

**ANALISIS SENTIMEN ALGORITMA K-NEAREST NEIGHBOR (KNN) PADA  
ULASAN GOOGLE PLAY APLIKASI TOKOPEDIA**

**SKRIPSI SARJANA INFORMATIKA**

Oleh

**DEGA PANDEGANI**

**197064516217**



**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI KOMUNIKASI DAN INFORMATIKA  
UNIVERSITAS NASIONAL**

**2023**

**ANALISIS SENTIMEN ALGORITMA K-NEARST NEIGHBOR (KNN) PADA  
ULASAN GOOGLE PLAY APLIKASI TOKOPEDIA**

**SKRIPSI SARJANAH INFORMATIKA**

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik Informatika dari  
Fakultas Teknologi Komunikasi dan Informatika

Oleh

DEGA PANDEGANI

197064516217



**PROGRAM STUDI INFORMATIKA**

**FAKULTAS TEKNOLOGI KOMUNIKASI DAN INFORMATIKA**

**UNIVERSITAS NASIONAL**

**2023**

HALAMAN PENGESAHAN

TUGAS AKHIR

ANALISIS SENTIMEN ALGORITMA K-NEAREST NEIGHBOR (KNN)  
PADA ULASAN GOOGLE PLAY APLIKASI TOKOPEDIA



Nama : Dega Pandegani  
NPM : 197064516217

Dosen Pembimbing 1

A handwritten signature in black ink, starting with a large 'A' and ending with a horizontal line.

(Andrianingsih, S. Kom., MMSI)

Dosen Pembimbing 2

A handwritten signature in black ink, appearing to be 'Albaar'.

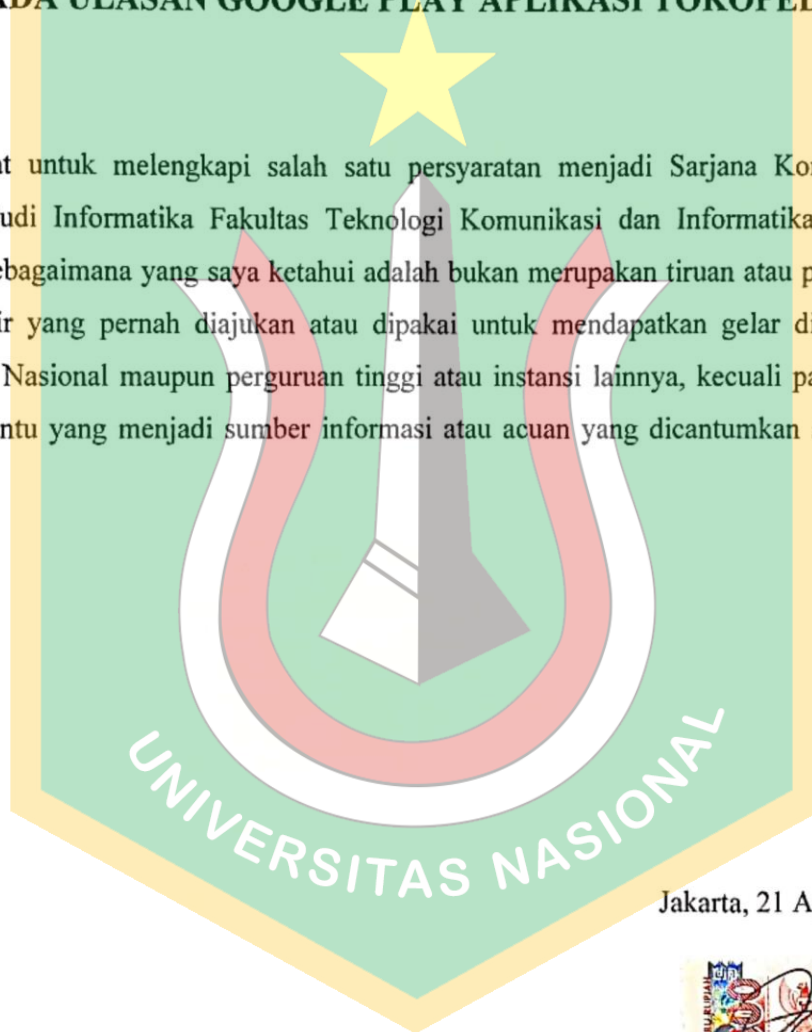
(Albaar Rubhasy, S.Si, MT1)

## PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

### **ANALISIS SENTIMEN ALGORITMA K-NEAREST NEIGHBOR (KNN) PADA ULASAN GOOGLE PLAY APLIKASI TOKOPEDIA**

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Informatika Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.



Jakarta, 21 Agustus 2023



Dega Pandegani

197064516217

## LEMBAR PERSETUJUAN TUGAS AKHIR

Tugas Akhir dengan judul:

# ANALISIS SENTIMEN ALGORITMA K-NEAREST NEIGHBOR (KNN) PADA ULASAN GOOGLE PLAY APLIKASI TOKOPEDIA

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Informatika Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Akhir Semester Genap 2022-2023 pada tanggal 15 Agustus Tahun 2023



**Dosen Pembimbing 1**

Andrianingsih, S. Kom., MMSI

NID 0111130826

**Ketua Program Studi**

Ratih Titi Komala Sari, ST, MM, MMSI

NID 0103150850

**LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI**

Nama : Dega Pandegani  
NPM : 197064516217  
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika  
Program Studi : Informatika  
Tanggal Sidang : 15 Agustus 2023

JUDUL DALAM BAHASA INDONESIA:

ANALISIS SENTIMEN ALGORITMA K-NEAREST NEIGHBOR (KNN) PADA  
ULASAN GOOGLE PLAY APLIKASI TOKOPEDIA

JUDUL DALAM BAHASA INGGRIS:

SENTIMENT ANALYSIS OF K-NEAREST NEIGHBOR (KNN) ALGORITHM ON GOOGLE  
PLAY REVIEWS ON TOKOPEDIA APPLICATION

**TANDA TANGAN DAN TANGGAL**

Pembimbing 1	Ka. Prodi	Mahasiswa
TGL : 23.08.2023	TGL :	TGL : 23 Agustus 2023
		



## LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Dega Pandegani  
NPM : 197064516217  
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika  
Program Studi : Informatika  
Tanggal Sidang : 15 Agustus 2023




JUDUL DALAM BAHASA INDONESIA:

ANALISIS SENTIMEN ALGORITMA K-NEAREST NEIGHBOR (KNN) PADA  
ULASAN GOOGLE PLAY APLIKASI TOKOPEDIA

JUDUL DALAM BAHASA INGGRIS:

SENTIMENT ANALYSIS OF K-NEAREST NEIGHBOR (KNN) ALGORITHM ON GOOGLE  
PLAY REVIEWS ON TOKOPEDIA APPLICATION

### TANDA TANGAN DAN TANGGAL

Pembimbing 2	Ka. Prodi	Mahasiswa
TGL :	TGL :	TGL : 20 Agustus 2023
		

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT yang telah melimpahkan rahmat, taufik serta hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul **“Analisis Sentimen Algoritma K-Nearest Neighbor (KNN) Pada Ulasan Google Play Aplikasi Tokopedia”** sebagai salah satu syarat kelulusan Program Studi Sarjana Informatika Fakultas Teknologi Komunikasi Dan Informatika.

Proses penelitian dan penulisan skripsi ini tidak terlepas dari bantuan berbagai pihak, oleh karena itu saya mengucapkan banyak terimakasih kepada dosen pembimbing Tugas Akhir, Ibu Andrianingsih, S. Kom., MMSI dan Bapak Albaar Rubhasy, S.Si.,M.T.I yang telah membimbing saya dalam penulisan dan penyusunan skripsi sebagai tugas akhir. Saya juga mengucapkan banyak terimakasih kepada:

1. Ibu dan Keluarga saya yang selalu memberikan banyak dukungan dan doa selama saya menyelesaikan penyusunan skripsi.
2. Seluruh Dosen dan Staf Karyawan Fakultas Teknologi Komunikasi dan Informatika yang telah memberikan bantuan dan pembelajaran selama masa perkuliahan.
3. Seluruh teman-teman seangkatan saya yang telah membantu dan memberikan dukungan.
4. Seluruh pihak yang terlibat langsung maupun tidak langsung menyelesaikan skripsi ini.

Akhir kata, semoga Tuhan Yang Maha Esa membalas semua kebaikan dan bantuan yang telah di berikan dengan hal yang baik. Saya mengharapkan kritikan dan saran yang bersifat membangun dan semoga skripsi ini dapat memberikan manfaat di bidang Teknologi Informatika

Jakarta, Maret 2023



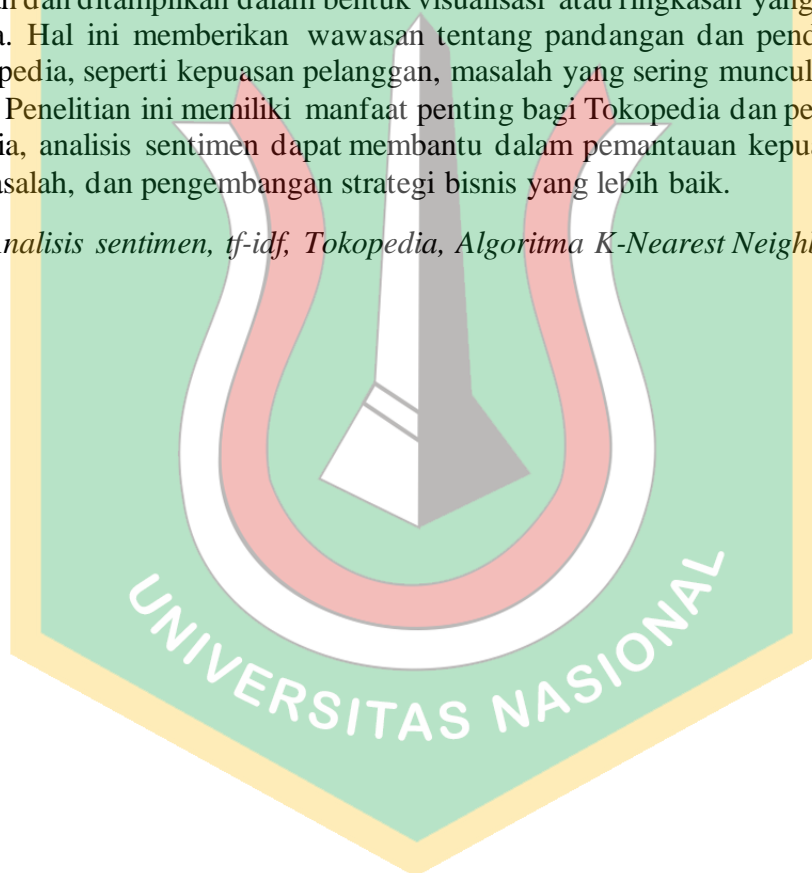
Dega Pandegani



## ABSTRAK

Sentimen pelanggan menjadi faktor penting dalam menilai kualitas dan kepuasan pelanggan kepada suatu layanan e-commerce. Dalam hal ini, Tokopedia ialah salah satu panggung bisnis online terbanyak di Indonesia. Buat mengenali rate yang didapat dalam mengelompokkan data survei negatif serta audit positif memakai perhitungan K- Nearest Neighbor ( KNN). Pertama, data ulasan berupa sentimen pelanggan terkait dengan Tokopedia dikumpulkan dari google paly menggunakan google scraper, Tahap pengolahan data meliputi pembersihan teks, pemisahan kata-kata, penghapusan stopwords, dan ekstraksi fitur menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency). Hal ini bertujuan untuk mengubah teks menjadi representasi numerik yang dapat digunakan oleh algoritma KNN. Selanjutnya, data yang telah diproses dijadikan sebagai input untuk algoritma KNN. Algoritma KNN digunakan untuk mengklasifikasikan sentimen pelanggan terhadap Tokopedia menjadi kategori: positif, dan negatif. Dalam algoritma KNN, sentimen pelanggan yang baru, diklasifikasikan berdasarkan keterdekatannya dengan tetangga terdekat Setelah proses klasifikasi, hasil analisis sentimen diolah dan ditampilkan dalam bentuk visualisasi atau ringkasan yang dapat dipahami oleh pengguna. Hal ini memberikan wawasan tentang pandangan dan pendapat pelanggan terhadap Tokopedia, seperti kepuasan pelanggan, masalah yang sering muncul, atau kelebihan yang dihargai. Penelitian ini memiliki manfaat penting bagi Tokopedia dan pengguna lainnya. Bagi Tokopedia, analisis sentimen dapat membantu dalam pemantauan kepuasan pelanggan, identifikasi masalah, dan pengembangan strategi bisnis yang lebih baik.

**Kata kunci:** Analisis sentimen, tf-idf, Tokopedia, Algoritma K-Nearest Neighbor (KNN).



## ABSTRACT

*Customer sentiment is an important factor in assessing the quality and customer satisfaction in an e-commerce service. In this case, Tokopedia is one of the largest e-commerce platforms in Indonesia. To find out the percentage obtained in classifying negative review data and positive reviews using the K-Nearest Neighbor (KNN) algorithm. First, review data in the form of customer sentiment regarding Tokopedia collected from Google Play using Google Scraper. The data processing stage includes text cleaning, word separation, stopword removal, and feature extraction using the TF-IDF (Term Frequency-Inverse Document Frequency) method. . It aims to convert text into a numeric representation usable by the KNN algorithm. Furthermore, the processed data is used as input for the KNN algorithm. The KNN algorithm is used to classify customer sentiment towards Tokopedia into categories: positive and negative. In the KNN algorithm, new customer sentiment will be classified based on its proximity to the nearest neighbor. After the classification process, the sentiment analysis results will be processed and displayed in the form of a visualization or summary that can be understood by the user. It provides insight into the views and opinions of Tokopedia customers, such as customer satisfaction, frequently occurring issues, or profit value. This research has important benefits for Tokopedia and other users. For Tokopedia, sentiment analysis can help monitor customer satisfaction, solve problems, and develop better business strategies.*

**Keywords:** *Sentimen analysis, tf-idf, tokopedia, K-Nearest Neighbor (KNN) algorithm.*



## DAFTAR ISI

<b>KATA PENGANTAR.....</b>	<b>Error! Bookmark not defined.</b>
<b>ABSTRAK.....</b>	<b>Error! Bookmark not defined.</b>
<b>ABSTRACK.....</b>	<b>Error! Bookmark not defined.</b>
<b>PENDAHULUAN.....</b>	<b>Error! Bookmark not defined.</b>
1.1 Latar Belakang.....	<b>Error! Bookmark not defined.</b>
1.2 Identifikasi Masalah.....	<b>Error! Bookmark not defined.</b>
1.3 Tujuan.....	<b>Error! Bookmark not defined.</b>
1.4 Batasan Masalah.....	<b>Error! Bookmark not defined.</b>
1.5 Kontribusi.....	<b>Error! Bookmark not defined.</b>
<b>TINJAUAN PUSATAKA.....</b>	<b>Error! Bookmark not defined.</b>
2.1 Tokopedia.....	<b>Error! Bookmark not defined.</b>
2.2 Google Play.....	<b>Error! Bookmark not defined.</b>
2.3 Analisis Sentimen.....	<b>Error! Bookmark not defined.</b>
2.4 Text mining.....	<b>Error! Bookmark not defined.</b>
2.5 Algoritma K-Nearest Neighbor (K-NN).....	<b>Error! Bookmark not defined.</b>
2.6 Studi Literatur.....	<b>Error! Bookmark not defined.</b>
<b>METODE</b>	
<b>PENELITIAN.....</b>	<b>Error! Bookmark not defined.</b>
3.1 Lokasi Penelitian.....	<b>Error! Bookmark not defined.</b>
3.2 Waktu Penelitian.....	<b>Error! Bookmark not defined.</b>
3.3 Fokus Penelitian.....	<b>Error! Bookmark not defined.</b>
3.4 Sumber Data.....	<b>Error! Bookmark not defined.</b>

3.5	Desain	
	Penelitian.....	<b>Error! Bookmark not defined.</b>
3.4.1	Tahapan penelitian.....	<b>Error! Bookmark not defined.</b>
3.4.2	Flowchart Algoritma K-Nearest Neighbor (KNN).....	<b>Error! Bookmark not defined.</b>
<b>HASIL DAN DISKUSI.....Error!</b>		
<b>Bookmark not defined.</b>		
4.1	Hasil Penelitian.....	<b>Error!</b>
<b>Bookmark not defined.</b>		
4.1.1	Implementasi Sistem.....	<b>Error! Bookmark not defined.</b>
4.1.2	Problem Formulation.....	<b>Error! Bookmark not defined.</b>
4.1.3	Conceptual Model Text Mining.....	<b>Error! Bookmark not defined.</b>
4.1.4	Conceptual Model Algoritma K-NN.....	<b>Error! Bookmark not defined.</b>
4.2	Pembahasan dan Pengujian.....	<b>Error! Bookmark not defined.</b>
<b>Error! Bookmark not defined.</b>		
4.2.1	Verification, validasi and experimentation.....	<b>Error! Bookmark not defined.</b>
4.2.2	Analisis implementasi data kinerja dari metode algoritma k-nn.....	<b>Error! Bookmark not defined.</b>
<b>Error! Bookmark not defined.</b>		
<b>KESIMPULAN DAN SARAN.....Error! Bookmark not defined.</b>		
<b>Bookmark not defined.</b>		
5.1	Kesimpulan.....	<b>Error! Bookmark not defined.</b>
5.2	Saran.....	<b>Error! Bookmark not defined.</b>
<b>DAFTAR PUSTAKA.....</b>		
<b>Error! Bookmark not defined.</b>		

