JURNAL KOMPREHENSIF
E ISSN: 3031-0970 | Vol. 4 No. 2 Juli-Desember 2026 | pp. 170-188
Available online at: https://ejournal.edutechjaya.com/index.php/komprehensif

Analisis Sentimen Pengguna *Iphone* Pada Platform *E-Commerce* Menggunakan Algoritma *Text Mining* dan *TF-IDF*

Tegar Yudhistyra[1], Zuli Agustina Gultom[2]

[1,2]Universitas Muhammadiyah Sumatera Utara, Indonesia

Email: tgryudhistyra@gmail.com[1], zuliagustina@umsu.ac.id[2]

ABSTRAK

Ulasan pengguna pada platform *e-commerce* merupakan sumber data teks yang sangat berharga untuk mengetahui persepsi konsumen terhadap suatu produk. Namun, jumlah ulasan yang besar dan bersifat tidak terstruktur menyulitkan proses pemahaman informasi secara manual. Penelitian ini bertujuan untuk menganalisis sentimen pengguna *iPhone* pada platform *e-commerce* Shopee menggunakan pendekatan *text mining* dan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Data yang digunakan berjumlah 250 ulasan berbahasa Indonesia yang diperoleh dari kolom ulasan produk *iPhone* di Shopee. Tahapan penelitian meliputi praproses teks (*case folding*, *tokenizing*, *filtering*, dan *stemming*), visualisasi *word cloud*, klasifikasi kata konotasi positif dan negatif, serta pembobotan kata menggunakan metode TF-IDF untuk menentukan kecenderungan sentimen. Hasil penelitian menunjukkan bahwa sentimen positif memperoleh persentase sebesar 77,80%, sedangkan sentimen negatif sebesar 22,20%, yang mengindikasikan bahwa mayoritas pengguna merasa puas terhadap produk *iPhone* yang dibeli melalui Shopee, terutama pada aspek keamanan pengiriman, kualitas, dan keaslian produk. Evaluasi kinerja model menggunakan *Confusion Matrix* menghasilkan nilai *Accuracy* sebesar 99,20%, *Precision* 99,47%, *Recall* 99,47%, *F1-Score* 99,47%, dan *Cohen's Kappa* sebesar 97,88% yang termasuk kategori *Almost Perfect Agreement*. Hasil ini membuktikan bahwa metode *text mining* dan TF-IDF efektif digunakan untuk menganalisis sentimen ulasan pengguna secara objektif dan akurat. Penelitian ini juga menghasilkan sistem analisis sentimen berbasis web yang menyajikan hasil analisis melalui dashboard interaktif berupa statistik sentimen, distribusi sentimen, *word cloud*, dan riwayat ulasan.
Kata Kunci: Analisis Sentimen, *iPhone*, *E-commerce*, *Text Mining*, *TF-IDF*, Shopee

ABSTRACT

User reviews on e-commerce platforms are a valuable source of text data for understanding consumer perceptions of a product. However, the large volume of unstructured reviews makes manual interpretation difficult. This study aims to analyze the sentiment of iPhone users on the Shopee e-commerce platform using a text mining approach combined with the Term Frequency-Inverse Document Frequency (TF-IDF) method. The dataset consists of 250 Indonesian-language reviews collected from the iPhone product review section on Shopee. The research stages include text preprocessing (case folding, tokenizing, filtering, and stemming), word cloud visualization, classification of positive and negative connotation words, and word weighting using the TF-IDF method to determine sentiment tendencies. The results show that positive sentiment accounted for 77.80%, while negative sentiment accounted for 22.20%, indicating that the majority of users were satisfied with the iPhone products purchased through Shopee, particularly regarding shipping safety, product quality, and authenticity. Model performance evaluation using a Confusion Matrix produced an Accuracy of 99.20%, Precision of 99.47%, Recall of 99.47%, an F1-Score of 99.47%, and a Cohen's Kappa value of 97.88%, which falls into the Almost Perfect Agreement category. These results demonstrate that the text mining and TF-IDF methods are effective for analyzing user review sentiment objectively and accurately. This research also produced a web-based sentiment

analysis system that presents the analysis results through an interactive dashboard featuring sentiment statistics, sentiment distribution, a word cloud, and review history
Keywords: Sentiment Analysis, iPhone, E-commerce, Text Mining, TF-IDF, Shopee

PENDAHULUAN

Analisis sentimen merupakan metode yang digunakan untuk mengetahui pendapat atau respon seseorang terhadap suatu objek berdasarkan teks yang ditulis. Pendapat tersebut dikelompokkan ke dalam kategori positif atau negatif sehingga lebih mudah untuk dipahami (Musfiroh et al., 2024). Sentimen positif bisa didapatkan dari hasil ulasan yang diberikan oleh pengguna karena merasa puas dengan produk yang didapatkan. Sedangkan sentimen negatif muncul akibat ketidakpuasan pengguna terhadap produk atau bahkan layanan yang didapatkan dari seller maupun pihak pengiriman. *E-commerce* merupakan sistem jual beli berbasis internet yang memungkinkan pengguna untuk mencari, membeli, dan memberikan penilaian terhadap suatu produk secara langsung tanpa harus bertatap muka. Berdasarkan data yang dirilis oleh Badan Pusat Statistik, jumlah transaksi *e-commerce* di Indonesia terus tumbuh secara signifikan, yang berdampak pada semakin banyaknya data ulasan pengguna yang tersedia di berbagai platform belanja *online* (Badan Pusat Statistik, 2023). Jumlah ulasan yang sangat banyak membuat informasi di dalamnya sulit dipahami jika dibaca satu per satu, sehingga diperlukan pendekatan pengolahan data yang tepat agar informasi yang terkandung di dalamnya dapat lebih mudah dipahami dan dimanfaatkan (Gultom et al., 2025).

Data mining adalah proses untuk menemukan pola atau informasi penting dari kumpulan data dalam jumlah besar. Dalam konteks penelitian ini, *data mining* digunakan untuk menggali informasi dari data ulasan pengguna yang tersedia di platform *e-commerce*. Karena data yang digunakan berbentuk teks bebas yang tidak terstruktur, maka pendekatan yang digunakan adalah *text mining*, yaitu teknik untuk mengolah dan menganalisis data teks agar dapat menghasilkan informasi yang bermakna. Setelah data teks diolah melalui *text mining*, diperlukan suatu metode lanjutan untuk memahami kecenderungan isi dari teks tersebut. Analisis sentimen adalah metode yang digunakan untuk mengetahui kecenderungan opini seseorang terhadap suatu objek berdasarkan teks yang ditulis, di mana ulasan pengguna dapat dikelompokkan menjadi sentimen positif atau negatif, sehingga data yang sebelumnya sulit dipahami dapat diubah menjadi informasi yang lebih sederhana dan mudah diinterpretasikan (Musfiroh et al., 2024).

Penelitian ini memilih produk *iPhone* sebagai objek kajian karena *iPhone* merupakan salah satu produk *smartphone* yang memiliki banyak pengguna di Indonesia dan secara konsisten mendapatkan ulasan dari konsumen di berbagai platform *e-commerce*. Ulasan pengguna *iPhone* cenderung memuat penilaian yang cukup spesifik terhadap beberapa aspek produk, seperti kondisi baterai, kualitas layar, kemampuan kamera, performa perangkat, pengiriman serta pengalaman terhadap

layanan penjual (*seller*). Keberagaman aspek ini menjadikan data ulasan *iPhone* sangat relevan untuk dianalisis menggunakan pendekatan analisis sentimen, karena dari data tersebut dapat diketahui aspek mana yang paling banyak mendapat respons positif maupun negatif dari konsumen. Platform *Shopee* dipilih sebagai sumber data dalam penelitian ini karena merupakan salah satu platform *e-commerce* dengan jumlah pengguna terbesar di Indonesia dan memiliki fitur ulasan yang aktif digunakan oleh konsumen setelah melakukan transaksi. *Shopee* secara konsisten menempati posisi teratas di antara platform *marketplace* di Indonesia berdasarkan jumlah pengunjung aktif, sehingga data ulasan yang tersedia di platform ini sangat representatif untuk digunakan dalam penelitian analisis sentimen karena mencerminkan pengalaman nyata dari banyak pengguna yang berbeda-beda (Sari & Kusuma, 2023).

Metode *text mining* dipilih karena mampu mengolah data teks yang tidak terstruktur menjadi data yang lebih teratur dan siap dianalisis. Selain itu, metode *Term Frequency-Inverse Document Frequency* (*TF-IDF*) digunakan karena mampu memberikan bobot pada kata berdasarkan tingkat kepentingannya dalam suatu dokumen, sehingga proses analisis sentimen menjadi lebih terarah dan didasarkan pada nilai bobot kata yang dihitung secara sistematis (Septiani & Isabela, 2022). Jika dibandingkan dengan penelitian terdahulu, penelitian ini memiliki beberapa perbedaan yang cukup jelas. Penelitian oleh Simatupang & Utomo (2019) menggunakan *text mining* dan *TF-IDF*, namun objek yang digunakan hanya terbatas pada satu toko *online*. Penelitian oleh Najiyah & Hariyanti (2021) juga menggunakan *TF-IDF*, tetapi fokus pada data media sosial terkait COVID-19, bukan pada ulasan produk. Sementara itu, penelitian yang menggunakan *Naïve Bayes* atau *Support Vector Machine* lebih menekankan pada proses klasifikasi, bukan pada pembobotan kata sebagai fokus utama. Berdasarkan hal tersebut, penelitian ini hadir untuk mengisi celah tersebut dengan mengkaji secara khusus sentimen pengguna *iPhone* pada platform *Shopee* menggunakan pendekatan *text mining* dan *TF-IDF*, sekaligus memperhatikan aspek-aspek spesifik yang paling berpengaruh dalam ulasan pengguna.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *text mining* untuk menganalisis sentimen ulasan pengguna iPhone pada platform *e-commerce* Shopee. Data penelitian berupa 250 ulasan pengguna yang dikumpulkan dari komentar pembeli terhadap produk iPhone. Ulasan yang digunakan hanya berbentuk teks agar proses analisis dapat dilakukan secara konsisten.

Tahapan penelitian meliputi pengumpulan data, *preprocessing* teks, pembobotan kata menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF), analisis sentimen, dan evaluasi hasil. Proses *preprocessing* terdiri atas *case folding*, *tokenizing*, *stopword removal*, dan *stemming* untuk menghasilkan data teks yang bersih dan siap diolah.

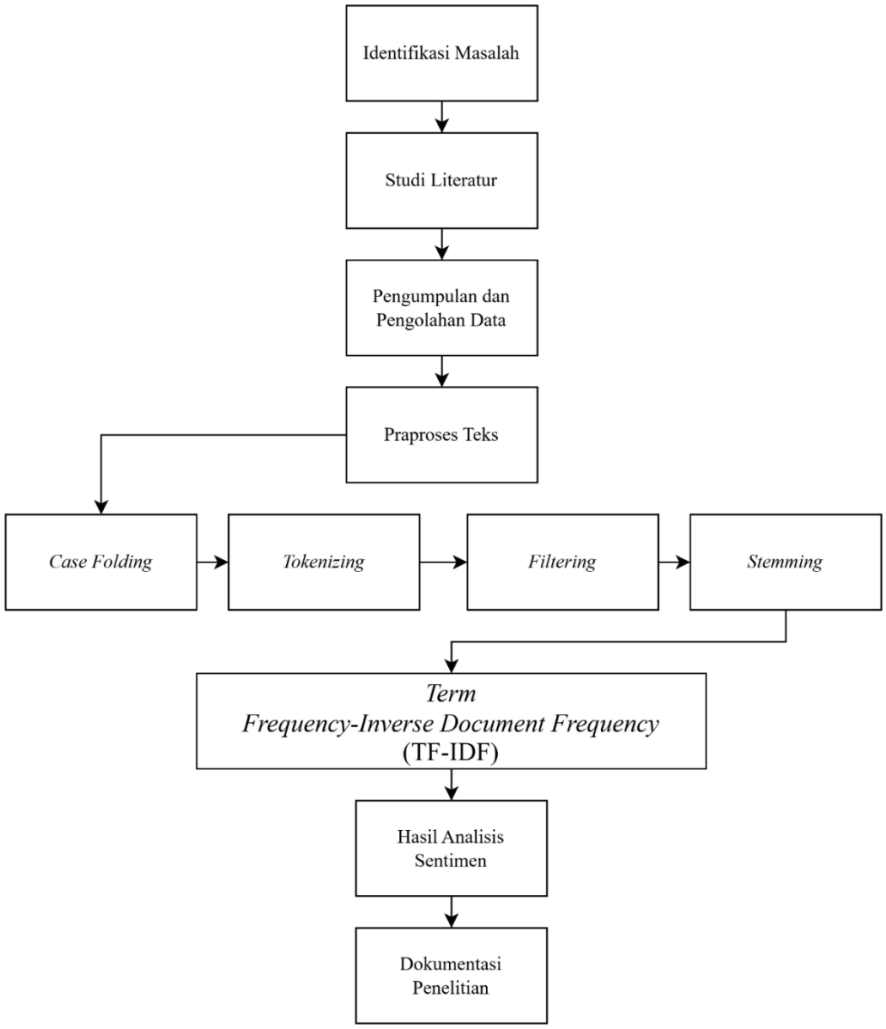
Selanjutnya, setiap kata diberi bobot menggunakan metode TF-IDF untuk mengidentifikasi kata-kata yang paling representatif dalam setiap ulasan. Hasil

JURNAL KOMPREHENSIF
E ISSN: 3031-0970 | Vol. 4 No. 2 Juli-Desember 2026 | pp. 170-188
Available online at: https://ejournal.edutechjaya.com/index.php/komprehensif

pembobotan digunakan untuk menentukan kecenderungan sentimen positif dan negatif pada data ulasan.

Kinerja model dievaluasi menggunakan *Confusion Matrix* dengan metrik *Accuracy, Precision, Recall*, dan *F1-Score*. Selain itu, penelitian ini juga menghitung Cohen's Kappa untuk mengukur tingkat kesepakatan antara hasil klasifikasi sistem dan pelabelan manual (*gold standard*), sehingga diperoleh gambaran yang lebih objektif mengenai performa model analisis sentimen.

Pada penelitian ini melalui kerangka kerja penelitian atau yang sering disebut dengan tahapan-tahapan dalam penelitian. Berikut merupakan kerangka kerja dalam proses penelitian ini.



Gambar 1. Kerangka Penelitian

HASIL DAN PEMBAHASAN

1. Penerapan *Case Folding*

Seluruh huruf dalam teks ulasan diubah menjadi huruf kecil dan tanda baca dihapus sehingga data menjadi seragam dan konsisten.

JURNAL KOMPREHENSIF
E ISSN: 3031-0970 | Vol. 4 No. 2 Juli-Desember 2026 | pp. 170-188
Available online at: https://ejournal.edutechjaya.com/index.php/komprehensif

Tabel 1 Hasil *Case Folding*

Ulasan Awal	**Hasil *Case Folding***
Kualitas dan respon sangat ok..video nya menyusul..karena blom sempet dibuka boxnya..	kualitas dan respon sangat okvideo nya menyusulkarena blom sempet dibuka boxnya

2. Penerapan *Tokenizing*

Teks ulasan yang telah melalui proses *case folding* dipecah menjadi satuan kata tunggal yang disebut *token*. Setiap kata dipisahkan satu sama lain sehingga dapat diproses secara individual pada tahapan berikutnya.

Tabel 2 Hasil *Tokenizing*

Ulasan Setelah *Case Folding*	**Hasil *Tokenizing***
kualitas dan respon sangat okvideo nya menyusulkarena blom sempet dibuka boxnya	[kualitas] [dan] [respon] [sangat] [ok] [video] [nya] [menyusul] [karena] [blom] [sempet] [dibuka] [boxnya]

3. Penerapan *Filtering*

Kata-kata yang tidak memiliki makna penting dalam analisis sentimen dihapus dari hasil *tokenizing* sebelumnya. Kata-kata yang dihapus merupakan *stopword*, yaitu kata umum seperti kata hubung dan kata depan yang tidak memberikan kontribusi terhadap penentuan sentimen.

Tabel 3 Hasil *Filtering*

Ulasan Setelah *Case Folding* dan *Tokenizing*	**Hasil *Filtering***
[kualitas] [dan] [respon] [sangat] [ok] [video] [nya] [menyusul] [karena] [blom] [sempet] [dibuka] [boxnya]	[kualitas] [respon] [video] [menyusul] [dibuka] [boxnya]

4. Penerapan *Stemming*

Setiap kata hasil *filtering* dikembalikan ke bentuk dasarnya dengan menghilangkan imbuhan berupa awalan maupun akhiran. Proses ini dilakukan agar variasi kata yang memiliki makna dasar yang sama dapat direpresentasikan dalam satu bentuk yang seragam.

JURNAL KOMPREHENSIF
E ISSN: 3031-0970 | Vol. 4 No. 2 Juli-Desember 2026 | pp. 170-188
Available online at: https://ejournal.edutechjaya.com/index.php/komprehensif

Tabel 4 Hasil *Stemming*

Ulasan Setelah *Case Folding*, *Tokenizing* dan *Filtering*	**Hasil *Stemming***
[kualitas] [respon] [video] [menyusul] [dibuka] [boxnya]	[kualitas] [respon] [video] [susul] [buka] [box]

Setelah data diolah dengan algoritma *text mining*, maka didapat hasil keseluruhannya. Proses pengolahan data sebanyak 250 data yang di proses penerapan preprocessing text.

Setelah melakukan penerapan algoritma *text mining* pada analisis ulasan produk *iPhone*, selanjutnya dilakukan pengolahan data untuk menghasilkan visualisasi berupa *word cloud* guna melihat kata-kata yang paling sering muncul dari 250 data ulasan yang diambil dari platform *e-commerce Shopee*. Pembuatan *word cloud* dilakukan menggunakan layanan *website* pembuat *word cloud* berbasis *online*. Berikut tampilan hasil *word cloud* dari data ulasan tersebut dapat dilihat pada gambar di bawah ini.

2. Penerapan *Word Cloud* pada Ulasan Produk *iPhone*

Setelah seluruh data melewati tahapan praproses *text mining*, hasil pengolahan kata kemudian divisualisasikan dalam bentuk *word cloud* untuk mengetahui kata-kata yang paling sering muncul dari 250 ulasan yang dikumpulkan. Diperoleh 10 kata dengan frekuensi tertinggi pada konotasi positif, yaitu *aman*, *bagus*, *sesuai*, *original*, *silver*, *bonus*, *kualitas*, *harga*, *kamera*, dan *garansi*. Sementara itu, 10 kata dengan frekuensi tertinggi pada konotasi negatif adalah *buruk*, *rusak*, *kendala*, *kecewa*, *tipu*, *palsu*, *hilang*, *cacat*, *lecet*, dan *repot*. Kata-kata tersebut selanjutnya digunakan sebagai dasar dalam proses pembobotan menggunakan metode *TF-IDF*.

Gambar 2 Hasil *Word Cloud* Ulasan Produk *iPhone* pada platform *Shopee*

3. Penerapan *Term Frequency-Inverse Document Frequency (TF-IDF)*

Setelah kata-kata konotasi positif dan negatif diperoleh dari hasil *word cloud*, tahap selanjutnya adalah menghitung bobot setiap kata menggunakan metode *TF-IDF*.

Untuk memperoleh nilai *TF-IDF*, terlebih dahulu dihitung nilai *Term Frequency* (*TF*), *Document Frequency* (*DF*), dan *Inverse Document Frequency* (*IDF*). Nilai *TF*

menunjukkan jumlah kemunculan suatu kata dalam seluruh ulasan, sedangkan *DF* menunjukkan jumlah dokumen yang mengandung kata tersebut. Jumlah dokumen yang digunakan pada penelitian ini sebanyak 250 ulasan sehingga nilai N = 250. Nilai *IDF* dihitung menggunakan rumus berikut.

$$IDF = Log\left(\frac{N}{DF}\right) + 1$$

Keterangan:

IDF = Nilai *Inverse Document Frequency*
N = Jumlah seluruh dokumen atau ulasan yang digunakan dalam penelitian
DF = Jumlah dokumen yang mengandung kata yang dihitung
Log = Logaritma basis 10

Sebagai contoh, perhitungan *IDF* pada kata *aman* adalah sebagai berikut.

$$IDF = Log\left(\frac{250}{39}\right) + 1$$
$$IDF = Log\ (6{,}410256) + 1$$
$$IDF = 0{,}806875 + 1$$
$$IDF = 1{,}806875$$

Setelah nilai *IDF* diperoleh, langkah berikutnya adalah menghitung bobot *TF-IDF* menggunakan rumus berikut.

$$TF\ IDF = TF \times IDF$$

Keterangan:
TF IDF = Bobot kata
TF = Frekuensi kemunculan kata pada dokumen
IDF = Tingkat keunikan kata pada seluruh dokumen

Contoh perhitungan bobot *TF-IDF* pada kata *aman* adalah sebagai berikut.

$$TF\ IDF = 44 \times 1{,}806875$$
$$TF\ IDF = 79{,}5025$$

Hasil perhitungan tersebut kemudian diterapkan pada seluruh kata konotasi positif dan negatif.

Berdasarkan hasil total bobot konotasi positif sebesar 350,1511 dan total bobot konotasi negatif sebesar 99,8985. Ringkasan hasil pembobotan ditampilkan berikut.

Tabel 6 Hasil Bobot Sentimen Positif dan Negatif

Sentimen	Bobot
Positif	350,1511
Negatif	99,8985

Persentase sentimen dihitung dengan membandingkan total bobot masing-masing konotasi terhadap keseluruhan bobot *TF-IDF* yang diperoleh.

Sentimen Positif

$$PresentasePositif = \frac{Total\ Bobot\ Positif}{Total\ Bobot\ Positif + Total\ Bobot\ Negatif} \times 100\%$$

$$Presentase\ Positif = \frac{350{,}1511}{(350{,}1511 + 99{,}8985)} \times 100\%$$
$$Presentase\ Positif = \frac{350{,}1511}{450{,}0496} \times 100\%$$
$$Presentase\ Positif = 77{,}80\%$$

Sentimen Negatif

$$PresentaseNegatif = \frac{Total\ Bobot\ Negatif}{Total\ Bobot\ Positif + Total\ Bobot\ Negatif} \times 100\%$$

$$Presentase\ Negatif = \frac{99{,}8985}{(350{,}1511 + 99{,}8985)} \times 100\%$$
$$Presentase\ Negatif = \frac{99{,}8985}{450{,}0496} \times 100\%$$
$$Presentase\ Negatif = 22{,}20\%$$

Hasil perhitungan persentase sentimen ditampilkan berikut ini.

Tabel 7 Hasil Sentimen

Sentimen	Hasil Sentimen	Jumlah Sentimen
Positif	77,80%	187
Negatif	22,20%	63
Total Hasil Sentimen	100%	250

Berdasarkan hasil analisis menggunakan metode *Term Frequency-Inverse Document Frequency* (*TF-IDF*), diperoleh sentimen positif sebesar 77,80% dan sentimen negatif sebesar 22,20%. Hasil tersebut menunjukkan bahwa sebagian besar pengguna memberikan penilaian positif terhadap produk yang diteliti. Tingginya persentase sentimen positif mengindikasikan bahwa pengguna merasa puas terhadap berbagai aspek produk, seperti keamanan, kualitas, kesesuaian produk, keaslian produk, harga, kamera, dan garansi. Sementara itu, sentimen negatif yang muncul umumnya berkaitan dengan kondisi produk yang rusak, kendala penggunaan, serta pengalaman yang kurang memuaskan setelah pembelian. Berdasarkan hasil tersebut, dapat disimpulkan bahwa produk yang diteliti memperoleh respons yang cenderung positif dari pengguna berdasarkan analisis terhadap 250 ulasan pada platform *Shopee*.

4. Evaluasi Ketepatan Akurasi Model Analisis Sentimen

Evaluasi kinerja model dilakukan untuk mengetahui tingkat kemampuan sistem dalam mengklasifikasikan sentimen ulasan pengguna *iPhone* ke dalam kategori positif dan negatif. Pengujian dilakukan menggunakan 250 data yang telah diberi label sentimen secara manual oleh peneliti. Hasil klasifikasi sistem kemudian dibandingkan dengan label aktual menggunakan *Confusion Matrix*.

JURNAL KOMPREHENSIF
E ISSN: 3031-0970 | Vol. 4 No. 2 Juli-Desember 2026 | pp. 170-188
Available online at: https://ejournal.edutechjaya.com/index.php/komprehensif

```
[42]
# Upload File Excel
from google.colab import files

uploaded = files.upload()
```

```
Pilih File  Akurasi.xlsx
Akurasi.xlsx(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 30812 bytes, last modified: 20/6/2026 - 100% done
Saving Akurasi.xlsx to Akurasi (1).xlsx
```

```
[4]
# Baca Data Excel
import pandas as pd

df = pd.read_excel("Akurasi.xlsx")

df.head()
```

	No	Tanggal	Pengguna	Ulasan	Hasil Sistem	Interpretasi Manual	Keterangan
0	1	03/01/2024	b*****2	Alhamdulillah bisa kebeli hp idaman	Positif	Positif	Sesuai
1	2	08/01/2024	j*****p	Yang dikirim pecahan keramik	Negatif	Positif	Tidak Sesuai
2	3	12/01/2024	a*****1	Barang sampai dgn aman	Positif	Positif	Sesuai
3	4	26/01/2024	s*****6	Kecepatan pengiriman baik, produk original, se...	Positif	Positif	Sesuai
4	5	02/02/2024	t*****1	Alhamdulillah barang smpai dengan aman ke aceh...	Positif	Positif	Sesuai

Langkah berikutnya: Buat kode dengan df New interactive sheet

```
[5]
# Install Library
!pip install scikit-learn
```

```
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.12/dist-packages (1.6.1)
Requirement already satisfied: numpy>=1.19.5 in /usr/local/lib/python3.12/dist-packages (from scikit-learn) (2.0.2)
Requirement already satisfied: scipy>=1.6.0 in /usr/local/lib/python3.12/dist-packages (from scikit-learn) (1.16.3)
Requirement already satisfied: joblib>=1.2.0 in /usr/local/lib/python3.12/dist-packages (from scikit-learn) (1.5.3)
Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/lib/python3.12/dist-packages (from scikit-learn) (3.6.0)
```

Gambar 3 *Upload*, Baca dan *Install Library Excel*

```
[40]
0 d
from sklearn.metrics import confusion_matrix
import pandas as pd

cm = confusion_matrix(y_true, y_pred, labels=['Positif','Negatif'])

cm_df = pd.DataFrame(
    cm,
    index=['Prediksi Positif','Prediksi Negatif'],
    columns=['Aktual Positif','Aktual Negatif']
)

print("CONFUSION MATRIX")
print("="*50)
print(cm_df)
```

```
CONFUSION MATRIX
==================================================
                  Aktual Positif  Aktual Negatif
Prediksi Positif             186               1
Prediksi Negatif               1              62
```

Gambar 4 *Confusion Matrix*

Keterangan:

True Positive (TP) = 186
True Negative (TN) = 62
False Positive (FP) = 1
False Negative (FN) = 1

Berdasarkan nilai pada *Confusion Matrix* tersebut, dilakukan perhitungan beberapa metrik evaluasi yaitu *Accuracy*, *Precision*, *Recall*, *F1-Score*, dan *Cohen's Kappa.*

a. Perhitungan *Accuracy*

Accuracy merupakan ukuran yang digunakan untuk mengetahui tingkat ketepatan model dalam melakukan klasifikasi sentimen. Nilai *Accuracy* menunjukkan

persentase jumlah prediksi yang benar dibandingkan dengan keseluruhan data yang diuji.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Keterangan:

Accuracy = Tingkat akurasi model
TP = *True Positive*
TN = *True Negative*
FP = *False Positive*
FN = *False Negative*

```
[41]
✓ 0 d
# Nilai TP, TN, FP dan FN
# Perhitungan Accuracy

from IPython.display import display, Math

TP = 186
TN = 62
FP = 1
FN = 1

display(Math(r'Accuracy = \frac{TP + TN}{TP + TN + FP + FN}'))
display(Math(
rf'Accuracy = \frac{{{TP}+{TN}}}{{{TP}+{TN}+{FP}+{FN}}}'
))
display(Math(
rf'Accuracy = \frac{{248}}{{250}}'
))
display(Math(
r'Accuracy = 0.992'
))
display(Math(
r'Accuracy = 99.20\%'
))
```

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$Accuracy = \frac{186 + 62}{186 + 62 + 1 + 1}$$
$$Accuracy = \frac{248}{250}$$
$$Accuracy = 0.992$$
$$Accuracy = 99.20\%$$

Gambar 5 Pehitungan *Accuracy*

Berdasarkan hasil perhitungan diperoleh nilai *Accuracy* sebesar 99,20%. Nilai tersebut menunjukkan bahwa sistem mampu mengklasifikasikan 99,20% data ulasan dengan benar dari keseluruhan data yang diuji.

b. Perhitungan *Precision*

Precision digunakan untuk mengukur tingkat ketepatan model dalam memprediksi data yang termasuk ke dalam kelas positif. Nilai *Precision* yang tinggi menunjukkan bahwa sebagian besar data yang diprediksi positif oleh sistem memang benar-benar positif.

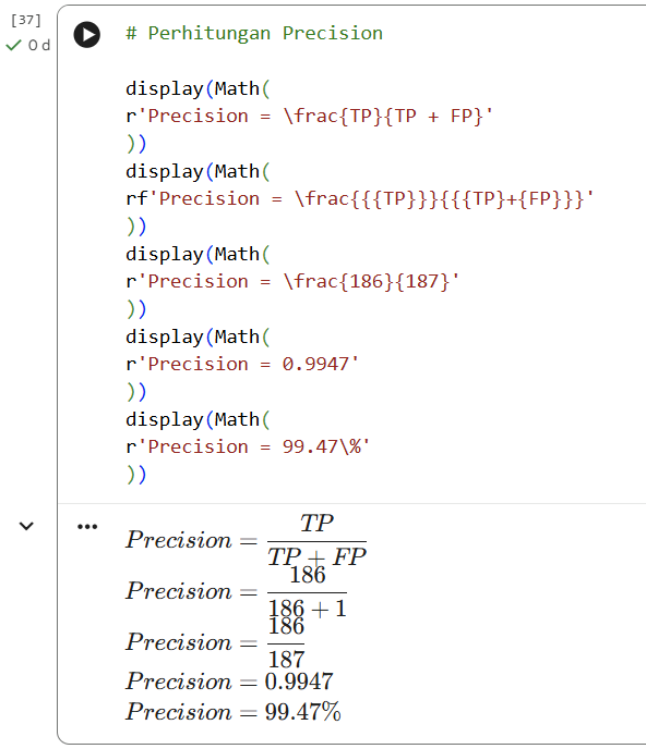
$$Precision = \frac{TP}{TP + FP}$$

Keterangan:

Precision = Tingkat ketepatan prediksi positif

TP = *True Positive*

FP = *False Positive*



Gambar 6 Perhitungan *Precision*

Berdasarkan hasil perhitungan diperoleh nilai *Precision* sebesar 99,47%. Hal ini menunjukkan bahwa dari seluruh ulasan yang diprediksi positif oleh sistem, sebanyak 99,47% merupakan ulasan yang benar-benar positif.

c. Perhitungan *Recall*

Recall digunakan untuk mengukur kemampuan model dalam menemukan seluruh data yang sebenarnya termasuk ke dalam kelas positif.

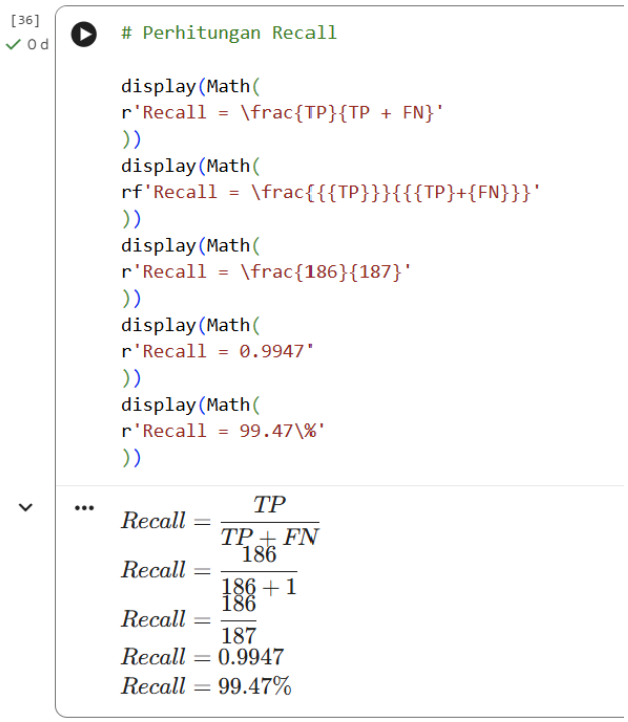
$$Recall = \frac{TP}{TP + FN}$$

Keterangan:

Recall = Tingkat keberhasilan menemukan data positif

TP = *True Positive*

FN = *False Negative*



Gambar 7 Perhitungan *Recall*

Berdasarkan hasil perhitungan diperoleh nilai *Recall* sebesar 99,47%. Nilai ini menunjukkan bahwa seluruh ulasan positif pada data uji berhasil dikenali dengan baik oleh sistem tanpa adanya kesalahan klasifikasi ke dalam kategori negatif.

d. Perhitungan *F1-Score*

F1-Score merupakan rata-rata harmonik antara *Precision* dan *Recall*. Metrik ini digunakan untuk memberikan gambaran keseimbangan antara ketepatan dan kemampuan model dalam menemukan data positif.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Keterangan:
F1-Score = Nilai keseimbangan antara *Precision* dan *Recall*
Precision = Tingkat ketepatan prediksi positif
Recall = Tingkat keberhasilan menemukan data positif

JURNAL KOMPREHENSIF
E ISSN: 3031-0970 | Vol. 4 No. 2 Juli-Desember 2026 | pp. 170-188
Available online at: https://ejournal.edutechjaya.com/index.php/komprehensif

```
[31]
✓ 0d
# Perhitungan F1-Scroe

display(Math(
r'F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}'
))
display(Math(
r'F1 = 2 \times \frac{0.9947 \times 0.9947}{0.9947 + 0.9947}'
))
display(Math(
r'F1 = 0.9947'
))
display(Math(
r'F1 = 99.47\%'
))
```

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$
$$F1 = 2 \times \frac{0.9947 \times 0.9947}{0.9947 + 0.9947}$$
$$F1 = 0.9947$$
$$F1 = 99.47\%$$

Gambar 8 Perhitungan *F1-Score*

Berdasarkan hasil perhitungan diperoleh nilai *F1-Score* sebesar 99,47%. Nilai tersebut menunjukkan bahwa model memiliki keseimbangan yang sangat baik antara kemampuan mengidentifikasi data positif dan ketepatan hasil prediksinya

e. Perhitungan *Cohen's Kappa*

Cohen's Kappa digunakan untuk mengukur tingkat kesesuaian antara hasil klasifikasi sistem dengan label aktual yang diberikan oleh peneliti. Berbeda dengan *Accuracy*, metode ini mempertimbangkan kemungkinan kesesuaian yang terjadi secara kebetulan.

1. Menghitung *Observed Agreement* (P_o)

Observed Agreement merupakan tingkat kesesuaian aktual antara hasil prediksi sistem dengan label yang diberikan peneliti.

$$P_o = \frac{TP + TN}{N}$$

Keterangan:

P_o= *Observed Agreement*
TP= *True Positive*
TN= *True Negative*
N= Jumlah data

JURNAL KOMPREHENSIF
E ISSN: 3031-0970 | Vol. 4 No. 2 Juli-Desember 2026 | pp. 170-188
Available online at: https://ejournal.edutechjaya.com/index.php/komprehensif

```
[29]
# Perhitungan Observed Agreement

from IPython.display import display, Math

TP = 186
TN = 62
FP = 1
FN = 1

N = TP + TN + FP + FN
Po = (TP + TN) / N

display(Math(r'P_o=\frac{TP+TN}{N}'))
display(Math(
rf'P_o=\frac{{{TP}+{TN}}}{{{N}}}'
))
display(Math(
r'P_o=\frac{248}{250}'
))
display(Math(
rf'P_o={Po:.4f}'
))
display(Math(
rf'P_o={Po*100:.2f}\%'
))
```

$$P_o = \frac{TP + TN}{N}$$
$$P_o = \frac{186 + 62}{250}$$
$$P_o = \frac{248}{250}$$
$$P_o = 0.9920$$
$$P_o = 99.20\%$$

Gambar 9 Perhitungan *Observed Agreement*

Berdasarkan hasil perhitungan yang telah dilakukan, diperoleh nilai *Observed Agreement* (P_o) sebesar 99,20%. Nilai tersebut menunjukkan tingkat kesesuaian yang sangat tinggi antara hasil klasifikasi yang dihasilkan oleh sistem dan hasil interpretasi manual yang digunakan sebagai acuan penelitian. Tingginya nilai *Observed Agreement* mengindikasikan bahwa sebagian besar data berhasil diklasifikasikan secara tepat oleh sistem, dengan jumlah ketidaksesuaian yang sangat kecil. Hasil ini mencerminkan bahwa proses analisis sentimen yang diterapkan memiliki tingkat kecocokan yang sangat baik terhadap data aktual yang digunakan dalam penelitian.

2. Menghitung *Expected Agreement* (P_e)

Expected Agreement merupakan tingkat kesesuaian yang mungkin terjadi secara kebetulan.

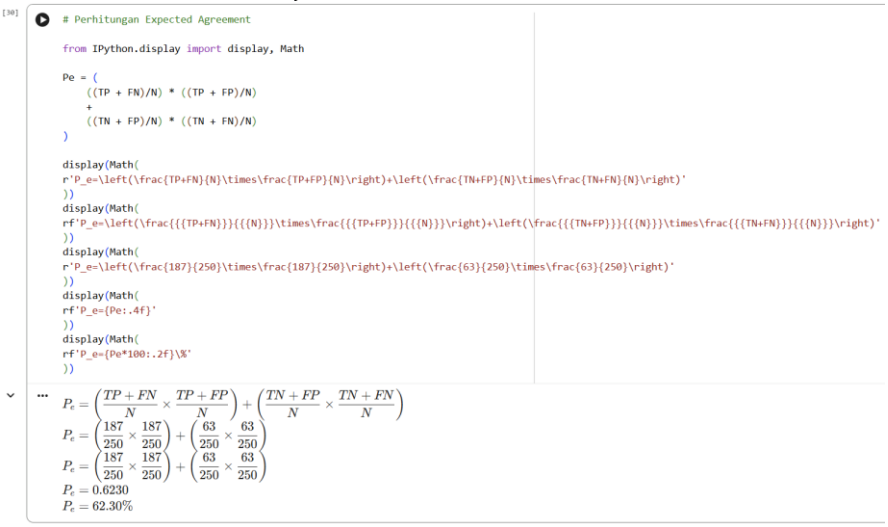
$$P_e = \left(\frac{(TP + FN)}{N} \times \frac{(TP + FP)}{N}\right) + \left(\frac{(TN + FP)}{N} \times \frac{(TN + FN)}{N}\right)$$

Keterangan:

P_e= *Expected Agreement*
TP= *True Positive*
TN= *True Negative*

JURNAL KOMPREHENSIF
E ISSN: 3031-0970 | Vol. 4 No. 2 Juli-Desember 2026 | pp. 170-188
Available online at: https://ejournal.edutechjaya.com/index.php/komprehensif

FP= *False Positive*
FN= *False Negative*
N= Jumlah seluruh data uji



Gambar 10 Perhitungan *Expected Agreement*

Berdasarkan hasil perhitungan yang telah dilakukan, diperoleh nilai *Expected Agreement* (P_e) sebesar 62,30%. Nilai tersebut menunjukkan tingkat kesesuaian yang berpotensi terjadi secara kebetulan antara hasil klasifikasi sistem dan label acuan yang digunakan dalam penelitian. *Expected Agreement* digunakan sebagai dasar untuk mengukur sejauh mana kesepakatan yang terbentuk bukan disebabkan oleh faktor acak. Nilai P_e sebesar 62,30% mengindikasikan bahwa sebagian tingkat kesesuaian yang ditemukan masih dapat terjadi secara kebetulan, sehingga diperlukan pengukuran lanjutan melalui koefisien *Cohen's Kappa* untuk memperoleh tingkat kesepakatan yang sesungguhnya setelah pengaruh kesepakatan acak diperhitungkan.

3. Menghitung Nilai *Cohen's Kappa*

Menghitung nilai *Cohen's Kappa* dapat menggunakan rumus berikut:

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

Keterangan:

κ= Nilai *Cohen's Kappa*
P_o= *Observed Agreement*
P_e= *Expected Agreement*

JURNAL KOMPREHENSIF
E ISSN: 3031-0970 | Vol. 4 No. 2 Juli-Desember 2026 | pp. 170-188
Available online at: https://ejournal.edutechjaya.com/index.php/komprehensif

```
[27]
# Perhitungan Nilai Cohen's Kappa

Kappa = (Po - Pe) / (1 - Pe)

display(Math(
r'\kappa=\frac{P_o-P_e}{1-P_e}'
))

display(Math(
rf'\kappa=\frac{{{Po:.4f}-{Pe:.4f}}}{{1-{Pe:.4f}}}'
))

display(Math(
rf'\kappa={Kappa:.4f}'
))

display(Math(
rf'\kappa={Kappa*100:.2f}\%'
))
```

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$
$$\kappa = \frac{0.9920 - 0.6230}{1 - 0.6230}$$
$$\kappa = 0.9788$$
$$\kappa = 97.88\%$$

Gambar 11 Perhitungan nilai *Cohen's Kappa*

Berdasarkan hasil perhitungan diperoleh nilai *Cohen's Kappa* sebesar 97,88%. Nilai tersebut termasuk dalam kategori *Almost Perfect Agreement* atau kesesuaian hampir sempurna. Hal ini menunjukkan bahwa hasil klasifikasi yang dihasilkan sistem memiliki tingkat kesesuaian yang sangat tinggi dengan interpretasi manual yang dilakukan oleh peneliti.

Tabel 8 Hasil Evaluasi Model

Metrik Evaluasi	**Nilai**
Accuracy	99,20%
Precision	99,47%
Recall	99,47%
F1-Score	99,47%
Cohen's Kappa	97,88%

Berdasarkan hasil evaluasi yang diperoleh, model analisis sentimen yang menerapkan metode *Text Mining* dan *TF-IDF* menunjukkan performa yang sangat baik. Nilai *Accuracy* sebesar 99,20% mengindikasikan bahwa sebagian besar data ulasan berhasil diklasifikasikan secara tepat sesuai dengan label acuan. Nilai *Precision* sebesar 99,47% mencerminkan tingkat ketepatan yang sangat tinggi dalam menghasilkan prediksi sentimen positif. Nilai *Recall* sebesar 99,47% menunjukkan kemampuan model yang sangat baik dalam mengidentifikasi data yang termasuk ke dalam kategori sentimen positif. Nilai *F1-Score* sebesar 99,47% menggambarkan kualitas kinerja klasifikasi yang sangat tinggi. Nilai *Cohen's Kappa* sebesar 97,88% menunjukkan tingkat kesepakatan yang sangat kuat antara hasil klasifikasi sistem dan interpretasi manual. Hasil evaluasi tersebut menunjukkan bahwa model yang

dikembangkan memiliki tingkat keandalan yang sangat baik dalam melakukan analisis sentimen terhadap ulasan pengguna *iPhone*.

KESIMPULAN

Berdasarkan hasil penelitian mengenai analisis sentimen pengguna iPhone pada platform *e-commerce* menggunakan metode *text mining* dan pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF) terhadap 250 data ulasan yang diperoleh dari Shopee, dapat disimpulkan bahwa proses *preprocessing* teks yang meliputi *case folding*, *tokenizing*, *filtering* (*stopword removal*), dan *stemming* berhasil diterapkan secara sistematis sehingga data ulasan yang semula tidak terstruktur dapat diubah menjadi data yang lebih bersih dan siap untuk dianalisis. Selanjutnya, metode TF-IDF berhasil digunakan untuk memberikan bobot pada setiap kata berdasarkan tingkat kepentingannya dalam kumpulan dokumen, sehingga mampu mengidentifikasi kata-kata yang paling representatif dalam ulasan pengguna. Berdasarkan hasil analisis sentimen yang dilakukan, diperoleh bahwa mayoritas ulasan pengguna iPhone di platform Shopee menunjukkan kecenderungan sentimen positif dengan persentase sebesar 77,80%**,** sedangkan sentimen negatif sebesar 22,20%**.** Temuan ini menunjukkan bahwa secara umum pengguna memberikan tanggapan yang positif terhadap produk iPhone yang dipasarkan melalui platform *e-commerce* Shopee.

DAFTAR PUSTAKA

Astuti, K. C., Firmansyah, A., & Riyadi, A. (2024). Implementasi text mining untuk analisis sentimen masyarakat terhadap ulasan aplikasi digital Korlantas Polri pada Google Play Store. *Riset Dan E-Jurnal Manajemen Informatika Komputer*, *8*(1), 384–394.

Badan Pusat Statistik. (2023). *Statistik e-commerce 2023*.

Bourequat, W., & Mourad, H. (2021). Sentiment analysis approach for analyzing iPhone release using support vector machine. *International Journal of Advances in Data and Information Systems*, *2*(1), 36–44.

Clara, S., Laksmi Prianto, D., Al Habsi, R., Friscila Lumbantobing, E., Chamidah, N., Kom, S., Kom, M., Informatika, J., Ilmu Komputer, F., Pembangunan Nasional Veteran Jakarta Jl Fatmawati Raya, U. R., Labu, P., Cilandak, K., & Depok, K. (2021). Implementasi Seleksi Fitur Pada Algoritma Klasifikasi Machine Learning Untuk Prediksi Penghasilan Pada Adult Income Dataset. In *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta-Indonesia*.

Fadli, M., & Saputra, R. A. (2023). *KLASIFIKASI DAN EVALUASI PERFORMA MODEL RANDOM FOREST UNTUK PREDIKSI STROKE Classification And Evaluation Of Performance Models Random Forest For Stroke Prediction*. *12*. http://jurnal.umt.ac.id/index.php/jt/index

Fithri, P., Muluk, A., & Rayhanda, R. H. (2024). Perancangan User Interface (UI) dan User Experience (UX) pada Sistem Informasi PT. XYZ. *Jurnal Nasional Teknologi Dan Sistem Informasi*, *9*(3), 280–289. https://doi.org/10.25077/TEKNOSI.v9i3.2023.280-289

Gultom, Z. A., Nazry, H. W., Syahra, Y., & Zulherry, A. (2025). Analisis pengaruh angka buta aksara di Indonesia. Jurnal Sistem Informasi Triguna Dharma. *Jurnal Sistem Informasi Triguna Dharma*, *4*(1), 82–89.

Hasugian, A. H., Fakhriza, M., & Zukhoiriyah, D. (2023). Analisis Sentimen Pada Review Pengguna *E-commerce* Menggunakan Algoritma Naïve Bayes. *J-SISKO TECH: Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD*, *6*(1), 98–107.

Hermawati, L., Berland, V., Rahmadiah, A., Hutabarat, E., & Dwi Saputra, D. (2023). Komparasi metode text mining terhadap masalah pengklasifikasian narasi informative & non informative pada Twitter @PLN_123. *Jurnal Sistim Informasi Dan Teknologi*, *5*(1), 109–120.

Jabar, A. A., Farta Wijaya, R., & Wahyuni, S. (2025). DASHBOARD VISUALISASI DATA KECERDASAN BISNIS UNTUK MENDUKUNG PENGAMBILAN KEPUTUSAN BISNIS PADA PT BMPT. *Jurnal Teknologi Informasi*, *6*(1). https://doi.org/10.46576/djtechno

Julianto, I. T., & Lindawati, L. (2022). Analisis Sentimen Terhadap Sistem Informasi Akademik Institut Teknologi Garut. *Jurnal Algoritma*, *19*(1), 458–468. https://doi.org/10.33364/algoritma/v.19-1.1112

Limbong, J. J. A., Sembiring, I., & Hartomo, K. D. (2022). Analisis Klasifikasi Sentimen Ulasan pada *E-commerce* Shopee Berbasis Word Cloud dengan Metode Naive Bayes dan K-Nearest Neighbor. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *9*(2), 347–356.

Musfiroh, M., Tholib, A., & Arifin, Z. (2024). Analisis Sentimen Terhadap Ulasan Aplikasi Shopee di Google Play Store Menggunakan Metode TF-IDF dan Long Short-Term Memory. *J. Electr. Eng. Comput*, *6*(2), 371–381.

Nafiisa B. L., Putri Y. N. W., & Ayunin Q. (2022). Dashboard Visualisasi Data UMK Sebagai Alat Pengambilan Keputusan Menggunakan Microsoft Power BI. *Jurnal Akuntansi Dan Manajemen*, *17*(2), 86–105.

Najiyah, I., & Hariyanti, I. (2021). Sentimen analisis covid-19 dengan metode probabilistic neural network dan tf-idf. *Jurnal Responsif: Riset Sains Dan Informatika*, *3*(1), 100–111.

Nugraha, F. A., & A. Susetyo, Y. (2023). ANALISIS PERBANDINGAN PERFORMA DATABASE DUCKDB DAN SQLITE PADA PENGOLAHAN BIG DATA. *JIPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, *8*(3), 1052–1060. https://doi.org/10.29100/jipi.v8i3.4032

Nurhayati, E., & Agussalim, A. (2023). Rancang Bangun Back-end API pada Aplikasi Mobile AyamHub Menggunakan Framework Node JS Express. *Jurnal Sistem Dan Teknologi Informasi (JustIN)*, *11*(3), 524. https://doi.org/10.26418/justin.v11i3.66823

Pradhisa K. C., & Fajriyah R. (2024). Analisis Sentimen Ulasan Pengguna *E-commerce* di Google Play Store Menggunakan Metode IndoBERT. *Building of Informatics, Technology and Science (BITS)*, *6*(1), 92–104. https://doi.org/10.47065/bits.v6i1.5247

Rifaldi, D., & Fadlil, A. (2023). Teknik preprocessing pada text mining menggunakan data tweet "mental health." *Decode: Jurnal Pendidikan Teknologi Informasi*, *3*(2), 161–171.

Rivaldi, R. C., & Wismarini, T. D. (2024). Analisis Sentimen Pada Ulasan Produk Dengan Metode Natural Language Processing (NLP):(Studi Kasus Zalika Store 88 Shopee). *Elkom: Jurnal Elektronika Dan Komputer*, *17*(1), 120–128.

Rosiana, P. S., Voutama, A., & Ridha, A. A. (2023). Perancangan Ui/Ux Sistem Informasi Pembelian Hasil Tani Berbasis Mobile Dengan Metode Design Thinking. *Jurnal Informatika Dan Teknik Elektro Terapan*, *11*(3).

Rosnelly, R., Wahyuni, L., Melvy Anggraini, G., & Lazuli, I. (2023). Implementasi Javascript Dalam Pembuatan Web Sederhana Javascript Implementation in Making a Simple Web. *Community Service Journal) e-ISSN*, *2*(1), 116–123. https://doi.org/10.22303/coral.2.1.2023.116-123

Sapanji, R. A. E. V. T., Hamdani, D., & Harahap, P. (2023). Sentiment Analysis of the Top 5 *E-commerce* Platforms in Indonesia using Text Mining and Natural Language Processing (NLP). *Journal of Applied Informatics and Computing*, *7*(2), 202–211.

Saputra, A. A., Pakpahan, A. G. S., Kurtubi, A., Amiruddin, A., Fridaniarta, B., Wicaksono, E. Y., Saputra, H., Putra, M. Y. A., & Azahra, R. Y. (2023). PELATIHAN DAN PEMBUATAN WEBSITE MENGGUNAKAN HTML DAN CSS. *Beujroh : Jurnal Pemberdayaan Dan Pengabdian Pada Masyarakat*, *1*(1), 119–125. https://doi.org/10.61579/beujroh.v1i1.41

Sari, D. B. I., & Kusuma, K. A. (2023). Perceptions of Shopee and Tokopedia: UMSIDA Management Students' Comparative Analysis: Analisis Komparatif *E-commerce* Shopee dan Tokopedia: Pandangan Mahasiswa Program Studi Manajemen UMSIDA. *Academia Open*, *8*(1), 10–21070.

Septiani, D., & Isabela, I. (2022). Analisis term frequency inverse document frequency (tf-idf) dalam temu kembali informasi pada dokumen teks. *Sistem Dan Teknologi Informasi Indonesia (SINTESIA)*, *1*(2), 81–88.

Simatupang, M. P., & Utomo, D. P. (2019). Analisa Testimonial Dengan Menggunakan Algoritma Text Mining Dan Term Frequency-Inverse Document Frequence (Tf-Idf) Pada Toko Allmeeart. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, *3*(1), 808–814.

Sinaga, A., & Nainggolan, S. P. (2023). Analisis perbandingan akurasi dan waktu proses algoritma stemming Arifin-Setiono dan Nazief-Adriani pada dokumen teks Bahasa Indonesia. *Sebatik*, *27*(1), 63–69.

Ubaidillah F. R, Anjani Arifiyanti, A., Luhur Indayanti Sugata Sistem Informasi, T., Pembangunan Nasional Veteran Jawa Timur JlRaya Rungkut Madya, U., & Anyar, G. (2025). ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI MIDI KRIING MENGGUNAKAN SUPPORT VECTOR MACHINE (SVM). In *Jurnal Mahasiswa Teknik Informatika)* (Vol. 9, Number 3).