

Analisis Emosi Wisatawan Menggunakan Metode *Lexicon Text Analysis*

<http://dx.doi.org/10.28932/jutisi.v10i1.6690>

Riwayat Artikel

Received: 12 Juni 2023 | Final Revision: 08 Januari 2024 | Accepted: 17 Januari 2024

Creative Commons License 4.0 (CC BY – NC)



Dea Caesy Rahmadani^{#1}, Siti Khomsah^{*2}, M Yoka Fathoni^{✉#3}

[#] Program Studi Sistem Informasi, Institut Teknologi Telkom Purwokerto
Jl. DI Panjaitan No.128 Karangreja, Kota, 53147, Indonesia

¹18103039@ittelkom-pwt.ac.id

³myokafathoni@ittelkom-pwt.ac.id

^{*}Program Studi Sains Data, Institut Teknologi Telkom Purwokerto

Jl. DI Panjaitan No.128 Karangreja, Kota, 53147, Indonesia

²siti@ittelkom-pwt.ac.id

[✉]Corresponding author: myokafathoni@ittelkom-pwt.ac.id

Abstrak — Komentar wisatawan terhadap objek wisata merupakan ungkapan pengalaman, pendapat, bahkan keluhan mereka. Banyaknya data komentar wisatawan di internet mengandung informasi berguna bagi stakeholder. Informasi tersebut dapat diekstrak dengan cara analisis sentiment/emosi. Ekspresi sentimen mendasar manusia yaitu positif dan negatif. Dalam ilmu psikologi, emosi manusia dapat dibagi menjadi 8: anger(marah), fear(takut), anticipation(antisipasi), disgust(jijik), sadness(sedih), surprise(terkejut), trust(yakin), dan joy(bahagia). Penelitian ini bertujuan mengetahui emosi wisatawan berdasarkan komentar-komentar mereka di akun Google Map sepuluh objek di Kabupaten Banjarnegara, Purbalingga, Banyumas, Cilacap, dan Kebumen. Metode yang digunakan yaitu *Lexicon Emotion Analysis*, analisis emosi berdasarkan makna leksikal setiap kata dalam komentar. Setiap komentar akan diurai menjadi token(kata tunggal), kemudian diberikan label emosi menggunakan kamus EmoLex. Emosi yang jumlahnya paling tinggi akan menjadi label komentar tersebut. Hasilnya, mayoritas wisatawan dari sepuluh objek wisata tersebut menunjukkan emosi “joy (bahagia)” dan “trust (yakin)”. Situasi objek wisata yang memicu komentar dengan mayoritas nada emosi “yakin” antara lain: Aquarium Purbasari Pancuran Mas dengan 33,52%, Lembah Asri Serang dengan 30,85%, Taman Reptil Sanggaluri sebanyak 30, 27%, Kebun Raya Baturaden dengan 27, 67%, dan Curug Jenggala sebesar 23,4%. Komentar dengan nada emosi “bahagia” yang tertinggi yaitu Benteng Pandem dengan 27, 41%, Candi Arjuna dengan 26,6%, Kawah Sikidang sebesar 20,71%, dan Pantai Menganti sebanyak 25, 74%. Hanya satu objek wisata yang ditanggapi dengan emosi “antisipatif” tertinggi yaitu Taman Miniatur Dunia. Emosi antisipatif diwakili oleh kata-kata tanggapan atau perhatian sehingga kata-kata tersebut bisa dikategorikan ke dalam emosi positif. Ekstraksi emosi berbasis makna leksikal perlu mempertimbangkan pembersihan data sehingga terbebas dari kata-kata yang tidak bermakna terkait emosi.

Kata kunci—Analisis Emosi; Google Map; leksikal; Objek Wisata.

Emotions Analysis of Tourist Using Lexicon Text Analysis

Abstract — Reviews from tourists are expression of their experiences, opinions, and even complaints. The large amount of tourist comments data contain valuable information for stakeholders. We can extract information and knowledge from user comments using sentiment/emotion analysis. The basic emotions/sentiments of human commonly are positive and negative. In psychology's domain, human emotions/sentiments divided into 8: anger, fear, anticipation, disgust, sadness, surprise, trust, and joy. This research aims to

determine tourists' emotions based on their comments on the GoogleMap account of ten objects in district Banjarnegara, Purbalingga, Banyumas, Cilacap, dan Kebumen. This research uses the Lexicon Emotion Analysis based on the lexical meaning of each word contained in the comment. Each comment tokenized into single words, then labeled using the EmoLex. The emotion with the highest weight will be the label. This study showed that majority emotions are "joy" and "trust". The circumstances of tourist sites that triggered comments with the majority emotional tone being "trust" include Purbasari Pancuran Mas Aquarium with 33.52%, Lembah Asri Serang with 30.85%, Sanggaluri Reptile Park with 30, 27%, Baturaden Botanical Gardens with 27, 67%, and Curug Jenggala at 23.4%. Comments with the highest "happy" tone were Pandem Fort at 27.41%, Arjuna Temple at 26.6%, Sikidang Crater at 20.71%, and Menganti Beach at 25.74%. Only one tourist site (the World Miniature Park) responded with the highest "anticipatory". Anticipatory is form of response or attention, so that considered positive emotion. Extraction of emotions based on lexical meaning needs to consider cleaning the data so that it is free from meaningless words related to emotions.

Keywords—Emotion Analysis; Google Map Review; Lexicon; Tourism Sites.

I. PENDAHULUAN

BARLINGMASCAKEB adalah akronim dari 5 nama Kabupaten yang terlibat yaitu Kabupaten Banjarnegara, Kabupaten Purbalingga, Kabupaten Banyumas, Kabupaten Cilacap, dan Kabupaten Kebumen [1]. Sektor pariwisata merupakan salah satu sumber pendapatan bagi lima Kabupaten tersebut. Menurut data pariwisata pada tahun 2018, wisatawan yang datang di wilayah ini terus meningkat, baik pada objek wisata yang dikelola pemerintah, swasta, maupun wisata yang dikelola perorangan [2]. Beberapa objek wisata di BARLINGMASCAKEB antara lain Aquarium Purbasari Pancuran Mas, Benteng Pendem, Candi Arjuna, Curug Jenggala, Kawah Sikidang, Kebun Raya Baturraden, Lembah Asri Serang, dan Pantai Menganti. Sepuluh objek wisata tersebut termasuk paling sering dikunjungi oleh wisatawan. Netizen banyak yang memberikan ulasan melalui akun Google Map pada setiap objek wisata tersebut. Ada komentar atau review yang bagus dan memberikan informasi sesuai pengalaman mereka saat di objek tersebut, namun ada pula yang menyatakan keluhan.

Google Map merupakan salah satu aplikasi yang digunakan untuk mencari suatu lokasi objek tertentu seperti tempat wisata. Google Map Review adalah salah satu fitur di Google Map yang dapat digunakan pengguna Internet untuk memberikan ulasan atau komentar terhadap objek di Google Map. Pengguna dapat menuliskan komentar dan memberikan rating pada objek di Google Map. Dalam bidang pariwisata, komentar atau review wisatawan di Google Map menjadi penting karena informasi tersebut dapat memberikan informasi kepuasan pengunjung [3]. Biasanya komentar wisatawan tersebut ditemukan dalam bentuk data teks dalam jumlah banyak. Untuk mengekstrak informasi dari komentar-komentar wisatawan diperlukan metode pengolahan teks seperti analisis sentimen. Ahli psikologi- Plutchik dalam [4], sentimen dasar setiap individu adalah positif dan negatif, sedangkan sentiment dasar tersebut dapat diperluas menjadi berbagai spektrum emosi. Emosi- emosi tersebut menunjukkan pengalaman wisatawan [5]

Penelitian-penelitian analisis sentiment [3], [5], [6], [7], [8], [9], [10] menunjukkan bahwa pengalaman wisatawan saat atau setelah mereka mengunjungi sebuah objek wisata dapat dideteksi dari komentar mereka di media sosial. Penelitian [8] menggali sentimen wisatawan Kalimantan Barat dari kumpulan review di Google Map, hasil penelitian berupa aplikasi website analisis sentimen dalam kelas sentimen positif dan negative, sedangkan data tidak mengandung sentimen diklasifikasikan menjadi kelas netral. Penelitian [10] menggunakan data tweet untuk membangun model machine learning yang dapat mengungkapkan pengalaman pelanggan dari komentarnya. Pengalaman pelanggan diklasifikasi ke dalam kelas sentimen positif dan negatif, sedangkan komentar tidak mengandung sentimen diklasifikasi menjadi kelas netral. Umumnya, penelitian analisis sentimen hanya membagi sentimen menjadi 2 atau tiga kelas yaitu positif, negatif, dan netral [5], [6], [7], [8], [9], [10], [11]. Menurut penelitian [10], sentiment pengunjung tidak terbatas pada kelas positif dan negatif, atau tanpa mengandung sentiment (netral). Sentimen dapat dikatakan sebagai perasaan. Dalam kamus bahasa Indonesia, sentimen bisa diartikan sebagai emosi yang berlebihan, bisa juga dimaknai pendapat atau pandangan. Namun dalam kajian analisis sentimen, kata sentimen tidak hanya bermakna negatif atau positif. Saif M. Mohammad [12] menjelaskan bahwa sentimen dapat berupa emosi dan dapat diukur berdasarkan sumber tekstual. Sumber tekstual dapat diperoleh dari media sosial seperti twitter. Saif M. Mohammad [12] mengukur emosi berdasarkan delapan emosi dasar dari Plutchik.

Selaras dengan cara pengukuran emosi oleh Saif M. Mohammad maka beberapa penelitian analisis emosi mulai dikembangkan [5], [13], [14]. Penelitian sebelumnya menguraikan 2 kelas sentimen dasar (positif dan negatif) dalam 8 jenis emosi yaitu "Joy", "trust", "Surprise", "anticipation", "Fear", "Anger", "Sadness", "Disgust". Penelitian analisis emosi pada berbagai domain antara lain oleh Amrita Mathur pada bidang Kesehatan [14], menganalisis emosi masyarakat India dengan selama situasi pandemi COVID-19. Data diambil dari Twitter, sedangkan pelabelan emosi menggunakan kamus NRC EmoLex menghasilkan delapan emosi. Tujuan analisis emosi tersebut untuk memahami kesehatan mental masyarakat, kemudian hasil analisis digunakan dalam pengambilan tindakan untuk memutuskan kebijakan saat memerangi virus COVID-19 di India [14]. Analisis teks merupakan penelitian yang unik karena model analisis yang dibangun biasanya hanya berlaku pada bahasa yang digunakan sebagai dataset. Analisis emosi dalam bahasa Indonesia antara lain oleh Aribowo [13], menghasilkan informasi tentang 7 emosi masyarakat Indonesia pada masa pemilu 2019. Emosi diekstrak dari komentar

netizen pada channel YouTube tentang debat calon presiden 2019. Pada domain pariwisata, penelitian [5] dapat mengungkapkan informasi preferensi wisatawan. Kesimpulannya, mayoritas wisatawan mengalami perasaan atau emosi atau pendapat mayoritas pada sentimen positif dan emosi “Joy (senang)” dan “Trust (percaya/yakin)”. Namun, pada penelitian [5] masih terbatas pada jumlah dataset, pra-pemrosesan teks yang digunakan dan sampel objek wisata yaitu hanya 5 objek wisata di Banyumas. Sedangkan pelabelan emosi masih menggunakan kamus lexicon (EmoLex) yang dikembangkan oleh Saif M. Mohammad [13].

Perbedaan-perbedaan pada penelitian-penelitian sebelumnya antara lain sumber data (Twitter dan Youtube. Metode yang digunakan adalah metode leksikon yang akan mendeteksi emosi dari review wisatawan. Pelabelan menggunakan EmoLex yang dimodifikasi. Penelitian [12] khusus ekstraksi emosi pada korpus bahasa Inggris. Kelemahan EmoLex adalah korpus emosi dalam bahasa Indonesia yang merupakan hasil translate dari Google sehingga banyak kata yang tidak sesuai maknanya. Maka, pada penelitian ini kamus emosi EmoLex diperbaiki sesuai kamus bahasa Indonesia. Pada tahap pre-processing, dikembangkan teknik penanganan kata-kata slang. Kata slang merupakan *noise* karena kata tersebut tidak dikenal dalam kamus bahasa Indonesia maupun dalam library *text-analytics* seperti sastrawi [5].

Berdasarkan penelitian-penelitian sebelumnya dan banyaknya komentar netizen pada akun Google Map objek-objek wisata di sekitar Banyumas, maka penelitian ini bertujuan mengetahui emosi pengunjung wisata di wilayah BARLINGMASCAKEB menggunakan metode *lexicon text analysis*. Dataset yang digunakan berasal dari komentar atau review netizen terkait objek wisata yang ada di BARLINGMASCAKEB pada laman Google Map Review. Ekstraksi emosi menggunakan kamus EmoLex yang telah diperbaiki tranlasinya sesuai kamus baku bahasa Indonesia. Pada tahap preprocessing ditambahkan tahap penanganan slang word.

II. METODE PENELITIAN

Metodologi penelitian ini terdiri dari tahapan pengumpulan data, pre-processing, ekstraksi emosi, dan analisis Hasil, ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

A. Pengumpulan Data

Langkah awal pada penelitian ini adalah pengumpulan data dengan melakukan *web crawling* komentar wisatawan pada laman Google Map Review, menggunakan aplikasi WebHarvy. Objek wisata yang digunakan sebanyak 10 objek wisata yang terletak di kawasan BARLINGMASCAKEB antara lain yaitu Taman Miniatur Dunia, Curug Jenggala, Pantai Menganti, Kebun Raya Baturraden, Kawah Sikidang, Candi Arjuna, Lembah Asri Serang, Taman Reptil Sanggaluri, Aquarium Purbasari Pancuran Mas, dan Benteng Pendem. Data komentar yang unduh adalah komentar yang ada di Google Map Review dalam kurun waktu 5 tahun terakhir (2017-2022). Tahapan *crawling* berhasil mengumpulkan data sebanyak 11.376 komentar, dan disimpan dalam format Ms Excel. Variabel data yang terkumpul terdiri nama user, waktu, rating, dan review. Namun untuk analisis emosi pada penelitian ini hanya menggunakan data review, hal ini disebabkan oleh nilai rating yang diberikan kadang berlawanan dengan kata-kata yang ditulis dalam review. Misalnya, netizen memberikan rating 4 (dalam skala 1-5) namun dalam komentarnya masih menulis kata-kata negatif seperti “kotor”, “menyesal”, dan sebagainya. Hasil seleksi data menghasilkan jumlah dataset setiap objek berbeda-beda, seperti Tabel 1.

TABEL 1.
DATASET PER OBJEK WISATA

No	Objek Wisata	Jumlah Komentar
1	Aquarium Purbasari Pancuran Mas	1518
2	Benteng Pendem	1930
3	Candi Arjuna	1452
4	Curug Jenggala	699
5	Kawah Sikidang	1409
6	Kebun Raya Baturraden	1063
7	Lembah Asri Serang	910
8	Pantai Menganti	437

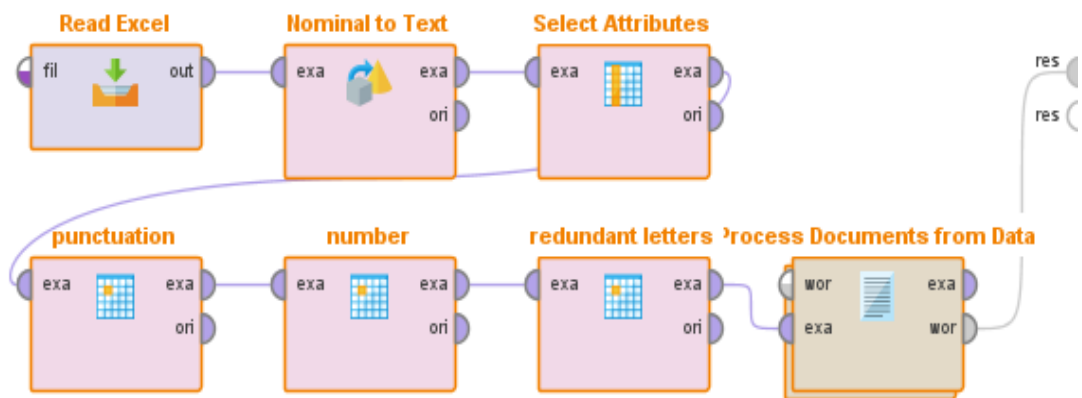
No	Objek Wisata	Jumlah Komentar
9	Taman Miniatur Dunia	732
10	Taman Reptil Sanggaluri	1226
Total Data		11.376

B. Data Preprocessing

Preprocessing bertujuan untuk menyiapkan data sebelum pemrosesan lebih jauh dilakukan Secara umum, data preprocessing ini dilakukan untuk menghilangkan karakter-karakter serta kata atau simbol-simbol yang tidak diperlukan dalam proses analisis. *Preprocessing* menggunakan aplikasi Rapidminer dan Python. Tahapan *preprocessing* mengadopsi langkah-langkah yang dilakukan pada pada penelitian Baskoro, Khomsah, dan Susanto [15] terdiri *data cleaning*, *stopword removing*, dan konversi *slangword*.

1. Data Cleaning

Data cleaning bertujuan menghapus seluruh karakter non-alphabet seperti angka dan tanda baca. Karakter, simbol seperti tanda baca tidak memiliki makna atau emosi. Proses ini dilakukan pada aplikasi RapidMiner dengan desain model dapat dilihat pada Gambar 2.



Gambar 2. Desain Model Data Cleaning

Pada Gambar 2, langkah pertama proses adalah mengubah data ke dalam bentuk text (*Nominal to Text*). Hal ini karena meskipun dataset yang diolah berupa tulisan, namun Rapidminer masih membacanya sebagai data bertipe nominal. Kemudian, seleksi atribut yaitu kolom Review. Selanjutnya proses menghilangkan tanda baca (*punctuation*), angka (*number*), dan karakter berulang (*redundant letters*). Selanjutnya proses *case folding* untuk mengubah seluruh teks menjadi huruf kecil atau *lowercase*.

2. Stopword Removing

Stopword removing bertujuan menghapus kata-kata yang tidak bermakna atau memiliki sentimen atau emosi. Kata yang memiliki sentimen atau emosi adalah kata sifat atau kata kerja. Proses tahap ini menggunakan bahasa pemrograman Python dan library sastrawi. Kata-kata yang termasuk *stopword* sudah ada dalam daftar *stopword* library sastrawi. Penambahan *stopword list* sastrawi perlu ditambahkan secara manual karena banyak kata-kata dalam dataset yang tidak dikenali oleh *stopword list* sastrawi. Proses *stopword removing* menggunakan kode program berikut.

```
#stopword removing
from Sastrawi.StopWordRemover.StopWordRemoverFactory import StopWordRemoverFactory
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
hasil_stopword_remove = []
factory = StopWordRemoverFactory()
#daftar tambahan stopwords
more_stopword = ['pt', 'gw', 'obyek', 'objek', 'banget', 'gaes', 'banyumas', 'purbalingga', 'blur', 'air', 'bgt', 'th', 'lh', 'rp', 'rb', 'yey', 'ribu', 'pancur', 'bs', 'jos', 'hahaha', 'haha', 'hehee', 'joss', 'josss', 'waterfall', 'view', 'nice', 'place', 'dll', 'banjarnegara', 'cilacap', 'kebumen', 'baturredn', 'candi', 'aquarium', 'akuarium', 'benteng', 'taman', 'bgd', 'beautiful', 'wow', 'andhang', 'pendem', 'pas', 'purbasari', 'pancuran', 'mas', 'arjuna', 'kawah', 'sikidang', 'kebun', 'raya', 'aturaden', 'what', 'lembah', 'serang', 'pantai', 'menganti', 'reptil', 'sanggaluri', 'nasi', 'purbayasa', 'ga', 'hehe', 'hehehe', 'wkwk', 'wkwkwk', 'sxn', 'gpp', 'mw', 'ehehe', 'wuih', 'curug', 'jengjala', 'purwokerto', 'the', 'is', 'josgandos', 'skrg', 'exit', 'sx', 'fyi', 'kulon']
stopwords = factory.get_stop_words() + more_stopword
```

```
temp = [t for t in re.findall(r'[a-z]+-[a-z]+\b',stopword_remove) if t not in stopwords]
stopword_remove = ''.join(temp)
dataclean.append(stopword_remove)
```

3. Konversi Slangwords

Proses ini dilakukan untuk mengubah kata-kata tidak baku atau singkatan menjadi kata yang sesuai, misalnya kata “gw” akan dikonversikan menjadi “saya”, kata “kzl” menjadi “kesal” dan lain sebagainya. Konversi *slangwords* dilakukan menggunakan Bahasa Python dengan algoritma berikut.

#Daftar slang dan konversinya

```
slangs={'yg':'yang', 'tdk':'tidak', 'gak':'tidak', 'pd':'pada', 'mlh':'malah', 'jgn':'jangan', 'jg':'juga', 'tp':'tapi', 'blkg':'belakang', 'dr':'dari',
'klo':'kalau', 'kl':'kalau', 'lg':'lagi', 'utk':'untuk', 'msh':'masih', 'jd':'menjadi', 'hy':'hanya', 'byk':'banyak', 'dg':'dengan',
'anakanak':'anak', 'bocil':'anak', 'gk':'tidak', 'utnk':'untuk', 'smg':'semoga', 'ckup':'cukup', 'tsbt':'tersebut', 'dgn':'dengan',
'asik':'seru', 'ckp':'cukup', 'dr':'dari', 'bgus':'bagus', 'bnyak':'banyak', 'fto':'foto', 'ga':'tidak', 'jln':'jalan', 'dpt':'dapat',
'krn':'karena', 'krna':'karena', 'bs':'bisa', 'sdh':'sudah', 'td':'tadi', 'sbg':'sebagai', 'kelg':'keluarga',
'ttg':'tentang', 'ruame':'ramai', 'jga':'juga', 'bwt':'untuk', 'mantaps':'mantap', 'mantapz':'mantap', 'bgs':'bagus',
'bljr':'belajar', 'rame':'ramai', 'seneng':'senang', 'tyt':'ternyata', 'blm':'belum', 'tmpt':'tempat', 'gx':'tidak',
'cape':'lelah', 'capek':'lelah', 'gud':'bagus', 'lpa':'lupa', 'good':'bagus', 'mantaf':'mantap', 'best':'baik', 'santuy':'santai'}
```

```
dataclean=[]
for sentence in dataset["Clean Review"]:
#proses konversi kosakata slang
stopword_remove = str(sentence)
words = stopword_remove.split()
rfrm=[slangs[word] if word in slangs else word for word in words]
stopword_remove=" ".join(rfrm)
```

C. Ekstraksi Emosi

Ekstraksi emosi dilakukan dengan memberi label pada komentar wisatawan. Pelabelan data dilakukan dengan menggunakan NRC Emotion Lexicon (EmoLex) dari Saif M. Mohammad [12] yang telah disesuaikan, yang terdiri dari delapan emosi yaitu *anger* (marah), *anticipation* (antisipatif), *fear* (takut), *disgust* (muak / jijik), *sadness* (sedih), *surprise* (terkejut), *joy* (gembira), *trust* (yakin) dengan roda emosi Plutchik. Setiap komentar akan dikenakan *tokenizing* atau memisahkan setiap kata dalam kalimat menjadi kata tunggal. Kemudian setiap kata dalam komentar tersebut akan dicocokkan dengan jenis emosi sesuai kamus EmoLex. Apabila ada kata dalam kalimat tersebut mewakili emosi marah, antisipatif, takut, muak, sedih, terkejut, gembira, atau yakin sesuai dengan yang ada pada kamus EmoLex, maka akan diberi bobot ‘1’, sedangkan apabila emosi tidak terwakilkan maka akan diberi bobot ‘0’. Jenis polaritas emosi [12] dan maknanya ada pada Tabel 2.

TABEL 2.
POLARITAS EMOSI

Polaritas	Informasi	Skor
<i>Anger</i>	Emosi Marah	0 – 1
<i>Fear</i>	Emosi Takut	0 – 1
<i>Anticipation</i>	Emosi Antisipatif	0 – 1
<i>Trust</i>	Emosi Yakin	0 – 1
<i>Surprise</i>	Emosi Terkejut	0 – 1
<i>Sadness</i>	Emosi Sedih	0 – 1
<i>Joy</i>	Emosi Gembira	0 – 1
<i>Disgust</i>	Emosi Muak	0 – 1

Contoh ekstraksi emosi dari komentar “ *perjalanan ke baturaden menyenangkan dan indah*”. Setelah dilakukan *cleaning* dan *tokenizing*, maka hasil ekstraksi emosi kalimat tersebut seperti pada Tabel 3.

TABEL 3.
CONTOH REPRESENTASI EMOSI DARI SEBUAH KOMENTAR

Hasil tokenizing	Anger (marah)	Fear (takut)	Anticipation (antisipasi)	Trust (Percaya)	Surprise (terkejut)	Sadness (sedih)	Joy (gembira)	Disgust (jijik)
Perjalanan	0	0	0	0	0	0	1	0
baturaden	0	0	0	0	0	0	0	0

Hasil tokenizing	Anger (marah)	Fear (takut)	Anticipation (antisipasi)	Trust (Percaya)	Surprise (terkejut)	Sadness (sedih)	Joy (gembira)	Disgust (jijik)
menyenangkan	0	0	0	0	0	0	1	0
indah	0	0	0	0	0	0	1	0
Total bobot	0	0	0	0	0	0	3	0

Maka kalimat “*perjalanan ke baturaden menyenangkan dan indah*” masuk dalam kelas label “joy (gembira)

III. HASIL DAN PEMBAHASAN

A. Hasil Ekstraksi Emosi

Proses ekstraksi emosi dilakukan dengan bahasa pemrograman Python untuk memberi label pada setiap komentar. Algoritma proses ekstraksi menggunakan emotion lexicon adalah berikut.

#labeling dengan NRC

```
dictionary=pd.read_excel('NRC_EMOLEX.xlsx', sheet_name='dict')
```

#Membuat variabel untuk menampung array emosi dari EmoLex

```
anger=[], anticipation=[], disgust=[], fear=[], joy=[], sadness=[], surprise=[], trust=[], teksbaru=""
```

#Proses memanggil kamus emolex dan menyimpannya dalam setiap variabel emosi yang disediakan.

```
for i in range(1,len(dictionary)):
```

```
    kata=dictionary.iloc[i,0]
```

```
    if (dictionary.iloc[i,1])==1:
```

```
        anger.append(kata)
```

```
    if (dictionary.iloc[i,2])==1:
```

```
        anticipation.append(kata)
```

```
    if (dictionary.iloc[i,3])==1:
```

```
        disgust.append(kata)
```

```
    if (dictionary.iloc[i,4])==1:
```

```
        fear.append(kata)
```

```
    if (dictionary.iloc[i,5])==1:
```

```
        joy.append(kata)
```

```
    if (dictionary.iloc[i,6])==1:
```

```
        sadness.append(kata)
```

```
    if (dictionary.iloc[i,7])==1:
```

```
        surprise.append(kata)
```

```
    if (dictionary.iloc[i,8])==1:
```

```
        trust.append(kata)
```

#proses ekstraksi emosi

```
emosi=[]
```

```
for kalimat in dataclean:
```

```
    baris=[]
```

```
    baris.append(kalimat)
```

```
    f_marah=0, f_antipiasipasi=0, f_takut=0, f_jijik=0, f_sedih=0, f_terkejut=0, f_bahagia=0, f_yakin=0, y=kalimat
```

#proses menentukan jenis emosi setiap kata dalam kalimat

```
for teks in y.split():
```

```
    for j in range(0,len(anger)-1):
```

```
        if teks==anger[j]: #== equal
```

```
            f_marah+=1
```

```
    for j in range(0,len(anticipation)-1):
```

```
        if teks==anticipation[j]:
```

```
            f_antipiasipasi+=1
```

```
    for j in range(0,len(fear)-1):
```

```
        if teks==fear[j]:
```

```
            f_takut+=1
```

```
    for j in range(0,len(disgust)-1):
```

```

if teks==disgust[j]:
    f_jijik+=1
for j in range(0,len(sadness)-1):
    if teks==sadness[j]:
        f_sedih+=1
for j in range(0,len(joy)-1):
    if teks==joy[j]:
        f_bahagia+=1
for j in range(0,len(surprise)-1):
    if teks==surprise[j]:
        f_terkejut+=1
for j in range(0,len(trust)-1):
    if teks==trust[j]:
        f_yakin+=1
#menyimpan semua hasil emosi ke dalam setiap variabel array emosi
baris.append(f_marah, f_antisipasi, f_takut, f_jijik, f_sedih, f_terkejut, f_bahagia, f_yakin)
emosi.append(baris)

```

Hasil ekstraksi emosi dengan algoritma tersebut diatas berupa jenis emosi beserta persentasenya, ditunjukkan Tabel 4.

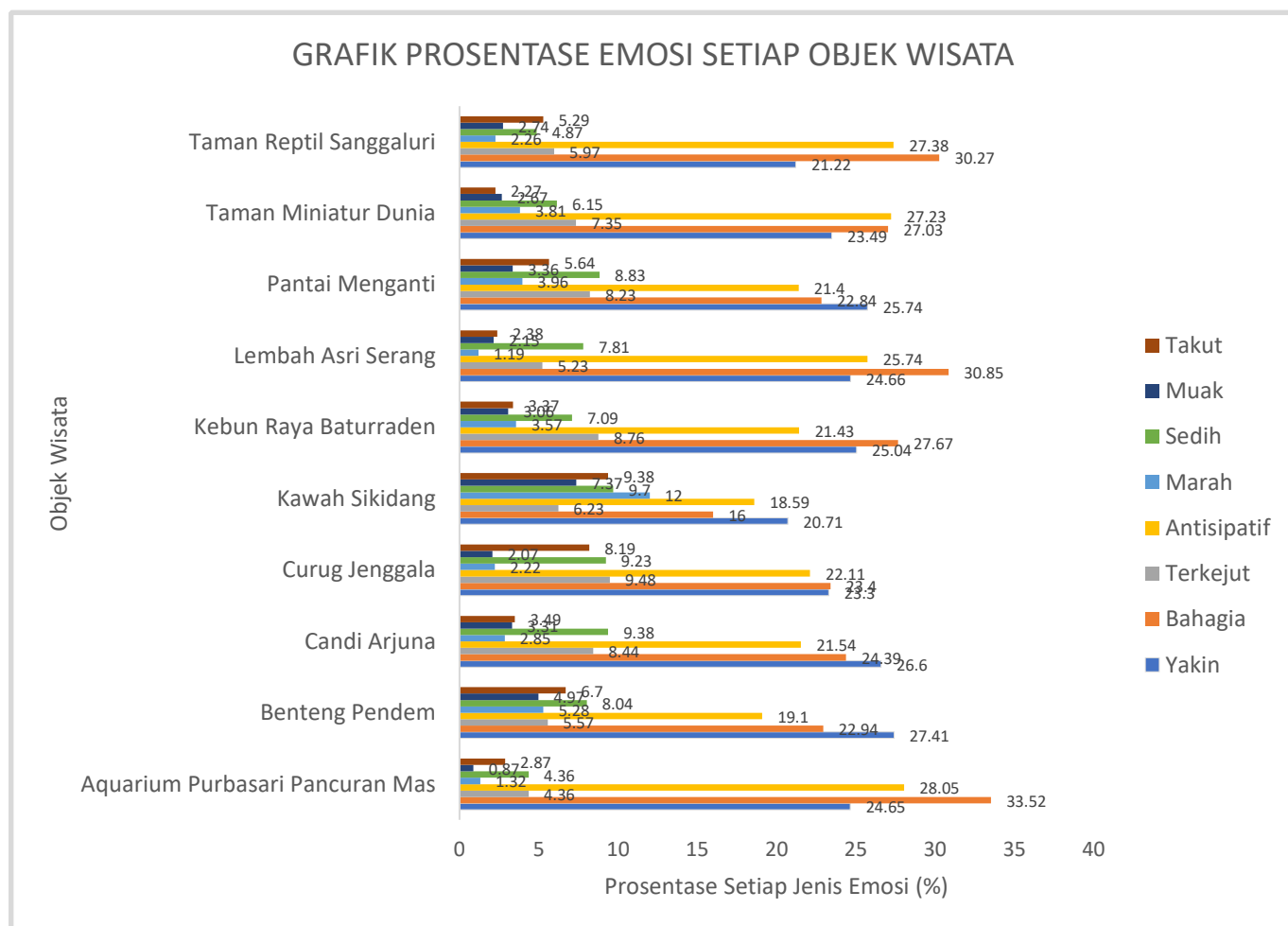
TABEL 4.
HASIL EKSTRAKSI EMOSI

Objek Wisata	Emosi							
	Yakin	Bahagia	Terkejut	Antisipatif	Marah	Sedih	Muak	Takut
Aquarium Purbasari Pancuran Mas	24,65%	33,52%	4,36%	28,05%	1,32%	4,36%	0,87%	2,87%
Benteng Pendem	27,41%	22,94%	5,57%	19,10%	5,28%	8,04%	4,97%	6,70%
Candi Arjuna	26,60%	24,39%	8,44%	21,54%	2,85%	9,38%	3,31%	3,49%
Curug Jenggala	23,30%	23,40%	9,48%	22,11%	2,22%	9,23%	2,07%	8,19%
Kawah Sikidang	20,71%	16,00%	6,23%	18,59%	12,00%	9,70%	7,37%	9,38%
Kebun Raya Baturraden	25,04%	27,67%	8,76%	21,43%	3,57%	7,09%	3,06%	3,37%
Lembah Asri Serang	24,66%	30,85%	5,23%	25,74%	1,19%	7,81%	2,15%	2,38%
Pantai Menganti	25,74%	22,84%	8,23%	21,40%	3,96%	8,83%	3,36%	5,64%
Taman Miniatur Dunia	23,49%	27,03%	7,35%	27,23%	3,81%	6,15%	2,67%	2,27%
Taman Reptil Sanggaluri	21,22%	30,27%	5,97%	27,38%	2,26%	4,87%	2,74%	5,29%

Tabulasi data pada Tabel 4 kemudian divisualisasikan. Visualisasi data bertujuan untuk mempermudah perbandingan emosi yang muncul pada setiap objek wisata [18]. Persentase setiap emosi pada kolom-kolom Tabel 4 merupakan perbandingan jumlah komentar yang mengandung emosi tertentu dibanding seluruh data komentar. Misalnya, data komentar pada objek wisata Aquarium Purbasari Pancuran mas sebanyak 1518 komentar. Hasil klasifikasi label dengan EmoLex, sebanyak 374 diklasifikasi sebagai emosi “yakin”. Maka, presentasi emosi “yakin” adalah 374/1518 atau 24,65%.

B. Analisis dan Pembahasan

Analisis informasi bisa menggunakan analisis visual berupa grafik dan *word cloud* [16]. Pada penelitian ini, hasil ekstraksi emosi dari seluruh komentar pada sepuluh objek wisata ditunjukkan oleh grafik pada Gambar 3. Grafik pada Gambar 3 menyajikan 8 emosi yang terkandung pada komentar- komentar yang dikumpulkan dari akun Google Map sepuluh objek wisata di BARLINGMASCAKEB.



Gambar 3. Grafik Emosi Tiap Objek Wisata

Gambar 3 menunjukkan perbandingan emosi yang dominan pada sepuluh objek wisata di BARLINGMASCAGEB. Secara keseluruhan, objek wisata di BARLINGMASCAGEB menghadirkan emosi dasar positif dari wisatawan yang ditunjukkan dengan mayoritas emosi yakin dan bahagia. Warna diagram diwakili mewakili jenis emosi. Mayoritas emosi pada setiap objek wisata adalah dominan warna emosi “bahagia (joy)” dan “yakin (trust)”. Kemudian emosi berikutnya yang dominan adalah emosi yakin dan antisipatif. Objek wisata dengan emosi mayoritas “bahagia” adalah Aquarium Purbasari Pancuran Mas, Kebun Raya Baturraden, Curug Jenggala, Lembah Asri Serang, dan Taman Reptil Sanggaluri. Objek wisata dengan emosi tertinggi “yakin” yaitu Benteng Pendem, Candi Arjuna, Kawah Sikidang, dan Pantai Menganti.

Jika dilihat dari sudut pandang jenis kelompok emosi positif dan negatif, maka emosi “bahagia”, paling tinggi pada Aquarium Purbasari Pancuran Mas dengan nilai 33,52%. Emosi “yakin” paling banyak pada Benteng Pendem yaitu 27,4 %. Emosi Antisipatif paling tinggi pada Taman Reptil Sanggaluri, sebesar 27,38%.

Jika memperhatikan makna leksikal kata dalam kamus bahasa Indonesia (KBI), kedua emosi (kata bahagia dan yakin) tersebut mengacu pada makna seperti “senang”, “bagus”, “berbahagia”, “percaya”, “sungguh-sungguh”, “merasa pasti” [17]. Sedangkan antisipatif mengacu pada makna leksikal yang netral artinya tidak mengacu pada makna yang menunjukkan rasa emosi positif atau negatif. Makna leksikal antisipatif dalam KBI adalah “tanggap terhadap sesuatu yang sedang/akan terjadi”, “memperhatikan/memperhitungkan hal-hal yang akan terjadi” [17]. Sedangkan emosi negatif yang diwakili oleh jenis emosi takut, muak, sedih, marah, dan terkejut secara umum tidak muncul dalam jumlah besar dari, hanya di bawah 13%. Komentar wisatawan dengan emosi terkejut juga cukup rendah dengan tingkat kemunculan dibawah 10%. Emosi terkejut, marah, sedih, takut, jijik dikonotasikan dengan kata-kata yang maknanya negatif. Terlihat pada objek wisata Candi Arjuna, Curug Jenggala, Kawah Sikidang emosi Sedih mencapai lebih dari 9%. Bahkan objek wisata Kawah Sikidang menghadirkan emosi marah sebesar 12,00%. Ekspresi takut terlihat hanya muncul paling banyak pada objek Kawah Sikidang juga dengan emosi muak sekitar 9% dan Curug Jenggala sekitar 8%. Jika dilihat dari jenis objek wisatanya, bisa jadi ekspresi

takut karena medan objek wisata. Curug biasanya medannya terjal dan sulit, sedangkan kawah tentu saja menakutkan karena mengandung unsur bahaya. Nilai emosi marah pada objek kawah Sikidang cukup tinggi hampir 10%. Hal ini perlu dikaji lebih dalam, apakah yang menyebabkan emosi tersebut.

Dari nilai tersebut bisa ditelusuri lebih lanjut, apakah kata-kata yang sering diungkapkan oleh wisatawan di Google Map Review. Dengan demikian akan lebih didalami kira-kira apa penyebab emosi wisatawan tersebut. Analisis menggunakan visualisasi *word cloud* bisa menyajikan informasi yang mudah dipahami melihat kata kunci yang paling sering muncul pada setiap emosi, lihat Gambar 4 – Gambar 13. Frekuensi kata dalam setiap dataset obyek wisata mencapai lebih dari 1500 kata, maka visualisasi setiap obyek wisata hanya diambil 20 kata dengan jumlah frekuensi teratas, ditunjukkan oleh Tabel 5 – Tabel 14.

TABEL 5.
FREKUENSI 20 KATA TERATAS PADA OBYEK WISATA AQUARIUM PURBASARI PANCURAN MAS

No	Word	Frekuensi	No	Word	Frekuensi
1	anak	344	11	liburan	88
2	keluarga	211	12	sangat	79
3	bagus	200	13	lengkap	72
4	edukasi	163	14	wahana	71
5	luas	105	15	murah	66
6	nyaman	105	16	berbagai	65
7	burung	97	17	tempatny	64
8	rekreasi	93	18	bersih	60
9	cocok	90	19	renang	59
10	kolam	89	20	tiket	58



Gambar 4. Wordcloud Aquarium Purbasari Pancuran Mas

TABEL 6.
FREKUENSI 20 KATA TERATAS PADA OBYEK WISATA BENTENG PENDEM

No	Word	Frekuensi	No	Word	Frekuensi
1	bagus	248	11	satu	95
2	kurang	234	12	jaman	94
3	terawat	208	13	masuk	86
4	peninggalan	188	14	sayang	82
5	lebih	126	15	baik	81
6	penyu	123	16	penjajahan	78
7	teluk	120	17	menarik	76
8	keluarga	103	18	nyaman	76
9	belanda	99	19	bersih	75
10	sangat	96	20	tempatny	75



Gambar 5. Wordcloud Benteng Pendem

TABEL 7.
FREKUENSI 20 KATA TERATAS PADA OBYEK WISATA CANDI ARJUNA

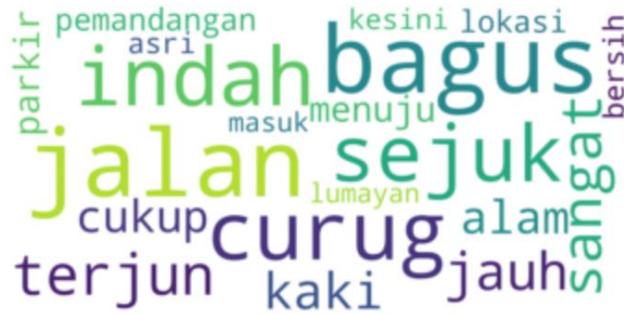
No	Word	Frekuensi	No	Word	Frekuensi
1	bagus	255	11	kesini	102
2	masuk	209	12	luas	101
3	sejuk	184	13	pemandangan	99
4	tiket	170	14	jadi	92
5	dieng	162	15	foto	91
6	sangat	161	16	kawasan	86
7	dingin	144	17	kalau	85
8	indah	143	18	cukup	82
9	satu	135	19	keluarga	75
10	bersih	127	20	lebih	73



Gambar 6. Wordcloud Candi Arjuna.

TABEL 8.
FREKUENSI 20 KATA TERATAS PADA OBYEK WISATA CURUG JENGGALA

Time	Word	Frekuensi	No	Word	Frekuensi
1	jalan	128	11	cukup	52
2	bagus	128	12	menuju	45
3	curug	94	13	parkir	44
4	indah	94	14	pemandangan	44
5	sejuk	88	15	lokasi	42
6	terjun	86	16	bersih	39
7	sangat	86	17	asri	37
8	jauh	72	18	kesini	35
9	kaki	67	19	masuk	32
10	alam	55	20	lumayan	32



Gambar 7. Wordcloud Curug Jenggala

TABEL 9
FREKUENSI 20 KATA TERATAS PADA OBYEK WISATA KAWAH SIKIDANG

No	Word	Frekuensi	No	Word	Frekuensi
1	bagus	387	11	jadi	153
2	jalan	335	12	kesini	153
3	sangat	249	13	lebih	134
4	keluar	245	14	jauh	132
5	belerang	216	15	jalur	130
6	foto	205	16	pintu	125
7	masker	194	17	kalau	125
8	bau	190	18	menyengat	122
9	masuk	182	19	dieng	122
10	spot	164	20	keluarnya	122



Gambar 8. Wordcloud Kawah Sikidang

TABEL 10.
FREKUENSI 20 KATA TERATAS PADA OBYEK WISATA KEBUN RAYA BATURADEN

No	Word	Frekuensi	No	Word	Frekuensi
1	sejuk	176	11	kurang	47
2	bagus	162	12	jalan	47
3	keluarga	100	13	segar	44
4	asri	89	14	terawat	43
5	indah	87	15	tanaman	42
6	alam	76	16	masuk	41
7	nyaman	72	17	hutan	37
8	sangat	68	18	udara	34
9	cocok	57	19	foto	33
10	tempatnya	49	20	bersama	33

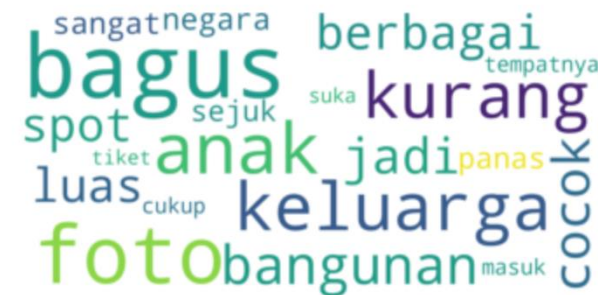
No	Word	Frekuensi	No	Word	Frekuensi
9	naik	50	19	kendaraan	38
10	menuju	49	20	tebing	37



Gambar 11. Wordcloud Pantai Menganti

TABEL 13
FREKUENSI 20 KATA TERATAS PADA OBYEK WISATA TAMAN MINIATUR DUNIA

No	Word	Frekuensi	No	Word	Frekuensi
1	bagus	162	11	spot	66
2	foto	132	12	negara	64
3	anak	100	13	sangat	64
4	keluarga	89	14	panas	60
5	kurang	82	15	sejuk	58
6	bangunan	80	16	tempatnya	58
7	cocok	74	17	masuk	58
8	jadi	67	18	cukup	56
9	berbagai	67	19	suka	50
10	luas	66	20	indah	48



Gambar 12. Wordcloud Taman Miniatur Dunia

TABEL 14.
FREKUENSI 20 KATA TERATAS PADA OBYEK WISATA TAMAN REPTIL SANGGALURI

No	Word	Frekuensi	No	Word	Frekuensi
1	anak	252	11	kurang	71
2	edukasi	164	12	sangat	56
3	keluarga	151	13	luas	55
4	wahana	145	14	belajar	54
5	bagus	145	15	uang	53
6	rekreasi	94	16	bersih	51
7	museum	86	17	baik	49

No	Word	Frekuensi	No	Word	Frekuensi
8	nyaman	80	18	lebih	48
9	cocok	78	19	cukup	42
10	bermain	72	20	terawat	42



Gambar 13. Wordcloud Taman Reptil Sanggaluri

Tabel 5 – Tabel 14 menunjukkan kata kunci yang paling sering muncul adalah pada kesepuluh objek tersebut adalah kata-kata positif seperti “bagus”. Kata “bagus” adalah komentar wisatawan paling banyak (teratas) pada objek wisata Candi Arjuna, Benteng Pendem, Kawah Sikidang, Pantai Menganti, dan Taman Miniatur Dunia. Kata kunci “sejuk” paling sering muncul dalam komentar pada objek Lembah Asri Serang, Kebun Raya Baturaden, Curug Jenggala, dan Candi Arjuna. Kata kunci yang sifatnya netral paling sering muncul pada objek Taman Reptil Sanggaluri dan Aquarium Purbasari Pancuran Mas, diwakili kata kunci “anak”.

Jika dihubungkan dengan persentase emosi negatif maka kata kunci sering muncul yaitu kata “jalan” pada objek Curug Jenggala dan Kawah Sikidang. Jika dikorelasikan dengan grafik emosi (Gambar 3) kedua obyek wisata, maka kata jalan ini berkorelasi dengan emosi marah. Sedangkan pada emosi negatif marah, diwakili oleh kata yang paling banyak yaitu “jalan”, “belerang”, dan “bau” dan “masker”. Kata kunci tersebut sangat berkorelasi erat dengan medan objek wisata kawah. Emosi dikorelasikan oleh kata sifat atau kata kerja. Kata benda seperti belerang, masker dan jalan sesungguhnya bermakna netral jika merujuk pada makna leksikal KBBI. Dengan demikian, perlu adanya *preprocessing* yang lebih tepat untuk memisahkan kata benda sehingga tidak dianggap sebagai kata sifat.

IV. SIMPULAN

Dapat disimpulkan bahwa mayoritas emosi yang dihadirkan oleh kesepuluh objek wisata di BARLINGMASCAKEB adalah emosi positif dengan jenis “bahagia (joy)”, “yakin (trust)”, “antisipatif (anticipation)”. Objek wisata yang menghadirkan emosi mayoritas “bahagia” adalah Aquarium Purbasari Pancuran Mas, Kebun Raya Baturraden, Curug Jenggala, Lembah Asri Serang, dan Taman Reptil Sanggaluri. Objek wisata yang memberikan pengalaman emosi tertinggi “yakin” yaitu Benteng Pendem, Candi Arjuna, Kawah Sikidang dan Pantai Menganti. Analisis kata kunci menggunakan *word cloud* menunjukkan kata-kata yang sering digunakan wisatawan dalam mengungkapkan pengalamannya atau pendapatnya terhadap objek wisata. Emosi diwakili oleh kata yang paling sering muncul. Pada Aquarium Purbasari Pancuran Mas emosi positif diwakili oleh kata paling banyak yaitu “bagus”. Benteng Pendem diwakili kata emosi “bagus”. Candi Arjuna diwakili kata “bagus” dan “dingin”. Curug Jenggala juga sama seperti Candi Arjuna, kata “bagus” menjadi kata yang dominan. Kawah Sikidang memunculkan kata “bagus”. Kebun Raya Baturraden diwakili kata “bagus” dan “sejuk”. Miniatur Dunia diwakili kata “bagus”, Taman Reptil Sanggaluri diwakili kata “bagus”. Pantai Menganti diwakili oleh kata “bagus” dan “indah”,Lembah Asri diwakili kata “bagus” dan “sejuk”. Sedangkan kata kunci negatif yang paling banyak muncul yaitu “bau”, “jalan”, “masker”, dan “belerang”. Kata kata belerang dan masker ini sifatnya sebenarnya bisa netral. Berdasarkan hasil- hasil analisis *wordcloud* dapat disimpulkan untuk mengetahui kata kunci yang benar-benar mengandung emosi yang tepat sesuai makna kata tersebut, perlu *pre-processing* yang lebih baik sehingga tidak ambigu antara kata dengan makna negatif dan netral.

DAFTAR PUSTAKA

- [1] R. Destiningsih, R. R. Sugiharti, and A. Achsa, “Food Commodity Competitiveness and Accessibility in Barlingmascakeb,” *Jejak (Journal Econ. Policy)*, vol. 12, no. 1, pp. 218–237, 2019.
- [2] S. M. Ibnu, “Representasi Pariwisata Purwokerto Di Instagram Studi Analisis Isi Pada Akun Instagram @INSTAPURWOKERTO,” p. 128, 2020.
- [3] F. U. Haq and H. Rachmat, “Penggunaan Google Review Sebagai Penilaian Kepuasan Pengunjung Dalam Pariwisata,” *Tornare J. Sustain. Tour.*

- Res., vol. 2, no. 1, p. 10, 2020.
- [4] A. B. Gumelar, E. M. Yuniarto, W. Anggraeni, I. Sugiarto, A. A. Kristanto, and M. H. Purnomo, "Kombinasi Fitur Multispektrum Hilbert dan Cochleagram untuk Identifikasi Emosi Wicara (Spectrum Features Combination of Hilbert and Cochleagram for Speech Emotions Identification)," *J. Nas. Tek. Elektro dan Teknol. Inf. /*, vol. 9, no. 2, pp. 180–189, 2020.
- [5] S. Khomsah, R. D. Ramadhani, and S. Wijayanto, "Big Data Analytics to Analyze Sentiment, Emotions, and Perceptions of Travelers (Case Study: Tourism Destination in Purwokerto Indonesia)," *J. E-Komtek*, vol. 5, no. 2, pp. 284–297, 2021.
- [6] A. R. Alaci, S. Becken, and B. Stantic, "Sentiment Analysis in Tourism: Capitalizing on Big Data," *J. Travel Res.*, vol. 58, no. 2, pp. 175–191, 2019.
- [7] F. Aziz, A. R. Thaha, and N. A. Ma'ruf, "Analisis Sentimen Destinasi Wisata Geopark Ciletuh," *J. Ilm. Pariwisata*, vol. 27, no. 1, p. 60, 2022.
- [8] A. Rifa'i, H. Sujaini, and D. Prawira, "Sentiment Analysis Objek Wisata Kalimantan Barat Pada Google Maps Menggunakan Metode Naive Bayes," *JEPIN (Jurnal Edukasi dan Penelit. Inform.)*, vol. 3, no. 7, pp. 400–407, 2021.
- [9] O. Somantri and Dairoh, "Analisis Sentimen Penilaian Tempat Tujuan Wisata Kota Tegal Berbasis Text Mining," *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 2, pp. 191–196, 2019.
- [10] R. Ferryawan, Kusriani, and W. W. Ferry, "Analisis Sentimen Wisata Jawa Tengah Menggunakan Naive Bayes," *INFORMA*, vol. 5, no. 3, pp. 2442–7942, 2019.
- [11] M. S. Saputri, R. Mahendra, and M. Adriani, "Emotion Classification on Indonesian Twitter Dataset," *Proc. 2018 Int. Conf. Asian Lang. Process. IALP 2018*, no. November, pp. 90–95, 2019.
- [12] S. M. Mohammad, "Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text," *Emot. Meas.*, pp. 201–237, 2016.
- [13] A. S. Aribowo, H. Basiron, N. S. Herman, and S. Khomsah, "An Evaluation of Preprocessing Steps and Tree-based Ensemble Machine Learning for Analysing Sentiment on Indonesian YouTube Comments," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 5, pp. 7078–7086, 2020.
- [14] A. Mathur, P. Kubde, and S. Vaidya, "Emotional analysis using twitter data during pandemic situation: Covid-19," in *Proc. 5th Int. Conf. Commun. Electron. Syst. ICCES 2020*, IEEE Explore, 2020, pp. 845–848.
- [15] S. Baskoro, B., Susanto, I., & Khomsah, "Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF (Studi Kasus: Ulasan Pelanggan Pada Situs TRIPADVISOR)," *J. Informatics Inf. Syst. Softw. Eng. Appl.*, vol. 3, no. 2, pp. 21–29, 2021.
- [16] I. D. Komang, N. N. U. Januhari, R. A. N. Diaz, I. P. Ramayasa, and I. M. A. W. Putra, "Visualisasi Data Opini Publik pada Media Sosial Twitter (Studi Kasus : Nusantara Sebagai IKN Indonesia)," *J. Tek. Inform. Unika ST. Thomas*, vol. 07, no. 02, pp. 1501–2657, 2022.
- [17] B. P. dan Pembinaan Bahasa, "KBBI Daring," *Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia*. 2016.