

**ANALISIS SENTIMEN APLIKASI MYPERTAMINA
MENGUNAKAN METODE *NAÏVE BAYES* DAN
*K-NEAREST NEIGHBOR***

SKRIPSI

Oleh:

Aufa Rofi

201910225395



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

LEMBAR PERSETUJUAN PEMBIMBING

Judul Skripsi : Analisis Sentimen Aplikasi MyPertamina
Menggunakan Metode *Naïve Bayes* dan *K-Nearest Neighbor*.
Nama Mahasiswa : Aufa Rofi
Nomor Pokok Mahasiswa : 201910225395
Program Studi /Fakultas : Informatika/ Ilmu Komputer
Tanggal Lulus Ujian Skripsi : 14/7/2023

Bekasi, 27 Juni 2023
MENYETUJUI,
Pembimbing I Pembimbing II



Adi Muhajirin, S.Kom., M.Kom., M.M. Herlawati, S.Si., M.M., M.Kom.
NIDN: 0318038501 NIDN: 0311097302

LEMBAR PENGESAHAN

Judul Skripsi : Analisis Sentimen Aplikasi MyPertamina
Menggunakan Metode *Naïve Bayes* dan *K-Nearest Neighbor*.
Nama Mahasiswa : Aufa Rofi
Nomor Pokok Mahasiswa : 201910225395
Program Studi/ Fakultas : Informatika/ Ilmu Komputer
Tanggal Lulus Ujian Skripsi : 14/07/2023

Bekasi, 27 Juli 2023

MENGESAHKAN,

Ketua Tim Penguji : Dr. Dra Tyastuti Sri Lestari, M.M.

NIDN : 0327036701

Penguji I : Mayadi, S.Kom., M.Kom.

NIDN : 0408087802

Penguji II : Adi Muhajirin, S.Kom., M.Kom., M.M.

NIDN : 0318038501

MENGETAHUI,

Ketua Program Studi Informatika

Dekan Fakultas Ilmu Komputer

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

NIP. 2012486

Dr. Dra Tyastuti Sri Lestari, M.M.

NIP. 1408206



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini:

Nama : Aufa Rofi
NPM : 201910225395
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Aplikasi MyPertamina
Menggunakan Metode *Naïve Bayes* dan
K-Nearest Neighbor

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, 27 Juni 2023

Peneliti

METERAI
TEMPEL
02694AKX561175502
Aufa Rofi

ABSTRAK

Aufa Rofi. 201910225395. Analisis Sentimen Aplikasi MyPertamina Menggunakan Metode *Naive Bayes* dan *K-Nearest Neighbor*. Bekasi: Fakultas Ilmu Komputer. Universitas Bhayangkara Jakarta Raya. 2023

Aplikasi MyPertamina adalah aplikasi *e-money* yang menerapkan sistem transaksi non-tunai dalam pembelian BBM bersubsidi yang sudah terintegrasi dengan aplikasi LinkAja. Aplikasi MyPertamina telah tersedia pada *google play store* dengan rating 2.5 dan mendapat lebih dari 307000 komentar pengguna per 9 Oktober 2022. Analisis sentimen diterapkan pada penelitian ini untuk mengklasifikasikan ulasan MyPertamina kedalam sentimen positif, negatif, dan netral menggunakan metode *Naïve Bayes* dan *K-Nearest Neighbor*. Pengambilan data ulasan aplikasi MyPertamina pada situs *google play store* dilakukan menggunakan teknik *web scraping*. Metode *text preprocessing* digunakan untuk mengolah data ulasan pengguna aplikasi MyPertamina. Per 23 Juni 2023 didapatkan hasil dari penelitian ini bahwasanya aplikasi MyPertamina kurang memenuhi harapan konsumen dengan rating 3.1 khususnya pada kasus pendaftaran pengguna penerimaan BBM Subsidi. Adapun hasil penelitian mengenai analisis sentimen aplikasi MyPertamina dengan menggunakan data ulasan sebanyak 2478 dan perbandingan data sebesar 70%:30% dengan metode *Naïve Bayes* menghasilkan skor akurasi, presisi dan recall sebesar 83% dan skor AUC sebesar 0.48. Sedangkan untuk algoritma *K-Nearest Neighbor* dengan pengujian menggunakan nilai k-8 menghasilkan skor akurasi, presisi, recall sebesar 89% dengan skor AUC sebesar 0.59.

Kata Kunci: MyPertamina, *E-Money*, Analisis Sentimen, *Naive Bayes Classifier*, *K-Nearest Neighbor*.

ABSTRACT

Aufa Rofi. 201910225395. *Sentiment Analysis of MyPertamina Application Using Naive Bayes and K-Nearest Neighbor Method.* Bekasi: Faculty of Computer Science. Jakarta Bhayangkara University. 2023

MyPertamina application is an e-money application that implements a non-cash transaction system in purchasing subsidized fuel that has been integrated with the LinkAja application. MyPertamina application has been available on google play store with a rating of 2.5 and received more than 307000 user comments as of October 9, 2022. Sentiment analysis is applied in this research to classify MyPertamina reviews into positive, negative, and neutral sentiments using Naïve Bayes and K-Nearest Neighbor methods. Retrieval of MyPertamina application review data on the Google Play Store site is done using web scraping techniques. Text preprocessing method is used to process MyPertamina application user review data. As of June 23, 2023, the results of this study showed that the MyPertamina application did not meet consumer expectations with a rating of 3.1, especially in the case of user registration of subsidized fuel recipients. The results of research on sentiment analysis of the MyPertamina application using 2478 review data and data comparison of 70%: 30% with the Naïve Bayes method resulted in an accuracy, precision and recall score of 83% and an AUC score of 0.48. While for the K-Nearest Neighbor algorithm with testing using the k-8 value resulted in an accuracy score, precision, recall of 89% with an AUC score of 0.59.

Keywords: *MyPertamina, E-Money, Sentiment Analysis, Naive Bayes Classifier, K-Nearest Neighbor.*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

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

Nama : Aufa Rofi
NPM : 201910225395
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 Aplikasi MyPertamina Menggunakan Metode *Naïve Bayes* dan *K-Nearest Neighbor*

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 : 27 Juli 2023
Yang Menyatakan


Aufa Rofi

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga skripsi yang berjudul “Analisis Sentimen Aplikasi MyPertamina Menggunakan Metode *Naive Bayes* dan *K-Nearest Neighbor*” dapat diselesaikan dengan tepat waktu. Dalam penyusunan skripsi ini tidak luput dari kesulitan dan hambatan. Namun, terdapat pihak – pihak yang memberikan doa dan selalu menyemangati sehingga penyusunan skripsi ini dapat diselesaikan. Oleh karena itu ucapan terima kasih disampaikan kepada:

1. Bapak Irjen Pol. (Purn) 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 Fathurrozi, S.E., M.M.S.I. Selaku Ketua Program Studi Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
4. Bapak Adi Muhajirin, S.Kom., M.Kom., M.M. Selaku Dosen Pembimbing satu, dan Ibu Herlawati, S.Si., M.M., M.Kom. Selaku Dosen Pembimbing dua yang telah memberikan bimbingan dalam penyusunan skripsi ini.
5. Keluarga tercinta terutama kedua orang tua yang selalu memberikan doa, semangat serta dukungan dan teman-teman yang turut memberikan dukungan dan semangat dalam proses penyusunan skripsi ini.

Peneliti menyadari bahwa masih terdapat banyak kesalahan dalam penyusunan skripsi ini. Maka dari itu diharapkan kritik dan saran yang membangun agar menjadi bahan evaluasi untuk penelitian selanjutnya. Peneliti juga berharap hasil penelitian ini bermanfaat bagi siapapun yang membacanya.

Bekasi, 27 Juni 2023

Peneliti



(Aufa Rofi)

201910225395

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	ix
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah.....	6
1.3 Rumusan Masalah	6
1.4 Batasan Masalah.....	6
1.5 Tujuan Penelitian	7
1.6 Manfaat Penelitian	7
1.7 Sistematika Penulisan.....	8
BAB II TINJAUAN PUSTAKA	10
2.1 <i>State of The Art</i>	10

2.2	Teori Pendukung	12
2.2.1	<i>Data Science</i>	12
2.2.2	<i>Data Mining</i>	13
2.2.3	<i>Text Mining</i>	14
2.2.4	<i>Machine Learning</i>	14
2.2.5	<i>Natural Language Processing (NLP)</i>	15
2.2.6	<i>Analisis Sentimen</i>	15
2.2.7	<i>Web Scraping</i>	16
2.2.8	<i>Text Preprocessing</i>	16
2.2.9	<i>Vektorisasi</i>	17
2.2.10	<i>Klasifikasi</i>	18
2.2.11	<i>Naïve Bayes Classifier</i>	18
2.2.12	<i>K-Nearest Neighbor</i>	19
2.2.13	<i>Multiclass Confusion Matrix</i>	20
2.2.14	<i>ROC (Receiver Operating Characteristic)</i>	21
2.2.15	<i>Slovin</i>	22
2.2.16	<i>Google Play Store</i>	23
2.2.17	<i>MyPertamina</i>	23
2.2.18	<i>Python</i>	24
2.2.19	<i>Google Colaboratory</i>	25
2.2.20	<i>Fishbone Diagram</i>	25
2.2.21	<i>Flowchart</i>	26
BAB III	METODOLOGI PENELITIAN	28
3.1	Objek Penelitian	28

3.2	Tempat dan Waktu Penelitian	28
3.3	Kerangka Penelitian	28
3.4	Pembuatan Model <i>Machine Learning</i>	29
3.5	Metode Pengumpulan Data	29
3.6	Sampel Penelitian	30
3.7	Analisis Sentimen	31
3.8	<i>Preprocessing</i>	31
3.8.1	<i>Cleansing</i>	32
3.8.2	<i>Case Folding</i>	32
3.8.3	<i>Tokenizing</i>	32
3.8.4	<i>Removing Stopword</i>	32
3.8.5	<i>Stemming</i> KBBI	33
3.9	Pelabelan Data	33
3.10	Vektorisasi	33
3.11	Klasifikasi Dengan <i>Naive Bayes</i> dan KNN	34
BAB IV HASIL DAN PEMBAHASAN		35
4.1	Pengumpulan Data	35
4.2	Manipulasi Data	40
4.3	Analisis Deskriptif	41
4.4	<i>Pre-processing Data</i>	42
4.4.1	<i>Data Cleansing</i>	43
4.4.2	<i>Case Folding</i>	45
4.4.3	<i>Tokenizing</i>	47
4.4.4	<i>Removing Stopword</i>	48

4.4.5	<i>Stemming</i> KBBI	51
4.5	Pelabelan Data.....	53
4.6	Visualisasi <i>Wordcloud</i>	58
4.7	Vektorisasi.....	60
4.8	Pembagian Data Latih dan Uji	61
4.9	Klasifikasi dengan Metode <i>Naïve Bayes</i>	62
4.10	Klasifikasi dengan Metode <i>K-Nearest Neighbor</i>	68
4.11	Perhitungan Manual <i>Naïve Bayes</i>	74
4.12	Perhitungan Manual <i>K-Nearest Neighbor</i>	77
4.13	Perbandingan Metode <i>Naïve Bayes</i> dan KNN	80
4.14	Infografis Analisis Sentimen Aplikasi MyPertamina	82
BAB V PENUTUP		83
5.1	Kesimpulan	83
5.2	Saran.....	84
DAFTAR PUSTAKA.....		86
LAMPIRAN		90

DAFTAR TABEL

Tabel 2.1 <i>State of The Art</i>	10
Tabel 2.2 <i>Multiclass Confusion Matrix</i>	20
Tabel 2.3 <i>Flowchart Symbols</i>	27
Tabel 4.1 Data Hasil <i>Scraping</i>	38
Tabel 4.2 Data Hasil <i>Cleansing</i>	45
Tabel 4.3 Data Hasil <i>Case Folding</i>	46
Tabel 4.4 Data Hasil <i>Tokenizing</i>	48
Tabel 4.5 Data Hasil <i>Removing Stopword</i>	50
Tabel 4.6 Data Hasil <i>Stemming KBBI</i>	53
Tabel 4.7 Perbandingan Data Latih dan Data Uji	62
Tabel 4.8 Evaluasi <i>Multiclass Confusion Matrix Naïve Bayes</i>	65
Tabel 4.9 Evaluasi <i>Multiclass Confusion Matrix K-Nearest Neighbor</i>	72
Tabel 4.10 Contoh Data Latih <i>Naïve Bayes</i>	74
Tabel 4.11 Contoh Data Uji <i>Naïve Bayes</i>	75
Tabel 4.12 Hasil Klasifikasi Data Uji <i>Naïve Bayes</i>	77
Tabel 4.13 Contoh Data Latih <i>K-Nearest Neighbor</i>	77
Tabel 4.14 Contoh Data Uji <i>K-Nearest Neighbor</i>	78
Tabel 4.15 Perhitungan Nilai Jarak	78
Tabel 4.16 Urutan Jarak Terdekat	79
Tabel 4.17 Jarak Terdekat Sesuai Nilai K.....	79
Tabel 4.18 Hasil Klasifikasi Data Uji KNN	80

DAFTAR GAMBAR

Gambar 1.1 Grafik Pendapat Warga	3
Gambar 2.1 Logo <i>Google Play Store</i>	23
Gambar 2.2 Logo Aplikasi MyPertamina	23
Gambar 2.3 Logo <i>Python</i>	24
Gambar 2.4 Logo <i>Google Colaboratory</i>	25
Gambar 3.1 Diagram <i>Fishbone</i>	28
Gambar 3.2 <i>Flowchart</i> Pembuatan Model <i>Machine Learning</i>	29
Gambar 3.3 <i>Dataset</i> Ulasan MyPertamina	30
Gambar 3.4 <i>Preprocessing Flowchart</i>	31
Gambar 4.1 <i>Scraper Library</i>	36
Gambar 4.2 Parameter <i>Scraper</i>	37
Gambar 4.3 Menampilkan Data Terteratas	37
Gambar 4.4 Menyimpan Data <i>Scraping</i>	38
Gambar 4.5 Manipulasi Data	41
Gambar 4.6 Distribusi Rating Pengguna MyPertamina	42
Gambar 4.7 <i>Data Cleansing</i>	44
Gambar 4.8 <i>Case Folding</i>	46
Gambar 4.9 <i>Tokenizing</i>	47
Gambar 4.10 <i>Removing Stopword</i>	49
Gambar 4.11 <i>Stemming</i> KBBI	51
Gambar 4.12 Aturan Pelabelan Data	55
Gambar 4.13 <i>Labelling Code</i>	55
Gambar 4.14 Hasil Proses Pelabelan	56

Gambar 4.15 Persentase Kelas Sentimen.....	57
Gambar 4.16 Histogram Distribusi Label Ulasan	57
Gambar 4.17 <i>Wordcloud</i> Seluruh Ulasan	58
Gambar 4.18 <i>Wordcloud</i> Ulasan Positif	59
Gambar 4.19 <i>Wordcloud</i> Ulasan Negatif.....	59
Gambar 4.10 <i>Wordcloud</i> Ulasan Netral.....	60
Gambar 4.21 Vektorisasi.....	60
Gambar 4.22 Pelatihan Model <i>Naïve Bayes</i>	63
Gambar 4.23 Evaluasi Model <i>Naïve Bayes</i>	65
Gambar 4.24 Kurva ROC Model <i>Naïve Bayes</i>	67
Gambar 4.25 Pelatihan Model KNN.....	69
Gambar 4.26 Grafik Akurasi KNN	70
Gambar 4.27 Evaluasi Model KNN.....	71
Gambar 4.28 Kurva ROC Model <i>K-Nearest Neighbor</i>	73
Gambar 4.29 Infografis Analisis Sentimen Aplikasi MyPertamina.....	82

DAFTAR LAMPIRAN

1. Hasil Plagiarisme *Checker Originality*
2. Biodata Mahasiswa
3. Kartu Bimbingan Tugas Akhir
4. Surat Rekomendasi Dari Pembimbing

