



**PENERAPAN ALGORITMA NAÏVE BAYES UNTUK ANALISIS SENTIMEN
ULASAN PADA *GOOGLE PLAY STORE* TERHADAP SALAH SATU APLIKASI
*MOBILE BANKING***

SKRIPSI

AINI CAHYANING PUTRI

1910511115

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

2023



**PENERAPAN ALGORITMA NAÏVE BAYES UNTUK ANALISIS SENTIMEN
ULASAN PADA *GOOGLE PLAY STORE* TERHADAP SALAH SATU APLIKASI
*MOBILE BANKING***

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Komputer**

AINI CAHYANING PUTRI

1910511115

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

2023

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri, dan sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Aini Cahyaning Putri

NIM : 1910511115

Tanggal : 17 Juli 2023

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 17 Juli 2023

Yang Menyatakan,



Aini Cahyaning Putri

PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan dibawah ini :


Nama : Aini Cahyaning Putri
NIM : 1910511115
Fakultas : Ilmu Komputer
Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non eksklusif (*Non exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul :

**PENERAPAN ALGORITMA NAÏVE BAYES UNTUK ANALISIS SENTIMEN
ULASAN PADA *GOOGLE PLAY STORE* TERHADAP SALAH SATU APLIKASI
*MOBILE BANKING***

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.
Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 17 Juli 2023
Yang Menyatakan,


Aini Cahyaning

LEMBAR PERSETUJUAN

Yang bertanda tangan di bawah ini :

Nama : Aini Cahyaning Putri
NIM : 1910511115
Program Studi : SI Informatika
Judul : Penerapan Algoritma Naïve Bayes untuk Analisis Sentimen
Ulasan pada *Google Play Store* Terhadap Salah Satu
Aplikasi *Mobile Banking*

Dinyatakan telah memenuhi syarat dan menyetujui untuk mengikuti sidang skripsi.

Jakarta, 05 Juni 2023

Mengetahui,

Menyetujui,



Dr. Widya Cholil, S.Kom., M.IT.
Ketua Program Studi



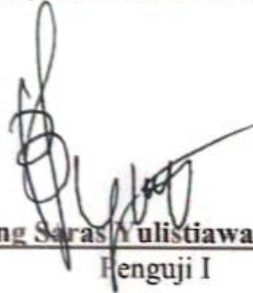
Dr. Widya Cholil, S.Kom., M.IT.
Dosen Pembimbing

LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa Skripsi berikut :

Nama : Aini Cahyaning Putri
NIM : 1910511115
Program Studi : S1 Informatika
Judul : Penerapan Algoritma Naïve Bayes untuk Analisis Sentimen
Ulasan pada *Google Play Store* Terhadap Salah Satu
Aplikasi *Mobile Banking*

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Dr. Bambang Saras Yulistiawan, S.T., M.Kom.
Penguji I



Iin Ernawati, S.Kom., M.Si.
Penguji II



Dr. Widva Cholil, S.Kom., M.IT.
Pembimbing



Dr. Widva Cholil, S.Kom., M.IT.
Kepala Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 3 Juli 2023

PENERAPAN ALGORITMA NAÏVE BAYES UNTUK ANALISIS SENTIMEN ULASAN PADA *GOOGLE PLAY STORE* TERHADAP SALAH SATU APLIKASI *MOBILE BANKING*

Aini Cahyaning Putri

ABSTRAK

Pada saat ini perkembangan teknologi semakin berkembang dengan pesat sehingga setiap penggunanya dapat dengan bebas memiliki aplikasi jenis apapun dan sebanyak apapun yang diinginkan seperti dengan aplikasi Google yang menyediakan layanan Play Store. Salah satunya fitur yang tersedia oleh *Google Play Store* yaitu fitur dimana pengguna dapat memberikan sebuah penilaian berupa *review* atau ulasan terhadap produk yang digunakan. Seperti pada saat ini sebuah aplikasi layanan Internet Banking dari PT Bank Negara Indonesia (Persero) Tbk yang saat ini dapat diakses melalui *smartphone* masing – masing oleh nasabah. Dari kondisi tersebut, maka dilakukan penelitian mengenai analisis sentimen pada ulasan pengguna terhadap sistem aplikasi salah satu *Mobile Banking* dengan menggunakan metode Naïve Bayes. Jumlah data yang di crawling sebanyak 1000 data yang belum memiliki label. Sebelum dilakukan proses klasifikasi data yang diperoleh, data harus melalui tahap pelabelan serta pembersihan data terlebih dahulu sebelum masuk pada tahap *text processing*, selanjutnya data diberikan bobot pada tiap kata dengan *Term Frequency-Inverse Document Frequency* (TF-IDF) yang nantinya kata tersebut akan dijadikan sebagai fitur. Lalu, karena data berlabel positif serta negatif mempunyai jumlah yang jauh berbeda, maka diterapkanlah metode *Synthetic Minority Oversampling Technique* (SMOTE) yang berfungsi sebagai penyeimbangan data. Tahap selanjutnya dilakukan pembagian data yang besarnya berjumlah 80% dan 20% serta diklasifikasikan dengan metode Naïve Bayes. Hasil yang diperoleh dari penelitian ini memberikan nilai bahwa data uji memiliki akurasi dengan presentase sebesar 95%, *precision* dengan presentase sebesar 95%, *recall* dengan presentase sebesar 95%.

Kata Kunci : Analisis Sentimen, Klasifikasi, *Mobile Banking*, Naïve Bayes

APPLICATION OF NAÏVE BAYES ALGORITHM FOR SENTIMENT ANALYSIS OF REVIEWS ON GOOGLE PLAY STORE ON ONE OF THE *MOBILE BANKING* APPLICATIONS

Aini Cahyaning Putri

ABSTRACT

At this time the development of technology is growing rapidly so that every user can freely have any type of application and as much as desired such as Google applications that provide Play Store services. One of the features provided by the Google Play Store is a feature where users can provide an assessment in the form of a review or review of the product used. As at this time the Internet Banking service application from PT Bank Negara Indonesia (Persero) Tbk which can now be accessed via the smartphone of each customer. From these conditions, research was conducted on sentiment analysis on user reviews of one of the *Mobile Banking* application systems using the Naïve Bayes method. The amount of data crawled is 1000 data that does not yet have a label. Before the data classification process is obtained, the data must go through the labeling and data cleaning stage first before entering the text processing stage, then the data is given a weight on each word with Term Frequency-Inverse Document Frequency (TF-IDF) which later the word will be used as a feature. Then, because the data labeled positive and negative has a much different amount, the *Synthetic Minority Oversampling Technique* (SMOTE) method is applied which functions as a data balancer. The next stage is to divide the data which amounts to 80% and 20% and classified with the Naïve Bayes method. The results obtained from this study provide a value that the test data has *accuracy* with a percentage of 95%, *precision* with a percentage of 95%, *recall* with a percentage of 95%.

Keyword : Sentiment Analysis, Classification, *Mobile Banking*, Naïve Bayes

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kehadirat Allah SWT karena berkat limpahan rahmat dan karunia-Nya, penulis dapat menyelesaikan Skripsi (Tugas Akhir) ini tepat pada waktunya. Skripsi ini disusun sebagai pemenuhan syarat kelulusan Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.

Dalam penyelesaian karya tulis ilmiah ini tidak lepas dari bantuan banyak pihak yang telah memberikan masukan kepada penulis. Untuk itu penulis mengucapkan terima kasih kepada :

1. Dr. Ermatita, M.Kom., selaku dekan Fakultas Ilmu Komputer
2. Dr. Widya Cholil, S.Kom., M.IT., selaku Ketua Program Studi Sarjana Jurusan S1 Informatika
3. Dr. Widya Cholil, S.Kom., M.IT., selaku dosen pembimbing dari pihak jurusan.
4. Orang tua yang telah memberikan dukungan baik secara moril maupun materil.
5. Seluruh pihak yang terlibat dalam kelancaran pembuatan makalah karya ilmiah ini dan yang belum disebutkan di atas, penulis mengucapkan terima kasih.

Penulis menyadari bahwa masih banyak terdapat kekurangan dari laporan ini, baik dari materi maupun teknik penyajian laporan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun agar penulis dapat memperbaikinya.

Jakarta, 06 Juni 2023

Penulis

DAFTAR ISI

PERNYATAAN ORISINALITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI	iv
LEMBAR PERSETUJUAN.....	v
LEMBAR PENGESAHAN	vi
ABSTRAK	vii
ABSTRACT.....	viii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
DAFTAR SIMBOL	xiv
BAB 1	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Ruang Lingkup.....	3
1.4 Tujuan	3
1.5 Luaran yang Diharapkan.....	3
1.6 Manfaat	3
1.7 Sistematika Penulisan	4
BAB 2	5
2.1 <i>Google Play Store</i>	5
2.2 Analisis Sentimen	5
2.3 Fleiss Kappa.....	6
2.4 <i>Text Analysis</i>	7
2.3.1 Text Mining	7
2.3.2 Pra Proses Data	8
2.3.2.1 Transform Cases	8
2.3.2.2 Pembersihan Data	8
2.3.2.3 Stopword Removal.....	8
2.3.2.4 Stemming	9

2.3.2.5	Tokenisasi	9
2.5	Metode Klasifikasi	9
2.5.1	Algoritma <i>Naïve Bayes</i>	9
2.5.2	Klasifikasi <i>Multiclass</i>	10
2.5.3	<i>Synthetic Minority Oversampling Technique (SMOTE)</i>	11
2.6	Evaluasi.....	11
2.7	<i>Python</i>	13
2.8	Riset Yang Relevan.....	13
BAB 3	14
3.1	Kerangka Pikir	14
3.1.1	Identifikasi Masalah.....	15
3.1.2	Studi Literatur	15
3.1.3	Akuisisi Data.....	15
3.1.4	Pelabelan Data.....	16
3.1.5	Pra Proses Data	16
3.1.5.1	Transform Cases	16
3.1.5.2	Cleaning	17
3.1.5.3	Stopword Removal.....	18
3.1.5.4	Stemming	18
3.1.5.5	Tokenisasi	18
3.1.5.6	Normalisasi	18
3.1.6	Pembobotan TF IDF.....	19
3.1.7	Klasifikasi	19
3.1.8	Hasil Klasifikasi.....	19
3.1.9	Evaluasi.....	19
3.2	Perangkat Penelitian.....	20
BAB 4	21
4.1	Pengumpulan Data	21
4.2	Pelabelan Data	21
4.3	Praproses.....	26
4.3.1.	<i>Case Folding</i>	26
4.3.2.	<i>Data Cleaning</i>	27
4.3.3.	<i>Normalization</i>	29

4.3.4.	<i>Stemming</i>	30
4.3.5.	<i>Stopwords Removal</i>	31
4.3.6.	<i>Tokenizing</i>	33
4.4	Pembobotan TF-IDF	34
4.5	SMOTE (<i>Synthetic Minority Oversampling Technique</i>).....	35
4.6	Klasifikasi Naïve Bayes	36
4.6.1	Data Latih.....	36
4.6.2	Data Uji	39
4.7	Evaluasi.....	43
4.8	Analisis Hasil Kegiatan.....	46
4.8.1	<i>Wordcloud</i> Sentimen Positif terhadap Data Ulasan BNI <i>Mobile Banking</i>	46
4.8.2	<i>Wordcloud</i> Sentimen Negatif terhadap Data Ulasan BNI <i>Mobile Banking</i>	47
BAB 5	48
5.1	Kesimpulan	48
5.2	Saran	48
DAFTAR PUSTAKA	50
RIWAYAT HIDUP	52
LAMPIRAN	53






DAFTAR GAMBAR

Gambar 3.1 Kerangka Berpikir	14
Gambar 4.1 Hasil Crawling Data Aplikasi.....	21
Gambar 4.2 Wordcloud Sentimen Positif.....	46
Gambar 4.3 Wordcloud Sentimen Negatif	47

DAFTAR TABEL

Tabel 3.1 Tabel <i>Flowchart</i>	xiv
Tabel 2.1 Interpretasi Kappa.....	7
Tabel 3.2 Tabel <i>Cleaning</i>	17
Tabel 3.3 Tabel <i>Stopword Removal</i>	18
Tabel 3.4 <i>Confusion Matrix</i>	11
Tabel 4.2 Review dan Hasil Pelabelan	22
Tabel 4.3 Hasil Penilaian Anotator.....	23
Tabel 4.4 Tabel Hasil Perhitungan Nilai <i>pi</i>	24
Tabel 4.5 Case Folding	26
Tabel 4.6 Data Cleaning	28
Tabel 4.7 Normalization	29
Tabel 4.8 Stemming	30
Tabel 4.9 Stopword Removal	31
Tabel 4.10 Tokenizing	33
Tabel 4.11 Contoh Data TF-IDF.....	34
Tabel 4.12 Penghitungan TF-IDF.....	34
Tabel 4.13 Jumlah Data Sebelum dan Sesudah SMOTE.....	35
Tabel 4.14 Sample Data Latih	36
Tabel 4.15 Nilai Probabilitas Pada Data Latih.....	39
Tabel 4.16 Contoh Data Uji	39
Tabel 4.17 Data Uji yang Setelah Praproses.....	40
Tabel 4.18 Probabilitas Sampel Data Uji.....	40
Tabel 4.19 Nilai Probabilitas Kelas Positif.....	41
Tabel 4.20 Nilai Probabilitas Kelas Negatif	42
Tabel 4.21 Confusion Matrix Model	44

DAFTAR SIMBOL

Simbol <i>Flowchart</i>			
No	Gambar	Nama	Keterangan
1		<i>Terminal (Start, End)</i>	Penggambaran ketika kegiatan yang akan dimulai maupun kegiatan yang berakhir
2		<i>Flow Direction</i>	Penggambaran mengenai hubungan antar simbol yang menyatakan suatu jalannya proses pada sistem
3		<i>Process</i>	Penggambaran mengenai deskripsi dari proses yang dijalankan
4		<i>Document</i>	Penggambaran bahwa masukkan (<i>input</i>) berasal dari sebuah data dokumen yang dapat berupa kertas atau keluaran (<i>output</i>) yang dicetak ke kertas
5		<i>Predefine Process</i>	Penggambaran pelaksanaan dari sebuah proses atau disebut sebagai subprogram

Tabel 3.1 Tabel *Flowchart*