
Analisis Sentimen Ulasan Penumpang Maskapai Penerbangan Indonesia Menggunakan Support Vector Machine , Naive Bayes, dan Random Forest

Daryanti¹, Tri Widodo²

daryantiyanti559@gmail.com¹, triwidodo@uty.ac.id²

^{1,2} Universitas Teknologi Yogyakarta

Informasi Artikel

Diterima : 19 Okt 2024

Direvisi : 26 Okt 2024

Disetujui : 30 Okt 2024

Kata Kunci

Analisis
Sentimen, SVM, Naive
Bayes, Random forest

Abstrak

Maskapai penerbangan merupakan perusahaan yang menyediakan layanan penerbangan untuk penumpang dan barang. Di Indonesia, terdapat dua jenis penerbangan, yaitu Low-Cost Carrier (LCC) dan full service. Penelitian ini menganalisis sentimen terhadap maskapai penerbangan Indonesia dengan menggunakan data ulasan dari X (Twitter) dan TripAdvisor, yang terdiri dari 6.469 ulasan. Tujuan utama penelitian adalah membandingkan kinerja tiga algoritma machine learning: Support Vector Machine (SVM) dengan kernel RBF, Naive Bayes, dan Random Forest. Model dievaluasi berdasarkan metrik akurasi, presisi, recall, dan F1-Score. Hasil menunjukkan bahwa Random Forest memberikan kinerja terbaik dengan akurasi 91%, presisi 91%, recall 91%, dan F1-Score 90%. Sementara itu, SVM dengan kernel RBF mencapai akurasi 89%, dan Naive Bayes memperoleh akurasi 79%. Dengan demikian, Random Forest terbukti lebih efektif dalam analisis sentimen pengguna maskapai penerbangan di Indonesia.

Keywords

Sentiment
Analysis, SVM, Naive
Bayes, Random Forest

Abstract

An airline is a company that provides flight services for passengers and goods. In Indonesia, there are two types of airlines, namely Low-Cost Carrier (LCC) and full service. This research analyzes sentiment towards Indonesian airlines using review data from X (Twitter) and TripAdvisor, which consists of 6,469 reviews. The main objective of the research is to compare the performance of three machine learning algorithms: Support Vector Machine (SVM) with RBF kernel, Naive Bayes, and Random Forest. The models were evaluated based on accuracy, precision, recall, and F1-Score metrics. The results show that Random Forest provides the best performance with 91% accuracy, 91% precision, 91% recall, and 90% F1-Score. Meanwhile, SVM with RBF kernel achieved 89% accuracy, and Naive Bayes obtained 79% accuracy. Thus, Random Forest proved to be more effective in analyzing the sentiment of airline users in Indonesia.

A. Pendahuluan

Maskapai penerbangan merupakan perusahaan yang menyediakan layanan penerbangan untuk penumpang dan barang dengan menggunakan armada pesawat serta tim yang terampil di berbagai bidang terkait penerbangan. Di Indonesia memiliki dua jenis penerbangan, yakni *Low Cost Carrier (LCC)* dan *full service* [1]. Beberapa maskapai yang paling relevan dan sering digunakan untuk penerbangan domestik meliputi Garuda Indonesia, Lion Air, Batik Air, dan AirAsia. Masing-masing maskapai ini menawarkan berbagai fasilitas dan layanan yang berbeda-beda, yang kemudian menghasilkan beragam ulasan dari penumpang. Layanan yang ditawarkan oleh setiap maskapai ini mencakup aspek kenyamanan, keamanan, ketepatan waktu, hingga pelayanan dari awak kabin, yang semuanya berkontribusi pada pengalaman keseluruhan penumpang selama perjalanan[2]. Salah satu cara paling mudah yang dapat dilakukan perusahaan untuk mengevaluasi kinerja layanan mereka adalah dengan memperhatikan ulasan-ulasan yang diberikan oleh konsumen[3]. Menurut survei CNN Indonesia, maskapai lokal terbaik adalah Garuda Indonesia yang dipilih oleh responden, diikuti oleh Citilink di posisi kedua, Lion Air di posisi ketiga, Batik Air di posisi keempat, AirAsia di posisi kelima, Sriwijaya Air di posisi keenam, dan Super Jet Air di posisi ketujuh. Survei ini dilakukan secara daring dengan menggunakan metode *multiple answers*. Namun, untuk hasil yang lebih akurat dan mendalam mengenai persepsi penumpang, diperlukan analisis sentimen yang dapat mengungkap nuansa opini publik dari berbagai platform digital[4].

Analisis sentimen merupakan bagian dari text mining yang bertujuan mendeteksi kecenderungan seseorang melalui tulisan atau teks atau mengekstrak data teks untuk mendapatkan informasi tentang sentimen yang biasanya dituliskan di dalam media sosial baik bermakna positif, negatif atau netral. Proses analisis ini biasanya dilakukan secara otomatis melalui beberapa sistem atau platform yang mendukung untuk dilakukannya analisis teks[5]. Analisis sentimen ulasan pengguna maskapai penerbangan membantu dalam mengoptimalkan kualitas layanan dan strategi pemasaran. Dengan menggunakan algoritma analisis sentimen, maskapai dapat mengidentifikasi pola dan tren dalam ulasan yang dapat membantu dalam memperbaiki aspek-aspek yang tidak memuaskan pelanggan. Analisis sentimen juga dapat membantu maskapai dalam mengembangkan strategi pemasaran yang lebih efektif, membangun reputasi merek yang positif, serta menciptakan pengalaman penerbangan yang lebih memuaskan bagi pelanggan.

Beberapa penelitian terkait analisis sentimen terhadap maskapai penerbangan telah dilakukan sebelumnya. Penelitian yang dilakukan oleh Husada & Paramita mengkaji analisis sentimen terhadap maskapai penerbangan di platform *Twitter* menggunakan algoritma *Support Vector Machine (SVM)* dan metode ekstraksi fitur *TF-IDF*. Beberapa *kernel* yang digunakan pada SVM meliputi *Linear*, *Radial Basis Function (RBF)*, *Sigmoid*, dan *Polynomial*. Hasil penelitian ini menunjukkan bahwa *SVM kernel RBF* mencapai akurasi sebesar 84,37%. Penelitian yang dilakukan oleh Rahmawati, dkk lebih difokuskan pada analisis sentimen terhadap ulasan pengguna mengenai penerbangan Lion Air di platform daring. Mereka membandingkan beberapa algoritma, termasuk *Naïve Bayes*, *Random Forest*, dan *Logistic Regression*, dengan ekstraksi fitur berbasis metode *Lexicon Based*. Hasil penelitian ini menunjukkan bahwa *Logistic Regression* mencapai

akurasi 82%, sementara itu, *Naïve Bayes* memperoleh akurasi 47% dan *Random Forest* 39% [6]. Selanjutnya, pada penelitian yang dilakukan oleh Handoko, dkk juga melakukan analisis sentimen terhadap ulasan penumpang maskapai penerbangan di Indonesia. Mereka menggunakan algoritma *Random Forest* dan *K-Nearest Neighbor (KNN)* dengan ekstraksi fitur menggunakan metode *TF-IDF*. Hasil penelitian menunjukkan bahwa *Random Forest* mencapai akurasi sebesar 83%, sedangkan KNN mencapai 82% [2].

Penelitian ini menanggapi keterbatasan penelitian-penelitian sebelumnya yang hanya menggunakan satu platform serta menunjukkan akurasi rendah pada algoritma Naive Bayes dan Random Forest. Dengan menganalisis opini masyarakat di media sosial, khususnya di platform X dan TripAdvisor, penelitian ini memperluas cakupan data agar lebih luas dan representatif. Opini masyarakat diklasifikasikan ke dalam tiga kategori: positif, negatif, dan netral. Algoritma yang digunakan dalam penelitian ini meliputi Support Vector Machine dengan kernel RBF, Naive Bayes, dan Random Forest. Penggunaan ketiga algoritma ini memungkinkan perbandingan akurasi yang lebih baik pada data yang kompleks. SVM efektif dalam menemukan hyperplane optimal yang memisahkan data non-linear, Naive Bayes memberikan model klasifikasi yang sederhana dan jelas [7], sementara Random Forest mengurangi risiko overfitting dengan melakukan perhitungan pada banyak pohon keputusan secara acak [8]. Dengan pendekatan ini, diharapkan informasi dari komentar masyarakat dapat membantu calon penumpang dalam memilih layanan maskapai penerbangan yang sesuai dan memberikan analisis sentimen yang akurat untuk membantu maskapai memahami opini pelanggan dengan lebih baik.

B. Metode Penelitian

Penelitian ini menggunakan desain kuantitatif dengan metode eksperimen untuk menganalisis sentimen ulasan penumpang maskapai Indonesia. Penelitian kuantitatif merupakan penelitian empiris yang menggunakan data berupa angka-angka. Metode kuantitatif meliputi penelitian eksperimen dan survei. Metode eksperimen bertujuan untuk menjelaskan hubungan sebab-akibat antara satu variabel dengan variabel lainnya [9]. Data penelitian berupa ulasan penumpang terhadap maskapai penerbangan di Indonesia yang dikumpulkan dari platform media sosial X (*Twitter*) dan *TripAdvisor*. Data yang terkumpul kemudian diberikan label dan diproses melalui tahap pra-pemrosesan, di mana data dibersihkan untuk menghilangkan atribut yang tidak relevan, duplikasi, atau informasi yang tidak diperlukan. Setelah itu, teks diubah ke dalam format yang dapat dianalisis dengan menghapus tanda baca dan angka, mengubah semua huruf menjadi huruf kecil, serta menghilangkan kata-kata umum (*stop words*) yang tidak memberikan informasi penting. Selain itu, dilakukan *stemming* untuk menyederhanakan bentuk kata. Tahap selanjutnya adalah pembobotan kata menggunakan metode *TF-IDF* (*Term Frequency-Inverse Document Frequency*). Setelah pembobotan, dilakukan implementasi algoritma klasifikasi, yaitu *SVM*, *Naive Bayes*, dan *Random Forest*. Untuk mengevaluasi kinerja masing-masing algoritma, dilakukan tahap pengujian data setelah klasifikasi. Alur penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

1. Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari platform media sosial X, yang dulu dikenal sebagai *Twitter*, dan *TripAdvisor*, yang terdiri dari ulasan penumpang tentang pengalaman di berbagai maskapai yang ada di Indonesia. Pengumpulan data dilakukan dengan metode *crawling data* menggunakan *API* dari kedua platform dan bahasa pemrograman Python. Berikut data yang digunakan dalam analisis sentimen dapat dilihat pada

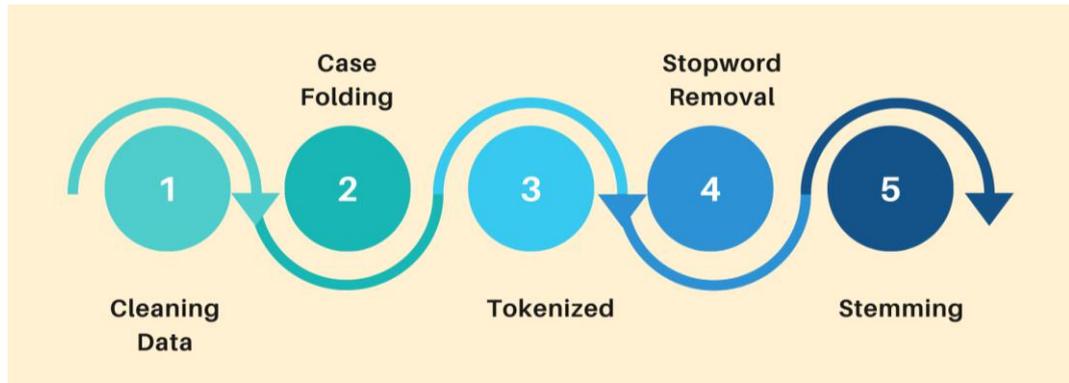
	Response
0	Air asia sangat bagus pelayanan terbaik harga ...
1	petugas check in counter di Yogyakarta atas na...
2	GOT NO HEART!!!! Not recommended!!!! \nDisclai...
3	mw check in..40 menit sebelum jam berangkat pe...
4	Kalau anda punya pilihan lain, jangan pernah b...
...	...
6464	EMANG JIR SUMPAH SUPER AIR JET INI GA OKE BGT ...
6465	@SuperAirJet Wkwkwkwk super air jet emang anji...
6466	sampe detik ini gue gaada sama sekali spill na...
6467	Super Air Jet cuman menang pramugarinya agak c...
6468	@pluvioophiles Perasaan direct flight dari/ke ...

Gambar 2. Dataset Analisis Sentimen

2. Pelabelan Data

Data yang sudah dikumpulkan, selanjutnya diberikan label sentimen berupa label positif, negatif, dan netral dengan menggunakan *TextBlob* [10].

3. Pemrosesan Data



Gambar 3. Tahapan pemrosesan data

Pemrosesan data dalam analisis ini dimulai dengan pembersihan data untuk menghapus komentar yang tidak relevan, duplikat, dan informasi yang tidak berguna, diikuti dengan penghapusan tanda baca dan angka, konversi teks menjadi huruf kecil, serta penghapusan *stop words* yang tidak memiliki nilai informasi penting. Selain itu, dilakukan *stemming* untuk menyederhanakan kata. Tahapan preprocessing mencakup proses cleaning (menghapus karakter non-alfabet) [5], *case folding* (mengubah huruf menjadi kecil) [11], *tokenizing* (memecah kalimat menjadi kata) [9], *stopword removal* (menghapus kata yang tidak bermakna) [12], dan *stemming* (menghapus imbuhan untuk menemukan kata dasar) [13].

4. Pembobotan Kata

Tahapan pembobotan kata menggunakan metode *TF-IDF*. Metode *TF-IDF* dikenal karena efisiensinya, kemudahan penggunaannya, dan hasil yang akurat. Pendekatan ini mengukur *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF) untuk setiap kata di seluruh dokumen, membantu dalam menentukan frekuensi kemunculan suatu kata dalam dokumen. Hasil dari metode ini adalah matriks fitur *TF-IDF* yang menggambarkan representasi numerik dari kata-kata dalam dokumen [14].

5. Klasifikasi dan Evaluasi Model

Setelah data siap diproses, dilakukan pemisahan data latih dan data uji untuk mengukur kinerja dari ketiga model klasifikasi, yaitu *SVM*, *Naive Bayes*, dan *Random Forest*. Pada evaluasi model, kinerja akan diukur menggunakan metrik seperti akurasi, presisi, *recall*, dan *f1-score* [15]. Evaluasi model dilakukan percobaan dengan pembagian data yang berbeda.

C. Hasil dan Pembahasan

Data yang diperlukan untuk penelitian ini berasal dari *crawling* data *TripAdvisor* dan X (*Twitter*) dengan kata kunci maskapai di Indonesia, antara lain Garuda Indonesia, Citilink, AirAsia Indonesia, Sriwijaya Air, Batik Air, Lion Air, dan Super Jet Air. Ada sebanyak 6.469 data yang digunakan dalam penelitian ini. Sebelum digunakan untuk analisis sentimen, data tersebut dipetakan sentimennya menjadi positif, negatif, dan netral menggunakan *library Textbox*. Berikut hasil pelabelan data dapat dilihat pada Gambar 4 dan 5.

	Response	Label
0	Air asia sangat bagus pelajanan terbaik harga ...	Netral
1	petugas check in counter di Yogyakarta atas na...	Positif
2	GOT NO HEART!!!! Not recommended!!!! \nDisclai...	Netral
3	mw check in..40 menit sebelum jam berangkat pe...	Netral
4	Kalau anda punya pilihan lain, jangan pernah b...	Netral
...
6464	EMANG JIR SUMPAAH SUPER AIR JET INI GA OKE BGT ...	Positif
6465	@SuperAirJet Wkwkwkwk super air jet emang anji...	Positif
6466	sampe detik ini gue gaada sama sekali spill na...	Positif
6467	Super Air Jet cuman menang pramugarinya agak c...	Positif
6468	@pluvioophiles Perasaan direct flight dari/ke ...	Positif

Gambar 4. Dataset setelah pelabelan data pada python

response	label
Mau nonton MotoGP di Mandalika? Beli tiket travelingnya pakai Mandiri Kartu Kredit aja! Kamu bisa dapatkan diskon s.d. Rp500 ribu* untuk transaksi di aplikasi FlyGaruda atau website Garuda Indonesia serta aplikasi https://t.co/uudu1BOA1n menggunakan Mandiri Kartu Kredit! https://t.co/LXD6SDfPp	Netral
@fajamugros Bener banget! Garuda Indonesia bikin bangga!	Netral
Dapatkan beragam penawaran menarik GarudaMiles di Hong Kong! Bergabunglah dengan program Diaspora Indonesia Hong Kong dari GarudaMiles dan nikmati beragam keuntungan eksklusif hanya untuk Anda! Info lengkap: https://t.co/KYZouaAbki #GarudaIndonesia #BecauseYouMatter https://t.co/CWZqrd0B	Netral
@AndiSimulingga Sejak kapan Garuda Indonesia menjadi lambang negara Du?!	Netral
Maskapai penerbangan nasional Garuda Indonesia konsisten memperkuat transformasi Perusahaan setelah melewati fase restrukturisasi. Fase transformasi tersebut turut ditandai dengan diumumkannya rekognisi dunia pada pertengahan September 2024 lalu. Garuda Indonesia masuk ke dalam https://t.co/c9CVAU24	Netral
Siap menyaksikan pembalap MotoGP favorit Anda di Sirkuit Mandalika? Terbang bersama Garuda Indonesia dan dapatkan potongan harga special hingga 20%! Info lengkap: https://t.co/dhNSCGU8y #GarudaIndonesia #BecauseYouMatter #Mandalika #MotoGP https://t.co/SE3v2wQND0	Positif
@yudhiiprakoza Seluruh penerbangan Garuda Indonesia dari dan menuju Malang dihentikan untuk sementara mulai tgl 23 September 2024. Mohon berkenan menghubungi contact center kami di tautan berikut: https://t.co/Z8rptcLY0u guna mengatur ulang jadwal penerbangannya. Tks. (2/2) - Nadine	Negatif
Nikmati kenyamanan dan keindahan penerbangan dari Jakarta ke Pekanbaru bersama Garuda Indonesia. Pesan tiketnya segera di aplikasi FlyGaruda atau https://t.co/uhNniEFck #GarudaIndonesia #BecauseYouMatter #Pekanbaru https://t.co/QjSIsAdtG	Netral

Gambar 5. Dataset setelah pelabelan data pada sistem

Tahap selanjutnya setelah pelabelan data ialah *preprocessing* data. Dalam proses *preprocessing* ada *case folding*, *cleaning data*, *tokenization*, *stopword removal*, dan *stemming*. Berikut hasil *preprocessing* dapat dilihat pada Gambar 6 dan 7.

	Response	Label	preprocessing_result
0	Delapan dari sepuluh orang Indonesia pasti bar...	Netral	delapan puluh orang indonesia tau nang bikin t...
1	@athwete @indosatim3 @airasia @traveloka @goje...	Netral	hai kak terima kasih tarik bekerjasama bni ya ...
2	Klik https://t.co/zxVRMiX5vq untuk info lebih ...	Netral	klik https co zxvrmix5vq info airasiainforms s...
3	Terbang saat hamil? No worries! Cek tipsnya yu...	Netral	terbang hamil no worries cek tipsnya yuk biar ...
4	Naik pesawat dari Hong Kong? Bisaaaaa..... In...	Netral	pesawat hong kong bisaaaaa indonesia airasia b...
...
658	Sebagai Allstars kami selalu berusaha memperse...	Netral	allstars usaha sembah baik 14 menang harga mas...
659	Min kalau beli tiket maskapai lain di airasia...	Netral	min beli tiket maskapai airasia superapp amp r...
660	Min kalau pembelian tiket dari OTA atau agen ...	Netral	min beli tiket ota agen travel gimana aju refu...
661	Dan jangan khawatir setelah kamu selesai menga...	Netral	khawatir selesai aju refund https co 3ilbglug5...
662	Jadi kalau penerbangan kamu terdampak perubaha...	Netral	terbang dampak ubah jadwal mengcancel terbang ...

Gambar 6. Hasil preprocessing pada python

Initial Process			
Response	Preprocessing	TF-IDF	Label
--Original Text--	--Processed Text--	--TF-IDF Weighting--	-- Label--
Mau nonton MotoGP di Mandalika Beli tiket travelingnya pakai Mandiri Kartu Kredit aja 65039 Kamu bisa dapatkan diskon sd Rp500 ribu untuk transaksi di aplikasi FlyGaruda atau website Garuda Indonesia serta aplikasi httpstcooudul80AIn menggunakan Mandiri Kartu Kredit httpstcoXDbSSDFp	mau nonton motogp mandalika beli tiket travelingnya pakai mandiri kartu kredit dapatkan diskon rp500 ribu transaksi aplikasi flygaruda website garuda indonesia serta aplikasi menggunakan mandiri kartu kredit	mau: 0.111 nonton: 0.111 motogp: 0.111 mandalika: 0.111 beli: 0.111 tiket: 0.111 travelingnya: 0.111 pakai: 0.111 mandiri: 0.2219 kartu: 0.2219 kredit: 0.2219 dapatkan: 0.111 diskon: 0.111 rp500: 0.111 ribu: 0.111 transaksi: 0.111 aplikasi: 0.2219 flygaruda: 0.111 website: 0.111 garuda: 0.111	Netral

Gambar 7. Hasil preprocessing pada sistem

Berikut ini peringkat maskapai terbaik berdasarkan ulasan positif, dapat dilihat pada gambar.

Peringkat	Maskapai	Total Ulasan	Ulasan Positif	Ulasan Netral	Ulasan Negatif
0	1 Super Jet Air	538	518	10	10
1	2 AirAsia-Indonesia	1452	282	1122	48
2	3 Citilink	1113	184	885	44
3	4 Lion Air	1317	181	1082	54
4	5 Garuda Indonesia	721	107	595	19
5	6 Sriwijaya Air	801	102	676	23
6	7 Batik Air	527	82	427	18

Gambar 8. Peringkat maskapai berdasarkan ulasan positif

Berdasarkan gambar peringkat maskapai berdasarkan ulasan positif, diperoleh bahwa maskapai terbaik adalah Super Jet Air dengan ulasan positif

518 dari total ulasan 538. Berikut Tabel 1 menunjukkan perbandingan hasil akurasi dari ketiga algoritma berdasarkan pembagian data.

Berikut tabel 1 menunjukkan perbandingan hasil akurasi dari ketiga algoritma berdasarkan pembagian data.

Tabel 1. Tabel Perbandingan Akurasi SVM, Naive Bayes dan Random Forest

No	Metode	50:50	60:40	70:30	80:20	90:10
1	SVM Kernel RBF	87%	87%	88%	88%	89%
2	Naive Bayes	77%	77%	78%	79%	79%
3	Random Forest	88%	88%	89%	90%	91%

Berdasarkan tabel perbandingan akurasi dari ketiga algoritma di peroleh pembagian data terbaik pada 90:10. Berikut ini Evaluasi Performa dari ketiga algoritma pada pembagian data 90:10.

Tabel 2. Tabel Hasil Evaluasi

No	Metode	Presisi	Recall	F1-Score	Akurasi
1	SVM Kernel RBF	90%	89%	87%	89%
2	Naive Bayes	79%	79%	72%	79%
3	Random Forest	91%	91%	90%	91%

D. Simpulan

Berdasarkan hasil *Analisis Sentimen Ulasan Penumpang Maskapai Penerbangan Indonesia Menggunakan Support Vector Machine, Naive Bayes, dan Random Forest*, dapat disimpulkan bahwa dari pengujian ketiga algoritma tersebut, metode *Random Forest* memiliki akurasi tertinggi, yaitu 91%, dengan proporsi data *testing* sebesar 10% dan data *training* sebesar 90%. Maskapai terbaik yang diperoleh adalah Super Jet Air, yang memiliki tingkat kepuasan tinggi dengan dominasi ulasan positif. Kompleksitas dataset dan proses pelabelan data berpengaruh besar pada evaluasi model, terutama pada metrik presisi, *recall*, *f1-score*, dan akurasi.

E. Referensi

- [1] R. A. Sari, "Analisis Pengaruh Pelayanan dan Produk Terhadap Pengambilan Keputusan Pemakaian Jasa Maskapai PT.Garuda Indonesia di Bandara Adi Soemarmo Boyolali," *J. Multidisiplin Madani*, vol. 2, no. 4, pp. 1833–1844, 2022, [Online]. Available: <https://journal.y3a.org/index.php/mudima/index>
- [2] Handoko, D. Ramadhansyah, A. Asrofiq, Rahmaddeni, and Y. Yunefri, "Analisis Sentimen Ulasan Penumpang Maskapai Penerbangan di Indonesia Dengan Algoritma Random Forest dan KNN," *Zo. J. Sist. Inf.*, vol. 6, no. 2, pp. 287–297, May 2024, doi: 10.31849/zn.v6i2.19177.
- [3] L. M. Khaerunnizar and Y. A. Nugroho, "Analisis Kualitas Layanan Maskapai Penerbangan Xyz Berdasarkan Sentimen Ulasan Pelanggan," *J. Cakrawala Ilm.*, vol. 1, no. 10, pp. 2493–2504, 2022, doi: 10.53625/jcijurnalcakrawalailmiah.v1i10.2580.
- [4] Tim, "Daftar Maskapai Favorit Orang Indonesia: Garuda Indonesia Teratas," CNN Indonesia. Accessed: Oct. 05, 2024. [Online]. Available:

- <https://www.cnnindonesia.com/gaya-hidup/20230113094846-269-899899/daftar-maskapai-favorit-orang-indonesia-garuda-indonesia-teratas>
- [5] Herwinsyah and A. Witanti, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm)," *J. Sist. Inf. dan Inform.*, vol. 5, no. 1, pp. 59–67, 2022, doi: 10.47080/simika.v5i1.1411.
- [6] I. Rahmawati, T. Rika Fitriani, A. No'eman, and A. Y. P. Yusuf, "Analisis Sentimen Menggunakan Algoritma Logistic Regression Pada Penerbangan Lion Air berdasarkan Ulasan Platform Online," *J. Ris. Inform. dan Teknol. Inf.*, vol. 1, no. 1, pp. 11–16, 2023, doi: 10.58776/jriti.v1i1.60.
- [7] F. Zamachsari, G. V. Saragih, Susafa'ati, and W. Gata, "Analysis of Sentiment of Moving a National Capital with Feature Selection Naive Bayes Algorithm and Support Vector Machine," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 3 SE-Information Technology Articles, Jun. 2020, doi: 10.29207/resti.v4i3.1942.
- [8] L. Barreñada, P. Dhiman, D. Timmerman, A.-L. Boulesteix, and B. Van Calster, "Understanding random forests and overfitting: a visualization and simulation study," vol. 2001, no. 1, pp. 1–60, 2024, [Online]. Available: <http://arxiv.org/abs/2402.18612>
- [9] H. Syahrizal and M. S. Jailani, "Jenis-Jenis Penelitian Dalam Penelitian Kuantitatif dan Kualitatif," *QOSIM J. Pendidikan, Sos. Hum.*, vol. 1, no. 1 SE-Articles, pp. 13–23, May 2023, doi: 10.61104/jq.v1i1.49.
- [10] R. W. Hardian, P. E. Prasetyo, U. Khaira, and T. Suratno, "Analisis Sentiment Kuliah Daring Di Media Sosial Twitter Selama Pandemi Covid-19 Menggunakan Algoritma Sentistrength," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 2, pp. 138–143, 2021, doi: 10.57152/malcom.v1i2.15.
- [11] Yuyun, Nurul Hidayah, and Supriadi Sahibu, "Algoritma Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Pemerintah Terhadap Penanganan Covid-19 Menggunakan Data Twitter," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 820–826, 2021, doi: 10.29207/resti.v5i4.3146.
- [12] E. Yuniar, D. S. Utsalinah, and D. Wahyuningsih, "Implementasi Scrapping Data Untuk Sentiment Analysis Pengguna Dompot Digital dengan Menggunakan Algoritma Machine Learning," *J. Janitra Inform. dan Sist. Inf.*, vol. 2, no. 1, pp. 35–42, 2022, doi: 10.25008/janitra.v2i1.145.
- [13] D. Nugraha and D. Gustian, "Analisis Sentimen Penggunaan Aplikasi Transportasi Online Pada Ulasan Google Play Store dengan Metode Naive Bayes Classifier," *J. Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 5, no. 1, pp. 326–335, 2024.
- [14] F. Baehaqi and N. Cahyono, "Analisis Sentimen Terhadap Cyberbullying Pada Komentar Di Instagram Menggunakan Algoritma Naïve Bayes," *Indones. J. Comput. Sci.*, vol. 13, no. 1, pp. 1051–1063, 2024, doi: 10.33022/ijcs.v13i1.3301.
- [15] D. Atmajaya, A. Febrianti, and H. Darwis, "Metode SVM dan Naive Bayes untuk Analisis Sentimen ChatGPT di Twitter," *Indones. J. Comput. Sci.*, vol. 12, no. 4, pp. 2173–2181, 2023, doi: 10.33022/ijcs.v12i4.3341.