

# Klasifikasi Sentimen Masyarakat Terhadap Pemberlakuan Pembatasan Kegiatan Masyarakat Menggunakan *Text Mining* Pada Twitter

Asdar Mustofa<sup>1</sup>, Rice Novita<sup>2</sup>

Sains dan Teknologi, Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: <sup>1</sup>11850314529@students.uin-suska.ac.id, <sup>2</sup>rice.novita@uin-suska.ac.id

Email Penulis Korespondensi: 11850314529@students.uin-suska.ac.id

Submitted: 08/06/2022; Accepted: 25/06/2022; Published: 30/06/2022

**Abstrak**—Corona Virus Disease 2019 (Covid-19) saat ini menjadi suatu pandemi yang ada di dunia, termasuk di Indonesia. Berbagai kebijakan telah dilakukan untuk memutus rantai penyebaran Covid-19 salah satunya kebijakan yang dilakukan pemerintah adalah Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM). PPKM menjadi salah satu topik yang banyak dibicarakan di sosial media termasuk Twitter. Tweets di Twitter yang diberikan masyarakat terhadap kebijakan PPKM yang diselenggarakan memunculkan sentimen, jika diolah dengan baik dan benar dapat menjadi bahan evaluasi penerapan PPKM, maka perlu dilakukan klasifikasi sentimen masyarakat menggunakan text mining, pada penelitian ini menggunakan algoritma K-Nearest Neighbor (KNN) dan Naïve Bayes Classifier (NBC) dengan data dari ciutan Twitter selama penyelenggaraan PPKM setahun terakhir dengan 3.516 data. Dimana didapatkan hasil akurasi bahwa algoritma NBC lebih baik daripada algoritma KNN dengan akurasi 79,67% berbanding 78,86%, didapatkan juga polaritas sentimen masyarakat terhadap PPKM dengan sentimen positif sebesar 36,83% dengan jumlah 1.295, sentimen netral tweets 54,15% dengan jumlah 1.902 tweets, dan sentimen negatif 9,02% dengan jumlah 317 tweets.

**Kata Kunci:** Covid-19; klasifikasi; text mining; PPKM; KNN; NBC; Twitter

**Abstract**—Corona Virus Disease 2019 (Covid-19) is currently a pandemic in the world, including in Indonesia. Various policies have been carried out to break the chain of the spread of Covid-19, one of which is the government's policy of implementing Community Activity Restrictions (PPKM). PPKM is one of the most discussed topics on social media, including Twitter. Tweets on Twitter given by the public to the PPKM policy that was held to evaluate the implementation of PPKM, it is necessary to classify public sentiment using text mining, in this study using the K-Nearest Neighbor (KNN) and Naïve Bayes Classifier (NBC) algorithms with data from tweets. Twitter during the PPKM last year with 3,516 data. Where the results are that the NBC algorithm is better than the KNN algorithm with an accuracy of 79.67% compared to 78.86%, the polarity of public sentiment towards PPKM is also obtained with positive sentiment of 36.83% with a total of 1,295, neutral sentiment of tweets 54.15% with the number of 1,902 tweets, and 9.02% negative sentiment with a total of 317 tweets.

**Keywords:** Covid-19; classification; text mining; PPKM; KNN; NBC; Twitter

## 1. PENDAHULUAN

Sosial media merupakan sarana yang efektif serta efisien untuk berbagi informasi kepada masyarakat umum, sosial media juga dapat memberikan dampak yang signifikan baik pada perubahan negatif maupun positif, sehingga sosial media juga memungkinkan komunikasi terbuka dari latar belakang yang berbeda dengan pemangku kepentingan yang berbeda [1]. Masyarakat Indonesia terkenal sebagai pengguna paling aktif bermedia sosial di dunia [2]. Sosial media yang banyak digunakan dikalangan masyarakat dewasa ini oleh masyarakat Indonesia salah satunya adalah Twitter [3].

Twitter ialah salah satu contoh sosial media yang sedang populer saat ini, Twitter menggunakan konsep *microblogging* yang memiliki *user* lebih dari 500 juta dan 400 *tweets* setiap harinya [4]. *Microblogging* seperti Twitter digunakan secara publik untuk memberikan pendapat maupun melakukan penilaian mengenai segala hal dan mengunggah postingan dengan berbagai opini melalui *tweets* [5]. Pengguna Twitter saat ini seperti orang dewasa, remaja, dan bahkan anak-anak.

Kelebihan sosial media untuk mengkomunikasikan informasi terbaru maupun informasi penting, sosial media sangat berperan sebagai alur penyebaran informasi bagi masyarakat umum disegala bidang [6], sosial media bisa menjadi sistem sensor bagi masyarakat umum dengan menggunakannya. Hal tersebut dapat terjadi dikarenakan media sosial memiliki *hashtag* atau tagar tertentu, *longitude*, dan memberikan informasi lokasi dimana postingan tersebut dibuat. Salah satu yang menjadi banyak perbincangan di media sosial termasuk Twitter saat ini ialah mengenai Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) yang diterapkan di seluruh kota di Indonesia.

Diselenggarakannya PPKM agar dapat membantu mengurangi laju dari kenaikan dari angka positif kasus *Corona Virus Disease 2019* (Covid-19) [7]. Sejak pandemi Covid-19 di Indonesia, pemerintah telah menetapkan bermacam istilah untuk menghadapi Covid-19. Hingga bulan April 2020, pemerintah memakai setidaknya ada tujuh penamaan yang berbeda-beda. Mulai dari pembatasan sosial berskala besar (PSBB) dari sejak April 2020 hingga yang terbaru yaitu PPKM hingga Juli 2021 [8]. PPKM diterapkan berdasarkan Intruksi Mendagri No.15 Tahun 2021 mengenai PPKM Darurat Covid-19 [9]. Dengan diselenggarakannya PPKM ini berdampak pada kalangan masyarakat, seperti: tutupnya pusat-pusat pemberlanjaan (1), bangkrutnya para pedagang pasar (2), sulitnya kondisi hotel dan restoran (3), dan jatuhnya omzet pedagang kuliner (4) [10]. Hal ini menimbulkan opinin-opini masyarakat yang ada terhadap kebijakan yang diterapkan, menurut mereka pemberlakuan PPKM yang dilakukan mulai dari pusat sampai ke daerah malah menyengsarakan masyarakat dan dianggap tidak berdampak terhadap pengendalian Covid-19 [11].

Penerapan PPKM ini juga banyak memunculkan mural-mural dinding seperti, “Tuhan Aku Lapar”, “404: not found”, “Dipaksa sehat di negara yang sakit” [12]. Opini-Opini juga banyak disampaikan media sosial seperti pada cuitan *tweets* yang rata-rata mengandung teks. Hal-hal semaca ini seringkali disebut dengan sentimen. Jika diolah dengan baik dan benar, maka dapat menjadi suatu informasi yang bermanfaat. Data-data berbentuk teks tersebut dapat diolah menggunakan *text mining* yang mana sebagai teknik mengekstrak informasi dari sekumpulan data yang tidak terstruktur berkualitas tinggi serta diperoleh data-data masalah-masalah dalam teks topik tertentu [13].

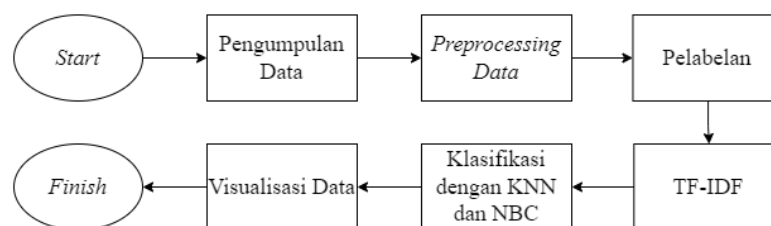
Pada penelitian ini menerapkan dua algoritma klasifikasi yaitu *K-Nearest Neighbor* (KNN) dan *Naïve Bayes Classifier* (NBC). Algoritma ini metode klasifikasinya sangat sederhana dan sangat berguna untuk mengklasifikasikan data baru berdasarkan atribut dan pola pelatihan yang dihasilkan dari titik pelatihan yang paling dekat dengan pertanyaan [14]. Selanjutnya algoritma NBC merupakan algoritma klasifikasi statistik yang digunakan untuk memprediksi probabilitas anggota suatu kelas [15]. Kelebihan dari algoritma ini adalah pada dataset yang berukuran besar dapat mengurangi data *noise* [16] seperti data cuitan pada Twitter.

Berdasarkan penelitian yang dilakukan oleh Hidayat dkk (2019), penelitian ini melakukan pengklasifikasian opini masyarakat terhadap toko *online* menggunakan KNN dan PNN (*Probabilistic Neural Network*) pada Twitter. Didapatkan hasil akurasi yang menunjukkan KNN lebih baik dibandingkan PNN dimana KNN memiliki nilai akurasi lebih tertinggi yaitu 71,57% pada data Lazada dan 66,71% pada data Blibli dibandingkan dengan metode PNN yang bernilai akurasi sebesar 68,29% pada data Lazada dan 65,29% pada data Blibli [17]. Kemudian penelitian mengenai Analisis Sentimen Masyarakat Terhadap PPKM Menggunakan Algoritma NBC pada *software* RStudio, dengan Hasil berupa hingga 99 sentimen positif, hingga 1% sentimen negatif, atau dapat dilihat dari hasil klasifikasi sentimen. Sebagian besar “tidak diketahui” atau tidak diketahui, dan dominan polaritas positif [18]. Kemudian penelitian mengenai Analisis Sentimen Masyarakat menggunakan Algoritma NBC dan *Particle Swarm Optimization*, dengan hasil akurasi sebesar 77,16% dari 302 data *tweets* [19]. Kemudian peneliti mengenai Sentimen Anallysis of User Review on Covid-19 Infomartion Application Using NBC, *Support Vector Machine*, and KNN didapatkan hasil diantaranya akurasi NBC lebih baik dibandingkan dengan KNN dengan akurasi 72,3% berbanding 59,1% [20]. Selanjutnya penelitian mengenai Penerapan Algoritma NBC dan PNN Untuk Klasifikasi Nasabah Bank Dalam Membayar Kredit, didapatkan hasil PNN memiliki nilai akurasi yang lebih baik dibandingkan dengan NBC, 93,58% berbanding 89,90% [16].

Penelitian ini bertujuan mendapatkan sentimen masyarakat terhadap kebijakan PPKM dan dapat menjadi bahan evaluasi pemerintah dalam penerapan kebijakan PPKM, serta mendapatkan hasil perbandingan algoritma KNN dan NBC dengan data yang didapatkan melalui *crawling data* Twitter dengan kata kunci “PPKM” selama setahun terakhir penyelenggaraan PPKM.

## 2. METODOLOGI PENELITIAN

Adapun metodologi penelitian dalam penelitian ini secara garis besar terdiri dari 6 tahapan yang dilakukan. Berikut adalah tahap-tahap dalam penelitian yang dapat dilihat pada Gambar 1.



**Gambar 1.** Metodologi Penelitian

Tahap awal pada penelitian ini dengan mengumpulkan data yang di dapat dengan metode *crawling*. Data yang di kumpulkan adalah data *tweet* pada Twitter dengan kata kunci “PPKM” dengan jumlah data yaitu 3.516 data. Selanjutnya dilakukan *preprocessing data*, karena data yang di dapatkan adalah data yang tidak terstruktur. Data yang sudah melalui tahapan *preprocessing data* selanjutnya dilakukan pelabelan untuk untuk menentukan sentimen, pelabelan ini dilakukan secara manual oleh tiga orang yaitu, Nuratika, S.Hum, M.Pd., Dina Hartanti Hermawan, S.Pd., dan Wahyudi, S.Pd.. Selanjutnya dilakukan pembobotan kata menggunakan *Term Frequency-Invers Document Frequency* (TF-IDF) untuk mendapatkan bobot nilai pada setiap kata. Pada proses klasifikasi pada penelitian ini menggunakan KNN dan NBC. Tahapan terakhir pada penelitian ini adalah melakukan analisis hasil klasifikasi dari algoritma yang digunakan, mendapatkan visualisasi kata, dan hasil perbandingan algoritma.

### 2.1 Text Mining

*Text mining* ialah sebagai teknik mengekstrak informasi dari sekumpulan data tidak terstruktur berkualitas tinggi serta diperoleh data-data masalah-masalah dalam teks topik tertentu [13]. *Text Mining* dapat menemukan informasi penting dari sumber data dengan mengidentifikasi dan memeriksa pola tertentu [21].

## 2.2 Algoritma *K-Nearest Neighbor* (KNN)

KNN merupakan salah satu metode yang memakai algoritma *supervised* dengan hasil dari *query instance* yang baru diklasifikasikan yang didapat dari data pembelajaran terdekat [22]. KNN termasuk dalam grup pembelajaran berbasis instans. Algoritma ini juga merupakan teknik pembelajaran yang malas. JST dapat dilakukan dengan mencari sekelompok *k* objek pada *data training* yang paling dekat dengan objek pada *data testing* atau data baru.

## 2.3 Algoritma *Naïve Bayes Classifier* (NBC)

NBC mempunyai tingkat akurasi yang tinggi, yaitu akurat, cepat, dan akurat ketika digunakan dalam database yang besar. Algoritma ini dapat mengklasifikasikan pohon keputusan dan jaringan saraf [23]. NBC menggunakan dua tahapan dalam proses klasifikasi, yaitu tahapan *trainig* dan tahapan *classification*. Pada tahapan *classification*, nilai kategori salah satu dokumen ditentukan didasarkan atas istilah-istilah yang muncul pada dokumen rahasia tersebut [24].

# 3. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini didapatkan dan diambil dari data *tweets* pada Twitter dengan jumlah data 3.516 data, pengambilan data pada Twitter dilakukan dengan teknik *crawling* menggunakan bahasa pemrograman *python* versi 3.6.0.. Data yang digunakan yaitu data *tweets* dengan kata kunci yang dipakai yaitu “PPKM” dalam rentang waktu April 2021 sampai Desember 2021. Selanjutnya data tersebut akan di proses di *preprocessing data* sebelum diolah menggunakan algoritma Klasifikasi KNN dan NBC, karena data tersebut masih berbentuk data tidak terstruktur. Selanjutnya dilakukan pelabelan dan TF-IDF untuk memberikan bobot nilai pada data-data tersebut. Kemudian dilakukan klasifikasi menggunakan algoritma KNN dan NBC.

### 3.1 Data Awal

Pada tahapan ini, data penelitian yang digunakan adalah data *tweets* yang didapatkan dari *crawling data tweets* periode April 2021 sampai Desember 2021. Berikut adalah data awal yang digunakan, dapat dilihat pada Tabel 1.

Tabel 1. Data Awal

No	username	Tweet
1	dkijakarta	(2/2) Jalankan terus protokol kesehatan 3M. Jangan keluar rumah bila tdk mendesak. Selalu gunakan masker dgn benar dan jaga jarak bila harus keluar rumah. Sering cuci tangan dgn sabun. Ingatkan sesama warga. #JagaJakarta #HadapiBersama #COVID19 #PSBBJakarta #PPKM <a href="https://t.co/ChBrZS7r36">https://t.co/ChBrZS7r36</a>
2	dkijakarta	[TERBARU] Penanganan #COVID19 di Jakarta. (1/2) Update data tes PCR/antigen, kasus dan vaksinasi DKI Jakarta 4 April 21. Strategi tes, lacak dan isolasi terus digencarkan untuk temukan sebanyaknya kasus positif agar tidak menularkan virus. #JagaJakarta #PSBBJakarta #PPKM <a href="https://t.co/9dgQoNr05U">https://t.co/9dgQoNr05U</a>
...	...	...
3515	pikobar_jabar	Ketahui informasi pembagian #PPKM di wilayah Jabar berdasarkan level 2 dan 1 di #PikoData <a href="https://t.co/qsYrXdHOPA">https://t.co/qsYrXdHOPA</a>
3516	redaksisuma	Polda Lampung mengingatkan para bupati dan wali kota menerapkan kebijakan PPKM Mikro untuk mencegah penyebaran covid-19. #poldalampung #ppkm #ppkmlampung #ppkmmikro #sumaid <a href="https://t.co/N7nrgwsZHh">https://t.co/N7nrgwsZHh</a>

### 3.2 Preprocessing

Pada tahapan *preprocessing* merupakan kegiatan untuk menyiapkan data mentah sebelum masuk ke proses selanjutnya. *Preprocessing* dapat dilakukan dengan menghapus data yang tidak terstruktur (*unsupervised*) atau mengubahnya menjadi format suatu data yang lebih mudah untuk diolah atau diproses oleh sistem.

#### 3.2.1 Case Folding

*Case folding* merupakan tahapan pertama dalam preprocessing, *case folding* dilakukan karena data awal yang didapat masih beragam tidak hanya teks. Pada tahapan *case folding* dilakukan penghapusan semua karakter menjadi huruf dan mengubah semua jenis teks maupun huruf menjadi huruf kecil. Hasil dari *case folding* dapat dilihat pada Tabel 2.

Tabel 2. Hasil Data Case Folding

No	Tweet
1	(2/2) jalankan terus protokol kesehatan 3m. jangan keluar rumah bila tdk mendesak. selalu gunakan masker dgn benar dan jaga jarak bila harus keluar rumah. sering cuci tangan dgn sabun. ingatkan sesama warga. #jagajakarta #hadapibersama #covid19 #psbbjakarta #ppkm <a href="https://t.co/chbrzs7r36">https://t.co/chbrzs7r36</a>

- 2 [terbaru] penanganan #covid19 di jakarta. (1/2) update data tes pcr/antigen, kasus dan vaksinasi dki jakarta 4 april 21. strategi tes, lacak dan isolasi terus digencarkan untuk temukan sebanyakny kasus positif agar tidak menularkan virus. #jagajakarta #psbbjakarta #ppkm <https://t.co/9dgqonr05u>
- ...
- 3515 ketahui informasi pembagian #ppkm di wilayah jabar berdasarkan level 2 dan 1 di #pikodata <https://t.co/qsyrdhopa>
- 3516 polda lampung mengingatkan para bupati dan wali kota menerapkan kebijakan ppkm mikro untuk mencegah penyebaran covid-19. #poldalampung #ppkm #ppkmlampung #ppkmmikro #sumaid <https://t.co/n7nrgwszhh>

### 3.2.2 Cleaning

Data *tweets* yang diproses dalam format teks terlebih dahulu melalui proses *cleaning*, dikarenakan pada data *tweets* dengan bahasa Indonesia masih banyak menggunakan simbol, kalimat maupun kata yang belum baku, angka, hashtag (#), link url, *username* maupun data kalimat yang *duplicate*. Hasil dari *data cleaning* dapat dilihat pada Tabel 3.

Tabel 3. Hasil *Data Cleaning*

No	Tweet
1	jalankan terus protokol kesehatan m jangan keluar rumah bila tdk mendesak selalu gunakan masker dgn benar dan jaga jarak bila harus keluar rumah sering cuci tangan dgn sabun ingatkan sesama warga
2	terbaru penanganan di jakarta update data tes pcr antigen kasus dan vaksinasi dki jakarta april strategi tes lacak dan isolasi terus digencarkan untuk temukan sebanyakny kasus positif agar tidak menularkan virus
...	...
3515	ketahui informasi pembagian di wilayah jabar berdasarkan level dan di
3516	polda lampung mengingatkan para bupati dan wali kota menerapkan kebijakan ppkm mikro untuk mencegah penyebaran covid

### 3.2.3 Filtering

*Filtering* merupakan pengambilan kata-kata yang penting dari sebuah kalimat untuk membuang atau menghilangkan kata-kata atau kalimat tidak penting seperti tanda baca maupun *stopword* [25]. Daftar kata yang digunakan adalah *stoplist* yang berjumlah 758 *stopwords*. Beberapa *stopword* yang digunakan dapat dilihat pada Tabel 4.

Tabel 4. Data Kata *Stopword*

No	Stopword	No	Stopword	No	Stopword	No	Stopword
1	ada	6	akhiri	11	antaranya	16	atas
2	adalah	7	akhirnya	12	apa	17	atau
3	adanya	8	aku	13	apaan	18	ataukah
4	adapun	9	akulah	14	apabila	...	...
5	akhir	10	antara	15	asalkan	758	yang

Hasil dari *preprocessing* dapat dilihat pada Tabel 5.

Tabel 5. Hasil *Preprocessing*

No	Tweet
1	jalankan protokol kesehatan m rumah tdk mendesak masker dgn jaga jarak rumah cuci tangan dgn sabun ingatkan warga
2	terbaru penanganan jakarta update data tes pcr antigen vaksinasi dki jakarta april strategi tes lacak isolasi digencarkan temukan sebanyakny positif menularkan virus
...	...
3515	ketahui informasi pembagian wilayah jabar berdasarkan level
3516	polda lampung bupati wali kota menerapkan kebijakan ppkm mikro mencegah penyebaran covid

### 3.2.4 Stemming

Pada tahap *stemming* merupakan tahap menrubah kata yang berimbuhan menjadi akar kata. Pada tahap ini dilakukan dengan mencari akar kata pada setiap kata menggunakan modul sastrawi pada pyhton. Hasil dari *stemming* dapat dilihat pada Tabel 6.

Tabel 6. Hasil *Stemming*

No	Stemming
1	jalan protokol sehat m rumah tdk desak masker dgn jaga jarak rumah cuci tangan dgn sabun ingat warga
2	baru tangan jakarta update data tes pcr antigen vaksinasi dki jakarta april strategi tes lacak isolasi gencar temu banyak positif tular virus

...  
 3515 tahu informasi bagi wilayah jabar dasar level  
 3516 polda lampung bupati wali kota terap bijak ppkm mikro cegah sebar covid

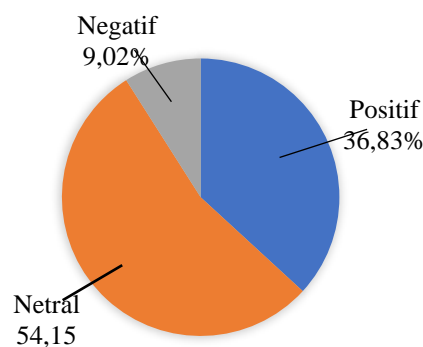
### 3.3 Pelabelan

Data *tweets* terhadap PPKM yang tersedia adalah *unsupervised* data atau yang tidak berlabel. Sehingga diperlukan suatu metode untuk dilakukan pelabelan pada data *tweets* tersebut. Pada penelitian ini dilakukan pelabelan data *tweets* terhadap PPKM dilakukan secara manual oleh tiga orang pakar, yaitu Nuratika, S.Hum, M.Pd., Dina Hartanti Hermawan, S.Pd., dan Wahyudi, S.Pd.

Hasil dari pelabelan data *tweets* adalah pada kelas positif yaitu 1295 *tweets*, pada kelas netral yaitu 1904 *tweets*, dan pada kelas negatif yaitu 317 *tweets*. Berikut adalah hasil dari pelabelan data *tweets* terhadap PPKM dapat dilihat pada Tabel 7 dan Gambar 2.

Tabel 7. Hasil Pelabelan

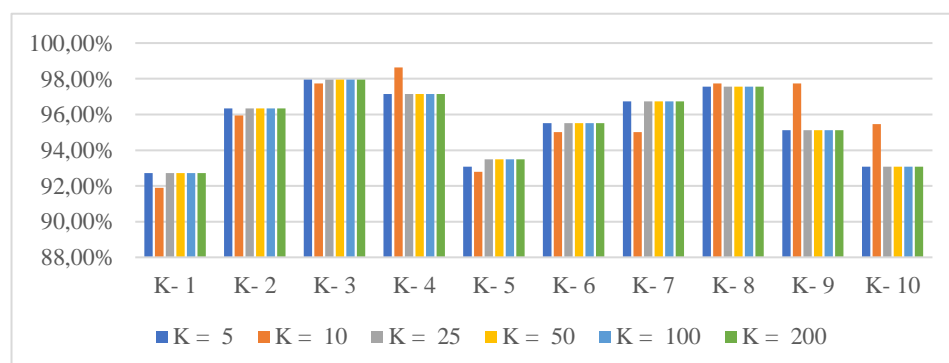
No	Tweet	Sentimen
1	jalankan terus protokol kesehatan m jangan keluar rumah bila tdk mendesak selalu gunakan masker dgn benar dan jaga jarak bila harus keluar rumah sering cuci tangan dgn sabun ingatkan sesama warga	Positif
2	terbaru penanganan di jakarta update data tes pcr antigen kasus dan vaksinasi dki jakarta april strategi tes lacak dan isolasi terus digencarkan untuk temukan sebanyaknya kasus positif agar tidak menularkan virus	Positif
...	...	...
3515	ketahui informasi pembagian di wilayah jabar berdasarkan level dan di	Positif
3516	polda lampung mengingatkan para bupati dan wali kota menerapkan kebijakan ppkm mikro untuk mencegah penyebaran covid	Positif



Gambar 2. Pelabelan Sentimen

### 3.4 Algoritma *K-Nearest Neighbor* (KNN)

Pada penelitian ini digunakan 10 split *K-Fold Cross Validation* pada pembagian *data training* dan *data testing* untuk menghasilkan sebuah akurasi yang akan dibentuk. Data yang digunakan berjumlah 3.516 data set dari data *tweets*. Setelah dibagi menjadi 10 k, maka nantinya akan dilakukan sebanyak enam kali percobaan menggunakan algoritma KNN dengan parameter K yang memiliki nilai berbeda-beda yaitu k = 5, k = 10, k = 25, k = 50, k = 100, dan k = 200. Hasil dari pembagian *data training* dan *data testing* menggunakan 10 *K-Fold Cross Validation* pada KNN dengan parameter k = 5, k = 10, k = 25, k = 50, k = 100, dan k = 200 dapat dilihat pada Gambar 3 dan Tabel 8.



Gambar 3. Hasil Analisis *K-Fold*

**Tabel 8.** Hasil dari Pembagian *Data Training* dengan *K-Fold Cross Validation* pada NBC

K-Fold	Parameter K					
	K = 5	K = 10	K = 25	K = 50	K = 100	K = 200
K- 1	69,63%	68,82%	67,61%	67,20%	65,99%	65,99%
K- 2	78,86%	78,45%	78,04%	78,04%	79,67%	78,04%
K- 3	62,19%	67,88%	66,66%	68,29%	66,66%	67,07%
K- 4	53,25%	48,37%	45,93%	39,83%	42,68%	42,27%
K- 5	55,28%	58,53%	56,91%	56,50%	58,13%	58,13%
K- 6	61,78%	69,10%	66,26%	68,69%	69,10%	69,10%
K- 7	62,60%	62,60%	65,44%	67,07%	67,47%	64,47%
K- 8	41,05%	44,71%	66,66%	67,88%	67,88%	67,88%
K- 9	62,19%	65,44%	63,82%	66,66%	66,26%	66,66%
K- 10	72,35%	67,47%	69,10%	69,51%	68,69%	68,69%
Nilai Max	78,86%	78,45%	78,04%	78,04%	79,67%	78,04%

Berdasarkan dari Gambar 3 dan Tabel 8 dapat disimpulkan hasil percobaan tertinggi berada pada Parameter k = 10 di *split* K-4 dengan nilai akurasi sebesar 98,64%.

### 3.5 Algoritma *Naïve Bayes Classifier* (NBC)

Pada penelitian ini pengujian 10 *K-Fold Cross Validation* pada NBC menggunakan 3.516 data set dari data *tweets*, pembagian data dibagi menjadi *data training* dan *data testing*. Dengan mengulang k-kali secara *random* untuk dibagi menjadi k himpunan bagian yang paling bebas dengan mengulang k-kali, setiap perulangan dibagi menjadi himpunan pelatihan dan himpunan. Hasil dari pembagian *data training* dan *data testing* menggunakan 10 *K-Fold Cross Validation* pada NBC dapat dilihat pada Tabel 9.

**Tabel 9.** Hasil dari Pembagian *Data Training* dengan *K-Fold Cross Validation* pada NBC

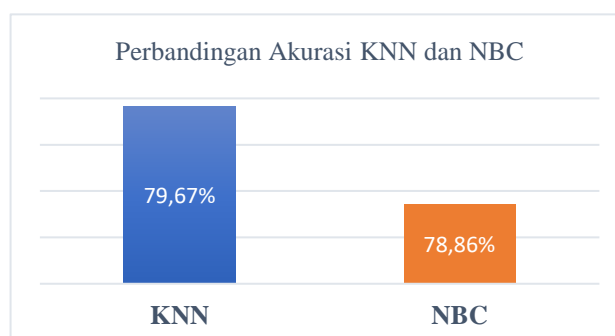
K-Fold	Akurasi
K-1	70,85%
K-2	78,86%
K-3	71,13%
K-4	45,52%
K-5	56,09%
K-6	67,07%
K-7	62,19%
K-8	65,04%
K-9	70,32%
K-10	70,73%
Nilai Max	78,86%

Berdasarkan dari Tabel 11 dapat disimpulkan bahwa hasil percobaan tertinggi berada pada *split* K-3 dengan nilai akurasi sebesar 78,86%.

### 3.6 Analisis dan Hasil Klasifikasi

#### 3.6.1 Perbandingan Akurasi

Percobaan pembagian data menggunakan metode 10 *K-Fold Cross Validation*, didapatkan hasil perbandingan akurasi antara KNN dan NBC terhadap data *tweets*. Dimana hasil akurasi dari KNN lebih baik dibandingkan dengan NBC, KNN dengan akurasi sebesar 79,67% berbanding 78,86%. Perbandingan akurasi KNN dan NBC dapat dilihat pada Gambar 4.

**Gambar 4.** Hasil Perbandingan Akurasi KNN dan NBC



## 4. KESIMPULAN

Berdasarkan dari rangkaian setiap tahapan yang dilakukan menggunakan bahasa *python* pada data *tweets* pada Twitter dengan jumlah 3.516 data mengenai tanggapan masyarakat terhadap kebijakan PPKM, diperoleh hasil yaitu perbandingan akurasi antara NBC dan KNN dengan akurasi terbaik yaitu KNN 79,67% sedangkan NBC 78,86% dengan pembagian data menggunakan 10 *K-Fold Cross Validation*. Selanjutnya didapatkan polaritas sentimen masyarakat terhadap PPKM dengan sentimen positif sebesar 36,83% dengan jumlah 1.295, sentimen netral *tweets* 54,15% dengan jumlah 1.902 *tweets*, dan sentimen negatif 9,02% dengan jumlah 317 *tweets*. Didapatkan juga kemunculan kata dengan 3 frekuensi tertinggi yaitu “Covid”, “Masyarakat”, dan “Kota”. Pada kelas positif kemunculan kata dengan 3 frekuensi kata tertinggi yaitu “Ppkm”, “Kota”, dan “Covid”. Pada kelas netral kemunculan kata dengan 3 frekuensi kata tertinggi yaitu “Ppkm”, “Level”, dan “Panjang”. Pada kelas negatif kemunculan kata dengan 3 frekuensi kata tertinggi yaitu “Ppkm”, “Covid”, dan “Panjang”.

## REFERENCES

- [1] D. R. Rahadi, “Perilaku Pengguna Dan Informasi Hoax Di Media Sosial,” *Jurnal Manajemen Dan Kewirausahaan*, vol. 5, no. 1, pp. 58–70, 2017, doi: 10.26905/jmdk.v5i1.1342.
- [2] Liputan6, “Jumlah Pengguna Instagram dan Facebook Indonesia Terbesar ke-4 di Dunia - Tekno Liputan6.com.” 2019. <https://www.liputan6.com/teknoread/3998624/jumlah-pengguna-instagram-dan-facebook-indonesia-terbesar-ke-4-di-dunia> (accessed Nov. 10, 2021).
- [3] suara.com, “Jumlah Pengguna Media Sosial Indonesia Capai 191,4 Juta per 2022 - Bagian 2.” <https://www.suara.com/teknoread/2022/02/23/191809/jumlah-pengguna-media-sosial-indonesia-capai-1914-juta-per-2022?page=2> (accessed Feb. 27, 2022).
- [4] M. S. Hadna, P. I. Santosa, and W. W. Winarno, “Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter,” *Seminar Nasional Teknologi Informasi dan Komunikasi*, vol. 2016, no. Sentika, pp. 57–64, 2016.
- [5] Hartanto, “Text Mining Dan Sentimen Analisis Twitter Pada Gerakan Lgbt,” *Intuisi : Jurnal Psikologi Ilmiah*, vol. 9, no. 1, pp. 18–25, 2017, doi: 10.15294/intuisi.v9i1.9561.
- [6] Y. Fitriani, “Analisis Pemanfaatan Berbagai Media Sosial sebagai Sarana Penyebaran Informasi bagi Masyarakat,” *Paradigma - Jurnal Komputer dan Informatika*, vol. 19, no. 2, p. 152, 2017.
- [7] Tribunnews.com, “Apa Itu PPKM? Simak Penjelasannya dan Ini Rincian Lengkap Aturannya,” 2021. <https://www.tribunnews.com/corona/2021/07/12/apa-itu-ppkm-simak-penjelasannya-dan-ini-rincian-lengkap-aturannya> (accessed Nov. 10, 2021).
- [8] Nasional Tempo.co, “Gonta-ganti Istilah Penanganan Covid-19: PSBB Hingga Terkini PPKM Level 4 ,” 2021. <https://nasional.tempo.co/read/1486390/gonta-ganti-istilah-penanganan-covid-19-psbb-hingga-terkini-ppkm-level-4/full&view=ok> (accessed Nov. 10, 2021).
- [9] H. Y. P. Sibuea, “Penegakan Hukum Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) Darurat Jawa dan Bali,” *Info Singkat*, vol. XIII, pp. 1–6, 2021.
- [10] detikFinance, “Begini Pahitnya Dampak dari PPKM Darurat - Halaman 2,” 2021. <https://finance.detik.com/berita-ekonomi-bisnis/d-5648999/begini-pahitnya-dampak-dari-ppkm-darurat/2> (accessed Nov. 10, 2021).
- [11] Top Jabar, “Demo Tolak PPKM: Si Miskin Teriak Lapar, Si Kaya Teriak Prokes,” 2021. <https://topjabar.co/2021/07/21/demo-tolak-ppkm-si-miskin-teriak-lapar-si-kaya-teriak-prokes/peristiwa/30543/> (accessed Nov. 10, 2021).
- [12] JawaPos.com, “Graffiti dan Mural: Curahan Hati Rakyat tentang PPKM,” 2021. <https://radarbanyuwangi.jawapos.com/kolom/27/08/2021/graffiti-dan-mural-curahan-hati-rakyat-tentang-ppkm> (accessed Nov. 29, 2021).
- [13] F. Ratnawati, “Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter,” *INOVTEK Polbang - Seri Informatika*, vol. 3, no. 1, p. 50, 2018, doi: 10.35314/isi.v3i1.335.
- [14] G. N. Bagaskoro, M. A. Fauzi, and P. P. Adikara, “Penerapan Klasifikasi Tweets Pada Berita Twitter Menggunakan Metode K-Nearest Neighbor Dan Query Expansion Berbasis Distributional Semantic,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, vol. 2, no. 10, pp. 3849–3855, 2018.
- [15] A. Tarigan, E. Wahyudi, and J. Adhiva, “Klasifikasi Status Kesejahteraan Rumah Tangga di Kabupaten Siak Menggunakan Algoritma Naive Bayes Classifier,” *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI) 11*, no. November, pp. 187–196, 2019.
- [16] T. Qurahman, “Klasifikasi Nasabah Dalam Membayar Kredit Bank Menggunakan Algoritma Naive Bayes Classifier dan Probabilistic Neural Network,” *Universitas Islam Negeri Sultan Syarif Kasim Riau*, 2019.
- [17] A. Hidayat, Mustakim, M. Z. Fauzi, and I. Syukra, “Implementasi Algoritma K-Nearest Neighbor dan Probabilistic Neural Network untuk Analisis Opini Masyarakat Terhadap Toko Online di Indonesia,” *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI) 11*, no. November, pp. 254–261, 2019.
- [18] T. Krisdiyanto, “Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naive Bayes Clasifiers,” *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, vol. 7, no. 1, pp. 32–37, 2021.
- [19] S. D. Pramukti, A. Nugroho, and A. S. Sunge, “Analisis Sentimen Masyarakat Dengan Metode Naive Bayes dan Particle Swarm Optimization,” *Techno.Com*, vol. 21, no. 1, pp. 61–74, 2022, doi: 10.33633/tc.v21i1.5332.
- [20] A. Salma and W. Silfianti, “Sentiment Analysis of User Reviews on COVID-19 Information Applications Using Naive Bayes Classifier , Support Vector Machine , and K-Nearest Neighbor,” vol. 6, no. 4, pp. 158–162, 2021.
- [21] J. A. Pratama, Y. Suprijadi, and Z. Zulhanif, “The Analisis Sentimen Sosial Media Twitter Dengan Algoritma Machine Learning Menggunakan Software R,” *Jurnal Fourier*, 2017, doi: 10.14421/fourier.2017.62.85-89.





- [22] N. D. Mentari, M. A. Fauzi, and L. Muflikhah, “Analisis Sentimen Kurikulum 2013 Pada Sosial Media Twitter Menggunakan Metode K-Nearest Neighbor dan Feature Selection Query Expansion Ranking,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, vol. 2, no. 8, pp. 2739–2743, 2018.
- [23] W. Muslehatin, M. Ibnu, and Mustakim, “Penerapan Naïve Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Sistem Informasi UIN Suska Riau,” *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI)*, pp. 250–256, 2017.
- [24] I. M. B. S. Darma, R. S. Perdana, and Indriati, “Penerapan Sentimen Analisis Acara Televisi Pada Twitter Menggunakan Support Vector Machine dan Algoritma Genetika sebagai Metode Seleksi Fitur,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2018.
- [25] W. Gata, “Akurasi Text Mining Menggunakan Algoritma K-Nearest Neighbour pada Data Content Berita SMS,” vol. 6, pp. 1–13, 2017.