

Pemanfaatan Analisis Sentimen Terhadap Kasus Bunuh Diri Mahasiswa Menggunakan *Naïve Bayes Classifier*

Ainnur Rafli¹, Kusnawi²

ainnurrafli31@students.amikom.ac.id, khusnawi@amikom.ac.id

^{1,2}Universitas Amikom Yogyakarta

Informasi Artikel

Diterima : 15 Des 2023

Direview : 30 Des 2023

Disetujui : 10 Feb 2024

Kata Kunci

Analisis Sentimen, *Naïve Bayes Classifier*, Klasifikasi, Bunuh Diri, Mahasiswa

Abstrak

Bunuh diri saat ini merupakan masalah serius yang terjadi di lingkungan pendidikan tinggi, khususnya di lingkungan mahasiswa, dan diperlukan pendekatan dan perhatian khusus untuk mencegahnya. Dengan perkembangan teknologi saat ini, teknik analisis emosi dapat menjadi cara yang efektif untuk memahami perasaan dan pikiran siswa yang mungkin mengarah pada perilaku bunuh diri atau mengindikasikan adanya risiko bunuh diri. Pada penelitian kali ini menggunakan data yang di-*scrapping* melalui Twitter sebanyak 1151 tweet, kemudian dilakukan cleaning data menjadi sebanyak 817 data. Yang mana terdapat 745 tweet negatif dan 72 tweet positif. Selanjutnya, data tersebut diimplementasikan ke dalam algoritma yang mana dilakukan splitting data dengan perbandingan 80:20 dengan hasil akurasi 90,24%. Yang mana dengan visualisasi **data** yang sering muncul adalah "Depresi". Dengan begitu banyaknya Tindakan bunuh diri yang terjadi khususnya di Indonesia berasal dari depresi. Tujuan dari penelitian ini adalah untuk memahami faktor-faktor yang berhubungan dengan kejadian bunuh diri siswa dan untuk mengetahui efektivitas dan keakuratan algoritma ini. Lebih jauh lagi, penelitian ini diharapkan dapat memberikan wawasan terhadap lingkungan pendidikan dan kesehatan mental untuk meningkatkan strategi pencegahan dan pendekatan yang lebih efektif.

Keywords

Sentiment Analysis, Naïve Bayes Classifier, Classification, Suicide, Student

Abstrak

Suicide is currently a serious problem in higher education, especially among university students, and special approaches and attention are required to prevent it. With today's advances in technology, emotion analysis techniques can be an effective way to understand students' feelings and thoughts that may lead to suicidal behavior or indicate a risk of suicide. For this study, we scraped the data for his 1,151 tweets on Twitter and cleaned it up to 817. Of these, there are 745 negative tweets and 72 positive tweets. Additionally, the data is implemented in an algorithm that performs a data split of 80:20 with an accuracy of 90,24%. That's the "depression" that often appears when visualizing data. Especially in Indonesia, there are many suicides due to depression. The purpose of this study is to understand the factors associated with student suicide and to determine the effectiveness and accuracy of this algorithm. Additionally, this study is expected to provide insights into educational and mental health settings to improve prevention strategies and more effective approaches.

A. Pendahuluan

Kesehatan mental mahasiswa menjadi aspek penting dari timbulnya dan maraknya kasus bunuh diri yang terjadi di Indonesia. Di Indonesia kejadian bunuh diri yang diakibatkan dari kesehatan mental yang mana menjadi tingkat kematian ke-2 yang mana menghasilkan rata-rata 1 jam 1 orang yang tewas dikarenakan bunuh diri pada rentan usia 15 sampai dengan 29, dengan salah satu populasi percobaan dan ide bunuh diri adalah kaum muda termasuk mahasiswa yang sedang mengalami masa transisi, memasuki pendidikan tinggi sering membawa masalah psikologi. Kenaikan depresi pada mahasiswa cenderung mengalami kenaikan dibandingkan dengan usia anak-anak dan dewasa. Dengan adanya depresi pada mahasiswa, berpengaruh terhadap keseharian mahasiswa tersebut. Sehingga depresi menjadi masalah sosial dengan tingkat bunuh diri yang tinggi dan cenderung meningkat [1]. Dari penyakit kejiwaan ini yang mana diderita lebih dari 300 juta orang di dunia, 96% penyebab bunuh diri adalah depresi, tercatat 800.000 orang mati setiap harinya karena bunuh diri [2]. WHO menyebut 79% masalah bunuh diri terjadi pada negara berpenghasilan rendah serta menengah pada tahun 2016 [3] Organisasi Kesehatan Dunia (WHO) menyatakan bahwa terdapat rata-rata satu orang yang meninggal dikarenakan bunuh diri di dunia setiap 40 detik. Dengan angka 3,7 itu, menempatkan Indonesia di peringkat 159 dalam kasus bunuh diri di dunia. Pada *Asian Journal of Psychiatry* warga Indonesia menderita depresi sebanyak 21.8% yang mana kebanyakan adalah remaja [4].

Platform online yang semakin banyak memberikan ruang kepada setiap orang untuk berbagi pengalaman mereka akan hal ini. Oleh karena itu, penggunaan analisis sentimen pada data digital dapat menjadi alat yang efektif untuk mendeteksi dan memahami perasaan mahasiswa terkait kasus bunuh diri. Dengan adanya berbagai macam aplikasi, pemberian informasi dapat menyebar dengan lebih luas dan cepat, dan pengguna dapat mengekspresikan pendapatnya dalam bentuk komentar, yang mana pada kolom komentar ini Masyarakat dapat dengan bebas berpendapat [5]. Pada penelitian [6] yang mana mengolah data komentar yang diambil dari media social *Facebook* dengan menggunakan *pre-processing* data untuk menghilangkan kata atau karakter yang tidak dibutuhkan, membangun aplikasi *prototype filter* komentar untuk filterisasi komentar negatif *cyberbullying*, dengan data komentar yang digunakan yaitu 300 *data training*, dan 100 *data testing*. Dan mendapatkan nilai akurasi sebesar 86% dengan menerapkan metode *confusion matrix* dengan jumlah komentar 100 komentar dari *data testing*.

Analisis sentimen adalah studi yang bertujuan untuk menganalisis opini sentimen dan emosi yang terdapat pada dokumen atau data [7]. Menganalisis ungkapan-ungkapan *online* mahasiswa merupakan hal yang akan diidentifikasi sebagai unsur dari adanya pemikiran bahkan sampai dengan tindakan bunuh diri. Dari banyaknya kasus yang terjadi di Indonesia dibutuhkan intervensi yang efektif dan efisien untuk melakukan proses klasifikasi sentimen tersebut. Pentingnya penelitian ini yaitu untuk mencegah kasus bunuh diri semakin marak terjadi di kalangan mahasiswa sebagai generasi bangsa. Di sisi lain, keterlibatan teknologi dan analisis sentimen juga dapat memberikan pengetahuan yang mendalam terhadap bagaimana teknologi dapat digunakan untuk mendeteksi masalah kesehatan mental dengan lebih efektif. Maka dari itu digunakanlah salah satu metode pengklasifikasian sederhana dalam bida *machine learning*, namun masih

mendukung mesin vector yaitu *Naïve Bayes Classifier*. Kegunaan dari metode klasifikasi dan metode statistik ini adalah untuk memprediksi peluang masa depan berdasarkan pada pengalaman teorema *Bayes* sebelumnya [8].

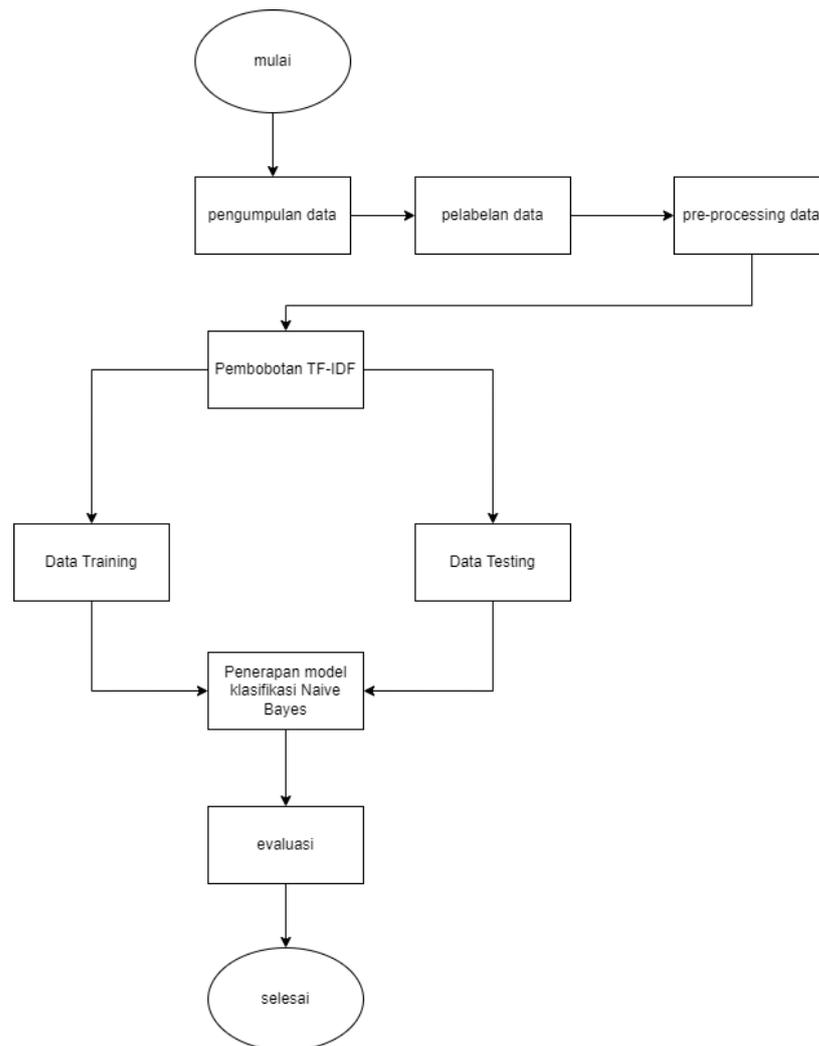
Beberapa penelitian yang meneliti topik ini seperti pada [9] yang dimana dilakukannya pembobotan TF-IDF dari total data yang didapatkan pada Twitter yakni sebanyak 1058 *tweet*. Kemudian digunakannya rasio *training* dan *testing* dengan perbandingan 70:30, maka klasifikasi dengan algoritma *Naïve Bayes* hanya akan memprediksi data dengan hanya berjumlah 318 *tweet* dan hal ini menghasilkan nilai akurasi yang cukup baik yaitu 89%.

Kemudian untuk penelitian [10] yang mana melakukan perancangan serta pengujian untuk aplikasi pendeteksi *cyberbullying* dari komentar pada unggahan di Instagram dengan metode *Naïve Bayes Classifier* sebanyak 200 data komentar dari suatu unggahan yang menunjukkan terdapat 78 komentar *bullying* dan 122 komentar bukan *bullying*. Kemudian data tersebut diujikan dengan menggunakan pengujian *confussion matrix* yang menunjukkan bahwa perhitungan yang dilakukan oleh aplikasi memiliki nilai akurasi sebesar 98,5% dan mendapatkan klasifikasi kata yang memiliki bobot tinggi, diketahui bahwa kata kunci “tolol” dengan nilai sebesar 13,82% .

Fokus dari penelitian ini yaitu pada pengembangan model dari algoritma yang akan digunakan oleh penulis yakni *Naïve Baye Classifier* dari adanya unggahan mahasiswa di *platform* digital *Twitter*. Karena, menurut situs situs pemeringkatan hosting, *Twitter* memiliki pengguna dengan mencapai 1,3 miliar akun harian, 330 juta pengguna aktif, dan 500 juta *tweet* dikirim [11] . Harapannya hasil analisis dari penelitian ini dapat memudahkan untuk pendekatan serta memberikan pengetahuan berdasarkan data akan pentingnya dukungan kesehatan mental yang lebih khususnya di lingkungan perguruan tinggi. Serta harapannya dapat memberikan kontribusi yang besar dalam pengembangan strategi pencegahan tindakan bunuh diri di kalangan mahasiswa melalui pendekatan berbasis teknologi yakni analisis sentimen dan metode klasifikasi *Naïve Bayes Classifier*.

B. Metode Penelitian

Metode penelitian yang dilakukan pada penelitian ini dengan menggunakan algoritma *Naïve Bayes Classifier* untuk melakukan pengklasifikasian sentiment yang mana hal ini sekaligus melihat banyaknya ungkapan-ungkapan dari media digital seperti *Twitter*, *Kaggle*. Kemudian untuk pembuatan model pada penelitian ini akan dilakukan dengan pemrograman *Python*. Berikut adalah alur penelitian yang akan diimplementasikan yang akan diimplementasikan.



Gambar 1. Alur Penelitian

1. Pengumpulan Data

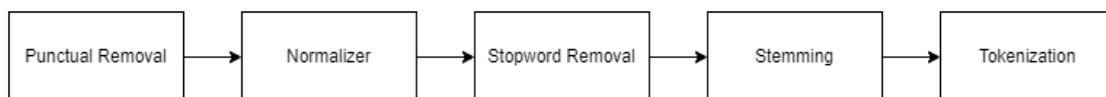
Pada penelitian kali ini penulis melakukan metode *scrapping data* pada sosial media yakni twitter. Proses *scrapping data* terhadap unggahan baik itu thread dilakukan dengan menggunakan *tweet-harvest* pada pemrograman *Python*. Pada proses pengumpulan data kali ini, penulis mendapatkan 1151 *tweet* yang berhasil diperoleh dari tahun 2022 sampai dengan 2023. Berikut adalah contoh dari hasil pengumpulan data.

2. Pelabelan Data

Kemudian setelah *tweet* yang berhasil disimpan pada bentuk format *.csv*, selanjutnya adalah memberikan label untuk setiap *tweet* dengan merubah parameter skor menjadi sentiment berdasarkan skor tersebut. Dalam hal ini, penulis menggunakan kode dalam Bahasa pemrogram *Python* dan menggunakan *tools* Google Colaboration [12]. Yang mana nantinya data akan diberi label negative dan positif.

3. Pre-processing

Pada tahapan ini merupakan proses penyiapan data yang nantinya digunakan dalam proses klasifikasi menggunakan algoritma. Data-data yang di dapatkan dari internet ini masih berupa data tak berstruktur yang mana perlu diterapkannya proses *pre-processing* supaya data yang dikumpulkan dapat digunakan sebagai data yang matang [5]. Yang mana tujuannya dari dilakukannya tahapan ini adalah agar mendapatkan hasil yang maksimal saat pengimplementasian klasifikasi [13]. Terdapat beberapa metode yang digunakan pada proses ini, berikut pada gambar 2. dijelaskan terkait alur proses.



Gambar 2. Alur Pre-Processing

3.1 Punctual Removal: Proses penghapusan tanda baca yang tidak penting bagi proses perhitungan kata atau pembobotan kata seperti *https*, *@*, *#*.

3.2 Normalizer : Normalisasi pada penelitian ini guna merubah kata-kata yang tidak baku menjadi baku supaya nantinya lebih mudah dimengerti.

3.3 Stopword Removal : Proses ini adalah menghilangkan kata-kata yang sering muncul namun sifatnya tidak informatif dan tidak memberikan perubahan yang signifikan dalam proses pengolahan sentiment.

3.4 Stemming : pada tahap ini adalah proses untuk mengurangi kata yang terdapat imbuhan di dalamnya dan kemudian diubah menjadi kata dasarnya [15].

3.5 Tokenization : Tahapan ini dilakukan dengan tujuan untuk membagi kalimat menjadi *array token* [14].

4. Term Frequency – Inverse Document Frequency (TF – IDF)

Record yang melalui proses preprocessing masih dalam format word dan harus diubah ke format numerik. Pada penelitian ini pembobotan dilakukan dengan menggunakan TF-IDF. Metode ini digunakan untuk menentukan hubungan antara kata (istilah) dan dokumen dengan memberikan nilai pada setiap kata. Tujuan dari metode ini adalah untuk mencari kata-kata dalam suatu kumpulan yang mungkin mewakili dokumen teks. TF-IDF adalah cara terbaik untuk mengumpulkan informasi.

$$W_{x,n} = TfIdf_{x,n} \times w_{x,n} - 1 \quad (1)$$

5. Naïve Bayes Classifier

Naive Bayes adalah salah satu metode klasifikasi yang populer dan digunakan secara luas. Naive Bayes merupakan metode yang sederhana untuk klasifikasi berbasis probabilitas yang menghitung serangkaian kemungkinan dengan menjumlahkan frekuensi dan kombinasi nilai dari kumpulan data yang diberikan [16]. Berikut adalah rumus dari algoritma ini.

$$P(A|B) = \frac{P(A) \times P(B|A)}{P(B)} \quad (2)$$

6. Evaluasi

Dalam penelitian ini, matriks konfusi atau *confusion matrix* digunakan untuk evaluasi. Dari hasil evaluasi diperoleh nilai kinerja untuk sistem yang dibuat berdasarkan hasil klasifikasi [17]. Berikut adalah perumusan umum yang digunakan pada *confusion matrix*.

Tabel 1. Confusion Matrix

	Prediksi Positif	Prediksi Negatif
Positif Aktual	TP	FN
Negatif Aktual	FP	TN

- True Positif (TP) : Jumlah sampel positif yang sudah diklasifikasikan dengan benar bersama model sebagai positif.
- True Negative (TN) : Jumlah sampel negatif yang sudah diklasifikasikan dengan benar Bersama model sebagai negative.
- False Positif (FP) : Jumlah sampel negative yang salah diklasifikasikan Bersama dengan model sebagai positif.
- False Negatif (FN) : Jumlah sampel positif yang salah diklasifikasikan oleh model sebagai negatif

C. Hasil dan Pembahasan

1. Scrapping Data

Pada penelitian kali ini penulis melakukan metode *scrapping data* pada sosial media yakni twitter. Proses *scrapping data* terhadap unggahan baik itu thread ataupun komenmtar dilakukan dengan menggunakan *tweet-harvest* pada pemrograman *Python*. Pada proses pengumpulan data kali ini, penulis mendapatkan 1151 *tweet* yang berhasil diperoleh dari tahun 2022 sampai dengan 2023. Berikut adalah contoh dari hasil pengumpulan data.

5	kelas mahasiswa skripsi yg galau galau depresi	-1
---	--	----

3. Pre-processing

Tahap kali ini adalah proses mempersiapkan data yang sudah di proses pada tahap sebelumnya, yang mana nantinya data akan diubah ke dalam bentuk yang lebih rinci lagi melalui tahapan pre-processing data yaitu seperti *Punctual Removal, Normalizer, Stopword Removal, Stemming, Tokenization*.

Tabel 4. Hasil Pre-processing Data

No	Pre-Processing	Result
1	Tweet	@lvusmuch Itulah, mending minta saran ke dokter..
2	Punctual Removal	itulah mending minta saran ke dokter..
3	Normalizer	itulah mending minta saran ke dokter..
4	Stopword Removal	mending minta saran dokter sungguhan jangan ma...
5	Stemming	mending minta saran dokter sungguh jangan maha...
6	Tokenization	mending, minta, saran, dokter, sungguh, jangan...

4. Pembobotan TF-IDF

Pada tahap ini penulis menggunakan *library sklearn* dengan mengimport *tfidfvectorizer* yang mana dengan memaksimalkan *frequency* atas *max_features* 800 yang selanjutnya akan digunakan pada pengujian *Data testing* dan *Data Training*. Dengan di *importnya tfidfvectorizer* dapat dilihat bobot dari setiap kata yang ada pada dataset yang digunakan. Berikut merupakan hasil yang di dapatkan beserta programnya.

```
[ ] tweet_suicide = pd.DataFrame(data=tfidf_mat, index=list(range(1, len(tfidf_mat[:,1])+1, )), columns=[a])
tweet_suicide
```

	abadi	aborsi	ada	adalah	adik	agak	ah	aib	air	ajak	...	wibu	wisuda	wkwk	woi	xxx	ya	yang	yogyakarta	you	zaman	
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.177381	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
...
813	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
814	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
815	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
816	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
817	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0

817 rows x 800 columns

Gambar 3. Hasil Pembobotan TF-IDF

5. Visualisasi Data

Pada tahap ini dilakukan proses visualisasi data terhadap hasil pembobotan dengan menggunakan *wordcloud*. Kemudian akan terlihat kata yang sering muncul pada tweet yang dilakukan proses pre-processing dengan tampil lebih mencolok daripada kata yang lain. Berikut hasil dari visualisasi data terhadap sentimen dataset.



Gambar 4. Visualisasi Data menggunakan WordCloud

6. Klasifikasi menggunakan Naive Bayes Classifier

Setelah melalui berbagai tahap proses, selanjutnya data dilakukan proses klasifikasi untuk menganalisis sentimen kasus bunuh diri yang terjadi pada mahasiswa dengan menerapkan algoritma Naive Bayes Classifier yang mana mampu dengan baik memprediksi hasil sentimen dari sebuah dataset.

Tabel 5. Akurasi Naive Bayes

	Precision	Recall	F1-Score	Support
-1	0.90	1.00	0.95	148
1	0.00	0.00	0.00	16

Akurasi Algoritma Naive Bayes Classifier : $0.9024 = 90,24\%$

D. Simpulan

Hasil dari data yang diperoleh sebanyak 1151 Tweet. Kemudian dari data tersebut penulis melakukan polaritas beberapa jenis sentiment yakni positif, negative. Selanjutnya dilakukan pembagian data dengan menggunakan rasio perbandingan 80:20, serta dilakukan proses pembobotan kata dengan menggunakan metode TF-IDF. Dari banyaknya data yang dikumpulkan masih kurang untuk dilakukan penelitian sehingga nilai akurasi dari penggunaan algoritma yakni bernilai 90,24%. Data yang dihasilkan sangat jauh berbeda jumlahnya antara sentiment negative dan positif dikarenakan Tindakan bunuh diri merupakan Tindakan yang negative secara sentiment, maka dari itu hasil sentiment positif yang didapatkan sangat kecil.

Dari hasil analisis yang dilakukan mendapatkan kata “Depresi” dengan kata yang paling banyak muncul pada dataset yang digunakan. Yang mana hal ini juga menjadi salah satu faktor yang menyebabkan terjadinya Tindakan bunuh diri yang dilakukan oleh mahasiswa.

Tentunya pada penelitian ini masih terdapat banyak kekurangan untuk mendukung hasil penelitian yang maksimal. Untuk penelitian kedepannya penulis menyarankan untuk melakukan pengumpulan data yang lebih banyak yang mana dapat diambil dari *platform* lain seperti *facebook* dan Instagram, guna mendapatkan kata-kata yang lebih banyak dan lebih akurat yang menjadi penyebab terjadinya kasus bunuh diri.

E. Referensi

Journal Article

- [1] I. Algoritma *et al.*, “Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi) 919,” Online, 2023.
- [2] R. Wajhillah, S. Bahri, and A. Wibowo, “Komparasi Metode Machine Learning pada Diagnosa Gangguan Kejiwaan Depresi,” 2020. [Online]. Available: <https://www.kaggle.com/everseek/depression>
- [3] T. Anjarsari, I. Ratna, I. Astutik,) Informatika, F. Sains, and D. Teknologi, “DETEKSI DINI GANGGUAN KECEMASAN MENGGUNAKAN METODE NAÏVE BAYES.”
- [4] “Analisis Berbasis Emosional pada Depresi di Media Sosial Menggunakan Pendekatan Convolutional Neural Network.”
- [5] S. Mulyani and R. Novita, “IMPLEMENTATION OF THE NAIVE BAYES CLASSIFIER ALGORITHM FOR CLASSIFICATION OF COMMUNITY SENTIMENT ABOUT DEPRESSION ON YOUTUBE,” *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 5, pp. 1355–1361, Oct. 2022, doi: 10.20884/1.jutif.2022.3.5.374.
- [6] Y. Salim and L. Budi Ilmawan, “Buletin Sistem Informasi dan Teknologi Islam Analisis Sentimen terhadap Komentar Negatif di Media Sosial Facebook dengan Metode Klasifikasi Naïve Bayes INFORMASI ARTIKEL ABSTRAK,” vol. 1, no. 4, pp. 259–265, 2020.
- [7] Y. Nurtikasari, Syariful Alam, and Teguh Iman Hermanto, “Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan

- Algoritma Naive Bayes,” *INSOLOGI: Jurnal Sains dan Teknologi*, vol. 1, no. 4, pp. 411–423, Aug. 2022, doi: 10.55123/insologi.v1i4.770.
- [8] M. Meinaldi Nurrochman, A. Luhur Prasasti, and R. Astuti Nugrahaeni, “IMPLEMENTASI MACHINE LEARNING UNTUK MENDETEKSI UNSUR DEPRESI PADA TWEET MENGGUNAKAN METODE NAÏVE BAYES (MACHINE LEARNING IMPLEMENTATION FOR DEPRESSION DETECTION IN TWEET USING NAÏVE BAYES METHOD).”
- [9] Y. Femilia Nugraini, R. Rohmat Saedudin, and R. Andreswari, “IMPLEMENTASI DATA MINING DALAM KASUS MENTAL HEALTH PADA SOSIAL MEDIA TWITTER MENGGUNAKAN METODE NAIVE BAYES IMPLEMENTATION OF DATA MINING IN THE CASE OF MENTAL HEALTH ON SOCIAL MEDIA TWITTER USING NAIVE BAYES METHOD.”
- [10] A. S. Hutagalung, A. B. P. Negara, and E. E. Pratama, “Aplikasi Pendeteksi Cyberbullying Terhadap Komentar Postingan Media Sosial Instagram dengan Metode Naïve Bayes Classifier Berbasis Website,” *Jurnal Sistem dan Teknologi Informasi (Justin)*, vol. 9, no. 3, p. 364, Aug. 2021, doi: 10.26418/justin.v9i3.44843.
- [11] K. Budiman, N. Zaatsiyah, U. Niswah, F. Muhanna, and N. Faizi, “Analysis of Sexual Harassment Tweet Sentiment on Twitter in Indonesia using Naïve Bayes Method through National Institute of Standard and Technology Digital Forensic Acquisition Approach,” *Journal of Advances in Information Systems and Technology*, vol. 2, no. 2, 2020, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/jaist>
- [12] J. Khatib Sulaiman, A. Shopee Salman Alfaris, and U. Amikom Yogyakarta, “Komparasi Metode KNN dan Naive Bayes Terhadap Analisis Sentimen Pengguna,” *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 5, pp. 2023–2766.
- [13] J. Khatib Sulaiman, P. Perppu Cipta Kerja Nur Muhammad Farhan, B. Setiaji, and U. Amikom Yogyakarta, “Komparasi Metode Naive Bayes dan SVM pada Sentimen Twitter Mengenai,” *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 5, pp. 2023–2718.
- [14] J. Muliawan, E. Dazki, and R. D. Kurniawan, “SENTIMENT ANALYSIS OF INDONESIA’S CAPITAL CITY RELOCATION USING THREE ALGORITHMS: NAÏVE BAYES, KNN, AND RANDOM FOREST,” *Jurnal Teknik Informatika (JUTIF)*, vol. 4, no. 5, pp. 1227–1236, 2023, doi: 10.52436/1.jutif.2023.4.5.347.
- [15] I. Yunanto and S. Yulianto, “TWITTER SENTIMENT ANALYSIS PEDULILINDUNGI APPLICATION USING NAÏVE BAYES AND SUPPORT VECTOR MACHINE,” *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 4, pp. 807–814, Aug. 2022, doi: 10.20884/1.jutif.2022.3.4.292.
- [16] S. Riadi, E. Utami, and A. Yaqin, “Comparison of NB and SVM in Sentiment Analysis of Cyberbullying using Feature Selection,” *sinkron*, vol. 8, no. 4, pp. 2414–2424, Oct. 2023, doi: 10.33395/sinkron.v8i4.12629.
- [17] F. Agus Maulana, I. Ernawati, and J. R. Fatmawati, *ANALISA SENTIMEN CYBERBULLYING DI JEJARING SOSIAL TWITTER DENGAN ALGORITMA NAÏVE BAYES*. 2020.

