

Analisis Sentimen Kinerja KPU di Pemilu 2024 pada Media Sosial X Menggunakan Metode Naïve Bayes

Zaiima Tsabitha Nabhandiningrat Nasihin^{1*)}, Syariful Alam²⁾, Mutiara Andayani Komara³⁾

¹⁾²⁾³⁾ Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Wastukencana

^{*)}Correspondence author: zaimatsabitha@gmail.com, Purwakarta, Indonesia

DOI: <https://doi.org/10.37012/jtik.v10i2.2249>

Abstrak

Bangsa Indonesia belum lama ini telah melakukan hajatan besar demokrasi yaitu PEMILU (Pemilihan Umum). Hal ini merupakan amanat Undang-Undang (UU) Nomor 42/2008 tentang Pilpres. Namun dalam pelaksanaannya banyak sekali hal-hal yang dinilai oleh para pemilih bahwa Pemilu ini berjalan tidak sesuai dengan aturan yang ada. Mulai dari aroma kecurangan banyak disuarakan netizen, ketidaknetralan para penegak hukum terhadap salah satu calon pasangan presiden dan wakil presiden, serta pelaksanaan yang tidak Jujur dan Adil. dari para sebagian pemilih. Komisi Pemilihan Umum (KPU) Indonesia memainkan peran penting dalam memastikan integritas pemilihan umum. Sentimen yang diberikan kepada KPU lewat media sosial merupakan gambaran kekecewaan tersebut. Penelitian ini berfokus pada analisis sentimen publik terhadap kinerja KPU dalam pemilu 2024. Dataset yang digunakan dalam penelitian ini terdiri dari 317 entri, yang dikurangi menjadi 244 entri setelah tahap pembersihan dan pelabelan. Pemrosesan teks dilakukan untuk mempersiapkan data untuk klasifikasi. Algoritma Klasifikasi Naïve Bayes, yang diimplementasikan menggunakan Python, digunakan dalam analisis ini. Kinerja model dievaluasi menggunakan matriks kebingungan (confusion matrix). Studi ini mengkonfirmasi bahwa algoritma Klasifikasi Naïve Bayes efektif untuk analisis sentimen data media sosial. Tingkat akurasi sebesar 57% menunjukkan kepercayaan publik yang kuat terhadap kinerja KPU. Penelitian ini menyoroti pentingnya memantau sentimen publik di media sosial untuk mengukur efektivitas badan pemilihan seperti KPU.

Kata Kunci: Kinerja KPU, Pemilu 2024, Analisis Sentimen, Klasifikasi Naïve Bayes, X

Abstract

The Indonesian nation has recently held a big celebration of democracy, namely PEMILU (General Election). This is a mandate of Law (UU) Number 42/2008 concerning Presidential Elections. However, in its implementation there were many things that voters considered that this election was not running in accordance with existing regulations. Starting from the smell of fraud that many netizens have voiced, the non-neutrality of law enforcers towards one of the presidential and vice presidential candidates, as well as implementation that is not honest and fair. from some voters. The Indonesian General Election Commission (KPU) plays an important role in ensuring the integrity of general elections. The sentiments given to the KPU via social media are an illustration of this disappointment. This research focuses on analyzing public sentiment towards the KPU's performance in the 2024 elections. The dataset used in this research consists of 317 entries, which were reduced to 244 entries after the cleaning and labeling stage. Text processing is performed to prepare data for classification. The Naïve Bayes Classification algorithm, implemented using Python, was used in this analysis. Model performance is evaluated using a confusion matrix. This study confirms that the Naïve Bayes Classification algorithm is effective for sentiment analysis of social media data. An accuracy rate of 57% shows strong public trust in the KPU's performance. This research highlights the importance of monitoring public sentiment on social media to measure the effectiveness of electoral bodies such as the KPU.

Keywords: KPU performance, 2024 election, sentiment analysis, Naïve Bayes Classification, X

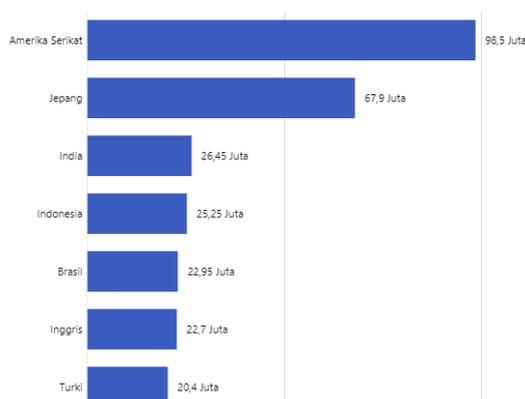
PENDAHULUAN

Bangsa Indonesia belum lama ini telah melakukan hajatan besar demokrasi yaitu PEMILU (Pemilihan Umum). Hal ini merupakan amanat Undang-Undang (UU) Nomor 42/2008 tentang Pilpres. Namun dalam pelaksanaannya banyak sekali hal-hal yang dinilai oleh para pemilih bahwa Pemilu ini berjalan tidak sesuai dengan aturan yang ada. Mulai dari aroma kecurangan banyak disuarakan netizen, ketidaknetralan para penegak hukum terhadap salah satu calon pasangan presiden dan wakil presiden, serta pelaksanaan yang tidak jujur dan adil dari para sebagian pemilih. KPU (Komisi Pemilihan Umum) selaku penyelenggara hajatan besar ini tak lepas dari tuduhan-tuduhan yang dilontarkan oleh sebagian pemilih yang merasa tidak puas dengan pelaksanaan Pemilu kali ini. Sentimen yang diberikan kepada KPU lewat media sosial merupakan gambaran kekecewaan tersebut (Putri & Lintang Muliawati, 2023).

Pada tahun 2024 ini Teknologi membuat jarak tidak lagi menjadi masalah dalam berkomunikasi, dan sekarang sosial media menjadi kebutuhan wajib bagi sebagian kalangan masyarakat mengatakan komentar mereka di web media sosial.

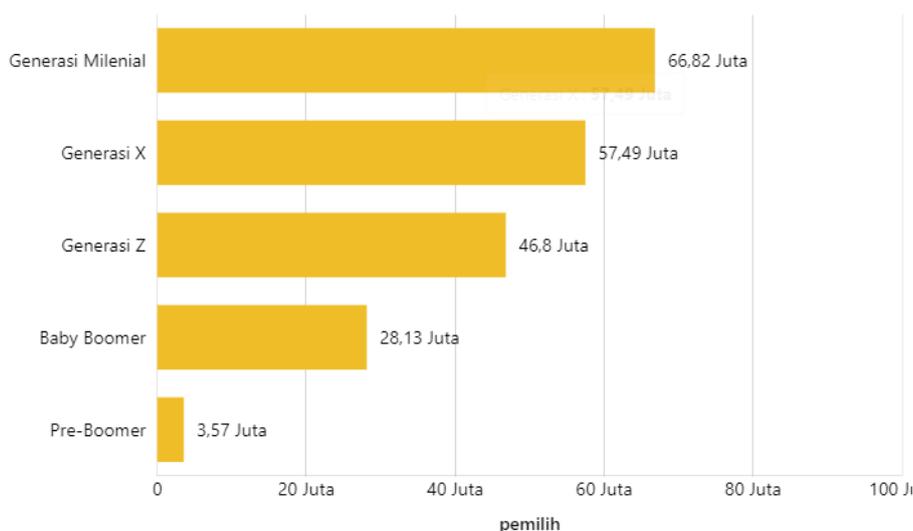
Semacam Facebook, X, Wechat, ataupun sejenisnya. Dan salah satu web media sosial yang kerap digunakan untuk memperdebatkan isu-isu politik serta sosial adalah X (Fuad et al., 2023).

Adapun jumlah pengguna X dapat dilihat pada gambar 1 di bawah ini:



Gambar 1. Jumlah Pengguna X

Di lihat dari detik news detik.com Komisi Pemilihan Umum (KPU) telah menetapkan Daftar Pemilih Tetap (DPT) untuk Pemilu 2024. Jumlahnya mencapai 204.807.222 pemilih. Pemilih - pemilih Sebanyak 66.822.389 atau 33,60% pemilih dari generasi milenial,” kata Komisioner KPU RI Betty Epsilon Idroos dalam Rapat Pleno Terbuka Rekapitulasi DPT di kantor KPU, Jakarta, Minggu (2/7/2023).Generasi milenial adalah sebutan untuk orang yang lahir pada 1980 hingga 1994.Sedangkan pemilih dari generasi Z adalah sebanyak 46.800.161 pemilih atau sebanyak 22,85% dari total DPT Pemilu 2024.Adapun sebutan generasi Z merujuk pada orang yang lahir mulai 1995 hingga 2000-an. Jika diakumulasikan, total pemilih dari kelompok generasi milenial dan generasi Z berjumlah lebih dari 113 juta pemilih.Kedua generasi ini mendominasi pemilih Pemilu 2024, yakni sebanyak 56,45% dari total keseluruhan pemilih.Selain itu, adapun kelompok pemilih dari generasi X yang menyusul di urutan berikutnya yaitu sebanyak 57.486.482 atau 28,07% dari total pemilih. Generasi X adalah orang kelahiran 1965 hingga 1979.Sisanya berasal dari kelompok generasi pre-boomer, atau orang yang lahir sebelum tahun 1944 dengan total sebanyak 3.570.850 atau 1,74% pemilih.



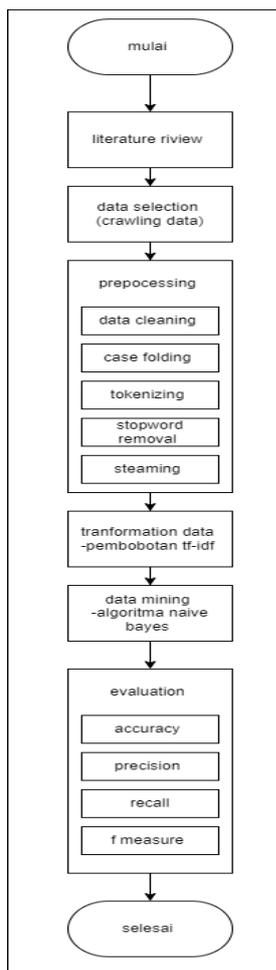
Gambar 2. Jumlah Daftar Pemilih 2024

Seluruh warga negara berhak menggunakan hak pilihnya dalam menentukan masa depan bangsa Indonesia. Maka untuk untuk mengetahui kinerja pada pemilihan tahun 2024 penulis tertarik mengangkat judul tentang “ANALISIS SENTIMEN KINERJA KPU DI

PEMILU 2024 PADA MEDIA SOSIAL X MENGGUNAKAN METODE NAÏVE BAYES”

METODE

Kerangka berpikir adalah logika penelitian yang diciptakan melalui pengamatan, fakta, dan tinjauan pustaka. Dalam penelitian ini, dijelaskan dalam kerangka berpikir sebagai berikut :



Gambar 3. Kerangka Berpikir Penelitian

Dari gambar kerangka berpikir di atas maka dapat diuraikan sebagai berikut:

1. *Data selection*

Tahap awal dalam penelitian ini adalah seleksi data, Dimana kami menggunakan teknik crawling data untuk mengumpulkan informasi relevan dari media sosial X dengan pertanyaan penelitian, terutama terkait Kinerja KPU 2024. Saya mengumpulkan tweet yang berkaitan dengan topik tersebut dengan melakukan pencarian menggunakan API.

2. *Pre processing*

Sebelum melakukan klasifikasi data perlu di atur dan dibersihkan selama tahap pre-processing melibatkan Langkah-langkah sebagai berikut:

1. Cleaning

Proses pembersihan data melibatkan penghapusan emotikon atau symbol tambahan yang tidak relevan atau tidak diperlukan

2. Case folding

Dalam Langkah ini dilakukan Upaya mengubah teks menjadi huruf kecil atau merendahkan teks

3. Tokenizing

Tahap berikutnya adalah tahap tekoning atau tekonisasi Dimana kalimat akan dipisahkan menjadi kata-kata secara individu. Dengan menggunakan metode ini kalimat dapat dipecah menjadi unit-unit kata yang terpisah,memungkinkan untuk analisis lebih lanjut pada Tingkat kata.

4. Stopword removal

Untuk melakukan pegghapusan stopwords, terdapat sebuah daftar yang berisi 32 stopwords yang telah disediakan. Stopword ini akan menghapus kata apapun dalam data yang juga terdaftar dalam daftar tersebut

5. Stemming

Prosuder yang dilakukan untuk menghilangkan imbuhan adalah dengan menghaous imbuhan setelah melakukan tekonisasi. Dalam Langkah ini imbuhan pada hasil token aan dihapus untuk mendapatkan bentuk dasar atau kata dasar dari setiap kata yang ada

3. *Transformation Data*

Setelah dokumen telah di proses dan di modifikasi, tahap selanjutnya adalah mengubah teks menjadi data numerik yang dapat secara akurat mewakili dokumen tersebut. Salah satu metode pembobotan yang digunakan adalah term frequency inverse document frequency (TF-IDF), yang mengindikasikan seberapa sering kata kunci atau istilah lain muncul dalam dokumen tersebut.

4. *Data Mining*

Langkah berikutnya adalah data mining,Dimana data mentah di ubah menjadi informasi yang berharga. Pada tahap ini tujuannya adalah untuk menemukan dan menganalisa pola tersembunyi dalam jumlah besar data guna mendapatkan wawasan yang berguna dan bermanfaat. Dalam Langkah ini digunakan pendekatan klasifikasi untuk mengkategorikan sekumpulan sentiment ke dalam kelompok netral, positif, dan negatif. Salah satu algoritma yang digunakan adalah naïve bayes berdasarkan teorema naïve bayes. Data yang telah disiapkan akan dibagi menjadi data pelatihan dan pengujian sebagai bagian dari proses klasifikasi.

5. *Evaluation*

Pada langkah terakhir, model yang telah dibentuk akan dievaluasi untuk memastikan konsistensi dengan hipotesis sebelumnya. Untuk mengukur akurasi algoritma, kami menggunakan confusion matrix. Confusion matrix digunakan untuk menggambarkan prediksi model dan seberapa baik metode tersebut diterapkan. Dengan menggunakan confusion matrix, kita dapat menganalisis kinerja model dalam empat dimensi dan mendapatkan pemahaman yang lebih jelas tentang performa yang dicapai:

- a. Akurasi: Untuk mengukur sejauh mana nilai yang diharapkan dan nilai aktual memiliki kemiripan satu sama lain.

b. Presisi : Untuk mengevaluasi keakuratan atau presisi suatu model berdasarkan prediksi positifnya. Presisi adalah metrik yang berguna saat model memiliki tingkat Positif Palsu yang tinggi.

c. Recall : Untuk menilai berapa banyak nilai Positif Aktual yang telah dikenali sebagai True Positif oleh model melalui pelabelan. Recall adalah metrik yang digunakan saat terdapat tingkat False Negatif yang tinggi. Recall menjadi metrik penting dalam menentukan model yang terbaik.

d. F-Measure: Perbandingan antara presisi dan recall dapat diukur menggunakan F1-score, yang merupakan perolehan hasil rata-rata tertimbang dari kedua metrik tersebut. F1-score memberikan gambaran keseluruhan tentang keseimbangan antara presisi dan recall dalam penilaian kinerja model.

Analisis sentimen merupakan penelitian berupa text mining yang memiliki tujuan sebagai analisis pendapat, evaluasi, sikap, sentimen seseorang yang berupa teks untuk ditinjau pada suatu objek apakah kalimat tersebut termasuk kedalam kalimat positif atau kedalam kalimat negatif (Romadhoni et al., 2022).

Analisis sentimen adalah suatu proses yang bertujuan untuk mengetahui apakah polaritas data berupa teks (dokumen, kalimat, paragraf) akan mengarah ke positif, negative, atau netral. Analisis sentiment terdiri dari pemrosesan Bahasa alami, analisis teks dan komputasi linguistic untuk mengidentifikasi sentiment dari suatu dokumen informasi tekstual secara umum dapat dibagi menjadi informasi fakta dan opini. Tugas besar dalam analisis sentiment adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, apakah pendapatan yang dikemukakan dalam dokumen bersifat positif, negative atau netral. Penelitian mengenai analisis sentimen telah berkembang sejak tahun 2003 dan merupakan bagian dari text mining yang merupakan penelitian komputasi berdasarkan sentimen, emoticon, pendapat, komentar dan setiap ekspresi yang diungkapkan oleh teks (Fitriyana et al., 2023).

Naïve Bayes adalah metode yang tidak memiliki aturan dan menggunakan cabang matematika yang disebut teori probabilitas untuk mendapatkan peluang setinggi mungkin

Tahap awal dalam penelitian ini adalah teknik crawling data yang didapatkan hasil crawling disimpan dalam format CSV, berikut hasil crawling yang didapatkan:

Gambar 5. Hasil Crawling

Dengan kolom-kolom tertentu yang mencakup informasi penting yang Anda kumpulkan dari media sosial Twitter. Kolom-kolom ini adalah

1. conversation_id_str: ID percakapan yang terkait dengan tweet.
2. created_at: Tanggal dan waktu tweet dibuat.
3. favorite_count: Jumlah favorit yang diberikan pada tweet.
4. full_text: Teks lengkap dari tweet.
5. id_str: ID unik untuk tweet.
6. image_url: URL gambar yang terlampir pada tweet (jika ada).
7. in_reply_to_screen_name: Nama pengguna yang ditunjukkan oleh tweet balasan.
8. lang: Bahasa dari tweet (misalnya, "id" untuk Bahasa Indonesia).
9. location: Lokasi pengguna yang membuat tweet (jika diaktifkan).
10. quote_count: Jumlah kutipan (retweet dengan komentar) yang diterima oleh tweet.
11. reply_count: Jumlah balasan yang diterima oleh tweet.
12. retweet_count: Jumlah retweet yang diterima oleh tweet.
13. tweet_url: URL tweet yang unik.
14. user_id_str: ID unik pengguna yang membuat tweet.
15. username: Nama pengguna Twitter yang membuat tweet.

Data yang didapat berjumlah 357 baris dan 15 kolom, CSV ini berisi sejumlah besar informasi yang relevan yang dapat di analisis ke tahapan selanjutnya.

Setelah melakukan crawling, lalu data yang dibutuhkan hanya data `full_text` dan dibersihkan selama tahap pre-processing melibatkan. Langkah-langkah sebagai berikut:

1. Cleaning

Pada tahap cleaning penghapusan emotikon atau simbol tambahan yang tidak relevan atau tidak diperlukan dalam teks. Hal ini juga dapat mencakup penghapusan URL yang tidak diinginkan.

Data `'full_text'` yang berisi teks asli sebelum cleaning, dan `'cleaned_text'` yang berisi teks setelah cleaning dilakukan. Setelah itu `@mentions` dan tagar (`#hashtags`) ada didalam data `full_text` dan akan dilakukan penghapusan karena mengganggu.

Fungsi `remove_mentions_and_hashtags(text)` juga menggunakan ekspresi reguler untuk menghapus semua kata yang dimulai dengan '@' dan '#'. Ini dilakukan untuk menghapus metadata seperti nama pengguna (`@mentions`) dan tagar (`#hashtags`) dari teks tweet, sehingga hanya konten utamanya yang tetap. Fungsi ini kemudian diterapkan pada kolom `'full_text'` dari DataFrame menggunakan metode `.apply()`, dan hasilnya ditambahkan ke kolom `'cleaned_text'` yang telah dibersihkan sebelumnya.

Tahapan selanjutnya adalah melakukan proses case folding. Case folding adalah proses merubah seluruh huruf dalam teks menjadi huruf kecil.

Dengan menambahkan fungsi `case_folding()` pada penelitian ini dan menerapkannya pada kolom `'cleaned_text'`, kita akan mendapatkan teks yang seluruhnya dalam huruf kecil.

2. Tokenizing

Setelah tahap case folding selesai, langkah selanjutnya dalam pre processing teks adalah tokenisasi. Tokenisasi adalah proses membagi teks menjadi potongan-potongan yang lebih kecil, yang disebut sebagai token.

Dalam penelitian ini, kita menggunakan library NLTK untuk melakukan tokenizing. Fungsi `tokenize_text()` menerima teks sebagai input dan mengembalikan daftar token menggunakan fungsi `word_tokenize()` dari NLTK. Kemudian, hasil tokenizing disimpan dalam kolom baru `'tokenized_text'` pada DataFrame.

3. Stopword removal

Setelah rokenizing, langkah selanjutnya dalam pre processing teks adalah stopwords removal. Stopwords adalah kata-kata umum yang sering muncul dalam teks dan tidak memberikan informasi yang signifikan tentang konten teks tersebut.

Dalam pembahasan diatas, kita menggunakan modul stopwords dari NLTK untuk mendapatkan daftar stopwords bahasa Indonesia. Kemudian, fungsi `remove_stopwords()` digunakan untuk menghapus stopwords dari setiap daftar token dalam kolom `'tokenized_text'` DataFrame.

4. Stemming

Setelah tahap stopwords removal, langkah selanjutnya adalah stemming. Stemming adalah proses mengubah kata-kata ke bentuk dasarnya atau akarnya. Ini dilakukan dengan menghapus imbuhan atau akhiran dari kata-kata.

Dalam penelitian ini, pada tahap stemming menggunakan `StemmerFactory()` dari library Sastrawi, untuk menggunakan algoritma stemming bahasa Indonesia yang telah disediakan. Lalu mendefinisikan sebuah fungsi bernama `stem_words` yang menerima daftar kata-kata (tokens) sebagai argumen. Fungsi ini menggunakan stemmer yang sudah dibuat sebelumnya untuk melakukan stemming pada setiap kata dalam daftar tersebut. Dan menerapkan fungsi `stem_words` pada kolom `tokenized_text` DataFrame menggunakan metode `apply()`. Hasil stemming akan disimpan dalam kolom baru yang disebut `stemmed_text`.

Setelah data hasil stemming menjadi kata-kata yang telah disederhanakan ke bentuk dasarnya. Misalnya, kata "dimainkan" telah diubah menjadi "main", "kemasukan" menjadi "masuk", "diremehkan" menjadi "remeh", dan seterusnya. Proses ini membantu dalam mengurangi dimensi dan variasi kata-kata dalam teks.

Transformation Data

Setelah data berhasil di pre processing, tahap selanjutnya adalah mengubah teks menjadi data numerik yang dapat secara akurat mewakili data tersebut. Salah satu metode pembobotan yang digunakan adalah term frequency inverse document frequency (TF-IDF).

Dalam penelitian ini tahapan transformation data menggunakan TF-IDF dengan menggunakan library `TfidfVectorizer` dari `sklearn.feature_extraction.text` untuk mengonversi kumpulan teks menjadi representasi numerik menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). `TfidfVectorizer` mengubah kumpulan dokumen teks menjadi matriks TF-IDF, di mana setiap baris mewakili dokumen, dan setiap kolom mewakili kata dalam seluruh korpus dokumen. Setelah proses vektorisasi, lalu menggunakan `get_feature_names_out()` untuk mendapatkan daftar fitur kata, yaitu kata-kata yang digunakan sebagai kolom dalam matriks TF-IDF.

Selanjutnya menampilkan kata-kata dengan nilai TF-IDF tertinggi untuk dokumen pertama dalam dataset. Dengan menampilkan 10 jumlah kata teratas dengan nilai TF-IDF untuk masing-masing kata.

Data mining

Langkah berikutnya adalah penerapan data mining. Dalam Langkah ini sebelum ke tahap model maka dilakukan tahap labelling dengan 3 kelas negatif, netral dan positif.

Dalam tahap pemberian kelas menggunakan `VaderSentiment` dengan mengimport library terlebih dahulu. Lalu inisialisasi analyzer dari `SentimentIntensityAnalyzer`.

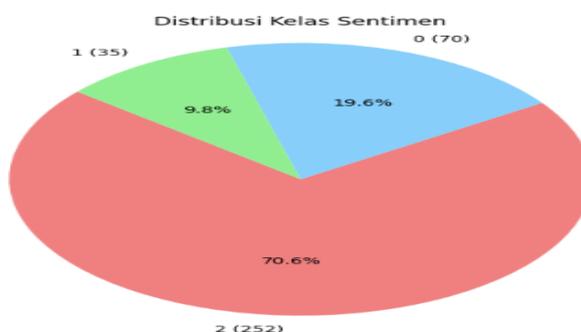
Selanjutnya perhitungan score sentiment dalam dataframe 'cleaned_data'. Lalu penambahan kolom `compound_Score` untuk skor sentimen, dengan menggunakan skor sentimen komposit, data dapat dengan mudah mengetahui apakah sebuah teks termasuk sentimen positif, negatif, atau netral secara keseluruhan.

Lalu menggunakan fungsi `map` pada kolom 'Sentiments' untuk mengubah nilai 'Negatif' menjadi 0, 'Netral' menjadi 1, dan 'Positif' menjadi 2. Setelah itu kolom 'Sentiments' akan berisi nilai numerik sesuai dengan label yang sudah ditentukan.

	cleaned_text	compound_Score	Sentiments
0	kpu kota tangerang rapat evaluasi kinerja pemi...	0.0000	1
1	bola panas mulai dimainkan siapa yg kemasukan ...	0.3400	2
2	cara team hukum 01 bekerja luar biasa tadinya ...	-0.0754	0
3	bawastu sama kpu kerja 5 tahun sekali tapi kin...	-0.2960	0
4	rapat koordinasi pengelolaan dokumentasi produ...	0.5106	2
...
352	kinerja kpu di pemilu 2024 menunjukkan komitmen...	0.8625	2
353	pleno rekapitulasi hasil pemilu berjalan kondus...	0.7184	2
354	ternyata si calon ini masih sodaraan dengan ca...	-0.4019	0
355	dan jika kalian berlaku adil demi allah saya b...	0.8738	2
356	sebaiknya untuk mengembalikan kepercayaan raky...	0.8442	2

Gambar 6. Nilai Sentiment

Selain itu, tahap selanjutnya visualisasi data yang sudah diberi label berikut hasilnya.



Gambar 7. Visualisasi Jumlah Distribusi Kelas Sentiment

Dapat dilihat bahwa distribusi kelas diatas tidak seimbang karena jumlah dalam setiap kelas tidak seimbang. Biasanya, ketidakseimbangan kelas dapat mempengaruhi kinerja model, terutama untuk kelas yang kurang mewakili (dalam penelitian ini, kelas 0 dan 1). Untuk mengatasi masalah ini, penulis menggunakan teknik undersampling yang dapat digunakan untuk mencoba menyeimbangkan kelas sebelum melatih model. Salah satu teknik oversampling yang digunakan adalah dengan menggunakan fungsi `resample` dari modul `sklearn.utils`. Fungsi ini memungkinkan kita untuk memilih jumlah sampel dari setiap kelas mayoritas yang ingin kita downsample agar seimbang dengan kelas minoritas.

Dalam penelitian ini, kedua kelas mayoritas (Positif dan Negatif) diundersample menjadi 35 sampel masing-masing untuk membuatnya seimbang dengan kelas minoritas

(Netral) yang memiliki 35 sampel. `replace=False` menandakan bahwa sampling dilakukan tanpa penggantian, artinya setiap sampel yang dipilih tidak akan dipilih lagi dalam proses selanjutnya. Setelah proses undersampling selesai, data-data yang telah di-undersample digabungkan kembali dengan kelas minoritas menggunakan fungsi `pd.concat()` untuk membentuk DataFrame baru yang seimbang secara proporsional lalu menyimpan ke dalam file CSV baru dengan nama `sentiment.csv`. Dengan menggunakan teknik ini, dapat menyeimbangkan kelas mayoritas dan minoritas dalam dataset dengan memilih jumlah sampel yang sesuai untuk setiap kelas mayoritas.

Setelah dataset seimbang, maka tahap selanjutnya adalah penerapan algoritma dengan menggunakan algoritma naïve bayes.

Import library untuk membuat model klasifikasi Naive Bayes menggunakan `from sklearn.naive_bayes import MultinomialNB`. Lalu menentukan variabel X dan y digunakan untuk menampung fitur (`cleaned_text`) dan label (Sentiments) dari DataFrame .Data yang telah disiapkan akan dibagi menjadi data pelatihan dan pengujian menggunakan `from sklearn.model_selection import train_test_split` dengan total data :

- a. Data pelatihan (train): 80% dari keseluruhan data.
- b. Data pengujian (test): 20% dari keseluruhan data.

Kemudian, menggunakan TF-IDF untuk mengonversi teks menjadi representasi numerik yang dapat dipahami oleh model. Selanjutnya, melatih model Naïve Bayes menggunakan data pelatihan.

Evaluation

Pada langkah terakhir, model yang telah latih akan dievaluasi. Untuk mengukur akurasi algoritma, menggunakan confusion matrix dan classification report.

Dalam penelitian ini, terdapat tiga kelas yang direpresentasikan dengan angka 0, 1, dan 2. Setiap baris dalam confusion matrix mewakili kelas sebenarnya, sedangkan setiap kolom mewakili kelas prediksi.

Dalam proses evaluasi pada penelitian ini, akurasi yang didapatkan adalah 0.57 atau 57%. Akurasi ini mengindikasikan bahwa sekitar 57% dari pengujian algoritma naïve bayes. Setelah evaluasi, membuat inputan baru untuk prediksi analisis sentimen diperlukan, berikut hasil inputan sesuai dengan kelas sentiment yang sudah di proses.

KESIMPULAN DAN REKOMENDASI

Berdasarkan hasil pengujian dalam penelitian yang berjudul "Implementasi GIS untuk monitoring jaringan pada unit layanan pengaduan jaringan di Diskominfo Kabupaten Kuantan Singingi," ditemukan bahwa implementasi GIS tersebut dapat dijalankan dengan efektif melalui aplikasi berbasis web. Dengan menggunakan GIS, sistem berhasil menampilkan traffic jaringan dengan memanfaatkan perangkat MikroTik dan Telegram sebagai media notifikasi kepada klien. Hasilnya menunjukkan bahwa sistem ini mampu memberikan solusi yang efisien dan efektif dalam memantau dan mengelola jaringan. Dalam konteks penelitian selanjutnya, disarankan untuk fokus pada deteksi serangan jaringan seperti DDoS, brute force, atau upaya hacking lainnya. Dalam hal ini, fitur-fitur firewall dan manajemen akses MikroTik dapat dieksplorasi lebih lanjut untuk meningkatkan keamanan dan kehandalan sistem. Langkah ini dapat memberikan kontribusi signifikan dalam menghadapi ancaman keamanan jaringan yang semakin kompleks.

REFERENSI

- Budiarto, S., & Suryadi, K. (2018). Geographic Information System (GIS) Application for Customer Complaint Handling in the Telecommunication Industry. *International Journal of Computer Science and Information Security*, 16(3), 94-99.
- Elgamar. (2020). *Buku Ajar Konsep Dasar Pemrograman Dengan PHP*. Guntoro, G., Gustiawan, M., Yudianto, R. J., Pratama, J., & Fauzi, A. (2021). Implementasi Jaringan Hotspot Di Perkantoran Guna Meningkatkan Keamanan Jaringan Komputer. *Jurnal Nasional Komputasi dan Teknologi Informasi*, 4(4), 244247.

- Hamuddin, B., & Febridi, B. (2019). Rancang Bangun Portal International Conference on Environment and Technology (Icetech) Universitas Lancang Kuning. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 4(1), 28. <https://doi.org/10.29100/jipi.v4i1.1363>
- Jamaludin, A., & Pradana, R. A. (2019). GIS-Based System for Complaint Handling and Service Quality Improvement of Internet Service Providers. *Journal of Physics: Conference Series*, 1243(1), 012077.
- Nugroho, S. A., Butar, B. B., & Mariskhana, K. (2019). Sistem Informasi Kegiatan Kemahasiswaan Berbasis Mobile pada Institut Sains dan Teknologi Pradita. *Jurnal Inovasi Informatika*, 4(2), 9–18.
- Rishav, S., Kumar, A., & Kumar, V. (2020). GIS-Based Complaint Management System: A Case Study on Telecommunication Industry. In *Proceedings of International Conference on Data Engineering and Communication Technology* (pp. 505-514). Springer, Singapore.
- Rusydi, I. (2019). Peranan Perkembangan Teknologi. *Journal of Chemical Information and Modeling*, 53(9), 1689-1699.
- Saputra, R. A., Parjito, P., & Wantoro, A. (2020). Implementasi Metode Jackson Network Queue Pada Pemodelan Sistem Antrian Booking Pelayanan Car Wash (Studi Kasus: Autoshine Car Wash Lampung). *Jurnal Teknologi Dan Sistem Informasi*, 1(2), 80-86.
- Setiawan, A., Rusdianto, D. S., & ... (2019). ... Pelaporan Gangguan Berbasis Web dengan menggunakan Teknologi Progressive Web Application (Studi Kasus: Unit IT PT. Kereta Api Indonesia Daerah Operasi III *Pengembangan Teknologi ...*, 5, 4194–4203. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/5159>
- Shukla, D., Shukla, A., & Sahu, R. (2019). Design and Implementation of GISBased Online Complaint Management System. *International Journal of Research in Engineering, Science and Management*, 2(2), 293-297.

- Sutopo, A., Hariyani, S., & Rahardjo, K. (2020). GIS-Based Complaint Handling System for Telecommunication Network Infrastructure. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 12(2-5), 55-59.
- Supriadi, K., Kebudayaan, D. A. N., Tinggi, S., Informatika, M., Komputer, D. A. N., & Lapangan, P. K. (2021). *Pt . Medion Ardhika Bhakti Cabang Palembang Berbasis Web Bayu Indrayana Pt . Medion Ardhika Bhakti Waterfall. InfoTekJar (Jurnal Nasional Informatika Dan Teknologi Jaringan)*, 2(2), 98–105. <https://doi.org/10.30743/infotekjar.v2i2.171>
- Siregar, M. H (2021). *Sms Gateway Menggunakan Metodologi Rational Unified Process.4(1)*.
<https://doi.org/10.36378/jtos.v4i1.1393>
- Wati, F. F., & Khasanah, U. (2019). Sistem Informasi Penjualan Berbasis Web pada UD Dwi Surya Aluminium dan Kaca Yogyakarta. *Paradigma – Jurnal Komputer Dan Informatika*, 21(2), 149–156. <https://doi.org/10.31294/p.v21i2.6026>
- Yani, A., Saputra, B., & Jurnal, R. T. (2018). *Rancang Bangun Sistem Informasi Evaluasi Siswa Dan Kehadiran Guru Berbasis Web. Petir*, 11(2).
<https://doi.org/10.33322/petir.v11i2.344>
- Supriadi, K., Kebudayaan, D. A. N., Tinggi, S., Informatika, M., Komputer, D. A. N., & Lapangan, P. K. (2021). *Pt . Medion Ardhika Bhakti Cabang Palembang Berbasis Web Bayu Indrayana Pt . Medion Ardhika Bhakti*.