

**OPTIMASI ANALISIS SENTIMEN PROGRAM
MERDEKA BELAJAR-KAMPUS MERDEKA (MBKM)
PADA MEDIA SOSIAL TWITTER MENGGUNAKAN
KOMBINASI *K-NEAREST NEIGHBOR* DAN SMOTE**



**OLEH
KOMANG PRAMAYASA
NIM 1915091014**

**PROGRAM STUDI SISTEM INFORMASI
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KEJURUAN
UNIVERSITAS PENDIDIKAN GANESHA
SINGARAJA**

2023



**OPTIMASI ANALISIS SENTIMEN PROGRAM
MERDEKA BELAJAR-KAMPUS MERDEKA (MBKM)
PADA MEDIA SOSIAL TWITTER MENGGUNAKAN
KOMBINASI *K-NEAREST NEIGHBOR* DAN SMOTE**

SKRIPSI



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

**Oleh
Komang Pramayasa
NIM 1915091014**

**PROGRAM STUDI SISTEM INFORMASI
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KEJURUAN
UNIVERSITAS PENDIDIKAN GANESHA**

SINGARAJA

2023

SKRIPSI

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

Menyetujui

Pembimbing I,

Pembimbing II,



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

NIP. 199005152019031008



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

NIP. 198907112020122004

Skripsi oleh Komang Pramayasa

Telah dipertahankan di depan dewan penguji

Pada tanggal 21 Juni 2023

Dewan Penguji,



I Gede Mahendra Darmawiguna, S.Kom., M.Sc.

(Ketua)

NIP. 198501042010121004



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

(Anggota)

NIP. 198307252008011008



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

(Anggota)

NIP. 199005152019031008



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

(Anggota)

NIP. 198907112020122004

Diterima oleh Panitia Ujian Fakultas Teknik dan Kejuruan
Universitas Pendidikan Ganesha
guna memenuhi syarat-syarat untuk mencapai gelar Sarjana Komputer

Pada:

Hari : Rabu

Tanggal : 21 Juni 2023

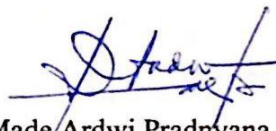
Mengetahui,

Ketua Ujian,

Sekretaris Ujian,



Prof. Dr. Ketut Agustini, S.Si., M.Si.
NIP. 197408012000032001



I Made Ardwi Pradnyana, S.T., M.T.
NIP. 198611182015041001

Mengesahkan,

Wakil Rektor Bidang Akademik dan Kerjasama



Prof. Dr. Gede Rasben Dantes, S.T., M.T.I.
NIP. 197502212003121001

PERNYATAAN

Dengan ini saya menyatakan bahwa karya tulis yang berjudul “OPTIMASI ANALISIS SENTIMEN PROGRAM MERDEKA BELAJAR-KAMPUS MERDEKA (MBKM) PADA MEDIA SOSIAL TWITTER MENGGUNAKAN KOMBINASI *K-NEAREST NEIGHBOR* DAN SMOTE” beserta 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, 21 Juni 2023

Yang membuat pernyataan,



Komang Pramayasa

NIM. 1915091014

KATA PERSEMBAHAN

“Om Swastyastu”

*“Om Dewa Suksma Parama Acintya ya namah swaha
Sarwa karya prasidhantam”*

:

*Om Sang Hyang Widhi Wasa dalam wujud Parama Acintya yang maha gaib
nan maha karya, hanya atas anugrahMu-lah maka pekerjaan ini berhasil dengan
baik.*

Dengan mengucapkan puji syukur atas berkat rahmat-Nya, skripsi ini saya
dedikasikan kepada:

KEDUA ORANG TUA TERCINTA

(Alm. Ketut Raka Adnyana & Luh Semi Arini)

Yang telah menyayangi, membimbing, dan mendukung penulis dengan penuh
kasih sayang. Terima kasih atas semangat, motivasi, dukungan, dan doa yang telah
diberikan dalam setiap langkah yang penulis ambil. Terima kasih sudah menjadi
obor penyemangat penulis untuk menyelesaikan jenjang pendidikan ini.

SAUDARA TERSAYANG

**(Gede Oka Brawida Uthama, Made Wahyu Putra Pratistha
& Luh Indah Sari Pratiwi)**

Yang selalu menghibur dan menyemangati penulis ketika sedang merasa lelah saat
proses pengerjaan skripsi ini dan semoga kita semua menjadi anak yang
membanggakan untuk kedua orang tua kita.

REKAN-REKAN SEPERJUANGAN

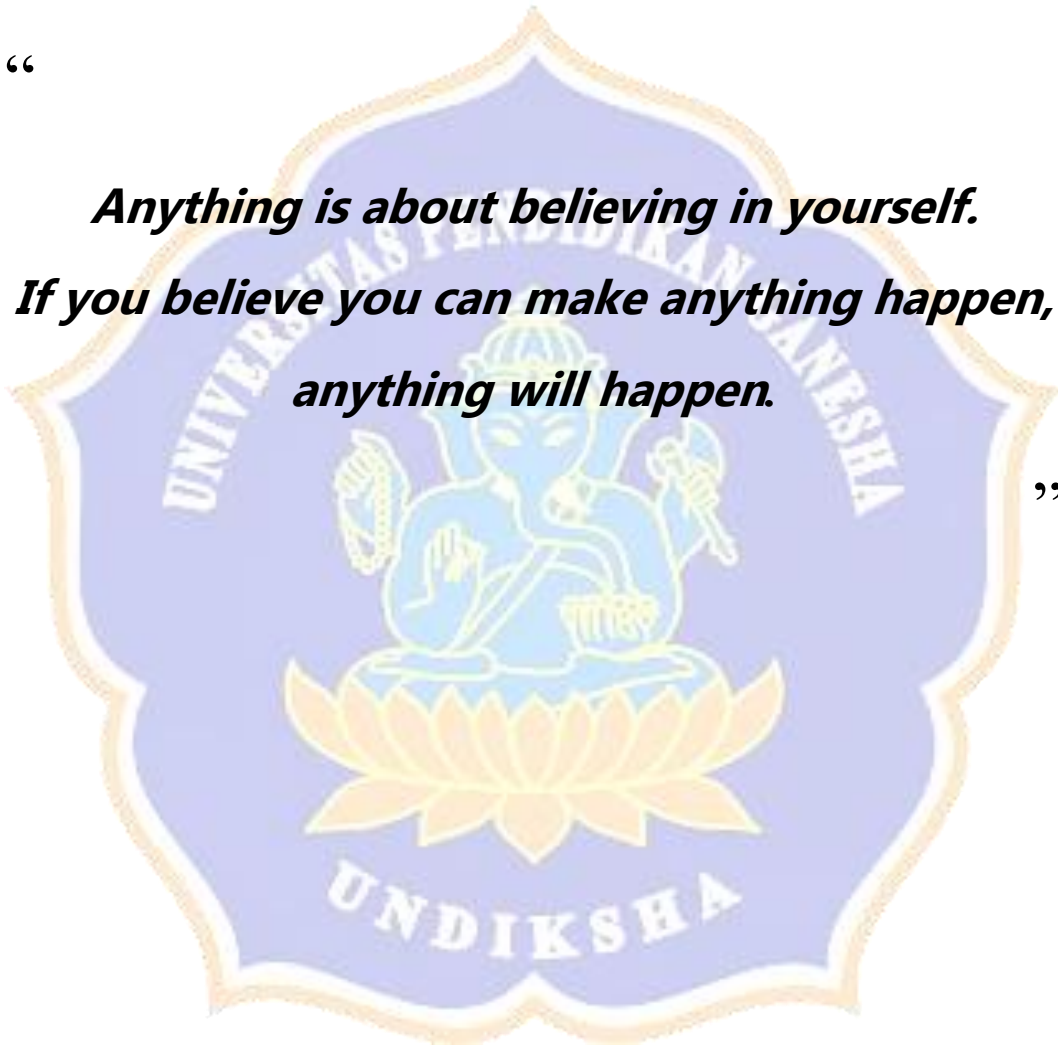
Mirah Pradnya, Regina, Dwi, Sangtu, Daniel, Yura, dan Wiweka yang selalu
menjadi teman seperjuangan dan tempat penulis berkeluh kesah selama penulis
menempuh pendidikan sarjana.

MOTTO

“

***Anything is about believing in yourself.
If you believe you can make anything happen,
anything will happen.***

”



PRAKATA

Puji syukur penyusun panjatkan ke hadapan Tuhan Yang Maha Esa karena berkat rahmatNya-lah, penulis dapat menyelesaikan skripsi yang berjudul “Optimasi Analisis Sentimen Program Merdeka Belajar-Kampus Merdeka (MBKM) Pada Media Sosial Twitter Menggunakan Kombinasi *K-Nearest Neighbor* dan SMOTE”. Skripsi ini disusun guna memenuhi persyaratan mencapai gelar sarjana komputer pada Universitas Pendidikan Ganesha.

Dalam menyelesaikan skripsi ini, penulis banyak mendapat bantuan baik berupa moral maupun material dari berbagai pihak. Untuk itu, dalam kesempatan ini, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. I Gede Sudirtha, S.Pd., M.Pd., selaku Dekan Fakultas Teknik dan Kejuruan atas motivasi dan fasilitas yang diberikan sehingga penulis bisa menyelesaikan studi sesuai dengan rencana.
2. Ibu Dr. Luh Joni Erawati Dewi, S.T., M.Pd., selaku Ketua Jurusan Teknik Informatika atas motivasi dan fasilitas yang diberikan dalam penyelesaian skripsi ini.
3. Bapak I Made Ardwi Pradnyana, S.T., M.T., selaku Koordinator Program Studi Sistem Informasi atas motivasi dan fasilitas yang diberikan dalam penyelesaian skripsi ini.
4. Bapak I Made Dendi Maysanjaya, S.Pd., M.Eng., selaku Pembimbing I yang telah memberikan bimbingan, arahan, petunjuk, dan motivasi kepada penulis dalam penyelesaian skripsi ini.
5. Ibu I Gusti Ayu Agung Diatri Indradewi, S.Kom., M.T., selaku Pembimbing II yang telah memberikan bimbingan, arahan, petunjuk, dan motivasi kepada penulis dalam penyelesaian skripsi ini.
6. Ibu Dra. Komang Ari Setyawati selaku Guru Mata Pelajaran Bahasa Indonesia pada SMA Negeri 1 Semarapura, Klungkung yang telah bersedia meluangkan waktunya untuk membantu penulis dalam melengkapi data penelitian dalam penyelesaian skripsi ini.
7. Ibu Putu Ayu Widya Astari, S.Pd. selaku Guru Mata Pelajaran Bahasa Indonesia pada SMK Negeri 1 Busungbiu, Buleleng yang telah bersedia

meluangkan waktunya untuk membantu penulis dalam melengkapi data penelitian dalam penyelesaian skripsi ini.

8. Secara khusus kepada kedua orang tua, adik, kakak, serta sahabat yang telah mendoakan dan memberikan semangat yang sangat mendalam kepada penulis dalam penyelesaian skripsi ini.
9. Rekan mahasiswa Program Studi Sistem Informasi yang telah memberikan motivasi kepada penulis dalam penyelesaian skripsi ini.
10. Semua pihak yang tidak bisa disebutkan satu per satu yang telah membantu penyusunan skripsi ini.

Penulis menyadari sepenuhnya bahwa apa yang tersaji dalam skripsi ini masih jauh dari sempurna karena keterbatasan kemampuan yang penulis miliki. Untuk itu demi kesempurnaan skripsi ini, penulis mengharapkan segala kritik maupun saran yang sifatnya membangun dari berbagai pihak. Penulis berharap skripsi ini dapat bermanfaat dan berguna bagi kita semua.

Singaraja, 21 Juni 2023

Penulis

DAFTAR ISI

HALAMAN SAMBUNG	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	xii
<i>ABSTRACT</i>	xiii
DAFTAR ISI	xiv
DAFTAR TABEL	xvii
DAFTAR GAMBAR	xix
BAB I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	5
1.3. Tujuan Penelitian	5
1.4. Batasan Masalah Penelitian	5
1.5. Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA	7
2.1. Studi Sebelumnya	7
2.2. Dasar Teori	11
2.2.1. Merdeka Belajar-Kampus Merdeka (MBKM)	11

2.2.2.	Twitter	13
2.2.3.	<i>Data Mining</i>	15
2.2.4.	<i>Text Mining</i>	17
2.2.5.	Analisis Sentimen	19
2.2.6.	Ekstraksi Fitur <i>Term Frequency-Inverse Document Frequency</i>	20
2.2.7.	<i>Resampling Synthetic Minority Oversampling Technique</i>	22
2.2.8.	Klasifikasi	24
2.2.9.	<i>K-Nearest Neighbor (KNN)</i>	24
2.2.10.	<i>Cross Validation</i>	27
2.2.11.	<i>Confusion Matrix</i>	28
BAB III METODOLOGI PENELITIAN.....		30
3.1.	Tahapan Penelitian.....	30
3.2.	Metode Pengumpulan Data	32
3.3.	Pengolahan dan Analisis Data	32
BAB IV HASIL DAN PEMBAHASAN.....		42
4.1.	Hasil.....	42
4.1.1.	Pengumpulan Data	42
4.1.3.	Pelabelan Data	47
4.1.4.	<i>Text Pre-processing</i>	48
4.1.5.	Ekstraksi TF-IDF	55
4.1.6.	Model Klasifikasi.....	57
4.2.	Pembahasan	71
4.2.1.	Analisis Performa Model Klasifikasi.....	71
4.2.2.	Visualisasi Sentimen Program MBKM.....	76
BAB V PENUTUP.....		82
5.1.	Simpulan.....	82

5.2. Saran.....	82
DAFTAR PUSTAKA.....	84
LAMPIRAN.....	91



DAFTAR TABEL

Tabel 2.1. <i>Confusion Matrix</i>	28
Tabel 3.1. Contoh <i>Dataset</i>	33
Tabel 3.2. Contoh Pelabelan <i>Dataset</i>	34
Tabel 3.3. Contoh Proses <i>Case Folding</i>	35
Tabel 3.4. Contoh Proses <i>Cleansing</i>	35
Tabel 3.5. Contoh Proses <i>Tokenizing</i>	36
Tabel 3.6. Contoh Proses <i>Normalizing</i>	36
Tabel 3.7 Contoh Proses <i>Stopword Removal</i>	37
Tabel 3.8 Contoh Proses <i>Stemming</i>	37
Tabel 3.9. Contoh Data TF-IDF	37
Tabel 3.10. Contoh Perhitungan TF-IDF	38
Tabel 3.11. Perhitungan Vektor	39
Tabel 3.12. Perhitungan Panjang Setiap Dokumen.....	39
Tabel 3.13. Hasil Perhitungan <i>Cosine Similarity</i>	40
Tabel 4.1 Kode Program <i>Crawling</i> Data Twitter	42
Tabel 4.2 Hasil Pengumpulan Data <i>Tweet</i> MBKM.....	45
Tabel 4.3 Contoh <i>Tweets</i> Tidak Sesuai Kriteria	46
Tabel 4.4 Hasil Pelabelan Data	47
Tabel 4.5 Kode Program Proses <i>Case Folding</i>	49
Tabel 4.6 Hasil Proses <i>Case Folding</i>	49
Tabel 4.7 Kode Program Proses <i>Cleansing</i>	49
Tabel 4.8 Hasil Proses <i>Cleansing</i>	51
Tabel 4.9 Kode Program Proses <i>Tokenizing</i>	51
Tabel 4.10 Hasil Proses <i>Tokenizing</i>	51
Tabel 4.11 Kode Program Proses <i>Normalizing</i>	52

Tabel 4.12 Hasil Proses <i>Normalizing</i>	52
Tabel 4.13 Kode Program Proses <i>Stopword Removing</i>	53
Tabel 4.14 Hasil Proses <i>Stopword Removing</i>	53
Tabel 4.15 Kode Program Proses <i>Stemming</i>	54
Tabel 4.16 Hasil Proses <i>Stemming</i>	55
Tabel 4.17 Kode Program Proses TF-IDF	55
Tabel 4.18 Hasil Proses Ekstraksi Fitur TF-IDF	56
Tabel 4.19 Nilai IDF <i>Top 10 Term</i>	56
Tabel 4.20 Kode Program Proses Klasifikasi.....	58
Tabel 4.21 Hasil Klasifikasi Model KNN	58
Tabel 4.22 Hasil Evaluasi Model KNN.....	62
Tabel 4.23 Kode Program Proses SMOTE	63
Tabel 4.24 Hasil <i>Sample</i> Proses <i>Resampling</i> SMOTE.....	64
Tabel 4.25 Hasil Klasifikasi Model KNN&SMOTE	65
Tabel 4.26 Hasil Evaluasi Model KNN&SMOTE	69
Tabel 4.27 <i>Dataset</i> Pengujian Model Klasifikasi.....	69
Tabel 4.28 Hasil Prediksi Masing-Masing Model Klasifikasi	70
Tabel 4.29 Perbandingan Performa Model.....	74
Tabel 4.30 Perbandingan Klasifikasi Model dan Ahli	75

DAFTAR GAMBAR

Gambar 2.1. Proses <i>Knowledge Discovery in Database</i> (KDD)	16
Gambar 2.2. Visualisasi Metode KNN.....	26
Gambar 2.3 Contoh Model <i>K-Fold Cross Validation</i>	27
Gambar 3.1. Tahapan Penelitian	30
Gambar 3.2. Tahapan Pengolahan dan Analisis Data.....	32
Gambar 3.3. Tahapan <i>Text Pre-processing</i>	35
Gambar 4.1 Persentase Jumlah Sentimen	48
Gambar 4.2 Grafik Rata-Rata <i>Accuracy</i> Parameter Model KNN.....	59
Gambar 4.3 Grafik Rata-Rata <i>Precision</i> Parameter Model KNN.....	60
Gambar 4.4 Grafik Rata-Rata <i>Recall</i> Parameter Model KNN.....	60
Gambar 4.5 Grafik Rata-Rata <i>F1-Score</i> Parameter Model KNN.....	61
Gambar 4.6 <i>Confusion Matrix</i> Parameter Optimal Model KNN	62
Gambar 4.7 Perbandingan Jumlah Sentimen Sebelum dan Sesudah <i>Resampling</i> .64	
Gambar 4.8 Grafik Rata-Rata <i>Accuracy</i> Parameter Model KNN&SMOTE	66
Gambar 4.9 Grafik Rata-Rata <i>Precision</i> Parameter Model KNN&SMOTE	66
Gambar 4.10 Grafik Rata-Rata <i>Recall</i> Parameter Model KNN&SMOTE	67
Gambar 4.11 Grafik Rata-Rata <i>F1-Score</i> Parameter Model KNN&SMOTE	67
Gambar 4.12 <i>Confusion Matrix</i> Parameter Optimal Model KNN&SMOTE	68
Gambar 4.13 Perbandingan Jumlah Klasifikasi Sentimen Model KNN.....	72
Gambar 4.14 <i>Word Cloud</i> Sentimen Positif	77
Gambar 4.15 <i>Word Frequency</i> Sentimen Positif.....	77
Gambar 4.16 <i>Word Cloud</i> Sentimen Netral.....	78
Gambar 4.17 <i>Word Frequency</i> Sentimen Netral	79
Gambar 4.18 <i>Word Frequency</i> Sentimen Negatif	79
Gambar 4.19 <i>Word Frequency</i> Sentimen Negatif	80

DAFTAR LAMPIRAN

Lampiran 01. Riwayat Hidup.....	91
Lampiran 02. Hasil Proses Ekstraksi TF-IDF	92
Lampiran 03. Data Hasil <i>Resampling</i> SMOTE.....	97
Lampiran 04. Surat Permohonan Pelabelan Data	102

