

**ANALISIS SENTIMEN ASPEK PADA OPINI WISATAWAN  
MENGUNAKAN METODE *NAÏVE BAYES* (STUDI KASUS:  
PERUM PERHUTANI)**

**SKRIPSI**

Untuk memenuhi sebagian persyaratan  
memperoleh gelar Sarjana Komputer

Disusun oleh:  
Fasya Ghassani Hadiyan  
NIM: 145150407111026



PROGRAM STUDI SISTEM INFORMASI  
JURUSAN SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS BRAWIJAYA  
MALANG  
2021

## PENGESAHAN

ANALISIS SENTIMEN ASPEK PADA OPINI WISATAWAN MENGGUNAKAN METODE  
NAÏVE BAYES (STUDI KASUS: PERUM PERHUTANI)

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan  
memperoleh gelar Sarjana Komputer

Disusun Oleh :

Fasya Ghassani Hadiyan

NIM: 145150407111026

Skripsi ini telah diuji dan dinyatakan lulus pada  
9 Juli 2021

Telah diperiksa dan disetujui oleh:

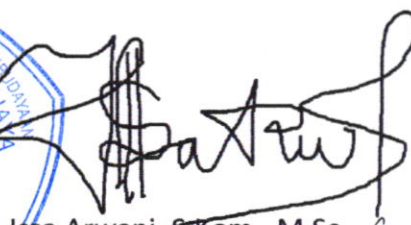
Dosen Pembimbing 1



Dr. Eng. Fitra Abdurrachman Bachtiar, S.T., M.Eng.

NIP: 198406282019031006

Mengetahui  
Ketua Jurusan Sistem Informasi



Issa Arwani, S.Kom., M.Sc.

NIP: 198309222012121003

## PERNYATAAN ORISINALITAS

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar referensi.

Apabila ternyata didalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur plagiasi, saya bersedia skripsi ini digugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70).

Malang, 24 Juni 2021



Fasya Ghassani Hadiyan

NIM: 145150407111026





## ABSTRAK

Fasya Ghassani Hadiyan, Analisis Sentimen Aspek Pada Opini Wisatawan Menggunkan Metode Naïve Bayes (Studi Kasus: Perum Perhutani)

Pembimbing: Dr.Eng. Fitra Abdurrachman Bachtiar, S.T., M.Eng.

Perum Perhutani merupakan BUMN (Badan Usaha Milik Negara) yang salah satu portofolio bisnisnya adalah mengelola dan mengembangkan 768 ekowisata di Jawa dan Madura. Untuk mengembangkan objek wisata yang dikelolanya Perhutani belum dapat melakukannya dengan optimal dikarenakan kurangnya data dan informasi mengenai preferensi wisatawan mengenai objek-objek wisata di samping upaya yang dilakukannya. Analisis sentimen menggunakan *Multinomial Naïve Bayes* dilakukan untuk mengetahui preferensi wisatawan dilihat dari tiga aspek terhadap objek wisata yang dikelola Perhutani. Ketiga aspek tersebut adalah *attractions*, *accessibilities*, dan *amenities*. Ulasan didapatkan dari TripAdvisor yang akan diklasifikasikan ke dalam tiga kelas, yaitu positif, negatif, dan netral. Tujuan dari penelitian ini, yaitu untuk mengetahui hasil implementasi metode *Multinomial Naïve Bayes* pada proses analisis sentimen, mengetahui nilai *precision*, *recall*, dan *f1*, dan terakhir mengetahui hasil usability testing dari dashboard. Hasil proses klasifikasi untuk aspek *attractions* berupa nilai *accuracy* sebesar 69,89%. Sedangkan untuk aspek *accessibilities* dan *amenities* didapatkan nilai *accuracy* sebesar 61,67% dan 63,01% untuk masing-masing aspek. Informasi berupa sentimen wisatawan tersebut ditampilkan ke dalam *dashboard* yang kemudian diujikan kepada pihak Perhutani menggunakan SUS (*System Usability Scale*) dan mendapatkan skor 65. Hasil skor tersebut menandakan dashboard memiliki *rating* "Good" dan termasuk ke dalam *range Marginal (High)* untuk penerimaan.

Kata kunci: analisis sentimen, *multinomial naïve bayes*, klasifikasi, objek wisata, *dashboard*

## ABSTRACT

**Fasya Ghassani Hadiyan, Aspect Sentiment Analysis on Tourists Opinions Using Naïve Bayes (Study Case: Perum Perhutani)**

**Supervisors: Dr.Eng. Fitra Abdurrachman Bachtiar, S.T., M.Eng.**

*Perum Perhutani is a BUMN (Badan Usaha Milik Negara) whose business portfolio is managing and developing 768 ecotourism in Java and Madura. Perhutani has not been able to develop the tourism objects optimally due to the lack of data and information regarding tourist preferences about tourism objects despite the effort they have done. Sentiment analysis using Multinomial Naïve Bayes was carried out to determine tourist preferences from three aspects of tourism objects managed by Perhutani. The three aspects are attractions, accessibilities, and amenities. Reviews obtained from TripAdvisor will be classified into three classes, namely positive, negative, and neutral. The purposes of this study are first, to find out the implementation results of Multinomial Naïve Bayes in sentiment analysis process, second, to find out the values of precision, recall, and f1, and lastly to know the results of usability testing from the dashboard. The results of the classification process for the attractions aspect is 69.89% of accuracy. As for accessibilities and amenities, the accuracy values are 61.67% and 63.01% for each aspect. Information about of tourist sentiment is displayed on a dashboard which is then tested by Perhutani using SUS (System Usability Scale) and gets a score of 65. That result means the dashboard has "Good" rating and in range Marginal (High) for acceptability.*

**Keywords: sentiment analysis, multinomial naïve bayes, classifications, tourism objects, dashboard**



## DAFTAR ISI

PENGESAHAN .....	2
PERNYATAAN ORISINALITAS .....	3
PRAKATA .....	4
ABSTRAK .....	5
ABSTRACT .....	6
DAFTAR ISI .....	7
DAFTAR TABEL .....	10
DAFTAR GAMBAR .....	11
DAFTAR LAMPIRAN .....	12
BAB 1 PENDAHULUAN .....	13
1.1 Latar belakang .....	13
1.2 Rumusan masalah .....	14
1.3 Tujuan .....	15
1.4 Manfaat .....	15
1.5 Batasan masalah .....	15
1.6 Sistematika pembahasan .....	15
BAB 2 LANDASAN KEPUSTAKAAN .....	17
2.1 Kajian Pustaka .....	17
2.2 Profil Perusahaan .....	18
2.3 <i>TripAdvisor</i> .....	19
2.4 <i>Web Scraping</i> .....	19
2.5 <i>Data Mining</i> .....	20
2.5.1 Definisi <i>Data Mining</i> .....	20
2.5.2 <i>Data Mining Task</i> .....	21
2.6 <i>Text Mining</i> .....	22
2.6.1 <i>Text Pre-Processing</i> .....	24
2.7 Pembobotan TF-IDF .....	24
2.8 Analisis Sentimen .....	25
2.9 <i>Machine Learning</i> .....	25
2.9.1 <i>Naïve Bayes Classifier</i> .....	26
2.10 Evaluasi .....	27

2.11 Visualisasi Data .....	28
2.11.1 <i>Dashboard</i> .....	28
2.12 <i>System Usability Scale (SUS)</i> .....	29
<b>BAB 3 METODOLOGI</b> .....	31
3.1 Identifikasi Masalah .....	31
3.2 Studi Literatur .....	32
3.3 Pengumpulan Data dan Informasi .....	32
3.4 Analisis Sentimen .....	32
3.5 Analisis dan Pengujian Hasil Klasifikasi .....	32
3.6 Visualisasi Data .....	33
3.7 Penarikan Kesimpulan dan Saran .....	33
<b>BAB 4 PERANCANGAN</b> .....	34
4.1 <i>Web Scraping</i> .....	34
4.2 Penentuan Aspek .....	39
4.3 Pelabelan Aspek dan Sentimen .....	39
4.4 Desain Antarmuka <i>Dashboard</i> .....	41
<b>BAB 5 IMPLEMENTASI</b> .....	42
5.1 <i>Text Preprocessing</i> .....	42
5.1.1 <i>Data Cleansing</i> .....	43
5.1.2 <i>Fix Typo</i> .....	43
5.1.3 <i>Stopwords Removal</i> .....	44
5.1.4 <i>Stemming</i> .....	45
5.2 TF-IDF dan Klasifikasi .....	45
5.3 Implementasi <i>Dashboard</i> .....	45
<b>BAB 6 HASIL DAN PEMBAHASAN</b> .....	48
6.1 Hasil Klasifikasi .....	48
6.1.1 Aspek <i>Attractions</i> .....	48
6.1.2 Aspek <i>Accessibilities</i> .....	49
6.1.3 Aspek <i>Amenities</i> .....	50
6.2 Hasil <i>System Usability Score (SUS) Dashboard</i> .....	51
<b>BAB 7 PENUTUP</b> .....	53
7.1 Kesimpulan .....	53



7.2 Saran .....	53
DAFTAR REFERENSI .....	55
LAMPIRAN A HASIL WAWANCARA .....	58
LAMPIRAN B PERSETUJUAN CONTOH LABEL DATA .....	60
LAMPIRAN C HASIL USABILITY TESTING .....	66





## DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i> .....	27
Tabel 4.1 Variabel dan tipe data pada proses <i>web scraping</i> .....	34
Tabel 4.2 Daftar <i>link</i> objek wisata pada TripAdvisor .....	36
Tabel 4.3 Contoh hasil <i>web scraping</i> .....	38
Tabel 4.4 Kategori aspek objek wisata .....	39
Tabel 4.5 Contoh kata kunci sentimen .....	40
Tabel 4.6 Total aspek pada kalimat .....	40
Tabel 5.1 Detail <i>dataset</i> .....	43
Tabel 5.2 Contoh perbaikan kata .....	44
Tabel 6.1 Hasil <i>accuracy, precision, recall, dan f1</i> aspek <i>attractions</i> .....	48
Tabel 6.2 Hasil <i>accuracy, precision, recall, dan f1</i> aspek <i>accessibilities</i> .....	49
Tabel 6.3 Hasil <i>accuracy, precision, recall, dan f1</i> aspek <i>amenities</i> .....	50
Tabel 6.4 Hasil <i>Usability Testing</i> .....	51



## DAFTAR GAMBAR

Gambar 2.1 Struktur Organisasi Perum Perhutani .....	19
Gambar 2.2 Contoh <i>Dashboard</i> .....	29
Gambar 2.3 Skor <i>System Usability Scale</i> .....	29
Gambar 3.1 Alur Metodologi Penelitian .....	31
Gambar 4.1 Mengakses situs web TripAdvisor pada WebHarvy .....	35
Gambar 4.2 Menyimpan data ke variabel .....	35
Gambar 4.3 Menambahkan URL objek wisata lainnya .....	36
Gambar 4.4 Proses <i>scraping</i> data yang telah dikonfigurasi .....	37
Gambar 4.5 Menyimpan data hasil <i>scraping</i> .....	37
Gambar 4.6 Tahap pelabelan data .....	39
Gambar 4.7 Contoh kata kunci aspek dalam dokumen <i>Tools Assessment</i> Wisata .....	40
Gambar 4.8 Desain <i>Dashboard</i> .....	41
Gambar 5.1 Alur Analisis Sentimen .....	42
Gambar 5.2 Proses <i>Text Preprocessing</i> .....	42
Gambar 5.3 Contoh daftar <i>stopwords</i> yang ditambahkan .....	45
Gambar 5.4 Grafik jumlah ulasan per aspek .....	46
Gambar 5.5 Grafik tren sentimen terhadap aspek .....	46
Gambar 5.6 Jumlah ulasan tiap objek wisata .....	47
Gambar 5.7 <i>Dashboard</i> ulasan wisatawan terhadap objek wisata .....	47
Gambar 6.1 <i>System Usability Score</i> (SUS) .....	52





## DAFTAR LAMPIRAN

LAMPIRAN A HASIL WAWANCARA.....	58
LAMPIRAN B PERSETUJUAN CONTOH LABEL DATA.....	60
LAMPIRAN C HASIL <i>USABILITY TESTING</i> .....	66



## BAB 1 PENDAHULUAN

### 1.1 Latar belakang

Indonesia sebagai negara kepulauan terbesar di dunia yang memiliki kurang lebih 17.000 pulau menjadikan Indonesia sebagai tempat wisata yang berpotensi untuk menarik wisatawan domestik dan mancanegara. Hutan, gunung, air terjun, pantai, dan situs-situs bersejarah yang dimiliki Indonesia dapat menjadi objek wisata. Hal ini dibuktikan dengan sektor pariwisata yang menjadi salah satu sumber utama pendapatan terbesar negara. Namun demikian, akibat pengelolaan dan pengembangan pariwisata yang tidak optimal Indonesia mendapat indeks daya saing pariwisata yang rendah berdasarkan data yang dikeluarkan WEF (*World Economic Forum*) pada tahun 2019. Indonesia mendapat peringkat 40 dari 136 negara yang masih berada di bawah negara tetangga Indonesia seperti Singapura pada peringkat 13, Malaysia di peringkat 26, dan Thailand di peringkat 34 (WEF, 2019).

Perum Perhutani atau singkatnya Perhutani merupakan salah satu Badan Usaha Milik Negara (BUMN) yang bertugas dalam pengelolaan hutan dan juga mengembangkan sekitar 768 objek wisata alam di Pulau Jawa meliputi wisata hutan, wisata pantai, air terjun, telaga, dan gua. Destinasi wisata andalan perhutani, antara lain Kawah Putih (Bandung), Cilemer (Bogor), Pantai Tanjung Papuma (Jember), dan Pantai Pulau Merah (Banyuwangi). Di dalam mengelola kawasan hutan, Perum Perhutani dituntut untuk dapat memberikan manfaat yang optimal dari kawasan hutan baik dari aspek ekonomi, sosial maupun lingkungan. Pengembangan bisnis wisata alam menjadi salah satu arah kebijakan portofolio bisnis perusahaan yang bertujuan untuk mengoptimalkan manfaat hutan dari ketiga aspek dimaksud. Dari sisi komersial arah pengembangan dan pengelolaan wisata Perhutani ditujukan untuk meningkatkan jumlah wisatawan, meningkatkan lama kunjungan wisatawan, dan meningkatkan pendapatan. Untuk mencapai tujuan tersebut pihak Perhutani perlu melakukan investasi terhadap objek wisata yang dikelolanya.

Saat ini Perhutani belum optimal dalam mengembangkan bisnis wisata alamnya. Salah satu persoalan yang dihadapi manajemen Perhutani dalam pengembangan objek wisata adalah kekurangan data dan informasi mengenai preferensi wisatawan untuk dijadikan acuan dalam berinvestasi. Pihak Perhutani telah melakukan beberapa upaya untuk mendapatkan data untuk kepentingan pengembangan objek wisata antara lain melakukan survei satu tahun sekali dengan mengirimkan ahli ekowisata serta meletakkan kotak saran di beberapa titik wisata dan mengelola media sosial untuk menampung kritik dan saran dari wisatawan. Akan tetapi, data yang didapatkan dari upaya-upaya tersebut dianggap belum bisa menjadi landasan pengambilan keputusan oleh pihak manajemen karena belum diolah menjadi informasi-informasi yang berguna.

Kekurangan data sebagai landasan pengambilan keputusan tersebut dapat dilengkapi oleh data dari media sosial. Media sosial selain digunakan untuk saling



berkomunikasi antar penggunanya, saat ini juga digunakan untuk kepentingan organisasi seperti pemasaran dan pelayanan pelanggan. Salah satu aplikasi web pada bidang pariwisata, yaitu TripAdvisor memfasilitasi para penggunanya untuk memberikan komentar serta foto terkait pengalaman mereka pada sebuah objek wisata (TripAdvisor, 2021). Objek-objek wisata pada TripAdvisor mencakup beberapa objek wisata yang dikelola Perum Perhutani. Data tersebut dapat dimanfaatkan oleh pihak Perhutani untuk mendukung pengambilan keputusan. Akan tetapi, data yang ada pada TripAdvisor belum langsung bisa digunakan untuk pengambilan keputusan. Meskipun data yang tersedia banyak, namun data yang berasal dari pengguna media sosial tersebut memiliki banyak *noise* dan tidak terstruktur (Liu, 2012). Salah satu metode yang bisa digunakan untuk mengatasi permasalahan tersebut adalah *data mining*.

Pada penelitian sebelumnya yang serupa oleh Prameswari et al. (2017) melakukan analisis sentimen aspek terhadap opini pengunjung hotel yang diambil dari TripAdvisor untuk membantu meningkatkan kualitas hotel-hotel yang terdapat di Bali. Penelitian selanjutnya oleh Geetha et al. (2017) bertujuan untuk mencari relasi antara rating dan review dari pengunjung hotel saling berkaitan satu sama lain atau tidak menggunakan analisis sentimen. Metode yang digunakan untuk melakukan analisis sentimen, yaitu algoritme *Naïve Bayes*.

Penelitian ini dilakukan analisis sentimen aspek terhadap data yang berisi opini wisatawan mengenai objek wisata yang dikelola Perhutani. Metode yang diusulkan untuk penelitian ini adalah algoritme *Naïve Bayes* untuk klasifikasi pada analisis sentimen. Metode tersebut dipilih karena dinilai efektif dan memiliki tingkat kesalahan paling minimal (Geetha et al., 2017).

Berdasarkan permasalahan yang sudah dipaparkan, maka penyusun mengajukan penelitian terhadap permasalahan pada Perum Perhutani dengan judul "Analisis Sentimen Aspek pada Opini Wisatawan Menggunakan Metode *Naïve Bayes* (Studi Kasus: Perum Perhutani)".

## 1.2 Rumusan masalah

Penelitian ini menerapkan pada ulasan data objek-objek wisata yang dikelola Perhutani yang diambil dari TripAdvisor untuk menjawab pertanyaan-pertanyaan:

1. Bagaimana hasil implementasi algoritme *Naïve Bayes* pada analisis sentimen aspek tentang ulasan objek wisata Perhutani yang berasal dari TripAdvisor?
2. Bagaimana hasil *precision*, *recall*, dan *f1* analisis sentimen aspek tentang ulasan objek wisata Perhutani yang berasal TripAdvisor?
3. Bagaimana hasil *usability testing dashboard* sentimen wisatawan yang dilakukan dengan pihak Perum Perhutani?

### 1.3 Tujuan

1. Mengetahui hasil implementasi algoritme *Naïve Bayes* pada analisis sentimen aspek tentang ulasan objek wisata Perhutani yang berasal dari TripAdvisor.
2. Mengetahui hasil akurasi analisis sentimen aspek tentang ulasan objek wisata Perhutani yang berasal TripAdvisor.
3. Mengetahui hasil *usability testing dashboard* sentimen wisatawan yang dilakukan dengan pihak Perhutani.

### 1.4 Manfaat

Manfaat yang didapatkan dari penelitian ini adalah pihak Perhutani dapat melihat sentimen wisatawan (positif, negatif, atau netral) mengenai objek-objek wisata yang dikelolanya terkait aspek *attractions*, *accessibilities*, dan *amenities*.

### 1.5 Batasan masalah

1. Penelitian dilakukan di Perum Perhutani Indonesia.
2. Analisis Sentimen dilakukan berdasarkan data ulasan berbahasa Indonesia yang didapat dari TripAdvisor terkait aspek *attractions*, *accessibilities*, dan *amenities* dari objek-objek wisata yang dikelola Perum Perhutani Indonesia di pulau Jawa dan Madura.
3. Metode *Naïve Bayes* yang digunakan, yaitu *Multinomial Naïve Bayes*.

### 1.6 Sistematika pembahasan

Sistematika pembahasan pada penelitian ini disusun menjadi beberapa bab berikut.

#### BAB 1 PENDAHULUAN

Bab ini menjelaskan latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, dan sistematika pembahasan dari penelitian yang dilakukan.

#### BAB 2 LANDASAN KEPUSTAKAAN

Bab ini menjelaskan pembahasan tentang penelitian-penelitian sebelumnya yang menjadi referensi pada penelitian ini dan dasar teori untuk mendukung penelitian yang berasal dari literatur ilmiah.

#### BAB 3 METODOLOGI

Bab ini menjelaskan tahapan-tahapan yang dilakukan penyusun untuk menyelesaikan penelitian.

#### BAB 4 PERANCANGAN

Bab ini menjelaskan tahapan-tahapan awal perancangan penelitian sebelum bisa dilakukan implementasi meliputi tahapan akuisisi data, penentuan aspek wisata, perancangan data, dan perancangan antarmuka *dashboard*.



## BAB 5 IMPLEMENTASI

Bab ini menjelaskan implementasi dari hasil perancangan yang telah dilakukan pada bab sebelumnya.

## BAB 6 HASIL DAN PEMBAHASAN

Bab ini menjelaskan hasil penelitian dan pembahasan yang didapat dari penelitian yang dilakukan.

## BAB 7 KESIMPULAN DAN SARAN

Bab ini menjelaskan kesimpulan yang ditarik dari penelitian yang telah dan saran untuk penelitian selanjutnya.



## BAB 2 LANDASAN KEPUSTAKAAN

Bab 2 membahas literatur-literatur yang digunakan sebagai landasan atau rujukan penelitian di skripsi ini, profil singkat dari Perum Perhutani, dan juga tentang TripAdvisor sebagai sumber data penelitian.

### 2.1 Kajian Pustaka

Penelitian yang dilakukan mengacu kepada beberapa penelitian sebelumnya. Penelitian yang digunakan sebagai acuan penelitian ini oleh Geetha et al. (2017) menganalisis relasi antara *review* dan *rating* yang diberikan pengunjung hotel serta menganalisis konsistensi *review* dan *rating* terhadap kategori hotel. Penelitian ini menggunakan algoritme *Naïve Bayes* untuk mengklasifikasikan *review* pengunjung hotel menjadi positif, negatif, atau netral. Selain itu pada penelitian ini juga dicari kategori-kategori yang sering muncul untuk mengetahui preferensi pengunjung terhadap layanan hotel menggunakan *hierarchical clustering*. Hasil dari penelitian ini membuktikan bahwa *review* dan *rating* yang diberikan pengunjung dapat membantu pihak manajerial dalam melakukan perbaikan.

Sementara itu analisis sentimen di bidang pariwisata yang lainnya dilakukan oleh Prameswari et al. (2017) yang meneliti *review* pengunjung hotel-hotel di Bali pada TripAdvisor dengan lebih memperhatikan aspek layanan dan fasilitas hotel. Analisis sentimen dilakukan menggunakan algoritme *Recursive Neural Tensor Network* (RNTN). Hasil akurasi yang didapatkan, yaitu sebesar 85%.

Penelitian oleh Chang (2017) juga masih meneliti analisis sentimen untuk bidang pariwisata. Penelitian ini bertujuan untuk menganalisis keterkaitan antara *review* hotel dan *rating* hotel Hilton di Amerika Serikat pada TripAdvisor. Penelitian ini bertujuan untuk memvisualisasikan *rating* dan *review* hotel Hilton pada *dashboard Google Trends* sehingga dapat membantu pihak pemasaran. Pada penelitian ini analisis sentimen dilakukan menggunakan metode *Convulsion Tree Kernel Classification* untuk mengklasifikasikan *review* ke dalam label positif dan negatif. Hasil penelitian ini, yaitu akurasi sebesar 89,3%.

Penelitian selanjutnya oleh Bachtiar et al. (2020) bertujuan untuk melakukan analisis sentimen mengenai aspek *location*, *room*, *food*, *price*, dan *service* pada industri hotel. Analisis sentimen pada tingkat aspek dilakukan untuk mengetahui kepuasan pelanggan pada *guest house* sehingga dapat menghasilkan E-WOM (*Electronic Word of Mouth*) yang positif. Data yang digunakan merupakan ulasan berbahasa Indonesia yang didapat dari situs web yang berfokus pada biro perjalanan seperti Agoda.com, Expedia, dan Booking.com. Hasil dari accuracy dari analisis sentimen ini yang menggunakan metode SVM adalah sebesar 93% untuk aspek *location*, 80% untuk aspek *room*, 68% pada aspek *food*, 90% *accuracy* aspek *price*, dan 92% pada aspek *service*.

Penelitian terbaru tentang analisis sentimen di bidang pariwisata dapat dilihat di Paolanti, et al. (2021) yang memaparkan permasalahan informasi geografis



terkait satu destinasi pariwisata yang sangat terkenal di Italia, yaitu Cilento. Metode yang digunakan pada penelitian ini adalah metode-metode *deep learning* dengan *dataset* yang diperoleh dari twitter. Penelitian ini bertujuan untuk membuat karakteristik dari pergerakan wisatawan secara spasial, temporal, dan demografis sepanjang daerah wisata pedesaan dan sepanjang pantai di Cilento.

Metode *Naïve Bayes* tidak hanya diterapkan untuk analisis sentimen di bidang pariwisata. Sebagai contoh penelitian oleh Arulselvi (2017) bertujuan untuk mengklasifikasikan sentimen mengenai buku pada topik-topik tertentu sehingga memudahkan pelajar untuk mengetahui buku-buku yang sering dibaca dari topik tertentu. Klasifikasi sentimen dilakukan menggunakan metode *Naïve Bayes Classifier* ke dalam kelas positif, negatif, dan netral. Pada penelitian ini juga dilakukan analisis tren untuk mengetahui judul buku apa saja yang populer pada rentang waktu tertentu. Akurasi klasifikasi *Naïve Bayes* pada penelitian ini sebesar 89.9%. Penelitian lainnya dilakukan oleh Xu et al. (2020) di mana di penelitian ini klasifikasi menggunakan metode *Naïve Bayes* dilakukan untuk mengklasifikasikan beberapa *dataset*, yaitu *movie*, *network*, dan *health* di beberapa *e-commerce*. Selain itu, penelitian ini juga membahas proses *continuous learning* agar model mampu melakukan klasifikasi secara otomatis ketika *review* tentang produk terus bertambah dan *e-commerce domain* juga bertambah.

Penelitian-penelitian tersebut dijadikan sebagai acuan pada penelitian yang dilakukan karena memiliki beberapa kesamaan. Untuk menggali sentimen dari wisatawan digunakan analisis sentimen dengan algoritme *Multinomial Naïve Bayes* untuk klasifikasi sama seperti dua penelitian yang sudah dipaparkan. Selain itu pada penelitian ini akan dicari kategori layanan objek wisata yang sering muncul.

## 2.2 Profil Perusahaan

Perum Perhutani adalah Badan Usaha Milik Negara pengelola hutan di pulau Jawa dan Madura yang memiliki peran strategis mendukung sistem kelestarian lingkungan, sosial budaya dan perekonomian masyarakat perhutanan nasional. Perhutani bersama Pemerintah menyelenggarakan kegiatan pengelolaan sumberdaya hutan yang dapat bersaing di tingkat internasional pada masa mendatang. Perum Perhutani didirikan pada tanggal 29 Maret tahun 1972 berdasarkan Peraturan Pemerintah Nomor 15 Tahun 1972. Visi dan Misi dari Perum Perhutani (Perhutani, 2021), yaitu:

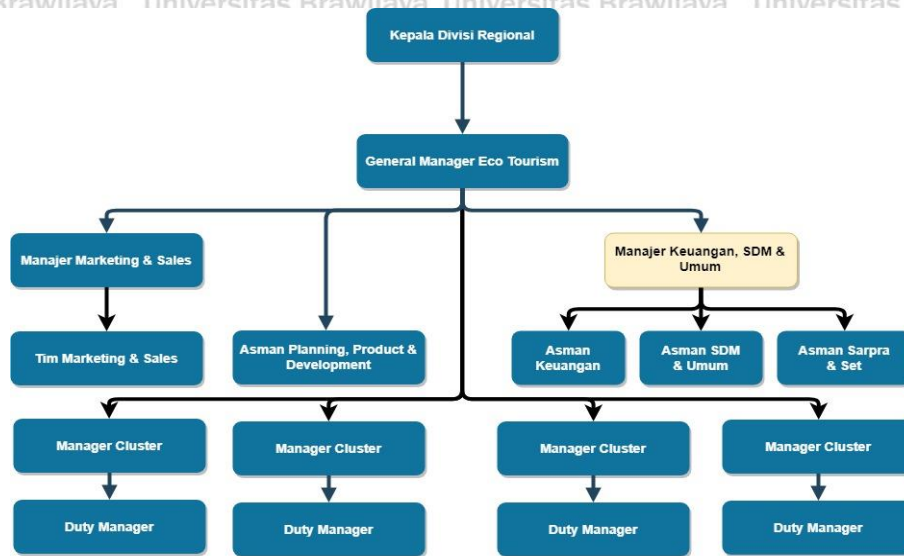
### Visi

Menjadi perusahaan unggul dalam pengelola hutan lestari

### Misi

- Mengelola sumberdaya hutan secara lestari.
- Meningkatkan manfaat pengelolaan sumber daya hutan bagi seluruh pemangku kepentingan.
- Menyelenggarakan bisnis kehutanan dengan prinsip *good corporate*.

Dalam rangka mendapatkan dan meningkatkan keuntungan perusahaan, Perum Perhutani melakukan beberapa upaya dalam pemanfaatan hutan dan hasil hutan. Salah satu portofolio bisnisnya, yaitu memanfaatkan hutan yang tidak termasuk ke dalam hutan lindung menjadi ekowisata. Berdasarkan hasil wawancara, Divisi *Ecotourism* dan *Agroforestry* yang berada langsung di bawah Kementerian Pariwisata dan Ekonomi Kreatif, menunjuk Divisi Keuangan, SDM, dan Umum untuk mengelola dan pengembangan objek-objek wisata di seluruh daerah yang tersebar di Jawa dan Madura. Divisi Keuangan, SDM, dan Umum berkolaborasi dengan Divisi-Divisi Regional menyangkut pengelolaan pariwisata dan berfokus untuk menentukan perbaikan dan pengembangan apa saja terkait objek-objek wisata. Struktur organisasi dapat dilihat pada Gambar 2.1.



Gambar 2.1 Struktur Organisasi Perum Perhutani

### 2.3 TripAdvisor

TripAdvisor merupakan situs web wisata terbesar yang memiliki tujuan untuk membantu penggunanya merencanakan perjalanan. Pada tahun 2018, TripAdvisor mengubah citranya menjadi *social travel networks* untuk mewedahi penggunanya untuk memberikan ulasan dan opini mengenai destinasi wisata, akomodasi (meliputi hotel, penginapan, dan vila), serta restoran yang telah dikunjungi dan pengguna juga dapat menambahkan foto layaknya media sosial pada umumnya. Hingga September 2020 terdapat 702 juta ulasan dan opini mengenai berbagai destinasi wisata, akomodasi, serta restoran di TripAdvisor. Ulasan dan opini tersebut digunakan pengguna untuk membandingkan harga dan pelayanan serta membantu mengenal destinasi wisata (TripAdvisor, 2021).

### 2.4 Web Scraping

*Web Scraping* adalah metode-metode untuk mengumpulkan informasi dari situs web secara otomatis. Sebuah *web scraper*, aplikasi untuk melakukan *web scraping*, dapat mengakses halaman-halaman situs web, menemukan data yang



relevan dengan informasi yang diinginkan, mengekstrak data, mentransformasi data, dan menyimpan data dalam *dataset* terstruktur. Kemampuan mengakses, menemukan, mengekstrak, mentransformasi, dan menyimpan data didapat dari suatu program yang terotomatisasi yang dapat mengimplementasikan *query* pada server web, mengirim permohonan data (pada umumnya dalam format *Hyper Text Markup Language* (HTML) atau format lainnya yang biasa digunakan untuk membangun situs web), dan melakukan *parsing* data untuk mengekstrak informasi (Aydin, 2018).

Walaupun sudah banyak situs web seperti Twitter, Facebook, LinkedIn, dan Google yang sudah menyediakan *Application Programming Interface* (APIs) yang bertujuan membuka akses bagi pihak luar untuk mengakses repositori dari situs web tersebut dengan cara yang terstruktur, Broucke dan Baesens (2018) mengemukakan beberapa alasan mengapa *web scraping* lebih baik digunakan dibandingkan penggunaan API. Alasan tersebut antara lain:

- Situs web yang datanya akan diekstrak belum tentu menyediakan API.
- API yang tersedia sifatnya berbayar.
- API yang tersedia hanya dapat diakses dalam waktu yang terbatas.
- API yang tersedia belum tentu menampilkan seluruh data yang ingin diekstrak.

## 2.5 Data Mining

### 2.5.1 Definisi Data Mining

*Data mining* merupakan proses untuk menemukan pola-pola unik dan pengetahuan dari sekumpulan besar data. *Data mining* dapat dianalogikan dengan proses penambangan emas di mana emas yang berharga dipisahkan dari pasir dan batu yang kurang berharga (Han et al., 2012). Turban et al. (2010) menjabarkan karakteristik utama dan tujuan dari *data mining* sebagai berikut.

- a. Data sering kali tersembunyi di dalam basis data yang besar, yang terkadang merupakan data yang sudah ada beberapa tahun. Pada banyak kasus, data dibersihkan dan dimasukkan ke dalam *data warehouse*.
- b. Lingkungan *data mining* biasanya merupakan sebuah arsitektur *client-server* atau sebuah sistem informasi web.
- c. *Tools* yang canggih, termasuk *tools* visualisasi termutakhir, dapat membantu menggali informasi yang terkubur di dalam berkas korporat atau arsip catatan publik. Untuk menemukan informasi tersebut melibatkan proses yang kompleks agar mendapatkan hasil yang tepat dari data yang ada.
- d. *Miner* biasanya merupakan *end-user* yang didukung oleh data latih dan *tools query* lainnya untuk mendapatkan hasil secara cepat dengan kemampuan pemrograman yang sedikit.

- e. Hasil dari *data mining* sering kali tidak terduga dan membutuhkan *miner* untuk berpikir kreatif selama proses *mining*, termasuk menginterpretasikan pengetahuan yang didapat.
- f. *Tools data mining* pada umumnya sudah dilengkapi dengan *spreadsheet* dan *tools* pengembangan perangkat lunak lainnya sehingga pengetahuan yang didapat bisa dianalisis dan diaplikasikan secara cepat dan mudah.
- g. Dikarenakan jumlah data yang besar dan dibutuhkan usaha yang besar untuk *data mining*, terkadang dilakukan pemrosesan secara paralel.

### 2.5.2 Data Mining Task

Turban et al. (2010) membagi *data mining task* ke dalam tiga kategori utama:

1. Prediksi termasuk ke dalam *supervised learning* yang bertujuan untuk mengetahui kemungkinan terjadinya sebuah kejadian di masa mendatang berdasarkan apa yang telah terjadi sebelumnya. Metode *data mining* yang umum digunakan adalah *Classification and Regression Trees*, ANN, SVM, dan metode lainnya yang termasuk dalam *supervised learning*.
  - a. Klasifikasi  
Klasifikasi merupakan proses untuk menemukan sebuah model yang mendeskripsikan dan membedakan kelas atau konsep. Model tersebut dibangun berdasarkan data latih yang sudah memiliki label. Model tersebut digunakan untuk memprediksi kelas label dari objek data yang kelasnya belum diketahui.
  - b. Regresi  
Regresi digunakan untuk memprediksi nilai data yang menghilang atau tidak tersedia di mana model yang dibentuk memiliki nilai kontinu.
2. Asosiasi termasuk ke dalam *unsupervised learning* yang bertujuan untuk menemukan kelompok barang yang paling sering muncul secara bersamaan. Metode *data mining* yang sering digunakan adalah Apriori, OneR, ZeroR, dan Eclat.
  - a. *Link Analysis*  
*Link analysis* digunakan untuk mengevaluasi hubungan atau koneksi antar objek yang berkepentingan.
  - b. *Sequence Analysis*  
*Sequence analysis* digunakan untuk memeriksa relasi berdasarkan urutan kejadian.
3. *Clustering* termasuk ke dalam *unsupervised learning* yang bertujuan untuk mengidentifikasi pengelompokan data berdasarkan karakteristiknya. Metode *data mining* yang umum digunakan adalah K-Means, ANN, atau SOM.
  - a. *Outlier Analysis*  
*Outlier* merupakan objek data yang tidak sesuai dengan model dari sebuah *dataset*. Metode *data mining* pada umumnya menganggap outlier sebagai



*noise* yang biasanya diabaikan. *Outlier analysis* bertujuan untuk menganalisis *outlier* tersebut.

## 2.6 Text Mining

*Text Mining* adalah proses semi otomatis ekstraksi informasi dan pengetahuan dari sekumpulan sumber data tidak terstruktur yang besar. *Text mining* memiliki tujuan dan proses yang sama seperti *data mining*, tetapi *text mining* memiliki masukan berupa data tidak terstruktur seperti dokumen Word, berkas PDF, berkas XML, kutipan teks, dan sumber data serupa lainnya. Secara umum *text mining* memiliki dua tahap utama dimulai dengan menentukan struktur terhadap sumber data teks kemudian mengekstrak informasi dan pengetahuan dari sumber data terstruktur teks menggunakan teknik dan tools *data mining* (Turban et al., 2010).

Tujuan utama dari *text mining* dalam konteks *knowledge discovery* adalah memproses data teks yang tidak terstruktur serta data terstruktur yang relevan terhadap permasalahan untuk mengekstrak pola-pola yang dapat digunakan dan bermanfaat untuk pengambilan keputusan (Turban et al., 2010). Proses *text mining* dapat dibagi ke dalam tiga tahap berurutan, yaitu mendirikan *corpus*, membentuk *term-document matrix*, dan mengekstrak *knowledge* yang tiap-tiap tahap menghasilkan keluaran spesifik. Jika untuk beberapa alasan keluaran dari salah satu tahap tidak sesuai apa yang diharapkan, maka penelusuran kembali pada eksekusi tahap sebelumnya dibutuhkan.

Tahap pertama, yaitu mendirikan *corpus* bertujuan untuk mengumpulkan seluruh dokumen yang terkait konteks permasalahan yang diteliti. *Corpus* sendiri merupakan kumpulan khusus dari materi tekstual yang dikumpulkan berdasarkan kriteria-kriteria tertentu (Manning et al., 2009). Dokumen-dokumen yang dikumpulkan dapat berupa dokumen teks, berkas XML, pos elektronik, halaman web, dan catatan pendek. Dokumen-dokumen yang sudah terkumpul tersebut akan ditransformasi dan diorganisasi ke dalam format representasi yang sama sehingga dapat diproses secara terkomputerisasi. Pengorganisasian dari dokumen-dokumen tersebut dari bisa berupa kumpulan kutipan teks di dalam folder atau bisa berupa daftar dari tautan ke beberapa halaman web.

Keluaran dari tahap sebelumnya, yaitu *corpus* digunakan untuk membuat *term-document matrix* (TDM). Baris dalam TDM merepresentasikan dokumen sedangkan kolom merepresentasikan istilah yang berkaitan dengan konteks permasalahan yang sedang diteliti. Hubungan antara istilah dan dokumen diberi tanda menggunakan indeks. Tujuan dari tahap ini adalah untuk mengkonversi *corpus* ke dalam TDM dimana setiap sel terisi dengan indeks yang sesuai. Daftar istilah yang dimasukkan ke dalam TDM disebut *stop terms* atau *stop words*, yaitu istilah-istilah yang sesuai dengan bidang yang sedang diteliti dan ditentukan oleh ahli dari bidang yang diteliti. Selain cara tersebut dapat ditentukan sebelumnya istilah-istilah yang ada di dalam dokumen untuk diberi indeks. Hal lain yang perlu dipertimbangkan adalah sinonim dan kalimat spesifik dapat ditambahkan ke dalam TDM sehingga indeks yang diberikan akurat. *Stemming* merupakan pengurangan kata hanya menjadi kata dasar saja juga harus dilakukan.

Tahap terakhir adalah mengekstrak *knowledge* menggunakan TDM yang terstruktur secara baik dan dapat digabungkan dengan sumber data terstruktur terkait permasalahan. Kategori-kategori utama dari metode ekstraksi pengetahuan antara lain, klasifikasi, *clustering*, asosiasi, dan analisis tren. Pada penelitian ini metode yang digunakan adalah klasifikasi.

*Text mining* memiliki banyak persamaan dengan *data mining* bersangkutan pada *high-level architecture*. Sebagai contoh, sistem *data mining* bergantung pada *preprocessing*, algoritme untuk menemukan pola pada data, dan elemen *presentation layer* seperti perangkat visualisasi untuk mempermudah penelusuran hasil *mining*. Selain itu, *text mining* mengadopsi berbagai tipe pola *knowledge discovery* yang sudah diperkenalkan dalam *data mining*.

Dalam *data mining* diasumsikan bahwa data yang dimiliki dalam format terstruktur sehingga *preprocessing* berfokus pada dua tugas utama normalisasi data dan membuat sejumlah besar *table join*. Pada *text mining preprocessing* berpusat dalam mengidentifikasi dan mengekstrak fitur representatif untuk dokumen bahasa alami. *Preprocessing* berfokus untuk mentransformasikan data tidak terstruktur berbentuk dokumen menjadi terstruktur (Feldman dan Sanger, 2007).

Berdasarkan Feldman dan Sanger (2007) pada tingkat fungsional *text mining* mengikuti model umum yang disediakan oleh aplikasi *data mining* dan maka dari itu terbagi ke dalam empat area, yaitu:

1. *Preprocessing* melibatkan seluruh rutinitas, proses, dan metode yang dibutuhkan untuk mempersiapkan data untuk operasi *knowledge discovery*. Tahap ini pada umumnya berpusat pada *preprocessing* sumber data dan mengkategorikan aktivitas. *Preprocessing* secara umum mengubah informasi dari sumber data ke dalam format terstruktur sebelum metode ekstraksi fitur dapat diaplikasikan pada dokumen.
2. *Core mining* operation merupakan pusat dari sistem text mining dan di dalamnya termasuk proses *pattern discovery*, analisis tren, dan algoritme *incremental knowledge discovery*.
3. *Presentation layer components* di dalamnya termasuk GUI dan fungsionalitas penelusuran *pattern* serta akses pada *query*. Perangkat visualisasi dan *editor query* juga termasuk ke dalam kategori arsitektur ini.
4. *Refinement techniques* di dalamnya termasuk metode untuk menyaring informasi yang tidak diperlukan dan melakukan *clustering* terhadap data yang saling erat terkait sebagai upaya untuk mengoptimalkan *knowledge discovery* pada sistem *text mining* yang diberikan.

## 2.6.1 Text Pre-Processing

### 2.6.1.1 Case Folding

Tahap *case folding* adalah tahap merubah huruf-huruf pada dokumen ke dalam huruf kecil. Proses ini penting dilakukan karena tidak semua dokumen konsisten dengan penggunaan huruf kapital. Pada tahap ini juga dilakukan proses pengeliminasian *delimiter* seperti tanda seru dan titik.

### 2.6.1.2 Cleansing

Pada *text preprocessing* tahap *cleansing* dilakukan untuk menghilangkan *link*, *tag html*, *script*, dan elemen lainnya yang tidak dibutuhkan.

### 2.6.1.3 Filtering

Tahap *filtering* merupakan tahapan yang dilakukan untuk mengambil kata-kata yang dianggap penting dan membuang kata-kata yang dianggap tidak penting. Kata-kata yang dihapuskan biasanya merupakan kata-kata yang terlalu sering muncul dan tidak memiliki arti atau tidak relevan (Manning et al., 2009). Pada tahap *filtering* dapat menggunakan algoritme *stoplist*, yaitu membuang kata yang kurang bermakna dalam pendekatan *bag-of-words*. *Stop words* dimasukkan ke dalam basis data dan akan dilakukan pengecekan terhadap setiap data latih yang apabila memiliki kata kurang bermakna yang telah didaftar, maka akan dihapus.

### 2.6.1.4 Stemming

Satu kata dapat memiliki beberapa bentuk seperti “kunjung”, “berkunjung”, dan “kunjungan”. Tahap *stemming* bertujuan untuk mengubah kata berimbuhan menjadi kata dasarnya dengan cara memotong imbuhan (Manning, 2009). Implementasi pada tahap *stemming* memiliki berbagai metode yang dapat digunakan bergantung pada bahasa dari dokumen. Salah satu contoh *stemming* dapat dilihat adalah merubah kata “kunjungan” dan “berkunjung” menjadi “kunjung”.

## 2.7 Pembobotan TF-IDF

TF-IDF atay *Term Frequency Inverse Document* merupakan algoritme untuk mendapatkan bobok komposit dengan cara menggabungkan *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF). *Term frequency* (TF) sendiri merupakan pembobotan dengan cara menghitung kemunculan kata (*term*) di dalam sebuah dokumen. Sedangkan *Inverse Document Frequency* (IDF) adalah sebuah mekanisme untuk meminimalisir efek dari kata (*term*) yang terlalu sering muncul di dalam keseluruhan dokumen sehingga menemukan relevansi antar kata (*term*) dengan lebih baik. Berikut merupakan persamaan dari pembobotan dengan mekanisme TF-IDF.

$$\sum tf_{t,d} \times \log \frac{N}{df_t} = \sum tfidf_{t,d} \quad (2.1)$$

Keterangan:



$tf_{t,d}$  : Kemunculan kata (t) pada sebuah dokumen (d)

$N$  : Jumlah total sebuah koleksi dokumen

$df_t$  : Frekuensi dokumen di mana kata (t) muncul

$idf_t$  : Logaritma dari perbandingan  $N$  dan  $df_t$

## 2.8 Analisis Sentimen

Analisis sentimen atau disebut juga *opinion mining* merupakan suatu teknik yang digunakan untuk menganalisis pendapat, sentimen, emosi, dan perilaku manusia terhadap entitas seperti, produk, layanan, organisasi, atau topik tertentu (Liu, 2012). Analisis sentimen dapat dilakukan untuk berbagai kepentingan, di antaranya mengetahui opini pembeli mengenai suatu produk, mengetahui opini masyarakat mengenai kandidat pemilu, dan berbagai kepentingan lainnya yang melibatkan opini manusia. Pendapat, sentimen, emosi, dan perilaku manusia didapatkan melalui berbagai cara seperti survey, wawancara, pemungutan suara, dan *review* yang terdapat di media sosial.

Terdapat tiga tingkatan analisis sentimen, yaitu *document level*, *sentence level*, dan *entity and aspect level*. Pada *document level* klasifikasi dilakukan untuk menentukan apakah keseluruhan opini dalam dokumen mengekspresikan sentimen positif atau negatif. Sebagai contoh, pada *review* produk sistem menentukan apakah *review* yang didapat secara keseluruhan mengekspresikan opini negatif atau positif mengenai produk. Maka dari itu, tingkat ini tidak dapat diaplikasikan untuk mengevaluasi atau membandingkan beberapa entitas. Pada *sentence level* ditentukan apakah setiap kalimat mengekspresikan opini positif, negatif, atau netral. Sedangkan pada *entity and aspect level* yang sebelumnya dikenal sebagai *feature level* tidak mencari susunan kata, tetapi langsung mencari opini itu sendiri. Hal tersebut didasarkan pada ide bahwa sebuah opini terdiri dari sentimen (positif atau negatif) dan sebuah target (Liu, 2012).

## 2.9 Machine Learning

*Machine learning* merupakan aplikasi dari algoritme yang dibuat agar komputer dapat mempelajari perilaku suatu sistem dan meningkatkan kemampuan belajar komputer secara otomatis (Mirjalili dan Raschka, 2018). Han et al. (2012) menggambarkan beberapa permasalahan dalam machine learning yang bersangkutan dengan *data mining*:

### a. Supervised Learning

*Supervised learning* dapat dikatakan sebagai sinonim dari klasifikasi. Proses belajar dilakukan dengan cara menggunakan data latih yang telah diberi label sehingga dapat dikatakan *supervised*.

### b. Unsupervised Learning

*Unsupervised learning* merupakan sinonim dari *clustering*. Proses belajar yang dilakukan tidak diawasi karena masukan data tidak memiliki label kelas. Biasanya *clustering* digunakan untuk menemukan kelas dalam data.

### c. Semi-supervised Learning

*Semi-supervised learning* adalah teknik *machine learning* yang menggunakan kedua data yang memiliki label dan data yang tidak memiliki label ketika mempelajari sebuah model. Data yang memiliki label pada umumnya digunakan untuk mempelajari model kelas dan data yang tidak memiliki label digunakan untuk memperbaiki batasan antar kelas.

#### d. *Active Learning*

*Active learning* memungkinkan pengguna untuk berperan secara aktif dalam proses pembelajaran. Pengguna diminta untuk memberi label pada data yang sebelumnya belum diberi label atau data yang dibentuk oleh program. Tujuan dari *active learning* adalah untuk mengoptimalkan kualitas model dengan cara mendapatkan pengetahuan dari pengguna dengan batasan banyaknya data yang akan diberi label oleh pengguna.

### 2.9.1 *Naïve Bayes Classifier*

*Naïve Bayes* merupakan salah satu algoritme *supervised learning* yang paling sederhana yang dibangun berdasarkan Teorema Bayes. Metode klasifikasi ini diklaim bekerja secara cepat, akurat dan handal, bahkan untuk *dataset* dengan ukuran besar Berdasarkan Teorema Bayes ini, efek dari fitur-fitur tertentu yang mungkin saling berelasi dalam suatu kelas, dipandang bersifat bebas atau independen dari fitur-fitur di kelas lainnya. Dari asumsi inilah kata *naive* digunakan (Navlani, 2018). Secara umum proses dari klasifikasi *Naïve Bayes* dapat dilihat pada Persamaan 2.1.

$$P(c_j|w_i) = \frac{P(c_j)P(w_i|c_j)}{P(w_i)} \tag{2.1}$$

Keterangan:

$P(c_j|w_i)$  : Peluang kelas  $j$  ketika kata  $i$  muncul

$P(w_i|c_j)$  : Peluang kata  $i$  masuk ke kelas  $j$

$P(c_j)$  : Peluang kemunculan kelas  $j$

$P(w_i)$  : Peluang kemunculan kata  $i$

Persamaan 2.1 dapat disederhanakan dengan asumsi fitur-fitur saling independen satu kelas dan lainnya, maka peluang kemunculan kata dapat dihilangkan. Proses klasifikasi dapat dilakukan seperti Persamaan 2.2.

$$P(c_j|w_i) = P(c_j) P(w_i|c_j) \tag{2.2}$$

Perhitungan *posterior probability* merupakan perhitungan yang dilakukan dengan mengalikan prior dengan total *conditional probability*. Rumus perhitungan *posterior probability* dapat dilihat pada Persamaan 2.3.

$$P(c_j|w_i) = P(c_j) \times P(w_1|c_j) \times P(w_2|c_j) \dots P(w_n|c_j) \tag{2.3}$$

### 2.9.2 Multinomial Naïve Bayes

Terdapat tiga metode *Naïve Bayes classifier*, yaitu *Gaussian Naïve Bayes*, *Bernoulli Naïve Bayes*, dan *Multinomial Naïve Bayes*. *Gaussian Naïve Bayes* dapat diaplikasikan pada data yang bersifat *continuous*, sementara itu *Bernoulli Naïve Bayes* diaplikasikan pada data biner. *Multinomial Naïve Bayes* sesuai digunakan untuk mengklasifikasikan fitur diskrit. Sebagai contoh *Multinomial Naïve Bayes* menghitung jumlah kemunculan data salah satunya kemunculan kata dalam suatu kalimat dalam *text classification* (Muller dan Guido, 2017). Persamaan *Multinomial Naïve Bayes* diberikan pada persamaan 2.4.

$$P(w_i|c_j) = \frac{\text{count}(w_i, c_j) + 1}{(\sum_{w \in V} \text{count}(w, c_j) + |V|)} \quad (2.4)$$

Keterangan:

- $P(w_i|c_j)$  : Peluang kata  $i$  masuk ke kelas  $j$
- $\text{count}(w_i, c_j)$  : Jumlah suatu kata yang muncul dalam suatu kelas dan penambahan angka 1 digunakan untuk menghindari nilai nol.
- $\sum_{w \in V} \text{count}(w, c_j)$  : Jumlah dari seluruh kata yang ada pada kelas  $c_j$ .
- $|V|$  : Jumlah dari seluruh kata unik pada seluruh kelas.

### 2.10 Evaluasi

**Tabel 2.1 Confusion Matrix**

	Terklasifikasi Positif	Terklasifikasi Negatif
Positif	<i>True positives (tp)</i>	<i>False positives (fp)</i>
Negatif	<i>False negatives (fn)</i>	<i>True negatives (tn)</i>

Sumber: Alpaydin (2010)

Hasil klasifikasi kemudian dievaluasi dengan cara membandingkan hasil dari pelabelan yang dilakukan secara manual dengan standar yang ditetapkan dengan cara menghitung *accuracy*, *precision*, *recall*, dan *F-measure* yang diukur melalui *confusion matrix* seperti Tabel 2.2 karena memiliki dua kelas, yaitu positif dan negatif (Manning et al., 2009). *Accuracy* digunakan untuk mengukur ketepatan klasifikasi yang dilakukan *text classifier*. Efektivitas dari *text classifier* dapat dievaluasi melalui *precision* ( $p$ ) dan *recall* ( $r$ ) yang dihasilkan. *Precision* diukur melalui persentase dari dokumen yang diklasifikasikan ke kelas  $c$ . Sedangkan *recall* diukur melalui persentase dokumen dari kelas  $c$  oleh *classifier*. Berdasarkan definisi tersebut *precision* dan *recall* berlawanan satu dengan lainnya, maka dari itu digunakan *F-measure* untuk menyeimbangkan keduanya (Berry et al., 2010).

$$\text{accuracy} = \frac{tp+tn}{tp+fp+fn+tn} \quad (2.6)$$



$$p = \frac{tp}{tp+fp} \tag{2.7}$$

$$r = \frac{tp}{tp+fn} \tag{2.8}$$

$$F_1 = \frac{2pr}{p+r} \tag{2.9}$$

Keterangan:

*True positive (tp)* : data yang terklasifikasi positif secara tepat oleh sistem.

*False positive (fp)* : data yang terklasifikasi positif secara salah oleh sistem.

*True negative (tf)* : data yang terklasifikasi negatif secara tepat oleh sistem.

*False negative (fn)* : data yang terklasifikasi negatif secara salah oleh sistem.

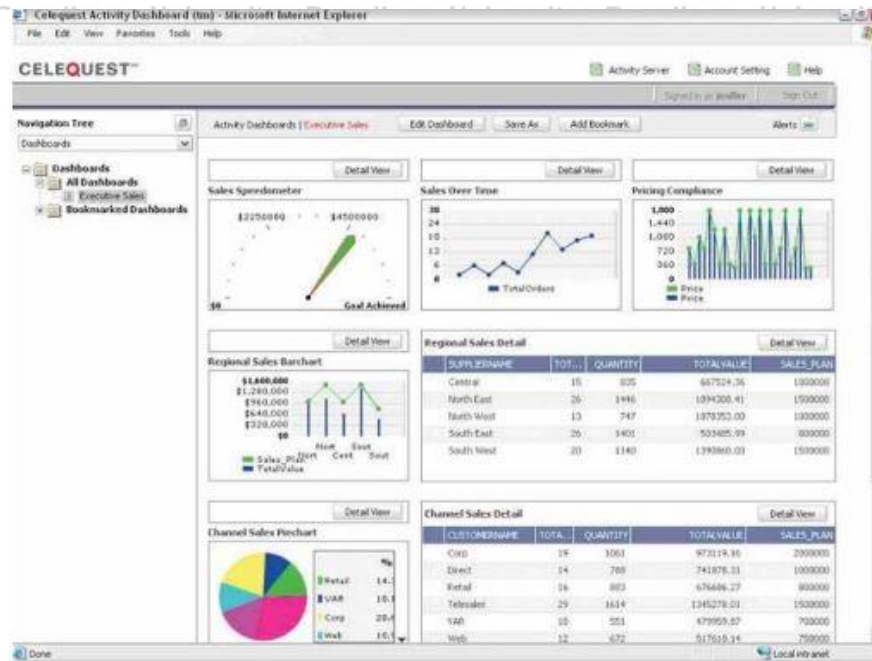
## 2.11 Visualisasi Data

### 2.11.1 Dashboard

*Dashboard* merupakan tampilan visual dari informasi-informasi terpenting yang diperlukan untuk mencapai satu atau lebih tujuan yang dapat diatur ke dalam satu layar komputer sehingga dapat dipantau dengan sekali lihat (Few, 2006).

*Dashboard* dipresentasikan dalam bentuk kumpulan grafik karena jika ditangani secara ahli *dashboard* dapat mengkomunikasikan informasi dengan efisiensi lebih baik dibandingkan hanya dengan tulisan saja seperti pada Gambar 2.2. Informasi yang didapat dari grafik-grafik pada *dashboard* harus dapat diekstrak dan dipahami secara tepat oleh otak manusia. Maka dari itu, untuk mendesain *dashboard* secara efektif berikut beberapa hal yang harus diperhatikan mengenai esensi *dashboard* (Few, 2006).

1. **Dashboard menampilkan informasi yang dibutuhkan untuk mencapai tujuan spesifik.** Untuk mencapai suatu tujuan sering kali dibutuhkan akses ke kumpulan informasi yang tidak berelasi langsung dengan tujuan dan berasal dari sumber yang berbeda yang berhubungan dengan serangkaian fungsi bisnis. Informasi yang dibutuhkan bisa dan seringkali merupakan KPI.
2. **Dashboard dapat ditampilkan dalam satu layar komputer.** Informasi harus dapat cukup ke dalam satu layar sehingga terjangkau dalam pengelihatannya pengguna *dashboard*.
3. **Dashboard digunakan untuk memonitor dalam sekali pandang.** Sebuah *dashboard* harus dapat menunjukkan informasi yang butuh perhatian khusus.

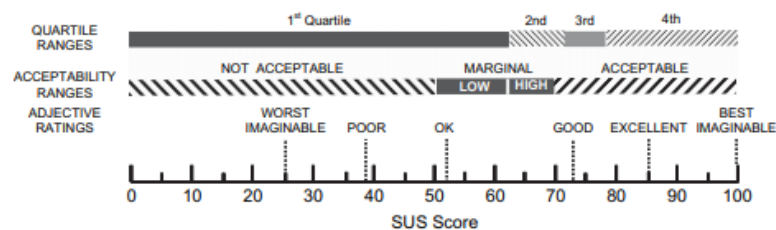


Gambar 2.2 Contoh Dashboard

Sumber: Few (2006)

## 2.12 System Usability Scale (SUS)

Untuk menguji kegunaan dari suatu sistem informasi dilakukan pengujian dengan menggunakan satu alat uji berupa *system usability scale* (SUS). SUS sendiri tidak hanya digunakan untuk menguji sistem informasi, tetapi dapat juga digunakan untuk menguji situs web, telepon seluler, aplikasi TV, dan lainnya (Bangor et al., 2008). Hasil pengujian menggunakan SUS berupa skala angka 0 sampai 100. Semakin besar nilai angka yang dihasilkan, maka semakin besar kegunaan dari produk yang diuji. Untuk mempermudah masyarakat umum dalam menginterpretasikan nilai hasil SUS, skala 0 sampai 100 dibagi ke dalam tiga kelas *Acceptability*, yaitu *not acceptable*, *marginal*, dan *acceptable* seperti pada Gambar 2.3.



Gambar 2.3 Skor System Usability Scale

Sumber: Bangor et al. (2008)

Nilai SUS didapatkan sebagai berikut.

1. Pengguna dari sistem diminta untuk mengisi kuesioner yang terdiri dari sepuluh pertanyaan tentang kegunaan dan kemudahan menggunakan sistem yang diuji. Sepuluh pertanyaan tersebut lima di antaranya bersifat pertanyaan negatif dan sisanya berupa pertanyaan positif. Untuk tiap pertanyaan disediakan lima pilihan jawaban angka yang bernilai bilangan bulat dari 0 sampai 4 (bisa diganti dengan interval lainnya). Pada kondisi ideal, di mana pengisi kuesioner menganggap sistem berguna dan mudah digunakan, pengisi kuesioner akan memberi nilai semakin besar untuk pertanyaan positif dan memberi nilai semakin kecil pada pertanyaan negatif.

2. Hasil pengisian kuesioner kemudian dijadikan masukan untuk rumus penentuan SUS yang sesuai dengan skala pilihan jawaban. Sebagai contoh untuk skala pilihan jawaban 0 sampai 4 digunakan rumus:

$$SUS = 2.5 * (\sum_{i=1}^5 (Pos_i - 1)) + \sum_{j=1}^5 (5 - Neg_j) \quad (2.5)$$

Keterangan:

$Pos_i$  : Nilai jawaban pertanyaan positif ke-i

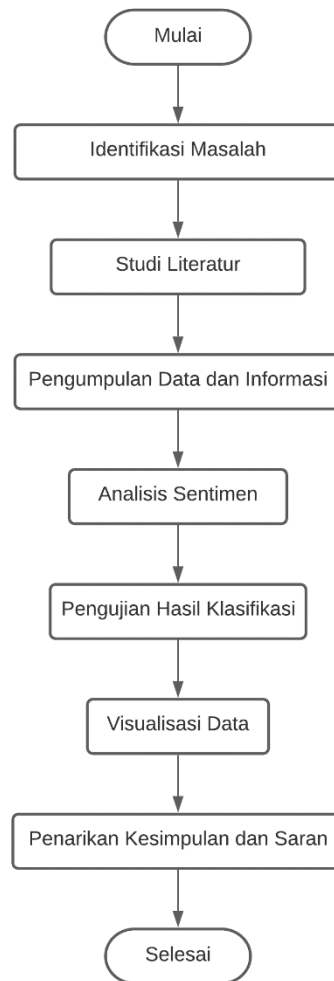
$Neg_i$  : Nilai jawaban pertanyaan negatif ke-i





## BAB 3 METODOLOGI

Penelitian dilakukan dengan mengikuti langkah-langkah yang terstruktur supaya dapat dilakukan secara terarah dan baik. Gambar 3.1 menunjukkan alur metodologi penelitian yang dilaksanakan. Penelitian dimulai dengan mengidentifikasi masalah, melakukan studi literatur, mengumpulkan data dan informasi, analisis sentimen, melakukan analisis dan pengujian hasil klasifikasi, visualisasi data, dan diakhiri dengan pengambilan kesimpulan dan saran.



Gambar 3.1 Alur Metodologi Penelitian

### 3.1 Identifikasi Masalah

Tahap awal yang dilakukan, yaitu melakukan identifikasi masalah. Pada tahap ini terjadi proses menemukan masalah-masalah terkait objek penelitian, yaitu ulasan terhadap objek wisata yang dikelola oleh Perum Perhutani. Selanjutnya permasalahan yang ditemukan disusun menjadi pertanyaan rumusan masalah. Permasalahan tersebut dijawab melalui penelitian ini dan menghasilkan kesimpulan dan saran di akhir penelitian.



### 3.2 Studi Literatur

Tahap selanjutnya dari penelitian ini adalah melakukan studi literatur. Literatur pada penelitian ini berperan sebagai sumber teori dan rujukan sehingga penelitian berjalan secara terarah. Sumber literatur yang digunakan, antara lain berasal dari jurnal terkait penelitian serupa yang sudah ada, buku, situs web, dan sumber literatur lainnya yang dibutuhkan selama berjalannya penelitian. Sumber literatur juga berfungsi untuk memperjelas teori ataupun istilah yang digunakan pada penelitian ini.

### 3.3 Pengumpulan Data dan Informasi

Setelah melakukan identifikasi masalah untuk melanjutkan penelitian dibutuhkan data dan informasi yang mendukung jalannya penelitian. Dilakukan wawancara dengan cara menyusun pertanyaan-pertanyaan terstruktur yang ditanyakan kepada narasumber. Narasumber pada penelitian kali ini adalah Manajer Keuangan, SDM, dan Umum Perum Perhutani yang berperan dalam pengelolaan objek-objek wisata. Data dan informasi yang dikumpulkan terkait dengan kepentingan berjalannya penelitian dan dokumentasi penelitian.

Selain data dan informasi yang terkumpul melalui wawancara narasumber, terdapat data penting lain yang dibutuhkan. Data tersebut berupa komentar-komentar wisatawan yang diambil dari salah satu situs web yang mengulas objek-objek wisata yang dikelola oleh Perum Perhutani, yaitu *TripAdvisor*. Ulasan wisatawan tersebut diambil menggunakan cara yang disebut *web scraping* menggunakan *tools WebHarvy*. Data ini digunakan agar bisa melakukan tahapan selanjutnya, yaitu analisis sentimen.

### 3.4 Analisis Sentimen

Pada tahap ini dibutuhkan masukan berupa data yang telah didapatkan pada tahap *web scraping*. Data mentah berupa ulasan tersebut melalui *preprocessing* dan *aspect extraction* sehingga didapatkan aspek-aspek apa saja yang ada pada ulasan dari tiga aspek (*Attractions*, *Accessibilities*, dan *Amenities*) yang sudah ditentukan berdasarkan hasil wawancara. Dikarenakan data yang dimiliki berupa teks, maka sebelum dilakukan klasifikasi dilakukan konversi teks menjadi angka supaya data bisa diolah untuk proses klasifikasi. Setelah itu dilakukan klasifikasi menggunakan metode *Multinomial Naïve Bayes* untuk menentukan sentimen wisatawan kepada kelas positif, netral, atau negatif.

### 3.5 Analisis dan Pengujian Hasil Klasifikasi

Hasil klasifikasi mengenai ulasan wisatawan dianalisis untuk mengetahui performa model yang dibentuk. Terdapat beberapa pengukuran yang dilakukan, yaitu *accuracy*, *precision*, *recall*, dan *f1 score*. *Accuracy* menunjukkan keseluruhan performa model. Sedangkan *precision* dan *recall* dibutuhkan untuk memilih model terbaik berdasarkan tingginya jumlah *false positive* untuk *precision* atau *false*

*negative* untuk *recall*. Terakhir, *f1 score* dibutuhkan untuk mencari keseimbangan antara *precision* dan *recall*.

### 3.6 Visualisasi Data

Data yang digunakan sebagai sumber penelitian dan data yang didapat dari penelitian disajikan ke dalam bentuk *dashboard*. Data-data tersebut disajikan ke dalam grafik-grafik yang membentuk informasi terkait objek-objek wisata yang dikelola oleh Perum Perhutani. *Dashboard* ini diperuntukkan oleh pihak manajemen Divisi SDM, Keuangan, dan Umum yang berperan dalam mengelola objek-objek wisata yang dikelola Perum Perhutani. Sehingga *dashboard* yang disusun sesuai dengan kebutuhan dengan pemegang kepentingan tersebut.

### 3.7 Penarikan Kesimpulan dan Saran

Tahap terakhir dari penelitian ini adalah penarikan kesimpulan dari hasil penelitian yang telah dilakukan. Penarikan kesimpulan dilakukan supaya bisa menjawab pertanyaan rumusan masalah yang sudah dikemukakan sebelumnya. Selain kesimpulan terdapat beberapa saran yang diberikan untuk penelitian serupa selanjutnya berdasarkan penelitian yang dilakukan kali ini.





## BAB 4 PERANCANGAN

Pada bab ini ditampilkan perancangan yang dibutuhkan untuk tahap implementasi, yaitu pengambilan data, penentuan aspek, dan desain antarmuka *dashboard*.

### 4.1 Web Scraping

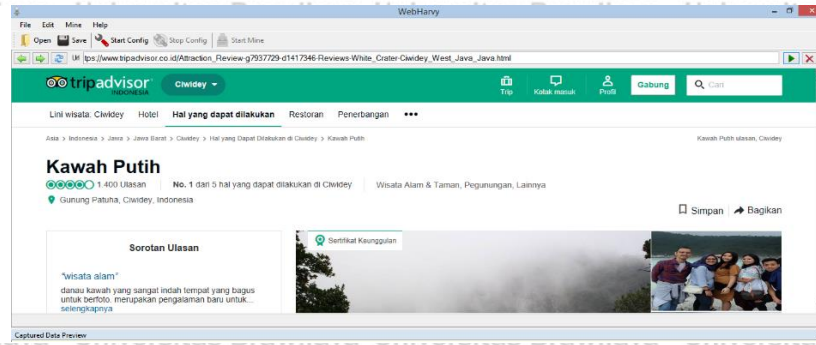
*Web scraping* dilakukan untuk mengumpulkan data-data yang dibutuhkan dari situs TripAdvisor. Detail data yang di-*scraping* tertera di Tabel 4.1, yaitu ulasan yang ditulis oleh pengguna TripAdvisor bertipe data *string* yang disimpan ke dalam variabel bernama Ulasan. Selain ulasan dibutuhkan variabel Nama dengan tipe data *string* digunakan untuk menyimpan nama-nama pengguna TripAdvisor yang menuliskan ulasan mengenai objek-objek wisata yang dikelola oleh Perum Perhutani. Variabel terakhir, yaitu variabel Tanggal digunakan untuk menyimpan tanggal dituliskannya ulasan yang didapat dari proses *web scraping*.

Tabel 4.1 Variabel dan tipe data pada proses *web scraping*

Variabel	Tipe Data	Keterangan
Ulasan	<i>String</i>	Teks mengenai ulasan objek wisata yang ditulis pengguna TripAdvisor
Nama	<i>String</i>	Nama pengguna yang tertera di situs web TripAdvisor
Objek	<i>String</i>	Nama objek wisata yang diulas
Tanggal	<i>Date</i>	Tanggal pengguna menuliskan ulasan di situs web TripAdvisor

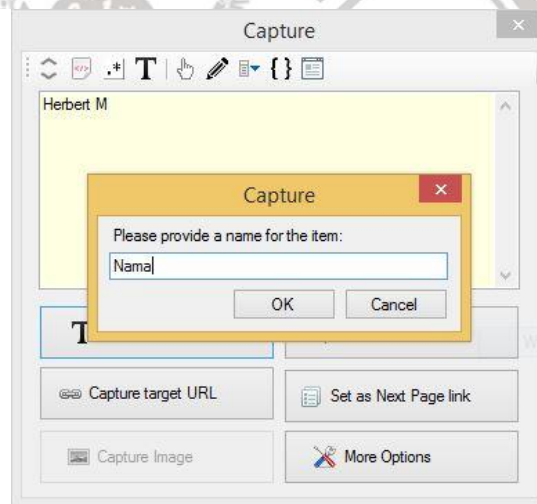
Pada penelitian ini *web scraping* dilakukan menggunakan bantuan *tools* WebHarvy. WebHarvy merupakan sebuah perangkat lunak yang dikembangkan oleh Sysnucleus yang mempermudah proses *web scraping* sehingga dapat dioperasikan melalui *Graphical User Interface* (GUI). Untuk mendapatkan data yang dibutuhkan menggunakan WebHarvy perlu dilakukan langkah-langkah tertentu. Berikut merupakan tahapan *web scraping* menggunakan WebHarvy.

1. Menuliskan **Uniform Resource Locator (URL)** dari halaman TripAdvisor yang datanya ingin diakuisisi pada kolom URL seperti pada gambar 4.1 sehingga bisa mengakses halaman situs web melalui WebHarvy.



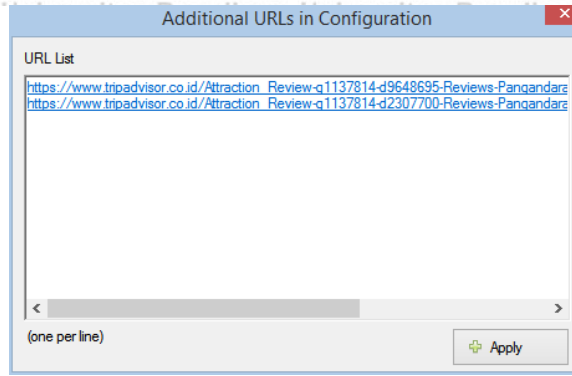
Gambar 4.1 Mengakses situs web TripAdvisor pada WebHarvy

2. Melakukan konfigurasi untuk mengatur elemen-elemen yang digunakan dari situs web TripAdvisor untuk kepentingan penelitian. Pada *web scraping* diambil nama pengguna, ulasan pengguna, nama objek wisata, dan tanggal menulis ulasan dari situs TripAdvisor. Konfigurasi dimulai dengan menekan *Start Config*. Selanjutnya untuk mendapatkan data-data tersebut klik elemen tempat data tertera dan menyimpannya ke dalam variabel data seperti pada Gambar 4.2.



Gambar 4.2 Menyimpan data ke variabel

3. Untuk mengambil data pada objek wisata lainnya, maka pilih *Add/Remove URLs from Configuration* dan akan muncul jendela seperti Gambar 4.6 untuk mengisi tautan dari halaman-halaman yang datanya ingin diakuisisi. Pilih *Apply* ketika semua URL telah tertulis. Berdasarkan hasil wawancara, Perhutani mencukupkan ulasan objek-objek wisata yang diambil adalah dari objek wisata yang tergolong kepada Standar Usaha Taman Rekreasi (wisata alam). Daftar tautan dari objek-objek wisata yang diambil dapat dilihat pada Tabel 4.2.



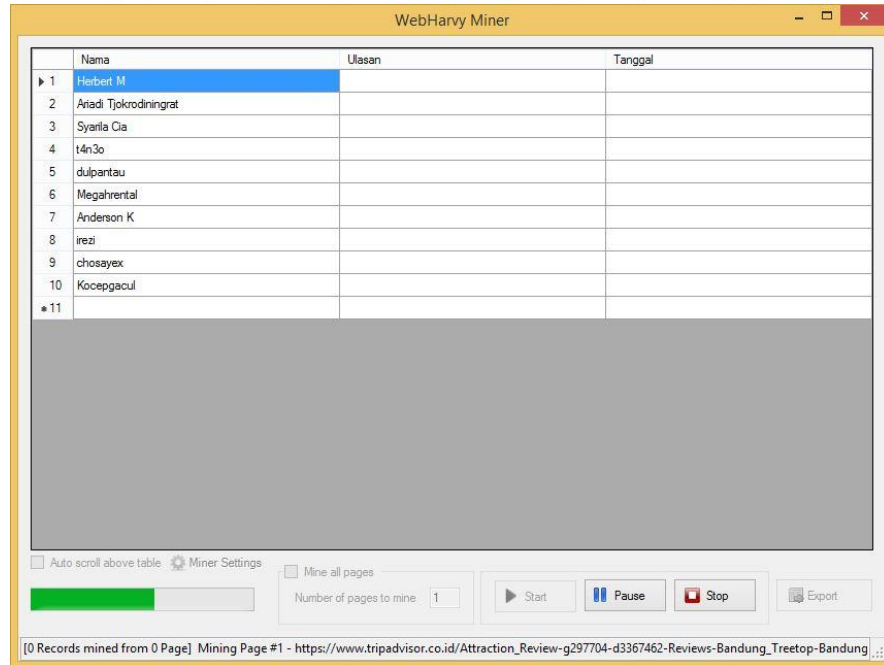
Gambar 4.3 Menambahkan URL objek wisata lainnya

Tabel 4.2 Daftar link objek wisata pada TripAdvisor

No.	Link	Objek Wisata
1.	<a href="https://www.tripadvisor.co.id/Attraction_Review-g1237079-d8746462-Reviews-Paralayang_Batu-Batu_East_Java_Java.html">https://www.tripadvisor.co.id/Attraction_Review-g1237079-d8746462-Reviews-Paralayang_Batu-Batu_East_Java_Java.html</a>	Paralayang
2.	<a href="https://www.tripadvisor.co.id/Attraction_Review-g7937729-d1417346-Reviews-White_Crater-Ciwidey_West_Java_Java.html">https://www.tripadvisor.co.id/Attraction_Review-g7937729-d1417346-Reviews-White_Crater-Ciwidey_West_Java_Java.html</a>	Kawah Putih
3.	<a href="https://www.tripadvisor.co.id/Attraction_Review-g297704-d7134067-Reviews-Dusun_Bambu_Family_Leisure_Park-Bandung_West_Java_Java.html">https://www.tripadvisor.co.id/Attraction_Review-g297704-d7134067-Reviews-Dusun_Bambu_Family_Leisure_Park-Bandung_West_Java_Java.html</a>	Dusun Bambu
4.	<a href="https://www.tripadvisor.co.id/Attraction_Review-g1144644-d1231258-Reviews-Madakaripura_Waterfall-Probolinggo_East_Java_Java.html">https://www.tripadvisor.co.id/Attraction_Review-g1144644-d1231258-Reviews-Madakaripura_Waterfall-Probolinggo_East_Java_Java.html</a>	Air Terjun Madakaripura
5.	<a href="https://www.tripadvisor.co.id/Attraction_Review-g6643743-d446941-Reviews-Mount_Bromo-Bromo_Tengger_Semeru_National_Park_East_Java_Java.html#REVIEWS">https://www.tripadvisor.co.id/Attraction_Review-g6643743-d446941-Reviews-Mount_Bromo-Bromo_Tengger_Semeru_National_Park_East_Java_Java.html#REVIEWS</a>	Gunung Bromo
...	...	...
10.	<a href="https://www.tripadvisor.co.id/Attraction_Review-g2699971-d3267606-Reviews-Taman_Safari_Prigen-Prigen_East_Java_Java.html">https://www.tripadvisor.co.id/Attraction_Review-g2699971-d3267606-Reviews-Taman_Safari_Prigen-Prigen_East_Java_Java.html</a>	Taman Safari Prigen
11.	<a href="https://www.tripadvisor.co.id/Attraction_Review-g1052738-d3295861-Reviews-Tanjung_Papuma-Jember_East_Java_Java.html">https://www.tripadvisor.co.id/Attraction_Review-g1052738-d3295861-Reviews-Tanjung_Papuma-Jember_East_Java_Java.html</a>	Tanjung Papuma

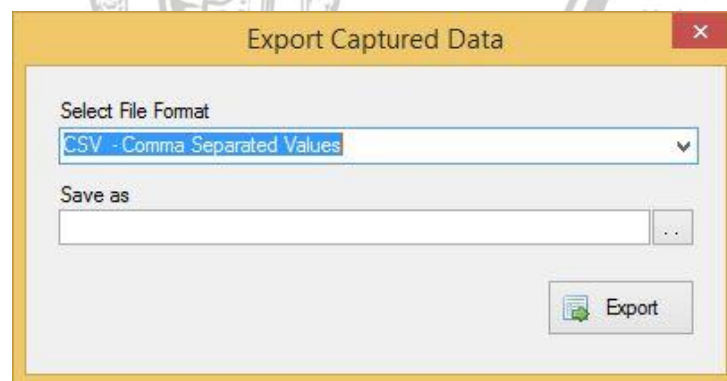


4. **Proses *scraping*** dapat dijalankan saat konfigurasi selesai dilakukan. Klik *Stop Config* ketika selesai lalu klik *Start-Mine* untuk memulai proses *web scraping* di WebHarvy. Pada Gambar 4.4 dapat dilihat proses *scraping* data yang sudah dikonfigurasi sebelumnya.



**Gambar 4.4 Proses *scraping* data yang telah dikonfigurasi**

5. **Simpan data** setelah proses *scraping* selesai. Klik *Export* untuk menyimpan data ke dalam ekstensi csv seperti Gambar 4.5 sehingga seluruh data yang tersimpan dalam variabel selama proses *scraping* dapat diakses untuk kepentingan tahap penelitian selanjutnya. Beberapa contoh ulasan dari hasil web *scraping* dapat dilihat pada Tabel 4.3.



**Gambar 4.5 Menyimpan data hasil *scraping***

Tabel 4.3 Contoh hasil *web scraping*

Nama	Ulasan	Tanggal	Objek
Yazchenko	Dari beberapa air terjun yang mudah dijangkau dari pusat kota Bandung, Curug Cimahi itu paling rekomen. Tiket masuknya murah, akses menuju sini mudah, fasilitas juga lengkap sudah, dan yang paling utama, pemandangan di sini indah! Kalau malam Curug Cimahi bertransformasi jadi Curug Pelangi. Kalau berkunjung ke sini, sekalian berkunjung ke Dusun Bambu, tempat rekomended lainnya!	Nov-15	Curug Cimahi
Neng Yofie	Buat kesini kita hrs melewati 500an lbh anak tangga.. Pas pergi sih turun ga kerasa.. Pulangnya lumayan menguras energi, jd yg mau bawa orang tua mungkin jng sampai bawah bgt.. Kasian.. Tapi ada tempat menunggu yg dibikin baru.. Ada maksimum orang yg bisa duduk disitu mengingat bawahnya jurang.. Pas sampai dibawah keren banget. Apalagi menuju malam.. Airnya super dingin.. Batunya aga licin jd hati2 ya..tiketnya murmer	Oct-15	Curug Cimahi
Priojati A	sangat recomended tuk pergi kesini, sangat tersa unsur perdesaan dan perkebunan serta mencicipi berbagai kuliner enak khas dari kota bandung	Dec-18	Dusun Bambu
abdillahhw	Krn aq suka icip2 makanan, jadi nya menurut aq wisatanya sih biasa aja. Kesan pertama, okeeeh sejuk dingin. Sisanya yg q ingat cemilan dan makanan nya enak ??????. Buat yg bawa anak dan keluarga.seru laah disini.	Oct-18	Dusun Bambu
armandhoFC	Akhirnya mencoba penginapan disini dengan tema "Glamping" ... tenda nya cukup bagus dan luas, juga terdapat kamar mandi dibawah dan juga tempat barbeque .... cocok buat banyak teman ... tempat tidur nya kuueerraasss .... bangun 2x siap 2x sakit pinggang ... wkwwkkk ... makan pagi di burarang, view enak .... makananya cukup variasi ... tentunya disini juga jangan lupa ikuti dengan permainan yg disediakan	Mar-18	Dusun Bambu

## 4.2 Penentuan Aspek

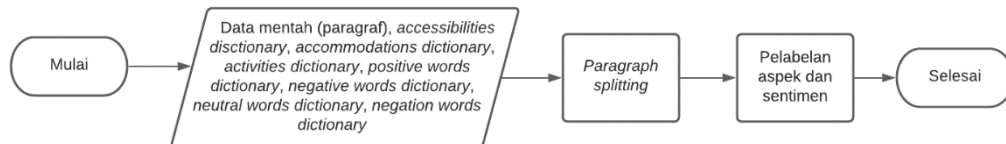
Data mentah yang didapatkan melalui proses *scraping* belum langsung bisa diimplementasikan ke tahap klasifikasi. Salah satu tahap yang harus dilakukan, yaitu pelabelan aspek. Pelabelan aspek dilakukan untuk mengidentifikasi aspek-aspek yang terdapat dalam sebuah ulasan. Terdapat lima aspek pariwisata, yaitu *attractions*, *activities*, *accessibilities*, *accommodation*, dan *amenities* (Cooper et al., 1993). Pada penelitian ini dipilih aspek *attractions*, *accessibilities*, dan *amenities* seperti yang terlihat pada Tabel 4.4 berdasarkan jumlah ulasan terbanyak.

**Tabel 4.4 Kategori aspek objek wisata**

Aspek	Keterangan
<i>Attractions</i>	Ulasan yang berkaitan dengan atraksi atau daya tarik utama objek wisata
<i>Accessibilities</i>	Ulasan yang berkaitan dengan akses yang menuju/berada di objek wisata
<i>Amenities</i>	Ulasan yang berkaitan dengan sarana pelengkap yang tersedia di lokasi objek wisata

Sumber: Cooper et al. (1993)

## 4.3 Pelabelan Aspek dan Sentimen



**Gambar 4.6 Tahap pelabelan data**

Proses pelabelan data dilakukan dengan algoritme yang dibuat sendiri oleh peneliti, yaitu diawali dengan pembuatan *dictionary*, *paragraph splitting*, dan diakhiri dengan pelabelan aspek seperti pada Gambar 4.6. Maka dari itu, suatu algoritme dibuat untuk melabeli *dataset* merujuk ke pedoman yang diberikan oleh NLP Stanford (2009). Pertama-tama, disusun beberapa *dictionary* untuk pelabelan aspek dan sentimen. *Dictionary* untuk mengidentifikasi aspek dilakukan berdasarkan *keywords* yang diberikan oleh pihak Perhutani. Saat melakukan wawancara, Perhutani memberikan dokumen *Tools Assessment* Wisata yang berisi kata kunci untuk aspek *attractions*, *accessibilities* dan *amenities* yang dapat dilihat pada Gambar 4.7.



Attractions		Accessibilities	Amenities
A	Naturally		
1	Beaches	Distance	Public toilet
2	Canyons and Gorges	Traveling time	Rest area
3	Caves	Transport	Parking
4	Cliffs	e Marketing	Health clinic
5	Climate	Booking online	sarana ibadah
6	Exceptionally clear night views for star-gazing	Accessible to informatio	Guest room
7	Farms	Roads	Spa
8	Fishing Streams	Lapangan terbang	Swimming pool
9	Gorges	Pelabuhan	Welcome drink
10	High velocity or unusual wind behavior		Fitness center
11	Lakes		Valet service
12	Light hitting		Room service
13	Marinas		Laundry service
14	Meteor Showers		Bar
15	Mountain view		Wifi
16	Mountains		Complimentary drink
17	Nature preserve		Toiletries
18	Nature Trails		Bike

**Gambar 4.7** Contoh kata kunci aspek dalam dokumen *Tools Assessment Wisata*

Selanjutnya dibuat kata kunci terkait sentimen berdasarkan program pelabelan sentimen oleh NLP Stanford (2009) yang diterjemahkan ke bahasa Indonesia. Dari hasil pendataan kata kunci didapatkan tujuh *dictionaries*, yaitu *attraction dictionary*, *accessibilities dictionary*, *amenities dictionary*, *positive words dictionary*, *negative words dictionary*, *neutral word dictionary*, dan *negation words dictionary* yang contohnya dapat dilihat pada Tabel 4.5. Kemudian dilakukan *paragraph splitting* yang bertujuan untuk memecah ulasan-ulasan berbentuk paragraph (terdiri lebih dari satu kalimat utama) ke dalam kalimat utama sehingga proses pelabelan diharapkan lebih akurat. Algoritme disusun untuk mendeteksi kata kunci pada *dictionary* yang terdapat dalam kalimat-kalimat ulasan. Hasil pelabelan aspek tertera pada Tabel 4.6, yaitu sebanyak 4647 kalimat tidak memiliki aspek, terdeteksi satu aspek pada 6493 kalimat, terdeteksi dua spek pada 2800 kalimat, dan 3 aspek pada 693 kalimat.

**Tabel 4.5** Contoh kata kunci sentimen

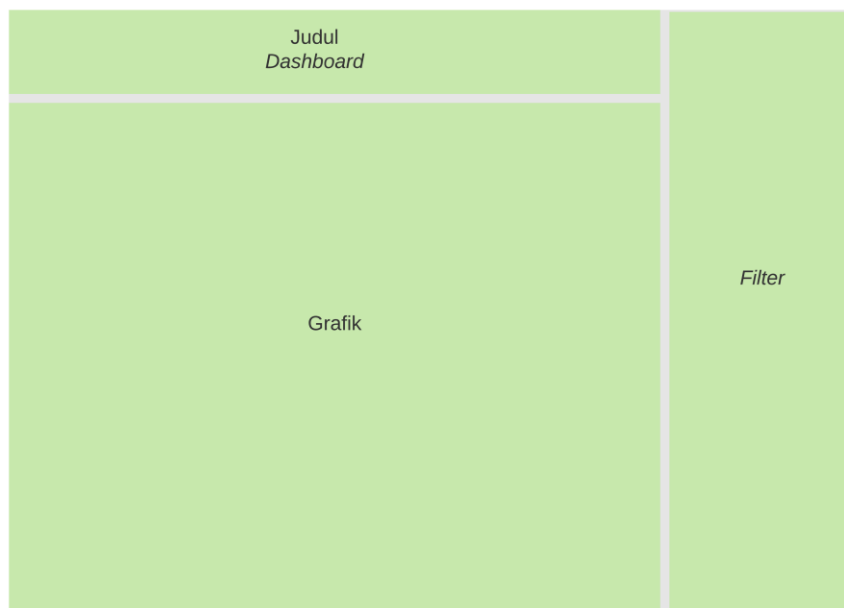
Positif	Negatif	Netral	Negasi
indah	dingin	biasa	tidak
bagus	jauh	biasa saja	tapi
baik	panas	biasa aja	namun
menikmati	mahal	overall biasa	
dingin	jauhnya	terdapat	

**Tabel 4.6** Total aspek pada kalimat

Tidak Ada Aspek	1 Aspek	2 Aspek	3 Aspek	Total
4647	6493	2800	693	14633

#### 4.4 Desain Antarmuka *Dashboard*

Visualisasi data dilakukan dengan membuat sebuah *dashboard*. Untuk membuat sebuah *dashboard* dilakukan desain antarmuka terhadap hal-hal yang akan dimasukkan ke dalam *dashboard*. Pada Gambar 4.8 terlihat terdapat tiga bagian utama yang dijadikan acuan dalam pembuatan *dashboard*. Judul *Dashboard* diperuntukkan untuk memberi keterangan terhadap pihak manajemen mengenai *dashboard* yang sedang dilihat. Grafik diperuntukkan untuk grafik-grafik yang berfungsi untuk menyampaikan informasi penting di *dashboard*. Sedangkan *filter* berfungsi untuk mengatur grafik yang ingin ditampilkan.

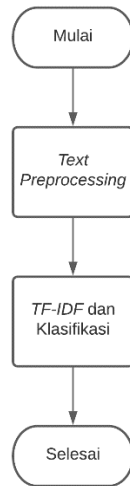


Gambar 4.8 Desain *Dashboard*



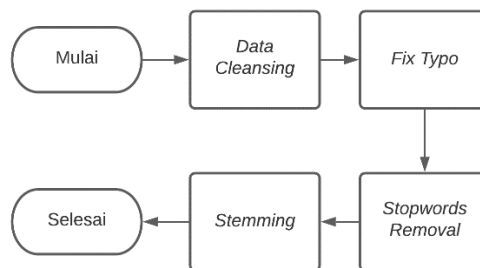
## BAB 5 IMPLEMENTASI

Bab 5 membahas implementasi rancangan analisis sentimen dan juga implementasi desain *dashboard*. Secara keseluruhan implementasi analisis sentimen mengikuti diagram alir pada Gambar 5.1.



**Gambar 5.1 Alur Analisis Sentimen**

### 5.1 Text Preprocessing



**Gambar 5.2 Proses Text Preprocessing**

*Text preprocessing* merupakan tahapan awal yang dibutuhkan dalam analisis sentimen pada penelitian ini. Terdapat empat tahap utama seperti pada Gambar 5.2, yaitu tahap *data cleansing*, *fix typo*, *stopwords removal*, dan *stemming*. *Data cleansing* dilakukan untuk menghapus dan mengubah komponen-komponen yang tidak diperlukan dari data mentah. Selanjutnya *fix typo* atau perbaikan kata yang salah ketik diperuntukkan untuk menyeragamkan kata. Tahap ketiga, yaitu *stopwords removal* atau *filtering* dilakukan untuk membuang kata-kata umum yang sering muncul yang tidak terkait dengan kebutuhan analisis sentimen. Tahap terakhir, *stemming*, dilakukan untuk mengubah kata menjadi kata dasar yang hasilnya akan dipergunakan untuk tahap selanjutnya.



### 5.1.1 Data Cleansing

Pada tahapan *data cleansing*, hal pertama yang dilakukan adalah menetapkan periode data dari data yang telah di-*scaping*. Berdasarkan hasil wawancara, Perhutani ingin melihat perkembangan hasil dari investasi dengan skala cukup besar yang dilakukannya pada tahun 2015 sehingga pada penelitian ini diambil data ulasan yang dimulai dari Januari 2015 hingga Desember 2020 untuk melihat perubahan sentimen wisatawan. Data ulasan yang sebelumnya berjumlah 3867 paragraf ulasan berkurang menjadi 2907 ulasan yang kemudian menghasilkan 14633 kalimat setelah proses *paragraph splitting*. Dari hasil tersebut dilakukan *data cleansing* kembali untuk memilih kalimat yang hanya memiliki satu aspek setelah hasil pelabelan data. Dari hasil data *cleansing* ini didapatkan *dataset* dengan detail seperti pada Tabel 5.1.

Tabel 5.1 Detail *dataset*

No.	Aspek	Jumlah Sentimen Positif	Jumlah Sentimen Netral	Jumlah Sentimen Negatif	Total
1.	<i>Attractions</i>	2078	787	682	3547
2.	<i>Accessibilities</i>	105	534	78	717
3.	<i>Amenities</i>	189	147	75	411

Memasuki algoritme pemrograman, dilakukan beberapa hal pada *dataset* untuk masing-masing aspek. Pertama, dilakukan penyeragaman seluruh bentuk huruf menjadi huruf kecil (*lower case*) dan menghapus tautan html. Selain itu dilakukan penghapusan tanda baca seperti tanda kutip dan *underscore*. Tahap terakhir, yaitu menghapus angka dan spasi berlebih (*whitespace*) dalam kalimat ulasan. Hasil dari pemrograman *data cleansing* kemudian masuk ke dalam tahapan *fix typo*.

```

Algoritme 1: Data Cleansing
1 df['review_lower'] = df['Ulasan'].apply(lambda x: " ".join(x.lower()
2 for x in x.split()))
3
4 df['review_nolink'] = df['review_lower'].str.replace(r"http\S+", '
5 ')
6 df['review_nopunct'] = df['review_nolink'].str.replace('[^\w\s]', '
7 ')
8 df['review_nonum'] = df['review_nopunct'].str.replace(r'\d+', ' ')
9 df['review_nounder'] = df['review_nonum'].str.replace('_', ' ')
10 df['review_nospace'] = df['review_nounder'].str.replace('[\s]+', '
11 ')
    
```

### 5.1.2 Fix Typo

*Fix typo* atau perbaikan kata merupakan tahapan yang bertujuan untuk merubah kata-kata yang memiliki kesalahan ejaan dan dilakukan penyeragaman beberapa bentuk kata. Untuk melakukan hal ini, disusun sebuah *dictionary* berisi kata yang akan diganti dan kata dengan perbaikan yang akan menggantikannya. Pada Tabel 5.2 diberikan beberapa contoh perubahan kata pada tahap ini. Sebagai contoh pada kolom “sebelum” terdapat kata-kata yang memiliki kesalahan

pengejaan seperti kata “baguuusss” yang diubah menjadi kata “bagus”. Selain itu, dilakukan penyeragaman beberapa kata dengan bentuk kata yang tidak baku menjadi bentuk baku seperti kata “kebayar” menjadi “terbayar”.

Tabel 5.2 Contoh perbaikan kata

No.	Sebelum	Sesudah
1.	baguuussss	bagus
2.	kebayar	terbayar
3.	kelokasi	ke lokasi
4.	segerrrr	segar
5.	wuow	wow

```

Algoritme 2: Fix Typo
1 #load dictionary file
dict1 = pd.read_csv('.././Documents/dictionarybaru.csv', encoding='latin1')
#set column for changing words in 'before' to 'after'
2 dict = dict1.set_index('before')['after'].to_dict()
3 d2 = {r'(\b){(\b)'.format(k):r'\1{\2}'.format(v) for k,v in dict.items()}
#replace typo words with words in d2
4 df['review_lemma'] = df['review_nospace'].replace(d2, regex=True)
    
```

### 5.1.3 Stopwords Removal

Tahapan selanjutnya, yaitu melakukan *stopwords removal*. Tahapan ini bertujuan untuk menghapus kata-kata yang tidak relevan untuk proses klasifikasi analisis sentimen. Pada penelitian ini tahapan *stopwords removal* dilakukan menggunakan bantuan, yaitu *library* pySastrawi. *Stoplist* dimodifikasi menyesuaikan dengan kepentingan penelitian dengan cara menambahkan dan mengurangi kata-kata secara manual pada berkas *StopWordRemoverFactory.py* yang terdapat di dalam *folder library* Sastrawi tersimpan. Contoh kata yang dikurangi adalah kata “tidak” dan “bukan” dikarenakan jika kata tersebut dihilangkan, maka hasil analisis sentimen tidak akan sesuai dengan data yang sudah terlabel dan mengurangi tingkat akurasi klasifikasi. Sedangkan contoh kata yang ditambahkan pada *stoplist*, yaitu nama hari seperti “Senin” dan nama bulan “februari” yang dapat dilihat pada Gambar 5.3 dikarenakan kata yang menunjukkan keterangan waktu pada penelitian ini belum dipertimbangkan.

```

Algoritme 3: Stopwords Removal
1 from Sastrawi.StopWordRemover.StopWordRemoverFactory import
StopWordRemoverFactory
2 factory = StopWordRemoverFactory()
3 stopwords = factory.create_stop_word_remover()
4 df['stopped'] = df.apply(lambda row:
5 stopwords.remove(row['review_lemma']), axis=1)
    
```

```
'dst', 'dtg', 'du', 'dua', 'dug', 'duh', 'duh', 'dulu', 'dun',
'dwi', 'dy', 'e', 'eco', 'edc', 'edi', 'eg', 'eh', 'ehehe', 'eight', 'eit',
'eits', 'elo', 'em', 'emg', 'empat', 'empat', 'emu', 'en', 'enak', 'enam',
'end', 'entah', 'entahlah', 'eo', 'era', 'es', 'esa', 'esok', 'esp', 'etc',
'eur', 'euy', 'ex', 'exy', 'eye', 'f', 'fal', 'fam', 'fan', 'far', 'fe',
'feb', 'februari', 'fed', 'fel', 'fet', 'fi', 'fit', 'five', 'fix', 'fj',
'flu', 'fly', 'fo', 'fod', 'for', 'fot', 'four', 'fox', 'ft', 'fto', 'fyi',
'gad', 'gaes', 'gag', 'gal', 'gan', 'gas', 'gb', 'gengs', 'geo', 'ger',
'get', 'gf', 'gha', 'gn', 'gng', 'go', 'god', 'gogle', 'gp', 'gt', 'gtu',
'gud', 'gue', 'gue', 'gun', 'guna', 'gunakan', 'guys', 'gw', 'gwd', 'h',
'ha', 'had', 'hadap', 'haha', 'hahaha', 'hahahaha', 'hahahahaha',
'hahahahay', 'hai', 'hai', 'hak', 'hal', 'halnya', 'halo', 'hampir', 'hanya',
'hanyalah', 'hari', 'haruslah', 'harusnya', 'has', 'hat', 'hau', 'hav',
'hdg', 'he', 'hehe', 'hehehe', 'hei', 'hel', 'helo', 'hem', 'hendak',
'hendaklah', 'hendaknya', 'hey', 'hft', 'hi', 'hih', 'hiks', 'hinga',
```

Gambar 5.3 Contoh daftar stopwords yang ditambahkan

### 5.1.4 Stemming

Kalimat-kalimat yang sudah diproses melalui tahapan-tahapan sebelumnya kemudian memasuki tahapan *stemming*. *Stemming* dilakukan untuk mengubah kata-kata berimbuhan menjadi kata dasar. Pada penelitian ini digunakan stemmer dari *library* dari *pySastrawi* yang menyediakan *stemmer* untuk bahasa Indonesia.

```

Algoritme 4: Stemming
1 from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
2 factory = StemmerFactory()
3 stemmer = factory.create_stemmer()
4 df['stemmed'] = df.apply(lambda row:stemmer.stem (row['stopped']),
axis=1)

```

### 5.2 TF-IDF dan Klasifikasi

Sebelum memasuki tahap klasifikasi dikarenakan mesin tidak bisa langsung mengolah teks untuk proses klasifikasi, maka dilakukan *feature extraction*. Pada penelitian ini digunakan TF-IDF dan metode Multinomial *Naïve Bayes* untuk klasifikasi. Selain itu juga digunakan bantuan dari *library* *ScikitLearn* untuk mempersingkat algoritme.

```

Algoritme 5: TF-IDF & Klasifikasi
1 from sklearn.pipeline import Pipeline
2 from sklearn.feature_extraction.text import TfidfVectorizer
3 from sklearn.naive_bayes import MultinomialNB
4 from sklearn import metrics
5 from sklearn.model_selection import cross_val_score
6 text_clf_nb = Pipeline([('tfidf', TfidfVectorizer(min_df=5)),
7 ('clf', MultinomialNB()),
8 ])
9 text_clf_nb.fit(X_train, y_train)

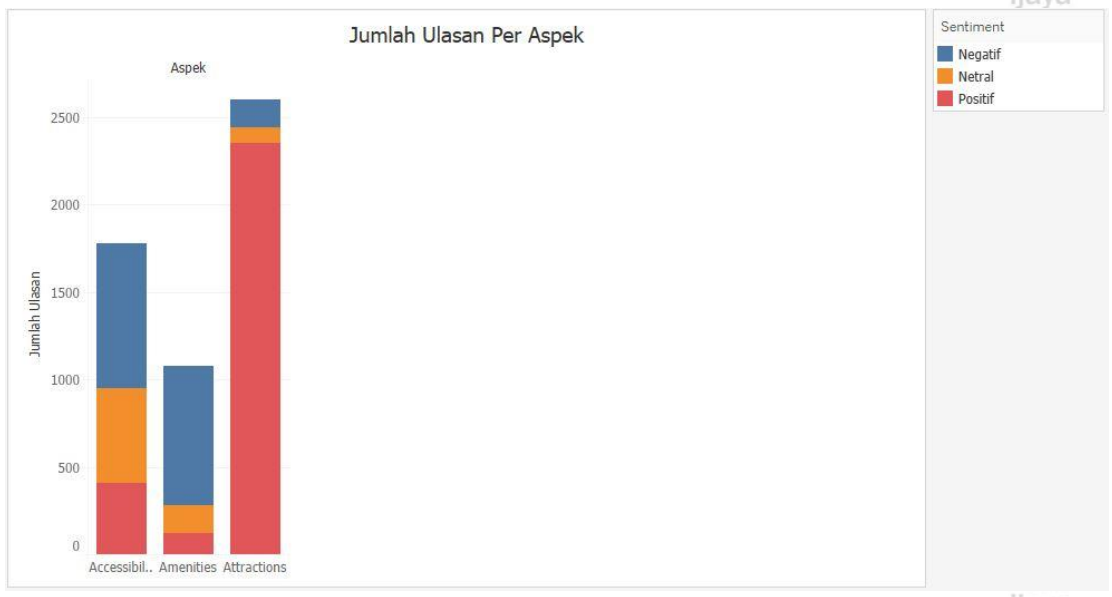
```

### 5.3 Implementasi Dashboard

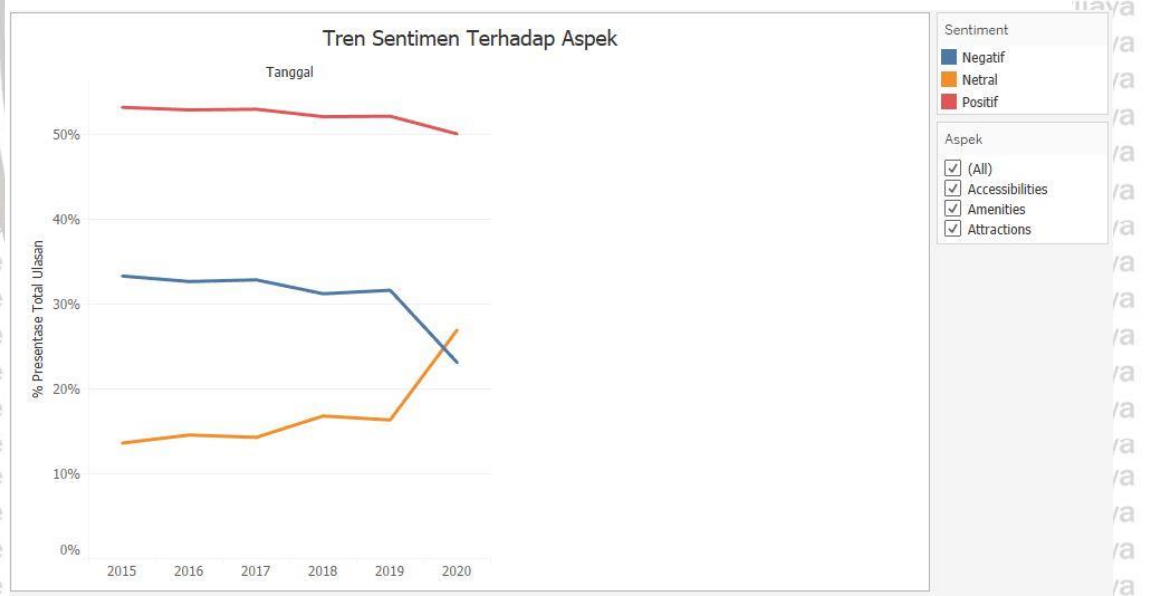
Data hasil analisis dipresentasikan ke dalam dashboard yang terdiri dari tiga grafik yang dapat dilihat pada Gambar 5.5. Grafik pertama berupa *bar chart* yang menampilkan jumlah sentimen pada ketiga aspek seperti pada Gambar 5.2. Grafik kedua berupa *line chart* yang menampilkan tren aspek dari tahun ke tahun seperti



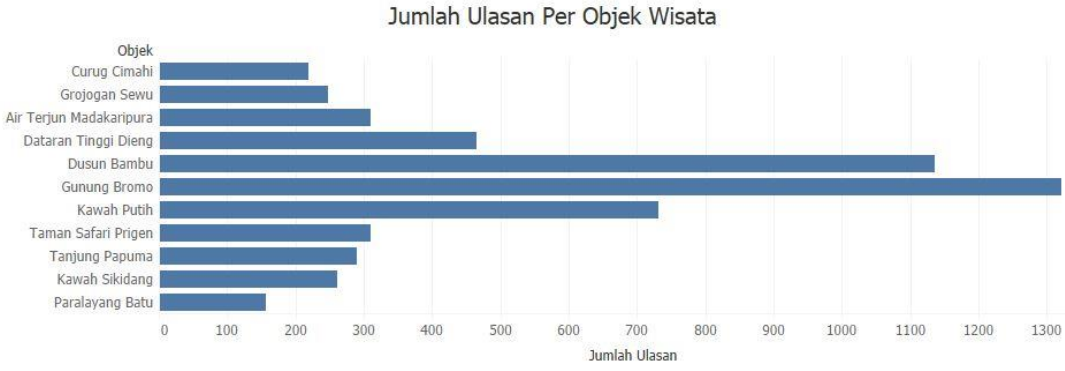
pada Gambar 5.3. Grafik terakhir pada Gambar 5.4 berupa histogram menampilkan objek-objek wisata.



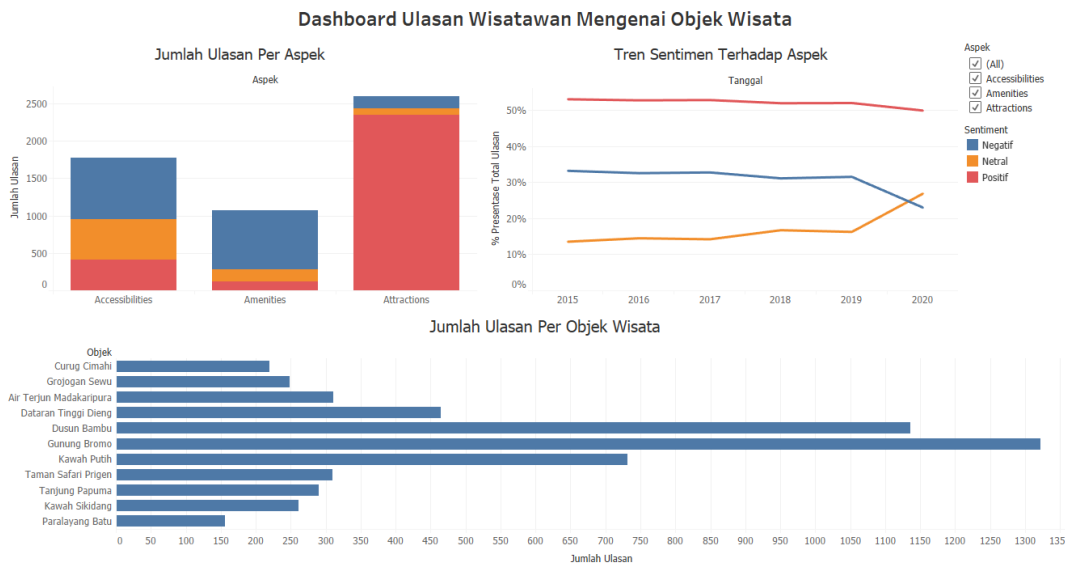
**Gambar 5.4 Grafik jumlah ulasan per aspek**



**Gambar 5.5 Grafik tren sentimen terhadap aspek**



Gambar 5.6 Jumlah ulasan tiap objek wisata



Gambar 5.7 Dashboard ulasan wisatawan terhadap objek wisata



## BAB 6 HASIL DAN PEMBAHASAN

### 6.1 Hasil Klasifikasi

#### 6.1.1 Aspek *Attractions*

Tabel 6.1 Hasil *accuracy*, *precision*, *recall*, dan *f1* aspek *attractions*

Attractions			
<i>Accuracy</i>	0,69898		
	<i>Precision</i>	<i>Recall</i>	<i>F1</i>
Negatif	0,84	0,44	0,58
Netral	0,57	0,28	0,38
Positif	0,7	0,97	0,81

Hasil evaluasi dari model klasifikasi untuk aspek *attractions* adalah nilai *accuracy* sebesar 0,69898. Selain nilai *accuracy*, didapatkan nilai *precision*, *recall* dan *f1* untuk kelas positif, negatif, dan netral. Hasil *precision*, *recall*, dan *f1* kelas negatif pada aspek *attractions* adalah sebesar 84%, 44%, dan 58% seperti yang tertera pada Tabel 6.1. Nilai *precision* untuk sentimen negatif pada aspek *attractions* adalah 0,84 yang menandakan ada 16% dari kalimat yang tidak memuat sentimen negatif diklasifikasikan menjadi kalimat dengan sentimen negatif. Nilai *recall* untuk sentimen negatif pada aspek *attractions* adalah 0,44 yang menandakan sebanyak 56% kalimat yang seharusnya diklasifikasikan menjadi negatif terklasifikasikan ke dalam kelas selain negatif. Nilai *f1* sebesar 0,58 untuk sentimen negatif didapatkan dari nilai *precision* lebih tinggi dari nilai *recall*.

Pada kelas netral didapatkan nilai sebesar 0,57 yang menandakan terdapat 43% kalimat dari kelas lain yang salah diklasifikasikan menjadi netral. Nilai *recall* untuk sentimen netral pada aspek *attractions* adalah 0,28. Ini berarti model belum mampu mengklasifikasikan 72% kalimat yang seharusnya masuk ke dalam kelas netral terklasifikasikan ke dalam kelas selain netral. Nilai 0,38 untuk sentimen netral menggambarkan nilai *precision* dua kali lebih besar dari *recall* sehingga harmoni antara keduanya berada di bawah 0,5. Hasil tersebut diduga dikarenakan persebaran kelas pada *dataset* yang tidak merata.

Nilai *precision* untuk sentimen positif pada aspek *attractions* adalah 0,7. Ini berarti ada 30% dari kalimat yang tidak memuat sentimen positif terklasifikasikan menjadi kalimat dengan sentimen positif. Nilai *recall* sebesar 0,97 mendandakan model yang dibuat mampu memutuskan suatu kalimat yang memuat sentimen positif diklasifikasikan dengan tepat menjadi kelas positif. Nilai *f1* yang menggambarkan harmoni antara *precision* dan *recall* dari sentimen positif adalah 0,81. Hasil *precision*, *recall*, dan *f1* pada kelas positif cenderung lebih tinggi dari dua kelas lainnya diduga dikarenakan jumlah kelas positif lebih banyak dibandingkan dengan kelas netral dan negatif sehingga model dapat mengklasifikasikan kelas positif lebih baik.



### 6.1.2 Aspek *Accessibilities*

Tabel 6.2 Hasil *accuracy, precision, recall, dan f1 aspek accessibilities*

Accessibilities			
Accuracy	0,61667		
	Precision	Recall	F1
Negatif	0,5	0,14	0,22
Netral	0,62	0,94	0,75
Positif	0,71	0,16	0,26

Hasil evaluasi dari model klasifikasi untuk aspek *accessibilities* adalah nilai *accuracy* sebesar 0,61667. Selain nilai *accuracy*, didapatkan nilai *precision, recall* dan *f1* untuk kelas positif, negatif, dan netral. Hasil *precision, recall*, dan *f1* kelas negatif pada aspek *attractions* adalah sebesar 50%, 14%, dan 0.22% seperti yang tertera pada Tabel 6.2. Nilai *precision* untuk sentimen negatif pada aspek *accessibilities* adalah 0,5 yang berarti terdapat setengah dari kalimat yang tidak memuat sentimen negatif diklasifikasikan menjadi kalimat dengan sentimen negatif. Nilai *recall* untuk kelas negatif adalah 0,14 yang menggambarkan bahwa model belum mampu mengklasifikasikan sebanyak 86% kalimat yang seharusnya diklasifikasikan menjadi negatif terklasifikasikan ke dalam kelas selain negatif. Nilai *f1* sebesar 0,22 untuk kelas negatif menggambarkan harmoni antara *precision* dan *recall* berada di bawah 0,5.

*Precision* kelas netral pada aspek *accessibilities* menandakan terdapat 38% kalimat yang salah diklasifikasikan menjadi netral. Nilai *recall* untuk sentimen netral pada aspek *accessibilities* adalah 0,94. Ini berarti terdapat 6% kalimat yang seharusnya masuk ke dalam kelas netral terklasifikasikan ke dalam kelas selain netral. Selain nilai *precision* dan *recall*, didapatkan *f1* pada kelas netral, yaitu sebesar 75%

Nilai *precision* untuk kelas positif pada aspek *accessibilities* adalah 0,71. Ini berarti ada 29% dari kalimat yang tidak memuat sentimen positif diklasifikasikan menjadi kalimat dengan sentimen positif. Nilai *recall* untuk sentimen positif pada aspek *accessibilities* adalah 0,16 yang berarti model yang dibuat belum mampu memutuskan, dengan peluang 84%, suatu kalimat yang memuat sentimen positif diklasifikasikan menjadi kelas lainnya. Nilai *f1* dari kelas positif adalah 26% yang disebabkan oleh nilai *recall* yang lebih rendah dari *precision*.

### 6.1.3 Aspek Amenities

Tabel 6.3 Hasil *accuracy*, *precision*, *recall*, dan *f1* aspek amenities

Amenities			
Accuracy	0,63106		
	Precision	Recall	F1
Negatif	1	0,21	0,34
Netral	0,68	0,65	0,67
Positif	0,57	0,87	0,69

Hasil evaluasi dari model klasifikasi untuk aspek amenities adalah nilai *accuracy* sebesar 0,63106. Selain nilai *accuracy*, didapatkan nilai *precision*, *recall* dan *f1* untuk kelas positif, negatif, dan netral. Hasil *precision*, *recall*, dan *f1* kelas negatif pada aspek attractions adalah sebesar 100%, 21%, dan 34% seperti yang tertera pada Tabel 6.3. Nilai *precision* untuk sentimen negatif pada aspek amenities adalah 100% yang berarti seluruh kalimat bersentimen negatif diklasifikasikan secara tepat. Nilai *recall* untuk sentimen negatif pada aspek amenities adalah 0,21. Ini berarti model belum mampu mengklasifikasikan sebanyak 79% kalimat yang seharusnya diklasifikasikan menjadi negatif terklasifikasikan ke dalam kelas selain negatif. Nilai *f1* untuk kelas negatif pada aspek amenities adalah 0,34. Hal tersebut dikarenakan nilai *recall* kelas negatif lebih rendah dari *precision*.

Nilai *precision* untuk kelas netral pada aspek amenities mengindikasikan sebesar 32% kalimat yang seharusnya terklasifikasikan menjadi kelas selain netral salah diklasifikasikan menjadi netral. Nilai *recall* untuk sentimen netral pada aspek amenities adalah 0,65. Ini berarti model belum mampu mengklasifikasikan 35% kalimat yang seharusnya masuk ke dalam kelas netral terklasifikasikan ke dalam kelas selain netral. Nilai *f1* untuk kelas netral adalah 0,67 yang menandakan harmoni antara *precision* dan *recall* kelas ini.

Nilai *precision* untuk sentimen positif pada aspek amenities menggambarkan sebanyak 43% dari kalimat yang tidak memuat sentimen positif diklasifikasikan menjadi kalimat dengan sentimen positif. Nilai *recall* kelas positif mengindikasikan model yang dibuat mampu memutuskan, dengan peluang 87%, kalimat yang memuat sentimen positif diklasifikasikan dengan tepat menjadi kelas positif. Hasil dari *f1* menggambarkan *precision* memiliki nilai *precision* lebih kecil dibandingkan nilai *recall* sehingga menghasilkan nilai sebesar 69%.

## 6.2 Hasil System Usability Score (SUS) Dashboard

Untuk menghitung nilai dari kemudahan penggunaan sistem atau *dashboard* yang telah dibuat penyusun membuat kuesioner untuk diisi oleh *end user* (Perum Perhutani). Kuesioner disusun dengan 10 pernyataan di mana 5 di antaranya merupakan pernyataan positif dan sisanya berupa pernyataan negatif. Tabel 6.4 menampilkan hasil *Usability Testing* dengan menggunakan SUS yang telah dilakukan oleh perwakilan pihak Perum Perhutani.

Pernyataan bernomor ganjil merupakan pernyataan positif sedangkan pernyataan bernomor genap berupa pernyataan negatif. Pihak Perhutani diminta untuk mengisi kuesioner dengan angka dengan interval 0 (sangat tidak setuju) hingga 4 (sangat setuju). Untuk mendapatkan skor SUS perlu dilakukan perhitungan dengan cara:

1. Untuk pernyataan bernomor ganjil, maka nilai - 1.
2. Untuk pernyataan bernomor genap, maka 5 - Nilai.
3. Hasil perhitungan dari setiap pertanyaan kemudian dijumlahkan lalu dikalikan 2,5 untuk mendapat hasil akhir.

**Tabel 6.4 Hasil Usability Testing**

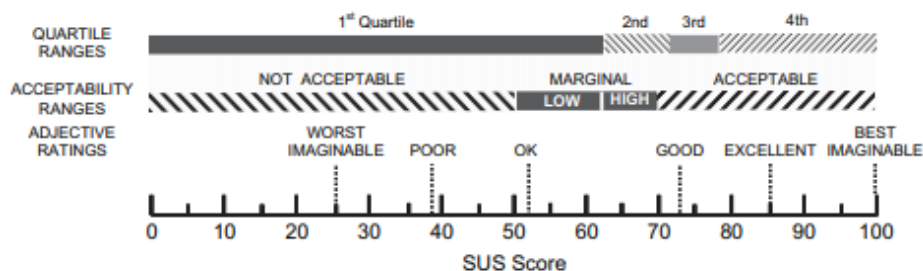
No.	Pernyataan	Nilai
1.	Saya pikir saya akan sering menggunakan sistem ini	3
2.	Saya rasa sistem ini terlalu kompleks	3
3.	Saya kira sistem ini mudah untuk digunakan	3
4.	Saya pikir saya akan membutuhkan bantuan teknis untuk dapat menggunakan sistem ini	1
5.	Saya kira ada berbagai fungsi dalam sistem ini terintegrasi dengan baik	3
6.	Saya pikir ada terlalu banyak ketidaksesuaian dalam sistem ini	1
7.	Saya bisa bayangkan bahwa kebanyakan orang akan mudah mempelajari dan menggunakan sistem ini	3
8.	Saya kira sistem ini sangat rumit untuk digunakan	1
9.	Saya merasa percaya diri dalam menggunakan sistem ini	3
10.	Saya perlu belajar banyak hal sebelum saya bisa memulai menggunakan sistem ini	3



Berikut merupakan hasil perhitungan SUS.

$$\begin{aligned} \text{SUS} &= ((3-1)+(5-3)+(3-1)+(5-1)+(3-1)+(5-1)+(3-1)+(5-1)+(3-1)+(5-3)) * 2,5 \\ &= (2+2+2+4+2+4+2+4+2+2) * 2,5 \\ &= 65 \end{aligned}$$

Bangor et al. (2008) memberikan cara untuk dapat mengategorikan hasil dari Usability Testing dengan menggunakan SUS. Gambar 6.1 menunjukkan bagaimana mengelompokkan hasil dari SUS Score. SUS Score bisa dikelompokkan dalam 3 parameter yakni *Acceptability Ranges*, *Grade Scale*, dan *Adjective Ratings*.



**Gambar 6.1 System Usability Score (SUS)**

Sumber: Bangor, et al. (2008)

Berdasarkan perhitungan yang telah dilakukan maka SUS Score untuk *Dashboard* sebesar 65 sehingga dapat disimpulkan bahwa *dashboard* tergolong ke dalam kategori *Marginal (high)* dan dalam rating “*Good*” yang artinya *Dashboard* memiliki tingkat penerimaan dalam kategori marjinal atau belum memenuhi kategori *acceptable* oleh pihak Perum Perhutani.

Setelah penyusun menganalisis mengapa nilai dari rancangan *dashboard* ini dinilai belum *acceptable* oleh Pelum Perhutani, terlihat tampaknya yang mengisi kuesioner tidak terlalu memahami seluruh pertanyaan-pertanyaan yang ada. Sebagai contoh, dari pertanyaan nomor dua dapat disimpulkan pengisi kuesioner menganggap bahwa sistem terlalu kompleks untuk dapat digunakan. Sementara itu pada pertanyaan nomor tiga pengisi kuesioner menganggap sistem mudah digunakan. Hal ini mengindikasikan pengisi kuesioner tidak terlalu memahami pengertian kompleks untuk suatu sistem.

Perlu digarisbawahi bahwa nilai 65 untuk SUS diperoleh ketika penyusun memberikan intervensi apapun kepada pengisi kuesioner dalam mengisi kuesioner. Berdasarkan pengamatan bahwa pengisi kuesioner tidak memahami makna kompleks dalam suatu sistem mungkin di masa yang akan datang penyusun akan melakukan tes dengan lebih banyak petunjuk.

## BAB 7 PENUTUP

### 7.1 Kesimpulan

Di skripsi ini telah diteliti analisis sentimen aspek terhadap tiga aspek pariwisata, yaitu *attractions*, *accessibilities*, dan *amenities* dari objek-objek wisata yang dikelola oleh Perum Perhutani. Berikut merupakan kesimpulan dari penelitian ini.

1. Analisis sentimen menggunakan metode *Multinomial Naïve Bayes* pada data hasil *web scraping* dari TripAdvisor. Didapatkan nilai *accuracy* sebesar 69,898% pada ulasan aspek *attraction*. Sedangkan untuk aspek *accessibilities* didapatkan nilai *accuracy* 61,667%. Terakhir, nilai *accuracy* untuk aspek *amenities* adalah 63,106%.
2. Pada aspek *attractions* untuk sentimen negatif didapatkan nilai *precision* sebesar 84%, *recall* 44%, dan *f1* sebesar 58%. Untuk sentimen netral pada aspek *attractions* didapatkan *precision* sebesar 57%, *recall* 28% dan *f1* 38%. Sedangkan untuk sentimen positif pada aspek *attractions* didapatkan nilai *precision* 70%, *recall* 97%, dan *f1* 81%. Pada aspek *accessibilities* untuk sentimen positif didapatkan nilai *precision* sebesar 71%, *recall* 16%, dan *f1* sebesar 26%. Untuk sentimen netral pada aspek *accessibilities* didapatkan *precision* sebesar 62%, *recall* 94% dan *f1* 75%. Sedangkan untuk sentimen negatif pada aspek *accessibilities* didapatkan nilai *precision* 50%, *recall* 14%, dan *f1* 22%. Untuk aspek *amenities* untuk sentimen negatif didapatkan nilai *precision* sebesar 100%, *recall* 21%, dan *f1* sebesar 34%. Untuk sentimen netral pada aspek *amenities* didapatkan *precision* sebesar 68%, *recall* 65% dan *f1* 67%. Sedangkan untuk sentimen positif pada aspek *amenities* didapatkan nilai *precision* 57%, *recall* 87%, dan *f1* 69%.
3. *Dashboard* diuji menggunakan alat uji berupa *System Usability Scale (SUS)* dilakukan dengan pihak manajemen Divisi SDM, Keuangan, dan Umum Perum Perhutani. Skor *SUS* yang didapatkan sebesar 65 sehingga dapat disimpulkan bahwa *dashboard* tergolong ke dalam kategori *Marginal (high)* dan dalam rating "Good" yang artinya *dashboard* memiliki tingkat penerimaan dalam kategori marjinal atau belum memenuhi kategori *acceptable* oleh pihak Perum Perhutani.

### 7.2 Saran

Nilai-nilai *accuracy* di skripsi ini masih pada kisaran 60%. Penyusun menyadari nilai ini masih dapat ditingkatkan dengan beberapa cara berikut ini.

1. Data yang diperoleh dari TripAdvisor adalah paragraf atau kalimat yang tertulis jauh dari tata bahasa baku. Proses melabeli kalimat-kalimat tersebut secara otomatis bukanlah perkara mudah, dengan kata lain memerlukan metode-metode tambahan di luar penggunaan *dictionary* seperti yang dilakukan di penelitian skripsi ini. Metode yang dimaksud adalah metode dengan banyak

tahapan, yang memuat tes keacakan seperti yang dilakukan oleh Vallantin (2019).

2. Dari data mentah yang diperoleh didapatkan kalimat-kalimat yang memuat aspek *attractions* jauh lebih banyak dibandingkan kalimat yang memuat kedua aspek lainnya. Kurang banyaknya data untuk kedua aspek ini penyusun duga menjadi penyebab mengapa nilai *accuracy* untuk aspek *accessibilities* dan *amenities* lebih kecil dibandingkan aspek *attractions*. Untuk penelitian berikutnya perlu dipertimbangkan keseimbangan banyaknya data untuk setiap aspek dengan menambah *dataset* dari sumber lain.
3. Mengingat terdapat banyak metode lain untuk analisis sentimen, penyusun menyarankan penerapan metode-metode lainnya selain metode yang digunakan di skripsi ini agar mendapatkan perbandingan nilai *accuracy* untuk ketiga aspek tersebut.
4. Sebenarnya terdapat aspek lain yang sering dipertimbangkan di bidang pariwisata, yaitu aspek *activities* dan *accommodation*. Untuk mendapatkan informasi yang lengkap Perum Perhutani dapat menerapkan analisis sentimen untuk menilai kedua aspek tersebut.
5. Mengingat ada indikasi pengisi kuesioner belum tentu memahami semua terminologi untuk suatu *dashboard* sistem informasi, maka di penelitian mendatang proses pelaksanaan SUS harus disertai dengan petunjuk-petunjuk yang lebih mendetail dan apabila memungkinkan metode pengambilan data dengan cara pengisian kuesioner bisa digantikan atau disertai dengan wawancara.





## DAFTAR REFERENSI

- Alaeri, A., Becken, S., & Stantic, B., 2017. Sentiment Analysis in Tourism: Capitalizing Big Data. *Journal of Travel Research*, [e-journal] 58(2), p.175-191. Tersedia melalui: Perpustakaan Universitas BX <<http://perpustakaan.ubx.ac.id>> [Diakses 14 Juni 2021]
- Alpaydin, E. (2010). *Introduction to Machine Learning*. 2nd ed. Cambridge: MIT Press.
- Arulselvi, A. C., Sendhilkumar, S., & Mahalakshmi, S., 2018. Classification of tweets for sentiment and trend analysis. *Proceedings of the 2017 International Conference on Intelligent Computing and Control Systems*, pp.566–573.
- Aydin, O., 2018. *R Web Scraping Quick Start Guide: Techniques and tools to crawl and scrape data from websites*. Brimingham: Packt Publishing.
- Bachtiar, F. A., Paulina, W., & Rusydi, A. N., 2020. Text Mining for Aspect Based Sentiment Analysis on Customer Review : a Case Study in the Hotel Industry. *5th International Workshop on Innovations in Information and Communication Science and Technology, March*.
- Bangor, A., Staff, T., Kortum, P., Miller, J., & Staff, T. (2008). Determining what individual SUS scores mean: adding an adjective rating scale. *Journal of Usability Studies*, 4(3), p.114–123.
- Berry, M. W., & Kogan, J., 2010. *Text Mining: Applications and Theory*. John Wiley & Sons, Inc.
- Chang, Y. C., Ku, C. H., & Chen, C. H., 2017. Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. *International Journal of Information Management*.
- Cooper, C., Fletcher, J., Gilbert D. & Wanhill, S., 1993. *Tourism: Principles and Practice*. Harlow: Longman Group Limited.
- Feldman, R., & Sanger, J., 2007. *The Text Mining Handbook*. Cambridge: Cambridge University Press.
- Ferdiana, R., 2019. Dataset Indonesia untuk Analisis Sentimen. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 8(4).
- Few, S. 2006. *Information Dashboard Design (The Effect)*. Italy: O'Reilly Media, Inc.
- Geetha, M., Singha, P., & Sinha, S., 2017. Relationships between Customer Sentiment and Online Customer Ratings for Hotels - An Empirical Analysis. *Tourism Management*, 61, p.43-54.

- Han, J., Kamber, M. & Pei, J., 2012. *Data Mining Concepts and Techniques*. 3rd ed. Massachusetts: Morgan Kaufmann.
- Liu, B., 2012. *Sentiment Analysis and Opinion Mining*. Massachusetts: Morgan & Claypool Publishers.
- Manning, C.D., Raghavan. P., & Schütze, H., 2009. *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.
- Mirjalili, V., & Raschka, S., 2018. *Python Machine Learning*. 3rd ed. Birmingham: Packt Publishing.
- Müller, A. C., & Guido, S., 2015. *Introduction to Machine Learning with Python and Scikit-Learn*. California: O'Reilly Media.
- Navlani, A., 2018. *Naive Bayes Classification Using Scikit-learn*. Tersedia di: <<https://www.datacamp.com/community/tutorials/naive-bayes-scikit-learn>> [Diakses tanggal 23 Juni 2021]
- NLP Stanford, 2009. *Extracting Common Sentiments from Review*. [pdf] NLP Stanford. Tersedia di: <<https://nlp.stanford.edu/courses/cs224n/2009/fp/14.pdf>> [Diakses 13 Juni 2021].
- Paolanti, M., et al., 2021. Tourism destination management using sentiment analysis and geo-location information: a deep learning approach. *Information Technology & Tourism*, 23, p.241–264.
- Perhutani, 2021. Visi, Misi, dan Tata Nilai. [online] Tersedia di: <https://perhutani.co.id/tentang-kami/visi-misi-dan-tata-nilai/> [Diakses 14 Juni 2021]
- Prameswari, P., Surjandari, I., & Laoh, E., 2017. Opinion mining from online reviews in Bali tourist area. *International Conference on Science in Information Technology: Theory and Application of IT for Education, Industry and Society in Big Data Era*, pp.226–230.
- TripAdvisor, 2021. *About TripAdvisor*. [online] Tersedia di: <<https://tripadvisor.mediaroom.com/us-about-us>> [Diakses 14 Juni 2021]
- Turban, E., Sharda, S., Delen, D. & King, D., 2010. *Business Intelligence A Managerial Approach*. 2nd ed. New Jersey: Prentice Hall.
- Vallantin, L., 2019. *How to label text for sentiment analysis-good practices*. [online] Tersedia di: <<https://towardsdatascience.com/how-to-label-text-for-sentiment-analysis-good-practises-2dce9e470708>> [Diakses 14 Juni 2021]
- WEF, 2019. *The Travel & Tourism Competitiveness Report 2019*. [pdf] World Economic Forum. Tersedia di: <[http://www3.weforum.org/docs/WEF\\_TTCR\\_2019.pdf](http://www3.weforum.org/docs/WEF_TTCR_2019.pdf)> [Diakses 14 April 2021]

Xu, F., et al., 2020. E-Commerce Product Review Sentiment Classification Based On Naïve Bayes Continuous Learning Framework. *Information Processing and Management*, 57.





## LAMPIRAN A HASIL WAWANCARA

Judul Penelitian : Analisis Sentimen Aspek pada Opini Wisatawan Menggunakan Metode Naïve Bayes (Studi Kasus: Perum Perhutani)

Peneliti : Fasya Ghassani Hadiyan

Dosen Pembimbing : Dr.Eng. Fitra Abdurrachman Bachtiar, S.T., M.Eng.

Narasumber : Afifah Yuseka

Jabatan : Junior Manajer Bisnis Kesatuan Pemangkuan Hutan (KPH) selaku Humas

Instansi : Perum Perhutani Indonesia

Tanggal Wawancara : 19 Februari 2019

Perihal : Wawancara terkait permasalahan penelitian

<p>Objek-objek wisata apa saja yang dikelola Perhutani?</p>	<p>Perhutani mengelola kurang lebih 768 objek wisata di Jawa dan Madura yang tergolong ke dalam beberapa standar usaha sesuai Peraturan Kementerian Pariwisata dan Ekonomi Kreatif. Standar usaha tersebut antara lain, taman rekreasi, bumi perkemahan, air panas, dan vila non bintang. Sebanyak 70% dari objek wisata tersebut tergolong ke dalam standar usaha taman rekreasi. Contoh dari objek wisata dari standar usaha taman rekreasi adalah Wana Wisata Kawah Putih, Curug Cilember, dan Tanjung Papuma.</p>
<p>Bagaimana mekanisme pengelolaan dan pengembangan objek-objek wisata yang tersebar di wilayah Jawa dan Madura?</p>	<p>Perum Perhutani bekerja di bawah Kementerian Pariwisata dalam mengelola dan mengembangkan 768 objek wisata di pulau Jawa dan Madura. Pada tiap-tiap Divisi Regional, di bawah seorang <i>General Manager Ecotourism</i>, Divisi SDM, Keuangan dan Umum bertugas melakukan perawatan rutin juga bertugas melaporkan beberapa hal seperti keluhan perbaikan yang dilakukan melalui telpon, tren wisata, dan jumlah pengunjung pada Perhutani Pusat (Jakarta) setiap bulannya. Untuk menilai kinerja pengelolaan dilakukan <i>assessment</i> terhadap objek-objek wisata dengan melakukan survei lapangan setahun sekali sehingga didapatkan</p>

	<p>rekomendasi investasi fasilitas-fasilitas terhadap objek-objek wisata. Hasil pelaporan dan <i>assessment</i> tersebut akan dijadikan analisis pasar, pertimbangan kelayakan bisnis, dan mitigasi risiko terkait objek wisata oleh Perhutani Pusat selaku pertimbangan untuk pengembangan objek wisata.</p>
<p>Apa saja upaya Perhutani dalam mengembangkan objek-objek wisata yang dikelolanya?</p>	<p>Pada tahun 2015 terjadi investasi dalam skala yang cukup besar antara lain menambahkan sistem <i>e-ticketing</i>, EDC (pembayaran <i>cashless</i>), <i>booking online</i>, dan membuat media sosial sebagai sarana promosi di beberapa objek wisata. Baru-baru ini, pada tahun 2017, dilakukan beberapa pengembangan fasilitas objek wisata salah satunya menambah <i>spot-spot selfie</i> pada Wana Wisata Kawah Putih. Pengembangan objek-objek wisata belum bisa dilakukan secara rutin dan menyeluruh akibat kurangnya SDM dengan <i>experience</i> yang terstandarisasi.</p>
<p>Apakah Perhutani sudah memanfaatkan media sosial sebagai sarana informasi dalam pengelolaan dan pengembangan objek-objek wisata?</p>	<p>Sejauh ini Perhutani baru menggunakan media sosial sebagai sarana untuk melakukan promosi. Belum terlalu banyak informasi yang bisa didapatkan Perhutani dari media sosial yang dimiliki, seperti Instagram dan Facebook, dikarenakan jumlah komentar yang dianggap belum memadai. Selama ini untuk mengetahui preferensi wisatawan sebagai acuan untuk mengembangkan objek wisata ke depannya dilakukan melalui <i>assessment</i> objek wisata.</p>

Jakarta, 19 Februari 2019



Afifah Yuseka



## LAMPIRAN B PERSETUJUAN CONTOH LABEL DATA

Judul Penelitian : Analisis Sentimen Aspek pada Opini Wisatawan Menggunakan Metode Naïve Bayes (Studi Kasus: Perum Perhutani)

Peneliti : Fasya Ghassani Hadiyan

Dosen Pembimbing : Dr.Eng. Fitra Abdurrachman Bachtiar, S.T., M.Eng.

Validator : Afifah Yuseka

Jabatan : Junior Manajer Bisnis Kesatuan Pemangkuan Hutan (KPH) selaku Humas

No.	Ulasan	Aspek	Sentimen
1	perjalanan mobil cukup nyaman	Accessibilities	Positif
2	akses jalan baik	Accessibilities	Positif
3	menuju lokasi diaspal aman	Accessibilities	Positif
4	indah banget..meskipun tetap ada perbaikan jalan masuk menuju obyek wisata..peningkatan fasilitas umum keamanan..	Accessibilities	Positif
5	tempat lokasi sangat bersih sejuk	Accessibilities	Positif
6	jalan mendapatkan sana mudah	Accessibilities	Positif
7	panduan mangannya hadir membantu di beberapa karang-karang indah	Accessibilities	Positif
8	benar-benar membantu memiliki seseorang membantu sepanjang jalan	Accessibilities	Positif
9	waktu terbaik berkunjung saat musim liburan puasa	Accessibilities	Positif
10	panduan (wajib) sangat ramah dengan senang hati menampung barang-barang sementara menjelajahi kelimun foto-foto	Accessibilities	Positif
11	bila waktu singkat mudah	Accessibilities	Positif
12	jalan cukup baik	Accessibilities	Positif
13	memiliki pemandu lokal (shalim? demikian) sangat tekun mencari kami	Accessibilities	Positif
14	adalah sedikit perjalanan indah	Accessibilities	Positif
15	jalan setapak dilewati bagus teertimpa longsor	Accessibilities	Positif
16	sy baca beberapa blog lebih baik bilang awal hanya mau 1 guide	Accessibilities	Positif
17	perjalanan terbaik dilakukan mobil transportasi sendiri (seperti sepeda motor)	Accessibilities	Positif
18	tukang ojek berseragam biar tamu lebih enak	Accessibilities	Positif
19	wajib bahasa inggris melayani tamu asing lebih baik	Accessibilities	Positif
20	supir taksi tahu terbaik bermain musik bagus sepanjang jalan	Accessibilities	Positif
21	mending pulang aja..gak perlu takut kotor..cukup menantanglah..hanya akses menuju lokasi cukup sunyi..tapi banyak petunjuk jalan sangat membantu..	Accessibilities	Netral
22	better pake jasa guide kalo ga mau kesasar	Accessibilities	Netral
23	sebenarnya tidak perlu panduan	Accessibilities	Netral



24	pemandu lokal sangat dianjurkan	Accessibilities	Netral
25	tidak lagi jalan setapak	Accessibilities	Netral
26	akan membutuhkan sebuah mobil menyimpan kamera waterpruf handphonean	Accessibilities	Netral
27	jalannya udah lumayan bagus.	Accessibilities	Netral
28	sini - perhentian tidak direncanakan saran panduan	Accessibilities	Netral
29	tidak perlu pemandu lokal jalur sepenuhnya dikelola hanya membawa ke air terjun.	Accessibilities	Netral
30	tidak perlu panduan	Accessibilities	Netral
31	sebenarnya tidak perlu panduan sebenarnya.	Accessibilities	Netral
32	biarkan tahu anda dipadam perlu panduan tidak membayar layanan nya	Accessibilities	Netral
33	disarankan mengajak pemandu	Accessibilities	Netral
34	bukan jalur utama jadi transportasi umum disana ada	Accessibilities	Netral
35	tidak mendapatkan semua jalan bawah	Accessibilities	Netral
36	terlalu banyak lalu lintas	Accessibilities	Netral
37	fasilitas ada shelter terdapat pinggir jalan	Accessibilities	Netral
38	jalan tangga kebawahnya setapak jadi pelan-pelan hati-hati menuruni tangganya	Accessibilities	Netral
39	layak perjalanan hanya menatap air jatuh membuat di kesurupan.	Accessibilities	Netral
40	perjalanan mulai terminal bis wonosobo.ada banyak opsi bisa di terminal bis	Accessibilities	Netral
41	tempat terpencil & jauh jalan utama	Accessibilities	Negatif
42	tempat cukup jauh & jalan kecil	Accessibilities	Negatif
43	perjln macet uk smp sana tp smp sna msh hrs berjuang uk dpt tiket msk	Accessibilities	Negatif
44	di lokasi jam 10 keluar lokasi sktr jam 16.30 rasanya kurang aja.	Accessibilities	Negatif
45	perjalanan ditempuh kurang lebih 25 menit pusat kota lembang	Accessibilities	Negatif
46	sayangnya waktu kesana	Accessibilities	Negatif
47	capek karena 1.000-an anak tangga haha!	Accessibilities	Negatif
48	memerlukan upaya ekstra menjangkaunya harus menuruni ratusan anak tangga	Accessibilities	Negatif
49	perjalanan cukup melelahkan	Accessibilities	Negatif
50	perjalanan cukup jauh memakan waktu 5 jam	Accessibilities	Negatif
51	kalo lewat loket 1 jalannya sebaliknya turun bawah dulu turunan curam	Accessibilities	Negatif
52	harus menuruni 1000an anak tangga terlebih	Accessibilities	Negatif
53	kaki dibuat lelah dulu harus turun melalui lebih lima ratus anak tangga	Accessibilities	Negatif
54	kaki dibuat lebih lelah karena menaiki lebih tuju ratus anak tangga	Accessibilities	Negatif
55	cuma jalannya cukup jauh lokasi selalu ramai dihari hari libur.	Accessibilities	Negatif
56	jalan sudah menurun (mungkin sedikit licin hujan)	Accessibilities	Negatif

57	pas berangkat sih gak capek turun tangga	Accessibilities	Negatif
58	siapkan tenaga kalau kesini untuk menuju grojogan sewu harus naik turun tangga sejumlah 1250 anak tangga	Accessibilities	Negatif
59	ternyata harus mengganti kendaraan dengan menggunakan mobil jeep yg khusus kan perjalanan	Accessibilities	Negatif
60	ditumpang menuju penanjakan harus melewati trek pasir khas bromo amat menyulitkan	Accessibilities	Negatif
61	ada restoran arimbi pemandangannya indah	Amenities	Positif
62	tempat nya nyaman sejuk makanan nya emmmmm enak banget jadi ketagihan suana rame karyawan nya ramah ramah	Amenities	Positif
63	tempat makan berada "atas pohon" sungguh menarik dicoba	Amenities	Positif
64	makanan cukup enak	Amenities	Positif
65	konsepnya cukup unik membeli makan minum menggunakan uang mainan monopoli	Amenities	Positif
66	makanan sini enak enak harga terjangkau.	Amenities	Positif
67	tempat makan lantai 2 nya sangat nyaman	Amenities	Positif
68	pelayanan yg cukup ramah	Amenities	Positif
69	makanan d foodcourt dgn menu soto bening sangat enak	Amenities	Positif
70	restoran prasmanan (lupa nama) memberikan makanan baik	Amenities	Positif
71	saya pikir adalah tempat bagus makan siang keluarga pasangan.	Amenities	Positif
72	makanan sangat baik	Amenities	Positif
73	sangat merekomendasikan 'restoran burangrang' anda memilih setiap restoran lainnya	Amenities	Positif
74	makanan baik harga baik	Amenities	Positif
75	juga menikmati makan siang restoran burangrang	Amenities	Positif
76	makanan yg dijual variatif belanjanya pake kartu itu bagi anak anak saiya ???? liburan keluarga yg menyenangkan	Amenities	Positif
77	sesekali bolehlah menikmati makanan disini bercengkrama	Amenities	Positif
78	tempat makan yg bagus	Amenities	Positif
79	petugas ramah	Amenities	Positif
80	aneka makanan tersebar dsini rata" harganya 25-40k cukup worth it asal masak	Amenities	Positif
81	makanan nya enak	Amenities	Positif
82	disana banyak menjual jas hujan/payung	Amenities	Netral
83	menjual ponco sandal tersedia dalam kasus tidak mempersiapkan	Amenities	Netral
84	banyak orang menjual ponco sandal	Amenities	Netral
85	untuk makan sebaiknya bawa rumah	Amenities	Netral
86	jadi jangan letakan bawaan khususnya makanan minuman tempat parkir	Amenities	Netral





87	jangan lupa bawa makanan snacks di bawah sana ada penjual makan minum	Amenities	Netral
88	perlu penataan serius tidak merusak tatanan daerah	Amenities	Netral
89	didepannya terdapat banyak warung menjual topi	Amenities	Netral
90	banyak blog-blog mengulas restoran / warung dieng dapat coba lebih	Amenities	Netral
91	makanan mie ongklok khan wonosobo jangan terlewatkan	Amenities	Netral
92	tak lupa membeli buah tangan bunga edolweis (yaitu bunga abadi) dasarnya bunga sangat dilindungi di semua gunung.	Amenities	Netral
93	tidak menu jujur	Amenities	Netral
94	jangan lupa membeli produk olahan buah ciraca	Amenities	Netral
95	bawa jaket tebal makanan banyak	Amenities	Netral
96	warung tradisional banyak d sini	Amenities	Netral
97	sebelah kiri terdapat restoran ala premium sebelah kanan berupa makanan tenant2 disana	Amenities	Netral
98	banyak pilihan restoran sini	Amenities	Netral
99	akhirnya biasa makan burangrang.	Amenities	Netral
100	tempat makan banyak macam bisa coba	Amenities	Netral
101	saat tiba disini banyak booth-booth makanan belum buka	Amenities	Netral
102	orang menjual juga rusak	Amenities	Negatif
103	sepulangnya pasti tukang parkir mencuci motor/mobil meminta imbalan nya hahaha antyara 10-15ribu heheh	Amenities	Negatif
104	sekarang mobilnya parkir jauh	Amenities	Negatif
105	harus membeli jas hujan memakai	Amenities	Negatif
106	menunggu makanan kita....	Amenities	Negatif
107	50 sgd) langkah-langkah bawah licin hujan mereka punya pegangan tangan sisi berada kotak centang	Amenities	Negatif
108	toilet kurang	Amenities	Negatif
109	harus membayar parkir biasanya adalah 20 p	Amenities	Negatif
110	sayang ga bawa makan sendiri	Amenities	Negatif
111	kesan pertama sedikit kesal memasuki gerbang mendapat info petugas per orang dikenakan ttiket masuk rp 15 rb	Amenities	Negatif
112	gift maupun mini foodcourt saja membeli pakai kupon saya malas deposit dulu harus mengantri dulu kasir	Amenities	Negatif
113	makanan ada kurang asik	Amenities	Negatif
114	jaraknya terlalu jauh sebenarnya tempat parkir	Amenities	Negatif
115	30 menit menunggu duduk restoran	Amenities	Negatif
116	dipaksa makan restoran kelas atas	Amenities	Negatif
117	sangat sulit menemukan tempat makan	Amenities	Negatif
118	aku menunggu 1 jam mendapatkan duduk kafe burangrang	Amenities	Negatif



119	aku menunggu orang menyelesaikan makanan dan meninggalkan meja	Amenities	Negatif
120	aku menunggu cukup lama dilayani untuk makan	Amenities	Negatif
121	harus membeli kupon	Amenities	Negatif
122	air terjun bisa dikunjungi rangkaian wisata gunung bromo letaknya cukup dekat	Attractions	Positif
123	yg keren air yg keluar batu ..	Attractions	Positif
124	air terjun madakaripura merupakan air paling indah tumpak sewu di kota sebelahnya lumajang	Attractions	Positif
125	view pertama sungai yg jernih	Attractions	Positif
126	trip mengasyikan sekaligus menjengkelkan.	Attractions	Positif
127	wisata bagus sekali semua orang mudah dicapai	Attractions	Positif
128	recommended pecinta alam lepas...	Attractions	Positif
129	bersih lingkungannya	Attractions	Positif
130	salah satu air terjun terbaik pernah lihat	Attractions	Positif
131	ketinggian 183 meter merupakan air terjun tertinggi kedua sipiso-piso sumut	Attractions	Positif
132	paling menarik untuk menuju airterjun utama	Attractions	Positif
133	cuaca sejuk	Attractions	Positif
134	sangat menakjubkan pemandangan air terjun madakaripura	Attractions	Positif
135	airnya sangat jernih pemandangannya sangat menakjubkan	Attractions	Positif
136	tersedia tempat penitipan barang pembungkus alat elektronik ongkos terjangkau barang aman	Attractions	Positif
137	air terjun indah	Attractions	Positif
138	setinggi 200 m merupakan air terjun tertinggi nomor 2 indonesia	Attractions	Positif
139	akses menuju air terjun sudah mudah	Attractions	Positif
140	air terjun 1 ok banget ..	Attractions	Positif
141	sesungguhnya indah sekali pemandangan dieng plateau	Attractions	Positif
142	kawasan lumayan sejuk.	Attractions	Netral
143	pengunjung diperhatikan tidak berlama-lama didaerah kawah putih	Attractions	Netral
144	tak salah kami sebut kawah selebriti	Attractions	Netral
145	jangan datang kawah sore hari	Attractions	Netral
146	menuju kawah dari pintu masuk terdapat 2 pilihan	Attractions	Netral
147	kawah mempunyai pemandangan alami instagrammable	Attractions	Netral
148	hampir tak pernah luput itinery untuk selalu memasukkan kawah putih tujuan wisata area bandung selatan	Attractions	Netral
149	kawah sikidang bukan sekedar kawah	Attractions	Netral
150	memasuki daerah kawasan sikidang disarankan menggunakan masker	Attractions	Netral
151	kawah sikidang bukan sekedar kawah	Attractions	Netral

152	justru gas tidak berbau lah tergolong area berbahaya terindikasi beracun bisa mengakibatkan kematian	Attractions	Netral
153	sini banyak kawah kecil-kecil yang besar	Attractions	Netral
154	kawah sikidang banyak kawah aktif	Attractions	Netral
155	banyak spot foto bagus.	Attractions	Netral
156	mencapai kawah sikidang tidak perlu mendaki	Attractions	Netral
157	carilah pijakan tepat tidak terpeleset tanah kapur berkerikil kecil membuat tergelincir kanan kiri sudah jurang kawah	Attractions	Netral
158	kawah terlalu dikomersialkan telah berhasil mempertahankan kecantikan alami	Attractions	Netral
159	tempat tidak terlalu ramai beberapa kios souvenir tersedia 24/7	Attractions	Netral
160	ditempat terdapat kawah naman sikidang	Attractions	Netral
161	banyak tempat spot foto berbayar	Attractions	Netral
162	harganya terbilang terlalu mahal / standart	Attractions	Negatif
163	bahkan yang sering mengunjungi pantai papuma pernah bosan mengunjungi	Attractions	Negatif
164	sayang ada sampah-sampah plastik pantai	Attractions	Negatif
165	tiketnya sendiri begitu mahal	Attractions	Negatif

Disetujui oleh,

  
Afifah Yuseka



## LAMPIRAN C HASIL USABILITY TESTING

Nama: Adrian Bestari	
Jabatan: Kepala Divisi Umum Perhutani Pusat	
Pernyataan	Jawaban
1. Saya pikir saya akan sering menggunakan sistem ini	3 (Setuju)
2. Saya rasa sistem ini terlalu kompleks	3 (Setuju)
3. Saya rasa sistem ini mudah untuk digunakan	3 (Setuju)
4. Saya pikir saya akan membutuhkan bantuan teknis untuk dapat menggunakan sistem ini	1 (Tidak Setuju)
5. Saya pikir ada berbagai fungsi dalam sistem ini terintegrasi dengan baik	3 (Setuju)
6. Saya pikir ada terlalu banyak ketidaksesuaian dalam sistem ini	1 (Tidak Setuju)
7. Saya bisa bayangkan bahwa kebanyakan orang akan mudah mempelajari dan menggunakan sistem ini	3 (Setuju)
8. Saya kira sistem ini sangat rumit untuk digunakan	1 (Tidak Setuju)
9. Saya merasa percaya diri dalam menggunakan sistem ini	3 (Setuju)
10. Saya perlu belajar banyak hal sebelum saya bisa memulai menggunakan sistem ini	3 (Setuju)

Jakarta, 31 Mei 2021

  
Adrian Bestari