

**ANALISIS SENTIMEN ULASAN APLIKASI *SP4N*
LAPOR! DI *GOOGLE PLAY STORE* MENGGUNAKAN
ALGORITMA *SUPPORT VECTOR MACHINE***

SKRIPSI

Oleh:

Oki Akbar Ramadhan

201910225020



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2023**

LEMBAR PERSETUJUAN PEMBIMBING

Judul Tugas Akhir : Analisis Sentimen Ulasan Aplikasi *SP4N LAPOR!*
di *Google Play Store* Menggunakan Algoritma
Support Vector Machine

Nama Mahasiswa : Oki Akbar Ramadhan

Nomor Pokok Mahasiswa : 201910225020

Program Studi / Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 01 Agustus 2023

Jakarta, 11 Agustus 2023

MENYETUJUI,

Pembimbing I

Pembimbing II

Prima Dina Atika, S.Kom., M.Kom.

NIDN : 0311037107

Siti Setiawati, S.Pd., M.Pd.

NIDN : 0313107904

LEMBAR PENGESAHAN

Judul Skripsi : ANALISIS SENTIMEN ULASAN APLIKASI
SP4N LAPOR! DI GOOGLE PLAY STORE
MENGUNAKAN ALGORITMA SUPPORT
VECTOR MACHINE

Nama Mahasiswa : Oki Akbar Ramadhan

Nomor Pokok Mahasiswa : 201910225020

Program Studi / Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 01 Agustus 2023

Jakarta, 11 Agustus 2023

MENGESAHKAN,

Ketua Tim Penguji : Wowon Priatna, S.T., M.T.I.

NIDN : 0429118007

Penguji I : Asep Ramdhani Mahbub, S.Kom., M.Kom.

NIDN : 0329087703

Penguji II : Prima Dina Atika, S.Kom., M.Kom.

NIDN : 0311037107

MENGETAHUI,

Ketua

Dekan

Program Studi Informatika

Fakultas Ilmu Komputer



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

NIP. 2012486



Dr. Dra. Tyastuti Sri Lestari, M.M.

NIP. 1408206



LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Oki Akbar Ramadhan
NPM : 201910225020
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Ulasan Aplikasi *SP4N LAPOR!* di
*Google Play Store Menggunakan Algoritma 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, 11 Agustus 2023

Penulis



Oki Akbar Ramadhan

ABSTRAK

Oki Akbar Ramadhan. 201910225020. Analisis Sentimen Ulasan Aplikasi *SP4N LAPOR!* di *Google Play Store* menggunakan Algoritma *Support Vector Machine*. 2023.

Salah satu upaya yang dilakukan oleh pemerintah dalam menyediakan sarana pengaduan masyarakat terhadap pelayanan publik, pemerintah Indonesia meluncurkan aplikasi *SP4N LAPOR!* yang digunakan untuk menerima pengaduan dan aspirasi masyarakat Indonesia terhadap pelayanan publik melalui aplikasi *play store*. Namun Semenjak perilisannya, beberapa pengguna mengeluh ketidakpuasan terhadap penggunaan aplikasi tersebut. Beberapa jenis keluhan pengguna *SP4N LAPOR!* diantaranya server yang error, respon yang lambat dalam menanggapi aduan, dan sebagainya. Berdasarkan masalah tersebut, peneliti melakukan analisis sentimen pada Aplikasi *SP4N LAPOR!* melalui ulasan pengguna menggunakan metode *Support Vector Machine (SVM)*. Data yang diperoleh setelah dilakukan pemilahan dan pelabelan adalah 1189 data ulasan dengan ulasan negatif sebanyak 632 data dan ulasan positif sebanyak 557 data. Data yang sudah diperoleh dilanjutkan dengan tahapan *text processing*, kemudian data diberikan bobot dengan *Term Frequency-Inverse Document Frequency (TF-IDF)*. Tahapan selanjutnya dilakukan 4 skenario *splitting* data yang dikombinasikan dengan *10-fold validation* yaitu 60%:40%, 70%:30%, 80%:20%, 90%:10%, kemudian diklasifikasikan dengan algoritma *SVM*. Hasil evaluasi kinerja algoritma *Support Vector Machine* diperoleh hasil terbaik pada skenario pertama dengan rasio perbandingan data *training* 60% dan data *testing* 40% yang menghasilkan *accuracy* sebesar 85,55%, *precision* 78,02%, *recall* 92,0%, dan *f1-score* 84,46%. Hasil analisis menunjukkan bahwa sentimen masyarakat terhadap aplikasi *SP4N LAPOR!* cenderung negatif.

Kata kunci: *Text Mining, Analisis Sentimen, Klasifikasi Teks, Support Vector Machine, SP4N LAPOR!*

ABSTRACT

Okki Akbar Ramadhan. 201910225020. Sentiment Analysis of SP4N LAPOR! Application Reviews on Google Play Store using Support Vector Machine Algorithm. 2023.

One of the efforts made by the government to provide a means for public complaints regarding public services, the Indonesian government launched the SP4N LAPOR! application, which is used to receive complaints and aspirations from the Indonesian public regarding public services through the Play Store application. However, since its release, some users have expressed dissatisfaction with the application. Some of the complaints from SP4N LAPOR! users include server errors, slow response in addressing complaints, and so on. Based on these issues, the researcher conducted sentiment analysis on the SP4N LAPOR! application through user reviews using the Support Vector Machine (SVM) method. The data obtained after sorting and labeling consisted of 1189 review data, with 632 data labeled as negative reviews and 557 data labeled as positive reviews. The obtained data then underwent text processing, and the data was weighted using the Term Frequency-Inverse Document Frequency (TF-IDF) method. The next step involved performing 4 data splitting scenarios combined with 10-fold validation, namely 60%:40%, 70%:30%, 80%:20%, and 90%:10%, and classified using the SVM algorithm. The evaluation of the Support Vector Machine algorithm's performance revealed the best results in the first scenario, with a data training to testing ratio of 60% to 40%, resulting in an accuracy of 85.55%, precision of 78.02%, recall of 92.0%, and an F1-score of 84.46%. The analysis results indicate that public sentiment towards the SP4N LAPOR! application tends to be negative.

Keywords: Text Mining, Sentiment Analysis, Text Classification, Support Vector Machine, SP4N LAPOR!

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK**

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

Nama : Oki Akbar Ramadhan
NPM : 201910225020
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 APLIKASI SP4N LAPOR! DI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA 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 : 11 Agustus 2023
Yang Menyatakan



Oki Akbar Ramadhan

KATA PENGANTAR

Dengan memanjatkan puja dan puji syukur atas kehadiran Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga skripsi dengan judul **“Analisis Sentimen Ulasan Aplikasi *SP4N LAPOR!* di *Google Play Store* menggunakan Algoritma *Support Vector Machine*”** ini dapat terselesaikan. Skripsi ini ditulis untuk memenuhi persyaratan untuk menyelesaikan pendidikan Strata Satu (S1) pada Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.

Disadari bahwa skripsi ini tidak mungkin dapat terselesaikan tanpa adanya bantuan, dukungan, bimbingan, dan nasehat dari berbagai pihak selama penyusunan skripsi ini. Oleh karena itu, diucapkannya kata terima kasih setulus-tulusnya kepada kedua orang tua yang telah mendoakan serta memberikan dukungan dalam penulisan skripsi ini dan kata terima kasih diucapkan juga setulus-tulusnya kepada:

1. Bapak Irjen Pol. (Purn) Prof. Dr. Drs. H. 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;
4. Ibu Prima Dina Atika, S.Kom., M.Kom., selaku dosen pembimbing I penyusunan skripsi;

5. Ibu Siti Setiawati, S.Pd., M.pd., selaku dosen pembimbing II penyusunan skripsi;
6. Bapak Muhammad Yasir, S.Si., M.Kom., selaku mentor di bidang *data mining*;
7. Ibu Hafizah, S.S., M.Pd., selaku validator ahli bahasa dalam proses *labelling data*;
8. Teman-teman yang sudah membantu dan memberi dukungan dalam penulisan skripsi ini.

Disadari bahwa skripsi ini jauh dari kesempurnaan, untuk itu mohon maaf apabila terdapat kekurangan dan kesalahan baik dari pembahasan maupun penulisan. Oleh karena itu, diharapkan masukan berupa kritik dan saran yang membangun.

Jakarta, 17 Juli 2023



Oki Akbar Ramadhan

NPM: 201910225020

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI.....	iii
ABSTRAK	iv
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	6
1.3 Rumusan Masalah	7
1.4 Batasan Masalah.....	7
1.5 Tujuan Penelitian.....	8
1.6 Manfaat Penelitian.....	9
1.7 Sistematika Penulisan.....	9
BAB II TINJAUAN PUSTAKA.....	11
2.1 Penelitian Terdahulu.....	11
2.2 Tabel <i>Related Reseach / State of The Art</i>	15
2.3 Sentimen.....	23
2.4 <i>Google Play Store</i>	23

2.5	<i>SP4N LAPOR!</i>	24
2.6	<i>Text Mining</i>	25
2.7	<i>Data Mining</i>	26
2.7.1	<i>KDD (Knowledge Discovery in Databases)</i>	27
2.7.2	<i>Text Preprocessing</i>	29
2.7.3	Pembobotan <i>TF-IDF</i>	30
2.7.4	Algoritma <i>Support Vector Machine</i>	31
2.7.5	<i>Confusion Matrix</i>	36
2.8	Analisis Sentimen.....	39
2.9	<i>Python</i>	39
2.10	<i>Google Colaboratory</i>	41
BAB III METODOLOGI PENELITIAN		42
3.1	Obyek Penelitian	42
3.2	Variabel dan Definisi Operasional Variabel.....	42
3.3	Kerangka Penelitian	43
3.4	Analisis Kebutuhan Sitem.....	46
3.4.1	Spesifikasi Perangkat Keras (<i>Hardware</i>).....	46
3.4.2	Spesifikasi Perangkat Lunak (<i>Software</i>)	46
BAB IV HASIL DAN PEMBAHASAN		47
4.1	<i>Data Selection</i>	47
4.1.1	Pemilahan Data	48
4.1.2	Pelabelan Data.....	52
4.2	<i>Preprocessing</i>	55
4.2.1	<i>Stop Word Removal</i>	56
4.2.2	<i>Stemming</i>	58

4.2.3	<i>Punctuation Mark</i>	61
4.3	<i>Transformation</i>	63
4.3.1	<i>Splitting Dataset</i>	63
4.3.2	Pembobotan <i>TF-IDF</i>	68
4.4	<i>Data Mining</i>	70
4.4.1	Klasifikasi <i>Support Vector Machine</i>	71
4.5	<i>Evaluation</i>	81
4.5.1	Evaluasi Model Skenario Pembagian Data 60:40.....	81
4.5.2	Evaluasi Model Skenario Pembagian Data 70:30.....	84
4.5.3	Evaluasi Model Skenario Pembagian Data 80:20.....	86
4.5.4	Evaluasi Model Skenario Pembagian Data 90:10.....	89
4.5.5	Perbandingan Evaluasi Model Setiap Skenario	92
4.6	<i>Knowledge Presentation</i>	94
4.6.1	<i>Diagram Pie</i>	94
4.6.2	Diagram Batang	95
4.6.3	<i>Word Cloud</i>	96
BAB V PENUTUP		98
5.1	Kesimpulan.....	98
5.2	Saran	99
DAFTAR PUSTAKA		101
LAMPIRAN		104

DAFTAR TABEL

Tabel 2.1 <i>Related Reseach / State of The Art</i>	15
Tabel 2.2 <i>Confusion Matrix</i>	37
Tabel 3.1 Variabel Definisi Operasional.....	42
Tabel 4.1 Contoh Ulasan Berbahasa Non-Indonesia	49
Tabel 4.2 Contoh Ulasan yang Rancu.....	50
Tabel 4.3 Contoh Ulasan yang Tidak Relevan.....	51
Tabel 4.4 Contoh Hasil Labeling Manual Dataset.....	53
Tabel 4.5 Hasil Pelabelan Sentimen	54
Tabel 4.6 Contoh <i>Stop Word Removal</i>	58
Tabel 4.7 Contoh <i>Stemming</i>	60
Tabel 4.8 Contoh <i>Punctuation Mark</i>	62
Tabel 4.9 Skenario Pembagian Data	64
Tabel 4.10 <i>Stratified Shuffle Split</i>	64
Tabel 4.11 Sampel Data yang di Vektorisasi	70
Tabel 4.12 Hasil Akurasi Klasifikasi <i>SVM</i>	71
Tabel 4.13 Nilai <i>TF-IDF</i> D1-D4.....	73
Tabel 4.14 Tabel Fungsi Kernel.....	73
Tabel 4.15 Hasil Perhitungan Fungsi kernel data D1-D3	75
Tabel 4.16 Hasil Perhitungan Fungsi Kernel data D4.....	76
Tabel 4.17 Tabel Matriks	77
Tabel 4.18 Hasil Perhitungan Nilai <i>Error</i>	77
Tabel 4.19 Hasil Perhitungan <i>Delta Alpha</i>	78
Tabel 4.20 Nilai <i>Alpha</i> Baru	79
Tabel 4.21 Nilai <i>TF-IDF</i> D4	80
Tabel 4.22 Hasil Pembobotan data D4 dengan data Latih	80
Tabel 4. 23 <i>Confusion Matrix</i> Pembagian Data 60:40.....	82
Tabel 4.24 <i>Confusion Matrix</i> Pembagian Data 70:30.....	85
Tabel 4.25 <i>Confusion Matrix</i> Pembagian Data 80:20.....	87

Tabel 4.26 Hasil Confusion Matrix.....	90
Tabel 4.27 Evaluasi Model Setiap Skenario	92



DAFTAR GAMBAR

Gambar 1.1 Aplikasi <i>SP4N LAPOR!</i>	4
Gambar 2.1 <i>The data processing pipeline</i>	26
Gambar 2.2 Tahapan <i>KDD</i>	28
Gambar 2.3 <i>Hyperplane</i> pada <i>SVM</i>	32
Gambar 2.4 <i>Support Vector</i> dan <i>Margin</i>	34
Gambar 3.1 Kerangka Penelitian	43
Gambar 4.1 Proses <i>Scraping</i> Ulasan dengan <i>Python</i>	47
Gambar 4.2 Dataset Ulasan Aplikasi <i>SP4N LAPOR!</i>	48
Gambar 4.3 Dataset dengan Bahasa Non-Indonesia.....	48
Gambar 4.4 Dataset Bersih dari Bahasa Non-Indonesia.....	49
Gambar 4.5 Dataset dengan Ulasan yang Rancu	49
Gambar 4.6 Dataset Bersih dari Ulasan yang Rancu	50
Gambar 4.7 Dataset dengan Ulasan yang Tidak Relevan.....	50
Gambar 4.8 Dataset Bersih dari Ulasan yang Tidak Relevan.....	51
Gambar 4.9 <i>New Dataset</i> Ulasan.....	52
Gambar 4.10 Prosedur Pelabelan Dataset.....	52
Gambar 4.11 Validasi Labeling Dataset	54
Gambar 4.12 Dataset yang Sudah diberi Label.....	54
Gambar 4.13 <i>Flowchart Preprocessing</i>	56
Gambar 4.14 <i>Flowchart Stop Word Removal</i>	57
Gambar 4.15 Dataset Sebelum <i>Stop Word</i>	57
Gambar 4.16 Proses <i>Stop Word</i> pada <i>Python</i>	58
Gambar 4.17 Dataset Hasil <i>Stop Word</i>	58
Gambar 4.18 Flowchart Stemming	59
Gambar 4.19 Dataset Sebelum Stemming	59
Gambar 4.20 Proses <i>Stemming</i> menggunakan <i>Python</i>	60
Gambar 4.21 Dataset Hasil <i>Stemming</i>	60
Gambar 4.22 <i>Flowchart Punctuation Mark</i>	61

Gambar 4.23 Dataset Sebelum <i>Punctuation Mark</i>	61
Gambar 4.24 Proses <i>Punctuation Mark</i> menggunakan <i>Python</i>	62
Gambar 4.25 Dataset <i>Clean</i>	62
Gambar 4.26 Proses <i>Splitting</i> Dataset 60:40 pada <i>Python</i>	65
Gambar 4.27 Dataset rasio 60%:40%	65
Gambar 4.28 Proses <i>Splitting</i> Data 70:30 pada <i>Python</i>	66
Gambar 4.29 Dataset rasio 70%:30%	66
Gambar 4.30 Proses <i>Splitting</i> Data 80:20 pada <i>Python</i>	67
Gambar 4.31 Dataset rasio 80%:20%	67
Gambar 4.32 Proses <i>Splitting</i> Data 90:10 pada <i>Python</i>	68
Gambar 4.33 Dataset rasio 90%:10%	68
Gambar 4.34 Hasil <i>TF-IDF</i>	69
Gambar 4.35 Hasil Klasifikasi <i>SVM</i>	72
Gambar 4.36 Plot <i>Confusion Matrix</i> Pembagian Data 60:40.....	82
Gambar 4.37 Plot <i>Confusion Matrix</i> Pembagian Data 70:30.....	84
Gambar 4.38 Plot <i>Confusion Matrix</i> Pembagian Data 80:20.....	87
Gambar 4.39 Plot <i>Confusion Matrix</i> Pembagian Data 90:10.....	89
Gambar 4.40 Diagram <i>Pie</i> Distribusi Sentimen.....	95
Gambar 4.41 Diagram Batang <i>Top 5</i> Frekuensi Kata.....	96
Gambar 4.42 <i>Word Cloud</i> Ulasan <i>SP4N LAPOR!</i>	97
Gambar 4.43 <i>Word Cloud</i> Ulasan Negatif dan Positif.....	97

DAFTAR LAMPIRAN

1. Surat Rekomendasi dari Pembimbing
2. Form Validasi Instrumen
3. Lembar Plagiarisme
4. Biodata Mahasiswa
5. Kartu Konsultasi Bimbingan Skripsi

