

**ANALISIS SENTIMEN PENGGUNA E-WALLET
PADA TWITTER MENGGUNAKAN *ALGORITMA*
*NAÏVE BAYES CLASSIFIER***

SKRIPSI

Oleh :

RAFI ADYATMA PUTRA

201910225277



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

LEMBAR PERSETUJUAN PEMBIMBING

Judul Tugas Akhir : Analisis Sentimen Pengguna *E-wallet* pada
Twitter Menggunakan Algoritma *Naïve Bayes*
Classifier

Nama Mahasiswa : Rafi Adyatma Putra

Nomor Pokok Mahasiswa : 201910225277

Program studi/Fakultas : Informatika/Illmu Komputer

Tanggal Lulus Ujian Skripsi : 16 Februari 2024

Jakarta, 16 Februari 2024

MENYETUJUI,

Pembimbing I

Allan Desi Alexander, S.T., M.Kom.

NIDN: 0305127404

Ketua Program Studi

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

NIP: 2012486

**Program Studi Informatika
Fakultas Ilmu Komputer
Universitas Bhayangkara Jakarta Raya
2024**

LEMBAR PENGESAHAN

Judul Skripsi : Analisis Sentimen Pengguna *E-wallet* pada
Twitter Menggunakan Algoritma *Naïve Bayes Classifier*

Nama Mahasiswa : Rafi Adyatma Putra

Nomor Pokok Mahasiswa : 201910225277

Program Studi /Fakultas : Informatika/ Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 16 Februari 2024

Jakarta, 16 Februari 2024
MENGESAHKAN,

Ketua Tim Penguji : Abrar Hiswara, S.T., M.M., M.Kom.

NIDN : 0324028101

Penguji II : Rafika Sari, S.Si., M.Si.

NIDN : 0329098902

Penguji III : Allan Desi Alexander, S.T., M.Kom.

NIDN : 00305127404

MENGETAHUI,

Ketua Program Studi Informatika

Dekan Fakultas Ilmu Komputer

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

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



LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Rafi Adyatma Putra

NPM : 201910225277

Program Studi : Informatika

Fakultas : Ilmu Komputer

Judul Tugas Akhir : Analisis Sentimen Pengguna *E-wallet* pada *Twitter*
Menggunakan Algoritma *Naïve Bayes Classifier*

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 bertanggungjawabkan 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.

Jakarta, 16 Februari 2024

Penulis



Rafi Adyatma Putra

ABSTRAK

Rafi Adyatma Putra. 201910225277. Analisis Sentimen Pengguna *E-wallet* pada *Twitter* Menggunakan Algoritma *Naïve Bayes Classifier*.

Twitter merupakan *microblog* yang memungkinkan penggunanya dapat mengirim dan membaca informasi ataupun opini terhadap isu-isu yang bermacam-macam dalam bentuk 280 karakter pesan *twitter* yang dinamakan *tweet*. Pada tahun 2023 dicapai oleh *e-wallet* yaitu dengan persentase sebesar 81% pada tahun 2023 dan diikuti oleh *virtual account* dengan 60%. Selain itu, ada juga metode transfer bank dan *cash/COD (cash on delivery)* dengan persentase masing-masing mencapai 55%. Maka dari itu perlu diadakannya penelitian analisis sentimen terhadap ulasan komentar aplikasi dompet digital agar predikat aplikasi terbaik menjadi relevan. Tujuan dari penelitian ini adalah untuk melakukan analisis sentimen terhadap aplikasi dompet digital berdasarkan ulasan komentar yang ada di *Twitter* menggunakan algoritma *Naïve Bayes Classifier*. Hal – hal tersebut selalu dianggapi oleh pengguna *twitter*. *Tweet* yang disebar oleh pengguna *twitter* ini mengandung konten yang berbeda-beda sesuai dengan kebiasaan dan gaya menulis pengguna. Dan hasil dari *lexicon based* menunjukkan *e-wallet* Dana sebanyak 3980 dengan kelas positif sebanyak 1780 *tweet* dan kelas negatif sebanyak 2200 *tweet*. Sedangkan *e-wallet* OVO mendapatkan data bersih sebanyak 3976 *tweet* dengan kelas positif sebanyak 953 *tweet* dan kelas negatif sebanyak 3023 *tweet*. Dan yang terakhir *e-wallet* LinkAja mendapatkan data bersih sebanyak 3972 *tweet* dengan kelas positif sebanyak 1429 *tweet* dan kelas negatif sebanyak 2543 *tweet* dari penelitian ini. Dan pada algoritma *Naïve Bayes Classifier* dalam menganalisis sentimen *tweet* dengan *model split data* dengan 4 rasio menunjukkan bahwa hasil akurasi terbaik dari pembagian data tersebut pada ketiga *e-wallet* tersebut yang mendapatkan akurasi terbesar didapat pada aplikasi *e-wallet* LinkAja dengan akurasi 85.17%, kedua didapat aplikasi *e-wallet* OVO dengan akurasi 81.97%, sedangkan terakhir didapat pada aplikasi *e-wallet* Dana dengan akurasi 79.39%.

Kata Kunci : *Twitter, E-wallet, Lexicon Based, Naïve Bayes Classifier*

ABSTRACT

Rafi Adyatma Putra. 201910225277. *Analysis of E-wallet User Sentiment on Twitter Using the Naïve Bayes Classifier Algorithm.*

Twitter is a microblog that allows users to send and read information or opinions on various issues in the form of 280 character Twitter messages called tweets. In 2023 it will be achieved by e-wallets, namely with a percentage of 81% in 2023 and followed by virtual accounts with 60%. Apart from that, there are also bank transfer and cash/COD (cash on delivery) methods with each percentage reaching 55%. Therefore, it is necessary to conduct sentiment analysis research on digital wallet application comment reviews so that the title of best application becomes relevant. The aim of this research is to conduct sentiment analysis of digital wallet applications based on reviews of comments on Twitter using the Naïve Bayes Classifier algorithm. Twitter users always respond to these things. Tweets distributed by Twitter users contain different content according to the user's writing habits and style. And the results from the lexicon based show there are 3980 Dana e-wallets with a positive class of 1780 tweets and a negative class of 2200 tweets. Meanwhile, the OVO e-wallet received clean data of 3976 tweets with 953 tweets in the positive class and 3023 tweets in the negative class. And finally, the LinkAja e-wallet received clean data of 3972 tweets with 1429 positive class tweets and 2543 tweets in the negative class from this research. And the Naïve Bayes Classifier algorithm in analyzing tweet sentiment with a split data model with 4 ratios shows that the best accuracy results from dividing the data on the three e-wallets which got the greatest accuracy were obtained on the LinkAja e-wallet application with an accuracy of 85.17%, second obtained on the OVO e-wallet application with an accuracy of 81.97%, while the last one was obtained on the Dana e-wallet application with an accuracy of 79.39%.

Keyword : *Twitter, E-wallet, Lexicon Based, Naïve Bayes Classifier*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

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

Nama : Rafi Adyatma Putra
NPM : 201910225277
Program Studi : Ilmu Komputer
Fakultas : Informatika
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 Pengguna *E-wallet* pada *Twitter* Menggunakan Algoritma *Naïve Bayes Classifier*”

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 : Jakarta
Pada tanggal : 16 Februari 2024
Yang Menyatakan



Rafi Adyatma Putra

KATA PENGANTAR

Puji syukur saya ucapkan kepada Allah SWT karena atas nikmat dan rahmat Nya, penulis dapat menyelesaikan skripsi ini tugas akhir skripsi yang berjudul “Analisis Sentimen Pengguna *E-wallet* pada *Twitter* Menggunakan Algoritma *Naïve Bayes Classifier*” sebagai syarat untuk mencapai gelar Sarjana Komputer Program Studi Informatika Fakultas Ilmu Komputer Univeritas Bhayangkara Jakarta Raya. Selama proses penyusunan skripsi ini, tidak terlepas dari beberapa bantuan, saran, kritik dan dukungan. Dalam kesempatan ini, penulis berterima kasih kepada:

1. Bapak Irjen Polisi (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 Fathurrozi, S.E., M.M.S.I. Selaku ketua Program Studi Informatika Universitas Bhayangkara Jakarta Raya.
4. Bapak Allan Desi Alexander, S.T., M.Kom. Selaku pembimbing I yang telah memberikan arahan serta motivasi yang mudah dipahami saat sesi bimbingan.
5. Kedua orang tua yang selalu memberi semangat dan dukungan, selalu mendoakan setiap harinya agar diberikan kesehatan dan kemudahan dalam menyelesaikan Seminar Proposal.
6. Seluruh teman-teman Program Studi Informatika terkhusus Yusuf Bintang Mulyana, Muhammad Anwar Lirboyo, Fahry Ramadhani, Muhammad

Saddam Rahmat, Sultan Zaki Firdaus, Rafly Cahya Pradana (Pabeng Pala Gepeng), Sadam Rahmat dan Muhammad Fahri dengan atas bantuan selalu memberi semangat, dukungan, dan kerja samanya dalam perkuliahan.

Jakarta, 16 Februari 2024

Rafi Adyatma Putra

201910225277



DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
1.6 Batasan Masalah	6
1.7 Sistematika Penulisan	6
BAB II LANDASAN TEORI	8
2.1 Tinjauan Pustaka	8
2.2 <i>E-Wallet</i>	13
2.3 Analisis Sentimen	16
2.4 <i>Twitter</i>	17
2.5 Data Mining	18
2.6 <i>Text Mining</i>	20
2.7 <i>Web Crawler</i>	21
2.8 <i>Knowledge Discovery in Database (KDD)</i>	21

2.9 <i>Text Preprocessing</i>	24
2.10 <i>Lexicon Based</i>	26
2.11 <i>Term Frequency–Inverse Document Frequency (TF-IDF)</i>	27
2.12 <i>Naïve Bayes Classifier</i>	29
2.13 <i>Confusion Matrix</i>	31
2.14 <i>Python</i>	33
2.15 <i>Google Colaboratory</i>	33
2.16 <i>Diagram Alir (Flowchart)</i>	33
BAB III METODOLOGI PENELITIAN	36
3.1 <i>Obyek Penelitian</i>	36
3.2 <i>Kerangka Penelitian</i>	36
3.4 <i>Analisis Data</i>	39
3.5 <i>Metode Pengumpulan Data</i>	42
3.6 <i>Perangkat Penelitian</i>	43
BAB IV HASIL DAN PEMBAHASAN	44
4.1 <i>Pengumpulan Data (Data Selection)</i>	44
4.2 <i>Text Preprocessing</i>	47
4.2.1 <i>Case Folding</i>	47
4.2.2 <i>Cleaning/Cleasng</i>	48
4.2.3 <i>Normalisasi</i>	50
4.2.4 <i>Tokenizing</i>	52
4.2.5 <i>Filtering</i>	52
4.2.6 <i>Stemming</i>	54
4.3 <i>Tranformation</i>	55
4.3.1 <i>Pelabelan Kamus Lexicon Based</i>	56
4.3.2 <i>Visualisasi</i>	58
4.3.3 <i>Pembobotan Term Frequency–Inverse Document Frequency</i>	65
4.4 <i>Data Mining</i>	69
4.4.1 <i>Data Latih Dan Data Uji</i>	69
4.4.2 <i>Algoritma Naïve Bayes Classifier</i>	70
4.5 <i>Evaluasi</i>	82
4.5.1 <i>Klasifikasi Rasio 60:40 (Aplikasi E-Wallet Dana)</i>	82

4.5.2 Klasifikasi Rasio 70:30 (Aplikasi <i>E-Wallet</i> Dana)	83
4.5.3 Klasifikasi Rasio 80:20 (Aplikasi <i>E-Wallet</i> Dana)	84
4.5.4 Klasifikasi Rasio 90:10 (Aplikasi <i>E-wallet</i> Dana)	85
4.5.5 Klasifikasi Rasio 60:40 (Aplikasi <i>E-Wallet</i> OVO)	85
4.5.6 Klasifikasi Rasio 70:30 (Aplikasi <i>E-Wallet</i> OVO)	86
4.5.7 Klasifikasi Rasio 80:20 (Aplikasi <i>E-Wallet</i> OVO)	87
4.5.8 Klasifikasi Rasio 90:10 (Aplikasi <i>E-Wallet</i> OVO)	88
4.5.9 Klasifikasi Rasio 60:40 (Aplikasi <i>E-Wallet</i> LinkAja)	89
4.5.10 Klasifikasi Rasio 70:30 (Aplikasi <i>E-Wallet</i> LinkAja)	90
4.5.11 Klasifikasi Rasio 80:20 (Aplikasi <i>E-Wallet</i> LinkAja)	90
4.5.12 Klasifikasi Rasio 90:10 (Aplikasi <i>E-Wallet</i> LinkAja)	91
4.5.13 Hasil Evaluasi <i>Confusion Matriks</i>	92
BAB V PENUTUP	94
5.1 Kesimpulan	94
5.2 Saran	95
DAFTAR PUSTAKA	97
LAMPIRAN	ci

DAFTAR GAMBAR

Gambar 1.1 Jumlah Metode Pembayaran Digunakan Di Indonesia Tahun 2023 ...	2
Gambar 2.1 Grafik <i>E-wallet</i> yang paling banyak digunakan.....	14
Gambar 2.2 Perubahan Logo <i>Twitter</i> (2010 – 2023)	17
Gambar 2.3 Tahapan KDD	22
Gambar 3.1 Kerangka Penelitian	38
Gambar 3.2 Kerangka Tahap <i>Data Selection</i>	39
Gambar 3.3 Kerangka Tahap <i>Text Preprocessing</i>	40
Gambar 3.4 Kerangka Tahap <i>Transformation</i>	41
Gambar 3.5 Kerangka Tahap Data Mining	42
Gambar 3.6 Hasil <i>Crawling</i> Data <i>Tweet E-wallet</i>	43
Gambar 4.1 <i>Script/Source Code</i> Proses <i>Case Folding</i>	47
Gambar 4.2 <i>Script/Source Code</i> Mendownload <i>Slang Words</i>	50
Gambar 4.3 <i>Script/Source Code</i> menormalisasi <i>Tweet</i> Setiap Dokumen	51
Gambar 4.4 <i>Script/Source Code</i> Proses <i>Tokenizing</i>	52
Gambar 4.5 <i>Script/ Source Code Library</i> Pada <i>Filtering</i> Dan <i>Stopwords</i>	53
Gambar 4.6 Membuat File Dataset Data <i>Stopwords</i> Untuk Tahap <i>Filtering</i>	53
Gambar 4.7 <i>Script/Source Code</i> Proses <i>Filtering</i>	54
Gambar 4.8 <i>Script/Source Code</i> Proses <i>Stemming</i>	55
Gambar 4.9 Hasil Grafik Kelas Sentimen <i>Tweet E-wallet</i> Dana	60
Gambar 4.10 <i>Wordcloud</i> Ulasan Positif <i>E-wallet</i> Dana	60
Gambar 4.11 <i>Wordcloud</i> Ulasan Negatif <i>E-wallet</i> Dana	61
Gambar 4.12 Hasil Grafik Kelas Sentimen <i>Tweet E-wallet</i> OVO	62
Gambar 4.13 <i>Wordcloud</i> Ulasan Positif <i>E-wallet</i> OVO	62
Gambar 4.14 <i>Wordcloud</i> Ulasan Negatif <i>E-wallet</i> OVO.....	63
Gambar 4.15 Hasil Grafik Kelas Sentimen <i>Tweet E-wallet</i> LinkAja	64
Gambar 4.16 <i>Wordcloud</i> Ulasan Positif <i>E-wallet</i> LinkAja.....	64
Gambar 4.17 <i>Wordcloud</i> Ulasan Negatif <i>E-wallet</i> OVO.....	65

DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka.....	8
Tabel 2.2 <i>Confusion Matrix</i>	31
Tabel 2.3 Simbol Simbol <i>Flowchart</i>	34
Tabel 3.1 Spesifikasi <i>Hardware</i> Dan <i>Software</i>	43
Tabel 4.1 <i>Script Web Crawler Tweet Harvest E-wallet</i> Dari <i>Twitter</i>	45
Tabel 4.2 Data Hasil <i>Web Crawler Tweet Harvest E-wallet</i>	46
Tabel 4.3 Hasil Proses <i>Case Folding</i>	48
Tabel 4.4 <i>Script/Source Code</i> Proses <i>Cleaning/Cleasing</i>	48
Tabel 4.5 Hasil Proses <i>Cleaning/Cleaning</i>	49
Tabel 4.6 Hasil Proses Normalisasi	51
Tabel 4.7 Hasil Proses <i>Tokenezing</i>	52
Tabel 4.8 Hasil Proses <i>Filtering</i>	54
Tabel 4.9 Hasil Proses <i>Stemming</i>	55
Tabel 4.10 Jumlah Aktual Data Set.....	56
Tabel 4.11 <i>Script/Source Code</i> Tahapan Proses <i>Labeling Lexicon Based</i>	56
Tabel 4.12 Hasil Proses <i>Labeling Dengan Lexicon Based</i>	57
Tabel 4.13 <i>Script Visualisasi Data Ulasan Grafik Dan Worlcloud</i>	58
Tabel 4.14 <i>Script/Source Code</i> Tahapan Proses Pembobotan TF-IDF.....	66
Tabel 4.15 Hasil Pembobotan Kata TF-IDF (Sampel Teks Urutan Ke-22).....	67
Tabel 4.16 Hasil Proses Pembobotan TF-IDF	68
Tabel 4.17 Perbandingan Rasio Data Latih dan Data Uji NBC	69
Tabel 4.18 <i>Script/Source Code</i> Pembagian Data <i>Train</i> Dan Data <i>Test</i> Rasio	70
Tabel 4.19 Pembagian Data Sesuai Rasio Dari 3890 Tweet.....	71
Tabel 4.20 Komentar Data <i>Training</i>	71
Tabel 4.21 Pembobotan TF (<i>Term Frequency</i>).....	72
Tabel 4.22 Dokumen Data Latih	79
Tabel 4.23 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 60:40 <i>E-Wallet</i> Dana	82
Tabel 4.24 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 70:30 <i>E-Wallet</i> Dana	83
Tabel 4.25 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 80:20 <i>E-Wallet</i> Dana	84
Tabel 4.26 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 90:10 <i>E-Wallet</i> Dana	85

Tabel 4.27 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 60:40 <i>E-Wallet</i> OVO.....	86
Tabel 4.28 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 70:30 <i>E-Wallet</i> OVO.....	87
Tabel 4.29 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 80:20 <i>E-Wallet</i> OVO.....	87
Tabel 4.30 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 90:10 <i>E-Wallet</i> OVO.....	88
Tabel 4.31 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 60:40 <i>E-Wallet</i> LinkAja .	89
Tabel 4.32 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 70:30 <i>E-Wallet</i> LinkAja .	90
Tabel 4.33 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 80:20 <i>E-Wallet</i> LinkAja .	91
Tabel 4.34 Hasil <i>Confusion Matriks</i> Klasifikasi Rasio 90:10 <i>E-Wallet</i> LinkAja .	91
Tabel 4.35 Hasil Evaluasi 3 Aplikasi <i>E-wallet</i> Klasifikasi 4 Rasio.....	92



DAFTAR LAMPIRAN

Lampiran 1 : Hasil Plagiarism Checker Originality Report.....	cii
Lampiran 2 : Biodata Mahasiswa.....	ciii
Lampiran 3 : Kartu Bimbingan Skripsi	civ
Lampiran 4 : Bimbingan Dosen Pembimbing.....	cv
Lampiran 5 : Surat Rekomendasi dari Pembimbing	cvi

