

ANALISIS SENTIMEN MASYARAKAT TERHADAP PEMERINTAH DI ERA KABINET JOKO WIDODO BERDASARKAN SOSIAL MEDIA X MENGGUNAKAN NAÏVE BAYES DAN K-NEAREST NEIGHBOR (KNN)

ANALYSIS OF PUBLIC SENTIMENT TOWARDS THE GOVERNMENT IN THE JOKO WIDODO CABINET ERA BASED ON SOCIAL MEDIA X USING NAÏVE BAYES AND K-NEAREST NEIGHBOR (KNN)

Fadhel Naufal Rahman¹, Sri Lestari²

^{1,2}Sekolah Tinggi Ilmu Komputer (STIKOM) Cipta Karya Informatika Jakarta, Indonesia

¹fadhelnauf@gmail.com, ²sri.lestari1203@gmail.com

ABSTRACT

During his term of office, President Joko Widodo and the cabinet he formed often made new policies that caused reactions in society, either positive or negative. Public reactions also usually vary from supporting to criticizing the policy. If these comments are processed properly, they can be used as a reference for the government that will take office in the next period to evaluate the government's decisions regarding policy making. Seeing this, a sentiment analysis is needed which is aimed at classifying the various comments circulating so that they can be presented as useful information. In this research, data processing was carried out using the Naïve Bayes Classifier and K-Nearest Neighbor (KNN) algorithms with data taken from tweets on the social media platform (X). These two methods are used as the final result of the accuracy of this research and make a comparison between each other. The data used amounted to 5329 tweets and then went through a cleansing stage to total 3437 data. After processing the data, we got an accuracy value from the Naïve Bayes algorithm of 93.02%%, whereas if we used the K-Nearest Neighbor (KNN) algorithm we got an accuracy value of 96.33%. So it can be said that the K-Nearest Neighbor (KNN) algorithm can classify data better than Naïve Bayes.

Keywords: Data Mining, Social Media X, Jokowi Cabinet, Naïve Bayes, K-Nearest Neighbor.

ABSTRAK

Selama masa jabatannya, Presiden Joko Widodo beserta kabinet bentukannya seringkali membuat kebijakan baru yang menimbulkan reaksi dimasyarakat baik positif, atau negatif. Reaksi masyarakat pula biasanya beragam mulai dari yang mendukung hingga mengkritisi kebijakan tersebut. Jika komentar tersebut diolah dengan baik maka dapat dijadikan acuan pemerintah yang akan menjabat di periode berikutnya guna mengevaluasi keputusan – keputusan pemerintah terhadap pengambilan kebijakan. Melihat hal ini maka diperlukan sebuah analisis sentimen yang ditujukan untuk melakukan klasifikasi dari berbagai komentar yang beredar agar dapat disajikan menjadi sebuah informasi yang bermanfaat. Pada penelitian ini dilakukan pengolahan data dengan algoritma Naïve Bayes Classifier dan K-Nearest Neighbor (KNN) dengan data yang diambil dari cuitan pada platform media sosial (X). Dari kedua metode ini digunakan sebagai hasil akhir akurasi dari penelitian ini dan menjadikan perbandingan nya antara satu sama lain. Data yang digunakan berjumlah 5329 tweets lalu melalui tahapan cleansing menjadi berjumlah 3437 data. Setelah diolah datanya didapatkan nilai accuracy dari algoritma Naïve Bayes sebesar 93.02%%, sedangkan jika menggunakan algoritma K-Nearest Neighbor (KNN) mendapatkan nilai accuracy 96.33%. Sehingga dapat dikatakan bahwa algoritma K-Nearest Neighbor (KNN) dapat mengklasifikasikan data secara lebih baik dibandingkan Naïve Bayes.

Kata Kunci: Data Mining, Sosial Media X, Kabinet Jokowi, Naïve Bayes, K-Nearest Neighbor.

PENDAHULUAN

Pada era globalisasi dan dinamika politik yang terus berkembang, peran pemerintah dalam mengimplementasikan kebijakan menjadi semakin penting. Kebijakan publik merupakan instrumen utama yang digunakan pemerintah untuk mengatur berbagai aspek kehidupan

masyarakat, mulai dari ekonomi, pendidikan, kesehatan, hingga lingkungan. Di tengah perubahan cepat dalam masyarakat dan tantangan yang kompleks, kebijakan pemerintah menjadi fokus utama dalam upaya mencapai tujuan pembangunan yang berkelanjutan.

Banyak cara masyarakat menuangkan opininya terkait kebijakan – kebijakan baru yang dikeluarkan pemerintah, dimulai dari aksi unjuk rasa sampai mengutarakan opininya melalui berbagai platform media sosial salah satunya X. Opini yang disampaikan melalui platform tersebut masih belum dapat dibedakan bersentimen positif ataupun negatif, sehingga pemerintah kesulitan dalam melihat penilaian publik apabila adanya kebijakan baru yang dikeluarkan pemerintah. Karena penelitian ini berfokus pada pendapat masyarakat di media sosial X, maka evaluasi ini menggunakan text mining untuk memproses informasi yang ditinjau dan perasaan individu terhadap objek, orang, organisasi, dan isu.[1]

Pada tahun 2023 pengguna media sosial sebanyak 167 juta di Indonesia. Jumlah pengguna Twitter (X) di Indonesia pada awal tahun 2023 mencapai 24 juta pengguna.[2] dengan begitu media sosial Twitter dapat dijadikan sebagai acuan untuk mengetahui sentiment dan menentukan kecenderungan opini masyarakat terhadap suatu kejadian yang terjadi disekitar masyarakat. Salah satu topik yang masih hangat sampai sekarang yaitu sentimen masyarakat terkait Kabinet Jokowi.[3] Pemerintahan Ir. H. Joko Widodo telah dimulai sejak dilantikannya untuk kedua kalinya pada Minggu, 20 Oktober 2019 dalam Sidang Paripurna Majelis Permusyawaratan Rakyat (MPR) RI yang digelar di Gedung Nusantara, Kompleks Parlemen, Senayan. Semasa pemerintahannya Presiden Joko Widodo telah banyak mengeluarkan kebijakan yang cukup dinilai kontroversional. Oleh karena itu, diperlukan analisis sentimen untuk mengetahui bagaimana sentimen yang ada pada media sosial mengenai kepuasan masyarakat semasa pemerintahan Kabinet Presiden Ir. H. Joko Widodo.[3]

Pemerintahan Ir. H. Joko Widodo telah dimulai sejak dilantikannya untuk kedua kalinya pada Minggu, 20 Oktober 2019

dalam Sidang Paripurna Majelis Permusyawaratan Rakyat (MPR) RI yang digelar di Gedung Nusantara, Kompleks Parlemen, Senayan. Semasa pemerintahannya Presiden Joko Widodo telah banyak mengeluarkan kebijakan yang cukup dinilai kontroversional. Oleh karena itu, diperlukan analisis sentimen untuk mengetahui bagaimana sentimen yang ada pada media sosial mengenai kepuasan masyarakat semasa pemerintahan Kabinet Presiden Ir. H. Joko Widodo.[3]

METODE

Pada penelitian ini algoritma yang digunakan adalah Naïve bayes dan K-Nearest Neighbor (KNN). Data yang digunakan ialah tweet yang diambil dari aplikasi media sosial twitter (X) dengan pemrograman python melalui Google Colab setelah data tersebut di dapatkan lalu masuk ke proses penerapan CRISP-DM yaitu pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi dan penyebaran (Deployment). Data tersebut akan diproses dalam pengujian algoritma Naïve bayes dan K-Nearest Neighbor (KNN) menggunakan tools Rapidminer supaya dapat nilai akurasi klasifikasi yang baik dan optimal. Pemrosesan data ini dilakukan dengan menggunakan metode Cross Industry Standard Process for Data Mining (CRISP-DM) yaitu :

1) Pemahaman Bisnis (*Business Understanding*)

Dinamika kebijakan pemerintahan era Kabinet Presiden Joko Widodo mempengaruhi berbagai aspek kehidupan masyarakat seperti dibidang industri, ekonomi dan lain sebagainya. Walaupun pemerintahan era Kabinet Presiden Joko Widodo dinilai sangat berpengaruh positif bagi masyarakat namun tidak sedikit masyarakat yang menilai negatif dari pemerintahan era Kabinet Presiden Joko Widodo. Dengan demikian maka akan dilakukan sentimen terhadap pemerintahan era Kabinet Presiden Joko Widodo.

2) Pemahaman Data (*Data Understanding*)

Pada tahapan ini maka peneliti mencoba untuk memahami data yang akan digunakan Data dari twitter (X) yang telah di ambil menggunakan pemrograman python berjumlah 5329 tweet dengan dua atribut yang akan dilakukan dengan cara pemilihan atribut menjadi dua atribut yaitu tweet dan sentimen. Data yang didapatkan dalam penelitian ini diperoleh tanggal 10 Oktober 2019 sampai 10 Juli 2024.

3) Data Preparation

Pada tahap data preparation dilakukan beberapa tahapan antara lain sebagai berikut :

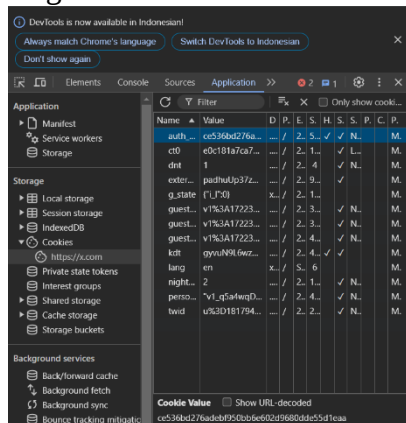
a. Pengumpulan Data

Untuk pengambilan data menggunakan pemrograman python dengan melalui google colab. Untuk tahapannya antara lain yaitu diawali instalasi package python, mengambil token dari twitter (X), lalu crawling data berdasarkan text, dan periode tanggal, dan yang terakhir data di save. Untuk instalasi package python bisa dilihat gambar dibawah ini.

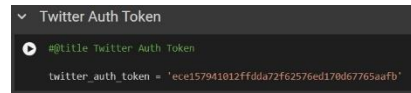


Gambar 1. Instalasi Package Phytion

Setelah melalui package python dapat terinstall dengan baik, selanjutnya mengambil token dari akun twitter (X), lalu setelahnya memasukan token seperti gambar dibawah ini.



Gambar 2. Pengambilan Token dari X



Gambar 3. Memasukan Token

Setelah memasukan token dan dapat terautentifikasi dengan baik maka selanjutnya tinggal memasukan keyword Kabinet Jokowi, limit yang dicari yaitu 5000 data dengan bentuk text, dan periode yang dicari yaitu dari 10 Oktober 2019 sampai 10 Juli 2024.



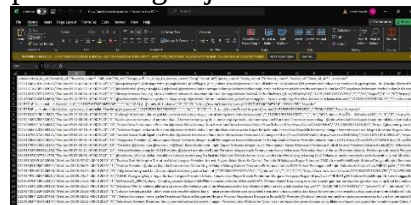
Gambar 4. Proses Crawling Data

Setelah proses crawling data, file yang sudah siap tinggal di save dan di export.



Gambar 5. Proses Export Data

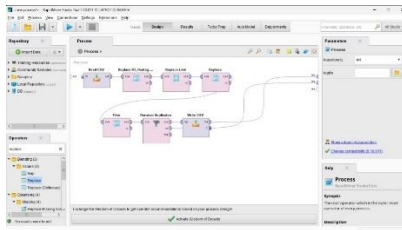
Setelah dapat di export data bisa didapatkan dengan jumlah 5329 data.



Gambar 5. Jumlah Data yang Didapat

b. Cleansing Data

Pada tahapan ini peneliti akan melakukan proses Cleansing dengan menggunakan rapidminer sebelum melakukan pemberian label pada sentimen, agar dataset pada file csv terbaru terhindar dari duplikasi data dan tanda yang tidak diperlukan. Pada proses ini akan dilakukan pembersihan dari berbagai noise seperti menghilangkan link URL, username, retweet, digit angka dan karakter. Setelah selesai menjalankan proses Cleansing dengan menggunakan rapidminer. Awalnya terdapat 5329 data tweet pada file csv, kemudian setelah melewati proses Cleansing maka data tweet berkurang menjadi 3437 data.



Gambar 6. Model Proses Cleansing



Gambar 7. Hasil Proses Cleansing

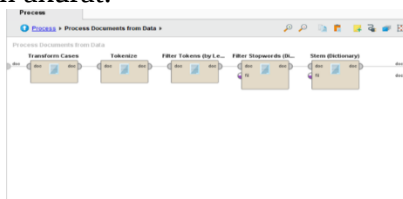
c. Labeling

Kemudian tahap selanjutnya adalah melakukan klasifikasi atau penentuan sentimen berdasarkan masing-masing tweet. Selama proses penentuan sentimen berlangsung, disini peneliti melakukan pemberian sentimen secara manual. Berikut ini data yang sudah di beri label.

Gambar 8. Hasil Labeling Data

d. Tokenizing

Pada proses ini adalah untuk mengubah teks dari kalimat-kalimat yang kompleks menjadi urutan token atau unit-unit yang lebih sederhana. Tujuannya untuk mempersiapkan data teks agar dapat diproses oleh algoritma analisis sentimen dengan lebih efektif dan akurat.



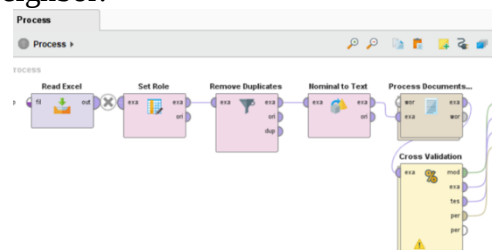
Gambar 9. Model Tokenizing

Gambar 10. Hasil Tokenizing Naïve Bayes

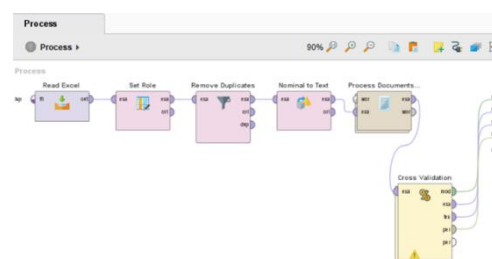
Gambar 11. Hasil Tokenizing K-Nearest Neighbor (KNN)

4) Modeling

Pada tahapan ini peneliti akan melakukan preprocessing pada dataset hal ini ditunjukkan untuk menyiapkan data yang bersih dan data yang bebas dari noise. Pada proses ini juga untuk menghitung pembobotan kata untuk keperluan pada proses modelling nanti. Setelah melewati tahapan text preparation, maka jumlah data tweet didapatkan menjadi 3437 data yang merupakan data bersih untuk dipakai pada tahap selanjutnya. Pada tahapan ini akan dilakukan pengukuran performa dua klasifikasi sekaligus dengan menggunakan algoritma Naïve bayes dan K-Nearest Neighbor.



Gambar 12. Model Proses Data Naïve bayes



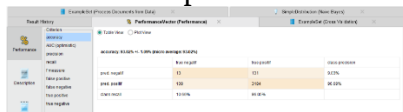
Gambar 13. Model Proses Data K-Nearest Neighbor

HASIL DAN PEMBAHASAN

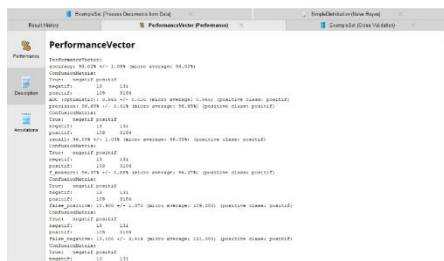
Pada penelitian ini akan dilakukan pengujian menggunakan algoritma Naïve bayes dan K-Nearest Neighbor dengan hasil sebagai berikut :

a. Algoritma Naïve bayes

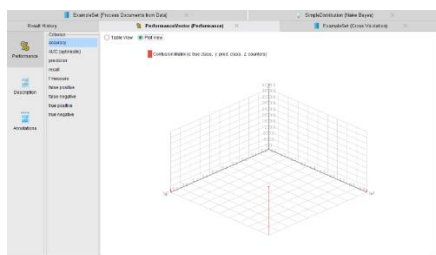
Total dataset yang dikumpulkan adalah 3437 data. Berikut ini adalah hasil dari tahapan modelling dengan menggunakan algoritma Naïve bayes yang dapat dilihat hasil dari perhitungan hasil Rapidminer dapat dilihat pada Gambar dibawah ini untuk menunjukkan dan membuktikan hasil prediksi.



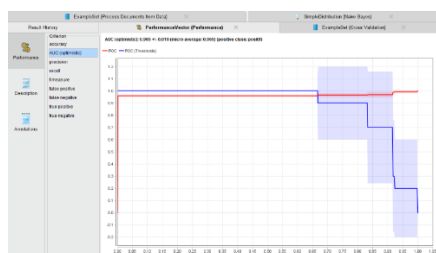
Gambar 14. Hasil Akurasi Naïve bayes



Gambar 15. Deskripsi Akurasi Naïve bayes



Gambar 16. Plot View Naïve bayes



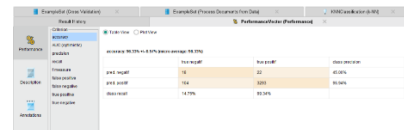
Gambar 17. Grafik AUC (Optimistic) Naïve bayes

Berdasarkan hasil diatas dapat disimpulkan bahwa accuracy menunjukkan data yang benar diprediksi positif yang

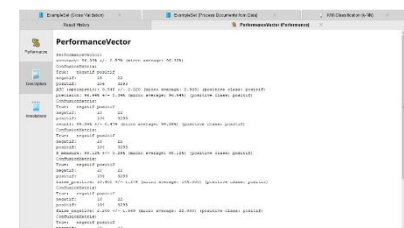
menghasilkan presentase ketepatannya adalah 96.69%, sedangkan untuk data bersentimen Negatif memiliki precision sebesar 9.03%. Untuk sentimen Positif memiliki recall (Specificity) adalah 96.05%, sedangkan pada sentimen negatif memiliki recall sebesar 10.66%. Nilai accuracy yang dihasilkan menggunakan model Naïve bayes adalah 93.02%, sehingga dapat disimpulkan bahwa algoritma Naïve bayes dapat mengklasifikasi sentiment cukup menggunakan data Kabinet Jokowi.

b. Algoritma K-Nearest Neighbor

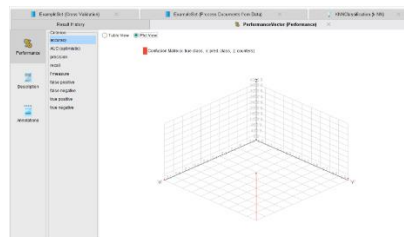
Total dataset yang dikumpulkan adalah 3437 data. Berikut ini adalah hasil dari tahapan modelling dengan menggunakan algoritma K-Nearest Neighbor (KNN) yang dapat dilihat hasil dari perhitungan hasil Rapidminer dapat dilihat pada Gambar dibawah ini untuk menunjukkan dan membuktikan hasil prediksi.



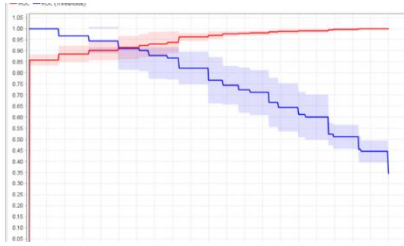
Gambar 18. Hasil Akurasi K-Nearest Neighbor



Gambar 19. Deskripsi Hasil Akurasi K-Nearest Neighbor



Gambar 20. Plot View K-Nearest Neighbor



Gambar 21. Grafik AUC (Optimistic) K-Nearest Neighbor

Berdasarkan hasil diatas dapat disimpulkan bahwa accuracy menunjukkan data yang benar diprediksi positif yang menghasilkan presentase ketepatannya adalah 96.94%, sedangkan untuk data bersentimen Negatif memiliki precision sebesar 45.00%. Untuk sentimen Positif memiliki recall (Specificity) adalah 99.34% , sedangkan pada sentimen negatif memiliki recall sebesar 14.75%. Nilai accuracy yang dihasilkan menggunakan model *K-Nearest Neighbor* adalah 96.33%, sehingga dapat disimpulkan bahwa algoritma *Naïve Bayes* dapat mengklasifikasi sentiment dengan baik menggunakan data Kabinet Jokowi.

PENUTUP

Kesimpulan

Berdasarkan penelitian dan penerapan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* dalam analisis sentimen terhadap Kabinet Jokowi dapat disimpulkan bahwa :

- Berdasarkan penelitian yang telah dilakukan mengenai klasifikasi opini masyarakat Indonesia terhadap Kabinet Jokowi dapat diambil kesimpulan yaitu dari 3437 data ditemukan 3315 atau 98% data bersentimen positif dan 122 data atau 2% bersentimen negatif.
- Dari hasil penelitian terdapat masyarakat Indonesia sebagian besar memiliki respon positif terhadap Kabinet Jokowi dan berdasarkan klasifikasi model algoritma *Naïve Bayes* terhadap dataset Kabinet Jokowi, didapatkan nilai accuracy sebesar 93.02%. Sedangkan jika menggunakan algoritma *K-Nearest Neighbor* mendapatkan nilai accuracy 96.33%.

- Sehingga dapat dikatakan bahwa algoritma *K-Nearest Neighbor* lebih baik dibandingkan algoritma *Naïve Bayes* dalam mengklasifikasikan data sentimen tentang Kabinet Jokowi

DAFTAR PUSTAKA

- [1] D. A. Lestari and D. Mahdiana, "Penerapan Algoritma K-Nearest Neighbor pada Twitter untuk Analisis Sentimen Masyarakat Terhadap Larangan Mudik 2021".
- [2] M. Hidayat and U. Budiyanto, "3 rd Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI) 30 Agustus 2023- Jakarta," 2023.
- [3] Z. Rahman Hakim and S. Sugiyono, "Analisa Sentimen Terhadap Kereta Cepat Jakarta – Bandung Menggunakan Algoritma *Naïve Bayes* Dan *K-Nearest Neighbor*," Jurnal Sains dan Teknologi, vol. 5, no. 3, pp. 939–945, Apr. 2024, doi: 10.55338/saintek.v5i3.1423.
- [4] M. Syarifuddin, "ANALISIS SENTIMEN OPINI PUBLIK TERHADAP EFEK PSBB PADA TWITTER DENGAN ALGORITMA DECISION TREE,KNN, DAN *NAÏVE BAYES*," INTI Nusa Mandiri, vol. 15, no. 1, pp. 87–94, Aug. 2020, doi: 10.33480/inti.v15i1.1433.
- [5] M. I. Ahmadi, D. Gustian, and F. Sembiring, "Analisis Sentiment Masyarakat terhadap Kasus Covid-19 pada Media Sosial Youtube dengan Metode *Naive bayes*," 2021.
- [6] H. Hilmi Zain, R. Maulana Awangga, and W. Isti Rahayu, "Perbandingan Model Svm, Knn Dan *Naïve Bayes* Untuk Analisis Sentiment Pada Data Twitter: Studi Kasus Calon Presiden 2024," JIMPS: Jurnal Ilmiah Mahasiswa Pendidikan Sejarah, vol. 8, no. 3, pp. 2083–2093, doi: 10.24815/jimps.v8i3.25342.

- [7] D. Sandi and E. Utami, "Analisis Sentimen Publik Terhadap Elektabilitas Ganjar Pranowo di Tahun Politik 2024 di Twitter dengan Algoritma KNN dan Naïve Bayes," vol. 7, no. 3, pp. 1097–1108, 2023, doi: 10.30865/mib.v7i3.6298.
- [8] D. Era, S. Andryana, and A. Rubhasy, "Perbandingan Algoritma Naïve Bayes Dan K-Nearest Neighbor pada Analisis Sentimen Pembukaan Pariwisata Di Masa Pandemi Covid 19," 2023.
- [9] D. Pramana, M. Afdal, and I. Permana, "Analisis Sentimen Terhadap Pemindahan Ibu Kota Negara Menggunakan Algoritma Naive Bayes Classifier dan K-Nearest Neighbors," vol. 7, no. 3, pp. 1306–1314, 2023, doi: 10.30865/mib.v7i3.6523.
- [10] Y. Nurtikasari, Syariful Alam, and Teguh Iman Hermanto, "Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes," *INSOLOGI: Jurnal Sains dan Teknologi*, vol. 1, no. 4, pp. 411–423, Aug. 2022, doi: 10.55123/insologi.v1i4.770.
- [11] M. Furqan, S. Mayang Sari, and P. Ilmu Komputer Fakultas Sains dan Teknologi, "Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia Sentiment Analysis using K-Nearest Neighbor towards the New Normal During the Covid-19 Period in Indonesia," 2022. [Online]. Available: www.tripadvisor.com
- [12] A. Sentimen, "Analisis Sentimen Pengguna Twitter terhadap Perpanjangan PPKM Menggunakan Metode K-Nearest Neighbor".
- [13] F. S. Pamungkas and I. Kharisudin, "Analisis Sentimen dengan SVM," vol. 4, pp. 628–634, 2021, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [14] N. Hardi et al., "SISTEMASI: Jurnal Sistem Informasi Analisis Sentimen Physical Distancing pada Twitter Menggunakan Text Mining dengan Algoritma Naive Bayes Classifier." [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [15] F. STMIK Syaikh Zainuddin Nahdlatul Wathan Anjani, "https://bufnets.tech https://doi.org/10.59688/bufnets BULLETIN OF NETWORK ENGINEER AND PERBANDINGAN ALGORITMA NAÏVE BAYES DAN KNN DALAM ANALISIS SENTIMEN MASYARAKAT TERHADAP PELAKSANAAN PPPK GURU COMPARISON OF NAÏVE BAYES AND KNN ALGORITHMS IN THE ANALYSIS OF PUBLIC SENTIMENT REGARDING THE IMPLEMENTATION OF GOVERNMENT EMPLOYEES WITH TEACHER WORK AGREEMENTS", doi: 10.59688/bufnets.
- [16] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *Jurnal KomtekInfo*, pp. 1–7, Jan. 2023, doi: 10.35134/komtekinf.v10i1.330.
- [17] S. H. Ramadhani and M. I. Wahyudin, "Analisis Sentimen Terhadap Vaksinasi Astra Zeneca pada Twitter Menggunakan Metode Naïve Bayes dan K-NN," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 6, no. 4, p. 2022, 2022, doi: 10.35870/jti.
- [18] K. Zuhri, N. Adha, and O. Saputri, "Analisis Sentimen Masyarakat Terhadap Pilpres 2019 Berdasarkan Opini Dari Twitter Menggunakan Metode Naive Bayes Classifier," 2020. [Online]. Available:

<https://journal-computing.org/index.php/journal-cisa/index>

- [19] A. Y. Permana and H. Noviyani, "Tahun 2022 Call for papers dan Seminar Nasional Sains dan Teknologi Ke-1 2022 Fakultas Teknik, Universitas Pelita Bangsa," vol. 1, no. 1, p. 128, 2022.
- [20] J. Supriyanto, D. Alita, and A. R. Isnain, "Penerapan Algoritma K-Nearest Neighbor (K-NN) Untuk Analisis Sentimen Publik Terhadap Pembelajaran Daring," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 4, no. 1, pp. 74–80, Mar. 2023, doi: 10.33365/jatika.v4i1.2468