

PENERAPAN METODE KLASIFIKASI *SUPPORT VECTOR MACHINE* PADA DATASET KELULUSAN MAHASISWA

Hezkia Vialen Anrastra
(180411100116)

Pembimbing 1 : Dr. Fika Hastarita Rachman, S.T., M.Eng.

Pembimbing 2 : Fifin Ayu Mufarroha, S.Kom., M.Kom.

ABSTRAK

Penelitian ini bertujuan untuk menerapkan metode klasifikasi *Support Vector Machine (SVM)* pada dataset kelulusan mahasiswa Program Studi Teknik Informatika Universitas Trunojoyo Madura guna memprediksi status kelulusan mahasiswa berdasarkan berbagai fitur yang relevan. Tiga jenis *kernel SVM* yang diuji dalam penelitian ini adalah *Polynomial*, *Radial Basis Function (RBF)*, dan *Linear*, dengan variasi proporsi *splitting* data yaitu 70%:30%, 80%:20%, dan 90%:10%. Hasil evaluasi model menunjukkan bahwa *kernel Polynomial* memberikan performa terbaik di antara ketiga *kernel* yang diuji, dengan akurasi tertinggi mencapai 74% pada *splitting* data 70%:30% dan 80%:20%. Sementara itu, *kernel RBF* dan *Linear* memiliki performa yang lebih rendah, dengan akurasi tertinggi masing-masing mencapai 62% dan 56%. Untuk metrik lain seperti *precision*, *recall*, dan *F1-score*, *kernel Polynomial* menunjukkan hasil yang lebih konsisten, terutama pada kelas positif (1), dengan nilai *F1-score* tertinggi 83% pada *splitting* data 70%:30%. Secara keseluruhan, penelitian ini menunjukkan bahwa *kernel Polynomial* lebih efektif dalam memprediksi kelulusan mahasiswa dibandingkan dengan *kernel RBF* dan *Linear* pada dataset ini, meskipun ketiga *kernel* memberikan hasil yang dapat diterima untuk aplikasi prediksi kelulusan.

Kata Kunci: *Support Vector Machine*, Klasifikasi, Kelulusan Mahasiswa

KATA PENGANTAR

Puji syukur kehadiran Allah SWT berkat limpahan rahmat dan hidayah-Nya, penulis mampu menuntaskan skripsi berjudul “PENERAPAN METODE KLASIFIKASI *SUPPORT VECTOR MACHINE* PADA DATASET KELULUSAN MAHASISWA”. Penulis juga menyampaikan terimakasih yang sebesar-besarnya kepada:

1. Terima kasih kepada Tuhan atas kasih karunia dan berkat-Nya bagi penulis hingga saat ini. Selain itu, Yesus Kristus adalah Juru Selamat umat manusia yang membawa terang dan keselamatan bagi dunia.
2. Keluarga dan kerabat yang selalu menyayangi dan mendukung penulis secara emosional dan materil.
3. Ibu Dr. Fika Hastarita Rachman, S.T., M.Eng. sebagai dosen pembimbing 1 telah memberi masukan serta motivasi untuk penulis selama penelitian berlangsung
4. Ibu Fifin Ayu Mufarroha, S.Kom., M.Kom. sebagai dosen pembimbing 2 telah memberi kritik beserta saran untuk penulis selama penelitian berlangsung.
5. Bapak dan Ibu dosen Teknik Informatika UTM yang sudah berbagi pengetahuan mereka selama kuliah yang singkat ini.
6. *I wanna thank me for believing in me. I wanna thank me for doing all this hard work. I wanna thank me for having no days off. I wanna thank me for never quitting. I wanna thank me for being me at all times.*

Penulis sadar bahwa pasti terdapat kesalahan dalam pengerjaan skripsi ini. Maka dari itu, penulis memohon maaf atas segala kesalahan yang ada. Penulis, dengan segala kerendahan hati, menerima komentar dan saran dari pembaca untuk membantu memperbaiki penelitian di masa mendatang.

Bangkalan, 14 Januari 2025

Hezkia Vialen Anrastra

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN ORISINALITAS.....	iii
ABSTRAK	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	viii
DAFTAR GAMBAR	ix
DAFTAR KODE PROGRAM.....	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.2.1 Permasalahan.....	3
1.2.2 Solusi Permasalahan.....	3
1.2.3 Pertanyaan Penelitian	3
1.3 Batasan Masalah.....	3
1.4 Tujuan dan Manfaat	4
1.4.1 Tujuan	4
1.4.2 Manfaat	4
1.5 Sistematika Laporan.....	4
1.6 Jadwal Penelitian.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Landasan Teori.....	6
2.1.1 Data Mining	6
2.1.2 Klasifikasi	7
2.1.3 <i>Support Vector Machine</i>	7
2.1.4 <i>SMOTE (Synthetic Minority Oversampling Technique)</i>	9
2.1.5 <i>Confusion Matrix</i>	9
2.1.6 Kelulusan.....	10
2.2 Penelitian Terdahulu	11
BAB III METODE USULAN.....	19
3.1 Pengambilan <i>Dataset</i>	19

3.2	<i>Preprocessing Data</i>	20
3.2.1	Pemisahan Fitur dan Label	20
3.2.2	Data Transformation	21
3.2.3	Normalisasi Min-Max Scaler	21
3.2.4	<i>Splitting Data</i>	22
3.2.5	Proses <i>SMOTE (Synthetic Minority Over-sampling Technique)</i>	23
3.3	Proses <i>Support Vector Machine</i>	23
3.4	Evaluasi Sistem	25
3.5	Skenario Uji Coba	26
BAB IV HASIL DAN PEMBAHASAN		28
4.1	Dataset.....	28
4.2	<i>Preprocessing Data</i>	28
4.2.1	Pemisahan Fitur dan Label.....	28
4.2.2	Data Transformation	30
4.2.3	Normalisasi Data	31
4.2.4	<i>Splitting Data</i>	32
4.2.5	Penyeimbangan Data Dengan Proses <i>SMOTE</i>	32
4.3	Penerapan Metode <i>Support Vector Machine</i>	33
4.4	Evaluasi Sistem	34
4.4.1	Hasil Evaluasi <i>Kernel Polynomial</i>	34
4.4.2	Hasil Evaluasi <i>Kernel RBF</i>	38
4.4.3	Hasil Evaluasi <i>Kernel Linear</i>	41
4.5	Pembahasan Hasil	44
4.5.1	<i>Kernel Polynomial</i>	45
4.5.2	<i>Kernel RBF (Radial Basis Function)</i>	45
4.5.3	<i>Kernel Linear</i>	46
4.5.4	Perbandingan Antara <i>Kernel</i>	47
BAB V KESIMPULAN DAN SARAN.....		48
5.1	Kesimpulan	48
5.2	Saran.....	48
DAFTAR PUSTAKA		49

DAFTAR TABEL

Tabel 1. 1 Jadwal Penelitian.....	5
Tabel 2. 1 <i>Confusion Matrix</i>	9
Tabel 2. 2 Penelitian Terdahulu	13
Tabel 3. 1 Atribut Dalam Data	19
Tabel 3. 2 Contoh Dataset.....	20
Tabel 3. 3 Nilai <i>Minimum</i> dan <i>Maximum</i>	21
Tabel 3. 4 Hasil Normalisasi	21
Tabel 3. 5 Data <i>Training</i>	22
Tabel 3. 6 Data <i>Testing</i>	23
Tabel 3. 7 Jarak <i>Euclidean</i> Dan Kernel <i>RBF</i> Setiap Pasangan	24
Tabel 3. 8 Nilai <i>Langrangian</i> dan Nilai Bias	25
Tabel 3. 9 Hasil Prediksi dan Evaluasi Model	25
Tabel 3. 10 Skenario Uji Coba	26
Tabel 4. 1 Hasil Evaluasi Skenario Uji Coba.....	44

DAFTAR GAMBAR

Gambar 2. 1 SVM Linear dan Non Linear.....	8
Gambar 3. 1 Arsitektur Sistem.....	19
Gambar 3. 2 Label Prediksi.....	26
Gambar 4. 1 Tampilan Dataset.....	28
Gambar 4. 2 Fitur yang Digunakan.....	29
Gambar 4. 3 Label.....	30
Gambar 4. 4 Hasil Transformasi	31
Gambar 4. 5 Hasil Normalisasi	31
Gambar 4. 6 Hasil Penyeimbangan dengan <i>SMOTE</i>	33
Gambar 4. 7 <i>Confusion Matrix Kernel Polynomial Split 70%:30%</i>	35
Gambar 4. 8 Metrik Evaluasi <i>Kernel Polynomial Split 70%:30%</i>	35
Gambar 4. 9 <i>Confusion Matrix Kernel Polynomial Split 80%:20%</i>	36
Gambar 4. 10 Metrik Evaluasi <i>Kernel Polynomial Split 80%:20%</i>	36
Gambar 4. 11 <i>Confusion Matrix Kernel Polynomial Split 90%:10%</i>	37
Gambar 4. 12 Metrik Evaluasi <i>Kernel Polynomial Split 90%:10%</i>	37
Gambar 4. 13 <i>Confusion Matrix Kernel RBF Split 70%:30%</i>	38
Gambar 4. 14 Metrik Evaluasi <i>Kernel RBF Split 70%:30%</i>	38
Gambar 4. 15 <i>Confusion Matrix Kernel RBF Split 80%:20%</i>	39
Gambar 4. 16 Metrik Evaluasi <i>Kernel RBF Split 80%:20%</i>	39
Gambar 4. 17 <i>Confusion Matrix Kernel RBF Split 90%:10%</i>	40
Gambar 4. 18 Metrik Evaluasi <i>Kernel RBF Split 90%:10%</i>	40
Gambar 4. 19 <i>Confusion Matrix Kernel Linear Split 70%:30%</i>	41
Gambar 4. 20 Metrik Evaluasi <i>Kernel Linear Split 70%:30%</i>	41
Gambar 4. 21 <i>Confusion Matrix Kernel Linear Split 80%:20%</i>	42
Gambar 4. 22 Metrik Evaluasi <i>Kernel Linear Split 80%:20%</i>	42
Gambar 4. 23 <i>Confusion Matrix Kernel Linear Split 90%:10%</i>	43
Gambar 4. 24 Metrik Evaluasi <i>Kernel Linear Split 90%:10%</i>	43
Gambar 4. 25 Grafik <i>Polynomial</i>	45
Gambar 4. 26 Grafik <i>RBF</i>	45
Gambar 4. 27 Grafik <i>Linear</i>	46

DAFTAR KODE PROGRAM

Kode Program 4. 1 Dataset	28
Kode Program 4. 2 Pemisahan Fitur dan Label	29
Kode Program 4. 3 Pemanggilan Label	29
Kode Program 4. 4 Transformasi Data	30
Kode Program 4. 5 Normalisasi	31
Kode Program 4. 6 <i>Split</i> Data 70%:30%.....	32
Kode Program 4. 7 <i>Split</i> Data 80%:20%.....	32
Kode Program 4. 8 <i>Split</i> Data 90%:10%.....	32
Kode Program 4. 9 Proses <i>SMOTE</i>	32
Kode Program 4. 12 Model <i>SVM Kernel Linear</i>	33
Kode Program 4. 11 Model <i>SVM Kernel RBF</i>	33
Kode Program 4. 10 Model <i>SVM Kernel Polynomial</i>	33
Kode Program 4. 13 Prediksi dengan Data <i>Testing</i>	33
Kode Program 4. 14 Evaluasi Sistem.....	34