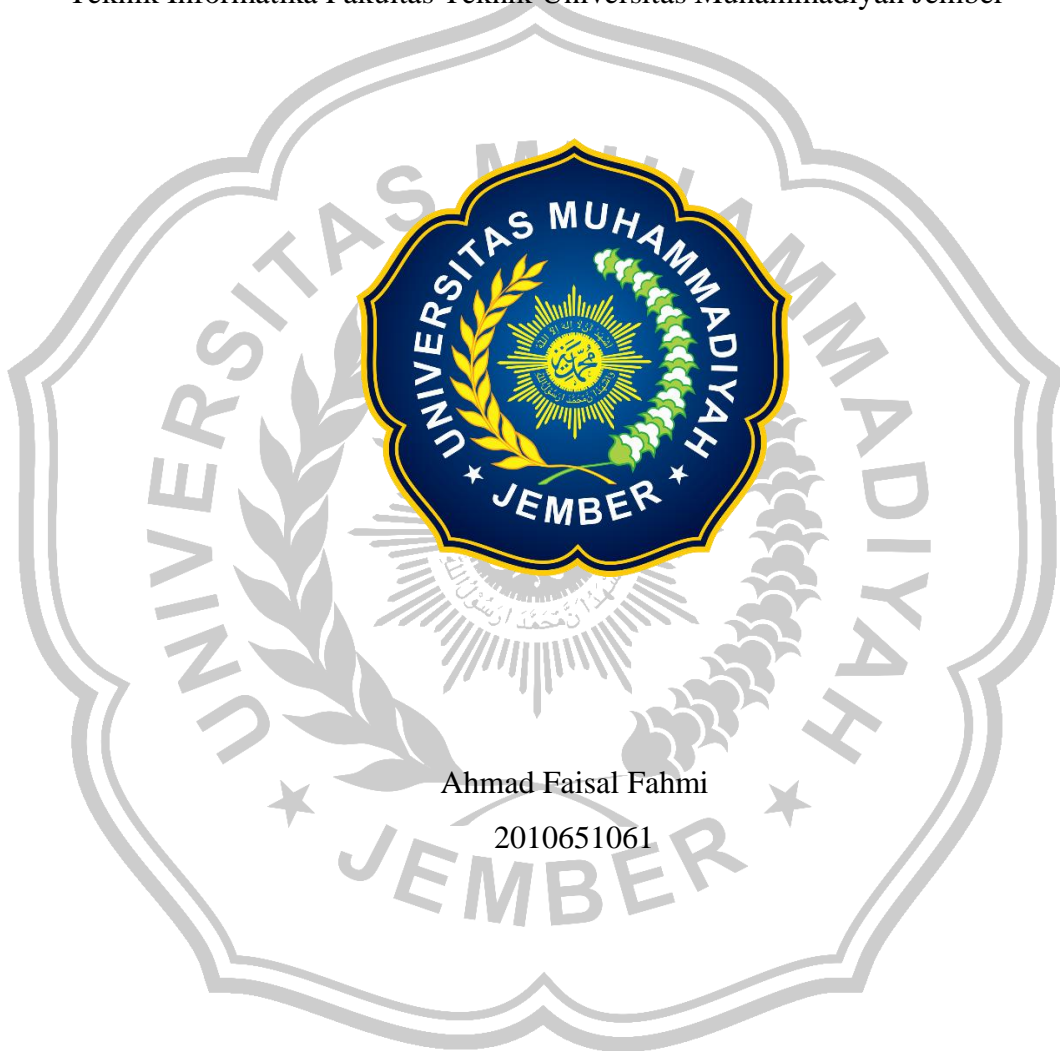


TUGAS AKHIR

KLASIFIKASI UNIFORM RESOURCE LOCATOR BERBAHAYA MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)

Diajukan sebagai salah satu syarat untuk kelulusan Strata Satu (S-1) Prodi
Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember



Ahmad Faisal Fahmi

2010651061

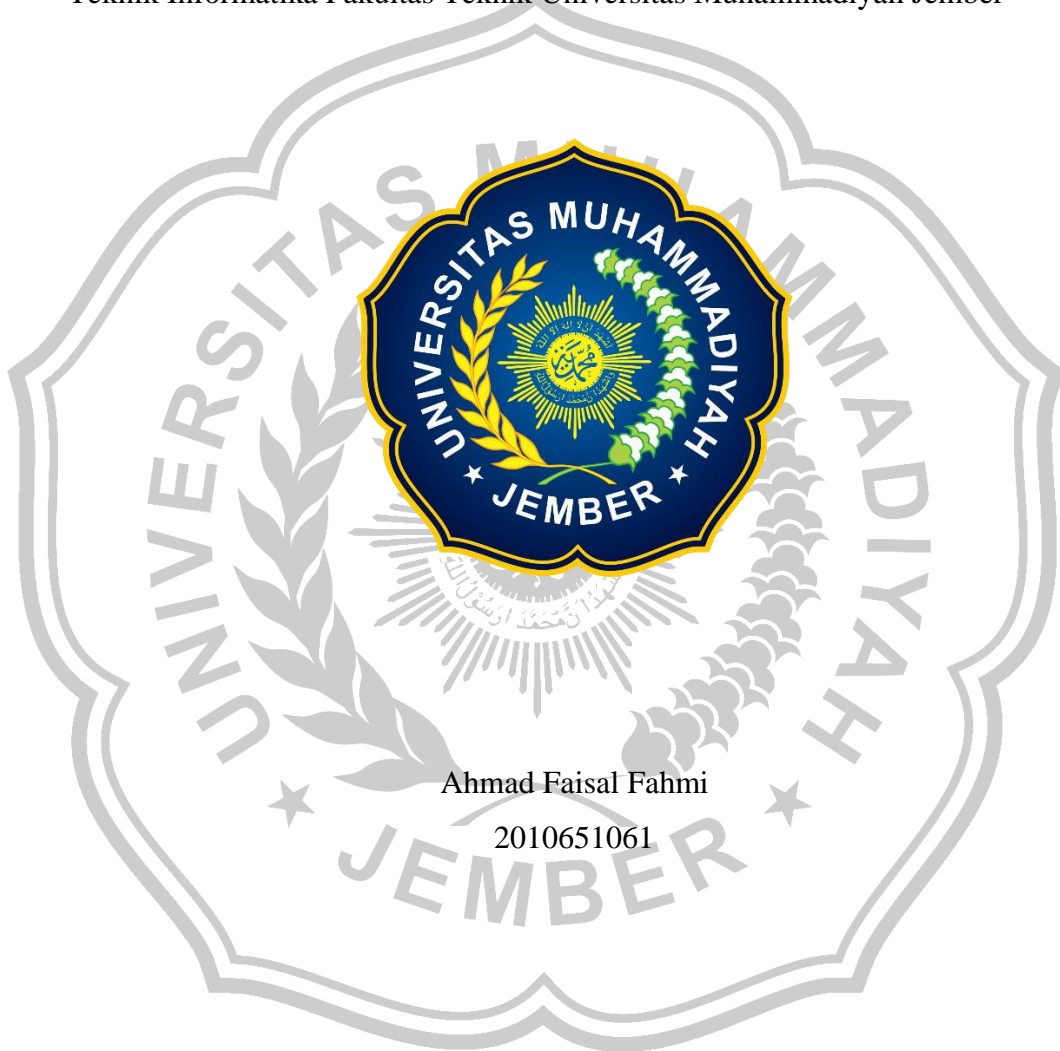
**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER**

2026

TUGAS AKHIR

KLASIFIKASI UNIFORM RESOURCE LOCATOR BERBAHAYA MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)

Diajukan sebagai salah satu syarat untuk kelulusan Strata Satu (S-1) Prodi
Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember



Ahmad Faisal Fahmi

2010651061

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER**

2026

LEMBAR PERSETUJUAN
KLASIFIKASI UNIFORM RESOURCE LOCATOR BERBAHAYA
MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE
(SVM)

Ahmad Faisal Fahmi

2010651061

Telah disetujui bahwa Laporan Tugas Akhir ini untuk diajukan pada sidang Tugas Akhir sebagai salah satu syarat kelulusan dan mendapat gelar Sarjana Komputer (S.Kom.)

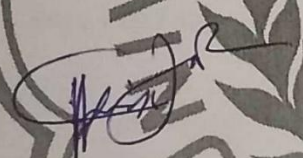
di

Universitas Muhammadiyah Jember

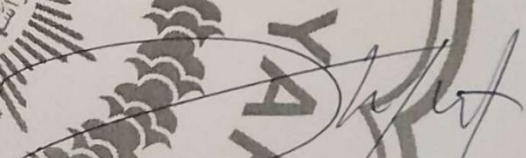
Disetujui Oleh:

Pembimbing I

Pembimbing II


Ginanjar Abdurrahman, S.Si., M.Pd.

NIDN. 0714078704


Ari Eko Wardoyo, ST., M.Kom.

NIDN. 0014027501

SURAT PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : AHMAD FAISAL FAHMI
Nomor Induk Mahasiswa : 2010651061
Program Studi : Teknik Informatika
Perguruan Tinggi : Universitas Muhammadiyah Jember

Menyatakan dengan sesungguhnya karya ilmiah berupa tugas akhir yang berjudul **“KLASIFIKASI UNIFORM RESOURCE LOCATOR BERBAHAYA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)”** adalah benar merupakan karya sendiri. Hal-hal yang bukan karya saya, dalam skripsi tersebut diberi tanda sitasi dan ditunjukkan dalam daftar pustaka. Apabila di kemudian hari terbukti pernyataan saya tidak benar dan ditemukan pelanggaran atas karya Tugas Akhir ini, saya bersedia menerima sanksi akademik berupa pencabutan Tugas Akhir dan gelar yang saya peroleh dari Tugas Akhir tersebut.

Jember, 10 Februari 2026



Ahmad Faisal Fahmi

NIM. 2010651061

LEMBAR PENGESAHAN

KLASIFIKASI UNIFORM RESOURCE LOCATOR BERBAHAYA
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)

Oleh:

AHMAD FAISAL FAHMI

2010651061

Telah mempertanggungjawabkan Laporan Tugas Akhirnya pada sidang Tugas Akhir tanggal sebagai salah satu syarat kelulusan dan mendapat gelar Sarjana

Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

Disetujui Oleh:

Penguji I,

Pembimbing I,

Miftahur Rahman, M.Kom

NIDN. 0724039201

Ginjar Abdurrahman, S.Si., M.Pd.

NIDN. 0714078704

Penguji II,

Pembimbing II,

Moh. Dasuki, M.Kom

NIDN. 0722109103

Ari Eko Wardoyo, ST., M.Kom.

NIDN. 0014027501

Mengesahkan,

Dekan Fakultas Teknik,



Prof. Dr. Ir. Muhtar, ST., MT., IPM

NIDN. 0010067301

Mengetahui,

Ketua Program Studi Teknik Informatika,



Rosita Yanuarti, S.Kom., M.Cs

NIDN. 0629018601

KATA PENGANTAR

Segala puji Syukur penulis panjatkan kehadiran Allah SWT atas limpahan Rahmat, karunia dan anugerah-Nya, sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “Klasifikasi Uniform Resource Locator Berbahaya Menggunakan Algoritma Support Vector Machine (SVM)”. Penyusunan tugas akhir ini dilakukan sebagai salah satu persyaratan untuk memperoleh gelar sarjana pada Program Studi Teknik Informatika.

Penulis mengucapkan terimakasih yang sebesar-besarnya kepada Bapak Ginanjar Abdurrahman, S.Si.,M.Pd., M.Kom dan Bapak Ari Eko Wardoyo, ST., M.Kom. selaku dosen pembimbing atas bimbingannya, arahan serta perhatian yang telah diberikan selama proses penyusunan tugas akhir ini.

Penulis menyadari bahwa tugas akhir ini masih memiliki berbagai keterbatasan dan kekurangan, mengingat keterbatasan kemampuan serta pengetahuan yang dimiliki. Oleh karena itu, penulis memohon maaf sebesar-besarnya. Akhir kata penulis berharap semoga Tugas Akhir ini dapat memberikan manfaat dan kontribusi yang positif bagi para pembaca.

PERSEMBAHAN

Alhamdulillah segala puji bagi Allah SWT yang telah melimpahkan berbagai rahmat anugerah dan karunia yang luar biasa diberikan, sehingga penulis dapat menyelesaikan tugas akhir. Maka dari itu pada kesempatan ini penulis menyampaikan rasa terimakasih yang sedalam-dalamnya diberikan kepada:

1. Allah SWT berkat segala ridho, rahmat dan hidayah-Nya penulis dapat menyelesaikan segala urusan dalam menyusun laporan Tugas Akhir dan diberikan kesempatan mendapat gelar Sarjana Komputer.
2. Keluarga saya yang senantiasa selalu mendoakan, memberikan dukungan materi dan memberi nasihat.
3. Diri saya sendiri yang telah berjuang menyelesaikan tugas akhir ini.
4. Bapak Prof. Dr. Ir. Muhtar, ST., M.T., IPM., selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
5. Ibu Rosita Yanuarti, S.Kom., M.Cs. selaku ketua progam Studi Teknik Informatika Universitas Muhammadiyah Jember.
6. Bapak Ginanjar Abdurrahman, S.Si.,M.Pd. selaku dosen pembimbing I dan Bapak Ari Eko Wardoyo, ST., M.Kom. selaku dosen pembimbing II, yang telah meluangkan waktu untuk memberikan bimbingan/arahan serta masukan dalam menyusun penulisan tugas akhir ini sampai selesai.
7. Bapak Miftahur Rahman, M.Kom selaku dosen penguji I dan Bapak Moh. Dasuki, M.Kom selaku dosen penguji II, yang telah memberikan saran serta masukan dalam menyelesaikan tugas akhir ini.
8. Teman seperjuangan teknik informatika senantiasa membantu dan mendukung saya dalam menyusun tugas akhir saya ini.

MOTTO

“Don't cry because it's over, smile because it happened.”

*“Mereka yang berdiri setelah dihantam badai tidak akan terusik oleh gerimis”
(Natanael Jansudin Siregar)*



DAFTAR ISI

HALAMAN JUDUL..	i
LEMBAR PERSETUJUAN	iii
SURAT PERNYATAAN	iv
LEMBAR PENGESAHAN	v
KATA PENGANTAR	vi
PERSEMBAHAN.....	vii
ABSTRAK.....	viii
<i>ABSTRACT</i>	ix
MOTTO	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL.....	xv
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah.....	2
1.3. Tujuan Penelitian.....	3
1.4. Manfaat Penelitian.....	3
1.5. Batasan Penelitian	3
BAB 2 TINJAUAN PUSTAKA	4
2.1. Penelitian Terdahulu	4
2.2. URL.....	5
2.3. Metode Klasifikasi	6
2.4. <i>Preprosesing</i> Data.....	6
2.4.1. <i>Missing Value</i>	6
2.4.2. <i>Cleaning</i>	7
2.4.3. <i>Case Folding</i>	7
2.4.4. <i>Tokenizing</i>	7

2.4.5. <i>Filtering</i>	7
2.4.6. <i>Stemming</i>	8
2.4.7. <i>TF-IDF</i>	8
2.4.8. <i>Pembagian Data</i>	9
2.4.9. <i>K-Fold Cross Validation</i>	9
2.5. <i>Oversampling</i>	10
2.6. <i>Confusion Matrix</i>	10
2.7. <i>Akurasi</i>	11
2.8. <i>Presisi</i>	12
2.9. <i>Recall</i>	12
2.10. <i>SVM (Support Vector Machine)</i>	12
BAB 3 METODOLOGI PENELITIAN	14
3.1. <i>Studi Literatur</i>	14
3.2. <i>Pengumpulan Data</i>	16
3.3. <i>Preprosesing Data</i>	16
3.3.1. <i>Proses Penghapusan Data Kosong</i>	17
3.3.2. <i>Proses Cleaning</i>	17
3.3.3. <i>Proses Case Folding</i>	17
3.3.4. <i>Proses Tokenizing</i>	18
3.3.5. <i>Proses Filtering</i>	18
3.3.6. <i>Proses Stemming</i>	19
3.4. <i>Perhitungan TF-IDF</i>	19
3.5. <i>Oversampling</i>	22
3.6. <i>Support Vector Machine</i>	22
3.7. <i>Perhitungan hyperplane linear</i>	27
3.7.1. <i>Menghitung hyperplane dari label benign dan label lainnya</i>	
28	
3.9.1. <i>Menghitung Hyperplane dari Label Malware, Phising dan Defacement</i>	29
3.9. <i>Pembagian K-Fold</i>	32

3.10. Hasil Akurasi, Presisi Dan <i>Recall</i>	33
BAB 4 HASIL PENELITIAN DAN PEMBAHASAN	35
4.1. Persiapan Data.....	35
4.2. <i>Preprocessing</i> Data	35
4.2.1. <i>Missing Value</i>	35
4.2.2. <i>Cleaning</i>	37
4.2.3. <i>Case Folding</i>	37
4.2.4. <i>Tokenizing</i>	39
4.2.5. <i>Filtering</i>	40
4.2.6. <i>Stemming</i>	41
4.2.7. TF-IDF	42
4.2.8. <i>Oversampling</i>	44
4.2.9. <i>K-fold</i>	46
4.2.10. Algoritma Support Vector Machine (SVM).....	48
4.3. Hasil Percobaan dan Analisis	51
4.4. Confusion Matrix	52
4.4.1. Perhitungan Akurasi, Presisi, dan <i>Recall</i>	55
4.4.2. Perbandingan Hasil	58
BAB 5 KESIMPULAN.....	59
5.1. Kesimpulan.....	59
5.2. Saran.....	60
DAFTAR PUSTAKA	61
LAMPIRAN.....	67

DAFTAR GAMBAR

Gambar 3. 1 Diagram alur studi literatur.....	14
Gambar 3. 2 Word Cloud Dari Label Phishing	20
Gambar 3. 3 Word Cloud Dari Label Malware	20
Gambar 3. 4 Word Cloud Dari Label Defacement.....	21
Gambar 3. 5 Word Cloud Dari Label Benign.....	21
Gambar 3. 6 Contoh Penempatan Titik Data.....	23
Gambar 3. 7 Contoh Hyperplane.....	24
Gambar 3. 8 Visualisasi Kernel Linear	24
Gambar 3. 9 Visualisasi Kernel Poly	25
Gambar 3. 10 Visualisasi Kernel RBF	26
Gambar 3. 11 Visualisasi Kernel RBF 3 dimensi.....	27
Gambar 3. 12 Visualisasi Titik Kordinat Data	27
Gambar 4. 1 Data yang termasuk dalam missing value	35
Gambar 4. 2 Gambar setelah missing value dihapus.....	36
Gambar 4. 3 Data yang belum melalui proses cleaning	37
Gambar 4. 4 Contoh data.....	37
Gambar 4. 5 Perbandingan jumlah data saat oversampling.	45
Gambar 4. 6 pembagian data menggunakan K-fold.....	47
Gambar 4. 7 putaran pertama K-fold.....	47
Gambar 4. 8 putaran kedua K-fold.....	48
Gambar 4. 9 Koordinat titik data.....	49
Gambar 4. 10 Support vector.....	49
Gambar 4. 11 Data setelah dipetakan	50
Gambar 4. 12 Confusion matrix	54
Gambar 4. 13 Confusion matrik pada kernel RBF	56
Gambar 4. 14 Confusion matrik pada kernel linear	56
Gambar 4. 15 Confusion matrik pada kernel polynomial	57

DAFTAR TABEL

Tabel 3. 1 Data Sampel.....	16
Tabel 3. 2 Hasil <i>Preprocessing</i>	17
Tabel 3. 3 Perbandingan Data Pada Proses <i>Cleaning</i>	17
Tabel 3. 4 Perbandingan Data pada Proses <i>Case Folding</i>	18
Tabel 3. 5 Perbandingan Data Pada Proses <i>Tokenizing</i>	18
Tabel 3. 6 Perbandingan Data Pada Proses <i>Filering</i>	18
Tabel 3. 7 Perbandingan Data Pada Proses <i>Stemming</i>	19
Tabel 3. 8 Perbandingan Data Pada Proses <i>Oversampling</i>	22
Tabel 3. 9 Koordinat <i>Support Vector</i>	28
Tabel 3. 10 Koordinat Titik Data Sampel.....	30
Tabel 3. 11 Pembagian <i>Fold Data</i>	32
Tabel 3. 12 Hasil Prediksi dan Data Aktual <i>Fold 1</i>	33
Tabel 3. 13 Hasil Prediksi dan Data Aktual <i>Fold 2</i>	33
Tabel 3. 14 Hasil Prediksi dan Data Aktual <i>Fold 3</i>	34
Gambar 4. 1 Data yang termasuk dalam missing value.....	35
Gambar 4. 2 Gambar setelah missing value dihapus.....	36
Gambar 4. 3 Data yang belum melalui proses cleaning.....	37
Gambar 4. 4 Contoh data.....	37
Gambar 4. 5 Perbandingan jumlah data saat oversampling.....	45
Gambar 4. 6 pembagian data menggunakan K-fold.....	47
Gambar 4. 7 putaran pertama K-fold.....	47
Gambar 4. 8 putaran kedua K-fold.....	48
Gambar 4. 9 Koordinat titik data.....	49
Gambar 4. 10 Support vector.....	49
Gambar 4. 11 Data setelah dipetakan.....	50
Gambar 4. 12 Confusion matrix.....	54
Gambar 4. 13 Confusion matrik pada kernel RBF.....	56

Gambar 4. 14 Confusion matrik pada kernel linear.....56
Gambar 4. 15 Confusion matrik pada kernel polynomial57

