

**ANALISIS SENTIMEN MENGENAI KESADARAN MASYARAKAT  
INDONESIA TERHADAP KEAMANAN SIBER DALAM  
MENGHADAPI KEBOCORAN DATA MENGGUNAKAN  
ALGORITMA NAÏVE BAYES CLASSIFIER**

**SENTIMENT ANALYSIS ABOUT INDONESIAN PEOPLE'S  
AWARENESS ABOUT CYBER SECURITY IN DETERMINING  
DATA LEAKAGE USING NAÏVE BAYES CLASSIFIER  
ALGORITHM**

**Dinna Nurfadlillah**

Universitas Alma Ata Yogyakarta  
[\*\*dinnanf1503@gmail.com\*\*](mailto:dinnanf1503@gmail.com)

**Abstrak**

Kebocoran data pribadi merupakan tantangan serius, dampak buruk dari insiden kebocoran data dapat menimbulkan denda, hilangnya reputasi hingga kepercayaan publik yang telah dirugikan. Keamanan siber berperan dalam meminimalisir tingkat terjadinya resiko ancaman serangan siber yang mengganggu keamanan seluruh komponen sistem seperti infrastruktur, software, hardware dan data. Penelitian ini dilakukan untuk menganalisis sentimen masyarakat Indonesia terhadap query “keamanan siber” dan “kebocoran data” dari tweet pengguna Twitter. Klasifikasi data pada penelitian ini menggunakan metode algoritma Naïve Bayes Classifier. Tujuan dari penelitian ini untuk mengetahui performa presentase dari sentimen positif dan negatif. Hasil dari performa pemodelan yaitu memperoleh nilai presentase accuracy 88,80%, Precision 29,27%, dan Recall 30,77%. Hasil visualisasi data menunjukkan bahwa masyarakat Indonesia khususnya pengguna Twitter banyak memberikan pendapat negatif dalam menanggapi pembahasan ini. Sehingga presentase dari sentimen negatif lebih tinggi dari sentimen positif.

**Kata kunci : Analisis sentimen, Twitter, Naïve Bayes Classifier, Kebocoran data, Keamanan Siber**

**Abstract**

Leakage of personal data is a serious challenge, the adverse effects of incidents of data leakage can result in fines, loss of reputation and public trust that has been harmed. Cyber security plays a role in minimizing the level of risk of cyberattack threats that disrupt the security of all system components such as infrastructure, software, hardware and data. This research was conducted to analyze the sentiments of the Indonesian people towards the queries "cyber security" and "data leaks" from Twitter user tweets. The data classification in this study uses the Naïve Bayes Classifier algorithm method. The purpose of this study is to determine the percentage performance of positive and negative sentiments. The results of the modeling performance are obtaining accuracy percentage values of 88.80%, Precision 29.27%, and Recall 30.77%. The results of the data visualization show that Indonesian people, especially Twitter users, give many negative

opinions in response to this discussion. So the percentage of negative sentiment is higher than positive sentiment.

**Keywords: Sentiment analysis, Twitter, Naïve Bayes Classifier, Data leakage, Cybersecurity**

## 1. PENDAHULUAN

Perkembangan digital dan kemajuan teknologi informasi telah memberikan dampak positif dan banyak manfaat bagi kehidupan manusia. Sehingga kebutuhan dan gaya hidup masyarakat semakin bergantung terhadap teknologi [1]. Seiring dengan perkembangannya pun mengakibatkan berbagai macam kejahatan baru yang disebut cybercrime yang dimana kejahatannya dilakukan dalam jaringan internet [1]. Beberapa kasus cybercrime yang terjadi di Indonesia antara lain penipuan, spamming, hacking, hingga penyadapan data pribadi penting milik orang lain [2].

Sebuah insiden kebocoran data terjadi pada tahun 2020 pada marketplace Tokopedia yang menyebabkan puluhan juta akun yang berisikan data-data penting pengguna bocor [3]. Tokopedia merupakan marketplace terkenal di Indonesia dan memiliki cukup banyak pengguna yang mencakup hampir seluruh lapisan masyarakat [4]. Pada September 2022 lalu Indonesia dihebohkan dengan bocornya data mulai dari data Indihome sampai PLN. Salah satu hacker juga mengklaim memiliki data 1 miliar nomor telepon seluler Indonesia yang dijual bebas dengan harga 50.000 dolar atau setara dengan Rp 743 juta [5]. Kebocoran data pribadi merupakan tantangan serius, dampak buruk dari insiden kebocoran data dapat menimbulkan denda, hilangnya reputasi hingga kepercayaan publik yang telah dirugikan [6]. Karena kebocoran data yang berisikan data pribadi kemungkinan akan disalahgunakan oleh para oknum yang tidak bertanggungjawab [7]. Keamanan siber (Cyber Security) merupakan hal yang sangat penting menjadi fokus utama bagi setiap penyedia layanan karena menyimpan berbagai informasi pengguna yang bersifat pribadi [8]. Hasil Indeks Keamanan Informasi (KAMI) juga menunjukkan sebagian besar pengelola keamanan siber tidak memiliki kompetensi yang memadai [9].

Twitter adalah salah satu media jejaring sosial terbesar di dunia, didirikan oleh Jack Dorsey pada Maret 2006 dan resmi diluncurkan pada 15 Juli 2006 ini menjadi salah satu situs yang paling sering dikunjungi di internet dan kerap dipakai oleh masyarakat umum termasuk masyarakat Indonesia dalam menyampaikan segala aspirasi maupun sentimen pendapat di dunia maya [10]. Dengan melihat pembahasan mengenai penggunaan Twitter, maka Twitter sendiri dapat menjadi sumber data yang berisi pendapat dan sentimen masyarakat yang dapat digunakan untuk pemasaran maupun studi sosial [11].

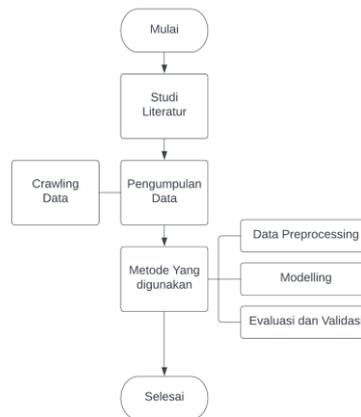
Analisis sentimen merupakan proses mengolah data berupa teks yang mengandung opini atau penilaian pada suatu objek [12]. Dataset yang digunakan pada penelitian merupakan dataset yang diambil dari media sosial Twitter yang berisi tweet dan komentar pendapat yang bersifat positif maupun negatif [13]. Dalam penyusunan penelitian analisis sentimen ini dilakukan dengan menggunakan metode algoritma Naïve Bayes Classifier yang merupakan algoritma klasifikasi machine learning yang umumnya digunakan untuk bidang text mining dan analisis sentimen [14]. Penerapan metode Naïve Bayes Classifier diharapkan mampu untuk menghasilkan performa yang baik dalam melakukan pemodelan klasifikasi.

Pada penelitian ini, struktur data yang digunakan terdiri dari variabel independen dan variabel dependen. Dimana variabel independen (X) dari penelitian ini merupakan tweet yang berisikan opini masyarakat mengenai topik yang dibahas. Variabel dependen (Y) dari penelitian ini adalah klasifikasi sentimen tweet (positif dan negatif). Sehingga hasil dari analisis dapat digunakan untuk mengetahui bagaimana sentimen masyarakat dalam menyikapi pentingnya keamanan siber

dalam menghadapi kebocoran data melalui tweet berisikan opini yang mereka tuangkan di media sosial Twitter.

## 2. METODOLOGI

Penelitian mengenai Analisis Sentimen Mengenai Kesadaran Masyarakat Indonesia Terhadap Keamanan Siber Dalam Menghadapi Kebocoran Data Menggunakan Algoritma Naïve Bayes Classifier ini merupakan jenis metode analisis kuantitatif statistik deskriptif. Statistik deskriptif adalah teknik metode yang digunakan untuk tujuan menganalisis data dengan cara mendeskripsikan data yang telah tersedia tanpa membuat sebuah kesimpulan yang berlaku secara luas dan umum [15]. Teknik metode ini juga biasa digunakan dalam penelitian yang bertujuan untuk mengetahui sentimen masyarakat di Twitter mengenai suatu isu atau informasi yang sedang terjadi dan sebagainya. Penelitian-penelitian jenis ini biasanya hanya mencoba untuk mengungkapkan dan mendeskripsikan hasil penelitiannya [16]. Selanjutnya peneliti menggambarkan metodologi yang digunakan dalam penelitian ini ke dalam sebuah skema alur penelitian pada Gambar 1:



Gambar 1 Alur Metode Penelitian

### 2.1 Studi Literatur

Studi literatur dilakukan untuk memperoleh materi-materi dan referensi agar penelitian menjadi baik. Adapun yang peneliti lakukan dalam studi kepustakaan seperti referensi buku, skripsi penelitian terdahulu, dan penelusuran data online. Studi pustaka dan literatur dilakukan agar peneliti dapat memahami mengenai topik penelitian yang dipilih dan materi maupun informasi yang dibutuhkan dapat dijadikan sebagai acuan dalam penelitian dan mendapatkan solusi dari permasalahan yang sedang dihadapi. Beberapa materi dan informasi yang digunakan dalam penyusunan penelitian ini seperti mengenai Analisis Sentimen, Twitter, Text Mining, Text Preprocessing, RapidMiner, dan Naïve Bayes Classifier.

### 2.2 Pengumpulan Data

Penelitian dilakukan dengan cara crawling yaitu metode untuk memperoleh semua data informasi yang dibutuhkan dari berbagai sumber internet maupun website [17]. Pada penelitian ini data yang diambil berupa tweet pengguna media sosial Twitter mengenai isu keamanan siber dan kebocoran data. Data yang diambil dari Twitter adalah data yang berbahasa Indonesia. Crawling dilakukan dengan memberikan kata kunci dalam periode tertentu. Sebanyak >1000 dataset dari Twitter akan dilakukan penelitian analisis sentimen. Proses crawling data dilakukan dengan

menggunakan sncrape. Pada penelitian ini data yang diambil bersumber dari Twitter dengan tweet yang berisikan opini dengan berbagai kategori sentimen. Pada pengambilan dataset keyword yang digunakan yaitu “keamanan siber” dan “kebocoran data” dengan rentang waktu dari tanggal 01 September 2022 sampai dengan 30 September 2022. Hasil pengambilan data akan dibagi menjadi 2 yaitu data untuk training atau data yang akan dijadikan pemodelan, kemudian data untuk testing atau data yang akan dijadikan pengujian klasifikasi.

### 2.3 Metode Yang Digunakan

Dalam penyusunan skripsi yang baik diperlukan suatu metode penelitian yang sesuai dengan permasalahan. Metode penelitian dipergunakan sebagai sarana untuk menentukan langkah-langkah apa saja dalam penyusunan penelitian. Pembahasan metode yang digunakan dalam penyusunan penelitian ini meliputi sebagai berikut:

Data Preprocessing merupakan tahap awal pengolahan yang digunakan agar data teks yang terdapat noise seperti simbol, duplicate, dan tanda baca lainnya dihilangkan. Ada beberapa tahap data preprocessing yang digunakan dalam penelitian ini, antara lain:

1. Pelabelan (labelling) pada data dilakukan untuk menentukan apakah dataset tweet termasuk ke dalam kelas sentimen positif atau kelas sentimen negatif, pelabelan dilakukan dengan menggunakan menggunakan TextBlob library dengan menentukan subjektivitas atau polaritas.
2. Cleansing, tahap ini dilakukan penghapusan data duplikasi (remove duplicate), hashtag (#), retweet (rt), mention (@), url, angka, simbol-simbol, dan emoji.
3. Case Folding, perubahan semua karakter huruf pada kalimat menjadi lowcase (huruf kecil).
4. Tokenizing, memisahkan kalimat atau dipecah menjadi perkata sehingga bisa mempermudah proses.
5. Stopwords, pada proses ini dilakukan penghapusan kata-kata yang kurang bermakna atau tidak memiliki kepentingan dalam analisis sentimen. Sehingga yang diperlukan hanyalah kata-kata yang penting atau hanya memiliki arti.
6. Pembobotan Kata, pembobotan kata dilakukan untuk menghitung frequency suatu term pada dokumen. Pembobotan kata dilakukan dengan menggunakan TF-IDF.

### 2.4 Modelling

Pemodelan (modelling) dilakukan terhadap dataset yang sudah dilakukan pelabelan dan sudah melalui tahap text preprocessing. Pemodelan menggunakan algoritma Naïve Bayes Classifier dan pada tahap modelling ini juga dilakukan uji validasi, dataset akan melalui tahap uji validasi untuk mengukur performa model dengan menggunakan K-Fold Cross Validation. Pada tahap cross validation, dataset akan dibagi secara acak kedalam 10 folds. Bentuk umum Naïve Bayes Classifier dapat di formulasikan dengan persamaan [18].

$$P(H|X) = \frac{P(H|X) P(H)}{P(X)} \quad (1)$$

Keterangan:

X = Data dengan class yang belum diketahui

H = Hipotesis data X merupakan suatu class spesifik

- $P(H | X)$  = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)
- $P(H)$  = Probabilitas hipotesis H (prior prob.)
- $P(H | X)$  = Probabilitas X berdasarkan kondisi tersebut
- $P(X)$  = Probabilitas dari X

**2.5 Evaluasi**

Dalam penelitian ini evaluasi dilakukan dengan mengukur performa hasil dari pemodelan dengan algoritma Naïve Bayes Classifier yang sudah dilakukan terhadap dataset yang diberi label. Evaluasi menggunakan pengukuran yang terdapat pada RapidMiner yaitu tabel confusion matrix. Confusion matrix pada penelitian ini dihitung dengan menggunakan Accuracy, Precision, dan Recall [20].

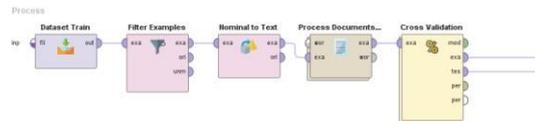
Tabel 1 Confusion Matrix

		Nilai Sebenarnya	
		TRUE	FALSE
Nilai Prediksi	TRUE	TP	FP
	FALSE	FN	TN

**3. PEMBAHASAN**

**3.1 Modelling**

Pada tahapan modelling ini, peneliti melakukan pemodelan menggunakan data training yang sudah diketahui kategori sentimennya dengan menggunakan metode Naïve Bayes Classifier (NBC). Pada penelitian ini peneliti menggunakan data training sebanyak 100 tweet yang sudah melalui tahap pelabelan dan data preprocessing.



Gambar 2 Modelling

Pada tahapan ini juga dilakukan pengukuran performa klasifikasi dengan menggunakan cross validation. Cross validation pada penelitian ini, data akan displit secara acak menjadi 10 folds. Dari data training yang sudah melewati cross validation akan diketahui pengukuran performa dari model yang digunakan yaitu Naïve Bayes Classifier.

**3.2 Hasil Evaluasi Confusion Matrix**

Pada hasil evaluasi dengan menggunakan confusion matrix yang diperoleh dari tahapan modelling dengan menggunakan algoritma Naïve Bayes Classifier dan menggunakan cross validation yaitu 10 folds. Selain itu, perhitungan confusion matrix pada penelitian ini dihitung dengan menggunakan Accuracy, Precision, dan Recall. Pada tahap evaluasi, sistem pemodelan yang telah dibuat menggunakan data yang sudah dilabeli, diperoleh nilai hasil confusion matrix yang mencatat hasil dari klasifikasi model. Pada tabel 2 diketahui confusion matrix yang diperoleh dari hasil evaluasi model Naïve Bayes Classifier:

Tabel 2 Confusion Matrix Naive Bayes

	<b>True Positif</b>	<b>True Negatif</b>	<b>Jumlah</b>
<b>Pred. Positif</b>	28	11	39
<b>Pred. Negatif</b>	15	46	61
<b>Jumlah</b>	43	57	100

Dari tabel 2 dapat diketahui TP (True Positive) adalah keseluruhan data dengan kelas sentimen positif terdapat ada 28 data. TN (True Negative) adalah seluruh data dengan kelas sentimen negatif yang berhasil diklasifikasikan atau diprediksi negatif terdapat ada 46 data. Adapun FP (False Positive) yaitu data yang bersentimen negatif namun diklasifikasi atau diprediksi positif terdapat ada 11 data. Sedangkan FN (False Negative) yaitu data yang bersentimen positif namun diklasifikasi atau diprediksi negatif terdapat ada 15 data.

Setelah melakukan pengumpulan data dan implementasi tahap-tahap yang digunakan dalam penelitian ini, didapatkan hasil modelling dan pengujian 10-folds cross validation. Berdasarkan hasil pengujian terhadap model yang telah dibuat dengan menggunakan metode 10-folds cross validation, dan dari hasil evaluasi confusion matrix diketahui nilai Accuracy, Precision dan Recall pada tabel 3.

Tabel 3 Hasil 10-Folds Cross Validation

<b>Parameter</b>	<b>Hasil (%)</b>
Accuracy	74,00%
Precision	77,15%
Recall	80,00%

Dari tabel 4.16, diketahui nilai Accuracy memiliki nilai 74,00%, Precision 77,15%, dan Recall 80,00%. Nilai Accuracy didapatkan dari jumlah banyaknya data yang diklasifikasikan dengan benar sesuai pelabelan sentimen dari jumlah seluruh data. Nilai Precision didapatkan dari jumlah data positif yang diklasifikasi benar sebagai data positif dibagi dengan jumlah data yang diklasifikasi sebagai data positif. Nilai Recall didapatkan dari jumlah data positif yang diklasifikasi benar sebagai data positif dibagi dengan jumlah data yang sebenarnya positif.

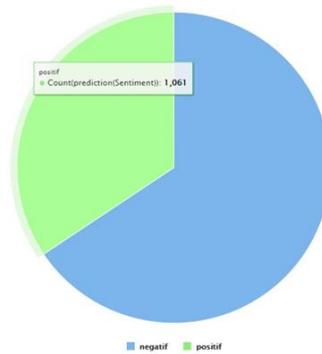
### 3.3 Pengujian Klasifikasi

Tahapan pengujian sistem pemodelan klasifikasi akan diterapkan pada data testing atau menggunakan data tweet baru yang telah dilakukan tahapan text preprocessing. Tujuan pengujian klasifikasi untuk mengetahui bagaimana respon atau sentimen masyarakat terhadap kebijakan topik keamanan siber dan kebocoran data.

Tabel 4 Presentase Prediksi Klasifikasi

<b>Sentimen</b>	<b>Jumlah</b>	<b>Presentase</b>
Positif	1061	34%
Negatif	2025	66%
<b>Jumlah</b>	3086	100%

Dari tabel 4 diketahui hasil yang diperoleh dari prediksi analisis sentimen pada data menunjukkan prediksi klasifikasi sentimen negatif yaitu 2025 data dengan presentase 66% lebih besar dibanding dengan prediksi klasifikasi sentimen positif yang memperoleh 1061 data dengan presentase 34%.



Gambar 3 Visualisasi Data Pie Chart

Dari gambar 3 diketahui hasil yang diperoleh dari prediksi analisis sentimen respon dan tanggapan masyarakat Indonesia khususnya pengguna Twitter mengenai kesadaran masyarakat Indonesia terhadap keamanan siber dalam menghadapi kebocoran data, dapat dilihat bahwa masih banyak tanggapan bersentimen negatif dibanding sentimen positif.



Gambar 4 Word Cloud Prediksi Sentimen Positif



Gambar 5 Word Cloud Prediksi Sentimen Negatif

Visualisasi data Word cloud mengenai prediksi sentimen pada gambar 4 dan 5 menunjukkan kata-kata yang sering muncul pada setiap kategori sentimen. Kata yang paling sering muncul adalah kata “data”, “kebocoran”, “keamanan”, dan “siber”. Hasil tersebut dikarenakan kata kunci yang digunakan sama yaitu mengenai kebocoran data dan keamanan siber.

#### 4. KESIMPULAN

Dari implementasi metode NBC dalam mengklasifikasikan sentimen mengenai kebocoran data dan keamanan siber menghasilkan performa dari hasil evaluasi confusion matrix, dimana memperoleh nilai presentase accuracy 74,00%, Precision 77,15%, dan Recall 80,00%. Berdasarkan

visualisasi data dari pie chart dan word cloud yaitu hasil prediksi klasifikasi pada pengujian sistem pemodelan dengan data tweet baru terdapat presentase sentimen negatif yaitu 66%. Sedangkan sentimen positif memperoleh sisanya yaitu 34%. Hal ini menunjukkan bahwa masyarakat Indonesia khususnya pengguna Twitter banyak memberikan pendapat negatif dalam menanggapi pembahasan ini yaitu keamanan siber dan kebocoran data. Sehingga presentase dari sentimen negatif lebih tinggi dari sentimen positif.

#### DAFTAR PUSTAKA

- [1] R. Sholehurrohman and I. Sabda Ilman, "Analisis Sentimen Tweet Kasus Kebocoran Data Penggunaan Facebook Oleh Cambridge Analytica," *J. Pepadun*, vol. 3, no. 1, pp. 140–147, 2022, doi: 10.23960/pepadun.v3i1.108.
- [2] M. R. T. R. Herryani and H. Njoto, "Perlindungan Hukum Terhadap Kebocoran Data Pribadi Konsumen Online Marketace," *Transparansi Huk.*, vol. 5, no. 1, pp. 110–133, 2022, [Online]. Available: <http://ojs.unik-kediri.ac.id/index.php/transparansihukum/article/view/2274>.
- [3] Djaenab and B. HB, "Tinjauan Yuridis Perlindungan Data Pribadi Terkait Kebocoran Data Dalam Ruang Cyber Crime," vol. 10, no. April, pp. 70–76, 2022.
- [4] N. I. Wibowo, T. A. Maulana, H. Muhammad, and N. A. Rakhmawati, "Perbandingan Algoritma Klasifikasi Sentimen Twitter Terhadap Insiden Kebocoran Data Tokopedia," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 6, no. 2, pp. 120–129, 2021, doi: 10.14421/jiska.2021.6.2.120-129.
- [5] N. Fitrianti, I. Marina, S. S. Qadrifa, P. Pascasarjana, I. Komunikasi, and I. Komunikasi, "Integrated Marketing Communication in E-commerce (a Study Case of Customers Data Leakage in Tokopedia)," *J. Interak. J. Ilmu Komun.*, vol. 6, no. 1, pp. 83–92, 2022, doi: 10.30596/interaksi.v6i1.8226.
- [6] Kompas, "Kilas Balik, Lima Kasus Kebocoran Data Pribadi di Indonesia...," *Kompas.com*, 2022. <https://www.kompas.com/cekfakta/read/2022/09/06/171100182/kilas-balik-limakasus-kebocoran-data-pribadi-di-indonesia->
- [7] H. B. Setiawan, & Fatma, and U. Najicha, "Perlindungan Data Pribadi Warga Negara Indonesia Terkait Dengan Kebocoran Data," *J. Kewarganegaraan*, vol. 6, no. 1, pp. 976–982, 2022.
- [8] R. N. Rohmah, "Cendekia Niaga Journal of Trade Development and Studies Upaya Membangun Kesadaran Keamanan Siber pada Konsumen E-commerce di Indonesia Ratri Nur Rohmah Abstrak," 2022.
- [9] T. Tan and B. Soewito, "Menggunakan Framework Nist Cybersecurity Di Universitas Zxc," vol. 6, no. 2, pp. 411–422, 2022, doi: 10.52362/jisamar.v6i2.781.
- [10] A. N. Aini and E. Wahyudi, "Model Perilaku Keamanan Siber Pada Pengguna Sosial Media Pada Masa Pandemi Covid-19," vol. 9, no. 1, pp. 37–40, 2022.
- [11] F. Ratnawati, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," 2018.
- [12] L. D. Utami, "Integrasi Metode Information Gain Untuk Seleksi Fitur dan Adaboost Untuk Mengurangi Bias Pada Analisis Sentimen Review Restoran Menggunakan Algoritma Naive Bayes," vol. 1, no. 2, pp. 120–126, 2015.

- [13] A. Rozaq, Y. Yunitasari, K. Sussolaikah, E. Resty, N. Sari, and R. I. Syahputra, "Analisis Sentimen Terhadap Implementasi Program Merdeka Belajar Kampus Merdeka Menggunakan Naïve Bayes , K-Nearest Neighbors Dan Decision Tree," vol. 6, no. April, pp. 746–750, 2022, doi: 10.30865/mib.v6i2.3554.
- [14] F. Yahyadi, Adi; Latifah, "Analisis Sentimen Twitter Terhadap Kebijakan Ppkm Di Tengah Pandemi Covid-19," J. Inf. Syst. Applied, Manag. Account. Res., vol. 6, no. 2, pp. 464–471, 2022, doi: 10.52362/jisamar.v6i2.791.
- [15] T. Wiratama Putra, A. Triayudi, and A. Andrianingsih, "Analisis Sentimen Pembelajaran Daring Menggunakan Metode Naïve Bayes, KNN, dan Decision Tree," J. JTIK (Jurnal Teknol. Inf. dan Komunikasi), vol. 6, no. 1, pp. 20–26, 2022, doi: 10.35870/jtik.v6i1.368.
- [16] R. Rudini, "Peranan Statistika Dalam Penelitian Sosial Kuantitatif," J. SAINTEKOM, vol. 6, no. 2, p. 53, 2017, doi: 10.33020/saintekom.v6i2.13.
- [17] A. Muhson, "Teknik Analisis Kuantitatif 1 TEKNIK ANALISIS KUANTITATIF," 2006.
- [18] A. Dewandaru and J. S. Wibowo, "Jurnal Teknologi dan Manajemen Informatika Analisis Sentimen dan Klasifikasi Tweet Terkait Mutasi COVID-19 Menggunakan Metode Naïve Bayes Classifier," vol. 8, no. 1, pp. 32–38, 2022.
- [19] A. Dewandaru and J. S. Wibowo, "Jurnal Teknologi dan Manajemen Informatika Analisis Sentimen dan Klasifikasi Tweet Terkait Mutasi COVID-19 Menggunakan Metode Naïve Bayes Classifier," vol. 8, no. 1, pp. 32–38, 2022.
- [20] R. Indra Borman and M. Wati, "Penerapan Data Mining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandarlampung Dengan Algoritma Naïve Bayes," J. Ilm. Fak. Ilmu Komput., vol. 09, no. 01, pp. 25–34, 2020.