

Implementasi *Lexicon Based* Untuk Analisis Sentimen Dalam Mengetahui *Trend* Wisata Pantai Di DI Yogyakarta Berdasarkan Data Twitter

Amalia Rachmadana Ismail^{1*}, Raden Bagus Fajriya Hakim¹

¹Universitas Islam Indonesia, Jl. Kaliurang KM 14.5, DI. Yogyakarta, Indonesia 55584

*Corresponding author: amalia.ismail@students.uii.ac.id



P-ISSN
E-ISSN

Riwayat Artikel

Dikirim

3 Januari 2023

Direvisi

8 Januari 2023

Diterima

17 Januari 2023

ABSTRAK

Saat ini, peningkatan sektor pariwisata Indonesia tidak lepas dari yang namanya teknologi. Selama berwisata, wisatawan akan berbagi pengalaman mereka secara *real time* melalui unggahan foto dan video serta komentar aktif di media sosial. Salah satu media sosial yang menjadi favorit di Indonesia adalah *twitter*. Dari *tweet* tersebut akan diketahui destinasi wisata yang sedang populer dan sedang ramai diperbincangkan oleh masyarakat. Oleh karena itu kita membutuhkan sistem untuk memberikan informasi tentang tren pariwisata di sebuah kota. Tentu informasi tersebut sangat bermanfaat bagi wisatawan, masyarakat, pemerintah, dan sektor pariwisata di kota tersebut. Dalam penelitian ini, sistem yang dapat menentukan destinasi wisata favorit melalui *hashtag* atau *keyword* tertentu dengan melakukan analisis sentimen. Salah satu dari pendekatan untuk melakukan analisis sentimen adalah menggunakan metode *lexicon based*. Berdasarkan analisis yang telah dilakukan dengan *keyword* 'wisata pantai jogja', diperoleh bahwa wisata pantai di DI Yogyakarta yang sedang ramai diperbincangkan adalah pantai slili dan pantai siung. Hasil analisis sentimen menunjukkan bahwa mayoritas masyarakat yang menulis *tweet* tentang wisata pantai jogja menunjukkan opini yang positif terhadap wisata pantai yang ada di DI Yogyakarta.

Kata Kunci: Twitter, Analisis Sentimen, Lexicon, Pantai, Jogja.

ABSTRACT

Nowadays, the increase in Indonesia's tourism sector cannot be separated from technology. During the trip, tourists will share their experiences in real time through photo and video also active comments on social media. One of the favorite social media in Indonesia is Twitter. From the tweet, it will be known that tourists destinations are popular and being discussed by public. Therefore, we need a system to provide information about tourism trends in a city. Such information is very useful for tourist, the public, the government, and the tourist sector in the city. In this study, a system that can determine favorite tourist destinations through certain hashtags or keywords by conducting sentiment analysis. One of the approaches to conducting sentiment analysis is to use the lexicon based method. Based on the analysis that has been done with keywords 'jogja beach tourism', obtained that beach tourism in DI Yogyakarta that is being discussed and popular is slili beach and siung beach. The result of the sentiment analysis showed that the majority of people

who wrote tweets about jogja beach tourism showed a positive opinion towards the beach tourism in DI Yogyakarta.

Keywords: *Twitter, Sentiment Analysis, Lexicon, Beach, Tourism, Jogja*

1. Pendahuluan

Salah satu masalah sosial yang menjadi tantangan bagi hampir seluruh negara, teknologi memainkan peran penting dalam sektor pariwisata yang perlahan-lahan mulai meningkat di Indonesia. Era digital saat ini sangat memudahkan masyarakat untuk mengakses informasi dengan mudah dan cepat. Hampir semua masyarakat negara Indonesia terhubung ke internet menggunakan berbagai perangkat, baik ponsel, pc atau perangkat lainnya. Perkembangan media sosial telah mengubah kondisi pasar untuk sektor pariwisata [1]. Dalam beberapa tahun terakhir, media sosial telah berkembang sangat cepat dikarenakan tingginya fasilitas yang ditawarkan oleh media sosial. Terdapat beberapa media sosial yang digunakan oleh masyarakat, salah satunya adalah *twitter* yang menjadi favorit bagi masyarakat Indonesia [2]. Dalam layanan *microblogging* seperti *twitter*, pengguna cenderung memposting *tweet* secara *real time*. Informasi yang sedang hangat dibicarakan sangat mudah dicari melalui *twitter*. Salah satunya adalah informasi tentang tujuan wisata, dimana kita dapat melihat opini publik yang terkait dengan pariwisata melalui tagar populer.

Proliferasi teknologi baru seperti perangkat seluler dan *platform* media sosial membuat masyarakat sangat mudah mengekspresikan pendapat mereka. Partisipasi jutaan orang di media sosial menghasilkan big data yang memberikan kesempatan bagi para peneliti dan bisnis untuk mengetahui pendapat masyarakat tentang suatu topik. Analisis sentimen merupakan salah satu alat untuk mengekstrak sentimen dari data media sosial. [3] Mendefinisikan analisis sentimen sebagai studi komputasi pendapat, penilaian, sikap, dan emosi masyarakat terhadap entitas, individu, masalah, peristiwa, dan topik mereka. Situs *web microblogging twitter* dengan ratusan juta pengguna dan lebih dari 500 juta *tweet* yang dikirim setiap hari, memegang posisi unik di antara *platform* media sosial. Mengingat banyaknya *database twitter* yang tersedia, data *twitter* telah banyak digunakan untuk menganalisis sentimen masyarakat tentang peristiwa, produk, layanan, dan pandangan politik.

Beberapa penelitian masih kesulitan tentang mengekstraks sentimen *twitter* dikarenakan kendala panjang pesan yang hanya terdiri dari maksimal 140 karakter serta kurangnya ortografi konvensional sehingga data *twitter* memberikan tantangan tambahan untuk melakukan analisis sentimen. Secara umum, terdapat dua jenis metode untuk analisis sentimen diantaranya *supervised learning* dan *unsupervised learning*. *Supervised learning* dilakukan dengan cara pengklasifikasian sentimen objek ke dalam salah satu kelas yaitu positif, negatif, dan netral. *Supervised learning* dalam analisisnya membutuhkan pelatihan dan pengujian data. Data pelatihan harus diberi label positif, negatif, atau netral. Metode *supervised learning* seperti klasifikasi *naïve bayes* dan *support vector machine* (SVM) paling sering digunakan oleh para peneliti. Umumnya, metode *supervised learning* menghasilkan *output* dengan akurasi yang tinggi karena pelabelan data dibutuhkan usaha yang besar. Sedangkan *unsupervised learning* menggunakan kata-kata frasa opini dari masing-masing teks yang diteliti dan masuk ke dalam daftar kata-kata atau *lexicon*, di mana setiap kata memiliki sentimen tertentu.

Sekarang sudah banyak kajian yang dilakukan tentang analisis sentimen dengan berbagai metode yang digunakan seperti yang dilakukan Lazuardi, D. R, pada tahun 2014

Penelitian ini menunjukkan bahwa penelitian tersebut menemukan tiga asosiasi kata yang paling dominan mengenai persepsi kualitas merek telkomsel, ada juga penelitian dari Fanani pada tahun 2017 yang membahas tentang Penggunaan Support Vector Machine dan Naïve Bayes Classifier untuk mengklasifikasikan review software menjadi SPAM dan bukan SPAM, dan terakhir ada penelitian yang dilakukan Abtohi tahun 2017 mengenai Penggunaan SVM untuk klasifikasi sentimen pengguna hotel. Dari penelitian sebelumnya yang menjadi pembeda dalam penelitian ini terdapat pada data yang diambil atau di peroleh serta ada yang menggunakan Metode Naïve Bayes.

Dalam makalah ini penulis menyajikan metode analisis sentimen dengan metode *lexicon based* mengenai opini masyarakat pada media sosial *twitter* tentang wisata pantai yang sedang *trend* di DI Yogyakarta. Hasil dari penelitian diharapkan dapat memberikan informasi yang berguna untuk mengetahui pantai yang sedang *trend* dan sering disebut pada media sosial *twitter* dan sentimen dari pantai tersebut.

2. Metodologi Penelitian

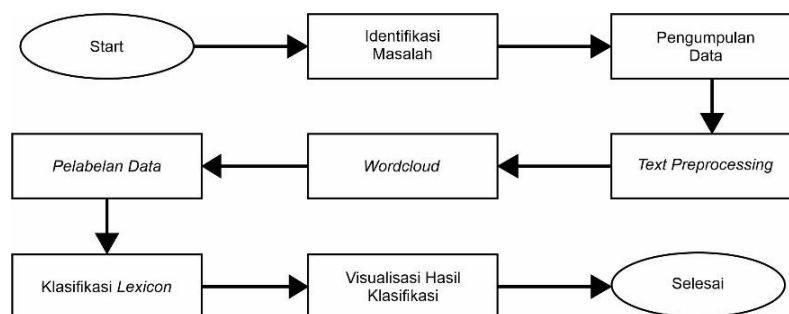
2.1. Data dan Metode

Data yang digunakan adalah data *tweet* pengguna *twitter* dengan *keyword* ‘wisata pantai jogja’. Pengumpulan data dilakukan dengan proses *web crawling* melalui Twitter API dengan memanfaatkan *library twitter* menggunakan bahasa pemrograman *RStudio*. Adapun variabel penelitian yang digunakan dapat dilihat pada 1.

Tabel 1 Variabel Penelitian

<i>Variabel</i>	Definisi Operasional
Tweet	Postingan berupa tulisan pengguna <i>twitter</i> .
Skor	Nilai yang diberikan pada setiap teks berdasarkan kamus kata positif dan negatif.
Label	Teks yang sudah terlabel dan memiliki kategori sentimen positif, negatif, dan netral.

Metode yang digunakan dalam penelitian ini adalah *text mining* dan analisis sentimen dengan menggunakan metode *Lexicon Based*. Tujuan penelitian ini ialah dapat mengetahui rekomendasi pantai di DI Yogyakarta serta opini masyarakat terhadap pantai tersebut.



Gambar 1. Diagram Alir Penelitian

Beberapa tahapan dilakukan dalam penelitian ini antara lain, (1) pengumpulan data, dilakukan dengan *web crawling* melalui *twitter* API, (2) *text preprocessing*, dilakukan *cleansing data* meliputi *remove punctuation*, *case folding*, *spelling normalization*, *filtering*, dan *tokenizing*, (3) *wordcloud*, dilakukan visualisasi untuk melihat kata yang paling sering muncul, (4) pelabelan data, dilakukan *text scoring* berdasarkan kamus kata negatif dan positif, (5) klasifikasi, dilakukan untuk menentukan kelas sentimen dari hasil *scoring*, (6) visualisasi hasil pelabelan, dilakukan visualisasi menggunakan *barplot* untuk melihat frekuensi hasil sentimen.

2.2. Text Preprocessing

Data *tweet* yang berhasil dikumpulkan memiliki banyak *noise* sehingga perlu adanya tahapan *text preprocessing* yang didalamnya akan dilakukan *cleansing data* meliputi *remove punctuation*, *case folding*, *spelling normalization*, *filtering*, dan *tokenizing* agar proses analisis sentimen menjadi lebih akurat.

2.3. Pelabelan Metode Lexicon

Setelah data dibersihkan pada tahapan *text preprocessing*, dilakukan pelabelan menggunakan metode *lexicon*. Penentuan dilakukan pada data teks berupa kalimat yang memiliki kata pada kamus *lexicon* yang terdiri dari kata negatif dan positif. Kata yang teridentifikasi dalam kamus *lexicon* akan dihitung skornya sesuai dengan jumlah kata pada setiap teks atau kalimat.

$$S_{positive} = \sum_{i \in t}^n positive\ score_i \quad (1)$$

$$S_{negative} = \sum_{i \in t}^n negative\ score_i \quad (2)$$

Dimana ($S_{positive}$) adalah bobot dari kalimat yang didapatkan melalui penjumlahan n skor polaritas kata opini positif dan ($S_{negative}$) adalah bobot dari kalimat yang didapatkan melalui penjumlahan n skor polaritas kata opini negatif. Dari persamaan nilai sentimen dalam satu kalimat maka diperoleh persamaan 3 untuk menentukan orientasi sentimen dengan perbandingan jumlah nilai positif, negatif dan netral.

$$Sentence_{sentiment} = \begin{cases} positive & \text{if } S_{positive} > S_{negative} \\ neutral & \text{if } S_{positive} = S_{negative} \\ negative & \text{if } S_{positive} < S_{negative} \end{cases} \quad (3)$$

Jika dalam suatu teks memiliki jumlah kata positif lebih banyak dari kata negatif, maka data teks tersebut akan dilabeli sentimen positif. Jika dalam suatu teks memiliki jumlah kata positif lebih sedikit dari kata negatif, maka data teks tersebut akan dilabeli sentimen negatif. Jika dalam suatu teks memiliki jumlah kata positif sama dengan kata negatif, maka data teks tersebut akan dilabeli sentimen netral.

2.4. Visualisasi

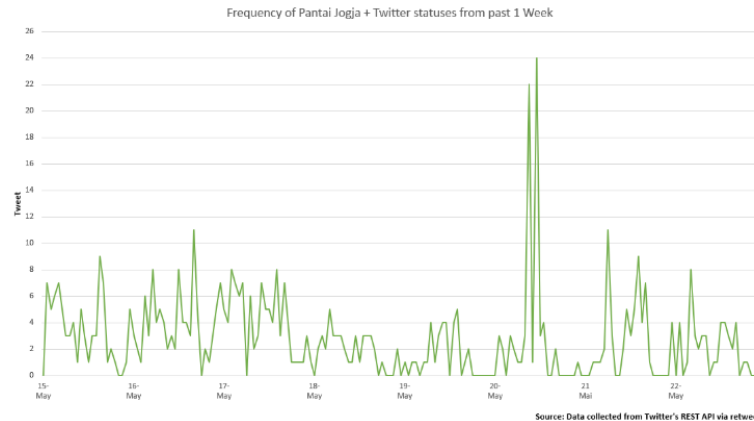
Visualisasi dalam penelitian ini akan berupa *wordcloud* dan diagram batang. Dimana *wordcloud* digunakan untuk memvisualisasikan kata-kata yang paling banyak muncul dalam data teks. Sedangkan diagram batang untuk memvisualisasikan frekuensi dari hasil analisis sentimen.

3. Hasil dan Pembahasan

Data yang digunakan pada penelitian ini diambil dari media sosial *twitter* sebanyak 540 *tweets* melalui kata kunci ‘wisata pantai jogja’ pada tanggal 15 Mei 2022 hingga 22 Mei 2022. Data yang diperoleh akan dilakukan analisis sentimen menggunakan metode *lexicon based* sesuai dengan tahapan yang sudah dijelaskan pada bab sebelumnya.

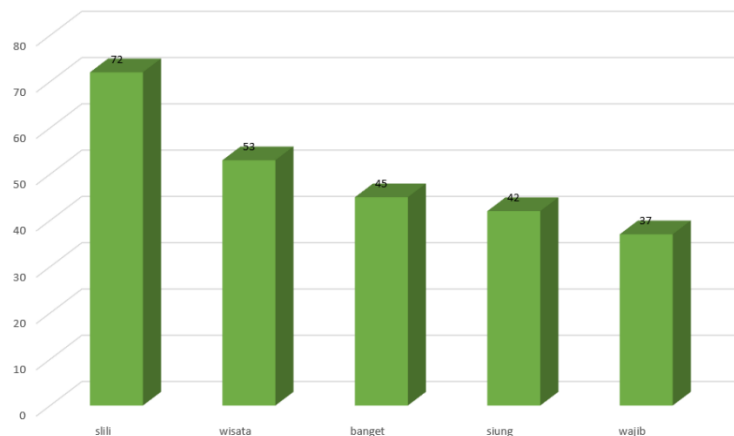
3.1. Analisis Deskriptif

Analisis deskriptif digunakan untuk mengetahui gambaran secara umum mengenai penelitian yang dilakukan guna mendeskripsikan data yang terkumpul dengan tujuan untuk membuat kesimpulan yang berlaku umum.



Gambar 2. Frekuensi *Tweets* dalam 1 Minggu

Pada **Gambar 2.** menunjukkan frekuensi *tweets* setiap jam pada tanggal 15 Mei 2022 hingga 22 Mei 2022. Dapat di lihat bahwa jumlah *tweets* mengalami fluktuasi hingga pada akhirnya meningkat drastis pada tanggal 20 Mei 2022 dan kembali menurun beberapa jam setelahnya, dan pada akhirnya terus mengalami fluktuasi hingga tanggal 22 Mei 2022.



Gambar 3. Frekuensi Kata Paling Sering Muncul

Pada **Gambar 3.** menunjukkan 5 kata teratas yang paling sering muncul pada cuitan jejaring sosial *twitter* tentang Wisata Pantai di DI Yogyakarta. Kata “slii” muncul paling muncul yaitu 72 kali, kata “wisata” muncul sebanyak 53 kali, kata “banget” muncul sebanyak 45 kali, kata siung muncul sebanyak 42 kali, dan kata “wajib” muncul sebanyak 37 kali.

Tuliskan analisis yang dilakukan. Gunakan TNR 11, spasi 1, *before and after spacing* 6, dan paragraph rata-kiri.

3.2. Pengumpulan Data

Pengumpulan data dilakukan dengan cara *web crawling* melalui Twitter API memanfaatkan *library* twitter menggunakan bahasa pemrograman *RStudio*. Adapun *keyword* yang digunakan yaitu ‘wisata pantai jogja’ dan berhasil didapatkan sebanyak 540 data *tweet* dari tanggal 15 Mei 2022 sampai dengan 22 Mei 2022. Data tersebut kemudian

disimpan kedalam format .csv untuk memudahkan tahapan selanjutnya. Contoh data yang diperoleh dari hasil *web crawling* dapat dilihat pada **Gambar 4**.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
514	513 wishlist yg ogn bgt bgt	FALSE	2	NA	5/15/2022 16:15	TRUE	NA	1.53E+18	NA		0	FALSE	FALSE	
515	514 Klo ditung2 banyak bgt	FALSE	0	NA	5/15/2022 16:15	TRUE	NA	1.53E+18	NA		0	FALSE	FALSE	
516	515 Baru saja mengirim foto	FALSE	0	NA	5/15/2022 15:40	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
517	516 @hiannananaa @jogjar	FALSE	0	hiannana	5/15/2022 15:38	TRUE	1.53E+18	1.53E+18	1.22E+18		0	FALSE	FALSE	
518	517 Gumuk Pasti Jeep	FALSE	70	NA	5/15/2022 15:33	TRUE	NA	1.53E+18	NA		2	FALSE	FALSE	
519	518 @kimeyewon jadi cerita	FALSE	0	kimeyew	5/15/2022 15:04	TRUE	1.53E+18	1.53E+18	8.77E+17		0	FALSE	FALSE	
520	519 Baru saja mengirim foto	FALSE	0	NA	5/15/2022 15:03	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
521	520 (Tribun jogja) Selama Lil	FALSE	0	NA	5/15/2022 14:30	TRUE	NA	1.53E+18	NA		0	FALSE	FALSE	
522	521 Selama Libur Lebaran 20	FALSE	1	NA	5/15/2022 14:04	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
523	522 muka leleh setelah 6 jar	FALSE	3	NA	5/15/2022 13:46	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
524	523 Liat story orang	FALSE	0	NA	5/15/2022 13:46	TRUE	NA	1.53E+18	NA		0	FALSE	FALSE	
525	524 Pantai Selatan Kigomuk	FALSE	0	NA	5/15/2022 12:26	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
526	525 Pantai Kukup https://t.c	FALSE	0	NA	5/15/2022 10:57	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
527	526 rekomendasi pantai Jog	FALSE	0	NA	5/15/2022 10:54	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
528	527 ke jogja mau ke pantai t	FALSE	0	NA	5/15/2022 10:12	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
529	528 @girugameshug Malio!	FALSE	0	girugame	5/15/2022 8:32	TRUE	1.53E+18	1.53E+18	1.23E+18		0	FALSE	FALSE	
530	529 raono apa2ne dibanding	FALSE	0	NA	5/15/2022 8:18	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
531	530 @jogja_Uncover @SARS	FALSE	0	jogja_Unc	5/15/2022 8:16	FALSE	1.53E+18	1.91E+09		0	FALSE	FALSE		
532	531 Detik detik Tim SAR	FALSE	121	NA	5/15/2022 7:27	TRUE	NA	1.53E+18	NA		47	FALSE	FALSE	
533	532 glak, jam 12 di pantai jo	FALSE	0	NA	5/15/2022 7:07	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
534	533 RT @ipoksamidub: di pa	FALSE	0	NA	5/15/2022 6:57	FALSE	NA	1.53E+18	NA		5	TRUE	FALSE	
535	534 @shuuaaangg kalo w	FALSE	0	shuuaaar	5/15/2022 6:36	FALSE	1.53E+18	1.53E+18	1.5E+18		0	FALSE	FALSE	
536	535 Ohh main ke pantai di gi	FALSE	0	NA	5/15/2022 5:43	TRUE	NA	1.53E+18	NA		0	FALSE	FALSE	
537	536 Saatnya meromantisasi	FALSE	0	NA	5/15/2022 5:28	FALSE	NA	1.53E+18	NA		0	FALSE	FALSE	
538	537 Anus lalu lintas menuju	FALSE	7	NA	5/15/2022 4:31	TRUE	NA	1.53E+18	NA		0	FALSE	FALSE	
539	538 @bayuandikaprtm, Gim	FALSE	0	bayuandik	5/15/2022 4:20	FALSE	1.52E+18	1.53E+18	3.79E+08		0	FALSE	FALSE	
540	539 @BLINKmf pantai jogja i	FALSE	0	BLINKmf	5/15/2022 4:15	FALSE	1.53E+18	1.53E+18	1E+18		0	FALSE	FALSE	
541	540 Gua tuh pengen banget	FALSE	0	NA	5/15/2022 4:15	TRUE	NA	1.53E+18	NA		0	FALSE	FALSE	

Gambar 4. Hasil Web Crawling Data Twitter

Data yang akan digunakan untuk analisis selanjutnya yaitu data teks yang terdapat pada kolom kedua. Data tersebut hanya berisi cuitan atau opini yang ditulis masyarakat dalam jejaring sosial *twitter*. Karena data teks tersebut masih dalam keadaan yang tidak terstruktur, akan dilakukan tahapan selanjutnya yaitu *text preprocessing*.

Tuliskan analisis yang dilakukan. Gunakan TNR 11, spasi 1, *before and after spacing* 6, dan paragraph rata kanan-kiri.

3.3. Text Preprocessing

Pada tahapan ini akan dilakukan transformasi data teks yang tidak terstruktur menjadi data teks yang terstruktur. Data teks tersebut akan diolah dengan menggunakan metode *text mining* untuk memperoleh informasi yang baik sehingga mudah untuk dianalisis lebih lanjut. Tahapan *text preprocessing* pada penelitian ini yaitu *remove punctuation, case folding, spelling normalization, filtering, dan tokenizing*.

3.3.1. Remove Punctuation

Pada tahapan *remove punctuation* akan dilakukan penghapusan karakter, tanda baca, dan simbol-simbol yang tidak memberikan pengaruh terhadap proses analisis sentimen, seperti menghapus tanda titik (.), koma (,), hastag (#), mention (@), *link URL*, dan karakter selain huruf dan angka.

Tabel 2 Proses Remove Punctuation

Sebelum	Sesudah
@kangsukron_ Dekat YIA,, pantai gkagah indah kulon progo	Dekat YIA pantai glagah indah kulon progo.

3.3.2. Case Folding

Pada tahapan *remove punctuation* akan dilakukan penghapusan karakter, tanda baca, dan simbol-simbol yang tidak memberikan pengaruh terhadap proses analisis sentimen, seperti menghapus tanda titik (.), koma (,), hastag (#), mention (@), *link URL*, dan karakter selain huruf dan angka.

Tabel 3 Proses Case Folding

Sebelum	Sesudah
Pantai Bugel Pesona Wisata Pantai Kulon Progo	pantai bugel pesona wisata pantai kulon progo

Hasil visualisasi *wordcloud* pada **Gambar 5**, menunjukkan bahwa kata yang paling sering muncul pada cuitan *twitter* adalah kata “slii” diikuti dengan kata “wisata”, “banget”, “siung”, dan “wajib”. Sehingga dapat diartikan bahwa pantai slii menjadi wisata pantai yang sedang *trend* di DI Yogyakarta. Adapun beberapa pantai yang sering disebutkan oleh masyarakat pada jejaring sosial *twitter* yaitu pantai siung, pantai midodaren, pantai kukup, pantai drini, pantai glagah, pantai parangtritis, dan pantai indrayanti. Untuk memastikan apakah opini masyarakat tentang pantai tersebut merupakan opini positif, netral, ataupun negatif maka dilakukan pelabelan data emosi menggunakan analisis sentimen dengan metode *lexicon based*.

3.5. Pelabelan Data

Kumpulan kata yang telah melalui tahap *text preprocessing* perlu diberikan label agar kumpulan kata tersebut bisa diolah dan dihitung. Perhitungan skor akan dilakukan pada jumlah yang terdeteksi berdasarkan kamus kata *lexicon*. Adapun rumus perhitungan skor sentimen yaitu dengan mengurangi jumlah kata positif dengan jumlah kata negatif.

$$\text{Skor} = (\text{jumlah kata positif}) - (\text{jumlah kata negatif}) \quad (4)$$

Contoh hasil pelabelan data ditunjukkan pada **Tabel 7**, dimana kata yang berhasil terdeteksi kamus kata akan memiliki skor dan dijumlahkan sesuai dengan rumus perhitungan skor.

Teks	Kata Positif	Kata Negatif
pantai gunungkidul sangat bagus bersih berbeda kota panas	bagus bersih	panas
Jumlah	2	1
Perhitungan	Skor = 2 – 1 Skor = 1	

Setelah melakukan perhitungan skor sentimen berdasarkan kamus *lexicon*, tahap selanjutnya yaitu melihat hasil pelabelan sentimen berdasarkan skor yang diperoleh. Pelabelan sentimen dilakukan kedalam tiga kelas sentimen yaitu sentimen negatif, netral, dan positif. Dimana kelas positif memiliki skor > 1, kelas negatif memiliki skor < 0, dan kelas netral memiliki skor = 0. Label netral diperoleh karena memiliki dua kemungkinan antara lain :

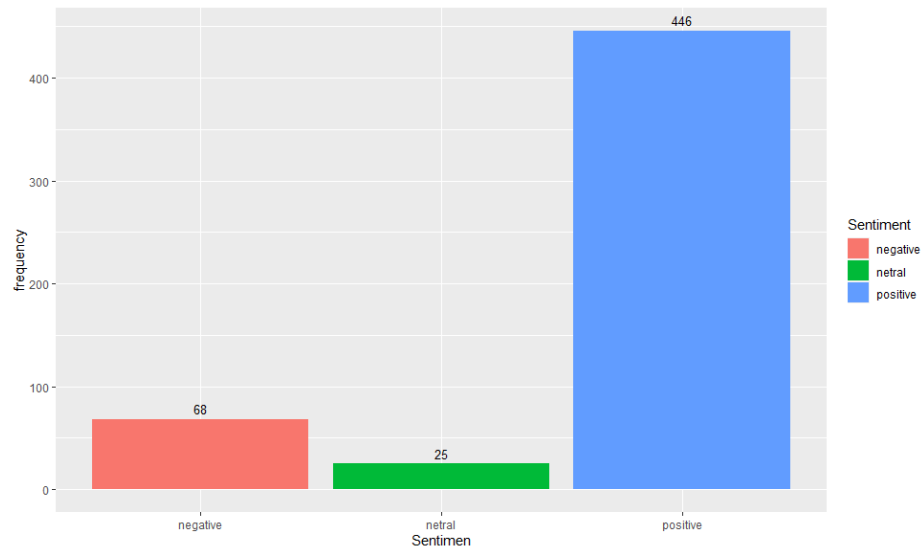
- Terdapat kata-kata baku yang tidak terdeteksi pada kamus *slangwords*.
- Tidak terdapat kata sentimen pada teks yang terdeteksi pada kamus *lexicon*.
- Jumlah skor kata positif sama dengan jumlah skor negatif atau 0.

Setelah semua tahapan analisis sentimen dilakukan, diperoleh hasil pelabelan kelas sentimen pada 540 data *tweet* tentang ‘wisata pantai jogja’ adalah sebagai berikut :

Sentimen	Frekuensi
Negatif	68
Netral	25
Positif	446

Dari hasil analisis sentimen yang diperoleh pada **Tabel 8**, dapat dilihat bahwa data teks telah masuk kedalam beberapa sentimen. Diantaranya adalah sentimen negatif dengan jumlah sekitar 68 data teks, sentimen netral dengan jumlah sekitar 25 data teks, dan sentimen positif dengan jumlah 446 data teks. Sehingga dapat disimpulkan bahwa mayoritas masyarakat yang menulis cuitan atau *tweet* dengan *keyword* ‘wisata pantai jogja’

menunjukkan opini yang positif terhadap wisata pantai yang ada di DI Yogyakarta. Adapun visualisasi yang diperoleh dari hasil pelabelan kelas sentimen ditunjukkan pada **Gambar 6**.



Gambar 6. Visualisasi hasil Pelabelan Kelas Sentimen

4. Kesimpulan

Berdasarkan hasil analisis sentimen yang telah dilakukan, maka didapatkan kesimpulan sebagai berikut :

1. 5 kata teratas yang paling sering muncul pada cuitan jejaring sosial *twitter* tentang Wisata Pantai di DI Yogyakarta yaitu 'slii', 'wisata', 'banget', 'siung, dan 'wajib'.
2. Berdasarkan hasil visualisasi *wordcloud* dapat ditarik kesimpulan bahwa pantai slii sangat menjadi wisata pantai yang sedang *trend* di DI Yogyakarta. Adapun beberapa pantai yang sering disebutkan oleh masyarakat pada media sosial *twitter* yaitu pantai siung, pantai midodaren, pantai kukup, pantai drini, pantai glagah, pantai parangtritis, dan pantai indrayanti.
3. Dari hasil analisis sentimen diperoleh bahwa mayoritas masyarakat yang menulis cuitan atau *tweet* dengan *keyword* 'wisata pantai jogja' menunjukkan opini yang positif terhadap wisata pantai yang ada di DI Yogyakarta.

5. Daftar Pustaka

- [1] R. Minazzi, "Social Media Marketing," in *Tourism and Hospitality*, Switzerland, Springer International Publishing, 2015, pp. 127-135.
- [2] S. D. Prihadi, "Berapa Jumlah Pengguna Facebook dan Twitter di Indonesia?," CNN Indonesia, 27 March 2015. [Online]. Available: <https://www.cnnindonesia.com/teknologi/20150327061134-185-42245/berapa-jumlah-pengguna-facebook-dan-twitter-di-indonesia>. [Accessed 21 May 2021].
- [3] L. Zhang and B. Liu, "A Survey of Opinion Mining and Sentiment Analysis," in *Mining Text Data*, Boston, Springer US, 2012, pp. 415-463.
- [4] D. Puntodi, *Menciptakan Penjualan Melalui Media Sosial*, Jakarta: PT Alex Media, 2011.

- [5] D. Zarella, *The Social Media Marketing Book*, Jakarta: Serambi, 2011.
- [6] R. G. Santosa, *Statistik*, Yogyakarta: Andi Offset, 2004.
- [7] Sugiyono, *Metode Penelitian*, Bandung: Alfabeta, 2004.
- [8] J. Han and M. Kamber, *Data Mining: Concepts and Techniques Third Edition*, Waltham: Morgan Kaufmann, 2012.
- [9] A. I and F. Hakim, "Implementasi Teks Mining pada Mesin pencarian Twitter Untuk Menganalisis Topik-Topik Terkait KPK dan Jokowi," in *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika UMS*, Surakarta, UMS, 2015, pp. 579-581.