

PERBANDINGAN KINERJA NAIVE BAYES DAN SVM BERBASIS SMOTE DALAM ANALISIS SENTIMEN KOMENTAR YOUTUBE MENGENAI WACANA REDENOMINASI RUPIAH

TESIS



**PROGRAM STUDI ILMU KOMPUTER (S2)
PROGRAM PASCASARJANA
UNIVERSITAS PENDIDIKAN GANESHA**



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.

TESIS

DIAJUKAN UNTUK MELENGKAPI TUGAS DAN MEMENUHI SYARAT-SYARAT UNTUK MENCAPAI GELAR MAGISTER KOMPUTER

Menyetujui

Pembimbing I	Dr. I Gede Aris Gunadi, S.Si. M.Kom. NIP.197703182008121004
Pembimbing II	Prof. Dr. Luh Joni Erawati Dewi, S.T., M.Pd. NIP.197606252001122001



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.

Tesis oleh Kadek Kusuma Wardana telah dipertahankan di depan tim penguji dan dinyatakan diterima sebagai salah satu persyaratan untuk memperoleh gelar Magister Komputer di Ilmu Komputer (S2), Program Pascasarjana, Universitas Pendidikan Ganesha
Pada tanggal 03 Juli 2026

Dewan 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. Luh Joni Erawati Dewi, S.T., M.Pd. NIP.197606252001122001
Anggota	Kadek Yota Ernanda Aryanto, S.Kom., M.T., Ph.D. NIP.197803242005011001

Mengetahui Direktur Program Pascasarjana
Universitas Pendidikan Ganesha,



Prof. Dr. I Nyoman Jampel, M.Pd.

NIP.195910101986031003



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.

LEMBAR PERNYATAAN

Saya menyatakan dengan sesungguhnya bahwa tesis yang saya susun sebagai syarat untuk memperoleh gelar Magister Pendidikan 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 bagian bagian 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, 12 Juli 2026

Yang memberi pernyataan,



Kadek Kusuma Wardana

PRAKATA

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas anugrah-Nya, sehingga tesis yang berjudul: <PERBANDINGAN KINERJA NAIVE BAYES DAN SVM BERBASIS SMOTE DALAM ANALISIS SENTIMEN KOMENTAR YOUTUBE MENGENAI WACANA REDENOMINASI RUPIAH=, dapat diselesaikan sesuai dengan yang direncanakan.

Tesis ini ditulis untuk memenuhi salah satu persyaratan untuk memperoleh gelar Magister Pendidikan 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. Dr. I Gede Aris Gunadi, S.Si. M.Kom. 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. Prof. Dr. Luh Joni Erawati Dewi, S.T., M.Pd. 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. Kadek Yota Ernanda Aryanto, S.Kom., M.T., Ph.D. dan Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs. sebagai penguji yang telah banyak memberikan masukan-masukan yang bermanfaat untuk penyempurnaan tesis ini.
4. Para guru di SMA Negeri 11 Denpasar dan Guru di SMA Negeri 1 Ubud yang telah memberikan masukan dan ikut dalam penelitian ini.
5. Ketua Program Studi Ilmu Komputer dan staf dosen pengajar yang telah banyak membantu dan memotivasi penulis selama penyusunan tesis ini.
6. Direktur Pascasarjana Undiksha dan staf, yang telah banyak membantu selama penulis menyelesaikan tesis ini.
7. Rektor Universitas Pendidikan Ganesha, yang telah memfasilitasi berbagai kepentingan penulis dalam menyelesaikan tesis ini.
8. Keluarga, sahabat, dan berbagai pihak terkait yang telah memberikan dukungan baik secara langsung maupun tidak langsung dalam penyelesaian 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

Singaraja, 3 Juli 2026

Penulis

DAFTAR ISI

LEMBAR PERNYATAAN	ii
PRAKATA	iii
ABSTRAK	iv
ABSTRACT	v
DAFTAR ISI	vi
DAFTAR TABEL	ix
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Identifikasi Masalah	6
1.3. Pembatasan Masalah	6
1.4. Rumusan Masalah	7
1.5. Tujuan Penelitian	8
1.6. Manfaat Penelitian	8
1.6.1. Manfaat Teoritis	8
1.6.2. Manfaat Praktis	8
BAB II PEMBAHASAN	10
2.1. Kajian Teori	10
2.1.1 Analisis sentimen	10
2.1.2 Youtube	10
2.1.3 Text Preprocessing	11
2.1.4 Ekstraksi Fitur Teks	12
2.1.5 Naïve Bayes Classifier (NBC)	13
2.1.6 Support Vector Machine (SVM)	15
2.1.7 Synthetic Minority Over-sampling Technique (SMOTE)	16
2.1.8 Confusion Matrix	16
2.1.9 K-Fold Validation	18
2.1.10 AUC	18
2.2. Kajian Penelitian Yang Relevan	19
2.3. Kerangka Berpikir	22
BAB III METODE PENELITIAN	24
3.1. Metode Penelitian	24
3.2. Pengumpulan Data	25
3.3. Pelabelan	26
3.4. Preprocessing	27
3.5. Ekstraksi Fitur	32
3.6. Split Data	32
3.7. Balancing data	32
3.8. Klasifikasi	34
3.9. Evaluasi Hasil	35
BAB IV HASIL DAN PEMBAHASAN	36

4.1.	Pengumpulan Data	36
4.2.	Pelabelan.....	38
4.3.	Preprocessing.....	40
4.3.1.	Cleansing	40
4.3.2.	Case Folding.....	41
4.3.3.	Tokenization	43
4.3.4.	Normalization	44
4.3.5.	Stopword Removal	45
4.3.6.	Stemming.....	47
4.4.	Ekstraksi Fitur	48
4.5.	Split Data.....	50
4.6.	Balancing Data	52
4.7.	Klasifikasi.....	59
4.7.1.	Naive Bayes Tanpa SMOTE	60
4.7.2.	SVM Tanpa SMOTE.....	61
4.7.3.	Naive Bayes Dengan SMOTE.....	62
4.7.4.	SVM Dengan SMOTE	63
4.7.5.	Naive Bayes Dengan SMOTE Rasio 20:80	63
4.7.6.	SVM Dengan SMOTE Rasio 20:80	64
4.7.7.	Naive Bayes Dengan SMOTE Rasio 30:70	65
4.7.8.	SVM Dengan SMOTE Rasio 30:70	66
4.7.9.	Naive Bayes Dengan SMOTE Rasio 40:60	67
4.7.10.	SVM Dengan SMOTE Rasio 40:60	68
4.7.11.	Naive Bayes Dengan SMOTE Rasio 60:40	69
4.7.12.	SVM Dengan SMOTE Rasio 60:40	70
4.7.13.	Naive Bayes Dengan SMOTE Rasio 70:30	71
4.7.14.	SVM Dengan SMOTE Rasio 70:30	72
4.7.15.	Naive Bayes Dengan SMOTE Rasio 80:20	73
4.7.16.	SVM Dengan SMOTE Rasio 80:20	74
4.8.	Evaluasi Hasil.....	75
4.8.1.	Perbandingan Model Klasifikasi	76
4.8.1.1.	Pengujian Klasifikasi Naive Bayes	76
4.8.1.2.	Pengujian Klasifikasi SVM.....	79
4.8.1.3.	Perbandingan Hasil Pengujian Klasifikasi	82
4.8.2.	ROC dan AUC.....	87
4.8.2.1.	Pengujian ROC Dan AUC Naive Bayes	87
4.8.2.2.	Pengujian ROC Dan AUC SVM.....	91
4.8.2.3.	Hasil Pengujian ROC dan AUC	95
4.8.3.	Pengujian Data Overfitting.....	98
4.8.4.	Hasil K-Fold Cross Validation	102
4.8.4.1.	Pengujian K-Fold Cross Validation Naive Bayes	102
4.8.4.2.	Pengujian K-Fold Cross Validation SVM.....	104
4.8.4.3.	Hasil Pengujian K-Fold Cross Validation	106
4.8.5.	Visualisasi data dengan WordCloud	107
BAB V	PENUTUP.....	110
5.1.	Rangkuman.....	110

5.2.	Simpulan.....	112
5.3.	Saran.....	114
DAFTAR PUSTAKA.....		115
LAMPIRAN		119



DAFTAR TABEL

Tabel 2. 1 Tabel Data Kamus Slangword.....	12
Tabel 2. 2 Tabel Confusion Matrix	16
Tabel 3. 1 Contoh Proses Tabel Cleansing.....	27
Tabel 3. 2 Contoh Proses Tabel Case Folding	28
Tabel 3. 3 Contoh Proses Tabel Tokenization.....	29
Tabel 3. 4 Contoh Proses Tabel Normalization.....	30
Tabel 3. 5 Contoh Proses Tabel Stopword	31
Tabel 3. 6 Contoh Proses Tabel Stemming	32
Tabel 4. 1 Pengumpulan Data	37
Tabel 4. 2 Hasil Pemberian label.....	39
Tabel 4. 3 Hasil Tahapan Cleansing.....	41
Tabel 4. 4 Hasil Tahapan Case Folding	42
Tabel 4. 5 Hasil Tahapan Tokenization.....	43
Tabel 4. 6 Data kata baku.....	44
Tabel 4. 7 Hasil Tahapan Normalization.....	45
Tabel 4. 8 Hasil tahapan Stopword Removal.....	46
Tabel 4. 9 Hasil Tahapan Stemming	47
Tabel 4. 10 Hasil Ekstraksi Fitur TF-IDF Indeks 1.....	48
Tabel 4. 11 Hasil Ekstraksi Fitur TF-IDF Indeks 2.....	49
Tabel 4. 12 Hasil Ekstraksi Fitur TF-IDF Indeks 3.....	49
Tabel 4. 13 Hasil Ekstraksi Fitur TF-IDF Indeks 4.....	49
Tabel 4. 14 Hasil Ekstraksi Fitur TF-IDF Indeks 5.....	49
Tabel 4. 15 Hasil Ekstraksi Fitur TF-IDF Indeks 6.....	49
Tabel 4. 16 Persebaran data sesuai rasio	58
Tabel 4. 17 hasil evaluasi performa Naive Bayes	78
Tabel 4. 18 hasil evaluasi performa SVM.....	81
Tabel 4. 19 hasil evaluasi AUC keseluruhan.....	96

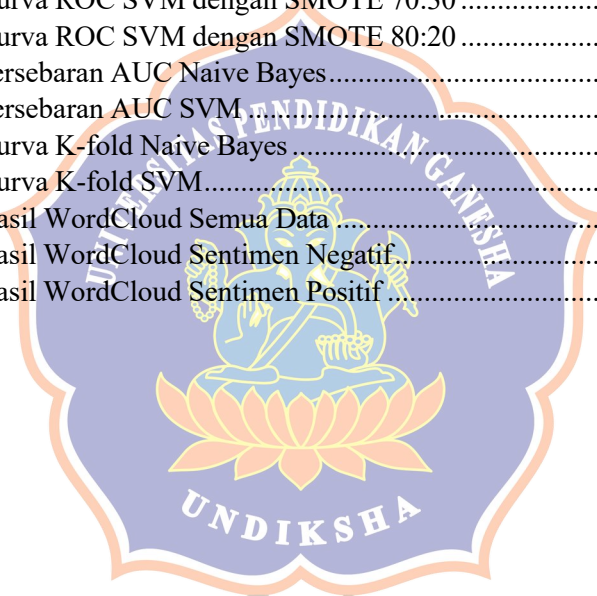
Tabel 4. 20 Hasil Evaluasi Data Training dan Testing.....	99
Tabel 4. 21 Hasil Pengujian K-Fold Naive Bayes.....	103
Tabel 4. 22 Hasil Pengujian K-Fold SVM	105



DAFTAR GAMBAR

Gambar 2. 1 Kerangka Berpikir	23
Gambar 3. 1 Tahapan Penelitian	25
Gambar 3. 2 Distribusi Sentimen	27
Gambar 3. 3 Rasio SMOTE	33
Gambar 4. 1 API Key Google Developer.....	36
Gambar 4. 2 distribusi sentimen.....	40
Gambar 4. 3 Jumlah Data Setelah Preprocessing	48
Gambar 4. 4 Hasil program TF-IDF	50
Gambar 4. 5 Hasil Split Data Latih	51
Gambar 4. 6 Hasil Split Data Uji	51
Gambar 4. 7 Hasil Resampling data SMOTE	53
Gambar 4. 8 Hasil Sebelum SMOTE	54
Gambar 4. 9 Hasil SMOTE dengan rasio 50:50.....	55
Gambar 4. 10 Hasil SMOTE dengan rasio 20 : 80.....	55
Gambar 4. 11 Hasil SMOTE dengan rasio 30 : 70.....	56
Gambar 4. 12 Hasil SMOTE dengan rasio 40 : 60.....	56
Gambar 4. 13 Hasil SMOTE dengan rasio 60 : 40.....	57
Gambar 4. 14 Hasil SMOTE dengan rasio 70 : 30.....	57
Gambar 4. 15 Hasil SMOTE dengan rasio 80 : 20.....	58
Gambar 4. 16 Confusion Matrix Naive Bayes	60
Gambar 4. 17 Confusion Matrix SVM.....	61
Gambar 4. 18 Confusion Matrix Naive Bayes dengan SMOTE	62
Gambar 4. 19 Confusion Matrix SVM dengan SMOTE	63
Gambar 4. 20 Confusion Matrix Naive Bayes SMOTE rasio 20:80	64
Gambar 4. 21 Confusion Matrix SVM SMOTE rasio 20:80.....	65
Gambar 4. 22 Confusion Matrix Naive Bayes SMOTE rasio 30:70	66
Gambar 4. 23 Confusion Matrix SVM SMOTE Rasio 30:70	67
Gambar 4. 24 Confusion Matrix Naive Bayes SMOTE rasio 40:60	68
Gambar 4. 25 Confusion Matrix SVM SMOTE 40:60	69
Gambar 4. 26 <i>Confusion Matrix Naive Bayes SMOTE Rasio 60:40</i>	70
Gambar 4. 27 <i>Confusion Matrix SVM SMOTE rasio 60:40</i>	71
Gambar 4. 28 <i>Confusion Matrix Naive Bayes SMOTE Rasio 70:30</i>	72
Gambar 4. 29 <i>Confusion Matrix SVM SMOTE Rasio 70:30</i>	73
Gambar 4. 30 <i>Confusion Matrix Naive Bayes SMOTE Rasio 80:20</i>	74
Gambar 4. 31 <i>Confusion Matrix SVM SMOTE rasio 80:20</i>	75
Gambar 4. 32 Hasil Evaluasi Performa Akurasi Naive Bayes	77
Gambar 4. 33 Hasil Evaluasi Performa Presisi Naive Bayes	77
Gambar 4. 34 Hasil Evaluasi Performa Recall Naive Bayes.....	77
Gambar 4. 35 Hasil Evaluasi Performa F1-Score Naive Bayes	77
Gambar 4. 36 Hasil Evaluasi Performa Akurasi SVM.....	80
Gambar 4. 37 Hasil Evaluasi Performa Presisi SVM.....	80
Gambar 4. 38 Hasil Evaluasi Performa Recall SVM	80
Gambar 4. 39 Hasil Evaluasi Performa F1-Score SVM.....	80

Gambar 4. 40 Kurva ROC Naive Bayes tanpa SMOTE	87
Gambar 4. 41 Kurva ROC Naive Bayes dengan SMOTE 50:50.....	88
Gambar 4. 42 Kurva ROC Naive Bayes dengan SMOTE 20:80.....	88
Gambar 4. 43 Kurva ROC Naive Bayes dengan SMOTE 30:70.....	89
Gambar 4. 44 Kurva ROC Naive Bayes dengan SMOTE 40:60.....	89
Gambar 4. 45 Kurva ROC Naive Bayes dengan SMOTE 60:40.....	90
Gambar 4. 46 Kurva ROC Naive Bayes dengan SMOTE 70:30.....	90
Gambar 4. 47 Kurva ROC Naive Bayes dengan SMOTE 80:20.....	91
Gambar 4. 48 Kurva ROC SVM tanpa SMOTE	91
Gambar 4. 49 Kurva ROC SVM dengan SMOTE	92
Gambar 4. 50 Kurva ROC SVM dengan SMOTE 20:80	92
Gambar 4. 51 Kurva ROC SVM dengan SMOTE 30:70	93
Gambar 4. 52 Kurva ROC SVM dengan SMOTE 40:60	93
Gambar 4. 53 Kurva ROC SVM dengan SMOTE 60:40	94
Gambar 4. 54 Kurva ROC SVM dengan SMOTE 70:30	94
Gambar 4. 55 Kurva ROC SVM dengan SMOTE 80:20	95
Gambar 4. 56 Persebaran AUC Naive Bayes.....	98
Gambar 4. 57 Persebaran AUC SVM.....	98
Gambar 4. 58 Kurva K-fold Naive Bayes	103
Gambar 4. 59 Kurva K-fold SVM.....	105
Gambar 4. 60 Hasil WordCloud Semua Data	108
Gambar 4. 61 Hasil WordCloud Sentimen Negatif.....	108
Gambar 4. 62 Hasil WordCloud Sentimen Positif	109



DAFTAR LAMPIRAN

Lampiran 1. Dataset CNBC Indonesia	120
Lampiran 2. Dataset Kompas Pontianak	120
Lampiran 3. Dataset Kumparan.....	120
Lampiran 4. Crawling Data Komentar Youtube Pada Python	120
Lampiran 5. Cleansing dan Case Folding Pada Python	122
Lampiran 6. Tokenize Pada Python	122
Lampiran 7. Normalize Pada Python	123
Lampiran 8. Stopword Pada Python.....	123
Lampiran 9. Stemming Pada Python.....	124
Lampiran 10. Ekstraksi Fitur TF-IDF Pada Python.....	124
Lampiran 11. Split Data Pada Python	124
Lampiran 12. Balancing Data Dengan SMOTE Pada Python.....	125
Lampiran 13. Fungsi Klasifikasi Pada Python.....	125
Lampiran 14. Klasifikasi Pada Python.....	125
Lampiran 15. Pengujian K-Fold Tanpa SMOTE Pada Python	126
Lampiran 16. Pengujian K-Fold Dengan SMOTE Pada Python.....	128
Lampiran 17. WordCloud Pada Python	129
Lampiran 18. Pengujian Overfitting Data Pada Python	129
Lampiran 19. Dokumentasi	130
Lampiran 20. Dataset yang digunakan	131