

Analisis Sentimen dan Pemeringkatan Popularitas Tempat Wisata dengan *Naïve Bayes* dan AHP

Rohmatul Hasanah¹, Wiyli Yustanti²

^{1,2}Program Studi Sistem Informasi, Universitas Negeri Surabaya

¹Rohmatulhasanah16051214013@mhs.unesa.ac.id

²Wiyliyustanti@unesa.ac.id

Abstrak— Kabupaten Banyuwangi terkenal akan keindahan alamnya. Kabupaten Banyuwangi juga memiliki banyak tempat wisata yang menarik wisatawan lokal maupun internasional. Penelitian ini dilakukan untuk melakukan analisa sentimen terhadap tanggapan wisatawan mengenai tempat wisata untuk mengetahui sentimen wisatawan terhadap tempat wisata. Tanggapan yang akan dianalisa diambil dari *Tripadvisor* dan *Twitter* menggunakan teknik *scrapping*. Data yang digunakan adalah 50 data tempat wisata, beserta data sumber mengenai 50 tempat wisata tersebut. Hasil dari analisa sentimen dapat dimanfaatkan dalam pemeringkatan popularitas tempat wisata. Algoritma yang digunakan untuk analisis sentimen dalam penelitian ini adalah *Naïve Bayes*. Nilai akurasi tertinggi dari model yang dihasilkan adalah 70%. Kemudian, dilakukan evaluasi terhadap model dan menghasilkan nilai akurasi sebesar 74%, presisi sebesar 61%, dan recall sebesar 55%. Pemeringkatan dilakukan menggunakan algoritma *Analytic Hierarchy Process*. Kriteria yang dinilai dalam pemeringkatan adalah tagar *Twitter*, tagar *Instagram*, rating *Tripadvisor*, review *Tripadvisor*, dan presentase sentimen. Kriteria presentase sentimen memiliki bobot paling tinggi yakni sebesar 43,8%. Sedangkan kriteria dengan bobot paling rendah adalah tagar *Twitter* dengan nilai sebesar 5,3%. Hasil Penghitungan rasio konsistensi adalah sebesar 0,09 dan dapat dinyatakan valid.

Kata Kunci— Analisis Sentimen, Naive Bayes, Analytic Hierarchy Process, Klasifikasi, Pemeringkatan

I. PENDAHULUAN

Kabupaten Banyuwangi terkenal akan keindahan alamnya. Selain itu, Banyuwangi juga memiliki banyak wisata budaya serta festival yang diselenggarakan tiap tahunnya. Kabupaten Banyuwangi telah menerima banyak penghargaan dalam hal pariwisata baik dari dalam maupun luar negeri. Keberhasilan Banyuwangi dalam hal pariwisata tentunya tidak didapat secara instan. Kabupaten Banyuwangi gencar melakukan pembangunan demi memperbaiki dan memperindah tempat wisata yang ada. Perbaikan dilakukan untuk meningkatkan kenyamanan serta menarik lebih banyak wisatawan. Tanggapan dari wisatawan dibutuhkan untuk mengetahui popularitas sebuah tempat wisata agar tempat wisata yang kurang populer dapat dikembangkan lagi untuk menarik lebih banyak wisatawan Analisis sentimen dapat digunakan untuk mengetahui bagaimana tanggapan wisatawan terhadap tempat wisata yang ada.

Analisis sentimen adalah bagian dari *Natural Language Processing* dan *Machine Learning*. Analisis sentimen melakukan klasifikasi terhadap opini, kemudian

mengelompokkannya kedalam kelompok tertentu yang disesuaikan dengan sentimen dari opini tersebut [1]. Analisis sentimen digunakan untuk menganalisa opini masyarakat terhadap suatu entitas. Entitas ini dapat berupa individu, peristiwa maupun topik tertentu [2]. Analisis sentimen sentimen dapat dilakukan secara manual. Namun, melihat semakin banyaknya wisatawan setiap waktu, analisis sentimen secara manual kurang efisien untuk dilakukan.

Tujuan dari penelitian ini adalah untuk mengetahui kemampuan algoritma *Naive Bayes* dalam mengelompokkan sentimen pada tanggapan wisatawan serta mengetahui kemampuan algoritma *Analytic Hierarchy Process* dalam melakukan pemeringkatan tempat wisata. Sedangkan yang menjadi batasan masalah dalam penelitian ini adalah:

- Peneliti menetapkan 50 tempat wisata di kabupaten Banyuwangi sebagai uji coba
- Data yang akan dianalisis hanyalah tanggapan berbahasa Indonesia mengenai 50 lokasi yang telah ditetapkan peneliti dan didapat dari media sosial melalui proses *scraping*
- Data yang digunakan hanya diambil dari *instagram*, *twitter*, dan *tripadvisor*
- Algoritma yang digunakan untuk mengelompokkan tanggapan wisatawan adalah *Naive Bayes*
- Pemeringkatan popularitas tempat wisata dilakukan menggunakan algoritma *Analytic Hierarchy Process*

Penelitian mengenai analisis sentimen dan pemeringkatan, sebelumnya telah dilakukan oleh Munarwan dan Ardiles Sinaga. Penelitian ini berjudul “Pemanfaatan Analisis Sentimen untuk Pemeringkatan Popularitas Tujuan Wisata” [3]. Penelitian ini dilakukan pada tahun 2017 dengan tujuan menghasilkan informasi peringkat tempat wisata paling populer berdasarkan hasil analisis sentimen terhadap komentar pada media sosial. Analisa sentimen pada penelitian ini dilakukan menggunakan metode *Naive Bayes*. Namun, tidak ada pembobotan pada komponen-komponen yang dinilai sehingga metode pemeringkatan yang digunakan terasa kurang akurat.

Billy Gunawan, Helen Sasty Pratiwi, dan Enda Esyudha Pratama melakukan penelitian berjudul “Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode *Naive Bayes*” pada tahun 2018 [4]. Penelitian ini dilakukan untuk mengetahui umpan balik terhadap suatu merk dagang. Penelitian dilakukan dengan cara mengelompokkan ulasan mengenai merk dagang sesuai dengan kelas sentimennya. Hasil uji dari penelitian ini membuktikan bahwa metode *Naive Bayes* dapat digunakan untuk memprediksi kelas sentimen

dari suatu ulasan. Nilai akurasi tertinggi yang dihasilkan pada penelitian ini adalah sebesar 77,78%

Sebuah penelitian lain berjudul “Pengembangan Aplikasi Analisis Sentimen *Twitter* Menggunakan Metode *Naive Bayes Classifier* (Studi Kasus SAMSAT Kota Malang)” dilakukan oleh Imam Fahrur Rozi, Elok Nur Hamdana, dan Muhammad Balya Iqbal Alfahmi pada tahun 2018 [5]. Penelitian ini bertujuan untuk melakukan analisis sentimen dengan menerapkan pendekatan *Naive Bayes Classifier* dalam klasifikasi kata-katanya. Pada penelitian ini metode *Naive Bayes Classifier* terbukti dapat digunakan dalam klasifikasi analisis sentimen. Hasil uji klasifikasi juga menghasilkan nilai akurasi yang tinggi.

Penelitian berjudul “Sistem Pendukung Keputusan dengan Menggunakan Metode *Analytical Hierarchy Process* (AHP) untuk Menentukan Peringkat Siswa Berdasarkan Nilai *Hard Skill* dan *Soft Skill*” dilakukan oleh Tineke Karunika dan I Made Sudana pada tahun 2015 [6]. Penelitian ini bertujuan untuk membantu wali kelas dalam menentukan peringkat siswa dengan menggunakan metode AHP. Validasi yang dilakukan mendapatkan hasil yang tinggi dan dapat dinyatakan valid.

II. METODE PENELITIAN

A. Pengumpulan Data

Data yang digunakan dalam penelitian ini berupa 50 data tempat wisata, dan data sumber mengenai 50 tempat wisata. Data sumber diambil melalui proses *scraping* pada *Instagram*, *Tripadvisor*, dan *Twitter* menggunakan *library selenium* dan *beautifulsoup4* pada bahasa pemrograman *python*. Rincian data yang akan diambil, yaitu:

- 1) *Tweet*: Jumlah unggahan pada *twitter* yang menyertakan tagar mengenai tempat wisata. *Tweet* kemudian akan di analisa sentimennya, kemudian hasil analisa sentimen akan dijadikan salah satu kriteria dalam pemeringkatan.
- 2) *Tripadvisor*: Jumlah *review* dan *rating* tempat wisata pada *Tripadvisor*. *Review* kemudian akan dianalisa sentimennya. Hasil analisa kemudian akan dijadikan salah satu kriteria dalam pemeringkatan.
- 3) *Instagram*: Jumlah unggahan pada *Instagram* yang menyertakan tagar sesuai dengan tempat wisata.

B. Pengolahan Data

Pegolahan data dilakukan dengan langkah-langkah sebagai berikut:

1) Pelabelan Data

Langkah pertama yang dilakukan sebelum melakukan analisis adalah menyiapkan data yang akan digunakan. Data yang akan digunakan ini diberi label secara manual sesuai dengan kategori sentimennya untuk kemudian digunakan sebagai data training. Kategori sentimen yang akan digunakan adalah netral, positif dan negatif.

TABEL I
PELABELAN DATA

No.	Tanggapan	Kategori
1	Sayang jalannya rusak....yg paling parahh masih jalan kaki mau ke teluknya	2
2	Wisatanya bagus namun untuk jalan menuju kesana kurang bagus jalannya tolong pihak terkait diperbaiki jalannya	0
3	surga tersembunyi dengan pemandangan fantastis. tempat terbaik yang pernah saya kunjungi.	1

Tabel I merupakan contoh data yang sudah dilabeli. Data tanggapan bersentimen netral diberi label angka nol, data positif dilabeli angka satu, sedangkan data dengan sentimen negatif diberi label angka dua.

2) Preprocessing

Ada dua *library python* yang digunakan dalam tahap preprocessing data. *Library* pertama adalah *NLTK Natural Language Toolkit*. *Library* ini digunakan pada proses *case folding*, *tokenizing*, dan *filtering*. *Library* kedua adalah *sastrawi* yang digunakan untuk proses *stemming*.

Berikut ini tahapan dalam *preprocessing*:

a. Case Folding

Pada tahap ini seluruh teks dalam dokumen akan di rubah menjadi huruf kecil, kemudian karakter selain huruf akan dihilangkan.

TABEL II
CASE FOLDING

No.	Sebelum	Sesudah
1	Jalan masih rusak parah, tempat kurang tertata.	jalan masih rusak parah tempat kurang tertata
2	masih banyak sampah dijalan menuju pantai.	masih banyak sampah dijalan menuju pantai

Tabel II menggambarkan contoh dari proses *case folding*. Huruf kapital pada awal kalimat diubah menjadi huruf kecil, kemudian tanda baca seperti koma dan titik dihilangkan.

b. Tokenizing

Pada tahap ini karakter-karakter dalam teks diubah menjadi satuan kata-kata. Contoh proses *tokenizing* dapat dilihat pada Tabel III.

TABEL III
TOKENIZING

No.	Sebelum	Sesudah
1	jalan masih rusak parah tempat kurang tertata	jalan masih rusak parah tempat kurang

No.	Sebelum	Sesudah
		tertata
2	masih banyak sampah dijalan menuju pantai	masih banyak sampah dijalan menuju pantai

Tabel III menunjukkan hasil *tokenizing* dari sebuah kalimat dengan banyak kata, yang dipecah tiap-tiap katanya.

c. *Filtering*

Tahap *filtering* digunakan untuk mengambil kata-kata penting dari hasil *tokenizing* atau menghilangkan kata-kata umum yang tidak penting atau tidak memiliki makna.

TABEL IV
FILTERING

No.	Sebelum	Sesudah
1	jalan masih rusak parah tempat kurang tertata	jalan rusak parah tertata
2	masih banyak sampah dijalan menuju pantai	sampah dijalan pantai

Tabel IV menggambarkan proses *filtering*. Sesudah melalui proses *filtering*, beberapa kata seperti “masih”, “kurang”, dan “menuju”, dihilangkan.

d. *Stemming*

Tahap ini dilakukan untuk menghilangkan imbuhan pada kata dan merubahnya kembali pada bentuk dasarnya.

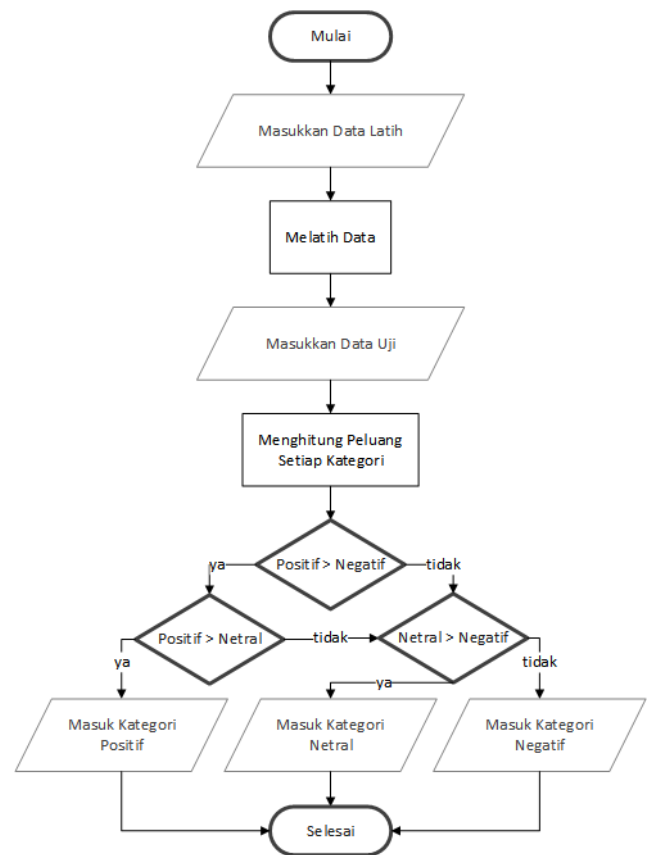
TABEL V
STEMMING

No.	Sebelum	Sesudah
1	jalan rusak parah tertata	jalan rusak parah tata
2	sampah dijalan pantai	sampah jalan pantai

Tabel V menggambarkan proses *stemming*. Setelah proses *stemming*, kata “tertata” diubah menjadi “tata”, dan kata “dijalan” diubah menjadi “jalan”.

3) *Modelling*

Tahap *modelling* dilakukan setelah data melalui *preprocessing*. Tahap *modelling* dilakukan menggunakan algoritma *Naive Bayes*. Gbr 1 menunjukkan tahap-tahap yang akan dilakukan.



Gbr. 1 Proses Naive Bayes

Dalam tahap ini setiap kata dalam kalimat akan dihitung probabilitas kemunculannya pada masing-masing kategori. nilai probabilitas tertinggi akan menunjukkan label kategori dari kalimat tersebut. *Library* yang digunakan dalam tahap ini adalah *library NLTK (Natural Language Toolkit)*.

4) *Evaluation*

Tahap *evaluation* dilakukan untuk mengetahui tingkat akurasi kinerja algoritma yang digunakan. *Evaluation* ini dilakukan dengan menghitung *accuracy*, *precision*, dan *recall* menggunakan *confusion matrix*. *Library* yang digunakan dalam tahap ini adalah *library scikit learn*.

TABEL VI
CONFUSION MATRIX

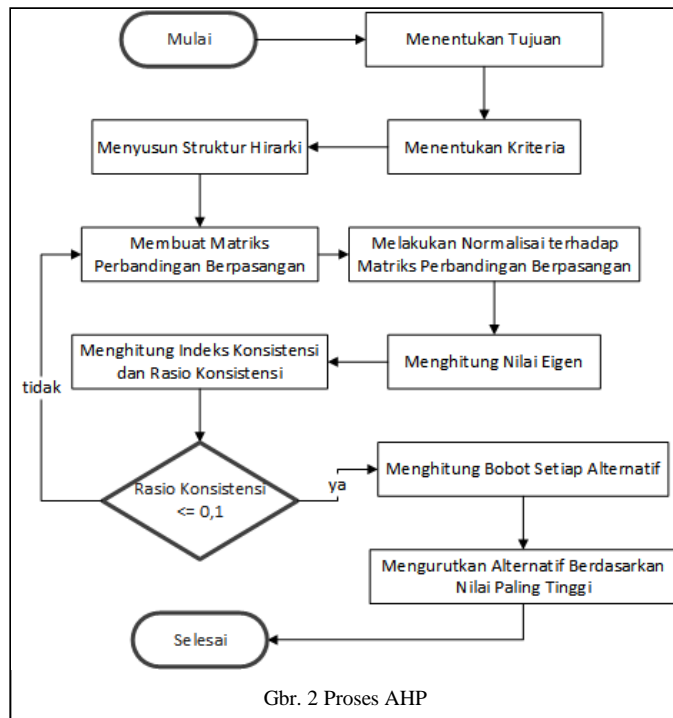
True Class	Predicted Class	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Keterangan :

- TP(true positive) : jumlah data positif yang terdeteksi positif
- TN(true negative) : jumlah data negatif yang terdeteksi negatif
- FP(false positive) : jumlah data negatif yang terdeteksi positif
- FN(false negative) : jumlah data positif yang terdeteksi negatif

5) Ranking

Tahap *ranking* dilakukan menggunakan algoritma AHP (Analytic Hierarchy Process). Data yang telah dikumpulkan diolah menggunakan algoritma AHP untuk mengetahui peringkat dari setiap tempat wisata.



Gbr. 2 Proses AHP

Tahapan dalam proses *ranking* menggunakan algoritma AHP adalah sebagai berikut:

- a. Mendefinisikan masalah keputusan, kemudian menyusunnya secara hierarkis dengan cara menguraikan masalah tersebut kedalam hirarki elemen keputusan yang saling berhubungan.
- b. Membandingkan elemen dalam level yang sama berpasangan, dengan kriteria tertentu yang berada dalam level yang lebih tinggi untuk disusun kedalam matriks perbandingan berpasangan.
- c. Menghitung eigen vektor menggunakan *eigenvector method*.
- d. Menguji konsistensi matriks perbandingan menggunakan *consistency ratio* (CR).
- e. Membuat keputusan dengan cara menjumlahkan prioritas akhir dari alternative.

III. HASIL DAN PEMBAHASAN

A. Data

Data pertama yang dikumpulkan adalah nama tempat wisata sejumlah 50 tempat. 50 nama tempat tersebut ditampilkan dalam Tabel VII

TABEL VII
NAMA TEMPAT WISATA

No.	Nama Tempat
1	Air Terjun Antogan
2	Air Terjun Jagir
3	Air Terjun Kembar
4	Air Terjun Legomoro
5	Air Terjun Lider
6	Air Terjun Selendang Arum
7	Air Terjun Wonorejo
8	Alam Indah Lestari
9	Bukit Mondoleko
10	Desa Wisata Osing
11	Gumuk Candi
12	Pantai Grajagan
13	Kalilo
14	Kawah Ijen
15	Mirah Fantasia
16	Padang Savana Sadengan
17	Pancoran
18	Pantai Bangsring
19	Pantai Blimbingsari
20	Pantai Boom
21	Pantai Cacalan
22	Pantai Cemara
23	Pantai Lampon
24	Pantai Plengkung
25	Pantai Rajegwesi
26	Pantai Sukamade
27	Pantai Wedi Ireng
28	Pulau Merah
29	Pulau Tabuhan
30	Gunung Raung
31	Puncak Asmoro
32	Wisata Pinus Songgon
33	Taman Blambangan
34	Taman Jawatan
35	Taman Nasional Alas Purwo
36	Taman Sritanjung
37	Taman Suruh
38	Taman Tirtawangi

No.	Nama Tempat
39	Teluk Biru
40	Teluk Hijau
41	Umbul Aga Puncak
42	Umbul Bening
43	Umbul Pule
44	Waduk Bajulmati
45	Waduk Sidodadi
46	Pantai Muncar
47	Watu Dodol
48	Watu Gedhek
49	X-Badeng
50	Air Terjun Telunjuk Raung

Pengumpulan data untuk analisis sentimen dilakukan menggunakan *library selenium* dan *beautifulsoup4*. Data yang digunakan dalam analisis sentimen diambil dari *Twitter* dan *Tripadvisor* berdasarkan pada data yang ada pada Tabel VII. Total data yang berhasil dikumpulkan, dapat dilihat pada Tabel VIII.

TABEL VIII
HASIL SCRAPING

Sumber Data	Jumlah	Total
<i>Tripadvisor</i>	1834	3715
<i>Twitter</i>	1871	

Seluruh data kemudian disimpan dalam sebuah *file* berformat *.csv* untuk kemudian dilabeli kaegorinya. Proses pelabelan dilakukan untuk mengetahui kategori sentimen dari data tersebut. Sebelum proses pelabelan, terlebih dahulu dilakukan penghapusan terhadap data-data yang ganda. Total data setelah melalui proses penghapusan dan pelabelan data dapat dilihat pada Tabel IX

TABEL IX
HASIL PELABELAN

No.	Kategori	Jumlah	Total
1	<i>Positive</i>	1396	3391
2	<i>Negative</i>	121	
3	<i>Neutral</i>	1874	

B. Modeling

Tahap Klasifikasi diawali dengan melakukan *training* terhadap data. Pada proses *training*, dilakukan beberapa kali uji coba dengan jumlah data latih dan data uji yang berbeda untuk mendapatkan model dengan nilai akurasi yang paling tinggi. Dari beberapa uji coba yang dilakukan, berikut hasil yang didapatkan

TABEL X
AKURASI MODEL

No.	Data Latih	Data Uji	Akurasi
1	2500	891	68.76%
2	2750	641	70.00%
3	3000	391	67.69%

Berdasarkan isi dari Tabel X, dapat dilihat bahwa akurasi tertinggi adalah pada percobaan kedua dengan nilai akurasi sebesar 70%. Maka model hasil dari *training* percobaan kedua yang akan digunakan dalam proses klasifikasi data.

TABEL XI
HASIL KLASIFIKASI

No.	Kategori	Jumlah	Total
1	<i>Positive</i>	1021	3391
2	<i>Negative</i>	78	
3	<i>Neutral</i>	2292	

Tabel XI merupakan hasil dari klasifikasi data menggunakan model yang telah dibuat.

C. Evaluation

Tahap pertama dalam proses evaluasi membandingkan data hasil dari uji model dengan data yang telah dilabeli secara manual. Proses ini akan menghasilkan *confusion matrix* berikut:

1667	178	28
558	811	27
66	32	23

Setelah *confusion matrix* tersebut didapatkan, kemudian dilakukan penghitungan terhadap nilai akurasi, presisi, serta *recall* untuk mengetahui performa dari model yang telah dibuat. Hasil yang didapatkan adalah nilai akurasi sebesar 0.74, presisi sebesar 0.61, dan *recall* sebesar 0.55. Berdasarkan hasil perhitungan tersebut, dapat disimpulkan bahwa akurasi dari model adalah sebesar 74%, nilai presisi atau ketepatan model dalam memberikan informasi yang sesuai dengan permintaan adalah sebesar 61%, dan nilai keberhasilan model dalam melabeli data sesuai dengan label aslinya adalah sebesar 55%. Selanjutnya, performa klasifikasi model untuk masing-masing kategori, dapat dilihat Tabel XII

TABEL XII
PERFORMA KATEGORI

No.	Kategori	Presisi	Recall
1	<i>Negative</i>	0.29	0.19
2	<i>Neutral</i>	0.73	0.89
3	<i>Positive</i>	0.79	0.58

Berdasarkan Tabel XII, dapat dilihat bahwa nilai presisi dari kategori *negative* adalah sebesar 29%, kategori *neutral* sebesar 73%, dan kategori *positive* sebesar 79%. sedangkan nilai *recall* dari kategori *negative* adalah sebesar 19%, kategori *neutral* sebesar 89%, dan kategori *positive* sebesar 58%. Maka, dapat diartikan bahwa ketepatan model dalam memberikan informasi

yang sesuai dengan permintaan dan keberhasilan model dalam melabeli data sesuai dengan label aslinya masih rendah.

D. Ranking

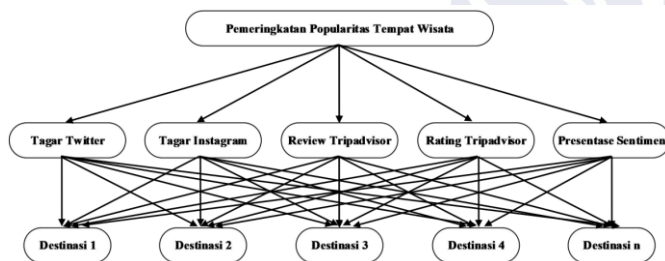
1) Menentukan riteria dan struktur hirarki

Dalam hal ini terdapat 5 kriteria yang akan digunakan. Kriteria-kriteria tersebut di jelaskan pada Tabel XIII

TABEL XIII
KRITERIA

No.	Kriteria	Keterangan
1.	Tagar <i>Twitter</i>	Jumlah Unggahan pada <i>Twitter</i> yang menyertakan tagar mengenai tempat wisata
2.	Tagar <i>Instagram</i>	Jumlah Unggahan pada <i>Instagram</i> yang menyertakan tagar mengenai tempat wisata
3.	<i>Rating Tripadvisor</i>	Rata-rata <i>Rating</i> tempat wisata pada <i>Tripadvisor</i>
4.	<i>Review Tripadvisor</i>	Jumlah Ulasan mengenai tempat wisata pada <i>Tripadvisor</i>
5.	Presentase Sentimen	Total presentase ulasan berkategori positif ditambah dengan setengah dari ulasan berkategori netral

Kriteria-kriteria yang sudah ditentukan, kemudian disusun dalam sebuah struktur hirarki bersama dengan alternatif-alternatif yang ada.



Gbr. 3 Struktur Hirarki.

Berdasarkan Gbr. 3, Level paling atas dari hirarki merupakan tujuan utama yang ingin dicapai. Level dibawahnya, merupakan kriteria-kriteria yang dijadikan pertimbangan dalam penilaian. Kemudian, level paling bawah merupakan alternatif yang akan dinilai

2) Membuat matriks perbandingan berpasangan

Masing-masing kriteria kemudian dibandingkan satu sama lain untu mendapatkan matriks perbandingan berpasangan. Tabel XIV merupakan hasil dari perbandingan masing-masing kriteria.

TABEL XIV
MARIKS PERBANDINGAN BERPASANGAN

	x1	x2	x3	x4	x5
x1	1	1/3	1/5	1/3	1/5
x2	3	1	1/3	3	1/3

	x1	x2	x3	x4	x5
x3	5	3	1	3	1/3
x4	3	1/3	1/3	1	1/5
x5	5	3	3	5	1

Keterangan:

- x1 : *Review Tripadvisor*
- x2 : *Rating Tripadvisor*
- x3 : Tagar *Instagram*
- x4 : Tagar *Twitter*
- x5 : Presentase Sentimen

Matriks tersebut kemudian dinormalisasi dan mendapatkan hasil seperti pada Tabel XV

TABEL XV
NORMALISASI MARIKS PERBANDINGAN BERPASANGAN

	x1	x2	x3	x4	x5
x1	1.00	0.33	0.20	0.33	0.20
x2	3.00	1.00	0.33	3.00	0.33
x3	5.00	3.00	1.00	3.00	0.33
x4	3.00	0.33	0.33	1.00	0.20
x5	5.00	3.00	3.00	5.00	1.00

3) Menghitung eigen vektor

Setelah Matriks dinormalisasi, kemudian dilakukan penghitungan eigen vektor untuk mendapatkan bobot prioritas dari masing-masing kriteria. Tabel XVI merupakan hasil penghitungan bobot untuk masing-masing kriteria.

TABEL XVI
BOBOT PRIORITAS SETIAP KRITERIA

No.	Kriteria	Bobot Prioritas
1	x1	0.0534397802535406
2	x2	0.155986417470562
3	x3	0.259087002552719
4	x4	0.0932594548838518
5	x5	0.438227344839327

4) Menguji konsistensi matriks perbandingan

Penghitungan rasio konsistensi mendapatkan hasil sebesar 0,09 dan dapat dinyatakan valid karena nilai yang dihasilkan adalah kurang dari 0,1.

5) Membuat keputusan

Tahap terakhir yang dilakukan adalah mengalikan nilai kriteria tiap alternatif terhadap bobot prioritas. Hasil perkalian tersebut kemudian diurutkan untuk menghasilkan peringkat dari tempat wisata. Tabel XVII menampilkan 10 tempat wisata dengan peringkat paling tinggi.

TABEL XVII
PERINGKAT TEMPAT WISATA

No.	Destinasi	Peringkat	Nilai
1	Kawah Ijen	1	118470.19
2	Pulau Merah	2	18224.97

No.	Destinasi	Peringkat	Nilai
3	Gunung Raung	3	7985.20
4	Taman Jawatan	4	6084.99
5	Pulau Tabuhan	5	5442.45
6	Teluk Hijau	6	4493.00
7	Pantai Boom	7	3945.03
8	Watu Dodol	8	3669.26
9	Waduk Bajulmati	9	2004.31
10	Pantai Lampon	10	1592.12

Sedangkan pada aplikasi ditampilkan seperti pada Gbr 4 berikut ini.



Gbr. 4 Hasil pada Aplikasi.

Sebagai contoh dalam perhitungan pada Kawah Ijen, nilai masing-masing kriterianya ditampilkan pada Tabel XVIII

TABEL XVIII
NILAI KRITERIA KAWAH IJEN

x1	x2	x3	x4	x5
836	4.616028 708	457,038	127	0.79

Nilai masing-masing kriteria kemudian dikalikan dengan bobot prioritas sehingga perhitungannya menjadi sebagai berikut

$$\begin{aligned}
 &= (836 \times 0.0534397802535406) + (4.616028708 \times 0.155986417470562) + (457,038 \times 0.259087002552719) + \\
 & (127 \times 0.0932594548838518) + (0.79 + 0.438227344839327) \\
 &= 44.67565629 + 0.720037781 + 118412.60547269 + \\
 & 11.8439507702492 + 0.348246495036895 \\
 &= 118470.1934
 \end{aligned}$$

IV. KESIMPULAN

Berdasarkan hasil dari penelitian yang telah dilakukan, dapat diambil kesimpulan bahwa algoritma *Naive Bayes* dapat digunakan untuk melakukan analisis sentimen terhadap tanggapan wisatawan mengenai tempat wisata dikarenakan menghasilkan akurasi sebesar 74%, presisi sebesar 61%, dan recall sebesar 55%. Nilai performa keseluruhan model yg dihasilkan cukup besar, namun, model kurang mampu mendeteksi tanggapan bersentimen negatif. Hal ini dikarenakan jumlah data latih untuk tanggapan bersentimen negatif yang jauh lebih sedikit jika dibandingkan tanggapan bersentimen positif dan netral. Selanjutnya, untuk algoritma *ahp*, berdasarkan hasil yang telah disajikan, matriks perbandingan dapat dinyatakan valid karena memiliki nilai konsistensi dibawah 0,1. Kriteria dengan bobot paling besar adalah presentase sentimen sedangkan kriteria dengan bobot paling sedikit adalah tagar twitter.

Berdasarkan penelitian yang telah dilakukan, saran yang dapat diberikan untuk penelitian adalah;

- 1) Menambah jumlah data yang digunakan
- 2) Mengupayakan agar jumlah data tiap kategorinya tidak jauh berbeda
- 3) Melakukan normalisasi kata sebelum *preprocessing*
- 4) Menambah kriteria dalam tahap pemeringkatan

REFERENSI

- [1] Y. Cahyono, "Analisis sentiment pada sosial media twitter menggunakan Naive Bayes classifier dengan feature selection particle swarm optimization dan term frequency," *J. Inform. Univ. pamulang*, vol. 2, no. 1, pp. 14–19, 2017.
- [2] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms adn application: A survey," *Ain Shams Eng. J.*, vol. 5, no. 4, pp. 1093–1113, 2014.
- [3] S. Murnawan, "Pemanfaatan analisis sentimen untuk pemeringkatan popularitas tujuan wisata," *JPPJ*, vol. 7, no. 2, pp. 109–120, 2017.
- [4] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, "Sistem Analisis Sentimen pada Ulasan Produk menggunakan Metode Naive Bayes," *J. Edukasi dan Penelit. Inform.*, vol. 4, no. 2, pp. 113–118, 2018.
- [5] I. F. Rozi, E. N. Hamdana, and M. B. I. Alfahmi, "Pengembangan aplikasi analisis sentimen twitter menggunakan metode Naive Bayes classifier (studi kasus: SAMSAT kota Malang)," *J. Inform. Polinema*, vol. 4, no. 2, pp. 149–154, 2018.
- [6] T. Karunika and I. M. Sudana, "Sistem pendukung keputusan dengan menggunakan metode analytical hierarchy process (AHP) untuk menentukan peringkat siswa berdasarkan nilai hard skill dan soft skill," *EDUKOM*, vol. 2, no. 1, pp. 1–6, 2015.
- [7] R. Umar, A. Fadlil, and Yuminah, "Sistem Pendukung Keputusan dengan Metode AHP untuk Penilaian Kompetensi Soft Skill Karyawan," *J. Ilmu Komput. dan Inform.*, vol. 4, no. 1, pp. 27–34, 2018.