

**ANALISIS SENTIMENT ULASAN PENGGUNA
LINKEDIN PADA GOOGLE PLAYSTORE DENGAN
METODE NAÏVE BAYES DAN KNN**

SKRIPSI

**Oleh :
Aditya Putra Prastyo
202010225259**



**FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2024**

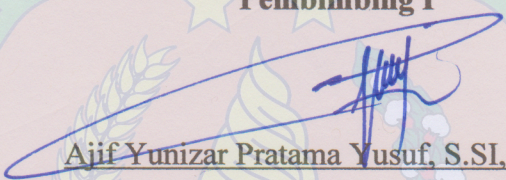
LEMBAR PERSETUJUAN PEMBIMBING

Judul Proposal Tugas : Analisis Sentiment Ulasan Pengguna LinkedIn Pada
Google Playstore Dengan Metode Naïve Bayes Dan
KNN
Nama Mahasiswa : Aditya Putra Prastyo
Nomor : 202010225259
Pokok
Mahasiswa
Program Studi/Fakultas : Informatika / Ilmu Komputer

Jakarta, 23/01/2024

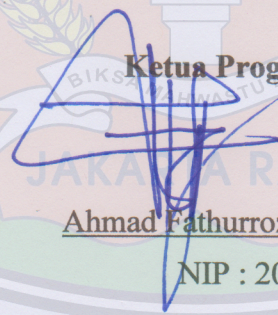
MENYETUJUI,

Pembimbing I


Ajif Yunizar Pratama Yusuf, S.SI, M. Eng

NIDN : 0328068603

Ketua Program Studi


Ahmad Fathurrozi, S.E., M.M.S.

NIP : 2012786

**Program Studi Informatika
Fakultas Ilmu Komputer
Universitas Bhayangkara Jakarta Raya**

2024

LEMBAR PENGESAHAN

Judul Tugas akhir : ANALISIS SENTIMENT ULASAN
PENGGUNA LINKEDIN
PADA GOOGLE PLAYSTORE
DENGAN
METODE NAÏVE BAYES DAN KNN

Nama Mahasiswa : Aditya Putra Prastyo

Nomor Pokok Mahasiswa : 202010225259

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Tugas akhir : 6 Februari 2024

Jakarta, 12/02/2024

Ketua Tim Penguji : Wowon Priatna, ST., MTI

NIDN : 0429118007

Penguji I : Allan D. Alexander, ST, M. Kom

NIDN : 0305127404

Penguji II : Ajif Yunizar Pratama Yusuf, S.
SI, M. Eng

NIDN : 0328068603

MENGETAHUI,

Ketua

Program Studi Informatika

Ahmad Fathurrozi, S.E., M.M.S.I
NIP. 2012486

Dekan

Fakultas Ilmu Komputer

Dr. Dra. Tyastuti Sri Lestari, M.M
NIP. 1408206



LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Aditya Putra Prastyo
NPM : 202010225259
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Ulasan Pengguna LinkedIn Pada Google Playstore Dengan Metode Naïve Bayes dan KNN

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Bekasi, 16 Januari 2024
Penulis



Aditya Putra Prastyo

ABSTRAK

Aditya Putra Prastyo, 202010225259 Analisis Sentiment Ulasan Pengguna Google Playstore Dengan Metode Naïve Bayes Dan KNN. LinkedIn merupakan salah satu platform jaringan sosial yang berfokus pada profesional dan karier. keberhasilan dalam bisnis sering kali melibatkan pemahaman dan penanganan dengan baik terhadap kepuasan, kepercayaan dan kenyamanan pengguna. Identifikasi masalah dari ulasan pelanggan menjadi kunci, meskipun membaca dan mengklasifikasikan setiap ulasan secara manual dianggap tidak efisien dan sulit. Oleh karena itu, penelitian ini fokus pada analisis sentimen ulasan pengguna aplikasi LinkedIn untuk mengatasi tantangan ini. Data yang digunakan berasal dari ulasan dan rating pengguna di Google Play Store dan Survey, Ulasan pada google playstore dikumpulkan melalui teknik Scraping. Algoritma Naïve Bayes dan K Nearest Neighbors sebagai pembandingan digunakan dalam penelitian ini, dengan penerapan metodologi CRISP-DM untuk menggali informasi dari data teks. Hasil klasifikasi menggunakan Naïve Bayes menunjukkan tingkat akurasi sebesar 90% dan pada K Nearest Neighbors mendapatkan 89% . Penelitian ini memberikan kontribusi dengan menunjukkan bahwa teknik analisis sentimen dapat mengidentifikasi kekurangan dalam aplikasi, yang dapat dijadikan dasar evaluasi untuk optimalisasi produk dan layanan.

Kata Kunci : Naïve Bayes, K Nearest Neighbors, Analisis Sentimen, LinkedIn

ABSTRACT

Aditya Putra Prastyo, 202010225259 *Sentiment Analysis of LinkedIn User Reviews on Google Playstore Using Naïve Bayes and KNN Methods.* LinkedIn is a social networking platform that focuses on professionals and careers. Success in business often involves understanding and properly addressing user satisfaction, trust, and comfort. Identifying problems from customer reviews is key, although reading and classifying each review manually is considered inefficient and difficult. Therefore, this study focuses on sentiment analysis of LinkedIn app user reviews to address this challenge. The data used comes from user reviews and ratings on Google Play Store and Surveys. Reviews on Google Play Store are collected using scraping techniques. The Naïve Bayes and K Nearest Neighbors algorithms as comparisons were used in this research, with the application of the CRISP-DM methodology to extract information from text data. Classification results using Naïve Bayes show an accuracy rate of 90% and K Nearest Neighbors gets 89.6%. This research contributes by showing that sentiment analysis techniques can identify deficiencies in applications, which can be used as a basis for evaluation for optimizing products and services.

Keywords: *Naïve Bayes Method, Sentiment Analysis, LinkedIn and Data Mining*

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK**

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Aditya Putra Prastyo
NPM : 202010225259
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (Non-Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul :

Analisis Sentimen Ulasan Pengguna LinkedIn Pada Google Playstore Dengan Metode Naïve Bayes dan KNN

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi
Pada tanggal : 16 Januari 2024
Yang Menyatakan



Aditya Putra Prastyo

KATA PENGANTAR

Puji dan syukur penulis sampaikan kehadiran Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penyusunan proposal skripsi dengan judul “Analisis Sentiment Ulasan Pengguna LinkedIn pada Google Play Store dengan Metode Naïve Bayes dan KNN”. Proposal skripsi ini dibuat sebagai syarat untuk mencapai gelar Sarjana di Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya. Penyusunan proposal skripsi ini dapat terlaksana dengan baik berkat dukungan dan doa dari berbagai pihak. Pada kesempatan ini, penulis ingin mengungkapkan rasa terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam penyelesaian skripsi ini, yaitu kepada:

1. Bapak Irjen Pol.(Purn) Prof. Dr. Drs. Bambang Karsono, S.H., M.M. Selaku Rektor Universitas Bhayangkara Jakarta Raya.
 2. Ibu Dr. Dra. Tyastuti Sri Lestari., M.M. Selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
 3. Bapak Ahmad Fathurozi, S.E., MMSI. Selaku Ketua Program Studi Informatika Universitas Bhayangkara Jakarta Raya.
 4. Bapak Ajif Yunizar Pratama Yusuf, S.Si., M.Eng Selaku Dosen Pembimbing
 5. Harry Mardhiono dan Nuryunita Selaku Kedua Orang Tua dan Pemberi Semangat
- Penulis menyadari bahwa proposal skripsi ini mungkin memiliki kekurangan. Oleh karena itu, penulis memohon kritik dan saran yang membangun untuk dapat memperbaiki skripsi ini.

Bekasi, 18 September 2023

Peneliti

Aditya Putra Prastyo

202010225259

DAFTAR ISI

LEMBAR PERSETUJUAN	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
ABSTRACT	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Identifikasi Masalah	5
1.3 Rumusan Masalah	6
1.4 Batasan Masalah.....	6
1.5 Tujuan Penelitian.....	6
1.6 Manfaat Penelitian.....	7
1.7 Sistem Penulisan.....	7

BAB II LANDASAN TEORI	9
2.1 Landasan Teori (<i>State of Art</i>)	9
2.2 <i>Linkedin</i>	13
2.3 Sentimen Analisis.....	14
2.4 <i>Data Mining</i>	15
2.4.1 Pengertian <i>Data Mining</i>	15
2.4.2 Metode <i>Data Mining</i>	15
2.5 <i>Cross-Industry Standard for Data Mining (CRISP-DM)</i>	16
2.5.1 Pengertian <i>CRISP-DM</i>	16
2.5.2 Tahapan <i>CRISP-DM</i>	16
2.6 <i>Text Mining</i>	19
2.6.1 Pengertian <i>Text Mining</i>	19
2.6.2 Tahapan <i>Text Mining</i>	19
2.7 <i>Naïve Bayes</i>	21
2.8 <i>K-Nearest Neighbors</i>	21
2.9 <i>TF - IDF</i>	22
2.10 <i>Confusion Matrix</i>	23
BAB III METODOLOGI PENELITIAN	26
3.1 Kerangka Penelitian	26
3.2 Metode Pengumpulan Data	34
3.3 Kebutuhan Perangkat Lunak	35

3.4	Kebutuhan Perangkat Keras	36
BAB IV HASIL DAN PEMBAHASAN		37
4.1	<i>Scrapping Data</i>	37
4.2	<i>Bussiness Understanding</i>	39
4.3	<i>Data Understanding</i>	39
4.4	<i>Data Preparation</i>	40
4.4.1	<i>Casefolding</i>	41
4.4.2	Normalisasi	43
4.4.3	<i>Stopwords</i>	54
4.4.4	<i>Tokenizing</i>	59
4.4.5	<i>Stemming</i>	61
4.4.6	Label	63
4.4.7	Pembagian Data	69
4.4.8	<i>TF – IDF</i>	70
4.5	Naïve Bayes	77
4.5.1	Hitung Manual Naïve Bayes	77
4.5.2	Impelementasi Naïve Bayes	82
4.5.3	Evaluasi Naïve Bayes	82
4.6	<i>KNN (K Nearest Neighbor)</i>	85
4.6.1	Hitung Manual <i>KNN</i>	86
4.6.2	Impelementasi <i>KNN</i>	94

4.6.3	Evaluasi <i>KNN</i>	95
4.7	Visualisasi	98
4.7.1	Sentimen Positif	99
4.7.2	Sentimen Negatif.....	100
4.8	Hasil Penelitian.....	103
4.8.1	Hasil <i>Google Playstore</i>	103
4.8.2	Hasil Survey	104
BAB V PENUTUP.....		113
5.1	Kesimpulan.....	113
5.2	Saran.....	114
DAFTAR PUSTAKA.....		116
LAMPIRAN.....		121



DAFTAR TABEL

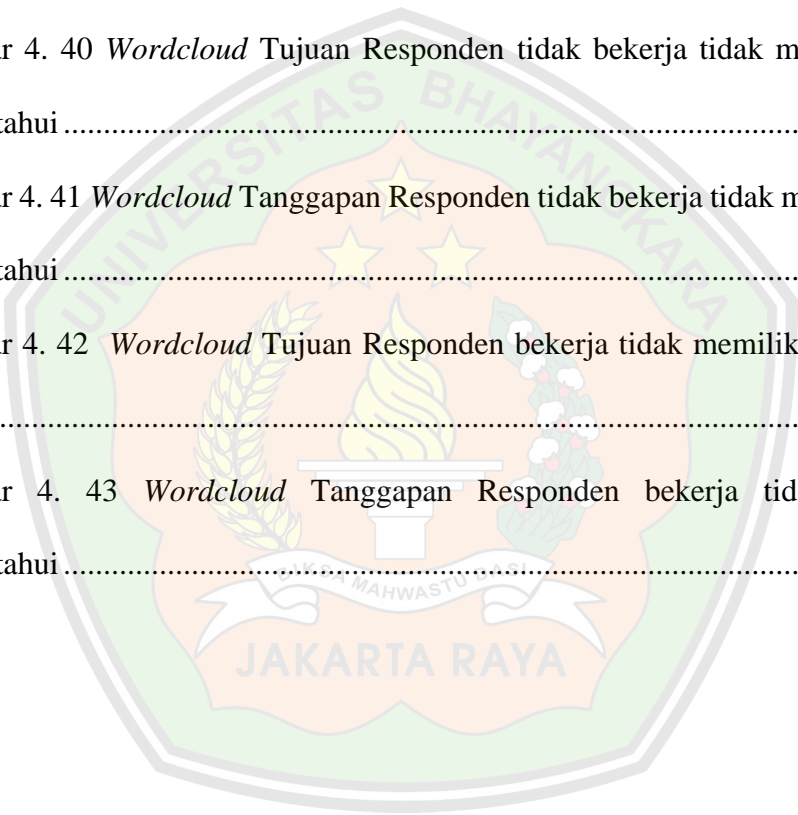
Tabel 2. 1 Studi Literatur	9
Tabel 2. 2 <i>Confusion Matrix</i>	24
Tabel 4. 1 Penerapan <i>Casefolding</i> pada data	42
Tabel 4. 2 Penerapan <i>Normalize</i> pada data	53
Tabel 4. 3 Penerapan <i>stopwords</i> pada data	57
Tabel 4. 4 Penerapan <i>Tokenizing</i> pada data	60
Tabel 4. 5 Penerapan <i>Stemming</i> pada data	62
Tabel 4. 6 Pelabelan pada data	68
Tabel 4. 7 Perhitungan <i>TF</i> dan <i>IDF</i>	72
Tabel 4. 8 Hasil <i>TF - IDF</i>	76
Tabel 4. 9 Kemunculan kata pada dokumen	78
Tabel 4. 10 <i>Confusion Matrix Naive Bayes</i>	83
Tabel 4. 11 <i>Confusion Matrix KNN</i>	96
Tabel 4. 12 Perbandingan kedua algoritma	103

DAFTAR GAMBAR

Gambar 1. 1 Rasio pengguna <i>Linkedin</i> di ketiga negara.....	2
Gambar 1. 2 Persentase responden menggunakan dan memiliki <i>Linkedin</i>	3
Gambar 2. 1 Logo <i>Linkedin</i>	13
Gambar 2. 2 Tahapan <i>CRISP - DM</i>	17
Gambar 3. 1 Kerangka Penelitian	26
Gambar 3. 2 Tahap Data <i>Preparation</i>	28
Gambar 3. 3 Proses <i>Case Folding</i>	29
Gambar 3. 4 Proses <i>TF – IDF</i>	31
Gambar 3. 5 Proses <i>Naive Bayes</i>	32
Gambar 3. 6 Proses <i>K- Nearest Neighbors</i>	33
Gambar 4. 1 Ulasan Pengguna <i>Linkedin</i> Pada <i>Google Playstore</i>	37
Gambar 4. 2 Penggunaan <i>Library</i> untuk <i>scrapping</i> data	38
Gambar 4. 3 <i>Scrapping</i> data.....	38
Gambar 4. 4 data dijadikan <i>dataframe</i>	38
Gambar 4. 5 Mengambil fitur <i>content</i> dan <i>score</i>	40
Gambar 4. 6 Mengecek Seluruh Kolom.....	40
Gambar 4. 7 Proses <i>Casefolding</i> Pada <i>Python</i> dengan bantuan <i>Library re</i>	41
Gambar 4. 8 Implementasi data yang tidak memiliki spasi	46
Gambar 4. 9 Menerapkan perbaikan pada kata <i>typo</i> dan <i>slangword</i>	52
Gambar 4. 10 Deklarasi <i>library nltk.corpus.stopwords</i>	54
Gambar 4. 11 Menggabungkan <i>stopwords nltk</i> dengan <i>stopwords</i> manual.....	57
Gambar 4. 12 Proses <i>Tokenize</i> pada data	59

Gambar 4. 13 Proses <i>Stemming</i> pada data	61
Gambar 4. 14 <i>Wordcloud</i> pada rating 1	63
Gambar 4. 15 <i>Wordcloud</i> pada rating 2	64
Gambar 4. 16 <i>Wordcloud</i> pada rating 3	65
Gambar 4. 17 <i>Wordcloud</i> pada rating 4	65
Gambar 4. 18 <i>Wordcloud</i> pada rating 5	66
Gambar 4. 19 Implementasi Labeling pada data.....	67
Gambar 4. 20 Menentukan X sebagai fitur dan Y sebagai target	67
Gambar 4. 21 Persentase Label Positif dan Negatif.....	68
Gambar 4. 22 Pembagian data dengan <i>library sklearn</i>	69
Gambar 4. 23 Jumlah antara data latih dan data uji	70
Gambar 4. 24 Implementasi <i>TF - IDF</i> pada python.....	71
Gambar 4. 25 Implementasi Naive Bayes.....	82
Gambar 4. 26 Implementasi Evaluasi pada model Naïve bayes	82
Gambar 4. 27 Implementasi model knn pada data.....	94
Gambar 4. 28 Implementasi evaluasi model <i>knn</i>	95
Gambar 4. 29 <i>Wordcloud</i> sentimen positif	99
Gambar 4. 30 Frekuensi kata pada sentimen positif	100
Gambar 4. 31 <i>Wordcloud</i> sentiment negatif	101
Gambar 4. 32 Frekuensi kata negatif	102
Gambar 4. 33 Diagram data hasil survey	105
Gambar 4. 34 <i>Wordcloud</i> Tujuan Tidak Bekerja Memiliki Mengetahui.....	106
Gambar 4. 35 <i>Wordcloud</i> Tanggapan Tidak Bekerja Memiliki Mengetahui.....	106

Gambar 4. 36 <i>Wordcloud</i> Tujuan Bekerja Memiliki Mengetahui	107
Gambar 4. 37 <i>Wordcloud</i> Tanggapan Responden Bekerja Memiliki Mengetahui	108
Gambar 4. 38 <i>Wordcloud</i> Tujuan Responden tidak bekerja tidak memiliki mengetahui	108
Gambar 4. 39 <i>Wordcloud</i> Tanggapan Responden tidak bekerja tidak memiliki mengetahui	109
Gambar 4. 40 <i>Wordcloud</i> Tujuan Responden tidak bekerja tidak memiliki tidak mengetahui	110
Gambar 4. 41 <i>Wordcloud</i> Tanggapan Responden tidak bekerja tidak memiliki tidak mengetahui	110
Gambar 4. 42 <i>Wordcloud</i> Tujuan Responden bekerja tidak memiliki mengetahui	111
Gambar 4. 43 <i>Wordcloud</i> Tanggapan Responden bekerja tidak memiliki mengetahui	111



DAFTAR LAMPIRAN

Lampiran 1 Plagiarisme	122
Lampiran 2 Biodata Mahasiswa.....	123
Lampiran 3 Kartu Bimbingan Skripsi.....	124

