



**ANALISIS SENTIMEN TERHADAP ULASAN APLIKASI FLIP.ID
MENGUNAKAN KLASIFIKASI NAÏVE BAYES**

TUGAS AKHIR

Evandro Malsi
41518010095

UNIVERSITAS
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022



**ANALISIS SENTIMEN TERHADAP ULASAN APLIKASI FLIP.ID
MENGUNAKAN KLASIFIKASI NAÏVE BAYES**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

Evandro Malsi
41518010095

UNIVERSITAS
PROGRAM STUDI TEKNIK INFORMATIKA
MERCU BUANA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022

LEMBAR PERNYATAAN ORISINALITAS

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41518010095

Nama : Evandro Malsi

Judul Tugas Akhir : Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id
Menggunakan Klasifikasi Naive Bayes

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 30 Juni 2022



Evandro Malsi



UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Evandro Malsi
NIM : 41518010095
Judul Tugas Akhir : Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id Menggunakan Klasifikasi Naïve Bayes

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 30 Juni 2022



Evandro Malsi

UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN LUARAN TUGAS AKHIR

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Evandro Malsi
NIM : 41518010095
Judul Tugas Akhir : Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id Menggunakan Klasifikasi Naïve Bayes

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis	Status
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi	Diajukan ✓
		Jurnal Nasional Terakreditasi	✓
		Jurnal International Tidak Bereputasi	Diterima
		Jurnal International Bereputasi	
Disubmit/dipublikasikan di :	Nama Jurnal	: Jurnal Informatika dan Teknologi Informasi (Telematika)	
	ISSN	: 1829-667X	
	Link Jurnal	: http://jurnal.upnyk.ac.id/	
	Link File Jurnal Jika Sudah di Publish		

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Mengetahui
Dosen Pembimbing TA

 UNIVERSITAS
MERCU BUANA

Dwiki Jatikusumo S.Kom.,M.Kom

Jakarta, 30 Juni 2022



Evandro Malsi

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010095
Nama : Evandro Malsi
Judul Tugas Akhir : Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id
Menggunakan Klasifikasi Naïve Bayes

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022



(Eliyani, Dr. Ir.)
Dosen Penguji 1
UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010095
Nama : Evandro Malsi
Judul Tugas Akhir : Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id
Menggunakan Klasifikasi Naïve Bayes

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022



(Wawan Anwar, S.Kom, MT)
Dosen Penguji 2
UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010095
Nama : Evandro Malsi
Judul Tugas Akhir : Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id
Menggunakan Klasifikasi Naïve Bayes

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022



UNIVERSITAS
MERCU BUANA
(Dwi Anindyani Rocmah, ST, MTI)
Dosen Penguji 3

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

NIM : 41518010095
Nama : Evandro Malsi
Judul Tugas Akhir : Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id
Menggunakan Klasifikasi Naïve Bayes

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022

Menyetujui,



(Dwiki Jatikusumo S.Kom.,M.Kom)
Dosen Pembimbing

UNIVERSITAS
MERCU BUANA

Mengetahui,



(Wawan Gunawan, S.Kom, MT)
Koord. Tugas Akhir Teknik Informatika



(Ir. Emil R. Kaburuan, Ph.D., IPM.)
Ka. Prodi Teknik Informatika

ABSTRAK

Nama : Evandro Malsi
NIM : 41518010095
Pembimbing TA : Dwiki Jatikusumo S.Kom.,M.Kom
Judul : Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id Menggunakan Klasifikasi Naïve Bayes

Bank-bank terbesar di Indonesia telah menetapkan biaya administrasi untuk melakukan transfer dana antar bank yaitu sebesar Rp6.500,00. Biaya tersebut menyebabkan masyarakat terkadang merasa keberatan jika harus membayar biaya administrasi ketika transfer ke bank yang berbeda. Namun dengan adanya kehadiran aplikasi *fintech* (*financial technology*) seperti OVO, DANA, OY! Indonesia dan lain sebagainya, masyarakat mulai beralih menggunakan beberapa aplikasi *fintech* tersebut yang membebaskan biaya administrasi. Salah satu aplikasi *fintech* yang banyak diminati dengan total unduhan lebih dari 10 juta dan mendapatkan rating 4.4 di *google play store* adalah aplikasi Flip.id. Seiring berjalannya waktu pengguna sering membahas dan memberikan perspektif maupun opini mereka melalui komentarnya pada ulasan aplikasi Flip.id yang telah disediakan oleh *google play store*. Opini yang diberikan oleh pengguna melalui ulasan di aplikasi Flip.id juga berfluktuasi. Selain itu, dalam beberapa kasus, bintang dan ulasan yang diberikan unik dan bisa juga berlebihan. Salah satu cara untuk mengevaluasi opini dan ulasan pengguna pada aplikasi ini adalah dengan menggunakan analisis sentimen. Sumber data penelitian ini berasal dari ulasan pengguna terhadap aplikasi Flip.id di *google play store* untuk dilakukan klasifikasi sentimen menggunakan metode klasifikasi naïve bayes. Berdasarkan hasil pengujian yang telah dilakukan, dapat disimpulkan bahwa metode naïve bayes berhasil dalam melakukan klasifikasi analisis sentimen terhadap ulasan aplikasi Flip.id dan mendapatkan hasil akurasi sebesar 79%.

Kata kunci:

Analisis Sentimen, Fintech, Flip.id, Klasifikasi, Naïve Bayes

ABSTRACT

Name : Evandro Malsi
Student Number : 41518010095
Counsellor : Dwiki Jatikusumo S.Kom.,M.Kom
Title : Sentiment Analysis Against Flip.id Application Reviews Using Naïve Bayes Classification

The largest banks in Indonesia have set an administrative fee for interbank fund transfers of IDR 6,500.00. These fees cause people to sometimes feel objections if they must pay administrative fees when transferring to different banks. However, with the presence of fintech (financial technology) applications such as OVO, DANA, OY! Indonesia and so on, people began to switch to using some of these fintech applications which waived administrative fees. One of the fintech applications that is in great demand with a total download of more than 10 million and getting a rating of 4.4 on the google play store is the Flip.id application. Over time users often discuss and provide their perspectives and opinions through their comments on the reviews of Flip.id applications that have been provided by the google play store. The opinions given by users through reviews on the app Flip.id also fluctuate. In addition, in some cases, the stars and reviews given are unique and can also be overwhelming. One way to evaluate user opinions and reviews on this app is to use sentiment analysis. The source of this research data comes from user reviews of Flip.id applications on the google play store to be classified for sentiment using the naïve bayes classification method. Based on the results of tests that have been carried out, it can be concluded that the naïve bayes method was successful in carrying out a sentiment analysis classification of Flip.id application reviews and obtained an accuracy result of 79%.

Key words:

Sentiment Analysis, Fintech, Flip.id, Classification, Naïve Bayes

KATA PENGANTAR

Segala puji dan syukur kehadirat ALLAH SWT atas berkah, rahmat, karunia serta hidayah-Nya yang senantiasa dilimpahkan kepada penulis, sehingga dapat menyelesaikan tugas akhir ini yang berjudul “*Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id Menggunakan Klasifikasi Naïve Bayes*” dengan lancar. Tugas akhir ini disusun untuk memenuhi syarat memperoleh gelar sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Penulis menyadari tugas akhir ini tidak akan terwujud tanpa adanya bantuan, bimbingan dan dorongan dari berbagai pihak baik secara moral maupun spiritual. Oleh karena itu, dengan segala kerendahan hati penulis mengucapkan terima kasih kepada:

1. Kedua orang tua yang selalu memberikan doa, dukungan, kasih sayang sehingga saya dapat menyelesaikan tugas akhir ini dengan lancar
2. Bapak Dwiki Jatikusumo S.Kom.,M.Kom selaku dosen pembimbing tugas akhir yang telah membantu dan memberikan masukan serta arahan kepada saya dalam menyelesaikan tugas akhir ini
3. Bapak Emil R. Kaburuan Ph.D selaku Kepala Program Studi Informatika yang telah memberikan arahan dan bimbingannya selama perkuliahan
4. Bapak Wawan Gunawan S.Kom.,MT selaku Koordinator Tugas Akhir Prodi Informatika yang telah memberikan arahan selama periode tugas akhir
5. Bapak Drs. Achmad Kodar, MT selaku dosen pembimbing akademik saya yang telah membimbing selama perkuliahan sampai pada tugas akhir
6. Teman-teman Informatika 2018, sahabat dan kerabat yang selalu memberikan dukungan dan motivasi dalam menyelesaikan tugas akhir ini.

Akhir kata, penulis berharap semoga tugas akhir ini dapat bermanfaat dan berguna bagi para pembaca serta menambah wawasan pengetahuan semua pihak.

Jakarta, 26 Juli 2022

Penulis

DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN JUDUL	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR ..	iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR	iv
LEMBAR PERSETUJUAN PENGUJI	v
LEMBAR PENGESAHAN	viii
ABSTRAK	ix
ABSTRACT.....	x
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xii
NASKAH JURNAL	1
KERTAS KERJA.....	12
BAB 1. LITERATUR REVIEW.....	13
BAB 2. ANALISIS DAN PERANCANGAN.....	18
BAB 3. SOURCE CODE.....	21
BAB 4. DATASET.....	29
BAB 5. TAHAPAN EKSPERIMEN.....	30
BAB 6. HASIL SEMUA EKSPERIMEN.....	35
DAFTAR PUSTAKA	37
LAMPIRAN DOKUMEN HAKI.....	40
LAMPIRAN KORESPONDENSI	42

NASKAH JURNAL

Telematika: Jurnal Informatika dan Teknologi Informasi
Vol. 18, No. 1, Juli 2022, pp.1-11

ISSN: 1829-667X / E-ISSN: 2460-9021
DOI: 10.31515/telematika.v18i3.xxxx

Sentiment Analysis Against Flip.id Application Reviews Using Naïve Bayes Classification

Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id Menggunakan Klasifikasi Naïve Bayes

Evandro Malsi¹, Dwiki Jatikusumo²

^{1, 2} Teknik Informatika, Universitas Mercu Buana Jakarta, Indonesia

¹41518010095@student.mercubuana.ac.id, ²dwiki.jatikusumo@mercubuana.ac.id

*corresponding author

Informasi Artikel

Received: 2022

Revised: 2022

Accepted: 2022

Published: 2022

Abstract

Purpose: Conducting sentiment analysis of application reviews Flip.id into 3 sentiment classifications, namely positive, neutral and negative and measuring the performance results and accuracy level of the naïve bayes method

Design/methodology/approach: Using the classification method of naïve bayes

Findings/result: Based on the results of tests that have been carried out, obtaining accuracy results of 79%

Originality/value/state of the art: The difference between this study and previous research is in the stage of dataset labeling, pre-processing and classification methods. This labeling is done because the supervised learning method requires sample data. Labeling is done by labeling negatively on 1 and 2 star data, neutral labels on data with 3 stars, and positive labels on data with 4 and 5 stars. After that, it is checked the match between the label and the review text manually, because there are reviews that are not relevant to the given star, and quite a lot of data is unique and ambiguous. Before modeling there were many slang words in user reviews, therefore slang words were changed to standard words using the alay dictionary dataset. After that the dataset is classified using the naïve bayes method

Keywords: Sentiment Analysis; Application; Flip.id; Classification; Naïve Bayes

Kata kunci: Analisis Sentimen; Aplikasi; Flip.id; Klasifikasi; Naïve Bayes

Abstrak

Tujuan: Melakukan analisis sentimen terhadap ulasan aplikasi Flip.id menjadi 3 klasifikasi sentimen yaitu positif, netral dan negatif serta mengukur hasil kinerja dan tingkat akurasi metode *naïve bayes*

Perancangan/metode/pendekatan: Menggunakan metode klasifikasi *naïve bayes*

Hasil: Berdasarkan hasil pengujian yang telah dilakukan, mendapatkan hasil akurasi sebesar 79%

Keaslian/ state of the art: Perbedaan penelitian ini dengan penelitian terdahulu sebelumnya adalah pada tahap labelisasi dataset, *preprocessing* dan metode klasifikasi. Labelisasi ini dilakukan karena metode *supervised learning* membutuhkan contoh data. Labelisasi dilakukan dengan memberi label negatif pada data berbintang 1 dan 2, label netral pada data dengan bintang 3, dan label positif pada data dengan bintang 4 dan 5. Setelah itu dilakukan pengecekan kecocokan antara label dengan teks ulasan secara manual, karena terdapat ulasan yang tidak relevan dengan bintang yang diberikan, dan cukup banyak data yang bersifat unik dan ambigu. Sebelum dilakukan pemodelan terdapat banyak kata gaul pada ulasan pengguna, maka dari itu kata gaul diubah menjadi kata baku menggunakan dataset kamus alay. Setelah itu dataset diklasifikasi menggunakan metode *naïve bayes*

1. Pendahuluan

Pada era digitalisasi saat ini, telah membawa perubahan signifikan pada perilaku masyarakat yang dahulu bersifat konvensional dan sekarang menjadi serba digital[1]. Selama masa pandemi, transaksi digital atau nontunai di Indonesia meningkat setiap bulannya. Berdasarkan data dari Databoks, per bulan Mei 2021 transaksi digital tumbuh sebesar 23,66 triliun rupiah dibandingkan dengan bulan Mei 2020 yang hanya sebesar 15,03 triliun rupiah. Dari banyaknya jumlah transaksi tersebut, pengguna transaksi digital melakukan transaksi menggunakan kartu ATM, *mobile banking* SMS dan *internet banking*[2]. Tidak hanya melakukan transfer ke rekening di bank yang sama, transfer juga dilakukan ke rekening di bank berbeda dengan biaya administrasi yang nominalnya berbeda sesuai kebijakan masing-masing bank.

Bank-bank terbesar di Indonesia telah menetapkan biaya administrasi untuk melakukan transfer dana antar bank yaitu sebesar Rp6.500,00[3]. Biaya tersebut menyebabkan masyarakat terkadang merasa keberatan jika harus membayar biaya administrasi ketika transfer ke bank yang berbeda. Namun dengan adanya kehadiran aplikasi *fintech (financial technology)* seperti OVO, DANA, OY! Indonesia dan lain sebagainya, masyarakat mulai beralih menggunakan beberapa aplikasi *fintech* tersebut yang membebaskan biaya administrasi[4]. Menurut laporan dari *Standard*

& Poor's yang bertema 'Southeast Asia E-Money Market Report', masyarakat Indonesia lebih memilih pembayaran melalui fintech ketimbang dari bank.

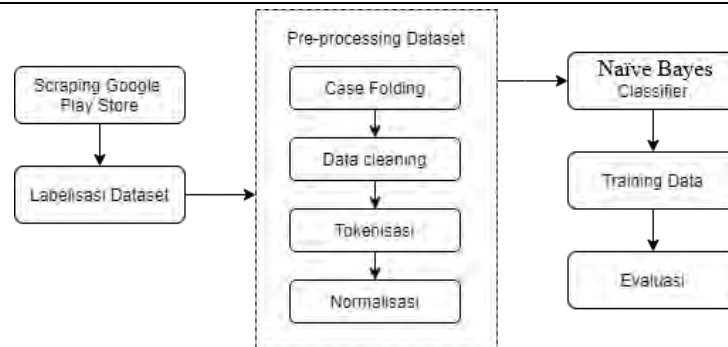
Fintech merupakan inovasi yang menyatukan jasa keuangan dengan teknologi guna mempercepat dan memudahkan dalam transaksi digital atau *digital payment*[5]. Salah satu aplikasi *fintech* yang banyak diminati dengan total unduhan lebih dari 10 juta dan mendapatkan rating 4.4 di *google play store* adalah aplikasi Flip.id. Flip.id (PT Fliptech Lentera Inspirasi Pertiwi) didirikan pada tahun 2015, dan telah tersedia di berbagai macam platform seperti *android* dan *ios*[6]. Flip.id adalah aplikasi pionir penyedia layanan transfer antar bank tanpa biaya administrasi dan telah berlisensi serta diawasi Bank Indonesia. Flip.id bekerja sebagai jembatan transaksi antar bank, pengirim dana hanya perlu melakukan transfer terlebih dahulu ke rekening Flip.id yang sama dengan bank yang dimiliki. Kemudian pihak Flip.id akan meneruskan uang ke rekening tujuan[7].

Seiring berjalannya waktu pengguna sering membahas dan memberikan perspektif maupun opini mereka melalui komentarnya pada ulasan aplikasi Flip.id yang telah disediakan oleh *google play store*. Opini yang diberikan oleh pengguna melalui ulasan diaplikasi Flip.id juga berfluktuasi. Ulasan dalam aplikasi Flip.id di *google play store* juga membawa berbagai macam dampak pada aplikasi Flip.id karena memuat ulasan positif, negatif dan netral. Namun, jumlah ulasan pengguna sangat banyak, membuat proses penanganan dan pengolahan informasi menjadi sangat lama. Selain itu, dalam beberapa kasus, bintang dan ulasan yang diberikan unik dan bisa juga berlebihan. Oleh karena itu peneliti mencari cara untuk mengevaluasi opini dan ulasan tanpa bergantung pada bintang yang diberikan oleh pengguna. Salah satu cara untuk mengevaluasi opini dan ulasan pengguna pada aplikasi ini adalah dengan menggunakan analisis sentimen[8].

Sumber data penelitian ini berasal dari ulasan pengguna terhadap aplikasi Flip.id di *google play store* untuk dilakukan klasifikasi sentimen menggunakan metode klasifikasi *naive bayes*. Pemilihan metode ini, didasari oleh penelitian terdahulu terkait sebelumnya[9], *naive bayes* memiliki hasil akurasi paling akurat sebesar 89% diikuti oleh *support vector machine* 88% dan *random forest* 85%. Penelitian lain[10] hasil *naive bayes* memperoleh nilai akurasi rata rata sebesar 89%. Penelitian lainnya[11] algoritma *naive bayes* yang mendapatkan hasil sebesar 80,67% sedangkan algoritma C4.5 83,33%. Dari penelitian terdahulu yang telah dilakukan sebelumnya, dapat disimpulkan bahwa analisis sentimen menggunakan metode klasifikasi *naive bayes* mendapatkan akurasi yang tinggi. Maka dari itu peneliti memutuskan untuk memilih metode ini untuk diimplementasikan ke analisis sentimen terhadap ulasan aplikasi Flip.id menggunakan klasifikasi *naive bayes*.

2. Metode

Penelitian diawali dengan melakukan pengumpulan data atau *scraping* data dari situs web *google play store* berupa ulasan pengguna aplikasi Flip.id. Kemudian hasil dari *scraping* dikumpulkan dan dibuat menjadi sebuah dataset. Setelah itu dataset dilabelisasi dengan label *positive*, *neutral*, dan *negative*. Lalu dataset masuk kedalam tahap *preprocessing*. Beberapa tahapan pada *preprocessing*, yaitu *case folding*, *data cleaning*, *tokenization*, *normalization*, *stemming*, dan *stopwords*. Setelah melalui tahapan proses-proses tersebut, dataset kemudian dilatih menggunakan *naive bayes* untuk diklasifikasi menjadi tiga kategori yaitu positif, negatif, dan netral. Data yang telah terklasifikasi kemudian dievaluasi untuk melihat keakuratan prediksi. Arsitektur dari metode penelitian ini terdapat pada gambar di bawah ini.



Gambar 1. Metodologi Penelitian

2.1. Pengumpulan Data

Pada penelitian ini menggunakan data yang tersedia di *google play store* berupa ulasan pengguna pada aplikasi Flip buatan PT Fliptech Lentera Inspirasi Pertiwi. *Scraping* dilakukan dengan menggunakan *library google-play-scraper*. Untuk mendapatkan dataset yang seimbang, *scraping* dilakukan per kategori bintang. Kemudian dilakukan penyeleksian fitur pada dataset.

2.2. Labelisasi Data

Pada analisis sentimen dengan metode *supervised learning*, diperlukan dataset yang sudah memiliki label. Labelisasi ini dilakukan karena metode *supervised learning* membutuhkan contoh data. Labelisasi dilakukan dengan memberi label negatif pada data berbintang 1 dan 2, label netral pada data dengan bintang 3, dan label positif pada data dengan bintang 4 dan 5.

2.3. Preprocessing Data

Tahap *preprocessing* dilakukan untuk mengolah data yang tidak terstruktur untuk membuatnya lebih mudah diproses klasifikasi yaitu dengan melakukan beberapa tahapan sebagai berikut[12]:

2.3.1 Case Folding

Case folding adalah langkah yang mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf (a-z) yang diterima. Karakter kecuali huruf dihilangkan dan dianggap sebagai pembatas.

2.3.2 Data Cleaning

Tahap ini digunakan untuk menghilangkan inkonsistensi data dan menghilangkan data yang tidak relevan angka, beberapa gambar, url, *username* (@*usernames*), *hashtags* (#), spasi berlebih, aksentuasi, *emoticon*, dan pengulangan karakter dalam kalimat[13].

2.3.3 Tokenizing

Tokenizing atau *parsing* adalah fase pemotongan string info tergantung pada setiap kata yang membuatnya. Juga, spasi digunakan untuk independen kata-kata.

2.3.4 Normalization

Normalization merupakan proses mengembalikan kata yang tidak baku ke dalam bahasa baku dalam kamus bahasa Indonesia atau sesuai ejaan. Proses ini dilakukan karena cukup banyak kalimat yang menggunakan kata tidak baku atau gaul seperti *apk*, *tf*, *jt*, *ga*, *ngga*, dan *gak*.

2.3.1 Stemming

Stemming adalah proses yang digunakan untuk mengamati ekspresi akar suatu kata dengan menghilangkan semua imbuhan besar yaitu awalan, sisipan, tambahan, dan campuran awalan dan akhiran pada kata[14].

2.3.2 Stopwords

Stopwords adalah proses penghilangan kata-kata yang sering muncul yang memiliki nilai kecil dan tidak banyak membantu saat memilih data yang dibutuhkan pengguna.

2.1. Klasifikasi Naive Bayes

Metode yang akan digunakan dalam investigasi opini eksplorasi ini adalah *naive bayes*. *Naive bayes* adalah teknik pengelompokan yang memanfaatkan kemungkinan dalam probabilitas[15]. *Naive bayes* sangat sederhana dalam eksekusi dan efektivitas terutama untuk berbagai aplikasi praktis seperti spam filtering atau klasifikasi artikel berita. Teorema bayes dirumuskan sebagai berikut:

$$P(D|E) = \frac{P(E|D) \times P(D)}{P(E)}$$

Keterangan:

- $P(D|E)$: Probabilitas bersyarat (*conditional probability*) dari Hipotesis D terjadi jika diberikan bukti (*evidence*) E terjadi.
 $P(E|D)$: Probabilitas bukti E terjadi akan memengaruhi hipotesis D.
 $P(D)$: Probabilitas awal hipotesis/kejadian D terjadi tanpa memandang bukti apapun.
 $P(E)$: Probabilitas awal bukti E tanpa memandang hipotesis/bukti yang lain.

2.2. Confusion Matrix

Confusion Matrix adalah metode untuk mengetahui tingkat ketepatan, dengan menghitung jumlah kuantitas prediksi benar dan salah dari nilai sesungguhnya atau aktual dan nilai prediksi.

		Predicted Class		
		Positive	Neutral	Negative
True Class	Positive	True Positive (TP)	False Neutral (FNt)	False Negative (FN)
	Neutral	False Positive (FP)	True Neutral (TNt)	False Negative (FN)
	Negative	False Positive (FP)	False Neutral (FNt)	True Negative (TN)

Gambar 2. Confusion Matrix

2.6. Evaluasi

Setelah mendapatkan nilai untuk *confusion matrix*, nilai untuk *accuracy*, *precision*, *recall*, dan *F1-score* juga dapat diperoleh. Rumus perhitungan nilai untuk mendapatkan *accuracy*, *precision*, *recall*, dan *F1-score* ditunjukkan pada persamaan berikut.

$$\text{Accuracy} = \frac{\text{TP} + \text{TnT} + \text{TN}}{\text{TP} + \text{FP} + \text{TnT} + \text{FNt} + \text{TN} + \text{FN}}$$

$$\text{Precision}_{\text{positive}} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall}_{\text{positive}} = \frac{\text{TP}}{\text{TP} + \text{FNt} + \text{FN}}$$

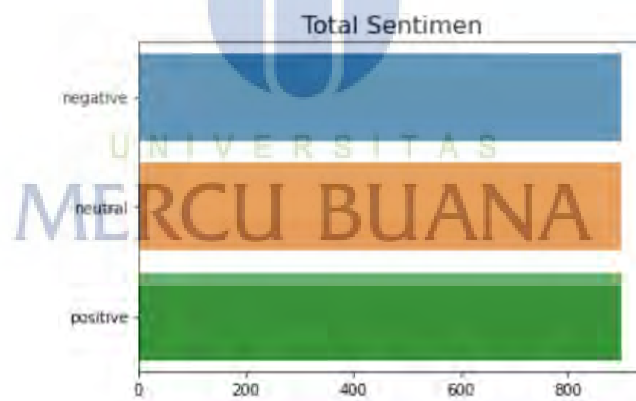
$$\text{F1score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Gambar 3. Perhitungan Persamaan *Classification Report*

2. Hasil dan Pembahasan

2.1. Pengumpulan data

Penyeleksian fitur dilakukan dengan membuang kolom *id*, *username*, tanggal karena dianggap tidak digunakan dalam proses analisis sentimen. Seluruh dataset selanjutnya digabung dan disimpan dalam format tab separated value (.tsv). Dari penyeleksian fitur didapatkan sebanyak 2700 data berupa 900 positif, 900 netral dan 900 negatif. Berikut total sentimen yang disajikan pada gambar di bawah ini



Gambar 4. Total Sentimen

2.2. Labelisasi Data

Pengecekan kecocokan antara label dengan teks ulasan dilakukan secara manual, karena terdapat ulasan yang tidak relevan dengan bintang yang diberikan, dan cukup banyak data yang bersifat unik dan ambigu. Berikut adalah hasil labelisasi data yang disajikan gambar di bawah ini

	komentar	sentimen
0	Busuk lemot parah	negative
1	LEMOT	negative
2	Aplikasi ribet	negative
3	Burukkkkkkk	negative
4	Error	negative
...
2695	bagus aplikasi mudah digunakan cm kekurangan ...	positive
2696	Untuk pertama kali pakai lumayan menghemat bia...	positive
2697	Hehe lumayan ngirit 6500 🤔👍👍	positive
2698	Bagus banget	positive
2699	Mantul	positive

Gambar 5. Hasil Labelisasi Data

2.1. Preprocessing

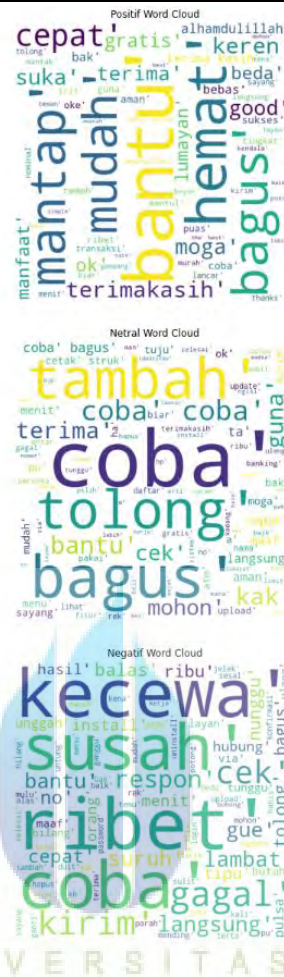
Tahap *preprocessing* teks dilakukan menggunakan *library* Sastrawi. *Stemming* memanfaatkan kelas *StemmerFactory()* menghasilkan string yang terdiri dari kata dasar dengan karakter huruf kecil. Pada proses ini tanda baca dihilangkan, *string* ini kemudian, dipisahkan dengan kelas *StopWordRemoverFactory()* yang akan dibuang dengan *stopwords* atau kata sambung yang tidak digunakan. Proses ini menghasilkan *string* kata dasar tanpa menggunakan *stopwords*. Selanjutnya jenis string diubah menjadi token dengan memanfaatkan kelas *TokenizerFactory()*. Kemudian dilakukan penghapusan karakter selain huruf sehingga hanya menghasilkan token yang berisi kata. Sebelum dilakukan pemodelan terdapat banyak kata gaul pada ulasan pengguna, maka dari itu kata gaul diubah menjadi kata baku menggunakan dataset kamus alay. Setiap hasil dari *preprocessing* teks ini disimpan di dalam dataset. Berikut hasil dari tahapan *preprocessing* yang disajikan di bawah ini

	komentar	sentimen	text_clean	text_preprocessed
0	Busuk lemot parah	negative	busuk lemot parah	['busuk', 'lambat', 'parah']
1	LEMOT	negative	lemot	['lambat']
2	Aplikasi ribet	negative	aplikasi ribet	['ribet']
3	Burukkkkkkk	negative	buruk	['buruk']
4	Error	negative	error	['error']
5	Zonk	negative	zonk	['zonk']
6	Leg ajujubilah	negative	leg ajujubilah	['leg', 'ajujubilah']
7	Transfer saya 6 jt belum diproses tlp cs susah...	negative	transfer saya jt belum diproses tlp cs susah g...	['juta', 'proses', 'telepon', 'susah']
8	Susah amat buat di hubungi uang saya gak di re...	negative	susah amat buat di hubungi uang saya gak di re...	['susah', 'hubung']
9	Saya sudah download ini apktp saya ga ngerti c...	negative	saya sudah download ini apktp saya ga ngerti c...	['apktp', 'erti', 'bgkata', 'orang', 'sales', ...]

Gambar 6. Hasil *Preprocessing*

2.2. Visualisasi Data

Data yang sudah dilakukan tahap *preprocessing*, kemudian dibuat visualisasi menggunakan *library* WordCloud. Berikut adalah visualisasi data dari masing-masing kategori sentimen yaitu positif, negatif dan netral yang disajikan di bawah ini



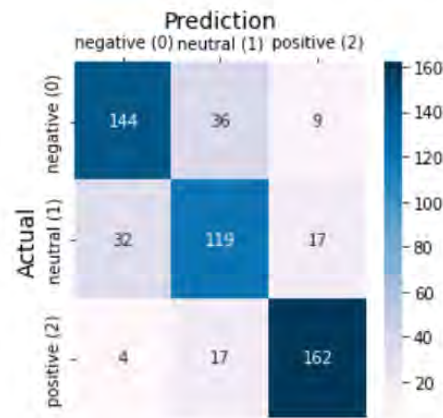
Gambar 7. Visualisasi Data

3.5. Klasifikasi Naïve Bayes

Sebelum melakukan klasifikasi terdapat dua proses penting yaitu *splitting* data dengan membagi data menjadi dua bagian secara acak, sebagian sebagai data *learning* (training) dan sebagian lainnya sebagai data testing. Pengujian ini menggunakan data training sebanyak 80% dari total data yaitu 2160 data, dan untuk data testing sebanyak 20% dari total data yaitu 540 data.

3.6. Confusion Matrix

Selanjutnya model diuji dengan menggunakan data *testing*. Hasil uji coba model dengan data uji akan ditampilkan dengan *confusion matrix* untuk mengukur seberapa banyak model yang berhasil memprediksi dengan benar semua sentimen pada data testing. Berikut di bawah ini adalah hasil dari *confusion matrix*.



Gambar 8. Hasil Confusion Matrix

Dari hasil *confusion matrix* diatas dapat dilihat bahwa model memiliki kesulitan dalam memprediksi data dengan sentimen netral, namun untuk sentimen positif dan negatif sudah cukup baik.

3.7. Evaluasi

Untuk menganalisis dan mengukur keakuratan hasil yang telah diperoleh dari model klasifikasi *naive bayes* dengan menggunakan *classification report*. Berikut di bawah ini adalah hasil dari *classification report*

Classification Report				
	precision	recall	f1-score	support
0	0.80	0.76	0.78	189
1	0.69	0.71	0.70	168
2	0.86	0.89	0.87	183
accuracy			0.79	540
macro avg	0.78	0.79	0.78	540
weighted avg	0.79	0.79	0.79	540

Gambar 9. Hasil Classification Report

Dari hasil *classification report* di atas model klasifikasi *naive bayes* mendapat hasil cukup tinggi pada sentimen negatif, dan hasil relatif lebih rendah pada sentimen positif dan netral.

4. Kesimpulan dan Saran

Berdasarkan hasil pengujian yang telah dilakukan, dapat disimpulkan bahwa metode *naive bayes* berhasil dalam melakukan klasifikasi analisis sentimen terhadap ulasan aplikasi Flip.id dan mendapatkan hasil akurasi sebesar 79%. Selanjutnya dalam mengklasifikasikan sentimen netral, metode *naive bayes* mengalami kesulitan dikarenakan jumlah dataset yang termasuk sedikit. Dan juga terdapat inkonsistensi pada dataset, karena ulasan yang tidak sesuai dengan bintang yang diberikan dan bersifat unik dan ambigu. Untuk penelitian selanjutnya dapat menggunakan

algoritma klasifikasi lainnya seperti *Support Vector Machine*, *Artificial Neural Network*, *Levenshtein Distance*, *K-Nearest Neighbor* dan *Maximum Entropy* untuk mendapatkan hasil akurasi yang lebih baik. Selain itu diharapkan untuk aplikasi Flip.id terus meningkatkan pelayanan yang sudah diberikan sebelumnya oleh Flip.id sehingga kualitas dari aplikasi Flip.id tetap terjaga dan semakin banyak pengguna yang menggunakannya.

Daftar Pustaka

- [1] D. Natalia and M. Shihab, "Public Relations Strategies to Built Financial Technology (Fintech) Awareness The Ikoinkworksr Way," Proceedings of the International Conference on Media and Communication Studies (ICOMACS 2018), 2018, doi: 10.2991/icomacs-18.2018.62.
- [2] S. W. Narayan and S. Sahminan, "HAS FINTECH INFLUENCED INDONESIA'S EXCHANGE RATE AND INFLATION?," Buletin Ekonomi Moneter dan Perbankan, vol. 21, no. 2, pp. 189–202, Oct. 2018, doi: 10.21098/bemp.v21i2.966.
- [3] F. Ardiansyah, M. E. Siregar, D. B. Hakim, and H. Siregar, "Determinant Factors Affecting Efficiency Intermediation and Production Approach (Case Study: Bank Buku II Go Public in Indonesia)," Proceedings of the Business Innovation and Engineering Conference 2020 (BIEC 2020), 2021, doi: 10.2991/aebmr.k.210727.025.
- [4] T. A. Safitri, "The Development of Fintech in Indonesia," Proceedings of the 1st Borobudur International Symposium on Humanities, Economics and Social Sciences (BIS-HESS 2019), 2020, doi: 10.2991/assehr.k.200529.139.
- [5] F. Amalia, "THE FINTECH BOOK: THE FINANCIAL TECHNOLOGY HANDBOOK FOR INVESTORS, ENTREPRENEURS AND VISIONARIES," Journal of Indonesian Economy and Business, vol. 31, no. 3, p. 345, Jan. 2016, doi: 10.22146/jieb.23554.
- [6] E. Wintarsih, R. Cahyadi, Z. Irfandi, M. Nazri, M. R. Shihab, and N. F. A. Budi, "Reducing Interbank Money Transfer Cost for Personal and Small Businesses using Rational Consumer Behavior Analysis: Case Study of Flip.id," 2019 International Conference on Information Management and Technology (ICIMTech), Aug. 2019, doi: 10.1109/icimtech.2019.8843782.
- [7] A. Purnamasari, F. Pranastara, I. S. Rosa, D. Riana, and S. Hadianti, "Analysis of the Effect of DTPB on User Intention to Keep Using the Flip Application," SISTEMASI, vol. 11, no. 1, p. 108, Jan. 2022, doi: 10.32520/stmsi.v11i1.1633.
- [8] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, "Sentiment Analysis of Social Media Twitter with Case of Anti-LGBT Campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest Algorithm," Procedia Computer Science, vol. 161, pp. 765–772, 2019, doi: 10.1016/j.procs.2019.11.181
- [9] A. Nayak and D. Natarajan, "Comparative study of Naïve Bayes , Support Vector Machine and Random Forest Classifiers in Sentiment Analysis of Twitter feeds," Int. J. Adv. Stud. Comput. Sci. Eng., vol. 5, no. 1, pp. 14–17, 2016.
- [10] M. Afzaal, M. Usman and A. Fong, "Tourism Mobile App With Aspect-Based Sentiment Classification Framework for Tourist Reviews," in IEEE Transactions on Consumer Electronics, vol. 65, no. 2, pp. 233-242, May 2019, doi: 10.1109/TCE.2019.2908944.

-
- [11] M. Sadikin and F. Alfiandi, "Comparative Study of Classification Method on Customer Candidate Data to Predict its Potential Risk," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 6, p. 4763, Dec. 2018, doi: 10.11591/ijece.v8i6.pp4763-4771.
- [12] F. Gunawan, M. A. Fauzi, and P. P. Adikara, "Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan Naïve Bayes dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile)," *Systemic: Information System and Informatics Journal*, vol. 3, no. 2, pp. 1–6, Dec. 2017, doi: 10.29080/systemic.v3i2.234.
- [13] A. Budiyanto and S. Dwiasnati, "The Prediction of Best-Selling Product Using Naïve Bayes Algorithm (A Case Study at PT Putradabo Perkasa)," *Ijctjournal.Org*, vol. 5, no. 6, pp. 68–74, 2018.
- [14] F. N. Zuhri and A. Alamsyah, "Menggunakan Naïve Bayes Classifier Di Forum Kaskus Publik Sentiment Analysis Of Smartfren Brand Using Naïve Bayes Classifier On Kaskus Forum," *e- Proceeding Manag.*, vol. 4, no. 1, pp. 242–251, 2017.
- [15] H. D. Wijaya and S. Dwiasnati, "Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat," *Jurnal Informatika*, vol. 7, no. 1, pp. 1–7, Apr. 2020, doi: 10.31311/ji.v7i1.6203.



KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan material kelengkapan artikel jurnal dengan judul “Analisis Sentimen Terhadap Ulasan Aplikasi Flip.id Menggunakan Klasifikasi Naïve Bayes”. Kertas kerja berisi semua material hasil penelitian Tugas Akhir yang tidak dimuat/atau disertakan di artikel jurnal. Di dalam kertas kerja ini disajikan: literature review, analisis dan perancangan, source code, dataset yang digunakan, tahapan eksperimen dan hasil eksperimen secara keseluruhan.

