

**KOMPARASI METODE *RANDOM FOREST* DENGAN
SUPPORT VECTOR MACHINE PADA SENTIMEN
ANALISIS KOMENTAR KUESIONER
KEPUASAN MAHASISWA (KKM)
ITB STIKOM BALI**



**PROGRAM STUDI ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS PENDIDIKAN GANESHA
2025**

**KOMPARASI METODE *RANDOM FOREST* DENGAN
SUPPORT VECTOR MACHINE PADA SENTIMEN
ANALISIS KOMENTAR KUESIONER
KEPUASAN MAHASISWA (KKM)
ITB STIKOM BALI**



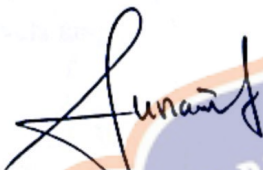
**PROGRAM STUDI ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS PENDIDIKAN GANESHA
2025**

LEMBAR PERSETUJUAN PEMBIMBING

Tesis oleh Purnama Sidik ini telah diperiksa dan disetujui untuk mengikuti Ujian Tesis

Singaraja, 31 Juli 2025

Pembimbing 1



Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs.

NIP: 198307252008011008

Pembimbing 2



Dr. I Gede Aris Gunadi, S.Si., M.Kom.

NIP: 19770318200812100

LEMBAR PERSETUJUAN TIM PENGUJI

Tesis oleh Purnama Sidik ini telah dipertahankan di depan tim penguji dan dinyatakan diterima sebagai salah satu persyaratan untuk memperoleh gelar Magister Komputer di Program Studi Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha.

Disetujui pada tanggal: 31 Juli 2025

Oleh

Tim Penguji

Ketua (Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs.)

NIP: 198307252008011008

Anggota (Dr. I Gede Aris Gunadi, S.Si. M.Kom.)

NIP: 197703182008121004

Anggota (Prof. Dr. I Made Candiasa, M.I.Kom.)

NIP: 196012311986011004

Anggota (Dr. Putu Hendra Suputra, S.Kom., M.Cs.)

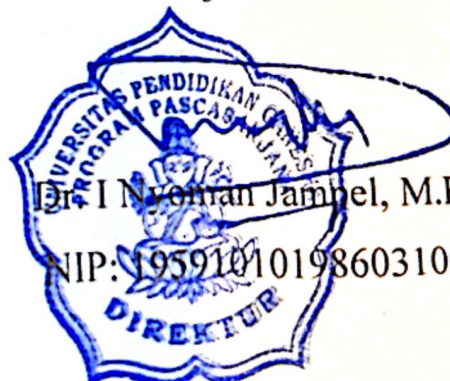
NIP: 198212222006041001

Mengetahui Direktur

Pascasarjana Undiksha,

Dr. I Nyoman Jambel, M.Pd.

NIP: 195910101986031003



LEMBAR PERNYATAAN

Saya menyatakan dengan sesungguhnya bahwa tesis yang saya susun sebagai syarat untuk memperoleh gelar Magister Komputer dari Pascasarjana Universitas Pendidikan Ganesha seluruhnya merupakan hasil karya saya sendiri. Bagian-bagian tertentu dalam penulisan tesis yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas dan sesuai dengan norma, kaidah, serta etika akademis.

Apabila di kemudian hari ditemukan seluruh atau sebagian tesis ini bukan hasil karya saya sendiri atau adanya plagiat dalam bagianbagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sandang dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku di wilayah Negara Kesatuan Republik Indonesia.

Singaraja, 31 Juli 2025

Yang membuat pernyataan


10000
METERAI
TEMPEL
0DD2DAMX449215099
(Purnama Sidik)

PRAKATA

Puji dan syukur dipanjatkan kepada Tuhan Yang Maha Esa atas anugrah-Nya, sehingga tesis yang berjudul: “Komparasi Metode *Random Forest* Dengan *Support Vector Machine* Pada Sentimen Analisis Komentar Kuesioner Kepuasan Mahasiswa (KKM) ITB STIKOM Bali”, dapat diselesaikan sesuai dengan yang direncanakan.

Tesis ini ditulis untuk memenuhi salah satu persyaratan untuk memperoleh gelar Magister Komputer Pascasarjana Universitas Pendidikan Ganesha pada Program Studi Ilmu Komputer. terselesaikannya tesis ini telah banyak memperoleh uluran tangan dari berbagai pihak. Untuk itu, izinkan penulis menyampaikan terima kasih dan penghargaan kepada pihak-pihak berikut.

1. Bapak Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs., sebagai pembimbing I yang telah dengan sabar membimbing, mengarahkan, dan memberikan motivasi yang demikian bermakna, sehingga penulis mampu melewati berbagai hambatan dalam perjalanan studi dan penyelesaian tesis ini;
2. Bapak Dr. I Gede Aris Gunadi, S.Si., M.Kom., sebagai pembimbing II, yang dengan gaya dan pola komunikasi yang khas, telah melecut semangat, motivasi, dan harapan penulis selama penelitian dan penulisan naskah laporan tesis ini, sehingga tesis ini dapat terwujud dengan baik sesuai harapan;
3. Bapak Prof. Dr. I Made Candiasa, M.I.Kom. dan Dr. Putu Hendra Suputra, S.Kom., M.Cs. sebagai penguji yang telah banyak memberikan masukan-masukan yang bermanfaat untuk penyempurnaan tesis ini;
4. Rektor Universitas Pendidikan Ganesha, yang telah memberikan bantuan secara moral dan memfasilitasi berbagai kepentingan penulis dalam menyelesaikan tesis ini;
5. Direktur Pascasarjana Undiksha dan staf, yang telah banyak membantu selama penulis menyelesaikan tesis ini;

6. Koordinator Program Studi Ilmu Komputer dan staf dosen pengajar yang telah banyak membantu dan memotivasi penulis selama penyusunan tesis ini;
7. Rekan-rekan seangkatan di Program Studi Ilmu Komputer yang dengan karakternya masing-masing telah banyak berkontribusi membentuk kedirian penulis selama menjalani studi dan penyelesaian tesis ini;
8. Bapak dan Ibu selaku orang tua penulis, yang telah banyak membantu secara material dan moral selama penyelesaian tesis ini;
9. Keluarga besar ITB STIKOM Bali yang telah banyak berkontribusi selama penulis menyelesaikan tesis ini.

Semoga semua bantuan yang telah mereka berikan dalam menyelesaikan studi ini, mereka diberkati imbalan yang sepadan oleh Tuhan Yang Maha Esa, kesehatan, dan keharmonian dalam menjalani kehidupan. Penulis menyadari bahwa tesis ini belum sempurna. Namun, kehadirannya dalam konstelasi masyarakat akademis akan menambah perbendaharaan ilmu dalam perkembangan ilmu pengetahuan. Semoga tesis ini bermanfaat bagi masyarakat akademis, terutama mereka yang menyatakan diri bernaung di bawah kebesaran panji-panji pendidikan.

Singaraja, 31 Juli 2025

Penulis

DAFTAR ISI

	(halaman)
LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PERSETUJUAN TIM PENGUJI.....	ii
LEMBAR PERNYATAAN.....	iii
PRAKATA	iv
ABSTRAK.....	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah.....	3
1.3 Pembatasan Masalah.....	4
1.4 Rumusan Masalah.....	4
1.5 Tujuan Penelitian.....	5
1.6 Manfaat Penelitian.....	5
1.6.1 Bagi Penulis.....	5
1.6.2. Bagi Akademik	5
1.6.3. Bagi Pihak Penjaminan Mutu	6
BAB II KAJIAN PUSTAKA	7
2.1 Kajian Teori.....	7
2.1.1 Analisis Sentimen	7
2.1.2 <i>Knowledge Discovery in Database (KDD)</i>	7
2.1.3 <i>Machine Learning</i>	8

2.1.4	<i>Data Mining</i>	9
2.1.5	<i>Text Mining</i>	10
2.1.6	<i>Term Frequency–Inverse Document Frequency (TF-IDF)</i>	10
2.1.7	<i>Preprocessing</i>	11
2.1.8	<i>Augmentation Text</i>	13
2.1.9	<i>K-Fold Cross Validation</i>	14
2.1.10	<i>Random Forest</i>	14
2.1.11	<i>Support Vector Machine</i>	16
2.1.12	<i>Confusion Matrix</i>	17
2.1.13	<i>Wordcloud</i>	18
2.1.14	<i>Kurva Receiver Operating Characteristics</i>	18
2.2	Kajian Hasil Penelitian yang Relevan	19
2.3	Kerangka Konsep	23
2.4	Hipotesis.....	24
BAB III METODE PENELITIAN		26
3.1	Rancangan Penelitian	26
3.2	Simulasi Penelitian.....	27
3.2.1	Pengumpulan Data.....	27
3.2.2	<i>Labeling Data</i>	27
3.2.3	<i>Balance/Imbalance Data</i>	27
3.2.4	<i>Preprocessing</i>	29
3.2.5	<i>Transformation</i>	31
3.2.6	<i>Data Mining</i>	33
3.2.7	<i>Data Pure dan Cross Validation</i>	36
3.2.8	<i>Evaluation</i>	37
3.2.9	Visualisasi.....	38
3.2.10	Pengujian Hipotesis	38
BAB IV HASIL PENELITIAN & PEMBAHASAN		39
4.1	Persiapan Penelitian	39
4.2	Pengumpulan Data	40

4.3	<i>Labeling Data</i>	40
4.4	<i>Balance/Imbalance Data</i>	40
4.4	<i>Preprocessing</i>	43
4.5	<i>Transformation</i>	48
4.6	<i>Data Mining, Evaluation dan Visualisasi</i>	54
4.6.1	<i>Random Forest</i>	54
4.6.2	<i>Support Vector Machine</i>	70
4.6.3	<i>WordCloud</i>	84
4.7	<i>Pembahasan</i>	86
4.7.1	<i>Pembahasan Perbandingan 70:30</i>	86
4.7.2	<i>Pembahasan Perbandingan 80:20</i>	88
4.7.3	<i>Pembahasan Perbandingan 90:10</i>	90
4.8	<i>Hipotesis</i>	91
4.9	<i>Kontribusi Penelitian Terhadap Ilmu Komputer</i>	92
BAB V PENUTUP		93
5.1	<i>Kesimpulan</i>	93
5.2	<i>Saran</i>	94
DAFTAR PUSTAKA		96
LAMPIRAN		101
DAFTAR RIWAYAT HIDUP		105

DAFTAR TABEL

	<i>(halaman)</i>
Tabel 2.1 <i>Confusion Matrix</i>	18
Tabel 2.2 Kajian Hasil Penelitian yang Relevan.....	19
Tabel 3.1 Contoh <i>Augmentation Text</i>	28
Tabel 3.2 Contoh <i>Filtering and Cleansing</i>	29
Tabel 3.3 Contoh <i>Tokenizing</i>	29
Tabel 3.4 Contoh <i>Case Folding</i>	30
Tabel 3.5 Contoh <i>Stopword</i>	30
Tabel 3.6 Contoh <i>Stemming</i>	30
Tabel 3.7 Contoh Hasil <i>Stemming</i>	31
Tabel 3.8 Hasil TF-IDF	33
Tabel 3.9 Contoh <i>Confusion Matrix</i>	37
Tabel 4.1 Contoh <i>Back Translation</i>	41
Tabel 4.2 Tabel <i>Filtering and Cleansing</i>	43
Tabel 4.3 Tabel <i>Case Folding</i>	44
Tabel 4.4 Tabel <i>Tokenizing</i>	45
Tabel 4.5 Tabel <i>Stopword</i>	46
Tabel 4.6 Tabel <i>Stemming</i>	47
Tabel 4.7 Dimensi TF-IDF 70:30.....	48
Tabel 4.8 Dimensi TF-IDF 80:20.....	49
Tabel 4.9 Dimensi TF-IDF 90:10.....	49
Tabel 4.10 Tabel Contoh Hasil <i>Stemming</i>	50
Tabel 4.11 Hasil TF-IDF	53
Tabel 4.12 Hasil Akurasi <i>Random Forest</i> 70:30.....	56
Tabel 4.13 <i>Cross Validation</i> Sebelum Augment <i>Random Forest</i> 70:30.....	57

Tabel 4.14 <i>Cross Validation</i> Setelah Augment <i>Random Forest</i> 70:30.....	58
Tabel 4.15 Hasil Akurasi <i>Random Forest</i> 80:20.....	61
Tabel 4.16 <i>Cross Validation</i> Sebelum Augment <i>Random Forest</i> 80:20.....	62
Tabel 4.17 <i>Cross Validation</i> Setelah Augment <i>Random Forest</i> 80:20.....	63
Tabel 4.18 Hasil Akurasi <i>Random Forest</i> 90:10.....	66
Tabel 4.19 <i>Cross Validation</i> Sebelum Augment <i>Random Forest</i> 90:10.....	67
Tabel 4.20 <i>Cross Validation</i> Setelah Augment <i>Random Forest</i> 90:10.....	68
Tabel 4.21 Hasil Akurasi SVM 70:30.....	71
Tabel 4.22 <i>Cross Validation</i> Sebelum Augment <i>Support Vector Machine</i> 70:30	72
Tabel 4.23 <i>Cross Validation</i> Setelah Augment <i>Support Vector Machine</i> 70:30	73
Tabel 4.24 Hasil Akurasi SVM 80:20.....	76
Tabel 4.25 <i>Cross Validation</i> Sebelum Augment <i>Support Vector Machine</i> 80:20	77
Tabel 4.26 <i>Cross Validation</i> Setelah Augment <i>Support Vector Machine</i> 80:20	78
Tabel 4.27 Hasil Akurasi SVM 90:10.....	81
Tabel 4.28 <i>Cross Validation</i> Sebelum Augment <i>Support Vector Machine</i> 90:10	82
Tabel 4.29 <i>Cross Validation</i> Setelah Augment <i>Support Vector Machine</i> 90:10	83



DAFTAR GAMBAR

	<i>(halaman)</i>
Gambar 2.1 Tahapan dalam melakukan <i>preprocessing</i>	12
Gambar 2.2 Model <i>Random Forest</i>	15
Gambar 2.3 <i>Support Vector Machine</i>	17
Gambar 2.4 Ilustrasi Kerangka Konsep	23
Gambar 3.1 <i>Flowchart</i> rancangan penelitian	26
Gambar 3.2 Contoh <i>Augmentation Text</i> sebelum dan sesudah	28
Gambar 4.1 Grafik Sebelum dan Sesudah Diseimbangkan	41
Gambar 4.2 <i>Confusion Matrix Before Augmentation Random Forest 70:30</i>	55
Gambar 4.3 <i>Confusion Matrix After Augmentation Random Forest 70:30</i>	55
Gambar 4.4 Kurva ROC 70:30 <i>Random Forest</i>	59
Gambar 4.5 <i>Confusion Matrix Before Augmentation Random Forest 80:20</i>	60
Gambar 4.6 <i>Confusion Matrix After Augmentation Random Forest 80:20</i>	60
Gambar 4.7 Kurva ROC 80:20 <i>Random Forest</i>	64
Gambar 4.8 <i>Confusion Matrix Before Augmentation Random Forest 90:10</i>	65
Gambar 4.9 <i>Confusion Matrix After Augmentation Random Forest 90:10</i>	65
Gambar 4.10 Kurva ROC <i>Random Forest 90:10</i>	69
Gambar 4.11 <i>Confusion Matrix Before Augmentation SVM 70:30</i>	70
Gambar 4.12 <i>Confusion Matrix After Augmentation SVM 70:30</i>	70
Gambar 4.13 Kurva ROC 70:30 <i>SVM</i>	74
Gambar 4.14 <i>Confusion Matrix Before Augmentation SVM 80:20</i>	75
Gambar 4.15 <i>Confusion Matrix After Augmentation SVM 80:20</i>	75
Gambar 4.16 Kurva ROC 80:20 <i>SVM</i>	79
Gambar 4.17 <i>Confusion Matrix Before Augmentation SVM 90:10</i>	80
Gambar 4.18 <i>Confusion Matrix After Augmentation SVM 90:10</i>	80

Gambar 4.19 Kurva ROC 90:10 SVM.....	84
Gambar 4.20 <i>WordCloud Positive</i> Sebelum <i>Augment</i>	85
Gambar 4.21 <i>WordCloud Negative</i> Sebelum <i>Augment</i>	85
Gambar 4.22 <i>WordCloud Positive</i> Setelah <i>Augment</i>	86
Gambar 4.23 <i>WordCloud Negative</i> Setelah <i>Augment</i>	86



DAFTAR LAMPIRAN

(halaman)

Lampiran 1 Validasi Ahli Psikologi.....	101
Lampiran 2 Validasi Ahli Bahasa Indonesia.....	102
Lampiran 3 Contoh <i>Back Translation</i>	103

