 ISSN: 2715-9426

KINERJA ALGORITMA MACHINE LEARNING DALAM ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI LIVIN' BY MANDIRI PADA GOOGLE PLAY STORE

Ovelina Devi Kurnia[1)], Erna Daniati[2)] , Aidina Ristyawan[3)]
Program Studi Sistem Informasi, Universitas Nusantara PGRI Kediri [1),2),3)]

ernadaniati@unpkediri.ac.id[2)]

Abstrak

Analisis sentimen menjadi pendekatan penting dalam memahami opini pengguna terhadap suatu produk digital, termasuk aplikasi mobile banking. Penelitian ini bertujuan untuk mengklasifikasikan sentimen ulasan pengguna terhadap aplikasi *Livin' by Mandiri* yang diperoleh dari *Google Play Store*. Data dikumpulkan melalui teknik scraping dan diproses menggunakan metode *stratified balancing* berbasis perhitungan Slovin untuk menangani ketidakseimbangan kelas. Sebanyak 1925 data ulasan digunakan sebagai sampel penelitian. Representasi fitur dilakukan dengan teknik *TF-IDF*, dan model klasifikasi dibangun menggunakan tiga algoritma pembelajaran mesin, yaitu *Support Vector Machine* (SVM), *Logistic Regression* (LR), dan *Naive Bayes* (NB). Hasil evaluasi menunjukkan bahwa algoritma SVM memberikan performa terbaik dengan akurasi 82%, diikuti oleh LR (81%), dan NB (62%). Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score. Temuan ini menunjukkan bahwa pemilihan algoritma yang tepat berperan besar dalam efektivitas klasifikasi sentimen, khususnya pada data teks yang besar dan tidak seimbang. Penelitian ini memberikan kontribusi dalam pengembangan sistem evaluasi otomatis berbasis ulasan pengguna untuk mendukung peningkatan kualitas layanan aplikasi perbankan digital.

Kata kunci: analisis sentimen*, TF-IDF, stratified balancing, Mechine Learning, confution matrix*

1. Pendahuluan

Di era digital saat ini, analisis sentimen memainkan peran penting dalam memahami persepsi pengguna terhadap suatu produk atau layanan, khususnya melalui ulasan di platform digital seperti Google Play Store. Salah satu aplikasi yang banyak digunakan di Indonesia adalah Livin' by Mandiri, yang merupakan layanan mobile banking dari Bank Mandiri. Aplikasi ini menerima ribuan ulasan dari penggunanya yang mencerminkan pengalaman mereka dalam menggunakan layanan tersebut. Ulasan ini berpotensi memberikan informasi berharga terkait kualitas layanan, kemudahan penggunaan, serta keluhan atau apresiasi dari pengguna.[1]

Analisis sentimen merupakan salah satu bidang yang berkembang pesat dalam ranah pemrosesan bahasa alami (Natural Language Processing/NLP) dan klasifikasi teks. Teknik ini memiliki peran penting dalam berbagai aplikasi modern, seperti dalam dunia politik untuk memantau opini publik, dalam bisnis untuk menilai kepuasan pelanggan, serta dalam

periklanan dan strategi pemasaran untuk memahami preferensi konsumen. Seiring dengan meningkatnya volume data teks dari berbagai platform digital, analisis sentimen menjadi semakin relevan dan dibutuhkan dalam pengambilan keputusan berbasis data.[2]

Tujuan penelitian ini adalah untuk melakukan analisis sentimen terhadap ulasan pengguna aplikasi Livin' by Mandiri guna mengklasifikasikan opini pengguna ke dalam kategori sentimen positif dan negatif. Untuk itu, digunakan representasi fitur berbasis TF-IDF, serta model pembelajaran mesin seperti *Support Vector Machine* (SVM)**,** *Logistic Regression* (LR)**,** dan *Naive Bayes* (NB**)**. Tantangan utama dalam analisis ini adalah adanya ketidakseimbangan kelas serta volume data yang besar, sehingga pendekatan seperti stratified balancing data atau balancing manual diterapkan untuk menjaga distribusi data yang proporsional dan representatif.

Penelitian ini memberikan kontribusi dalam bentuk model klasifikasi sentimen yang dapat membantu pengembang aplikasi memahami opini pengguna secara sistematis dan efisien. Selain itu, hasil penelitian juga dapat digunakan sebagai acuan dalam peningkatan kualitas layanan mobile banking berbasis ulasan publik.

Hasil penelitian ini akan dibandingkan secara internal dan dibandingkan dengan penelitian sebelumnya yang menggunakan salah satu algoritma serupa untuk mereview pengguna aplikasi bank di google playstore. Penelitian sebelumnya telah dilakukan oleh ini Garda Zidane Dhamara, Danu Nur Alamsyah, Priyo Wildan Saputro, Erna Daniati, Aidina Ristyawan.. dengan judul "Analisis Sentimen Aplikasi MyBCA Melalui Review Pengguna di Google Play Store Menggunakan Algoritma Naive Bayes". Bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi myBCA berdasarkan ulasan yang diambil dari *Google Play Store*. Berdasarkan hasil evaluasi model. bahwa hasil klasifikasi model *Naive Bayes* mencapai 46,00 %. Temuan ini memberikan masukan untuk pengembangan aplikasi myBCA, terutama dalam meningkatkan kualitas layanan dan pengalaman pengguna.[3]

Dalam penelitian terdahulu yang berjudul, "Perbandingan Algoritma Machine Learning untuk Analisis Sentimen Berbasis Aspek pada Review Female Daily" yang dilakukan oleh Muhammad Hadiyan Wicaksono, Mahendra Dwiefbri Purbolaksono, Said Al Faraby. Data dibagi 80% pelatihan dan 20% pengujian. Fitur diekstraksi dengan *TF-IDF* dan dipilih menggunakan *Chi-Square*. Klasifikasi dilakukan dengan SVM, KNN, dan Random Forest. SVM dengan kernel linear memberikan akurasi tertinggi (67,1%). Preprocessing lengkap meningkatkan performa, meski stemming kurang signifikan akibat banyaknya kata tidak baku dan bahasa asing. Pemilihan algoritma dan teknik preprocessing sangat memengaruhi hasil klasifikasi. [4]

Menurut jurnal lainya yang berjudul "Analisis Sentimen Ulasan Restoran Menggunakan Pendekatan *Machine Learning*" yang ditulis oleh Harini Burra, Pallavi Mishra. Membahas tentang analisis sentimen ulasan restoran berbahasa Inggris menggunakan pendekatan machine learning. Tujuan penelitian adalah membandingkan efektivitas algoritma *Regresi Logistik* dan *Support Vector Machine* (SVM) dalam mengklasifikasi ulasan sebagai positif atau negatif. Data yang digunakan terdiri dari 1001 ulasan dari Kaggle, yang diklasifikasikan sebagai positif atau negatif. Metodologi yang diterapkan meliputi pembersihan data, pembagian dataset menjadi 80% pelatihan dan 20% pengujian, serta pelatihan model menggunakan kedua algoritma tersebut. Hasil evaluasi menunjukkan

bahwa *Regresi Logistik* mencapai akurasi 76,40%, sedangkan *Support Vector Machine* mencapai akurasi 76,80%. Kesimpulan dari penelitian ini adalah kedua algoritma menunjukkan performa yang cukup baik, dengan SVM sedikit lebih unggul.[5]

2. Kajian Pustaka dan pengembangan hipotesis

Berikut ini merupakan kajian pustaka dan landasan teori dalam penelitian ini :

2.1.Analisa Sentimen

Dalam ilmu data mining, analisis sentimen bertujuan untuk menganalisis dan mengekstrak data tekstual berupa opini, evaluasi, sikap, emosi, penilaian, dan sentimen seseorang terhadap suatu objek. Pada penelitian ini, proses pelabelan sentimen pada respon dilakukan dengan menghitung jumlah kata negatif, netral, dan positif.[6]

2.2. *Machine Learning*

Machine learning adalah bidang yang mencakup banyak bidang, termasuk ilmu komputer, statistik, ilmu kognitif, teknik, teori pengoptimalan, dan banyak disiplin matematika dan sains. Banyak aplikasi pembelajaran mesin, tetapi penambangan data adalah yang paling penting. [6]

2.3.*Stratified Balancing Data*

Metode ini untuk mengatasi ketidakseimbangan data meningkatkan keandalan model, pembagian data dilakukan secara *stratified balancing data* atau balancingdata secara manual, sehingga distribusi kelas sentimen tetap seimbang pada data training maupun testing. Dalam metode ini, pengambilan sampel acak dilakukan secara terstruktur dengan membagi anggota populasi menjadi beberapa sub kelompok yang disebut strata yang kemudian dari masing-masing strata diambil secara acak sampel.[7] Model kemudian dievaluasi untuk membandingkan performa ketiganya dalam mengklasifikasikan sentimen ulasan pengguna terhadap aplikasi *Livin' by Mandiri*.

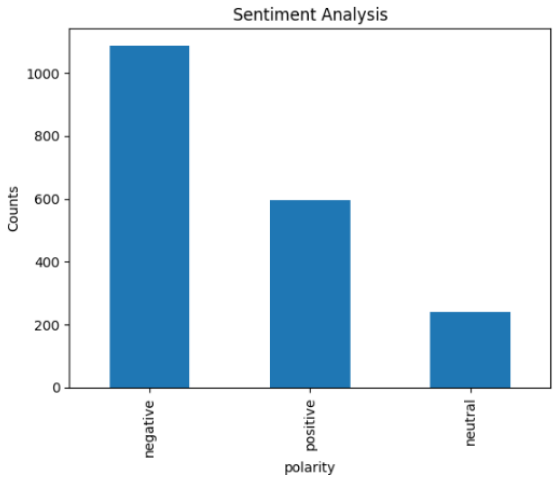
Proses ini bertujuan menyeimbangkan jumlah data dari setiap kelas rating (1 hingga 5) agar distribusinya merata untuk pelatihan model *machine learning*. Data difilter berdasarkan nilai *score*, lalu diambil secara acak masing-masing 385 sampel hasl perhitungan *slovin* dari setiap kelas. Hasil sampling dari lima kelas tersebut digabungkan ke dalam satu list *balanced_data*, data.menghasilkan total data sebanyak 1925 secara acak yang disimbolkan dengan id (5 kelas × 385 data).

2.4. *Preprocessing*

Pra-pemrosesan (*preprocessing*) merupakan langkah awal dalam pengolahan data masukan sebelum memasuki tahap utama. Tujuan dari preprocessing teks adalah untuk menyamakan format data agar lebih mudah dianalisis. Proses ini melibatkan beberapa tahap, antara lain: *case folding, tokenisasi, stopword removal*, dan *stemming.* [2]

2.5. *Labeling*

Data ulasan dilakukan pelabelan sentimen menjadi tiga kelas, yaitu positif, netral, dan negatif, berdasarkan nilai skor yang diberikan oleh pengguna. Hasil pelabelan ini digunakan sebagai target klasifikasi dalam pembangunan model machine learning. Berikut Gambar 1. Menunjukkan diagram hasil labeling yang menunjkkan performa negatif pada aplikasi lebih tinggi :



Gambar 1. *Labeling*

2.6. Pemobotan *TF-IDF*

Data hasil preprocessing berupa kata akan diubah menjadi bentuk numerik melalui proses pembobotan. Proses ini bertujuan untuk menghitung bobot setiap kata yang akan digunakan sebagai fitur. Semakin banyak dokumen yang diproses, semakin banyak pula fitur yang dihasilkan. Tahapan ini terdiri dari dua bagian utama, yaitu TF (*Term Frequency*) dan IDF (*Inverse Document Frequency*) [8].

Adapun rumus dari pembobotan kata TD-IDF adalah [8]:

$$W_{t,d} = tf_{t,d} \text{ x } idf_t = tf_{t,d} \text{ x } \log \frac{N}{\mathrm{d}ft} \quad (1)$$

Keterangan (1) :

$Wt_{,d}$ = Bobot TF-IDF

$tf_{t,d}$ = Jumlah frekuensi kata

idf_t = Jumlah inversefrekuensi dokumen tiap kata

df_t = Jumlah frekuensi dokumen tiap kata

N = Jumlah total dokumen

2.7.SVM

SVM (*Support Vector Machine*) adalah metode machine learning yang bekerja berdasarkan prinsip *Structural Risk Minimization* (SRM) dengan tujuan menemukan hyperplane terbaik yang memisahkan dua kelas dalam ruang input. [6]

2.8.*Naive Bayes*

Dengan menggunakan model fitur independen, metode pengklasifikasi *Naive Bayes* adalah prosedur klasifikasi probabilistik langsung yang didasarkan pada independensi yang kuat dari teorema *Bayes*, yang sering dikenal sebagai aturan Bayes.

Manfaat dari klasifikasi adalah bahwa estimasi parameter yang diperlukan untuk klasifikasi hanya membutuhkan sedikit data pelatihan.[9]

2.9.*Logistic Regression*

Dalam ilmu pengetahuan ilmiah dan sosial, *Logistic regression* adalah alat analisis yang penting. *Logistic regression* berfungsi sebagai dasar untuk algoritma *supervised machine learning* yang diawasi yang digunakan dalam klasifikasi dalam pemrosesan bahasa alami. Pengamatan dapat dibagi menjadi dua kelas atau lebih (seperti "sentimen positif" dan "sentimen negatif") dengan menggunakan *Logistic regression*.[10].

2.10. Confution Matrix

Confusion matrix digunakan sebagai dasar evaluasi utama dalam penelitian ini, dengan mengukur performa model melalui metrik akurasi, presisi, *recall*, dan *F1-score*. Evaluasi ini bertujuan untuk mengetahui kemampuan masing-masing algoritma dalam mengklasifikasikan sentimen ulasan pengguna secara tepat dan seimbang.[11]. Berikut rumus yang digunakan untuk menghitung metrik evaluasi adalah sebagai berikut:

$$\textbf{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\textbf{Presisi} = \frac{TP}{TP+FP} \quad (3)$$

$$\boldsymbol{Recall} = \frac{TP}{TP+FN} \quad (4)$$

$$\boldsymbol{F1\ score} = 12 \times \frac{Precision \times Recall}{Precission+Recall} \quad (5)$$

3. Metode Penelitian

Pada Penelitian ini penulis menggunakan jenis gabungan antara kuantitatif dan kualitatif. Menggunakan jenis kuatitatif karena menggunakan angka, statistik, dan algoritma untuk mengolah data dan menghasilkan hasil akurasi. Dan menggunakan Kualitatif karena fokus pada penafsiran menalam terhadap teks untuk menangkap aspek subjektif dari ulasan.

Proses pengumpulan data menggunakan ***Teknik Scraping*** yang dilakukan dengan menggunakan *Google Collab* dan bahasa pemrograman *Phyton.* Pengumpulan data ini bertujuan untuk mendapatkan representasi opini publik yang aktual dan autentik terhadap kinerja serta pengalaman pengguna terhadap aplikasi Livin' by Mandiri. Berikut gambar data hasi scrapping yang sudah di count menyisakan kolom "*content*", "*score*", "*id*"

	content	score	id
0	Semakin susah,tiba² keluar dan masuk lagi suru...	2	1
1	Susah log in, padahal udah update versi terbaru	1	2
2	Lemot banget	3	3
3	Abis update malah GK bisa transfer, gagal teru...	2	4
4	tidak bisa daftar, fitur verifikasi tidak jela...	1	5
...	...	...	...
148495	Good semakin bagus	5	148496
148496	Keren ... banyak fitur fitur baru... lebih pra...	5	148497
148497	Mantap cepat, gampang,praktis...	5	148498
148498	Keren. Cakep benar semakin canggih. Terdepan t...	5	148499
148499	Udah di coba, keren dan responsive, dengan tam...	5	148500

[148500 rows x 3 columns]

Gambar 2. Data Hasil *Scrapping*

4. Hasil dan Pembahasan

4.1. Hasil

Pada baian ini, fokus menampilkan hasil evaluasi dari ketiga algoritma *Mechine Learning* yang digunakan untuk analisis sentimen ulasan aplikasi *Livin by' Mandiri* menggunakan yang berupa akurasi, presisi, *recall*, *fi-score*.

Algoritma	Akurasi	Presisi	*Recall*	*F1-score*
SVM	82%	83%	82%	83%
LR	81%	81%	81%	81%
NB	62%	65%	62%	53%

Tabel 1. *Confution Matrix* Ketiga Algoritma

Berdasarkan data, terlihat bahwa SVM akurasi mencapai 82%, menunjukkan kinerja terbaik di antara ketiga pengklasifikasi. Secara khusus, metode ini mencapai Akurasi dan skor *F1-score* sebesar 83%, yang secara efektif mengidentifikasi makalah yang dihasilkan. *Logistic Regression* mengikuti dengan ketiga metrik 81%, menunjukkan akurasi yang relatif tinggi dalam mendeteksi teks. Terakhir, *Naive Bayes* (NB) mencapai Akurasi 62%, Presisi 65%, *Recall* 62% dan *F1-score* 53%, yang menunjukkan kemampuan terbatas dalam mengidentifikasi teks secara efektif dalam dataset ini. Di antara ketiga model tersebut, SVM memiliki kinerja terbaik, diikuti oleh *Logistic Regression*, dan terakhir *Naive Bayes*

4.2. Pembahasan

Untuk memberikan pemahaman yang lebih jelas mengenai data yang digunakan dalam penelitian ini, disajikan contoh data hasil akhir yang telah diproses. Berikut ini adalah cuplikan dari data hasil akhir yang telah melalui proses preprocessing dan pelabelan sentimen. Dari total 1925 data yang digunakan dalam penelitian, berikut Tabel 2. contoh dataset asli yang sudah terlabel :

no	content	score	id	polarity
1.	Gabisa verivikasi perlu verivikasi tambahan	2	103	positive
2.	Livin saya ndak bisa kebuka knapa ya	4	200	positive
3.	Apa aplikasi Livin ada gangguan yaaa Dari kemarin tidak bisa log in buka aplikasi	3	205	negative
4.	Ini saya masi gk bisa masuk ke livin mandiri dri kmren setelah verifikasi kode OTP no telp sistem error mulu	2	356	negative
5.	Tolong min aplikasi Livin Saya ngak bisa di buka	4	362	negative
6.	Setelah mengganti nomor telfon dan sudah verifikasi lewat bank . Kenapa sampai hari ini daftar ulang di aplikasih livin masih "fitur ini tidak tersedia ya".	4	393	negative
7.	Saya tidak bisa terverifikasi pada saat selvi sangat di sayangkan	4	505	negative
8.	Kenap ya kok verifikasi wajahnya gagal terus	1	592	neutral

9.	Ini gimana kok tambah buruk di install ulang lama bener Tlng di perbaiki lagi jangan bilang enak nya aja makan kuota aja	3	687	negative
...	.....	...	....	...
1925	Klo sebelumnya ga punya mbanking livin (hanya punya No rek dan kartu debit) apa bisa pake aplikasi new livin ?	2	148374	negative

Tabel 2. *Dataset* asli dan *Labeling*

Pada penelitian ini, dataset hasil *scrapping* sebanyak 148500 data, dan dihitung mengguanakan *Stratified balancing* data dan *Metode slovin* sebanyak 1925 data random ulasan pengguna Aplikasi *Livin by Mandiri* pada *Google Playstore.*

```
Rating 1 (sampled): 385 IDs, e.g., [53478, 53902, 75157, 75819, 124080]
Rating 2 (sampled): 385 IDs, e.g., [55525, 91492, 32468, 131102, 73549]
Rating 3 (sampled): 385 IDs, e.g., [45171, 143072, 55497, 93391, 17036]
Rating 4 (sampled): 385 IDs, e.g., [42239, 49944, 145620, 99200, 70509]
Rating 5 (sampled): 385 IDs, e.g., [127819, 134603, 65192, 97510, 84478]
Total balanced data: 1925 IDs
Index(['content', 'score', 'id'], dtype='object')
```

Gambar 2. Visualisasi Jumlah Data Acak Hasil Balanced Data

Hasil evaluasi performa dari ketiga algoritma machine learning yang digunakan, yaitu Multinomial *Logistic Regression*, *Support Vector Machine* (SVM), dan *Naive Bayes*. Evaluasi dilakukan dengan menggunakan confusion matrix, yang menggambarkan kemampuan model dalam mengklasifikasikan data ke dalam kategori sentimen negatif, netral, dan positif. Setelah melihat gambaran distribusi prediksi benar dan salah melalui *confusion matrix*, analisis dilanjutkan dengan evaluasi yang lebih mendalam menggunakan *classification report*. Berikut Tabel 3 *Confusion Matrix* dari ketiga algoritma tersebut :

Label	Precision	Recall	F1-score	Support
Negative	86%	77%	81%	104
Neutral	60%	68%	63%	68
Positive	89%	90%	89%	213
Accuracy	-	-	82%	385
Macro avg	78%	78%	78%	385
Weighted avg	83%	82%	83%	385

Tabel 4. Classification Report Logistic Regression

Label	Precision	Recall	F1-score	Support
Negative	86%	71%	78%	104
Neutral	57%	57%	57%	68
Positive	86%	93%	89%	213
Accuracy	-	-	81%	385
Macro avg	76%	74%	75%	385
Weighted avg	81%	81%	81%	385

Tabel 5. Classification Report Naive Bayes

Label	Precision	Recall	F1-score	Support
Negative	86%	77%	78%	104
Neutral	60%	60%	60%	68
Positive	89%	90%	89%	213
Accuracy	-	-	62%	385
Macro avg	78%	76%	77%	385
Weighted avg	83%	82%	83%	385

Berdasarkan hasil eksperimen dalam penelitian ini, di antara tiga model yang diuji *Logisitc Regression*, SVM, dan *Naive Bayes* (NB). Dan model SVM menunjukkan performa terbaik dalam mendeteksi teks yang hasil Scrapping.

Hasil *confusion matrix* dari tiga algoritma SVM, *Logistic Regression*, dan *Naive Bayes* menunjukkan bahwa ketiganya memiliki performa berbeda terhadap masing-masing kelas (negative, neutral, dan positive). SVM tampil unggul dalam mengenali kelas positive dengan recall tinggi (93%) serta presisi yang baik untuk kelas negative (86%), meskipun masih lemah pada kelas neutral. *Logistic Regression* menunjukkan performa yang hampir seimbang dengan SVM, terutama pada kelas positive, namun masih kesulitan di kelas neutral dengan precision dan recall hanya 60%. *Naive Bayes* juga cukup baik untuk kelas positive, tetapi memiliki kelemahan dalam membedakan kelas neutral dan sedikit lebih rendah akurasinya pada kelas negative dibanding dua algoritma lainnya. Secara umum, SVM dan *Logistic Regression* lebih unggul dalam mendeteksi sentimen kuat, sementara Naive Bayes cocok untuk pemrosesan cepat dengan kompleksitas rendah.

5. Kesimpulan dan Saran

5.1. Kesimpulan

Berdasarkan hasil penelitian, analisis sentimen terhadap ulasan pengguna aplikasi *Livin' by Mandiri* menggunakan *TF-IDF* dan pembagian data *stratified balancing* berbasis *Metode Slovin* menghasilkan model klasifikasi yang seimbang dan akurat yang berjumlah 1925 data dari data awal148.500. Dari tiga algoritma yang diuji, SVM menunjukkan performa terbaik dengan akurasi 82%, diikuti *Logistic Regression* (81%) dan *Naive Bayes* (62%). Hasil ini menunjukkan bahwa pemilihan algoritma berperan penting dalam efektivitas analisis sentimen, terutama untuk data teks yang tidak seimbang. Penelitian ini memberikan kontribusi dalam pengembangan sistem evaluasi otomatis yang bermanfaat bagi pengembang aplikasi dalam meningkatkan kualitas layanan.

5.2. Saran

Penelitian selanjutnya disarankan untuk menggunakan pendekatan representasi fitur yang lebih kompleks, guna meningkatkan akurasi model. Dan hasil analisis ini dapat dimanfaatkan oleh pengembang aplikasi sebagai dasar pengambilan keputusan dalam meningkatkan kualitas fitur dan layanan aplikasi *Livin' by Mandiri*, berdasarkan opini pengguna yang telah dianalisis secara otomatis dan sistematis.

Referensi

[1] A. A. Santoso and I. Rachmawati, "Analisis Minat Pengguna Layanan M-Banking Livin' by Mandiri di Indoneisa Menggunakan Model Modiffikasi UTAUT 2," *E-Proceeding Manag.*, vol. 8, no. 5, pp. 4316–4322, 2021.

[2] A. Pramono, R. Indriati, and A. Nugroho, "Sentiment Analysis Tokoh Politik Pada Twitter," *Semin. Nas. Inov. Teknol.* , pp. 195–200, 2017, [Online]. Available: https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/403/317

[3] G. Z. Dhamara, D. N. Alamsyah, P. W. Saputro, E. Daniati, and A. Ristyawan, "Analisis Sentimen Aplikasi Mybca Melalui Review Pengguna Di Google Play Store Menggunakan Algoritma Naive Bayes," *Agustus*, vol. 8, pp. 2549–7952, 2024.

[4] M. Zeng, "Research on AI-Generated Text Detection Based on Machine Learning Models," 2024.

[5] H. Burra and P. Mishra, "Restaurant Reviews Sentimental Analysis Using Machine Learning Approach," *Proc. 2024 Int. Conf. Emerg. Tech. Comput. Intell. ICETCI 2024*, pp. 414–417, 2024, doi: 10.1109/ICETCI62771.2024.10704184.

[6] F. S. Pamungkas and I. Kharisudin, "Analisis Sentimen dengan SVM, NAIVE BAYES dan KNNuntuk Studi Tanggapan Masyarakat Indonesia TerhadapPandemi Covid-19 pada Media Sosial Twitter," *Pros. Semin. Nas. Mat.*, vol. 4, pp. 1–7, 2021, [Online]. Available: https://journal.unnes.ac.id/sju/prisma/article/view/45038

[7] & F. Yogaswara, Y., Yusfida'ala, F., "Sistem Penimbangan Stratified Random Sampling Pada Pengangkutan Batubara PT. Adaro Indonesia," *Pros. TPT XXXII PERHAPI*, no. April, pp. 487–498, 2023.

[8] J. A. Septian, T. M. Fachrudin, and A. Nugroho, "Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor," *J. Intell. Syst. Comput.*, vol. 1, no. 1, pp. 43–49, 2019, doi: 10.52985/insyst.v1i1.36.

[9] F. Harahap, N. E. Saragih, E. T. Siregar, and H. Sariangsah, "Penerapan Data Mining Dengan Algoritma Naive Bayes Classifier Dalam Memprediksi Pembelian Cat," *J. Ilm. Inform.*, vol. 9, no. 01, pp. 19–23, May 2021, doi: 10.33884/JIF.V9I01.3702.

[10] N. L. P. C. Savitri, R. A. Rahman, R. Venyutzky, and N. A. Rakhmawati, "Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning," *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 1, pp. 47–58, 2021, doi: 10.28932/jutisi.v7i1.3216.

[11] I. H. Kusuma and N. Cahyono, "Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor," *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, pp. 302–307, 2023, doi: 10.30591/jpit.v8i3.5734.