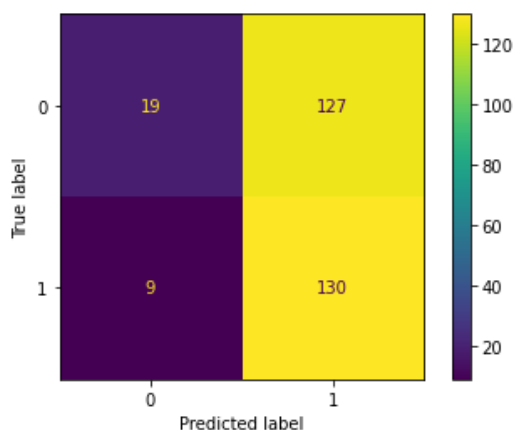
### 3.4. K-Nearest Neighbor

Selanjutnya, Metode yang diuji adalah metode KNN. Hyperparameter yang digunakan adalah n\_neighbors = 2, weights = 'distance', algorithm = 'brute', leaf\_size = 100, p = 2, metric = 'minkowski'. Setelah menerapkan hyperparameter, kemudian vektor latih x dan y dilakukan fit. Menggunakan fungsi score, dapat diketahui score vektor latih x dan y, yaitu sebesar 100%. Setelah menghitung nilai vektor, dilakukan pengujian *classification report* dan *confusion matrix* dengan hasil sebagai berikut.

Tabel 6. *Classification report* data uji KNN

	Precision	Recall	F1 Score	Support
0	0.68	0.13	0.22	146
1	0.51	0.94	0.66	139
accuracy			0.52	285
Macro avg	0.59	0.53	0.44	285
Weighted avg	0.59	0.52	0.43	285



Gambar 6. *Confusion matrix* data uji K-NN

Berdasarkan tabel 4 dan gambar 6, hasil *classification report* performanya masih lebih rendah dibandingkan *Logistic Regression*, SVM dan Multinomial NB. Hasil *confusion matrix*-nya kurang bagus. Hal tersebut dikarenakan angka *False positive* dan *False negative* yang diklasifikasikan oleh K-NN berjumlah 127 *False positive* dan 9 *False negative*.

#### 4. DISKUSI

Berikut merupakan tabel performa dari keempat algoritma yang kami teliti terhadap *dataset* ulasan aplikasi Gojek di Google Play Store. Terdapat *accuracy*, *precision*, *recall*, dan *f1-score* untuk perbandingan performa keempat metode, yaitu *Logistic Regression*, MultinomialNB, SVM, dan K-NN.

Tabel 7. Perbandingan performa keempat algoritma

Metode	Accuraton	Precisi on	Recal l	F1-Score
LR	82,45%	82,49%	82,45%	82,43%
Multinomia INB	76,41%	76,16%	76,14%	76,11%
SVM	76,49%	76,58%	76,49%	76,49%
K-NN	52,28%	59,43%	93,52%	65,65%

Berdasarkan pada tabel 1-5 dan gambar 3-6, diperoleh masing-masing performa yang telah diuji menggunakan *classification report* dan *confusion matrix*.

Terlihat perbedaan dari nilai *precision*, *recall*, F1 Score, dan *accuracy* yang diperoleh dari performa tiap metode. Performa yang dapat dilihat dari tabel 1 dan tabel 5, menunjukkan bahwa *Logistic Regression* yang nilai akurasi terendah seperti yang didapatkan dari penelitian dengan berjudul “Analisis Sentimen Masyarakat Terhadap Kebijakan Vaksinasi Covid-19 pada Media Sosial Twitter Menggunakan Metode *Logistic Regression*” dengan tingkat akurasi 67%, berbeda signifikan dari

penelitian yang kami uji ini menggunakan *dataset* yang berbeda, yaitu ulasan aplikasi Gojek.

Performa metode *K-Nearest Neighbor* yang mendapatkan tingkat akurasi yang tertinggi seperti yang telah diujikan pada penelitian dengan judul “Analisis Sentimen Menggunakan *K-Nearest Neighbor* Terhadap *New Normal* Masa Covid-19 Di Indonesia” diperoleh akurasi sebesar 100%. Pada penelitian yang kami lakukan ini, tingkat akurasi turun drastis menjadi yang terendah, yaitu hanya mendapatkan 52%, seperti yang dapat dilihat pada tabel 4 dan tabel 5.

#### 5. KESIMPULAN

Berdasarkan hasil penelitian di atas, dapat dilihat pada tabel 5 performa dari keempat algoritma yang diuji, yaitu *Logistic Regression*, *MultinomialNB*, *SVM*, dan *K-NN*, yang memiliki performa terbaik keseluruhan adalah metode *Logistic Regression*. Performa skor *Logistic Regression* secara berturut-turut adalah 82,45%, 82,49%, 82,45%, dan 82,43%.

Metode *K-NN* memiliki skor *recall* tertinggi, yaitu 93,52%, namun skor lainnya terbilang rendah. Sehingga karena alasan tersebut, metode *K-NN* tidak cocok untuk digunakan dalam mengklasifikasikan *dataset* ulasan aplikasi Gojek di *Google Play Store*. *MultinomialNB* dan *SVM* memiliki performa yang cukup baik, yaitu pada rentang sekitar 76% untuk keseluruhan skornya, mulai dari *accuracy*, *precision*, *recall*, serta *f1-score*. Kedua metode itu dapat digunakan kembali untuk *dataset* yang serupa untuk penelitian selanjutnya..

#### DAFTAR PUSTAKA

- [1] A. A. d. P. R. Adawia, "Analisis Perkembangan Industri Transportasi Online di Era Inovasi Disruptif (Studi Kasus PT Gojek Indonesia)", *Cakrawala-Jurnal Humaniora*, vol. 18, no. 2, p. 8, 2018.
- [2] M. N. M. d. I. Kharisudin, "Analisis Sentimen Aplikasi Gojek Menggunakan Support Vector Machine Dan K Nearest NEIGHBOR", *UNNES Journal of Mathematics*, vol. 2, no. 10, p. 23, 2021.
- [3] F. FANANI, "Klasifikasi Review Software Pada Google Play Menggunakan Pendekatan Analisis Sentimen", Program Studi Teknologi Informasi Fakultas Teknik UGM Yogyakarta, 2017.
- [4] M. Alwi. "Classification of Book Based on DDC 23 Using Text Mining Algorithm". UNIDA Gontor Library, 2021
- [5] M. I. Putri “Analisis Sentimen Pengguna Aplikasi Marketplace Tokopedia Pada Situs Google Play Menggunakan Metode Support Vector Machine(SVM), Naive Bayes, dan Logistic Regression”, 2022.

- [6] A. Shiddicky, S Agustian, "Analisis Sentimen Masyarakat Terhadap Kebijakan Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Metode Logistic Regression", *Teknik informatika, Sains dan Teknologi*, Universitas Islam Negeri Sultan Syarif Kasim Riau, 2022.
- [7] E. L. W. Ningrum, dan A. P. Widodo, "Implementasi Metode Multinomial Naïve Bayes Classifier Untuk Analisis Sentimen", *Journal of Fundamental Mathematics and Applications (JFMA)*, vol. 1, no. 2, pp. 128-137, Nov. 2018.
- [8] D. Darwis, E. S. Pratiwi, A. F. O. Pasaribu, "Penerapan Algoritma Svm untuk Analisis Sentimen pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia", *Jurnal Ilmiah Editor/Vol.7, No.1*, November 2020.
- [9] M. Furqan, S. Sriani, S. M. Sari, "Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia", *Jurnal Teknologi Informasi Vol.21, No.1*, 2022.
- [10] E. L. W. Ningrum, dan A. P. Widodo, "Implementasi Metode Multinomial Naïve Bayes Classifier Untuk Analisis Sentimen", *Journal of Fundamental Mathematics and Applications (JFMA)*, vol. 1, no. 2, pp. 128-137, Nov. 2018
- [11] "Understanding TF-IDF (Term Frequency-Inverse Document Frequency)", 2023. <https://www.geeksforgeeks.org/understanding-tf-idf-term-frequency-inverse-document-frequency> (accessed Jan. 26, 2023).
- [12] J. C. W. Pantouw, "Perbandingan Klasifikasi Rocchio Dan Multinomial Naïve Bayes Pada Analisis Sentimen Data Twitter Bahasa Indonesia", Departemen Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor, 2017.
- [13] J. B. Kelvin, E. Indra, S. H. Sinurat, "Analisis Perbandingan Sentimen Coronavirus Disease-2019 (Covid19) Pada Twitter Menggunakan Metode Logistic Regression Dan Support Vector Machine (Svm)", *Sistem Informasi, FTIK*, Universitas Prima Indonesia Jalan Sampul, Medan, 2022
- [14] R. I. Pristiyanti, M. A. Fauzi, L. Muflikhah, "Sentiment Analysis Peringkasan Review Film Menggunakan Metode Information Gain Dan K-Nearest Neighbor", *Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya*, 2018.
- [15] "Classification report in Machine Learning", 2021. <https://thecleverprogrammer.com/2021/07/07/classification-report-in-machine-learning> (accessed Jan. 27, 2023)
- [16] "*Confusion matrix*", 2020. <https://socs.binus.ac.id/2020/11/01/confusion-matrix> (accessed Jan. 27, 2023)