KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)
Terakreditasi Nomor 204/E/KPT/2022 | Vol. 5, No. 3, Juli (2024), pp. 1244-1252

Analisa Sentimen Twitter Vaksin Covid-19 di Indonesia dengan Metode Support Vector Machine

Sulastri[1], Fahad Abdul Nur[2]*
[1,2]Fakultas Teknologi Informasi dan Industri, Universitas Stikubank Semarang, Indonesia
E-mail : sulastri@edu.unisbank.ac.id[1],fahadabdulnur@gmail.com[2]

Abstract

In today's digital era, Twitter is very widely used by the public as a means of communication. These facilities provide freedom in expressing opinions and / or opinions on twitter social media. Opinions in social media twitter can be called tweets. Opinions that often arise are very diverse in expression and meaning in the context of the problem being discussed. Tweets analyzed in this study are related to the issue of the Coronavirus in Indonesia. The data used in this study, 500 tweet data with 350 training data and 150 test data. The tools used in the classification process is the Python programming language. Then for the method used in the classification is the vector support machine method with the sentiment data used, namely positive and negative. The results are given in the support vector machine method in a value of 66%, with a recall of 61% and a precision of 74%. Therefore, the support vector machine method is quite good in classifying.

Keywords: *Analisis Sentimen Twitter, Metode Support Vector Machine, Isu Vaksin Covid-19*

Abstrak

Di era digital saat ini, Twitter sangat banyak digunakan oleh masyarakat sebagai alat komunikasi. Fasilitas tersebut memberikan kebebasan dalam mengemukakan pendapat dan/atau pendapat di media sosial twitter. Pendapat di media sosial twitter bisa disebut dengan tweet. Pendapat-pendapat yang sering muncul sangat beragam ekspresi dan maknanya dalam konteks permasalahan yang dibicarakan. Tweet yang dianalisis dalam penelitian ini berkaitan dengan isu virus Corona di Indonesia. Data yang digunakan dalam penelitian ini sebanyak 500 data tweet dengan 350 data latih dan 150 data uji. Alat yang digunakan dalam proses klasifikasi adalah bahasa pemrograman Python. Kemudian untuk metode yang digunakan dalam klasifikasinya adalah metode vector support machine dengan data sentimen yang digunakan yaitu positif dan negatif. Hasil yang diberikan pada metode support vector machine dengan nilai 66%, recall 61%, dan presisi 74%. Oleh karena itu, metode support vector machine cukup baik dalam melakukan klasifikasi.

Kata Kunci: *Analisis Sentimen Twitter, Metode Support Vector Machine, Isu Vaksin Covid-19*

1. Pendahuluan

Pada 2020, seluruh dunia sedang digemparkan dengan pandemi virus corona. Hampir seluruh negara di dunia terdampak dengan pandemi ini, termasuk negara Indonesia. Timbulnya masalah ini masyarakat merasa khawatir dikarenakan sudah banyak kasus yang bermunculan, masalah tersebut yaitu berupa virus corona atau covid19. Dengan belum adanya vaksin/obat untuk mematikan virus tersebut, pemerintah melakukan tindakan seperti lockdown/PSBB (Pembatasan Sosial Berskala Besar) terhadap masyarakat untuk tinggal di rumah. Menurut Wikipedia total kasus virus corona di Indonesia menjadi 598.933 jiwa, untuk korban yang meninggal mencapai 18.336 jiwa,

ISSN: 2720-992X

sedangkan total pasien sembuh dari virus corona mencapai 492.000 jiwa per 12 Desember 2020[1]. Untuk itu perlu penelitian dalam permasalahan ini.

Media sosial saat ini memegang peranan penting dalam menyebarkan informasi, salah satunya media sosial twitter yang sangat banyak digunakan oleh masyarakat sebagai sarana komunikasi. Sarana tersebut memberikan kebebasan dalam mengutarakan opini atau pendapat didalamnya. Twitter sebagai media informasi yang akan digunakan untuk mengambil informasi *tweet* pengguna atau status yang ada ditwitter. Opini dan pendapat masyarakat dalam memecahkan permasalahan yang sedang dibicarakan sangat beragam mulai dari berkomentar secara baik atau positif maupun buruk atau negatif.

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam database. Data mining adalah proses yang menguraikan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengeksraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar[2].

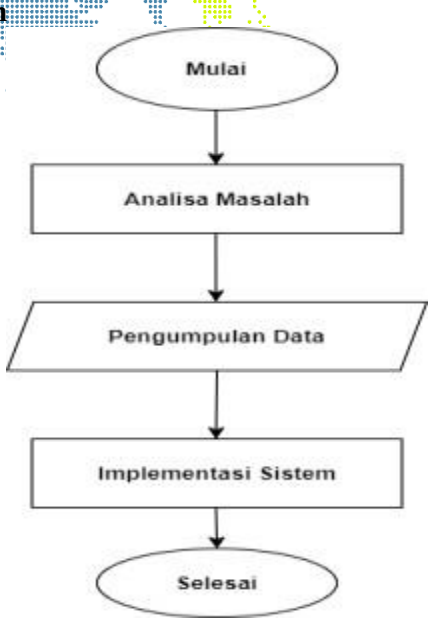
Text mining yang dikenal juga dengan *text data mining* atau pencarian pengetahuan di basis data textual adalah sebuah proses yang semi otomatis melakukan ekstraksi dari pola yang ada di *database*. *Text mining* mempunyai kesamaan dengan *data mining.* Keduanya memiliki tujuan yang sama yaitu untuk memperoleh informasi dan pengetahuan dari sekumpulan data sangat banyak. Data tersebut bisa berbentuk sebuah *database*[3].

Penelitian yang dilakukan oleh Novantirani Anita, Mira Kania Sabariah, danVeronika Effendu, (2015) menunjukkan sebuah proses klasifikasi dokumen tekstual ke dalam dua kelas, yaitu kelas sentimen positif dan negatif. Hasil dari penelitian dapat membantu usaha untuk melakukan riset pasar atas opini publik[4].·Penelitian berikutnya dilakukan oleh Ghulam Asrofi Buntoro (2017), penelitian ini diharapkan dapat mengandung sentimen positif, netral atau negatif. Dengan menggunakan metode dalam penelitian ini yaitu untuk proses klasifikasinya menggunakan metode *Naive Bayes Classifier (NBC)* dan *Support Vector Machine (SVM).* Untuk preprocessing data menggunakan tokenisasi, cleansing dan filtering, untuk menentukan class sentimen dengan metode Lexicon Based. Hasil dari penelitian ini adalah analisis sentimen terhadap calon gubernur DKI Jakarta 2017 dengan jumlah dataset sebanyak 300 tweet[5]. Adapun penelitian yang hampir sama dengan penelitian terkait yaitu penelitian yang dilakukan oleh Umi Rofiqoh, Rizal Setya Perdana dan M Ali Fauzi (2017) yang bermaksud untuk menganalisis sentimen salah satu cabang penelitian dari *Text Mining* yang berguna untuk mengklasifikasi dokumen teks berupa opini berdasarkan sentimen. Dokumen teks yang digunakan dalam penelitian berasal Twitter tentang opini masyarakat mengenai penyedia layanan telekomunikasi seluler. Metode yang digunakan adalah Support Vector Machine dengan menggunakan *Lexicon Based Features* sebagai pembaharuan fitur selain memakai fitur TF-IDF. Data yang digunakan pada penelitian ini sebanyak 300 data yang dibagi menjadi dua jenis data dengan perbandingan 70% untuk data latih dan 30% untuk data uji[6].

Permasalahan dalam penelitian ini adalah menganalisis, mengelola data opini pengguna twitter terkait isu vaksin covid-19 di Indonesia dan kemudian mengklasifikasikan sentimen data *tweet* tersebut dengan menggunakan metode *support vector machine* serta menghitung tingkat akurasi yang dihasilkan oleh metode *support vector machine*[7]. Maka dari itu penelitian ini bertujuan untuk mengimplementasikan algoritma *Support Vector Machine* dalam analisis sentimen pengguna sosial media twitter dengan topik virus corona di indonesia. Mengetahui isu apa yang terjadi dengan isu vaksin covid-19 di Indonesia, mentukan hasil akurasi dengan menggunakan metode *support vector machine*[8].

2. Metodelogi Penelitian

2.1. Flowchart Penelitian



Gambar 1. Flowchart Penelitian

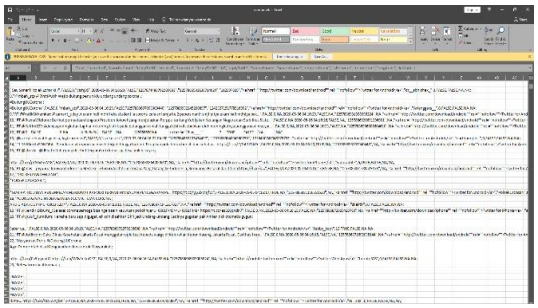
2.2. Analisis Masalah

Masalah yang dipakai dalam penelitian ini adalah terkait isu vaksin covid-19 di Indonesia dengan pengumpulan data *tweet* mengenai isu vaksin coivd-19 menggunakan Twitter API (*Application Programming Interface)*. Metode yang digunakan dalam penelitian ini yaitu metode *support vector machine* guna untuk mengklasifikasi data *tweet* terkait isu vaksin covid-19 di Indonesia[9]. Data *tweet* yang digunakan dalam penelitian ini diambil dari dua kelas sentimen positif dan negatif. Untuk menentukan dua kelas sentiment tersebut dilakukan secara manual. Sebelum melakukan proses mengklasifikasi data *tweet* harus dilakukan terlebih dahulu proses untuk menyiapkan data dengan proses *text preprocessing*[10].

2.3. Pengumpulan Data

Proses pengambilan data menggunakan crawling dengan menggunakan kata kunci "#corona dan #vaksin" di Indonesia. Pengambilan data menggunakan *api key* untuk autentifikasi ke media social *Twitter*. *Api key* merupakan suatu layanan yang disediakan untuk pengembang[11]. Proses *api key* dapat dilakukan sebagai berikut :

1. Masuk ke *twitter* developer atau buka *https://developer.twitter.com/en/apps*.
2. Lalu membuat app dengan memilih *create an app*



Gambar 2. Crawling Data Format Excel

3. Kemudian mengisi formulir yang disediakan untuk mendapatkan *api key*

Setelah mendapatkan *api key* dapat dilakukan proses crawling data dengan mengautentifikasi dahulu antara *twitter* dengan *R Studio*. Pencarian data dapat dilakukan setelah proses autentifikasi selesai dengan menggunakan kata kunci atau *keyword* yang sudah di tentukan[12]. Kemudian data yang terambil akan disimpan dengan bentuk format .xlsx karena memudahkan untuk proses selanjutnya. Penelitian ini terdapat dua jenis data diantaranya data *training* dan data *testing*. Untuk mengetahui data *training* analisis sentimen tersebut perlu mengklasifikasikan sentimen mengenai isu vaksin covid-19 yaitu dengan kelas positif dan kelas negatif. Berikut adalah contoh pengumpulan data pada Tabel 1.

Tabel 1. Contoh Pengumpulan Data

Sentimen	Tweet
Positif	Udah gak usah bahkan gak perlu nyalahin pemerintah lagi. Disaat ini saling bahu-membahu bersama adalah waktu yg tepat
Positif	Demi menghadapi pandemi virus #corona , Pemerintah #india mengambil kebijakan #lockdown . Selama menerapkan kebijakan tersebut
Positif	Terimakasih karena tetap bekerja untuk membantu melawan covid-19 di rumah sakit saat kami bekerja di rumah
Negatif	Kalau kaya gini terus, lama-lama kita lupa caranya bertegur sapa secara langsung.
Negatif	Hampir seluruh aspek kehidupan penduduk dunia terdampak pandemi corona, termasuk sektor penerbangan yg anjlok sejak viruscorona masuk
?	@jokowi Ojol tidak pernah takut mati karena Corona, tapi mereka takut anak istrinya mati kelaparan! Kalau memang banyak

Data diatas merupakan contoh dari data *training* dan data *testing* yang masing-masing terdiri dari 6 data training, 3 data positif dan 2 data negatif, kemudian 1 data yang tidak diketahui apakah data tersebut termasuk kedalam data sentimen positif atau negatif.

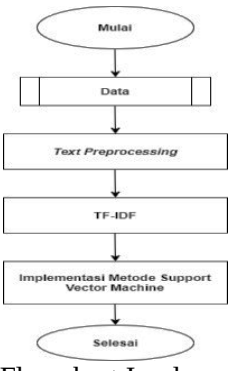
Tabel 2. Data Penelitian

Jenis Tweet Sentimen	Positif	Negatif
Isu Vaksin Covid-19	260	240
Data Training	157	193
Data Testing	83	67

Data untuk penelitian ini dengan tweet sentimen berjumlah 499 data dengan total sentimen positif 260 data dan sentimen negatif 240 data.

2.4. Implementasi Sistem

Pada tahap ini akan dijelaskan proses mengenai implementasi sistem dalam penelitian ini. Berikut adalah flowchart implementasi sistem dapat dilihat pada Gambar 3.



Gambar 3. Flowchart Implementasi Sistem

2.4.1. Penginputan Data

Data yang sudah disimpan dalam format excel, selanjutnya data akan diinputkan kedalam bahasa *Python.*

2.4.2. Text Preprocessing

Pada tahap *text preprocessing* adalah tahapan dimana aplikasi melakukan seleksi data yang akan diproses pada setiap dokumen. Proses preprocessing ini meliputi (1) case folding, (2) tokenizing, (3) filtering, dan (4) stemming. Tahapan tersebut juga merupakan proses analisis sentimen yang memerlukan beberapa tahap untuk mempersiapkan suatu teks dapat diubah secara terstruktur[13].

2.4.3. TF-IDF

TF (*Term Frequency)* adalah frekuensi dari kemunculan sebuah *term* dalam dokumen yang bersangkutan[14]. Sedangkan menurut Gieffari Satria Abdilah tahun 2019 term freq merupakan sebuah proses pembobotan kata dimana dalam proses ini dilakukan perhitungan *term* atau kata yang muncul dalam *tweet* (tf), menghitung data *tweet* yang mengandung *term* (df), menghitung inverse dokumen frekuensi (idf)[15]. Berikut rumus dari:

$W_{ij} = tf_{ij}.\text{Idf}$ (1)

2.4.4. Imlementasi Metode Support Vector Machine

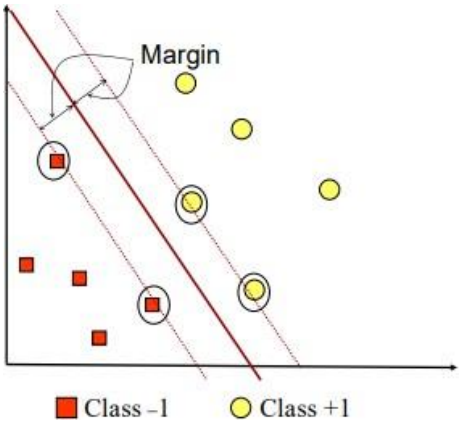
Support vector machine itu sendiri memiliki konsep klasifikasi dengan mencari hyperplane atau garis pembatas untuk memisahkan kedua data sentimen dengan baik. Garis pembatas tersebut berguna untukmemisahkan *tweet* bersentimen positif (+1) dengan *tweet* negatif (-1).

Persamaan SVM :

$f(x) = w.x + b$ atau $f(x) = \sum y_i a_i K(x, x_i) + bm_i = 1$ (2)

Keterangan :

w = Parameter hyperplane yang dicari (garis tegak lurus antara hyperplane dan titik support vector).
x = Titik data masukan support vector machine.
a_i = Nilai bobot setiap titik data.
$K(x, x_i)$ = Fungsi kernel.
b = Parameter hyperplane yang dicari (nilai bias).



Gambar 4. Struktur Metode SVM

2.4.5. Hitung Akurasi

Proses klasifikasi data dalam penelitian ini menggunakan confussion matrix. Dimana perhitungan akurasi, recall, presisinya menggunakan confussion matrix tersebut diambil dari jumlah perhitungan data positif dan data negatif. Berikut adalah hasil klasifikasi dari

confussion matrix akan dijelaskan pada tabel dibawah dengan data testing sebanyak 150 data *tweet*.

Tabel 3. Confussion Matrix

	Kelas Prediksi	
	Negatif	**Positif**
Positif	*TP*	*FP*
Negatf	*TN*	*FN*

Dalam menghitung menggunakan rumus sebagai berikut:

$$\text{Akurasi} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+TN} \quad (2)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (3)$$

3. Hasil Dan Pembahasan

Hasil penelitian dari analisis sentimen *twitter* yang berfokus pada topik isu terkait virus corona di Indonesia adalah:

3.1. Membaca Data *Tweet*

Out[4]:

	sentimen	tweet
0	0	Alhamdulillah, Pak Bos @PlateJohnny dah divaks...
1	0	Lebih dari 500 ribu tenaga kesehatan telah men...
2	1	Tindakan lebih tegas akan dikenakan kepada man...
3	0	Wakil Presiden Ma'ruf Amin mengatakan menerapk...
4	0	Vaksin Covid-19 Aman dan Halal
5	0	Vaksina telah teruji klinis dan Aman untuk dig...
6	0	Ayoo vaksinasi untuk mengurangi penularan covi...
7	1	Setelah mendapat #vaksin #Sinovac . seLama 2 h...
8	0	Tidak perlu takut divaksin
9	0	Indonesia menerima konfirmasi indikasi alokasi...

Gambar 5. Membaca data *Tweet*

Pada Gambar 5. yaitu script yang digunakan untuk membaca data yang sebelumnya di import maka akan menampilkan hasil tersebut. Untuk penelitian ini proses pembacaan data. Data yang digunakan dalam penelitian ini yaitu data sentimen twitter media sosial terkait isu vaksin covid-19 di Indonesia dengan klasifikasi positif dan negatif yang nantinya akan digunakan dalam proses *training* dan *testing*.

3.2. Stopword dengan bentuk Wordcloud

Pada Gambar 6. merupakan hasil proses dari memvisualisasikan suatu bentuk teks secara bebas. Pada *wordcloud* diatas tersebut merupakan imbuhan dari data sentimen tweet terkait isu vaksin covid-19 di Indonesia. Kemudian untuk data sentimen yang sering banyak muncul yaitu kata psbb, mrt, jakarta, tambah, penutupan, stasiun. Oleh karena itu proses visualisasi ini hanya menampilan imbuhan yang sering banyak muncul.

```
['pak', ',', 'bapak', 'sempat', 'kena', 'virus', 'corona', 'loh', '', 'sama', 'allah', '',
'dikasih', 'waktu']

<matplotlib.image.AxesImage at 0x7f3283fa3890>
```

Gambar 6. Hasil dalam bentuk *Wordcloud*

3.3. Hasil TF-IDF

Pada Gambar 7 merupakan tampilan hasil dari proses penghitungan tf-idf dengan diimplementasikan dalam bentuk numerik. Untuk 0 yaitu sentimen positif dan 1 negatif, lalu untuk angka setelah sentimen tersebut yaitu kata pada data yang dilakukan dalam perhitungan *tf-idf*, dan kemudian angka dengan 0. setelah kata data yaitu hasil perhitungan dari *tf-idf*.

```
In [10]: print(Train_X_Tfidf)

  (0, 1979)     0.13939423137377935
  (0, 1925)     0.1103390233655628
  (0, 1672)     0.29443296250293244
  (0, 1432)     0.21746646106393935
  (0, 1352)     0.2761268966573378
  (0, 1132)     0.21353804413080313
  (0, 1088)     0.26313853623966516
  (0, 664)      0.15806587260605912
  (0, 541)      0.29443296250293244
  (0, 438)      0.29443296250293244
  (0, 380)      0.29443296250293244
  (0, 351)      0.11965476534100114
  (0, 142)      0.29443296250293244
  (0, 129)      0.29443296250293244
  (0, 76)       0.3904639565704169
  (0, 12)       0.13456462057619029
  (1, 1924)     0.18544324459612915
  (1, 1923)     0.21353591991338103
  (1, 1593)     0.2643093549217823
  (1, 1343)     0.21353591991338103
  (1, 1112)     0.24787619331625518
  (1, 1069)     0.20812400428727862
  (1, 1042)     0.2643093549217823
  (1, 834)      0.2643093549217823
  (1, 784)      0.24787619331625518
```

Gambar 7. Hasil Proses Penghitungan TF-IDF

3.4. Hasil Implementasi Metode SVM dan Hasil Akurasi

Pada Gambar 8. yaitu menampilkan hasil klasifikasi SVM dengan proses pemanggilan data hasil dari test_y dan predictions_svm dimana hasil yang ditampilkan merupakan hasil akhir dari penelitian ini yang merupakan nilai ketepatan prediksi diatas sebesar 81%

```
In [12]: print(classification_report(Test_Y, predictions_SVM))

              precision    recall  f1-score   support

           0       0.92      0.72      0.81        83
           1       0.73      0.93      0.82        67

    accuracy                           0.81       150
   macro avg       0.83      0.82      0.81       150
weighted avg       0.84      0.81      0.81       150
```

Gambar 8. Implementasi Metode SVM dan Hasil Akurasi

Tabel 4. Hasil Prediksi

		Actual	
		Positif	**Negatif**
Kelas Prediksi	Positif	60	23
	Negatif	5	62

Berikut ini penjelasan dari tabel diatas :

1. *True* positif 60 data tweet.
2. *True* negatif 23 data tweet.
3. *False* positif 5 data tweet.
4. *False* negatif 62 data tweet.

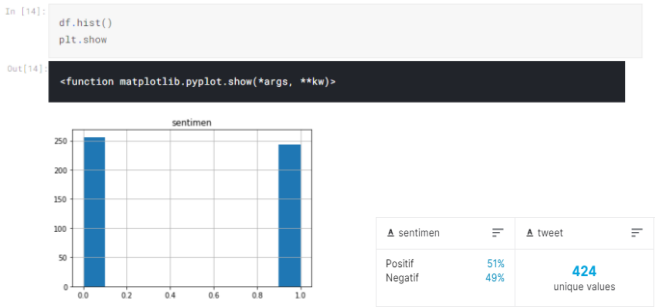
Dalam menghitung menggunakan rumus sebagai berikut:

$$\text{Akurasi} = \frac{TP + TN}{TP + FP + TN + FN} = \frac{60 + 62}{60 + 23 + 5 + 62} = 0.81 \times 100\% = 81\%$$

$$\text{Recall} = \frac{TP}{TP + TN} = \frac{60}{60 + 23} = 0.72 \times 100\% = 72\%$$

$$\text{Presisi} = \frac{TP}{TP + FP} = \frac{60}{60 + 5} = 0.92 \times 100\% = 92\%$$

Hasil perhitungan akurasi adalah 81 % artinya kinerja dari algoritma svm cukup baik dalam melakukan klasifikasi, nilai recall adalah 72 % artinya algortima svm dalam melakukan recall cukup mengenali sentimen data uji pada saat pengecekan kata bersentimen pada proses pembobotan data, dan untuk presisinya berjumlah 92 % yang artinya algortima svm ini cukup teliti dalam melakukan proses pengklasifikasian data dikarenakan nilai dari presisi lebih tinggi dibandingkan dengan yang lainnya. Kesimpulan dari perhitungan svm ini bahwa metode support vector machine cukup baik dalam mengklasifikasikan data.



Gambar 9. Histogram

Pada Gambar 9 merupakan salah satu visualisasi analisis sentimen dari klasifikasi terkait isu vaksin covid-19 di Indonesia dalam bentuk histogram yang dibagi 2 nilai dataset dengan diurutkan ke dalam interval data. Hasil dari proses tersebut yaitu 51 % untuk sentimen positif dan 49 % untuk sentimen negatif dengan positif (0.0) dan negatif (1.0).

4. Kesimpulan

Metode yang digunakan dalam penelitian ini yaitu *Support Vector Machine* yang dimana metode ini dapat melakukan klasifikasi *tweet* yang memiliki sentimen positif dan negatif dengan data sebanyak 350 data latih dan 150 data uji yang mendapatkan hasil akurasi sebesar 81 %, dengan hasil recall 72 %, dan hasil presisinya sebesar 92 %. Penelitian ini menghasilkan beberapa hasil plot Gambar berupa plot wordcloud dan histogram yang dapat memberikan sebuah informasi statistik data yang dihasilkan.

Daftar Pustaka

[1] Wikipedia, "Pandemi COVID-19 di Indonesia," 2020. https://id.wikipedia.org/wiki/Pandemi_COVID-19_di_Indonesia. (accessed Dec. 13, 2020).

[2] dkk. Turban, E., *Decision Support Systems and Intelligent Systems*. Yogyakarta: Andi Offset., 2005.

[3] Sumarno and M. A. Rosid, "Classification of Student Complaints with the Naive Bayes and Literature Methods," *JOINCS (Journal Informatics, Network, Comput. Sci.*, vol. 3, pp. 1–7, 2020, doi: 10.21070/joincs.v3i0.711.

[4] A. Novantirani, M. K. Sabariah, and V. Effendy, "Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat Dalam Kota dengan Metode Support Vector Machine," *e-Proceeeding Eng.*, vol. 2, no. 1, pp. 1–7, 2015.

[5] G. A. Buntoro, "Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter," *INTEGER J. Inf. Technol.*, vol. 2, no. 1, pp.32–41, 2017, [Online] .Available: https://www.researchgate.net/profile/Ghulam_Buntoro/publication/316617194_Analisis_Sentimen_Calon_Gubernur_DKI_Jakarta_2017_Di_Twitter/links/5907eee44585152d2e9ff992/Analisis-Sentimen-Calon-Gubernur-DKI-Jakarta- 2017-Di-Twitter.pdf.

[6] U. Rofiqoh, R. S. Perdana, and M. A. Fauzi, "Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature," *J.Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1725–1732, 2017, [Online]. Available: http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/628.

[7] A. V. Sudiantoro and E. Zuliarso, "Analisis Sentimen Twitter Menggunakan Text Mining Dengan Algoritma Naïve Bayes Classifier," *Pros. SINTAK 2018*, pp. 398–401, 2018.

[8] M. Syarifuddin, "Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naïve Bayes Dan Knn," *Inti Nusa Mandiri*, vol. 15, no. 1, pp. 23–28, 2020.

[9] L. Nurhajati, R. Sukandar, R. C. Oktaviani, and X. A. Wijayanto, "Perbincangan Isu Corona COVID-19 di Media Daring dan Media Sosial di Indonesia," *Lemb. Penelitian, Publ. dan Pengabdi. Masy.*, pp. 1–27, 2020, [Online]. Available: https://www.researchgate.net/publication/340916589_Perbincangan_Isu_Corona_COVID-19_di_Media_Daring_dan_Media_Sosial_di_Indonesia__Big_Data_Analysis/link/5ea3d8fe45851553faace361/download.

[10] B. M. Pintoko and K. M. L, "Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naïve Bayes Classifier," *e-Proceeding Eng.*, vol. 5, no. 3, pp. 8121–8130, 2018.

[11] F. Rahutomo, P. Y. Saputra, and M. A. Fidyawan, "Implementasi Twitter Sentiment Analysis Untuk Review Film Menggunakan Algoritma Support Vector Machine," *J. Inform. Polinema*, vol. 4, no. 2, p. 93, 2018, doi: 10.33795/jip.v4i2.152.

[12] Andrybrew, "Crawling – Mining Twitter Data menggunakan R," 2015. https://andrybrew.blog/2015/03/01/crawling- mining-twitter-data-menggunakan-r/ (accessed Dec. 13, 2020).

[13] INFORMATIKALOGI, "Text Preprocessing," 2016. https://informatikalogi.com/text-preprocessing/ (accessed Dec. 13, 2020).

[14] S. T. Yuliana, "Anaisis Sentimen Lirik Lagu Indonesia Dengan Metode K-Nearest Neighbor," Universitas Stikubank Semrang, 2018.

[15] G. S. Abdillah, "Analisis Sentimen Media Sosial Twitter Terkait isu yang Terjadi pada Papua Dengan Metode Support Vector Machine," Universitas Stikubank Semarang, 2019.