

**ANALISIS KLASIFIKASI SENTIMEN REVIEW PADA APLIKASI
PEDULI LINDUNGI MENGGUNAKAN METODE
K-NEAREST NEIGHBOR**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



diajukan oleh

CHAIRUNISA NUR AFIFAH

18.11.2448

Kepada

PROGRAM SARJANA

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2022

**ANALISIS KLASIFIKASI SENTIMEN REVIEW PADA APLIKASI PEDULI
LINDUNGI MENGGUNAKAN METODE K-NEAREST NEIGHBOR**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



diajukan oleh
CHAIRUNISA NUR AFIFAH
18.11.2448

Kepada
PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2022

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS KLASIFIKASI SENTIMEN REVIEW PADA APLIKASI PEDULI
LINDUNGI MENGGUNAKAN METODE K-NEAREST NEIGHBOR**

yang disusun dan diajukan oleh

Chalrunisa Nur Alfah

18.11.2448

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 9 September 2022

Dosen Pembimbing,

Kusnawi, S.Kom, M.Eng

NIK. 190302112

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS KLASIFIKASI SENTIMEN REVIEW PADA APLIKASI PEDULI
LINDUNGI MENGGUNAKAN METODE K-NEAREST NEIGHBOR**

yang disusun dan diajukan oleh

Chalrunisa Nur Aifah

18.11.2448

Telah dipertahankan di depan Dewan Penguji
pada tanggal 17 September 2022

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Windha Mega Pradnya D, M.Kom
NIK. 190302185

Subektiningsih
NIK. 190302413

Kusnawi, S.Kom, M.Eng
NIK. 190302112

Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 17 September 2022

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, S.Kom., M.Kom.
NIK. 19030209

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Chairunisa Nur Affah
NIM : 18.11.2448

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Klasifikasi Sentimen Review Pada Aplikasi Peduli Lindungi-
Meninggunkan Metode K-Nearest Neighbor**

Dosen Pembimbing : Kinawati S.Kon, M.Ling

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diberikan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam masalah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka 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, 17 September 2022

Yang Menyatakan,



Chairunisa Nur Affah

HALAMAN PERSEMBAHAN

Puji Syukur kepada Tuhan Yang Maha Esa Allah SWT atas berkat, rahmat dan hidayah-Nya sehingga penulis mampu menyelesaikan skripsi dengan baik. Penulis juga mengucapkan terima kasih kepada pihak yang membantu baik secara langsung maupun tidak. Skripsi ini saya persembahkan kepada :

1. Allah SWT untuk segala rahmat dan nikmat yang telah diberikan.
2. Kepada kedua orang tua yang sangat berarti bagi hidup saya. Terima kasih atas segala pengorbanan dan doa baik yang tidak pernah putus diberikan kepada saya, dan juga terima kasih telah menghadirkan saya ke dunia.
3. Kedua kakak saya, Mba Eki dan Mas Rizal yang selalu mendukung saya, selalu menyayangi dan mengerti saya.
4. Kepada diri saya sendiri, terima kasih telah berjuang, terima kasih telah bekerja keras, terima kasih telah bertahan hingga saat ini, terima kasih sekali lagi untuk diri saya sendiri.
5. Teman-teman kuncup yang telah bersama, berbagi canda tawa hingga kesedihan bersama, sehingga dapat tumbuh bersama.
6. Kepada sahabat-sahabat saya, Krisda, Sintiya, Harti yang telah banyak membantu, dan saling memberikan motivasi hidup.
7. Semua pihak yang membantu baik secara langsung maupun secara tidak langsung.

KATA PENGANTAR

Puji Syukur kepada Tuhan Yang Maha Esa atas berkat dan karunia-Nya, penulis dapat menyelesaikan skripsi yang berjudul “Analisis Klasifikasi Sentimen Review pada Aplikasi Peduli Lindungi Menggunakan Metode K-Nearest Neighbor” sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi Informatika.

Penulis menyadari bahwa tidaklah mudah dalam menyelesaikan skripsi ini tanpa bantuan dan bimbingan dari berbagai pihak yang senantiasa memberikan ilmu, motivasi, bantuan, bimbingan dan pengarahan dari berbagai pihak, untuk itu dengan segala hormat mengucapkan terima kasih khususnya kepada :

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Kusnawi, S. Kom, M. Eng. selaku dosen pembimbing yang dengan sabar dan tulus membimbing dan memberikan ilmu dan motivasi kepada penulis.
3. Bapak Ibu Dosen Universitas Amikom Yogyakarta yang telah memberikan ilmunya selama kuliah.
4. Teman-teman yang sudah berjuang bersama, berbagi ilmu dan juga kebersamaannya selama masa perkuliahan.
5. Kepada semua pihak yang memberikan bantuan baik moril maupun materil sehingga laporan skripsi ini dapat diselesaikan.

Penulis menyadari bahwa laporan skripsi ini masih jauh dari kata sempurna dan banyak kekurangan, sehingga kritik dan saran yang membangun sangat diharapkan. Akhir kata semoga tulisan ini bermanfaat bagi pihak yang membutuhkan dan khususnya bagi penulis sendiri.

Yogyakarta, 17 September 2022

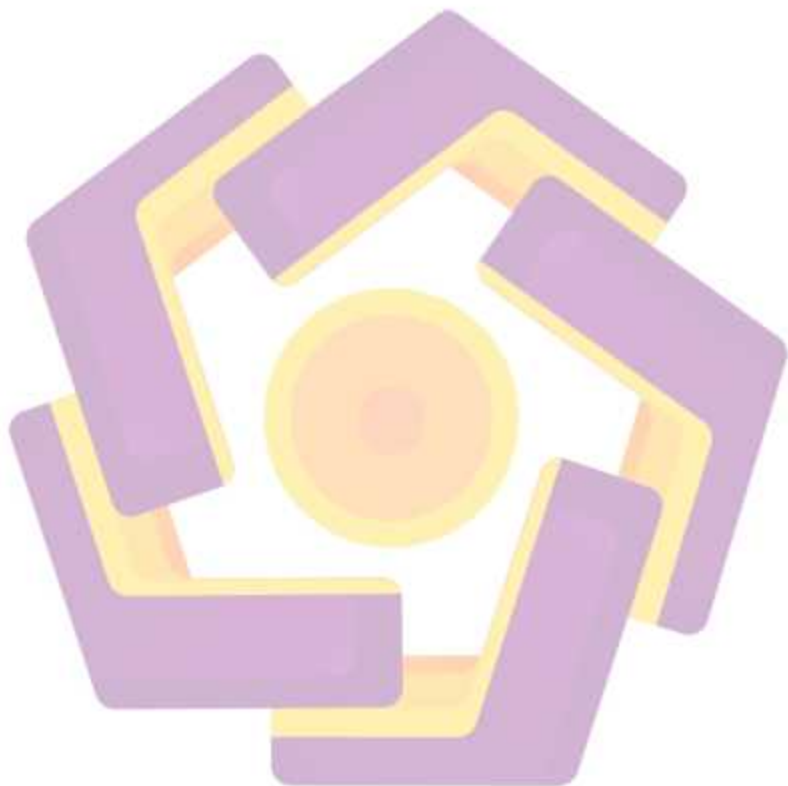
Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	v
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
INTISARI	xiii
Abstract	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan masalah	2
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Kajian Pustaka	5
2.2 Dasar Teori	9
2.2.1 Analisis sentimen	9
2.2.2 Klasifikasi	9
2.2.3 Peduli Lindungi	9
2.2.4 Text mining	10
2.2.5 Text Preprocessing	10
2.2.5.1 Case folding	11
2.2.5.2 Tokenizing	11
2.2.5.3 Cleansing	11
2.2.5.4 Stopword removal	11
2.2.5.5 Stemming	11

2.2.6	TF-IDF	11
2.2.7	K-nearest neighbor	12
2.2.8	Confusion matrix	13
2.2.9	K-fold cross validation	15
BAB III METODOLOGI PENELITIAN		16
3.1	Alat dan Bahan	16
3.1.1	Hardware (perangkat keras)	16
3.1.2	Software (perangkat lunak)	16
3.2	Alur Penelitian	17
3.2.1	Data Penelitian	18
3.2.2	Text Preprocessing	18
3.2.3	Translate Data	21
3.2.4	Pelabelan Data	21
3.2.5	Pembagian Data	21
3.2.6	Pembobotan Kata	22
3.2.7	K-Nearest Neighbors	23
3.2.8	Evaluasi dan Validasi	28
3.2.8.1	Confusion Matrix	28
3.2.8.2	K-Fold Cross Validation	29
BAB IV HASIL DAN PEMBAHASAN		31
4.1	Hasil Pengumpulan Data	31
4.2	Hasil Preprocessing Data	32
4.3	Hasil Translate Data	41
4.4	Hasil Pelabelan Data	42
4.5	Hasil Pembagian Data	44
4.6	Hasil Pembobotan Kata	45
4.7	Hasil K-Nearest Neighbors	46
4.8	Hasil Evaluasi	48
4.9	Hasil Validasi	49
4.10	Visualisasi	51
4.10.1	Review Positif	53
4.10.2	Review Negatif	54
BAB V KESIMPULAN DAN SARAN		56

5.1	Kesimpulan	56
5.2	Saran	57
DAFTAR PUSTAKA		58



DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	7
Tabel 3. 1 Hardware.....	16
Tabel 3. 2 Software.....	16
Tabel 3. 3 Data Sample	18
Tabel 3. 4 Perhitungan TF-IDF.....	22
Tabel 3. 5 Hasil Perhitungan TF-IDF	23
Tabel 3. 6 Contoh Kasus.....	24
Tabel 3. 7 Preprocessing Contoh Kasus	25
Tabel 3. 8 Perhitungan TF-IDF pada Contoh Kasus.....	26
Tabel 3. 9 Perhitungan Jarak.....	27
Tabel 3. 10 Hasil Mengurutkan Jarak	27
Tabel 3. 11 Hasil Tetangga Terdekat.....	28
Tabel 3. 12 Confusion Matrix	29
Tabel 3. 13 K-fold cross validation.....	29
Tabel 4. 1 Script Proses Pengumpulan Data.....	31
Tabel 4. 2 Scripts Proses Case Folding.....	33
Tabel 4. 3 Hasil Preprocessing Case Folding	33
Tabel 4. 4 Scripts Preprocessing Cleansing	34
Tabel 4. 5 Hasil Proses Cleansing.....	34
Tabel 4. 6 Scripts Proses Tokenizing.....	35
Tabel 4. 7 Hasil Proses Tokenizing	35
Tabel 4. 8 Scripts Preprocessing Normalisasi	36
Tabel 4. 9 Hasil Proses Normalisasi	37
Tabel 4. 10 Scripts Preprocessing Stopwords.....	39
Tabel 4. 11 Hasil Proses Stopwords	39
Tabel 4. 12 Script Proses Stemming	40
Tabel 4. 13 Hasil Stemming.....	41
Tabel 4. 14 Scripts Translate Data.....	41
Tabel 4. 15 Scripts Pelabelan Data	43
Tabel 4. 16 Scripts pembagian Data	44
Tabel 4. 17 Script Pembobotan Kata	45
Tabel 4. 18 Scripts Klasifikasi KNN	47
Tabel 4. 19 Hasil Prediksi Menggunakan KNN	47
Tabel 4. 20 Scripts Confusion Matrix.....	48
Tabel 4. 21 Hasil Akurasi, Presisi, Recall, F1-Score.....	49
Tabel 4. 22 Scripts K-Fold Cross Validation.....	49
Tabel 4. 23 Hasil K-Fold Cross Validation	50
Tabel 4. 24 Script Visualisasi	51
Tabel 4. 25 Script WordCloud.....	52

DAFTAR GAMBAR

Gambar 2. 1 Konsep Kerja K-Fold Cross Validation (Mustika, et al. 2021)	15
Gambar 3. 1 Diagram Alir Penelitian.....	17
Gambar 3. 2 Alur Preprocessing.....	19
Gambar 3. 3 Proses Case Folding.....	19
Gambar 3. 4 Proses Cleansing.....	19
Gambar 3. 5 Proses Tokenizing.....	20
Gambar 3. 6 Proses Normalisasi.....	20
Gambar 3. 7 Proses Stopwords.....	20
Gambar 3. 8 Proses Stemming.....	21
Gambar 3. 9 Proses Translate.....	21
Gambar 3. 10 Flowchart KNN.....	24
Gambar 4. 1 Hasil Scraping.....	32
Gambar 4. 2 Hasil Pengumpulan Dataset.....	32
Gambar 4. 3 Download Stopwords.....	38
Gambar 4. 4 List Stopwords Indonesia.....	38
Gambar 4. 5 Hasil Translate Dataset.....	42
Gambar 4. 6 Hasil Pelabelan Menggunakan TextBlob.....	43
Gambar 4. 7 Jumlah Data Label Menggunakan TextBlob.....	44
Gambar 4. 8 Hasil Pembagian Data.....	45
Gambar 4. 9 Hasil TF-IDF.....	46
Gambar 4. 10 Grafik akurasi.....	46
Gambar 4. 11 Confusion Matrix.....	49
Gambar 4. 12 Frekuensi Kata pada Data Review.....	51
Gambar 4. 13 WordCloud pada Data Review.....	52
Gambar 4. 14 Frekuensi Kata Positif.....	53
Gambar 4. 15 WordCloud Positif.....	54
Gambar 4. 16 Frekuensi Kata Negatif.....	54
Gambar 4. 17 WordCloud Negatif.....	55

INTISARI

Pada saat ini dunia sedang menghadapi wabah virus covid-19. Agar tetap bisa memantau data covid-19 dengan mudah diperlukan aplikasi yang berisi tentang informasi-informasi covid-19. Peduli Lindungi merupakan aplikasi yang dapat digunakan untuk memantau data covid-19, melakukan proses check in/check out di lokasi yang ramai seperti mall dan untuk mendapatkan sertifikat vaksin covid-19 bagi masyarakat yang telah melakukan vaksinasi. Berdasarkan informasi yang tercantum dalam halaman deskripsi aplikasi pada Google Play Store, terdapat lebih dari 90 juta unduhan aplikasi Peduli Lindungi sampai dengan April 2022. Data Penelitian ini diambil dari website google play store dan data yang diambil yaitu teks review dengan jumlah sekitar 5000 ulasan. Ulasan yang diberikan oleh pengguna terhadap aplikasi Peduli Lindungi sangat berguna sebagai masukan untuk perusahaan dalam meningkatkan kualitas aplikasi. Untuk itu, penelitian ini dibuat untuk mengklasifikasikan ulasan yang diberikan termasuk positif atau negatif terhadap aplikasi Peduli Lindungi. Metode yang dipakai dalam penelitian ini adalah metode eksperimen dengan menggunakan algoritma K-Nearest Neighbor. Hasil dari penelitian ini adalah menghasilkan tingkat akurasi dari metode yang digunakan dengan pelabelan TextBlob dan juga hasil analisis sentimen ulasan aplikasi Peduli Lindungi.

Kata kunci: K-Nearest Neighbor, Covid-19, Ulasan, Sentimen, Peduli Lindungi

Abstract

At this time the world is facing an outbreak of the Covid-19 virus. In order to be able to easily monitor Covid-19 data, an application that contains information about COVID-19 is needed. Peduli Lindungi is an application that can be used to monitor COVID-19 data, carry out the check-in/check-out process at busy locations such as malls and to obtain a certificate of Covid-19 vaccine for people who have been vaccinated. Based on the information contained in the application description page on the Google Play Store, there are more than 90 million downloads of the Peduli Lindungi application until April 2022. This research data was taken from the google play store website and the data taken was a review text with a total of around 5000 reviews. Reviews given by users on the Peduli Lindungi application are very useful as input for companies in improving the quality of the application. For this reason, this study was made to classify the reviews given as positive or negative and labelling by Textblob towards the Peduli Lindungi application. The method used in this study is an experimental method using the K-Nearest Neighbor algorithm. The results of these research is produce the level of accuracy of the method used with TextBlob labelling and result of sentiment analysis Peduli Lindungi application's review.

Keyword: *K-Nearest Neighbor, Covid-19, Review, Sentiment, Peduli Lindungi*