

REKOMENDASI PENGEMBANGAN FASILITAS WISATA TUGU PAHLAWAN SURABAYA MELALUI VISUALISASI *DASHBOARD* HASIL KLASIFIKASI ANALISIS SENTIMEN ULASAN PENGUNJUNG

Fawwaz Roja Mahardika^{1*}, Ahmad Afif Supianto², Nanang Yudi Setiawan³, Raden Sandra Yuwana⁴,
Endang Suryawati⁵

^{1,2,3}Universitas Brawijaya, Malang, ^{2,4,5}Pusat Riset Informatika, Badan Riset dan Inovasi Nasional, Bandung
Email: ¹fawwazrodika@gmail.com, ²afif.supianto@ub.ac.id / ahma083@lipi.go.id, ³nanang@ub.ac.id,
⁴rade014@lipi.go.id, ⁵enda029@lipi.go.id
*Penulis Korespondensi

(Naskah masuk: 07 Oktober 2021, diterima untuk diterbitkan: 17 Februari 2022)

Abstrak

Tugu Pahlawan Surabaya merupakan salah satu pariwisata andalan Kota Surabaya yang selalu berupaya memperhatikan ulasan pengunjung sebagai acuan evaluasi. Namun, pengelola tidak memiliki teknologi yang mampu mengumpulkan, mengolah, dan menganalisis seluruh data ulasan yang dapat menghasilkan informasi secara ringkas. Salah satu solusi dapat dilakukan melalui analisis sentimen pada level aspek terhadap aspek edukasi, fasilitas, kebersihan, pelayanan, dan umum dengan penyajian informasi dalam bentuk dashboard. Analisis sentimen dilakukan menggunakan *Support Vector Machine* terhadap 2180 data ulasan selama 2 tahun terakhir yang diambil dari *Google Review*. Ulasan terbanyak terdapat pada aspek fasilitas sebanyak 538 ulasan dengan sebaran sentimen 285 ulasan positif, 95 ulasan negatif, dan 158 ulasan netral. Rekomendasi berdasarkan kekuatan dan kelemahan saat ini adalah penyediaan lahan atau objek foto bernuansa sejarah pahlawan secara lebih nyata serta penyediaan ventilasi terbuka atau *standing cooler* di beberapa area. Berdasarkan *confusion matrix*, nilai F1-Score menjadi penentu seberapa baik model mengklasifikasi data daripada nilai *Accuracy* dikarenakan dataset yang dimiliki bersifat *imbalance* sehingga kesalahan prediksi pada *precision* atau *recall* sangat memungkinkan terjadi. Kesalahan prediksi banyak ditemukan pada kelas sentimen netral. Keseluruhan hasil klasifikasi disajikan dalam bentuk *dashboard* dengan nilai SUS Score 77,5, menandakan bahwa dashboard dapat diterima dengan baik oleh responden sebagai pengguna.

Kata kunci: Analisis Sentimen, Klasifikasi, SVM, Visualisasi Data, SUS

RECOMMENDATIONS FOR FACILITY DEVELOPMENT OF SURABAYA HERO MONUMENT THROUGH DASHBOARD VISUALIZATION OF CLASSIFICATION RESULTS ON VISITOR REVIEWS SENTIMENT ANALYSIS

Abstract

Tugu Pahlawan Surabaya is one of the mainstays of tourism in Surabaya city which always tries to pay attention to visitor reviews as a reference for evaluation. However, the managers do not have technology capable of collecting, processing and analyzing all review datas that can produce information in a concise manner. One solution can be done through sentiment analysis at the aspect level of education, facilities, cleanliness, service, and general aspects by presenting information in the form of a dashboard. Sentiment analysis was carried out using the *Support Vector Machine* on 2180 review datas for the last 2 years taken from *Google Reviews*. The most reviews were on the facility aspect in total of 538 reviews with a sentiment distribution of 285 positive reviews, 95 negative reviews and 158 neutral reviews. Recommendations based on current strengths and weaknesses are providing more real area or photo objects with historical nuances of heroes and providing open ventilation or *standing coolers* in several areas. Based on the *confusion matrix*, the F1-Score value determines how well the model classifies data rather than the *Accuracy* value because the dataset is *imbalance* so that prediction errors in *precision* or *recall* are very possible. Prediction errors are more likely to be found in the neutral sentiment class. The overall classification results are presented in the form of a dashboard with a SUS Score of 77.5, indicating that the dashboard is well received by respondents.

Keywords: Sentiment Analysis, Classification; SVM, Data Visualization, SUS

1. PENDAHULUAN

Berdasarkan Laporan Akhir: Kajian Dampak Sektor Pariwisata Terhadap Perekonomian Indonesia yang disusun oleh (Universitas Indonesia, 2018) serta hasil dari *World Economic Forum 2019* yang ditulis oleh (Calderwood & Soshkin, 2019), kepariwisataan Indonesia masih perlu dikembangkan lagi agar dapat mendorong perekonomian naik dan meningkatkan citra Indonesia. Kondisi tersebut mengakibatkan seluruh pemerintah tingkat daerah perlu melakukan pembenahan dalam bidang pariwisata setiap daerah guna memenuhi pencapaian kualitas pariwisata yang baik, tak terkecuali Pemerintah Kota Surabaya (Dinas Komunikasi dan Informatika Provinsi Jawa Timur, 2018).

Salah satu sektor pariwisata yang sangat melekat pada Surabaya ialah Tugu Pahlawan. Tugu Pahlawan Surabaya merupakan salah satu pariwisata andalan di bawah naungan Dinas Kebudayaan dan Pariwisata Kota Surabaya yang selalu berupaya untuk tetap memperhatikan ulasan pengunjung sebagai acuan evaluasi untuk memenuhi kepuasan dari perspektif pengunjung. (Dinas Kebudayaan dan Pariwisata Kota Surabaya, 2019). Namun hingga saat ini, pengelola kawasan wisata Tugu Pahlawan tidak memiliki teknologi yang dapat menyimpan dan menganalisis data yang mampu memberikan informasi tentang ulasan pengunjung. Belum terdapat visualisasi yang menggambarkan ulasan pengunjung wisata Tugu Pahlawan secara lebih ringkas guna mempermudah pengelola dalam menganalisis ulasan pengunjung.

Dari masalah yang dijelaskan sebelumnya, solusi yang dapat ditawarkan ialah menyajikan informasi ulasan pengunjung secara ringkas yang didapat dari internet khususnya Google Review melalui proses analisis sentimen. Analisis sentimen merupakan metode otomatis untuk menentukan apakah teks yang ada menyampaikan pandangan positif, negatif, atau umum dari suatu objek (Shathik & Karani, 2020). Hasil pengolahan dan analisis data divisualisasikan dalam bentuk yang ringkas dan mencukupi kebutuhan analisis agar dapat membantu dalam proses pengambilan keputusan strategis untuk pengelolaan ke depannya.

Penelitian terkait analisis sentimen pernah dilakukan sebelumnya dengan menggunakan *machine learning* oleh Al-Smadi et al., (2018). Penelitian tersebut membandingkan algoritme *machine learning Recurrent Neural Network (RNN)* dan *Support Vector Machine (SVM)* dengan berbagai Aspect-Based Sentiment Analysis (ABBA) *task*. Hasil penelitian menunjukkan bahwa pendekatan SVM mengungguli RNN dalam berbagai *task* seperti *aspect category identification*, *Opinion Target Extraction (OTE)*, dan *aspect sentiment polarity identification*. Penelitian *text classification* juga pernah dilakukan oleh Sulaeman et al. (2019), Salam et al. (2019), dan Vitandy et al. (2019)

mengenai klasifikasi sentimen saran masuk dari mahasiswa berdasarkan evaluasi mengajar dosen yang ada di Fakultas Ilmu Komputer, Universitas Brawijaya. Ketiganya menggunakan algoritma yang berbeda dan tentunya menghasilkan akurasi yang berbeda. Sulaeman et al. (2019) menggunakan SVM dan mendapatkan hasil akurasi tertinggi sebesar 82% dibandingkan dengan dua lainnya yaitu oleh Vitandy et al. (2019) menggunakan Naïve Bayes dengan akurasi 80,1% dan Salam et al. (2019) menggunakan K-Nearest Neighbour dengan akurasi sebesar 80%.

Keunggulan algoritme SVM dalam proses *machine learning* untuk kasus analisis sentimen tersebut didukung oleh penggunaan TF-IDF menjadi alasan pada penelitian ini untuk menerapkan algoritme SVM pula. Pemilihan SVM sebagai model klasifikasi juga turut dilandasi oleh kemampuannya dalam penerapan *Structural Risk Minimization (SRM)* dengan margin *hyperplane* optimal, sehingga dapat menjamin kesalahan klasifikasi yang rendah (Shafiabady, et al., 2016).

Analisis sentimen sendiri banyak menerapkan pengklasifikasian teks untuk mengklasifikasikan ribuan data komentar yang dibagi berdasarkan positif dan negatif dengan tambahan netral yang berarti tidak mengandung nilai positif maupun negatif. Hasil dari proses analisis sentimen banyak disajikan dalam *visual dashboard analytic* yang mempermudah pengguna data dalam mengolahnya menjadi informasi untuk mempermudah *stakeholder* memahami hasil yang diberikan. Contoh penelitian yang menerapkannya yakni penelitian oleh Pratama, Bachtiar, dan Setiawan (2018) yang juga menggunakan SVM dengan hasil *Accuracy* sebesar 85%, serta penelitian oleh Ramadhani dan Fajarianto (2020) dimana keduanya mengaplikasikan hasil dari analisis sentimen ke dalam *visual dashboard analytic*. Hasil pengolahan data menjadi informasi dapat dengan mudah dibaca oleh *stakeholder* ketika disajikan dalam *dashboard* dengan tampilan yang sesuai. Hal ini juga menjadi alasan untuk menyajikan hasil klasifikasi dalam bentuk *dashboard* pada penelitian kali ini.

Dashboard yang akan dibangun pada penelitian ini akan menyediakan tampilan pengurutan atau *rank* komentar berdasarkan ulasan yang paling sering menjadi perhatian pengunjung. Hal ini dapat membantu pengguna *dashboard* untuk mengetahui permasalahan yang sering muncul dan juga apa yang paling menarik perhatian pengunjung di Tugu Pahlawan. Pengurutan *rank* ulasan ini diterapkan ke masing-masing sentimen yang dapat di-*filter* berdasarkan aspek, sehingga pengguna dapat dengan mudah mencari dan menganalisis sesuai kebutuhan.

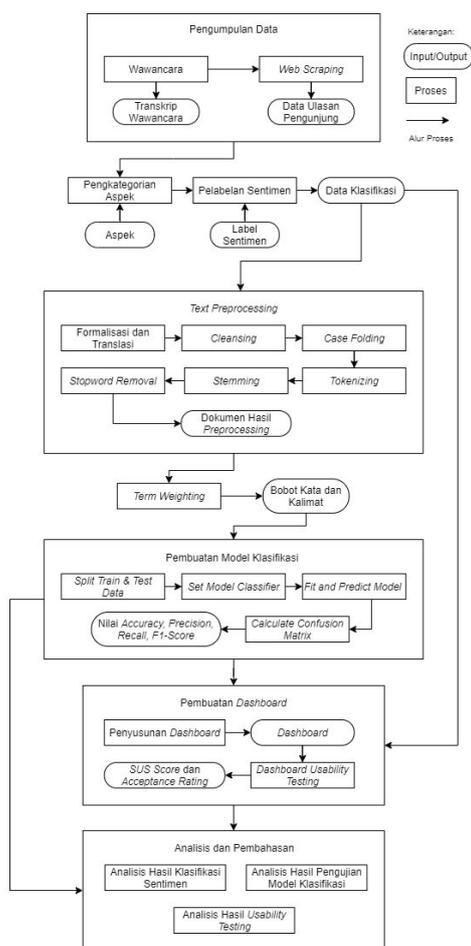
Berdasarkan latar belakang yang telah dijelaskan serta beberapa penelitian sejenis yang pernah dilakukan sebelumnya, terdapat beberapa rumusan masalah seperti (R1) bagaimana rekomendasi yang dapat diberikan untuk

pengembangan Tugu Pahlawan, (R2) bagaimana hasil pengukuran performa pengujian model klasifikasi analisis sentimen menggunakan SVM, dan (R3) bagaimana *dashboard* klasifikasi sentimen yang dibangun dapat diterima oleh pengguna berdasarkan hasil kuesioner *SUS*.

Penelitian kali ini akan membahas analisis sentimen dengan memanfaatkan data ulasan *online* dari pengunjung terhadap Tugu Pahlawan Surabaya dari Google Review yang didapatkan menggunakan teknik *web scraping*. Analisis sentimen dilakukan pada level aspek yang terbagi menjadi aspek edukasi, fasilitas, kebersihan, pelayanan, dan umum menggunakan metode klasifikasi SVM yang dikombinasikan dengan *term weighting* menggunakan TF-IDF. Hasil analisis akan divisualisasikan dalam bentuk *dashboard* menggunakan Google Data Studio guna memudahkan pengguna membaca hasil analisis yang dapat menentukan keputusan strategis kedepannya.

2. METODE PENELITIAN

Alur penelitian analisis sentimen pada Tugu Pahlawan Surabaya terdiri dari beberapa proses sesuai pada Gambar 1.



Gambar 1. Metode penelitian

2.1. Pengumpulan Data

Pengumpulan data ulasan pengunjung dilakukan melalui dua proses utama yaitu wawancara langsung dengan pihak pengelola Tugu Pahlawan dan akuisisi data ulasan pengunjung melalui Google Review terhadap data ulasan terbaru untuk seluruh bahasa dari Oktober 2018 hingga Oktober 2020. Proses akuisisi data dilakukan dengan *web scraping* agar dapat memilih elemen data pada halaman tersebut, lalu mengekstraksi dan mentransformasi data, yang kemudian menyimpan data yang diperoleh menjadi dataset terstruktur (Boeing & Waddell, 2016). *Web scraping* dilakukan menggunakan Puppeteer dengan hasil sebanyak 1714 data ulasan yang disimpan dalam *file CSV*.

2.2. Pengkategorian Aspek

Keseluruhan data ulasan akan dibagi ke dalam aspek edukasi, fasilitas, kebersihan, pelayanan, dan umum secara manual. Aspek-aspek tersebut disesuaikan berdasarkan aspek yang menjadi penilaian penting bagi pengelola Tugu Pahlawan yang telah ditetapkan sejauh ini yang mengacu pada perspektif pengunjung. Aspek yang diberikan pada satu data ulasan dapat lebih dari satu aspek. Sebagian dari pengkategorian aspek terdapat pada Tabel 1.

Tabel 1. Pengkategorian aspek

Ulasan	Aspek
Sebagai monumen kepahlawanan yang sangat membanggakan	Umum
Suka ke tempat ini untuk mengenang jasa para pahlawan dan mendidik anak-anak supaya lebih memahami tentang perjuangan para pahlawan dalam melawan penjajah	Edukasi
Tidak ada tempat parkirnya, bingung cari tempat parkir, karena saya bukan orang situ	Fasilitas
Toilet kotor, baunya tidak enak	Kebersihan
Pelayanannya bagus, pemandu wisatanya baik banget dan ramah	Pelayanan

2.3. Pelabelan Data

Seluruh data ulasan akan diberi label sentimen berupa positif atau negatif. Jika dalam suatu kalimat tidak mengandung opini positif atau negatif, maka kalimat opini tersebut dinyatakan netral (Kharde & Sonawane, 2016). Proses dilakukan dengan mengecek satu per satu data ulasan yang ada dan kemudian memberikan pelabelan sentimen secara manual pada data ulasan tersebut. Jumlah data ulasan bertambah menjadi 2180 data ulasan karena setiap terdapat ulasan yang memiliki lebih dari satu aspek dan/atau sentimen akan diduplikasi pada baris data baru dengan pemberian aspek dan/atau sentimen yang berbeda tergantung dari konteks ulasannya. Tabel 2 menunjukkan sebagian pelabelan sentiment yang dilakukan.

Tabel 2. Pelabelan sentimen

Ulasan	Aspek	Label
Sebagai monumen kepahlawanan yang sangat membanggakan	Umum	Netral
Suka ke tempat ini untuk mengenang jasa para pahlawan dan mendidik anak-anak supaya lebih memahami tentang perjuangan para pahlawan dalam melawan penjajah	Edukasi	Positif
Tidak ada tempat parkirnya, bingung cari tempat parkir, karena saya bukan orang situ	Fasilitas	Negatif
Toilet kotor, baunya tidak enak	Kebersihan	Negatif
Pelayanannya bagus, pemandu wisatanya baik banget dan ramah	Pelayanan	Positif

2.4. Text Mining & Preprocessing

Text mining merupakan proses menggali dan menemukan pengetahuan melalui kumpulan dokumen yang ada dari waktu ke waktu dengan menggunakan seperangkat alat analisis. Text mining banyak digunakan untuk menganalisis informasi lebih lanjut sebagai pendukung dalam pengambilan keputusan strategis kedepannya (Aggarwal & Zhai, 2012). Text preprocessing dilakukan menggunakan bahasa pemrograman Python untuk mengkondisikan data ulasan agar sesuai dengan kebutuhan penelitian. Tahap yang dilakukan antara lain melakukan formalisasi dan translasi kata secara manual, kemudian cleansing, case folding, tokenizing, stemming, dan stopword removal. Pada proses ini memanfaatkan library seperti pandas, NLTK, dan Sastrawi yang ada di Python. Proses text preprocessing terdapat pada Gambar 2 berikut.

Ulasan	Formal-Trans	Cleansing	Case Folding	Tokenizing	Stemming	Stopword Removal
Guide oke, yg pasti puas	Pemandu oke, yang pasti puas	Pemandu oke yang pasti puas	pemandu oke yang pasti puas	['pemandu', 'oke', 'yang', 'pasti', 'puas']	['pandu', 'oke', 'yang', 'pasti', 'puas']	['pandu', 'oke', 'pasti', 'puas']

Gambar 2. Text preprocessing

2.5. Term Weighting

Tahap term weighting atau pembobotan dilakukan menggunakan dokumen berisikan data teks hasil dari text preprocessing sebelumnya. Proses ini menerapkan Term Frequency-Inverse Document Frequency (TF-IDF) dengan memanfaatkan modul Scikit-learn pada Python. Pembobotan bertujuan untuk memberi nilai bobot pada masing-masing kata yang menjadi variabel penting dalam pengujian model klasifikasi agar setiap kata memiliki nilainya masing-masing. TF-IDF cocok digunakan dalam mengklasifikasi banyak kelas karena TF-IDF generate pada satu unique value untuk setiap kata yang unik (Dogan & Uysal, 2019). Sebagian hasil perhitungan TF-IDF dapat dilihat pada Tabel 3.

Tabel 3. Hasil TF-IDF

term (t)	TF-IDF				
	D1	D2	D3	D4	D5
jalan	0,471674	0	0	0	0,406147

term (t)	TF-IDF				
	D1	D2	D3	D4	D5
keluarga	0	0,187028	0	0	0,135382
nikmat	0	0	0	0	0,167803
panas	0	0,187028	0	0	0,135382
sejarah	0	0	0	0,269639	0,135382

2.6. Classification Implementation

Proses klasifikasi terdiri dari beberapa tahapan yakni split train & test data, setting model classifier, fitting model with train data, predicting test data, dan calculate confusion matrix. Model klasifikasi dibangun menggunakan library Support Vector Machine (SVM) dengan algoritma LinearSVC. SVM merupakan supervised learning model dengan asosiasi learning algorithms yang mampu menganalisa data dan menemukan pola, yang digunakan untuk menganalisa regresi dan klasifikasi (Frunza, 2016).

Validasi data pada proses pelatihan klasifikasi dilakukan menggunakan GridSearchCV yang juga menerapkan cross validation sehingga mampu mengalokasikan metode "fit" untuk menentukan hyperparameter terbaik dan "score" untuk memperoleh hasil akurasi terbaik (Scikit-learn, 2020). Output dari tahapan ini adalah mendapatkan hasil nilai berupa Accuracy, Precision, Recall, dan F1-Score sebagai nilai validasi performa dan keakuratan model klasifikasi yang digunakan. Setiap aspek akan mendapatkan confusion matrix dan nilai performanya masing-masing. Hasil yang terdapat pada Gambar 3 merupakan hasil pengujian klasifikasi sentimen terhadap dokumen aspek fasilitas.

```
Result for testing data (predicting)
Confusion Matrix
[[12  5  5]
 [ 0 31  9]
 [ 1  7 65]]
```

Classification Report	precision	recall	f1-score	support
negatif	0.92	0.55	0.69	22
netral	0.72	0.78	0.75	40
positif	0.82	0.89	0.86	73
accuracy			0.80	135
macro avg	0.82	0.74	0.76	135
weighted avg	0.81	0.80	0.80	135

Gambar 3. Output hasil pengujian klasifikasi

2.7. Pembuatan Dashboard

Dashboard merupakan suatu media yang dapat menyajikan informasi secara visual melalui tabel dan indikator visual lainnya yang dapat memudahkan pengguna dalam membaca, mengawasi serta mengelola bisnisnya (Wajong, 2015). Pembuatan dashboard dilakukan dengan memvisualisasikan hasil klasifikasi analisis sentimen yang menampilkan beberapa informasi yang diharapkan dapat berguna bagi pemangku kepentingan sebagai informasi pendukung untuk pengambilan keputusan strategis ke depannya. Kemudian, dashboard akan diuji menggunakan

usability testing sebagai bentuk validasi kesesuaian *dashboard* terhadap kebutuhan pengguna. Pembuatan *dashboard* dilakukan dengan menggunakan Google Data Studio dengan sumber data berupa *spreadsheets*.

Pengujian *usability testing* menggunakan *System Usability Scale* (SUS) yang merupakan skala survei yang dikembangkan oleh Brooke pada tahun 1996. SUS banyak digunakan oleh peneliti untuk menilai *usability* dari suatu layanan atau produk di beberapa bidang, seperti misalnya pada bidang pendidikan (Lia, 2021) dan pariwisata (Pratama, 2018). Pada instrumen SUS asli yang dirancang, terdapat 10 pernyataan yang masing-masing memiliki nilai skala 5 yang menunjukkan derajat persetujuan berdasarkan penilaian dari responden. Hasil akhir dari SUS memiliki nilai dengan rentang 0-100. Semakin tinggi nilainya, maka semakin baik pula *usability*-nya.

SUS pernah diadaptasi ke dalam versi Bahasa Indonesia oleh Sharfina dan Santoso (2016) melalui penelitian yang mendefinisikan *usability* sebagai kualitas pengalaman pengguna ketika berinteraksi dengan suatu sistem atau produk yang menilai seberapa mudah antarmuka pengguna digunakan. SUS dapat diproses melalui adaptasi lintas budaya untuk memastikan bahwa SUS versi Bahasa Indonesia memiliki status valid sehingga dapat digunakan dalam budaya dan populasi yang berbeda.

2.8. Analisis dan Pembahasan

Analisis dan pembahasan akan membahas implementasi yang telah dilakukan terkait hasil klasifikasi sentimen, hasil pengujian model klasifikasi, serta hasil *usability testing* terhadap dashboard. Dengan melakukan analisis dan pembahasan secara lebih mendalam, diharapkan dapat mengetahui dan menjelaskan temuan-temuan penelitian terkait hasil, solusi, rekomendasi, dan juga kendala yang ditemukan selama penelitian dilakukan.

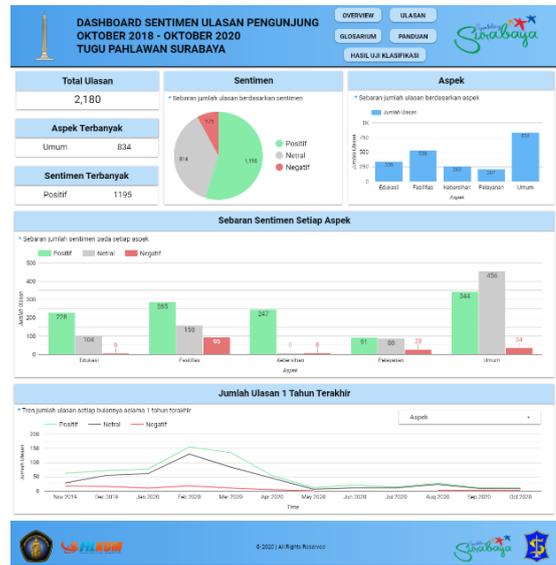
3. HASIL DAN PEMBAHASAN

3.1. Hasil Implementasi Dashboard

Dashboard disusun menjadi 3 halaman utama yang menyajikan kebutuhan analisis rekomendasi serta 2 halaman sebagai pendukung dan bantuan. Halaman pendukung merupakan Glosarium dan Panduan yang berisikan catatan yang dapat dibaca oleh pengguna sebagai bentuk pembelajaran mandiri untuk lebih memahami pemakaian dashboard.

Halaman pertama dari tiga halaman utama pada *dashboard* merupakan halaman *Overview* yang berisikan pemusatan dan penyederhanaan informasi hasil pengkategorian aspek dan klasifikasi sentimen ulasan pengunjung. Komponen-komponen informasi yang disajikan pada halaman ini mewakili informasi ulasan pengunjung berdasarkan kategori aspek dan sentimen ulasan pada setiap aspek, serta analisis tren

dalam satu tahun terakhir. Tampilan halaman pertama dapat dilihat pada Gambar 4.



Gambar 4. Dashboard halaman 'Overview'

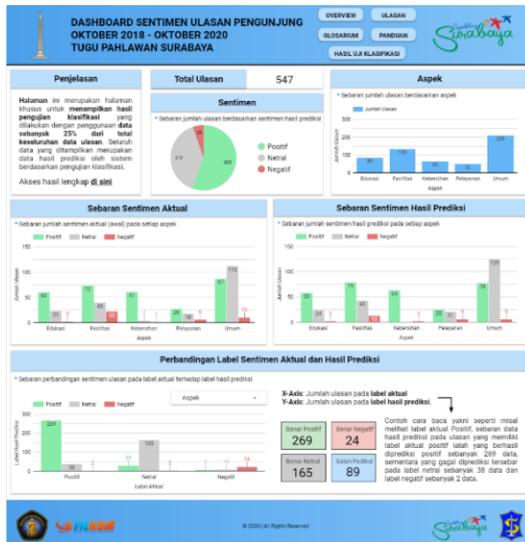
Halaman kedua dari dua halaman utama pada dashboard merupakan halaman *Ulasan* yang berisikan informasi terkait dua komponen utama. Terdapat bagian kata terbanyak yang sering digunakan oleh pengunjung dalam memberikan ulasan. Selain itu terdapat pengurutan ulasan berdasarkan konteks yang paling sering disampaikan oleh pemberi ulasan. Tampilan dari halaman kedua dapat dilihat pada Gambar 5.



Gambar 5. Dashboard halaman 'Ulasan'

Halaman ketiga pada *dashboard* merupakan halaman Hasil Uji Klasifikasi yang berisikan informasi hasil pengujian klasifikasi terhadap sentimen ulasan pengunjung. Data yang digunakan berjumlah 25% dari total keseluruhan data yang dimanfaatkan pada proses pengujian klasifikasi. Hasil klasifikasi sentimen yang ditampilkan pada

Tampilan dari halaman ketiga dapat dilihat pada Gambar 6.



Gambar 6. Dashboard halaman 'Hasil Uji Klasifikasi'

3.2. Analisis Hasil Klasifikasi Sentimen

Data yang digunakan sebagai analisis lebih lanjut untuk memberikan rekomendasi adalah data awal dimana klasifikasi yang digunakan adalah klasifikasi hasil dari pelabelan manual. Jumlah data ulasan yang diklasifikasikan berdasarkan aspek dan sentimennya berjumlah 2180 data ke dalam 3 sentimen yaitu 1195 ulasan positif, 171 ulasan negatif, dan 814 netral. Data ulasan tersebut telah dibagi menjadi 5 kategori aspek yakni aspek edukasi, fasilitas, kebersihan, pelayanan, dan umum.

Masing-masing aspek memiliki kekuatan dan kelemahan yang diperoleh berdasarkan hasil kata terbanyak beserta dengan konteks ulasan yang sering menjadi perhatian pengunjung dalam memberi ulasan. Setiap kekuatan dan kelemahan pada masing-masing aspek dapat menjadi acuan untuk memberikan rekomendasi pengembangan Tugu Pahlawan guna menjawab rumusan masalah pertama (R1).

3.2.1. Aspek Edukasi

Aspek edukasi menjelaskan nilai pembelajaran yang dapat diperoleh saat berkunjung di Tugu Pahlawan. Hasil klasifikasi sentimen dari 338 data ulasan pada aspek edukasi tersebar menjadi 228 positif, 104 netral, dan 6 negatif. Rekomendasi yang dapat diberikan adalah perlunya menyediakan media pembelajaran sejarah perjuangan pahlawan yang lebih interaktif dan bersifat dua arah, sehingga dapat menumbuhkan ketertarikan dan menghilangkan rasa mengantuk dan bosan. Perlu ditekankan bahwa jika dilihat berdasarkan kata terbanyak, pasar yang paling menjadi perhatian ialah anak-anak.

3.2.2. Aspek Fasilitas

Aspek fasilitas menjelaskan berbagai sarana dan prasarana yang terdapat pada Tugu Pahlawan.

Hasil klasifikasi sentimen dari 538 data ulasan pada aspek fasilitas tersebar menjadi 285 positif, 158 netral, dan 95 negatif. Rekomendasi yang dapat diberikan adalah perlu memperhatikan bahwa spot foto menjadi kunci utama dalam kepuasan pengunjung, sehingga perlu menyediakan lahan atau objek foto yang lebih menarik. Hal tersebut juga harus didukung dengan menyediakan lingkungan yang nyaman bagi pengunjung dengan memberi tindak lanjut terhadap keluhan rasa kepanasan pengunjung yang diterima. Pengoptimalan penggunaan AC perlu dipertimbangkan.

3.2.3. Aspek Kebersihan

Aspek kebersihan menjelaskan kondisi lingkungan Tugu Pahlawan Surabaya terkait kebersihan dan keasriannya yang dapat memberikan rasa nyaman bagi pengunjung atas bersihnya lingkungan tersebut. Hasil klasifikasi sentimen dari 263 data ulasan pada aspek kebersihan tersebar menjadi 247 positif, 8 netral, dan 8 negatif. Rekomendasi yang dapat diberikan adalah perlunya mempertahankan kebersihan dan keasrian lingkungan yang sudah ada saat ini, terutama pada kebersihan di area toilet dan museum. Beberapa hal yang dapat dilakukan yakni memberikan wangi-wangian di beberapa tempat khususnya toilet, memperketat pengawasan agar pengunjung tidak membuang sampah sembarangan, dan penyediaan tong sampah yang lebih banyak.

3.2.4. Aspek Pelayanan

Aspek pelayanan menjelaskan terkait layanan yang disediakan dan diberikan kepada pengunjung serta diimplementasikan oleh orang-orang yang bertugas di Tugu Pahlawan Surabaya. Hasil klasifikasi sentimen dari 207 data ulasan pada aspek pelayanan tersebar menjadi 91 positif, 88 netral, dan 28 negatif. Rekomendasi yang dapat diberikan adalah perlunya mempertahankan kebijakan yang ada terkait harga tiket masuk museum yang murah serta adanya akses gratis ketika memasuki area tugu pahlawan selain museumnya. Hal tersebut perlu diimbangi dengan tindak lanjut terhadap segala keluhan pengunjung, seperti pada jam operasional untuk mengakses wisata Tugu Pahlawan Surabaya.

3.2.5. Aspek Umum

Aspek umum menjelaskan hal-hal yang kurang spesifik atau tidak bisa dikategorikan ke dalam 4 aspek lainnya. Hasil klasifikasi sentimen dari 834 data ulasan pada aspek umum tersebar menjadi 344 positif, 456 netral, dan 34 negatif. Dalam kasus aspek umum, meskipun ulasan dengan sentimen netral pada aspek umum lebih banyak dari sentimen positif dan negatif, namun tidak banyak yang bisa dianalisis secara lebih lanjut karena informasi yang kurang spesifik. Selain itu, aspek umum juga tidak memiliki fokus kategori aspek secara khusus sehingga segala ulasan pada aspek ini bisa dijadikan

sebagai data tambahan yang bersifat opsional dan tidak mendesak.

3.3. Analisis Hasil Pengujian Model

Hasil pengujian klasifikasi sentimen didapatkan dari proses klasifikasi yang meliputi model (*data*) *training* dan model (*data*) *predicting* menggunakan model SVM dengan algoritme LinearSVC. Pengujian klasifikasi dilakukan pada sebanyak 25% data ulasan berjumlah 547 ulasan. Hasil pengujian pada aspek edukasi mengklasifikasikan sentimen menjadi 59 positif, 24 netral, dan 2 negatif dengan kesalahan prediksi 17 ulasan. Pada aspek fasilitas menjadi 79 positif, 43 netral, dan 13 negatif dengan kesalahan prediksi 27 ulasan. Pada aspek kebersihan menjadi 64 positif dan 2 negatif dengan kesalahan prediksi 4 ulasan. Pada aspek pelayanan menjadi 25 positif, 21 netral, dan 6 negatif dengan kesalahan prediksi 7 ulasan. Sementara pada aspek umum menjadi 78 positif, 125 netral, dan 6 negatif dengan kesalahan prediksi 34 ulasan. Tabel 4 menunjukkan hasil *confusion matrix* pada pengujian klasifikasi.

Tabel 4. Hasil confusion matrix keseluruhan data

Nilai Label Prediksi	Nilai Label Aktual			Jumlah Data
	Negatif	Netral	Positif	
Negatif	24	3	2	29
Netral	10	165	38	213
Positif	7	29	269	305
Total Data Uji				547

Kesalahan prediksi banyak terjadi pada ulasan netral, yang sebenarnya juga banyak bertukar kesalahan dengan ulasan positif. Berdasarkan analisis terhadap data ulasan yang mengalami kesalahan prediksi sentimen, hal yang dapat menjadi penyebab adalah *dataset* yang *imbalance* dengan sebaran sentimen positif jauh lebih banyak daripada sentimen netral dan negatif. Model klasifikasi akan cenderung lebih banyak melatih kelas label sentimen positif. Oleh karena itu, sangat memungkinkan bahwa model akan salah memprediksi karena kekuatan nilai pembobotan yang dihasilkan oleh label positif akan cenderung lebih besar daripada label netral dan negatif.

Faktor kedua adalah kelas label netral yang menjadi suatu kebingungan tersendiri karena sentimen netral tidak memiliki fokus sentimen yang jelas. Pada akhirnya bisa saja label tersebut salah prediksi, bahkan untuk dikategorikan sebagai label positif ataupun negatif. Hal ini disebabkan oleh kemampuan model klasifikasi yang berfokus pada *term-term* yang terdapat pada suatu ulasan yang dibandingkan terhadap keseluruhan ulasan untuk melatih dan menguji atau memprediksi klasifikasi.

Untuk memahami hasil pengujian model secara lebih lanjut, analisis terhadap nilai *Precision*, *Recall*, *F1-Score*, dan *Accuracy* berdasarkan *classification report* juga diperlukan. Tabel 5 menunjukkan hasil

classification report dari pengujian klasifikasi yang dilakukan.

Tabel 5. Hasil *classification report*

Aspek	Sentimen	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Accuracy</i>
Edukasi	Negatif	0.86	0.85	0.86	0.80
	Netral	0.62	0.65	0.64	
	Positif	1.00	1.00	1.00	
	Mean	0.83	0.83	0.83	
Fasilitas	Negatif	0.82	0.89	0.86	0.80
	Netral	0.72	0.78	0.75	
	Positif	0.92	0.55	0.69	
	Mean	0.82	0.74	0.77	
Kebersihan	Negatif	0.95	1.00	0.98	0.94
	Netral	0.00	0.00	0.00	
	Positif	0.50	1.00	0.67	
	Mean	0.48	0.67	0.55	
Pelayanan	Negatif	0.96	0.86	0.91	0.87
	Netral	0.76	0.89	0.82	
	Positif	0.83	0.83	0.83	
	Mean	0.85	0.86	0.85	
Umum (Lainnya)	Negatif	0.92	0.86	0.89	0.86
	Netral	0.79	0.90	0.84	
	Positif	0.92	0.71	0.80	
	Mean	0.88	0.82	0.84	

Dapat dilihat pada hasil dari perhitungan nilai *Precision*, *Recall*, *F1-Score*, dan *Accuracy* pada seluruh aspek bahwa keseluruhan aspek memiliki nilai *Accuracy* yang cukup tinggi. Sementara aspek kebersihan merupakan aspek dengan nilai *Accuracy* tertinggi sebesar 0.94 yang jauh lebih besar dibandingkan aspek lainnya. Jika dibandingkan dengan nilai *F1-Score* aspek kebersihan yang hanya menyentuh 0.55, model pengujian klasifikasi terhadap aspek kebersihan belum tentu dapat dikatakan baik jika hanya menilai dari hasil *Accuracy* saja.

Jika memperhatikan nilai-nilai *Precision*, *Recall*, dan *F1-Score* yang dihasilkan oleh kelas label netral pada setiap aspeknya, nilai-nilai tersebut lebih kecil daripada nilai kelas label positif dan negatif. Kondisi tersebut memperkuat fakta bahwa model klasifikasi yang digunakan belum bisa mendefinisikan kelas label netral secara pasti karena adanya kebingungan dalam menentukan *terms* apa saja yang dapat dikategorikan sebagai netral.

Pada pembobotan juga cenderung sulit menentukan apakah data ulasan netral tersebut pantas diprediksi sebagai data netral atau harus masuk ke data ulasan positif atau bahkan negatif. Pada dasarnya, kelas label netral juga cenderung tidak menunjukkan adanya emosi atau sentimen secara jelas sehingga ketika diimplementasikan dalam ilmu *machine learning* dapat berpotensi mengganggu pengklasifikasian sentimen positif dan negatif. Seluruh hasil dari pengujian baik berupa *confusion matrix*, *classification report*, dan analisis kesalahan pengujian dapat menjawab rumusan masalah kedua (R2).

3.4. Analisis Hasil Usability Testing Dashboard

Usability testing yang dilakukan pada *dashboard* dilakukan untuk mengetahui tingkat kemudahan pengguna ketika menggunakan *dashboard* dan seberapa diterima dan bermanfaat *dashboard* bagi pengguna. Seluruh tampilan *dashboard* tersebut diujikan pada *Usability Testing* dengan *output* keluaran nilai akhir SUS *Score*. Nilai akhir tersebut diperoleh dari hasil pengisian kuesioner SUS. Tabel 6 menunjukkan pertanyaan pada kuesioner SUS, sementara Tabel 7 menunjukkan hasil kuesioner SUS yang diisi oleh responden pengguna *dashboard*.

Tabel 6. Pernyataan SUS

Nomor Pernyataan	Pernyataan
1	Saya berpikir bahwa saya akan sering menggunakan <i>dashboard</i> ini
2	Saya merasa penggunaan <i>dashboard</i> ini kompleks
3	Saya merasa <i>dashboard</i> ini mudah untuk digunakan
4	Saya membutuhkan bantuan teknisi atau orang lain untuk menggunakan <i>dashboard</i> ini
5	Saya merasa bahwa fitur yang ada dalam <i>dashboard</i> ini terintegrasi dengan baik
6	Saya merasa terdapat banyak hal yang tidak konsisten dalam <i>dashboard</i> ini
7	Saya merasa bahwa orang lain akan mudah memahami dan menggunakan <i>dashboard</i> ini
8	Saya merasa <i>dashboard</i> ini membingungkan
9	Saya merasa tidak terdapat hambatan dalam menggunakan <i>dashboard</i> ini
10	Saya perlu belajar dan membiasakan diri sebelum saya menggunakan <i>dashboard</i> ini

Tabel 7. Hasil Kuesioner SUS

Nomor Pernyataan	Nilai	Perhitungan	Hasil
1	4	4-1	3
2	2	5-2	3
3	4	4-1	3
4	2	5-2	3
5	4	4-1	3
6	2	5-2	3
7	4	4-1	3
8	2	5-2	3
9	5	5-1	4
10	2	5-2	3
Total			31
Nilai Akhir (Total * 2,5)			77,5

Nilai SUS *dashboard* yang diperoleh sebesar 77,5, dimana nilai 77,5 pada parameter SUS *Adjective Ratings* dikategorikan sebagai 'Good' dengan *Acceptability Ranges* dikategorikan sebagai 'Acceptable'. Hal ini membuktikan bahwa *dashboard* yang dihasilkan sudah dapat digunakan dengan cukup baik oleh pengguna. Hasil dari penilaian SUS ini merupakan jawaban terhadap rumusan masalah ketiga (R3).

4. KESIMPULAN

Dengan mengesampingkan aspek umum karena sifatnya yang non-spesifik, aspek yang paling sering diperhatikan pengunjung adalah aspek fasilitas.

Berdasarkan hal tersebut rekomendasi berdasarkan kekuatan Tugu Pahlawan Surabaya adalah penyediaan lahan foto atau objek foto bernuansa sejarah pahlawan secara lebih nyata seperti gambar dinding tiga dimensi atau patung pahlawan. Sementara rekomendasi untuk mengatasi kelemahan yang sering dikeluhkan oleh pengunjung adalah penyediaan ventilasi terbuka atau AC yang lebih banyak di beberapa spot, berupa *standing cooler* yang lebih fleksibel dan efektif digunakan di berbagai tempat.

Sementara hasil pengujian model klasifikasi yang dihasilkan menunjukkan bahwa *dataset* yang *imbalance* menjadi pengaruh dalam penilaian *F1-Score* dan *Accuracy*, dimana pada kasus *imbalance dataset* lebih memperhatikan nilai *F1-Score*. Hasil *F1-Score* tertinggi dimiliki oleh aspek pelayanan sebesar 85% dan hasil terendah dimiliki aspek kebersihan sebesar 55%. Dua faktor utama yang menyebabkan nilai *F1-Score* cenderung kecil adalah karena *imbalance dataset* membuat model memberikan nilai tinggi pada sentimen yang memiliki sebaran jauh lebih banyak daripada yang lainnya. Yang kedua, keberadaan kelas sentimen netral cukup membingungkan terkait bagaimana menentukan standarnya karena tidak didasari oleh adanya emosi pada opini yang diberikan.

Berdasarkan hasil penilaian *Usability Testing* terhadap *dashboard*, meskipun *Grade Scale* yang diperoleh adalah 'C', namun *dashboard* sudah dapat diterima oleh pengguna dengan baik jika dilihat dari *Ajective Ratings* dan terutama *Acceptability Ranges*. Dengan adanya *dashboard* dengan penyajian informasi secara ringkas dapat menjadi inovasi baru dari sisi teknologi bagi pihak pengelola yang juga sebagai sumber analisis untuk menyusun rencana strategis kedepannya, khususnya dalam memperhatikan perspektif pengunjung.

Jika melihat dari temuan dari penelitian ini baik berupa hasil serta kendala, maka disarankan pada penelitian serupa di masa mendatang dapat menerapkan beberapa alternatif dari segi penerapan pelabelan sentimen secara otomatis seperti penggunaan *Lexicon-based* sehingga menghindari adanya subjektivitas, serta penggunaan teknik *oversampling* atau *undersampling* untuk mengatasi *imbalance dataset*. Selain itu juga disarankan untuk mempertimbangkan *tool* lain dalam pembangunan *dashboard* yang dapat melakukan *automate synchronization* dan mampu mengaplikasikan *privilege* akses informasi pada *dashboard* berdasarkan akun pengguna yang digunakan.

DAFTAR PUSTAKA

- AGGARWAL, C. C. & ZHAI, C., 2012. Mining Text Data. New York: Springer.
- AL-SMADI, M. ET AL., 2018. Deep Recurrent Neural Network vs. Support Vector Machine for Aspect-based Sentiment Analysis of Arabic

- Hotels' Reviews. *Journal of Computational Science*, vol. 27, pp. 386-393.
- AMALIA, N. L. R., SUPIANTO, A. A., SETIAWAN, N. Y., ZILVAN, V., YULIANI, A. R., RAMDAN, A., 2021. Student Academic Mark Clustering Analysis and Usability Scoring on Dashboard Development Using K-Means Algorithm and System Usability Scale. *Jurnal Ilmu Komputer dan Informasi (JIKI)*, 14(2), pp. 137-143.
- BANGOR, A., KORTUM, P. . T. & MILLER, . J. T., 2008. An Empirical Evaluation of the System Usability Scale. *Intl. Journal of Human-Computer Interaction*, 24(6), pp. 574-594.
- BOEING, G. & WADDELL, P., 2016. New Insights into Rental Housing Markets across the United States: Web Scraping and Analyzing Craigslist Rental Listings. *Journal of Planning Education and Research*, 37(4), pp. 457-476.
- CALDERWOOD, L. U. & SOSHKIN, M., 2019. The Travel & Tourism Competitiveness Report 2019, Geneva: World Economic Forum.
- DINAS KEBUDAYAAN DAN PARIWISATA KOTA SURABAYA, 2019. Hero Monument - Dinas Kebudayaan dan Pariwisata Kota Surabaya. Tersedia di: <<https://sparkling.surabaya.go.id/en/hero-monument/>> [Diakses 15 September 2020]
- DINAS KOMUNIKASI DAN INFORMATIKA PROVINSI JAWA TIMUR, 2018. Surabaya Kota Terbaik di Ajang Yokatta Wonderful Indonesia Tourism Awards 2018. Tersedia di: <<http://kominfo.jatimprov.go.id/read/umum/surabaya-kota-terbaik-di-ajang-yokatta-wonderful-indonesia-tourism-awards-2018>> [Diakses 18 November 2020]
- DOGAN, T. & UYSAL, A. K., 2019. Improved Inverse Gravity Moment Term Weighting for Text Classification. *Expert Systems with Applications*, vol. 130, 45-59.
- FRUNZA, M.-C., 2016. Solving Modern Crime in Financial Markets: Analytics and Case Studies. s.l.:Elsevier.
- KHARDE, V. A. & SONAWANE, S. S., 2016. Sentiment Analysis of Twitter Data: A Survey of Techniques. *International Journal of Computer Applications*, 139(11).
- PRATAMA, Y. T., BACHTIAR, F. A. & SETIAWAN, N. Y., 2018. Analisis Sentimen Opini Pelanggan Terhadap Aspek Pariwisata Pantai Malang Selatan Menggunakan TF-IDF dan Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(12), pp. 6244-6252.
- RAMADHANI, N. & FAJARIANTO, N. 2020. Sistem Informasi Evaluasi Perkuliahan dengan Sentimen Analisis Menggunakan Naïve Bayes dan Smoothing Laplace. *Jurnal Sistem Informasi Bisnis*, 10(2), pp. 238-234.
- SALAM, S. N., SUPIANTO, A. A., & PERDANAKUSUMA, A. R. 2019. Analisis Sentimen Opini Mahasiswa Terhadap Saran Kuesioner Penilaian Kinerja Dosen dengan Menggunakan TF-IDF dan K-Nearest Neighbor, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3(6), pp. 6148-6156.
- SCIKIT-LEARN, 2020. `sklearn.model_selection.GridSearchCV` - scikit-learn 0.24.0 documentation. Tersedia di: <https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html> [Diakses 18 November 2020]
- SHAFIABADY, N. ET AL., 2016. Using Unsupervised Clustering Approach to Train the Support Vector Machine for Text Classification. *Neurocomputing*, vol. 211, pp. 4-10.
- SHARFINA, Z. & SANTOSO, H. . B., 2016. An Indonesian adaptation of the System Usability Scale (SUS). Malang, Indonesia, International Conference on Advanced Computer Science and Information Systems (ICACISIS).
- SHATHIK, A., & KARANI, K. P, 2020. A Literature Review on Application of Sentiment Analysis Using Machine Learning Techniques. *International Journal of Applied Engineering and Management Letters*, 4(2), pp. 41-77.
- SULAEMAN, A. F., SUPIANTO, A. A., & BACHTIAR, F. A, 2019. Analisis Sentimen Opini Mahasiswa Terhadap Saran Evaluasi Kinerja Dosen Menggunakan TF-IDF dan Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3(6), pp. 5647-5655.
- UNIVERSITAS INDONESIA, 2018. Laporan Akhir: Kajian Dampak Sektor Pariwisata Terhadap Perekonomian Indonesia, Jakarta: Kementerian Pariwisata dan Ekonomi Kreatif Republik Indonesia.
- VITANDY, W. S., SUPIANTO, A. A., & BACHTIAR, F. A, 2019. Analisis Sentimen Evaluasi Kinerja Dosen menggunakan Term Frequency-Inverse Document Frequency dan Naïve Bayes Classifier. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3(6), pp. 6080-6088.
- WAJONG, A. M. R., 2015. Applying Performance Dashboard in Hospitals. *International Journal of Software Engineering and Its Applications*, 9(1), pp. 213-220.

Halaman ini sengaja dikosongkan