

**ASPECT-BASED SENTIMENT ANALYSIS PADA ULASAN
APLIKASI BCA MOBILE DENGAN MENGGUNAKAN
METODE TF-IDF DAN ALGORITMA NAÏVE BAYES
CLASSIFIER**

SKRIPSI



Oleh :

RIZQY AHSANA PUTRI

NPM. 19081010058

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"

JAWA TIMUR

2023

**ASPECT-BASED SENTIMENT ANALYSIS PADA ULASAN
APLIKASI BCA MOBILE DENGAN MENGGUNAKAN
METODE TF-IDF DAN ALGORITMA NAÏVE BAYES
CLASSIFIER**

SKRIPSI

Diajukan Untuk Memenuhi Sebagai Persyaratan Dalam Menempuh Gelar Sarjana
Komputer Program Studi Teknik Informatika



Oleh :

RIZOY AHSANA PUTRI

NPM. 19081010058

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER**

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"

JAWA TIMUR

2023

**LEMBAR PENGESAHAN
SKRIPSI**

Judul : ASPECT-BASED SENTIMENT ANALYSIS PADA
ULASAN APLIKASI BCA MOBILE DENGAN
MENGUNAKAN METODE TF-IDF DAN
ALGORITMA NAÏVE BAYES CLASSIFIER

Oleh : RIZQY AHSANA PUTRI

NPM : 19081010058

Telah Diseminarkan Dalam Ujian Skripsi, pada :
Hari Kamis, Tanggal 26 Januari 2023

Mengetahui

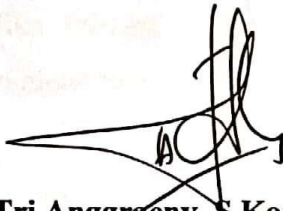
Dosen Pembimbing

1.



Hendra Maulana, S.Kom., M.Kom
NPT. 201198 31 223248

2.



Fetty Tri Anggraeny, S.Kom., M.Kom
NIP. 19820211 2021212 005

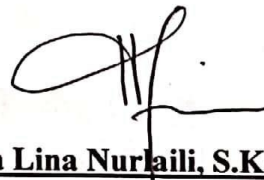
Dosen Penguji

1.



Fetty Tri Anggraeny, S.Kom., M.Kom
NIP. 19820211 2021212 005

2.



Afina Lina Nurlaili, S.Kom., M.Kom
NIP. 202199 31 213198

Menyetujui



**Dekan
Fakultas Ilmu Komputer**

Dr. Novirina Hendrasarie, S.T., MT.

NIP. 19681126 199403 2 001

**Koordinator Program Studi
Informatika**



Fetty Tri Anggraeny, S.Kom., M.Kom

NIP. 19820211 2021212 005

SURAT PERNYATAAN ANTI PLAGIAT

Saya, mahasiswa Program Studi Informatika UPN “Veteran” Jawa Timur, yang bertandatangan di bawah ini:

Nama : Rizqy Ahsana Putri

NPM : 19081010058

Menyatakan bahwa judul skripsi yang saya ajukan dan kerjakan, dengan judul:

**“ASPECT-BASED SENTIMENT ANALYSIS PADA ULASAN APLIKASI
BCA MOBILE DENGAN MENGGUNAKAN METODE TF-IDF DAN
ALGORITMA NAÏVE BAYES CLASSIFIER”**

Bukan merupakan plagiat dari skripsi/tugas akhir/penelitian orang lain dan juga bukan merupakan produk atau *software* yang saya beli dari pihak lain. Saya juga menyatakan bahwa skripsi ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam daftar pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di UPN “Veteran” Jawa Timur maupun di institusi pendidikan lain.

Jika ternyata di kemudian hari pernyataan ini terbukti tidak benar, maka saya siap menerima segala konsekuensinya.

Surabaya, 22 Januari 2023

Hormat saya,



RIZQY AHSANA PUTRI

NPM. 19081010058

ASPECT-BASED SENTIMENT ANALYSIS PADA ULASAN APLIKASI BCA MOBILE DENGAN MENGGUNAKAN METODE TF-IDF DAN ALGORITMA NAÏVE BAYES CLASSIFIER

Nama Mahasiswa : Rizqy Ahsana Putri
NPM : 19081010058
Program Studi : Informatika
Dosen Pembimbing : Hendra Maulana, S.Kom, M.Kom
Fetty Tri Anggraeny, S.Kom, M.Kom

ABSTRAK

BCA mobile merupakan salah satu aplikasi mobile banking yang populer di Indonesia. Aplikasi BCA mobile versi android dapat diunduh gratis melalui Google Playstore. Di google playstore, para calon user bisa melihat ulasan-ulasan untuk mengetahui pengalaman pengguna lain dalam menggunakan aplikasi tersebut. Selain berguna bagi calon user, ulasan ini juga berguna bagi developer aplikasi untuk dapat mengetahui aspek-aspek apa saja yang sudah baik dan aspek-aspek apa yang kurang baik sehingga bisa dijadikan bahan evaluasi untuk update kedepannya.

Penelitian ini akan mengimplementasikan algoritma Naïve Bayes Classifier dan metode TF-IDF untuk melakukan analisis sentimen berbasis aspek pada ulasan aplikasi BCA mobile, yang bertujuan untuk mengukur tingkat performansi algoritma Naïve Bayes dalam mengklasifikasi ulasan ke dalam dua kelas, yaitu kelas sentimen positif-negatif, sekaligus mengetahui kelebihan dan kelemahan aplikasi BCA mobile. Dataset yang digunakan sebanyak 1000 data ulasan yang di-*scrapping* dari Google Playstore.

Dari hasil pengujian didapat tingkat akurasi yang cukup tinggi, yaitu berkisar antara 82% sampai 96%. Selain itu, dari penelitian ini juga dapat diketahui bahwa kelebihan yang dimiliki aplikasi BCA *mobile* terdapat pada aspek *efficiency* dan *satisfaction*. Sedangkan untuk kelemahannya, BCA *mobile* masih banyak mendapat sentimen negatif terutama pada aspek *learnability* dan *error*.

Kata Kunci : Naïve Bayes, TF-IDF, Analisis Sentimen, BCA mobile

KATA PENGANTAR

Alhamdulillah, puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan inayah-Nya sehingga penulis dapat menyelesaikan laporan skripsi yang berjudul “*Aspect-Based Sentiment Analysis* Pada Ulasan Aplikasi BCA Mobile Dengan Menggunakan Metode TF-IDF dan Algoritma Naïve Bayes Classifier”.

Terima kasih saya ucapkan kepada seluruh pihak yang telah memberikan banyak dukungan kepada penulis sehingga penulis bisa menyelesaikan laporan skripsi ini dengan baik tepat waktu.

Penulis menyadari, bahwa laporan skripsi ini masih jauh dari kata sempurna baik segi penyusunan, bahasa, maupun penulisannya. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari semua pembaca agar penulis bisa menjadi lebih baik lagi di masa mendatang.

Semoga laporan skripsi ini bisa menambah wawasan para pembaca dan bisa bermanfaat untuk perkembangan dan peningkatan ilmu pengetahuan.

Surabaya, Januari 2023

Penulis

UCAPAN TERIMA KASIH

Laporan skripsi ini dapat diselesaikan dengan baik atas berkat dan rahmat Allah SWT yang telah memberikan kesehatan dan kelancaran untuk penulis selama proses menyelesaikan skripsi. Proses penyelesaian skripsi ini juga tak lepas dari bantuan dan dukungan berbagai pihak, untuk itu penulis ingin menyampaikan ucapan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Dr. Novirina Hendrasarie, S.T., M.T. selaku Dekan Fakultas Ilmu Komputer di Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny S.Kom., M.Kom. selaku Koordinator Program Studi Informatika Fakultas Ilmu Komputer UPN “Veteran” Jawa Timur sekaligus dosen pembimbing II yang telah banyak memberikan arahan dan bimbingan kepada penulis selama proses penyelesaian skripsi.
4. Bapak Hendra Maulana, S.Kom., M.Kom., selaku dosen pembimbing I dan Bapak Dr. Basuki Rahmat, S.Si, MT. selaku dosen wali penulis yang selalu memberikan arahan selama proses penyelesaian skripsi dan perkuliahan.
5. Staff dan seluruh bapak ibu dosen program studi Informatika UPN “Veteran” Jawa Timur yang telah mengajar dan memberikan ilmu serta pengalaman selama masa perkuliahan.
6. Kedua orang tua penulis, Bapak Kholilur Rochman dan Ibu Ulfatun Nihayah, serta adik-adik yang selalu memberikan motivasi dan doa untuk penulis.
7. Teman-teman seperjuangan Informatika angkatan 2019, kakak dan adik tingkat, serta seluruh pihak yang telah membantu dan mendukung penulis dalam menyelesaikan skripsi dan perkuliahan ini.

Penulis hanya bisa berharap, semoga Allah SWT selalu memberikan perlindungan dan balasan yang lebih di kemudian hari. Amin.

DAFTAR ISI

| | |
|--|-----|
| LEMBAR PENGESAHAN SKRIPSI | i |
| SURAT PERNYATAAN ANTI PLAGIAT | ii |
| ABSTRAK | iii |
| KATA PENGANTAR | iv |
| UCAPAN TERIMA KASIH..... | v |
| DAFTAR ISI..... | vi |
| DAFTAR GAMBAR | ix |
| DAFTAR TABEL..... | xi |
| DAFTAR KODE PROGRAM..... | xii |
| BAB I PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 4 |
| 1.3 Tujuan Penelitian..... | 5 |
| 1.4 Manfaat Penelitian..... | 5 |
| 1.5 Batasan Masalah..... | 5 |
| BAB II TINJAUAN PUSTAKA..... | 7 |
| 2.1 Penelitian Terdahulu..... | 7 |
| 2.2 <i>BCA Mobile</i> | 9 |
| 2.3 Analisis Sentimen..... | 10 |
| 2.4 Aspect-Based Sentiment Analysis (ABSA) | 10 |
| 2.5 <i>Text Preprocessing</i> | 10 |
| 2.5.1 <i>Case-folding</i> | 11 |
| 2.5.2 <i>Cleaning</i> | 11 |
| 2.5.3 <i>Tokenizing</i> | 12 |

| | |
|---|----|
| 2.5.4 <i>Stopword Removal</i> | 12 |
| 2.5.5 <i>Stemming</i> | 13 |
| 2.6 Term Frequency-Inverse Document Frequency (TF-IDF)..... | 14 |
| 2.7 Naïve Bayes Classifier | 15 |
| 2.8 Confusion Matrix | 16 |
| BAB III METODOLOGI PENELITIAN | 19 |
| 3.1 Tahapan Penelitian | 19 |
| 3.2 Studi Literatur..... | 19 |
| 3.3 Analisis dan Desain | 21 |
| 3.4 Akuisisi dan Labelling Data | 23 |
| 3.5 Preprocessing Data | 27 |
| 3.5.1 <i>Case-folding</i> | 27 |
| 3.5.2 <i>Cleaning</i> | 28 |
| 3.5.3 <i>Tokenizing</i> | 28 |
| 3.5.4 <i>Stopword Removal</i> | 29 |
| 3.5.5 <i>Stemming</i> | 29 |
| 3.6 Word Embedding | 30 |
| 3.7 Klasifikasi Data | 31 |
| 3.8 Skema Pengujian | 37 |
| BAB IV HASIL DAN PEMBAHASAN | 39 |
| 4.1 Akuisisi dan Labelling Data | 39 |
| 4.2 <i>Preprocessing</i> Data | 43 |
| 4.2.1 <i>Case-folding</i> | 45 |
| 4.2.2 <i>Cleaning</i> | 46 |
| 4.2.3 <i>Tokenizing</i> | 47 |
| 4.2.4 <i>Stopword Removal</i> | 48 |

| | |
|---|-----------|
| 4.2.5 Stemming..... | 49 |
| 4.3 Word Embedding | 50 |
| 4.4 Pembagian Data..... | 54 |
| 4.5 Klasifikasi Data | 55 |
| 4.6 Hasil Pengujian..... | 58 |
| 4.6.1 Hasil Pengujian Skenario 1 | 61 |
| 4.6.2 Hasil Pengujian Skenario 2..... | 63 |
| 4.6.3 Hasil Pengujian Skenario 3..... | 65 |
| 4.6.4 Hasil Pengujian Skenario 4..... | 67 |
| 4.6.5 Hasil Pengujian Skenario 5..... | 69 |
| 4.6.6 Hasil Pengujian Skenario 6..... | 71 |
| 4.6.7 Hasil Pengujian Skenario 7..... | 73 |
| 4.6.8 Hasil Pengujian Skenario 8..... | 74 |
| 4.6.9 Hasil Pengujian Skenario 9..... | 76 |
| 4.6.10 Hasil Pengujian Skenario 10..... | 78 |
| 4.6.11 Hasil Pengujian Skenario 11..... | 80 |
| 4.6.12 Hasil Pengujian Skenario 12..... | 82 |
| 4.6.13 Perbandingan Hasil Skenario Pengujian..... | 84 |
| 4.7 Analisis Sentimen Aspek..... | 86 |
| BAB V PENUTUP..... | 89 |
| 5.1 Kesimpulan..... | 89 |
| 5.2 Saran | 89 |
| DAFTAR PUSTAKA | 90 |
| BIODATA PENULIS | 94 |

DAFTAR GAMBAR

| | |
|--|----|
| Gambar 3.1 Tahapan Penelitian | 19 |
| Gambar 3.2 Diagram Alur Sistem..... | 21 |
| Gambar 3.3 Diagram Alur Tahapan Akuisisi dan Labelling Data..... | 23 |
| Gambar 3.4 Scrapping Dataset..... | 24 |
| Gambar 3.5 Labelling Dataset..... | 25 |
| Gambar 3.6 Contoh Labelling Dataset..... | 26 |
| Gambar 3.7 Diagram Alur Tahapan Preprocessing Data..... | 27 |
| Gambar 3.8 Diagram Alur Klasifikasi Data..... | 31 |
| Gambar 4.1 Dataset Hasil Scrapping Data..... | 41 |
| Gambar 4.2 Download Scraped Data..... | 42 |
| Gambar 4.3 Dataset Hasil Pelabelan | 42 |
| Gambar 4.4 Upload File Untuk Preprocessing Data..... | 44 |
| Gambar 4.5 Dataset Hasil Upload File | 44 |
| Gambar 4.6 Dataset Hasil Case-folding..... | 45 |
| Gambar 4.7 Dataset Hasil Cleaning | 47 |
| Gambar 4.8 Dataset Hasil Tokenizing | 48 |
| Gambar 4.9 Dataset Hasil Stopword Removal | 49 |
| Gambar 4.10 Dataset Hasil Stemming..... | 50 |
| Gambar 4.11 Hasil Label Encoding | 51 |
| Gambar 4.12 Hasil Pembobotan TF-IDF..... | 52 |
| Gambar 4.13 Potongan Hasil Pembobotan TF-IDF Versi Excel..... | 53 |
| Gambar 4.14 Confusion Matrix Skenario 1 | 61 |
| Gambar 4.15 Classification Report Skenario 1..... | 62 |
| Gambar 4.16 Confusion Matrix Skenario 2 | 63 |
| Gambar 4.17 Classification Report Skenario 2..... | 64 |
| Gambar 4.18 Confusion Matrix Skenario 3 | 65 |
| Gambar 4.19 Classification Report Skenario 3..... | 66 |
| Gambar 4.20 Confusion Matrix Skenario 4 | 67 |
| Gambar 4.21 Classification Report Skenario 4..... | 68 |
| Gambar 4.22 Confusion Matrix Skenario 5 | 69 |
| Gambar 4.23 Classification Report Skenario 5..... | 70 |

| | |
|--|----|
| Gambar 4.24 Confusion Matrix Skenario 6 | 71 |
| Gambar 4.25 Classification Report Skenario 6..... | 72 |
| Gambar 4.26 Confusion Matrix Skenario 7 | 73 |
| Gambar 4.27 Classification Report Skenario 7 | 74 |
| Gambar 4.28 Confusion Matrix Skenario 8 | 75 |
| Gambar 4.29 Classification Report Skenario 8 | 75 |
| Gambar 4.30 Confusion Matrix Skenario 9 | 77 |
| Gambar 4.31 Classification Report Skenario 9..... | 77 |
| Gambar 4.32 Confusion Matrix Skenario 10 | 79 |
| Gambar 4.33 Classification Report Skenario 10..... | 79 |
| Gambar 4.34 Confusion Matrix Skenario 11 | 81 |
| Gambar 4.35 Classification Report Skenario 11 | 81 |
| Gambar 4.36 Confusion Matrix Skenario 12 | 83 |
| Gambar 4.37 Classification Report Skenario 12..... | 83 |
| Gambar 4.38 Diagram Perbandingan Akurasi Hasil Skenario Pengujian..... | 85 |
| Gambar 4.39 Jumlah Sentimen Masing-Masing Aspek..... | 86 |

DAFTAR TABEL

| | |
|--|----|
| Tabel 2.1 Proses Case-folding | 11 |
| Tabel 2.2 Proses Cleaning..... | 11 |
| Tabel 2.3 Proses Tokenizing | 12 |
| Tabel 2.4 Proses Stopword Removal | 13 |
| Tabel 2.5 Proses Stemming..... | 13 |
| Tabel 2.6 Confusion Matrix | 16 |
| Tabel 3.1 Contoh Case-folding | 27 |
| Tabel 3.2 Contoh Cleaning | 28 |
| Tabel 3.3 Contoh Tokenizing..... | 28 |
| Tabel 3.4 Contoh Stopword Removal..... | 29 |
| Tabel 3.5 Contoh Stemming | 29 |
| Tabel 3.6 Ringkasan Perhitungan Probabilitas Term..... | 35 |
| Tabel 3.7 Skenario Pengujian | 37 |
| Tabel 4.1 Skenario Pengujian | 58 |
| Tabel 4.2 Perbandingan Hasil Skenario Pengujian | 84 |

DAFTAR KODE PROGRAM

| | |
|--|----|
| Kode Program 4.1 Scrapping Data | 40 |
| Kode Program 4.2 Import Library Preprocessing Data | 43 |
| Kode Program 4.3 Upload File Untuk Preprocessing Data | 43 |
| Kode Program 4.4 Case-folding..... | 45 |
| Kode Program 4.5 Cleaning..... | 46 |
| Kode Program 4.6 Tokenizing | 47 |
| Kode Program 4.7 Stopword Removal | 48 |
| Kode Program 4.8 Stemming..... | 49 |
| Kode Program 4.9 Label Encoding..... | 51 |
| Kode Program 4.10 TF-IDF..... | 52 |
| Kode Program 4.11 Split Data | 54 |
| Kode Program 4.12 Multinomial Naive Bayes..... | 57 |
| Kode Program 4.13 Memanggil Class Multinomial Naive Bayes..... | 58 |
| Kode Program 4.14 Confusion Matrix..... | 60 |
| Kode Program 4.15 Classification Report | 61 |