



Analisis Sentimen *Twitter* Terhadap Tren Penyebaran Informasi Pelaku Kejahatan Menggunakan Algoritma *Naives Bayes*

Awwaliyah Aliyah

Institut Teknologi Sepuluh Nopember

Email: aliyahawwaliyah@gmail.com

Nailah Azzahra

Institut Teknologi Sepuluh Nopember

Email: nailahazhr@gmail.com

Aliffia Isma Putri

Institut Teknologi Sepuluh Nopember

Email: aliffiaisma28@gmail.com

Nur Aini Rakhmawati

Institut Teknologi Sepuluh Nopember

Email: nur.aini@is.its.ac.id

Institut Teknologi Sepuluh Nopember, Jl. Teknik Kimia, Keputih, Kec. Sukolilo, Surabaya

Korespondensi Penulis : aliyahawwaliyah@gmail.com

Abstract

In the rapidly developing digital era, social media such as Twitter has become part of everyday life and facilitates the rapid dissemination of information, including information about criminals. This research aims to analyze public sentiment towards information about criminals spread on Twitter using the Naive Bayes algorithm. This algorithm was chosen because of its simplicity and effectiveness in text classification. Data was collected through a crawling process from Twitter, followed by a preprocessing stage to remove noise. The research results show that public sentiment towards information about criminals on Twitter is divided into three categories: positive, neutral and negative. After classification, it was found that neutral sentiment increased significantly to 63.4%, while positive and negative sentiment decreased to 10.5% and 26.1%. These findings indicate that people tend to be more careful in reacting to sensitive information. This research provides important insights for related parties in managing information about criminals on social media and can be a reference for developing further policies and strategies.

Keywords: *Criminals, Naive Bayes Algorithm, Sentiment Analysis, Social Media, Twitter*

Abstrak

Dalam era digital yang semakin berkembang pesat, media sosial seperti Twitter telah menjadi bagian dari kehidupan sehari-hari dan memfasilitasi penyebaran informasi dengan cepat, termasuk informasi tentang pelaku kejahatan. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap informasi tentang pelaku kejahatan yang tersebar di Twitter dengan menggunakan algoritma Naive Bayes. Algoritma ini dipilih karena kesederhanaannya dan efektivitasnya dalam klasifikasi teks. Data dikumpulkan melalui proses crawling dari Twitter, diikuti dengan tahap preprocessing untuk menghilangkan noise. Hasil penelitian menunjukkan bahwa sentimen masyarakat terhadap informasi pelaku kejahatan di Twitter terbagi dalam tiga kategori: positif, netral, dan negatif. Setelah dilakukan klasifikasi, ditemukan bahwa sentimen netral meningkat secara signifikan menjadi 63,4%, sementara sentimen positif dan negatif menurun menjadi 10,5% dan 26,1%. Temuan ini mengindikasikan bahwa masyarakat cenderung lebih berhati-hati dalam bereaksi terhadap informasi sensitif. Penelitian ini memberikan wawasan penting bagi pihak terkait dalam mengelola informasi tentang pelaku kejahatan di media sosial dan dapat menjadi acuan untuk pengembangan kebijakan dan strategi lebih lanjut.

Kata kunci: Algoritma Naive Bayes, Analisis Sentimen, Media Sosial, Pelaku Kejahatan, Twitter

Received April 30, 2024; Accepted Mei 24, 2024; Published Mei 31, 2024

* Awwaliyah Aliyah, aliyahawwaliyah@gmail.com

LATAR BELAKANG

Dalam era digital yang semakin berkembang pesat, media sosial telah menjadi bagian dari kehidupan sehari-hari. Perkembangan teknologi informasi terutama pada media sosial, telah mengubah proses penyebaran informasi dan interaksi antar manusia menjadi semakin cepat dan terbuka. Kecepatan informasi ini membuat isu ataupun permasalahan menyebar secara luas dalam waktu beberapa jam saja, termasuk informasi tentang pelaku kejahatan (Fitriani & Pakpahan, 2020).

Informasi tentang pelaku kejahatan yang tersebar di media sosial sering kali mendapatkan perhatian besar dari masyarakat (Marwan, 2017, dikutip dalam Rahmah, 2023). Sentimen yang muncul dari berbagai postingan dan komentar bisa sangat beragam, mulai dari yang negatif, netral, hingga positif. Analisis sentimen terhadap informasi ini menjadi penting untuk memahami bagaimana masyarakat bereaksi terhadap isu-isu kejahatan yang berkembang.

Twitter sebagai salah satu *platform* media sosial terbesar, menjadi sumber data yang berharga untuk melakukan analisis sentimen. Menurut laporan *We Are Social*, ada sekitar 27,5 juta pengguna *Twitter* alias X di Indonesia per Oktober 2023, jumlah itu menempatkan Indonesia di peringkat keempat global (Annur, 2023). Dengan jumlah pengguna yang sangat besar dan beragam, *Twitter* menyediakan data yang kaya untuk memahami dinamika opini publik. Analisis sentimen pada *platform* ini dapat memberikan wawasan penting tentang persepsi masyarakat terhadap pelaku kejahatan dan bagaimana informasi tersebut mempengaruhi opini publik.

Dalam penelitian ini, digunakan algoritma *Naive Bayes* untuk melakukan analisis sentimen terhadap *tweet-tweet* yang berkaitan dengan pelaku kejahatan. Algoritma *Naive Bayes* dipilih karena kesederhanaannya dan kemampuannya yang cukup efektif dalam tugas klasifikasi teks (Darwis dkk., 2021). Dengan menggunakan algoritma ini, diharapkan dapat mengidentifikasi pola-pola sentimen yang dominan dan memahami bagaimana tren penyebaran informasi tentang pelaku kejahatan berkembang di *Twitter*.

Penelitian ini juga bertujuan untuk mengevaluasi akurasi algoritma *Naive Bayes* dalam mengklasifikasikan sentimen dari data *Twitter*. Hasil dari analisis ini diharapkan dapat memberikan kontribusi dalam memahami bagaimana informasi tentang pelaku kejahatan menyebar di media sosial dan bagaimana masyarakat bereaksi terhadap informasi tersebut. Dengan demikian, penelitian ini dapat menjadi acuan bagi pihak-pihak terkait dalam mengelola dan menyikapi informasi tentang pelaku kejahatan yang beredar di media sosial.

Beberapa penelitian yang sebelumnya telah dilakukan berkaitan dengan analisis sentimen pada *Twitter*. Penelitian Darwis, dkk. (2021) yang mengekstraksi data dari *Twitter* BMKG Nasional dan menganalisis sentimen menggunakan *Naive Bayes*, mencapai akurasi 68,97%. Penelitian Fahmi, dkk. (2021) yang melakukan analisis sentimen terhadap kasus penembakan anggota FPI oleh Polri menggunakan *Naive Bayes* menghasilkan akurasi tinggi (98,51%) dan menunjukkan bahwa banyaknya komentar positif mengindikasikan permintaan masyarakat untuk investigasi lebih lanjut, namun perlu perbandingan dengan metode lain, seleksi fitur, dan penambahan kosakata Bahasa Inggris. Penelitian Toy, dkk. (2021) yang menggunakan *Naive Bayes* untuk analisis sentimen opini masyarakat di *Twitter* tentang kebijakan *New Normal* mencapai akurasi tertinggi 73,3%, dengan rata-rata akurasi meningkat dari 62,6% menjadi 65,3% setelah seleksi fitur *Relevance Frequency Feature Selection*.

KAJIAN TEORITIS

Twitter

Twitter adalah platform media sosial dan layanan microblogging daring yang memungkinkan pengguna untuk mengirim dan membaca pesan teks singkat yang dikenal dengan sebutan "tweet" (Surapati & Zulkarnain, 2023). Pengguna Twitter dapat mengirim tweet untuk berbagi informasi, pemikiran, berita, atau opini dengan pengikut mereka. Selain itu, pengguna juga dapat mengunggah gambar, video, dan tautan ke dalam tweet mereka. Twitter memungkinkan interaksi yang cepat dan langsung antara pengguna serta memfasilitasi berbagai topik pembicaraan yang sedang tren atau viral.

Pelaku Kejahatan

Dalam Pasal 55 Ayat 1 Kitab Undang-Undang Hukum Pidana (KUHP) Indonesia, pelaku kejahatan didefinisikan sebagai berikut:

1. Mereka yang melakukan, yang menyuruh melakukan, dan yang turut serta melakukan perbuatan.
2. Mereka yang dengan memberi atau menjanjikan sesuatu, dengan menyalahgunakan kekuasaan atau martabat, dengan kekerasan, ancaman atau penyesatan, atau dengan memberi kesempatan, sarana atau keterangan, sengaja menganjurkan orang lain supaya melakukan perbuatan.

Dari uraian tersebut dapat disimpulkan bahwa pelaku kejahatan tidak hanya terbatas pada mereka yang secara langsung melakukan perbuatan, tetapi juga mereka yang memfasilitasi atau mendorong perbuatan tersebut.

Privasi

Privasi mencakup pada hak asasi setiap individu untuk mengontrol informasi dan aktivitas pribadi mereka. Hal ini memungkinkan mereka untuk mengetahui dengan jelas apa yang terjadi dengan data pribadi mereka, siapa yang mengaksesnya, tujuan penggunaan data tersebut, serta bagaimana data tersebut diolah dan disimpan (Suari & Sarjana, 2023).

Dalam konteks penyebaran informasi pelaku kejahatan di media sosial, perlindungan privasi menjadi sangat penting karena informasi yang tersebar mungkin mengandung data sensitif atau informasi pribadi tentang individu yang belum terbukti bersalah.

Perlindungan Data Pribadi

Menurut Pasal 1 Undang-Undang Nomor 27 Tahun 2022 tentang Perlindungan Data Pribadi (UU PDP), dijelaskan bahwa “Data Pribadi adalah data tentang orang perseorangan yang teridentifikasi atau dapat diidentifikasi secara tersendiri atau dikombinasi dengan informasi lainnya baik secara langsung maupun tidak langsung melalui sistem elektronik atau nonelektronik”. Definisi ini menjadi landasan penting dalam konteks perlindungan data pribadi karena menegaskan bahwa data pribadi individu harus dijaga dari akses, penggunaan, atau pengungkapan yang tidak sah atau tidak diinginkan (Fikri & Rusdiana, 2023).

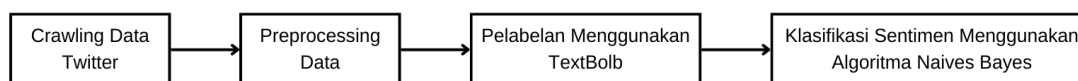
Dengan memahami bahwa data pribadi mencakup informasi yang dapat mengidentifikasi individu, baik secara langsung maupun tidak langsung, maka perlindungan data pribadi menjadi krusial untuk mencegah potensi penyalahgunaan atau akses yang tidak sah terhadap informasi sensitif tersebut. Melalui implementasi kebijakan dan praktik perlindungan data yang sesuai dengan prinsip-prinsip yang berlaku, seperti yang diatur dalam UU PDP, diharapkan bahwa data pribadi individu dapat terlindungi dengan baik dan privasi mereka tetap terjaga.

Algoritma Naive Bayes

Algoritma *Naive Bayes* adalah salah satu metode klasifikasi yang banyak digunakan dalam *machine learning*, terutama dalam bidang data mining dan text mining (Wijaya et al., 2016, seperti dikutip dalam Darwis, 2021). Algoritma ini didasarkan pada teorema *Bayes* dengan asumsi "*naive*" atau sederhana, yaitu asumsi bahwa setiap fitur dalam dataset adalah independen satu sama lain. Secara umum, algoritma *Naive Bayes* digunakan untuk memprediksi kelas atau label dari data berdasarkan probabilitas dan statistik dari fitur-fitur yang ada, sehingga memungkinkan pengklasifikasian data ke dalam kategori yang telah ditentukan sebelumnya (Amolik et al., 2016, seperti dikutip dalam Darwis, 2021).

METODE PENELITIAN

Penelitian ini menggunakan analisis sentimen untuk mengkategorikan tweet ke dalam sentimen positif, netral, dan negatif. Data yang dikumpulkan dianalisis menggunakan analisis sentimen TextBlob Python. Adapun tahapan metode penelitian dirinci dalam gambar 1 berikut.



Gambar 1. Tahapan Penelitian

Crawling Data Twitter

Crawling merupakan proses untuk memperoleh informasi tertentu dari halaman *website* dan menyimpannya secara offline (Priansya, 2017). Dalam penelitian ini, *crawling* dilakukan pada laman media sosial Twitter menggunakan *keyword* #Pelaku dengan memanfaatkan Application Programming Interface (API) yang dikembangkan oleh twitter. Data-data yang diambil merupakan data publik berupa tweet, waktu pemosting, username pengguna, dan keterangan lain yang diperlukan.

Preprocessing Data

Data yang telah melewati proses *crawling* disimpan dalam bentuk CSV untuk digunakan pada tahap *preprocessing data*. Tahap *preprocessing data* dilakukan untuk menghilangkan *noise* untuk memberikan hasil optimal. Fitur-fitur NLTK digunakan sebab NLTK mendukung berbagai proses tahapan *preprocessing* seperti tokenisasi, stemming, dan filtering (Parlika, dkk., 2020). Berikut rincian dari setiap sub-tahapan dalam tahap *preprocessing*.

1. *Translate*: Proses untuk mengubah bahasa data hasil *crawling* menjadi bahasa Inggris untuk memudahkan dalam tahap pelabelan dan klasifikasi.
2. *Cleaning Data*: Proses untuk menghilangkan *noise* data hasil *crawling* dengan menghapus karakter, tanda baca, hashtag, alamat URL, angka, dan username.
3. *Text Normalization*: Proses untuk mengubah teks pada sebuah dokumen dari kata yang tidak tepat atau kata singkatan bahasa gaul dan salah ketik dan yang disingkat menjadi kata yang memiliki arti.
4. *Case Folding*: Proses untuk menyeragamkan semua huruf dalam dokumen menjadi huruf kecil.
5. *Tokenizing*: Proses untuk memecah sebuah teks berdasarkan pada tiap kata yang menyusunnya.

6. *Filtering*: Proses untuk mengambil kata-kata penting dari hasil token.
7. *Stemming*: Proses untuk mencari *root* kata dari setiap kata hasil dari tahap *filtering*.

Pelabelan Data dengan TextBlob

Setelah tahap preprocessing selesai, langkah selanjutnya adalah melakukan pelabelan data. Penelitian ini menggunakan pelabelan otomatis dengan bantuan TextBlob yang menyediakan API untuk pemrosesan NLP. TextBlob memungkinkan berbagai fungsi NLP seperti ekstraksi frase kata benda, analisis sentimen, klasifikasi, penerjemahan, dan lainnya (Parlika, dkk., 2020).

Klasifikasi dengan Algoritma Naive Bayes

Metode Naïve Bayes Classification (NBC) adalah teknik yang menggunakan statistik sederhana berdasarkan teorema Bayes, dengan asumsi bahwa kehadiran atau ketiadaan suatu kelas tidak bergantung pada fitur lainnya. Metode NBC digunakan untuk klasifikasi dengan tujuan menentukan nilai *f1_score*, akurasi, recall, dan presisi (Azhar, Surahman, & Juliane, 2022).

Fitur-fitur NLTK digunakan untuk klasifikasi text mining menggunakan metode Naïve Bayes. NLTK mendukung berbagai proses pengolahan bahasa alami seperti klasifikasi, tokenisasi, stemming, tagging, parsing, dan filtering. Kemudian untuk pelabelan, dalam penelitian ini menggunakan Python TextBlob yang menyediakan API yang untuk analisis sentimen dan klasifikasi.

Visualisasi

Visualisasi digunakan untuk mengidentifikasi pola dan menginterpretasi data dengan lebih mudah. Dalam penelitian ini, visualisasi data dilakukan menggunakan bar chart memanfaatkan *library* seaborn. Visualisasi dilakukan dua kali, sebelum dan setelah melakukan klasifikasi untuk membandingkan perubahan data yang terjadi.

HASIL DAN PEMBAHASAN

Hasil Crawling Data

Data yang terkumpul dari hasil crawling sebanyak 506 cuitan berdasarkan pengambilan data pada tanggal 16 Mei 2024. Terdapat batasan jumlah cuitan yang dapat diambil hingga 600 cuitan per hari berdasarkan kebijakan baru Twitter yang diterapkan oleh Elon Musk, sehingga hasil crawling data tidak mencapai angka 1000.

	full_text	username	created_at
0	Rada sangsi anjir sama kasus ini kalo misal p...	Mokasellingacc	Thu May 16 21:19:10 +0000 2024
1	@RadioElshinta apa karena Vina bukan orang kay...	FoodaholicPr	Thu May 16 21:17:21 +0000 2024
2	@humaspoldajbr Teman Pelaku udah tertangkap.....	MPrimbon	Thu May 16 21:14:54 +0000 2024
3	Ketika tindakan pelaku bully di sekolah terlal...	Enji_PW	Thu May 16 21:14:41 +0000 2024
4	@donaldarmen @bacottetangga_ Yg kakak adik itu...	gula_jawwa	Thu May 16 21:06:54 +0000 2024
...
501	gw selalu akan berpikir kalo orang yg permisif...	criminace	Thu May 16 10:36:55 +0000 2024
502	Solusi parkir liar: Memberantas pelaku Dikelol...	haryonoagung	Thu May 16 10:35:42 +0000 2024
503	Rp. 35.520.000.000. Atas perbuatannya para pel...	BeltimSi	Thu May 16 10:35:20 +0000 2024
504	Selanjutnya baby lobster tersebut masuk kesala...	BeltimSi	Thu May 16 10:35:18 +0000 2024
505	udah ada berapa ya pelaku sexual harassment yg...	peacecespiscis	Thu May 16 10:35:01 +0000 2024

506 rows x 3 columns

Gambar 2. Hasil *Crawling* Data

Hasil *Preprocessing*

Setelah melakukan *preprocessing*, yang mencakup tokenisasi, penghapusan stopwords, dan stemming, data hasil *preprocessing* ditampilkan dengan mengambil satu teks (satu baris data) sebagai sampel .

Tabel 1. Hasil *Preprocessing*

Sub-tahapan <i>Preprocessing</i>	Cuitan
Data hasil <i>Crawl</i>	@opposite6892 Kasus yg di bogor juga blm terungkap siswi yg nama noven yg tusuk sampai tewas ..udah lama bgt itu juga pelaku ya blm tertangkap
<i>Cleaning</i>	Kasus yg di bogor juga blm terungkap siswi yg nama noven yg tusuk sampai tewas udah lama bgt itu juga pelaku ya blm tertangkap
<i>Translate</i>	The case in Bogor has not yet been revealed the female student whose name was Noven who was stabbed to death it's been a long time since the perpetrator was caught
<i>Case Folding</i>	the case in bogor has not yet been revealed the female student whose name was noven who was stabbed to death its been a long time since the perpetrator was caught

<i>Normalizing</i>	the case in bogor has not yet been revealed the female student who was stabbed to death its been a long time since the perpetrator was caught
<i>Tokenizing</i>	['the', 'case', 'in', 'bogor', 'has', 'not', 'yet', 'been', 'revealed', 'the', 'female', 'student', 'who', 'was', 'stabbed', 'to', 'death', 'its', 'been', 'a', 'long', 'time', 'since', 'the', 'perpetrator', 'was', 'caught']
<i>Filtering</i>	['case', 'bogor', 'yet', 'revealed', 'female', 'student', 'stabbed', 'death', 'long', 'time', 'since', 'perpetrator', 'caught']
<i>Stemming</i>	case bogor yet reveal female student stab death long time since perpetr caught

Pada tahap *cleaning*, baris data direduksi dengan penghapusan data yang duplikasi maupun tidak relevan dengan topik sehingga menghasilkan baris data akhir berjumlah 459 baris. Kemudian data diterjemahkan ke dalam bahasa Inggris untuk mempermudah peneliti dalam melakukan pelabelan dan klasifikasi.

Hasil Pelabelan

Pelabelan dilakukan dengan menganalisis sentimen dari setiap tweet dalam dataset, mengklasifikasikannya sebagai positif, netral, atau negatif berdasarkan polaritas yang dihitung oleh TextBlob. Berikut merupakan tabel dari pelabelan klasifikasi sentimen untuk setiap tweet, yang terbagi menjadi 3 sentimen, 'Positif', 'Netral', atau 'Negatif'.

	processed_text	klasifikasi
0	im littl doubt case exampl perpetr arrest impr...	Positif
1	girl rich perpetr son polic offic	Positif
2	perpetr friend caught clear name clear address...	Positif
3	action bulli school often condon parent teache...	Positif
4	brother sister victim wife perpetr wife husban...	Negatif
...
454	would alway think person permiss behavior like...	Netral
455	illeg park solut erad statemanag perpetr	Netral
456	rp 3552 billion action perpetr charg articl 92...	Positif
457	next babi lobster enter one hous contract perp...	Netral
458	mani perpetr sexual harass kebraon peopl tri e...	Positif

459 rows x 2 columns

Gambar 3. Tabel Hasil Pelabelan Data menggunakan TextBlob

Persebaran klasifikasi juga dicatat dan ditampilkan dalam bentuk ringkasan jumlah tweet untuk setiap kategori sentimen serta total tweet yang dianalisis.

```

Hasil Analisis Data:
Positif = 155
Netral = 148
Negatif = 156

Total Data : 459

```

Gambar 3. Ringkasan Hasil Pelabelan Data

Ringkasan hasil menunjukkan pelabelan sentimen positif sebanyak 155 *tweet* atau sebesar 33,8%, sentimen netral sebanyak 148 *tweet* atau sebesar 32,2%, dan sentimen negatif sebanyak 156 *tweet* atau sebesar 34%. Hasil pelabelan sebelum melakukan klasifikasi menunjukkan sentimen masyarakat Twitter terhadap penyebaran informasi pelaku kejahatan publik menyebar secara merata sebab tidak terdapat perbedaan yang signifikan antara setiap klasifikasi sentimen.

Klasifikasi Menggunakan *Naive Bayes*

Klasifikasi dilakukan dengan mengelompokkan data berdasarkan sentimen, kemudian melakukan sampling acak untuk menyeimbangkan data, dan melatih model *Naive Bayes* menggunakan `TextBlob`. Akurasi model kemudian diuji terhadap dataset asli untuk mengevaluasi kinerja model. Berikut merupakan akurasi yang diperoleh dari hasil *train model* menggunakan *Naive Bayes*.

```

# Cek Akurasi
from textblob.classifiers import NaiveBayesClassifier
cl = NaiveBayesClassifier(train_set)
print("Akurasi Test:", cl.accuracy(dataset))

Akurasi Test: 0.7647058823529411

```

Gambar 4. Akurasi dari Klasifikasi *Naive Bayes*

Akurasi yang diberikan melalui klasifikasi dengan algoritma *Naive Bayes* sebesar 0.76. Setelah melakukan klasifikasi dengan algoritma *Naive Bayes*, dilakukan pelabelan kembali untuk menganalisis perubahan sentimen dari data. Persebaran sentimen ditampilkan dalam bentuk ringkasan jumlah tweet untuk setiap kategori sentimen serta total tweet yang dianalisis.

Hasil Analisis Data:
Positif = 48
Netral = 291
Negatif = 120

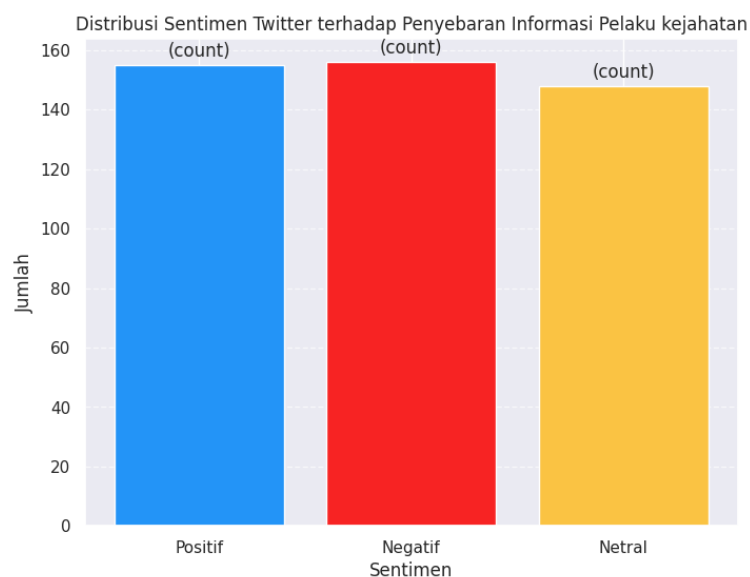
Total Data : 459

Gambar 5. Ringkasan Pelabelan Data Setelah Klasifikasi

Pelabelan kedua menunjukkan bahwa sentimen positif turun menjadi 48 tweet atau sebesar 10,5%, sentimen netral naik menjadi 291 tweet atau sebesar 63,4%, dan sentimen negatif turun menjadi 120 tweet atau sebesar 26,1%. Hasil pelabelan setelah melakukan klasifikasi menunjukkan bahwa sentimen masyarakat Twitter terhadap penyebaran informasi pelaku kejahatan publik cenderung lebih netral. Penurunan dalam sentimen positif dan negatif, serta kenaikan signifikan dalam sentimen netral, mengindikasikan bahwa masyarakat mungkin lebih berhati-hati dalam bereaksi terkait fenomena ini.

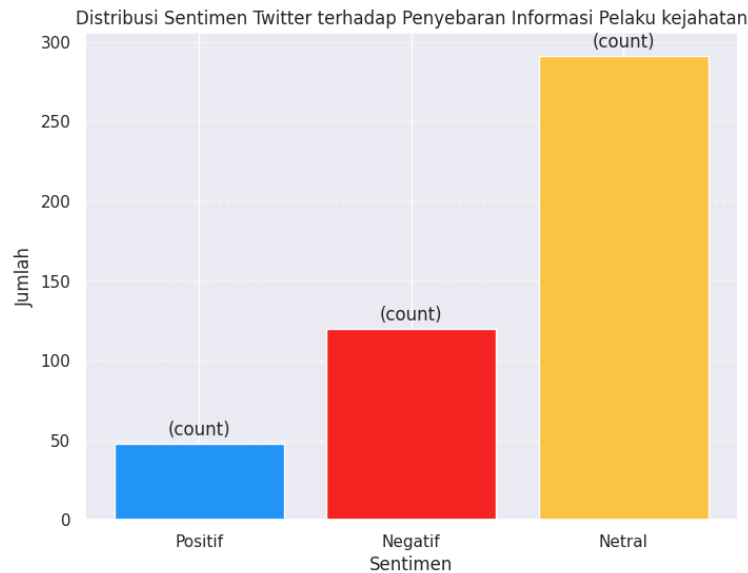
Visualisasi Bar Chart

Dalam penelitian ini, hasil pelabelan sentimen divisualisasikan menggunakan bar chart dari library Seaborn. Pelabelan dibandingkan antara pelabelan sebelum melakukan klasifikasi dengan setelah melakukan klasifikasi untuk memahami perubahan sentimen masyarakat Twitter terhadap penyebaran informasi pelaku kejahatan publik. Berikut bar chart pelabelan sebelum klasifikasi.



Gambar 6. Bar Chart Pelabelan Data Sebelum Klasifikasi

Pelabelan yang dilakukan sebelum klasifikasi menunjukkan bahwa sentimen masyarakat Twitter terhadap penyebaran informasi cenderung merata. Sentimen positif sebanyak 155 *tweet* atau sebesar 33,8%, sentimen netral sebanyak 148 *tweet* atau sebesar 32,2%, dan sentimen negatif sebanyak 156 *tweet* atau sebesar 34%. Pelabelan sebelum klasifikasi kemudian dibandingkan dengan pelabelan setelah klasifikasi yang dilampirkan pada gambar 7.



Gambar 7. Bar Chart Pelabelan Data Setelah Klasifikasi

Pelabelan setelah klasifikasi menunjukkan bahwa masyarakat Twitter cenderung netral terhadap fenomena penyebaran informasi pelaku kejahatan. Sentimen positif turun menjadi 48 *tweet* atau sebesar 10,5%, sentimen netral naik menjadi 291 *tweet* atau sebesar 63,4%, dan sentimen negatif turun menjadi 120 *tweet* atau sebesar 26,1%.

Hasil pelabelan menunjukkan perubahan yang signifikan dalam persepsi masyarakat Twitter terhadap penyebaran informasi pelaku kejahatan setelah klasifikasi, dimana sebelum dilakukan klasifikasi, sentimen masyarakat cenderung merata dengan distribusi yang hampir seimbang antara positif, netral, dan negatif. Sedangkan ketika klasifikasi telah dilakukan, sentimen netral meningkat tajam, sedangkan sentimen positif dan negatif menurun. Hal ini mengindikasikan bahwa setelah klasifikasi, masyarakat Twitter lebih cenderung bersikap netral terhadap penyebaran informasi pelaku kejahatan publik. Perubahan ini mungkin mencerminkan kelelahan informasi, ketidakpastian, atau sikap yang lebih berhati-hati terhadap fenomena ini.

KESIMPULAN DAN SARAN

Simpulan dari penelitian ini menunjukkan bahwa setelah dilakukan klasifikasi, masyarakat Twitter cenderung bersikap lebih netral dalam menanggapi penyebaran informasi pelaku kejahatan publik. Hal ini dapat diartikan bahwa terdapat kecenderungan masyarakat untuk lebih berhati-hati dan netral dalam menyikapi informasi yang sensitif seperti ini.

Adapun prospek pengembangan dari hasil penelitian ini meliputi peningkatan teknik crawling untuk memperoleh lebih banyak data, serta pengembangan model analisis sentimen dengan menggunakan algoritma *machine learning* lainnya seperti Support Vector Machine (SVM) atau *deep learning*. Selain itu, penerapan lebih lanjut dari penelitian ini dapat digunakan untuk memonitor dan menganalisis tren sentimen masyarakat terhadap isu-isu sensitif lainnya di media sosial.

DAFTAR REFERENSI

- Aliyah, A., Azzahra, N., Putri, A. I., & Rakhmawati, N. A. (2024). Dataset for Twitter Sentiment Analysis on Criminal Data Propagation using Naive Bayes Algorithm (1.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.11243533>.
- Annur, C., M. (2023). Ada 27 Juta Pengguna Twitter di Indonesia, Terbanyak ke-4 Global. Databoks. Available at: <https://databoks.katadata.co.id/datapublish/2023/11/28/ada-27-juta-pengguna-twitter-di-indonesia-terbanyak-ke-4-global>. Diakses tanggal 19 Mei 2024.
- Azhar, R., Surahman, A., & Juliane, C. (2022). Analisis Sentimen Terhadap Cryptocurrency Berbasis Python TextBlob Menggunakan Algoritma Naïve Bayes. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 6(1): 267-281.
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, 15(1), 131-145. <https://doi.org/10.33365/jtk.v15i1.744>.
- Fahmi, R., N., Nursyifa, & Primajaya, A. (2021). Analisis Sentimen Pengguna Twitter Terhadap Kasus Penembakan Laskar FPI oleh Polri dengan Metode Naive Bayes Classifier. *Jurnal Informatika dan Komputer*, 5(2). <https://dx.doi.org/10.26798/jiko.v5i2.437>.
- Fikri, M. & Rusdiana, S. (2023). RUANG LINGKUP PERLINDUNGAN DATA PRIBADI: KAJIAN HUKUM POSITIF INDONESIA. *GANESHA LAW REVIEW*, 5(1).
- Fitriani, Y. & Pakpahan, R. (2020). Analisa Penyalahgunaan Media Sosial untuk Penyebaran Cybercrime di Dunia Maya atau Cyberspace. *Cakrawala: Jurnal Humaniora Bina Sarana Informatika*, 20(1), 21-27. <https://doi.org/10.31294/jc.v19i2>.
- Parlika, R., Pradika, S. I., Hakim, A. M., & K. R. N. M. (2020). Analisis Sentimen Twitter Terhadap Bitcoin dan Cryptocurrency Berbasis Python TextBlob. *Jurnal Ilmiah Teknologi Informasi dan Robotika*, 2(2): 33–37.

- Priansya, S. (2017). *NORMALISASI TEKS MEDIA SOSIAL MENGGUNAKAN WORD2VEC, LEVENSHTTEIN DISTANCE, DAN JARO-WINKLER DISTANCE*. Tugas Akhir. Surabaya: Institut Teknologi Sepuluh Nopember.
- Suari, K. R. A. & Sarjana, I M. (2023). Menjaga Privasi di Era Digital: Perlindungan Data Pribadi di Indonesia. *JURNAL ANALISIS HUKUM*, 6(1): 132-146.
- Surapati, U. & Zulkarnain, A. Y. (2023). IMPLEMENTASI METODE NAÏVE BAYES UNTUK MENDETEKSI HATE SPEECH PADA TWITTER. *Journal of Information Technology and Computer Science (INTECOMS)*, 6(2): 830-837.
- Rahmah Tamhidah, M. A. (2023). Pengaruh Media Sosial Terhadap Penyebaran Informasi Palsu dan Kejahatan Siber. *Innovative: Journal Of Social Science Research*, 3(6), 9133–9147. <https://doi.org/10.31004/innovative.v3i6.7197>.
- Toy, K. V. S., Sari, Y. A., & Cholissodin, I. (2021). Analisis Sentimen Twitter menggunakan Metode Naive Bayes dengan Relevance Frequency Feature Selection (Studi Kasus: Opini Masyarakat mengenai Kebijakan New Normal). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(11), 5068–5074. Diambil dari <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/10172>.