

**KOMPARASI ALGORITMA *NAÏVE BAYES* DAN
SUPPORT VECTOR MACHINE (SVM) PADA ANALISIS
SENTIMEN IMPLEMENTASI PROGRAM KARTU
INDONESIA PINTAR KULIAH**



OLEH:

NI PUTU ANA RAINITA

NIM 2115091075

PROGRAM STUDI SISTEM INFORMASI

JURUSAN TEKNIK INFORMATIKA

FAKULTAS TEKNIK DAN KEJURUAN

UNIVERSITAS PENDIDIKAN GANESHA

SINGARAJA

2025



**KOMPARASI ALGORITMA *NAÏVE BAYES* DAN
SUPPORT VECTOR MACHINE (SVM) PADA ANALISIS
SENTIMEN IMPLEMENTASI PROGRAM KARTU
INDONESIA PINTAR KULIAH**

SKRIPSI

**Diajukan kepada
Universitas Pendidikan Ganesha
Untuk Memenuhi Salah Satu Persyaratan dalam Menyelesaikan
Program Sarjana Komputer**



Oleh:

**Ni Putu Ana Rainita
NIM 2115091075**

PROGRAM STUDI SISTEM INFORMASI

JURUSAN TEKNIK INFORMATIKA

FAKULTAS TEKNIK DAN KEJURUAN

UNIVERSITAS PENDIDIKAN GANESHA

SINGARAJA

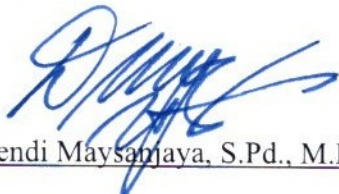
2025

SKRIPSI

**DIAJUKAN UNTUK MELENGKAPI TUGAS
DAN MEMENUHI SYARAT-SYARAT UNTUK
MENCAPAI GELAR SARJANA KOMPUTER**

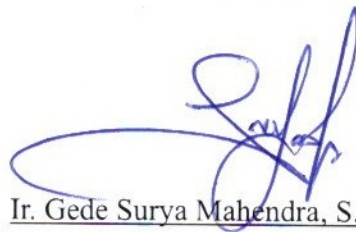
Menyetujui

Pembimbing I,



Ir. I Made Dendi Maysamaya, S.Pd., M.Eng.
NIP.199005152019031008

Pembimbing II,



Ir. Gede Surya Mahendra, S.Pd., M.Kom.
NIP.199003132022031009

Skripsi oleh Ni Putu Ana Rainita
Telah dipertahankan di depan dewan penguji

Pada tanggal 25 Juni 2025

Dewan Penguji,



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

(Ketua)

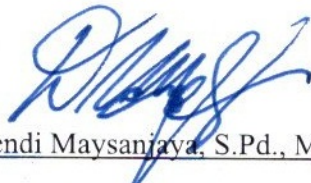
NIP.198307252008011008



Ir. I Gusti Ayu Agung Diatri Indradewi, S.Kom., M.T.

(Anggota)

NIP.198907112020122004



Ir. I Made Dendi Maysanjaya, S.Pd., M.Eng.

(Anggota)

NIP.199005152019031008



Ir. Gede Surya Mahendra, S.Pd., M.Kom.

(Anggota)

NIP.199003132022031009

Diterima oleh Panitia Ujian Fakultas Teknik dan Kejuruan

Universitas Pendidikan Ganesha

Guna memenuhi syarat-syarat untuk mencapai gelar Sarjana Komputer

Pada

Hari

: Senin

Tanggal

: 12.1 JUL 2025



Mengetahui,

Ketua Ujian,

Sekretaris Ujian,

Made Windu Antara Kesiman, S.T., M.Sc., Ph.D.

NIP.198211112008121001

Ir. I Made Dendi Maysarjaya, S.Pd., M.Eng.

NIP.199005152019031008

Mengesahkan,



Prof. Dr. Kadek Rihendra Dantes, S.T., M.T.

NIP.197912012006041001

PERNYATAAN

Dengan ini saya menyatakan bahwa karya tulis yang berjudul “KOMPARASI ALGORITMA *NAÏVE BAYES* DAN *SUPPORT VECTOR MACHINE* (SVM) PADA ANALISIS SENTIMEN IMPLEMENTASI PROGRAM KARTU INDONESIA PINTAR KULIAH” berserta seluruh isinya adalah benar benar karya sendiri dan saya tidak melakukan penjiplakan dan pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung risiko/sanksi yang dijatuhkan kepada saya apabila kemudian ditemukan adanya pelanggaran atas etika keilmuan dalam karya saya ini atau ada klaim terhadap keaslian karya saya ini.

Singaraja, 12 Juni 2025

Yang membuat pernyataan,



Ni Putu Ana Rainita

NIM.2115091075

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN_LOGO	ii
HALAMAN JUDUL.....	iii
LEMBAR PERSETUJUAN PEMBIMBING	iv
LEMBAR PERSETUJUAN DOSEN PENGUJI	v
LEMBAR PERSETUJUAN DAN PENGESAHAN PANITIA UJIAN	vi
PERNYATAAN KEASLIAN TULISAN	vii
KATA PERSEMBAHAN	viii
MOTTO.....	ix
PRAKATA.....	x
ABSTRAK	xiii
<i>ABSTRACT</i>	xiv
DAFTAR ISI	xv
DAFTAR TABEL.....	xvii
DAFTAR GAMBAR	xviii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	5
1.4 Ruang Lingkup Penelitian.....	5
1.5 Manfaat Penelitian	6
BAB II KAJIAN PUSTAKA DAN LANDASAN TEORI	8
2.1 Kajian Pustaka.....	8
2.2 Dasar Teori	11
2.2.1 Program Kartu Indonesia Pintar Kuliah.....	11
2.2.2 Media Sosial X.....	12
2.2.3 <i>Data Mining</i>	13
2.2.4 <i>Text Mining</i>	15
2.2.5 Analisis Sentimen.....	17
2.2.6 Ekstraksi Fitur <i>Term Frequency-Inverse Document Frequency</i>	17
2.2.7 Klasifikasi	18
2.2.8 <i>Naive Bayes Classifier</i>	18

2.2.9 <i>Support Vector Machine</i>	20
2.2.10 <i>Resampling Synthetic Minority Oversampling Technique</i>	22
2.2.11 <i>K-Fold Cross Validation</i>	25
2.2.12 <i>Confusion Matrix</i>	26
BAB III METODOLOGI PENELITIAN	28
3.1 Tahapan Penelitian	28
3.1.1 Pendahuluan	29
3.1.2 Pengumpulan Data	29
3.1.3 Pemrosesan Data	30
3.1.4 Pembangunan Model	39
3.1.5 Interpretasi Hasil dan Kesimpulan	47
BAB IV HASIL DAN PEMBAHASAN	48
4.1 Hasil	48
4.1.1 Pengumpulan Data	48
4.1.2 Seleksi Data	49
4.1.3 Pelabelan Data	51
4.1.4 <i>Text Pre-processing</i>	52
4.1.5 Pembobotan Kata TF-IDF	55
4.1.6 Implementasi Model	58
4.1.7 Hasil Visualisasi Analisis Sentimen	70
4.2 Pembahasan	72
4.2.1 Analisis Hasil Komparasi Model	72
4.2.2 Analisis Hasil Sentimen Program Kartu Indonesia Pintar Kuliah	75
BAB V PENUTUP	77
5.1 Simpulan	77
5.2 Saran	78
DAFTAR PUSTAKA	79
LAMPIRAN	79

DAFTAR TABEL

Tabel 2.1 Ukuran Tingkat Ketidakseimbangan Data	23
Tabel 2.2 Confusion Matrix	26
Tabel 3.1 Hasil Pengumpulan Data <i>Tweet KIP Kuliah</i>	29
Tabel 3.2 Contoh Pelabelan.....	31
Tabel 3.3 Contoh <i>Case Folding</i>	32
Tabel 3.4 Contoh <i>Cleaning</i>	33
Tabel 3.5 Contoh <i>Tokenization</i>	33
Tabel 3.6 Contoh <i>Normalization</i>	34
Tabel 3.7 Contoh <i>Stopword Removal</i>	35
Tabel 3.8 Contoh <i>Stemming</i>	35
Tabel 3.9 Contoh Data TF-IDF	36
Tabel 3.10 Contoh TF IDF	37
Tabel 3.11 Contoh <i>Prior Probabilities</i>	39
Tabel 3.12 Contoh <i>Term Frquent</i>	40
Tabel 3.13 <i>Probabilities Term</i>	41
Tabel 3.14 Contoh Dokumen Testing	42
Tabel 3.15 Contoh TF-IDF	44
Tabel 3.16 Contoh Tabel <i>Confusion Matrix Binary Class</i>	47
Tabel 4.1 Hasil Pengumpulan Data <i>Tweet KIP-K</i>	48
Tabel 4.2 Contoh Data <i>Tweet Tidak Sesuai Kriteria</i>	50
Tabel 4.3 Hasil Pelabelan Data	51
Tabel 4.4 Hasil Proses <i>Case Folding</i>	53
Tabel 4.5 Hasil Proses <i>Cleaning</i>	53
Tabel 4.6 Hasil Proses <i>Tokenize</i>	54
Tabel 4.7 Hasil Proses <i>Normalize</i>	54
Tabel 4.8 Hasil Proses <i>Stopword Removal</i>	55
Tabel 4.9 Hasil Proses <i>Steeming</i>	55
Tabel 4.10 Nilai IDF Top 10 <i>Term</i>	57
Tabel 4.11 Hasil Spesifikasi Google Colaboratory	58
Tabel 4.12 Skenario Pengujian Metode Klasifikasi dan SMOTE.....	59
Tabel 4.13 Hasil Sampel Proses <i>Resampling SMOTE</i>	66

DAFTAR GAMBAR

Gambar 2.1 Proses <i>Knowledge Discovery In Database</i> (KDD)	14
Gambar 2.2 Ilustrasi <i>Metode Support Vector Machine</i>	20
Gambar 2.3 <i>K-Fold Cross Validation</i>	25
Gambar 3.1 Diagram Alir Tahapan Penelitian.....	28
Gambar 3.2 Tahapan <i>Text Pre-Processing</i>	32
Gambar 4.1 Jumlah Masing-Masing Kelas Sentimen	52
Gambar 4.2 Hasil <i>Run</i> Kode Ekstraksi TF-IDF	56
Gambar 4.3 Grafik Akurasi Per <i>Fold Naive Bayes</i>	60
Gambar 4.4 <i>Confusion Matrix</i> dan <i>Classification Report Naive Bayes</i>	61
Gambar 4.5 Hasil Evaluasi dari Pengujian Kombinasi Parameter.....	63
Gambar 4.6 Grafik Akurasi Per <i>Fold SVM</i>	63
Gambar 4.7 <i>Confusion Matrix</i> dan <i>Classification Report SVM</i>	64
Gambar 4.8 Perbandingan Jumlah Sentimen Sebelum dan Sesudah SMOTE.....	65
Gambar 4.9 Grafik Akurasi Per <i>Fold Naive Bayes + SMOTE</i>	66
Gambar 4.10 <i>Confusion Matrix</i> dan <i>Classification Report Naive Bayes + SMOTE</i>	67
Gambar 4.11 Hasil Rata-rata 10 <i>Fold Cross Validation</i> dengan SMOTE.....	68
Gambar 4.12 Grafik Akurasi Per <i>Fold SVM + SMOTE</i>	69
Gambar 4.13 <i>Confusion Matrix</i> dan <i>Classification Report SVM + SMOTE</i>	69
Gambar 4.14 <i>Word Cloud</i> pada Sentimen Positif.....	70
Gambar 4.15 <i>Word Frequency</i> Sentimen Positif.....	71
Gambar 4.16 <i>Word Cloud</i> Sentimen Negatif.....	71
Gambar 4.17 <i>Word Frequency</i> Sentimen Negatif	72



DAFTAR LAMPIRAN

Lampiran 1. Riwayat Hidup	83
Lampiran 2. Kode <i>Crawling Data</i>	84
Lampiran 3. Kode <i>Pre-Processing Data</i>	85
Lampiran 4. Kode <i>Pre-Processing Data</i>	86
Lampiran 5. Kode <i>Pre-Processing Data</i>	87
Lampiran 6. Kode <i>Pre-Processing Data</i>	88
Lampiran 7. Kode TF-IDF	89
Lampiran 8. Kode Implementasi <i>Naive Bayes</i>	90
Lampiran 9. Kode Implementasi <i>Support Vector Machine</i>	91
Lampiran 10. Kode Implementasi SMOTE	94
Lampiran 11. Surat Permohonan Pelabelan Data.....	95
Lampiran 12. Dokumentasi Pelabelan Bersama ibu Ida Ayu Candra Dewi, S.Pd., M.Pd.	96
Lampiran 13. Dokumentasi Pelabelan Bersama ibu Ni Putu Yuna Martika, S.Pd.	97

