



Analisis Sentimen Ulasan Wisatawan Terhadap Alun-Alun Kota Batam: Perbandingan Kinerja Metode *Naive Bayes* dan *Support Vector Machine*

John Friadi^{a*}, Dwi Ely Kurniawan^b

^aDepartment Information System, Universitas Batam

^bDepartment of Informatic Engineering, Politeknik Negeri Batam

Naskah masuk: 14 Desember 2023; Diterima untuk publikasi: 8 Agustus 2024
DOI: 10.21456/vol14iss4pp403-407

Abstract

Batam City, as a rapidly developing tourism destination in Indonesia, continues to strive to enhance the potential of its tourist attractions to attract more visitors. The assessment of reviews from tourists is crucial in identifying necessary development measures to improve the quality of tourist attractions. This research aims to analyze the sentiment of reviews for the Alun-Alun Kota Batam tourist destination by leveraging data from Google Maps. Two classification methods, Naive Bayes and Support Vector Machine, are employed for sentiment analysis, and their performances are compared. From 1140 collected reviews, the data is categorized into three labels: positive, negative, and neutral. The research results indicate that the Support Vector Machine method achieves higher accuracy (94%) compared to Naive Bayes (83%). This study contributes insights into visitor sentiments towards Alun-Alun Kota Batam, with implications for policy development and more effective actions in enhancing local tourism appeal.

Keywords: Sentiment Analysis; Tourist Reviews; Alun-Alun Kota Batam; Naive Bayes; Support Vector Machine.

Abstrak

Kota Batam, sebagai destinasi pariwisata yang berkembang pesat di Indonesia, terus berupaya meningkatkan potensi objek wisatanya untuk menarik lebih banyak pengunjung. Evaluasi ulasan dari wisatawan menjadi kunci dalam mengidentifikasi langkah-langkah pengembangan yang dibutuhkan untuk meningkatkan kualitas objek wisata. Penelitian ini bertujuan untuk menganalisis sentimen ulasan objek wisata Alun-Alun Kota Batam dengan memanfaatkan data dari *Google Maps*. Dua metode klasifikasi, yaitu *Naive Bayes* dan *Support Vector Machine*, digunakan untuk analisis sentimen, dan kinerja keduanya dibandingkan. Dari 1140 ulasan yang dikumpulkan, data dibagi menjadi tiga label: positif, negatif, dan netral. Hasil penelitian menunjukkan bahwa metode *Support Vector Machine* memberikan akurasi lebih tinggi (94%) dibandingkan dengan *Naive Bayes* (83%). Penelitian ini memberikan kontribusi dalam memberikan wawasan tentang sentimen pengunjung terhadap Alun-Alun Kota Batam, dengan implikasi untuk pengembangan kebijakan dan tindakan yang lebih efektif dalam meningkatkan daya tarik wisata lokal.

Kata kunci : Analisis Sentimen; Ulasan Wisatawan; Alun-Alun Kota Batam; Naive Bayes; Support Vector Machine.

1. Pendahuluan

Pariwisata telah menjadi sektor unggulan Indonesia dalam meningkatkan pendapatan devisa negara. Pentingnya meningkatkan dan mengembangkan daya tarik pariwisata menjadi kunci untuk menarik lebih banyak wisatawan, baik dari dalam maupun luar negeri. Tingkat kenyamanan suatu objek wisata menjadi faktor kunci dalam mempercepat pertumbuhan industri pariwisata nasional (thaha and Aziz, 2020). Evaluasi tingkat kenyamanan dapat dilihat melalui ulasan yang diberikan oleh para wisatawan, yang kemudian menjadi umpan balik penting bagi pemangku kepentingan untuk pengembangan kualitas objek wisata.

*) *Corresponding author:* john.friadi@gmail.com

Dalam era teknologi informasi dan internet, akses cepat dan mudah terhadap informasi dan data ulasan objek wisata menjadi lebih memungkinkan. Kota Batam, sebagai salah satu kota di Indonesia yang terus mengembangkan potensi objek wisatanya, memiliki tantangan terkait kurangnya promosi dan keterbukaan informasi terkait kualitas objek wisata di Batam. Berkembangnya teknologi informasi, seperti peta daring dan *Google Maps*, memungkinkan wisatawan untuk merencanakan rute dan memberikan ulasan serta rating pada tempat yang mereka kunjungi (Rifa'i, 2021).

Dari pengamatan yang dilakukan, banyak pengunjung objek wisata Batam, termasuk Alun-Alun Engku Putri, memberikan ulasan di *Google Maps*. Alun-Alun Engku Putri di Batam memiliki peran penting sebagai ikon kota dengan berbagai aktivitas

yang dapat dilakukan, seperti berkumpul dengan komunitas, menikmati jajanan khas, berolahraga, dan berlibur bersama keluarga. Ulasan wisatawan di *Google Maps* menjadi sumber informasi berharga untuk meningkatkan kualitas pariwisata di Batam (thaha and Aziz, 2020).

Dalam konteks ini, penelitian ini akan memanfaatkan metode *text mining* untuk menganalisis sentimen ulasan wisatawan terhadap Alun-Alun Engku Putri di Batam. *Text mining* merupakan proses eksplorasi dan analisis data teks tidak terstruktur untuk mengidentifikasi pola, informasi, atau pengetahuan yang berguna (thaha and Aziz, 2020). Analisis sentimen terhadap ulasan wisatawan dapat memberikan wawasan tentang kelebihan dan kekurangan objek wisata, serta memberikan pemahaman mengenai sentimen pengunjung. Dalam hal ini, algoritma klasifikasi seperti *naive bayes* akan diterapkan untuk meningkatkan akurasi analisis sentimen (Somantri et al., 2023).

Kontribusi utama dari penelitian ini adalah memberikan informasi yang cepat dan akurat melalui analisis sentimen ulasan wisatawan. Hal ini dapat membantu pemangku kepentingan, terutama pengelola Alun-Alun Engku Putri di Batam, untuk mengambil tindakan yang tepat dalam meningkatkan kualitas objek wisata. Dengan memahami sentimen pengunjung, stakeholder dapat merencanakan langkah-langkah strategis untuk memperbaiki dan meningkatkan daya tarik pariwisata di kota Batam.

2. Kerangka Teori

2.1. Analisis Sentimen

Analisis sentimen adalah bidang ilmu yang memanfaatkan kecerdasan buatan untuk dapat memberikan dukungan keputusan dalam menilai sentimen dengan mengkategorikan apakah sentimen itu positif, negatif, atau netral (Nasreen, 2021). Pengukuran analisis sentimen diperlukan sebuah metode, diantaranya adalah metoda algoritma klasifikasi *naive bayes* dan *metode support vector machine*. Analisis sentimen merupakan studi komputasi yang diambil dari opini, sentimen, dan emosi masyarakat melalui entitas dan atribut yang dimilikinya yang direpresentasikan dalam bentuk teks (Jumanto et al., 2023). Analisis sentimen akan mengklasifikasikan polaritas teks dalam sebuah kalimat atau dokumen untuk mengetahui apakah salah satu upaya untuk meningkatkan tingkat akurasi yang dihasilkan model review sentimen analisis.

2.2. Alun-alun Kota Batam

Hampir setiap kota di Indonesia memiliki alun-alun kota yang tentunya menjadi kebanggaan kota tersebut, tidak terkecuali kota Batam. Alun-alun kota Batam adalah ikon kota Batam dan merupakan salah satu objek wisata yang terletak di Batam, Kepulauan Riau. Alun-alun Kota Batam selain sebagai ikon dari

kota Batam itu sendiri juga dapat digunakan untuk beragam aktifitas mulai dari sekadar berkumpul dengan satu komunitas tertentu, bercengkrama, menikmati jajanan khas Batam, atau bahkan berolahraga ringan, agenda liburan bersama keluarga menjadi rutinitas yang sangat ditunggu-tunggu dan paling menyenangkan tentunya, tempat olah raga jogging yang nyaman, bersih dan terawat. Selain memiliki *jogging track*, ada juga lapangan futsal yang lumayan bagus. Siang maupun malam tempat enak buat olah raga, apabila malam lampunya cukup terang sehingga terasa nyaman banget. Tempatnya yang luas sehingga lebih leluasa untuk bermain dan sangat cocok untuk anak-anak bermain. Banyak yang bisa dilakukan disini, mulai dari berolah raga sampai berfoto-fotonya. Di Alun-alun kota Batam juga terdapat masjid raya kota Batam, pusat pemerintahan kota Batam seperti gedung wali kota dan gedung BP Batam, gedung DPRD, termasuk restoran dan pusat perbelanjaan seperti *mall*, dan pelabuhan feri terminal internasional untuk feri tujuan Batam ke Singapura dan sebaliknya.

2.3. Metode Naive Bayes

Algoritma yang sering digunakan menurut kelebihannya untuk analisis sentimen, khususnya klasifikasi, adalah algoritma *neural network* dan algoritma *naive bayes* (Soemantri et al, 2023). *Naive bayes* biasanya digunakan untuk mengaplikasikan data skala kecil untuk pelatihan, selain itu banyak digunakan untuk pengolahan data khususnya *text mining* karena memiliki tingkat akurasi yang lebih baik (Blanquero et al., 2021).

Naive bayes merupakan metode yang termasuk dalam algoritma klasifikasi yang berasal dari konsep pengolahan data statistik dan probabilitas yang dapat memprediksi peluang tersebut berdasarkan pengalaman sebelumnya. Persamaan *naive bayes* ditunjukkan pada persamaan dibawah ini (Solanki et al., 2021)

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

$P(H|X)$: Probabilitas hipotesis H berdasarkan kondisi X (posteriori probabilitas)

$P(X|A)$: Probabilitas X berdasarkan kondisi tersebut
H : Hipotesis bahwa data X adalah suatu kelas spesifik

$P(H)$: Probabilitas hipotesis H (prior probabilitas)

$P(X)$: Probabilitas dari X

2.4. Metode Support Vector Machine

Selain *naive bayes*, *support vector machine* merupakan salah satu model dalam algoritma klasifikasi teknik pemodelan *supervised learning* yang relatif baru untuk melakukan sebuah prediksi. *Support vector machine* berusaha untuk menentukan fungsi pemisah yang optimal yang bisa memisahkan

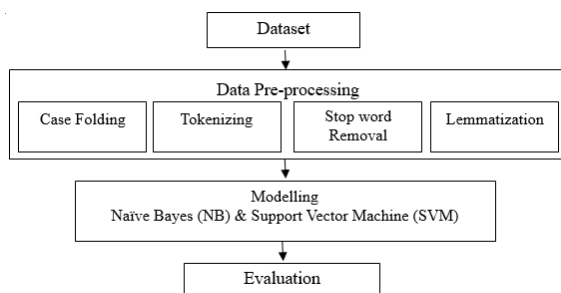
multi set data dari berbagai *class* yang berbeda (Pisner, et al., 2020). *Support vector machine* dapat menghasilkan akurasi dari hasil klasifikasi pengujian model (Darwis et al, 2020). *Support vector machine* merupakan model yang belajar dari teori pembelajaran statistika yang akan memberikan hasil yang lebih baik. *Support vector machine* adalah metode klasifikasi dan regresi yang biasa digunakan untuk masalah linier dan non linier. *Support vector machine* memiliki keuntungan dalam menerapkan pemisahan linier pada data masukan non linier berdimensi tinggi dan ini dicapai dengan menggunakan fungsi yang diperlukan. Efektivitas *support vector machine* sangat dipengaruhi oleh jenis fungsi yang dipilih dan diterapkan berdasarkan karakteristik data (Anreaja et al., 2022). Dalam analisis data banyak sekali jenis algoritma yang dapat digunakan. Masing-masing algoritma memiliki keunggulan tersendiri dalam menganalisa.

3. Metode

Penelitian ini bertujuan untuk mengevaluasi akurasi kinerja sentimen analisis ulasan objek wisata Alun-alun kota Batam dengan menggunakan metode *naive bayes* dan *support vector machine*. Metodologi penelitian ini diimplementasikan secara sistematis melalui beberapa tahap penelitian yang mencakup pengumpulan dataset, prapemrosesan data, pemodelan, dan tahap evaluasi model.

Pertama, proses pengumpulan dataset dilakukan untuk mengamati ulasan wisatawan terkait Alun-alun kota Batam pada platform *Google Maps*. Selanjutnya, tahap prapemrosesan data dilakukan untuk membersihkan dan menyusun data agar siap untuk proses analisis. Proses pemodelan dilanjutkan dengan penerapan metode *naive bayes* dan *support vector machine* untuk mengklasifikasikan sentimen ulasan menjadi kategori positif, negatif, atau netral.

Gambar 1. di bawah ini mengilustrasikan secara visual tahapan penelitian yang telah dijelaskan, memberikan pandangan yang lebih rinci tentang alur kerja metodologi penelitian ini.



Gambar 1. Tahapan Penelitian

Langkah terakhir dari penelitian ini adalah evaluasi model, yang mencakup pengukuran akurasi kinerja masing-masing model. Evaluasi dilakukan

dengan membandingkan hasil klasifikasi model dengan label sentimen yang sebenarnya pada data pengujian. Hasil evaluasi tersebut kemudian memberikan wawasan tentang sejauh mana masing-masing model mampu memprediksi sentimen ulasan dengan akurat.

4. Hasil dan Pembahasan

Hasil dan pembahasan memuat hasil penelitian dan pembahasannya yang ditulis secara terpisah.

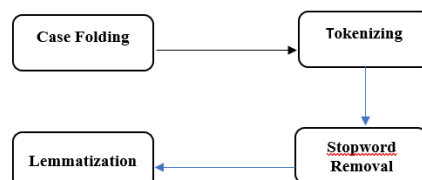
4.1. Hasil

4.1.1. Dataset

Pengumpulan dataset dilakukan dengan teknik *web scraping* yang diambil dari *website google map* <https://www.google.com/maps> menggunakan *instant data scraper* dengan memasukkan kata kunci objek wisata, tempat wisata, alun-alun kota Batam yang kemudian hasil pencariannya berisi ulasan komentar positif atau negatif dengan data yang diambil berupa data teks bahasa Indonesia mulai tahun 2018 sampai dengan tahun 2023 dengan memperoleh data sebanyak 1140 data, setelah dilakukan data preprocessing dan data cleaning diperoleh 578 data yang kemudian diolah menjadi model yang diinginkan ulasan dan hasilnya langsung bisa disimpan dalam bentuk *file excel*.

4.1.2. Prapemrosesan Data

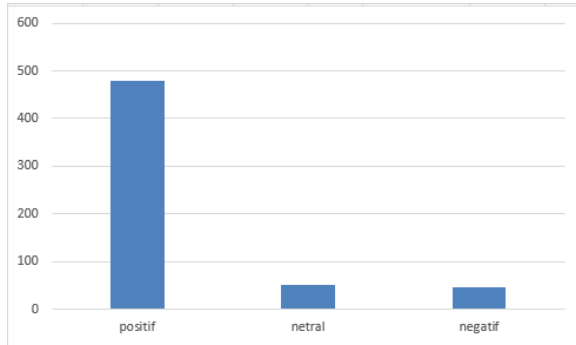
Prapemrosesan data merupakan tahap penting untuk dilakukan pada data sebelum analisis sehingga memperoleh informasi yang dibutuhkan. Tahapan prapemrosesan data untuk mendapatkan data teks yang diharapkan yaitu dengan melakukan data *cleaning* untuk memperbaiki dan mengoreksi data yang tidak akurat atau tidak lengkap dengan format yang sesuai. Tahapan prapemrosesan data dilakukan dengan beberapa tahapan yaitu *case folding*, *tokenization*, *stopword removal*, *lemmatization*. Tahapan *case folding* untuk merubah huruf kapital menjadi huruf kecil, *tokenization* yaitu proses seluruh kata pada setiap dokumen dipisahkan dan tanda baca dihilangkan begitu juga dengan karakter khusus yang dianggap sebagai pemisah kata atau tidak, serta jika ada simbol atau apapun yang bukan huruf, *stopword removal* dipakai untuk menghapus kata-kata yang ingin diabaikan atau difilter dari teks seperti kata-kata sangat umum seperti 'di', 'yang', dan *lemmatization* merupakan proses mengubah suatu kata menjadi kata dasar dengan mengetahui konteks dari kata tersebut (Zulfikar, et al., 2023)



Gambar 2. Tahapan Prapemrosesan Teks

4.1.3. Labeling

Proses penentuan label pada prapemrosesan data ini dilakukan secara manual berdasarkan penilaian ulasan positif, negatif dan netral seperti yang terlihat pada Gambar 3.



Gambar 3. Visualisasi Label Analisis Sentimen

4.2. Pembahasan

4.2.1. Evaluasi Pemodelan

Evaluasi pemodelan dalam penelitian ini menggunakan *confusion matrix* dan *classification report*. *Confusion matrix* berfungsi memvisualisasikan dan merangkum kinerja dan merangkum kerja algoritma klasifikasi di perbandingan label dan hasil prediksi. *Classification report* untuk mengukur kualitas prediktif dari klasifikasi tertentu algoritma sehingga menunjukkan presisi, recall dan akurasi dari penerapan algoritma model (Anreaja *et al.*, 2022). Tujuannya adalah untuk melihat dan membandingkan presisi, *recall*, dan akurasi *naive bayes* dan *support vector machine* dalam menganalisis sentimen. Dalam evaluasi model *naive bayes* dengan *confusion matrix*, hasil yang diperoleh hasil *True Positive* (TP) = 384, *False Positive* (FP) = 56, *True Negative* (TN) = 0, *False Negative* (FN) = 0, *True Neutral* (TN) = 1, dan *False Neutral* (FN) = 0. Hasil dari *confusion Matrix* dapat dilihat pada Tabel 1.

Tabel 1. *Confusion Matrix Naive Bayes*

Naive Bayes		Predict Class		
		Negative	Neutral	Positive
Actual Class	Negative	0	0	38
	Neutral	0	1	38
	Positive	0	0	384

Classification report adalah digunakan untuk menentukan *precision*, *recall*, dan *accuracy* pada sebuah model yang sedang dijalankan. *Classification report* metode *naive bayes* menunjukkan *precision* negatif, netral, dan positif sebesar 0,00%, 100,00%, dan 83,00%, sedangkan kelas *recall* negatif, netral dan positif sebesar 00,00%, 3,00%, dan 100,00% jadi hasil yang diperoleh rata-rata *precision* sebesar 61,00%, rata-rata *recall* 34,33% dengan *accuracy* sebesar 84,00%, seperti yang terlihat pada Gambar 4. dan Tabel 2.

	precision	recall	f1-score	support
negatif	0.00	0.00	0.00	38
netral	1.00	0.03	0.05	39
positif	0.83	1.00	0.91	384
accuracy			0.84	461
macro avg	0.61	0.34	0.32	461
weighted avg	0.78	0.84	0.76	461

Gambar 4. *Classification Report Naive Bayes*

Tabel 2. *Classification Report Naive Bayes*

Klasifikasi	Accuracy	Precision	Recall	F1-Score
Negative	0,84	0,00	0,00	0,00
Neutral	0,84	1,00	0,03	0,05
Positive	0,84	0,83	1,00	0,91

Evaluasi model *support vector machine* dengan *confusion matrix*, hasil yang diperoleh hasil *True Positive* (TP) = 384, *False Positive* (FP) = 9, *True Negative* = 31, *False Negative* (FN) = 0, *True Neutral* (TN) = 37, dan *False Neutral* (FN) = 0. Hasil dari *confusion matrix* dapat dilihat pada Tabel 3.

Tabel 3. *Confusion Matrix Support Vector Machine*

Naive Bayes		Predict Class		
		Negative	Neutral	Positive
Actual Class	Negative	31	0	7
	Neutral	0	37	2
	Positive	0	0	384

Classification report metode *support vector machine* menunjukkan *precision* negatif, netral, dan positif sebesar 100,00%, 100,00%, dan 98,00%, sedangkan *recall* negatif, netral dan positif, yaitu 82,00%, 95,00%, dan 100,00%. jadi hasil yang diperoleh rata-rata *precision* sebesar 99,33%, rata-rata *recall* sebesar 92,33% dengan *accuracy* sebesar 98,00%, seperti yang terlihat pada Gambar 5. dan Tabel 4.

	precision	recall	f1-score	support
negatif	1.00	0.82	0.90	38
netral	1.00	0.95	0.97	39
positif	0.98	1.00	0.99	384
accuracy			0.98	461
macro avg	0.99	0.92	0.95	461
weighted avg	0.98	0.98	0.98	461

Gambar 5. *Classification Report Model Support Vector Machine*

Tabel 4. *Classification Report Support Vector Machine*

Klasifikasi	Accuracy	Precision	Recall	F1-Score
Negative	0,98	1,00	0,82	0,90
Neutral	0,98	1,00	0,95	0,97
Positive	0,98	0,98	1,00	0,99

Berdasarkan hasil perbandingan evaluasi metode *naive bayes* dan *support vector machine* seperti yang terlihat pada table dibawah ini, metode *naive bayes* memberikan hasil *precision* sebesar 87,311%, *recall* sebesar 71,02%, dan *accuracy* sebesar 89,81%.

Metode support vector machine memberikan hasil rata-rata precision sebesar 94,23%, rata-rata *recall* sebesar 71,96%, dan rata-rata *accuracy* sebesar 98,00% dengan hasilnya dapat dilihat pada Tabel 5.

Tabel 5. Perbandingan Evaluasi Kinerja Metode *Naive Bayes* dan *Support Vector Machine*

Metode	Accuracy	Precision	Recall
<i>Naive Bayes</i>	84,00%	61,00%	34,33%
<i>Support Vector Machine</i>	98,00%	94,96%	71,96%

5. Kesimpulan

Hasil penelitian menunjukkan bahwa metode *Support Vector Machine* memberikan akurasi lebih tinggi (94%) dibandingkan dengan *Naive Bayes* (83%). Penelitian ini memberikan kontribusi dalam memberikan wawasan tentang sentimen pengunjung terhadap Alun-Alun Kota Batam, dengan implikasi untuk pengembangan kebijakan dan tindakan yang lebih efektif dalam meningkatkan daya tarik wisata lokal.

Daftar Pustaka

- Anreaja, L.J., Harefa, N.N., Negara, J.G.P., Priyantara, V.N.H., Prasetyo, A.B., 2022. Naive Bayes and Support Vector Machine Algorithm for Sentiment Analysis Opensea Mobile Application Users in Indonesia. *JISA: Jurnal Informatika dan Sains*, 5(1), 62-68. <https://doi.org/10.31326/jisa.v5i1.1267>
- Blanquero, R., Carrizosa, E., Ramírez-Cobo, P., Sillero-Denamiel, M.R., 2021. Variable Selection for Naive Bayes Classification. *Computers & Operations Research*, 135, 105456. <https://doi.org/10.1016/j.cor.2021.105456>
- Darwis, D., Pratiwi, E. S., Pasaribu, A.F O., 2020. Penerapan Algoritma SVM untuk Analisis Sentimen pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. *Jurnal Ilmiah Edutic: Pendidikan dan Informatika*, 7(1), 1-11. <http://dx.doi.org/10.21107/edutic.v7i1.8779>
- Han, J., Pei, J., Kamber, M., 2011. Data mining: concepts and techniques. Elsevier.
- Jumanto, Muslim, M.A., Dasril, Y., Mustaqim, T., 2023. Accuracy of Malaysia Public Response to Economic Factors During the Covid-19 Pandemic Using Vader and Random Forest. *Journal of Information System Exploration and Research*, 1(1), 49-70. <https://doi.org/10.52465/joiser.v1i1.104>
- Ririanti, N.P., Purwinarko, A., 2021. Implementation of Support Vector Machine Algorithm with Correlation-Based Feature Selection and Term Frequency Inverse Document Frequency for Sentiment Analysis Review Hotel. *Scientifics Journal of Informatics*, 8(2), 297-303. <https://doi.org/10.15294/sji.v8i2.29992>
- Rifa'i, A., Sujaini, H., Prawira, D., 2021. Sentiment Analysis Objek Wisata Kalimantan Barat pada Google Maps Menggunakan Metode Naive Bayes. *Jurnal Edukasi dan Penelitian Informatika*, 7(3), 400-407. <https://dx.doi.org/10.26418/jp.v7i3.48132>
- Raharjo, R.A., Sunarya, I.M.G., Divayana, D.G.H., 2022. Perbandingan Metode Naive bayes Classifier dan Support Vector Machine Pada Kasus Analisis Sentimen Terhadap Data Vaksin Covid-19 di Twitter. *Jurnal Ilmiah Elektronika dan Komputer*, 15(2) 456-464.
- Solanki, S.D., Solanki, A.D., Borah, S., 2021. Assimilate Machine Learning Algorithms in Big Data Analytics: Review, in Applications of Machine Learning in Big-Data Analytics and Cloud Computing. *River Publishers*, 81-114.
- Somantri, O., Maharrani, R.H., Purwaningrum, S., 2023. Coastal Sentiment Review Using Naive bayes with Feature Selection Genetic Algorithm. *Scientific Journal of Informatics*, 10(3), 127-136. <https://doi.org/10.15294/sji.v10i3.43988>
- Sun, S., Luo, C., Chen, J., 2017. A Review of Natural Language Processing Techniques for Opinion Mining Systems. *Information Fusion*, 36, 10-25. <https://doi.org/10.1016/j.inffus.2016.10.004>
- Surahman, A., Octaviansyah, A.F., Darwis, D., 2020. Ekstraksi Data Produk E-Marketplace Sebagai Strategi Pengolahan Segmentasi Pasar Menggunakan Web Crawler. *SISTEMASI: Jurnal Sistem Informasi*, 9(1), 73-81. <https://doi.org/10.32520/stmsi.v9i1.580>
- Thaha, A.R., Aziz, F., 2020. Text Mining on Tourism Destinations in Bandung Raya (Case Study: Tangkuban Perahu and Kawah Putih). *Jurnal Sekretaris dan Administrasi Bisnis*, 4(2), 146-156. <https://doi.org/10.31104/jsab.v4i2.172>
- Wardani, F.K., 2019. Analisis Sentimen untuk Peningkatan Popularitas Situs Belanja Online di Indonesia Menggunakan Metode Naive bayes (Studi Kasus Data Sekunder). Surabaya: Institut Bisnis dan Informatika STIKOM.
- Marwanta, Y.Y., Badiyanto, 2023. Analisis Sentimen Pencitraan Perguruan Tinggi di Yogyakarta Menggunakan Metode Naive Bayes Classifier. *Journal of Applied Informatic and Computing (JAIC)*, 7(1), 21-27.
- Zhang, F., Fleyeh, H., Wang, X., Lu, M., 2019. Construction Site Accident Analysis using Text Mining and Natural Language Processing Techniques. *Automation in Construction*, 99, 238-248. <https://doi.org/10.1016/j.autcon.2018.12.016>
- Zulfikar, W.B., Atmadja, A.R., Pratama, S.F., 2023. Sentiment Analysis on Social Media Against Public Policy Using Multinomial Naive Bayes. *Scientific Journal of Informatics*. *Scientific Journal of Informatics*, 10(1), 25-34. <https://doi.org/10.15294/sji.v10i1.39952>