

ANALISIS SENTIMEN BERBASIS ASPEK PADA TEMPAT WISATA DI KALIMANTAN TENGAH DENGAN MEMANFAATKAN MODEL *DEEP LEARNING*

Bima Faru Rochkim Putra¹⁾, Viktor H Pranatawijaya²⁾, Putu Bagus A.A.P³⁾

¹⁾²⁾³⁾Teknik Informatika, Fakultas Teknik, Universitas Palangka Raya
Jl. Hendrik Timang Kampus Tunjung Nyaho, Palangka Raya

¹⁾ bimafaru04@mhs.eng.upr.ac.id

²⁾ viktorhp@it.upr.ac.id

³⁾ putubagus@it.upr.ac.id

Abstrak

Penyampaian informasi mengenai tempat wisata tentu perlu media untuk menyampaikan. Pada perkembangan era digital saat ini kita dimudahkan untuk mencari atau melihat bagaimana pandangan dari pengunjung atau wisatawan dari lokasi wisata yang ingin kita kunjungi dengan melihat ulasan pada Google Maps. Memungkinkan indentifikasi masalah dan pengembangan suatu lokasi wisata tersebut. Dengan adanya Teknik Aspect Based Sentiment Analysis atau ABSA, guna membantu kita bagaimana pandangan pengunjung lain terhadap lokasi wisata tersebut dengan mudah dipahami.

Tujuan utama dari penelitian ini untuk menganalisis sentimen berdasarkan aspek-aspek yang diungkap dalam ulasan pengunjung yang telah berkunjung ke suatu lokasi wisata yang berada di ruang lingkup Provinsi Kalimantan Tengah pada Google Maps. Dengan mencari pola sentimen yang akurat dan mengidentifikasi aspek-aspek. Dalam pencarian aspek-aspek menggunakan metode *Named Entity Recognition* atau biasa disebut dengan NER dengan dibantu dengan teknik pemodelan topik *Latent Dirichlet Allocation* (LDA) yang berguna untuk mengidentifikasi topik-topik relevan untuk dimasukkan kedalam kamus NER. Pelabelan sentimen dilakukan dengan menggunakan BERT, sementara evaluasi sentimen dan aspek dilakukan dengan menggunakan model BERT, LSTM, dan CNN.

Penelitian ini menganalisis pola sentimen dan aspek dalam ulasan lokasi wisata di Kalimantan Tengah menggunakan model BERT, LSTM, CNN. Hasil menunjukkan bahwa model BERT lebih unggul dengan mencapai akurasi sentimen (98.02%) kemudian pada aspek *Environment* (98.21%), *Historical* (97.55%), *Experience* (99%), *Attraction* (97.81%), *Location* (98%), *Nature* (95.06%).

Kata kunci: Kalimantan Tengah, *Aspect Based Sentimen Analysis* (ABSA), NER, Sentimen, Akurasi

Abstract

Conveying information about tourist attractions certainly requires media to convey it. In today's digital era, it is easy for us to search or see what visitors or tourists view the tourist location we want to visit by looking at reviews on Google Maps. Enables identification of problems and development of a tourist location. With the Aspect Based Sentiment Analysis or ABSA technique, it can help us understand how other visitors view the tourist location easily.

The main aim of this research is to analyze sentiment based on aspects expressed in reviews of visitors who have visited a tourist location within the scope of Central Kalimantan Province on Google Maps. By looking for accurate sentiment patterns and identifying aspects. In searching for aspects using the Named Entity Recognition method or commonly called NER, assisted by the Latent Dirichlet Allocation (LDA) topic modeling technique which is useful for identifying relevant topics to be included in the NER dictionary. Sentiment labeling is carried out using BERT, while sentiment and aspect evaluation is carried out using BERT, LSTM and CNN models.

This research analyzes sentiment patterns and aspects in reviews of tourist locations in

Central Kalimantan using BERT, LSTM, CNN models. The results show that the BERT model is superior by achieving sentiment accuracy (98.02%) then in the aspects of Environment (98.21%), Historical (97.55%), Experience (99%), Attraction (97.81%), Location (98%), Nature (95.06%).

Keywords: *Central Kalimantan, Aspect-Based Sentiment Analysis (ABSA), NER, Sentiment, Accuracy*

1. PENDAHULUAN

Tempat wisata merupakan salah satu tempat yang dimana sering dikunjungi orang untuk mencari kesenangan diluar aktivitas sehari – hari. Tempat wisata biasanya banyak dikunjungi oleh wisatawan ketika hari libur untuk menghabiskan waktu bersama orang – orang terdekat. Turis menggunakan website untuk mencari informasi terbaru tentang destinasi wisata, membaca ulasan tentang paket wisata secara online [1].

Data ulasan Google Maps menunjukkan bahwa destinasi wisata di Kalimantan Tengah memiliki beragam ulasan pada setiap lokasinya. Total ulasan mencapai lebih dari 10 ribu untuk berbagai destinasi wisata di Kalimantan Tengah. Dalam memahami pandangan wisatawan, penulis menggunakan teknik Aspect Based Sentiment Analysis (ABSA) dimana teknik ini merupakan Teknik sentiment analis untuk mendekati sentiment dari suatu teks berdasarkan aspek yang ada. pada level aspek memiliki tujuan untuk menentukan sentimen positif, negatif ataupun netral berdasarkan atribut dari suatu entitas [2].

Beberapa penelitian terbaru telah memfokuskan pada analisis sentimen berbasis aspek ulasan atraksi wisata menggunakan LSTM. Dalam penelitian oleh [3] mereka memperoleh hasil pengujian klasifikasi aspek dan klasifikasi sentimen secara berurutan memperoleh 90,63% dari presisi rata-rata dan 70,4% dari skor f1.

Penelitian terbaru telah mengeksplorasi analisis sentimen dan klasifikasi aspek dalam berbagai konteks. Misalnya, dalam penelitian [4], analisis sentimen berbasis aspek di Taman Nasional Bromo Tengger Semeru, Indonesia berdasarkan ulasan pengguna google maps. Dengan menggunakan aspek atraksi, fasilitas, akses, dan harga. Model klasifikasi sentimen yang digunakan adalah model machine learning yang terdiri dari SVM, Complement Naïve Bayes, Logistic Regression, dan transfer learning dari pre-trained BERT, IndoBERT dan mBERT. Berdasarkan hasil eksperimen, transfer learning dari model IndoBERT mencapai performa terbaik dengan akurasi dan F1-Score masing-masing sebesar 91.48% dan 71.56%.

Penelitian ini menggunakan ulasan wisatawan di Google Maps pada lokasi wisata di Kalimantan Tengah sebagai sumber data. Teknik pemodelan topik untuk mengidentifikasi topik-topik relevan sebagai aspek, dan juga analisis sentimen dilakukan dengan memanfaatkan model Deep Learning yaitu BERT, LSTM, dan CNN. Tujuan utama penelitian ini adalah menganalisis sentimen berdasarkan aspek-aspek pada ulasan Google Maps dan membandingkan kinerja BERT, LSTM, dan juga CNN.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen Berbasis Aspek

Analisis Sentimen Berbasis Aspek merupakan suatu teknik analisis sentimen yang digunakan untuk mendeteksi sentimen dari suatu teks berdasarkan aspek yang ada pada teks tersebut [5]. Analisis sentimen berbasis aspek melakukan ekstraksi aspek dan sentimen dalam sebuah kalimat sehingga jumlah aspek dalam kalimat dapat ditentukan dan polaritas sentimen dari setiap aspek dapat ditentukan [6].

2.2. Scraping Data

Scraping adalah salah satu teknik untuk mengumpulkan data dari sebuah situs web secara otomatis, kemudian data yang diperoleh dapat disimpan dalam format tertentu, data yang diperoleh melalui scraping tidak memiliki label yang menunjukkan apakah data tersebut positif atau negatif [7].

2.3. Text Preprocessing

Text preprocessing merupakan tahap awal dalam pengolahan teks sebelum dianalisis atau diproses lebih lanjut, dan memiliki peran penting dalam berbagai aplikasi seperti text mining, analisis sentimen, dan ekstraksi informasi dari dokumen teks yang tidak terstruktur, untuk menghindari kesalahan dalam ekstraksi fitur atau atribut yang dapat secara signifikan mengurangi kinerja analisis sentimen [8].

2.3.1 Data Cleansing

Tahapan pertama adalah pembersihan teks dari karakter yang tidak relevan, seperti n ama pengguna, spasi kosong, karakter khusus, dan URL, Tujuannya agar dokumen yang akan diolah menjadi lebih bersih dan relevan. Salah satu cara untuk membersihkan dokumen adalah dengan menghapus simbol dan tanda baca yang tidak berkontribusi dalam analisis teks [9].

2.3.2 Normalization

Normalisasi data adalah proses membuat beberapa variabel memiliki rentang nilai yang sama, tidak ada yang terlalu besar maupun terlalu kecil sehingga dapat membuat analisis statistik menjadi lebih mudah [10].

2.3.3 Case Folding

Case folding mengonversi teks menjadi huruf kecil sehingga data memiliki format yang sama, serta menghapus tanda baca yang tidak diperhitungkan[11].

2.3.4 Tokenizing

Tokenizing adalah proses mengubah dokumen teks menjadi serangkaian token atau unit-token, seperti kata – kata atau frasa, yang lebih mudah diolah oleh komputer [12].

2.3.5 Stop Word

Stopword adalah kata-kata yang kurang berarti dalam analisis dan mengurangi kualitas data. Penghilangan stopwords mempercepat proses analisis [7].

2.3.6 Lemmatization

Lemmatization salah satu teknik preprocessing yang umum digunakan dalam Natural Language Processing. Proses ini bertujuan untuk mengubah kata dalam bentuk infleksional sehingga dapat dikenali sebagai satu elemen tunggal [5].

2.4 Topic Modelling Lattent Dirichlet Allocation (LDA)

Topic modeling adalah proses analisis teks untuk mengelompokkan dokumen berdasarkan topik tertentu. Tujuannya adalah membangun model topik abstrak dari dokumen, memberikan tinjauan berdasarkan topik, dan mengelompokkan dokumen. Sebuah topik terdiri dari sejumlah kata yang merepresentasikan dokumen, dan dokumen dapat memiliki beberapa topik dengan peluang masing-masing. Hasilnya adalah kumpulan topik yang sering muncul dalam dokumen berdasarkan pola tertentu [13].

2.5 Named Entity Recognition (NER)

NER (Named Entity Recognition) adalah teknik untuk mengenali kata benda khusus seperti nama orang, tempat, dan organisasi. Awalnya, NER digunakan dalam konferensi pemahaman pesan (MUC) yang berfokus pada ekstraksi informasi dari data tidak terstruktur. Kini, NER sering digunakan di bidang NLP (Natural Language Processing) [14].

2.6 Long Short-Term Memory (LSTM)

LSTM merupakan salah satu metode yang banyak digunakan untuk mengolah kata pada NLP (Neural Language Processing). LSTM banyak digunakan pada data yang bersifat kata atau text base karena sifatnya yang sequential [15].

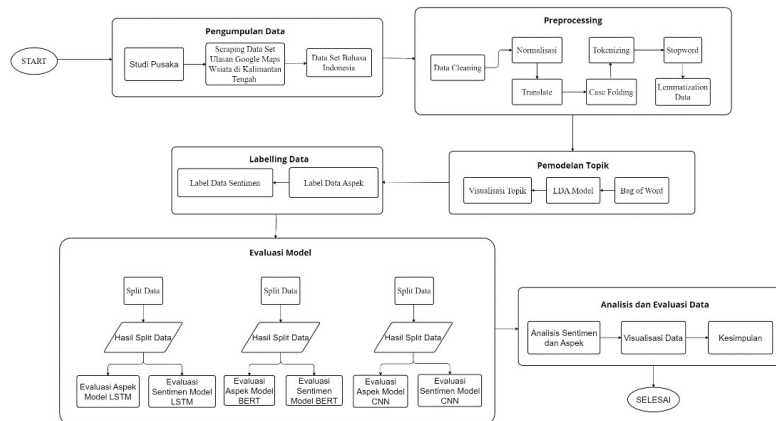
2.7 Bidirectional Encoder Representations from Transformers (BERT)

BERT adalah model bahasa yang menggunakan mekanisme encoder pada Transformer. Istilah "bidirectional" menjelaskan bahwa pada fase pretraining, encoder pada BERT membaca seluruh urutan kata masukan secara dua arah, berbeda dengan model sebelumnya yang membaca secara searah (kiri ke kanan atau kanan ke kiri saja)[16].

2.8 Convolutional Neural Network (CNN)

Convolutional Neural Network atau CNN merupakan kumpulan jaringan saraf yang memberikan informasi yang cukup untuk bidang pemrosesan urutan data seperti teks dengan menggunakan konvolusi pada kalimat, paragraph, atau seluruh dokumen teks [17].

3. METODE PENELITIAN



Gambar 1. Alur Metode Penelitian

Pada gambar menggambarkan visualisasi dari alur metode penelitian yang digunakan penulis dalam studi ini. Alur metode penelitian tersebut disusun untuk mengilustrasikan hubungan antara metode-metode dan variabel yang digunakan dalam penelitian. Terdapat enam tahapan dalam alur metode tersebut, yaitu pengumpulan data, preprocessing data, pemodelan topik, pelabelan data, analisis dan visualisasi data. Pada bagian berikutnya, akan dijelaskan lebih lanjut mengenai metode-metode yang tercantum dalam gambar sebelumnya.

4. PEMBAHASAN

4.1.1 Scraping Data

Data ulasan diambil dari rentang waktu dari tahun 2017 sampai dengan tahun 2024. Jumlah total ulasan yang diambil dari hasil pengumpulan dataset ialah 19.315 ulasan dari 75 lokasi wisata yang ada di 14 Kabupaten/Kota di Provinsi Kalimantan Tengah. Contoh hasil *scraping* data ulasan dapat dilihat pada Tabel 1.

Tabel 1. Hasil Scraping Ulasan Lokasi Wisata

No	Kabupaten	Lokasi	Pengguna	Tahun	Ulasan
2	Kota Palangka Raya	bukit baranahu	PRATAMA RAHMATULLAH	setahun lalu	Lokasi buat pendakian cukup terjal Pamandangan di atas tower sangat bagus Tapi sayang jalannya banyak yang licin dan terjal dan diatas nggak ada tempat buat berlindung kalau hujan nya kencang ...
		Barito Selatan	Taman Iring Witu	Sapdianto	11 bulan lalu
3	Barito Timur	Museum Lewu Hante	Kornelius Agustinus	2 tahun lalu	Batang Pasir Panas, ini pas terletak perbatasan antara Kalteng dan Kalsel di Pasir Panas, ini adalah bangunan

					baru, jadi bukan bangunan asli yang usianya ratusan tahun. ...
4	Gunung Mas	Taman Kota Kuala Kurun	Adi Suci Guntoro Office	11 bulan lalu	Bisa jadi alternatif hiburan permainan anak di Kuala kurun, pilihannya banyak dengan harga yang terjangkau 👍 ...

4.1.2 Preprocessing Data

Setelah memperoleh dataset, dilakukan tahapan preprocessing untuk membersihkan dan memproses data teks. Tahapan preprocessing dapat dilihat pada Tabel 3.

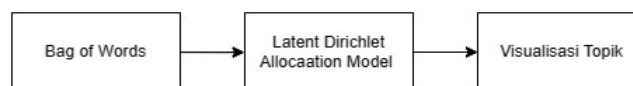
Tabel 3. Hasil Preprocessing Data

No	Pemrosesan Awal	Data Mentah	Pembersihan Data
1	Data Cleansing	Lokasi buat pendakian cukup terjal Pamandangan di atas tower sangat bagus Tapi sayang jalannya banyak yang licin dan terjal dan diatas nggak ada tempat buat berlindung kalau hujan nya kencang	Lokasi buat pendakian cukup terjal Pamandangan di atas tower sangat bagus Tapi sayang jalannya banyak yang licin dan terjal dan diatas nggak ada tempat buat berlindung kalau hujan nya kencang
2	Normalization	Lokasi buat pendakian cukup terjal Pamandangan di atas tower sangat bagus Tapi sayang jalannya banyak yang licin dan terjal dan diatas nggak ada tempat buat berlindung kalau hujan nya kencang	Lokasi buat pendakian cukup terjal Pamandangan di atas tower sangat bagus Tapi sayang jalannya banyak yang licin dan terjal dan diatas nggak ada tempat buat berlindung kalau hujan nya kencang
3	translate	Lokasi buat pendakian cukup terjal Pamandangan di atas tower sangat bagus Tapi sayang jalannya banyak yang licin dan terjal dan diatas nggak ada tempat buat berlindung kalau hujan nya kencang	The location for climbing is quite steep. The view at the top of the tower is very good. But unfortunately, many of the roads are slippery and steep and there is no place at the top to take shelter if it rains hard.
4	Case Folding	The location for climbing is quite steep. The view at the top of the tower is very good. But unfortunately, many of the roads are slippery and steep and there is no place at the top to take shelter if it rains hard.	the location for climbing is quite steep. the view at the top of the tower is very good. but unfortunately, many of the roads are slippery and steep and there is no place at the top to take shelter if it rains hard.
5	Tokenizing	the location for climbing is quite steep. the view at the top of the tower is very good. but unfortunately, many of the roads are slippery and steep and there is no place at the top to take shelter if it rains hard.	['the', 'location', 'for', 'climbing', 'is', 'quite', 'steep', ',', 'the', 'view', 'at', 'the', 'top', 'of', 'the', 'tower', 'is', 'very', 'good', ',', 'but', 'unfortunately', ',', 'many', 'of', 'the', 'roads', 'are', 'slippery', 'and', 'steep', 'and', 'there', 'is', 'no', 'place', 'at', 'the', 'top', 'to', 'take', 'shelter', 'if', 'it', 'rains', 'hard', '.']
6	Stopword Removal	['the', 'location', 'for', 'climbing', 'is', 'quite', 'steep', ',', 'the', 'view', 'at', 'the', 'top', 'of', 'the', 'tower', 'is', 'very', 'good', ',', 'but', 'unfortunately', ',', 'many', 'of', 'the', 'roads', 'are', 'slippery', 'and', 'steep', 'and', 'there', 'is', 'no', 'place', 'at', 'the', 'top', 'to',	['location', 'climbing', 'quite', 'steep', ',', 'view', 'top', 'tower', 'good', ',', 'unfortunately', ',', 'many', 'roads', 'slippery', 'steep', 'place', 'top', 'take', 'shelter', 'rains', 'hard', '.']

		'take', 'shelter', 'if', 'it', 'rains', 'hard', ']	
7	Lemmatization	['location', 'climbing', 'quite', 'steep', ,', 'view', 'top', 'tower', 'good', ',', 'unfortunately', ',', 'many', 'roads', 'slippery', 'steep', 'place', 'top', 'take', 'shelter', 'rains', 'hard', ',']	location climb quite steep view top tower good unfortunately many road slippery steep place top take shelter rain hard

4.1.3 Pemodelan Topik

Pemrosesan dilakukan menggunakan model Latent Dirichlet Allocation (LDA) untuk pengelompokkan topik. Langkah-langkah pembuatan model LDA terlihat pada Gambar 2.



Gambar 2. Alur Topic Modelling

Dari tahap tersebut, topik yang dihasilkan mengandung kata-kata kunci yang dapat digunakan untuk menentukan entitas-entitas relevan pada setiap topik. Hasilnya adalah enam entitas utama: *environment*, *location*, *nature*, *experience*, *historical*, *attraction*. Masing-masing dengan kata-kata yang sering muncul dalam topiknya. Ini dapat dilihat dalam Tabel 4.

Tabel 4. Sebaran Kata Tiap Jumlah Topik

Topik	Aspek	Kata Per Topik
1	Environment	<i>cool, relax, clean, maintain, well, really, people, get, need, facility, even, many, unfortunately, garden, look, quiet, rubbish, building, atmosphere, hopefully.</i>
2	Location	<i>place, family, comfortable, still, lot, suitable, beach, great, pretty, spacious, cultural, preserve, holiday, night, friend, forest, recreation, fish, snack, religious</i>
3	Nature	<i>visit, child, take, interesting, stop, play, photo, natural, tour, destination, make, enjoy, water, learn, orangutan, add, playground, bring, break, nature</i>
4	Experience	<i>palace, fun, bad, amazing, experience, extraordinary, far, shame, village, blue, hang, become, give, damage, kotawaringin, dirty, put, district, strategic, expensive</i>
5	Historical	<i>historical, beautiful, nice, history, park, city, view, location, pangkalan, see, site, bun, long, road, historic, close, know, use, locate, excellent</i>
6	Attraction	<i>good, tourist, attraction, area, quite, tourism, also, spot, time, come, go, parking, afternoon, yellow, food, want, visitor, local, river, day</i>

4.1.4 Pelabelan Sentimen

Pelabelan sentimen dilakukan dengan BERT. Hasil BERT menunjukkan 9722 data (8259 positif, 1463 negatif). Contoh dapat dilihat pada Tabel 5.

Tabel 5. Pelabelan BERT

Teks	Sentimen	score
<i>location climb quite steep view top tower nice unfortunately many road slippery steep place take shelter rain strong</i>	Negative	0.949403
<i>quite tiring climb top worth view tangkile top view tower even hot lucky rain road slippery</i>	Negative	0.929863

*place great view extraordinary good camping nature night
atmosphere also pronounce* Positive 0.999771

beautiful place photo spot still clean trash Positive 0.999830

4.1.5 Split Data

Pada dataset ulasan lokasi wisata di Kalimantan Tengah, yang telah memiliki label entitas dan sentimen. Selanjutnya dibagi menjadi data latih dan data uji. Detailnya dapat dilihat dalam Tabel

6.

Tabel 6. Pembagian Data Ulasan

Metode	Klasifikasi	Data Latih (60%)	Data Uji (40%)
BERT	Positive	4955	3304
	Negative	878	585
	Total	5833	3889
Metode	Klasifikasi	Data Latih (70%)	Data Uji (30%)
LSTM	Positive	5781	2478
	Negative	1024	439
	Total	6805	2917
Metode	Klasifikasi	Data Latih (85%)	Data Uji (15%)
CNN	Positive	7020	1239
	Negative	1244	219
	Total	8264	1458

4.1.6 Evaluasi Sentimen

Pada model BERT yang digunakan ialah distilbert-base-uncased-finetuned-sst-2. Pada model BERT ini, objek tokenizer dibuat untuk memecah teks menjadi token – token kecil. Pada model LSTM ini menggunakan library nn dari torch untuk membangun model jaringan syaraf. Pada model CNN ini menggunakan library nn dari torch untuk membangun model jaringan syaraf. Kemudian menggunakan library sklearn model selection untuk membagi dataset menjadi set training dan validasinya nanti, kemudian menggunakan sklearn metrics untuk menghitung metrik evaluasi dari akurasi. Evaluasi model terdapat pada Tabel 7.

Tabel 7. Evaluasi Sentimen Model SVM

Dataset	Pelabelan	Sentimen			
		Accuracy	Precision	Recall	F1-Score
Ulasan Lokasi Wisata Kalimantan Tengah	BERT	98,02%	98,92%	98,71%	98,82%
	LSTM	84%	84%	100%	91%
	CNN	86,70%	86,76%	99,43%	92,66%

4.1.7 Evaluasi Aspek

Pada evaluasi aspek menggunakan model BERT, LSTM, dan CNN. Pada proses ini dihasilkan data hasil evaluasi untuk aspek dari dataset "Ulasan Lokasi Wisata Kalimantan Tengah". Berikut adalah data evaluasi terdapat pada Tabel 8, Tabel 9, dan Tabel 10.

Tabel 8. Evaluasi Aspek Model BERT

Dataset	Model	Aspek	Accuracy	Precision	Recall	F1-Score
	BERT	Environment	98,21%	99,19 %	98,39%	98,79%

Ulasan Lokasi Wisata Kalimantan Tengah	<i>Historical</i>	97,55 %	97,89 %	99,43%	98,65%
	<i>Experience</i>	99,00 %	100%	98,60%	99,29%
	<i>Attraction</i>	97,81 %	97,88 %	99,69%	98,78%
	<i>Location</i>	98,00 %	98,78%	98,98%	98,88%
	<i>Nature</i>	95,06 %	95,40%	98,64%	97,00%
	<i>ORDINAL</i>	100%	100%	100%	100%
	<i>ORG</i>	100%	100%	100%	100%
	<i>NORP</i>	100%	100%	100%	100%
	<i>DATE</i>	100%	100%	100%	100%
	<i>TIME</i>	87,50%	83,33%	100%	90,90%
	<i>PERSON</i>	100%	100%	100%	100%
	<i>GPE</i>	100%	100%	100%	100%
	<i>CARDINAL</i>	100%	0.000%	0.000%	0.000%

Tabel 9. Evaluasi Aspek Model LSTM

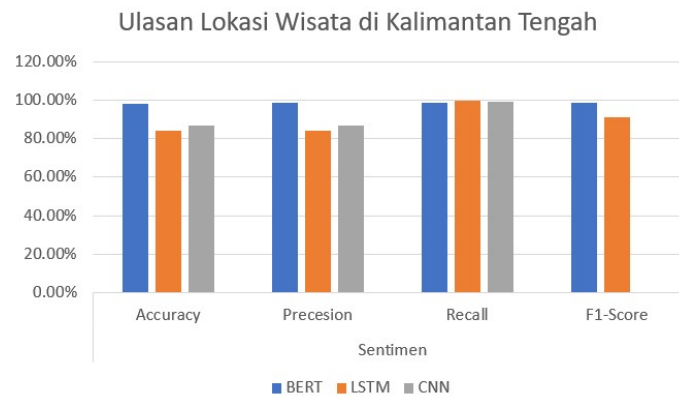
Dataset	Model	Aspek	Accuracy	Precision	Recall	F1-Score
Ulasan Lokasi Wisata Kalimantan Tengah	LSTM	<i>Environment</i>	77,18%	78,49%	96,05%	86,39%
		<i>Historical</i>	89,55%	89,55%	100%	94,48%
		<i>Experience</i>	68,87%	68,87%	100%	81,56%
		<i>Attraction</i>	88,88%	88,88%	100%	94,11%
		<i>Location</i>	89,85%	89,85%	100%	94,65%
		<i>Nature</i>	81,02%	81,02%	100%	89,51%
		<i>ORDINAL</i>	100%	100%	100%	100%
		<i>ORG</i>	80,00%	75,00%	100%	85,71%
		<i>NORP</i>	66,66%	66,66%	100%	80,00%
		<i>DATE</i>	83,33%	83,33%	100%	90,90%
		<i>TIME</i>	50,00%	50,00%	100%	66,66%
		<i>PERSON</i>	62,50%	62,50%	100%	76,92%
		<i>GPE</i>	100%	100%	100%	100%
<i>CARDINAL</i>	0.000%	0.000%	0.0%	0.000%		

Tabel 10. Evaluasi Aspek Model CNN

Dataset	Model	Aspek	Accuracy	Precision	Recall	F1-Score
Ulasan Lokasi Wisata Kalimantan Tengah	CNN	<i>Environment</i>	88,88%	87,85%	98,94%	93,06%
		<i>Historical</i>	90,06%	90,49%	99,22%	94,65%
		<i>Experience</i>	78,94%	77,77%	100%	87,50%
		<i>Attraction</i>	90,54%	90,45%	99,57%	94,80%
		<i>Location</i>	91,78%	91,87%	99,73%	95,64%
		<i>Nature</i>	84,67%	85,03%	98,18%	91,13%
		<i>ORDINAL</i>	100%	100%	100%	100%
		<i>ORG</i>	66,66%	66,66%	100%	80,00%
		<i>NORP</i>	100%	100%	100%	100%
		<i>DATE</i>	100%	100%	100%	100%
		<i>TIME</i>	66,66%	0%	0%	0%
		<i>PERSON</i>	50,00%	50,00%	100%	66,66%
		<i>GPE</i>	100%	100%	100%	100%
<i>CARDINAL</i>	0%	0%	0%	0%		

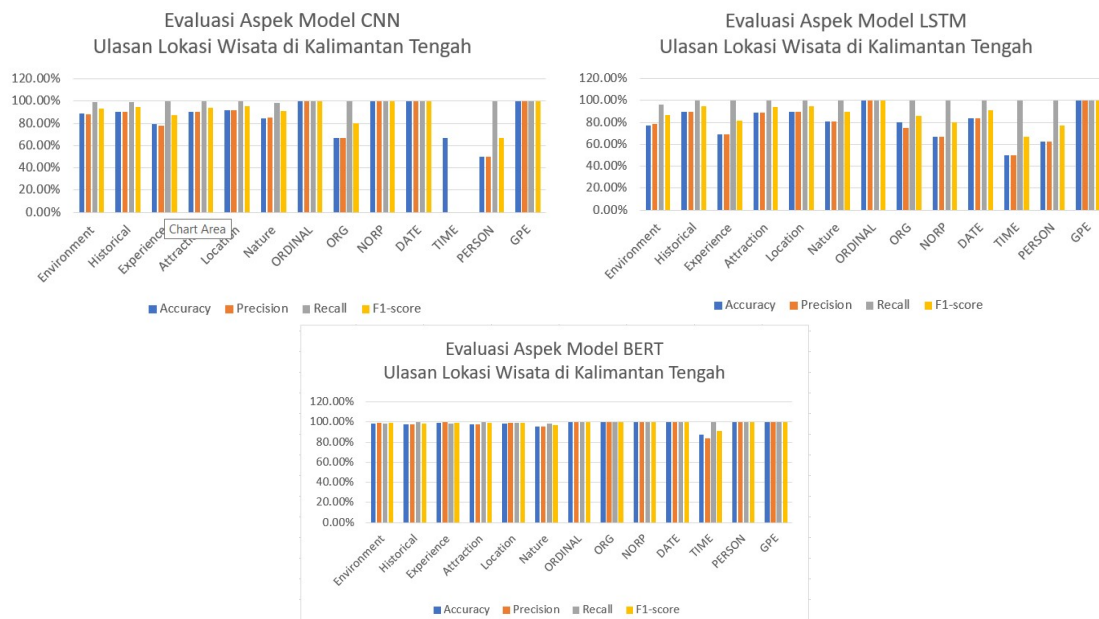
4.1.8 Analisis dan Visualisasi Data

Hasil evaluasi dengan menggunakan model BERT, LSTM, dan CNN. Hasil yang didapatkan di visualisasikan dengan menggunakan chart seperti pada Gambar 3.



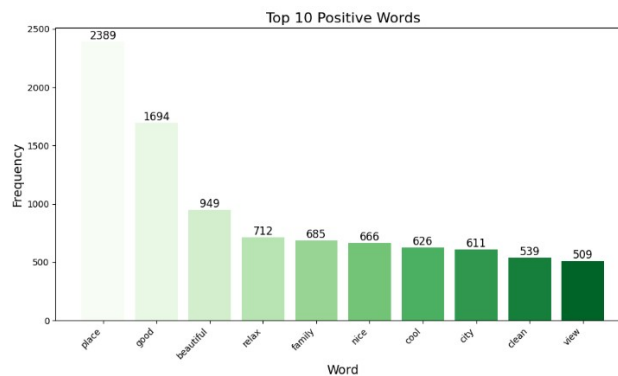
Gambar 3. Evaluasi Sentimen Model BERT, LSTM, dan CNN

Hasil evaluasi aspek dengan menggunakan model BERT, LSTM, dan CNN. Hasil yang didapatkan di visualisasikan dengan menggunakan chart seperti pada Gambar 4.

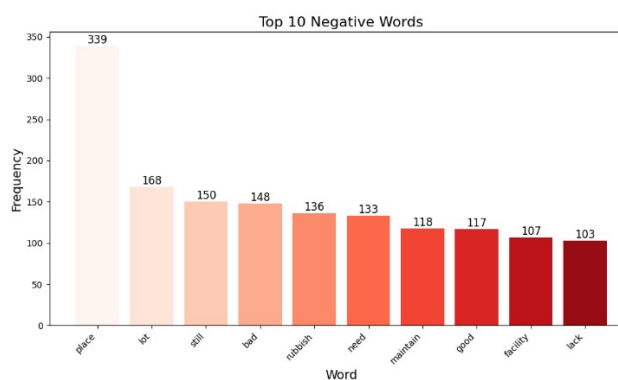


Gambar 4. Evaluasi Aspek Model SVM dan BERT

Selanjutnya pada analisis sentimen berbasis aspek ini didapat daftar kata – kata yang sering muncul pada ulasan yang merepresentasikan kata positif dan negatif. Dari ulasan tersebut dianalisis 10 kata positif dan negatif yang sering muncul dapat dilihat pada gambar dibawah.



Gambar 5. Kata Positif Yang Sering Muncul



Gambar 6. Kata Negatif Yang Sering Muncul

5. KESIMPULAN

Hasil penelitian ini berdasarkan data yang diambil pada Google Maps sejak tahun 2017 sampai dengan 2024. Dari data yang berjumlah 19315 data ulasan, kemudian dilakukan preproses data sehingga menghasilkan data berjumlah 9722 data ulasan. Dari data tersebut dilakukan proses Rule Base NER yang dimana Teknik ini berguna untuk mengidentifikasi dan mengklasifikasi entitas dari teks ulasan, dan dibantu dengan model LDA untuk mencari kata – kata yang dominan nantinya dijadikan topik dari entitas. Dimana model LDA yang telah dibuat mendapatkan enam topik dengan coherence score sebesar 0.429. Kemudian daftar kata yang ada didalam keenam topik tadi dimasukkan kedalam kamus untuk dijadikan entitas NER, sehingga dari proses tersebut mendapatkan lima belas entitas aspek yaitu location, historical, attraction, environment, nature, experience, person, time, date, org, norp, ordinal, gpe, cardinal. Ulasan dianalisis menggunakan metode BERT untuk menentukan sentimen positif dan negatif. Hasilnya menunjukkan bahwa terdapat sebanyak 9722 data, dengan 8259 data positif dan 1463 Data negatif.

Dari hasil evaluasi sentimen menggunakan model BERT berhasil mendapatkan nilai akurasi sebesar 98,02%, presisi sebesar 98,92%, recall sebesar 98,71%, dan f1-score sebesar 98,82%. Kemudian yang kedua menggunakan model LSTM mendapatkan akurasi 84%, presisi 84%, recall 100%, dan f1-score 91%. Selanjutnya yang terakhir menggunakan model CNN dengan mendapatkan akurasi sebesar 86,70%, presisi sebesar 86,76%, recall 99,43%, dan f1 -score 92,66%.

DAFTAR PUSTAKA

[1] C. L. H. S. Nivak, “Sentiment Analysis of Ten Tourism Destinations in Indonesia,”

- Journal of Indonesian Tourism and Development Studies*, vol. 9, no. 3, pp. 162–168, Sep. 2021, doi: 10.21776/ub.jitode.2021.009.03.03.
- [2] H. Chyntia Morama, D. E. Ratnawati, and I. Arwani, “Analisis Sentimen berbasis Aspek terhadap Ulasan Hotel Tentrem Yogyakarta menggunakan Algoritma Random Forest Classifier,” 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [3] D. I. Anggraeni, P. D. Rizki, M. B. Setiawan, and A. B. Handayani, “Aspect-Based Sentiment Analysis for Indonesian Tourist Attraction Reviews Using Bidirectional Long Short-Term Memory,” 2023.
- [4] C. A. Bahri and L. H. Suadaa, “Aspect-Based Sentiment Analysis in Bromo Tengger Semeru National Park Indonesia Based on Google Maps User Reviews,” *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 17, no. 1, p. 79, Feb. 2023, doi: 10.22146/ijccs.77354.
- [5] Novirianto, “ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI MYBLUEBIRD DENGAN IMPLEMENTASI N-GRAM DAN ALGORITMA LOGISTIC REGRESSION,” 2023. Accessed: Mar. 20, 2024. [Online]. Available: <https://repository.uinjkt.ac.id/dspace/bitstream/123456789/76656/1/IQBAL%20FARIZ%20NOVIRIANTO-FST.pdf>
- [6] Syaiful Imron, E. I. Setiawan, Joan Santoso, and Mauridhi Hery Purnomo, “Aspect Based Sentiment Analysis Marketplace Product Reviews Using BERT, LSTM, and CNN,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 3, pp. 586–591, Jun. 2023, doi: 10.29207/resti.v7i3.4751.
- [7] A. Komarudin, A. Meutia Hilda, and C. Author, “Analisis Sentimen Ulasan Aplikasi Identitas Kependudukan Digital Pada Play Store Menggunakan Metode Naïve Bayes,” 2024. [Online]. Available: <http://jurnal.bsi.ac.id/index.php/co-science>
- [8] I. Maulana, W. Apriandari, and A. Pambudi, “ANALISIS SENTIMEN BERBASIS ASPEK TERHADAP ULASAN APLIKASI MYPERTAMINA MENGGUNAKAN SUPPORT VECTOR MACHINE,” 2023. [Online]. Available: <http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index>
- [9] C. Villavicencio, J. J. Macrohon, X. A. Inbaraj, J. H. Jeng, and J. G. Hsieh, “Twitter sentiment analysis towards covid-19 vaccines in the Philippines using naïve bayes,” *Information (Switzerland)*, vol. 12, no. 5, May 2021, doi: 10.3390/info12050204.
- [10] M. R. Kusnaldi, T. Gulo, and S. Aripin, “Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal,” *Journal of Computer System and Informatics (JoSYC)*, vol. 3, no. 4, pp. 330–338, Sep. 2022, doi: 10.47065/josyc.v3i4.2112.
- [11] M. S. Nor Hasliza, Z. A. Alya Syahirah, Z. Yaacod, M. M. az Mohd Ali, and Z. Kun, “Twitter Sentiment Analysis on Meta: A Lexicon-Based Analysis using Rapidminer,” *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 6, pp. 259–270, Jun. 2023, doi: 10.17762/ijritcc.v11i6.7561.
- [12] A. Hidayat, A. Yani, P. Studi Sistem Informasi, and S. Mahakarya, “MEMBANGUN WEBSITE SMA PGRI GUNUNG RAYA RANAU MENGGUNAKAN PHP DAN MYSQL,” 2019.
- [13] W. A. Nawang Sari and H. Dwi Purnomo, “TOPIC MODELING USING THE LATENT DIRICHLET ALLOCATION METHOD ON WIKIPEDIA PANDEMIC COVID-19 DATA IN INDONESIA,” *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 5, pp. 1223–1230, Oct. 2022, doi: 10.20884/1.jutif.2022.3.5.321.
- [14] D. S. Rachmad, “Review Named Entity Recognition dengan Menggunakan Machine Learning,” *Jurnal Sains dan Informatika*, vol. 6, no. 1, pp. 28–33, Jun. 2020, doi: 10.34128/jsi.v6i1.204.
- [15] A. A. Mudding and Arifin A Abd Karim, “ANALISIS SENTIMEN MENGGUNAKAN ALGORITMA LSTM PADA MEDIA SOSIAL,” *Jurnal Publikasi Ilmu Komputer dan Multimedia*, vol. 1, no. 3, pp. 181–187, Sep. 2022, doi: 10.55606/jupikom.v1i3.517.

- [16] A. N. Azhar and M. L. Khodra, "Fine-tuning Pretrained Multilingual BERT Model for Indonesian Aspect-based Sentiment Analysis," in *2020 7th International Conference on Advanced Informatics: Concepts, Theory and Applications, ICAICTA 2020*, Institute of Electrical and Electronics Engineers Inc., Sep. 2020. doi: 10.1109/ICAICTA49861.2020.9428882.
- [17] U. D. Gandhi, P. Malarvizhi Kumar, G. Chandra Babu, and G. Karthick, "Sentiment Analysis on Twitter Data by Using Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM)," *Wirel Pers Commun*, 2021, doi: 10.1007/s11277-021-08580-3.