

Analisis Sentimen Opini Publik terhadap Fenomena Bunuh Diri Mahasiswa di *Twitter* Menggunakan Algoritma *Naive Bayes*

Fenilinas Adi Artanto

Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Muhammadiyah Pekajangan Pekalongan,
Kabupaten Pekalongan, Indonesia

Email: fenilinasadi@gmail.com

ABSTRACT

Currently, *Twitter* or *X* social media has become a popular social media for interacting and providing opinions on trending events. The opinion that is still widely discussed is the phenomenon of student suicide. With the many opinions on *Twitter* social media about the phenomenon of student suicide, this creates a dilemma about the phenomenon of student suicide, whether student suicide is a positive or negative behavior. For this reason, sentiment analysis was carried out to group opinions on *Twitter* social media. In analyzing this sentiment, the *Naive Bayes* algorithm is used, which is a data mining algorithm that has the advantage of being easy in the calculation stage. The *rapidminer* application is used to analyze opinions on *Twitter* social media. The results obtained by the *Naive Bayes* algorithm provide an accuracy of 60%. Sentiment analysis shows that there are 41% of opinions that have positive sentiment values and there are 59% of opinions that have negative sentiment values. So the majority of opinions about student suicide on *Twitter* social media have negative sentiment values. In collecting opinion data, it also shows that the words *news* *people* and *semarang* are words that are often written on *Twitter* social media.

Keywords: Sentimen Analysis, Text Mining, Naive Bayes, Rapidminer, Student Suicide, Twitter.

ABSTRAK

Saat ini media sosial *twitter* atau *X* telah menjadi media sosial yang populer untuk berinteraksi dan saling memberikan opini terhadap kejadian yang menjadi tren. Opini yang masih banyak dibicarakan adalah fenomena tentang bunuh diri mahasiswa. Dengan banyaknya opini di media sosial *twitter* tentang fenomena bunuh diri mahasiswa, menyebabkan dilema tentang fenomena bunuh diri mahasiswa, apakah bunuh diri mahasiswa merupakan sebuah perilaku yang bersifat positif atau negatif. Untuk itu dilakukan analisis sentimen dalam mengelompokkan opini di media sosial *twitter*. Dalam menganalisis sentimen tersebut digunakan algoritma *naive bayes* yang merupakan algoritma data mining yang memiliki keunggulan kemudahan dalam tahap perhitungannya. Digunakan aplikasi *rapidminer* dalam menganalisis opini di media sosial *twitter*. Hasil yang didapatkan algoritma *naive bayes* memberikan akurasi sebesar 60%. Analisis sentimen menunjukkan bahwa terdapat 41% opini yang memiliki nilai sentimen positif dan terdapat 59% opini yang memiliki nilai sentimen negatif. Sehingga mayoritas opini tentang bunuh diri mahasiswa yang berada di media sosial *twitter* memiliki nilai sentimen negatif. Dalam pengumpulan data opini juga menunjukkan kata *berita* *orang* dan *semarang* yang menjadi kata yang banyak dituliskan dalam media sosial *twitter*.

Kata Kunci: Analisis Sentimen, Text Mining, Naive Bayes, Rapidminer, Bunuh Diri Mahasiswa, Twitter.

1. Pendahuluan

Perkembangan teknologi dan informasi saat ini telah menjadi sangat pesat dengan adanya internet. Penggunaan internet menjadi hal yang digunakan sebagai sarana penyebaran informasi seperti opini ataupun opini dari masyarakat melalui media, yaitu media sosial. Media sosial yang saat ini sering digunakan sebagai media untuk penyebaran informasi dan opini yang populer adalah *twitter*. Saat ini menjadi

topik yang mengundang banyak opini adalah peristiwa tentang mulai banyaknya mahasiswa yang bunuh diri. Fenomena bunuh dirinya mahasiswa menjadi problema di antara masyarakat Indonesia.

Organisasi Kesehatan Dunia (WHO) menyatakan bahwa terdapat rata-rata satu orang meninggal bunuh diri di dunia pada setiap 40 detik, dan Indonesia menempati urutan ke 159 dalam kasus bunuh diri dengan kisaran umur 15 tahun sampai 29 tahun [1].

Dengan meningkatnya bunuh diri di Indonesia, terutama pada rentang umur 15 sampai 29 th yang dimana merupakan usia para mahasiswa menyebabkan banyak pertanyaan dan akhirnya menjadikan topik bunuh diri mahasiswa menjadi topik yang hangat dibahas di media sosial twitter. Banyaknya komentar dan opini dalam media sosial twitter dapat dimanfaatkan dengan mengklasifikasikan tweet sehingga dapat memperoleh informasi. Metode dalam menganalisa sebuah opini atau komentar yang berupa teks merupakan sebuah rumpun ilmu dari analisis sentimen. Sebuah komentar atau opini merupakan sebuah data yang nantinya diperoleh dari tweet dan akan diproses dengan metode *text mining* yang merupakan bagian dari data mining dan proses dilanjutkan dengan mengklasifikasi data kometar atau opini ke dalam dua kelas sentimen yaitu positif, atau negatif [2].

Untuk dapat mengklasifikasikan tweet digunakan algoritma *naive bayes*. Algoritma *naive bayes* merupakan salah satu metode data mining yang memiliki alur perhitungan yang mudah dan tidak panjang [3]. Metode klasifikasi dengan menggunakan algoritma *naive bayes* merupakan metode sederhana yang memiliki performa yang cepat dalam melatih data, mudah dalam implementasinya dan memiliki efektifitas yang tinggi [4] [5].

Dalam penelitian T. Widowati, dan M. Sadikin (2020) melakukan analisis sentiment pada opini di media sosial twitter tentang tokoh publik, dalam metode klasifikasi dengan membandingkan 2 algoritma yaitu *naive bayes* dan *support vector machine*. Dimana hasil penelitian tersebut menunjukkan bahwa *naive bayes* memiliki nilai akurasi sebesar 91.48% dalam melakukan analisis sentimen pada komentar di twitter, sedangkan algoritma *support vector machine* hanya memberikan nilai akurasi sebesar 85.47% dalam melakukan analisis sentimen pada komentar di twitter, sehingga dapat disimpulkan bahwa *naive bayes* lebih unggul dari *support vector machine*.

Dalam penelitian S. Alfaris dan Kusnawi (2023) melakukan penelitian dengan membandingkan metode algoritma *K-Nearest Neighbor* dengan algoritma *naive bayes* untuk melakukan analisis sentiment pada opini pengguna aplikasi shopee. Hasil yang didapatkan menunjukkan bahwa algoritma *K-Nearest Neighbor* menghasilkan nilai akurasi 70% sedangkan algoritma *naive bayes* menghasilkan nilai akurasi sebesar 71%. Hal ini menunjukkan bahwa algoritma *naive bayes* lebih baik dalam menganalisis komentar opini pada aplikasi shopee.

Oleh karena itu dalam mengklasifikasikan opini komentar di media sosial twitter terhadap fenomena bunuh diri mahasiswa digunakanlah algoritma *naive bayes*.

2. Metode Penelitian

2.1. Tinjauan Pustaka

Saputra (2023) melakukan penelitian tentang analisis sentimen dalam membandingkan algoritma *naive bayes* dan *support vector machine* untuk menganalisa piala dunia Fifa 2022. Hasil yang didapatkan menunjukkan bahwa *support vector machine* lebih baik daripada *Naive bayes*.

Fransiska (2019) dalam penelitiannya menggunakan *naive bayes* dalam menganalisa sentimen pelanggan toko online JD.id hasil yang didapatkan menunjukkan bahwa *naive bayes* memberikan nilai akurasi 96,44%. Dan merupakan algoritma yang baik dalam mengklasifikasikan sentiment analisis.

Susanti (2019) menggunakan *naive bayes* dalam mengklasifikasikan opini di twitter tentang review film. Hasil yang didapatkan menunjukkan bahwa *naive bayes* dapat digunakan dalam mengklasifikasikan review film dengan hasil pengklasifikasian sentimen review film menunjukkan bahwa lebih banyak sentimen negatif.

R. Apriani, D. Gustian (2019) menggunakan algoritma *naive bayes* dalam mengklasifikasi komentar pada pengguna aplikasi Tokopedia. Penelitian tersebut menunjukkan bahwa algoritma *naive bayes* merupakan algoritma yang handal dan baik untuk mengklasifikasikan komentar pada penggunaan aplikasi tokopedia dan menghasilkan nilai akurasi sebesar 97.13%.

Pada penelitian R. Fahmi dan N. Nursyfa (2021) dilakukan penelitian dalam menganalisis komentar pengguna twitter tentang sentiment kasus penembakan laskar FPI dengan menggunakan metode klasifikasi *naive bayes*. Pada penelitian tersebut didapatkan bahwa *naive bayes* memberikan performa yang baik yang tergolong pada kategori *excellent agreement* karena *naive bayes* memberikan nilai akurasi sebesar 98.51%, selain itu penelitian tersebut menunjukkan bahwa terdapat 131 komentar positif, 119 komentar netral dan 19 komentar negatif terkait opini di twitter tentang kasus penembakan laskar FPI.

Dari penelitian terdahulu metode *naive bayes* telah banyak digunakan dalam mengklasifikasikan komentar dan opini dalam analisis sentimen dan juga menunjukkan keunggulan algoritma *naive bayes* dalam melakukan klasifikasi sentiment analisis. Dalam penelitian ini, algoritma *naive bayes* akan digunakan dalam menganalisa sebuah fenomena yang dianggap tabu, dalam hal ini adalah bunuh diri, yang merupakan sebuah fenomena yang jarang dibicarakan secara umum. Seperti hal nya penelitian terdahulu milik R. Fahmi dan N. Nursyfa (2021) yang membahas tentang kasus penembakan.

2.2. Landasan Teori

2.2.1. Media Sosial

Media sosial merupakan sebuah situs yang digunakan sebagai tempat bagi masyarakat untuk dapat

berkomunikasi dengan rekan mereka, yang mereka kenal dalam dunia nyata dan juga dunia maya [12][13]. media sosial merupakan sebuah tempat, ataupun layanan yang digunakan individu penggunanya untuk dapat mengekspresikan diri mereka, selain itu juga digunakan untuk dapat bertemu dan juga berbagi dengan rekan lainnya dengan melalui teknologi internet [14]. Media sosial terdiri dari 2 kata, yaitu media dan sosial. Media adalah sebuah perantara, sedangkan sosial adalah suatu hubungan masyarakat. Sehingga, media sosial adalah media perantara untuk dapat saling menghubungkan atau berkomunikasi antar individu atau kelompok tertentu [15].

2.2.2. X (Twitter)

X atau dahulu disebut sebagai Twitter adalah sebuah layanan jejaring sosial yang digunakan sebagai alat komunikasi dalam bentuk media sosial. X atau Twitter dimanfaatkan sebagai media sosial yang dapat dimanfaatkan untuk sarana protes ataupun berkomentar dengan mengirimkan sebuah kalimat berupa opini pengguna yang biasa disebut sebagai ngetweet [16]. X atau Twitter mempunyai fitur yang seperti *tweet*, *mention*, *reply*, *retweet*, *hashtag*, *following*, *follower*, *direct message*, dan *search*. Tweet atau biasa disebut ngetweet adalah sebuah kegiatan untuk post dan posting yang dilakukan pengguna Twitter untuk mengirimkan opini mereka dalam media sosial twitter. Saat ,mengirimkan sebuah postingan yang biasa disebut dengan twett dapat menyampaikan atau mengirimkan postingan juga ke user lainnya dengan menggunakan @username. Username tersebut biasa disebut dengan mention, mention adalah sebuah tools yang memungkinkan pengirim pesan memberikan tanda kepada pengguna user lainnya untuk berbagi pesan, opini atau komentar yang sama dalam twitter [17].

2.2.3. Text Mining

Dalam Data Mining ada sebuah bentuk data yang diproses yaitu data text, atau biasa disebut tex mining. *Text Mining*. *Text mining* merupakan bagian dari proses penambangan data yang digunakan dalam mendapatkan sesuatu hal yang baru yang sebelumnya belum diketahui dan untuk menemukan informasi yang tersirat secara implisit. Temuan dari informasi di ekstrak dari sumber data yang berupa text yang berupa opini atau komentar dalam bentuk kata maupun kalimat [9] [13][18] [19] [20].

2.2.4. Naive Bayes

Salah satu metode dalam mengklasifikasi pada data mining adalah algoritma *Bayesian classification* Dimana pengklasifikasian statistik pada *Bayesian* dapat digunakan dalam melakukan prediksi probabilitas dari keanggotaan suatu class. *Bayes classification* didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan *decision tree* dan *neural network*. *Bayesian classification* terbukti memiliki akurasi dan kecepatan yang tinggi saat

diaplikasikan ke dalam database dengan data yang besar [3] [17]. *Naive Bayes* merupakan salah satu algoritma klasifikasi yang sederhana namun memiliki kemampuan dan akurasi tinggi [21] [22].

$$P(Ci|X) = \frac{P(X|Ci).P(Ci)}{P(X)}$$

Dengan

X= data dengan class yang belum diketahui

Ci= hipotesis data x merupakan suatu class spesifik

$P(Ci|X)$ = probabilitas hipotesis Ci berdasarkan kondisi X (posteriori probability)

$P(Ci)$ = probabilitas hipotesis Ci

$P(X|Ci)$ = probabilitas X berdasarkan kondisi pada hipotesis Ci

$P(X)$ = Probabilitas dari X

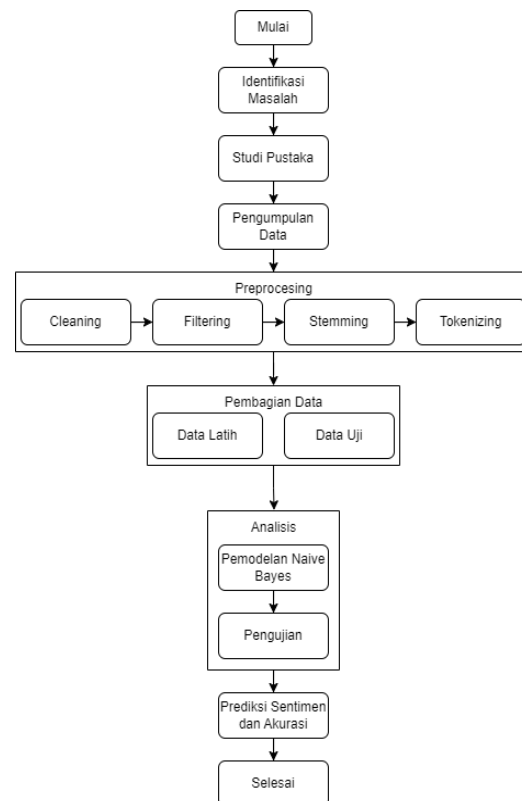
Naïve Bayesian Classifier mengasumsikan bahwa keberadaan sebuah atribut (variabel) tidak ada kaitannya dengan beradaan atribut (variabel) yang lain karena asumsi atribut tidak saling terkait (*conditionally independent*), ditulis dengan rumus:

$$P(X|Ci) = \sum_{k=1}^n P(Xk|Ci)$$

Setelah diperoleh hasil dari seluruh data pada setiap *class*, maka hasil akhirnya dapat menggunakan rumus:

$$P(X|Ci) = \arg \max P(X|Ci) * P(Ci)$$

3. Metode Penelitian



Gambar 1 Tahapan Penelitian

Berikut ini merupakan penjelasan dari alur tahapan penelitian:

1. Identifikasi Masalah

Dalam mengidentifikasi masalah dilakukan proses mengamati dan menemukan sebuah permasalahan yang terjadi dan merupakan permasalahan yang banyak dibahas di media sosial twitter. Dalam mengidentifikasi masalah ditemukan pada media sosial X atau twitter sebuah opini maupun komentar dari sebuah fenomena bunuh diri mahasiswa yang menjadi tren dalam pembahasan di media sosial X atau Twitter.

2. Studi Pustaka

Dalam studi pustaka dilakukan untuk memperdalam teori dari permasalahan tentang *text mining* terutama metode klasifikasi dengan algoritma *Naive Bayes* dan juga tren tentang bunuh diri pada mahasiswa yang ramai dibahas di media sosial X atau twitter.

3. Pengumpulan Data

Dalam proses pengumpulan data pada media sosial x terutama pada pokok pembahasan bunuh diri mahasiswa dilakukan melalui *web scraping* menggunakan bahasa pemrograman python dengan memberikan keyword bunuh diri mahasiswa.

4. Preprocessing

Tahapan *preprocessing* merupakan sebuah tahapan yang dilakukan untuk membersihkan data dari data komentar atau opini yang telah didapat dari media sosial X atau Twitter tentang bunuh diri mahasiswa, agar nantinya data komentar atau opini yang berupa text dapat mudah dia analisis dengan menghilangkan data yang mengganggu seperti *missing value*, data *redundant*, *outliers* ataupun data yang formatnya tidak sesuai dengan kebutuhan dari sistem. Pada tahapan *preprocessing* ini digunakan aplikasi rapidminer dengan versi 10.1.

5. Pembagian Data

Setelah data dibersihkan maka data akan dibagi menjadi 2 kelompok data yaitu data latih dan data uji dimana data latih akan diberikan pelabelan sentiment berupa sentiment negatif maupun positif. Sedangkan pada data uji nantinya akan di uji dengan menggunakan *naive bayes* bagaimana klasifikasi dari sentiment bunuh diri mahasiswa.

6. Analisis

Pada tahapan analisis dilakukan klasifikasi dengan menggunakan data latih yang telah dibagi dengan menggunakan algoritma *naive bayes* menggunakan aplikasi rapidminer.

7. Prediksi

Setelah data latih dilakukan analisa dengan menggunakan algoritma *Naive Bayes* maka hasil

analisis akan dibandingkan dengan data uji sehingga didapatkan hasil dari prediksi sentimen bunuh diri mahasiswa, mana komentar yang bernilai sentimen positif dan mana yang negatif, lalu juga akan diperoleh nilai akurasi dari algoritma *naive bayes* dalam melakukan klasifikasi pada analisis sentimen.

4. Hasil dan Pembahasan

4.1. Pengumpulan Data

Dalam proses pengumpulan data pada media sosial media twitter atau X dilakukan dengan metode *web scraping* dengan menggunakan bahasa pemrograman python dengan meng import library pandas, dan juga note js dimana node JS digunakan untuk menjalankan tweet harvest yang nantinya media yang digunakan adalah google collabs dengan menggunakan sintaks berikut ini:

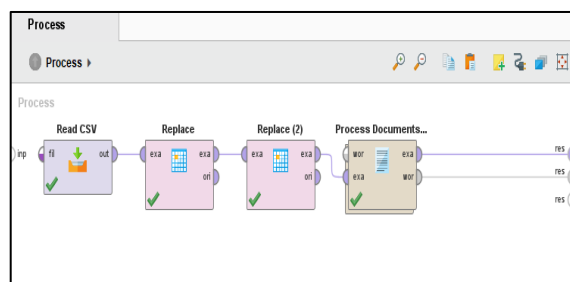
```
filename = 'bunuhdiri.csv'
search_keyword = 'bunuh diri
mahasiswa until:2020-10-30
since:2023-10-30 lang:id'
limit = 500

!npx --yes tweet-harvest@latest -o
"{filename}" -s "{search_keyword}" -
l {limit} --token ""
```

Dengan maksud dari sintaks menunjukan bahwa akan diambil data dari twitter dengan *keyword* bunuh dari mahasiswa yang akan diambil dari 30 oktober 2020 sampai dengan 30 oktober 2023 dengan batasan data yang diambil hanya data bahasa Indonesia dengan limit data sebanyak 500 tweet. Setelah proses *web scraping* didapatkan 511 data tweet.

4.2. Preprocessing

Setelah data sentimen bunuh diri mahasiswa didapatkan maka akan dilakukan proses *preprocessing* atau membersihkan data yang telah didapat menggunakan aplikasi rapidminer, agar saat analisis dapat menjadi lebih akurat, karena data-data yang mengganggu telah dihilangkan, proses *preprocessing* dilakukan dengan desain rapidminer sebagai berikut:



Gambar 2 Desain Preprocessing Rapidminer

Desain tersebut dapat dijelaskan sebagai berikut:

Tabel 1. Hasil Accuracy Naive Bayes

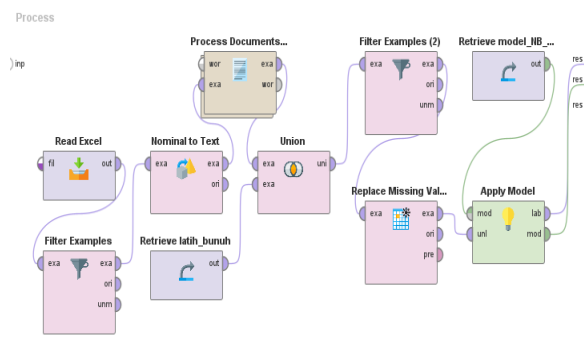
Accuracy: 60.00% +/- 13.33% (micro average: 60.00%)

	true Negatif	true Positif	class percision
pred Negatif	34	23	59.65%
pred Positif	17	26	60.47%
calss recall	66.67%	53.06%	

Dari tabel tersebut memperlihatkan bahwa algoritma *naive bayes* memberikan tingkat akurasi sebesar 60%.

4.5. Prediksi Sentimen Bunuh Diri Mahasiswa

Setelah data latih dilakukan analisis dengan menggunakan *naive bayes* maka data latih tersebut akan diterapkan sebagai dasar untuk membentuk prediksi dari data uji yang belum memiliki label sentiment dengan menggunakan desain rapidminer sebagai berikut:



Gambar 7 Prediksi dengan Naive Bayes

Dari proses rapidminer didapatkan hasil klasifikasi sentiment bunuh diri mahasiswa menggunakan algoritma *naive bayes* sebagai berikut:

Tabel 2. Hasil Sentiment Algoritma Naive Bayes

Jumlah Data	Sentiment
225	Positif
155	Negatif

Dari hasil rapidminer didapatkan hasil prediksi dari analisis sentiment media sosial X atau twitter dengan jumlah 59% twitt yang memiliki sentiment negatif pada bunuh diri mahasiswa dan 41% twitt yang memiliki sentiment positif bunuh diri mahasiswa.

4.6. Pembahasan

Dari proses dalam menganalisa kometar pada twitter tentang bunuh diri mahasiswa, hasil analisis yang telah dilakukan dengan mengambil data dari twitter dan didapatkan 511 data tentang bunuh diri mahasiswa. data tersebut dilakukan proses *preprocessing* data dimana data akan dibersihkan dari data-data komentar atau opini yang tidak diperlukan dan juga memish data yang berupa kalimat menjadi kata-per kata. Dalam pembersihan data terlihat bahwa kata Semarang dan berita menjadi kata terbanyak yang dibahas dan didapatkan dari *twitter* yang membahas tentang bunuh diri mahasiswa. Setelah melalui tahap *preprocessing* data yang telah dibersihkan

tersisa 480 data yang siap untuk di uji. Dari 480 data dibagi menjadi 2 kelompok data yaitu data uji dan data latih. Dimana dibagi menjadi 100 data latih dan 380 data uji. 100 data latih tersebut diberikan label sentiment positif dan negatif sesuai dengan opini peneliti. Dari hasil pelabelan data latih dilakukan analisis menggunakan algoritma *naive bayes* dan didapatkan nilai akurasi sebesar 60%. Setelah itu hasil dari pengelompokan dengan menggunakan *naive bayes* diterapkan untuk dapat memprediksi data uji yang berjumlah 380 data dan didapatkan bahwa terdapat 59% twitt yang memiliki sentiment negatif atau sebanyak 225 komentar pada twitter yang bersentiment negatif dan 41% twitt yang memiliki sentiment positif atau sebanyak 155 opini pada twitter yang memiliki sentiment positif.

5. Kesimpulan dan Saran

5.1. Kesimpulan

Berdasarkan hasil dari analisa dapat ditarik kesimpulan sebagai berikut:

1. Algoritma *naive bayes* memberikan akurasi sebesar 60 % pada analisis sentiment media sosial twitter terhadap bunuh diri mahasiswa.
2. Dari hasil uji didapatkan 59% atau 225 data berupa opini twitter yang bernilai sentiment negatif dan 41 % atau 155 data berupa opini twitter yang bernilai positif tentang fenomena bunuh diri mahasiswa di media sosial twitter.
3. Terdapat kata terbanyak yang ada pada opini bunuh diri mahasiswa adalah kata Semarang dan berita dimana semarang menjadi kata yang menunjukkan sebuah kota dimana opini bunuh diri mahasiswa banyak dibahas.
4. Dari seluruh opini masyarakat di media sosial twitter dapat ditarik kesimpulan bahwa bunuh diri bukanlah sebuah perilaku yang baik, dan lebih berkonotasi negatif. Dimana kota Semarang merupakan kota paling banyak pembahasan tentang fenomena bunuh diri. Sehingga diharapkan ada sebuah kegiatan yang dapat membangun motivasi kepada mahasiswa terutama di Semarang, bahwa bunuh diri bukanlah solusi dalam menyelesaikan masalah.

5.1. Saran

Dari penelitian yang telah dilakukan saran yang dapat diberikan adalah:

1. Dalam proses pengambilan data hanya didapatkan 511 data saja di sebabkan adanya pembatasan pengambilan data pada twitter. Sehingga jumlah data dapat ditambahkan dengan metode *web scrapping* yang lainnya.
2. Pengambilan data hanya dilakukan pada periode 30 oktober 2020 sampai dengan 30 oktober 2023, sehingga untuk dapat mendapatkan data yang lebih lengkap dan banyak batasan pada waktu pengambilan bisa diperluas.

3. Dalam metode analisis dapat menggunakan algoritma klasifikasi lain selain *naive bayes* seperti *support vector machine*, *Decision Tree* atau algoritma yang lain.
4. Dalam melakukan tahapan *Filter Stopwords* jumlah kata yang terfilter baru sedikit, masih terdapat kata-kata seperti singkatan singkatan yang masih terproses dalam analisis, sehingga dapat meningkatkan proses *preprocessing* untuk dalam meningkatkan akurasi metode klasifikasi pada analisis sentiment.

SUMBER RUJUKAN

Referensi

- [1] I. T. Mukaromah, "Problem dan Ide Bunuh Diri Pada Mahasiswa," Universitas Muhammadiyah Surakarta, 2020. [Online]. Available: <http://journal.um-surabaya.ac.id/index.php/JKM/article/view/2203%0Ahttp://mpoc.org.my/malaysian-palm-oil-industry/>
- [2] S. Fransiska and Yolanda, "Analisis sentiment twitter untuk review film menggunakan algoritma naive bayes classifier (nbc) pada sentiment r programming," *J. Siliwangi*, vol. 5, no. 2, pp. 68–71, 2019.
- [3] H. F. Putro, R. T. Vlandari, and W. L. Y. Saptomo, "Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan," *J. Teknol. Inf. dan Komun.*, vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.500.
- [4] S. K. Wardani and Y. A. Sari, "Analisis Sentimen menggunakan Metode Naive Bayes Classifier terhadap Review Produk Perawatan Kulit Wajah menggunakan Seleksi Fitur N-gram dan Document Frequency Thresholding," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 12, pp. 5582–5590, 2021.
- [5] I. Farida and S. W. H. L. Hendric, "Prediksi Pola Kelulusan Mahasiswa Menggunakan Teknik Data Mining Classification Emerging Pattern," *Petir*, vol. 12, no. 1, pp. 1–17, 2019, doi: 10.33322/petir.v12i1.414.
- [6] T. T. Widowati and M. Sadikin, "Analisis Sentimen Twitter terhadap Tokoh Publik dengan Algoritma Naive Bayes dan Support Vector Machine," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 11, no. 2, pp. 626–636, 2021, doi: 10.24176/simet.v11i2.4568.
- [7] S. Alfaris and Kusnawi, "Komparasi Metode KNN dan Naive Bayes Terhadap Analisis Sentimen Pengguna Aplikasi Shopee," *Indones. J. Comput. Sci.*, vol. 12, no. 5, pp. 2766–2776, 2023, doi: 10.33022/ijcs.v12i5.3304.
- [8] P. R. Saputra Andri, Subing Mulia, "Perbandingan Metode Naive Bayes Classifier Dan Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter Mengenai Piala Dunia Fifa 2022," *Teknomatika*, vol. 13, no. 01, pp. 22–31, 2023, [Online]. Available: <http://ojs.palcomtech.ac.id/index.php/teknomatika/article/view/616>
- [9] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naive Bayes Classifier Berbasis Konversi Ikon Emosi," *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.
- [10] R. Apriani and D. Gustian, "Analisis Sentimen dengan Naive Bayes Terhadap Komentar Aplikasi Tokopedia," *J. Rekayasa Teknol. Nusa Putra*, vol. 6, no. 1, pp. 54–62, 2019, [Online]. Available: <https://rekayasa.nusaputra.ac.id/article/view/86>
- [11] R. N. Fahmi, N. Nursyifa, and A. Primajaya, "Analisis Sentimen Pengguna Twitter Terhadap Kasus Penembakan Laskar Fpi Oleh Polri Dengan Metode Naive Bayes Classifier," *JIKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 2, pp. 61–66, 2021, [Online]. Available: <https://ejournal.akakom.ac.id/index.php/jiko/article/view/437/0>
- [12] M. Masril and F. W. Lubis, "Analisis Penggunaan Media Sosial dan Penyebaran Hoax Di Kota Medan," *J. SIMBOLIKA Res. Learn. Commun. Study*, vol. 6, no. 1, pp. 11–22, 2020, doi: 10.31289/simbollika.v6i1.2937.
- [13] H. H. Kusumawardani, I. Rosyadi, F. A. Artanto, F. I. Arzha, and N. A. Rachmayani, "Analisis Decision Tree dalam Pengaruh Digital Marketing terhadap Penerimaan Siswa Baru," *Remik*, vol. 6, no. April, pp. 225–231, 2022.
- [14] A. Setiadi, "Pemanfaatan media sosial untuk efektifitas komunikasi," *J. Ilm. Matrik*, vol. 16, no. 1, 2014.
- [15] A. Fatkhudin, F. A. Artanto, N. A. Safli, and D. Wibowo, "Decision Tree Berbasis SMOTE dalam Analisis Sentimen Penggunaan Artificial Intelligence untuk Skripsi," *Remik Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 8, no. April, pp. 494–505, 2024.
- [16] F. A. Artanto, "Support Vector Machine Berbasis Particle Swarm Optimization Pada Analisis Sentimen Anggota KPPS," *J. FASILKOM (teknologi Inf. dan Ilmu KOMputer)*, vol. 14, no. 1, pp. 75–79, 2024, doi: <https://doi.org/10.37859/jf.v14i1.6795>.
- [17] A. Wahyuningtyas, I. S. Sitanggang, and H. Khotimah, "Deteksi Spam pada Twitter Menggunakan Algoritme Naive Bayes," *J. Ilmu Komput. dan Agri-Informatika*, vol. 7, no. 1, pp. 31–40, 2020, doi: 10.29244/jika.7.1.31-40.
- [18] A. Fatkhudin, A. Khambali, F. A. Artanto, and N. A. P. Zade, "Implementasi Algoritma Clustering K-Means Dalam Pengelompokan Mahasiswa Studi Kasus (Prodi Manajemen Informatika)," *J. Minfo Polgan*, vol. 12, no. 2, pp. 777–783, 2023, doi: 10.33395/jmp.v12i2.12494.
- [19] I. Rosyadi, F. A. Artanto, S. E. Rahmawati, H. Tri, and B. Joyo, "Decision Tree Dalam Analisis Keputusan Pembelian Program Pada Perkumpulan Penggiat Programmer Indonesia," *J. Fasilkom*, vol. XII, no. III, pp. 141–144, 2022.
- [20] M. F. I. Haq, I. Rosyadi, M. Nasir, and A. Khambali, "Sentiment Analisis Ulasan Aplikasi Livin Pada Google Play Store," *J. Surya Inform.*, vol. 14, no. 1, pp. 24–29, 2024.
- [21] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 160–165, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.
- [22] S. Hadiani and F. Yosep Tember, "Analisis Sentiment Covid-19 Di Twitter Menggunakan Metode Naive Bayes Dan SVM," *J. Teknol. Inf.*, vol. 6, no. 1, pp. 58–63, 2022, [Online]. Available: www.kaggle.com.