

**ANALISIS SENTIMEN PROGRAM NATURALISASI  
TIM NASIONAL SEPAK BOLA INDONESIA PADA  
MEDIA SOSIAL X MENGGUNAKAN INDOBERT  
EMBEDDING DAN CONVOLUTIONAL NEURAL  
NETWORK**

**SKRIPSI**



**OLEH  
I PUTU ARYA VIDYANANTA  
NIM 2215091079**

**PROGRAM STUDI SISTEM INFORMASI (S1)  
JURUSAN TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KEJURUAN  
UNIVERSITAS PENDIDIKAN GANESHA  
SINGARAJA**

**2026**

**ANALISIS SENTIMEN PROGRAM NATURALISASI  
TIM NASIONAL SEPAK BOLA INDONESIA PADA  
MEDIA SOSIAL X MENGGUNAKAN INDOBERT  
EMBEDDING DAN CONVOLUTIONAL NEURAL  
NETWORK**

**SKRIPSI**

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



**Oleh:  
I Putu Arya Vidyananta  
NIM 2215091079**

**PROGRAM STUDI SISTEM INFORMASI (S1)  
JURUSAN TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KEJURUAN  
UNIVERSITAS PENDIDIKAN GANESHA  
SINGARAJA**

**2026**

# SKRIPSI

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



Pembimbing I	I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs. NIP.198910262019031004
Pembimbing II	Ir. Gede Surya Mahendra, S.Pd., M.Kom. NIP.199003132022031009

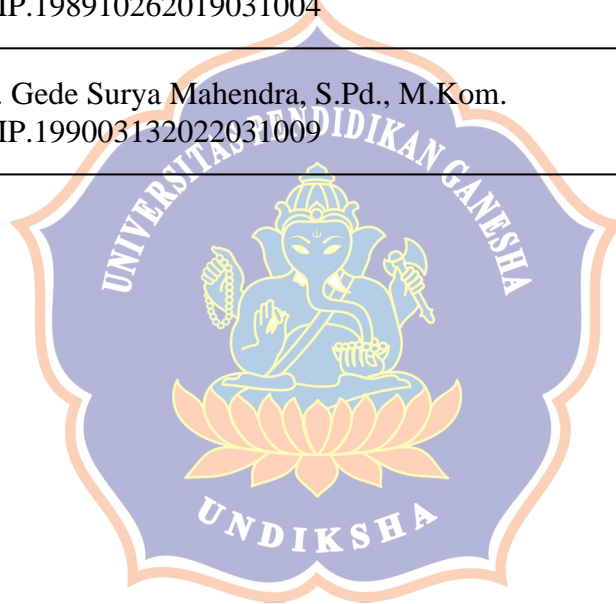


- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi 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.

Skripsi oleh I PUTU ARYA VIDYANANTA ini  
telah dipertahankan di depan dewan penguji  
Pada tanggal 06 Maret 2026

### Dewan Penguji

Ketua	Ir. I Gusti Ayu Agung Diatri Indradewi, S.Kom., M.T. NIP.198907112020122004
Anggota	Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs. NIP.198307252008011008
Anggota	I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs. NIP.198910262019031004
Anggota	Ir. Gede Surya Mahendra, S.Pd., M.Kom. NIP.199003132022031009



- 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 BSR - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSR
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.

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

### Menyetujui

Ketua Ujian	Ir. Made Windu Antara Kesiman, S.T., M.Sc., Ph.D. NIP.198211112008121001
Sekretaris Ujian	Ir. I Made Dendi Maysanjaya, S.Pd., M.Eng. NIP.199005152019031008



- 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.

## PERNYATAAN

Dengan ini saya menyatakan bahwa karya tulis yang berjudul “Analisis Sentimen Program Naturalisasi Tim Nasional Sepak Bola Indonesia Pada Media Sosial X Menggunakan Indobert Embedding Dan Convolutional Neural Network” beserta seluruh isinya adalah benar-benar karya saya sendiri dan saya tidak melakukan penjiplakan serta pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku. Atas pernyataan ini, saya siap menanggung risiko atau sanksi yang dijatuhkan kepada saya apabila di kemudian hari ditemukan adanya pelanggaran atas etika keilmuan dalam karya saya ini atau ada klaim terhadap keaslian karya saya.

Singaraja, 26 Maret 2026

Yang membuat pernyataan,



I Putu Arya Vidyananta

NIM. 2215091079

## **KATA PERSEMBAHAN**

Puja dan puji syukur penulis panjatkan kepada

**Ida Sang Hyang Widhi Wasa**

Berkat segala Rahmat-Nya penulis dapat menyelesaikan skripsi ini dengan baik melalui kelancaran dan kemudahan yang telah diberikan.

Skripsi ini penulis dedikasikan kepada:

**Segenap Keluarga Tercinta**

**Kedua Orang Tua Tersayang**

(I Made Sumerta dan Ni Kadek Suliasmini)

Terima kasih atas doa yang tiada henti, kasih sayang, dukungan moral maupun material, serta segala pengorbanan yang telah diberikan selama perjalanan pendidikan penulis.

**Saudara Terkasih**

(I Kadek Wira Satya)

Terima kasih atas kebersamaan, semangat, serta dukungan yang selalu memberikan kekuatan bagi penulis dalam menghadapi berbagai tantangan selama masa studi

**Seluruh Staf Dosen Prodi Sistem Informasi**

Terima kasih telah memberikan ilmu pengetahuan, bimbingan, arahan, serta motivasi sejak awal perkuliahan hingga penulis mampu menyelesaikan tugas akhir ini.

**Rekan-Rekan Seperjuangan**

Terima kasih kepada rekan-rekan seperjuangan mahasiswa Jurusan Teknik Informatika, khususnya mahasiswa SIFORS 2022, atas kebersamaan, dukungan, dan semangat yang telah diberikan sehingga perjalanan akademik ini dapat dilalui dengan penuh kenangan dan pembelajaran berharga.

## MOTTO

**“Tidak masalah seberapa lambat kamu berjalan, asalkan  
kamu tidak berhenti”  
- Confucius**



## PRAKATA

Puji Syukur penulis panjatkan kepada Tuhan yang Maha Esa Ida Sang Hyang Widhi Wasa karena atas berkat dan Rahmat-Nya penulis dapat menyelesaikan Skripsi yang berjudul “**Analisis Sentimen Program Naturalisasi Tim Nasional Sepak Bola Indonesia pada Media Sosial X Menggunakan IndoBERT Embedding dan Convolutional Neural Network**”. Skripsi ini merupakan salah satu syarat dalam menempuh gelar Sarjana Komputer di Universitas Pendidikan Ganesha.

Dalam penyusunan skripsi ini, penulis banyak mendapatkan bimbingan, motivasi, bantuan, kritik, dan saran yang bersifat membangun dari berbagai pihak. Pada kesempatan ini, penulis mengucapkan terima kasih sebagai rasa syukur dan hormat kepada:

1. Bapak Prof. Dr. I Wayan Lasmawan, M.Pd., selaku Rektor Universitas Pendidikan Ganesha yang telah menyediakan fasilitas serta kebijakan yang mendukung proses perkuliahan.
2. Bapak Prof. Dr. Kadek Rihendra Dantes, S.T., M.T., selaku Dekan Fakultas Teknik dan Kejuruan yang telah memberikan fasilitas dan kebijakan sehingga penulis dapat menyelesaikan studi dengan baik.
3. Bapak Ir. I Made Dendi Maysanjaya, S.Pd., M.Eng., selaku Koordinator Program Studi Sistem Informasi yang telah memberikan arahan, motivasi, kesempatan, serta pengalaman kepada penulis selama menempuh studi.
4. Bapak I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs., selaku Pembimbing I yang telah memberikan bimbingan, arahan, serta motivasi kepada penulis dalam menyelesaikan skripsi ini.
5. Bapak Ir. Gede Surya Mahendra, S.Pd., M.Kom., selaku Pembimbing II yang telah membantu penulis selama perkuliahan serta memberikan bimbingan, arahan, dan motivasi dalam penyelesaian skripsi ini.
6. Ibu Ir. I Gusti Ayu Agung Diatri Indradewi, S.Kom., M.T., selaku Penguji I sekaligus Pembimbing Akademik yang telah memberikan banyak arahan, motivasi, serta bimbingan kepada penulis.

7. Bapak Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs., selaku Penguji II yang telah memberikan bimbingan, arahan, serta evaluasi kepada penulis dalam penyelesaian skripsi ini.
8. Bapak Ir. Kadek Teguh Dermawan, S.Pd., M.Kom., yang telah banyak membantu, membimbing, memberikan kesempatan dan pengalaman, serta rekomendasi kepada penulis selama menempuh studi.
9. Seluruh staf dosen dan pegawai Fakultas Teknik dan Kejuruan yang telah membantu kelancaran penulis selama perkuliahan serta membantu dalam proses administrasi akademik.
10. Ibu I Gusti Ayu Putri Dwi Epayani, S.Pd., M.Pd., Ibu Dra. Ni Wayan Sumiasih, dan Bapak Gede Sukalima Aksirnaka, S.Pd., M.Pd., selaku guru Bahasa Indonesia di SMA Negeri 1 Singaraja yang telah membantu penulis dalam melengkapi data penelitian untuk penyelesaian skripsi.
11. Kedua orang tua tercinta, adik, serta seluruh keluarga yang telah memberikan doa, dukungan moral maupun material, serta pengorbanan selama perjalanan pendidikan penulis.
12. Pasangan penulis yang telah memberikan dukungan emosional, semangat, serta pengertian selama proses penyelesaian skripsi ini.
13. Teman-teman penulis, yaitu Devan, Krisna, Gandi, dan seluruh personil Grup Makan Baru Suka yang telah memberikan dukungan dan kebersamaan dalam menjalani perkuliahan. Serta seluruh rekan SIFORS angkatan 2022 dan teman-teman penulis lainnya, yang telah menjalin pertemanan yang baik selama masa perkuliahan.
14. Teman-teman Sulawesi Pride, yaitu Abdi, Viona, Dewi, dan Prini yang telah memberikan dukungan serta bantuan kepada penulis selama menempuh pendidikan di Universitas Pendidikan Ganesha.
15. Terima kasih kepada diri sendiri atas ketekunan, kerja keras, dan perjuangan selama menempuh perkuliahan hingga mampu menyelesaikan skripsi ini

Penulis menyadari bahwa skripsi ini masih memiliki berbagai keterbatasan dan belum sepenuhnya sempurna. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari berbagai pihak guna penyempurnaan penelitian di masa mendatang. Penulis berharap karya ini dapat memberikan manfaat bagi berbagai pihak, khususnya dalam pengembangan bidang Analisis Sentimen.

Singaraja, Maret 2026

Penulis



## DAFTAR ISI

	Halaman
KATA PERSEMBAHAN .....	iv
MOTTO.....	v
PRAKATA.....	vi
ABSTRAK .....	ix
ABSTRACT .....	x
DAFTAR ISI .....	xi
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR .....	xv
DAFTAR LAMPIRAN .....	xvii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah Penelitian .....	5
1.3 Rumusan Masalah .....	5
1.4 Tujuan.....	6
1.5 Batasan Masalah.....	6
1.6 Manfaat Penelitian .....	7
BAB II TINJAUAN PUSTAKA .....	9
2.1 Kajian Pustaka.....	9
2.2 Landasan Teori .....	15
2.2.1 Persatuan Sepak Bola Seluruh Indonesia (PSSI) .....	15
2.2.2 Naturalisasi.....	15
2.2.3 Media Sosial.....	17
2.2.4 Aplikasi X (Twiter) .....	18
2.2.5 <i>Cross-Industry Standard Process for Data Mining</i> .....	20
2.2.6 <i>Data Mining</i> .....	22
2.2.7 <i>Text Mining</i> .....	24
2.2.8 Analisis Sentimen.....	25
2.2.9 BERT .....	26
2.2.10 IndoBERT .....	29
2.2.11 <i>Convolutional Neural Network</i> .....	29

2.2.12	<i>K-Fold Cross-Validation</i> .....	33
2.2.13	<i>Confusion Matrix</i> .....	34
2.2.14	<i>Synthetic Minority Oversampling Technique (SMOTE)</i> .....	36
BAB III METODE PENELITIAN.....		38
3.1	Jenis Penelitian.....	38
3.2	Tahapan Penelitian .....	38
3.2.1	Tahapan Business Understanding .....	39
3.2.2	Tahapan Data <i>Understanding</i> .....	40
3.2.3	Tahapan Data Preparation .....	46
3.2.4	Tahapan <i>Modeling</i> .....	49
3.2.5	Tahapan <i>Evaluation</i> .....	54
3.2.6	Tahapan <i>Deployment</i> .....	55
BAB IV HASIL DAN PEMBAHASAN.....		56
4.1	Hasil .....	56
4.1.1	<i>Business Understanding</i> .....	56
4.1.2	<i>Data Understanding</i> .....	56
4.1.3	<i>Data Preparation</i> .....	60
4.1.4	<i>Modeling</i> .....	62
4.1.5	<i>Evaluation</i> .....	69
4.1.6	<i>Deployment</i> .....	94
4.2	Pembahasan.....	99
4.2.1	Analisis Kecenderungan Sentimen Program Naturalisasi Tim Nasional Indonesia .....	99
4.2.2	Analisis Hasil Performa Model.....	100
BAB V PENUTUP.....		105
5.1	Kesimpulan .....	105
5.2	Saran.....	106
DAFTAR PUSTAKA .....		107
LAMPIRAN.....		114

## DAFTAR TABEL

Tabel	Halaman
Tabel 3.1 Hastag dan Kata Kunci yang akan digunakan dalam Penelitian .....	41
Tabel 3.2 Contoh Dataset X .....	42
Tabel 3.3 Contoh Pelabelan Dataset X.....	44
Tabel 3.4 Contoh Proses Cleaning .....	46
Tabel 3.5 Contoh Proses Case Folding .....	47
Tabel 3.6 Contoh Proses tokenizing.....	47
Tabel 3.7 Contoh Proses normalization .....	48
Tabel 3.8 Contoh Proses Stemming .....	48
Tabel 3.9 Contoh Input Proses IndoBERT Embedding.....	49
Tabel 3.10 Contoh Proses Tokenisasi IndoBERT .....	50
Tabel 3.11 Contoh Proses Token Tokenisasi IndoBERT.....	50
Tabel 3.12 Contoh Hasil Vektorisasi embedding dari setiap ID numerik .....	51
Tabel 3.13 Contoh Hasil Final Embedding Seluruh Token.....	52
Tabel 3.14 Contoh Output Proses IndoBERT Embedding.....	52
Tabel 3.15 Contoh Proses Pooling Layer.....	53
Tabel 3.16 Contoh Hasil Perhitungan Softmax.....	54
Tabel 3.17 Contoh Hasil Tabel Confussion Matrix Multiclass .....	54
Tabel 4.1 Sampel Data Tweet.....	56
Tabel 4.2 Sampel Hasil Seleksi Data Tweet.....	58
Tabel 4.3 Sampel Hasil Pelabelan Data .....	59
Tabel 4.4 Hasil Cleaning Data .....	60
Tabel 4.5 Hasil Case Folding .....	60
Tabel 4.6 Hasil Tokenizing.....	61
Tabel 4.7 Hasil Normalization .....	62
Tabel 4.8 Hasil Stemming.....	62
Tabel 4.9 Hasil Embedding.....	63
Tabel 4.10 Spesifikasi Perangkat .....	66
Tabel 4.11 Skenario Pengujian Model .....	67
Tabel 4.12 Parameter Model CNN.....	68
Tabel 4.13 Hasil 12 Skenario Pengujian berdasarkan Kernal dan K-Fold.....	100
Tabel 4.14 Hasil Recall dari 12 Skenario.....	101

Tabel 4.15 Hasil Precision dari 12 Skenario ..... 102  
Tabel 4.16 Hasil f1-Score dari 12 Skenario ..... 103



## DAFTAR GAMBAR

Gambar	Halaman
Gambar 2.1 Tahapan Cross-Industry Standart Process for Data Mining .....	21
Gambar 2.2 Proses Data Mining .....	23
Gambar 2.3 Proses Pelatihan pada BERT .....	27
Gambar 2.4 Proses Embedding Pada BERT .....	29
Gambar 2.5 Arsitektur Convolutional Neural Network .....	31
Gambar 2.6 Proses K-Fold Validation .....	34
Gambar 3.1 Tahapan atau Alur Penelitian.....	39
Gambar 3.2 Tahapan Text Preprocessing.....	46
Gambar 4.1 Distribusi Hasil Pelabelan .....	60
Gambar 4.2 Grafik Performa kernel 2 dengan K-Fold 5 .....	69
Gambar 4.3 Hasil Confusion Matrix Kernel 2 dengan K-Fold 5.....	70
Gambar 4.4 Hasil Classification Report Kernel 2 dengan K-Fold 5.....	70
Gambar 4.5 Grafik Performa Kernel 2 dengan K-Fold 10.....	71
Gambar 4.6 Hasil Confusion Matrix Kernel 2 dengan K-Fold 10.....	72
Gambar 4.7 Hasil Classification Report Kernel 2 dengan K-Fold 10.....	72
Gambar 4.8 Grafik Performa kernel 3 dengan K-Fold 5 .....	73
Gambar 4.9 Hasil Confusion Matrix Kernel 3 dengan K-Fold 5.....	74
Gambar 4.10 Hasil Clasification Report Kernel 3 dengan K-Fold 5 .....	74
Gambar 4.11 Grafik Performa kernel 3 dengan K-Fold 10.....	75
Gambar 4.12 Hasil Confusion Matrix Kernel 3 dengan K-Fold 10.....	76
Gambar 4.13 Hasil Classification Report Kernel 3 dengan K-Fold 10.....	76
Gambar 4.14 Grafik Performa kernel 5 dengan K-Fold 5 .....	77
Gambar 4.15 Hasil Confusion Matrix Kernel 5 dengan K-Fold 5.....	78
Gambar 4.16 Hasil Clasification Report Kernel 5 dengan K-Fold 5 .....	78
Gambar 4.17 Grafik Performa kernel 5 dengan K-Fold 10 .....	79
Gambar 4.18 Hasil Confusion Matrix Kernel 5 dengan K-Fold 10.....	80
Gambar 4.19 Hasil Classification Report Kernel 5 dengan K-Fold 10.....	80
Gambar 4.20 Distribusi Label Data Asli .....	81
Gambar 4.21 Distribusi Data Traning Sebelum dan Setelah SMOTE.....	82
Gambar 4.22 Grafik Performa kernel 2 dengan K-Fold 5 dan SMOTE .....	83
Gambar 4.23 Hasil Confusion Matrix Kernel 2 dengan K-Fold 5 dan SMOTE...	83

Gambar 4.24 Hasil Clasification Report Kernel 2 dengan K-Fold 5 dan SMOTE	84
Gambar 4.25 Grafik Performa kernel 2 dengan K-Fold 10 dan SMOTE .....	85
Gambar 4.26 Hasil Confusion Matrix Kernel 2 dengan K-Fold 10 dan SMOTE.	85
Gambar 4.27 Hasil Clasification Report Kernel 2 dengan K-Fold 10 dan SMOTE .....	86
Gambar 4.28 Grafik Performa kernel 3 dengan K-Fold 5 dan SMOTE .....	87
Gambar 4.29 Hasil Confusion Matrix Kernel 3 dengan K-Fold 5 dan SMOTE...	87
Gambar 4.30 Hasil Clasification Report Kernel 3 dengan K-Fold 5 dan SMOTE	88
Gambar 4.31 Grafik Performa kernel 3 dengan K-Fold 10 dan SMOTE .....	89
Gambar 4.32 Hasil Confusion Matrix Kernel 3 dengan K-Fold 10 dan SMOTE.	89
Gambar 4.33 Hasil Clasification Report Kernel 3 dengan K-Fold 10 dan SMOTE .....	90
Gambar 4.34 Grafik Performa kernel 5 dengan K-Fold 5 dan SMOTE .....	91
Gambar 4.35 Hasil Confusion Matrix Kernel 5 dengan K-Fold 5 dan SMOTE...	91
Gambar 4.36 Hasil Clasification Report Kernel 5 dengan K-Fold 5 dan SMOTE	92
Gambar 4.37 Grafik Performa kernel 5 dengan K-Fold 10 dan SMOTE .....	93
Gambar 4.38 Hasil Confusion Matrix Kernel 5 dengan K-Fold 10 dan SMOTE.	93
Gambar 4.39 Hasil Clasification Report Kernel 5 dengan K-Fold 10 dan SMOTE .....	94
Gambar 4.40 Wordcloud Dataset Sentimen Positif.....	95
Gambar 4.41 Wordcloud Dataset Sentimen Netral .....	96
Gambar 4.42 Wordcloud Dataset Sentimen Negatif .....	96
Gambar 4.43 Tampilan Utama DASHBOARD Analisis Sentimen .....	97
Gambar 4.44 Tampilan Hasil Analisis Sentimen .....	98
Gambar 4.45 Tampilan Hasil Pemrosesan Teks .....	98
Gambar 4.46 Tampilan Lihat Detail Hasil Embedding IndoBERT .....	99

## DAFTAR LAMPIRAN

Lampiran	Halaman
Lampiran 1. Riwayat hidup.....	114
Lampiran 2. Surat Permohonan Pelabelan Data .....	115
Lampiran 3. Surat Validasi Pelabelan Data.....	116
Lampiran 4. Sertifikasi Pendidik Para Ahli.....	119
Lampiran 5. Dokumentasi Pelabelan Data.....	120
Lampiran 6. Code Crawling Data .....	120
Lampiran 7 Kode Pre-Processing Data.....	121
Lampiran 8. Kode Modeling.....	123
Lampiran 9. Kode Evaluasi.....	125

