

**ANALISIS SENTIMEN ULASAN PENGGUNA
APLIKASI FLO PADA *GOOGLE PLAY STORE*
MENGUNAKAN METODE *NAÏVE BAYES*
DAN *SUPPORT VECTOR MACHINE***

SKRIPSI

Oleh:

Prilia Hashifah Syafina

202010225060



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

LEMBAR PERSETUJUAN PEMBIMBING

Judul Proposal Tugas Akhir : Analisis Sentimen Ulasan Pengguna Aplikasi FLO
Pada *Google Play Store* Menggunakan Metode
Naïve Bayes Dan *Support Vector Machine*

Nama Mahasiswa : Prilia Hashifah Syafina

Nomor Pokok Mahasiswa : 202010225060

Program Studi/Fakultas : Informatika / Ilmu Komputer

Jakarta, 05 Juli 2024

MENYETUJUI,

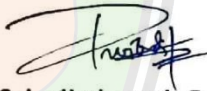
Pembimbing I

Pembimbing II



(Dr. Herlawati, S.Si., M.M., M.Kom.)

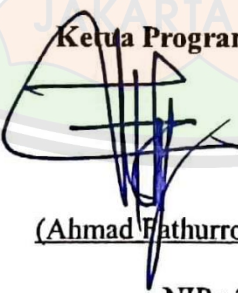
NIDN : 0311097302



(Dwi Budi Srisulistiowati, S.Kom., M.M.)

NIDN : 0323057701

Ketua Program Studi/Sesprodi



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

NIP : 2012486

Program Studi Informatika

Fakultas Ilmu Komputer

Universitas Bhayangkara Jakarta Raya

2024

LEMBAR PENGESAHAN

Judul Tugas akhir : Analisis Sentimen Ulasan Pengguna
: Aplikasi FLO Pada *Google Play Store*
Menggunakan Metode *Naïve Bayes* Dan
Support Vector Machine

Nama Mahasiswa : Prilia Hashifah Syafina
Nomor Pokok Mahasiswa : 202010225060
Program Studi/Fakultas : Informatika / Ilmu Komputer
Tanggal Lulus Ujian Tugas akhir : 27 Juni 2024

Jakarta, 05 Juli 2024

MENGESAHKAN,

Ketua Tim Penguji : Ratna Salkiawati, S.T., M.Kom.
NIDN : 0310038006

Penguji I : Ajif Yunizar Pratama Yusuf, S.Si., M.Eng.
NIDN : 0328068603

Penguji II : Dr. Herlawati, S.Si., M.M., M.Kom.
NIDN : 0311097302

MENGETAHUI,

Ketua
Program Studi Informatika

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

Dekan
Fakultas Ilmu Komputer

Dr. Dra. Tyastut 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 : Prilia Hashifah Syafina
NPM : 202010225060
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Ulasan Pengguna Aplikasi FLO Pada
*Google Play Store Menggunakan Metode Naïve Bayes Dan
Support Vector Machine*

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.

Jakarta, 05 Juli 2024

Penulis



Prilia Hashifah Syafina

ABSTRAK

Prilia Hashifah Syafina. 202010225060. Analisis Sentimen Ulasan

Pengguna Aplikasi FLO Pada *Google Play Store* Menggunakan Metode *Naïve Bayes* Dan *Support Vector Machine*

Kemajuan teknologi yang cepat telah membawa dampak signifikan pada industri teknologi informasi dan telekomunikasi, khususnya dalam pengembangan aplikasi mobile. Salah satu aplikasi terkemuka dalam bidang kesehatan reproduksi adalah FLO, yang digunakan oleh lebih dari 350 juta orang untuk melacak siklus menstruasi, ovulasi, dan aspek kesehatan reproduksi lainnya. Memahami pengalaman pengguna aplikasi ini melalui analisis ulasan yang diberikan pada *Google Play Store* adalah hal yang penting. Penelitian ini menggunakan metode Cross-Industry Standard Process for Data Mining (CRISP-DM) untuk menganalisis sentimen ulasan pengguna aplikasi FLO. Algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM) digunakan untuk mengklasifikasikan sentimen ulasan menjadi positif dan negatif. Data yang dianalisis meliputi 1.393 ulasan, di mana 796 ulasan menunjukkan sentimen positif dan 597 ulasan menunjukkan sentimen negatif. Hasil penelitian menunjukkan bahwa baik *Naïve Bayes* maupun SVM memiliki tingkat akurasi yang sama, yaitu sebesar 74%. Namun, metrik lainnya seperti recall dan precision menunjukkan bahwa SVM memiliki performa yang sedikit lebih baik dibandingkan *Naïve Bayes*. SVM memiliki nilai recall sebesar 74% dan precision sebesar 71%, sedangkan *Naïve Bayes* memiliki nilai recall sebesar 88% dan precision sebesar 67%.

Kata Kunci: analisis sentimen, FLO, *Google Play Store*, *Naïve Bayes*, *Support Vector Machine*.

ABSTRACT

Prilia Hashifah Syafina. 202010225060. *Sentiment Analysis of FLO App User Reviews on Google Play Store Using Naïve Bayes Method and Support Vector Machine.*

Rapid technological advancements have brought significant impact to the information technology and telecommunications industry, particularly in the development of mobile applications. One of the leading apps in the field of reproductive health is FLO, which is used by more than 350 million people to track their menstrual cycle, ovulation, and other aspects of reproductive health. Understanding the user experience of this app through analyzing reviews given on the Google Play Store is important. This research uses the Cross-Industry Standard Process for Data Mining (CRISP-DM) method to analyze the sentiment of FLO app user reviews. Naïve Bayes and Support Vector Machine (SVM) algorithms are used to classify review sentiment into positive and negative. The analyzed data includes 1,393 reviews, of which 796 reviews show positive sentiment and 597 reviews show negative sentiment. The results show that both Naïve Bayes and SVM have the same accuracy rate of 74%. However, other metrics such as recall and precision show that SVM performs slightly better than Naïve Bayes. SVM has a recall value of 74% and a precision of 71%, while Naïve Bayes has a recall value of 88% and a precision of 67%.

Keywords: *sentiment analysis, FLO, Google Play Store, Naïve Bayes, Support Vector Machine.*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

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

Nama : Prilia Hashifah Syafina
NPM : 202010225060
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 Ulasan Pengguna Aplikasi FLO Pada *Google Play Store* Menggunakan Metode *Naïve Bayes* Dan *Support Vector Machine*

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 : 05 Juli 2024
Yang Menyatakan


PRILIA HASHIFAH SYAFINA

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT yang telah melimpahkan rahmat dan hidayah nya sehingga telah memberikan kesempatan bagi penulis untuk dapat menyelesaikan tugas akhir skripsi yang berjudul “**Analisis Sentimen Ulasan Pengguna Aplikasi FLO Pada Google Play Store Menggunakan Metode Naïve Bayes Dan Support Vector Machine**” dengan tepat waktu. Maksud dan tujuan penyusunan tugas akhir skripsi ini adalah sebagai bentuk hasil kegiatan yang dilakukan selama kuliah di Universitas Bhayangkara Jakarta Raya.

Tidak lupa penulis juga mengucapkan terimakasih kepada semua pihak yang secara langsung maupun tidak langsung terlibat dalam penelitian ini. Penulis mengucapkan terima kasih sebesar-besarnya 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, SE., MMSI. selaku Ketua Program studi Informatika Universitas Bhayangkara Jakarta Raya.
4. Ibu Dr. Herlawati, S.Si., M.M., M.Kom. selaku Dosen Pembimbing I yang telah membimbing, serta memberikan masukan dan motivasi dalam penulis dalam menyusun penelitian ini.
5. Ibu Dwi Budi Srisulistiwati, S.Kom., M.M. selaku Dosen Pembimbing II yang telah membimbing, serta memberikan masukan dan motivasi dalam penulis dalam menyusun penelitian ini.

6. Bapak Kusdarnowo Hantoro, S.Kom., M.Kom. selaku Dosen Pembimbing Akademik penulis selama menjadi mahasiswa di Universitas Bhayangkara Jakarta Raya.
7. Kepada seluruh Bapak dan Ibu Dosen Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya yang selama ini memberikan ilmu pengetahuan.
8. Kepada seluruh Staf Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya atas bantuan dalam urusan administrasi.
9. Kedua Orang Tua yang selalu memberikan dukungan dengan tulus, doa yang tiada hentinya untuk penulis, membantu dalam hal apapun, dan selalu memberikan yang terbaik bagi penulis.
10. Para sahabat Syafira Cessa Agustin, Nida Rachmatin, Silviana Sevira, serta teman-teman mahasiswa/i Universitas Bhayangkara Jakarta Raya yang selalu membantu dan memberi semangat dalam kondisi apapun.

Penulis menyadari bahwa masih banyak kekurangan dari tugas akhir skripsi ini, baik dari materi maupun teknik penyajiannya, mengingat kurangnya pengetahuan dan pengalaman menulis. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun. Semoga penelitian ini bermanfaat bagi penulis dan pembaca.

Jakarta, 05 Juli 2024



Prilia Hashifah Syafina
202010225060

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	9
1.3 Rumusan Masalah	9
1.4 Tujuan Penelitian	9
1.5 Manfaat Penelitian	10
1.6 Batasan Masalah	10
1.7 Sistematika Penulisan	11
BAB II TINJAUAN PUSTAKA	13

2.1	Penelitian Terdahulu	13
2.2	Teori Pendukung.....	20
2.2.1	<i>Machine Learning</i>	20
2.2.2	<i>Data Mining</i>	21
2.2.3	<i>CRoss-Industry Process for Data Mining (CRISP-DM)</i>	24
2.2.4	<i>Text Mining</i>	27
2.2.5	<i>Term Frequency – Invers Document Frequency (TF-IDF)</i>	29
2.2.6	<i>Analisis Sentimen</i>	31
2.2.7	<i>Web Scraping</i>	33
2.2.8	<i>Jupyter Notebook</i>	34
2.2.9	<i>Google Play Store</i>	36
2.2.10	<i>FLO</i>	37
2.2.11	<i>Python</i>	39
2.2.12	<i>Naïve Bayes</i>	40
2.2.13	<i>Support Vector Machine</i>	42
2.2.14	<i>Confusion Matrix</i>	43
2.2.15	<i>Synthetic Minority Over-sampling Technique (SMOTE)</i>	45
2.2.16	<i>Lexicon</i>	47
2.2.17	<i>Word Cloud</i>	48
BAB III	METODOLOGI PENELITIAN.....	49
3.1	Objek Penelitian	49
3.2	Kerangka Penelitian	50

3.2.1	Pemahaman Bisnis (<i>Business Understanding</i>)	51
3.2.2	Pemahaman Data (<i>Data Understanding</i>)	52
3.2.3	Persiapan Data (<i>Data Preparation</i>)	54
3.2.4	Pemodelan (<i>Modeling</i>)	54
3.2.5	Evaluasi (<i>Evaluation</i>)	57
3.2.6	Penerapan (<i>Deployment</i>)	57
BAB IV HASIL DAN PEMBAHASAN		58
4.1	Pemahaman Bisnis (<i>Business Understanding</i>)	58
4.2	Pemahaman Data (<i>Data Understanding</i>)	59
4.3	Persiapan Data (<i>Data Preparation</i>)	61
4.3.1	<i>Case Folding</i>	61
4.3.2	<i>Tokenizing</i>	62
4.3.3	<i>Normalize</i>	62
4.3.4	<i>Stopword Removal</i>	63
4.3.5	<i>Stemming</i>	63
4.3.6	Pelabelan	64
4.3.7	TF-IDF	65
4.3.8	SMOTE	69
4.4	Pemodelan (<i>Modeling</i>)	70
4.5.1	<i>Splitting Data</i>	70
4.5.2	<i>Naïve Bayes</i>	71
4.4.3	<i>Support Vector Machine</i>	77

4.5	Evaluasi (<i>Evaluation</i>)	82
4.5.1	<i>Naïve Bayes</i>	82
4.5.2	<i>Support Vector Machine</i>	86
4.6	Penerapan (<i>Deployment</i>)	92
BAB V PENUTUP		97
5.1	Kesimpulan.....	97
5.2	Saran	98
DAFTAR PUSTAKA		101
LAMPIRAN		111



DAFTAR TABEL

Table 2. 1 Penelitian Terdahulu	13
Tabel 4. 1 Tahapan <i>Scraping</i>	59
Tabel 4. 2 Tahapan Penghapusan Kolom yang Tidak Relevan.....	60
Tabel 4. 3 Hasil Penerapan <i>Case Folding</i>	62
Tabel 4. 4 Hasil Penerapan <i>Tokenizing</i>	62
Tabel 4. 5 Hasil Penerapan <i>Normalize</i>	62
Tabel 4. 6 Hasil Penerapan <i>Stopword Removal</i>	63
Tabel 4. 7 Hasil Penerapan <i>Stemming</i>	64
Tabel 4. 8 Tahapan TF-IDF	65
Tabel 4. 9 Sebagian Data Perhitungan TF-IDF	66
Tabel 4. 10 Hasil TF Normalisasi.....	67
Tabel 4. 11 Hasil <i>Document Frequency</i>	67
Tabel 4. 12 Hasil <i>Inverse Document Frequency (IDF)</i>	68
Tabel 4. 13 Hasil TF-IDF	69
Tabel 4. 14 <i>Data Training</i>	73
Tabel 4. 15 <i>Data Testing</i>	76
Tabel 4. 16 Hasil Probabilitas <i>Data Testing</i>	77
Tabel 4. 17 Hasil TF-IDF dari Beberapa Ulasan	78
Tabel 4. 18 Hasil <i>Confusion Matrix Naïve Bayes</i> dan SVM	89

DAFTAR GAMBAR

Gambar 1. 1 Perbandingan Ulasan Pengguna Dari Aplikasi FLO, Clue dan Clover	4
Gambar 1. 2 Kelebihan Aplikasi FLO	4
Gambar 1. 3 Contoh Ulasan Positif dan Negatif	5
Gambar 1. 4 Hasil Kuesioner Terhadap Aplikasi FLO	8
Gambar 2. 1 Gambaran Dalam Artificial Intelligence.....	20
Gambar 2. 2 Tahapan Data Mining.....	23
Gambar 2. 3 Tahapan CRISP-DM	25
Gambar 2. 4 Tahapan Text Mining	28
Gambar 2. 5 Logo Jupyter Notebook	34
Gambar 2. 6 Logo Google Play Store.....	36
Gambar 2. 7 Logo Aplikasi FLO	37
Gambar 2. 8 Confusion matrix	44
Gambar 3. 1 Tampilan Beranda Aplikasi FLO	49
Gambar 3. 2 Kerangka Penelitian.....	50
Gambar 3. 3 Hasil <i>Scraping</i> Data Ulasan Aplikasi FLO	53
Gambar 3. 4 Download Hasil <i>Scraping</i> Data	53
Gambar 3. 5 Proses Algoritma <i>Naïve Bayes</i>	55
Gambar 3. 6 Proses Algoritma <i>Support Vector Machine</i>	56
Gambar 4. 1 Hasil <i>Scraping</i> Ulasan Pengguna Aplikasi FLO.....	60
Gambar 4. 2 Hasil Penghapusan Kolom yang Tidak Relevan.....	61

Gambar 4. 3 Tahapan Pelabelan	65
Gambar 4. 4 Jumlah Ulasan Label Positif dan Negatif	65
Gambar 4. 5 Hasil TF-IDF Seluruh Dokumen	66
Gambar 4. 6 Contoh Kode dan Hasil SMOTE	69
Gambar 4. 7 Hasil Grafik Setelah SMOTE	70
Gambar 4. 8 Tahapan <i>Spliting</i> Data.....	71
Gambar 4. 9 Hasil <i>Modeling Multinomial Naïve Bayes</i>	72
Gambar 4. 10 Hasil <i>Modeling Gaussian Naïve Bayes</i>	72
Gambar 4. 11 Hasil <i>Modeling Support Vector Machine</i>	77
Gambar 4. 12 Hasil Klasifikasi <i>Support Vector Machine</i> dengan Kernel <i>Linear</i>	78
Gambar 4. 13 Hasil Evaluasi <i>Multinomial Naïve Bayes</i>	83
Gambar 4. 14 <i>Confusion Matrix Naïve Bayes</i> Dengan <i>Multinomial Naïve Bayes</i>	83
Gambar 4. 15 Hasil Evaluasi <i>Gaussian Naïve Bayes</i>	84
Gambar 4. 16 <i>Confusion Matrix Naïve Bayes</i> Dengan <i>Gaussian Naïve Bayes</i>	85
Gambar 4. 17 Hasil Evaluasi <i>Support Vector Machine</i> Dengan Kernel <i>Linear</i>	86
Gambar 4. 18 <i>Confusion Matrix Support Vector Machine</i> Dengan Kernel <i>Linear</i>	87
Gambar 4. 19 <i>Word Cloud</i> Pada Ulasan Positif.....	90
Gambar 4. 20 Jumlah Kemunculan Kata dalam Ulasan Positif	91

Gambar 4. 21 <i>Word Cloud</i> Pada Ulasan Negatif	91
Gambar 4. 22 Jumlah Kemunculan Kata Pada Ulasan Negatif.....	92
Gambar 4. 23 Tampilan Utama <i>Website</i>	93
Gambar 4. 24 Memasukkan Ulasan Untuk Analisis Sentimen.....	94
Gambar 4. 25 Hasil Analisis Sentimen Positif	95
Gambar 4. 26 Memasukkan Ulasan Untuk Analisis Sentimen.....	96
Gambar 4. 27 Hasil Analisis Sentimen Negatif.....	96



DAFTAR LAMPIRAN

Lampiran 1 : Plagiarisme

Lampiran 2 : Biodata Mahasiswa

Lampiran 3 : Kartu Bimbingan Mahasiswa

Lampiran 4 : Data Mentah Ulasan Pengguna Aplikasi FLO

Lampiran 5 : Hasil Kuesioner

