

Analisis Sentimen Berdasarkan Hasil Review Lokasi Google Map Menggunakan Natural Language Toolkit TextBlob dan Naïve Bayes

Sri Lestanti^{1*}, Saiful Nur Budiman², Erwan³

^{1,2,3} Program Studi Teknik Informatika, Universitas Islam Balitar, Blitar, Jawa Timur - Indonesia

lestanti85@gmail.com¹, sync.saifulnb@gmail.com², erwanblitar86@gmail.com³

Correspondence: sync.saifulnb@gmail.com²

ABSTRAK

Metode analisa sentimen adalah metode yang digunakan untuk memberikan pemahaman tentang perasaan atau opini yang terkandung dalam suatu teks. Seiring dengan perkembangan teknologi di era digital saat ini, data yang dihasilkan semakin melimpah salah satunya merupakan data yang dihasilkan oleh platform review seperti Google Maps. Google Maps adalah salah satu aplikasi yang populer dengan kemampuannya mencari lokasi, memberikan rute hingga memberikan ulasan atau review terhadap tempat yang dikunjungi. Dengan adanya review yang bisa diberikan pada Google Maps maka dapat dilakukan analisa sentimen yang bertujuan untuk mengetahui bagaimana perasaan atau opini dari para pengguna terhadap suatu tempat atau layanan. Metode yang umum digunakan untuk menganalisis sentimen adalah dengan menggunakan Natural Language Toolkit dan Naive Bayes. Natural Language Toolkit adalah toolkit untuk pemrosesan bahasa alami yang dibangun dengan python, sedangkan Naive Bayes adalah salah satu algoritma simple tetapi powerful untuk analisis sentimen Naive bayes bekerja dengan asumsi setiap fitur independent satu sama lain yaitu setiap kata independent dan dapat dijadikan prediksi sentimen dengan kata kunci tertentu. Studi ini membahas mengenai analisis sentimen terhadap review pengguna google maps terhadap Cafe berdasarkan review article menggunakan kombinasi Natural Language Toolkit, TextBlob, dan Naive Bayes. Studi ini bertujuan untuk mengenali sentimen ulasan pengguna menjadi positif, negatif, dan netral. Dataset yang digunakan berjumlah 253 data latih dan 64 data uji dari hasil review pelanggan pada platform Google Maps . Hasil penelitian menunjukkan bahwa model yang dibangun memiliki akurasi sebesar 94% dan metode yang digunakan cukup efektif dalam menjalankan pengklasifikasian ulasan pelanggan. Studi ini dapat memberikan informasi kepada manajemen Cafe dalam memahami persepsi pelanggan, studi ini juga dapat menjadi referensi untuk melakukan analisis sentimen sentimen pada review lokasi lainnya.

Kata Kunci

Analisis Sentimen; NLTK; Google Map Review; Naïve Bayes

ABSTRACT

Sentiment analysis method is a method used to provide an understanding of the feelings or opinions contained in a text. With the development of technology in the current digital era, the data generated is becoming increasingly abundant, one of which is the data produced by review platforms such as Google Maps. Google Maps is one of the popular applications with its ability to search for locations, provide routes, and give reviews or ratings of places visited. With the reviews that can be given on Google Maps, sentiment analysis can be conducted to understand the feelings or opinions of users towards a place or service. The methods commonly used for sentiment analysis are Natural Language Toolkit and Naive Bayes. The Natural Language Toolkit is a toolkit for natural language processing built with Python, while Naive Bayes is a simple yet powerful algorithm for sentiment analysis. Naive Bayes works on the assumption that each feature is independent of one another, meaning each word is independent and can be used to predict sentiment with certain keywords. This study discusses sentiment analysis of Google Maps user reviews of cafes based on review articles using a combination of Natural Language Toolkit, TextBlob, and Naive Bayes. This study aims to identify user review sentiments as positive, negative, and neutral. The dataset used consists of 253 training data and 64 test data from customer reviews on the Google Maps platform. The research results show that the model built has an accuracy of 94% and the method used is quite effective in classifying customer reviews. This study can provide information to Cafe management in understanding customer perceptions, and it can also serve as a reference for conducting sentiment analysis on reviews of other locations.

Key Words

Sentiment Analysis; NLTK; Google Map Review; Naïve Bayes

Received: 20th November 2024

Accepted: 18th December 2024

Published: 31st December 2024

Citation: -

10.46510/jami.v5i2.311

ISSN 2722-4414 (p)/ 2722-4406 (e)

114

<https://journal.akb.ac.id/>

I. PENDAHULUAN

Penggunaan teks untuk mengumpulkan data dari berbagai sumber dari internet dan berbagai platform media sosial dikenal sebagai analisis sentimen. Tujuannya adalah untuk mendapatkan pendapat dari pengguna yang terlibat dalam platform tersebut. Dalam konteks bisnis dan pemasaran, analisis sentimen dapat membantu perusahaan untuk memahami pandangan dan opini konsumen terhadap produk atau layanan yang mereka tawarkan (Darwis dkk., 2020). Seiring dengan perkembangan teknologi informasi dan popularitas media sosial, semakin banyak data teks yang dihasilkan setiap hari oleh pengguna internet. Oleh karena itu, penting bagi perusahaan untuk dapat menganalisis data ini secara efisien dan cepat untuk dapat mengambil keputusan yang tepat dalam strategi pemasaran mereka. Salah satu alasan mengapa analisis sentimen menjadi semakin penting adalah karena faktor kecepatan dalam pengambilan keputusan. Dengan analisis sentimen, perusahaan dapat memantau secara real-time respons konsumen terhadap produk atau layanan tertentu. Hal ini memungkinkan perusahaan untuk segera menyesuaikan strategi pemasaran mereka jika terdapat sentimen negatif yang tersebar luas di media online. Selain itu analisis sentimen juga dapat digunakan untuk memahami reputasi brand suatu perusahaan. Dengan menganalisis sentimen yang terkait dengan brand perusahaan, perusahaan dapat mengetahui apakah brand mereka memiliki image yang positif di mata konsumen atau tidak. Hal ini penting dalam membangun hubungan yang baik dengan konsumen dan mempertahankan loyalitas konsumen terhadap brand perusahaan.

Analisis sentimen pada Google Map review memiliki peran yang sangat penting dalam kehidupan sehari-hari, terutama dalam era digital seperti sekarang ini (Ardiansyah dkk., 2023). Dengan semakin berkembangnya teknologi, banyak orang yang mengandalkan Google Map untuk mencari informasi tentang lokasi, tempat makan, tempat wisata, dan berbagai layanan lainnya. Oleh karena itu, analisis sentimen pada review di Google Map dapat memberikan gambaran yang jelas tentang pengalaman orang-orang terhadap suatu tempat atau layanan (Haq dan Rachmat, 2020). Pentingnya analisis sentimen pada Google Map review dapat dilihat dari berbagai aspek. Pertama, analisis sentimen dapat membantu pengguna dalam membuat keputusan. Dengan melihat review yang diberikan oleh pengguna lain, seseorang dapat menilai kualitas suatu tempat atau layanan sebelum mengunjunginya. Hal ini tentu sangat membantu dalam menghindari pengalaman negatif dan memilih tempat yang memenuhi kebutuhan dan harapan. Kedua, analisis sentimen juga penting bagi pemilik usaha. Dengan mengetahui sentimen positif dan negatif dari review yang diberikan oleh pengguna, pemilik usaha dapat memperbaiki atau meningkatkan kualitas layanan mereka (Utomo dkk., 2019). Mereka dapat mengetahui apa yang disukai dan tidak disukai oleh pengguna, sehingga dapat melakukan perbaikan yang diperlukan untuk meningkatkan kepuasan pelanggan. Namun, dalam melakukan analisis sentimen pada Google Map review, perlu diingat bahwa tidak semua review dapat dijadikan acuan. Ada kemungkinan adanya review yang tidak objektif atau mungkin disebabkan oleh faktor subjektif. Oleh karena itu, perlu dilakukan analisis lebih mendalam dan teliti agar hasil yang didapatkan dapat dijadikan acuan yang akurat.

Naive Bayes merupakan salah satu algoritma klasifikasi yang sering digunakan dalam analisis sentimen (Rahayu dan Indra, 2022). Algoritma ini bekerja dengan asumsi bahwa semua fitur dalam data adalah independen, meskipun dalam kenyataannya mungkin tidak selalu demikian. Meskipun kesederhanaan metode ini, Naive Bayes telah terbukti efektif dalam melakukan analisis sentimen pada berbagai jenis data, termasuk review Google Maps. Naive Bayes dapat digunakan sebagai solusi untuk menganalisis sentimen dari ulasan Google Maps secara otomatis. Algoritma ini bekerja dengan menghitung probabilitas bahwa suatu ulasan termasuk dalam kategori yang diinginkan berdasarkan kata-kata yang terdapat dalam ulasan tersebut. Dengan mengasumsikan bahwa setiap kata adalah independen dari satu sama lain, Naive Bayes dapat menghasilkan prediksi sentimen dari ulasan dengan akurasi yang cukup tinggi (Masripah dan Utami, 2020). Jadi pada penelitian ini, penggunaan Naive Bayes ditawarkan sebagai solusi untuk analisis sentimen review Google Maps. Algoritma tersebut dapat membantu pengguna untuk dengan cepat mengetahui sentimen umum dari ulasan suatu tempat tanpa harus membaca setiap ulasannya satu per satu. Dengan demikian, pengguna dapat dengan mudah menentukan apakah suatu tempat tersebut layak untuk dikunjungi atau tidak berdasarkan ulasan yang ada. Selain itu, penggunaan Naive Bayes juga memungkinkan untuk adanya analisis yang konsisten dan objektif. Algoritma ini tidak terpengaruh oleh emosi atau pendapat pribadi, sehingga hasil analisis sentimen yang dihasilkan tidak terpengaruh oleh faktor subjektif. Dengan demikian, pengguna dapat mendapatkan informasi yang lebih obyektif tentang suatu tempat berdasarkan ulasan pengguna lain.

Pendekatan pemecahan masalah analisis sentimen terhadap review Google Maps menggunakan Natural Language Toolkit (NLTK) dan algoritma Naive Bayes adalah salah satu metode yang dapat digunakan untuk mengidentifikasi dan menganalisis sentimen yang terkandung dalam review bahasa Indonesia. Penelitian ini bertujuan untuk mencari pola-pola dari review pengguna Google Maps yang dapat digunakan untuk menentukan sejauh mana kepuasan pengguna terkait dengan suatu tempat atau lokasi tertentu. Dalam melakukan analisis sentimen, langkah pertama yang dilakukan adalah mengumpulkan data review dari Google Maps. Data yang telah terkumpul kemudian akan dibersihkan dari noise dan di preprocessing menggunakan library Sastrawi untuk memastikan bahwa teks yang akan dianalisis dalam bahasa Indonesia sudah dalam format yang tepat (Agastya, 2018). Penggunaan library Sastrawi ini penting untuk melakukan stemming

dalam bahasa Indonesia, sehingga kata-kata yang memiliki akar kata yang sama akan menjadi satu. Setelah data telah dibersihkan dan dipreprocessing, langkah selanjutnya adalah melakukan tokenisasi teks menjadi kata-kata yang lebih kecil. Hal ini dilakukan agar teks dapat diubah menjadi vektor fitur yang dapat digunakan oleh algoritma Naive Bayes untuk melakukan klasifikasi sentiment. Algoritma Naive Bayes merupakan salah satu algoritma machine learning yang populer dalam melakukan klasifikasi teks berdasarkan probabilitas kemunculan kata-kata dalam teks. Selain itu, penggunaan NLTK sebagai toolkit dalam pengolahan bahasa alami juga sangat penting dalam analisis sentimen (Yao, 2019). NLTK menyediakan berbagai fitur dan algoritma yang dapat digunakan untuk melakukan pemrosesan teks seperti tokenisasi, stemming, dan stopword removal. Dengan menggunakan NLTK, peneliti dapat dengan mudah memanfaatkan algoritma-algoritma yang sudah tersedia untuk membantu dalam memproses teks ke dalam bentuk yang dapat digunakan oleh algoritma klasifikasi. Dalam melakukan eksperimen, peneliti dapat membagi data review menjadi dua bagian, yaitu data training dan data testing. Data training digunakan untuk melatih model klasifikasi Naive Bayes sementara data testing digunakan untuk menguji seberapa baik model tersebut dalam mengklasifikasikan sentimen dari review-review yang belum pernah dilihat sebelumnya. Melalui pendekatan ini, peneliti dapat mengidentifikasi pola-pola sentiment yang terkandung dalam review Google Maps dengan lebih akurat. Hal ini dapat membantu pemilik bisnis atau pengelola tempat untuk memahami sejauh mana kepuasan pengguna terhadap layanan atau produk yang mereka tawarkan. Selain itu, hasil analisis sentimen juga dapat digunakan untuk mengidentifikasi area-area yang perlu diperbaiki atau ditingkatkan dalam rangka meningkatkan pengalaman pengguna dalam menggunakan suatu layanan atau produk. Dengan demikian, pendekatan pemecahan masalah ini dapat memberikan manfaat yang besar bagi pemilik bisnis dan pengelola dalam meningkatkan kualitas layanan yang mereka tawarkan. Dengan menggabungkan NLTK, Naive Bayes classifier, dan library Sastrawi, peneliti dapat membuat sebuah state of the art dalam analisis sentimen terhadap review Google Maps dalam bahasa Indonesia. Dengan menggunakan penelitian ini, perusahaan atau bisnis dapat memperoleh wawasan yang berharga dari ulasan pengguna untuk meningkatkan kualitas produk atau layanan mereka.

II. MATERIAL DAN METODE

Dalam era digital yang semakin berkembang saat ini, analisis sentimen menjadi salah satu metode yang penting dalam memahami opini dan perasaan masyarakat terhadap suatu topik atau peristiwa tertentu. Analisis sentimen memungkinkan para peneliti dan perusahaan untuk memantau dan menganalisis pandangan publik secara real-time, sehingga dapat digunakan untuk menginformasikan keputusan dan strategi yang lebih baik. Analisis sentimen adalah proses pengumpulan, pengolahan, dan interpretasi opini atau evaluasi yang diekspresikan dalam bentuk teks (Muzaki dan Witanti, 2021). Metode analisis sentimen ini sering digunakan dalam berbagai bidang, seperti pemasaran, politik, dan pelayanan pelanggan. Dalam konteks pemasaran, analisis sentimen dapat membantu perusahaan untuk memahami bagaimana konsumen merespons suatu produk atau layanan, sehingga dapat meningkatkan strategi pemasaran mereka.

Terdapat beberapa pendekatan yang digunakan dalam analisis sentimen, antara lain adalah pendekatan berbasis aturan, pendekatan berbasis statistik, dan pendekatan berbasis machine learning. Pendekatan berbasis aturan menggunakan aturan atau kamus kata-kata untuk mengidentifikasi sentimen positif, negatif, atau netral dalam teks. Sedangkan pendekatan berbasis statistik menggunakan metode statistik seperti analisis regresi untuk mengukur hubungan antara kata-kata dengan sentimen yang diekspresikan. Sementara itu, pendekatan berbasis machine learning menggunakan algoritma machine learning untuk mengklasifikasikan teks ke dalam sentimen yang sesuai. Dengan perkembangan teknologi yang pesat, analisis sentimen juga semakin berkembang dengan munculnya platform-platform analisis sentimen yang memanfaatkan data dari media sosial dan platform online lainnya. Misalnya, penelitian yang dilakukan oleh Zhang (2021) menggunakan data dari Twitter untuk menganalisis sentimen terkait pandemi COVID-19 (Zhang dkk., 2023). Mereka menemukan bahwa analisis sentimen dapat membantu dalam memahami respon masyarakat terhadap kebijakan lockdown dan vaksinasi.

Pada penelitian ini, penggunaan SerpAPI untuk melakukan crawling data (LLC, 2024). Crawling data merupakan proses pengambilan data secara otomatis dari website oleh mesin pencari. Salah satu tool yang bisa digunakan untuk melakukan crawling data adalah SerpAPI. Dengan menggunakan SerpAPI, pengguna dapat mengambil data hasil pencarian tanpa perlu membangun scraper mereka sendiri. Layanan ini memungkinkan pengguna untuk mengakses informasi yang relevan dengan cepat dan mudah, sehingga memudahkan dalam melakukan analisis data dan riset pasar. Dalam teori SerpAPI, ada beberapa konsep penting yang perlu dipahami. Pertama, adalah API (Application Programming Interface) yang merupakan antarmuka yang memungkinkan komunikasi antara perangkat lunak. SerpAPI menggunakan API untuk menyediakan pengguna dengan hasil pencarian yang diinginkan dengan cepat dan efisien. Kemudian, konsep lain yang penting dalam teori SerpAPI adalah web scraping. Web scraping adalah proses ekstraksi informasi dari sebuah website dengan cara otomatis. SerpAPI menggunakan teknik web scraping untuk mengambil data hasil pencarian dan menyajikannya kepada pengguna dalam format yang mudah dimengerti (Ma'ady dkk., 2023). SerpAPI juga memanfaatkan konsep HTML (Hypertext Markup Language) untuk menampilkan hasil pencarian dalam bentuk yang dapat dibaca oleh mesin. Dengan demikian, pengguna dapat dengan mudah mengakses dan menggunakan informasi yang diberikan oleh SerpAPI.

Prapemrosesan dan TF-IDF (Term Frequency-Inverse Document Frequency) adalah dua tahap yang penting dalam analisis sentiment, terutama dalam konteks analisis teks. Prapemrosesan data adalah langkah awal yang harus dilakukan sebelum menerapkan metode analisis sentimen apa pun. Tujuan dari prapemrosesan adalah untuk membersihkan, mengubah, dan mempersiapkan data mentah agar dapat diolah dengan benar oleh algoritma analisis sentimen. Beberapa teknik yang sering digunakan dalam prapemrosesan data termasuk tokenisasi, menghapus stop words, stemming, dan lemmatization.

Sebagai contoh, penelitian oleh Dianda (2023) menjelaskan pentingnya prapemrosesan data dalam analisis sentimen Twitter. Mereka menunjukkan bahwa melalui tahap prapemrosesan yang tepat, hasil analisis sentimen dapat menjadi lebih akurat dan dapat diandalkan. Selain itu, prapemrosesan data juga dapat membantu dalam mengatasi masalah noise dan redundansi dalam data teks yang mungkin mempengaruhi hasil analisis sentimen (Rifaldi dkk., 2023). Setelah prapemrosesan data, langkah selanjutnya adalah menerapkan TF-IDF untuk menghitung pentingnya sebuah kata dalam sebuah dokumen atau kelompok dokumen. TF-IDF menggabungkan dua metrik yaitu term frequency (TF) dan inverse document frequency (IDF). TF mengukur seberapa sering sebuah kata muncul dalam sebuah dokumen, sedangkan IDF mengukur seberapa unik sebuah kata dalam seluruh korpus dokumen.

Dalam penelitian terbaru oleh Widyassari dkk. (2022), mereka menunjukkan bahwa penggunaan TF-IDF dalam analisis sentimen dapat membantu dalam mengekstraksi informasi yang relevan dan penting dari setiap dokumen teks (Widyassari dkk., 2022). Dengan menggunakan TF-IDF, informasi yang lebih berharga dapat diprioritaskan dan digunakan untuk mengidentifikasi sentimen positif atau negatif dalam sebuah teks. Secara keseluruhan, prapemrosesan data dan TF-IDF merupakan dua tahap yang krusial dalam analisis sentimen. Dengan menerapkan kedua tahap ini dengan benar, hasil analisis sentimen dapat menjadi lebih akurat dan bermakna. Oleh karena itu, penelitian lebih lanjut dan pengembangan metode analisis sentimen yang mengintegrasikan prapemrosesan data dan TF-IDF dapat membawa manfaat besar dalam berbagai bidang, termasuk pemasaran, layanan pelanggan, dan kecerdasan bisnis.

Naive Bayes adalah salah satu metode pembelajaran mesin yang sangat populer dalam analisis sentimen. Metode ini menggunakan teorema Bayes untuk mengklasifikasikan data ke dalam kategori yang berbeda berdasarkan probabilitas dari fitur-fitur yang ada (Huda Mustakim, 2022). Rumus dasar teori Naive Bayes untuk analisis sentimen adalah sebagai berikut:

$$P(\text{kategori} | \text{fitur}) = P(\text{kategori}) * P(\text{fitur} | \text{kategori}) / P(\text{fitur}) \quad (1)$$

Keterangan:

- $P(\text{kategori} | \text{fitur})$ adalah probabilitas bahwa data termasuk ke dalam suatu kategori tertentu berdasarkan fitur yang ada
- $P(\text{kategori})$ adalah probabilitas sebelum melihat data
- $P(\text{fitur} | \text{kategori})$ adalah probabilitas fitur tersebut muncul di dalam kategori tersebut
- $P(\text{fitur})$ adalah probabilitas dari fitur yang ada

Untuk mengaplikasikan rumus tersebut dalam analisis sentimen, pertama perlu mengumpulkan data training yang berisi teks-teks atau dokumen-dokumen yang sudah diberi label sentimen (positif, negatif, atau netral). Selanjutnya, fitur-fitur dari teks tersebut harus diekstraksi, misalnya dengan menggunakan metode TF-IDF. Setelah itu, model Naive Bayes dapat dibuat menggunakan data training tersebut. Dalam konteks analisis sentimen bahasa Indonesia, metode Naive Bayes dapat digunakan untuk mengklasifikasikan teks-teks ke dalam kategori sentimen yang sesuai.

Confusion matrix digunakan dalam analisis sentimen untuk mengevaluasi performa dari model klasifikasi. Pada analisis sentimen bahasa Indonesia, confusion matrix sangat penting karena membantu mengidentifikasi seberapa baik model telah mengklasifikasikan sentimen positif, negatif, atau netral dari teks yang dianalisis (Dwi Normawati, 2021). Confusion matrix terdiri dari empat metrik utama: true positive (TP), true negative (TN), false positive (FP), dan false negative (FN). True positive adalah jumlah data yang benar-benar positif dan diklasifikasikan sebagai positif oleh model. True negative adalah jumlah data yang benar-benar negatif dan diklasifikasikan sebagai negatif oleh model. False positive adalah jumlah data negatif yang salah diklasifikasikan sebagai positif, sedangkan false negative adalah jumlah data positif yang salah diklasifikasikan sebagai negatif. Bentuk dari confusion matrik ditunjukkan pada gambar 1.

Dengan menggunakan confusion matrix, peneliti dapat menghitung metrik evaluasi penting seperti akurasi, presisi, recall, dan F1 score. Akurasi mengukur seberapa sering model mengklasifikasikan sentimen dengan benar, sedangkan presisi mengukur seberapa banyak sentimen positif yang benar-benar positif. Recall mengukur seberapa banyak sentimen positif yang terdeteksi oleh model, dan F1 score menggabungkan presisi dan recall untuk memberikan gambaran yang lebih komprehensif tentang performa model. Selain itu, confusion matrix juga membantu dalam menemukan masalah yang mungkin terjadi dalam model klasifikasi. Misalnya, jika terdapat banyak false positive pada confusion matrix, hal ini mungkin menandakan bahwa model cenderung terlalu sensitif terhadap sentimen positif. Sebaliknya, jika terdapat banyak

false negative, model mungkin terlalu konservatif dalam mengklasifikasikan sentimen positif. Dengan demikian, confusion matrix merupakan alat yang sangat berguna dalam analisis sentimen bahasa Indonesia karena membantu peneliti tidak hanya mengukur performa model klasifikasi, tetapi juga menemukan dan mengatasi masalah yang mungkin terjadi dalam proses analisis sentimen. Dengan memahami fungsi confusion matrix, pada peneliti ini dapat meningkatkan kualitas analisis sentimen dan membuat keputusan yang lebih tepat berdasarkan evaluasi yang lebih akurat dari sentimen yang terdapat dalam teks yang dianalisis.

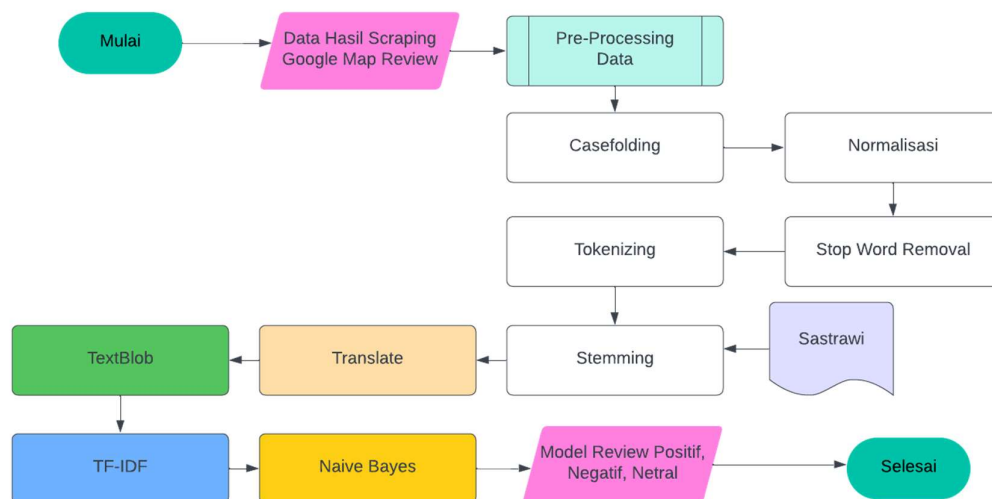
Data yang dikumpulkan merupakan hasil scarping dari SerpAPI terhadap aplikasi google map review pada lokasi Café yang telah ditentukan. Dari hasil scarping menghasilkan 317 data untuk analisis sentimen supaya mengetahui apakah review tersebut bersifat positif, negatif atau netral. Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian deskriptif kualitatif. Metode deskriptif kualitatif adalah sebuah metode penelitian yang memanfaatkan data kualitatif dan dijabarkan sejarah deskriptif. Penggunaan Naïve Bayes sebagai classifier untuk menentukan apakah review yang ditulis oleh pengguna pada Google Map Review bernilai positif, negative atau netral. Untuk mengujik akurasinya digunakanlah perhitungan dari tabel confusion matrix. Penggunaan NLTK dan Sastrawi sebagai stemmer supaya hasilnya dalam bentuk dasar kata indonesia secara lebih baik. Flowchart penelitian digambarkan pada gambar 2.

Hal pertama yang harus dilakukan pada penelitian ini adalah melakukan scarping data menggunakan SerpAPI dan mengambil kolom review dari setiap ulasan pengguna aplikasi Google Map Review. Dilakukan prapemrosesan data dimana untuk tahap casefolding akan mengubah semua huruf pada kalimat menjadi seragam huruf kecil semua. Tokenizing akan mengubah kalimat yang menjadi bentuk array untuk setiap kata. Penggunaan library NLTK untuk mempermudah tahap tokenizing. Filtering atau stopword digunakan untuk menghapus kata-kata yang bersifat kurang penting atau umum sebagai contoh kata ‘dan’, ‘yang’, ‘di’ dll. Dari hasil setiap langkah pada tahap praproses sebagai contoh ditunjukkan pada gambar 3.

Tahap selanjutnya melakukan stemming data untuk mengubah kata yang berimbuhan menjadi kata dasar. Library Sastrawi digunakan karena proses stemming dari bahasa Indonesia, jika menggunakan NLTK hasilnya sangat kurang sekali walaupun sudah ada kamusnya namun tidak selengkap Sastrawi. Dari hasil stemming akan dilakukan pembobotan untuk setiap kata dasar menggunakan TF-IDF. Semakin banyak kata yang berulang maka nilai bobotnya akan semakin tinggi. Naïve Bayes akan membentuk model berdasarkan data pembobotan TF-IDF. Model yang telah terbentuk nantinya akan digunakan untuk menentukan klasifikasi dari review komentar pengguna Google Map Review apakah positif, negative atau netral.

		Nilai Aktual	
		Positive	Negative
Nilai Prediksi	Positive	TP	FP
	Negative	FN	TN

Gambar 1. Confusion Matrix



Gambar 2. Diagram Alir Analisis Sentimen

review	text_cleaning	tokens
Tempat nongkrong salah satu ternyaman di kota ...	tempat nongkrong salah satu ternyaman di kota ...	[tempat, nongkrong, salah, satu, ternyaman, di...
stop	stemmed	
[tempat, nongkrong, salah, satu, ternyaman, ko...	[tempat, nongkrong, salah, satu, nyaman, kota,...	

Gambar 3. Hasil Prapemrosesan Data

III. HASIL

SerpAPI mengambil data komentar dari Google Map Review dengan cara memasukkan query nama lokasi tempat yang ingin dicari komentarnya serta api key seperti yang ditunjukkan pada gambar 4. Berdasarkan API Key tersebut, peneliti dapat mengambil data dalam jumlah besar namun dibatasi setiap hasil searchingnya yang berkisar 100 pencarian dalam waktu satu bulan untuk user dengan layanan gratis. Hasil pencarian akan menghasilkan snippet dimana berisikan informasi mengenai komentar user terhadap review lokasi tersebut. Untuk mempermudah membaca datanya, peneliti mengubah formatnya ke dalam file .CSV agar mudah dikelola. Gambar 5 menampilkan data ulasan dari Google Maps Review dengan total sebanyak 317 data, yang disimpan dalam bagian snippet. Kemudian dilakukan pre-processing.

```
# Cari di Google Maps Gelato n Coffe
params = {
    "engine": "google_maps",
    "q": "De Classe Gelato & Coffee",
    "location": "Blitar, Indonesia",
    "hl": "id",
    "gl": "id",
    "google_domain": "google.com",
    "api_key": "dac1cbcbb2a33046fad7ed78e58f800da50ca253fde48688fd379d1b2489765f"
}
search = GoogleSearch(params)
results = search.get_dict()
```

Gambar 4. Hasil Ulasan Google Map

```
data = pd.read_csv('/content/drive/MyDrive/Colab Notebooks/deClasseGelato(317).csv', encoding='latin-1')
[6] data.shape
(317, 4)
[8] data.head()
```

	username	rating	date	snippet
0	REVANANDO	5	2 bulan lalu	Makan makanan dingin di tengah panasnya Kota B...
1	Larasati Niken	5	2 bulan lalu	Gelato nya masih juara sih kalo di Blitar, ban...
2	Indra Ahmad	5	2 bulan lalu	Selalu membuat mood tenang
3	Bamboed	5	3 bulan lalu	De Classe Gelato & Coffee di Blitar adalah tem...
4	donny robyanto	5	3 bulan lalu	Tempat sangat nyaman dengan menu yang kompli...

Gambar 5. Tampilan Data Ulasan

Tahapan awal dalam pre-processing data ulasan dimulai dengan case folding, yang ditunjukkan pada tabel 1. Proses ini mengubah seluruh huruf pada kalimat ulasan menjadi huruf kecil secara seragam. Selain itu digunakan juga reguler expression (Regex) untuk mengurangi variasi teks tanpa mengubah maknanya. Pada penelitian ini regex yang digunakan antara lain menghapus angka dan simbol, menghapus white space atau spasi berlebih serta menghapus tab berlebih kemudian menggantinya dengan satu spasi.

Tahap selanjutnya melakukan normalisasi terhadap hasil case folding. Normalisasi memastikan bahwa data teks memiliki format yang konsisten sehingga algoritma analisis teks dapat bekerja dengan lebih efektif dan efisien. Kata-kata dengan arti yang sama tetapi berbebeda dapat diseragamkan, misalkan pada penelitian ini kata ‘yg’, ‘tdk’, ‘jg’, ‘bgt’, ‘bbrp’, ‘udah’ diubah menjadi ‘yang’, ‘tidak’, ‘juga’, ‘banget’, ‘beberapa’, ‘sudah’. Hasil dari proses normalisasi ditunjukkan pada tabel 2. Untuk menghapus kata-kata umum yang sering muncul digunakanlah tahapan ketiga yang disebut dengan stop word removal. Proses ini menghapus kata yang dianggap tidak memberikan banyak informasi dalam analisis teks. Beberapa kata-kata secara semantik yang tidak penting pada penelitian ini antarlain ‘tidak’, ‘dengan’, ‘ia’, ‘oleh’, ‘nya’, ‘ny’, ‘lo’ dan ‘sih’. Hasil dari tahapan normalisasi ditunjukkan pada tabel 3.

Tabel 1. Hasil Case Folding Teks

Snippet	Case Folding
Makan makanan dingin di tengah panasnya Kota Blitar dan ada beberapa part kenangan yang tertinggal disini, sangat recommended karena pilihan gelatonya jg beragam apalagi ada dark cocholatenya dan juga desain interiornya yg sangat mendukung, ciamik.	makan makanan dingin di tengah panasnya kota blitar dan ada beberapa part kenangan yang tertinggal disini sangat recommended karena pilihan gelatonya jg beragam apalagi ada dark cocholatenya dan juga desain interiornya yg sangat mendukung ciamik
Gelato nya masih juara sih kalo di Blitar, banyak varian rasa enak, nggk yg tipe gampang luber, tempatnya jg nyaman, luas, oiya satu lg, muraah lo drpd di gelato lain yg aku pernah coba. Menu lainnya oke udh bnyak bgt berkembang nya ð,	gelato nya masih juara sih kalo di blitar banyak varian rasa enak nggk yg tipe gampang luber tempatnya jg nyaman luas oiya satu lg muraah lo drpd di gelato lain yg aku pernah coba menu lainnya oke udh bnyak bgt berkembang nya
Selalu membuat mood tenang	selalu membuat mood tenang
De Classe Gelato & Coffee di Blitar adalah tempat yang sangat direkomendasikan bagi pecinta es krim dan kopi. Menyajikan berbagai macam gelato dengan rasa yang autentik dan lezat, tempat ini juga menawarkan berbagai pilihan kopi yang dibuat dengan biji berkualitas tinggi. Suasana kafe yang nyaman dan estetis membuatnya cocok untuk bersantai atau mengerjakan tugas. Pelayanan yang ramah dan harga yang terjangkau menambah nilai plus tempat ini. Lokasinya yang strategis di pusat kota membuatnya mudah diakses. Data De Classe Gelato & Coffee adalah destinasi yang sempurna untuk menikmati saat-saat santai bersama teman atau keluarga.	de classe gelato coffee di blitar adalah tempat yang sangat direkomendasikan bagi pecinta es krim dan kopi menyajikan berbagai macam gelato dengan rasa yang autentik dan lezat tempat ini juga menawarkan berbagai pilihan kopi yang dibuat dengan biji berkualitas tinggi suasana kafe yang nyaman dan estetis membuatnya cocok untuk bersantai atau mengerjakan tugas pelayanan yang ramah dan harga yang terjangkau menambah nilai plus tempat ini lokasinya yang strategis di pusat kota membuatnya mudah diakses data de classe gelato coffee adalah destinasi yang sempurna untuk menikmati saat saat santai bersama teman atau keluarga
Tempat sangat nyaman dengan menu yang komplit dan berkelas ð	tempat sangat nyaman dengan menu yang komplit dan berkelas

Tabel 2. Hasil Normalisasi Teks

Case Folding	Normalisasi
makan makanan dingin di tengah panasnya kota blitar dan ada beberapa part kenangan yang tertinggal disini sangat recommended karena pilihan gelatonya jg beragam apalagi ada dark cocholatenya dan juga desain interiornya yg sangat mendukung ciamik	makan makanan dingin di tengah panasnya kota blitar dan ada beberapa part kenangan yang tertinggal disini sangat recommended karena pilihan gelatonya juga beragam apalagi ada dark cocholatenya dan juga desain interiornya yang sangat mendukung ciamik
gelato nya masih juara sih kalo di blitar banyak varian rasa enak nggk yg tipe gampang luber tempatnya jg nyaman luas oiya satu lg muraah lo drpd di gelato lain yg aku pernah coba menu lainnya oke udh bnyak bgt berkembang nya	gelato nya masih juara sih kalo di blitar banyak varian rasa enak nggk yang tipe gampang luber tempatnya juga nyaman luas oiya satu lg muraah lo daripada di gelato lain yang aku pernah coba menu lainnya oke sudah bnyak banget berkembang nya
selalu membuat mood tenang	selalu membuat mood tenang
de classe gelato coffee di blitar adalah tempat yang sangat direkomendasikan bagi pecinta es krim dan kopi menyajikan berbagai macam gelato dengan rasa yang autentik dan lezat tempat ini juga menawarkan berbagai pilihan kopi yang dibuat dengan biji berkualitas tinggi suasana kafe yang nyaman dan estetis membuatnya cocok untuk bersantai atau mengerjakan tugas pelayanan yang ramah dan harga yang terjangkau menambah nilai plus tempat ini lokasinya yang strategis di pusat kota membuatnya mudah diakses data de classe gelato coffee adalah destinasi yang sempurna untuk menikmati saat saat santai bersama teman atau keluarga	de classe gelato coffee di blitar adalah tempat yang sangat direkomendasikan bagi pecinta es krim dan kopi menyajikan berbagai macam gelato dengan rasa yang autentik dan lezat tempat ini juga menawarkan berbagai pilihan kopi yang dibuat dengan biji berkualitas tinggi suasana kafe yang nyaman dan estetis membuatnya cocok untuk bersantai atau mengerjakan tugas pelayanan yang ramah dan harga yang terjangkau menambah nilai plus tempat ini lokasinya yang strategis di pusat kota membuatnya mudah diakses data de classe gelato coffee adalah destinasi yang sempurna untuk menikmati saat saat santai bersama teman atau keluarga
tempat sangat nyaman dengan menu yang komplit dan berkelas	tempat sangat nyaman dengan menu yang komplit dan berkelas

Tokenizing pada tahapan berikutnya untuk memecah teks menjadi unit-unit kecil yang disebut sebagai token. Unit tersebut bisa berupa frasa atau bahkan karakter. Proses pemisahan teks dilakukan berdasarkan kriteria spasi dan tanda baca. Untuk hasil daftar token tersebut ditunjukkan pada tabel 4. Tahap terakhir dari pre-processing adalah mengubah kata ke dalam bentuk dasar atau stem. Hal yang dilakukan di penelitian ini dengan cara menghapus imbuhan, akhiran ataupun awalan.

Proses ini tidak mempertimbangkan makna kontekstual kata, namun berfokus pada penghapusan morfologi kata untuk mendapatkan bentuk dasar. Hasil dari proses stemming ditunjukkan pada tabel 5.

Tabel 3. Hasil Stop Word Removal

Normliasasi	Stop Word Removal
makan makanan dingin di tengah panasnya kota blitar dan ada beberapa part kenangan yang tertinggal disini sangat recommended karena pilihan gelatonya juga beragam apalagi ada dark cocholatenya dan juga desain interiornya yang sangat mendukung ciamik	makan makanan dingin tengah panasnya kota blitar ada beberapa part kenangan tertinggal disini sangat recommended pilihan gelatonya beragam ada dark cocholatenya juga desain interiornya sangat mendukung ciamik
gelato nya masih juara sih kalo di blitar banyak varian rasa enak nggk yang tipe gampang luber tempatnya nyaman luas oiya satu lg muraahh lo daripada di gelato lain yang aku pernah coba menu lainnya oke sudah bnyak banget berkembang nya	gelato masih juara kalo blitar banyak varian rasa enak nggk tipe gampang luber tempatnya nyaman luas oiya satu lg muraahh daripada gelato yang aku pernah coba menu lainnya oke bnyak banget berkembang
selalu membuat mood tenang	selalu membuat mood tenang
de classe gelato coffee di blitar adalah tempat yang sangat direkomendasikan bagi pecinta es krim dan kopi menyajikan berbagai macam gelato dengan rasa yang autentik dan lezat tempat ini juga menawarkan berbagai pilihan kopi yang dibuat dengan biji berkualitas tinggi suasana kafe yang nyaman dan estetis membuatnya cocok untuk bersantai atau mengerjakan tugas pelayanan yang ramah dan harga yang terjangkau menambah nilai plus tempat ini lokasinya yang strategis di pusat kota membuatnya mudah diakses data de classe gelato coffee adalah destinasi yang sempurna untuk menikmati saat saat santai bersama teman atau keluarga	de classe gelato coffee blitar tempat sangat direkomendasikan pecinta es krim kopi menyajikan berbagai macam gelato rasa autentik lezat tempat juga menawarkan berbagai pilihan kopi dibuat biji berkualitas tinggi suasana kafe nyaman estetis membuatnya cocok bersantai mengerjakan tugas pelayanan ramah harga terjangkau menambah nilai plus tempat lokasinya strategis pusat kota membuatnya mudah diakses data de classe gelato coffee destinasi sempurna menikmati saat santai bersama teman keluarga
tempat sangat nyaman dengan menu yang komplit dan berkelas	tempat sangat nyaman menu komplit berkelas

Tabel 4. Hasil Tokenizing

Normliasasi	Tokenizing
makan makanan dingin tengah panasnya kota blitar ada beberapa part kenangan tertinggal disini sangat recommended pilihan gelatonya beragam ada dark cocholatenya juga desain interiornya sangat mendukung ciamik	makan,makanan,dingin,tengah,panasnya,kota,blitar,ada,beberapa,part,kenangan,tertinggal,disini,sangat,recommended,pilihan,gelatonya,beragam,ada,dark,cocholatenya,juga,desain,interiornya,sangat,mendukung,ciamik
gelato masih juara kalo blitar banyak varian rasa enak nggk tipe gampang luber tempatnya nyaman luas oiya satu lg muraahh daripada gelato yang aku pernah coba menu lainnya oke bnyak banget berkembang	gelato,masih,juara,kalo,blitar,banyak,varian,rasa,enak,nggk,tipe,gampang,lu,ber,tempatnya,nyaman,luas,oiya,satu,lg,muraahh,daripada,gelato,yang,aku,pernah,coba,menu,lainnya,oke,bnyak,banget,berkembang
selalu membuat mood tenang	selalu,membuat,mood,tenang
de classe gelato coffee blitar tempat sangat direkomendasikan pecinta es krim kopi menyajikan berbagai macam gelato rasa autentik lezat tempat juga menawarkan berbagai pilihan kopi dibuat biji berkualitas tinggi suasana kafe nyaman estetis membuatnya cocok bersantai mengerjakan tugas pelayanan ramah harga terjangkau menambah nilai plus tempat lokasinya strategis pusat kota membuatnya mudah diakses data de classe gelato coffee destinasi sempurna menikmati saat santai bersama teman keluarga	de,classe,gelato,coffee,blitar,tempat,sangat,direkomendasikan,pecinta,es,krim,kopi,menyajikan,berbagai,macam,gelato,rasa,autentik,lezat,tempat,juga,menawarkan,berbagai,pilihan,kopi,dibuat,biji,berkualitas,tinggi,suasana,kafe,nyaman,estetis,membuatnya,cocok,bersantai,mengerjakan,tugas,pelayanan,ramah,harga,terjangkau,menambah,nilai,plus,tempat,lokasinya,strategis,pusat,kota,membuatnya,mudah,diakses,data,de,classe,gelato,coffee,destinasi,sempurna,menikmati,saat,santai,bersama,teman,keluarga
tempat sangat nyaman menu komplit berkelas	tempat,sangat,nyaman,menu,komplit,berkelas

Tabel 5. Hasil Stemming

Tokenizing	Stemming
makan,makanan,dingin,tengah,panasnya,kota,blitar,ada,beberapa,part,kenangan,tertinggal,disini,sangat,recommended,pilihan,gelatonya,beragam,ada,dark,cocholatenya,juga,desain,interiornya,sangat,mendukung,ciamik	makan makan dingin tengah panas kota blitar ada beberapa part kenang tinggal sini sangat recommended pilih gelatonya agam ada dark cocholatenya juga desain interior sangat dukung ciamik
gelato,masih,juara,kalo,blitar,banyak,varian,rasa,enak,nggk,tipe,gampang,luber,tempatnya,nyaman,luas,oiya,satu,lg,muraahh,daripada,gelato,yang,aku,pernah,coba,menu,lainnya,oke,bnyak,banget,berkembang	gelato masih juara kalo blitar banyak varian rasa enak nggk tipe gampang luber tempat nyaman luas oiya satu lg muraahh daripada gelato yang aku pernah coba menu lain oke bnyak banget kembang
selalu,membuat,mood,tenang	selalu buat mood tenang

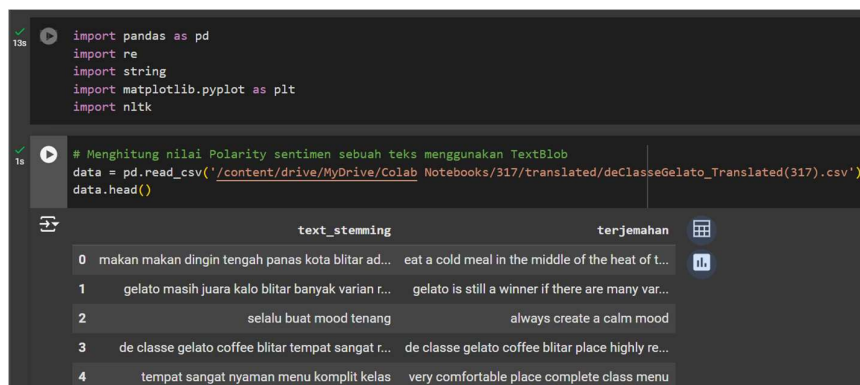
de,classe,gelato,coffee,blitar,tempat,sangat,direkomendasikan,pecinta,es,kr
im,kopi,menyajikan,berbagai,macam,gelato,rasa,autentik,lezat,tempat,juga
,menawarkan,berbagai,pilihan,kopi,dibuat,biji,berkualitas,tinggi,suasana,k
afe,nyaman,estetis,membuatnya,cocok,bersantai,mengerjakan,tugas,pelaya
nan,ramah,harga,terjangkau,menambah,nilai,plus,tempat,lokasinya,strategi
s,pusat,kota,membuatnya,mudah,diakses,data,de,classe,gelato,coffee,destin
asi,sempurna,menikmati,saat,santai,bersama,teman,keluarga

de classe gelato coffee blitar tempat sangat rekomendasi cinta es krim
kopi saji bagai macam gelato rasa autentik lezat tempat juga tawar bagai
pilih kopi buat biji kualitas tinggi suasana kafe nyaman estetis buat
cocok santa kerja tugas layan ramah harga jangkau tambah nilai plus
tempat lokasi strategis pusat kota buat mudah akses data de classe
gelato coffee destinasi sempurna nikmat saat santai sama teman
keluarga

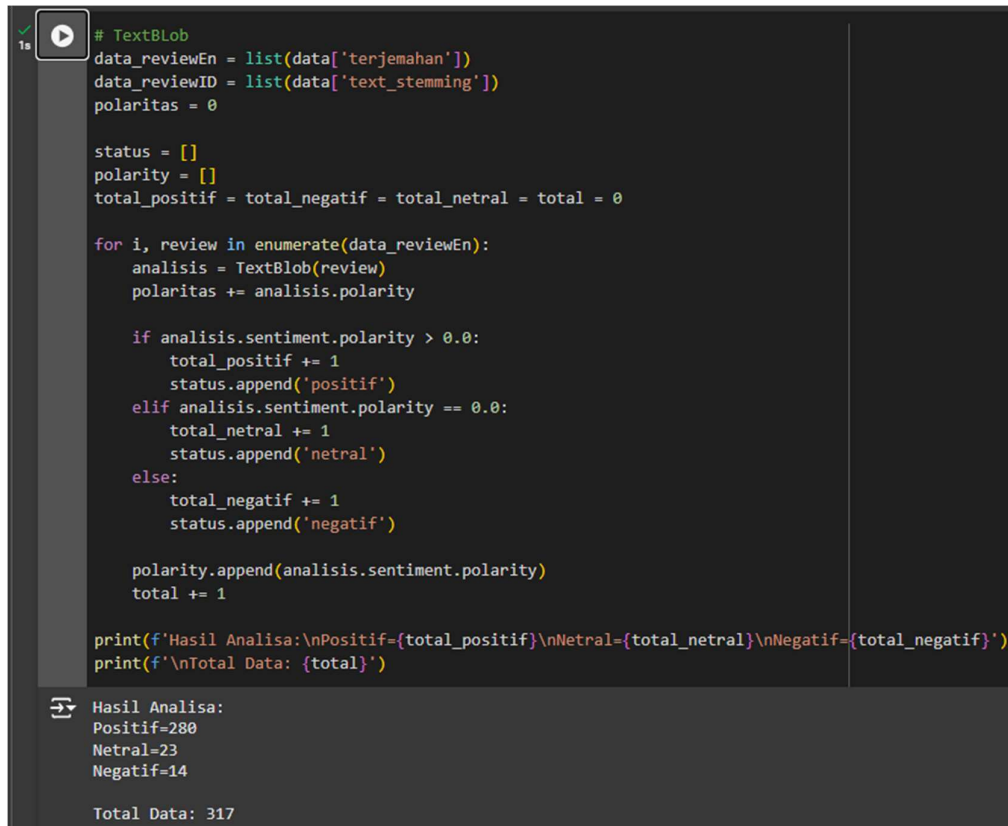
tempat,sangat,nyaman,menu,komplit,berkelas

tempat sangat nyaman menu komplit kelas

Hasil dari proses stemming akan menjadi inputan untuk tahapan translate. Pada penelitian ini teks indonesia harus di translate terlebih dahulu ke bahasa Inggris. Ini berkaitan dengan metode yang digunakan untuk memberikan label sentimen secara default pada komentar yang telah diperoleh yaitu dengan TextBlob. Library dari python tersebut mempermudah analisis sentimen positif, negatif ataupun netral berdasarkan analisis polarity atau tingkat kepositifan atau kenegatifan dan subjectivity seberapa subjektif atau objektif suatu teks. Hasil dari tahap penerjemahan ditunjukkan pada gambar 6.



Gambar 6. Hasil terjemahan ulasan

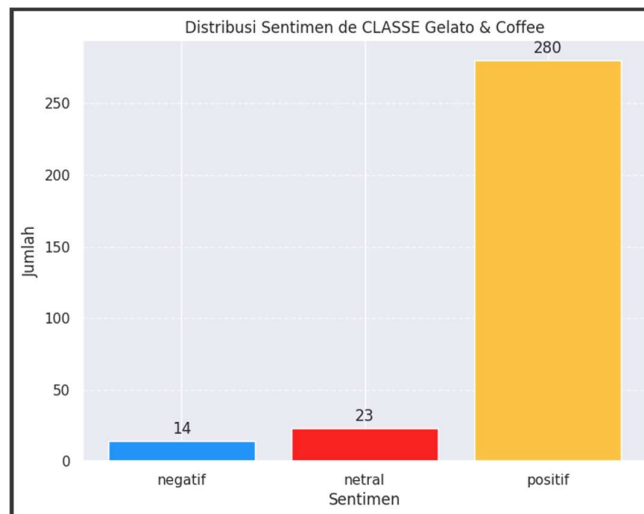


Gambar 7. Hasil Analisa Sentimen

Hasil analisis sentimen ditampilkan pada Gambar 7 menunjukkan jumlah ulasan dengan sentimen positif sebanyak 280 data, netral sebanyak 23 data, dan negatif sebanyak 14 data. Text dianggap memiliki ulasan positif jika nilai polarity > 0 sebagai contoh pada teks tersebut menunjukkan perasaan atau opini yang mendukung, puas, baik ataupun nilai positif secara umum. Sentimen bernilai negatif jika ulasa memiliki nilai polarity < 0 yang mana artinya pada ulasan tersebut mengandung perasaan atau opini tidak puas, buruku, kritik ataupun marah. Sentimen bernilai netral jika polarity yang dimilikinya bernilai 0. Teks tidak menunjukkan sentimen yang jelas, baik positif ataupun negatif. Biasanya ini berlaku untuk fakta atau pernyataan secara objektif. Hasil distribusi dari sentimen negatif, positif dan netral serta seringnya kata yang muncul ditunjukkan pada gambar 8 dan 9.



Gambar 8. Kata Yang Sering Muncul pada Ulasa



Gambar 9. Tampilan Grafik Hasil Ulasan

TF-IDF (Term Frequency - Inverse Document Frequency) pada penelitian ini digunakan untuk menilai seberapa penting sebuah kata dalam sebuah dokumen atau korpus (kumpulan dokumen). Dalam konteks analisis sentimen, TF-IDF dapat juga digunakan untuk mengubah teks menjadi representasi numerik yang dapat diproses oleh algoritma machine learning, sehingga memungkinkan model untuk memahami dan mengklasifikasikan sentimen dalam teks. Pada gambar 10 ditunjukkan hasil perhitungan TF-IDF.

Atribut max_features digunakan untuk menentukan jumlah kata yang paling sering muncul yang digunakan dalam pembuatan vektor fitur. Kata-kata yang jarang muncul tidak memiliki pengaruh signifikan dalam klasifikasi tidak akan dihitung, oleh karena itu hanya 1000 kata yang paling sering muncul dalam dataset sajalah yang disimpan. Nilai min_df sebesar 5 mengartikan bahwa kata tersebut harus muncul di setidaknya pada 5 dokumen, sementara nilai max_df sebesar 0,7 menetapkan bahwa kata tersebut tidak boleh muncul di lebih dari 70 persen di dokumen. Penetapan batas 70 persen ini bertujuan untuk menghindari penggunaan kata yang terlalu umum, yang cenderung tidak berkontribusi dalam klasifikasi sentimen.

Setelah teks dianalisis menggunakan TF-IDF, Naive Bayes dapat digunakan untuk melakukan klasifikasi sentimen berdasarkan vektor fitur yang dihasilkan oleh TF-IDF. Algoritma Naive Bayes bekerja dengan prinsip probabilistik, dan dapat memanfaatkan distribusi kata-kata dalam kelas yang berbeda (misalnya, sentimen positif dan negatif) untuk menentukan kelas teks baru (misalnya, positif atau negatif). Pada gambar 11 ditunjukkan proses algoritma Naive Bayes dengan proporsi data training sebesar 80% dan data testing sebesar 20% dari total 317 data.

```

from sklearn.feature_extraction.text import TfidfVectorizer

# banyaknya term yang akan digunakan,
# di pilih berdasarkan top max_features
# yang diurutkan berdasarkan term frequency seluruh corpus
tfidfconverter = TfidfVectorizer(max_features=1000, min_df=5, max_df=0.7, stop_words=stopwords.words('english'), ngram_range=(1,3))
X1 = tfidfconverter.fit_transform(x)
    
```

Gambar 10. Hasil TF-IDF

```

# Membagi Data ke Training dan Test Set
# 20% untuk testing -> test_size=0.2
# 80% untuk training
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(X1, y, test_size=0.2, random_state=0)

# predictions_nb = text_classifier_nb.predict(x_test.toarray())
predictions_proba_nb = text_classifier_nb.predict_proba(x_test.toarray()) # Get probability predictions
predictions_nb = predictions_proba_nb.argmax(axis=1) # Convert probabilities to class predictions
t2_nb = time.time()
time_linear_train_nb = t1_nb-t0_nb
time_linear_predict_nb = t2_nb-t1_nb

# hasil
print("Naive Bayes Training time: %fs; Prediction time: %fs" % (time_linear_train_nb, time_linear_predict_nb))
    
```

Naive Bayes Training time: 0.008339s; Prediction time: 48.940117s

Gambar 11. Prediksi Test Case Naive Bayes

Sebagai langkah akhir untuk mengevaluasi model klasifikasi pada penelitian ini digunakanlah confusion matrix yang ditunjukkan pada gambar 12. Diperoleh akurasi sebesar 0.94 atau 94%. Precision merupakan rasio prediksi positif yang benar dengan total prediksi positif. Untuk kelas negatif precision-nya 0%, kelas netral precision-nya 100%, kelas positif precision-nya 97%. Recall digunakan untuk prediksi positif yang benar dengan total contoh positif sebenarnya. Kelas negatif bernilai 0%, netral 75%, positif 98%. F1-Score merupakan rata-rata harmonis dari precision dan recall yang memberikan gambaran seimbang dari kedua metrik tersebut. Kelas negatif bernilai 0%, netral 86%, positif 97%. Support menunjukkan contoh data testing sebenarnya. Untuk uji kelas negatif ada 2, netral, 4, positif 58 sehingga total data testing 64 sesuai dengan skenario yang ditentukan sebelumnya. Untuk hasil confusion matrix nya ditunjukkan pada gambar 13 yang mana merupakan hasil evaluasi model menggunakan data uji.

```

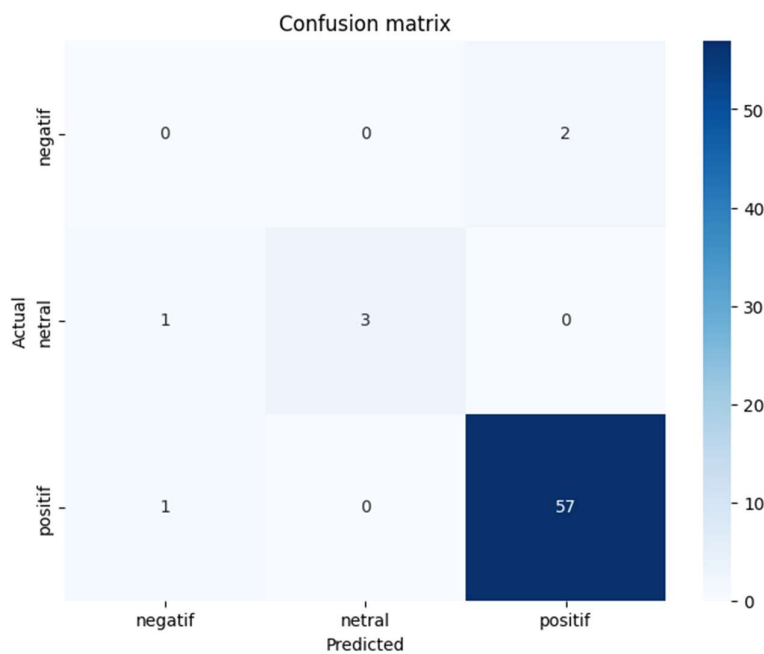
Naive Bayes Training time: 0.002716s; Prediction time: 1016.249185s
Accuracy: 0.94

Classification Report:
              precision    recall  f1-score   support

0               0.00         0.00         0.00         2
1               1.00         0.75         0.86         4
2               0.97         0.98         0.97         58

 accuracy                   0.94         0.94         64
 macro avg                  0.66         0.58         0.61         64
 weighted avg               0.94         0.94         0.94         64
    
```

Gambar 12. Hasil Klasifikasi Sentimen Analisis



Gambar 13. Hasil Confusion Matrix

Kualitas data ulasan lokasi dari Google Maps sering kali mengandung bahasa informal, slang, dan kesalahan pengetikan, yang menyulitkan proses pemuatan dataset dan klasifikasinya, terutama disebabkan oleh ketidakseimbangan data. Akibatnya, distribusi sentimen menjadi tidak seimbang, dengan jumlah ulasan positif yang lebih banyak dibandingkan ulasan negatif. Hal ini berdampak pada kinerja klasifikasi menggunakan algoritma Naïve Bayes yang menjadi kurang optimal.

IV. KESIMPULAN

Penelitian ini berhasil menganalisis sentimen dari ulasan Google Maps menggunakan metode TextBlob dan algoritma Naive Bayes. Dengan menggunakan dataset yang terdiri dari 253 data latih dan 64 data uji, hasil analisis menunjukkan bahwa model yang dikembangkan memiliki tingkat akurasi sebesar 94%. Keberhasilan ini menunjukkan bahwa pendekatan yang digunakan mampu mengklasifikasikan sentimen positif, negatif, dan netral dengan baik. Hasil ini juga mengindikasikan bahwa kombinasi TextBlob untuk ekstraksi sentimen dan Naive Bayes sebagai algoritma klasifikasi adalah metode yang efektif untuk analisis sentimen pada ulasan Google Maps. Namun, penelitian ini dapat ditingkatkan dengan menggunakan dataset yang lebih besar untuk menguji generalisasi model pada berbagai jenis data ulasan. Berdasarkan kendala yang telah diidentifikasi, penelitian selanjutnya perlu mempertimbangkan beberapa langkah, antara lain melakukan koreksi ejaan secara otomatis atau memanfaatkan library SymSpell. Selain itu, diperlukan deteksi dan penghapusan kata slang menggunakan kamus untuk memetakan kata slang ke dalam bentuk formal. Menyeimbangkan jumlah ulasan positif dengan ulasan negatif juga perlu dipertimbangkan, namun metode ini harus diterapkan secara hati-hati agar tidak menghilangkan informasi penting. Pengembangan model Machine Learning juga diperlukan untuk meningkatkan akurasi dalam analisis sentimen.

V. DAFTAR PUSTAKA

Agastya, I Made Artha. (2018). Pengaruh Stemmer Bahasa Indonesia Terhadap Performa Analisis Sentimen Terjemahan Usaha Film. *Jurnal Tekno Kompak*, 12. <https://doi.org/http://dx.doi.org/10.33365/jtk.v12i1.70>

Ardiansyah, Widagdo, Adika Sri, Qodri, Krisna Nuresa, Saputro, Fachruddin Edi Nugroho, & P, Nisrina Akbar Rizky. (2023). Analisis sentimen terhadap pelayanan Kesehatan berdasarkan ulasan Google Maps menggunakan BERT. *Jurnal Fasilkom*, 13(2), 326-333. <https://doi.org/http://dx.doi.org/10.37859/jf.v13i02.5170>

Darwis, Dedi, Pratiwi, Eka Shintya, & Pasaribu, A.Ferico Octaviansyah. (2020). Penerapan Algoritma SVM Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. *Jurnal Ilmiah Eductic*, 7. <https://doi.org/https://doi.org/10.21107/edutic.v7i1.8779>

Dwi Normawati, Surya Allit Prayogi. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2). <https://doi.org/http://dx.doi.org/10.30645/j-sakti.v5i2.369>

- Haq, Fathia Uqimul, & Rachmat, Heryandi. (2020). Penggunaan Google Review Sebagai Penilaian Kepuasan Pengunjung Dalam Pariwisata. *Journal of Sustainable Tourism Research*, 2. <https://doi.org/http://dx.doi.org/10.24198/tornare.v2i1.25826>
- Huda Mustakim, Sigit Priyanta. (2022). Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM. *Indonesian Journal of Computing and Cybernetics Systems (IJCCS)*, 16(2). <https://doi.org/https://doi.org/10.22146/ijccs.68903>
- LLC, SerpAPI. (2024). *SerpAPI*. <https://serpapi.com/>
- Ma'ady, Mochamad Nizar Palefi, Rizaldy, Denny Daffa, Satria, Rahul Fahmi, & Anaking, Purnama. (2023). SPARRING: Sistem Rekomendasi Peneliti Terintegrasi Google Scholar via SerpAPI dan Latent Dirichlet Allocation pada Konteks Perguruan Tinggi. *Jurnal Teknologi dan Manajemen Informatika*, 9. <https://doi.org/https://doi.org/10.26905/jtmi.v9i2.11111>
- Masripah, Siti, & Utami, Lula Dini. (2020). Algoritma Klasifikasi Naive Bayes Untuk Analisa Sentimen Aplikasi Shopee. *Jurnal SWABUMI*, 8. <https://doi.org/http://dx.doi.org/10.31294/swabumi.v8i2.8444>
- Muzaki, Akhmad, & Witanti, Arita. (2021). Sentiment Analysis of The Community in The Twitter to The 2020 Election in Pandemic Covid-19 By Method Naive Bayes Classifier. *Jurnal Teknik Informatika*, 2. <https://doi.org/https://doi.org/10.20884/1.jutif.2021.2.2.51>
- Rahayu, Irma Putri, & Indra, Ahmad Fauzi Jamaludin. (2022). Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes dan Support Vector Machine. *Jurnal Sistem Komputer dan Informatika (JSON)*, 4. <https://doi.org/http://dx.doi.org/10.30865/json.v4i2.5381>
- Rifaldi, Dianda, Fadlil, Abdul, & Herman. (2023). Teknik Preprocessing pada Text Mining Menggunakan Data Tweet "Mental Health". *Jurnal Pendidikan Teknologi Informasi*, 3. <https://doi.org/http://dx.doi.org/10.51454/decode.v3i2.131>
- Utomo, Pradita Eko Prasetyo, Manaar, Khaira, Ulfa, & Suratno, Tri. (2019). Analisis Sentimen Online Review Pengguna Bukalapak Menggunakan Metode Algoritma TF-IDF. *Jurnal Sains dan Sistem Informasi (JUSS)*, 2. <https://doi.org/https://doi.org/10.22437/juss.v2i2.8469>
- Widyassari, Adhika Pramita, Rustad, Supriadi, Shidik, Guruh Fajar, Noersasongko, Edi, Syukur, Abdul, Affandy, & Setiadi, De Rosal Ignatius Moses. (2022). Review of Automatic Text Summarization Techniques & Methods. *Journal of King Saud University – Computer and Information Sciences*, 34(4). <https://doi.org/https://doi.org/10.1016/j.jksuci.2020.05.006>
- Yao, Jiawei. (2019). Automated Sentiment Analysis of Text Data with NLTK. *Journal of Physics: Conference Series*, 1187. <https://doi.org/10.1088/1742-6596/1187/5/052020>
- Zhang, Qihuang, Yi, Grace Y., Chen, Li-Pang, & He, Wenqing. (2023). Sentiment Analysis and Casual Learning of COVID-19 Tweets Prior to The Rollout of Vaccines. *PLOS ONE*. <https://doi.org/https://doi.org/10.1371/journal.pone.0277878>