

ANALISIS SENTIMEN REVIEW PRODUK MOTOR HONDA PCX DAN YAMAHA N-MAX PADA TWITTER MENGGUNAKAN ALGORITMA NAÏVE BAYES

Rizqho Sayid Ali Al-Zaelani, Yudhi Raymond Ramadhan, Mutiara Andayani Komara

Program Studi Teknik Informatika S1, Sekolah Tinggi Teknologi Wastukencana

Jl. Cikopak No.53 Sadang Purwakarta, Jawa Barat, Indonesia

rizqhosaqid00@wastukencana.ac.id

ABSTRAK

Berkembangnya dunia teknologi informasi dan komunikasi yang pesat tidak jauh dari para penyedia layanan web yang menyajikan informasi beragam, contoh salah satunya yaitu data *text* yang didapat dari *twitter*. Namun, karena jumlah dan variasi komentar yang begitu banyak, membacanya satu per satu akan memakan waktu yang lama. Jika seorang pengunjung hanya membaca sedikit komentar, informasi yang diperoleh dapat menjadi tidak objektif. Untuk mengatasi masalah tersebut, perlu dilakukan analisis sentimen pada *twitter* menggunakan algoritma *naïve bayes*. *Tweet* diklasifikasi menjadi 3 kategori yaitu sentimen positif, negatif dan netral. Proses analisis sentimen terdiri dari tahapan *crawling*, *labelling*, *text processing*, pembobotan menggunakan perhitungan TF-IDF, implementasi *naïve bayes*, evaluasi menggunakan *confusion matrix*. Hasil penelitian menunjukkan bahwa honda pcx menghasilkan akurasi sebesar 0,79%, presisi positif 80%, presisi negatif 25%, presisi netral 90%, *recall* positif 80%, *recall* negatif 50%, *recall* netral 76%. Untuk yamaha n-max menghasilkan akurasi sebesar 72%, presisi positif 71%, presisi negatif 21%, presisi netral 93%, *recall* positif 71%, *recall* negatif 75%, *recall* netral 56%.

Kata kunci: *twitter, honda pcx, yamaha n-max, naïve bayes, teknologi, analisis, sentimen*

1. PENDAHULUAN

Berkembangnya dunia teknologi informasi dan komunikasi yang pesat tidak jauh dari para penyedia layanan web yang menyajikan informasi beragam. Informasi yang mengakibatkan peningkatan data yang mayoritas berwujud data teks mampu menjadi sumber sangat terkandung untuk menelusuri lebih bermakna, contoh salah satunya yaitu data *text* yang didapat dari *Twitter* [1].

PT Astra Honda Motor adalah sebuah perusahaan manufaktur yang aktif di sektor produksi kendaraan seperti mobil, motor, dan peralatan kendaraan. Sebagai perusahaan yang berukuran besar, Astra Honda Motor menerapkan berbagai strategi manajemen untuk mengembangkan, meningkatkan, dan menjaga nilai produknya. Hal ini bertujuan agar produk-produk yang dihasilkan memiliki kualitas tinggi dan memberikan manfaat yang besar bagi pengguna [2]. Penjualan sepeda motor mereka selama acara Gaikindo Indonesia *International Auto Show* (GIIAS) 2022. Selama acara tersebut, sebanyak 1.882 unit sepeda motor Honda terjual. Berdasarkan laporan AHM, penjualan mereka di GIIAS 2022 masih didominasi oleh PCX 160 dan ADV 160. Penjualan PCX 160 mencapai 349 unit, sementara ADV 160 terjual sebanyak 208 unit.

Yamaha Indonesia Motor *Manufacturing* (YIMM), yang juga dikenal sebagai perusahaan utama di Indonesia, merupakan bagian dari Yamaha Motor Company yang berpusat di Iwata, Shizuoka, Jepang. Yamaha yaitu perusahaan aktif dalam industri otomotif, terkhusus dalam produksi sepeda motor yang akhirnya didistribusikan ke dealer resmi Yamaha. Pada awalnya, Indonesia hanya melakukan perakitan sepeda motor Yamaha dengan menggunakan

komponen yang diimpor dari Jepang. Menurut Astam Sari Handayani dari Divisi Promosi Yamaha Sulselbar PT Suracojaya Abadimotor (PT SJAM), pada bulan November 2022, PT Suracojaya Abadimotor (SJAM) berhasil menjual sebanyak 5.441 unit. Dari jumlah tersebut, Nmax series menjadi produk terlaris dengan total penjualan sebanyak 1.953 unit.

Beberapa pengguna *Twitter* dan pembeli motor matic Honda PCX dan Yamaha N-MAX pasti akan mengungkapkan pendapat mereka tentang kedua produk tersebut melalui *Twitter*. Pendapat yang muncul di *Twitter* dapat dilihat oleh pengunjung melalui berbagai ulasan yang ada. Namun, karena jumlah dan variasi komentar yang begitu banyak, membacanya satu per satu akan memakan waktu yang lama. Jika seorang pengunjung hanya membaca sedikit komentar, informasi yang diperoleh dapat menjadi tidak objektif. Dari permasalahan tersebut perlu dilakukan analisis sentimen terhadap motor Honda PCX dan Yamaha N-MAX pada *Twitter* menggunakan algoritma *naïve bayes*, dengan tujuan untuk mengetahui hasil analisis sentimen terhadap motor Honda PCX dan Yamaha N-MAX pada *Twitter*.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen merupakan sebuah proses yang dimanfaatkan untuk mengevaluasi pendapat atau opini seseorang terhadap suatu topik. Tujuan utama dari analisis sentimen adalah mengelompokkan teks-teks dari dokumen, kalimat, atau fitur menjadi kategori positif, negatif, atau netral [3].

2.2. Motor Matic

Sepeda motor matic adalah jenis sepeda motor otomatis yang beroperasi tanpa perlu mengganti gigi secara manual, hanya dengan menggunakan satu tuas gas [4].

2.3. Media Sosial

Media sosial ialah platform daring yang memungkinkan pengguna untuk dengan mudah berinteraksi, berkolaborasi, dan berkontribusi dalam menciptakan berbagai konten, seperti blog, jejaring sosial, wiki, forum, dan dunia maya [5]

2.4. Twitter

Twitter adalah sebuah platform media sosial yang berbeda dengan format unik. Konsep utama

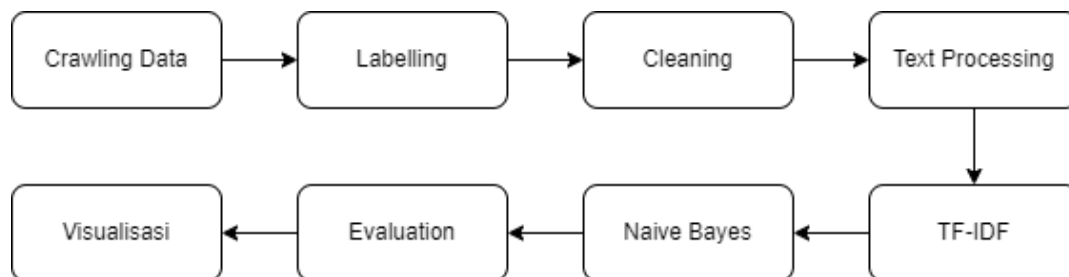
Twitter adalah menyebarkan informasi atau pesan dengan cepat, singkat, dan dalam waktu nyata (real-time) menggunakan kalimat yang tidak lebih dari 140 karakter. Dengan cakupan globalnya, Twitter menjadi sarana yang efektif untuk menyebarluaskan informasi kepada seluruh dunia dan semua orang [6]

2.5. Naïve Bayes

Naïve Bayes adalah metode klasifikasi yang sederhana namun mampu menghitung probabilitas dari berbagai kombinasi dan frekuensi nilai dalam basis data yang diperoleh [7]

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

3. METODE PENELITIAN



Gambar 1. Kerangka Pemikiran

a. Crawling Data

Proses ini melibatkan pengumpulan data *tweet* pengguna Twitter dari platform media sosial Twitter melalui proses pengambilan (*crawling*). Hal ini dilakukan dengan memanfaatkan kunci pelanggan (*customer key*), rahasia pelanggan (*customer secret*), token akses (*access token*), dan kunci akses (*access key*) yang diperoleh dari API Twitter [8]

b. Labelling

Pelabelan data adalah langkah untuk menempatkan label pada data sesuai dengan *kategori* yang sudah ditetapkan sebelumnya [9]

c. Text Processing

Preprocessing merupakan langkah awal dalam pengolahan data yang bertujuan untuk menghilangkan data yang mengganggu (*noise*) atau tidak konsisten. Tahap ini bertujuan untuk mengubah data mentah menjadi data berkualitas agar dapat diproses pada langkah selanjutnya dengan baik [10] *Text Processing* memiliki 4 tahapan yaitu :

a. Case Folding

Dalam proses ini, semua huruf kapital diubah menjadi huruf kecil. Dalam melakukan proses tersebut, karakter-karakter 'A'-'Z' dalam data diubah menjadi karakter 'a'-'z' [11]

b. Tokenizing

Tokenizing adalah langkah yang dilakukan untuk memisahkan teks, entah itu berupa

kalimat, paragraf, atau dokumen, menjadi bagian-bagian kecil yang disebut token. Proses ini melibatkan penghapusan tanda baca dan mengubah huruf kapital menjadi huruf kecil (*lower case*) pada teks [12]

c. Stopword Removal

Stopword Removal adalah langkah yang dilakukan dalam tahap *preprocessing* teks untuk menghapus kata-kata yang tidak penting dalam suatu kalimat berdasarkan daftar *stopword* [13]

d. Stemming

Stemming adalah suatu proses yang digunakan untuk menghubungkan beragam variasi morfologis dari kata-kata ke dalam bentuk dasar yang serupa. Proses ini dilakukan dengan menghapus semua afiksasi, termasuk awalan, sisipan, akhiran, dan kombinasi antara awalan dan akhiran pada kata-kata tersebut [14]

e. Term Frequency – Inverse Document Frequency (TF-IDF)

Metode TF-IDF adalah sebuah teknik yang digunakan untuk mengkalkulasi tingkat kepentingan setiap kata yang sering digunakan dalam pengambilan informasi. Metode ini dikenal karena efektivitas, kemudahan penggunaannya, serta kemampuannya dalam memberikan hasil yang akurat [15]

f. Naïve Bayes

Naïve Bayes merupakan metode klasifikasi yang sederhana namun mampu menghitung kemungkinan dari berbagai kombinasi dan frekuensi nilai berdasarkan data yang ada [7]

g. *Evaluation*

Confusion Matrix merupakan metode yang digunakan untuk menghitung akurasi, *recall*, *precision*. *Precision* mengukur kemampuan sistem dalam menemukan peringkat yang paling relevan dan diartikan sebagai persentase dokumen yang ditemukan dan benar-benar relevan terhadap *query*. *Recall* mengukur kemampuan sistem dalam menemukan semua item yang relevan dari koleksi dokumen dan diartikan sebagai persentase dokumen yang relevan terhadap *query*. *Accuracy* adalah perbandingan antara jumlah kasus yang diidentifikasi dengan benar terhadap seluruh kasus, sedangkan *error rate* adalah kasus yang diidentifikasi secara salah dibandingkan dengan jumlah seluruh kasus [16]

4. HASIL DAN PEMBAHASAN

4.1. Crawling Data

Crawling data dilakukan menggunakan *tools RapidMiner* dengan rentang waktu yaitu sejak tanggal 01-05-2023 sampai dengan 09-05-2023. Data yang didapat dari *Twitter* sebanyak 2502 data yang dibagi menjadi dua, 1500 data dari motor Yamaha N-MAX dan 1002 data dari motor Honda PCX.

4.2. Labelling

Pelabelan pada data tweet dilakukan secara otomatis menggunakan *Python* dan *library VaderSentiment*, dapat dilihat pada gambar dibawah ini :

```
#Compound Score Lexicon based
tweet_df.loc[tweet_df['Compound_Score'] < 0, 'Sentiments'] = 'Negatif'
tweet_df.loc[tweet_df['Compound_Score'] == 0, 'Sentiments'] = 'Netral'
tweet_df.loc[tweet_df['Compound_Score'] > 0, 'Sentiments'] = 'Positif'
tweet_df.head()
```

	Text	Compound_Score	Sentiments
0	based on the description of the perpetrators, ...	-0.2500	Negatif
1	present to you cat for nmax full mod modif sil...	-0.4019	Negatif
2	nmax vario mio beat f neck neck stainless diam...	0.0000	Netral
3	price comparison for Honda PCX and Yamaha Nmax...	0.0000	Netral
4	If after that I throw nmax don't get angry htt...	0.4023	Positif

Gambar 2. Labelling

4.3. Cleaning

Pada tahap ini dilakukan penghapusan data duplikat, dapat dilihat pada gambar berikut:

Tabel 1. Hasil *Cleaning*

Data	Sebelum Cleaning	Sesudah Cleaning	Data Terhapus
Honda PCX	1002	566	436
Yamaha N-MAX	1500	506	994

4.4. Text Processing

Pada tahap text processing ini menggunakan *tools Anaconda* serta bahasa *Python* dan memiliki 4 tahapan, yaitu:

a. *Case Folding*

Tabel 2. Hasil *Case Folding*

Sebelum	Sesudah
CMS Tutup Knalpot Tameng MK Pelindung Cover Knalpot Nmax	cms tutup knalpot tameng mk pelindung cover knalpot nmax

b. *Tokenizing*

Tabel 3. Hasil *Tokenizing*

Sebelum	Sesudah
Sebagai pengguna pcx meminta maaf atas silau yang diterima oleh orang orang yang telah melewati jalan dengan saya	['sebagai', 'pengguna', 'pcx', 'meminta', 'maaf', 'atas', 'silau', 'yang', 'diterima', 'oleh', 'orang', 'orang', 'yang', 'telah', 'melewati', 'jalan', 'dengan', 'saya']

c. *Stopword Removal*

Tabel 4. Hasil *Stopword Removal*

Sebelum	Sesudah
['sebagai', 'pengguna', 'pcx', 'meminta', 'maaf', 'atas', 'silau', 'yang', 'diterima', 'oleh', 'orang', 'orang', 'yang', 'telah', 'melewati', 'jalan', 'dengan', 'saya']	['pengguna', 'pcx', 'maaf', 'silau', 'diterima', 'melewati', 'jalan']

d. *Stemming*

Tabel 5. Hasil *Stemming*

Sebelum	Sesudah
['pengguna', 'pcx', 'maaf', 'silau', 'diterima', 'melewati', 'jalan']	['guna', 'pcx', 'maaf', 'silau', 'terima', 'lewat', 'jalan']

e. *Term Frequency – Inverse Document Frequency (TF-IDF)*

Tabel 6. TF-IDF

TERM	D1	D2	TF	D4	D5
			D3		
honda	1	0	0	0	0
pcx	1	1	1	1	1
standar	1	0	0	0	0
bagus	1	0	0	0	0
guna	0	1	0	0	0
maaf	0	1	0	0	0
silau	0	1	0	0	0
terima	0	1	0	0	0
lewat	0	1	0	0	0
jalan	0	1	0	0	0
elegan	0	0	1	0	0
ekonomis	0	0	0	1	0
kuat	0	0	0	0	1
sepeda	0	0	0	0	1
motor	0	0	0	0	1
Panjang Dokumen	4	7	2	2	4

Tabel 6. TF-IDF (Sambungan)

DF	IDF			TF-IDF		
		D1	D2	D3	D4	D5
1	0,699	0,699	0	0	0	0
5	0	0	0	0	0	0
1	0,699	0,699	0	0	0	0
1	0,699	0,699	0	0	0	0
1	0,699	0	0,699	0	0	0
1	0,699	0	0,699	0	0	0
1	0,699	0	0,699	0	0	0
1	0,699	0	0,699	0	0	0
1	0,699	0	0,699	0	0	0
1	0,699	0	0	0,699	0	0
1	0,699	0	0	0	0,699	0
1	0,699	0	0	0	0	0,699
1	0,699	0	0	0	0	0,699
1	0,699	0	0	0	0	0,699

f. *Naïve Bayes*

Dokumen ulasan yang telah melewati proses sebelumnya dikelompokkan dalam kategori tertentu berdasarkan kata yang terkandung di dalamnya. *Tools* yang digunakan untuk klasifikasi perhitungan *naïve bayes* yaitu *Ms. Excel*.

g. *Evaluation*• *Confusion Matrix Honda PCX*

```

prediksi benar: 45 data
prediksi salah: 12 data
Akurasi Algoritme: 78.94736842105263 %
TRUE NEGATIVE (TN): 2
FALSE NEGATIVE (FN1): 4
FALSE NEGATIVE (FN2): 2
TRUE POSITIVE (TP): 24
FALSE POSITIVE (FP1): 2
FALSE POSITIVE (FP2): 4
TRUE NETRAL (TI): 19
FALSE NETRAL (FI1): 2
FALSE NETRAL (FI2): 0
PRECISION1: 80.0 %
PRECISION2: 25.0 %
PRECISION0: 90.47619047619048 %
RECALL1: 80.0 %
RECALL2: 50.0 %
RECALL0: 76.0 %
AKURASI: 78.94736842105263 %
[[19 2 0]
 [ 4 24 0]
 [ 2 4 2]]

```

	precision	recall	f1-score	support
0	0.76	0.90	0.83	21
1	0.80	0.86	0.83	28
2	1.00	0.25	0.40	8

accuracy			0.79	57
macro avg	0.85	0.67	0.68	57
weighted avg	0.81	0.79	0.77	57

Gambar 3. Hasil Confusion Matrix Honda PCX

Berikut perhitungan akurasi (accuracy), presisi (precision) dan recall pada data Honda Pcx.

Tabel 7. Perhitungan Evaluasi Honda PCX

Sentimen	Akurasi	Presisi	Recall
Positif (1)	79%	80%	80%
Negatif (2)		25%	50%
Netral (0)		90%	76%

• *Confusion Matrix Yamaha N-MAX*

```

prediksi benar: 37 data
prediksi salah: 14 data
Akurasi Algoritme: 72.54901960784314 %
TRUE NEGATIVE (TN): 3
FALSE NEGATIVE (FN1): 7
FALSE NEGATIVE (FN2): 4
TRUE POSITIVE (TP): 20
FALSE POSITIVE (FP1): 1
FALSE POSITIVE (FP2): 7
TRUE NETRAL (TI): 14
FALSE NETRAL (FI1): 1
FALSE NETRAL (FI2): 0
PRECISION1: 71.42857142857143 %
PRECISION2: 21.428571428571427 %
PRECISION0: 93.33333333333333 %
RECALL1: 71.42857142857143 %
RECALL2: 75.0 %
RECALL0: 56.00000000000001 %
AKURASI: 72.54901960784314 %
[[14 1 0]
 [ 2 20 0]
 [ 4 7 3]]

```

	precision	recall	f1-score	support
0	0.70	0.93	0.80	15
1	0.71	0.91	0.80	22
2	1.00	0.21	0.35	14

accuracy			0.73	51
macro avg	0.80	0.69	0.65	51
weighted avg	0.79	0.73	0.68	51

Gambar 4. Hasil Confusion Matrix Yamaha N-MAX

Berikut perhitungan akurasi (accuracy), presisi (precision) dan recall pada data Yamaha N-MAX.

Tabel 8. Perhitungan Evaluasi Yamaha N-MAX

Sentimen	Akurasi	Presisi	Recall
Positif (1)	73%	71%	71%
Negatif (2)		21%	75%
Netral (0)		93%	72%

5. KESIMPULAN DAN SARAN

Hasil penelitian diatas mengenai Analisis Sentimen *Review* Motor Honda Pcx dan Yamaha N-Max Pada *Twitter* Menggunakan Algoritma *Naïve Bayes*, dapat disimpulkan bahwa pada penelitian ini untuk data Honda Pcx dengan jumlah sentimen positif 227, negatif 110 dan netral 229 menghasilkan nilai akurasi 79%, presisi positif 80%, presisi negatif 25%, presisi netral 90%, *recall* positif 80%, *recall* negatif 50%, *recall* netral 76%. Dan untuk data Yamaha N-Max dengan jumlah sentimen positif 189, negatif 120, netral 197 menghasilkan nilai akurasi 72%, presisi positif 71%, presisi negatif 21%, presisi netral 93%, *recall* positif 71%, *recall* negatif 75%, *recall* netral 56%. Pada nilai diatas menunjukkan bahwa data Honda Pcx memiliki hasil lebih baik dibanding Yamaha N-Max, Saran yang dapat diberikan peneliti untuk penelitian analisis sentimen selanjutnya yaitu menambahkan jumlah dataset yang lebih banyak agar menghasilkan hasil yang lebih maksimal lagi. Serta mengguna algoritma lainnya dalam menganalisis sentimen seperti Algoritma SVM.

DAFTAR PUSTAKA

- [1] F. Ratnawati, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film

- Pada Twitter,” *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 50, 2018, doi: 10.35314/isi.v3i1.335.
- [2] F. Cuandra, C. Angeline, J. F. Herwanto, and S. T. Putri, “Penerapan Manajemen Operasional PT Astra Honda Motor Sesuai Perspektif Teori Manajemen di Masa Pandemi,” *Optima*, vol. 6, no. 2, pp. 1–13, 2022.
- [3] M. R. Fahlevvi, “Analisis Sentimen Terhadap Ulasan Aplikasi Pejabat Pengelola Informasi Dan Dokumentasi Kementerian Dalam Negeri Republik Indonesia Di Google Playstore Menggunakan Metode Support Vector Machine,” *J. Teknol. dan Komun. Pemerintah.*, vol. 4, no. 1, pp. 1–13, 2022, doi: 10.33701/jtkp.v4i1.2701.
- [4] Y. Wiyandra, F. Yenila, and R. A. Mahessya, “Sistem Pakar Kerusakan Sepeda Motor Matic dengan Metoda Hybrid,” *J. KomtekInfo*, vol. 8, no. 2, pp. 145–153, 2021, doi: 10.35134/komtekinfo.v8i2.110.
- [5] T. Liedfray, F. J. Waani, and J. J. Lasut, “Peran Media Sosial Dalam Mempererat Interaksi Antar Keluarga Di Desa Esandom Kecamatan Tombatu Timur Kabupaten Tombatu Timur Kabupaten Minasa Tenggara,” *J. Ilm. Soc.*, vol. 2, no. 1, p. 2, 2022.
- [6] E. Annisa, B. Bara, K. Amelia, R. Zahara, and G. Kartini, “Penelitian tentang Twitter,” *J. Edukasi Nonform.*, vol. 3, no. 2, pp. 1–7, 2022, [Online]. Available: <https://ummaspul.ejournal.id/JENFOL/article/view/4531/1856>
- [7] R. Rachman and R. N. Handayani, “Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM,” *J. Inform.*, vol. 8, no. 2, pp. 111–122, 2021, doi: 10.31294/ji.v8i2.10494.
- [8] N. Legiawati, T. I. Hermanto, and Y. R. Ramadhan, “Analisis Sentimen Opini Pengguna Twitter Terhadap Perusahaan Jasa Ekspedisi Menggunakan Algoritma Naïve Bayes Berbasis PSO,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, p. 930, 2022, doi: 10.30865/jurikom.v9i4.4629.
- [9] R. Prabowo, H. Sujaini, T. Rismawan, J. H. Rekayasa Sistem Komputer Universitas Tanjungpura Jl Hadari Nawawi, K. Barat, and J. H. Hadari Nawawi, “Analisis Sentimen Pengguna Twitter Terhadap Kasus COVID-19 di Indonesia Menggunakan Metode Regresi Logistik Multinomial The Sentiment Analysis of Twitter User about COVID-19 Cases in Indonesia Using the Multinomial Logistics Regression Method,” vol. 11, no. 1, pp. 85–90, 2023, doi: 10.26418/justin.v11i1.57450.
- [10] F. Alghifari and D. Juardi, “Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes,” *J. Ilm. Inform.*, vol. 9, no. 02, pp. 75–81, 2021, doi: 10.33884/jif.v9i02.3755.
- [11] D. Juang, “Analisis Spam Dengan Menggunakan Naïve Bayes,” *J. Teknovasi*, vol. 03, no. 2, pp. 51–57, 2016, [Online]. Available: <http://www.ejurnal.plm.ac.id/index.php/Teknovasi/article/view/69%0Ahttps://www.ejurnal.plm.ac.id/index.php/Teknovasi/article/download/69/58>
- [12] G. R. Yunarfi, Ricky Simdy, and Jackson, “Implementasi Text Mining untuk Mengetahui Kata Abreviasi dalam Percakapan Media Sosial,” *J. Digit. Ecosyst. Nat. Sustain.*, vol. 1, no. 2, pp. 78–83, 2021.
- [13] R. Rinandyaswara, Y. A. Sari, and M. T. Furqon, “Pembentukan Daftar Stopword Menggunakan Term Based Random Sampling Pada Analisis Sentimen Dengan Metode Naïve Bayes (Studi Kasus: Kuliah Daring Di Masa Pandemi),” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 4, pp. 717–724, 2022, doi: 10.25126/jtiik.2022934707.
- [14] M. S. H. Simarangkir, “Studi Perbandingan Algoritma - Algoritma Stemming Untuk Dokumen Teks Bahasa Indonesia,” *J. Inkofar*, vol. 1, no. 1, pp. 40–46, 2017, doi: 10.46846/jurnalinkofar.v1i1.2.
- [15] R. Melita, V. Amrizal, H. B. Suseno, and T. Dirjam, “Penerapan Metode Term Frequency Inverse Document Frequency (Tf-Idf) Dan Cosine Similarity Pada Sistem Temu Kembali Informasi Untuk Mengetahui Syarah Hadits Berbasis Web (Studi Kasus: Hadits Shahih Bukhari-Muslim),” *J. Tek. Inform.*, vol. 11, no. 2, pp. 149–164, 2018, doi: 10.15408/jti.v11i2.8623.
- [16] A.- Arini, L. K. Wardhani, and D.- Octaviano, “Perbandingan Seleksi Fitur Term Frequency & Tri-Gram Character Menggunakan Algoritma Naïve Bayes Classifier (Nbc) Pada Tweet Hashtag #2019gantipresiden,” *Kilat*, vol. 9, no. 1, pp. 103–114, 2020, doi: 10.33322/kilat.v9i1.878.