



**ANALISIS SENTIMEN PADA ULASAN APLIKASI JAKARTA
TERKINI (JAKI) DI GOOGLE PLAY STORE MENGGUNAKAN
METODE SUPPORT VECTOR MACHINE**

SKRIPSI

**DANIEL DWI ERYANTO MANURUNG
1810511035**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN
JAKARTA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
2022**



**ANALISIS SENTIMEN PADA ULASAN APLIKASI JAKARTA
TERKINI (JAKI) DI GOOGLE PLAY STORE
MENGUNAKAN METODE SUPPORT VECTOR MACHINE**

SKRIPSI

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

**DANIEL DWI ERYANTO MANURUNG
1810511035**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN
JAKARTA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
2022**

PERNYATAAN ORISINALITAS

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

Nama : Daniel Dwi Eryanto Manurung

NIM : 1810511035

Tanggal : 22 Juli 2022

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

Jakarta, 22 Juli 2022

Yang menyatakan,



(Daniel Dwi Eryanto Manurung)

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama : Daniel Dwi Eryanto Manurung

NIM : 1810511035

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:

***Analisis Sentimen pada Ulasan Aplikasi Jakarta Terkini (JAKI) di Google
Play Store Menggunakan Metode Support Vector Machine***

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 Tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 22 Juli 2022

Yang menyatakan,



(Daniel Dwi Eryanto Manurung)

LEMBAR PENGESAHAN

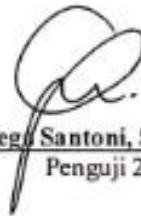
Dengan ini dinyatakan bahwa Tugas Akhir berikut:

Nama : Daniel Dwi Eryanto Manurung
NIM : 1810511035
Program Studi : S1 Informatika 2018
Judul : Analisis Sentimen pada Ulasan Aplikasi Jakarta Terkini (JAKI)
di *Google Play Store* Menggunakan Metode *Support Vector Machine*

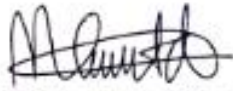
Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Henki Bayu Seta, S.Kom., MTI.
Penguji 1



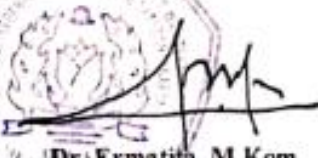
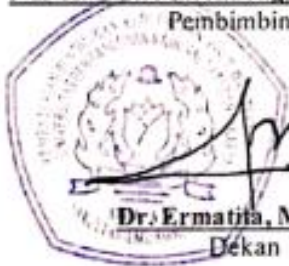
Mayanda Megi Santoni, S.Komp., M.Kom.
Penguji 2



Nur Hafifah Matondang, S.Kom., MTI.
Pembimbing 1



Desta Sandya Prasvita, S.Kom., M.Kom.
Pembimbing 2

Dr. Ermatifa, M.Kom.
Dekan



Desta Sandya Prasvita, S.Kom., M.Kom.
Ketua Program Studi

Ditetapkan di : Jakarta
Tanggal pengesahan : 08 Juli 2022



ANALISIS SENTIMEN PADA ULASAN APLIKASI JAKARTA TERKINI (JAKI) DI GOOGLE PLAY STORE MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

Daniel Dwi Eryanto Manurung

ABSTRAK

Seiring perkembangan waktu, aplikasi Jakarta Terkini (JAKI) terus berkembang dengan baik. Namun beberapa pengguna mengeluh ketidakpuasan mereka terhadap layanan yang disediakan oleh aplikasi yang dilihat melalui ulasan pengguna. Beberapa jenis keluhan pengguna JAKI diantaranya sertifikat vaksin yang tidak tersedia, kuota vaksin selalu penuh, kesulitan dalam pengisian data dan sebagainya. Berdasarkan masalah tersebut, penelitian ini melakukan analisis sentimen pada Aplikasi JAKI melalui ulasan pengguna untuk memberikan informasi kepada masyarakat terkait kinerja aplikasi dengan menggunakan metode *Support Vector Machine* (SVM) dan metode *chi-square* untuk seleksi fitur. Jumlah dataset yang diperoleh sebanyak 1000 data dan belum terlabelkan. Dalam mengklasifikasikan data yang diperoleh perlu dilakukan pelabelan data dan pembersihan data terlebih dahulu sebelum tahapan text processing, kemudian data diberikan bobot pada setiap kata dengan *Term Frequency-Inverse Document Frequency* (TF-IDF) yang akan dijadikan sebagai fitur, kemudian dilakukan seleksi fitur dengan *chi-square*. Tahapan selanjutnya dilakukan pembagian data sebesar 80% 20% dan diklasifikasikan dengan metode SVM. Hasil yang diperoleh dari penelitian ini adalah fitur terseleksi sebanyak 66 fitur dan pemodelan menggunakan kernel RBF, $C = 40$ dan $\gamma = 0.1$ dari SVM, diperoleh bahwa 120 data uji (*testing*) mendapatkan *accuracy* sebesar 97%, *precision* sebesar 100%, *recall* sebesar 93%, dan *specificity* sebesar 100%.

Kata Kunci: Analisis Sentimen, Klasifikasi, JAKI, SVM, *Chi-Square*.

SENTIMENT ANALYSIS OF JAKARTA TERKINI (JAKI) APPLICATION REVIEWS ON GOOGLE PLAY STORE USING SUPPORT VECTOR MACHINE METHOD

Daniel Dwi Eryanto Manurung

ABSTRACT

Over time, the Jakarta Terkini (JAKI) application continues to develop well. However some users have complained of their dissatisfaction with the services provided by the app which is seen through the comments review. Several types of complaints from JAKI users include vaccine certificates that are not available, vaccine quotas are always full, difficulties in data entry and so on. Based on these problems, this study conducted a sentiment analysis on the JAKI application through commentary reviews to provide information to the public regarding application performance using the Support Vector Machine (SVM) method and the chi-square method for feature selection. The number of datasets obtained is 1000 data and has not been labeled. In classifying the data obtained, it is necessary to label the data and clean the data first before the text processing stage, then the data is given a weight for each word with Term Frequency–Inverse Document Frequency (TF-IDF) which will be used as a feature, then feature selection is carried out with chi -square. The next stage is the distribution of data by 80% 20% and classified by the SVM method. The results obtained from this study are 66 features selected and modeling using the RBF kernel, $C = 40$ and $\gamma = 0.1$ from SVM, it is found that 120 test data (testing) get 97% accuracy, 100% precision, 93% recall, and a specificity of 100%.

Keyword: Sentiment analysis, Classification, JAKI, SVM, Chi-Square.

KATA PENGANTAR

Puji serta syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa atas segala limpahan rahmat, karunia dan ridha-Nya, sehingga Skripsi dengan judul “Analisis Sentimen pada Ulasan Aplikasi Jakarta Terkini (JAKI) di *Google Play Store* Menggunakan Metode *Support Vector Machine*” berhasil diselesaikan. Tak lupa, penulis ingin mengucapkan banyak terima kasih kepada:

1. Tuhan Yesus Kristus, yang sudah memberikan kesehatan, kekuatan, dan berkat kepada penulis dalam menyelesaikan skripsi ini.
2. Kedua orang tua penulis, Mustar (bapak) dan Erina (ibu), serta keluarga yang selalu memberikan doa dan motivasi sehingga penulis dapat menyelesaikan skripsi ini.
3. Ibu Nur Hafifah Matondang, S.Kom., M.T.I. dan Bapak Desta Sandya Prasvita, S.Kom, M.Kom., selaku dosen pembimbing yang telah membimbing serta memberikan saran yang bermanfaat.
4. Ibu Dr. Ermatita, M. Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
5. Bapak Desta Sandya Prasvita, S.Kom, M.Kom., selaku Ketua Program Studi Informatika Universitas Pembangunan Nasional Veteran Jakarta.
6. Bapak Bambang Tri Wahyono, S.Kom., M.Si. selaku dosen pembimbing akademik.
7. Bapak/Ibu dosen Informatika Universitas Pembangunan Nasional Veteran Jakarta yang telah memberikan ilmu yang bermanfaat.
8. Amang Pdt. Paulus Simamora, Kak Tama, Irfan, Hilda, Bobby, Hari, Cindy, Warung Jabar serta teman-teman seperjuangan yang sudah membantu penulis dalam menyelesaikan skripsi ini.
9. Seluruh pihak yang terlibat dalam kelancaran pembuatan proposal tugas akhir ini dan yang belum disebutkan di atas, penulis ucapkan terima kasih.

Akhir kata penulis ucapkan terima kasih dan semoga skripsi ini dapat bermanfaat bagi pembaca.

Depok, 25 Mei 2022

Penulis

DAFTAR ISI

PERNYATAAN ORISINALITAS	ii
PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	iii
LEMBAR PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Ruang Lingkup	3
1.6 Luaran Yang Diharapkan	4
1.7 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Analisis Sentimen	6
2.2 <i>Text Mining</i>	6
2.3 Jakarta Terkini	6
2.4 <i>Fleiss Kappa</i>	7

2.5	<i>Text Processing</i>	8
2.5.1	<i>Case Folding</i>	9
2.5.2	<i>Data Cleaning</i>	9
2.5.3	<i>Spelling Normalization</i>	9
2.5.4	<i>Stemming</i>	9
2.5.5	<i>Stopword Removal</i>	9
2.5.6	<i>Tokenizing</i>	9
2.6	Pembobotan Kata.....	10
2.6.1	TF (<i>Term Frequency</i>).....	10
2.6.2	IDF (<i>Inverse Document Frequency</i>)	10
2.7	<i>Chi-Square</i>	11
2.8	<i>Support Vector Machine</i>	11
2.9	Evaluasi	15
2.10	Penelitian Terkait.....	16
BAB III METODE PENELITIAN.....		20
3.1	Kerangka Pikir.....	20
3.1.1	Identifikasi Masalah	21
3.1.2	Studi Literatur	21
3.1.3	<i>Data Scraping</i>	21
3.1.4	Pelabelan Data.....	22
3.1.5	Praproses Data.....	22
3.1.6	Pembobotan Kata	23
3.1.7	Seleksi Fitur	24
3.1.8	Pembagian Data	24
3.1.9	Klasifikasi	24
3.1.10	Evaluasi.....	24

3.1.11 Analisis Hasil Klasifikasi	25
3.2 Perangkat Penelitian	25
BAB IV HASIL DAN PEMBAHASAN	26
4.1 Pengumpulan Data.....	26
4.2 Pelabelan Data	26
4.3 Praproses.....	31
4.3.1 <i>Case Folding</i>	32
4.3.2 <i>Data Cleaning</i>	33
4.3.3 <i>Spelling Normalization</i>	34
4.3.4 <i>Stemming</i>	35
4.3.5 <i>Stopwords Removal</i>	36
4.3.6 <i>Tokenizing</i>	38
4.4 Pembobotan TF-IDF.....	38
4.5 Seleksi Fitur <i>Chi-square</i>	40
4.6 Klasifikasi <i>Support Vector Machine</i>	42
4.6.1 Pembagian Data	42
4.6.2 Pemodelan Data	42
4.7 Evaluasi	44
4.8 Analisis Hasil Klasifikasi	46
4.8.1 <i>Wordcloud</i> sentimen positif terhadap aplikasi Jakarta Terkini (JAKI) 46	
4.8.2 <i>Wordcloud</i> sentimen negatif terhadap aplikasi Jakarta Terkini (JAKI) 48	
BAB V PENUTUP.....	50
5.1 Kesimpulan.....	50
5.2 Saran	50

DAFTAR PUSTAKA	52
RIWAYAT HIDUP	55
LAMPIRAN	56

DAFTAR TABEL

Tabel 2. 1 Interpretasi Kappa	8
Tabel 2. 2 Confusion Matrix	15
Tabel 4. 1 Ulasan dengan label hasil.....	27
Tabel 4. 2 Gabungan hasil penilaian annotator.....	29
Tabel 4. 3 Hasil pencarian nilai Pi	29
Tabel 4. 4 Case Folding	32
Tabel 4. 5 Data Cleaning.....	33
Tabel 4. 6 Spelling Normalization	34
Tabel 4. 7 Stemming	36
Tabel 4. 8 Stopwords Removal	37
Tabel 4. 9 Tokenizing	38
Tabel 4. 10 Perhitungan TF-IDF	39
Tabel 4. 11 Contoh TF-IDF	40
Tabel 4. 12 Hasil perhitungan chi-square	41
Tabel 4. 13 Jumlah fitur setelah diseleksi	42
Tabel 4. 14 Pembagian Data	42
Tabel 4. 15 Proses GridSearchCV	43
Tabel 4. 16 Hasil Pencarian Parameter Terbaik.....	44
Tabel 4. 17 Hasil Confusion Matrix Model	44
Tabel 4. 18 Data input wordcloud berlabel positif.....	47
Tabel 4. 19 Data input wordcloud berlabel negatif.....	48

DAFTAR GAMBAR

Gambar 2. 1 Linear Separable.....	11
Gambar 2. 2 non-Linear Separable	12
Gambar 2. 3 Hyperplane Class -1 dan +1	12
Gambar 2. 4 Polynomial kernel derajat satu	14
Gambar 2. 5 Polynomial kernel derajat enam.....	14
Gambar 2. 6 Penyebaran data tidak merata.....	15
Gambar 3. 1 Tahapan Penelitian	20
Gambar 4. 1 Hasil dari Scraping data pada Google Play Store	26
Gambar 4. 2 Diagram pie label	31
Gambar 4. 3 Wordcloud sentimen positif	47
Gambar 4. 4 Wordcloud sentiment negatif	49

DAFTAR LAMPIRAN

Lampiran 1 Kamus Normalisasi.....	57
Lampiran 2 Stopwords Bahasa Indonesia	71
Lampiran 3 Tabel Chi-Square.....	76
Lampiran 4 Perhitungan sample TF-IDF	77
Lampiran 5 Fitur yang digunakan.....	79
Lampiran 6 Experiment Tuning Parameter.....	80
Lampiran 7 Similarity Index	85