

THE PERCEPTIONS OF SEMARANG FIVE STAR HOTEL TOURISTS WITH SUPPORT VECTOR MACHINE ON GOOGLE REVIEWS

**Muhammad Haikal Aufan^{*1}, Maya Rini Handayani², Afifah Basmah Nurjanna³, Nur Cahyo Hendro
Wibowo⁴, Khotibul Umam⁵**

^{1,3,4,5}Department of Information Technology, Science and Technology Faculty, Universitas Islam Negeri
Walisongo Semarang, Indonesia

²Department of Da'wah and Communication, Science and Technology Faculty, Universitas Islam Negeri
Walisongo Semarang, Indonesia

Email: ¹2108096044@student.walisongo.ac.id, ²maya@walisongo.ac.id, ³2108096061@student.walisongo.ac.id,
⁴nur_cahyo@walisongo.ac.id, ⁵khotibul_umam@walisongo.ac.id

(Article received: April 30, 2024; Revision: June 05, 2024; published: October 20, 2024)

Abstract

Travelers on the road sometimes need a hotel to rest. In choosing a hotel, they refer to the ratings or reviews written by users through reviews on Google. This is because not all star hotels provide facilities in accordance with user assessments. This study discusses the analysis of the opinions of tourists who have stayed in 5-star hotels in Semarang through a review of commentary data on Google. The 5-star hotels used as the research are Padma, Gumaya, Tentrem, Grand Candi, Ciputra, and PO. The dataset of the six hotels was obtained through a scraping process then followed by data pre-processing. The data was retrieved from Google Maps using the Chrome Instant Data Scraper extension. Data preprocessing begins with case folding, tokenizing, filtering, and ends with stemming. Support Vector Machine (SVM) is implemented for sentiment classification process. The results from this study are the majority of 5-star hotel reviews in Semarang tend to have positive rather than negative sentiments. Our model was able to produce an accuracy of 0.87 to 0.98. The highest accuracy was achieved by Ciputra Hotel at 0.98 with 543 positive reviews.

Keywords: 5-star hotel, Semarang, SVM.

PERSEPSI WISATAWAN HOTEL BINTANG LIMA SEMARANG DENGAN SUPPORT VECTOR MACHINE PADA ULASAN GOOGLE

Abstrak

Wisatawan dalam melakukan perjalanan wisata kadang membutuhkan sebuah hotel sebagai tempat beristirahat. Dalam memilih hotel mereka mengacu pada ulasan yang dituliskan oleh para pengguna melalui ulasan di Google. Hal ini disebabkan tidak semua hotel berbintang fasilitas yang disediakan sesuai dengan penilaian pengguna. Studi ini membahas tentang analisa pendapat para wisatawan yang telah menginap di hotel bintang 5 di Semarang melalui tinjauan data ulasan di Google. Hotel bintang 5 yang dijadikan penelitian adalah Padma, Gumaya, Tentrem, Grand Candi, Ciputra, dan PO. Dataset dari ke enam hotel tersebut diperoleh dari proses *scraping* yang dilanjutkan dengan proses *pre-processing* data. Data diambil dari Google Maps dengan menggunakan extension Chrome Instan Data Scraper. Untuk *pre-processing* data diawali dengan *case folding*, *tokenizing*, *filtering*, dan diakhiri *stemming*. Support Vector Machine (SVM) diimplementasikan untuk proses klasifikasi sentimen. Hasil yang diperoleh dari penelitian ini adalah mayoritas ulasan hotel bintang 5 di Semarang cenderung bersentimen positif daripada negatif. Model mampu menghasilkan akurasi sebesar 0.87 hingga 0.98. Akurasi tertinggi diperoleh hotel Ciputra sebesar 0.98 dengan review positif sebanyak 543 data dan accuracy sebesar 0.98.

Kata kunci: hotel bintang 5, Semarang, SVM.

1. PENDAHULUAN

Industri pariwisata telah menunjukkan pertumbuhan yang signifikan dan menjadi salah satu kontributor utama bagi ekonomi di banyak kota di dunia, termasuk Semarang, salah satu kota yang

menonjol di Indonesia[1]. Pariwisata dan perhotelan memiliki hubungan yang sangat erat. Hotel adalah sarana pokok pariwisata yang banyak tergantung oleh jumlah wisatawan yang datang[2]. Dalam konteks ini, hotel-hotel di Semarang tidak hanya berfungsi

sebagai tempat menginap bagi wisatawan, tetapi juga sebagai elemen penting dalam menentukan pengalaman wisata yang memuaskan[3].

Keberadaan ulasan online di mana-mana dan keluasannya, seperti *Google Maps*, menjadi sumber informasi penting bagi wisatawan dalam mencari hotel[4]. Ulasan dari pengguna platform tersebut berperan penting dalam membantu calon wisatawan menentukan akomodasi yang sesuai[5]. Hal ini penting karena banyaknya pilihan hotel yang tersedia dapat menimbulkan kebingungan. Hotel memiliki rencana pemasaran yang bisa dikelompokkan berdasarkan beragam faktor, termasuk jenis produk, harga, lokasi, promosi, proses, staf, fasilitas fisik, dan fokus pada kebutuhan pelanggan [6]. Persepsi positif dari pengunjung dapat meningkatkan minat calon pengunjung untuk memilih hotel tersebut. Di sisi lain, persepsi negatif dapat membuat mereka mempertimbangkan pilihan lain.

Pada fitur ulasan di berbagai situs sering kali tidak tersedia opsi untuk memilah ulasan antara yang bersifat positif (ulasan dengan pandangan positif) atau negatif (ulasan dengan pandangan negatif), sehingga diperlukan analisis sentimen untuk membantu mengklasifikasikan ulasan dari pengunjung sebelumnya. *Opinion Mining* atau analisis sentimen adalah proses otomatis yang digunakan untuk mengerti, mengambil, dan mengolah data teks untuk mendapatkan informasi tentang sentimen yang terkandung dalam sebuah kalimat opini[7]. Analisis sentimen para wisatawan ditentukan oleh sejumlah faktor, termasuk volume data teks, sumber data, dan efisiensi algoritma yang digunakan. Wisatawan memiliki penilaian terhadap pelayanan yang diberikan oleh hotel di suatu tempat tujuan, sehingga data tersebut dapat dikelompokkan atau diklasifikasikan untuk menentukan apakah penilaiannya bersifat positif atau negatif[8].

Klasifikasi teks dapat dilakukan dengan berbagai metode, salah satunya *Support Vector Machine* (SVM). SVM adalah metode klasifikasi yang kuat dan efektif dengan tingkat keberhasilan tinggi di berbagai bidang. SVM mampu mengenali pola klasifikasi dalam teks dengan baik, sehingga cocok untuk berbagai aplikasi[9]. Akurasi yang tinggi menjadi keunggulan SVM dalam klasifikasi data teks[10]. Keberhasilan SVM dalam klasifikasi teks membuatnya menjadi pilihan populer di berbagai bidang.

Pada penelitian yang dilakukan oleh Cahyo dan Arief[11] dengan mengambil objek penelitian pada ulasan pengunjung Keraton Yogyakarta di *Google Maps*, hasil evaluasi SVM menunjukkan akurasi 85,67% dengan 1607 data dan rasio data training dan testing 80:20. Penelitian yang dilakukan Saraswati dan Indradewi[12] dengan penelitian Polaritas Ulasan Hotel menggunakan SVM, menghasilkan akurasi 98,98%

Pada penelitian yang dilakukan oleh Singgalen[8] dengan mengambil objek penelitian

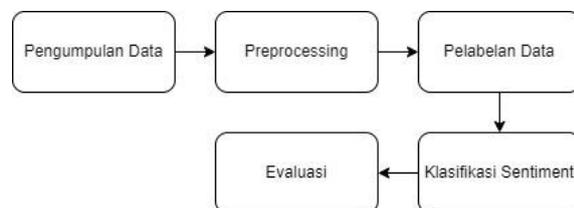
pada ulasan pengunjung terhadap 10 Traveler Ranked Hotel di Kota Makassar, hasil evaluasi SVM menunjukkan akurasi 98,98% dengan 1138 data dan rasio data training dan testing 70:30. Penelitian lainnya yang dilakukan oleh Singgalen[13] dengan objek penelitian yang berbeda, yakni pada Candi Borobudur menunjukkan evaluasi SVM dengan akurasi 98,73%

Pada penelitian yang dilakukan oleh Zaenal dan Astutik[14] dengan mengambil objek penelitian pada ulasan Aplikasi OYO, hasil evaluasi SVM menunjukkan akurasi 80,79% dengan 2000 data dan rasio data training dan testing 80:20.

Claudia dkk.[15] Menggunakan SVM Hasil percobaan menunjukkan bahwa model klasifikasi SVM dengan kernel RBF memiliki performa yang tinggi untuk semua kombinasi data latihan dan data uji pada kejadian dengan rasio ukuran yang sama maupun berbeda.

Dari ke enam penelitian yang telah disebutkan memiliki dua persamaan yaitu penggunaan SVM sebagai metodenya dan akurasi yang diperoleh cukup tinggi yaitu 80%. Berdasarkan hasil tersebut kami tertarik untuk menerapkan metode SVM untuk menganalisis ulasan hotel bintang 5 di *Google* dengan tujuan untuk mengetahui hotel mana yang memiliki *review* yang paling baik menggunakan metode SVM.

2. METODE PENELITIAN



Gambar 1 Diagram alir penelitian

2.1. Pengumpulan Data

Pengumpulan data yang diperoleh berasal dari ulasan pengguna *Google Maps*. Pengumpulan data dilakukan dengan teknik *Web Scraping* menggunakan *extension chrome Instant Data Scraper*. *Web scraping* adalah metode ekstraksi data yang dilakukan dengan cara membuka situs web dan mengambil data yang diinginkan secara otomatis menggunakan program. Terdapat empat informasi yang diperoleh dari hasil *scraping* dari *Google* yaitu *username*, *rating*, *review*, dan waktu. Waktu menunjukkan saat dimana pengguna menuliskan ulasan di *Google*. Hasil *scraping* disimpan dalam bentuk format XLSX, kemudian diambil data *username*, *reviews*, dan *rating* saja. File XLSX kemudian di konversi ke format CSV.

2.2. Pre-processing Data

Setelah melakukan pengambilan data ulasan hotel melalui proses *scraping*, langkah selanjutnya

adalah *pre-processing* data. Dalam tahap ini, beberapa langkah dilakukan yaitu, *case folding*, *tokenizing*, *filtering*, dan *stemming*[16].

2.2.1. Case Folding

Pada langkah ini, data ulasan yang telah diperoleh akan disesuaikan agar semua hurufnya menjadi huruf kecil dan karakter yang tidak relevan seperti tanda baca dan angka akan dihapus. Dengan demikian, teks akan berupa huruf kecil saja tanpa adanya tanda baca.

2.2.2. Tokenizing

Tokenizing adalah proses membagi teks menjadi bagian-bagian terkecil berupa kata. Dengan demikian, setiap kata yang sudah melalui tahap *case folding* akan dipisahkan menjadi entitas tersendiri.

2.2.3. Filtering

Kata-kata yang tidak relevan dengan penelitian akan disaring untuk kemudian dihilangkan.

2.2.4. Stemming

Setelah proses *filter*, kata-kata tersebut akan disesuaikan agar menjadi kata dasar dengan memangkas awalan, akhiran, atau sisipan yang masih ada. Kata dasar inilah yang akan digunakan untuk melakukan pembobotan kata sebelum masuk ke tahap klasifikasi.

2.3. Pelabelan Data

Pelabelan data dilakukan dengan mempertimbangkan ulasan dan rating pengguna di Google Maps. Proses ini mengkategorikan ulasan menjadi dua kelompok: positif untuk rating antara 3 hingga 5, dan negatif untuk rating antara 1 hingga 2.

2.4. Klasifikasi Sentimen

Setelah topik dari setiap ulasan diidentifikasi, langkah berikutnya adalah menentukan sentimen dari

setiap ulasan. Metode *Support Vector Machine* digunakan untuk mengetahui apakah pendapat bersifat positif atau negatif.

2.5. Evaluasi

Evaluasi pada penelitian ini menggunakan *Confusion Matrix*, sehingga nilai performansi dari sistem yang sudah dibuat berdasarkan hasil dari klasifikasi dapat diketahui.

Evaluasi kinerja model klasifikasi pada penelitian ini menghasilkan nilai perhitungan *accuracy*, *precision*, *recall*, dan *f1-score*. Berikut persamaan (1), (2), (3), (4), rumus perhitungan evaluasi kinerja model klasifikasi[17].

Accuracy merupakan prediksi benar dari *true positif* dan *true negatif*.

$$Accuracy = \frac{TP+TN}{(TP+FP+FN+TN)} \times 100\% \tag{1}$$

Precision merupakan nilai true positif dari nilai sebuah nilai positif.

$$Precision = \frac{TP}{TP+FP} \times 100\% \tag{2}$$

Recall merupakan presentase prediksi positif dengan *true positif*.

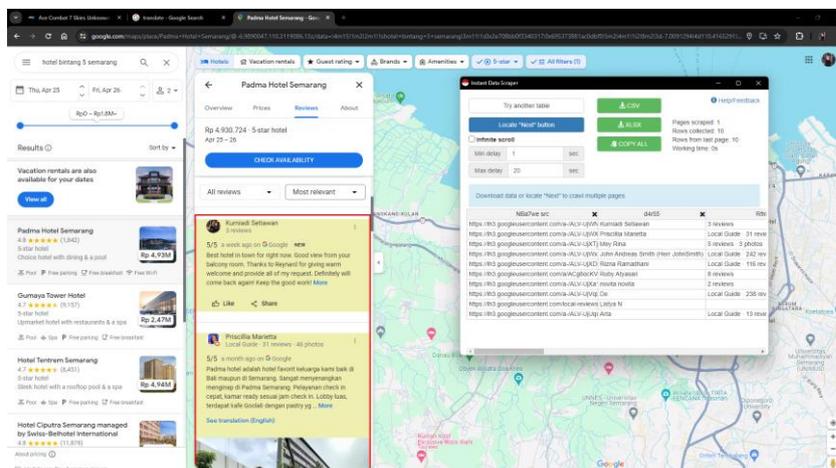
$$Recall = \frac{TP}{TP+FN} \times 100\% \tag{3}$$

F1-Score merupakan perbandingan rata-rata *precision* dan *recall*.

$$F1 - Score = \frac{2 \times precision \times recall}{precision + recall} \times 100\% \tag{4}$$

3. HASIL DAN PEMBAHASAN

Tahap pertama penelitian ini adalah pengambilan data, kami menggunakan data ulasan google map tentang hotel Bintang 5 yang ada di Semarang, kemudian kami jadikan 2 bagian: pertama data pelatihan (*training*) dan yang kedua data pengujian (*testing*).



Gambar 2. Proses *scraping data* dengan *instant web scraper*

Gambar 2 merupakan proses *scraping* data dengan melakukan *scraping* menggunakan ekstensi Chrome Instant Web Scraper, data awal yang diperoleh dari proses *scraping* mencakup sekitar 550 data per hotel.

Tabel 1. Sampel Data Ulasan Hotel Padma

index	Username	Rating	Reviews
0	Priscillia Marietta	5.0	Padma hotel adalah hotel favorit keluarga kami baik di Bali maupun di Semarang. Sangat menyenangkan menginap di Padma Semarang. Pelayanan check in cepat, kamar ready sesuai jam check in. Lobby luas, terdapat kafe Goolali dengan pastry yg ...
1	Mey Rina	5.0	Saya ingin membagikan pengalaman saya menginap 1244ersama keluarga di Hotel Padma Semarang. ...

Tabel 1 merupakan sampel data yang diperoleh dari hasil *scraping* menggunakan ekstensi *Chrome Instant Web Scraper* dari hotel Padma, selanjutnya akan diolah kembali pada tahap *pre-processing* dan klasifikasi sentimen.

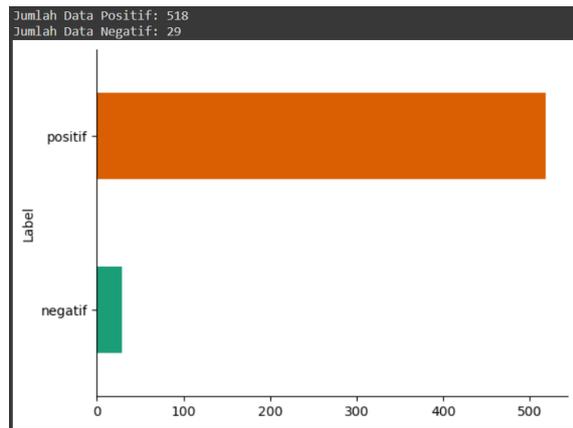
Tahap kedua merupakan *pre-processing data* melibatkan beberapa tahapan, termasuk *case folding*, *tokenizing*, *filtering*, dan *stemming*. Proses ini dilakukan menggunakan *Natural Language Toolkit* (NLTK) di Python. *Output* dari tahapan ini adalah data yang telah dipecah menjadi kata-kata dan diubah menjadi kata dasar.

Tabel 2. Sampel Data Clean Reviews Ulasan Hotel Padma

index	Username	Reviews	CleanReview
0	Priscillia Marietta	Padma hotel adalah hotel favorit keluarga kami baik di Bali maupun di Semarang. Sangat menyenangkan menginap di Padma Semarang. Pelayanan check in cepat, kamar ready sesuai jam check in. Lobby luas, terdapat kafe Goolali dengan pastry yg ...	Padma hotel favorit keluarga bali semarang menyenangkan menginap Padma semarang pelayanan check cepat kamar ready sesuai jam check lobby luas kafe goolali pastry
1	Mey Rina	Saya ingin membagikan pengalaman saya menginap 1244ersama keluarga di Hotel Padma Semarang. ...	membagikan pengalaman menginap keluarga hotel Padma semarang

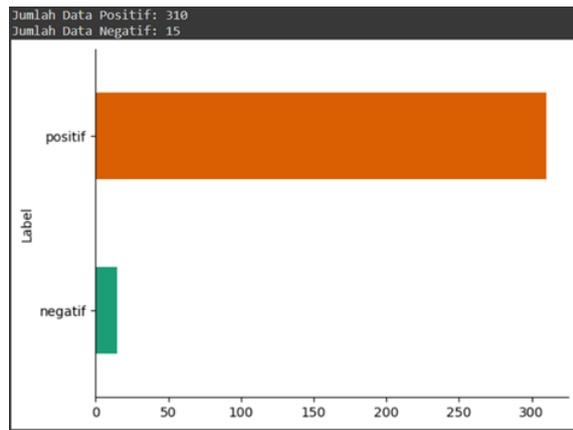
Tabel 2 merupakan sampel data hotel Padma yang telah menghasilkan tabel *clean review* yang merupakan hasil *case folding*, *tokenizing*, *filtering*, dan *stemming*.

Tahap ketiga adalah pemberian label, dalam penelitian ini, pelabelan dilakukan dengan mempertimbangkan ulasan dan rating pengguna di *Google Maps*. Proses ini mengkategorikan ulasan menjadi dua kelompok: positif untuk rating antara 3 hingga 5, dan negatif untuk rating antara 1 hingga 2. Berikut adalah hasil pelabelan untuk enam hotel bintang lima di Semarang:



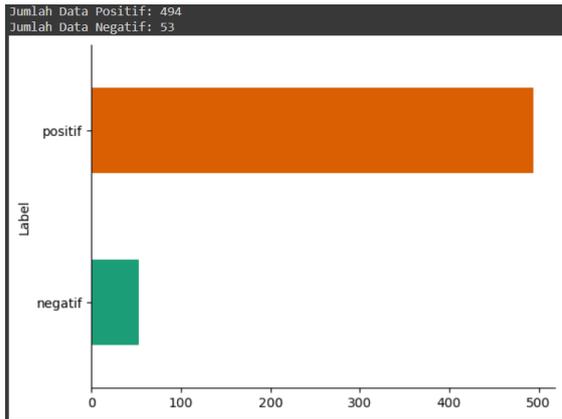
Gambar 2. Hasil pelabelan Hotel Padma

Gambar 2 menunjukkan hasil pelabelan dari Hotel Padma yang telah melewati proses *cleaning data* dengan *men-drop data kosong* dari tabel sehingga data yang terdeteksi hanya 547. Dari 547 data terdapat 518 data positif dan 29 negatif.



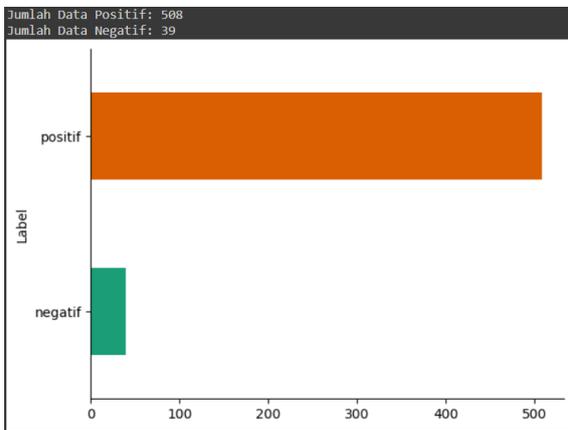
Gambar 3. Hasil pelabelan Hotel Gumaya

Gambar 3 menunjukkan hasil pelabelan dari Hotel Gumaya, data yang terdeteksi hanya 325 dan terbagi menjadi 310 data positif dan 15 data negatif, maka Hotel Gumaya memiliki ulasan yang positif.



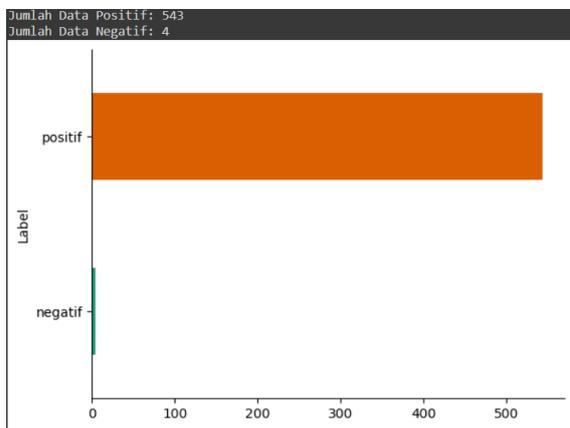
Gambar 4. Hasil pelabelan Hotel Tentrem

Gambar 4 menunjukkan hasil pelabelan dari Hotel Tentrem, data yang terdeteksi hanya 547 dan terbagi menjadi 494 data positif dan 53 data negatif, maka Hotel Tentrem memiliki ulasan yang positif.



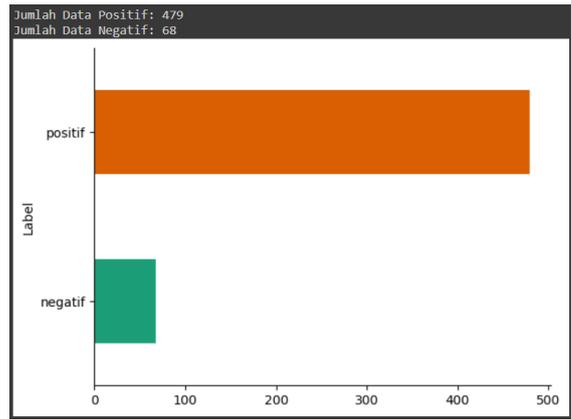
Gambar 5. Hasil pelabelan Hotel Grand Candi

Gambar 5 menunjukkan hasil pelabelan dari Hotel Grand Candi, data yang terdeteksi hanya 547 dan terbagi menjadi 508 data positif dan 39 data negatif, maka Hotel Tentrem memiliki ulasan yang positif.



Gambar 6. Hasil pelabelan Hotel Ciputra

Gambar 6 menunjukkan hasil pelabelan dari Hotel Ciputra, data yang terdeteksi hanya 547 dan terbagi menjadi 543 data positif dan 4 data negatif, maka Hotel Tentrem memiliki ulasan yang positif.



Gambar 7. Hasil pelabelan Hotel PO

Gambar 7 menunjukkan hasil pelabelan dari Hotel Ciputra, data yang terdeteksi hanya 547 dan terbagi menjadi 479 data positif dan 68 data negatif, maka Hotel Tentrem memiliki ulasan yang positif.

Tahapan berikutnya adalah proses SVM menggunakan parameter C dengan nilai 0.01, 0.05, 0.25, 0.5, 0.75, dan 1. Pada kasus ini kami memberikan sampel data pada Hotel Padma dengan akurasi berikut:

- Akurasi untuk $c = 0.01$: 0.97.
- Akurasi untuk $c = 0.05$: 0.97.
- Akurasi untuk $c = 0.25$: 0.97.
- Akurasi untuk $c = 0.5$: 0.97.
- Akurasi untuk $c = 0.75$: 0.97.
- Akurasi untuk $c = 1$: 0.97.

Tahapan terakhir akan diambil nilai C tertinggi untuk digunakan pada evaluasi kinerja SVM. Akurasi yang dihasilkan hotel Padma adalah sama maka kita bisa memilih salah satu akurasi, contoh pada hotel Padma adalah 0.5: 0.97 dengan rasio data training dan testing 80:20.

Tabel 3. Evaluasi SVM pada Hotel Padma

	Precision	Recall	f1-score	Support
negative	0.00	0.00	0.00	3
positive	0.97	1.00	0.99	107
accuracy			0.97	110
macro avg	0.49	0.50	0.49	110
weighted avg	0.95	0.97	0.96	110

Tabel 3 menunjukkan performa model klasifikasi. Model ini memiliki dua label: positif dan negatif. Untuk label positif, model ini memiliki nilai presisi 0.97, nilai recall 1.00, dan nilai F1-Score 0.99. Untuk label negatif, model ini memiliki nilai presisi 0.00, nilai recall 0.00, dan nilai F1-Score 0.00. Akurasi model ini secara keseluruhan adalah 0.97 dengan rasio data training dan data testing 80:20.

Kemudian berikut hasil parameter C pada hotel lainnya dengan nilai c tertinggi:

- Hotel Gumaya = 0.5: 0.97.
- Hotel Tentrem = 0.25: 0.90.
- Hotel Grand Candi = 0.25: 0.92.
- Hotel Ciputra = 0.5: 0.98.
- Hotel PO = 0.5: 0.87.

Berikut hasil evaluasi SVM pada hotel lainnya:

Tabel 4. Evaluasi SVM pada Hotel Gumaya

	Precision	Recall	f1-score	Support
negative	0.00	0.00	0.00	2
positive	0.97	1.00	0.98	63
accuracy			0.97	65
macro avg	0.48	0.50	0.49	65
weighted avg	0.94	0.97	0.95	65

Sama dengan tabel sebelumnya, tabel 4 menunjukkan hasil positif dan negatif. Untuk label positif, model ini memiliki nilai presisi 0.97, nilai recall 1.00, dan nilai F1-Score 0.98. Untuk label negatif, model ini memiliki presisi 0.00, recall 0.00, dan F1-Score 0.00. Akurasi model ini secara keseluruhan adalah 0.97 dengan rasio data testing 80:20.

Tabel 5. Evaluasi SVM pada Hotel Tentrem

	Precision	Recall	f1-score	Support
negative	0.50	0.10	0.17	10
positive	0.92	0.99	0.95	100
accuracy			0.91	110
macro avg	0.71	0.55	0.56	110
weighted avg	0.88	0.91	0.88	110

Untuk label positif pada tabel 5 memiliki nilai presisi 0.92, nilai recall 0.99, dan nilai F1-Score 0.95. dan label negatifnya memiliki nilai presisi 0.50, nilai recall 0.10, dan nilai F1-Score 0.17. Akurasi model ini secara keseluruhan adalah 0.91 dengan rasio data training dan data testing 80:20.

Tabel 6. Evaluasi SVM pada Hotel Grand Candi

	Precision	Recall	f1-score	Support
negative	0.00	0.00	0.00	8
positive	0.93	1.00	0.96	102
accuracy			0.93	110
macro avg	0.46	0.50	0.48	110
weighted avg	0.86	0.93	0.89	110

Pada tabel 6, label positif memiliki nilai presisi 0.93, nilai recall 1.00, dan nilai F1-Score 0.96. Untuk label negatifnya, memiliki nilai presisi 0.00, nilai recall 0.00, dan nilai F1-Score 0.00. Akurasi model ini secara keseluruhan adalah 0.93 dengan rasio data training dan data testing 80:20.

Tabel 7. Evaluasi SVM pada Hotel Ciputra

	Precision	Recall	f1-score	Support
negative	0.00	0.00	0.00	2
positive	0.98	1.00	0.99	108
accuracy			0.98	110
macro avg	0.49	0.50	0.50	110
weighted avg	0.96	0.98	0.97	110

Tabel 7 menunjukkan label positif memiliki nilai presisi 0.98, nilai recall 1.00, dan nilai F1-Score 0.99. Untuk label negatifnya memiliki nilai presisi 0.00, nilai recall 0.00, dan nilai F1-Score 0.00. Akurasi model ini secara keseluruhan adalah 0.98 dengan rasio data training dan data testing 80:20

Tabel 8. Evaluasi SVM pada Hotel PO

	Precision	Recall	f1-score	Support
negative	0.80	0.24	0.36	17
positive	0.88	0.99	0.93	93

accuracy			0.87	110
macro avg	0.84	0.61	0.65	110
weighted avg	0.86	0.87	0.84	110

Tabel 8 menunjukkan label positif memiliki nilai presisi 0.88, nilai recall 0.99, dan nilai F1-Score 0.93. Untuk label negatifnya memiliki nilai presisi 0.80, nilai recall 0.24, dan nilai F1-Score 0.36. Akurasi model ini secara keseluruhan adalah 0.87 dengan rasio data training dan data testing 80:20.

4. DISKUSI

Tabel 9 merupakan hasil keseluruhan perbandingan evaluasi terhadap hotel bintang 5 Semarang

Tabel 9. Keseluruhan evaluasi SVM hotel bintang 5 Semarang.

Hotel	Data Label	Positive	Negative	Accuracy
Padma	547	518	29	0.97
Gumaya	325	315	10	0.97
Tentrem	547	494	53	0.91
Grand Candi	547	508	39	0.93
Ciputra	547	543	4	0.98
PO	547	479	68	0.87

Data label telah mengalami pengapusan data kosong sehingga menjadi 547 dan 325. kemudian pelabelan data positif dan negatif berdasarkan rating dari para pengguna di google maps.

1. Hotel Padma memiliki 518 review positif dan 29 review negatif dengan accuracy 0.97.
2. Hotel Gumaya memiliki 315 review positif dan 10 review negatif dengan accuracy 0.97.
3. Hotel Tentrem memiliki 494 review positif dan 53 review negatif dengan accuracy 0.91.
4. Hotel Grand Candi memiliki 508 review positif dan 39 review negatif dengan accuracy 0.93.
5. Hotel Ciputra memiliki 543 review positif dan 4 review negatif dengan accuracy 0.98.
6. Hotel PO memiliki 479 review positif dan 68 review negatif dengan accuracy 0.87.

Dalam penelitian ini, hasil accuracy SVM masih mempunyai nilai diatas 80% sama halnya dengan penelitian-penelitian sebelumnya.

Hotel Ciputra memiliki data review positif paling banyak sebesar 543 data dengan accuracy sebesar 0.98.

5. KESIMPULAN

Kami melakukan kegiatan analisis sentimen terhadap 6 data hotel bintang 5 di Semarang menggunakan metode SVM dengan total data 550 dan setelah dilakukan cleaning, menyisakan 547 data valid pada data hotel Padma, Tentrem, Grand Candi, Ciputra, dan PO dan 325 pada data Hotel Gumaya.

Rasio data training dan data testing adalah 80:20, pada hotel Padma, Tentrem, Grand Candi, Ciputra dan PO melakukan testing pada 110 data dan Hotel Gumaya 65 data.

Hasil evaluasi SVM menunjukkan *accuracy* yang tinggi pada tiap hotel yakni, Padma, Gumaya,

Tentrem, Grand Candi, Ciputra dan PO dengan nilai 0,97, 0,97, 0,91, 0,93, 0,98, 0,87.

Hotel Ciputra menjadi hotel yang memiliki review paling baik diantara Hotel bintang 5 Semarang lainnya.

DAFTAR PUSTAKA

- [1] L. Sri Damayanti, "Peranan Keterampilan Berbahasa Inggris Dalam Industri Pariwisata," *Journey J. Tour. Culinary, Hosp. Conv. Event Manag.*, vol. 2, no. 1, pp. 71–82, Dec. 2019, doi: 10.46837/journey.v2i1.42.
- [2] E. K. Sugiyanto and F. Kurniasari, "DIMENSI KUALITAS PELAYANAN SEBAGAI UPAYA PENINGKATAN KEPUASAN PELANGGAN (Studi Pada Pelanggan Hotel X Semarang)," *Bus. Manag. Anal. J.*, vol. 3, no. 2, pp. 112–125, 2020, doi: 10.24176/bmaj.v3i2.5372.
- [3] J. Adinegara *et al.*, "Antecedents And Consequences Of Tourist Satisfaction: A Literature Review," *ASEAN Mark. J.*, vol. 9, no. 2, doi: 10.21002/amj.v9i2.5686.
- [4] J. A. Aryadi, Y. A. Aidil Basith, M. Munawir, and D. A. Rimadhani Agustini, "Analisis Data Review Hotel di Google Maps Melalui Text Mining (Studi Kasus: Kabupaten Bandung)," *JIKO (Jurnal Inform. dan Komputer)*, vol. 7, no. 2, p. 312, 2023, doi: 10.26798/jiko.v7i2.938.
- [5] R. Alhamdi, "PENGARUH ONLINE REVIEW DAN HARGA TERHADAP KEPUTUSAN PEMESANAN KAMAR HOTEL DI ONLINE TRAVEL AGENT (STUDI KASUS KOTA BATAM)," *J. Manaj. Perhotelan*, vol. 9, no. 2, pp. 63–70, Sep. 2023, doi: 10.9744/jmp.9.2.63-70.
- [6] M. F. Hattah, M. Asdar, and I. Nursyamsi, "Analisis Pengaruh Bauran Pemasaran terhadap Daya Tarik Pelanggan pada Hotel Claro di Kota Makassar," *SEIKO J. Manag. Bus.*, vol. 6, no. 1, pp. 208–224, 2023, doi: 10.37531/sejaman.v6i1.3845.
- [7] R. Fatmasari, V. M. Ayu, H. Anto, W. Gata, and L. D. Yulianto, "Analisis Sentimen Dalam Pengkategorian Komentar Youtube Terhadap Layanan Akademik dan Non-Akademik Universitas Terbuka Untuk Prediksi Kepuasan," *Build. Informatics, Technol. Sci.*, vol. 4, no. 2, pp. 395–404, Sep. 2022, doi: 10.47065/bits.v4i2.1738.
- [8] Y. A. Singgalen, "Analisis Sentimen Top 10 Traveler Ranked Hotel di Kota Makassar Menggunakan Algoritma Decision Tree dan Support Vector Machine," *Media Online*, vol. 4, no. 1, pp. 323–332, 2023, doi: 10.30865/klik.v4i1.1153.
- [9] M. A. Fauzi, "Word2Vec model for sentiment analysis of product reviews in Indonesian language," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 1, p. 525, Feb. 2019, doi: 10.11591/ijece.v9i1.pp525-530.
- [10] N. Pratiwi, J. Syahfitri, and M. Andesta, "PENYULUHAN SISTEM PERTANIAN TERPADU DAN PEMANFAATAN LAHAN KOSONG DI PEKARANGAN RUMAH BAGI MASYARAKAT DI KABUPATEN BENGKULU TENGAH Nurul," no. 2, pp. 69–73, 2021.
- [11] C. Prakoso and A. Hermawan, "Perbandingan Model Machine Learning dalam Analisis Sentimen Ulasan Pengunjung Keraton Yogyakarta pada Google Maps," *Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 3, pp. 1292–1302, 2023, doi: 10.30865/klik.v4i3.1419.
- [12] N. W. S. Saraswati and I. G. A. A. Diatri Indradewi, "Recognize The Polarity of Hotel Reviews using Support Vector Machine," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 22, no. 1, pp. 25–36, 2022, doi: 10.30812/matrik.v22i1.1848.
- [13] Y. A. Singgalen, "Analisis Performa Algoritma NBC, DT, SVM dalam Klasifikasi Data Ulasan Pengunjung Candi Borobudur Berbasis CRISP-DM," *Build. Informatics, Technol. Sci.*, vol. 4, no. 3, Dec. 2022, doi: 10.47065/bits.v4i3.2766.
- [14] Z. Zaenal and I. R. I. Astutik, "Sentiment Analysis of OYO App Reviews Using the Support Vector Machine Algorithm," *Procedia Eng. Life Sci.*, vol. 3, no. December, 2023, doi: 10.21070/pels.v3i0.1338.
- [15] C. Mei Serin Sitio, Y. Sibaroni, and S. Suryani Prasetiyowati, "Identifying Possible Rumor Spreaders on Twitter Using the Svm and Feature Level Extraction," *J. Tek. Inform.*, vol. 4, no. 3, pp. 611–618, 2023, doi: 10.52436/1.jutif.2023.4.3.868.
- [16] H. Mustofa and A. A. Mahfudh, "Klasifikasi Berita Hoax Dengan Menggunakan Metode Naive Bayes," *Walisongo J. Inf. Technol.*, vol. 1, no. 1, p. 1, 2019, doi: 10.21580/wjit.2019.1.1.3915.
- [17] K. Tri Putra, M. Amin Hariyadi, and C. Crysdiyan, "Perbandingan Feature Extraction Tf-Idf Dan Bow Untuk Analisis Sentimen Berbasis Svm," *J. Cahaya MAndalika*, p. 1449, 2023.