

ANALISIS SENTIMEN PENDAPAT MASYARAKAT TERHADAP PPKM DARURAT PADA MEDIA SOSIAL *TWITTER* MENGUNAKAN METODE *NAÏVE BAYES CLASSIFIER (NBC)*

Angga Aditya Permana¹, Arsanah², Dinar Ajeng Kristiyanti³, Manorang Sihotang⁴

^{1,2} Program Studi Informatika, Fakultas Teknik, Universitas Muhammadiyah Tangerang Jalan Perintis
Kemerdekaan 1/33 Cikokol Kota Tangerang

³ Program Studi Teknologi Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana
Informatika, Jl. Kramat Raya No. 98, Senen, Jakarta Pusat, Indonesia

⁴ Universitas Pramita Indonesia, Jl. Kampus Pramita, Tangerang 15810, Indonesia

Co Responden Email: ¹ anggaamt@gmail.com, ² arsanah.19@gmail.com, ³ dinar@nusamandiri.ac.id ⁴
sihotangparbahal@gmail.com

Abstract

Article history

Received October 1, 2021

Revised October 26, 2021

Accepted November 5, 2021

Available online November 10,
2021

Keywords

analisis sentiment
ppkm darurat,
naïve bayes classifier

The development of information is very fast and advanced, social media can also be used as a very popular source of information and communication tool from various internet users, one of which is Twitter, In government actions regarding emergency PPKM, many people tweet on social media platforms. their twitter, tweets containing criticism, opinions, or suggestions. Then an analysis is carried out to reveal the sentiment of the tweet, using the help of rstudio because with this analysis it will determine the opinion or input from the community including positive, negative, neutral or feeling sentiments, this sentiment is to classify 413 tweet data into training data and 83 tweet data into testing data resulted in twitter tweets about emergency PPKM getting more “neutral” sentiment responses.

Abstrak

Riwayat

Diterima 1 Oktober 2021

Revisi 26 Oktober 2021

Disetujui 5 November 2021

Terbit 10 November 2021

Kata Kunci

Analisis Sentimen,
PPKM Darurat,
naïve bayes classifier

Perkembangan informasi sangat cepat dan maju, pada media sosial pun bisa di jadikan sumber informasi dan alat komunikasi yang sangat populer dari berbagai pengguna internet, salah satu media sosial *twitter*. Dalam tindakan pemerintah mengenai ppkm darurat banyak masyarakat yang *mentweet* pada *platform* media sosial *twitter* mereka, *tweet* berisi kritikan, pendapat, atau pun saran. Maka di lakukan analisis untuk mengungkap sentimen dari *tweet*, menggunakan bantuan *rstudio* karena dengan adanya analisis ini akan menentukan kata pendapat atau masukan dari masyarakat itu termasuk sentimen positif, negatif, netral atau perasaan, sentimen ini untuk mengklasifikasi data *tweet* sebanyak 413 menjadi data *training* dan 83 data *tweet* menjadi data *testing* menghasikan *uitan twitter* tentang ppkm darurat lebih banyak mendapatkan respon sentimen “*netral*”.

PENDAHULUAN

Sekarang sangat mudah memperoleh informasi dan membagikannya, secara cepat dengan adanya internet dan media social, memberikan kemudahan akses bagi pengguna internet untuk menuangkan opini – opini *public* masyarakat mulai dari hasil kinerjanya, tindakan yang telah di lakukan.

Masyarakat indonesia di kenal sangat aktif di

internet dan media sosial untuk mencari dan berbagi informasi, internet juga memiliki jumlah pengguna yang besar menurut *survei Asosiasi Penyelenggara Jasa Internet Indonesia*

(APJII) hingga kuartal II tahun 2020, jumlah pengguna internet Indonesia ada 196,7 juta orang atau 73,7 persen dari total populasi Indonesia 266,91 juta. (Pengguna & Indonesia, 2020). Media sosial memberi pengaruh kuat bagi masyarakat mulai dari pola berpikir, gaya hidup,

memunculkan ide, perdebatan bahkan opini hal itu

sangat menarik perhatian saat mengambil kebijakan dengan cara menganalisis sentimen salah satu media social yaitu *twitter* itu jejaring social atau layanan komunikasi *real time* di rilis pada tahun 2006 hingga sekarang digunakan oleh jutaan orang untuk mengakses, membagikan informasi dengan unggahan tweet. (Statistika et al., 2020)



Gambar 1. Pengguna Internet

TINJAUAN PUSTAKA

Analisis Sentimen adalah salah satu metode yang berfungsi menganalisis opini atau pendapat masyarakat, sentimen, penilaian, perilaku dan emosi melalui layanan umum, organisasi atau pun media sosial, fokus pada kata-kata untuk mendapatkan opini yang mengandung pesan negatif atau positif. (Muzaki, 2021)

Twitter di dirikan oleh *Jack Dorsey* pada bulan maret 2006 sebuah *situs* jejaring media sosial di luncurkan pada juli . Sejak *twitter* di luncurkan telah menjadi salah satu situs yang paling sering di kunjungi di internet dan banyak *user* yang memakai *twitter*. (Zuhdi, 2019)

Twitter API (*Application Programming Interface*) atau antar muka pemograman aplikasi untuk berbagi informasi seluas mungkin di *twitter*, menyediakan akses programatik ke data *twitter* kepada perusahaan, pengembang, dan pengguna lewat *API twitter* mengizinkan akses ke bagian dari layanan melalui *API* untuk memungkinkan orang - orang membangun perangkat lunak yang terintegrasi dengan *twitter* seperti solusi yang membantu sebuah perusahaan menjawab umpan balik pelanggan di *twitter*.

Text preprocessing merupakan tahap awal pada analisis di mana data teks, akan di *preprocessing* ini mencakup tahapan pengambilan data tweet dari *twitter* hanya berupa data mentah yang belum di lakukan *preprocessing* guna mendapatkan data bersih agar memudahkna ke tahap berikutnya,. (Zuhdi, 2019) Pembobotan TF-IDF merupakan metode untuk menghitung bobot setiap kata yang paling banyak digunakan. Metode ini juga terkenal efisien, mudah dan memiliki hasil yang akurat.

Naïve Bayes Classifier

Naïve Bayes Classifier merupakan salah satu metode bertujuan dalam klasifikasi dengan probabilitas sederhana yang mengaplikasikan dengan rumus Metode *Naïve Bayes Classifier* di gunakan karna terlalu banyak dataset yang akan dipakai harus membutuhkan metode yang mempunyai performansi yang cepat dalam pengklasifikasian serta keakuratan yang baik dan tinggi. (Rusdianan & Rosiyadi, 2021) dinyatakan dalam persamaan dengan rumus *Teorema Naïve Bayes* :

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

X : Data dengan class yang belum diketahui
H : Hipotesis data merupakan suatu *class* spesifik

P(H|X) : Probabilitas hipotesis H berdasarkan kondisi X (*posterior probabilitas*)

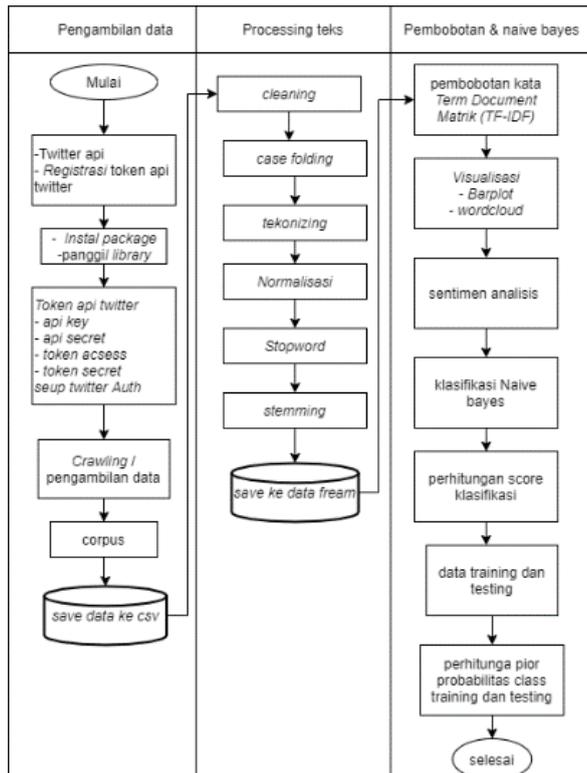
P(H) : Probabilitas hipotesis H (*prior probabilitas*)

P(X|H) : Probabilitas X Berdasarkan kondisi pada hipotesis H

P(X) : Probabilitas X

METODE PENELITIAN

Perancangan alur pengerjaan analisis sentimen menggunakan *rstudio* merupakan gambaran umum terkait alur analisis yang akan di lakukan mulai dari pengambilan data hingga klasifikasi *naïve bayes*.



Gambar 2. Langkah-langkah Penelitian

HASIL DAN PEMBAHASAN

A. Pengambilan Data

Pada proses pengambilan data di lakukan dengan *crawling* menggunakan *software R.3.5.1* untuk melakukan proses *crawling* data pada twitter, Analisis ini mengenai tweet yang berkaitan terhadap ppkm darurat dan hasil dari *crawling* pada media social twitter sebanyak 413 *tweet* data yang dari *tweet* dan data *retweet*. melakukan pemanggilan kode token dan *setup twitter oauth* pada *rstudio* , hasil yang di dapatkan carwling data seperti gambar dibawah ini

B. Preprocessing

Data hasil *crawling* yang masih berupa susunan kata tidak terstruktur dengan memiliki karakter- karakter, tanda baca, penggunaan huruf kapital dan ejaan singkatan yang tidak sesuai dengan bahasa indonesia, kata- kata yang tidak memiliki arti. Dari hal tersebut akan sangat mempengaruhi dalam data yang menjadi sebuah informasi.

C. Pembobotan Trem Frequency dan Inverse

Document Frequency (TF-IDF)

Tahapan untuk memberikan nilai pada setiap kata yang telah di bersihkan sebelumnya di tahapan *preprocessing*. Pada pembobotan term berdasarkan jumlah kemunculan nya dalam data analisis. Semakin besar jumlah kemunculan kata tersebut maka makin tinggi juga jumlah frekuensi kemunculan.

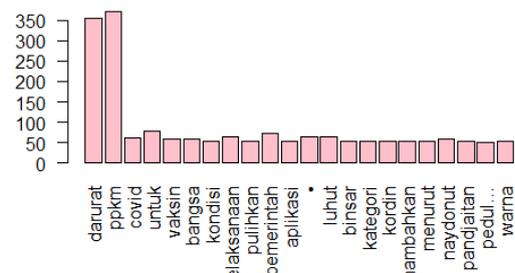
```

> tdm<-as.matrix(tom)
> tdm[1:10, 1:20]
      Docs
Terms 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
ani    1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
darurat 1 1 1 1 1 0 1 1 0 1 1 1 0 1 0 1 1 1 1 1
kek    1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
level  2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
lelur  1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
lockdown 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
odp    1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
pdp    1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
ppkm   3 1 1 1 1 0 1 1 0 1 1 1 0 1 0 1 1 1 1 1
secgroun 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  
```

Gambar 3. Pembobotan

D. Visualisasi Barplot dan Wordcloud

Dari hasil pembobotan *trm* akan menampilkan visualisasi berupa *barplot* dan *wordcloud* sesuai dari pembobotan kata pada *barplot* menunjukkan yang paling tinggi mengartikan kata paling banyak muncul dari kata – kata yang lain



Gambar 4. Visualisasi Barplot Hasil Pembobotan

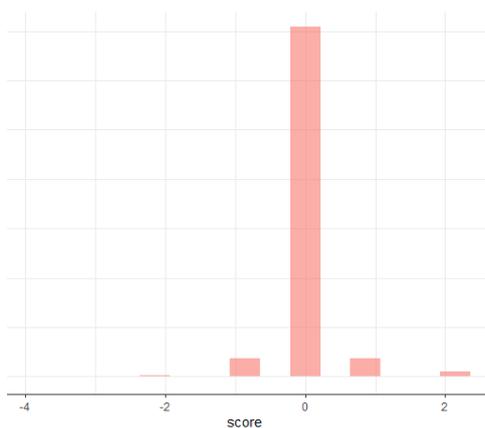


Gambar 5. Visualisasi Wordcloud Hasil Pembobotan

E. Klasifikasi *Naïve bayes Classification*

Dalam mengklasifikasi dengan metode *naïve bayes classification* proses klasifikasi menentukan kata tersebut masuk kedalam klasifikasi model kata positif, negatif, dan netral. Jika dalam klasifikasi data mendapatkan score 0 itu menandakan hasil klasifikasi termasuk netral, artinya kata – kata tersebut lebih banyak netral dari pada kata positif atau negatif. Jika dalam klasifikasi mendapatkan score 1 itu artinya termasuk klasifikasi positif, dan jika score -1 artinya negatif. Setelah itu menentukan data *training* dan data *testing* yang sudah di lakukan tahap *text preprocessing*. Data *training* berjumlah 330 kata dan data *testing* berjumlah 83 kata, yang akan di Selanjutkan dan menghitung *prior probabilitas* nya dalam data *training* dan *testing*. Maka telah di hasilkan bahwa hasil klasifikasi *naïve bayes* lebih banyak menghasilkan klasifikasi netral.

Dari *data testing* dengan teks “suasana hari sama ppkm darurat medan cukup baik” telah di klasifikasi termasuk model “netral” dengan prior probabilitas yang dihasilkan 0.01204819. Perhitungan Manual *Naive bayes*, data diambil dari hasil pembobotan yang memunculkan frekuensi data yang paling banyak muncul.



Gambar 6. Hasil *Score* Pembobotan

KESIMPULAN

Berdasarkan hasil analisis sentimen data twitter yang telah penulis lakukan , berikut merupakan kesimpulan pada analisis sentimen ini :

- Tahapan Analisis sentimen pertama *crawling* , *proprocessing* data, pembobotan term, sentimen, klasifikasi *naïve bayes*, menentukan data *training* dan *testing*, membuat class untuk *prior probabilitas*
- Hasil klasifikasi sentimen dari 413 data tweet menggunakan *algoritma naïve bayes* dengan seleksi fitur *Trem-Frequency* maka di dapatkan jumlah sentimen netral yang lebih dominan di dibandingkan sentimen positif dan netral dan dengan perhitungan score pun lebih banyak bernilai 0 yang artinya data analisis memiliki sentimen netral

REFRENSI

- Ardiani, L., Sujaini, H., & Tursina, T. (2020). Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak. *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)*, 8(2), 44–52. <https://doi.org/20.26428/justin.v8i2.36776>
- Darurat, P., Pilihan, O., & Bertahan, U. (2021). *PPKM Darurat : Onlinenisasi Pilihan UMKM Bertahan. 2020–2022*.
- Dharma Suriyanto, A., & Endah Pangesti, W. (2021). *ANALISA SENTIMEN TERHADAP TAGAR #dirumahaja MELALUI TWITTER DI INDONESIA*. 2(09), 9–27.
- Muzaki, A., & Witanti, A. (2021). Sentiment Analysis of the Community in the Twitter To the 2020 Election in Pandemic Covid-29 By Method Naive Bayes Classifier. *Jurnal Teknik Informatika (Jutif)*, 2(2), 202–207. <https://doi.org/20.20884/2.jutif.2022.2.2.52>
- Pengguna, P., & Indonesia, I. (2020). *Laporan survei internet apjii 2019 –*

2020. 2020.

Rusdian, D., & Rosiyadi, D. (2019).
Metode Naïve Bayes Classifier Dan
Support Vector Machine. *Jornal of
Computer Engineering System and
Science*, 4(2), 230–235.

Satori dan Komariyah (2018: 200),.Analisis
sentimen *Science*, 4(2), 230–235.

Statistika, J., Matematika, F., Ilmu, D. A. N.,
Alam, P., & Indonesia, U. I. (2020).
*TENTANG KESEHATAN MENTAL
SELAMA PANDEMI COVID-19 DI
MEDIA SOSIAL TWITTER
MENGUNAKAN NAIVE BAYES
CLASSIFIER DAN SUPPORT VECTOR.*

Syaifudin, Y. W., & Irawan, R. A. (2020).
Implementasi Analisis Clustering Dan
Sentimen Data Twitter Pada Opini
Wisata Pantai Menggunakan Metode K-
Means. *Jurnal Informatika Polinema*,
4(3), 289.
<https://doi.org/20.33795/jip.v4i3.205>

Zuhdi, A. M., Utami, E., & Raharjo, S.
(2019). Analisis sentiment twitter
terhadap capres Indonesia 2029 dengan
metode K-NN. *Jurnal Informa Politeknik
Indonusa Surakarta*, 5, 2–7.