

**ANALISIS SENTIMEN PADA ULASAN APLIKASI
GOJEK DI *GOOGLE PLAY STORE* DENGAN
METODE *SUPPORT VECTOR MACHINE*
DAN *NAÏVE BAYES***

SKRIPSI

Oleh:

Bhagaskara Farhan Wiguna

201910225401



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2023**

LEMBAR PERSETUJUAN PEMBIMBING

Judul Tugas akhir : Analisis Sentimen Pada Ulasan
Aplikasi Gojek di *Google Play Store*
Dengan Metode *Support Vector*
Machine dan *Naïve Bayes*

Nama Mahasiswa : Bhagaskara Farhan Wiguna

Nomor Pokok Mahasiswa : 201910225401

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 18 Juli 2023



Bekasi, 2 Agustus 2023

MENYETUJUI,

Pembimbing I

Pembimbing II

Herlawati, S.Si., M.M., M.Kom.

NIDN: 0311097302

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

NIDN: 0328068603

LEMBAR PENGESAHAN

Judul Tugas akhir : Analisis Sentimen Pada Ulasan
Aplikasi Gojek di *Google Play Store*
Dengan Metode *Support Vector*
Machine dan *Naïve Bayes*

Nama Mahasiswa : Bhagaskara Farhan Wiguna

Nomor Pokok Mahasiswa : 201910225401

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 18 Juli 2023

Bekasi, 2 Agustus 2023

MENGESAHKAN,

Ketua Tim Penguji : Prima Dina Atika, S.Kom., M.Kom.

NIDN: 0311037107

Penguji I : Dr. Rakhmi Khalida, S.T., M.M.S.I.

NIDN: 0304099201

Penguji II : Herlawati, S.Si., M.M., M.Kom.

NIDN: 0311097302

MENGETAHUI,

Ketua Program Studi

Dekan

Informatika

Fakultas Ilmu Komputer

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

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

NIP. 2012486

NIP. 1408206



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Bhagaskara Farhan Wiguna
NPM : 201910225401
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Pada Ulasan Aplikasi Gojek Di *Google Play Store* Dengan Metode *Support Vector Machine* Dan *Naïve Bayes*

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, 2 Agustus 2023

Penulis



Bhagaskara Farhan Wiguna

ABSTRAK

Bhagaskara Farhan Wiguna. 201910225401. Analisis Sentimen Pada Ulasan Aplikasi Gojek Di *Google Play Store* Dengan Metode *Support Vector Machine* dan *Naïve Bayes*.

Gojek merupakan salah satu transportasi *online* paling populer di Indonesia yang didirikan tahun 2010. Hingga saat ini Aplikasi Gojek telah diunduh mencapai 142 juta kali dengan jumlah *driver* lebih dari dua juta dan 400 ribu mitra pada layanan pesan-antar makanan. Karena meningkatnya penggunaan aplikasi Gojek dan pentingnya mengetahui pandangan pengguna tentang layanan yang diberikan oleh aplikasi tersebut. Pada penelitian ini menggunakan analisis sentimen dengan menggunakan *Support Vector Machine* dan *Naïve Bayes* untuk mengklasifikasikan sentimen positif dan sentiment negatif. Target label difokuskan pada label positif dan negatif dengan tujuan untuk menghindari bias yang ada pada ulasan yang berlabel netral pada Aplikasi Gojek. Proses penelitian meliputi pengumpulan data, pra-pemrosesan data, pembobotan dengan *Term Frequency-Inverse Document Frequency*, pelatihan *Support Vector Machine* dan *Naïve Bayes* dengan membagi data menjadi 90% data *training* dan 10% data *testing*, setelah itu dilakukan evaluasi hasil menggunakan *confusion matrix*. Hasil dari pengujian menggunakan algoritma *Support Vector Machine* menghasilkan *accuracy* 90%, *recall* 94%, *precision* 91%, dan *f1-score* 94%, sementara algoritma *Naïve Bayes* menghasilkan *accuracy* 77%, *recall* 96%, *precision* 77%, dan *f1-score* 85%.

Kata Kunci: Gojek, Analisis Sentimen, *Support Vector Machine*, *Naïve Bayes*, *Google Play Store*

ABSTRACT

Bhagaskara Farhan Wiguna. 201910225401. *Sentiment Analysis of Gojek Application Reviews in Google Play Store with Support Vector Machine and Naïve Bayes Method.*

Gojek is one of Indonesia's most popular online transportation, founded in 2010. The Gojek application has been downloaded one hundred forty-two million times with more than two million drivers and four hundred thousand partners in food delivery services. Due to the increasing use of the Gojek application and the importance of knowing user views about the services provided by the application. In this research, the sentiment analysis is using Support Vector Machine and the Naïve Bayes method to classify positive sentiment and negative sentiment. The target label focus on positive and negative labels to aims avoid the bias that exists in neutrally labeled reviews on the Gojek Application. The research process includes data collection, pre-processing the data, weighting with Term Frequency-Invers Document Frequency, Support Vector Machine, and Naïve Bayes training by dividing the data into 90% training data and 10% testing data and then evaluating the results using a confusion matrix. The results of testing using the Support Vector Machine algorithm resulted in 90% accuracy, 94% recall, 91% precision, and 94% f1-score, therefore the Naïve Bayes algorithm produces 77% accuracy, 96% recall, 77% precision, and 85% f1-score.

Keywords: *Gojek, Sentiment Analysis, Support Vector Machine, Naïve Bayes, Google Play Store*

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

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

Nama : Bhagaskara Farhan Wiguna
NPM : 201910225401
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 Pada Ulasan Aplikasi Gojek Di *Google Play Store* Dengan Metode *Support Vector Machine* Dan *Naïve Bayes*.

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 : 2 Agustus 2023
Yang Menyatakan



Bhagaskara Farhan Wiguna

KATA PENGANTAR

Puji syukur kami panjatkan kehadirat Allah SWT yang telah memberikan rahmat serta karunia-Nya kepada kami sehingga ini kami menyampaikan terimakasih kepada semua pihak terkait yang telah membantu terselesaikannya penelitian ini.

1. Bapak Irjen Pol (Purn) Prof. Dr. Drs. Bambang Karsono, SH., 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 Fathurrozi, S.E., M.M.S.I. Selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
4. Ibu Herlawati, S.Si., M.M., M.Kom. Selaku Dosen Pembimbing I Skripsi yang memberikan masukan serta saran dan motivasi dalam mengerjakan tugas akhir ini.
5. Bapak Ajif Yunizar Pratama Yusuf, S.Si., M.Eng. Selaku Dosen Pembimbing II Skripsi yang memberikan masukan serta saran dan motivasi dalam mengerjakan tugas akhir ini.
6. Kedua orang tua yang selalu mendoakan, memberi semangat serta motivasi untuk mendukung pengerjaan penelitian ini.
7. Teman-teman, sahabat, dan setiap orang yang entah sengaja atau tidak sengaja saya temui untuk berbagi cerita dan referensi, terima kasih atas waktunya.
8. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah dengan tulus ikhlas memberikan do'a dan dukungan sehingga penulis dapat menyelesaikan penelitian ini.

Saya menyadari bahwa tugas akhir ini masih jauh dari kesempurnaan. Oleh karena itu, segala saran dan kritik yang membangun dari semua pihak diharapkan demi perbaikan pada laporan selanjutnya. Semoga penelitian ini bermanfaat khususnya bagi penyusun dan bagi pembaca lain pada umumnya.

Bekasi, 19 Maret 2023

Bhagaskara Farhan Wiguna



DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
PERNYATAAN ORISINALITAS.....	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PUBLIKASI	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah	8
1.3 Rumusan Masalah	8
1.4 Tujuan Penelitian	9
1.5 Manfaat Penelitian.....	9
1.6 Batasan Masalah.....	10
1.7 Sistematika Penulisan	11
BAB II TINJAUAN PUSTAKA	13
2.1 Kajian Literatur.....	13
2.2 <i>Artificial Intelegent</i>	18
2.3 <i>Machine Learning</i>	19
2.4 <i>Data Mining</i>	20
2.5 <i>Text Mining</i>	21
2.5.1 <i>Knowledge Discovery in Database</i>	22

2.5.2 <i>Support Vector Machine</i>	30
2.5.3 <i>Naïve Bayes</i>	32
2.6 <i>Python</i>	33
2.7 Analisis Sentimen	33
2.7.1 <i>Scraping</i>	35
2.7.2 Ulasan	35
2.8 <i>Google Play Store</i>	36
2.9 Gojek	36
BAB III METODOLOGI PENELITIAN	39
3.1 Objek Penelitian	39
3.2 Kerangka Penelitian	40
3.3 Metode Pengumpulan Data	42
3.4 Metode Analisis Kebutuhan	42
3.4.1 Analisis Kebutuhan Sistem	44
3.4.2 Kebutuhan Perangkat Keras	44
3.4.3 Kebutuhan Perangkat Lunak	45
BAB IV HASIL DAN PEMBAHASAN	46
4.1 <i>Data Selection</i>	46
4.1.1 Implementasi <i>Scraping</i>	46
4.2 Pelabelan Data	49
4.3 <i>Preprocessing</i>	51
4.3.1 <i>Cleansing</i>	52
4.3.2 <i>Case Folding</i>	52
4.3.3 <i>Tokenizing</i>	53
4.3.4 <i>Normalization</i>	53

4.3.5 <i>Stopword Removal</i>	54
4.3.6 <i>Stemming</i>	54
4.4 <i>Transformation</i>	55
4.5 <i>Text Mining</i>	59
4.6 <i>Evaluation</i>	66
BAB V PENUTUP	73
5.1 Kesimpulan	73
5.2 Saran	74
DAFTAR PUSTAKA	76
LAMPIRAN	83



DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	15
Tabel 2.2 <i>Confusion Matrix</i>	29
Tabel 4.1 Penerapan <i>Cleansing</i>	52
Tabel 4.2 Penerapan <i>Case Folding</i>	52
Tabel 4.3 Penerapan <i>Tokenizing</i>	53
Tabel 4.4 Penerapan <i>Normalization</i>	53
Tabel 4.5 Penerapan <i>Stopword Removal</i>	54
Tabel 4.6 Penerapan <i>Stemming</i>	54
Tabel 4.7 Sebagian Dokumen Perhitungan TF-IDF	55
Tabel 4.8 Hasil TF Normalisasi	56
Tabel 4.9 Hasil Dari <i>Document Frequency</i>	56
Tabel 4.10 Hasil Dari <i>Inverse Document Frequency (IDF)</i>	57
Tabel 4.11 Hasil Dari TF-IDF	58
Tabel 4.12 Hasil TF-IDF Dari Beberapa Ulasan.....	60
Tabel 4.13 Data <i>Training</i>	63
Tabel 4.14 Hasil Perhitungan Probabilitas Pada Ulasan 1	64
Tabel 4.15 Hasil Perhitungan Probabilitas Pada Ulasan 2	64
Tabel 4.16 Tabel Untuk Uji Data Testing	64
Tabel 4.17 Hasil Probabilitas Pada Data <i>Testing</i>	66
Tabel 4.18 Jumlah Kemunculan Kata Pada Ulasan Positif	71
Tabel 4.19 Jumlah Kemunculan Kata Pada Ulasan Negatif.....	72
Tabel 4.20 Hasil Perbandingan Antara SVM Dengan <i>Naïve Bayes</i>	72

DAFTAR GAMBAR

Gambar 1.1 Contoh Ulasan Aplikasi Gojek Di <i>Google Play Store</i>	3
Gambar 1.2 Grafik Hasil Jawaban Kuesioner Masyarakat Sekitar	7
Gambar 2.1 Proses <i>Knowledge Discovery in Database</i> (KDD).....	24
Gambar 2.2 <i>Support Vector Machine</i>	31
Gambar 3.1 Data Ulasan Aplikasi Gojek	39
Gambar 3.2 Data Ulasan Setelah Dilakukan Penghapusan Beberapa Kolom.....	40
Gambar 3.3 Kerangka Penelitian.....	41
Gambar 3.4 Diagram Alir Analisis Data	44
Gambar 4.1 Tahapan <i>Scrapping</i>	47
Gambar 4.2 Hasil Scrapping Ulasan Aplikasi Gojek Di <i>Google Play Store</i>	48
Gambar 4.3 Tahapan Penghapusan Pada Kolom Yang Tidak Relevan	49
Gambar 4.4 Data Ulasan Aplikasi Gojek Setelah Penghapusan	49
Gambar 4.5 Tahapan Untuk Penghapusan Data Ulasan Dengan Score 3	50
Gambar 4.6 Tahapan Proses Pelabelan Data	51
Gambar 4.7 Tahapan Proses <i>Transformation</i>	55
Gambar 4.8 Tahapan <i>Text Mining</i> Dengan SVM	59
Gambar 4.9 Tahapan Text Mining Dengan <i>Naïve Bayes</i>	62
Gambar 4.10 Tahapan Evaluation Pada Algoritma SVM	67
Gambar 4.11 <i>Confusion Matrix</i> Dengan Skenario 90:10 Dengan SVM	67
Gambar 4.12 Tahapan Evaluation Pada Algoritma <i>Naïve Bayes</i>	68
Gambar 4.13 <i>Confusion Matrix</i> Dengan Skenario 90:10 Dengan <i>Naïve Bayes</i> ..	69
Gambar 4.14 <i>Word Cloud</i> Pada Ulasan Positif	70
Gambar 4.15 <i>Word Cloud</i> Pada Ulasan Negatif.....	71

DAFTAR LAMPIRAN

1. Hasil Plagiarisme *Checker Originality*
2. Biodata Mahasiswa
3. Kartu Konsultasi Tugas Akhir Pembimbing 1
4. Kartu Konsultasi Tugas Akhir Pembimbing 2
5. Surat Rekomendasi Pembimbing
6. Surat Keterangan Melakukan Riset
7. *Dataset* Penelitian

