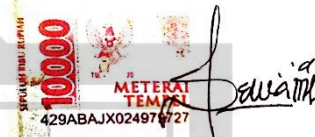


## PERNYATAAN KEASLIAN TULISAN

Saya menyatakan bahwa skripsi ini bebas plagiat, dan apabila di kemudian hari terbukti terdapat plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.

Bojonegoro, 15 Juni 2024



Selvia Dila Pevita

NIM : 2120200545



# UNUGIRI

## LEMBAR PERSETUJUAN


Nama : Selvia Dila Pevita

NIM : 2120200545

Judul : Analisis Sentimen Ulasan Aplikasi *Threads* di *Google Play Store*  
Menggunakan Metode *Support Vector Machine*


Telah disetujui dan dinyatakan memenuhi syarat untuk diajukan dalam sidang skripsi.

Pembimbing I

  
Muhammad Jauhar Vikri, M.Kom.

NIDN : 0712078803

Pembimbing II

  
Sunu Wahyudhi, M.Pd.

NIDN : 0709058902

# UNUGIRI

## LEMBAR PENGESAHAN

Nama : Selvia Dila Pevita

NIM : 2120200545

Judul : Analisis Sentimen Ulasan Aplikasi *Threads* di *Google Play Store*  
Menggunakan Metode *Support Vector Machine*

Telah dipertahankan di hadapan penguji pada tanggal 20 Juni 2024.

Dewan Penguji

Penguji I



Zakki Alawi, S.Kom., M.M.

NIDN : 0709068906

Tim Pembimbing

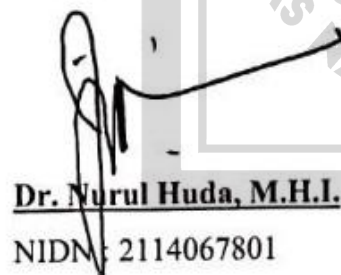
Pembimbing I



Muhammad Jauhar Vikri, M.Kom.

NIDN : 0712078803

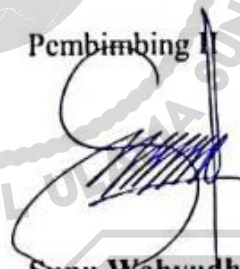
Penguji II



Dr. Nurul Huda, M.H.I.

NIDN : 2114067801

Pembimbing II



Sunu Wahyudhi, M.Pd.

NIDN : 0709058902

Mengetahui,

Dekan Fakultas Sains dan Teknologi



Muhammad Jauhar Vikri, M.Kom.

NIDN : 0712078803

Mengetahui,

Ketua Program Studi



Uca Pradema Sanjaya, M.Kom.

NIDN : 0729128903

## MOTTO DAN PERSEMBAHAN

### MOTTO

Everyone has their own process. So, just trust Your God and yourself.

~ Selvia Dila Pevita ~

### PERSEMBAHAN

Persembahan tugas akhir ini dan rasa terima kasih saya ucapkan untuk :

1. Tuhan Yang Maha Esa yang telah melancarkan semua proses mulai dari bimbingan, seminar proposal, pengerjaan skripsi, dan sidang.
2. Keluarga tercinta khususnya Bapak Jatomo, Ibu Anik Suprapti, dan adik Satriya Bilda Andrasta yang telah memberikan dukungan mental dan finansial sehingga penulis dapat melewati proses pengerjaan tugas akhir.
3. Sahabat seperjuangan yang selalu berbagi semangat, berbagi ilmu tanpa mengenal pelit, dan bersedia mendengarkan keluh kesah saat pengerjaan tugas akhir ini.
4. Terakhir, terima kasih kepada diri sendiri yang telah mampu berusaha dan berjuang sejauh ini, ingat perjalanan ini masih panjang.

UNUGIRI

## KATA PENGANTAR

Puji syukur kami panjatkan kehadirat Allah SWT yang telah memberikan rahmat dan ridhoNya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Ulasan Aplikasi *Threads* di *Google Play Store* Menggunakan Metode *Support Vector Machine*”. Dalam penulisan skripsi ini, penulis mendapatkan banyak motivasi, inspirasi, dukungan, dan bimbingan dari berbagai pihak sehingga penulis dapat menyelesaikan skripsi ini tepat waktu. Penulis ingin menggunakan kesempatan ini untuk mengucapkan terima kasih kepada :

1. Bapak M. Jauharul Ma'arif, M. Pd. I. selaku Rektor Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
2. Bapak Muhammad Jauhar Vikri, M. Kom. selaku Dekan Fakultas Sains dan Teknologi sekaligus Dosen Pembimbing 1 yang selalu membimbing penulis dan memberi masukan dalam penulisan skripsi ini.
3. Bapak Ucta Pradema Sanjaya, M. Kom. selaku Kaprodi Teknik Informatika Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
4. Bapak Sunu Wahyudhi, M.Pd. selaku Dosen Pembimbing 2 yang selalu membimbing penulis dan memberi masukan dalam penulisan skripsi ini.
5. Kedua orangtua dan keluarga besar yang selalu mendoakan dan mendukung, dan memotivasi penulis dalam menyelesaikan skripsi ini.
6. Sahabat dan teman-teman yang selalu mendukung penulis dalam menyelesaikan skripsi ini
7. Semua pihak yang telah membantu yang tidak dapat penulis sebutkan namanya satu per satu.

Penulis menyadari, bahwa skripsi ini masih banyak kekurangan dan belum sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan agar penulis dapat menyempurnakan skripsi ini. Penulis berharap skripsi ini dapat bermanfaat bagi pembaca dan berbagai pihak.

Bojonegoro, 15 Juni 2024

Selvia Dila Pevita

## ABSTRACT

*Pevita, Selvia Dila. Sentiment Analysis of Threads Application Reviews on the Google Play Store Using the Support Vector Machine Method. Thesis. Department of Informatics Engineering, Faculty of Science and Technology, Sunan Giri Nahdlatul Ulama University. First Supervisor Muhammad Jauhar Vikri, M.Kom. and Second Supervisor Sunu Wahyudhi, M.Pd.*

*This research aims to conduct sentiment analysis of Threads application reviews on the Google Play Store using the Support Vector Machine (SVM) method with two sentiment labels, namely positive and negative. The dataset used amounted to 2872 reviews. The research process includes several steps, namely Scraping data carried out using the Python programming language through Google Colab, then Preprocessing which includes Cleaning, Case Folding, Normalizing, Stopword Removal, Tokenizing, and Stemming. Next, Feature Extraction using Term Frequency-Inverse Document Frequency (TF-IDF). This SVM model applies validation using K-Fold Cross-Validation with scenarios  $k = 2, 3, 4, 5, 6, 7, 8, 9,$  and 10. Grid Search is also used to determine the best parameters, and the results obtained are Radial Basic Function Kernel, Complexity = 1 and Gamma = 1. Furthermore, model performance evaluation is carried out using Confusion Matrix. The results showed that the highest accuracy was achieved in the  $k = 8$  scenario with an accuracy value of 82,521%, precision 82,296%, and recall 82,869%. This shows that 82.521% of the Support Vector Machine Method can predict the class correctly. Thus, the performance of the Support Vector Machine Method for sentiment analysis of the Threads app is good as it can predict many classes of data correctly with a breakdown of 1190 truly positive data and 1180 truly negative data. In addition, the model only incorrectly predicts 256 data in the positive class that should be negative, and 246 data in the negative class that should be positive.*

## ABSTRAK

Pevita, Selvia Dila. *Analisis Sentimen Ulasan Aplikasi Threads di Google Play Store Menggunakan Metode Support Vector Machine*. Skripsi. Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Nahdlatul Ulama Sunan Giri. Pemimbing I Muhammad Jauhar Vikri, M.Kom. dan Pembimbing II Sunu Wahyudhi, M.Pd.

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan aplikasi *Threads* di *Google Play Store* menggunakan metode *Support Vector Machine* (SVM) dengan dua label sentimen, yaitu positif dan negatif. *Dataset* yang digunakan berjumlah 2872 ulasan. Proses penelitian mencakup beberapa tahap, yaitu *Scraping* data yang dilakukan menggunakan bahasa pemrograman *Python* melalui *Google Colab*, lalu *Preprocessing* yang meliputi *Cleaning*, *Case Folding*, *Normalizing*, *Stopword Removal*, *Tokenizing*, dan *Stemming*. Selanjutnya ekstraksi fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Model SVM ini menerapkan validasi menggunakan *K-Fold Cross-Validation* dengan skenario  $k = 2, 3, 4, 5, 6, 7, 8, 9$ , dan 10. *Grid Search* juga digunakan untuk menentukan parameter terbaik, dan hasil yang didapatkan yaitu *Kernel Radial Basic Function*, *Complexity* = 1 dan *Gamma* = 1. Selanjutnya evaluasi kinerja model dilakukan menggunakan *Confusion Matrix*. Hasil penelitian menunjukkan bahwa akurasi tertinggi dicapai pada skenario  $k = 8$  dengan nilai akurasi sebesar 82,521%, *precision* 82,296%, dan *recall* 82,869%. Hal ini menunjukkan bahwa 82,521% Metode Support Vector Machine dapat memprediksi kelas dengan benar. Dengan demikian, Performa Metode Support Vector Machine untuk analisis sentimen aplikasi *Threads* baik karena dapat memprediksi banyak kelas data dengan benar dengan rincian 1190 data benar-benar positif dan 1180 data benar-benar negatif. Selain itu, model hanya salah memprediksi 256 data dalam kelas positif yang harusnya negatif, dan 246 data dalam kelas negatif yang harusnya positif.

## DAFTAR ISI

HALAMAN SAMBUT DALAM .....	i
PERNYATAAN KEASLIAN TULISAN .....	ii
LEMBAR PERSETUJUAN.....	iii
LEMBAR PENGESAHAN .....	iv
MOTTO DAN PERSEMBAHAN .....	v
KATA PENGANTAR .....	vi
<i>ABSTRACT</i> .....	vii
ABSTRAK .....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR BAGAN .....	xiv
DAFTAR GAMBAR .....	xv
DAFTAR LAMPIRAN.....	xvii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat.....	4
1.5.1 Manfaat Praktis .....	4
1.5.2 Manfaat Teoritis .....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Penelitian Sebelumnya .....	5
2.1.1 Penelitian Kusnawi dkk <i>Neobank Digital Banking</i> .....	5
2.1.2 Penelitian Panjalu dkk <i>MOBA Game</i> .....	5
2.1.3 Penelitian Pratama dkk <i>Beauty Product E-Commerce</i> .....	6
2.1.4 Penelitian Riyanto dan Azis <i>Covid-19 Vaccine Issue in Indonesia</i> ..	6
2.1.5 Penelitian Sahara dkk <i>Online Marketplace Product</i> .....	6
2.1.6 Penelitian Prabowo dan Azizah dkk <i>Cyberbullying</i> .....	7
2.1.7 Penelitian Bourequat dan Mourad <i>iPhone Release</i> .....	7



2.1.8	Penelitian Fransiska dkk <i>Provider by.U</i> .....	7
2.1.9	Penelitian Ardhani dkk Kartu Prakerja .....	7
2.1.10	Penelitian Kurniawati dkk <i>Maxim</i> .....	8
2.2	Perbedaan dengan penelitian sebelumnya .....	18
2.3	Dasar Teori .....	20
2.3.1	<i>Threads</i> .....	20
2.3.2	Analisis Sentimen .....	20
2.3.3	<i>K-Fold Cross Validation</i> .....	21
2.3.4	<i>Support Vector Machine</i> .....	22
2.3.5	<i>Web Scraping</i> .....	24
2.3.6	<i>Text Preprocessing</i> .....	24
2.3.7	TF-IDF .....	25
2.3.8	<i>Confusion Matrix</i> .....	26
<b>BAB III METODE PENELITIAN</b> .....		28
3.1	Jenis Penelitian .....	28
3.2	Identifikasi Masalah .....	28
3.3	<i>Pipeline Research</i> .....	28
3.4	<i>Dataset</i> .....	29
3.4.1	Jenis dan Sumber Data .....	29
3.4.2	Populasi dan Sampel Penelitian .....	29
3.4.3	Metode Pengumpulan Data.....	29
3.4.4	Pelabelan Data.....	30
3.4.5	Data Latih dan Data Uji .....	31
3.5	<i>Preprocessing</i> .....	31
3.5.1	<i>Cleaning</i> .....	32
3.5.2	<i>Case Folding</i> .....	32
3.5.3	<i>Normalizing</i> .....	33
3.5.4	<i>Stopword Removal</i> .....	33
3.5.5	<i>Tokenizing</i> .....	34
3.5.6	<i>Stemming</i> .....	34
3.6	<i>Feature Extraction</i> .....	35
3.6.1	Pembobotan TF-IDF .....	35

3.7	<i>Algorithm dan Validation</i> .....	39
3.7.1	<i>Support Vector Machine</i> .....	39
3.7.3	<i>Validation</i> .....	46
3.8	<i>Evaluation</i> .....	46
3.8.1	<i>Confusion Matrix</i> .....	46
3.9	Skenario Pengujian .....	47
3.10	Metode Pengembangan Sistem .....	48
3.10.1	<i>Analysis</i> .....	48
3.10.2	<i>Design</i> .....	50
3.10.3	<i>Implementation</i> .....	53
3.10.4	<i>Testing</i> .....	53
BAB IV HASIL DAN PEMBAHASAN .....		56
4.1	Pengujian <i>Support Vector Machine</i> Menggunakan <i>Google Colab</i> .....	56
4.1.1	<i>Dataset</i> .....	56
4.1.2	<i>Preprocessing</i> .....	58
4.1.3	<i>Feature Extraction</i> .....	63
4.1.4	<i>Algorithm and Validation</i> .....	64
4.1.5	<i>Evaluation</i> .....	66
4.2	Tampilan <i>Website</i> .....	68
4.2.1	Tampilan <i>Login</i> .....	68
4.2.2	Tampilan Beranda .....	69
4.2.3	Tampilan <i>Dataset</i> .....	69
4.2.4	Tampilan Klasifikasi .....	70
4.2.5	Tampilan Evaluasi .....	71
BAB V KESIMPULAN DAN SARAN .....		73
5.1	Kesimpulan .....	73
5.2	Saran .....	73
DAFTAR PUSTAKA .....		75
LAMPIRAN .....		80

## DAFTAR TABEL

Tabel 2. 1 Penelitian Kusnawi dkk <i>Neobank Digital Banking</i> .....	9
Tabel 2. 2 Penelitian Panjalu dkk <i>MOBA Games</i> .....	10
Tabel 2. 3 Penelitian Pratama dkk <i>Beauty Product E-Commerce</i> .....	11
Tabel 2. 4 Penelitian Riyanto dan Azis <i>Covid-19 Vaccine Issue in Indonesia</i> .....	12
Tabel 2. 5 Penelitian Sahara dkk <i>Marketplace Online</i> .....	13
Tabel 2. 6 Penelitian Prabowo dan Azizah <i>Cyberbullying</i> .....	14
Tabel 2. 7 Penelitian Bourequat dan Mourad <i>iPhone Release</i> .....	15
Tabel 2. 8 Penelitian Fransiska dkk <i>Provider by.U</i> .....	16
Tabel 2. 9 Penelitian Ardhani dkk <i>Kartu Prakerja</i> .....	17
Tabel 2. 10 Penelitian Kurniawati dkk <i>Maxim</i> .....	18
Tabel 3. 1 Sampel Data .....	31
Tabel 3. 2 <i>Dataset</i> manualisasi .....	32
Tabel 3. 3 <i>Cleaning</i> .....	32
Tabel 3. 4 <i>Case Folding</i> .....	33
Tabel 3. 5 <i>Normalizing</i> .....	33
Tabel 3. 6 <i>Stopword Removal</i> .....	33
Tabel 3. 7 <i>Tokenizing</i> .....	34
Tabel 3. 8 <i>Stemming</i> .....	34
Tabel 3. 9 Data Latih dan Data Uji 1 .....	35
Tabel 3. 10 Data Latih dan Data Uji 2 .....	35
Tabel 3. 11 IDF Data Latih 1 .....	36
Tabel 3. 12 IDF Data Latih 2 .....	36
Tabel 3. 13 TF-IDF Data Latih 1 .....	36
Tabel 3. 14 TF-IDF Data Latih 2 .....	37
Tabel 3. 15 WTF Data Uji 1 .....	38
Tabel 3. 16 WTF Data Uji 2 .....	38
Tabel 3. 17 TF-IDF Data Uji 1 .....	38
Tabel 3. 18 TF-IDF Data Uji 2 .....	39
Tabel 3. 19 Kernel RBF Data Latih .....	41
Tabel 3. 20 <i>Matrix Hessian</i> Data Latih.....	41
Tabel 3. 21 Nilai <i>Alpha</i> Data Latih .....	43
Tabel 3. 22 Kernel yang digunakan untuk Nilai Bias .....	43
Tabel 3. 23 Inisialisasi Variabel.....	44
Tabel 3. 24 Kernel RBF Data Uji .....	45
Tabel 3. 25 Hasil Pengujian 1 .....	45
Tabel 3. 26 Hasil Pengujian 2 .....	45
Tabel 3. 27 Kebutuhan Fungsional dan Non-Fungsional.....	48
Tabel 3. 28 Kebutuhan Perangkat Lunak.....	49
Tabel 3. 29 Kebutuhan Perangkat Keras.....	49

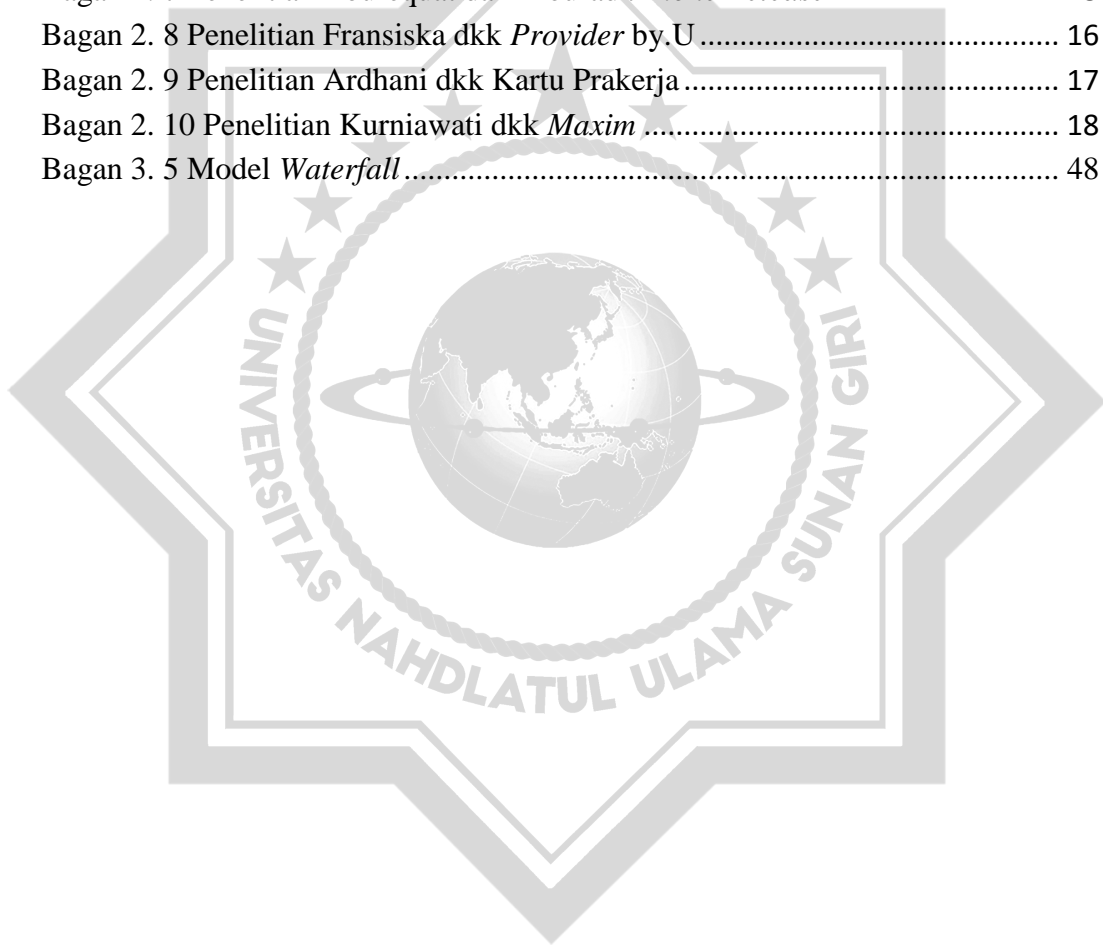
Tabel 3. 30 Rencana Pengujian <i>Blackbox</i> .....	53
Tabel 4. 1 Perbandingan Data Latih dan Data Uji .....	58
Tabel 4. 2 Hasil cleaning dan case folding .....	62
Tabel 4. 3 Hasil Normalizing dan Stopword Removal .....	62
Tabel 4. 4 Hasil Tokenizing dan Stemming .....	63
Tabel 4. 5 Hasil Grid Search .....	65
Tabel 4. 6 Hasil Klasifikasi .....	66
Tabel 4. 7 Confusion Matrix keseluruhan .....	67
Tabel 4. 8 Precision, Recall, dan Accuracy .....	67
Tabel 4. 9 Perbandingan akurasi .....	68



# UNUGIRI

## DAFTAR BAGAN

Bagan 2. 1 Penelitian Kusnawi dkk <i>Neobank Digital Banking</i> .....	9
Bagan 2. 2 Penelitian Panjalu dkk <i>MOBA Games</i> .....	10
Bagan 2. 3 Penelitian Pratama dkk <i>Beauty Product E-Commerce</i> .....	11
Bagan 2. 4 Penelitian Riyanto dan Azis <i>Covid-19 Vaccine Issue in Indonesia</i> ....	12
Bagan 2. 5 Penelitian Sahara dkk <i>Marketplace Online</i> .....	13
Bagan 2. 6 Penelitian Prabowo dan Azizah <i>Cyberbullying</i> .....	14
Bagan 2. 7 Penelitian Bourequat dan Mourad <i>iPhone Release</i> .....	15
Bagan 2. 8 Penelitian Fransiska dkk <i>Provider by.U</i> .....	16
Bagan 2. 9 Penelitian Ardhani dkk <i>Kartu Prakerja</i> .....	17
Bagan 2. 10 Penelitian Kurniawati dkk <i>Maxim</i> .....	18
Bagan 3. 5 Model <i>Waterfall</i> .....	48

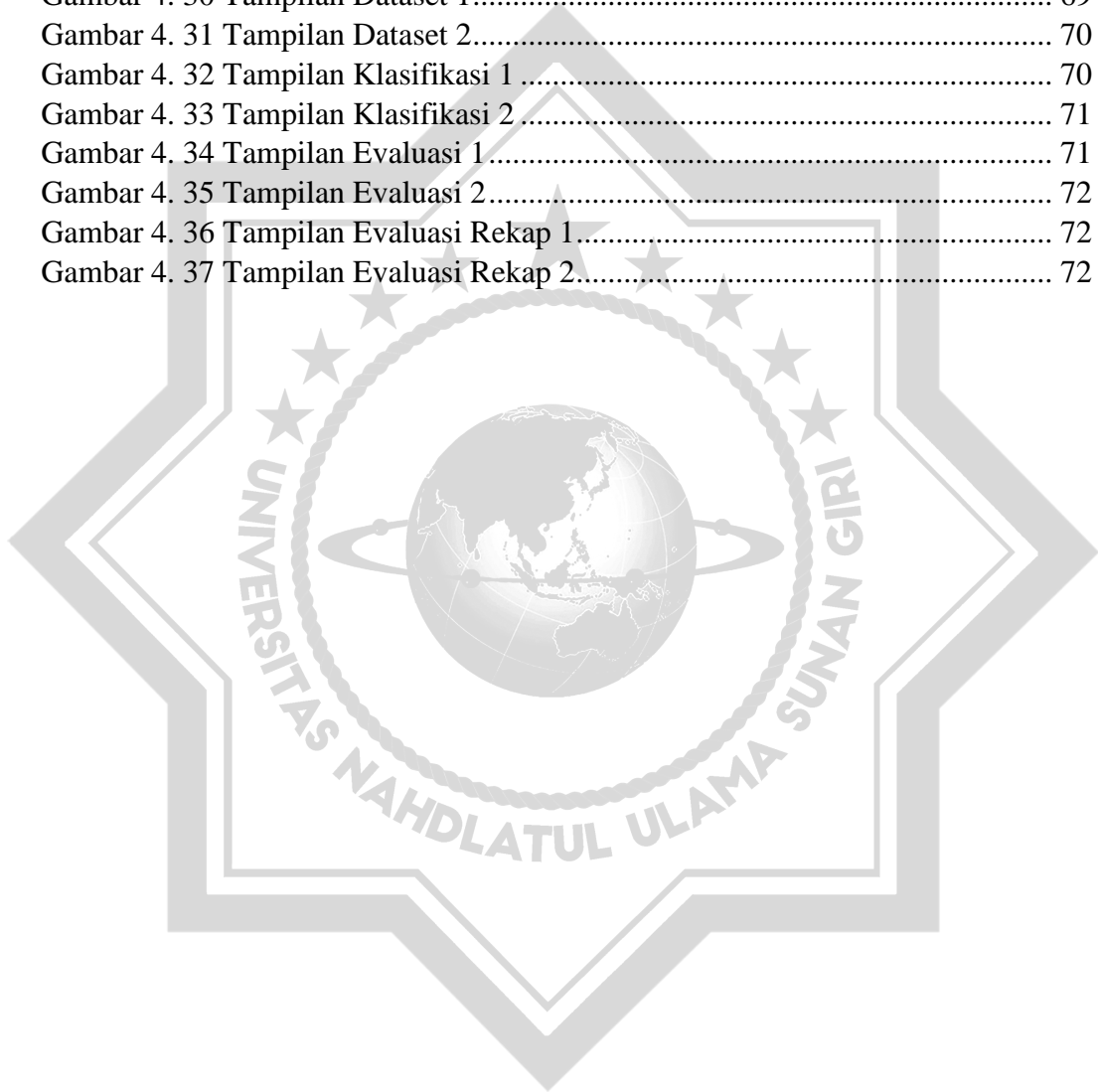


# UNUGIRI

## DAFTAR GAMBAR

Gambar 2. 1 <i>Threads</i> .....	20
Gambar 2. 2 <i>K-Fold Cross Validation</i> .....	21
Gambar 2. 3 Support Vector Machine .....	22
Gambar 2. 4 Confusion Matrix .....	26
Gambar 3. 1 <i>Pipeline Research</i> .....	28
Gambar 3. 2 Proses <i>Web Scraping</i> .....	29
Gambar 3. 3 Hasil <i>Scraping Data</i> .....	30
Gambar 3. 4 Pelabelan Data.....	30
Gambar 3. 5 <i>Text Preprocessing</i> .....	31
Gambar 3. 6 Skenario Pengujian.....	47
Gambar 3. 7 <i>Use Case Diagram</i> .....	50
Gambar 3. 8 Tampilan <i>Login</i> .....	51
Gambar 3. 9 Tampilan Beranda .....	51
Gambar 3. 10 Tampilan <i>Dataset</i> .....	51
Gambar 3. 11 Tampilan klasifikasi 1 .....	52
Gambar 3. 12 Tampilan Klasifikasi 2 .....	52
Gambar 3. 13 Tampilan Evaluasi 1 .....	52
Gambar 3. 14 Tampilan Evaluasi 2.....	53
Gambar 4. 1 Proses <i>scraping</i> .....	56
Gambar 4. 2 Source Code <i>Scraping</i> .....	56
Gambar 4. 3 Source Code Hapus Netral .....	57
Gambar 4. 4 Penghapusan label netral .....	57
Gambar 4. 5 Proses <i>Split Data</i> .....	57
Gambar 4. 6 Source Code <i>Split Data</i> .....	58
Gambar 4. 7 Proses <i>Cleaning</i> .....	58
Gambar 4. 8 Source Code <i>Cleaning</i> .....	59
Gambar 4. 9 Proses <i>Case Folding</i> .....	59
Gambar 4. 10 Source Code <i>Case Folding</i> .....	59
Gambar 4. 11 Proses <i>Normalizing</i> .....	59
Gambar 4. 12 Source Code <i>Normalizing</i> .....	60
Gambar 4. 13 Proses <i>Stopword Removal</i> .....	60
Gambar 4. 14 Source Code <i>Stopword Removal</i> .....	60
Gambar 4. 15 Proses <i>Tokenizing</i> .....	61
Gambar 4. 16 Source Code <i>Tokenizing</i> .....	61
Gambar 4. 17 Proses <i>Stemming</i> .....	61
Gambar 4. 18 Source Code <i>Stemming</i> .....	62
Gambar 4. 19 Proses <i>TF-IDF</i> .....	63
Gambar 4. 20 Source Code <i>TF-IDF</i> .....	63
Gambar 4. 21 Hasil <i>TF-IDF</i> .....	64
Gambar 4. 22 Proses <i>Grid Search</i> .....	64
Gambar 4. 23 Source Code <i>Grid Search</i> .....	65

Gambar 4. 24 Proses klasifikasi dan validasi.....	65
Gambar 4. 25 Source Code Klasifikasi dan Validasi.....	66
Gambar 4. 26 Proses Evaluasi.....	66
Gambar 4. 27 Source Code Evaluasi .....	66
Gambar 4. 28 Tampilan Login .....	68
Gambar 4. 29 Tampilan Beranda .....	69
Gambar 4. 30 Tampilan Dataset 1.....	69
Gambar 4. 31 Tampilan Dataset 2.....	70
Gambar 4. 32 Tampilan Klasifikasi 1 .....	70
Gambar 4. 33 Tampilan Klasifikasi 2 .....	71
Gambar 4. 34 Tampilan Evaluasi 1.....	71
Gambar 4. 35 Tampilan Evaluasi 2.....	72
Gambar 4. 36 Tampilan Evaluasi Rekap 1.....	72
Gambar 4. 37 Tampilan Evaluasi Rekap 2.....	72



**UNUGIRI**

## DAFTAR LAMPIRAN

Lampiran 1 Surat Pernyataan Pengujian Aplikasi .....	80
Lampiran 2 Angket Uji <i>Blackbox</i> .....	81



# UNUGIRI