

ANALISIS SENTIMEN PADA WISATA DIENG DENGAN ALGORITMA K-NEAREST NEIGHBOR (K-NN)

Hidayatus Sibyan, Nur Hasanah

Universitas Sains Al-Qur'an (UNSIQ) Wonsoobo

Email : hsibyan@unsiq.ac.id

085291041812

INFO ARTIKEL

Riwayat Artikel :

Diterima : 27 November 2021

Disetujui : 29 Desember 2021

Kata Kunci :

analisis sentimen; ulasan; wisata dieng; k-nearest neighbor.

ABSTRAK

Tujuan penelitian ini adalah untuk melakukan identifikasi dan klasifikasi ulasan pengunjung wisata dieng melalui analisis sentimen. Dengan adanya analisis sentimen ini diharapkan bisa bermanfaat untuk pengelola wisata sebagai bahan pertimbangan dalam melakukan perbaikan atas kekurangan pada tempat wisata yang dikelola. Serta bisa bermanfaat bagi para wisatawan sebelum memutuskan untuk berkunjung ke destinasi wisata tanpa harus membaca ulasan satu per satu. Metode K-Nearest Neighbor digunakan untuk melakukan klasifikasi data sentimen pada wisata dieng. Klasifikasi data dibagi ke dalam sentimen positif, negatif.

Penelitian ini menghasilkan sebuah sistem analisis sentimen yang dapat mengidentifikasi dan mengklasifikasi ulasan pengunjung wisata dieng dengan menggunakan algoritma K-Nearest Neighbor (K-NN) dengan hasil dari pengujian akurasi sebesar 86%.

ARTICLE INFO

Article History :

Received : 27 November 2021

Accepted : 29 Desember 2021

Keywords:

sentiment analysis; review; dieng tourism; k-nearest neighbor

ABSTRACT

The purpose of this study is to identify and classify reviews of visitors to Dieng tourism through sentiment analysis. With this sentiment analysis, it is hoped that it can be useful for tourism managers as a consideration in making improvements to the deficiencies in managed tourist attractions. And it can be useful for tourists before deciding to visit tourist destinations without having to read reviews one by one. The K-Nearest Neighbor method is used to classify sentiment data on Dieng tourism. The classification of data is divided into positive, negative sentiment. This study produces a sentiment analysis system that can identify and classify reviews of visitors to Dieng tourism using the K-Nearest Neighbor (K-NN) algorithm with the results of an accuracy test of 86%.

1. PENDAHULUAN

Wonosobo sebagai salah satu kabupaten di Jawa Tengah memiliki banyak destinasi wisata unggulan. Salah satu yang menjadi andalan pariwisata wonosobo dan menjadi unggulan jawa tengah adalah Dataran Tinggi Dieng yang kerap juga disebut sebagai Negeri di Atas Awan. Dataran Tinggi Dieng atau yang dikenal juga dengan Dieng Plateu merupakan kawasan wisata pegunungan di kabupaten Wonosobo dan Banjarnegara yang memiliki banyak pilihan wisata yang indah sehingga menjadi daerah kunjungan yang direkomendasikan. Alam pegunungan yang indah, udara yang dingin, sunrise yang mempesona, hingga budaya dan kuliner yang khas menjadi daya tarik tersendiri dari Dataran Tinggi Dieng ini. Apalagi setiap tahunnya di daerah ini digelar event Dieng Culture Festival yang merupakan sebuah acara tahunan yang menampilkan pagelaran seni dan budaya khas Dataran Tinggi Dieng sehingga membuat daerah ini banyak dikunjungi oleh para wisatawan baik lokal maupun mancanegara.

Dari sekian banyak pengunjung yang datang, tidak sedikit dari mereka meninggalkan kesan-kesan berupa review atau ulasan terhadap tempat yang telah dikunjunginya. Ulasan ini bisa menjadi informasi yang sangat penting bagi pengelola wisata maupun bagi wisatawan lainnya. Bagi pengelola wisata informasi ini sebagai bahan evaluasi dan perbaikan terhadap kekurangan-kekurangan yang ada guna meningkatkan kualitas tempat wisata. Sedangkan bagi wisatawan lain informasi ini bisa bermanfaat bagi mereka yang dijadikan sebagai bahan pertimbangan sebelum memutuskan untuk berkunjung ke tempat wisata tersebut.

Namun banyaknya jumlah ulasan dan beragamnya ulasan yang ada membuat sulit untuk menyimpulkan isi dari ulasan-ulasan tersebut karena harus membaca satu per satu yang tentunya akan membutuhkan waktu yang cukup lama. Sehingga diperlukan sebuah alat bantu analisis yang bisa digunakan untuk menyimpulkan ulasan dengan cepat yaitu dengan analisis sentimen.

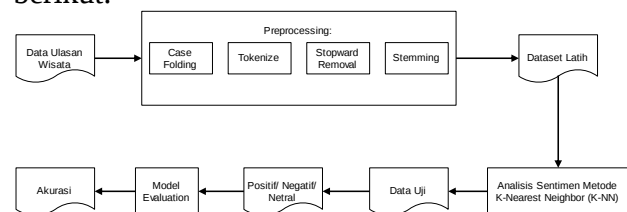
Analisis sentimen merupakan suatu bidang ilmu untuk menganalisis suatu sentimen, opini, pendapat, emosi, sikap, evaluasi, penilaian seseorang terhadap suatu produk, individu,

organisasi, masalah, peristiwa ataupun topik tertentu. [1]

Selanjutnya sentimen akan diklasifikasikan menjadi tiga yaitu positif, negatif, dan netral. Untuk mengklasifikasikannya tersebut bisa menggunakan metode K_Nearest Neighbor (K-NN) yang merupakan sebuah metode machine learning yang dapat melakukan klasifikasi terhadap objek berdasarkan dataset atau data latih berdasarkan jarak yang paling dekat dengan objek tersebut. [2]

2. METODE

Tahapan penelitian analisis sentimen ini menggunakan metode K-Nearest Neighbor (K-NN) dengan tahapan seperti pada gambar berikut:



Gambar 1. Model Penelitian

1. Data ulasan adalah data penelitian yang telah terkumpul
2. Proses preprocessing atau pengolahan data awal yaitu suatu tahapan untuk mengolah data mentah agar bisa digunakan sebagai dataset latih dan uji
3. Metode K-Nearest Neighbor digunakan untuk melakukan klasifikasi data opini/ sentimen pada wisata dieng. Klasifikasi data dibagi ke dalam sentimen positif, negatif, dan netral
4. Selanjutnya akan dilakukan pengujian sehingga menghasilkan nilai akurasi dari hasil klasifikasi yang telah dilakukan.

Metode Pengumpulan Data

Data yang dibutuhkan pada penelitian ini adalah data ulasan dari pengunjung tentang objek wisata yang telah dikunjunginya.

1. Tinjauan Langsung

Data dikumpulkan secara langsung dari objek wisata melalui metode observasi, angket, dan wawancara terhadap pengelola wisata dan wisatawan yang sedang berkunjung di tempat wisata. Selain itu data juga bisa diambil secara online melalui ulasan yang terdapat di google maps, dan ulasan di website tripadvisor.

2. Tinjauan Pustaka

Pengumpulan data referensi berupa literatur yang berhubungan dengan topik penelitian baik berupa buku maupun jurnal

Teknik Analisis Data

Pengolahan data pada penelitian ini dimulai dari pengumpulan data, kemudian preprocessing dan selanjutnya akan dianalisis dengan menggunakan bahasa pemrograman python.

3. HASIL DAN PEMBAHASAN

1. Text Preprocessing

Text processing merupakan suatu proses untuk menyeleksi data text agar menjadi lebih terstruktur lagi, sehingga data menjadi siap untuk dilakukan analisis. Pada proses text processing ini terdapat beberapa tahapan diantaranya:

a. Case Folding

Merupakan tahapan untuk merubah bentuk kata menjadi sama bentuknya yaitu menjadi lower case semua atau upper case semua. Pada contoh data di bawah ini akan diubah menjadi lower case semua.

Tabel 1. Data Awal dan Hasil Case Folding

Kode	Data Awal	Hasil Case Folding
U1	cukup bagus dan informatif ada film dokumentasi tentang dieng dan segala keindahannya	cukup bagus dan informatif ada film dokumentasi tentang dieng dan segala keindahannya
U2	terlalu bnyak tiket yg harus di bayar. Malesin	terlalu bnyak tiket yg harus di bayar. malesin
U3	Dieng sangat indah dan nyaman untuk dikunjungi..masyarakatnya juga ramah dan menyenangkan	dieng sangat indah dan nyaman untuk dikunjungi..masyar akatnya juga ramah dan menyenangkan
U4	Tempat yg mengasyikan..udara bersih segar..pemandangann ya bgt indah..pantas sbg destinasi utk tempat liburan yg nyaman	tempat yg mengasyikan..udara bersih segar..pemandanga nnya bgt indah..pantas sbg destinasi utk tempat liburan yg nyaman

U5	Tempat terbaik untuk mendapatkan spot foto sunrise dan sering dijuluki sebagai Golden Sunrire Terbaik di Asia Tenggara. Akses sangat mudah dan jalan bagus ber aspal, akses menuju ke bukit sangat aman karena sudah mengalami perbaikan	tempat terbaik untuk mendapatkan spot foto sunrise dan sering dijuluki sebagai golden sunrire terbaik di asia tenggara. akses sangat mudah dan jalan bagus ber aspal, akses menuju ke bukit sangat aman karena sudah mengalami perbaikan
U6	semakin lama semakin KUMUH. Sampah dimana Mana, parkir tidak teratur Dan krn main panas debu dimana Mana.	semakin lama semakin kumuh. sampah dimana mana, parkir tidak teratur dan krn main panas debu dimana mana.
U7	Terakhir kesana, , kliatan kurang terawat lingkungannya,, naik ke batu datar banyak pos pungli... Karcis bodong 10.000. Padahal udah masuk bayar... Ko di atas ada yg mintain lagi...	terakhir kesana, , kliatan kurang terawat lingkungannya,, naik ke batu datar banyak pos pungli... karcis bodong 10.000. padahal udah masuk bayar... ko di atas ada yg mintain lagi...
U8	Agak kotor,pas datang telaga kering jadi warna juga kurang terlihat,dan bau	agak kotor,pas datang telaga kering jadi warna juga kurang terlihat,dan bau
U9	Alam yg asri,udara sehat membuat nyaman untuk bersantai bersama keluarga	alam yg asri,udara sehat membuat nyaman untuk bersantai bersama keluarga

b. Tokenizing

Tahapan tokenizing adalah mengidentifikasi kata dalam teks menjadi beberapa urutan yang terpotong oleh spasi atau karakter special dan menghilangkan tanda baca dan angka pada teks. Berikut contoh hasil dari tokenizing

Tabel 2. Hasil Tokenizing

Kode	Case Folding	Hasil Tokenizing
U1	cukup bagus dan informatif ada film dokumentasi tentang dieng dan segala keindahannya	cukup bagus dan informatif ada film dokumentasi tentang dieng dan segala keindahannya

U2	terlalu bnyak tiket yg harus di bayar. malesin	terlalu bnyak tiket yg harus di bayar malesin
U3	dieng sangat indah dan nyaman untuk dikunjungi..masyarakat nya juga ramah dan menyenangkan	dieng sangat indah dan nyaman untuk dikunjungi masyarakatnya juga ramah dan menyenangkan
U4	tempat yg mengasyikan..udara bersih segar..pemandangannya bgt indah..pantas sbg destinasi utk tempat liburan yg nyaman	tempat yg mengasyikan udara bersih segar pemandangannya bgt indah pantas sbg destinasi utk tempat liburan yg nyaman
U5	tempat terbaik untuk mendapatkan spot foto sunrise dan sering dijuluki sebagai golden sunrire terbaik di asia tenggara. akses sangat mudah dan jalan bagus ber aspal, akses menuju ke bukit sangat aman karena sudah mengalami perbaikan	tempat terbaik untuk mendapatkan spot foto sunrise dan sering dijuluki sebagai golden sunrire terbaik di asia tenggara akses sangat mudah dan jalan bagus ber aspal akses menuju ke bukit sangat aman karena sudah mengalami perbaikan
U6	semakin lama semakin kumuh. sampah dimana mana, parkir tidak teratur dan krn main panas debu dimana mana.	semakin lama semakin kumuh sampah dimana mana parkir tidak teratur dan krn main panas debu dimana mana
U7	terakhir kesana, , kliatan kurang terawat lingkungannya,,, naik ke batu datar banyak pos pungli... karcis bodong 10.000. padahal udah masuk bayar... ko di atas ada yg mintain lagi...	terakhir kesana kliatan kurang terawat lingkungannya naik ke batu datar banyak pos pungli karcis bodong padahal udah masuk bayar ko di atas ada yg mintain lagi
U8	agak kotor,pas datang telaga kering jadi warna juga kurang terlihat,dan bau	agak kotor pas datang telaga kering jadi warna juga kurang terlihat dan bau
U9	alam yg asri,udara sehat membuat nyaman untuk bersantai bersama keluarga	alam yg asri udara sehat membuat nyaman untuk bersantai bersama keluarga

c. Stop words removal

Membuang kata-kata yang sering muncul dan bersifat umum, kurang menunjukkan relevansinya dengan teks. Kata-kata yang akan dibuang didefinisikan dalam stopword list.

Tabel 3. Hasil Stop words Removal

Kode	Tokenizing	Hasil Stopword Removal
U1	cukup bagus dan informatif ada film dokumentasi tentang dieng dan segala keindahannya	bagus informatif film dokumentasi dieng segala keindahannya
U2	terlalu bnyak tiket yg harus di bayar malesin	terlalu banyak tiket harus bayar malesin
U3	dieng sangat indah dan nyaman untuk dikunjungi masyarakatnya juga ramah dan menyenangkan	dieng sangat indah nyaman untuk dikunjungi masyarakatnya juga ramah menyenangkan
U4	tempat yg mengasyikan udara bersih segar pemandangannya bgt indah pantas sbg destinasi utk tempat liburan yg nyaman	tempat mengasyikan udara bersih segar pemandangannya begitu indah pantas sebagai destinasi untuk tempat liburan nyaman
U5	tempat terbaik untuk mendapatkan spot foto sunrise dan sering dijuluki sebagai golden sunrire terbaik di asia tenggara akses sangat mudah dan jalan bagus ber aspal akses menuju ke bukit sangat aman karena sudah mengalami perbaikan	tempat terbaik untuk mendapatkan spot foto sunrise sering dijuluki sebagai golden sunrire terbaik di asia tenggara akses sangat mudah dan jalan bagus ber aspal akses menuju bukit sangat aman sudah mengalami perbaikan
U6	semakin lama semakin kumuh sampah dimana mana parkir tidak teratur dan krn main panas debu dimana mana	semakin lama semakin kumuh sampah dimana mana parkir tidak teratur panas debu dimana mana

U7	terakhir kesana kliatan kurang terawat lingkungannya naik ke batu datar banyak pos pungli karcis bodong padahal udah masuk bayar ko di atas ada yg mintain lagi	terakhir kesana kliatan kurang terawat lingkungannya naik batu datar banyak pos pungli karcis bodong 10000 padahal udah masuk bayar di atas ada yg mintain lagi
U8	agak kotor pas datang telaga kering jadi warna juga kurang terlihat dan bau	agak kotor pas datang telaga kering jadi warna juga kurang terlihat bau
U9	alam yg asri udara sehat membuat nyaman untuk bersantai bersama keluarga	alam asri udara sehat membuat nyaman untuk bersantai bersama keluarga

U6	semakin lama semakin kumuh sampah dimana mana parkir tidak teratur panas debu dimana mana	lama, kumuh, sampah, parkir, atur, panas, debu
U7	terakhir kesana kliatan kurang terawat lingkungannya naik batu datar banyak pos pungli karcis bodong 10000 padahal udah masuk bayar di atas ada yg mintain lagi	kurang, rawat, lingkungan, pungli, karcis, bodong, bayar, minta
U8	agak kotor pas datang telaga kering jadi warna juga kurang terlihat bau	kotor, datang, telaga, kering, warna, lihat, bau
U9	alam asri udara sehat membuat nyaman untuk bersantai bersama keluarga	alam, asri, udara, sehat, nyaman, santai, keluarga

d. Stemming

Melakukan pengelompokan kata-kata lain yang memiliki kata dasar dan arti yang serupa namun memiliki bentuk atau form yang berbeda karena mendapatkan imbuhan yang berbeda

Tabel 4. Hasil Stemming

Kode	Hasil Stopword Removal	Hasil Stemming
U1	bagus informatif film dokumentasi dieng segala keindahannya	bagus, informatif, film, dokumentasi, indah
U2	terlalu banyak tiket harus bayar malesin	banyak, tiket, bayar, males
U3	dieng sangat indah nyaman untuk dikunjungi masyarakatnya juga ramah menyenangkan	indah, nyaman, kunjung, masyarakat, ramah, senang
U4	tempat mengasyikan udara bersih segar pemandangannya begitu indah pantas sebagai destinasi untuk tempat liburan nyaman	asyik, bersih, segar, indah, destinasi, liburan, nyaman
U5	tempat terbaik untuk mendapatkan spot foto sunrise sering dijuluki sebagai golden sunire terbaik di asia tenggara akses sangat mudah dan jalan bagus ber aspal akses menuju bukit sangat aman sudah mengalami perbaikan	baik, spot, foto, sunrise, golden, sunrise, baik, mudah, jalan, bagus, aspal, bukit, aman, baik

2. Analisis Sentimen dengan KNN

Setelah data diproses melalui text preprocessing, langkah selanjutnya adalah melakukan analisis sentiment dengan menggunakan metode K-Nearest Neighbor (KNN).

Kode	Ulasan	Sentimen
U1	bagus informatif film dokumentasi dieng segala keindahannya	??? (akan diprediksi)
U2	terlalu banyak tiket harus bayar malesin	Negatif
U3	dieng sangat indah nyaman untuk dikunjungi masyarakatnya juga ramah menyenangkan	Positif
U4	tempat mengasyikan udara bersih segar pemandangannya begitu indah pantas sebagai destinasi untuk tempat liburan nyaman	Positif
U5	tempat terbaik untuk mendapatkan spot foto sunrise sering dijuluki sebagai golden sunire terbaik di asia tenggara akses sangat mudah dan jalan bagus ber aspal akses menuju bukit sangat aman sudah mengalami perbaikan	Positif

U6	semakin lama semakin kumuh sampah dimana mana parkir tidak teratur panas debu dimana mana	Negatif
U7	terakhir kesana kliatan kurang terawat lingkungannya naik batu datar banyak pos pungli karcis bodong 10000 padahal udah masuk bayar di atas ada yg mintain lagi	Negatif
U8	agak kotor pas datang telaga kering jadi warna juga kurang terlihat bau	Negatif
U9	alam asri udara sehat membuat nyaman untuk bersantai bersama keluarga	Positif

	22	foto					1					1	9	0,954
	23	sunrise					2					1	9	0,954
	24	golden					1					1	9	0,954
	25	mudah					1					1	9	0,954
	26	jalan					1					1	9	0,954
	27	aspal					1					1	9	0,954
	28	bukit					1					1	9	0,954
	29	aman					1					1	9	0,954
	30	lama						1				1	9	0,954
	31	kumuh						1				1	9	0,954
	32	sampah						1				1	9	0,954
	33	parkir						1				1	9	0,954
	34	atur						1				1	9	0,954
	35	panas						1				1	9	0,954
IDF (log D/d _f)	36	debu						1				1	9	0,954
	37	kurang							1			1	9	0,954
	38	rawat							1			1	9	0,954
	39	lingkungan							1			1	9	0,954
	40	pungli							1			1	9	0,954
	41	karcis							1			1	9	0,954
	42	bodong							1			1	9	0,954
	43	minta							1			1	9	0,954
	44	kotor								1		1	9	0,954
	45	datang								1		1	9	0,954
	46	telaga								1		1	9	0,954
	47	kering								1		1	9	0,954
48	warna								1		1	9	0,954	
49	lihat								1		1	9	0,954	
50	bau								1		1	9	0,954	
51	alam									1	1	9	0,954	
52	asri									1	1	9	0,954	
53	udara									1	1	9	0,954	
54	sehat									1	1	9	0,954	
55	santai									1	1	9	0,954	
56	keluarga									1	1	9	0,954	

a. Menghitung bobot kata (TF-IDF)

No.	Term	tf									df	D/df	IDF ₃₆ (log D/df)
		U1	U2	U3	U4	U5	U6	U7	U8	U9			
1	bagus	1				1					2	4,5	0,653
2	informatif	1									1	9	0,954
3	film	1									1	9	0,954
4	dokumentasi	1									1	9	0,954
5	indah	1		1	1						3	3	0,477
6	banyak		1								1	9	0,954
7	tiket		1								1	9	0,954
8	bayar		1					1			2	4,5	0,653
9	malas		1								1	9	0,954
10	nyaman			1	1					1	3	3	0,477
11	kunjung			1							1	9	0,954
12	masyarakat			1							1	9	0,954
13	ramah			1							1	9	0,954
14	senang			1							1	9	0,954
15	asyik				1						1	9	0,954
16	bersih				1						1	9	0,954
17	segar				1						1	9	0,954
18	destinasi				1						1	9	0,954
19	liburan				1						1	9	0,954
20	baik					3					1	9	0,954
21	spot				1						1	9	0,954

No.	W								
	U1	U2	U3	U4	U5	U6	U7	U8	U9

1	0,653	0,000	0,000	0,000	0,653	0,000	0,000	0,000	47000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000
2	0,954	0,000	0,000	0,000	0,000	0,000	0,000	0,000	48000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000
3	0,954	0,000	0,000	0,000	0,000	0,000	0,000	0,000	49000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000
4	0,954	0,000	0,000	0,000	0,000	0,000	0,000	0,000	50000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000
5	0,477	0,000	0,477	0,477	0,000	0,000	0,000	0,000	51000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954
6	0,000	0,954	0,000	0,000	0,000	0,000	0,000	0,000	52000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954
7	0,000	0,954	0,000	0,000	0,000	0,000	0,000	0,000	53000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954
8	0,000	0,653	0,000	0,000	0,000	0,000	0,653	0,000	54000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954
9	0,000	0,954	0,000	0,000	0,000	0,000	0,000	0,000	55000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954
10	0,000	0,000	0,477	0,477	0,000	0,000	0,000	0,000	56477	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954
11	0,000	0,000	0,954	0,000	0,000	0,000	0,000	0,000	0,000									
12	0,000	0,000	0,954	0,000	0,000	0,000	0,000	0,000	No.	WU1 *WDi								
13	0,000	0,000	0,954	0,000	0,000	0,000	0,000	0,000		U2	U3	U4	U5	U6	U7	U8	U9	
14	0,000	0,000	0,954	0,000	0,000	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,427	0,000	0,000	0,000	0,000	
15	0,000	0,000	0,000	0,954	0,000	0,000	0,000	0,000	2,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
16	0,000	0,000	0,000	0,954	0,000	0,000	0,000	0,000	3,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
17	0,000	0,000	0,000	0,954	0,000	0,000	0,000	0,000	4,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
18	0,000	0,000	0,000	0,954	0,000	0,000	0,000	0,000	5,000	0,000	0,228	0,228	0,000	0,000	0,000	0,000	0,000	
19	0,000	0,000	0,000	0,954	0,000	0,000	0,000	0,000	6,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
20	0,000	0,000	0,000	0,000	2,863	0,000	0,000	0,000	7,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
21	0,000	0,000	0,000	0,000	0,954	0,000	0,000	0,000	8,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
22	0,000	0,000	0,000	0,000	0,954	0,000	0,000	0,000	9,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
23	0,000	0,000	0,000	0,000	1,908	0,000	0,000	0,000	10,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
24	0,000	0,000	0,000	0,000	0,954	0,000	0,000	0,000	11,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
25	0,000	0,000	0,000	0,000	0,954	0,000	0,000	0,000	12,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
26	0,000	0,000	0,000	0,000	0,954	0,000	0,000	0,000	13,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
27	0,000	0,000	0,000	0,000	0,954	0,000	0,000	0,000	14,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
28	0,000	0,000	0,000	0,000	0,954	0,000	0,000	0,000	15,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
29	0,000	0,000	0,000	0,000	0,954	0,000	0,000	0,000	16,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
30	0,000	0,000	0,000	0,000	0,000	0,954	0,000	0,000	17,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
31	0,000	0,000	0,000	0,000	0,000	0,954	0,000	0,000	18,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
32	0,000	0,000	0,000	0,000	0,000	0,954	0,000	0,000	19,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
33	0,000	0,000	0,000	0,000	0,000	0,954	0,000	0,000	20,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
34	0,000	0,000	0,000	0,000	0,000	0,954	0,000	0,000	21,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
35	0,000	0,000	0,000	0,000	0,000	0,954	0,000	0,000	22,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
36	0,000	0,000	0,000	0,000	0,000	0,954	0,000	0,000	23,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
37	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000	24,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
38	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000	25,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
39	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000	26,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
40	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000	27,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
41	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000	28,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
42	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000	29,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
43	0,000	0,000	0,000	0,000	0,000	0,000	0,954	0,000	30,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
44	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954	31,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
45	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954	32,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	
46	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,954	33,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	

34	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,91	0,000	0,00	0,00	0,00	0,00
35	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,91	0,000	0,00	0,00	0,00	0,00
36	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,91	0,000	0,00	0,00	0,00	0,00
37	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,91	0,000	0,00	0,00	0,00	0,00
38	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,91	0,000	0,00	0,00	0,00	0,00
39	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,91	0,000	0,00	0,00	0,00	0,00
40	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	8,195	0,00	0,00	0,00	0,00
41	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,911	0,00	0,00	0,00	0,00
42	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,911	0,00	0,00	0,00	0,00
43	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,911	0,00	0,00	0,00	0,00
44	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	3,642	0,00	0,00	0,00	0,00
45	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,911	0,00	0,00	0,00	0,00
46	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,911	0,00	0,00	0,00	0,00
47	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,911	0,00	0,00	0,00	0,00
48	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,911	0,00	0,00	0,00	0,00
49	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,911	0,00	0,00	0,00	0,00
50	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,911	0,00	0,00	0,00	0,00
51	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,911	0,00	0,00	0,00	0,00
52	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,91	0,91	0,00	0,00	0,00
53	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,000	1	0	0	0
54	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,000	1	0	0	0
55	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,000	1	0	0	0
56	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,000	1	0	0	0
	0,000	0,228	0,228	0,427	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,00	0,000	0,91	0,00	0,00	0,00

b. Menghitung tingkat kemiripan (Cosine Similarity)

No	Panjang Vektor									U9	U10	U11	U12	U13	U14	U15	U16
	U1	U2	U3	U4	U5	U6	U7	U8									
1	0,427	0,000	0,000	0,000	0,427	0,000	0,000	0,000	0,008	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
2	0,911	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,009	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
3	0,911	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
4	0,911	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
5	0,228	0,000	0,228	0,228	0,000	0,000	0,000	0,000	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
6	0,000	0,911	0,000	0,000	0,000	0,000	0,000	0,000	0,003	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
7	0,000	0,911	0,000	0,000	0,000	0,000	0,000	0,000	0,004	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
8	0,000	0,427	0,000	0,000	0,000	0,000	0,427	0,000	0,005	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
9	0,000	0,911	0,000	0,000	0,000	0,000	0,000	0,000	0,006	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
10	0,000	0,000	0,228	0,228	0,000	0,000	0,000	0,000	0,227	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
11	0,000	0,000	0,911	0,000	0,000	0,000	0,000	0,000	0,008	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
12	0,000	0,000	0,911	0,000	0,000	0,000	0,000	0,000	0,009	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,000
13	0,000	0,000	0,911	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,910	0,910
14	0,000	0,000	0,911	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,910	1

52	0,00 0	0,00 0	0,00 0	0,00 0	0,000 0	0,00 0	0,00 0	0,00 0	0,91 1
53	0,00 0	0,00 0	0,00 0	0,00 0	0,000 0	0,00 0	0,00 0	0,00 0	0,91 1
54	0,00 0	0,00 0	0,00 0	0,00 0	0,000 0	0,00 0	0,00 0	0,00 0	0,91 1
55	0,00 0	0,00 0	0,00 0	0,00 0	0,000 0	0,00 0	0,00 0	0,00 0	0,91 1
56	0,00 0	0,00 0	0,00 0	0,00 0	0,000 0	0,00 0	0,00 0	0,00 0	0,91 1
57	3,38 6	3,15 8	4,09 8	5,00 8	19,54 9	6,37 4	6,80 1	6,37 4	5,69 1
	1,84 0	1,77 7	2,02 4	2,23 8	4,421	2,52 5	2,60 8	2,52 5	2,38 6

c. Menentukan Nilai Sentimen

$$\begin{aligned} \text{Cos}(U1, U2) &= 0 / (1,840 * 1,777) = 0 \\ \text{Cos}(U1, U3) &= 0,228 / (1,840 * 2,024) = 0,061 \\ \text{Cos}(U1, U4) &= 0,228 / (1,840 * 2,238) = 0,055 \\ \text{Cos}(U1, U5) &= 0,427 / (1,840 * 4,421) = 0,052 \\ \text{Cos}(U1, U6) &= 0 / (1,840 * 2,525) = 0 \\ \text{Cos}(U1, U7) &= 0 / (1,840 * 2,608) = 0 \\ \text{Cos}(U1, U8) &= 0 / (1,840 * 2,525) = 0 \\ \text{Cos}(U1, U9) &= 0 / (1,840 * 2,386) = 0 \end{aligned}$$

Dari hasil tersebut dapat diurutkan nilai similarity U1 dari yang tertinggi ke yang terendah yaitu :

1. U3 (Positif)
2. U4 (Positif)
3. U5 (Positif)

Jika dipilih nilai k untuk KNN adalah 3 maka akan dipilih 3 nilai similarity yang tertinggi. Dari 3 nilai tertinggi tersebut, kelas sentimen yang paling banyak muncul adalah positif sehingga sistem akan mengklasifikasikan U1 ke dalam sentimen **Positif**.

3. Pengujian Akurasi

Pengujian dilakuann dengan menggunakan dataset sebanyak 147 ulasan untuk data training yang terbagi menjadi klasifikasi 74 positif, dan 73 data negatif. Sedangkan dataset yang digunakan untuk data testing sebanyak 37 ulasan. Uji akurasi dengan menggunakan bahasa pemrograman python dan didapat hasil sebesar 86% seperti terlihat pada gambar di bawah ini.

```
Administrator: C:\Windows\system32\cmd.exe

D:\SentimentAnalysis>"3. knn_dieng.py"

=====
APLIKASI SENTIMEN ULASAN WISATA DIENG DENGAN ALGORITMA K-NEAREST NEIGHBOR(
=====

Data training:
147
Counter({1: 75, 0: 72})

Data testing:
37
Counter({0: 20, 1: 17})

Akurasi: 0.8648648648648649

=====
TEST PREDIKSI ULASAN
Masukkan Ulasan yang akan diprediksi: dieng pemandangannya bagus sekali
Hasil Prediksi : Positive

=====
D:\SentimentAnalysis>
```

4. PENUTUP

4.1. Kesimpulan

Berdasarkan hasil penelitian yang sudah dilakukan, dapat disimpulkan bahwa Penelitian ini menghasilkan sebuah sistem analisis sentimen yang dapat mengidentifikasi dan mengklasifikasi ulasan pengunjung wisata dieng dengan menggunakan algoritma K-Nearest Neighbor (K-NN). Hasil dari pengujian, sistem mampu mengklasifikasikan sentimen menggunakan algoritma K-Nearest Neighbor (K-NN) dengan akurasi sebesar 86%.

5. DAFTAR PUSTAKA

1. Syaifudin, Y. W., & Irawan, R. A. (2018). Implementasi Analisis Clustering Dan Sentimen Data Twitter Pada Opini Wisata Pantai Menggunakan Metode K-Means. Jurnal Informatika Polinema, 4(3), 189-189.
2. Putri, P. (2013). Implementasi Metode Improved K-Nearest Neighbor pada Analisis Sentimen Twitter Berbahasa Indonesia (Doctoral dissertation, Universitas Brawijaya).
3. Wilianto, L., Pudjiantoro, T. H., & Umbara, F. R. (2017). Analisis Sentimen Terhadap Tempat Wisata Dari Komentar Pengunjung Dengan Menggunakan Metode Naive Bayes Classifier Studi Kasus Jawa Barat. Prosiding SNATIF, 439-448.
4. Azmi, M., Huda, A. K., & Setyanto, A. (2020). PEMANFAATAN DATA INSTAGRAM UNTUK MENGETAHUI

REPUTASI TEMPAT WISATA DI LOMBOK. TEKNIMEDIA: Teknologi Informasi dan Multimedia, 1(1), 39-46.

5. Deviyanto, A., & Wahyudi, M. D. R. (2018). Penerapan analisis sentimen pada pengguna twitter menggunakan metode K-Nearest Neighbor. JISKA (Jurnal Informatika Sunan Kalijaga), 3(1), 1-13.
6. Febrianti, Y. M., & Indriati, A. W. W. (2018). Analisis Sentimen Pada Ulasan “Lazada” Berbahasa Indonesia Menggunakan K-Nearest Neighbor (K-NN) Dengan Perbaikan Kata Menggunakan Jaro Winkler Distance. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN, 2548, 964X.
7. Sari, R. (2020). Analisis Sentimen Pada Review Objek Wisata Dunia Fantasi menggunakan Algoritma K-Nearest Neighbor (K-NN). EVOLUSI: Jurnal Sains dan Manajemen, 8(1).