

# Membandingkan Analisis Sentimen Review Pelanggan Shopee Dan Tokopedia Menggunakan Google's NLP API

Revi Valen Sumendap<sup>a1</sup>, Ida Bagus Made Mahendra<sup>a2</sup>

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana  
Badung, Bali, Indonesia

<sup>1</sup>revisumendap1@gmail.com

<sup>2</sup>ibm.mahendra@unud.ac.id

## Abstract

Dengan teknologi yang sudah sangat maju di jaman sekarang mempermudah kita manusia dalam berbelanja. Yang dulunya kita harus berbelanja secara langsung namun sekarang kita bisa berbelanja secara online dengan E-Commerce yang ada. Transaksi e-commerce di Indonesia semakin meningkat, hal tersebut memberikan peluang pada produsen untuk memasarkan produk dan memudahkan konsumen untuk berbagi aktivitas, salah satu faktor yang sangat sering kita jumpai adalah memberikan ulasan produk. Ulasan produk berperan penting untuk membangun kepercayaan konsumen ketika menentukan keputusan dalam pembelian produk. Dengan meningkatnya jumlah ulasan, membuat calon konsumen kesulitan untuk menarik kesimpulan yang tepat. Oleh karena itu, diperlukan analisis sentimen untuk membantu calon konsumen untuk menarik kesimpulan. Analisis sentimen bertujuan untuk menyimpulkan, mengidentifikasi sentimen pada data yang ada. Disini kami akan melakukan Analisis Sentimen menggunakan Google Natural Language Processing dan membandingkan hasil ulasan aplikasi Tokopedia dan Shopee pada Google Play Store

**Keywords:** *Sentiment Analysis, Shopee, Tokopedia, Perbandingan, Natural Language Processing*

## 1. Introduction

Kemajuan teknologi informasi dan komunikasi jelas memberi dampak pada perubahan gaya hidup masyarakat Indonesia. Yang dulunya kita harus berbelanja secara langsung namun sekarang kita bisa berbelanja secara online karena berkembangnya e-commerce di Indonesia. Transaksi online atau e-commerce di Indonesia mengalami pertumbuhan yang pesat, menurut Kementerian Koordinator (Kemenko) Bidang Perekonomian mencatat nilai transaksi e-commerce di Indonesia, baik domestik dan luar negeri, mencapai Rp 108,54 triliun sepanjang kuartal I-2022. Realisasi itu tumbuh 23 persen dibandingkan periode yang sama tahun lalu. Peningkatan tersebut memberikan peluang besar bagi produsen untuk memasarkan produk dan memudahkan konsumen dalam berbagi aktivitas, salah satunya memberikan ulasan pada produk

Ulasan merupakan media komunikasi pemasaran dan periklanan yang berperan dalam proses keputusan pembelian konsumen untuk mengidentifikasi dan mengevaluasi produk yang dibutuhkan. Ulasan juga penting untuk membangun kepercayaan konsumen dengan cara menilai pengalaman pembeli sebelumnya dan dapat membantu produsen mengembangkan bisnis sesuai dengan permintaan pasar (1)

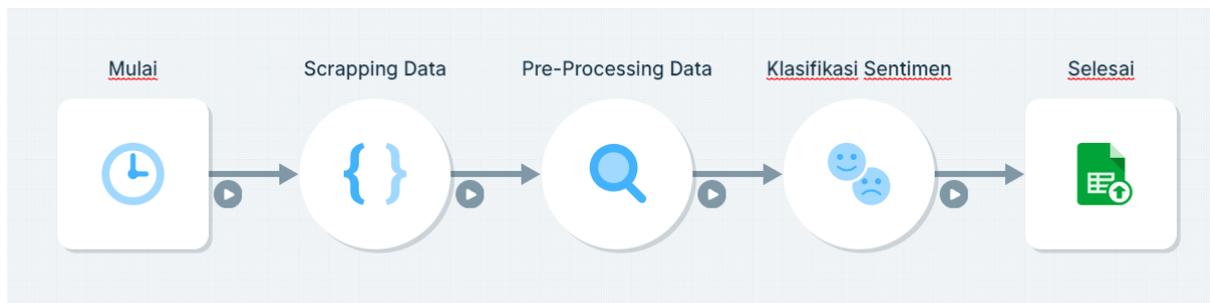
Di sisi lain, data mining dapat diartikan sebagai metode yang digunakan untuk menemukan pengetahuan dari sebuah database atau untuk menemukan pengetahuan yang berguna dari data. Ini adalah proses untuk mendapatkan desain dan hubungan yang menarik sehingga data dapat dilayani dalam volume yang besar. Dengan demikian data mining dapat digunakan sebagai analisis sentimen karena dapat mengolah data ulasan dalam jumlah banyak (2)

## 2. Metode Penelitian

### 2.1 Datasets

Datasets yang digunakan dalam penelitian ini adalah data yang diambil dari review user di google play store tentang aplikasi Tokopedia dan aplikasi Shopee. Review diambil secara acak dan merupakan review yang berbahasa inggris dan paling relevant sehingga bisa mendapatkan hasil yang lebih akurat. Untuk pengambilan data ini penulis menggunakan aplikasi Google Collab dan memanfaatkan Bahasa Python karena bisa dengan cepat mengambil hasil review di Google Play Store. Setelah data data ini diambil maka dilanjutkan dengan proses data pre-processing dengan menggunakan aplikasi Google Collab.

### 2.2 Alur Sistem

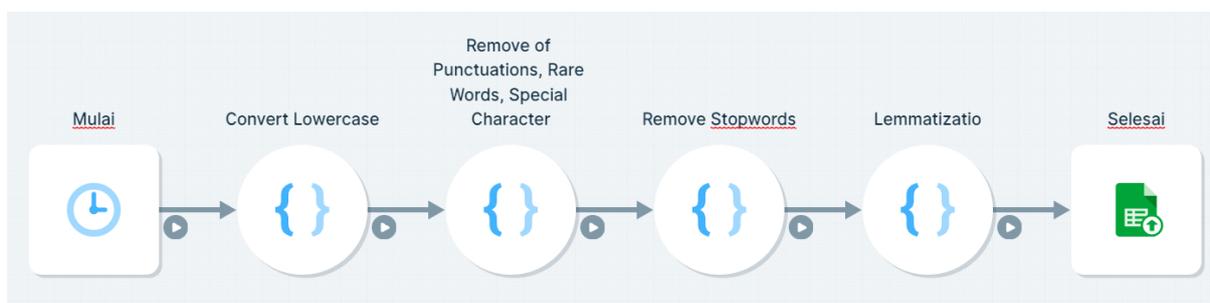


Gambar 1. Alur Sistem

Bisa kita lihat alur yang akan dilakukan dalam penelitian ini. Yang pertama adalah mengumpulkan datasets dari review google play store Menggunakan Google Collab dan menggunakan bahasa pemograman phyton. Setelah itu masuk ke tahap Pre-Processing datasets menggunakan aplikasi Google Collab yang juga menggunakan Bahasa phyton yang akan melalui Convert Lowercase, Removal of Punctuations, Removal of Stopwords, Removal of Rare Words, Removal of Special characters, Lemmatization. Setelah Datasets sudah melewati langkah langkah tersebut maka dilanjutkan dengan proses Validation Sentiment analysis dimana kita mengklasifikasi apakah Review itu merupakan review positif atau negatif dengan menggunakan aplikasi Parabola.io yang memanfaatkan Google's Natural Language Processing API.

### 2.3 Pre Processing Data

Pre Processing adalah Proses perlakuan terhadap data sebelum diolah pada pembelajaran mesin termasuk di dalamnya proses pembersihan dan penormalan data untuk memudahkan proses selanjutnya (3) Disini kita menggunakan Pre Processing Data yang berguna agar datasets kita yang sebelumnya belum terstruktur menjadi lebih terstruktur dan siap untuk diolah kedalam sentiment analysis dan kita bisa mendapat hasil yang lebih akurat. Untuk tahap Pre Processing Data ini penulis menggunakan sebanyak 6 tahap. Yaitu, Convert Lowercase, Removal of Punctuations, Removal of Stopwords, Removal of Rare Words, Removal of Special characters, Lemmatization. Dan akan dijelaskan fungsi dari proses proses ini sebagai berikut.



Gambar 2. Alur Pre-Processing Data

- Convert Lowercase

Dalam sebuah kalimat, biasanya ada bentuk huruf yang berbeda, tahap ini adalah proses keseragaman huruf, berkaitan dengan huruf kapital atau bukan.(2) Jadi disini kita akan membuat semua huruf menjadi huruf kecil.

- Removal of Punctuations, Rare Words, Special Character

Disini kita menghapus tanda baca seperti !"#\$%&'()\*+,-./:;<=>?@[]^\_`{|}~ menghapus kata kata yang jarang dipakai didalam datasets, dan juga menghapus input yang bukan Alfabet,nomor,atau kata penghubung karena semua hal ini bisa membuat data kita lebih susah dipahami oleh computer.

- Removal of Stopwords

Filter stopwords removal adalah proses menghilangkan kata-kata yang sering muncul namun tidak memiliki pengaruh apapun dalam ekstraksi sentimen suatu review. Kata yang termasuk seperti kata penunjuk waktu, kata tanya (4) contohnya seperti i, me, my, myself, we, our, ours

- Lemmatization

Lemmatization adalah algoritma yang memanfaatkan analisis morfologi dan aturan penulisan pemisahan dan penggabungan kata yang berguna sebagai pemeriksaan kebenaran ejaan (5). Jadi inti dari proses ini mengembalikan kata kepada kata dasarnya yang telah disesuaikan.

## 2.4 Validation

Ini adalah tahap terpenting atau tahap utama dari penelitian ini. Dimana kita akan mengklasifikasi sentiment hasil Review yang sudah diolah tadi, Apakah sentimennya positif atau sentiment negatif. Kita akan menggunakan Aplikasi Parabola.io untuk klasifikasi ini. Parabola.io merupakan sebuah aplikasi yang bisa melakukan sentiment analysis dengan memanfaatkan Google's Natural Language Processing API

Jadi langkah awal kita perlu memasukan dataset yang sudah melewati proses Pre-Processing tadi setelah itu kita perlu memasukannya kedalam menu Sentiment Analysis dan Parabola.io akan menerapkan Google's Natural Language Processing API untuk menentukan hasil klasifikasi yang kita inginkan.

Untuk output yang dikeluarkan nanti ada dua yaitu Sentiment Score dan Sentiment Confidence. Untuk sentiment score bisa angka positif dan juga angka negatif. Dimana angka positif diklasifikasikan menjadi Sentiment Positif dan angka negatif menjadi Sentimen Negatif. Untuk sentiment confidence adalah seberapa yakin system tentang score yang telah diberikan, outputnya bisa dari 0 sampai tak hingga.

## 3. Result and Discussion

### 3.1 Scrapping Data

Disini kami menggunakan Google Collab dan menggunakan Bahasa Phyton dan memakai tools tambahan seperti pandas dan numpy untuk scrapping data dari Review Google Playstore.

## Sumendap, Mahendra, Membandingkan Analisis Sentimen Review Pelanggan Shopee Dan Tokopedia Menggunakan Google's NLP API

reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent
0	f3d04d1c-9ed4-468b-b294-eb4af62d545a	Shafira Natasha	https://play-lh.googleusercontent.com/a-/ACNPE...	3	304	2.92.08	2022-08-29 00:29:34	Yaah maaf banget kalau kanu masih belum nyaman...
1	de3f8662-4107-45de-a2f1-7411aee9e13e	Surya Maulana	https://play-lh.googleusercontent.com/a-/ACNPE...	3	18	2.93.16	2022-09-28 13:09:00	None
2	543f4bb3-35df-4cda-8482-d20b9526f194	Pricylia W.	https://play-lh.googleusercontent.com/a-/ACNPE...	1	91	2.92.27	2022-08-22 01:35:05	None
3	74044e8a-3999-4fc1-af89-8bb0eea303b3	muichirou's lover	https://play-lh.googleusercontent.com/a-/ACNPE...	3	391	2.91.26	2022-08-09 16:21:04	Hi Mrs., sorry for the problem. I suggest you ...
4	250221bc-8e50-4179-8e2b-7d31829098e	mario billy	https://play-lh.googleusercontent.com/a-/ACNPE...	3	33	2.92.27	2022-09-23 07:53:26	I'm sorry for your trouble 🙏 Mimin advises to m...

**Gambar 3.** Hasil Scrapping

### 3.2 Merapikan Data

Data yang diambil dan akan melewati proses ini adalah Content atau tulisan dari Reviewer di Google Play Store. Disini content atau tulisan akan melewati 5 tahap.

#### 1. Convert Lowercase

Pada langkah pertama semua huruf dibuat menjadi huruf kecil

	content	clean_text
0	I like this app for the first time usage,it ha...	i like this app for the first time usage,it ha...
1	Can you make more ads page at the app? So my p...	can you make more ads page at the app? so my p...
2	i have enjoyed this marketplace and easy to use	i have enjoyed this marketplace and easy to use
3	Experienced frequent warning to update the app...	experienced frequent warning to update the app...
4	user friendly interface. Best e-commerce all t...	user friendly interface. best e-commerce all t...

**Gambar 4.** Hasil Convert Lowercase

#### 2. Removal of Punctuations

Pada langkah kedua kita menghapus tanda tanda baca seperti !"#\$%&'()\*+,-./:;<=>@[ ]^\_`{|}~

	content	clean_text
0	I like this app for the first time usage,it ha...	i like this app for the first time usageeit has...
1	Can you make more ads page at the app? So my p...	can you make more ads page at the app so my ph...
2	i have enjoyed this marketplace and easy to use	i have enjoyed this marketplace and easy to use
3	Experienced frequent warning to update the app...	experienced frequent warning to update the app...
4	user friendly interface. Best e-commerce all t...	user friendly interface best ecommerce all the...

**Gambar 5.** Hasil Removal of Puncutations

### 3. Removal of Stopwords

Pada langkah ini kita membuang kata kata yang tidak terlalu penting seperti i, me, my, myself, we, our, ours

	content	clean_text
0	I like this app for the first time usage,it ha...	like app first time usageeit speed delivery cer...
1	Can you make more ads page at the app? So my p...	make ads page app phone go lagi love ads
2	i have enjoyed this marketplace and easy to use	enjoyed marketplace easy use
3	Experienced frequent warning to update the app...	experienced frequent warning update app mandat...
4	user friendly interface. Best e-commerce all t...	user friendly interface best ecommerce time

**Gambar 6.** Hasil Removal Of Stopwords

### 4. Removal of Rare Words and Special Character

Pada langkah ini kita membuang kata kata yang jarang muncul dan membuang kata kata yang bukan Alfabet,nomor,atau kata penghubung karena dapat merusak dataset

	content	clean_text
0	I like this app for the first time usage,it ha...	like app first time usageeit speed delivery cer...
1	Can you make more ads page at the app? So my p...	make ads page app phone go lagi love ads
2	i have enjoyed this marketplace and easy to use	enjoyed marketplace easy use
3	Experienced frequent warning to update the app...	experienced frequent warning update app mandat...
4	user friendly interface. Best e-commerce all t...	user friendly interface best ecommerce time

**Gambar 7.** Hasil Removal Rare Words and Special Character

### 5. Lemmatization

Di tahap akhir ini kita menyerdehanakan kata kata yang ada menjadi bentuk paling simple seperti Enjoyed menjadi enjoy Dll.

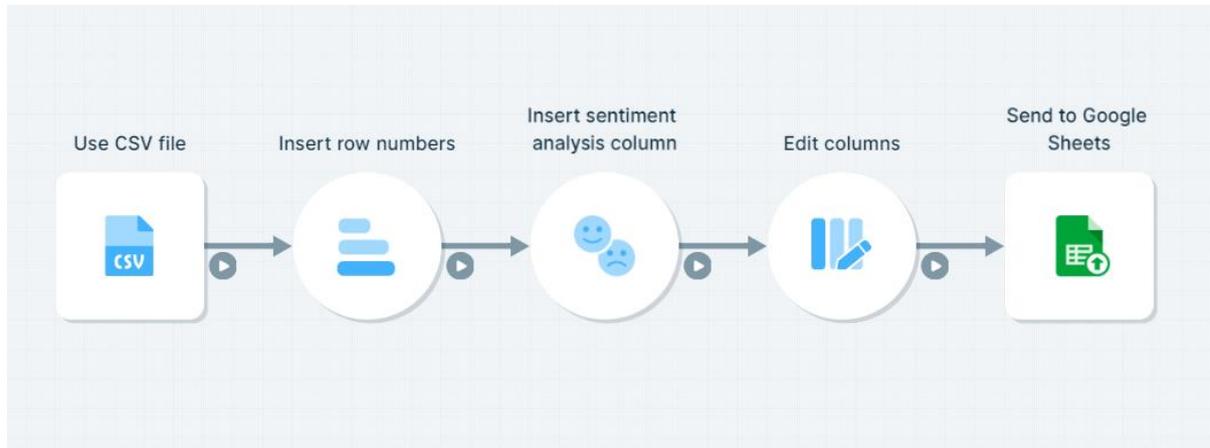
	content	clean_text	lemmatized_text
0	I like this app for the first time usage,it ha...	like app first time usageeit speed delivery cer...	like app first time usageeit speed delivery cer...
1	Can you make more ads page at the app? So my p...	make ads page app phone go lagi love ads	make ad page app phone go lagi love ad
2	i have enjoyed this marketplace and easy to use	enjoyed marketplace easy use	enjoy marketplace easy use
3	Experienced frequent warning to update the app...	experienced frequent warning update app mandat...	experienced frequent warning update app mandat...
4	user friendly interface. Best e-commerce all t...	user friendly interface best ecommerce time	user friendly interface best ecommerce time

**Gambar 8.** Hasil Lemmatization

## 3.3 Pengkategorian Sentimen Positif dan Negatif

Data yang sudah diproses tadi akan masuk ke tahap Klasifikasi Sentimen Positiv atau Negatif dengan menggunakan Parabola.io yang memanfaatkan Google's Natural Language Processing API.

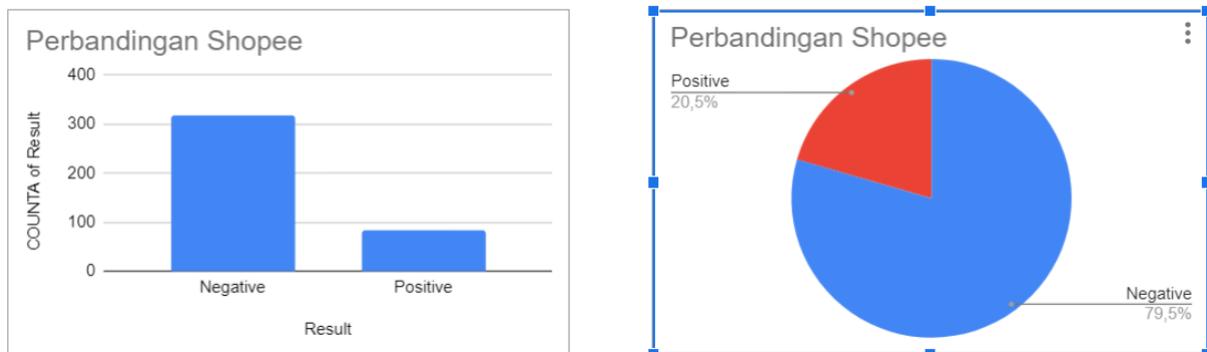
Untuk Langkahnya kita perlu memasukan data yang sudah di Pre-Processing tadi setelah itu memasukan kedalam Sentiment Analysis Column. Disinilah sistem akan melakukan Klasifikasi menggunakan Google Natural Language Processing API. Setelah itu kita hanya perlu export data tersebut, Disini penulis export ke Google Sheets agar bisa menganalisis dan menyimpan data data tersebut.



Gambar 9. Proses Pengkategorian Sentimen

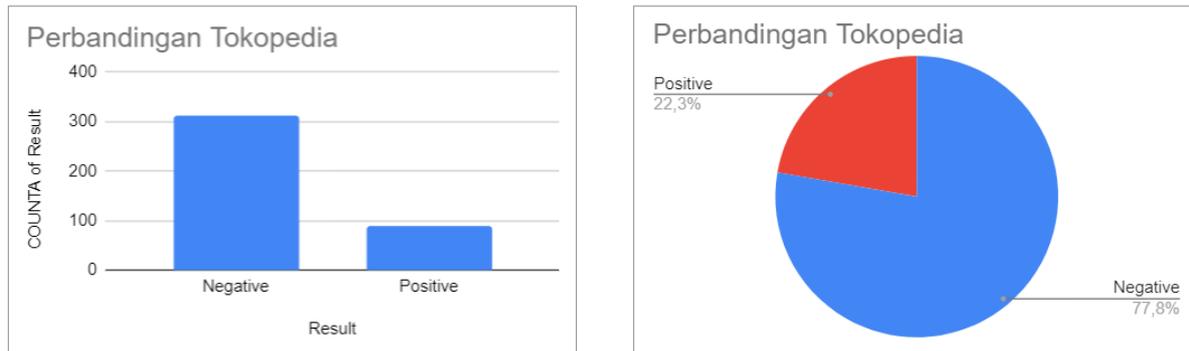
### 3.4 Hasil

Untuk Shopee mendapatkan 82 Review Positif dan 318 Review Negatif Dengan rata rata confidence 61%



Gambar 10. Hasil Sentimen Shopee

Untuk Tokopedia mendapatkan 89 Review Positive dan 311 Review Negatif Dengan rata rata confidence 61%



**Gambar 11.** Hasil Sentimen Tokopedia

#### 4. Conclusion

Berdasarkan pengujian Sentiment dari Review user Tokopedia dan user Shopee di Google Play Store dengan tahap tahap

1. Metode scrapping diambil secara acak dan merupakan review berbahasa inggris dan yang paling relevant sebanyak 400 review tokopedia dan 400 review shopee.
2. Metode Pre Processing yang melewati tahap tahap Convert Lowercase, Removal of Punctuations, Removal of Stopwords, Removal of Rare Words, Removal of Special characters, Lemmatization.
3. Dan Metode Klasifikasi Sentiment yang menggunakan Google's Natural Language Processing API

Didapatkan hasil Shopee mendapatkan 82 Review Positif dan 318 Review Negatif Dengan rata rata confidence 61% dan Tokopedia mendapatkan 89 Review Positif dan 311 Review Negatif Dengan rata rata confidence 61%. Ternyata hasil review dari kedua E-Commerce ini di playstore bisa dikatakan kurang lebih memiliki Sentiment yang sama. Mayoritas Bersentiment Negatif dengan skor sekitar 80%. Tentu saja hal ini tidak baik bagi kedua Perusahaan ini yang harus lebih membenahi dan melihat keluhan Kostumer di Review Play Store.

#### Daftar Pustaka

- [1] Handayani RN. OPTIMASI ALGORITMA SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PADA ULASAN PRODUK TOKOPEDIA MENGGUNAKAN PSO. Media Inform Vol20 No2. 2021;20(2):97–108.
- [2] Saputra I, AJI PAMBUDI RS, DARONO HE, AMSURY F, FAHDIA MR, RAMADHAN B, et al. Analisis Sentimen Pengguna Marketplace Bukalapak dan Tokopedia di Twitter Menggunakan Machine Learning. Fakt Exacta. 2021;13(4):200.
- [3] Resa M, Yudianto A, Rahim A, Sukmasetya P, Hasani RA. Perbandingan Metode Support Vector Machine Dengan Metode Lexicon Dalam Analisis Sentimen Bahasa Indonesia. J Teknol Informasi [Internet]. 2022;6(1). Available from: <https://github.com/fajri91/InSet>.
- [4] Lilyani Asri Utami. Analisis Sentimen Opini Publik Berita Kebakaran Hutan Melalui Komparasi Algoritma Support Vector Machine Dan K-Nearest Neighbor Berbasis Particle Swarm Optimization. J Pilar Nusa Mandiri [Internet]. 2017;13(1):103–12. Available from: <https://ejournal.nusamandiri.ac.id/index.php/pilar/article/view/153>
- [5] Yusup M, Jasman P, Renita D. Penerapan Algoritma Lemmatization pada Dokumen Bahasa Indonesia. MIND J. 2018;3(2):47–56.

