

**TUGAS AKHIR**

**PERBANDINGAN AKURASI ANOTASI SENTIMEN  
ANTARA PRETRAINED INDOBERT DAN MODEL  
BILSTM PADA ULASAN PENGGUNA APLIKASI BRIMO**



**RIO GHANIY SEPTIANSYAH**

20110016

**PROGRAM STUDI S1 SAINS DATA**

**FAKULTAS INFORMATIKA**

**INSTITUT TEKNOLOGI TELKOM PURWOKERTO**

**2024**

**TUGAS AKHIR**

**PERBANDINGAN AKURASI ANOTASI SENTIMEN  
ANTARA PRETRAINED INDOBERT DAN MODEL  
BILSTM PADA ULASAN PENGGUNA APLIKASI BRIMO**

***COMPARISON OF SENTIMENT ANNOTATION  
ACCURACY BETWEEN PRETRAINED INDOBERT AND  
BILSTM MODELS ON BRIMO APP USER REVIEWS***

Disusun Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Sains Data



RIO GHANIY SEPTIANSYAH

20110016

**PROGRAM STUDI S1 SAINS DATA**

**FAKULTAS INFORMATIKA**

**INSTITUT TEKNOLOGI TELKOM PURWOKERTO**

**2024**

**LEMBAR PERSETUJUAN PEMBIMBING**

**PERBANDINGAN AKURASI ANOTASI SENTIMEN  
ANTARA INDOBERT DAN MODEL BILSTM PADA  
ULASAN PENGGUNA APLIKASI BRIMO**

***COMPARISON OF SENTIMENT ANNOTATION  
ACCURACY BETWEEN INDOBERT AND BILSTM  
MODELS ON BRIMO APP USER REVIEWS***

Dipersiapkan dan Disusun oleh

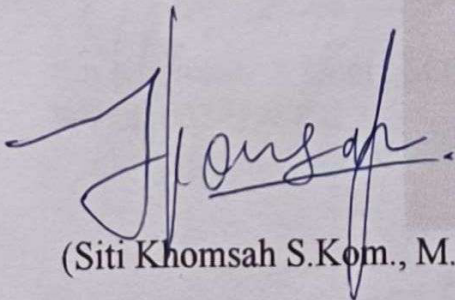
Rio Ghaniy Septiansyah  
20110016

**Fakultas Informatika**

**Institut Teknologi Telkom Purwokerto**

**Pada Tanggal: 7 Juni 2024**

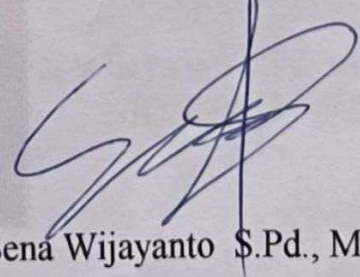
Pembimbing Utama,



(Siti Khomsah S.Kom., M.Cs)

NIDN: 0517108101

Pembimbing Pendamping,



(Sena Wijayanto S.Pd., M.T.)

NIDN: 0613109201

**PERBANDINGAN AKURASI ANOTASI SENTIMEN  
ANTARA PRETRAINED INDOBERT DAN MODEL  
BILSTM PADA ULASAN PENGGUNA APLIKASI BRIMO**

***COMPARISON OF SENTIMENT ANNOTATION  
ACCURACY BETWEEN PRETRAINED INDOBERT AND  
BILSTM MODELS ON BRIMO APP USER REVIEWS***

Disusun oleh

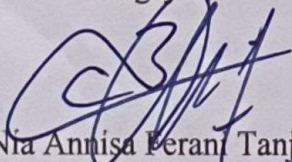
Rio Ghaniy Septiansyah

20110016

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas Akhir Pada

19 Juni 2024

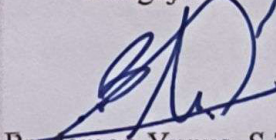
Penguji I



Nia Annisa Perani Tanjung, S.Si.,  
M.Sc.

NIDN: 0630049203

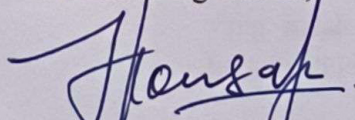
Penguji II



Andi Pademon Yunus, S.T.,  
M.Sc.Eng., Ph.D

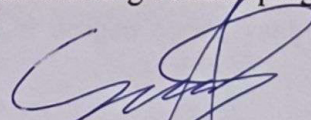
NIDN: 0616129301

Pembimbing Utama,



Siti Khomsah, S.Kom., M.Cs  
NIDN: 0517108101

Pembimbing Pendamping,



Sena Wijayanto, S.Pd. M.T.  
NIDN: 0613109201



Dekan  
Auliya Burhanuddin, S.Si., M.Kom.  
NIK: 19820008

## HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama mahasiswa : Rio Ghaniy Septiansyah  
NIM : 20110016  
Program Studi : S1 Sains Data

Menyatakan bahwa Tugas Akhir dengan judul berikut:  
***PERBANDINGAN ANOTASI SENTIMEN ANTARA INDOBERT  
DAN MODEL BILSTM PADA PENGGUNA APLIKASI BRIMO***

Dosen Pembimbing Utama : Siti Khomsah, S.Kom., M.Cs.  
Dosen Pembimbing Pendamping : Sena Wijayanto, S.Pd., M.T.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Institut Teknologi Telkom Purwokerto maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan, dan penelitian Saya Sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing.
3. Dalam Karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggungjawab Saya, bukan tanggungjawab Institut Teknologi Telkom Purwokerto.
5. Pernyataan ini Saya buat dengan sesungguhnya, apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima Sanksi Akademik dengan pencabutan gelar yang sudah diperoleh serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Purwokerto, 7 Juni 2024,  
Yang Menyatakan,



(Rio Ghaniy Septiansyah)

## KATA PENGANTAR

Puji syukur syukur kehadirat Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya penulis dapat menyelesaikan tugas akhir yang berjudul **“PERBANDINGAN AKURASI ANOTASI SENTIMEN ANTARA PRETRAINED INDOBERT DAN MODEL BILSTM PADA ULASAN PENGGUNA APLIKASI BRIMO”**. Selama proses penyusunan tugas akhir ini, penulis telah mendapatkan banyak bimbingan, dorongan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Allah SWT. karena atas izinnya penulis dapat menyelesaikan tugas akhir ini.
2. Ibu Dr. Tenia Wahyuningrum, S.Kom., M.T. selaku Rektor Institut Teknologi Telkom Purwokerto.
3. Bapak Auliya Burhanuddin, S.Si., M.Kom. selaku Dekan Fakultas Informatika.
4. Ibu Siti Khomsah, S.Kom., M.Cs. selaku Ketua Program Studi Sains Data sekaligus dosen pembimbing utama yang telah memberikan bimbingan, arahan dan dukungan yang sangat berharga selama proses penyusunan tugas akhir ini.
5. Bapak Sena Wjayanto, S.Pd., M.T. selaku dosen pembimbing pendamping yang telah memberikan masukan dan saran yang sangat bermanfaat.
5. Ibu Nia Annisa Ferani Tanjung, S.Si., M.Sc. selaku dosen penguji yang telah memberikan masukan dan saran yang sangat bermanfaat.
6. Bapak Andi Prademon Yunus, S.T., M.Sc.Eng., Ph.D selaku dosen penguji yang telah memberikan masukan dan saran yang sangat bermanfaat.
8. Kedua orang tua penulis yang selalu memberikan dukungan, doa, dan semangat dalam menyelesaikan tugas akhir ini.
7. Seluruh teman-teman seperjuangan yang telah memberikan dukungan dan semangat dalam menyelesaikan tugas akhir.

8. Seluruh Staff Dosen Fakultas Informatika, Institut Teknologi Telkom Purwokerto yang telah memberikan bimbingan, dukungan, dan pengajaran pada segenap ilmu yang diberikan.

Penulis berharap tugas akhir ini dapat bermanfaat bagi semua pihak yang membacanya dan mohon maaf jika dalam penyusunan tugas akhir ini terdapat kesalahan. Penulis sangat mengharapkan kritik dan saran yang membangun demi perbaikan kualitas tugas akhir ini di masa yang akan datang. Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Namun dengan adanya dukungan dari berbagai pihak, penulis berharap tugas akhir ini dapat bermanfaat bagi semua pihak yang membacanya.

## DAFTAR ISI

TUGAS AKHIR .....	ii
LEMBAR PERSETUJUAN PEMBIMBING .....	iii
HALAMAN PENGESAHAN TUGAS AKHIR .....	iv
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI .....	viii
DAFTAR GAMBAR .....	xiii
ABSTRAK .....	xiv
<i>ABSTRACT</i> .....	xv
BAB I PENDAHULUAN .....	1
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	3
1.3. Pertanyaan Penelitian .....	3
1.4. Batasan Masalah .....	3
1.5. Tujuan Penelitian .....	4
BAB II TINJAUAN PUSTAKA .....	5
2.1. Tinjauan Pustaka .....	5
2.2. Landasan Teori .....	17
2.2.1. Anotasi Sentimen .....	17
2.2.2. <i>Text Preprocessing</i> .....	17
2.2.3. <i>Word Embedding</i> .....	18
2.2.4. IndoBERT .....	24
2.2.5. LSTM .....	29



2.2.6.	BILSTM .....	33
2.2.7.	<i>Classification Report</i> .....	35
2.2.8.	<i>Cross entropy</i> .....	37
BAB III METODOLOGI PENELITIAN .....		38
3.1.	Subjek dan Objek Penelitian .....	38
3.2.	Alat dan Bahan Penelitian .....	38
3.2.1.	Perangkat Keras / <i>Hardware</i> .....	38
3.2.2.	Perangkat Lunak / <i>Software</i> .....	38
3.2.3.	Bahan / Data .....	38
3.3.	Diagram Alir Penelitian .....	38
3.3.1	Identifikasi Masalah .....	39
3.3.2	Studi Literatur .....	39
3.3.3	Pengumpulan Dataset .....	40
3.3.4	<i>Text Preprocessing</i> .....	40
3.3.5	Pelabelan Data .....	43
3.3.6	<i>Word Embedding</i> .....	43
3.3.7	Pembangunan Model BiLSTM .....	47
3.3.8	Pengujian model BiLSTM dan IndoBERT .....	48
3.3.9	Evaluasi dan Analisis Hasil .....	48
BAB IV HASIL DAN PEMBAHASAN .....		50
4.1	Hasil Penelitian .....	50
4.2	Pengambilan Dataset .....	50
4.3	<i>Text Preprocessing</i> .....	50
4.4	Pelabelan Data .....	53
4.5	<i>Word Embedding</i> .....	53

4.6	Pembangunan Model BiLSTM .....	56
4.7	Pengujian model BiLSTM dan IndoBERT .....	59
4.8	Evaluasi dan Analisis Hasil .....	60
BAB V KESIMPULAN DAN SARAN .....		64
5.1	Kesimpulan .....	64
5.1	Saran .....	64
DAFTAR PUSTAKA .....		65
LAMPIRAN .....		69

## DAFTAR TABEL

Tabel 2. 1 Penelitian terdahulu .....	8
Tabel 2. 2 One-hot encoded vector .....	21
Tabel 3. 1 Dataset.....	40
Tabel 3. 2 Case folding .....	41
Tabel 3. 3 Removing noise .....	41
Tabel 3. 4 Replacement of slang and abbreviations .....	42
Tabel 3. 5 Removing stopwords .....	42
Tabel 3. 6 Tokenization .....	43
Tabel 3. 7 Pelabelan Data .....	43
Tabel 3. 8 Encoding hasil kata .....	44
Tabel 3. 9 Nilai vektor serta output kata .....	46
Tabel 3. 10 Word output dan vektor kalimat x .....	47
Tabel 3. 11 Word output dan vektor kalimat y .....	47
Tabel 3. 12 Parameter tuning model BiLSTM .....	48
Tabel 4. 1 Dataset Brimo.....	50
Tabel 4. 2 Hasil case folding .....	51
Tabel 4. 3 Hasil removing noise .....	51
Tabel 4. 4 Replacement of slang and abbreviations .....	52
Tabel 4. 5 Hasil removing stopwords .....	52
Tabel 4. 6 Hasil tokenization .....	53
Tabel 4. 7 Label data .....	53
Tabel 4. 8 Word Embedding .....	53
Tabel 4. 9 Representasi kata dalam angka .....	54
Tabel 4. 10 Representasi vektor kalimat .....	55
Tabel 4. 11 Representasi vektor kata .....	55
Tabel 4. 12 Parameter BiLSTM .....	56
Tabel 4. 13 Hasil parameter tuning model BiLSTM .....	57
Tabel 4. 14 Hasil perbandingan pelabelan sentimen .....	59
Tabel 4. 15 Hasil perbandingan model BiLSTM dan pretrained IndoBERT .....	62

Tabel 4. 16 Hasil Inference Consumption .....	62
---	----

## DAFTAR GAMBAR

Gambar 2. 1 Arsitektur word2vec sederhana .....	19
Gambar 2. 2 Ilustrasi windowing skip-gram .....	20
Gambar 2. 3 Ilustrasi struktur neural network skip-gram .....	20
Gambar 2. 4 Feed forward neural network skip-gram .....	21
Gambar 2. 5 Skip-gram .....	22
Gambar 2. 6 Ilustrasi windowing CBOW .....	23
Gambar 2. 7 Ilustrasi neural network CBOW .....	23
Gambar 2. 8 Ilustrasi feedforward Word2vec CBOW .....	24
Gambar 2. 9 Arsitektur transformer encoder .....	25
Gambar 2. 10 Ilustrasi proses positional encoding .....	26
Gambar 2. 11 Ilustrasi matriks self attention mechanism .....	27
Gambar 2. 12 Ilustrasi hasil perhitungan .....	27
Gambar 2. 13 Contoh kalimat next sentence prediction .....	29
Gambar 2. 14 Arsitektur LSTM .....	29
Gambar 2. 15 Struktur forget layer pada LSTM .....	30
Gambar 2. 16 Struktur input gate fase 1 .....	31
Gambar 2. 17 Struktur input gate fase 2 .....	32
Gambar 2. 18 Struktur output gate pada LSTM .....	33
Gambar 2. 19 Arsitektur BiLSTM .....	34
Gambar 2. 20 Arsitektur hidden layer BiLSTM .....	35
Gambar 2. 21 Konfusi matriks .....	36
Gambar 3. 1 Diagram alir penelitian.....	39
Gambar 3. 2 Skip-gram dua window .....	44
Gambar 4. 1 Konfusi matriks hasil pemodelan.....	58
Gambar 4. 2 Konfusi matriks pengujian BiLSTM .....	61
Gambar 4. 3 Konfusi matriks pengujian IndoBERT .....	61