



Analisis Sentimen Kegiatan Pembersihan Sampah Pada Media Sosial X Menggunakan SVM dan Naïve Bayes

Dendy Aprilianto Nugroho, Firman Noor Hasan*

Fakultas Teknologi Industri dan Informatika, Teknik Informatika, Universitas Muhammadiyah Prof. Dr. Hamka, DKI Jakarta, Indonesia

Email: ¹dendyapriyanto98@gmail.com, ^{2,*}firman.noorhasan@uhamka.ac.id

Email Penulis Korespondensi: firman.noorhasan@uhamka.ac.id

Abstrak—Aktivitas kegiatan sehari-hari manusia tidak luput dari sampah, hampir semua kegiatan yang dilakukan manusia akan menimbulkan sampah. Kebiasaan buruk manusia dengan membuang sampah sembarangan yang mengakibatkan pengaruh terhadap keseimbangan lingkungan. Akibat permasalahan ini, ada komunitas anak muda bernama Pandawara Grup yang ingin membantu membersihkan sampah di Pantai Sukabumi, namun aksi tersebut ditolak oleh Kepala Desa dan Karang Taruna setempat yang membuat banyak opini-opini yang menghasilkan sentimen masyarakat yang terdapat pada tweet di media sosial X. Karena hal ini, penelitian ini dilakukan agar dapat menganalisis sentimen terhadap kegiatan pembersihan sampah yang dilakukan Pandawara Grup di pantai Sukabumi dengan menggunakan metode Support Vector Machine dan Naïve Bayes. Tujuan dari dibuatnya penelitian ini yaitu untuk mengetahui sentimen negatif maupun positif masyarakat terhadap permasalahan ini, serta menganalisis perbandingan akurasi dari metode Support Vector Machine dan Naïve Bayes. Pada analisis sentimen ini dengan menggunakan 2.339 dataset hingga menghasilkan accuracy tertinggi pada metode Support Vector Machine sebesar 91,67%, sedangkan pada metode Naïve Bayes hanya sebesar 63,89%, sehingga ditarik kesimpulan bahwa metode Support Vector Machine lebih baik untuk mengklasifikasikan sentimen terhadap kegiatan pembersihan sampah yang dilakukan Pandawara Grup di pantai Sukabumi dibandingkan metode Naïve Bayes.

Kata Kunci: Analisis Sentimen; Kegiatan Pembersihan Sampah; Media Sosial X; Support Vector Machine; Naïve Bayes

Abstract—Human daily activities inevitably produce waste, which negatively impacts environmental balance due to the bad habit of indiscriminately disposing of waste. As a result of this issue, there is a youth community named Pandawara Group that wants to help clean up trash on Sukabumi Beach. However, their initiative faced rejection from the local village chief and youth organization, sparking various opinions on social media platform X. Consequently, this research seeks to analyze public sentiment towards Pandawara Group's waste cleanup efforts at Sukabumi Beach using Support Vector Machine and Naïve Bayes methods. The objective is to gauge positive and negative sentiments and compare the accuracy of Support Vector Machine and Naïve Bayes. In this sentiment analysis using 2,339 datasets, the highest accuracy was achieved using the Support Vector Machine method at 91.67%, whereas the Naïve Bayes method only achieved 63.89%. Thus, it can be concluded that Support Vector Machine is superior in classifying sentiments regarding Pandawara Group's waste cleanup activities at Sukabumi Beach compared to Naïve Bayes.

Keywords: Sentiment Analysis; Trash Cleaning Activity; Social Media X; Support Vector Machine; Naïve Bayes

1. PENDAHULUAN

Era teknologi modern ini membuat internet semakin berkembang begitu pesat dengan mudahnya mengakses suatu internet bagi manusia dan perkembangan teknologi yang semakin canggih, [1] seperti contohnya teknologi smartphone dan jaringan internet yang sudah menuju sinyal 5G untuk memudahkan pengguna dalam mengakses suatu internet agar mendapatkan informasi dengan cepat dalam berbagai aspek seperti aspek kehidupan masyarakat, bisnis, pendidikan, dan masih banyak lainnya [2]. Hal ini membuat masyarakat lebih sering menggunakan media sosial yang bisa diakses oleh internet untuk mengakses sebuah informasi dan menjalin komunikasi dari jarak dekat maupun jarak jauh daripada menggunakan media cetak, seperti koran, majalah, atau media cetak lainnya yang dahulunya lebih sering digunakan [3].

Berkembangnya internet dan pengguna media sosial yang semakin banyak, serta kemudahan yang diberikan seperti salah satu contohnya yaitu media sosial X [4]. Pengguna internet pastinya sering menggunakan platform jejaring sosial, platform jejaring sosial yang banyak diakses pengguna internet salah satunya yaitu media sosial X, berdasarkan hasil riset We Are School tahun 2024, pengguna media sosial telah mencapai 167 juta jiwa dengan total keseluruhan populasi jiwa sebesar 276,4 juta jiwa [5]. Pada media sosial X, pengguna biasanya membagikan sebuah postingan atau membagikan sebuah pesan singkat terkait aktivitas sehari-hari antar pengguna ke pengguna lainnya yang disebut dengan tweet [6]. Oleh karena itu, tweet inilah yang akan digunakan untuk melihat bagaimana sentimen dari penggunaannya apakah menghasilkan opini-opini yang negatif atau positif mengenai isu tertentu yang terdapat pada media sosial X, selain itu media sosial X juga merupakan salah satu media sosial yang dapat mengutarakan pendapatnya secara bebas dengan opini tanpa batas yang dapat memudahkan para peneliti untuk melakukan analisis sentimen dan juga media sosial X merupakan sarana terbaik [7].

Pada aktivitas kegiatan sehari-hari manusia tidak luput dari sampah, hampir semua kegiatan yang dilakukan manusia akan menimbulkan sampah. Hingga saat ini, permasalahan global yang belum terpecahkan yaitu terkait sampah yang dihasilkan melalui aktivitas manusia [8]. Seiring berjalannya waktu dengan populasi yang akan semakin meningkat, jika tidak seimbang lahan pembuangan juga akan mempengaruhi sampah yang akan semakin bertambah [9]. Kebiasaan buruk manusia yang diawali dengan membuang sampah sembarangan yang



mengakibatkan pengaruh terhadap keseimbangan lingkungan, maka diperlukan langkah yang lebih mendalam dalam diberlakukannya kebijakan mengenai pengelolaan sampah [8].

Salah satu negara dengan populasi terpadat di dunia yaitu Indonesia dengan jumlah sampah meningkat pada awal tahun 2021 menjadi 28,5 juta ton/hari hingga tahun 2022 menjadi 37,4 juta ton/hari [10]. Sesuai hukum Pemerintahan Daerah dalam Undang-Undang Nomor 23 Tahun 2014, masalah tentang sampah termasuk tanggungjawab pemerintahan yang diserahkan ke daerah [11]. Namun, ada hal yang terjadi belakangan ini mengenai sebuah komunitas anak muda bernama Pandawara Grup yang ingin membantu membersihkan sampah di Pantai Loji tepatnya di daerah Sukabumi, namun aksi pembersihan sampah yang akan dilakukan tersebut ditolak oleh Kepala Desa dan Karang Taruna setempat dikarenakan Pandawara Grup mengatakan bahwa pantai tersebut merupakan pantai terkotor ke-4 di Indonesia dan tidak meminta izin terlebih dahulu yang membuat Kepala Desa Sangrawayang tidak terima atas pemberian predikat pantai terkotor ke-4 di Indonesia dan merasa bahwa pihak pemerintah seolah-olah tidak dilibatkan dalam kegiatan pembersihan sampah tersebut [12]. Kejadian tersebut yang viral, lalu menimbulkan banyak opini-opini sentimen masyarakat yang terdapat di tweet pada media sosial X untuk dilakukannya analisis sentimen untuk kepuasan terhadap suatu layanan [13].

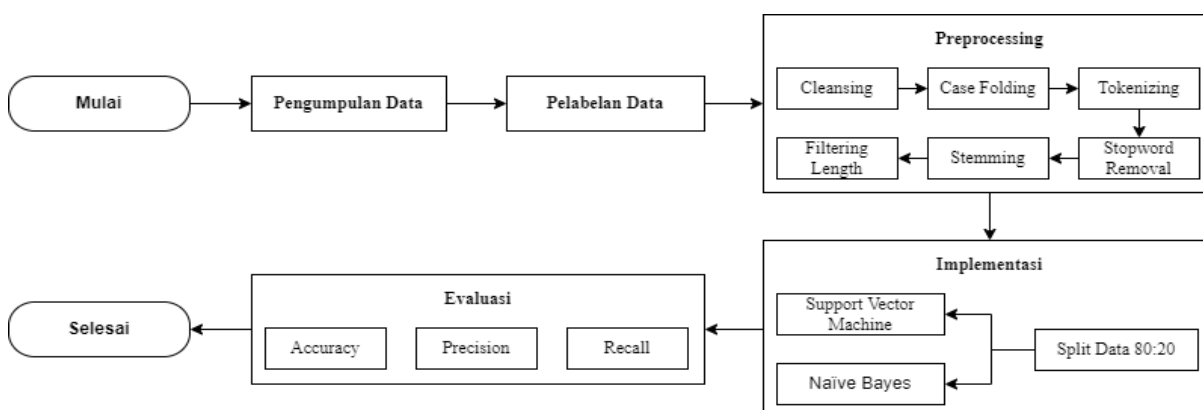
Analisis sentimen merupakan suatu teknik mendapatkan data komentar terhadap masyarakat untuk membantu menganalisis isi dari data yang didapatkan berupa pandangan ataupun opini negatif maupun positif terhadap suatu kejadian [14]. Data yang sudah didapatkan berupa komentar atau opini, kemudian dilakukan analisis sentimen untuk mengetahui sentimen perorangan yang dapat dijadikan tolak ukur untuk mengevaluasi suatu isu atau suatu layanan ke depannya [15].

Pada penelitian sebelumnya mengenai analisis sentimen terhadap pembobolan dan kebocoran data menggunakan metode Naïve Bayes dan Support Vector Machine dengan 1.083 dataset yang digunakan hingga menghasilkan akurasi 97,67% pada metode Naïve Bayes dan 73,75% pada metode Support Vector Machine [16]. Penelitian lainnya yaitu membandingkan metode Support Vector Machine dan Naïve Bayes terkait kebijakan penghapusan kewajiban skripsi dengan 1.303 dataset menghasilkan akurasi dari algoritma Naïve Bayes sebesar 75,00% dan Support Vector Machine sebesar 80,00% [17]. Penelitian lainnya mengenai penggunaan layanan BPJS dengan membandingkan ketiga algoritma yaitu KNN, Decision Tree, dan Naïve Bayes dengan data yang di mining sebanyak 903 dataset, lalu dari ketiga metode tersebut metode yang menghasilkan akurasi paling tinggi yaitu Decision Tree sebesar 96,13%, kedua KNN sebesar 95,58%, dan terakhir yaitu Naïve Bayes 89,14% [18]. Penelitian lainnya mengenai gadget berbagai aspek dengan membandingkan ketiga algoritma yaitu Naïve Bayes, SVM, dan k-NN dengan total 1.391 dataset hingga mendapatkan metode yang menghasilkan akurasi tertinggi yaitu metode SVM sebesar 96,43% [19]. Penelitian lainnya yaitu mengkomparasi metode Support Vector Machine dan Naïve Bayes dengan 2.334 dataset hingga menghasilkan akurasi paling tinggi pada algoritma Support Vector Machine sebesar 97,56% sedangkan Naïve Bayes sebesar 96,52% [20].

Dalam penelitian ini, peneliti memilih untuk membandingkan metode Support Vector Machine dan Naïve Bayes mengenai kegiatan pembersihan sampah yang dilakukan Pandawara Grup di pantai Sukabumi, dikarenakan pada penelitian terdahulu belum pernah ada terkait analisis sentimen pada kegiatan pembersihan sampah yang dilakukan Pandawara Grup di pantai Sukabumi. Tujuan dari penelitian ini yaitu, dari data yang sudah didapatkan, lalu dilakukan analisis sentimen untuk mengetahui sentimen negatif maupun positif, serta menganalisis perbandingan akurasi dari kedua metode Support Vector Machine dan Naïve Bayes. Penelitian ini berkontribusi salah satunya adalah terkait perbandingan dengan membandingkan Support Vector Machine dan Naïve Bayes, peneliti mengetahui berapa akurat dari kedua metode ini.

2. METODOLOGI PENELITIAN

Tahap-tahap pada metode penelitian yang digunakan pada penelitian ini yaitu seperti yang terdapat pada gambar 1.



Gambar 1. Metodologi Penelitian



Tahapan pertama pada metodologi penelitian ini yaitu pengumpulan data, dimana data didapatkan melalui komentar tweet media sosial X yang di crawling menggunakan tools Google Colab, kemudian data yang sudah didapatkan dijadikan file berformat csv, selanjutnya data yang sudah berformat csv dilakukan pelabelan secara negatif maupun positif secara manual oleh peneliti, lalu data yang sudah berlabel sentimen dilanjutkan dalam tahapan preprocessing. Pada tahap preprocessing merupakan tahapan pembersihan data yang semula data berisi data yang sangat kotor menjadi data yang bersih dan lebih terstruktur agar dapat dengan mudah untuk diolah, pada tahapan ini berupa menghilangkan mention, hastag, link, spasi, simbol-simbol dan data duplikat, tidak hanya itu tetapi juga merubah semua sentimen menjadi huruf kecil, kemudian memisahkan sentimen menjadi kata per kata, lalu menghapus kata-kata yang tidak relevan atau tidak baku, selanjutnya kata-kata yang relevan atau baku dijadikan kata-kata dasar dengan menghilangkan imbuhan-imbuhannya, terakhir menghapus kata yang tidak terdiri dari 4 – 25 karakter. Sentimen yang sudah melalui tahap preprocessing, lalu dibagi menjadi 80% data uji dan 20% data latih yang kemudian dilanjutkannya tahapan pengimplementasian kedua metode yaitu metode Support Vector Machine dan Naïve Bayes. Metode Support Vector Machine adalah teknik pembelajaran mesin yang mencari hyperplane terbaik untuk memisahkan kelas dalam data dengan margin maksimum dan dapat menangani data non-linear menggunakan fungsi kernel [21], sedangkan Naïve Bayes adalah salah satu algoritma klasifikasi dalam pembelajaran mesin berdasarkan teorema Bayes dengan menyederhanakan perhitungan probabilitas kelas dalam data [22]. Data yang sudah dilakukan tahapan pengujian metode, terakhir yaitu pengujian evaluasi untuk melihat dari sisi accuracy, precision, recall dari metode Support Vector Machine dan Naïve Bayes.

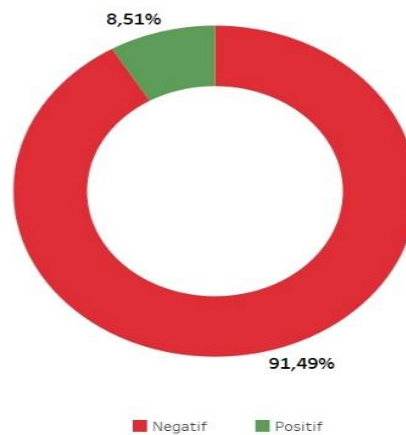
3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Tahap awal pada penelitian ini yaitu mengumpulkan data yang didapatkan melalui komentar tweet media sosial X dengan kata kunci pencarian “Pandawara Sukabumi” pada tanggal 6 Oktober 2023, kemudian data di crawling menggunakan bahasa pemrograman Python pada tools Google Colab, lalu dataset di export ke dalam file dengan format csv, seperti gambar 2 di bawah yang menunjukkan proses pengumpulan data.



Gambar 2. Proses Pengumpulan Data



Gambar 3. Presentase Keseluruhan Sentimen

3.2 Pelabelan Data

Data yang sudah didapatkan, kemudian tahap selanjutnya yaitu melabelkan dataset secara manual pada tools Microsoft Excel dengan label sentimen negatif dan positif. Pada proses pelabelan dihasilkan 2.876 sentimen negatif dan 251 sentimen positif seperti yang terdapat pada gambar 3 dari total keseluruhan 3.127 data, serta terdapat contoh dari beberapa data sentimen komentar tweet media sosial X yang terdapat pada tabel 1.

Tabel 1. Proses Pelabelan Dataset

Text	Sentimen
@dkristiawan_ malu kali tuh ketauan gitu....???? Ketauan selama ini si karang taruna ama Kadesnya kayak becus kerjanya	Negatif



Text	Sentimen
Gimana nggak di cap jadi pantai terkotor kalo sampahnya udh sebanyak itu. Kalo gamau dicap jadi pantai terkotor ya bersihin dong pak! Giliran ada yg mau bantuin malah ga terima. Anehh banget kepala desa dan ketua karang tarunannya	Negatif
Perkara pantainya diklaim jadi pantai terkotor Alhamdulillah Pandawara berhasil menginspirasi dan meningkatkan kesadaran pemerintah. tidak perlu ajakan ekstrim cukup aksi nyata dan akan ada pihak2 yg tersentil olehnya	Positif
Hebat sih pandawara grup berkat jasa mereka memviralkan pantai kotor tersebut jadi ada gerakan untuk membersihkannya. Segenap rakyat indonesia pastinya sangat berterimakasih kepada PG	Positif

3.3 Preprocessing

Data yang telah terlabeli, lalu dilakukan tahap selanjutnya yaitu preprocessing. Pada tahapan preprocessing dataset diolah yang semula data masih sangat kotor menjadi data yang bersih dan lebih terstruktur menggunakan tools RapidMiner untuk memudahkan dilakukan analisis sentimen. Tahap preprocessing seperti contohnya yang terdapat pada gambar 4, 5, dan 6, diantaranya yaitu:

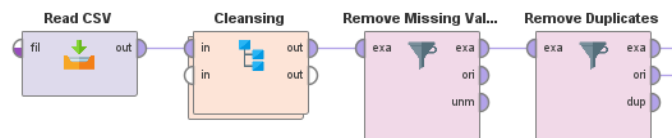
a. Cleansing



Gambar 4. Proses Cleansing Dataset

Pada proses cleansing terbagi lagi menjadi bagian-bagian seperti yang terdapat pada gambar 4, yaitu:

- 1. Replace URL**
Proses cleansing pertama yaitu replace URL untuk menghilangkan link-link yang terdapat pada komentar tweet.
- 2. Replace Simbol**
Proses cleansing kedua yaitu replace simbol untuk menghilangkan simbol-simbol yang terdapat pada komentar tweet.
- 3. Replace Mention**
Proses cleansing ketiga yaitu replace mention untuk menghilangkan mention atau tag kepada pengguna media sosial lainnya yang terdapat pada komentar tweet.
- 4. Replace Hastag**
Proses cleansing keempat yaitu replace hastag untuk menghilangkan kalimat bertagar yang terdapat pada komentar tweet.
- 5. Trim**
Proses cleansing kelima yaitu trim untuk menghilangkan spasi-spasi yang tidak berguna yang terdapat pada komentar tweet.



Gambar 5. Proses Missing Values dan Remove Duplicates

- 6. Remove Missing Values**
Proses cleansing keenam yaitu remove missing values untuk menghilangkan komentar tweet yang tidak memiliki nilai atau isi berupa komentar tweet di dalamnya.
- 7. Remove Duplicates**
Proses cleansing ketujuh yaitu remove duplicates untuk menghilangkan komentar tweet yang sama atau lebih dari satu tetapi memiliki isi yang sama pada komentar tweet.

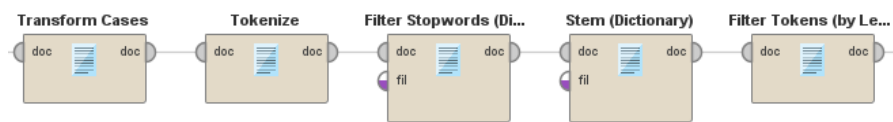
Setelah dilakukan proses cleansing yang semula data masih sangat kotor menjadi data yang bersih dan lebih terstruktur agar lebih mudah untuk diolah analisis sentimennya, seperti contohnya yang terdapat pada tabel 2.

Tabel 2. Proses Cleansing Dataset

Sebelum	Sesudah
@dkristiawan_ malu kali tuh ketauan gitu....???? Ketauan selama ini si karang taruna ama Kadesnya kagak becus kerjanya	malu kali tuh ketauan gitu Ketauan selama ini si karang taruna ama Kadesnya kagak becus kerjanya



Sebelum	Setelah
<p>Gimana nggak di cap jadi pantai terkotor kalo sampahnya udh sebanyak itu. Kalo gamau dicap jadi pantai terkotor ya bersihin dong pak! Giliran ada yg mau bantuin malah ga terima. Anehh banget kepala desa dan ketua karang tarunannya Perkara pantainya diklaim jadi pantai terkotor</p> <p>Alhamdulillah Pandawara berhasil menginspirasi dan meningkatkan kesadaran pemerintah. tidak perlu ajakan ekstrim cukup aksi nyata dan akan ada pihak2 yg tersentil olehnya</p> <p>Hebat sih pandawara grup berkat jasa mereka memviralkan pantai kotor tersebut jadi ada gerakan untuk membersihkannya. Sege nap rakyat indonesia pastinya sangat berterimakasih kepada PG</p> <p>https://t.co/qbz7n4Fcm t</p>	<p>Gimana nggak di cap jadi pantai terkotor kalo sampahnya udh sebanyak itu Kalo gamau dicap jadi pantai terkotor ya bersihin dong pak Giliran ada yg mau bantuin malah ga terima Anehh banget kepala desa dan ketua karang tarunannya Perkara pantainya diklaim jadi pantai terkotor</p> <p>Alhamdulillah Pandawara berhasil menginspirasi dan meningkatkan kesadaran pemerintah tidak perlu ajakan ekstrim cukup aksi nyata dan akan ada pihak yg tersentil olehnya</p> <p>Hebat sih pandawara grup berkat jasa mereka memviralkan pantai kotor tersebut jadi ada gerakan untuk membersihkannya Sege nap rakyat indonesia pastinya sangat berterimakasih kepada PG</p>



Gambar 6. Proses Case Folding sampai Filtering Length

b. Case Folding

Data yang sudah dilakukan proses cleansing, selanjutnya dilakukan proses case folding. Case folding merupakan proses mengubah seluruh huruf pada dataset menjadi huruf kecil menggunakan operator transform cases pada tools RapidMiner [23], seperti contohnya yang terdapat pada tabel 3.

Tabel 3. Proses Case Folding

Sebelum	Setelah
<p>malu kali tuh ketauan gitu Ketauan selama ini si karang taruna ama Kadesnya kagak becus kerjanya</p> <p>Gimana nggak di cap jadi pantai terkotor kalo sampahnya udh sebanyak itu Kalo gamau dicap jadi pantai terkotor ya bersihin dong pak Giliran ada yg mau bantuin malah ga terima Anehh banget kepala desa dan ketua karang tarunannya Perkara pantainya diklaim jadi pantai terkotor</p> <p>Alhamdulillah Pandawara berhasil menginspirasi dan meningkatkan kesadaran pemerintah tidak perlu ajakan ekstrim cukup aksi nyata dan akan ada pihak yg tersentil olehnya</p> <p>Hebat sih pandawara grup berkat jasa mereka memviralkan pantai kotor tersebut jadi ada gerakan untuk membersihkannya Sege nap rakyat indonesia pastinya sangat berterimakasih kepada PG</p>	<p>malu kali tuh ketauan gitu ketauan selama ini si karang taruna ama kadesnya kagak becus kerjanya</p> <p>gimana nggak di cap jadi pantai terkotor kalo sampahnya udh sebanyak itu kalo gamau dicap jadi pantai terkotor ya bersihin dong pak giliran ada yg mau bantuin malah ga terima ane hh banget kepala desa dan ketua karang tarunannya perkara pantainya diklaim jadi pantai terkotor</p> <p>alhamdulillah pandawara berhasil menginspirasi dan meningkatkan kesadaran pemerintah tidak perlu ajakan ekstrim cukup aksi nyata dan akan ada pihak yg tersentil olehnya</p> <p>hebat sih pandawara grup berkat jasa mereka memviralkan pantai kotor tersebut jadi ada gerakan untuk membersihkannya sege nap rakyat indonesia pastinya sangat berterimakasih kepada pg</p>

c. Tokenizing

Proses selanjutnya yaitu Tokenizing, Tokenizing merupakan proses memisahkan komentar tweet yang seolah sebuah kalimat menjadi bagian kata per kata [24], seperti contohnya yang terdapat pada tabel 4.

Tabel 4. Proses Tokenizing

Sebelum	Setelah
<p>malu kali tuh ketauan gitu ketauan selama ini si karang taruna ama kadesnya kagak becus kerjanya</p> <p>gimana nggak di cap jadi pantai terkotor kalo sampahnya udh sebanyak itu kalo gamau dicap jadi pantai terkotor ya bersihin dong pak giliran ada yg mau bantuin malah ga terima ane hh banget kepala desa dan ketua karang</p>	<p>['malu', 'kali', 'tuh', 'ketauan', 'gitu', 'ketauan', 'selama', 'ini', 'si', 'karang', 'taruna', 'ama', 'kadesnya', 'kagak', 'becus', 'kerjanya']</p> <p>['gimana', 'nggak', 'di', 'cap', 'jadi', 'pantai', 'terkotor', 'kalo', 'sampahnya', 'udh', 'sebanyak', 'itu', 'kalo', 'gamau', 'dicap', 'jadi', 'pantai', 'terkotor', 'ya', 'bersihin', 'dong', 'pak', 'giliran', 'ada', 'yg', 'mau', 'bantuin', 'malah', 'ga', 'terima', 'anehh', 'banget', 'kepala', 'desa', 'dan',</p>



Sebelum	Sesudah
tarunannya perkara pantainya diklaim jadi pantai terkotor	‘ketua’, ‘karang’, ‘tarunannya’, ‘perkara’, ‘pantainya’, ‘diklaim’, ‘jadi’, ‘pantai’, ‘terkotor’]
alhamdulillah pandawara berhasil menginspirasi dan meningkatkan kesadaran pemerintah tidak perlu ajakan ekstrim cukup aksi nyata dan akan ada pihak yg tersentil olehnya	[‘alhamdulillah’, ‘pandawara’, ‘berhasil’, ‘menginspirasi’, ‘dan’, ‘meningkatkan’, ‘kesadaran’, ‘pemerintah’, ‘tidak’, ‘perlu’, ‘ajakan’, ‘ekstrim’, ‘cukup’, ‘aksi’, ‘nyata’, ‘dan’, ‘akan’, ‘ada’, ‘pihak’, ‘yg’, ‘tersentil’, ‘olehnya’]
hebat sih pandawara grup berkat jasa mereka memviralkan pantai kotor tersebut jadi ada gerakan untuk membersihkannya segenap rakyat indonesia pastinya sangat berterimakasih kepada pg	[‘hebat’, ‘sih’, ‘pandawara’, ‘grup’, ‘berkat’, ‘jasa’, ‘mereka’, ‘memviralkan’, ‘pantai’, ‘kotor’, ‘tersebut’, ‘jadi’, ‘ada’, ‘gerakan’, ‘untuk’, ‘membersihkannya’, ‘segenap’, ‘rakyat’, ‘indonesia’, ‘pastinya’, ‘sangat’, ‘berterimakasih’, ‘kepada’, ‘pg’]

d. Stopword Removal

Sentimen yang sudah di Tokenizing, lalu dilakukan proses Stopword Removal yang merupakan proses penghapusan kata-kata yang tidak relevan atau tidak baku yang bereferensi dari file acuan pada website kaggle [25], seperti contohnya yang terdapat pada tabel 5.

Tabel 5. Proses Stopword Removal

Sebelum	Sesudah
[‘malu’, ‘kali’, ‘tuh’, ‘ketauan’, ‘gitu’, ‘ketauan’, ‘selama’, ‘ini’, ‘si’, ‘karang’, ‘taruna’, ‘ama’, ‘kadesnya’, ‘kagak’, ‘becus’, ‘kerjanya’]	[‘malu’, ‘kali’, ‘tuh’, ‘ketauan’, ‘gitu’, ‘ketauan’, ‘si’, ‘karang’, ‘taruna’, ‘ama’, ‘kadesnya’, ‘kagak’, ‘becus’, ‘kerjanya’]
[‘gimana’, ‘nggak’, ‘di’, ‘cap’, ‘jadi’, ‘pantai’, ‘terkotor’, ‘kalo’, ‘sampahnya’, ‘udh’, ‘sebanyak’, ‘itu’, ‘kalo’, ‘gamau’, ‘dicap’, ‘jadi’, ‘pantai’, ‘terkotor’, ‘ya’, ‘bersihin’, ‘dong’, ‘pak’, ‘giliran’, ‘ada’, ‘yg’, ‘mau’, ‘bantuin’, ‘malah’, ‘ga’, ‘terima’, ‘anehh’, ‘banget’, ‘kepala’, ‘desa’, ‘dan’, ‘ketua’, ‘karang’, ‘tarunannya’, ‘perkara’, ‘pantainya’, ‘diklaim’, ‘jadi’, ‘pantai’, ‘terkotor’]	[‘gimana’, ‘nggak’, ‘cap’, ‘pantai’, ‘terkotor’, ‘kalo’, ‘sampahnya’, ‘udh’, ‘kalo’, ‘gamau’, ‘dicap’, ‘pantai’, ‘terkotor’, ‘ya’, ‘bersihin’, ‘giliran’, ‘yg’, ‘bantuin’, ‘ga’, ‘terima’, ‘anehh’, ‘banget’, ‘kepala’, ‘desa’, ‘ketua’, ‘karang’, ‘tarunannya’, ‘perkara’, ‘pantainya’, ‘diklaim’, ‘pantai’, ‘terkotor’]
[‘alhamdulillah’, ‘pandawara’, ‘berhasil’, ‘menginspirasi’, ‘dan’, ‘meningkatkan’, ‘kesadaran’, ‘pemerintah’, ‘tidak’, ‘perlu’, ‘ajakan’, ‘ekstrim’, ‘cukup’, ‘aksi’, ‘nyata’, ‘dan’, ‘akan’, ‘ada’, ‘pihak’, ‘yg’, ‘tersentil’, ‘olehnya’]	[‘alhamdulillah’, ‘pandawara’, ‘berhasil’, ‘menginspirasi’, ‘meningkatkan’, ‘kesadaran’, ‘pemerintah’, ‘ajakan’, ‘ekstrim’, ‘aksi’, ‘nyata’, ‘yg’, ‘tersentil’]
[‘hebat’, ‘sih’, ‘pandawara’, ‘grup’, ‘berkat’, ‘jasa’, ‘mereka’, ‘memviralkan’, ‘pantai’, ‘kotor’, ‘tersebut’, ‘jadi’, ‘ada’, ‘gerakan’, ‘untuk’, ‘membersihkannya’, ‘segenap’, ‘rakyat’, ‘indonesia’, ‘pastinya’, ‘sangat’, ‘berterimakasih’, ‘kepada’, ‘pg’]	[‘hebat’, ‘sih’, ‘pandawara’, ‘grup’, ‘berkat’, ‘jasa’, ‘memviralkan’, ‘pantai’, ‘kotor’, ‘gerakan’, ‘membersihkannya’, ‘segenap’, ‘rakyat’, ‘indonesia’, ‘pastinya’, ‘berterimakasih’, ‘pg’]

e. Stemming

Data sentimen yang sudah tidak relevan, selanjutnya yaitu dilakukan proses stemming, dimana proses stemming merupakan menghapus kata-kata menjadi kata dasar dari kata-kata yang mengandung kata depan, kata sambung, kata ganti [26]. Proses stemming ini juga menggunakan referensi dari file acuan pada website gituhub [27], seperti contohnya yang terdapat pada tabel 6.

Tabel 6. Proses Stemming

Sebelum	Sesudah
[‘malu’, ‘kali’, ‘tuh’, ‘ketauan’, ‘gitu’, ‘ketauan’, ‘si’, ‘karang’, ‘taruna’, ‘ama’, ‘kadesnya’, ‘kagak’, ‘becus’, ‘kerjanya’]	[‘malu’, ‘kali’, ‘tuh’, ‘tau’, ‘gitu’, ‘tau’, ‘si’, ‘karang’, ‘taruna’, ‘ama’, ‘kades’, ‘kagak’, ‘becus’, ‘kerja’]
[‘gimana’, ‘nggak’, ‘cap’, ‘pantai’, ‘terkotor’, ‘kalo’, ‘sampahnya’, ‘udh’, ‘kalo’, ‘gamau’, ‘dicap’, ‘pantai’, ‘terkotor’, ‘ya’, ‘bersihin’, ‘giliran’, ‘yg’, ‘bantuin’, ‘ga’, ‘terima’, ‘anehh’, ‘banget’, ‘kepala’, ‘desa’, ‘ketua’, ‘karang’, ‘tarunannya’, ‘perkara’, ‘pantainya’, ‘diklaim’, ‘pantai’, ‘terkotor’]	[‘gimana’, ‘nggak’, ‘cap’, ‘pantai’, ‘kotor’, ‘kalo’, ‘sampah’, ‘udh’, ‘kalo’, ‘gamau’, ‘cap’, ‘pantai’, ‘kotor’, ‘ya’, ‘bersihin’, ‘giliran’, ‘yg’, ‘bantuin’, ‘ga’, ‘terima’, ‘anehh’, ‘banget’, ‘kepala’, ‘desa’, ‘ketua’, ‘karang’, ‘tarunannya’, ‘perkara’, ‘pantai’, ‘klaim’, ‘pantai’, ‘kotor’]
[‘alhamdulillah’, ‘pandawara’, ‘berhasil’, ‘menginspirasi’, ‘meningkatkan’, ‘kesadaran’, ‘aksi’, ‘nyata’, ‘yg’, ‘sentil’]	[‘alhamdulillah’, ‘pandawara’, ‘hasil’, ‘inspirasi’, ‘tingkat’, ‘sadar’, ‘pemerintah’, ‘ajak’, ‘ekstrim’, ‘aksi’, ‘nyata’, ‘yg’, ‘sentil’]



Sebelum	Sesudah
['pemerintah', 'ajakan', 'ekstrim', 'aksi', 'nyata', 'yg', 'tersentil'] ['hebat', 'sih', 'pandawara', 'grup', 'berkat', 'jasa', 'memviralkan', 'pantai', 'kotor', 'gerakan', 'membersihkannya', 'segenap', 'rakyat', 'indonesia', 'pastinya', 'berterimakasih', 'pg']	['hebat', 'sih', 'pandawara', 'grup', 'berkat', 'jasa', 'memviralkan', 'pantai', 'kotor', 'gerak', 'bersih', 'segenap', 'rakyat', 'indonesia', 'pasti', 'berterimakasih', 'pg']

f. Filtering Length

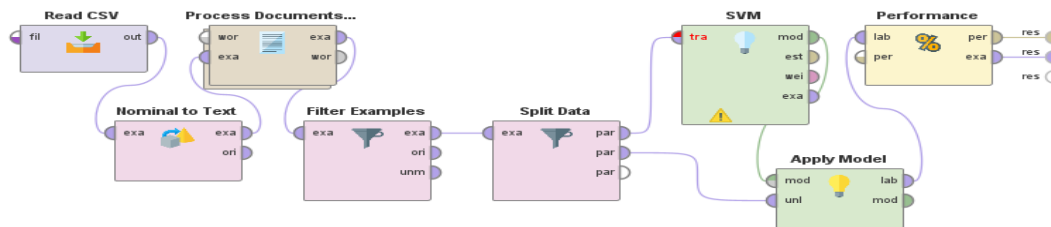
Proses terakhir dari tahapan preprocessing yaitu filtering length, pada proses ini kata yang tidak terdiri dari 4 sampai 25 huruf akan dihapus untuk memudahkan proses analisis sentimen selanjutnya [28], seperti contohnya yang terdapat pada tabel 7.

Tabel 7. Proses Filtering Length

Sebelum	Sesudah
['malu', 'kali', 'tuh', 'tau', 'gitu', 'tau', 'si', 'karang', 'taruna', 'ama', 'kades', 'kagak', 'becus', 'kerja'] ['gimana', 'nggak', 'cap', 'pantai', 'kotor', 'kalo', 'sampah', 'udh', 'kalo', 'gamau', 'cap', 'pantai', 'kotor', 'ya', 'bersihin', 'giliran', 'yg', 'bantuin', 'ga', 'terima', 'anehh', 'banget', 'kepala', 'desa', 'ketua', 'karang', 'tarunannya', 'perkara', 'pantai', 'klaim', 'pantai', 'kotor'] ['alhamdulillah', 'pandawara', 'hasil', 'inspirasi', 'tingkat', 'sadar', 'pemerintah', 'ajak', 'ekstrim', 'aksi', 'nyata', 'yg', 'sentil'] ['hebat', 'sih', 'pandawara', 'grup', 'berkat', 'jasa', 'memviralkan', 'pantai', 'kotor', 'gerak', 'bersih', 'segenap', 'rakyat', 'indonesia', 'pasti', 'berterimakasih', 'pg']	['malu', 'kali', 'gitu', 'karang', 'taruna', 'kades', 'kagak', 'becus', 'kerja'] ['gimana', 'nggak', 'pantai', 'kotor', 'kalo', 'sampah', 'kalo', 'gamau', 'pantai', 'kotor', 'bersihin', 'giliran', 'bantuin', 'terima', 'anehh', 'banget', 'kepala', 'desa', 'ketua', 'karang', 'tarunannya', 'perkara', 'pantai', 'klaim', 'pantai', 'kotor'] ['alhamdulillah', 'pandawara', 'hasil', 'inspirasi', 'tingkat', 'sadar', 'pemerintah', 'ajak', 'ekstrim', 'aksi', 'nyata', 'sentil'] ['hebat', 'pandawara', 'grup', 'berkat', 'jasa', 'memviralkan', 'pantai', 'kotor', 'gerak', 'bersih', 'segenap', 'rakyat', 'indonesia', 'pasti', 'berterimakasih']

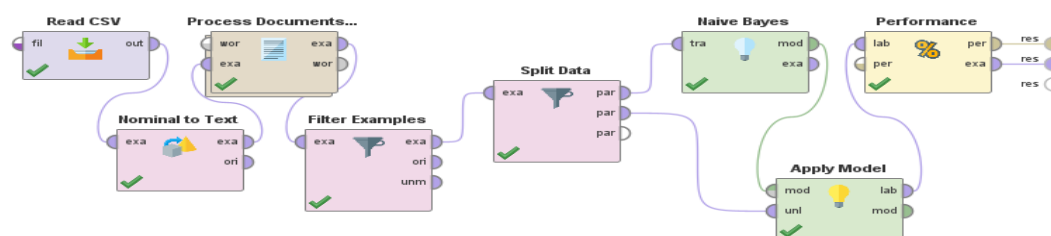
3.4 Implementasi

Data yang sudah melalui tahap preprocessing, selanjutnya yaitu tahapan implementasi. Tahapan implementasi pertama yaitu membagi dataset menjadi dua bagian, data uji dan data latih. Data uji dan data latih dibagi dengan persentase 80% untuk data uji dan 20% untuk data latih, semakin banyak data uji, semakin bagus juga metode akan mempelajari pola data untuk pengujian data latih untuk menghasilkan akurasi yang tinggi [29]. Total data yang sudah melalui tahap preprocessing hingga mendapatkan data sebanyak 2.339, dengan dibagi menjadi 1.871 data uji dan 468 data latih.



Gambar 7. Implementasi Metode Support Vector Machine

Data yang sudah dibagi menjadi data uji dan data latih, selanjutnya dilanjutkan pengimplementasian metode Support Vector Machine dengan melatih data sebanyak 468, setelah melatih data menggunakan metode Support Vector Machine nantinya akan menghasilkan confusion matrix yang berisi accuracy, precision, dan recall seperti contohnya proses implementasi dari metode Support Vector Machine yang terdapat pada gambar 7.



Gambar 8. Implementasi Metode Naive Bayes



Data yang sudah dibagi menjadi data uji dan data latih, selanjutnya dilanjutkan pengimplementasian metode Naïve Bayes dengan melatih data sebanyak 468, setelah melatih data menggunakan metode Naïve Bayes nantinya akan menghasilkan confusion matrix yang berisi accuracy, precision, dan recall seperti contohnya proses implementasi dari metode Naïve Bayes yang terdapat pada gambar 8. Data yang sudah melalui proses implementasi metode Support Vector Machine dan Naïve Bayes, selanjutnya yaitu melihat metode mana yang paling bagus dalam analisis sentimen ini dari hasil confusion matrix kedua metode yang digunakan.

3.5 Evaluasi

Data yang sudah dilakukan implementasi kedua algoritma yaitu Support Vector Machine dan Naïve Bayes seperti contohnya yang terdapat pada gambar 7 dan 8, kemudian akan menghasilkan hasil berupa confusion matrix yang berisi accuracy, precision, dan recall untuk melihat metode mana yang paling baik untuk digunakan dalam analisis sentimen terhadap kegiatan pembersihan sampah yang dilakukan Pandawara Grup di pantai Sukabumi.

a. Evaluasi Metode Support Vector Machine dan Naïve Bayes

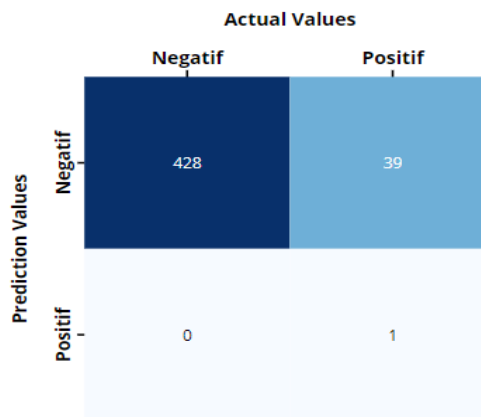
1. Support Vector Machine

Data yang sudah melalui tahapan implementasi metode Support Vector Machine, kemudian akan menghasilkan confusion matrix. Pada confusion matrix akan menghasilkan True Positif (TP), True Negatif (TN), False Positif (FP), dan False Negatif (FN) seperti yang terdapat pada tabel 8.

Tabel 8. Hasil Confusion Matrix Metode Support Vector Machine

	True Negatif	True Positif
Pred. Negatif	428	39
Pred. Positif	0	1

Berdasarkan implementasi dari metode Support Vector Machine yang terdapat pada gambar 9, menunjukkan bahwa accuracy yang dihasilkan sebesar 91,67%, precision sebesar 100%, dan recall sebesar 2,5%.



Gambar 9. Visualisasi Confusion Matrix Metode Support Vector Machine

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} = \frac{(1+428)}{(1+428+0+39)} = \frac{429}{468} = 0,9167 \text{ atau } 91,67\% \tag{1}$$

$$\text{Precision} = \frac{TP}{(TP+FP)} = \frac{1}{(1+0)} = \frac{1}{1} = 1 \text{ atau } 100\% \tag{2}$$

$$\text{Recall} = \frac{TP}{(TP+FN)} = \frac{1}{(1+39)} = \frac{1}{40} = 0,025 \text{ atau } 2,5\% \tag{3}$$

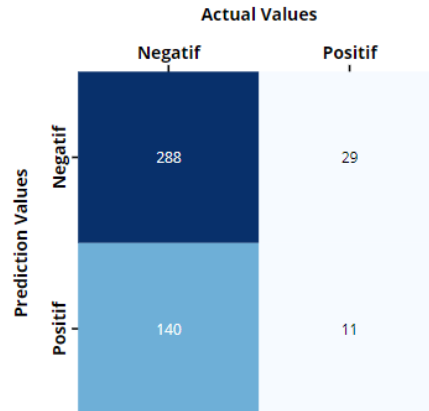
2. Naïve Bayes

Data yang sudah melalui tahapan implementasi metode Naïve Bayes, kemudian akan menghasilkan confusion matrix. Pada confusion matrix akan menghasilkan True Positif (TP), True Negatif (TN), False Positif (FP), dan False Negatif (FN) seperti yang terdapat pada tabel 9.

Tabel 9. Hasil Confusion Matrix Metode Naïve Bayes

	True Negatif	True Positif
Pred. Negatif	428	39
Pred. Positif	0	1

Berdasarkan implementasi dari metode Naïve Bayes yang terdapat pada gambar 10, menunjukkan bahwa accuracy yang dihasilkan sebesar 63,89%, precision sebesar 7,28%, dan recall sebesar 27,50%.



Gambar 10. Visualisasi Confusion Matrix Metode Naïve Bayes

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} = \frac{(11+288)}{(11+288+140+29)} = \frac{299}{468} = 0,6389 \text{ atau } 63,89\% \quad (1)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} = \frac{11}{(11+140)} = \frac{11}{151} = 0,07284 \text{ atau } 7,28\% \quad (2)$$

$$\text{Recall} = \frac{TP}{(TP+FN)} = \frac{11}{(11+29)} = \frac{11}{40} = 0,275 \text{ atau } 27,5\% \quad (3)$$

b. Visualisasi Wordcloud

Visualisasi wordcloud merupakan memperlihatkan hasil kata yang paling sering muncul pada analisis sentimen yang telah dilakukan, semakin besar ukuran kata maka semakin sering juga kata tersebut muncul [30]. Pada keseluruhan dataset sebesar 2.339 data dihasilkan visualisasi wordcloud pada data bersentimen positif berwarna hijau sedangkan data bersentimen negatif berwarna merah, seperti yang terdapat pada gambar 11 dan gambar 12.



Gambar 11. Wordcloud Sentimen Positif

Berdasarkan wordcloud yang terdapat pada gambar 11, memperlihatkan kata yang paling sering muncul pada sentimen positif yaitu kata “pandawara”, “pantai”, dan “bisa”.



Gambar 12. Wordcloud Sentimen Positif

Berdasarkan wordcloud yang terdapat pada gambar 12, memperlihatkan kata yang paling sering muncul pada sentimen negatif yaitu berupa bentuk kekecewaan masyarakat terhadap kegiatan pembersihan sampah yang



dilakukan Pandawara Grup di pantai Sukabumi namun dibatalkan oleh kepala desa dan karang taruna setempat, seperti kata “sampah”, “kades”, dan “bersihin”.

4. KESIMPULAN

Berdasarkan analisis sentimen yang dilakukan terhadap kegiatan pembersihan sampah yang dilakukan Pandawara Grup di pantai Sukabumi, dihasilkan dataset akhir yang mampu diproses hingga akhir ke dalam implementasi metode Support Vector Machine dan Naïve Bayes sebanyak 2.339 data, dengan 199 data bersentimen positif dan 2.140 bersentimen negatif yang berisi banyak kekecewaan terhadap kebijakan pemerintahan yang berada di wilayah Sukabumi terhadap kegiatan pembersihan sampah yang dilakukan Pandawara Grup di pantai Sukabumi. Hasil dari pengujian metode Support Vector Machine dan Naïve Bayes menghasilkan accuracy sebesar 91,67% pada metode Support Vector Machine dan 63,89% pada metode Naïve Bayes, mengartikan bahwa accuracy metode Support Vector Machine lebih tinggi dan lebih baik untuk mengklasifikasikan sentimen terhadap kegiatan pembersihan sampah yang dilakukan Pandawara Grup di pantai Sukabumi dibandingkan metode Naïve Bayes. Banyaknya sentimen negatif dibandingkan sentimen positif pada analisis sentimen ini, seharusnya dapat menjadi bahan pertimbangan bagi pemerintahan daerah setempat dalam mengambil keputusan mengenai kebijakan-kebijakan agar tidak mengecewakan masyarakat terkait suatu layanan dan dapat meningkatkan kinerja pemerintahan ke depannya. Rekomendasi saran dari peneliti untuk penelitian selanjutnya, diharapkan peneliti selanjutnya dapat lebih memperhatikan terhadap sentimen pendapat pengguna pada tahapan pelabelan dataset agar pada sentimen label positif dan negatif mendapatkan hasil yang lebih baik dan akurat.

REFERENCES

- [1] Alfandi Safira and F. N. Hasan, “Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier,” *ZONASI: Jurnal Sistem Informasi*, vol. 5, no. 1, pp. 59–70, Jan. 2023, doi: 10.31849/zn.v5i1.12856.
- [2] U. Yudatama et al., *Memahami Teknologi Informasi*. Kaizen Media Publishing, 2023. Accessed: Nov. 09, 2023. [Online]. Available: https://books.google.co.id/books?hl=en&lr=&id=PIHcEAAAQBAJ&oi=fnd&pg=PA67&ots=IbIdz291Sh&sig=PpRhRy3NzUL5g_DRmADqLutTNVM&redir_esc=y#v=onepage&q&f=false
- [3] Y. W. S. Putra et al., *Pengantar Aplikasi Mobile*. Penerbit Haura Utama, 2023. Accessed: Nov. 09, 2023. [Online]. Available: https://books.google.co.id/books?hl=en&lr=&id=2tLcEAAAQBAJ&oi=fnd&pg=PA35&ots=6KrSgdfjRC&sig=4RreBQDatMWWGzhSD6baVSyPLkQ&redir_esc=y#v=onepage&q&f=false
- [4] A. R. Abdillah and F. N. Hasan, “Analisis Sentimen Terhadap Kandidat Calon Presiden Berdasarkan Tweets Di Sosial Media Menggunakan Naive Bayes Classifier,” *SMATIKA JURNAL*, vol. 13, no. 01, pp. 117–130, Jul. 2023, doi: 10.32664/smatika.v13i01.750.
- [5] A. D. Riyanto, “Hootsuite (We are Social): Data Digital Indonesian 2024,” *andi.link*. Accessed: Nov. 05, 2023. [Online]. Available: <https://andi.link/hootsuite-we-are-social-data-digital-indonesia-2024/>
- [6] H. Dhery, A. Assyam, and F. N. Hasan, “Analisis Sentimen Twitter Terhadap Perpindahan Ibu Kota Negara Ke IKN Nusantara Menggunakan Orange Data Mining,” *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 1, pp. 341–349, 2023, doi: 10.30865/klik.v4i1.957.
- [7] Ankit and N. Saleena, “An Ensemble Classification System for Twitter Sentiment Analysis,” *Procedia Comput Sci*, vol. 132, no. 1, pp. 937–946, 2018, doi: 10.1016/j.procs.2018.05.109.
- [8] K. A. Puspita, I. Rachmawati, and H. Sampurna, “Pengaruh Implementasi Kebijakan Pengelolaan Sampah Terhadap Partisipasi Masyarakat di Kota Sukabumi,” *Kebijakan : Jurnal Ilmu Administrasi*, vol. 14, no. 1, pp. 1–11, Jan. 2023, doi: 10.23969/kebijakan.v14i1.5807.
- [9] O. Ramdhani, I. Yustiana, and A. Fergina, “Rancang Bangun Sistem Informasi Bank Sampah Menggunakan Metode Prototype (Studi Kasus di Kampung Lembur Sawah, Sukabumi),” *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 7, no. 3, pp. 757–767, Aug. 2022, doi: 10.29100/jipi.v7i3.3080.
- [10] Kementerian Lingkungan Hidup, “Capaian Kinerja Pengelolaan Sampah,” *SIPSN: Sistem Pengelolaan Sampah Nasional*. Accessed: Nov. 06, 2023. [Online]. Available: <https://sipsn.menlhk.go.id/sipsn/public/data/timbulan>
- [11] Ditjen PP, “Database Peraturan PerUndang-Undangan,” *Peraturan.go.id*. Accessed: Nov. 06, 2023. [Online]. Available: <https://peraturan.go.id/>
- [12] BBC News, “Aksi Pandawara Grup bersihkan Pantai Loji sempat ditolak kepala desa - Mengapa sampah di pesisir atau sungai tak diurus?,” *BBC News Indonesia*. Accessed: Nov. 06, 2023. [Online]. Available: <https://www.bbc.com/indonesia/articles/cd1jd17p9j7o#:~:text=Pandawara%20Group%20adalah%20sekumpulan%20anak,Group%20dengan%20mengajak%20warga%20setempat>
- [13] M. Dwijayanti, F. Noor Hasan, and R. Zein Adam, “Analisis Sentimen Pada Ulasan Pelanggan Menggunakan Metode Naïve Bayes Classifier (Studi Kasus: Grab Indonesia),” *Prosiding Seminar Nasional Teknoka*, vol. 6, no. 1, pp. 93–99, Jan. 2022, doi: 10.22236/teknoka.v6i1.441.
- [14] A. Syakir and F. N. Hasan, “Analisis Sentimen Masyarakat Terhadap Perilaku Korupsi Pejabat Pemerintah Berdasarkan Tweet Menggunakan Naive Bayes Classifier,” *Jurnal Media Informatika Budidarma*, vol. 7, no. 1, pp. 1796–1805, 2023, doi: 10.30865/mib.v7i4.6648.
- [15] M. I. Fikri, T. S. Sabrila, and Y. Azhar, “Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter,” *SMATIKA JURNAL*, vol. 10, no. 02, pp. 71–76, Dec. 2020, doi: 10.32664/smatika.v10i02.455.



- [16] A. T. Zy and W. Hadikristanto, "Implementasi Algoritma Metode Naive Bayes dan Support Vector Machine Tentang Pembobolan dan Kebocoran Data di Twitter," *Bulletin of Information Technology (BIT)*, vol. 4, no. 1, pp. 49–56, 2023, doi: 10.47065/bit.v3i1.
- [17] R. Yunita and M. Kamayani, "Perbandingan Algoritma SVM Dan Naive Bayes Pada Analisis Sentimen Penghapusan Kewajiban Skripsi," *Indonesian Journal of Computer Science*, vol. 12, no. 5, pp. 2879–2890, Oct. 2023, doi: 10.33022/ijcs.v12i5.3415.
- [18] R. Puspita and A. Widodo, "Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS," *Jurnal Informatika Universitas Pamulang*, vol. 5, no. 4, pp. 646–654, Dec. 2021, doi: 10.32493/informatika.v5i4.7622.
- [19] J. W. Iskandar and Y. Nataliani, "Perbandingan Naive Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 6, pp. 1120–1126, Dec. 2021, doi: 10.29207/resti.v5i6.3588.
- [20] A. Puji Astuti, S. Alam, and I. Jaelani, "Komparasi Algoritma Support Vector Machine dengan Naive Bayes Untuk Analisis Sentimen Pada Aplikasi BRImo," *Jurnal Bangkit Indonesia*, vol. 11, no. 2, pp. 1–6, Oct. 2022, doi: 10.52771/bangkitindonesia.v11i2.196.
- [21] F. P. Herlambang and D. Avianto, "Analisis Sentimen Opini Pengguna Twitter Terhadap Tragedi Kanjuruhan Malang dengan Metode Support Vector Machine," *Jurnal Media Informatika Budidarma*, vol. 7, no. 4, pp. 1727–1739, Oct. 2023, doi: 10.30865/mib.v7i4.6332.
- [22] S. D. Prasetyo, S. S. Hilabi, and F. Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naive Bayes dan KNN," *Jurnal KomtekInfo*, vol. 10, no. 1, pp. 1–7, Jan. 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [23] L. A. Sari, N. F. Ramadhita, and F. N. Hasan, "Analysis Of Public Sentiment On Google Play Store TIJE Application Users Using Naive Bayes Classifier Method," *JUTIF Jurnal Teknik Informatika*, vol. 5, no. 1, pp. 243–251, Feb. 2024, doi: 10.52436/1.jutif.2024.5.1.1648.
- [24] D. Normawati and S. A. Prayogi, "Implementasi Naive Bayes Classifier dan Confusion Matrix pada Analisis Sentimen Berbasis Teks pada Twitter," *Jurnal Sains Komputer dan Informatika (J-Sakti)*, vol. 5, no. 2, pp. 697–711, 2021, doi: 10.30645/j-sakti.v5i2.369.
- [25] O. R. Hartono, "Indonesian Stoplist," Kaggle. Accessed: Feb. 25, 2024. [Online]. Available: <https://www.kaggle.com/datasets/oswinrh/indonesian-stoplist>
- [26] A. R. Isnain, A. I. Sakti, D. Alita, and N. S. Marga, "Sentimen Analisis Publik terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma SVM," *Jurnal Data Mining dan Sistem Informasi*, vol. 2, no. 1, pp. 31–37, Feb. 2021, doi: 10.33365/jdmsi.v2i1.1021.
- [27] A. Librian, "Stemming Sastrawi," Github. Accessed: Mar. 02, 2024. [Online]. Available: <https://github.com/sastrawi/sastrawi>
- [28] A. T. Mukti and F. N. Hasan, "Analisis Sentimen Warganet Terhadap Keberadaan Juru Parkir Liar Menggunakan Metode Naive Bayes Classifier," *Jurnal Media Informatika Budidarma*, vol. 8, no. 1, pp. 644–653, Jan. 2024, doi: 10.30865/mib.v8i1.6982.
- [29] R. Wati, S. Ernawati, and H. Rachmi, "Pembobotan TF-IDF Menggunakan Naive Bayes pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 13, no. 1, pp. 84–93, Apr. 2023, doi: 10.34010/jamika.v13i1.9424.
- [30] M. G. Pradana, "Penggunaan Fitur Wordcloud dan Document Term Matrix dalam Text Mining," *Jurnal Ilmiah Informatika*, vol. 8, no. 01, pp. 38–43, 2020, doi: 10.33884/jif.v8i01.1838.