



**ANALISIS SENTIMEN PENGGUNA APLIKASI DANA
BERDASARKAN ULASAN PADA GOOGLE PLAY STORE
MENGUNAKAN METODE SUPPORT VECTOR MACHINE**

SKRIPSI

**ABITDAVY ATHALLAH MUHAMMAD
1810511079**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN
JAKARTA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
2021**



**ANALISIS SENTIMEN PENGGUNA APLIKASI DANA
BERDASARKAN ULASAN PADA GOOGLE PLAY STORE
MENGUNAKAN METODE SUPPORT VECTOR MACHINE**

SKRIPSI

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

ABITDAVY ATHALLAH MUHAMMAD

1810511079

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN

JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI INFORMATIKA

2021

PERNYATAAN ORISINALITAS

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

Nama : Abitdavy Athallah Muhammad

NIM : 1810511079

Tanggal : 17 Juli 2022

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

Jakarta, 17 Juli, 2022

Yang menyatakan,



(Abitdavy Athallah Muhammad)

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini :

Nama : Abitdavy Athallah Muhammad

NIM : 1810511079

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 PENGGUNA APLIKASI DANA BERDASARKAN
ULASAN PADA 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 : Depok

Pada tanggal : 17 Juli 2022

Yang menyatakan,



(Abitdavy Athallah Muhammad)

PENGESAHAN

Skripsi diajukan oleh :

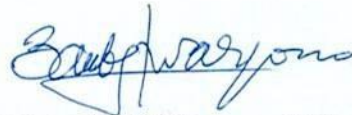
Nama : Abitdavy Athallah Muhammad
NIM : 1810511079
Program Studi : S1 Informatika
Judul Skripsi : Analisis Sentimen Pengguna Aplikasi DANA Berdasarkan Ulasan Pada Google Play Store Menggunakan Metode Support Vector Machine

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



Jayanta, S.Kom., M.Si

Penguji 1




Bambang Tri Wahyono, S.Kom. M.Si

Penguji 2



Dr. Ermatita, M. Kom.

Pembimbing 1



Desta Sandya Prasvita, S.Kom, M.Kom

Pembimbing 2



Dr. Ermatita, M. Kom.

Dekan



Desta Sandya Prasvita, S.Kom, M.Kom

Kaprodi



Ditetapkan di : Jakarta

Tanggal Ujian : Senin, 13 Juli 2022

ANALISIS SENTIMEN PENGGUNA APLIKASI DANA BERDASARKAN ULASAN PADA GOOGLE PLAY STORE MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

Abitdavy Athallah Muhammad

ABSTRAK

Perkembangan teknologi pada era digital ini sangat pesat, salah satu bidang yang sangat berkembang ialah finansial. Produk dari finansial digital salah satunya adalah uang elektronik. Beberapa perusahaan dari aplikasi fintech mulai bermunculan, dan salah satunya adalah DANA. Dari kemudahan yang ditawarkan, banyak orang yang mulai menggunakan aplikasi DANA. Dan sebab semakin banyak pengguna, maka semakin banyak juga pengguna yang memberikan pengalaman menggunakannya berupa kepuasan dan keluhan dalam ulasan pada *Google Play Store*. Ulasan dari pengguna dapat menjadi bahan untuk dianalisa agar menjadi masukan atau *feedback* kepada perusahaan yang berkepentingan sehingga dapat menjadi perbaikan kedepannya. Pada penelitian ini data didapat dari ulasan pengguna aplikasi dompet digital DANA pada *Google Play Store*. Data ulasan dibagi menjadi 2 kategori yaitu data bersifat positif dan negatif berdasarkan pelabelan manual oleh 3 anatator. Metode *support vector machine* dan seleksi fitur *chi square* digunakan dalam pembuatan model. Hasil evaluasi diperoleh akurasi sebesar 87,58%, presisi sebesar 91,20%, dan recall sebesar 90,21% untuk model SVM, sedangkan untuk model SVM – *chi square* diperoleh akurasi sebesar 89,41%, presisi sebesar 93,29%, dan recall sebesar 90,76%.

Kata Kunci : Klasifikasi, *Support Vector Machine*, Sentimen

ABSTRACT

The development of technology in this era is very rapid, one of the fields that is very developed are finance. One of the digital financial products are electronic money. Several companies from fintech applications are starting to emerge, and one of them is DANA. From the convenience offered, many oepople have started using apllication DANA. And because there are more users, more and more users provide the experience of using it in the form of satisfaction and complaints in reviews on the Google Play Stores. Reviews from users can be used as material for analysis so that they can be input or feedback to interested companies so that they can be improved in the future. In this study the data obtained from user reviews of the DANA digital wallet application on the Google Play Store. The data review is divied into 2 categories, namely positive and negative data based on manual labeling by 3 anatators. Support Vector Machine method and chi square feature selection are used in model making. The evaluation results are 87,58%, precision is 91,20%, and recall is 90,21% for the SVM model, while for the SVM – chi square model, the accuracy is 89,41%, precision is 93,29% and recall by 90,76%.

Keywords : *Classification, Support Vector Machine, Sentiment*

KATA PENGANTAR

Puji syukur saya ucapkan kepada Allah SWT atas karunia dan ridha-Nya yang diberikan, sehingga Skripsi dengan judul “Analisis Sentimen Pengguna Aplikasi DANA Berdasarkan Ulasan pada *Google Play Store* Menggunakan Metode *Support Vector Machine*” dapat dikerjakan tepat dalam waktunya. Penulis ingin mengucapkan terima kasih banyak kepada:

1. Ibu Dr. Ermatita, M.Kom. Selaku Dekan Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.
2. Bapak Desta Sandya Prasvita, S.Kom, M.Kom, selaku Ketua Program Studi Informatika Universitas Pembangunan Nasional Veteran Jakarta.
3. Kedua orang tua serta keluarga yang terus mendoakan dan memberi dukungan dalam berbagai hal
4. Ibu Dr. Ermatita, M.Kom. dan Bapak Desta Sandya Prasvita, S.Kom, M.Kom selaku dosen pembimbing tugas akhir yang telah bersedia meluangkan waktunya dan memberi arahan.
5. Seluruh dosen dan pegawai Fakultas Ilmu Komputer.
6. Orang terdekat penulis dan teman – teman Informatika angkatan 2018 yang sudah membantu dan memotivasi penulis untuk menyelesaikan tugas akhir ini, terutama Aditya Samiaji.

Penulis menyadari masih banyak memiliki kekurangan dan masih jauh dari kata sempurna dalam penyusunan skripsi ini, sehingga penulis mengharapkan saran dan kritik yang bersifat membangun untuk penulis. Akhir kata penulis ucapkan terima kasih.

Jakarta, 25 Juni
2022.



(Abitdavy Athallah
Muhammad)

DAFTAR ISI

PERNYATAAN ORISINALITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	iv
LEMBAR PENGESAHAN	v
ABSTRAK	vi
ABSTRACT	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xiv
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1 <i>Data Mining</i>	7
2.2 <i>Machine Learning</i>	7
2.2.1 <i>Supervised Learning</i>	8
2.2.2 <i>Unsupervised Learning</i>	8
2.2.3 <i>Reinforcement Learning</i>	8
2.3 Analisis Sentimen	8
2.4 Web Scraping	9
2.5 DANA	9
2.5 Klasifikasi Teks	10
2.5.1 <i>Preprocessing</i>	10
2.5.2 Evaluasi Model Klasifikasi.....	10
2.6 Pembobotan Kata.....	12
2.6.1 Term Frequency Inverse Document Frequency (TF – IDF).....	12

2.7	Seleksi Fitur.....	13
2.7.1	Chi-Square.....	13
2.8	Support Vector Machine	14
2.9	Penelitian Terkait.....	17
BAB III METODOLOGI PENELITIAN		19
3.1	Tahapan Penelitian.....	19
3.2	Studi Pustaka	20
3.3	Identifikasi Masalah.....	20
3.4	Pengumpulan Data.....	20
3.5	Pelabelan Data	21
3.6	Data <i>Preprocessing</i>	22
3.6.1	<i>Case Folding</i>	22
3.6.2	<i>Cleansing</i>	23
3.6.3	<i>Tokenizing</i>	24
3.6.4	<i>Normalization</i>	24
3.6.5	<i>Filtering</i> atau <i>Stopword Removing</i>	25
3.6.6	<i>Stemming</i>	25
3.7	Pembobotan Kata.....	26
3.8	Seleksi Fitur.....	27
3.9	Klasifikasi	28
3.10	Evaluasi Hasil Klasifikasi	29
3.11	Hasil.....	29
3.12	Alat Penelitian	29
Bab IV Hasil dan Pembahasan.....		31
4.1	Pengumpulan Data.....	31
4.2	Pelabelan Data	33
4.3	<i>Preprocessing Data</i>	35
4.3.1	<i>Case Folding</i>	35
4.3.2	<i>Cleansing</i>	36
4.3.3	<i>Tokenizing</i>	37
4.3.4.	<i>Normalization</i>	38
4.3.5.	<i>Filtering</i> atau <i>stopword removing</i>	39
4.3.6	<i>Stemming</i>	39
4.4	Pembobotan Kata.....	40

4.5	Seleksi Fitur Chi Square.....	43
4.6	Klasifikasi	45
4.6.1	Pembagian Data	46
4.6.2	Pemodelan Tanpa Seleksi Fitur (SVM).....	46
4.6.3	Pemodelan Dengan Seleksi Fitur (SVM – <i>chi square</i>).....	47
4.7	Evaluasi dan Analisis Hasil Evaluasi.....	48
4.7.1	Evaluasi Model SVM tanpa Seleksi Fitur	48
4.7.2	Evaluasi Model SVM dengan Seleksi Fitur <i>chi square</i>	49
4.7.3	Analisis Hasil Evaluasi.....	50
4.8	Visualisasi	51
4.8.1	<i>WordCloud</i> Sentimen Negatif.....	51
4.8.2	<i>WordCloud</i> Sentimen Positif	53
BAB V PENUTUPAN		56
5.1	Kesimpulan.....	56
5.2	Saran	57
DAFTAR PUSTAKA		58
RIWAYAT HIDUP.....		61
LAMPIRAN		62
	Lampiran 1 Kamus Normalisasi <i>Slangword</i>	63
	Lampiran 2 Kamus <i>Stopword</i>	79
	Lampiran 3 Script <i>scraping google-play-scrapers</i>	86
	Lampiran 4 Script data <i>preprocessing</i>	87
	Lampiran 5 Script Pembobotan Kata	90
	Lampiran 6 Seleksi Fitur <i>Chi Square</i>	91
	Lampiran 7 Script klasifikasi SVM tanpa seleksi fitur dan evaluasi.....	92
	Lampiran 8 Script Klasifikasi SVM Seleksi Fitur <i>Chi Square</i> dan Evaluasi.....	93
	Lampiran 9 Script <i>Wordcloud</i>	95
	Lampiran 10 Similiarity Index	96

DAFTAR TABEL

Tabel 2. 1. <i>Confusion Matrix</i>	11
Tabel 3. 1. Tabel kontingensi <i>observed value</i>	27
Tabel 3. 2. Tabel kontingensi <i>expected value</i>	28
Tabel 3. 3. Perangkat keras	29
Tabel 3. 4. Jadwal penelitian	30
Tabel 4. 1. Hasil <i>scraping</i> data ulasan	33
Tabel 4. 2. Hasil pelabelan data.....	33
Tabel 4. 3. Hasil <i>case folding</i>	35
Tabel 4. 4. Hasil data cleansing	36
Tabel 4. 5. Hasil <i>tokenizing</i>	37
Tabel 4. 6. Hasil data <i>normalization</i>	38
Tabel 4. 7. Hasil <i>filtering</i> atau <i>stopword removing</i>	39
Tabel 4. 8. Hasil <i>stemming</i>	40
Tabel 4. 9. Sampel dokumen pembobotan kata.....	40
Tabel 4. 10. Perhitungan pembobotan kata TF - IDF	41
Tabel 4. 11. Sampel dokumen seleksi fitur	43
Tabel 4. 12. Simulasi perhitungan nilai <i>chi square</i> fitur “saldo”	44
Tabel 4. 13. Hasil seleksi fitur <i>chi square</i>	45
Tabel 4. 14. Hasil pembagian data.....	46
Tabel 4. 15. Hasil <i>GridSearch</i> SVM.....	47
Tabel 4. 16. Hasil <i>GridSearch</i> SVM – <i>chi square</i>	48
Tabel 4. 17. Evaluasi model SVM.....	48
Tabel 4. 18. Evaluasi model SVM – <i>chi square</i>	49
Tabel 4. 19. Perbandingan hasil evaluasi model.	50
Tabel 4. 20. Kemunculan kata negatif	52
Tabel 4. 21. Asosiasi kata negatif.....	53
Tabel 4. 22. Kemunculan kata positif	54
Tabel 4. 23. Asosiasi kata positif.....	55

DAFTAR GAMBAR

Gambar 2. 1. Ilustrasi Metode <i>Support Vector Machine</i>	15
Gambar 2. 2. Konsep Kernel SVM	16
Gambar 3. 1. Tahapan penelitian	19
Gambar 3. 2. Proses <i>scraping</i>	21
Gambar 3. 3. Proses <i>case folding</i>	23
Gambar 3. 4. Proses <i>cleansing</i>	23
Gambar 3. 5. Proses <i>tokenizing</i>	24
Gambar 3. 6. Proses <i>normalization</i>	24
Gambar 3. 7. Proses <i>stemming</i>	26
Gambar 3. 8. Proses pembobotan kata TF - IDF	26
Gambar 3. 9. Proses seleksi fitur <i>chi square</i>	27
Gambar 4. 1. Program <i>google-play-scrapers</i>	31
Gambar 4. 2. ID aplikasi DANA	32
Gambar 4. 3. Hasil pengonversian data aplikasi DANA	32
Gambar 4. 4. Visualiasi hasil pelabelan data	35
Gambar 4. 5. Grafik performa model	51
Gambar 4. 6. <i>Wordcloud</i> sentimen negatif	52
Gambar 4. 7. <i>Wordcloud</i> sentimen positif	54