

Perbandingan Algoritma Klasifikasi dalam Analisis Sentimen Opini Masyarakat tentang Kenaikan Harga BBM

Ngabdul Basidt¹, Eko Supriyadi², Agus Susilo³

Prodi Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas An Nuur Purwodadi, Indonesia

DOI: <https://doi.org/10.31331/joined.v3i1.kodeartikel>

Info Articles

Keywords:
K-NN; Naive Bayes;
Opinion;, Support Vector
Machine, , social media,

Abstrak

Kenaikan harga bahan bakar minyak (BBM) telah menjadi permasalahan yang cukup kompleks dan kontroversial . Peningkatan harga BBM memengaruhi berbagai aspek ekonomi dan sosial, termasuk inflasi, biaya produksi, dan tarif transportasi di Indonesia. Klasifikasi sentimen menggunakan algoritma *Naive Bayes*, *Support Vector Machine*, dan *K-Nearest Neighbors* untuk menentukan algoritma klasifikasi sentimen manakah yang terbaik. Dengan melakukan perbandingan metode algoritma *Naive Bayes*, *Support Vector Machine*, dan *K-Nearest Neighbors* untuk menentukan algoritma klasifikasi sentimen manakah yang terbaik. Dengan melakukan perbandingan algoritma klasifikasi sentimen menghasilkan akurasi yang paling tinggi didapatkan oleh algoritma *Naive Bayes* dengan akurasi sebesar 80,28%. Kedua adalah algoritma *Support Vector Machine* (SVM) dengan akurasi sebesar 73,89%. Algoritma yang memiliki nilai akurasi paling kecil adalah algoritma *K-Nearest Neighbor* (KNN) dengan akurasi sebesar 50,00%.

Abstract

The rise in fuel prices has been a very complex and controversial problem. As time went on, people's economies in Indonesia grew worse as they grew grew. . Sentient classification USES the Naive Bayes, Support Vector Machine, and K-Nearest Neighbors to determine which sentiments are the best. the algorithm had the highest accuracy so far as Naive Bayes algorithm with an accuracy of 80.28%. The second is an algorithm Support Vector Machine (SVM) with an accuracy of 73.89%. The algorithm that has the smallest accuracy value is an algorithmic K-Nearest Neighbors (KNN) with 50% accuracy.

✉ Alamat Korespondensi:
E-mail: abdol.basedt54@gmail.com

PENDAHULUAN

Kenaikan harga bahan bakar minyak (BBM) telah menjadi permasalahan yang cukup kompleks dan kontroversial di banyak negara di seluruh dunia [1]. Beberapa negara mengalami kenaikan harga BBM karena fluktuasi harga minyak dunia yang tinggi [2]. Oleh karena itu, beberapa negara memilih untuk menaikkan harga BBM dan mengurangi subsidi pemerintah untuk mengurangi defisit anggaran dan meningkatkan kestabilan ekonomi [3]. Kenaikan BBM dapat mempengaruhi distribusi, transportasi, biaya produksi sehingga mempengaruhi harga-harga barang yang lain. Kebijakan dalam kenaikan harga BBM selalu menjadi fenomena dalam berbagai media yang menyebabkan terjadi pro dan kontra di masyarakat [4]. Banyak media seperti tweet yang menyinggung tentang kenaikan BBM [5].

Di antara tweet-tweet yang di sampaikan oleh masyarakat ada yang bersifat positive, negative, dan netral. Untuk menentukan sifat dari tweet tersebut kita tidak bisa asal dalam melakukannya. Salah satu cara untuk menyelesaikan masalah tersebut menggunakan teknik klasifikasi sentimen dengan menganalisis teks digital untuk menentukan apakah nada emosional pesan tersebut positif, negatif, atau netral [6]. Metode yang digunakan dalam penelitian ini yaitu klasifikasi sentimen analisis. Ada banyak metode klasifikasi yang dapat digunakan, diantaranya metode *Naïve Bayes* [7]., *Support Vector Machine* [8]., dan *K-Nearest Neighbors* [9]. Adapun tujuan dari penelitian ini adalah untuk menentukan suatu kalimat apakah kalimat tersebut bersifat negative atau positive supaya membantu pemerintah untuk membuat keputusan yang lebih baik dan mengurangi dampak kenaikan BBM bagi masyarakat.

METODE

A. Studi Literatur

Text mining merupakan salah satu cabang dalam ilmu komputer yang mempelajari tentang ekstraksi dokumen teks secara otomatis, sehingga informasi yang terdapat di dalamnya dapat dimunculkan [10]. Dalam penerapannya, *text mining* banyak dilakukan untuk kebutuhan marketing, perencanaan produksi, pemasaran digital, dengan sumber datanya berupa teks dari media sosial [11]. Pemanfaatan *text mining* dengan menggunakan data dari media sosial dikarenakan tulisan dan ulasan dari masyarakat memiliki pengaruh di dunia marketing, sehingga dengan mengolah data teks menggunakan teknik *text mining* dapat menghasilkan informasi dan pengetahuan yang bisa dijadikan dasar dalam pengambilan keputusan [12].

Dalam *text mining* terdapat tahap *preprocessing* yang merupakan tahap yang cukup menentukan terhadap hasil akhir performa model yang terbentuk [13]. Adapun tahapan-tahapan dari *praprocessing* yang di lakukan adalah.

1. *Cleaning* merupakan tahapan untuk membersihkan kalimat dari kata yang tidak diperlukan untuk mengurangi noise seperti karakter HTML, RT, ikon emosi, hashtag (#), username (@), url (<http://situs.com>), email (nama@situs.com), simbol dan tanda baca.
2. *case folding* merupakan tahapan merubah bentuk huruf menjadi huruf kecil atau disebut juga penyeragaman bentuk huruf.
3. *Tokenizing* merupakan tahap pemotongan string inputan berdasarkan kata yang menyusunnya.

4. *Stopword Removal* merupakan proses menghilangkan kata-kata yang tidak memiliki arti seperti kata “yang”, “di”, “itu” dan lain sebagainya.
5. *Remove Emotion* merupakan proses penghilangan emoticon pada tweet. Ketika sedang menulis status (tweet) seseorang kadang salah atau kurang tepat dalam penggunaan emoticon, entah disengaja atau tidak banyak yang melakukannya.
6. *Stemming* merupakan proses perubahan kata menjadi kata dasar yang sesuai dengan aturan bahasa Indonesia. Proses stemming menggunakan bantuan algoritma Sastrawi.
7. *Normalisasi bertujuan* mempermudah proses analisis sentimen terhadap entitasentitas dikarenakan banyaknya kata yang tidak baku didalam tweet seperti singkatan, tanggal, jumlah mata uang, dan akronim.

K-NN adalah salah satu algoritma paling sederhana untuk memecahkan masalah klasifikasi. Algoritma ini sering digunakan untuk klasifikasi teks dan data. Pada metode ini dilakukan klasifikasi terhadap obyek berdasarkan data yang jaraknya paling dekat dengan obyek tersebut [14]. Tujuan dari algoritma ini adalah untuk mengklasifikasikan obyek berdasarkan atribut dan training sample. Algoritma ini menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari query instance yang baru. Adapun rumus yang digunakan sebagai berikut :

$$D_{xy} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Keterangan :

- D : jarak terdekat
- S : data training
- E : data testing
- G : jumlah antar atribut individu 1 s/d n.
- F : fungsi similitary atribut i antara kasus X dan kasus Y
- G : Atribut individu antara 1 sampai dengan n Langkah dalam

Naïve Bayes salah satu metode machine learning yang memanfaatkan perhitungan probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, cara kerja *Naïve Bayes* yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya [15]. Adapun rumus dasar dari *naïve bayes* adalah :

$$P(A|B) = (P(B|A) * P(A)) / P(B) \tag{1}$$

Keterangan:

Peluang kejadian A sebagai B ditentukan dari peluang B saat A, peluang A, dan peluang B. Pada pengaplikasiannya nanti rumus pada persamaan (1) berubah menjadi persamaan (2).

$$P(C_i|D) = (P(D|C_i) * P(C_i)) / P(D) \tag{2}$$

Naïve Bayes atau bisa disebut sebagai *Multinomial Naïve Bayes* merupakan model penyederhanaan dari Metoda Bayes yang cocok dalam pengklasifikasian teks atau dokumen. Pada persamaan (3) merupakan persamaan model penyederhanaan dari Metoda Bayes.

$$V_{MAP} = \arg \max P(V_j | a_1, a_2, \dots, a_n) \quad (3)$$

Berdasarkan persamaan (3), maka persamaan (1) dapat ditulis seperti yang terdapat pada persamaan(4)

$$V_{MAP} = \arg \max (V_j \in V) P(a_1, 2, \dots, a_n | P(V_j) | P(a_1, a_2, \dots, a_n)) \quad (4)$$

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*) [16]. 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 linear maupun non linear.

Penelitian terkait dengan klasifikasi sudah banyak dilakukan oleh para peneliti sebelumnya, baik untuk data teks [17], data citra atau gambar [18], serta data terstruktur [5]. Klasifikasi merupakan proses untuk menemukan sekumpulan model yang dijelaskan kelas-kelas data, sehingga model tersebut dapat digunakan untuk memprediksi nilai suatu kelas yang belum diketahui pada sebuah objek [2].

Pada penelitian [19] membahas tentang pemakaian algoritma *Naive Bayes Classifier*, KNN dan *Decision Tree* untuk Klasifikasidaerah kemacetan dikota Jakarta. Dari penelitian ini menghasilkan *Naive Bayes* akurasi sebesar 80%, Pada metode KNN akurasi sebesar 80%, dan pada metode *Decision Tree* akurasi sebesar 100.

Pada penelitian [20] membahas tentang mengimplementasikan algoritma KNN (*K - Nearest Neighbor*) dalam analisis sentimen pengguna Twitter tentang topik Pilkada DKI 2017. Dari penelitian ini menghasilkan Nilai K-NN akurasi terbesar adalah 67,2% ketika k=5, presisi tertinggi 56,94% ketika k=5, dan recall 78,24% dengan k=15.

Pada penelitian [16] membahas tentang pemakaian algoritma *Naive Bayes Classifier* (NBC) Dan *Support Vector Machine* (SVM) untuk menganalisis tweet yang membicarakan tentang isu public yaitu pilkada 2020 di tangan Pandemic COVID-19. Dari penelitian ini *Naive Bayes* memiliki performansi yang baik untuk melakukan klasifikasi tweet dengan tingkat akurasi 92,2%..

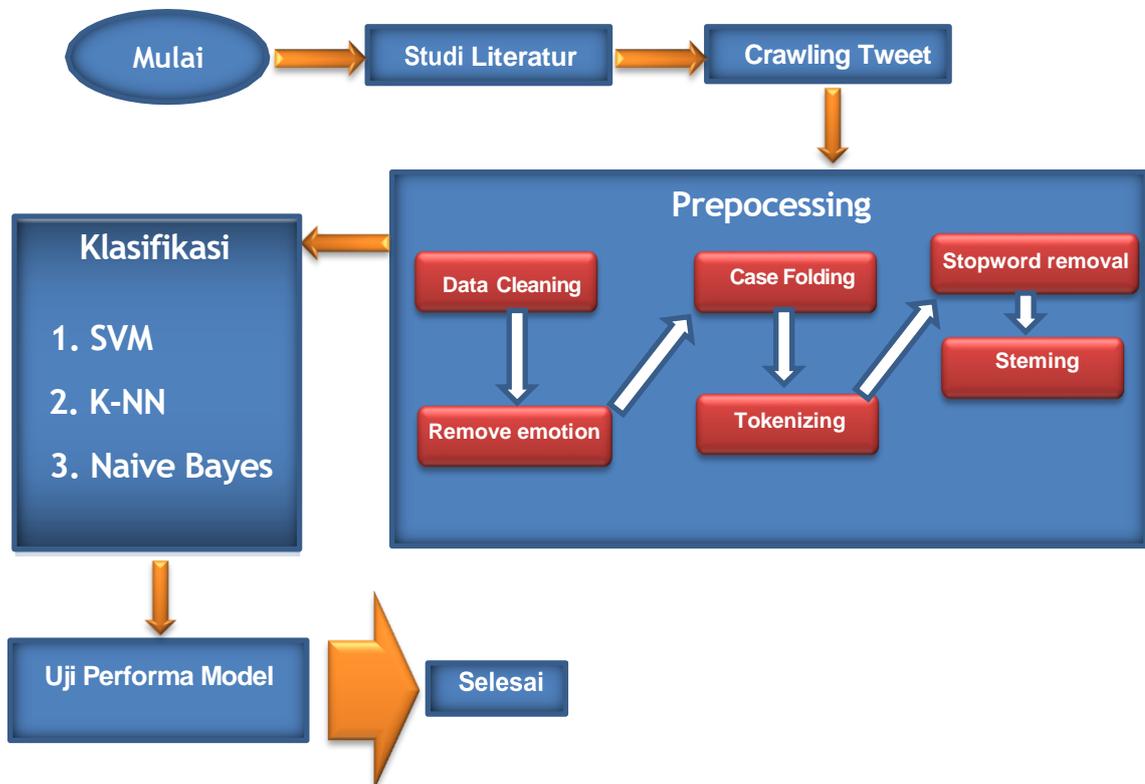
Dari beberapa penelitian yang telah dilakukan peneliti sebelumnya, belum ditemukan perbandingan dokumen teks dengan menggunakan algoritma *Naive Bayes*, KNN dan *Support Vector Machine*.

B. Database

Data yang digunakan dalam penelitian ini merupakan data yang diambil dari Twitter melalui Search API dengan parameter query “BBM naik”. Proses pengumpulan data dilakukan pada rentang bulan september 2022 sampai desember 2022 dengan database berjumlah 3.856 data.

C. Metodologi Penelitian

Metode penelitian ini dimulai dari studi literatur yang dilanjutkan dengan pengambilan dataset. Dataset diambil dengan cara crawling data twitter dengan melalui Search API dengan parameter query “BBM naik” seperti yang telah disebutkan pada poin sebelumnya. Dari 3.856 data yang telah dikumpulkan yang bisa digunakan hanya 1.280 data dikarenakan tidak semua data antara bulan September 2022 hingga Desember 2022 dapat dikategorikan sebagai komentar yang bisa diklasifikasikan sebagai tweet positif atau negatif, seperti contoh tweet tentang promo atau tweet yang digunakan sebagai syarat mendapatkan give away. Data yang telah diambil diolah terlebih dahulu di tahap *preprocessing*. Tahap *preprocessing* ini dilakukan dengan beberapa proses, dimulai data *cleaning*, *Remove Emotion*, *case folding*, *tokenizing*, *stopword removal*, *stemming* dan *normalisasi*. Setelah *preprocessing* selesai dilakukan, maka kita akan lanjut ketahap selanjutnya yaitu pengujian model dari database tersebut dengan menggunakan algoritma klasifikasi yaitu algoritma *naive bayes*, *K-NN*, dan *Support Vector Machine (SVM)* kemudian dilihat performa model dengan mencari nilai Cross Validation. Semakin kecil nilai Cross Validation maka semakin baik pula performa model klasifikasi yang terbentuk. Metode penelitian ini dapat dilihat pada gambar 1 berikut :



Gambar 1. Metode Penelitian

Gambar 1 di atas menunjukkan alur metode penelitian yang dilakukan pada penelitian ini. Selanjutnya dilakukan proses penelitian dan dibahas serta memunculkan hasil seperti yang dijelaskan pada bab selanjutnya.

HASIL DAN PEMBAHASAN

Pengumpulan database yang telah dilakukan dengan cara *crawling* pada sosial media berbasis teks twitter mendapatkan 1.280 dataset. Agar dataset ini siap untuk dilakukan pengolahan, perlu perlakuan awal atau tahap *preprocessing* yaitu data *cleaning*, *Remove Emotion*, *case folding*, *tokenizing*, *stopword removal*, *stemming* dan *normalisasi*. Tahap *preprocessing* ini dapat dilihat pada tabel 1 berikut ini.

Tabel 1. Tahap Prapreprocessing dataset

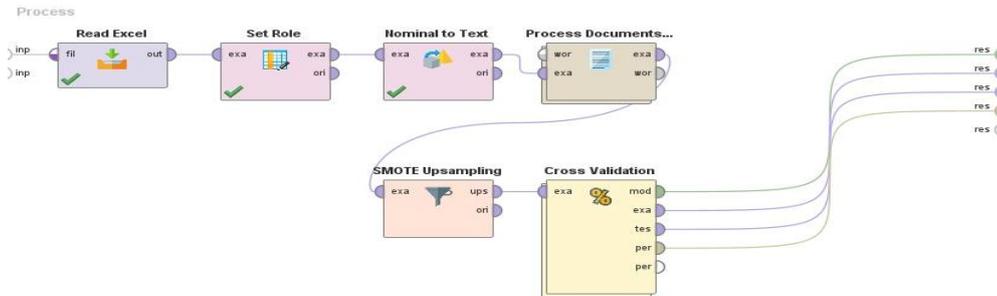
Data Awal	Data Cleaning	Remove Emotion	Case Folding	Tokenizing	Stopword Removal	Stemming	Normalisasi
1. IniSleman Ya paling 1-2 detik Kalo antrinya 15 motor berarti total beda 30-45 detik. kalo saya masih maklum daripada banyak bensin tumpah sia?????. Tapi nek nyetater n naik motor tanpa memajukan dulu itu emang bikin nyesek karena biasanya lama.	1. IniSleman Ya paling 1-2 detik. Kalo antrinya 15 motor berarti total beda 30-45 detik. kalo saya masih maklum daripada banyak bensin tumpah sia?????. Tapi nek nyetater n naik motor tanpa memajukan dulu itu emang bikin nyesek karena biasanya lama.	1. IniSleman Ya paling 1-2 detik. Kalo antrinya 15 motor berarti total beda 30-45 detik. kalo saya masih maklum daripada banyak bensin tumpah sia?????. Tapi nek nyetater n naik motor tanpa memajukan dulu itu emang bikin nyesek karena biasanya lama.	1. inisleman ya paling 1-2 detik. kalo antrinya 15 motor berarti total beda 30-45 detik. kalo saya masih maklum daripada banyak bensin tumpah sia?????. tapi nek nyetater naik motor tanpa memajukan dulu itu emang bikin nyesek karena biasanya lama.	1. Inisleman ya paling 1-2detik kalo antrinya 15motor berarti total beda 30-45 detik Kalo saya masih maklum daripada banyak bensin tumpah Sia tapi nek nyetater n naik motor tanpa memajukan dulu itu emang bikin Nyesek karena biasanya lama	1. Contoh nyata Harga hp bukanya turun Malah naik bensin 3. Inisleman paling 1-2detik antrinya 15motor berarti total beda 30-45 detik masih maklum daripada banyak bensin tumpah motor tanpa memajukan dulu bikin Nyesek biasa lama	1. sleman paling detik antrinya motor berarti total beda detik masih maklum daripada banyak bensin tumpah nyetater naik motor tanpa maju dulu bikin Nyesek biasa lama	1. sleman paling detik antrinya motor berarti total beda detik masih maklum daripada banyak bensin tumpah nyetater naik motor tanpa maju dulu bikin Nyesek biasa lama
2. Contoh nyata adlh harga hp bukannya	2. contoh nyata adlh harga hp bukannya turun	2. contoh nyata adlh harga hp bukannya turun malah naik,	2. contoh nyata adlh harga hp bukannya turun malah naik,	2. contoh nyata adlh harga	2. Contoh nyata Harga hp	2. Contoh nyata Harga hp	2. Contoh nyata Harga hp

Turun malah naik, pantek. bensin aplg	malah naik, pantek. bensin aplg	pantek. bensin aplg	pantek. bensin aplg	hp bukannya turun malah naik pantek bensin aplg	bukanya turun Malah naik bensin	bukan turun Malah naik bensin	bukan turun Malah naik bensin
---	---------------------------------------	------------------------	------------------------	--	---	---	---

3.	3.	3.	3.	3.	3.	3.	3.
1.280.	1.280.	1.280.	1.280.	1.280.	1.280.	1.280.	1.280.

Setelah tahap *preprocessing* dilakukan, dilanjutkan dengan ketahap selanjutnya yaitu pengujian model dari database tersebut dengan menggunakan algoritma klasifikasi yaitu algoritma *naive bayes*, *K-NN*, dan *Support Vector Machine (SVM)*. dalam pengujian model peneliti menggunakan Rapidminer untuk menguji database tersebut.

Pengujian permodelan pertama yang di lakukan oleh peneliti adalah algoritma *Support Vector Machine (SVM)* dengan menggunakan *K-fold cross-validation*.



Gambar 2. Pemodelan Uji Klasifikasi

Gambar 2 adalah pemodelan yang sudah dibuat oleh peneliti. Adapun hasil dari pengujian algoritma *Support Vector Machine (SVM)* sebagai berikut :

● Table View ○ Plot View

accuracy: 73.89% +/- 1.80% (micro average: 73.89%)

	true POSITIF	true NEGATIF	class precision
pred. POSITIF	870	440	66.41%
pred. NEGATIF	30	460	93.88%
class recall	96.67%	51.11%	

Gambar 3. Hasil Uji SVM

Gambar 3 adalah gambaran hasil performa yang di dapatkan oleh algoritma *Support Vector Machine (SVM)*. Dari gambar tersebut kita dapat melihat bahwa algoritma *Support Vector Machine (SVM)* memiliki accuracy sebanyak 73,89% . Nilai

recall masing-masing kelas sebesar 96,67% untuk kelas positif dan 51, 11% untuk kelas negatif. Sedangkan nilai presisi masing-masing kelas adalah sebesar 66,41% untuk kelas positif dan 93,88% untuk kelas negatif. Nilai akurasi ini disebabkan adanya beberapa data yang salah diklasifikasikan oleh sistem yaitu 30 data positif diklasifikasikan data negatif dan 440 data negatif diklasifikasikan menjadi data positif.

Pengujian model Algoritma yang kedua adalah Algoritma *K-Nearest Neighbor* (KNN) dengan permodelan yang sama dengan Algoritma *Support Vector Machine* (SVM). Adapun hasil dari pengujian Algoritma *K-Nearest Neighbor* (KNN) sebagai berikut :

Table View Plot View

accuracy: 50.00% +/- 0.00% (micro average: 50.00%)

	true POSITIF	true NEGATIF	class precision
pred. POSITIF	900	900	50.00%
pred. NEGATIF	0	0	0.00%
class recall	100.00%	0.00%	

Gambar 4. Hasil Uji K-NN

Gambar 4 adalah gambaran hasil performa yang di dapatkan oleh Algoritma *K-Nearest Neighbor* (KNN). Dari gambar tersebut kita dapat melihat bahwa Algoritma *K-Nearest Neighbor* (KNN) memiliki accurasy sebanyak 50,00% . Nilai recall masing-masing kelas sebesar 100,00% untuk kelas positif dan 0,00% untuk kelas negatif. Sedangkan nilai presisi masing-masing kelas adalah sebesar 50,00% untuk kelas positif dan 0,00% untuk kelas negative.

Pengujian model Algoritma yang terakhir adalah Algoritma *Naive Bayes*. hasil dari pengujian Algoritma *Naive Bayes* sebagai berikut :

Table View Plot View

accuracy: 80.28% +/- 3.08% (micro average: 80.28%)

	true POSITIF	true NEGATIF	class precision
pred. POSITIF	856	311	73.35%
pred. NEGATIF	44	589	93.05%
class recall	95.11%	65.44%	

Gambar 5. Hasil Uji Naive Bayes

Gambar 5 adalah gambaran hasil performa yang di dapatkan oleh algoritma *Naive Bayes*. Dari gambar tersebut kita dapat melihat bahwa algoritma *Naive Bayes* memiliki accurasy sebanyak 80,28% . Nilai recall masing-masing kelas sebesar 95,11% untuk kelas positif dan 65,44% untuk kelas negatif. Sedangkan nilai presisi masing-masing kelas adalah sebesar 73,35% untuk kelas positif dan 93,05% untuk kelas negatif.

Nilai akurasi ini disebabkan adanya beberapa data yang salah diklasifikasikan oleh sistem yaitu 44 data positif diklasifikasikan data negatif dan 311 data negatif diklasifikasikan menjadi data positif.

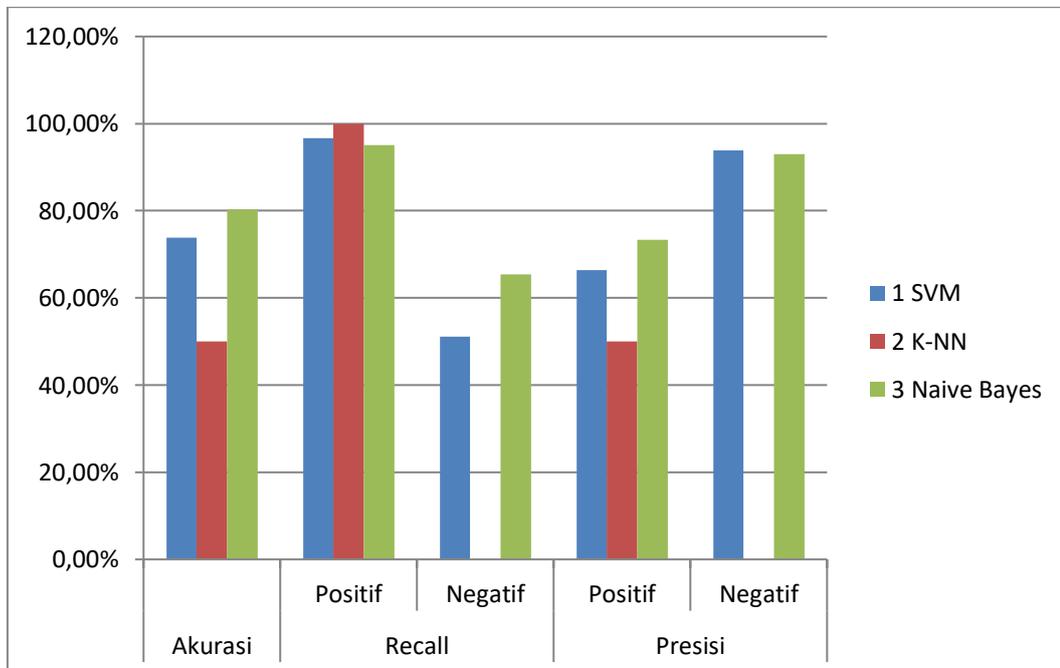
Berdasarkan hasil imlementasi masing-masing algoritma klasifikasi, adapun hasil nilai performance semua algoritma ditunjukkan pada tabel berikut ini :

Tabel 2. Hasil Nilai Peforma Semua Algoritma

No	Algoritma	Akurasi	Recall		Presisi	
			Positif	Negatif	Positif	Negatif
1	SVM	73,89%	96,67%	51,11%	66,41%	93,88%
2	K-NN	50,00%	100,00%	0,00%	50,00%	0,00%
3	Naive Bayes	80,28%	95,11%	65,44%	73,35%	93,05%

Dari Tabel 2 tersebut terlihat masing-masing algoritma memiliki nilai akurasi yang tidak sama. Akurasi yang paling tinggi didapatkan oleh algoritma *Naive Bayes* dengan akurasi sebesar 80,28%. Kedua adalah algoritma *Support Vector Machine* (SVM) dengan akurasi sebesar 73,89%. Algoritma yang memiliki nilai akurasi paling kecil adalah algoritma *K-Nearest Neighbor* (KNN) dengan akurasi sebesar 50,00%.

Tabel 3. Hasil Nilai Peforma Semua Algoritma



SIMPULAN

Model klasifikasi yang terbentuk dari penelitian ini pada 1.280 dataset sampel dengan menggunakan algoritma *naive bayes*, *K-NN*, dan *Support Vector Machine* (SVM). Pada Algoritma *Support Vector Machine* (SVM) didapatkan akurasi sebesar 73,89%, Algoritma *K-Nearest Neighbor* (KNN) didapatkan akurasi sebesar 50,00% Dan Algoritma *Naive Bayes* didapatkan akurasi sebesar 80,28%. Kesimpulan lain yang didapatkan pada penelitian ini menunjukkan bahwa Algoritma klasifikasi yang memiliki tingkat Akurasi yang paling tinggi adalah Algoritma *Naive Bayes* dengan Akurasi sebesar 80,28% dan Algoritma klasifikasi yang memiliki tingkat Akurasi yang paling rendah adalah Algoritma *K-Nearest Neighbor* (KNN) didapatkan akurasi sebesar 50,00%. Algoritma *K-Nearest Neighbor* (KNN) kurang baik diterapkan, terutama untuk data teks yang singkat seperti yang didapatkan melalui platform media sosial Twitter.

DAFTAR PUSTAKA

- [1] N. Normah, "Naïve Bayes Algorithm For Sentiment Analysis Windows Phone Store Application Reviews," *Sinkron*, vol. 3, no. 2, p. 13, 2019, doi: 10.33395/sinkron.v3i2.242.
- [2] D. Goldblatt, "ANALISIS KOMPARASI ALGORITMA KLASIFIKASI DATA MINING UNTUK PREDIKSI MAHASISWA NON AKTIF," *Semin. Neurol.*, vol. 14, no. 1, pp. 241–249, 1994, doi: 10.2307/j.ctv11hppt6.3.
- [3] H. Parveen and S. Pandey, "Sentiment analysis on Twitter Data-set using Naive Bayes algorithm," *Proc. 2016 2nd Int. Conf. Appl. Theor. Comput. Commun. Technol. iCATccT 2016*, pp. 416–419, 2017, doi: 10.1109/ICATCCT.2016.7912034.
- [4] N. Tri Romadloni, I. Santoso, and S. Budilaksono, "Perbandingan Metode Naive Bayes, Knn Dan Decision Tree Terhadap Analisis Sentimen Transportasi Krl Commuter Line," *J. IKRA-ITH Inform.*, vol. 3, no. 2, pp. 1–9, 2019.
- [5] W. T. Handoko, E. Supriyanto, D. I. Purwadi, Z. Budiarmo, and H. Listiyono, "Klasifikasi Opini Pengguna Media Sosial Twitter Terhadap JNT Di Indonesia dengan Algoritma Decision Tree," *J. Sains Komput. Inform.*, vol. 6, no. 2, pp. 790–799, 2022.
- [6] K. Jasa, "Bab 2 Landasan Teori," *Apl. dan Anal. Lit. Fasilkom UI*, vol. m, no. 1998, pp. 7–34, 2000, [Online]. Available: <http://elib.unikom.ac.id/files/disk1/655/jbptunikompp-gdl-supriadini-32740-6-12.unik-i.pdf>
- [7] S. I. Sari, W. Safitri, and R. D. P. Utami, "Pengaruh Pendidikan Kesehatan Dengan Metode Demonstrasi Terhadap Praktik Pertolongan Pertama Luka Bakar Pada Ibu Rumah Tangga Di Garen Rt.01/Rw.04 Pandean Ngemplak Boyolali," *J. Kesehat. Kusuma Husada*, pp. 98–105, 2018, doi: 10.34035/jk.v9i1.266.
- [8] M. Ichwan, I. A. Dewi, and Z. M. S, "Klasifikasi Support Vector Machine (SVM)

- Untuk Menentukan Tingkat Kemanisan Mangga Berdasarkan Fitur Warna,” *MIND J.*, vol. 3, no. 2, pp. 16–23, 2019, doi: 10.26760/mindjournal.v3i2.16-23.
- [9] G. A. Rosso, “Milton,” *William Blake Context*, no. September, pp. 184–191, 2019, doi: 10.1017/9781316534946.021.
- [10] K. V. S. Toy, Y. A. Sari, and I. Cholissodin, “Analisis Sentimen Twitter menggunakan Metode Naive Bayes dengan Relevance Frequency Feature Selection (Studi Kasus: Opini Masyarakat mengenai Kebijakan New Normal),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 11, pp. 5068–5074, 2021, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [11] S. R. I. Rezeki, “Penggunaan sosial media twitter dalam komunikasi organisasi (studi kasus pemerintah provinsi DKI Jakarta dalam penanganan covid-19),” *J. Islam. Law Stud.*, vol. 04, no. 02, pp. 63–78, 2020.
- [12] M. M. Rybarczyk *et al.*, “Revue systématique des cas de blessures par brûlure dans les pays à revenu faible et intermédiaire: épidémiologie dans la région africaine de l’OMS,” *African J. Emerg. Med.*, vol. 7, no. 1, pp. 30–37, 2017, doi: 10.1016/j.afjem.2017.01.006.
- [13] F. Handayani and S. Pribadi, “Implementasi Algoritma Naive Bayes Classifier dalam Pengklasifikasian Teks Otomatis Pengaduan dan Pelaporan Masyarakat melalui Layanan Call Center 110,” *J. Tek. Elektro*, vol. 7, no. 1, pp. 19–24, 2015.
- [14] A. Herdhianto, *Sentiment Analysis Menggunakan Naive Bayes Classifier (NBC) Pada Tweet Tentang Zakat*. 2020. [Online]. Available: <http://repository.uinjkt.ac.id/dspace/handle/123456789/53661>
- [15] 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.
- [16] V.A.R.Barao, R.C.Coata, J.A.Shibli, M.Bertolini, and J.G.S.Souza, “ANALISIS SENTIMEN OPINI MASYARAKAT PENGGUNA TWITTER TERHADAP PARIWISATA LAMPUNG MENGGUNAKAN SUPPORT VECTOR MACHINE DAN NAIVE BAYES,” *Braz Dent J.*, vol. 33, no. 1, pp. 1–12, 2022.
- [17] Yuli Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD) . Jurnal Edik Informatika,” *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2019.
- [18] M. C. Wujaya and L. W. Santoso, “Klasifikasi Pakaian Berdasarkan Gambar Menggunakan Metode YOLOv3 dan CNN,” *J. INFA*, vol. 9, no. 1, pp. 2–7, 2021.
- [19] N. Tri Romadloni, I. Santoso, S. Budilaksono, and M. Ilmu Komputer STMIK Nusa Mandiri Jakarta, “PERBANDINGAN METODE NAIVE BAYES, KNN

DAN DECISION TREE TERHADAP ANALISIS SENTIMEN
TRANSPORTASI KRL COMMUTER LINE.”

- [20] A. Heriyanto, “Penerapan Metode K-Nearest Neighbor (KNN) Untuk Klasifikasi Stanting Pada Balita,” *Publ. Ilm. Univ. Muhammadiyah Jember*, 2021.