

**DETEKSI BERITA PALSU POLITIK INDONESIA MENGGUNAKAN
INDOBERT-BI-GRU-ATTENTION**

TESIS



Oleh :

JULIANA DAMAYANTI MANURUNG

NIM. 221231098



PROGRAM STUDI S-2 TEKNOLOGI INFORMASI

FAKULTAS INFORMATIKA

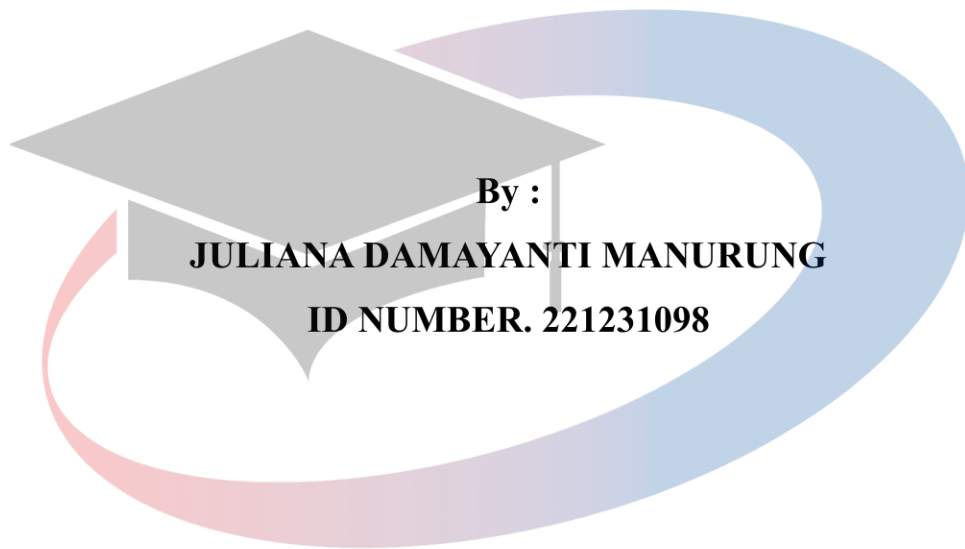
UNIVERSITAS MIKROSKIL

MEDAN

2024

**INDONESIAN POLITICAL FAKE NEWS DETECTION USING
INDOBERT-BI-GRU-ATTENTION**

THESIS



**UNIVERSITAS
MIKROSKIL**

MAJOR OF S-2 INFORMATION TECHNOLOGY

FACULTY OF INFORMATICS

UNIVERSITAS MIKROSKIL

MEDAN

2024

LEMBARAN PENGESAHAN

DETEKSI BERITA PALSU POLITIK INDONESIA MENGGUNAKAN
INDOBERT-BI-GRU-ATTENTION

TESIS

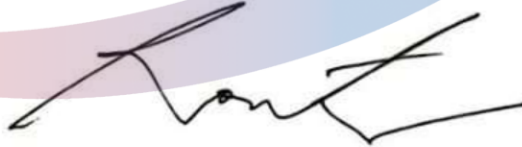
Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Magister
Program Studi S-2 Teknologi Informasi

Oleh:

JULIANA DAMAYANTI MANURUNG
NIM. 221231098

Disetujui Oleh:

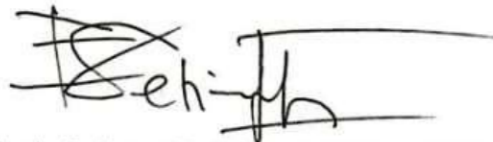
Dosen Pembimbing,



Dr. Ronsen Purba, M.Sc.

Medan, 5 Agustus 2024
Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-2 Teknologi Informasi,



Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D.

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-2 Teknologi Informasi Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Juliana Damayanti Manurung

NIM : 221231098

Saya telah melaksanakan penelitian dan penulisan Tesis dengan judul dan tempat penelitian sebagai berikut:

Judul Tesis : Deteksi Berita Palsu Politik Indonesia Menggunakan IndoBERT-Bi-GRU-Attention

Tempat Penelitian : -

Alamat Tempat Penelitian : -

No. Telp. Tempat Penelitian : -

Sehubungan dengan Tesis tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Tesis tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Tesis saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tesis saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tesis saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 23 Juli 2024

Saya yang membuat pernyataan,



Juliana Damayanti Manurung

DETEKSI BERITA PALSU POLITIK INDONESIA MENGGUNAKAN INDOBERT-BI-GRU-ATTENTION

Abstrak

Penyebaran berita palsu yang instan dan masif di media sosial memberikan dampak negatif pada kepercayaan masyarakat terhadap media dan lembaga berita. Dalam bidang politik, berita palsu sering dimanfaatkan oleh politisi untuk meraih dukungan menjelang Pemilu. Pendeteksian berita palsu di Indonesia merupakan tantangan besar, terutama bagi masyarakat yang rentan terhadap informasi palsu. Penelitian ini bertujuan membangun model terbaru yang menggabungkan IndoBERT dengan Bi-GRU dan Attention. Selain itu, dilakukan perbandingan model utama dengan dua model word embedding, FastText dan GloVe. Pengujian dilakukan terhadap dataset judul dan narasi berita secara terpisah. Data berasal dari CNN, Tempo.co, Kompas, dan TurnBackHoax.ID. Hasil penelitian menunjukkan bahwa model IndoBERT-Bi-GRU-Attention dengan FastText unggul pada dataset judul dengan akurasi 99.76% dan F1-Score 99.61%, sedangkan model utama IndoBERT-Bi-GRU-Attention unggul pada dataset narasi dengan akurasi 99.08% dan F1-Score 98.40%. Penelitian ini membuktikan bahwa IndoBERT dapat dikombinasikan dengan Bi-GRU dan memberikan kontribusi signifikan dalam pengembangan model pendeteksian berita palsu.

Kata kunci: *hoaks, IndoBERT, pembelajaran mendalam, representasi makna kata*

Abstract

The instant and massive spread of fake news on social media negatively impacts public trust in the media and news agencies. In politics, fake news is often used by politicians to gain support ahead of elections. Detecting fake news in Indonesia poses a significant challenge, especially for communities vulnerable to misinformation. This study aims to develop a new model that combines IndoBERT with Bi-GRU and Attention. Additionally, a comparison is made between the main model and two word embedding models, FastText and GloVe. The tests were conducted on datasets of headlines and news narratives separately. Data was sourced from CNN, Tempo.co, Kompas, and TurnBackHoax.ID. The results show that the IndoBERT-Bi-GRU-Attention model with FastText excelled on the headline dataset with an accuracy of 99.76% and an F1-Score of 99.61%, while the main IndoBERT-Bi-GRU-Attention model excelled on the narrative dataset with an accuracy of 99.08% and an F1-Score of 98.40%. This research demonstrates that IndoBERT can be combined with Bi-GRU, significantly contributing to the development of fake news detection models.

Keywords: *hoax, IndoBERT, deep learning, word embedding*

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan Tesis yang berjudul “DETEKSI BERITA PALSU POLITIK INDONESIA MENGGUNAKAN INDOBERT-BI-GRU-ATTENTION”.

Tesis ini dibuat untuk melengkapi persyaratan kurikulum pada Program Studi S-2 Teknologi Informasi, Universitas Mikroskil Medan.

Penulis juga menyampaikan banyak terima kasih kepada:

1. Bapak Dr. Ronsen Purba, M.Sc., selaku Dosen Pembimbing yang telah membimbing penulis dalam proses penyelesaian tesis ini.
2. Ibu Dr. Sofiana Nurjanah S.Kom., M.T.I., selaku Dosen Pendamping Pembimbing yang telah membimbing penulis dalam proses penyelesaian tesis ini.
3. Bapak Hardy, S.Kom, M.Sc, Ph.D., selaku Rektor Universitas Mikroskil Medan.
4. Bapak Ng Poi Wong, S.Kom, M.T.I., selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
5. Bapak Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D., selaku Ketua Program Studi S-2 Teknologi Informasi Universitas Mikroskil Medan.
6. Bapak atau Ibu Dosen Universitas Mikroskil Medan yang telah membantu proses penulisan tesis ini.

Penulis menyadari bahwa masih terdapat banyak kekurangan yang ada. Oleh sebab itu, kritik dan saran yang bersifat membangun sangat diterima. Akhir kata, semoga Tesis ini dapat bermanfaat bagi masyarakat. Terima kasih.

Medan, 25 Juli 2024

Penulis,

Juliana Damayanti Manurung

DAFTAR ISI

Abstrak	i
KATA PENGANTAR	ii
DAFTAR ISI	iii
DAFTAR GAMBAR	iv
DAFTAR TABEL	v
DAFTAR LAMPIRAN	vi
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan	2
1.4. Manfaat Penelitian	3
1.5. Ruang Lingkup	3
BAB II KAJIAN LITERATUR	4
2.1. Tinjauan Pustaka	4
2.1.1. Berita Palsu	4
2.1.2. Pendeteksian Berita Palsu.....	5
2.1.3. Word Embedding	5
2.1.4. IndoBERT.....	8
2.1.5. Bidirectional Gated Recurrent Unit.....	10
2.1.6. Attention Mechanism	11
2.1.7. Confusion Matrix	12
2.1.8. Adam optimizer	14
2.1.9. Binary Crossentropy.....	14
2.2. Penelitian Terdahulu	15
2.3. Kerangka Konsep	17
BAB III METODOLOGI PENELITIAN	20
3.1. Analisis Masalah	20
3.2. Rancangan Penelitian.....	21
3.3. Data yang Digunakan	25
3.4. Evaluasi Model	27
3.5. Alat-alat Penelitian	27
3.6. Jadwal Penelitian.....	28
Abstrak	i
KATA PENGANTAR	ii
DAFTAR ISI	iii
DAFTAR GAMBAR	iv

DAFTAR TABEL	v
DAFTAR LAMPIRAN	vi
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan	2
1.4. Manfaat Penelitian	3
1.5. Ruang Lingkup	3
BAB II KAJIAN LITERATUR	4
2.1. Tinjauan Pustaka	4
2.1.1. Berita Palsu	4
2.1.2. Pendeteksian Berita Palsu	5
2.1.3. Word Embedding	5
2.1.4. IndoBERT	8
2.1.5. Bidirectional Gated Recurrent Unit	10
2.1.6. Attention Mechanism	11
2.1.7. Confusion Matrix	12
2.1.8. Adam optimizer	14
2.1.9. Binary Crossentropy	14
2.2. Penelitian Terdahulu	15
2.3. Kerangka Konsep	17
BAB III METODOLOGI PENELITIAN	20
3.1. Analisis Masalah	20
3.2. Rancangan Penelitian.....	21
3.3. Data yang Digunakan	25
3.4. Evaluasi Model	27
3.5. Alat-alat Penelitian	27
3.6. Jadwal Penelitian.....	28
BAB IV HASIL DAN PEMBAHASAN	29
4.1. Hasil.....	29
4.1.1. Membaca Dataset.....	29
4.1.2. Pra-proses Data.....	31
Abstrak	i
KATA PENGANTAR	ii
DAFTAR ISI	iii
DAFTAR GAMBAR	iv
DAFTAR TABEL	v
DAFTAR LAMPIRAN	vi

BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Tujuan.....	2
1.4. Manfaat Penelitian.....	3
1.5. Ruang Lingkup.....	3
BAB II KAJIAN LITERATUR.....	4
2.1. Tinjauan Pustaka.....	4
2.1.1. Berita Palsu.....	4
2.1.2. Pendeteksian Berita Palsu.....	5
2.1.3. Word Embedding.....	5
2.1.4. IndoBERT.....	8
2.1.5. Bidirectional Gated Recurrent Unit.....	10
2.1.6. Attention Mechanism.....	11
2.1.7. Confusion Matrix.....	12
2.1.8. Adam optimizer.....	14
2.1.9. Binary Crossentropy.....	14
2.2. Penelitian Terdahulu.....	15
2.3. Kerangka Konsep.....	17
BAB III METODOLOGI PENELITIAN.....	20
3.1. Analisis Masalah.....	20
3.2. Rancangan Penelitian.....	21
3.3. Data yang Digunakan.....	25
3.4. Evaluasi Model.....	27
3.5. Alat-alat Penelitian.....	27
3.6. Jadwal Penelitian.....	28
BAB IV HASIL DAN PEMBAHASAN.....	29
4.1. Hasil.....	29
4.1.1. Membaca Dataset.....	29
4.1.2. Pra-proses Data.....	31
4.1.3. Pemodelan.....	32

4.2. Pembahasan.....	36
4.2.1. Perbandingan Kinerja Model pada Dataset Judul Berita	37
4.2.2. Perbandingan Kinerja Model pada Dataset Narasi Berita.....	38
BAB V PENUTUP	40
5.1. Kesimpulan	40
5.2. Saran.....	40
DAFTAR PUSTAKA	41
DAFTAR RIWAYAT	
HIDUP.....	46



UNIVERSITAS MIKROSKIL

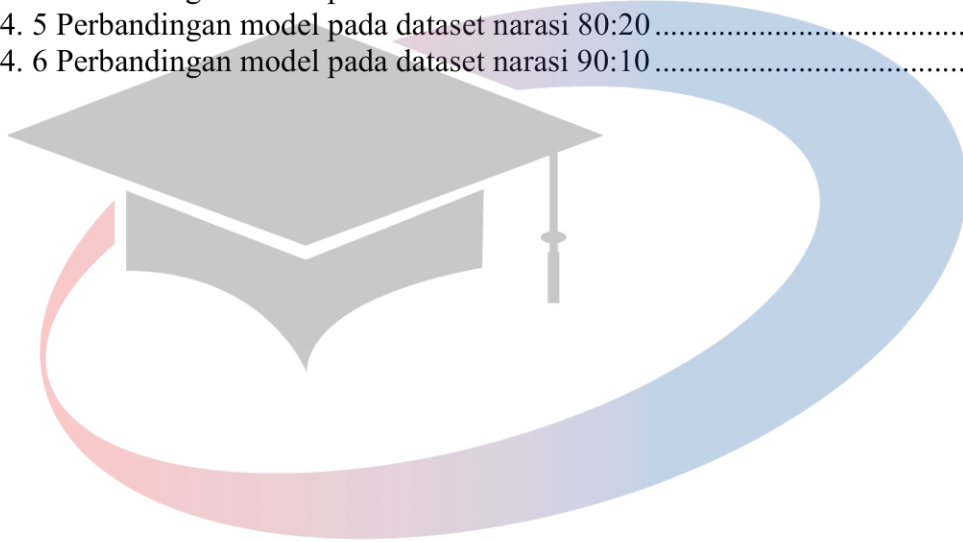
DAFTAR GAMBAR

Gambar 2. 1 Model Arsitektur FastText[22]	7
Gambar 2. 2 FastText Learning Method (window size = 5)[19]	7
Gambar 2. 3 Rasio Ko-okurensi GloVe	8
Gambar 2. 4 Arsitektur Model IndoBERT[28]	9
Gambar 2. 5 Detail Arsitektur Model IndoBERT[27]	10
Gambar 2. 6 Arsitektur Bidirectional GRU	10
Gambar 2. 7 Confusion Matrix [28]	13
Gambar 2. 8 Hasil Penelitian Ryan Pradana et al (2021)	16
Gambar 2. 9 Kerangka Konsep Pemecahan Masalah	18
Gambar 3. 1 Neural Network Training Word2Vec Model[21]	6
Gambar 3. 2 Metode Penelitian	21
Gambar 3. 3 Model IndoBERT-BiGRU-Attention	22
Gambar 3. 4 Struktur IndoBERT-Bi-GRU-Attention	25
Gambar 3. 5 Dataset Fakta	26
Gambar 3. 6 Dataset Hoax	26
Gambar 3. 7 Data yang telah melalui preprocessing	27
Gambar 4. 1 Dataset Narasi	29
Gambar 4. 2 Distribusi Dataset Narasi	30
Gambar 4. 3 Syntax baca data berita	30
Gambar 4. 4 Data yang telah melalui pre-processing	31
Gambar 4. 5 Syntax pre-processing data	31
Gambar 4. 6 Pemodelan IndoBERT-Bi-GRU-Attention	33
Gambar 4. 7 Pemodelan IndoBERT-Bi-GRU-Attention dengan FastText	34
Gambar 4. 8 Pemodelan IndoBERT-Bi-GRU-Attention dengan GloVe	36

UNIVERSITAS
MIKROSKIL

DAFTAR TABEL

Tabel 3. 1 Tahapan Case Folding.....	22
Tabel 3. 2 Tahapan Remove Punctuation	23
Tabel 3. 3 Tahapan Remove Stopwords	23
Tabel 3. 4 Tahapan Stemming	24
Tabel 4. 1 Perbandingan kinerja model pada dataset judul 70:30	37
Tabel 4. 2 Perbandingan kinerja model pada dataset judul 80:20	37
Tabel 4. 3 Perbandingan kinerja model pada dataset judul 90:10	37
Tabel 4. 4 Perbandingan model pada dataset narasi 70:30	38
Tabel 4. 5 Perbandingan model pada dataset narasi 80:20	38
Tabel 4. 6 Perbandingan model pada dataset narasi 90:10	38



UNIVERSITAS
MIKROSKIL

DAFTAR LAMPIRAN

Lampiran 1 Gambar44

