



SENTIMENT ANALYSIS UNTUK REVIEW DESTINASI WISATA UNGGULAN GUNUNG KIDUL MENGGUNAKAN METODE LEXICON DAN PIVOT

Larasati¹, Shalsa Nabilla², Eri Haryanto³

^{1,2,3}Program Studi Informatika, Fakultas Teknik, Universitas Janabadra.

09larasati@gmail.com , shalsanbl12@gmail.com , eri@janabadra.ac.id

Jalan Tentara Rakyat Mataram 55 Yogyakarta 55231

Keywords:

*sentimen analisis,
Gunungkidul,
pivot, lexicon based,
Twitter*

Abstract

Social media in providing freedom of access to information has a rapid effect on tourist destinations in Indonesia. A tweet may also contain information or conditions about tourist destinations that they will or have visited, such as visitor experiences in traveling, visitors' opinions of tourist attractions, and other tourist attractions. The condition of tourism globally, especially in the Gunungkidul area after the Covid-19 pandemic, has experienced a significant decline. This condition motivated researchers to conduct research with the aim of contributing to post-Covid-19 tourism development, in particular by conducting a sentiment analysis of comments by tourists visiting tourist areas in Gunungkidul Regency. Sentiment analysis is the process of using text analytics to derive various data sources from the internet and various social media platforms. Comment data obtained through social media Twitter. This classification is used to find out how comments about tourist attractions in Gunungkidul are. Lexicon based and pivoting are methods used to classify public opinion into three classes, namely positive, negative and neutral sentiments. This method is used to classify the results of community comment data written on the form of community satisfaction with tourism in Gunungkidul. This study aims to find out how people's opinions regarding tourism in Gunungkidul during the Covid-19 pandemic took place. The results of this study are used to determine the classification of community commentary data so that services can be easily developed through comments given by the public. The results of the sentiment analysis show that there are 51% positive emotion categories, 1% negative and 48% neutral.

Kata Kunci:

*sentimen analisis,
Gunungkidul,
pivot, lexicon based,
Twitter*

Abstrak

Media sosial dalam memberikan kebebasan akses informasi berpengaruh pesat terhadap destinasi wisata di Indonesia. Sebuah tweet juga kemungkinan berisi informasi atau kondisi tentang destinasi wisata yang akan atau pernah mereka kunjungi, seperti pengalaman pengunjung dalam berwisata, pendapat pengunjung terhadap suatu tempat wisata, dan tempat wisata lainnya. Kondisi pariwisata secara global khususnya di daerah Gunungkidul pasca pandemi Covid-19 mengalami penurunan yang signifikan. Kondisi ini memicu motivasi peneliti untuk melakukan penelitian dengan tujuan berkontribusi pada pengembangan pariwisata pasca pandemi Covid-19 khususnya dengan melakukan sentimen analisis terhadap komentar wisatawan yang berkunjung ke daerah wisata di Kabupaten Gunungkidul. Sentimen analisis adalah proses penggunaan *text analytics* untuk mendapatkan berbagai sumber data dari internet dan beragam platform media sosial. Data komentar didapat melalui sosial media Twitter. Klasifikasi ini digunakan untuk mengetahui bagaimana komentar mengenai tempat wisata yang ada di Gunungkidul. *Lexicon based* dan *pivoting* merupakan metode yang digunakan untuk mengelompokkan opini

masyarakat ke dalam tiga kelas yaitu sentiment positif, negatif maupun netral. Metode ini digunakan untuk mengklasifikasikan hasil data komentar masyarakat yang tertulis pada bentuk kepuasan masyarakat terhadap wisata yang ada di Gunungkidul. Penelitian ini bertujuan untuk mengetahui bagaimana opini masyarakat mengenai wisata yang berada di Gunungkidul selama pandemi Covid-19 berlangsung. Hasil dari penelitian ini digunakan untuk menentukan klasifikasi data komentar masyarakat sehingga pelayanan dapat dengan mudah dikembangkan melalui komentar yang diberikan masyarakat. Hasil analisis sentimen terdapat kategori emosi positif 51%, negatif 1% dan netral 48%.

Pendahuluan

Media sosial dapat digunakan untuk melakukan komunikasi tanpa adanya interaksi langsung antara manusia satu dengan yang lainnya. Pada era modern seperti saat ini, sosialisasi antar individu dapat dilakukan dengan komunikasi tidak langsung yaitu melalui media sosial [1]. Berdasarkan studi yang telah dilakukan sebelumnya, penelitian dari Lu (2021) [2] memaparkan temuan dari Lisi (2013) [3]. Di era informasi digital saat ini, Twitter adalah *platform* yang tak tertandingi bagi orang-orang untuk berbagi pendapat dan pengalaman mereka secara publik. Ini karena, tidak seperti situs media sosial populer lainnya seperti Facebook, Twitter secara eksplisit terbuka untuk umum dan ditujukan untuk *multicasting* interaktif besar-besaran dan *real-time*.

Twitter merupakan jenis media sosial *microblogging* yang memfasilitasi pengguna untuk menulis dan mempublikasikan aktivitas serta pendapatnya [4].

Tujuan dari penelitian ini adalah untuk mencerminkan hasil penelitian big data dan *smart tourism*. Tujuan khusus dari penelitian ini adalah untuk mengilustrasikan manfaat Big Data bagi pengelola pariwisata dan untuk memahami dinamika dan kemungkinan pengelolaan pariwisata yang lebih baik [5].

Berdasarkan uraian diatas, maka dapat dirumuskan sebuah masalah yaitu bagaimana menerapkan sentimen analisis dalam melakukan komparasi sebelum dan sesudah adanya pandemi Covid-19 pada objek wisata di Gunungkidul.

Analisis Sentimen atau sering juga disebut penambangan opini adalah studi komputasi dari pendapat, penilaian, dan emosi orang melalui entitas, peristiwa, dan atribut yang dimiliki [6]. Berdasarkan studi yang telah dilakukan sebelumnya [6] memaparkan temuan terdahulu [7], Analisis sentimen telah menjadi topik yang populer untuk digunakan karena analisis sentimen dapat diterapkan di banyak sektor industri, termasuk industri pariwisata.

Sektor pariwisata di Indonesia sendiri selalu mengalami perbaikan setiap tahunnya dan sektor pariwisata ini kini menduduki peringkat ke-42 dunia [8].

Analisis sentimen dapat mengelompokkan pendapat pada kalimat untuk mengetahui sentimen positif, negatif dan netral. Untuk analisis sentimen itu sendiri kami menggunakan metode Lexicon Based dan Pivot.

Penggunaan metode Lexicon Based ditujukan mampu untuk menangkap informasi sentimen dari media sosial seperti twitter. Sedangkan penggunaan metode pivot merupakan alternatif untuk melakukan analisis terhadap data terkait, mengelompokkan dan juga meringkas data.

Big data adalah istilah dari zaman kontemporer yang mengacu pada kumpulan data yang memiliki volume besar, kecepatan tinggi, dan nilai tinggi. Ini berasal dari berbagai aspek kehidupan kita, termasuk media sosial, kesehatan masyarakat, pariwisata, bisnis, dan pertanian.

Untuk mengekstrak, menyimpan, menganalisis, dan menunjukkan data, perlu menggunakan beberapa teknik statistik mutakhir. Big data adalah volume data terbesar yang diterjemahkan ke dalam penciptaan teknologi baru yang

memfasilitasi pertumbuhan dan perkembangan informasi.

Data dan algoritma membentuk dunia baru yang terdiri dari bentuk kulminasi untuk komputasi dan, lebih tepatnya, cara baru untuk mengendalikan informasi. Dengan lebih dari 95% kumpulan data dunia telah dibuat dalam beberapa tahun terakhir, penting untuk mengetahui bahwa bukan orang yang memiliki algoritma terbaik menang, tetapi orang yang memiliki lebih banyak data; dan itu bukan sembarang jenis data, tetapi hanya data yang dapat diandalkan yang dihitung. Akibatnya, sejumlah besar data akan terakumulasi karena kami memiliki algoritma yang bekerja sangat efisien berdasarkan data yang kami proses [9].

Gunungkidul merupakan kabupaten yang sedang gencar melakukan pengembangan destinasi wisata. Kawasan yang strategis dan banyak destinasi wisata membuat wisatawan berbondong-bondong ingin mengunjungi.

Metode

Pada penelitian ini dilakukan pengumpulan komentar yang bersumber dari Twitter. Tujuannya adalah untuk mengetahui kepuasan wisatawan dengan menggunakan aplikasi pendukung yaitu Orange sebagai aplikasi pengambilan komentar wisatawan dan Rstudio untuk mengolah atau mengelompokkan sehingga mendapatkan hasil positif, negatif, dan netral. Adapun secara sistematis alur penelitian adalah sebagai berikut:

a. Pengambilan Data Penelitian (*Crawling Data*)

Proses pengambilan data, peneliti mengumpulkan komentar dari media sosial Twitter dengan memanfaatkan fasilitas Application Program Interface (API) yang sudah disediakan oleh Twitter dilakukan dengan cara *meng-crawling* data komentar Twitter menggunakan aplikasi orange. Total data yang didapatkan yakni sebanyak 278.117 data dengan kata kunci 6 top wisata di Gunungkidul seperti, #Gunungkidul, #wisata, #pantaiindrayanti, #pantaidrini, #pantaisepanjang, #hehasky, #nglanggeran, #bukitbintang. Data dikumpulkan dari bulan Februari 2022 sampai bulan Mei 2022. Setelah di olah menjadi 30.432 data.

berbagai wisata yang ada di Kabupaten Gunungkidul. Namun sejak pandemi Covid-19, kondisi pariwisata di Gunungkidul juga terdampak cukup signifikan seperti daerah lainnya [10].

Kondisi ini memicu motivasi peneliti untuk melakukan penelitian dengan tujuan berkontribusi pada pengembangan pariwisata pasca pandemi Covid-19 khususnya dengan melakukan sentimen analisis terhadap komentar wisatawan yang berkunjung ke daerah wisata di Kabupaten Gunungkidul.

Penelitian ini bertujuan untuk mengetahui opini masyarakat terhadap wisata di Kabupaten Gunungkidul pasca pandemi covid-19 dengan menarikan komentar di Twitter menggunakan metode *lexicon base* dan *pivot*. Secara lebih dalam penelitian ini ingin mengetahui apakah terdapat lebih banyak komentar positif, negatif atau netral terkait kebijakan pemerintah untuk penerapan pembatasan kunjungan wisata di Gunungkidul



Gambar 1. *Crawling Data*

b. Pengolahan dan Analisis Data (*Preprocessing Data*)

Preprocessing sendiri merupakan langkah yang menghilangkan noise yang tidak perlu dari data sehingga informasi yang berkualitas baik dapat diperoleh dan tersedia untuk diproses lebih lanjut. Hasil yang diharapkan dari proses pengolahan data adalah kumpulan data akhir berkualitas sehingga dapat digunakan dengan benar pada proses data mining selanjutnya. Pada tahap ini terdapat beberapa proses seperti *cleansing*, *case folding & remove character*, *remove duplicate*, *word normalization*, *stemming*, dan *stopword*. Selanjutnya melakukan tokenizing dan stemming. Setelah itu dilakukan normalisasi/perbaikan kata menggunakan metode *Lexicon Based* [11].

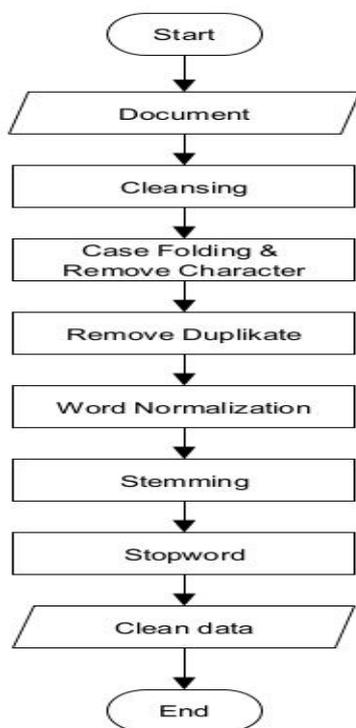
Berikut setiap tahapan preprocessing yang dilakukan pada penelitian ini, *Flowchart preprocessing data* dapat dilihat pada Gambar 2.

1. *Cleansing Data*

Proses *cleansing* dilakukan sesuai dengan tahapan-tahapan proses data mining sehingga dimungkinkan tidak semua tabel yang ada dimasing-masing basis data tersebut akan digunakan. Dan menghilangkan kata-kata yang tidak diperlukan seperti, emotikon, URL, *hashtag* (#), dan *mention* (@).

2. *Case Folding & Remove Character*

Case folding, yaitu memproses data dan mengubah seluruh data huruf kapital menjadi huruf kecil. *Remove character*, yaitu menghilangkan karakter atau simbol pada komentar seperti, angka dan tanda baca.



Gambar 2. *Preprocessing* Data

3. *Remove Duplicate*

Remove duplicate, yaitu menghapus komentar *tweet* yang menggunakan kalimat yang sama persis.

4. *Word Normalization*

Word normalization, yaitu suatu proses untuk menormalisasikan kata yang tidak baku menjadi kata baku sesuai kamus KBBI. Peneliti akan membuat kamus dalam bentuk excel untuk proses *word normalization*.

5. *Stemming*

Stemming merupakan proses untuk merubah seluruh kata ke dalam bentuk kata dasar.

6. *Stopword*

Stopword merupakan proses menghapus kata yang tidak diperlukan atau tidak memiliki sentimen seperti nama orang, tempat yang tidak bersangkutan. Objek penelitian yakni kata Gunungkidul, pantai, bersih, cantik, indah, kotor, mahal dan sebagainya. Peneliti akan membuat kamus dalam bentuk excel untuk proses *stopword*.

7. Data bersih

Data hasil akhir yang sudah terkategori positif, negatif dan netral.

8. *Pivoting*

Proses selanjutnya dilakukan *pivot* pada Excel yang nantinya akan diketahui hasil dari data komentar yang berupa data yang sudah diolah, menampilkan grafik dan menunjukkan kategori positif, negatif dan netral.

c. Pengumpulan data komentar Twitter wisatawan Gunungkidul

Data yang dikumpulkan yaitu data komentar Twitter yang diberikan oleh wisatawan tentang seputar obyek wisata Gunungkidul. Data ini didapatkan dengan cara mengunduh langsung dari data komentar di situs *website* Twitter Indonesia.

Berikut merupakan contoh beberapa data komentar wisatawan terhadap Wisata Gunungkidul.

Tabel 1. Contoh data komentar *tweet* wisatawan terhadap wisata Gunungkidul dalam bahasa Inggris

Content	Author
♥ saran pantai mana yang lebih bagus, Parangtritis atau Pantai di daerah Gunungkidul https://t.co/BHicRiW7ab	@tanyakanrl
@catatantravel Gunungkidul provinsi DIY kak 😊	@listianidl
Ini adalah salah satu kata bijak tentang pendidikan yang kalimatnya singkat namun isinya cukup mengena, dan blogger Gunungkidul paling suka dengan petikan bijak dari John Dewey ini.	@SDNMangunjaya1
Kapan kamu terlahir hiking disini?? 📍 Nglageran Gunungkidul, Jawa	@catatantravel

Tengah <https://t.co/1Z0eFgqSnm>

Karena aku selalu menikmati senja @catatantravel

dan deburan ombak yang membuat kita candu , udah pada sempet foto - foto disini belum?

👉Pantai indrayanti ,Gunungkidul Yogyakarta

<https://t.co/UIJpaOkSN0>

Kasus DBD di Kabupaten @jpnncm

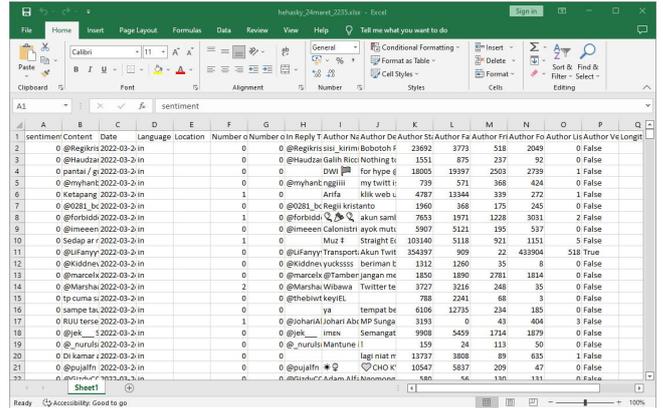
Gunungkidul sedang tinggi-tingginya. Pada awal tahun ini sudah dua orang meninggal dunia.

#DBD <https://t.co/qpLV38llyu>

Wisata Jeep Aventure di Goa @GTVID_News

Pindul, Gunungkidul, Yogyakarta menawarkan wisata susur hutan kayu putih menggunakan mobil offroad. CIL

#GTVNews #Gunungkidul #HutanKayuPutih #MobilOffroad #Yogyakarta #GoaPindul <https://t.co/8ojQbi1xDb>



Gambar 4. Contoh hasil crawling dari Orange yang disimpan Excel

b. Pre-Processing

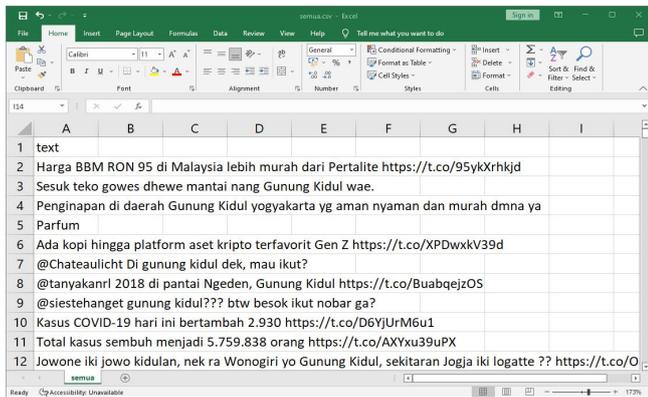
Pre-Processing dilakukan agar data menjadi lebih bersih dan siap untuk dianalisis. Sebelum melakukan proses pre-processing, file excel terlebih dahulu di ubah dalam bentuk CSV. Tahapan persiapan pengolahan data menggunakan Rstudio yaitu sebagai berikut :

1. Proses Cleansing

Disini dapat kita lihat bahwa data yang tersimpan dalam proses cleansing sudah sangat berbeda dengan data asli, sudah banyak data yang tidak digunakan dihilangkan oleh sistem RStudio.

Tabel 2. Hasil cleansing menggunakan RStudio dalam bentuk Excel dalam bahasa Inggris

1	" ??Heha Sky View"
2	" Terima kasih dukungan dari di event di Gunung Kidul "
3	" Secuil Surga di Gunung Kidul Yogyakarta "
4	" Air Terjun Sri Gethuk Gunung Kidul Jogjakarta"
5	" ini di pantai Drini nder "
6	" Warga Gunungkidul Tewas Terjatuh di Laut "



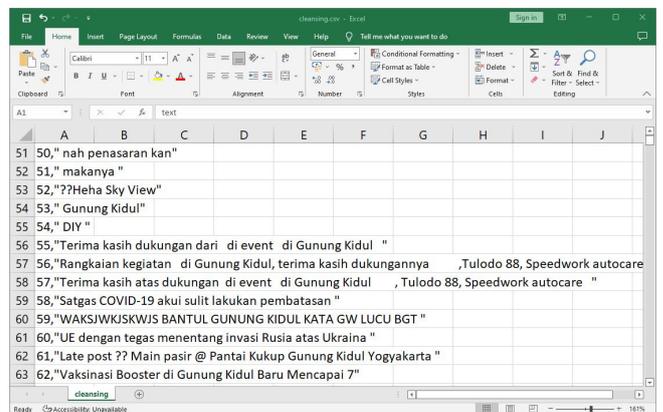
Gambar 3. Contoh Data Komentar Tweet Wisatawan Terhadap Wisata Gunung Kidul

Hasil dan Pembahasan

a. Proses Crawling Data (Crawling Data)

Tahapan pengumpulan data menggunakan Orange yaitu sebagai berikut :

1. Mencari API dengan mendaftarkan diri sebagai pengguna Twitter.
2. Mengajukan proposal permohonan untuk mendapatkan API yang akan di gunakan untuk pengambilan data dari Twitter.
3. Setelah API di dapatkan maka memasukan API pada aplikasi yang kita gunakan untuk crawling (disini kita menggunakan Orange).
4. Melakukan crawling data.
5. Hasil komentar didapatkan.



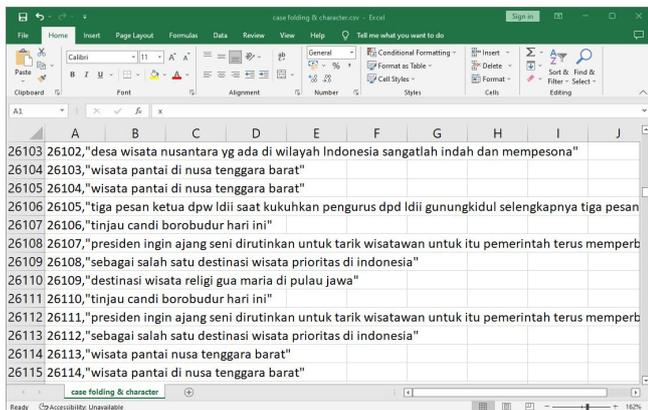
Gambar 5. Hasil cleansing menggunakan RStudio dalam bentuk Excel

1. Case Folding & Remove character

Disini dapat kita lihat pada proses *case folding & remove character* terdapat perubahan pada kata huruf besar berubah menjadi huruf kecil semua dan karakter dihilangkan.

Tabel 3. Hasil *case folding & remove character* menggunakan rstudio dalam bentuk Excel dalam bahasa Inggris

1	"harga bbm ron di malaysia lebih murah dari pertalite"
2	"gunung kidul"
3	"iseng jelajah google maps modelnya private beach gt pantai ngetun"
4	"sama ke gunung kidul banyak pantai"
5	"sundak gunung kidul relatif sepi linya ruame poll"
6	"pindahan dari ngopi in the sky gunung kidul jogja kyanya sayang udah rental setahun sama keluar modal"

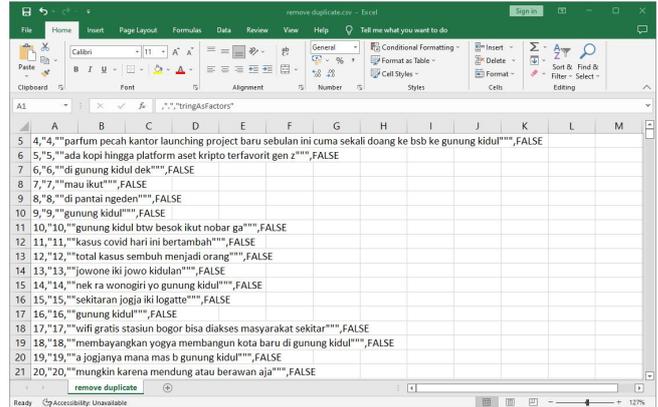


Gambar 6. Hasil *case folding & remove character* menggunakan RStudio dalam bentuk Excel

2. Remove Duplicate

Tabel 4. Hasil *remove duplicate* menggunakan RStudio dalam bentuk Excel dalam bahasa Inggris

1	"heha sky view",FALSE
2	"gunung kidul",FALSE
3	"iseng jelajah google maps modelnya private beach gt pantai ngetun",FALSE
4	"sama ke gunung kidul banyak pantai",FALSE
5	"sundak gunung kidul relatif sepi linya ruame poll",FALSE
6	"pindahan dari ngopi in the sky gunung kidul jogja kyanya sayang udah rental setahun sama keluar modal",FALSE



Gambar 7. Hasil *remove duplicate* menggunakan RStudio dalam bentuk Excel

3. Word Normalization

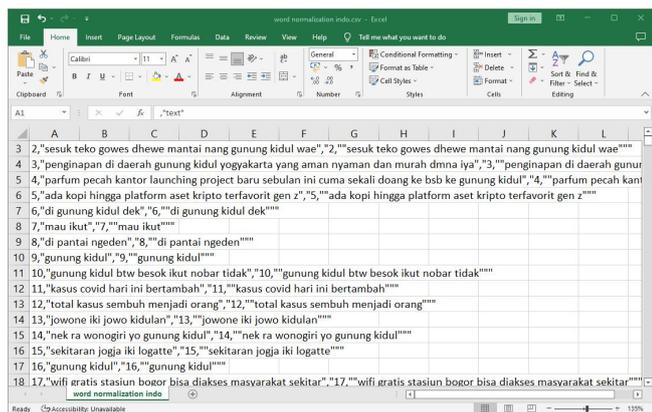
Sebelum melakukan proses *word normalization* terlebih dahulu membuat kamus. Kamus dalam penelitian ini dibuat dengan melihat kata-kata yang perlu diubah, seperti pada gambar 6. Hasil proses *word normalization* dapat dilihat ada Gambar 8.

Tabel 5. Kamus kata dalam bahasa Inggris

Ejaan lama	Kata Baku	Case
Uenak	enak	TRUE
Bgs	Bagus	TRUE
Lu	kamu	TRUE
Kab	kabupaten	TRUE
Gt	gitu	TRUE

Tabel 6. Hasil *word normalization* menggunakan RStudio dalam bentuk Excel dalam bahasa Inggris

1	pengen naik keatas ke hehaskey
2	sundak gunung kidul relatif sepi linya rame sekali
3	iseng jelajah google maps modelnya private beach gitu pantai ngetun
4	kamu kalo main ke Gunungkidul bilang-bilang lah



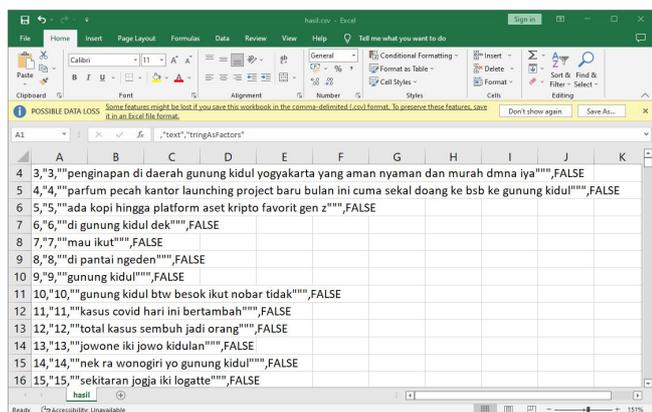
Gambar 8. Hasil *word normalization* menggunakan RStudio dalam bentuk Excel

4. Tahap *Stemming*

Setelah proses *stemming* data akan dibersihkan lagi melalui *word normalization*. Hasil dari proses *stemming* dapat dilihat pada Gambar 9.

Tabel 7. Hasil *Stemming* menggunakan RStudio dalam bentuk Excel dalam bahasa Inggris

1	""pengen naik keatas ke hehasky """"""""",FALSE
2	""sundak gunung kidul relatif sepi linya rame sekali""""""""",FALSE
3	""iseng jelajah google maps modelnya private beach gitu pantai ngetun""""""""",FALSE
4	""kamu kalo main ke Gunungkidul bilang-bilang lah""""""""",FALSE



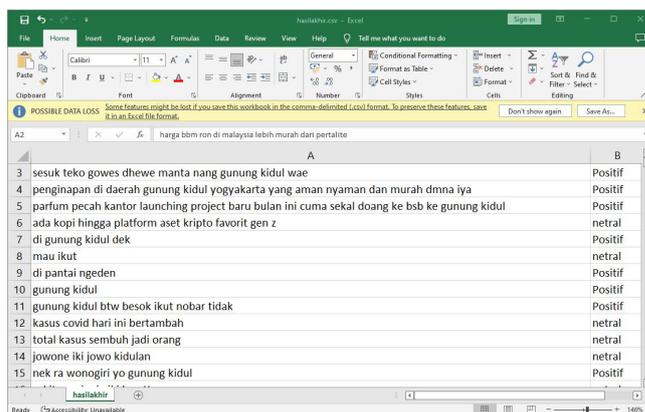
Gambar 9. Hasil *stemming* menggunakan RStudio dalam bentuk Excel

5. Hasil akhir (data bersih)

Hasil akhir ini menampilkan data bersih dengan menyajikan kesimpulan positif, negatif dan netral yang sudah di proses menggunakan RStudio.

Tabel 8. Hasil akhir data bersih menggunakan RStudio dalam Bentuk Excel dalam bahasa Inggris

1	""pengen naik keatas ke hehasky """"""""",FALSE
2	""sundak gunung kidul relatif sepi linya rame sekali""""""""",FALSE
3	""iseng jelajah google maps modelnya private beach gitu pantai ngetun""""""""",FALSE
4	""kamu kalo main ke Gunungkidul bilang-bilang lah""""""""",FALSE



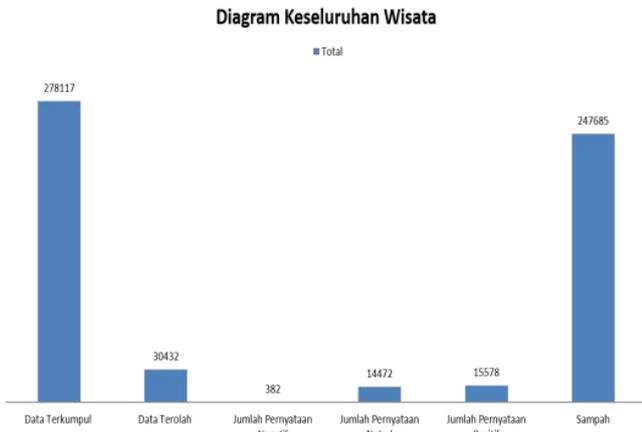
Gambar 10. Hasil akhir data bersih menggunakan RStudio dalam bentuk Excel

6. *Pivoting*

Pivot tabel bertujuan untuk merangkum, mengeksplorasi, mempresentasikan, menghitung dan menganalisa data.

Tabel 9. Keterangan gambar diagram hasil pivot komentar *tweet* wisatawan Gunungkidul dalam bahasa Inggris

278.117	Data terkumpul
30.423	Data Terolah
15.578	Jumlah pernyataan Positif
382	Jumlah pernyataan Negatif
14.472	Jumlah pernyataan Netral
247.685	Data Sampah



Gambar 11. Gambar diagram hasil pivot komentar *tweet* wisatawan Gunungkidul



Gambar 13. Gambar hasil *WordCloud* menggunakan Orange

7. Menampilkan kata yang sering muncul

Tabel 10. Hasil kata yang sering muncul menggunakan Orange dalam bahasa Inggris

Weight	Word
10046	pantai
5463	wisata
3617	Gunungkidul
1483	desa
1446	Jalan



Gambar 12. Gambar hasil kata yang sering muncul menggunakan Orange

8. Menampilkan *WordCloud*

WordCloud merupakan gambar yang terdiri dari kumpulan kata, dimana besarnya kata mempresentasikan kemunculan kata. Semakin besar suatu kata, maka semakin sering kata tersebut muncul dalam dokumen. *WordCloud* dapat dilihat pada *Orange* yaitu pantai, wisata, Gunungkidul, desa, jalan.

9. Menampilkan hasil klasifikasi positif, negatif dan netral dari Rstudio.

Tabel 11. Hasil klasifikasi positif, negatif dan netral dari Rstudio dalam bahasa Inggris

	Text	Klasifikasi
1	Pantai Gunungkidul	Positif
2	iseng jelajah private beach gitu pantai ngetun	Netral
3	sundak gunung kidul relatif sepi linya rame sekali	Negatif

6	1	6,"di gunung kidul dek"	Positif
7	0	7,"mau ikut"	netral
8	1	8,"di pantai ngeden"	Positif
9	1	9,"gunung kidul"	Positif
10	1	10,"gunung kidul btw besok ikut nobar tidak"	Positif
11	0	11,"kasus covid hari ini bertambah"	netral
12	0	12,"total kasus sembuh jadi orang"	netral
13	0	13,"jowone iki jowo kidulan"	netral
14	1	14,"neka ra wonogiri yo gunung kidul"	Positif
15	0	15,"sekitaran Jogja iki logatte"	netral
16	1	16,"gunung kidul"	Positif
17	0	17,"wifi gratis stasiun bogor bisa akses masyarakat sekitar"	netral
18	1	18,"membayangkan yoga bangun kota baru di gunung kid..."	Positif
19	1	19,"a jogjanya mana mas b gunung kidul"	Positif
20	0	20,"mungkin karena mendung atau awan aja"	netral
21	3	21,"aku aja tidak selalu ke pantai bisa liat sunsetsunrest kalo..."	Positif
22	1	22,"jos persig gunung kidul otw liga"	Positif
23	1	23,"gunung kidul"	Positif
24	0	24,"mas"	netral

Gambar 14. Gambar hasil klasifikasi positif, negatif dan netral dari Rstudio.

10. Menampilkan contoh kalimat positif, negatif dan netral.

Table 12. Contoh kalimat positif dari Rstudio dalam bahasa Inggris

No	Content	Analyst	Kategori
----	---------	---------	----------

1	ih seru banget ke bukit bintang.	Positif	Pujian
2	minat pantai Gunungkidul aja	Positif	Ketertarikan
3	piknik ke pantai Gunungkidul	Positif	Kejadian
4	Seriusan, Gunungkidul memang kawasa yang paket wisatanya lengkap banget. Ada wisata alam, wisata sejarah, sampai wisata religi. Mantap.	Positif	Wisata Lengkap
5	Pantai Sepanjang, Daya Tarik Bagi Pecinta Alam Terbuka Gunungkidul	Positif	Pujian Pesona

Tabel 13. Contoh kalimat negatif dari Rstudio dalam bahasa Inggris

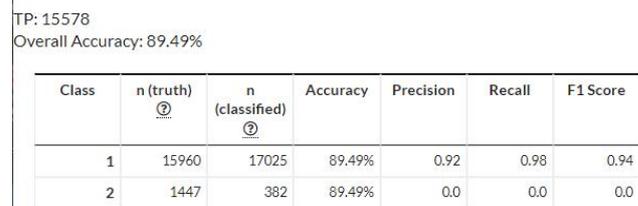
No	Content	Analyst	Kategori
1	Wisatawan yang Hilang di Pantai Siung Gunungkidul Belum Ditemukan	Negatif	Tragedi
2	Korban Kecelakaan Laut Di Pantai Siung Belum Ditemukan, Petugas Gabungan Perluas Pencarian Hingga Radius 4 Mil.	Negatif	Kecelakaan
3	Pantai Gunubg Kidul masih ada sampah plastic Bertebaran	Negatif	Keluhan
4	Asyik Berfoto Seorang Wisatawan Terseret Ombak, Hingga Kini Masih Dalam Pencarian	Negatif	Kecelakaan
5	Korban Jatuh Dari Tebing Watu Tumpeng, Ditemukan Meninggal Dunia.	Negatif	Kecelakaan

Tabel 14. Contoh kalimat netral dari Rstudio dalam bahasa Inggris

No	Content	Analyst	Kategori
1	Ngunggah beach  Giriwungu , Panggang , Gunungkidul .	Netral	Penyebutan
2	Piknik ke Gunungkidul boleh, abai prokes jangan. Sebelum berangkat piknik perhatikan hal - hal berikut ya sobat wisata, kami tunggu kunjungannya stay safe and stay healthy	Netral	Iklan
3	 saran pantai mana yang lebih bagus, Parangtritis atau Pantai	Netral	Membandingkan

4	di daerah Gunungkidul @m00ncherry Kopi klotok jakal, heha sky view, pantai Gunungkidul	Netral	Menyebut
5	Tempat piknik Gunungkidul	Netral	Penyebutan

11. Menguji akurasi metode *Lexicon Based* Setelah mendapatkan hasil pivoting, Langkah selanjutnya adalah pengukuran performasi dengan menggunakan Confusion matrix.



Gambar 15. Gambar hasil pengukuran performasi dengan dengan confusion matrix

Kesimpulan dan Saran

Kesimpulan dari penelitian analisis sentimen yang telah dilakukan adalah dapat memberikan informasi mengenai pendapat masyarakat tentang bagaimana wisata yang ada di Kabupaten Gunungkidul. Masyarakat rata-rata memberikan komentar positif, hal ini dapat dibuktikan dengan melihat prosentase komentar positif 51%, komentar negatif 1% dan komentar netral 48%. Tidak hanya itu saja komentar masyarakat juga dilakukan pengukuran menggunakan confusion matrix dan menghasilkan nilai akurasi sebesar 89,49%. Penelitian ini memberikan kontribusi kepada pengelola wisata yang ada di Kabupaten Gunungkidul untuk melakukan pemanfaatan sosial media khususnya dalam hal promosi dikarenakan pengaruh sosial media dalam perkembangan tempat wisata sangat besar. Dengan memaksimalkan sosial media sebagai media promosi maka jumlah wisatawan yang tertarik untuk mengunjungi tempat wisata juga akan meningkat. Dengan demikian ekonomi warga masyarakat juga akan meningkat.

Referensi

[1] Sabna E, Mustika B, Fonda H, Irfan D, Ambiyar. Text Mining Uses K-Means Clustering Algorithm To Predict. Inf

- Technol Comput Sci. 2020;3(2):380-6.
- [2] Lu Y, Zheng Q. Twitter public sentiment dynamics on cruise tourism during the COVID-19 pandemic. *Curr Issues Tour.* 2021;24(7):892-8.
- [3] Lisi J. Twitter: social communication in the Twitter age. *Choice Rev Online.* 2013;50(10):50-5424-50-5424.
- [4] Husnusyifa A. Twitter merupakan jenis media sosial microblogging yang memfasilitasi pengguna untuk menulis pendapatnya . Secara historis , kehadiran dan munculnya media sosial Twitter yang menyediakan ruang tertentu atau media sosial lainnya , di Twitter pengguna lain. *IDEA J Hum.* 2019;2(2):120-33.
- [5] Ardito L, Cerchione R, Del Vecchio P, Raguseo E. Big data in smart tourism: challenges, issues and opportunities. *Curr Issues Tour.* 2019;22(15):1805-9.
- [6] Steven C, Wella W. The Right Sentiment Analysis Method of Indonesian Tourism in Social Media Twitter. *IJNMT (International J New Media Technol.* 2020;7(2):102-10.
- [7] Ciric M, Stanimirovic A, Petrovic N, Stoimenov L. Comparison of different algorithms for sentiment classification. 2013 11th Int Conf Telecommun Mod Satell Cable Broadcast Serv TELSIS 2013. 2013;2:567-70.
- [8] Hermanto DT, Ziaurrahman M, Bianto MA, Setyanto A. Twitter Social Media Sentiment Analysis in Tourist Destinations Using Algorithms Naive Bayes Classifier. *J Phys Conf Ser.* 2018;1140(1).
- [9] Sedkaoui S, Khelfaoui M, Kadi N. Big data analytics: harnessing data for new business models. In: Sedkaoui S, Khelfaoui M, Kadi N, editors. *Apple Academic Press;* 2022. p. 3-13.
- [10] Dinas Pariwisata Gunungkidul.
- [11] Sentimen A, Menggunakan VC-, Naive A. and Levenshtein Distance Word Correction. *Inform dan Teknol Inf.* 2022;19(1):91-104.