

## Analisa Sentimen Pengunjung Hotel Menggunakan Metode Naïve Bayes

Tengku Ardiansyah Saputra<sup>1</sup>, Ahmad Zamsuri<sup>2</sup>, Yogo Turnandes<sup>3</sup>

<sup>1,2,3</sup> Teknik Informatika, Fakultas Ilmu Komputer Universitas  
Lancang Kuning

<sup>1,2,3</sup> Jl. Yossudarso KM, 8 Rumbai, Pekanbaru, Riau

e-mail: [tengkuardiansyah1601@gmail.com](mailto:tengkuardiansyah1601@gmail.com)<sup>1</sup>, [ahmadzamsuri@unilak.ac.id](mailto:ahmadzamsuri@unilak.ac.id)<sup>2</sup>,  
[turnandes@unilak.ac.id](mailto:turnandes@unilak.ac.id)

### Abstrak

*Review diberikan para pengunjung sangat penting ditelaah lebih lanjut agar kualitas hotel lebih meningkat. Banyaknya Review para pengunjung hotel membutuhkan waktu lama untuk mengulas satu per satu. Untuk mengumpulkan ulasan dan menganalisis dengan jumlah data berskala besar diperlukan suatu teknik. Tujuan penelitian ini untuk menghasilkan analisis sentiment pada review pengunjung Hotel dengan menggunakan metode algoritma Naïve Bayes. Tools yang digunakan Google Colaboratory dan menggunakan Algoritma Naïve Bayes untuk mencari akurasi hotel. Dengan menggunakan Algoritma Naïve Bayes menghasilkan akurasi dari Hotel Arena memiliki akurasi 0.74, K K Hotel George 0.94, Hotel Claridge Paris 0.92, Suite Hotel 900 m zur Oper 0.84, Atlantis Hotel Vienna 0.86.*

**Kata Kunci:** Review, Analisis, Naïve Bayes

### 1. PENDAHULUAN

Teknologi informasi terdapat perangkat lunak dan perangkat keras, merupakan alat untuk mendukung dan meningkatkan kualitas informasi secara cepat dan efektif bagi setiap orang di masyarakat. Karena kemajuan teknologi informasi, pengguna sekarang harus memeriksa untuk melihat apakah pengelolaan teknologi informasi dan fungsi bisnis mencapai hasil yang diharapkan.

Berdasarkan pertumbuhan internet dan media sosial telah berkembang pesat sehingga menyediakan berbagai informasi opini dari orang lain. Sebuah survey dilakukan oleh asosiasi Pengusaha Jasa Internet (APJII) pada tahun 2014 menunjukkan bahwa jumlah pengguna internet di Indonesia mengalami kenaikan dalam beberapa tahun terakhir. (Darwis dan Siskawati, 2021)

Indrustri perhotelan merupakan tempat pariwisata yang banyak dikunjungi. Dalam perhotelan pengunjung ingin mengetahui kualitas hotel tanpa harus berkunjung di hotel tersebut. Dengan adanya ulasan ulasan pada sebuah website yang menyediakan sebuah informasi mengenai akomodasi dan permasalahan hotel yang ada diberbagai dunia. Banyaknya ulasan yang diberikan membuat para pengunjung sulit mengetahui kualitas yang ada di hotel tersebut.

Penelitian ini bertujuan untuk memberikan saran supaya memperbaiki fasilitas dan pelayanan yang diberikan hotel. Review diberikan para pengunjung sangat penting untuk ditelaah lebih lanjut agar kualitas hotel menjadi lebih meningkat.

Penulis memilih judul "Analisa Sentimen Pengunjung Hotel Menggunakan Metode Naïve Bayes". Algoritma Naïve Bayes merupakan metode yang memanfaatkan probabilitas ketika kita mengetahui probabilitas tertentu lainnya.

---

## 2. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah salah satu tugas data mining untuk klasifikasi data, yaitu memetakan (mengklasifikasikan) data ke dalam satu kelas atau beberapa kelas yang sudah didefinisikan sebelumnya, sehingga memungkinkan untuk penentuan review hotel.

Pengumpulan data dilakukan dengan mengidentifikasi masalah penelitian yang dihadapi. Selanjutnya, melakukan pengumpulan data melalui observasi, dan kajian pustaka.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Analisa Data

Proses analisa data dalam penelitian ini menggunakan metode *Naïve Bayes* dengan mencari nilai sentimen dan nilai akurasi dari mengklasifikasi review hotel dengan menggunakan google colab.

### 3.2 Perhitungan Algoritma Naïve Bayes

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

$P(H|X)$  = Probabilitas berdasarkan kondisi (posteriori probability)

$P(H)$  = Probabilitas hipotesis (prior probability)

$P(X|H)$  = Probabilitas berdasarkan kondisi hipotesis H

$P(X)$  = Probabilitas x

Dengan menggunakan rumus diatas nilai kalimat review dapat dihitung. Selanjutnya melakukan perhitungan kalimat review, hasil perhitungan kalimat sebagai berikut:

Tabel 1 Data Riview

Doc	Komentar	Label
D1	staf yang luar biasa kamar yang bersih nyaman akan inap lagi	Positif
D2	staf senang tempat tidur besar sangat akomodatif tidak jauh dari pusat kota	Positif
D3	staf tipu kami dengan harga	Negatif
D4	kamar dingin tidak privasi kamar mandi	Negatif
D5	kamar pertama milik tangga yang curam tempat tidur loteng tidak praktis bahaya	?

Jumlah Term Negatif (D1, D2) = 18

Jumlah Term Positif (D3, D4) = 11

Berikut adalah cara perhitungan manual metode Naïve Bayes:

a. Peluang Kemunculan kategori Positif dan Negatif (Probabilitas Kelas) :

$$P(V_j) = \frac{|docs_j|}{|D|}$$

$$P(Positif) = \frac{\text{Jumlah data training positif}}{\text{Jumlah semua data training}} = \frac{2}{4} = 0,5$$

$$P(Negatif) = \frac{\text{Jumlah data training negatif}}{\text{Jumlah semua data training}} = \frac{2}{4} = 0,5$$


---

b. Perhitungan Probabilitas kata setiap kelas :

$$P(X_i | Y_j) = \frac{nk+1}{n+|kata|}$$

$$P(\text{"kamar"}|\text{positif}) = \frac{1+1}{18} = 0,11$$

$$P(\text{"kamar"}|\text{negatif}) = \frac{2+1}{11} = 0,27$$

$$P(\text{"tempat"}|\text{positif}) = \frac{1+1}{18} = 0,11$$

$$P(\text{"tempat"}|\text{negatif}) = \frac{0+1}{11} = 0,09$$

$$P(\text{"tidur"}|\text{positif}) = \frac{1+1}{18} = 0,11$$

$$P(\text{"tidur"}|\text{negatif}) = \frac{0+1}{11} = 0,09$$

$$P(\text{"tidak"}|\text{positif}) = \frac{1+1}{18} = 0,11$$

$$P(\text{"tidak"}|\text{negatif}) = \frac{1+1}{11} = 0,18$$

c. Perhitungan Klasifikasi Kategori :

$$= P(\text{Positif}) + P(\text{"kamar"}|\text{positif}) + P(\text{"tempat"}|\text{positif}) + P(\text{"tidur"}|\text{positif}) + P(\text{"tidak"}|\text{positif}) = 0,94$$

$$= P(\text{Negatif}) + P(\text{"kamar"}|\text{negatif}) + P(\text{"tempat"}|\text{negatif}) + P(\text{"tidur"}|\text{negatif}) + P(\text{"tidak"}|\text{negatif}) = 1,13$$

Berdasarkan perhitungan pada komentar Hotel Arena mendapatkan akurasi klasifikasi positif sebesar 0,94 sedangkan untuk akurasi klasifikasi negatif sebesar 1,13. Dimana pada setiap kata memiliki nilai probabilitas positif dan negatif. Dapat dilihat dari jumlah perkalian klasifikasi kategori positif lebih kecil dari pada kategori klasifikasi negatif maka komentar dikategorikan positif. Peluang nilai klasifikasi kategori kata pada label positif dan negatif sebesar  $0,94 < 1,13$  maka dapat disimpulkan bahwa komentar pada D5 termasuk label positif.

### 3.3 Pengujian Accuracy Naïve Bayes Menggunakan Google Colaboratory

#### 3.3.1 Hasil Pengujian Preprocessing

	nama hotel	alamat	sentimen	review
0	arena hotel	irlandia	positif	tidak keluh nyata hotel bagus lokasi yang bagu...
1	arena hotel	belgium	negatif	meski gambar tunjuk kamar yang sangat bersih k...
2	arena hotel	italia	positif	hotel sedang renovasi dengan sangat hati denga...
3	arena hotel	italia	positif	hotel luar biasa saya ambil dengan tulus karen...
4	arena hotel	belanda	positif	area umum indah kamar bagus tapi jendela rusak...
...	...	...	...	...
109	arena hotel	israel	negatif	jalan jauh dari pusat kota
110	arena hotel	britania raya	positif	ruang indah staf sangat bantu
111	arena hotel	britania raya	positif	ruang cerah lapang staf sangat ramah
112	arena hotel	britania raya	positif	staf sangat ramah sangat bantu
113	arena hotel	britania raya	positif	hotel takjub bangun yang indah kamar sangat mo...

114 rows x 4 columns

**Gambar 1.** Hasil Preprocessing

### 3.3.2 Pengaturan Kinerja

Hasil perhitungan data di bagi menjadi 80% data training dan 20% data testing. Hasil perhitungan dapat dilihat pada gambar 2 dan 3 di bawah ini.

	nama hotel	alamat	review	sentimen
0	arena hotel	australia	lokasi bagus belah trem tetapi sedang dalam ba...	positif
1	arena hotel	britannia raya	sarap	positif
2	arena hotel	britannia raya	ruang cerah lapang staf sangat ramah	positif
3	arena hotel	britannia raya	staf sangat sopan sedia bantu tiap saat hotel ...	positif
4	arena hotel	irlandia	staf tahan adalah terima yang bagus agak loyo ...	positif
...	...	...	...	...
86	arena hotel	austria	transparan buruk	negatif
87	arena hotel	britannia raya	kena biaya untuk paket ulang tahun ketika ikla...	negatif
88	arena hotel	italia	hotel bintang biasa tidak yang istimewa tidak ...	positif
89	arena hotel	britannia raya	kamar yang luar biasa hotel yang tarik lokasi ...	positif
90	arena hotel	turki	bangun dengan suasana luar biasa dengan desain...	positif

91 rows x 4 columns

**Gambar 2.** Hasil Data Training

Dalam konteks pembelajaran mesin dan pengembangan model, pembagian data menjadi dua bagian utama, yaitu data training (80%) dan data testing (20%), bertujuan untuk melatih dan menguji kinerja model. Proses ini dikenal sebagai pembagian data training-testing, dan berikut adalah penjelasan mengenai konsep ini:

a. Data Training (80%):

Fungsi Utama: Data training digunakan untuk melatih model mesin.

Proses: Model diberikan akses ke data training untuk memahami pola dan hubungan dalam data tersebut. Parameternya disesuaikan secara iteratif untuk meminimalkan kesalahan prediksi model terhadap data training.

Tujuan: Maksimalkan kemampuan model untuk menangkap pola yang mungkin ada dalam data, sehingga dapat membuat prediksi yang akurat.

b. Data Testing (20%):

Fungsi Utama: Data testing digunakan untuk menguji kinerja model yang telah dilatih.

Proses: Model yang sudah dilatih diuji dengan menggunakan data testing yang tidak pernah dilihat selama proses pelatihan. Tujuannya adalah untuk mengevaluasi sejauh mana model dapat melakukan prediksi yang baik pada data yang belum pernah dilihat sebelumnya.

Tujuan: Mengukur keberlanjutan model terhadap data yang baru dan tidak terlihat selama pelatihan. Memastikan bahwa model mampu generalisasi dengan baik dan tidak hanya "menghafal" data training.

	nama hotel	alamat	review	sentimen
0	arena hotel	britannia raya	staf yang luar biasa kamar yang bersih nyaman ...	positif
1	arena hotel	belanda	area umum indah kamar bagus tapi jendela rusak...	positif
2	arena hotel	argentina	kami suka kami inap arena hotel hotel sangat b...	positif
3	arena hotel	britannia raya	staf senang tempat tidur besar sangat akomodat...	positif
4	arena hotel	britannia raya	seluruh hotel sangat bersih staf sangat ramah ...	positif
5	arena hotel	britannia raya	hotel super gaya belah taman yang indah cukup ...	positif
6	arena hotel	britannia raya	staf tipu kami dengan harga	negatif
7	arena hotel	britannia raya	kami tinggal kamar studio kamar luas bersih te...	positif
8	arena hotel	rusia	kamar standar adalah lokasi yang luar biasa ge...	positif
9	arena hotel	britannia raya	hotel sangat bagus staf sangat ramah bantu say...	positif
10	arena hotel	britannia raya	akan suka tempat tidur ganda	positif
11	arena hotel	spanyol	saya suka desain dekorasi hotel staf sangat ba...	positif
12	arena hotel	britannia raya	staf fantastis yang tidak bisa lebih bantu ram...	positif
13	arena hotel	britannia raya	staf sangat ramah sangat bantu	positif
14	arena hotel	amerika serikat	kamar pertama milik tangga yang curam tempat t...	negatif
15	arena hotel	irlandia	tidak keluh nyata hotel bagus lokasi yang bagu...	positif
16	arena hotel	britannia raya	lokasi nuansa hotel	positif
17	arena hotel	britannia raya	hotel luar biasa interior yang indah kami tida...	positif
18	arena hotel	irlandia	kamar dingin tidak privasi kamar mandi	negatif
19	arena hotel	britannia raya	staf sangat bantu hotel lokasi yang bagus	positif
20	arena hotel	taiwan	sangat dekat dengan area museum gedung konser	positif
21	arena hotel	britannia raya	kerja bangun kamar	negatif
22	arena hotel	britannia raya	kami telah tinggal sini beberapa kali selalu n...	positif

**Gambar 3.** Hasil Data Testing

### 3.4 Hasil Accuracy Naïve Bayes

#### 3.4.1 Hasil accuracy hotel arena

Hasil pengujian hotel arena dengan split persentase 80% data training dan 20% data testing menunjukkan tingkat accuracy sebesar 74% atau 0,74 seperti gambar 4 di bawah ini.

	precision	recall	f1-score	support
negatif	0.50	0.33	0.40	3
positif	0.76	1.00	0.86	16
negatif	0.00	0.00	0.00	1
positif	0.00	0.00	0.00	3
accuracy			0.74	23
macro avg	0.32	0.33	0.32	23
weighted avg	0.60	0.74	0.65	23

**Gambar 4.** Hasil Accuracy Hotel Arena

### 3.4.2 Hasil accuracy KK hotel George

Pengujian hotel KK hotel George dengan split persentase 80% data training dan 20% data testing hasil tingkat accuracy sebesar 94% atau 0,94 seperti gambar 5 dibawah ini.

	precision	recall	f1-score	support
negatif	1.00	0.75	0.86	8
positif	0.93	1.00	0.97	28
accuracy			0.94	36
macro avg	0.97	0.88	0.91	36
weighted avg	0.95	0.94	0.94	36

**Gambar 5.** Hasil accuracy KK hotel George

### 3.4.3 Hasil Accuracy Hotel Claridge Paris

Hasil pengujian hotel Claridge paris di bawah ini pada gambar 6 dengan split persentase 80% data training dan 20% data testing hasil tingkat accuracy sebesar 92% atau 0,92.

	precision	recall	f1-score	support
negatif	0.00	0.00	0.00	0
positif	1.00	0.92	0.96	12
accuracy			0.92	12
macro avg	0.50	0.46	0.48	12
weighted avg	1.00	0.92	0.96	12

**Gambar 6.** Hasil Accuracy Hotel Claridge Paris

### 3.4.4 Hasil Accuracy Suite Hotel 900 M Zur Oper

Pengujian hotel Suite Hotel 900 M Zur Oper dengan split persentase 80% data training dan 20% data testing hasil tingkat accuracy sebesar 84% atau 0,84 seperti gambar 7 di bawah ini.

	precision	recall	f1-score	support
negatif	0.67	0.33	0.44	6
positif	0.86	0.96	0.91	25
accuracy			0.84	31
macro avg	0.76	0.65	0.68	31
weighted avg	0.82	0.84	0.82	31

**Gambar 7.** Hasil Accuracy Suite Hotel 900 M Zur Oper

### 3.4.5 Hasil Accuracy Atlantis Hotel Vienna

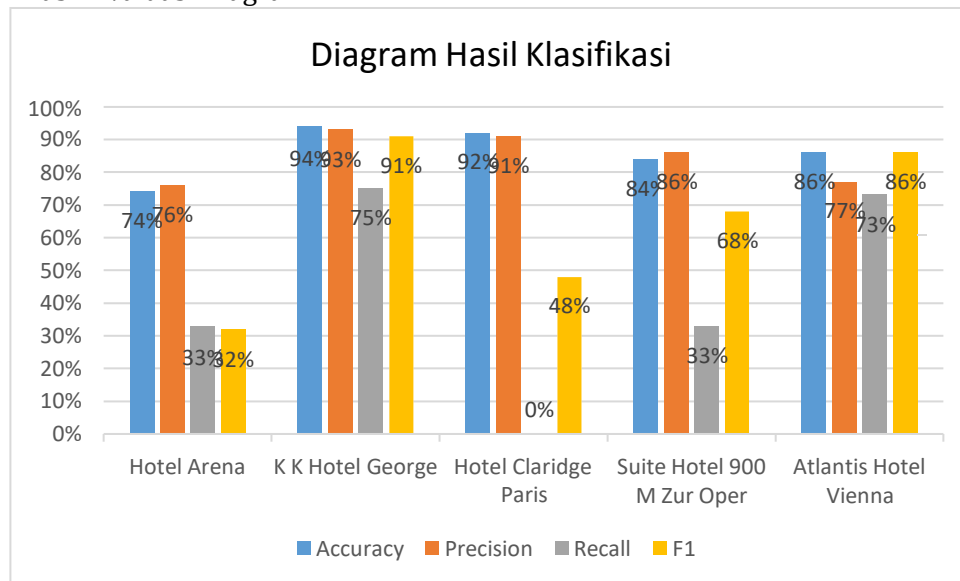
Pengujian hotel Atlantis Hotel Vienna di bawah ini pada gambar 8 dengan split persentase 80% data training dan 20% data testing hasil tingkat accuracy sebesar 86% atau 0,86.

---

	precision	recall	f1-score	support
negatif	1.00	0.73	0.84	11
positif	0.77	1.00	0.87	10
accuracy			0.86	21
macro avg	0.88	0.86	0.86	21
weighted avg	0.89	0.86	0.86	21

**Gambar 8.** Hasil Accuracy Atlantis Hotel Vienna

### 3.4.6 Hasil Evaluasi Diagram



**Gambar 9.** Hasil Evaluasi Diagram

Berdasarkan Gambar 9 diagram dari *Accuracy* klasifikasi *naïve bayes* dapat di lihat pada diagram di atas dengan hasil bentuk persen dengan nilai *Accuracy* berwarna biru, *Precision* berwarna jingga, *Recall* berwarna abu-abu, dan F1 berwarna kuning.

## 4. KESIMPULAN

Analisa review hotel menghasilkan nilai klasifikasi *naïve bayes* dengan *accuracy* pada hotel dengan nilai Hotel Arena memiliki akurasi 0.74, K K Hotel George 0.94, Hotel Claridge Paris 0.92, Suite Hotel 900 m zur Oper 0.84, Atlantis Hotel Vienna 0.86. Untuk penelitian selanjutnya, Penulis menyarankan agar pada penelitian dilakukan uji coba menggunakan Algoritma lainnya untuk mendapatkan hasil yang kurang atau lebih baik dan juga melakukan penerapan pada tools lainnya. Untuk penelitian selanjutnya, penulis menyarankan untuk melakukan pengujian di tempat wisata, restoran dan lainnya dengan klasifikasi sentimen yang tidak hanya sentimen positif ataupun negatif.

### UCAPAN TERIMAKASIH

Saya ingin menyampaikan ekspresi terima kasih yang mendalam atas bantuan, dukungan, dan bimbingan yang luar biasa dari dosen pembimbing dan seluruh program studi selama perjalanan penyusunan jurnal ini. Pada kesempatan ini, saya merasa perlu untuk mengungkapkan sejauh mana peran beliau dalam mencapai kesuksesan dalam penelitian ini. Dosen pembimbing telah menjadi sumber inspirasi dan penuntun yang tak ternilai selama seluruh proses penelitian. Bimbingan yang diberikan tidak hanya bersifat akademis, tetapi juga melibatkan dorongan positif, motivasi, dan arahan yang konstruktif. Beliau tidak hanya berperan sebagai pengajar, tetapi juga sebagai mentor yang peduli dengan pengembangan diri mahasiswa. Sikap kesabaran dan dedikasi dosen pembimbing dalam membimbing kami melalui tantangan dan kompleksitas penelitian ini benar-benar luar biasa.

### DAFTAR PUSTAKA

- [1] Amrullah, A. Z., Sofyan Anas, A., Adrian, M., & Hidayat, J. (2020). Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square. *Jurnal*, 2(1). <https://doi.org/10.30812/bite.v2i1.804>
  - [2] Azza Amira, S., Utama, S., & Muhammad Hanif Fahmi, dan. (2020). Edu Komputika Journal Penerapan Metode Support Vector Machine untuk Analisis Sentimen pada Review Pelanggan Hotel. *Edu Komputika*, 7(2). <http://journal.unnes.ac.id/sju/index.php/edukom>
  - [3] Bijaksana, A., Negara, P., Muhardi, H., & Putri, I. M. (2020). ANALISIS SENTIMEN MASKAPAI PENERBANGAN MENGGUNAKAN METODE NAIVE BAYES DAN SELEKSI FITUR INFORMATION GAIN SENTIMENT ANALYSIS ON AIRLINES USING NAÏVE BAYES METHOD AND FEATURE SELECTION INFORMATION GAIN. 7(3), 599–606. <https://doi.org/10.25126/jtiik.202071947>
  - [4] Cahyani, R., & Pandu Adikara, P. (2019). *Analisis Sentimen terhadap Ulasan Hotel menggunakan Boosting Weighted Extreme Learning Machine* (Vol. 3, Issue 8). <http://j-ptiik.ub.ac.id>
  - [5] Darwis, D., Siskawati, N., & Abidin, Z. (n.d.). *Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional*. 15(1).
  - [6] Dini Utami, L., Rachmi, H., Nurlaela, D., & Informasi Kampus Kota Bogor, S. (2018). *KOMPARASI ALGORITMA KLASIFIKASI PADA ANALISIS REVIEW HOTEL*. 14(2), 261. [www.bsi.ac.id](http://www.bsi.ac.id)
  - [7] Khofifah, W., Rahayu, D. N., & Yusuf, A. M. (2022). Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 16(4). <https://doi.org/10.35969/interkom.v16i4.192>
  - [8] Kusuma, A., & Nugroho, A. (2021). Analisa Sentimen Pada Twitter Terhadap Kenaikan Tarif Dasar Listrik Dengan Metode Naïve Bayes. *Jurnal Ilmiah Teknologi Informasi Asia*, 15(2). <https://doi.org/10.32815/jitika.v15i2.557>
  - [9] Nasution, N., Djahara, K., & Zamsuri, A. (2015). Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naïve Bayes (Studi Kasus: Fasilkom Unilak). *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, 6(2), 1-11.
  - [10] Program, A. T., Sistem, S., Stmik, I., Mandiri, N., Jl, J., Damai, N., Barat, W. J., & Selatan, J. (2017). Optimasi Particle Swarm Optimization Sebagai Seleksi Fitur Pada Analisis Sentimen Review Hotel Berbahasa Indonesia Menggunakan Algoritma Naïve Bayes. *Jurnal Teknik Komputer*, III(2). <http://www.tripadvisor.com>
-



- [11] Taufik STMIK Nusa Mandiri Jakarta Jl Damai No, A., Jati Barat Jakarta Selatan, W., & Sitasi, C. (2018). Komparasi Algoritma Text Mining Untuk Klasifikasi Review Hotel. *Jurnal Teknik Komputer*, 2, 112–118. <https://doi.org/10.31294/jtk.v4i2.3461>
- [12] Thomas, V. W. D., & Rumaisa, F. (2022). Analisis Sentimen Ulasan Hotel Bahasa Indonesia Menggunakan Support Vector Machine dan TF-IDF. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 6(3), 1767. <https://doi.org/10.30865/mib.v6i3.4218>
- [13] Utami, D. S., Erfina, A., & Id, M. A. (2022). Analisis Sentimen Ulasan Terkait UNESCO Global Geopark Di Google Maps dengan Algoritma Naive Bayes. In *Jurnal Sains Komputer & Informatika (J-SAKTI)* (Vol. 6, Issue 2).



*Prosiding- SEMASTER: Seminar Nasional Teknologi Informasi & Ilmu Komputer* is licensed under a [Creative Commons Attribution International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/)

---