

**DETEKSI POTENSI DEPRESI DARI UNGGAHAN MEDIA  
SOSIAL X MENGGUNAKAN INDOBERT**

**TESIS**

**Oleh:**

**GILBERT FERNANDO SITUMORANG  
NIM. 231231029**



**UNIVERSITAS  
MIKROSKIL**

**PROGRAM STUDI S-2 TEKNOLOGI INFORMASI  
FAKULTAS INFORMATIKA  
UNIVERSITAS MIKROSKIL  
MEDAN  
2024**

**POTENTIAL DEPRESSION DETECTION FROM SOCIAL  
MEDIA X POSTS USING INDOBERT**

**THESIS**

**By:**

**GILBERT FERNANDO SITUMORANG  
ID NUMBER. 231231029**



**UNIVERSITAS  
MIKROSKIL**

**MAJOR OF S-2 INFORMATION TECHNOLOGY  
FACULTY OF INFORMATICS  
UNIVERSITAS MIKROSKIL  
MEDAN  
2024**

## ABSTRAK

# DETEKSI POTENSI DEPRESI DARI UNGGAHAN MEDIA SOSIAL X MENGGUNAKAN INDOBERT

### Abstrak

Gangguan mental seperti depresi meningkat dan menjadi masalah kesehatan masyarakat yang serius. Banyak orang terdampak tidak mencari dukungan profesional karena stigma sosial. Media sosial seperti X memberikan peluang besar untuk mempelajari kesehatan mental, karena pengguna sering berbagi pengalaman dan emosi pribadi. Namun, memahami pola bahasa dan konteks unggahan menjadi tantangan, sehingga diperlukan teknik dan model yang tepat untuk mendeteksi potensi depresi. Penelitian ini menggunakan teknik Natural Language Processing (NLP) untuk menganalisis 37.554 teks dari unggahan media sosial, bertujuan mendeteksi potensi depresi. Model IndoBERT, adaptasi BERT yang dilatih pada teks berbahasa Indonesia, digunakan untuk mengidentifikasi potensi depresi. Data dikumpulkan melalui scraping dengan kata kunci berkonotasi negatif dan positif yang telah dikonsultasikan dengan psikiater. Proses pembersihan teks meliputi case folding, text cleaning, spell normalization, stopword removal, dan stemming. Data kemudian diberi label menggunakan model klasifikasi emosi IndoBERT, yang mengkategorikan emosi negatif sebagai depresi dan positif sebagai normal. Model dilatih dan dievaluasi menggunakan metrik akurasi, precision, recall, dan F1-score, dengan hasil terbaik menunjukkan akurasi 94,91%. Hasil menunjukkan bahwa IndoBERT efektif dalam mendeteksi potensi depresi dari teks media sosial, meskipun ada keterbatasan terkait representasi emosional pengguna.

**Kata kunci:** Deep Learning, Deteksi Depresi, IndoBERT, Media Sosial, Natural Language Processing

### Abstract

Mental health disorders like depression have been on the rise, becoming a serious public health concern. Many affected individuals avoid seeking professional help due to social stigma. Social media platforms like X offer valuable opportunities to study mental health, as users frequently share personal experiences and emotions. However, understanding the language patterns and context of posts poses challenges, necessitating appropriate techniques and models to effectively detect potential depression. This study utilizes Natural Language Processing (NLP) techniques to analyze 37,554 texts from social media posts, aiming to detect potential depression. The IndoBERT model, an adaptation of BERT trained on Indonesian-language texts, is employed to identify potential depression. Data was collected through scraping, using negative and positive connotation keywords that were consulted with psychiatrists. The text cleaning process included case folding, text cleaning, spell normalization, stopword removal, and stemming. The data was then labeled using the IndoBERT emotion classification model, categorizing negative emotions as depression and positive emotions as normal. The model was trained and evaluated using accuracy, precision, recall, and F1-score metrics, with the best results showing 94.91% accuracy. The findings indicate that IndoBERT is effective in detecting potential depression from social media texts, though limitations exist regarding the emotional representation of users.

**Keywords:** Deep Learning, Depression Detection, IndoBERT, Natural Language Processing, Social Media

## KATA PENGANTAR

Ucapan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, atas berkat dan Rahmat-Nya penulis dapat menyelesaikan Tesis yang berjudul “ Deteksi Potensi Depresi dari Unggahan Media Sosial X Menggunakan Teknik NLP dan Model IndoBERT.

1. Bapak Dr. Ronsen Purba M.Sc., selaku Dosen Pembimbing I yang telah membimbing dan memberikan saran serta motivasi dalam penyusunan Tesis ini.
2. Bapak Hardy, S.Kom, M.Sc., Ph.D., selaku Rektor Universitas Mikroskil.
3. Bapak Ng Poi Wong, S.Kom., M.T.I., selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
4. Bapak Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D., selaku Ketua Program Studi S-2 Teknologi Informasi Fakultas Informatika Universitas Mikroskil.
5. Bapak dan Ibu Dosen Program Studi S-2 Teknologi Informasi Universitas Mikroskil yang telah membimbing dan memberikan pengetahuan yang sangat berarti kepada penulis.
6. Orang tua dan keluarga yang senantiasa memberikan dukungan, nasihat dan doa yang selalu dipanjatkan agar ilmu yang penulis dapatkan mampu memberikan manfaat bagi penulis dan orang banyak.
7. Teman-teman dan rekan kerja penulis yang tiada henti dalam memberikan dorongan dan semangat kepada penulis untuk menyelesaikan penulisan Tesis ini.

Penulis menyadari bahwa Tesis ini masih terdapat banyak kekurangan karena keterbatasan kemampuan dari penulis. Oleh karena itu, kritik dan saran dari pembaca sangat diharapkan untuk penulisan yang lebih baik di masa yang akan datang dan semoga Tesis ini dapat memberikan manfaat bagi pembaca.

Medan, 09 Juli 2024

Penulis,

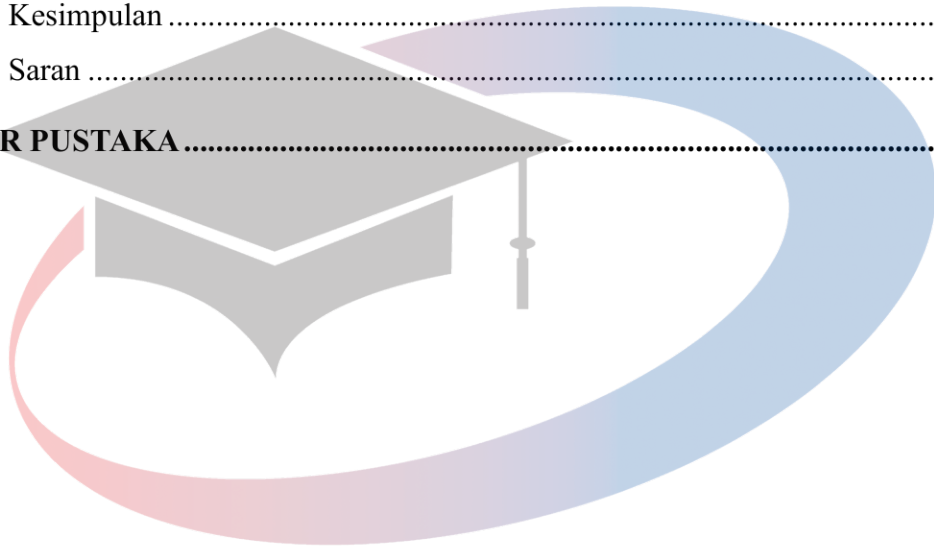


Gilbert Fernando Situmorang

# DAFTAR ISI

<b>ABSTRAK</b> .....	<b>I</b>
<b>KATA PENGANTAR</b> .....	<b>II</b>
<b>DAFTAR ISI</b> .....	<b>III</b>
<b>DAFTAR GAMBAR</b> .....	<b>V</b>
<b>DAFTAR TABEL</b> .....	<b>VI</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	3
1.3. Tujuan .....	3
1.4. Manfaat .....	3
1.5. Ruang Lingkup .....	4
<b>BAB II KAJIAN LITERATUR</b> .....	<b>5</b>
2.1. Tinjauan Pustaka .....	5
2.1.1. Deteksi Potensi Depresi.....	5
2.1.2. <i>Natural Language Processing</i> .....	5
2.1.3. BERT ( <i>Bidirectional Encoder Representations from Transformers</i> ) .....	6
2.1.4. BERT <i>Tokenizer</i> .....	7
2.1.5. IndoBERT .....	7
2.1.6. Evaluasi Model.....	8
2.2. Penelitian Terdahulu .....	9
<b>BAB III METODOLOGI PENELITIAN</b> .....	<b>11</b>
3.1. Analisis Masalah.....	11
3.2. Pengumpulan Data .....	11
3.3. Metode Penelitian .....	12
3.4. Alat – alat Penelitian.....	16
3.5. Teknik Analisis Data. ....	16
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	<b>17</b>
4.1. Hasil .....	17

4.1.1.	Normalisasi <i>Dataset</i> .....	17
4.1.2.	Pemrosesan Teks.....	18
4.1.3.	Pelabelan Data .....	18
4.1.4.	Pembuatan Model.....	22
4.1.5.	Evaluasi Model.....	23
4.1.6.	Implementasi Model.....	26
4.2.	Pembahasan .....	27
<b>BAB V PENUTUP .....</b>		<b>29</b>
5.1.	Kesimpulan .....	29
5.2.	Saran .....	29
<b>DAFTAR PUSTAKA .....</b>		<b>30</b>



# UNIVERSITAS MIKROSKIL

## DAFTAR GAMBAR

Gambar 2.1 Arsitektur Dasar Model BERT [17].....	6
Gambar 2.2 Proses Tokenisasi dalam Model BERT [18].....	7
Gambar 3.1 Alur Metode Penelitian.....	12
Gambar 3.2 Alur Pelabelan Data.....	14
Gambar 4.1 Proses <i>Preprocessing Text</i> .....	18
Gambar 4.2 Distribusi Label Emosi.....	20
Gambar 4.3 Distribusi Label Depresi.....	20
Gambar 4.4 <i>Word Cloud</i> dengan Label Normal.....	21
Gambar 4.5 <i>Word Cloud</i> dengan label berpotensi Depresi.....	21
Gambar 4.6 Proses Tokenisasi pada Teks <i>Input</i> Menggunakan <i>IndoBERT Tokenizer</i> .....	22
Gambar 4.7 <i>Confusion Matrix</i> dari Model <i>IndoBERT</i> dengan Performa Terbaik.....	26



UNIVERSITAS  
MIKROSKIL

## DAFTAR TABEL

Tabel 2.1 Varian Model IndoBERT .....	8
Tabel 2.2 Evaluasi Performa Menggunakan <i>Confusion Matrix</i> .....	9
Tabel 3.1 Kata Kunci Penelusuran Data .....	12
Tabel 3.2 Sampel <i>Spell Normalization</i> .....	13
Tabel 4.1 Sample Data Teks dari Sosial Media X .....	17
Tabel 4.2 Sampel Data dengan Label Emosi dan Depresi .....	19
Tabel 4.3 Hasil Eksperimen 70% <i>Training</i> , 15% Validasi dan 15% <i>Testing</i> .....	23
Tabel 4.4 Hasil Eksperimen 80% <i>Training</i> , 10% Validasi dan 10% <i>Testing</i> .....	24
Tabel 4.5 Hasil Prediksi Model Terhadap Data Baru .....	26



UNIVERSITAS  
MIKROSKIL