

TUGAS AKHIR

PENGARUH METODE *RANDOM UNDER SAMPLING* *RANDOM OVER SAMPLING* DAN *SMOTE* TERHADAP KINERJA *SUPPORT VECTOR MACHINE* UNTUK ANALISIS SENTIMEN



AGIL DWI SAPUTRA
2110651047

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
2025**

TUGAS AKHIR

PENGARUH METODE *RANDOM UNDER SAMPLING* *RANDOM OVER SAMPLING* DAN *SMOTE* TERHADAP KINERJA *SUPPORT VECTOR MACHINE* UNTUK ANALISIS SENTIMEN

Disusun Untuk Melengkapi Dan Memenuhi Syarat Kelulusan
Guna Meraih Gelar Sarjana Komputer
Teknik Informatika Universitas Muhammadiyah Jember



AGIL DWI SAPUTRA
2110651047

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
2025**

HALAMAN PERSETUJUAN

PENGARUH METODE *RANDOM UNDER SAMPLING* *RANDOM OVER SAMPLING* DAN *SMOTE* TERHADAP KINERJA *SUPPORT VECTOR MACHINE* UNTUK ANALISIS SENTIMEN

Agil Dwi Saputra

2110651047

Telah disetujui bahwa laporan Tugas Akhir ini untuk diajukan pada
Sidang Tugas Akhir sebagai salah satu syarat kelulusan dan
mendapatkan gelar Sarjana Komputer (S.Kom) di
Universitas Muhammadiyah Jember

Universitas Muhammadiyah Jember

Jember, 19 Juli 2025

Menyetujui,

Pembimbing I

Pembimbing II



Deni Arifianto. M.Kom.

NIDN. 0718068103



Dr. Reni Umilasari. S.Pd. M.Si

NIDN. 0728079101

HALAMAN PENGESAHAN

PENGARUH METODE *RANDOM UNDER SAMPLING* *RANDOM OVER SAMPLING* DAN *SMOTE* TERHADAP KINERJA *SUPPORT VECTOR MACHINE* UNTUK ANALISIS SENTIMEN

Oleh:
Agil Dwi Saputra
2110651047

Telah mempertanggung jawabkan Laporan Tugas Akhir pada sidang
Tugas Akhir 19 Juli 2025 sebagai salah satu syarat kelulusan dan mendapatkan
gelar Sarjana Komputer (S.Kom)
di
Universitas Muhammadiyah Jember

Disetujui oleh,

Dosen Penguji:
Penguji I



Dudi Irawan, S.T., M.Kom
NIDN. 0730037703

Dosen Pembimbing:
Pembimbing I



Deni Arifianto, M.Kom.
NIDN. 0718068103

Penguji II



Luluk Handayani, S.Si., M.Si.
NIDN. 0725108003

Pembimbing II



Dr. Reni Umilasari, S.Pd., M.Si
NIDN. 0728079101

Mengesahkan,
Dekan Fakultas Teknik



Dr. Ir. Muhtar. S.T., M.T., IPM
NIDN. 0010067301

Mengetahui,
Ketua Program Studi Teknik
Informatika



Rosita Yandarti, S.Kom., M.Cs
NIDN. 0629018601

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

NAMA : AGIL DWI SAPUTRA
NIM : 2110651047
INSTITUSI : Strata-1 Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember

Dengan ini saya menyatakan yang sesungguhnya bahwa karya ilmiah Tugas Akhir berjudul **“PENGARUH METODE *RANDOM UNDER SAMPLING* *RANDOM OVER SAMPLING* DAN *SMOTE* TERHADAP KINERJA *SUPPORT VECTOR MACHINE* UNTUK ANALISIS SENTIMEN”** adalah karya ilmiah saya sendiri, kecuali beberapa kutipan yang sudah saya sebutkan sumbernya. Sepanjang pengetahuan saya, karya ilmiah ini belum pernah diajukan pada institusi mana pun, serta bukan karya plagiat atau jiplakan. Saya bertanggung jawab atas keaslian, keabsahan, dan kebenaran isinya sesuai dengan sikap yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa adanya tekanan dan paksaan dari pihak mana pun, serta saya bersedia menerima sanksi akademik apabila di kemudian hari pernyataan yang saya buat adalah tidak benar.

Jember, 19 Juli 2025

Yang membuat pernyataan



Agil Dwi Saputra

NIM. 2110651047

MOTTO

“Tidak perlu sempurna, tapi harus lebih baik”

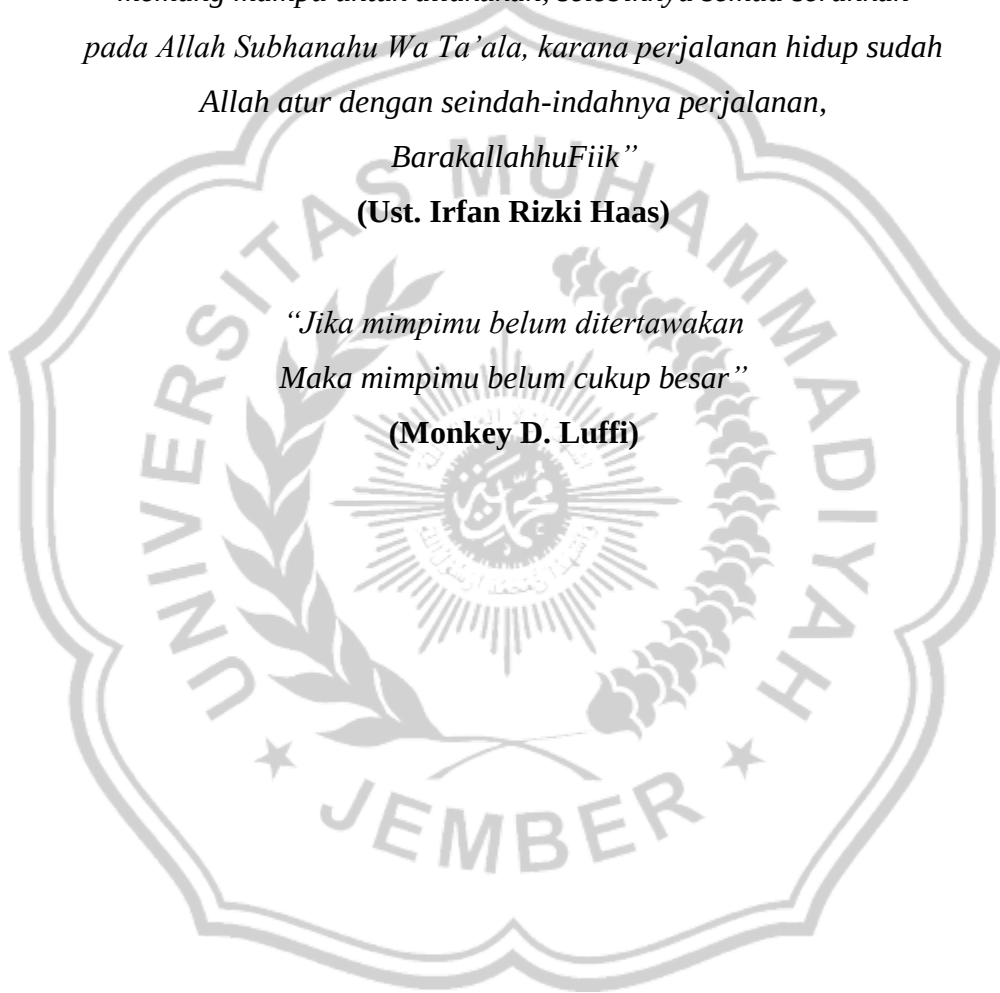
(Kalimasada)

“Biarkan semua berjalan apa adanya, lakukan sesuatu yang memang mampu untuk dilakukan, selebihnya semua serahkan pada Allah Subhanahu Wa Ta’ala, karena perjalanan hidup sudah Allah atur dengan seindah-indahnya perjalanan, BarakallahhuFiik”

(Ust. Irfan Rizki Haas)

*“Jika mimpimu belum ditertawakan
Maka mimpimu belum cukup besar”*

(Monkey D. Luffi)



KATA PENGANTAR

Segala puji dan syukur ke hadirat Allah Subhaanahu wa Ta'ala atas limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi ini dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana di Universitas Muhammadiyah Jember. Penulis menyadari bahwa penyusunan tugas akhir ini tidak akan berjalan lancar tanpa bantuan, dukungan, dan doa dari berbagai pihak. Oleh karena itu, penulis menyampaikan rasa terima kasih yang mendalam kepada semua pihak yang telah memberikan kontribusinya, diantaranya:

1. Kepada Allah Subhaanahu wa Ta'ala, atas segala limpahan rahmat, berkah, serta petunjuk-Nya yang senantiasa menyertai penulis hingga tugas akhir ini dapat diselesaikan dengan baik.
2. Kepada Nabi Muhammad Shalallahu 'alaihi wa sallam, suri teladan sepanjang zaman, atas segala ajaran dan tuntunan yang tak pernah terputus. Semoga kita senantiasa mendapat syafaat beliau di akhir nanti. Aaamiinn.
3. Kepada kedua orang tua tercinta, Bapak Kiswanto dan Ibu Yanti Ariyani, yang selalu setia mendampingi dengan doa, kasih sayang, semangat, serta dukungan moral dan spiritual yang tak pernah surut. Terima kasih atas segala pengorbanan dan kekuatan yang telah menjadi pendorong utama bagi saya dalam menghadapi berbagai rintangan selama menyelesaikan tugas akhir ini.
4. Kepada saudara dan keluarga tercinta, yang selalu hadir dengan doa, dukungan, dan semangat yang tak henti-hentinya. Terima kasih atas perhatian, kebersamaan, dan motivasi yang telah diberikan selama proses penyusunan tugas akhir ini.
5. Kepada pemilik NIM 2110411109 (EH), terimakasih atas dukungan yang tak henti-hentinya kamu berikan selama proses pengerjaan tugas akhir ini. Di setiap lelah dan keraguan, kehadiranmu menjadi sumber motivasi yang membuat saya terus berjuang dan tidak menyerah. Terima kasih telah setia membersamai proses ini dengan doa dan semangatmu. Semoga segala kebaikanmu selalu dibalas dengan yang terbaik oleh Allah Subhaanahu wa

Ta'ala, dan semoga langkah-langkah kita selalu seirama bukan hanya dalam perjuangan ini, tetapi juga dalam perjalanan panjang kehidupan ke depannya.

6. Kepada seluruh teman dan sahabat Teknik Informatika Universitas Muhammadiyah Jember angkatan 21, terimakasih atas kebersamaan, kerja sama, tawa, dan semangat yang telah kita bagi selama masa perkuliahan. Dukungan serta motivasi kalian menjadi bagian penting dalam perjalanan saya hingga sampai pada tahap ini. Semoga persaudaraan dan silaturahmi ini tetap terjaga dengan baik di masa yang akan datang
7. Kepada Dosen pembimbing, Bapak Deni Arifianto M.Kom. dan Ibu Dr. Reni Umilasari S.Pd, M.Si, terima kasih atas bimbingan, arahan, dan waktu yang telah diberikan selama proses penyusunan skripsi ini. Bantuan dan masukan yang diberikan sangat berarti dan membantu saya dalam menyelesaikan tugas akhir ini dengan lebih baik.
8. Semua pihak yang tidak mungkin disebutkan satu per satu, yang telah memberikan dukungan, bantuan, serta doa dalam bentuk apa pun selama proses penyusunan skripsi ini. Terima kasih atas segala kontribusi yang telah diberikan.

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN.....	v
MOTTO.....	vi
KATA PENGANTAR	vii
ABSTRAK	ix
ABSTRACT.....	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian.....	3
1.4 Batasan Penelitian	3
1.5 Manfaat Penelitian.....	3
BAB II TINJAUAN PUSTAKA.....	5
2.1 Analisis Sentimen.....	5
2.2 <i>Text mining</i>	5
2.3 <i>TF-IDF</i>	6
2.4 Klasifikasi.....	8
2.5 <i>Support Vector Machine (SVM)</i>	8
2.6 <i>Imbalance Data</i>	10
2.5.1 <i>Random Under Sampling (RUS)</i>	10
2.5.2 <i>Random Over Sampling (ROS)</i>	11
2.5.3 <i>Synthetic Minority Oversampling Technique (SMOTE)</i>	11
2.7 <i>Confusion matrix</i>	12
2.8 <i>K-Fold Cross Validation</i>	13
2.9 Fungsi Kernel	14
2.10 <i>X (Twitter)</i>	14
2.11 Naturalisasi.....	15
2.12 Sepak Bola.....	15
2.13 Penelitian Terdahulu.....	16
BAB III METODE PENELITIAN.....	18
3.1 Tahapan penelitian.....	18
3.2 Pengumpulan Data.....	18
3.3 Pelabelan Data	19
3.4 <i>Text Preprocessing</i>	20
3.5 Pembobotan <i>TF-IDF</i>	23
3.6 <i>Balancing Data</i> dengan <i>RUS</i>	27
3.7 <i>Balancing Data</i> dengan <i>ROS</i>	28
3.8 <i>Balancing Data</i> dengan <i>SMOTE</i>	29

3.9	Implementasi <i>Support Vector Machine</i> tanpa Proses <i>Balancing</i>	31
3.10	Pengujian <i>Support Vector Machine</i> tanpa Proses <i>Balancing</i>	37
3.11	Evaluasi Model	38
BAB IV IMPLEMENTASI DAN HASIL		40
4.1	Pengumpulan Data Dengan <i>Scraping</i>	40
4.2	Pelabelan Data	40
4.3	Text Preprocessing	41
4.4	Penerapan <i>TF-IDF</i>	43
4.5	Implementasi <i>SVM</i> Dengan <i>Balancing RUS</i>	45
4.6	Implementasi <i>SVM</i> Dengan <i>Balancing ROS</i>	50
4.7	Implementasi <i>SVM</i> Dengan <i>Balancing SMOTE</i>	56
4.8	Implementasi <i>SVM</i> Tanpa <i>Balancing</i>	61
4.9	Evaluasi dan Perbandingan Kinerja Model	66
BAB V KESIMPULAN		71
5.1	Kesimpulan	71
5.2	Saran	72
DAFTAR PUSTAKA		73
LAMPIRAN		76



DAFTAR GAMBAR

Gambar 3. 1 Tahapan Penelitian	18
Gambar 3. 2 <i>Flowchart Preprocessing</i>	20
Gambar 3. 3 <i>Flowchart RUS</i>	28
Gambar 3. 4 <i>Flowchart ROS</i>	29
Gambar 3. 5 <i>Flowchart SMOTE</i>	30
Gambar 4. 1 <i>Script proses Scraping</i>	40
Gambar 4. 2 Hasil <i>Scraping</i>	40
Gambar 4. 3 <i>Script Proses Cleaning</i>	41
Gambar 4. 4 Hasil <i>Cleaning</i>	42
Gambar 4. 5 <i>Script Proses Case Folding</i>	42
Gambar 4. 6 Hasil <i>Case Folding</i>	42
Gambar 4. 7 <i>Script Proses Tokenizing</i>	42
Gambar 4. 8 Hasil <i>Tokenizing</i>	42
Gambar 4. 9 <i>Script Proses Stopword Removal</i>	43
Gambar 4. 10 Hasil <i>Stopword Removal</i>	43
Gambar 4. 11 <i>Script Proses Stemming</i>	43
Gambar 4. 12 Hasil <i>Stemming</i>	43
Gambar 4. 13 <i>Script Proses TF</i>	43
Gambar 4. 14 Hasil <i>TF</i>	44
Gambar 4. 15 <i>Script Proses IDF</i>	44
Gambar 4. 16 Hasil <i>IDF</i>	44
Gambar 4. 17 <i>Script Proses TF-IDF</i>	44
Gambar 4. 18 Hasil <i>TF-IDF</i>	45
Gambar 4. 19 <i>Proses Balancing Dengan RUS</i>	45
Gambar 4. 20 Perbandingan <i>F1-score, Precision, Recall</i> Pada <i>RUS</i>	50
Gambar 4. 21 Grafik Perbandingan <i>Accuracy</i> Langkah Uji Pada <i>RUS</i>	50
Gambar 4. 22 <i>Proses Balancing Dengan ROS</i>	51
Gambar 4. 23 Perbandingan <i>F1-score, Precision, Recall</i> Pada <i>ROS</i>	55
Gambar 4. 24 Grafik Perbandingan <i>Accuracy</i> Langkah Uji Pada <i>ROS</i>	55
Gambar 4. 25 <i>Proses Balancing Dengan SMOTE</i>	56
Gambar 4. 26 Perbandingan <i>F1-score, Precision, Recall</i> Pada <i>SMOTE</i>	61
Gambar 4. 27 Grafik Perbandingan <i>Accuracy</i> Langkah Uji Pada <i>SMOTE</i>	61
Gambar 4. 28 <i>Proses SVM Tanpa Balancing</i>	62
Gambar 4. 29 Perbandingan <i>F1-score, Precision, Recall</i> Pada <i>SVM</i>	66
Gambar 4. 30 Grafik Perbandingan <i>Accuracy</i> Langkah Uji Pada <i>SVM</i>	66
Gambar 4. 31 <i>Confusion matrix</i> Model <i>SVM</i> Dengan <i>ROS</i>	68

DAFTAR TABEL

Tabel 2. 1 Rumus <i>Confusion Matrix</i>	12
Tabel 2. 2 Penelitian Terdahulu	16
Tabel 3. 1 Data Hasil <i>Scraping</i>	19
Tabel 3. 2 Pelabelan Data Latih	19
Tabel 3. 3 Data Uji	20
Tabel 3. 4 Data Hasil <i>Cleaning</i>	21
Tabel 3. 5 Data Hasil <i>Case Folding</i>	21
Tabel 3. 6 Data Hasil <i>Tokenizing</i>	22
Tabel 3. 7 Data Hasil <i>Stopword Removal</i>	22
Tabel 3. 8 Data Hasil <i>Stemming</i>	23
Tabel 3. 9 Nilai <i>Term Frequency</i>	24
Tabel 3. 10 Nilai <i>TF norm</i>	25
Tabel 3. 11 Nilai <i>DF, IDF, dan TF-IDF</i>	26
Tabel 3. 12 Data Sebelum <i>RUS</i>	28
Tabel 3. 13 Data Setelah <i>RUS</i>	28
Tabel 3. 14 Data Sebelum <i>ROS</i>	29
Tabel 3. 15 Data Setelah <i>ROS</i>	29
Tabel 3. 16 Data Sebelum <i>SMOTE</i>	30
Tabel 3. 17 Data Setelah <i>SMOTE</i>	31
Tabel 3. 18 Data Latih	31
Tabel 3. 19 Data Vektorisasi	31
Tabel 3. 20 Perhitungan Fungsi Kernel	32
Tabel 3. 21 Hasil Perhitungan Fungsi Kernel	32
Tabel 3. 22 Hasil Perhitungan <i>Matrix</i>	33
Tabel 3. 23 Hasil Perhitungan Nilai Error	33
Tabel 3. 24 Hasil Perhitungan <i>Delta Alpha</i>	34
Tabel 3. 25 Hasil Nilai <i>Alpha</i> Baru	34
Tabel 3. 26 Nilai <i>W Positif, Negatif, dan Netral</i>	35
Tabel 3. 27 Hasil Nilai <i>Bias</i>	36
Tabel 3. 28 Hasil Perhitungan Sentimen	36
Tabel 3. 29 Data Uji	37
Tabel 3. 30 Data Vektorisasi Data Uji	37
Tabel 3. 31 Hasil Perhitungan Fungsi Kernel Data Uji	37
Tabel 3. 32 Nilai Perhitungan Sentimen Data Uji	38
Tabel 3. 33 Hasil Perhitungan Sentimen Data Uji	38
Tabel 4. 1 Hasil Pelabelan Data	41
Tabel 4. 2 Data Sebelum <i>RUS</i> dan Sesudah <i>RUS</i>	46
Tabel 4. 3 Hasil Evaluasi Model SVM Dengan <i>Balancing RUS</i>	47
Tabel 4. 4 Data Sebelum <i>ROS</i> dan Sesudah <i>ROS</i>	51
Tabel 4. 5 Hasil Evaluasi Model SVM Dengan <i>Balancing ROS</i>	52
Tabel 4. 6 Data Sebelum <i>SMOTE</i> dan Sesudah <i>SMOTE</i>	56
Tabel 4. 7 Hasil Evaluasi Model SVM Dengan <i>Balancing SMOTE</i>	57
Tabel 4. 8 Data Sebelum SVM Tanpa <i>Balancing</i>	62
Tabel 4. 9 Hasil Evaluasi Model SVM Tanpa <i>Balancing</i>	63
Tabel 4. 10 Hasil Evaluasi Untuk Setiap Metode Perlabel	67

DAFTAR LAMPIRAN

Lampiran 1. Hasil <i>Confusion Matrix</i>	76
Lampiran 2. Dokumentsai.....	78
Lampiran 3. Data Validasi	79

