

**ANALISIS SENTIMEN MASYARAKAT TERHADAP  
PENGUNAAN E-COMMERCE MENGGUNAKAN  
ALGORITMA K-NEAREST NEIGHBOR**

**JALUR SCIENTIST**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Informatika



disusun oleh

**IKHSAN HABIB KUSUMA**

**20.11.3721**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2024**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP  
PENGUNAAN E-COMMERCE MENGGUNAKAN  
ALGORITMA K-NEAREST NEIGHBOR**

**JALUR SCIENTIST**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Informatika



disusun oleh

**IKHSAN HABIB KUSUMA**

**20.11.3721**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2024**

**HALAMAN PERSETUJUAN**

**JALUR SCIENTIST**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP PENGGUNAAN  
E-COMMERCE MENGGUNAKAN ALGORITMA K-NEAREST  
NEIGHBOR**

yang disusun dan diajukan oleh

**Ikhsan Habib Kusuma**

**20.11.3721**

telah disetujui oleh Dosen Pembimbing  
pada tanggal 21 Februari 2024

**Dosen Pembimbing,**

  
**Nuri Cahyono, M.Kom**

**NIK. 190402278**

**HALAMAN PENGESAHAN**  
**JALUR SCIENTIST**  
**ANALISIS SENTIMEN MASYARAKAT TERHADAP PENGGUNAAN**  
**E-COMMERCE MENGGUNAKAN ALGORITMA K-NEAREST**  
**NEIGHBOR**

yang disusun dan diajukan oleh

**Ikhsan Habib Kusuma**

20.11.3721

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 21 Februari 2024

**Susunan Dewan Penguji**

**Nama Penguji**

Ainul Yaqin, M.Kom  
NIK. 190302255

Yuli Astuti, M.Kom  
NIK. 190302146

Nuri Cahyono, M.Kom  
NIK. 190302278

**Tanda Tangan**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 21 Februari 2024

**DEKAN FAKULTAS ILMU KOMPUTER**



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.  
NIK. 190302096

## HALAMAN PERNYATAAN KONTRIBUSI KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Ikhsan Habib Kusuma

NIM : 20.11.3721

Menyatakan bahwa karya dengan judul berikut:

### **ANALISIS SENTIMEN MASYARAKAT TERHADAP PENGGUNAAN E-COMMERCE MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR**

Dosen Pembimbing: Nuri Cahyono, M.Kom

1. Karya adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya merupakan gagasan, rumusan maupun penelitian yang orisinal dan SAYA memiliki KONTRIBUSI terhadap karya tersebut.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka atau Referensi pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 21 Februari 2024

Yang Menyatakan,



Ikhsan Habib Kusuma

## HALAMAN PERSEMBAHAN

Dengan Puji syukur atas rahmat dan hidayah-Nya yang diberikan Allah SWT, penulis dapat menyajikan halaman persembahan ini sebagai ungkapan terima kasih kepada semua pihak yang telah memberikan dorongan, dukungan dan bimbingan selama proses penulisan skripsi ini untuk menyelesaikan masa studi di Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta. Ucapan terima kasih ini secara khusus dipersembahkan kepada:

1. Allah SWT Tuhan Yang Maha Esa karena atas karunia-Nya skripsi ini dapat selesai pada waktunya.
2. Kedua orang tua atas doa dan dukungannya sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
3. Keluarga yang memberikan semangat dan dorongan agar penulis dapat menyelesaikan skripsi dengan tepat pada waktunya.
4. Seluruh teman-teman Universitas Amikom Yogyakarta yang telah mendukung dan menemani selama masa perkuliahan.
5. Dan semua pihak yang tidak dapat penulis sebutkan satu per satu yang telah membantu.

## KATA PENGANTAR

Segala Puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul "Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor" dengan baik.

Penyusunan skripsi ini diajukan untuk memenuhi syarat kelulusan mata kuliah skripsi agar memperoleh gelar Sarjana Komputer pada Program Studi Informatika Fakultas Ilmu Komputer di Universitas Amikom Yogyakarta. Penulis juga menyadari bahwa skripsi ini dapat tersusun berkat semua pihak yang telah memberikan dukungan, bimbingan, dan arahan selama penulisan. Pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, MM, selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Windha Mega Pradnya Dhuhita, M.Kom, selaku Kepala Program Studi Informatika Universitas Amikom Yogyakarta.
4. Bapak Nuri Cahyono, M.Kom, selaku Dosen Wali sekaligus Dosen Pembimbing.
5. Segenap Dosen Program Studi Informatika yang telah mendidik serta memberikan ilmu selama kuliah dan seluruh staff yang bersedia melayani segala administrasi.

Yogyakarta, 21 Februari 2024

Penulis

## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PERSETUJUAN .....	ii
HALAMAN PENGESAHAN .....	iii
HALAMAN PERNYATAAN KONTRIBUSI KARYA .....	iv
HALAMAN PERSEMBAHAN .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI .....	vii
DAFTAR TABEL .....	x
DAFTAR GAMBAR .....	xi
ABSTRAK .....	xii
ABSTRACTION .....	xiii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
BAB II LANDASAN TEORI .....	3
2.1 Penelitian Terkait .....	3
2.2 Dasar Teori .....	5
2.2.1. E-Commerce .....	5
2.2.2. Shopee .....	5
2.2.3. Komentar .....	5
2.2.4. Google Play Store .....	5
2.2.5. Analisis Sentimen .....	5
2.2.6. Excel .....	6
2.2.7. Algoritma NLP .....	6

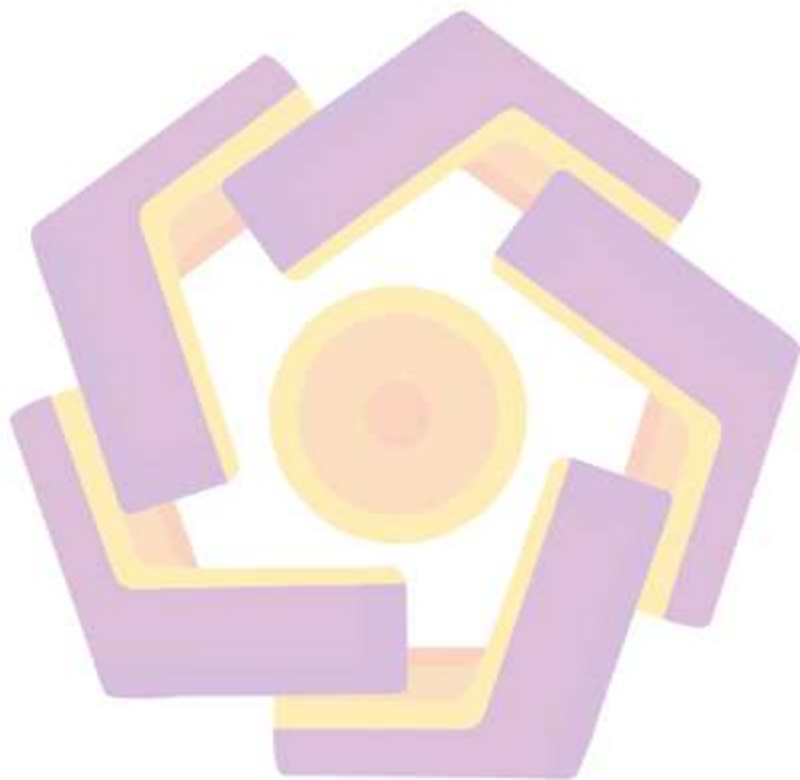


2.2.8.	Python .....	6
2.2.9.	Scraping Data .....	6
2.2.10.	WebHarvy .....	6
2.2.11.	Undersampling Dataset .....	7
2.2.12.	Preprocessing Data .....	7
2.2.13.	Algoritma K-Nearest Neighbor .....	7
2.2.14.	Confusion Matrix .....	8
2.2.15.	Word Cloud .....	8
<b>BAB III METODE PENELITIAN .....</b>		<b>9</b>
3.1	Web Scraping .....	9
3.2	Dataset .....	11
3.3	Preprocessing Dataset .....	11
3.3.1.	Lowercase Folding .....	11
3.3.2.	Menghilangkan Tanda Baca dan Angka .....	12
3.3.3.	Penanganan Tag .....	12
3.3.4.	Word Tokenization .....	12
3.3.5.	Word Normalization .....	12
3.3.6.	Stopwords Removal .....	13
3.3.7.	Stemming .....	13
3.4	Algoritma .....	14
3.4.1.	Natural Language Processing .....	14
3.4.2.	Preprocessing .....	14
3.4.3.	TF-IDF .....	14
3.4.4.	Cosine Similarity .....	15
3.4.5.	K-Nearest Neighbor .....	15

3.5	Evaluasi.....	16
3.5.1.	Confusion Matrix.....	16
3.5.2.	Metrik Evaluasi.....	17
3.5.3.	Cross Validation Score.....	18
BAB IV HASIL DAN PEMBAHASAN.....		19
4.1	Deskripsi Data.....	19
4.2	Preprocessing Data.....	19
4.3	Model Sentiment Analysis.....	19
4.4	Evaluasi Model.....	19
4.5	Visualisasi.....	20
BAB V KESIMPULAN.....		21
REFERENSI.....		22
LAMPIRAN BUKTI PENDUKUNG.....		25

## DAFTAR TABEL

Tabel 1 Confusion Matrix.....	17
Tabel 2 Metrik Evaluasi.....	17
Tabel 3 Identitas Publikasi.....	25



## DAFTAR GAMBAR

Gambar 1 Proses memulai web scraping dengan menyalin link ke WebHarvy.....	9
Gambar 2 Proses pengumpulan dataset dengan Teknik web scraping.....	9
Gambar 3 Proses menyimpan dan preprocessing dataset hasil scraping.....	10
Gambar 4 Tahap pra-pemrosesan NLP dengan cleaning simbol dataset.....	12
Gambar 5 Tahap pra-pemrosesan NLP dengan cleaning word dataset.....	13
Gambar 6 Wordcloud frekuensi kata yang sering muncul.....	20
Gambar 7 Paper Jurnal Halaman 1.....	26
Gambar 8 Paper Jurnal Halaman 2.....	27
Gambar 9 Paper Jurnal Halaman 3.....	28
Gambar 10 Paper Jurnal Halaman 4.....	29
Gambar 11 Paper Jurnal Halaman 5.....	30
Gambar 12 Paper Jurnal Halaman 6.....	31
Gambar 13 <i>Letter of Acceptance</i> .....	32
Gambar 14 Bukti Review.....	33
Gambar 15 Catatan Revisi 1.....	33
Gambar 16 Catatan Revisi 2.....	34
Gambar 17 Catatan Revisi 3.....	34
Gambar 18 Catatan Revisi 4.....	35
Gambar 19 Catatan Revisi 5.....	35
Gambar 20 Catatan Revisi 6.....	36
Gambar 21 Catatan Revisi 7.....	36
Gambar 22 Catatan Revisi 8.....	37
Gambar 23 Bukti Terbit Jurnal.....	37
Gambar 24 Bukti Daftar Terbit.....	38
Gambar 25 Bukti SINTA.....	38
Gambar 26 Bukti Index ROAD ISSN.....	39
Gambar 27 Bukti jurnal terbit Google Scholar.....	39

## ABSTRAK

Perkembangan e-commerce yang pesat telah menyebabkan peningkatan transaksi online dan perubahan perilaku konsumen. Di Indonesia, penggunaan e-commerce tumbuh pesat dengan banyak platform online bermunculan. Muncul masalah yakni bagaimana memahami sentimen Masyarakat secara aktual terhadap kualitas e-commerce di Indonesia sangat penting bagi bisnis untuk meningkatkan layanan dan menjaga kepuasan pelanggan. Oleh karena itu, dalam penelitian ini peneliti mengusulkan metodologi dengan menggunakan algoritma *K-Nearest Neighbor*. Pengumpulan data dilakukan dengan *Web Scraping* dari website Google Play Store platform Shopee. Setelah data terkumpul, dilakukan proses preprocessing untuk *Data Cleaning*, termasuk *Case Folding*, menghilangkan tanda baca dan angka, penanganan tag, *Word Tokenization*, *Word Normalization*, *Stopwords Removal*, dan *Stemming*. Setelah itu, algoritma *K-Nearest Neighbor* (K-NN) digunakan untuk mengklasifikasikan sentimen menjadi positif atau negatif. Evaluasi dilakukan dengan menggunakan *Confusion Matrix* dan *Classification Report* untuk menilai keakuratan algoritma. Hasil penelitian dari 895 baris data menunjukkan bahwa pendekatan yang diusulkan efektif dalam menganalisis sentimen masyarakat terhadap e-commerce di Indonesia, dengan nilai akurasi model 82%, nilai presisi model 86%, nilai recall model 79%, nilai F1 Score model 82% dan nilai dari *cross-validation score* sebesar 80%.

**Kata kunci:** Sentimen Analisis, E-Commerce, Supervised Learning, Machine Learning, NLP, KNN

## ABSTRACTION

*The rapid development of e-commerce has led to an increase in online transactions and changes in consumer behavior. In Indonesia, the use of e-commerce has grown rapidly with many online platforms emerging. A problem arises in understanding the actual sentiment of the community towards the quality of e-commerce in Indonesia, which is crucial for businesses to improve services and maintain customer satisfaction. Therefore, this study proposes a methodology using the K-Nearest Neighbor algorithm. Data collection was conducted through web scraping from the Google Play Store platform Shopee. After data collection, a preprocessing process for data cleaning was performed, including case folding, removing punctuation and numbers, handling tags, word tokenization, word normalization, stopwords removal, and stemming. Thereafter, the K-Nearest Neighbor (K-NN) algorithm was used to classify sentiments into positive or negative. Evaluation was conducted using the Confusion Matrix and Classification Report to assess the accuracy of the algorithm. The results from 895 rows of data show that the proposed approach is effective in analyzing the sentiment of the community towards e-commerce in Indonesia, with an accuracy rate of 82%, a precision rate of 86%, a recall rate of 79%, an F1 Score of 82%, and a cross-validation score of 80%.*

**Keyword:** *Sentiment Analysis, E-Commerce, Supervised Learning, Machine Learning, NLP, KNN*