

# Sentiment Analisis pada moda Transportasi BUS untuk Meningkatkan Kualitas Layanan pada Perusahaan Otobus Terendah

Asep Saphari\*, Leonard Goermanto

Informatika, Universitas Mercu Buana  
\*41517120025@student.mercubuana.ac.id

**Abstrak**— Transportasi BUS merupakan alat transportasi masal yang digunakan untuk berpindah dari satu tempat ke tempat lain atau dari satu kota ke kot lain dengan jumlah lebih dari satu penumpang. Kualitas layanan yang diberikan oleh masing-masing perusahaan penyedia layanan transportasi berbeda-beda. Proses evaluasi dan peningkatan layanan dibutuhkan untuk meningkatkan minat masyarakat dalam memilih moda transportasi dalam perjalanannya. Penggunaan metode algoritma K-Nearest Neighbor (KNN) dan Naive Bayes dapat digunakan untuk mengklasifikasikan perusahaan otobus yang ramai digunakan oleh masyarakat dalam melakukan aktifitas perjalanan dan memilih fasilitas terbaik pilihan masyarakat. Data yang digunakan merupakan data yang dikumpulkan melalui Pusat Data dan Informasi (PUSDATIN) DKI Jakarta yang terkoordinasi di terminal Pulogebang Jakarta Timur dan dengan survey yang disebarkan kepada masyarakat melalui media sosial. Dalam penelitian ini mendapatkan *Accuracy* 0.12, *Macro Avg* 0.02 dan *Weight Avg* 0.11 dengan jumlah data 2831.

**Kata Kunci**— Bus; Clustering; K-NN; Naive Bayes; Transportasi masal.

DOI: 10.22441/jitkom.v7i1.010

## Article History:

Received: Feb 2, 2022

Revised: Jun 23, 2023

Accepted: July 3, 2023

Published: July 10, 2023

## I. PENDAHULUAN

Transportasi masal sangat beragam yang digunakan di Indonesia diantaranya Transportasi Darat, Laut, Udara. [1] Dalam penelitian ini berfokuskan dalam menelitian penggunaan alat Transportasi darat yaitu Bus yang lebih banyak digunakan oleh masyarakat menengah kebawah. Kualitas layanan yang diberikan oleh penyedia jasa layanan Otobus ini sangat beragam mulai dari penyediaan fasilitas seperti kamar kecil, bantal, *usb port* dan lain-lain. Pengguna transportasi Bus pasti akan memilih bus yang nyaman, fasilitas banyak dan harga terjangkau. Tidak sedikit masyarakat yang memilih Bus untuk perjalanannya yang mengutamakan fasilitas dan ketepatan waktu. [2]

Pertumbuhan perusahaan Otobus baru semakin tahun semakin bertambah dan semakin hari semakin banyak pembukaan rute baru dalam menjalankan bisnis busnya. Banyak sekali perusahaan Bus yang hanya membuka rute baru namun dengan hasil yang tidak maksimal karena minimnya penelitian terlebih dahulu. Tercatat dari bulan Februari 2020 – Juni 2021 Jumlah PO Bus yang masuk ke teminal Pulogebang mencapai 109 PO Bus yang beroperasi membawa penumpang.

Secara teknis kenaikan penumpang dan penurunan penumpang pada bus memang tidak dapat dipastikan karena tidak adanya pemberhentian khusus yang mewajibkan penumpangnya untuk naik di terminal A dan turun di terminal B. namun dengan data yang didapatkan peneliti dapat menganalisa dengan kedatangan jumlah penumpang yang ada

di terminal pulogebang sebagai acuan dalam penelitian selanjutnya.

Sumber data yang diberikan oleh PusDatIn sangat berguna untuk meninjau banyaknya penumpang yang turun dari terminal asal ke terminal Pulogebang Jakarta Timur, data yang terupdate adalah data yang berisikan Bulan, Tahun, Nama\_PO, Tujuan terminal dan jumlah PNP.

Dengan data yang didapat, peneliti merasa masih membutuhkan data yang dapat menghasilkan sentiment analisis sebagai kepuasan penumpang dalam memilih moda transportasi Bus sebagai pilihannya, maka dari itu peneliti melakukan survey random di media sosial yang berisikan pertanyaan seputar dengan kenyamanan dan fasilitas pada masing-masing PO.

Dengan terkumpulnya data yang diberikan oleh PusDatIn dan Survey yang telah dilakukan maka peneliti menyimpulkan pengolahan data dengan Algoritma *K-Nearest Neighbor* (K-NN) dan Algoritma *Naive Bayes* untuk mengklasifikasikan Bus ramai dan Sepi serta untuk menjadikan pertimbangan bagi pengusaha Otobus untuk membuka rute perjalanan baru. [3]

Peneliti memilih Algoritma K-NN karena data yang diperoleh adalah berupa data kedatangan dari terminal asal dengan jumlah penumpang dan pengelompokan sepi dan ramai berdasarkan kapasitas yang ada pada Bus. Peneliti juga memilih Algoritma *Naive Bayes* untuk menyimpulkan dan membuat *clustering* pada survey yang dibuat untuk menunjukkan *positive* dan *negative* pada respon masyarakat terhadap Bus yang beroperasi di Indonesia. Metode yang digunakan adalah dengan

melakukan *preprocessing* dan *labeling* pada data untuk merapihkan data yang sudah ada

## II. LITERATURE REVIEW

Pada penelitian sebelumnya, [4] melakukan penelitian yang mendapatkan hasil peningkatan dengan menambahkan fitur naive bayes berbasis algoritma genetika, dengan akurasi yang dihasilkan dari hasil pengujian algoritma Naive Bayes dengan metode pengujian confusion matrix dan kurva AUC dan merekomendasikan bahwa algoritma klasifikasi Naive Bayes berbasis Algoritma Genetika sebagai algoritma klasifikasi yang baik dibandingkan dengan Naive Bayes Pada penelitian kali ini peneliti mencari referensi sebagai landasan teori dan sebagai dasar penelitian untuk pengembangan Metode yang sudah dilakukan. Kemudian [5] juga melakukan sentimen analisis yang diterapkan pada Optimalisasi Pemasaran Destinasi Pariwisata.

### A. Clustering

Clustering merupakan pembagian data ke dalam grup-grup yang memiliki karakteristik yang sama [6]. Sekelompok data akan terbagi menjadi kumpulan data dan dimasukkan ke dalam cluster tertentu sesuai dengan kriteria pada cluster yang dibentuk.

### B. Machine Learning

Machine learning merupakan algoritma matematika dalam sebuah aplikasi komputer yang didasari dengan pembelajaran yang berasal dari data dan akan menghasilkan sebuah prediksi untuk masa yang akan datang [7]. machine learning digunakan oleh perusahaan dalam menganalisa kegiatan atau perkembangan pada sebuah pergerakan dan dijadikan pembelajaran untuk memprediksi masa yang akan datang.

### C. Text Mining

Text mining merupakan proses untuk menemukan informasi baru dan tidak dikenal menggunakan mesin komputer dengan secara otomatis mengekstrak informasi dari berbagai sumber [8]. Text mining diperlukan dalam proses algoritma untuk machine learning.

### D. Naive Bayes

Naive Bayes merupakan metode machine learning yang memanfaatkan perhitungan statistik dan probabilitas yang ditemukan oleh seorang ilmuwan Inggris bernama Thomas Bayes. Dasar dari pemrograman Naive Bayes yang dipakai pada persamaan sebagai berikut.

$$P(A|B) = (P(B|A) * P(A)) / P(B) \quad (1)$$

Keterangan:

Peluang kejadian A sebagai B ditentukan dari peluang B saat A, peluang A, dan peluang B [9].

Naive Bayes diartikan sebagai salah satu metode dalam pengklasifikasian dengan probabilitas sederhana yang bersumber pada Teorema Bayes dan juga memiliki asumsi ke tidak tergantung yang tinggi pada masing-masing kondisi

[10]. Naive bayes memiliki ketepatan sangat baik pada semua data dan tanpa aturan pada transparasinya.

### E. K-Nearest Neighbor (K-NN)

Algoritma KNN merupakan salah satu metode untuk melakukan sebuah klasifikasi terhadap data atau object berdasarkan data atau object yang saling berdekatan antara satu dengan object lainnya [11]. Pada fase klasifikasinya fitur data yang sama dihitung untuk melakukan data test.

### F. Data Training

Data training merupakan sekumpulan data yang mempunyai atribut label atau kelas yang digunakan oleh *machine learning* untuk mengenal karakter dari data sehingga dapat menghasilkan sebuah model data [12].

### G. Data Testing

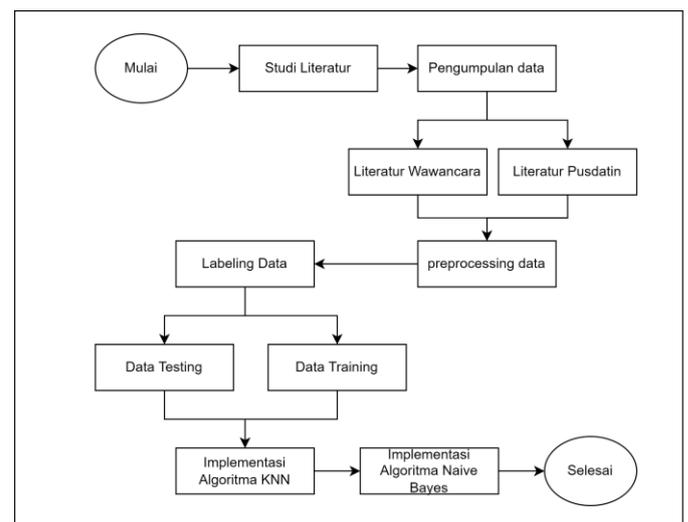
Data testing merupakan sekumpulan data yang memiliki label atau kelas yang biasanya digunakan untuk menguji ketepatan model dalam melakukan proses algoritma klasifikasi [12].

### H. Data Klasifikasi

Klasifikasi data merupakan metode untuk mengelompokan data berdasarkan ketentuan yang berlaku atau berdasarkan variable untuk memprediksi suatu objek yang kategorinya belum diketahui [13].

## III. METODOLOGI PENELITIAN

Metode yang digunakan dalam penelitian ini yaitu dengan melakukan beberapa proses seperti pada gambar 1.



Gambar. 1 Diagram Kerja

### A. Studi Literatur

Peneliti mengumpulkan informasi dan literatur dalam penelitian ini dari berbagai sumber dan referensi yang telah dilakukan sebelumnya seperti jurnal ilmiah dan karya ilmiah yang berkaitan dengan penelitian ini.

**B. Pengumpulan Data**

Pada proses ini peneliti mengumpulkan data melalui PusDatIn DKI Jakarta yang mengumpulkan data di terminal Pulogebang dan data tambahan yang didapatkan melalui sosial media seperti WhatsApp, Twitter, Telegram dan media lainnya.

**C. Preprocessing Data**

Preprocessing merupakan proses perapihan data menjadi lebih terstruktur dan menjadi lebih rapih[14]. Maka dari itu *Preprocessing* sangat diperlukan dalam algoritma clustering.

**D. Labeling Data**

Labeling data merupakan proses yang dilakukan untuk menandakan suatu data dan diberikan label sebelum melanjutkan ke dalam proses *machine learning*.

**E. Data Testing**

Data testing merupakan sekumpulan data yang memiliki label yang dapat digunakan untuk menguji ketepatan model dalam mengkalsifikasikan data[15]. Pada proses testing atribut pada label pada data testing disembunyikan selama proses klasifikasi berlangsung. Pada data testing bertujuan untuk melihat keakurasian dan melihat performancenya.

**F. Data Training**

Data training merupakan sekumpulan data yang sudah mempunyai label/kelas yang digunakan machine learning untuk mengenal karakter dari data dan dapat menghasilkan sebuah model data[12]. Pada bagian data training dataset yang dimiliki dilatih untuk membuat prediksi atau menjalankan fungsi dari sebuah algoritma *machine learning* pada algoritma naïve bayes dan KNN. Dalam data training ditujukan untuk melatih dapat mencari korelasinya sendiri.

**G. Algoritma K-Nearst Neighbor (K-NN)**

Algoritma KNN merupakan metode untuk melakukan pengklasifikasi terhadap objek berdasarkan data testing dan data training yang jaraknya mendekati dengan object tujuan[11]. KNN melakukan pencarian kelompok object k dalam data training yang mendekati dengan object tujuan pada data baru atau yang akan datang [16]. Pada penelitian ini penggunaan KNN digunakan untuk mencari rute teramai dan tersepi pada rute perjalanan dengan keberangkatan terminal Pulogebang.

**H. Algoritma Naive Bayes**

Algoritma naïve bayes merupakan salah satu algoritma klasifikasi yang dapat digunakan untuk dokumen teks[17]. Pada penelitian ini dengan menggunakan algoritma *naïve bayes* peneliti dapat mengelompokkan nilai positif dan negatif pada setiap bus penumpang yang menjadi akan menjadi acuan dalam peningkatan layanan bus penyedia layanan.

Algoritma Naïve Bayes memiliki kelebihan dibandingkan algoritma lain diantaranya sederhana, cepat, dan memiliki akurasi tinggi. Dalam penelitian ini menggunakan algoritma *naïve bayes* untuk menyimpulkan hasil positif dan negatif pada review masyarakat terhadap fasilitas yang diberikan oleh pendiri PO Bus di Indonesia.

Pada metode *naïve bayes* memiliki dua tahapan proses klasifikasi teks yang pertama tahap training dan kedua tahap klasifikasi[18]. Teorema Bayes menyatakan tentang probabilitas bersyarat menyatakan:

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \tag{2}$$

Keterangan:

P(H|X) : Probabilitas hipotesis H berdasarkan kondisi X (*posteriori probability*)

Dan untuk mencari keakurasian pada penelitian ini maka menggunakan perhitungan pada data testing. Nilai akurasi yang didapatkan dengan perhitungan berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \tag{3}$$

Keterangan:

- TP : *True positif*
- TN : *True negative*
- FP : *False positif*
- FN : *False negative*

**IV. HASIL DAN ANALISA**

Pada penelitian ini memberikan informasi dan acuan untuk perusahaan otobus yang akan membuka rute perjalanan baru dengan memperhatikan hasil yang akan disimpulkan ramai atau sepi dan hasil dari nilai positif negative dari implementasi algoritma naïve bayes sebagai acuan bus mana yang akan menjadi contoh untuk fasilitasnya.

Pada tahap awal yaitu melihat data yang akan di proses pada penelitian ini. Berikut merupakan dataset yang akan digunakan pada penelitian ini.

BULAN	TAHUN	NAMA_PO	KODE_PO	TUJUAN	KODE_KOTA	PNP
0	10	2020	EKA SARI LORENA TRANSPORT	22	BADUNG	1001 0
1	11	2020	EKA SARI LORENA TRANSPORT	22	BADUNG	1001 13
2	12	2020	EKA SARI LORENA TRANSPORT	22	BADUNG	1001 7
3	12	2020	EKA SARI LORENA TRANSPORT	22	BADUNG	1001 5
4	12	2020	KRAMATDJATI	44	BADUNG	1001 5

Gambar. 2 Tampilan Dataset PusDatIn

Data dari Pusdatin yang akan digunakan dalam metode KNN untuk menentukan nilai ramai dan sepi. Selanjutnya menampilkan data yang akan digunakan untuk algoritma *naïve bayes*.

Waktu	Nama_Lengkap	Alamat	Bus_Pilihan	Tujuan_Asal	Review	Fasilitas	Masakan	Tambahan_Rute	Rute_Tambahan	Keputusan
2021/11/30 8:05:54 AM GMT+7	Kassilto	Harjamukti Cirebon	Dewi Siti	Cirebon Jakarta	Nyaman aman dan terjangkau	AC, bantal	Rumah makan	Ya	Jakarta - Dongkal - Belik	5
2021/11/30 8:20:26 AM GMT+7	Retmon	Perum GPA id 19 ngijo karang pesisir malang	Gunung haris	Jawa Bali	Lebih nyaman	Tempat duduk lebih besar	Tarif	Ya	Jogya malang	5
2021/11/30 8:33:50 AM GMT+7	Rey	Bangka Baling	Bus Pariklata	Bangka Belitung - Sumatra Selatan	Karena membantu dalam bepergian jika mau ke lu.	Tempat duduk Bantal AC	Perlu disoroti menjadi lebih bagus dan nyaman	Tidak	Solo-Jakarta	3
2021/11/30 8:54:13 AM GMT+7	Karyadi Muhammad Rizca	Dusun karang sambung desa jebid selatan kecamatan.	Dewi sri	Pemalang Jakarta	Karena sudah nyaman dan pelayanannya bagus	AC	Perumahan body	Ya	Jakarta - soto	5
2021/11/30 8:56:41 AM GMT+7	Agus Ifan mustika	Tangerang	PO Dewi Siti	Tangerang-pemalang	Langganan dari kecil	AC	Service penumpang	Ya	Jakarta-jogjakarta	5

Gambar. 3 Tampilan Dataset hasil Survey

Dataset ini akan digunakan untuk algoritma naïve bayes sebagai acuan nilai positif dan negatif.

Setelah data diketahui maka langkah selanjutnya kita melakukan labeling data pada jumlah penumpang.

	BULAN	TAHUN	KODE_PO	KODE_KOTA	PNP	CLASS
0	10	2020	22	1001	0	Sepi
1	11	2020	22	1001	13	Sepi
2	12	2020	22	1001	7	Sepi
3	12	2020	22	1001	5	Sepi
4	12	2020	44	1001	5	Sepi
5	2	2020	2	1002	22	Sepi
6	3	2020	2	1002	25	Sepi
7	8	2020	2	1002	3	Sepi
8	9	2020	2	1002	15	Sepi
9	10	2020	2	1002	28	Ramai
10	11	2020	2	1002	31	Ramai
11	12	2020	2	1002	17	Sepi
12	10	2020	6	1002	2	Sepi
13	11	2020	6	1002	5	Sepi
14	11	2020	6	1002	1	Sepi

Gambar. 4 Tampilan Labeling Data

Setelah melakukan labeling peneliti mulai melakukan *data testing* dan *data training* sehingga menghasilkan nilai akurasi seperti pada Gambar 5.

[[ 515 924]					
[ 487 3736]]					
	precision	recall	f1-score	support	
	Ramai	0.51	0.36	0.42	1439
	Sepi	0.80	0.88	0.84	4223
	accuracy			0.75	5662
	macro avg	0.66	0.62	0.63	5662
	weighted avg	0.73	0.75	0.73	5662

Gambar. 5 Hasil Akurasi Data Test

Setelah mendapatkan hasil test dari KNN selanjutnya menentukan hasil dari algoritma naïve bayes untuk menjadikan acuan nilai positif dan negatif pada bus pesaing.

Pada tahap ini selanjutnya peneliti mencari nilai positif dari masing-masing bus yang tersedia pada survey masyarakat untuk menentukan seberapa banyak nilai positif dan negatif pada salah satu bus yang ada di Indonesia. Peneliti melakukan *tokenizer* untuk menghilangkan kata yang tidak diperlukan pada survey dan mengatur stopword. Pada tahap selanjutnya peneliti melakukan *lowercase* untuk membersihkan kata dari huruf yang tidak beraturan dan memberikan nilai kolom baru bernama "Clean\_Review" sebagai hasilnya.

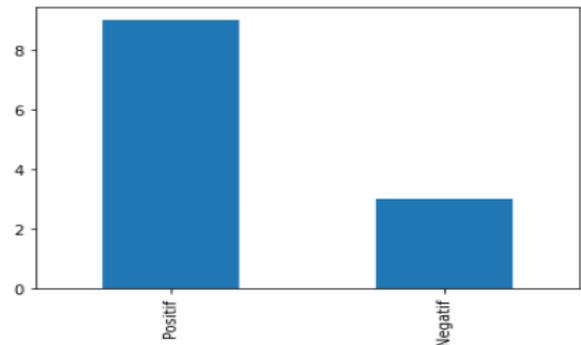
	Waktu	Nama_Lengkap	Alamat	Bus_Pilihan	Tujuan_Asal	Review	Fasilitas	Masukan	Tambahan_Rate	Rate_Tambahan	Kepuasan	Clean_Review
0	2021/11/30 8:09:54 AM GMT+7	Kaswito	Hajarumukti Cirebon	Dewi Sri	Cirebon Jakarta	Nyaman, aman dan terjangkau	AC, bantal	Rumah makan	Ya	Jakarta - Donglal - Belik	5	nyaman aman dan terjangkau
1	2021/11/30 8:20:26 AM GMT+7	Retmon	Pecur GPA id 19 ngglo kang plosok malang	Gunung huta	Jawa Bali	Lebih nyaman	Tempat duduk lebih besar	Tarif	Ya	Jogya malang	5	lebih nyaman
2	2021/11/30 8:33:50 AM GMT+7	Rey	Bangka Belitung	Bus Pariklata	Bangka Belitung - Sumatra Selatan	Karena membantu dalam bepergian jika mau ke lu.	Tempat duduk, Bantal, AC	Parlu direnovasi menjadi lebih bagus dan nyaman	Tidak	Solo-Jakarta	3	karena membantu dalam bepergian jika mau ke lu.
3	2021/11/30 8:54:13 AM GMT+7	Karyadi Muhammad Reza	Dusun Karang Sembung desa jebod seitan kecama.	Dewi sri	Pemalang - Jakarta	Karena sudah nyaman dan pelayanannya bagus	AC.	Pemajaan body	Ya	Jakarta - solo	2	karena sudah nyaman dan pelayanannya bagus

Gambar. 6 Hasil Cleaning Text

Selanjutnya melakukan proses algoritma naïve bayes untuk menentukan nilai positif dan negative pada salah satu bus yang

ada di Indonesia. Dan setelah melakukan proses algoritma maka menghasilkan nilai Positif dan Negative pada salah satu Bus yang dipilih. Dalam penelitian ini mengambil contoh pada Bus "Dewi Sri" yang menghasilkan nilai Positif lebih banyak dan nilai Negatif lebih sedikit.

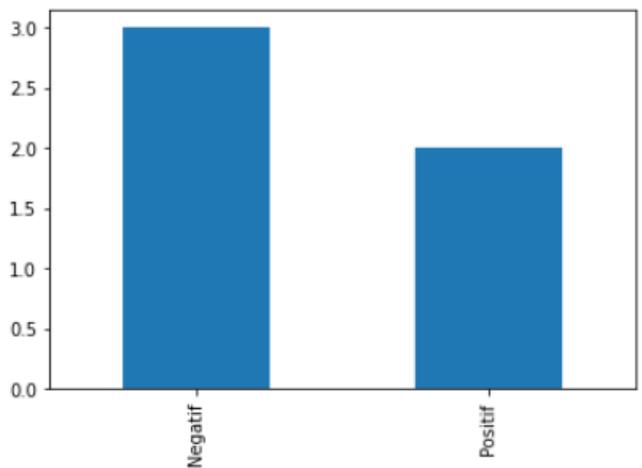
```
newdata['CLASS'].value_counts().plot(kind='bar');
```



Gambar.7 Hasil Nilai PO "Dewi Sri"

Selanjutnya peneliti mengambil sample pada "Sinar Jaya" yang memiliki nilai negatif lebih banyak daripada positif.

```
newdata['CLASS'].value_counts().plot(kind='bar');
```



Gambar.7 Hasil Nilai PO "Sinar Jaya"

Dari penelitian tersebut maka pemilik PO Bus yang akan membuka rute baru maka bisa mengambil referensi yang ada pada penelitian ini baik dari jumlah penumpang maupun dari hasil survey masyarakat.

## V. KESIMPULAN

Dari penelitian ini maka peneliti dapat menyimpulkan dengan menggunakan dua metode algoritma sebagai kolaborasi dalam menentukan rute penumpang ramai dan nilai positif pada bus yang sudah berlalu akan menjadikan referensi baik bagi para pelaku bisnis PO Bus yang ada di Indonesia. Penggunaan Algoritma KNN sangat baik digunakan pada kasus ini karena memiliki akurasi yang cukup tinggi karena lebih dari 60% dan penggunaan naïve bayes dalam klasifikasi nilai positif dan

negative sangat membantu untuk referensi dalam fasilitas yang akan digunakan.

#### DAFTAR PUSTAKA

- [1] S. N. Rahmatunnisa, A. Utami and A. Y. Nurhidayat, "Probabilitas Perpindahan Penumpang Transportasi Massal Berbasis Rel (Studi Kasus Kereta Api Argo Parahyangan Terhadap Kereta Cepat Jakarta – Bandung)," *Ge-STRAM: Jurnal Perencanaan dan Rekayasa Sipil*, vol. 4, no. 2, pp. 91–96, 2021, doi: <https://doi.org/10.25139/jprs.v4i2.4056>.
- [2] R. Setiawan, "Perlindungan Konsumen Transportasi Otobus: Studi Yuridis Kelayakan Transportasi Otobus di Kota Surakarta," *Jurnal Jurisprudence*, vol. 6, no. 2, pp. 100–105, 2013.
- [3] R. K. Dinata, H. Akbar, and N. Hasdyna, "Algoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus," *ILKOM Jurnal Ilmiah*, vol. 12, no. 2, Aug. 2020, doi: <https://doi.org/10.33096/ilkom.v12i2.539.104-111>.
- [4] R. Aryanti, A. Saepudin, E. Fitriani, R. Permana, and Dede Firmansyah Saefudin, "Komparasi Algoritma Naive Bayes Dengan Algoritma Genetika Pada Analisis Sentimen Pengguna Busway," *Jurnal Teknik Komputer*, vol. 5, no. 2, pp. 227–234, Aug. 2019, doi: <https://doi.org/10.31294/jtk.v5i2.5406>.
- [5] Y. A. Singgalen, "Analisis Sentimen dan Pemodelan Topik dalam Optimalisasi Pemasaran Destinasi Pariwisata Prioritas di Indonesia," *Journal of Information Systems and Informatics*, vol. 3, no. 3, pp. 459–470, Oct. 2021, doi: <https://doi.org/10.51519/journalisi.v3i3.171>
- [6] Asroni and R. Adrian, "Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang," *J. Ilm. Semesta Tek.*, vol. 18, no. 1, pp. 76–82, 2015.
- [7] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, pp. 75–82, 2020, doi: [10.31294/ijcit.v5i1.7951](https://doi.org/10.31294/ijcit.v5i1.7951).
- [8] S. F. Diba and J. Nugraha, "Implementation of Naive Bayes Classification Method for Sentiment Analysis on Community Opinion to Indonesian Criminal Code Draft," vol. 474, no. Isstec 2019, pp. 186–192, 2020, doi: [10.2991/assehr.k.201010.027](https://doi.org/10.2991/assehr.k.201010.027).
- [9] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: [10.33365/jtk.v15i1.744](https://doi.org/10.33365/jtk.v15i1.744).
- [10] A. Sari, F. V., & Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd. Id Menggunakan Metode Naive Bayes Classifier Berbasis Konversi Ikon Emosi," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 2, no. 2, pp. 681–686, 2019.
- [11] W. Yustanti, "Algoritma K-Nearest Neighbour untuk Memprediksi Harga Jual Tanah," *J. Mat. Stat. dan komputasi*, vol. 9, no. 1, pp. 57–68, 2012.
- [12] W. Musu, A. Ibrahim, and Heriadi, "Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4 . 5," *Pros. Semin. Ilm. Sist. Inf. Dan Teknol. Inf.*, vol. X, no. 1, pp. 186–195, 2021.
- [13] T. N. P. Dicki Pajri, Yuyun Umidah, "Implementation of K-Nearest Neighbor ( K-NN ) Algorithm For Public Sentiment Analysis of Online Learning," *Tek. Inform. dan Sist. Inf.*, vol. 15, no. 2, pp. 121–130, 2021, doi: [10.22146/ijccs.65176](https://doi.org/10.22146/ijccs.65176).
- [14] F. S. Jumeilah, "Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 1, no. 1, pp. 19 - 25, Jul. 2017.
- [15] S. Novita R, P. Harsani, and A. Qur'ania, "Penerapan K-Nearest Neighbor (KNN) untuk Klasifikasi Anggrek Berdasarkan Karakter Morfologi Daun dan Bunga," *Komputasi*, vol. 15, no. 1, pp. 118–125, 2018, [Online]. Available: <https://journal.unpak.ac.id/index.php/komputasi/article/view/1267>.
- [16] H. Leidiyana, "Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor," *J. Penelit. Ilmu Komputer, Syst. Embed. Log.*, vol. 1, no. 1, pp. 65–76, 2013
- [17] M. H. Rifqo and A. Wijaya, "Implementasi Algoritma Naive Bayes Dalam Penentuan Pemberian Kredit," *Pseudocode*, vol. 4, no. 2, pp. 120–128, 2017, doi: [10.33369/pseudocode.4.2.120-128](https://doi.org/10.33369/pseudocode.4.2.120-128)
- [18] S. Afrizal, H. N. Irmanda, N. Falih, and I. N. Isnainiyah, "Implementasi Metode Naive Bayes untuk Analisis Sentimen Warga Jakarta Terhadap," *Inform. J. Ilmu Komput.*, vol. 15, no. 3, p. 157, 2020, doi: [10.52958/iftk.v15i3.1454](https://doi.org/10.52958/iftk.v15i3.1454)
- [19] K. A. Padhana, "Analisis Sentimen Masyarakat Terhadap Kondisi Perekonomian di Indonesia Pada Masa Pandemi 2020," *Jurnal Ilmu Teknik dan Komputer*, vol. 5, no. 2, pp. 268–276, 2020.