

**KLASIFIKASI BERITA HOAX KONFLIK PALESTINA-ISRAEL
MENGUNAKAN MULTINOMIAL NAÏVE BAYES DENGAN
OPTIMASI PARTICLE SWARM OPTIMIZATION**

SKRIPSI



OLEH :

SALMA DIAN APRILIA

NPM. 20081010151

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR**

2024

LEMBAR PENGESAHAN
SKRIPSI

**Judul : KLASIFIKASI BERITA HOAX KONFLIK PALESTINA-
ISRAEL MENGGUNAKAN MULTINOMIAL NAIVE BAYES
DENGAN OPTIMASI PARTICLE SWARM OPTIMIZATION**

Oleh : SALMA DIAN APRILIA

NPM : 20081010151

Telah Diseminarkan Dalam Ujian Skripsi Pada :

Hari Rabu, Tanggal 22 Mei 2024

Mengetahui

Dosen Pembimbing

Dosen Penguji

1.



Dr. Basuki Rahmat, S.Si MT.

NIP. 19690723 2021211 002

1.



Dr. Rr. Ani Dajah Rahajoe, ST., M.Cs.

NIP. 19730512 200501 2003


2.



Hendra Maulana, S.Kom., M.Kom.

NPT. 201198 31 223248

2.

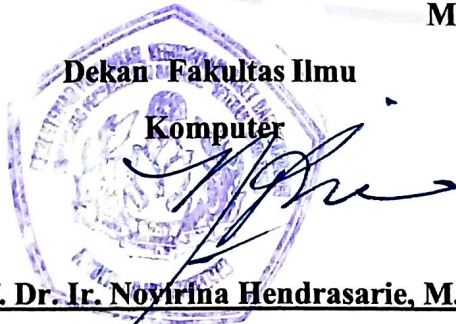


Fawwaz Ali Akbar, S.Kom., M.Kom

NIP. 19920317 2018031 002

Menyetujui

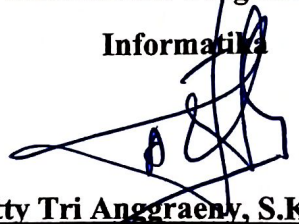
**Dekan Fakultas Ilmu
Komputer**



Prof. Dr. Ir. Novirina Hendrasarie, M.T

NIP. 19681126 199403 2 001

**Koordinator Program Studi
Informatika**



Fetty Tri Anggraeny, S.Kom. M.Kom

NIP. 19820211 2021212 005

SURAT PERNYATAAN BEBAS DARI PLAGIASI

Saya, mahasiswa Program Studi Sarjana Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur, yang bertanda tangan di bawah ini:

Nama : Salma Dian Aprilia

NPM : 20081010151

Menyatakan dengan sesungguhnya bahwa Skripsi/Tugas Akhir yang saya kerjakan berjudul:

**“KLASIFIKASI BERITA HOAX KONFLIK PALESTINA-ISRAEL
MENGUNAKAN MULTINOMIAL NAÏVE BAYES DENGAN
OPTIMASI PARTICLE SWARM OPTIMIZATION”**

bukan merupakan plagiasi sebagian atau keseluruhan dari Skripsi/Tugas Akhir/Penelitian orang lain dan juga bukan merupakan produk dan software yang saya beli dari pihak lain. Saya juga menyatakan bahwa Skripsi/Tugas Akhir ini secara keseluruhan adalah pekerjaan Saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional “Veteran” Jawa Timur maupun di Institut Pendidikan lain. Bukti hasil pengecekan plagiasi dokumen ini dapat ditelusuri melalui QR Code di bawah.

Apabila di kemudian hari terbukti bahwa dokumen ini merupakan plagiasi karya orang lain, saya sanggup menerima sanksi sesuai aturan yang berlaku.

Demikian atas perhatiannya disampaikan terima kasih.

Surabaya, 12 Mei 2024

Hormat saya,



Salma Dian Aprilia

NPM. 20081010151



**KLASIFIKASI BERITA HOAX KONFLIK PALESTINA-ISRAEL
MENGUNAKAN MULTINOMIAL NAÏVE BAYES DENGAN
OPTIMASI PARTICLE SWARM OPTIMIZATION**

Nama Mahasiswa : Salma Dian Aprilia

NPM : 20081010151

Program Studi : Informatika

Dosen Pembimbing : Dr. Basuki Rahmat, S.Si. MT.

Hendra Maulana, S.Kom., M.Kom

ABSTRAK

Berita memiliki kedudukan yang penting sebagai suatu sumber informasi dalam masyarakat. Berkembangnya era digitalisasi juga mendukung munculnya portal berita online. Dari yang awalnya berita hanya dalam media cetak, radio atau televisi, pada saat ini berita dapat disebarkan melalui media sosial, termasuk dengan penyebaran melalui website berita online. Dengan semakin terbukanya akses penyebaran informasi membuat banyak kasus penyebaran berita hoax yang kebenarannya tidak dapat dibuktikan. Topik yang sedang marak diperbincangkan saat ini adalah konflik Palestina-Israel. Untuk membuat suatu pengklasifikasian berita hoax, diperlukan suatu algoritma yang mendukung adanya hal tersebut. Algoritma tersebut adalah *Multinomial Naïve Bayes* yang efektif dalam pengklasifikasian teks. Untuk membuat agar tingkat akurasi yang dihasilkan semakin tinggi, maka ditambahkan suatu metode optimasi, yaitu *Particle Swarm Optimization* yang berperan sebagai *hyperparameter tuning* dalam penelitian ini. Hasil pengujian menunjukkan jika percobaan klasifikasi dengan algoritma *Multinomial Naïve Bayes* menghasilkan akurasi tertinggi dengan pembagian data 90%:10%, yaitu sebesar 76%, sedangkan percobaan dengan tambahan *Particle Swarm Optimization* dengan pembagian data 90%:10%, tingkat akurasinya meningkat menjadi 84%. Dengan hal ini membuktikan jika metode optimasi *Particle Swarm Optimization* mampu dalam hal meningkatkan hasil tingkat akurasi.

Kata kunci : *Optimasi, Particle Swarm Optimization, Multinomial Naïve Bayes, Hoax, Akurasi*

KATA PENGANTAR

Puji Syukur kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Klasifikasi Berita Hoax Konflik Palestina-Israel Menggunakan Multinomial Naïve Bayes dengan Optimasi Particle Swarm Optimization”

Dukungan dan bantuan dari pihak-pihak sekitar yang didapat oleh penulis selama pengerjaan skripsi ini membuat penulis akhirnya mampu untuk menyelesaikannya. Dengan rasa hormat, penulis berterima kasih pada seluruh pihak terkait yang terlibat dalam penyusunan skripsi ini dari awal hingga akhir. Penulis berharap agar skripsi ini dapat menambah wawasan baru bagi pembaca.

Penulis menyadari jika laporan skripsi ini masih memiliki kekurangan, oleh karena itu penulis sangat terbuka dalam hal masukan, saran, dan kritik dari seluruh pihak yang dapat membangun untuk mengembangkan laporan skripsi ini.

Surabaya, 12 Mei 2024

Penulis

UCAPAN TERIMA KASIH

Penyusunan skripsi ini tidak terlepas dari bantuan dan dukungan dari beberapa pihak terkait. Dalam kesempatan ini, penulis ingin menyampaikan terima kasih kepada Allah SWT yang selalu memberi kelancaran dan petunjuk-Nya dari segala kesulitan. Penulis juga ingin menyampaikan rasa terima kasih kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, MMT selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, M.T selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom, M.Kom. selaku Koordinator Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Bapak Dr. Basuki Rahmat., S.Si. MT. dan Bapak Hendra Maulana S.Kom., M.Kom selaku dosen pembimbing 1 dan dosen pembimbing 2 penulis yang telah bersedia dan ikhlas memberikan bimbingan, motivasi, dan waktunya kepada penulis selama penyusunan skripsi ini.
5. Seluruh dosen-dosen Informatika UPN “Veteran” Jawa Timur yang bersedia memberikan ilmu dan pengalamannya selama perkuliahan.
6. Kedua orang tua penulis, Doni Pamungkas, S.E., M.M. dan Lusy Dian Anggraeni S.E yang selalu memberikan support dan motivasinya untuk penulis dalam penyusunan skripsi ini.
7. Kedua saudara penulis, Sahrul Dzaky dan Satria Devin yang selalu memberikan semangat kepada penulis dalam penyusunan laporan skripsi ini.
8. Rizqitha Putri, Talitha Maritza, dan I Nengah Budiarta selaku sahabat terdekat penulis yang membantu memberikan support moril dan sarannya selama perkuliahan hingga penyusunan laporan akhir.
9. Kang Younghyun, Kim Wonpil, Yoon Dowoon, Park Sungjin sebagai penyemangat penulis untuk menjalankan perkuliahan dari awal hingga penyusunan laporan skripsi ini.

10. Bianca, Salsa, Olivia, Pinta, dan Shinta selaku sahabat penulis yang memberikan semangat kepada penulis dan berjuang bersama sedari SMA hingga bangku perkuliahan.
11. Teman-teman Pers Mahasiswa Periode Full Senyum yang mengisi dan menjadi bagian dalam kisah perkuliahan penulis, serta membantu penulis dalam memperoleh ilmu *soft skill* selama masa perkuliahan.
12. Seluruh teman-teman Informatika angkatan 2020 yang menemani perjuangan penulis selama menempuh perkuliahan di UPN “Veteran” Jawa Timur.
13. Semua pihak yang tidak dapat dituliskan satu per satu yang turut menyemangati dan mendoakan penulis dalam menyelesaikan skripsi ini.

Akhir kata, semoga laporan skripsi ini dapat memberikan inspirasi dan manfaat, serta menambah wawasan bagi pembaca. Terima kasih atas segala dukungan, support, dan doa-doanya bagi pihak-pihak yang telah membantu penulis hingga penyelesaian penelitian skripsi ini dengan baik.

Surabaya, 12 Mei 2024

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	i
SURAT PERNYATAAN BEBAS DARI PLAGIASI.....	ii
ABSTRAK	iii
KATA PENGANTAR	iv
UCAPAN TERIMA KASIH.....	v
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terdahulu	6
2.2 Berita Hoax	8
2.3 British Broadcasting Corporation	8
2.4 Klasifikasi	9
2.5 Teks Mining	9
2.6 Text Preprocessing	9
2.6.1 Cleaning	10
2.6.2 Case Folding	10
2.6.3 Tokenizing.....	11
2.6.4 Stopword Removal.....	11
2.6.5 Stemming	12
2.7 Term Frequency-Inverse Document Frequency (TF-IDF)	13
2.8 Naïve Bayes Classifier	13
2.8.1 Multinomial Naïve Bayes	14
2.9 <i>Particle Swarm Optimization</i>	15
2.10 <i>Confusion Matrix</i>	16
BAB III METODOLOGI PENELITIAN.....	18

3.1	Tahapan Penelitian	18
3.2	Studi Literatur.....	19
3.3	Analisis Desain.....	20
3.4	Akuisisi dan Pelabelan Data.....	22
3.5	Preprocessing Data	24
3.5.1	Cleaning	24
3.5.2	Case Folding	25
3.5.3	Tokenizing	25
3.5.4	Stopword Removal	26
3.5.5	Stemming	26
3.6	Pembobotan Data	27
3.7	Pemodelan Klasifikasi dan Optimasi	29
3.8	Skenario Pengujian.....	37
3.9	Penyimpanan Model.....	38
BAB IV HASIL DAN PEMBAHASAN		39
4.1	Akuisisi dan Pelabelan Data	39
4.2	<i>Preprocessing Data</i>	42
4.2.1	<i>Cleaning</i>	43
4.2.2	<i>Case Folding</i>	44
4.2.3	<i>Tokenizing</i>	45
4.2.4	<i>Stopword Removal</i>	45
4.2.5	<i>Stemming</i>	45
4.3	Pembobotan Data	47
4.4	Pembagian Data	49
4.5	Model <i>Multinomial Naïve Bayes</i>	50
4.6	Model <i>Particle Swarm Optimization</i>	52
4.7	Hasil Pengujian	54
4.7.1	Hasil Pengujian Skenario 1	55
4.7.2	Hasil Pengujian Skenario 2	57
4.7.3	Hasil Pengujian Skenario 3	59
4.7.4	Hasil Pengujian Skenario 4	60
4.7.5	Hasil Pengujian Skenario 5	62
4.7.6	Hasil Pengujian Skenario 6	64
4.8	Perbandingan Hasil Skenario Pengujian	65

4.9 Penyimpanan Model.....	67
BAB V KESIMPULAN	70
5.1 Kesimpulan	70
5.2 Saran.....	70
DAFTAR PUSTAKA	72
LAMPIRAN.....	77

DAFTAR TABEL

Tabel 2.1 Text Preprocessing Tahap Cleaning	10
Tabel 2.2 Text Preprocessing Tahap Case Folding.....	10
Tabel 2.3 Text Preprocessing Tahap Tokenizing.....	11
Tabel 2.4 Text Preprocessing Tahap Stopword Removal.....	12
Tabel 2.5 Text Preprocessing Tahap Stemming	12
Tabel 2.6 Confusion Matrix	16
Tabel 3.1 Pelabelan Dataset	23
Tabel 3.2 Proses Cleaning.....	24
Tabel 3.3 Proses Case Folding.....	25
Tabel 3.4 Proses Tokenizing	25
Tabel 3.5 Proses Stopword Removal	26
Tabel 3.6 Proses Stemming.....	26
Tabel 3.7 Inisialisasi Partikel PSO.....	32
Tabel 3.8 Hasil Evaluasi Fungsi.....	32
Tabel 3.9 Hasil Inisialisasi Nilai Pbest dan Gbest	32
Tabel 3.10 Ringkasan Perhitungan Manual MNB	36
Tabel 3.11 Skenario Pengujian	37
Tabel 4.1 Perbandingan Hasil Skenario Pengujian	65
Tabel 4.2 Hasil Skenario Pengujian dengan Standardisasi	65

DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian	18
Gambar 3.2 Diagram Alur Sistem.....	19
Gambar 3.3 Diagram Alur Akuisisi dan Pelabelan Data	21
Gambar 3.4 Hasil Scraping Dataset	22
Gambar 3.5 Diagram Alur Preprocessing Data.....	24
Gambar 3.6 Diagram Alur Pembobotan Data	28
Gambar 3.7 Diagram Alur Pemodelan Klasifikasi dan Optimasi	30
Gambar 3.8 Diagram Alur Penyimpanan Model	38
Gambar 4.1 Hasil Scraping Dataset	40
Gambar 4.2 Unduh CSV Scraping Dataset	41
Gambar 4.3 Pelabelan Dataset Secara Manual	42
Gambar 4.4 Output Hasil Dataset yang Berhasil di-upload.....	43
Gambar 4.5 Output Dataset Tahap Cleaning	44
Gambar 4.6 Output Dataset Tahap Case Folding	44
Gambar 4.7 Output Dataset Tahap Tokenizing.....	45
Gambar 4.8 Output Dataset Tahap Stopword Removal.....	46
Gambar 4.9 Output Dataset Tahap Stemming	47
Gambar 4.10 Output Proses Label Encoding	48
Gambar 4.11 Output Proses Pembobotan Sebelum Standardisasi Data	49
Gambar 4.12 Output Proses Pembobotan Setelah Standardisasi Data.....	49
Gambar 4.13 Confusion Matrix Skenario Pengujian 1	56
Gambar 4.14 Classification Report Skenario Pengujian 1	56
Gambar 4.15 Confusion Matrix Skenario Pengujian 2	57
Gambar 4.16 Classification Report Skenario Pengujian 2.....	58
Gambar 4.17 Confusion Matrix Skenario Pengujian 3	59
Gambar 4.18 Classification Report Skenario Pengujian 3.....	59
Gambar 4.19 Confusion Matrix Skenario Pengujian 4	61
Gambar 4.20 Classification Report Skenario Pengujian 4.....	61
Gambar 4.21 Confusion Matrix Skenario Pengujian 5	62
Gambar 4.22 Classification Report Skenario Pengujian 5.....	63
Gambar 4.23 Confusion Matrix Skenario Pengujian 6	64

Gambar 4.24 Classification Report Skenario Pengujian 6.....	64
Gambar 4.25 Hasil Output GUI Utama.....	69
Gambar 4.26 Hasil Prediksi Kategori	69