

**KLASIFIKASI MULTI LABEL BERITA MEDIA MASSA
ONLINE BAHASA INDONESIA MENGGUNAKAN BI-LSTM
DAN EMBEDDING INDOBERT**

SKRIPSI



Oleh:

EDO HANIFAUZAN SATRIA

200411100058

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS TRUNODJOYO MADURA
BANGKALAN
2026**

HALAMAN JUDUL

**KLASIFIKASI MULTI LABEL BERITA MEDIA MASSA
ONLINE BAHASA INDONESIA MENGGUNAKAN BI-LSTM
DAN EMBEDDING INDOBERT**

SKRIPSI

**Diajukan untuk Memenuhi Persyaratan Penyelesaian Studi Strata Satu (S1)
dan Memperoleh Gelar Sarjana Komputer (S.Kom)
Di Universitas Trunodjoyo Madura**

EDO HAIFAUZAN SATRIA

200411100058

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS TRUNODJOYO MADURA
2026**

HALAMAN PENGESAHAN

KLASIFIKASI MULTI LABEL BERITA MEDIA MASSA ONLINE BAHASA INDONESIA MENGGUNAKAN BI-LSTM DAN EMBEDDING INDOBERT

Skripsi ini diajukan sebagai salah satu syarat untuk menyelesaikan pendidikan
Program Studi S1 Teknik Informatika Universitas Trunodjojo Madura

Oleh:

Nama : Edo Hanifauzan Satria

NPM : 200411100058

Disetujui oleh Tim Penguji Skripsi:

Tanggal Sidang: 09 Januari
2026

Dr. Firdaus Solihin, S.Kom. M.Kom.

NIP. 197606272008011008

Dr. Hermawan, S.T. M.Kom.

NIP. 1979082820050110002

Ika Oktavia Suzanti, S.Kom, M.Cs.

NIP. 198810182015042004

Dr. Rika Yunitarini, S.T., M.T.

NIP. 198306072006042001

Prof. Dr. Aeri Rachmad, S.T. M.T.

NIP. 198002232008121001

(Pembimbing 1)

(Pembimbing 2)

(Penguji 1)

(Penguji 2)

(Penguji 3)

Bangkalan, 01 Oktober 2026

Mengetahui,

Ketua Jurusan Teknik Informatika

Husni, S.Kom, M.T.

NIP. 197707222003121006

HALAMAN PERNYATAAN ORISINALITAS

Saya yang bertanda tangan dibawah ini, menyatakan bahwa skripsi saya dengan judul:

“KLASIFIKASI MULTI LABEL BERITA MEDIA MASSA ONLINE BAHASA INDONESIA MENGGUNAKAN BI-LSTM DAN EMBEDDING INDOBERT”

1. Adalah asli, bukan merupakan karya pihak lain serta belum pernah diajukan untuk mendapatkan gelar akademik Sarjana Komputer baik di Universitas Trunojoyo Madura maupun Perguruan Tinggi yang lain di Indonesia.
2. Tidak terdapat karya atau pendapat pihak lain yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis telah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Apabila dikemudian hari terbukti skripsi ini sebagian atau seluruhnya merupakan hasil plagiasi atau terdapat hal-hal yang tidak sesuai dengan pernyataan diatas, maka saya sanggup menerima sanksi akademis yang berlaku dengan segala akibat hukumnya sesuai peraturan Universitas Trunojoyo Madura dan atau peraturan perundang-undangan yang berlaku.

Bangkalan, 01 Oktober 2026
Yang Menyatakan,

Edo hanifauzan Satria
NIM. 200411100058

KATA PENGANTAR

Alhamdulillah Puji Syukur penulis ucapkan kepada Allah Subhanahu wa ta'ala atas limpahan berkat dan rahmat-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “KLASIFIKASI MULTI LABEL BERITA MEDIA MASSA ONLINE BAHASA INDONESIA MENGGUNAKAN BI-LSTM DAN EMBEDDING INDOBERT”. Skripsi ini ditulis dalam rangka memenuhi syarat untuk mencapai gelar Sarjana Komputer pada Program Teknik Informatika Fakultas Teknik universitas Trunodjoyo Madura.

Selama proses penyelesaian skripsi ini, banyak bantuan, saran, dan motivasi yang diberikan oleh berbagai pihak sangat membantu kami sehingga kami dapat menyelesaikan laporan ini. Oleh karena itu, kami mengucapkan terimakasih atas seluruh bantuan yang diberikan kepada :

1. Allah SWT, Tuhan Yang Maha Esa yang telah melimpahkan kemudahan kepada penulis sehingga mampu menyelesaikan skripsi ini.
2. Bapak Dr. Firdaus Solihin, S.Kom., M.Kom., dan bapak Dr. Hermawan, S.T. M.Kom., selaku dosen pembimbing yang telah memberikan bimbingan, kesabaran serta ilmu yang sangat berharga dalam proses pengerjaan skripsi ini.
3. Ibu Ika Oktavia Suzanti, S.Kom, M.Cs., ibu Dr. Rika Yunitarini, S.T., M.T., dan bapak Prof. Dr. Aeri Rachmad, S.T., M.T. selaku dosen penguji yang telah memberikan masukan, kritik serta saran yang membangun untuk perbaikan dan penyempurnaan skripsi ini.
4. Kedua orang tua tercinta yang selalu mendukungnya saya dengan mencurahkan kasih sayang, doa, dan dukungan dalam setiap langkah perjalanan hidup penulis. Tak lupa pula terhadap saudara dan saudari penulis yang selalu membantu dan mendukung penulis dalam penulisan skripsi ini.
5. Kepada teman-teman penulis yang telah membantu dengan menyemangati dan memberikan saran kepada penulis dalam penyelesaian skripsi ini. Terimakasih telah hadir dalam perjalanan penulis dalam pengerjaan skripsi ini
6. Kepada diri sendiri yang telah berjuang sejauh ini dengan berbagai cobaan dan ujian yang dilewati. Tetap tegar dan teruslah berjuang sampai kapanpun dan dimanapun

Sebagai manusia biasa penulis menyadari penyusunan skripsi jauh dari kata sempurna karena keterbatasan kemampuan dan ilmu pengetahuan yang dimiliki oleh penulis. Oleh karena itu, atas segala kesalahan dan kekurangan dalam penulisan skripsi ini, penulis memohon maaf sebesar-besarnya dan bersedia menerima kritik yang membangun

Bangkalan, 01 Oktober 2026

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN ORISINALITAS.....	iii
KATA PENGANTAR.....	iv
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	3
1.2.1 Masalah	3
1.2.2 Metode Usulan.....	3
1.2.3 Pertanyaan Penelitian.....	3
1.3 Tujuan dan Manfaat.....	4
1.3.1 Tujuan	4
1.3.2 Manfaat	4
1.4 Batasan Penelitian	4
1.5 Sistematika penulisan.....	5
BAB II KAJIAN PUSTAKA	6
2.1 Pemrosesan Bahasa Alami / Natural Language Processing (NLP).....	6
2.2 Deep Learning untuk NLP	6
2.3 Word Embedding dan Representasi Kontekstual	7
2.4 IndoBERT	8
2.5 Klasifikasi Teks	10
2.6 Klasifikasi multi label	10
2.7 Bidirectional LSTM (Bi-LSTM).....	11
2.7.1 Rumus LSTM forward.....	13
2.7.2 Rumus LSTM Backward.....	13
2.7.3 Penggabungan Forward dan Backward.....	14
2.8 Evaluasi Model	14
2.8.1 Precision	15

2.8.2 Recall	15
2.8.3 F1 Score.....	16
2.8.4 Hamming loss	16
2.8.5 Waktu Pelatihan.....	17
2.9 Penelitian terdahulu	17
BAB III METODE USULAN	20
3.1 Metode Penelitian.....	20
3.2 Arsitektur Sistem.....	21
3.3 Dataset penelitian	23
3.3.1 Deskripsi Data	23
3.3.2 Pelabelan Data	25
3.3.3 Eksplorasi Data Awal (EDA).....	26
3.4 Pembagian Data	26
3.5 Representasi Teks IndoBERT	26
3.5.1 Tokenisasi IndoBERT.....	26
3.5.2 Embedding IndoBERT.....	27
3.6 Pelatihan Model Bi-LSTM	28
3.7 Evaluasi Model	31
3.7.1 Tujuan Evaluasi	32
3.7.2 Prosedur Evaluasi.....	32
3.8 Skenario Pengujian	33
3.8.1. Skenario 1: Perbandingan Threshold pelabelan	34
3.3.2 Skenario 2: Perbandingan Headline dan Isi Berita	34
3.3.3 Skenario 3: Variasi Panjang token (64, 128)	34
3.3.4 Skenario 4: Variasi Nilai Threshold (0.3–0.7)	35
BAB IV HASIL DAN PEMBAHASAN	36
4.1 Dataset Penelitian.....	36
4.2 Training Model	39
4.3 Evaluasi Model	41
4.3.1 Evaluasi Data <i>Threshold</i> Pelabelan 0.3	41
4.3.2 Evaluasi Data <i>Threshold</i> Pelabelan 0.5	42
4.3.2 Evaluasi Data <i>Threshold</i> Pelabelan 0.7	43
4.3.1 Evaluasi Data Isi Berita.....	45
4.4 Implementasi Sistem.....	50

BAB V PENUTUP	51
5.1 Kesimpulan	51
5.2 Saran	51
DAFTAR PUSTAKA.....	52