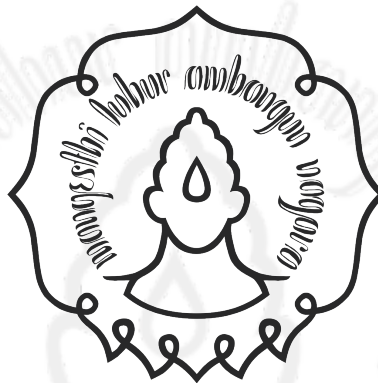


**DETEKSI ANOMALI PADA KLASIFIKASI EMOSI DARI UTAS  
TWITTER BERBASIS PERILAKU PENGGUNA DAN  
FINETUNING INDOBERTWEET**

**SKRIPSI**

Disusun untuk Memenuhi Sebagian Persyaratan Mendapatkan Gelar Strata Satu  
Program Studi Informatika



**Disusun oleh**

M. Fikri Ghazali Mukti

NIM M0518030

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS DATA  
UNIVERSITAS SEBELAS MARET  
SURAKARTA  
2022**

## DAFTAR ISI

<b>HALAMAN PERSETUJUAN</b> .....	<b>ii</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>iii</b>
<b>MOTO</b> .....	<b>iv</b>
<b>PERSEMBAHAN</b> .....	<b>v</b>
<b>ABSTRAK</b> .....	<b>vi</b>
<b>ABSTRACT</b> .....	<b>vii</b>
<b>KATA PENGANTAR</b> .....	<b>viii</b>
<b>DAFTAR ISI</b> .....	<b>ix</b>
<b>DAFTAR TABEL</b> .....	<b>xii</b>
<b>DAFTAR GAMBAR</b> .....	<b>xiv</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xvi</b>
<b>BAB 1 PENDAHULUAN</b> .....	<b>17</b>
1.1. Latar Belakang Masalah.....	17
1.2. Rumusan Masalah .....	21
1.3. Batasan Masalah.....	21
1.4. Tujuan Penelitian.....	21
1.5. Manfaat Penelitian.....	21
<b>BAB 2 TINJAUAN PUSTAKA</b> .....	<b>23</b>
2.1. Dasar Teori.....	23
2.1.1. Deep Learning .....	23
2.1.2. Natural Language Processing .....	33
2.1.3. Text Mining.....	34
2.1.4. Perilaku Pengguna pada Media Sosial .....	36
2.1.5. <i>Mahalanobis Distance</i> .....	37
2.2. Penelitian Terkait.....	39
2.2.1. <i>Attention is All You Need</i> .....	39
2.2.2. <i>BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding</i> .....	40
2.2.3. <i>IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization</i> .....	40
2.2.4. <i>Twitter, time, and emotions</i> .....	41
2.2.5. <i>Detecting Emotions in Social Media: A Constrained Optimization Approach</i> .....	42
2.2.6. <i>Exploiting Social Relations for Sentiment Analysis in Microblogging</i> ....	43
2.2.7. <i>A Survey on Social Media Anomaly Detection</i> .....	43

2.2.8. Mahalanobis Distance and Its Application for Detecting Multivariate Outliers.....	44
<b>BAB 3 METODOLOGI PENELITIAN .....</b>	<b>50</b>
3.1. Pengumpulan dan Pembuatan <i>Dataset</i> .....	50
3.2. Pra-pemrosesan Data .....	51
3.2.1. <i>Duplicate Removal</i> .....	52
3.2.2. <i>Case Folding</i> .....	52
3.2.3. <i>Noise Removal</i> .....	52
3.2.4. Normalisasi.....	52
3.3. Implementasi IndoBERTweet .....	52
3.3.1. Tokenisasi.....	53
3.3.2. <i>Embedding</i> .....	53
3.3.3. <i>Fine-Tuning</i> .....	53
3.3.4. Prediksi Emosi.....	54
3.4. Penyusunan Algoritma Deteksi Anomali .....	54
3.4.1. Penentuan <i>Threshold</i> Pertama .....	55
3.4.2. Penentuan <i>Threshold</i> Kedua.....	55
3.5. Pengujian Algoritma Deteksi Anomali .....	55
3.6. Evaluasi Algoritma Deteksi Anomali.....	56
3.6.1. Evaluasi Algoritma Deteksi Anomali dengan <i>Mahalanobis Distance</i> .....	56
3.6.2. Evaluasi Anomali oleh Ahli Bahasa.....	56
<b>BAB 4 HASIL DAN PEMBAHASAN .....</b>	<b>58</b>
4.1. Pengumpulan dan Pembuatan <i>Dataset</i> .....	58
4.2. Pra-pemrosesan <i>Dataset</i> .....	64
4.2.1. <i>Duplicate Removal</i> .....	65
4.2.2. <i>Case Folding</i> .....	65
4.2.3. <i>Noise Removal</i> .....	66
4.2.4. Normalisasi.....	70
4.3. Implementasi IndoBERTweet .....	72
4.3.1. Tokenisasi.....	72
4.3.2. <i>Embedding</i> .....	73
4.3.3. <i>Fine-Tuning</i> .....	77
4.3.4. Prediksi.....	81
4.4. Penyusunan Algoritma Deteksi Anomali .....	84
4.5. Pengujian Algoritma Deteksi Anomali .....	89
4.5.1. Skenario Percobaan .....	89

4.6. Evaluasi Algoritma Deteksi Anomali.....	91
4.6.1. Evaluasi Algoritma Deteksi Anomali dengan <i>Mahalanobis Distance</i> .....	91
4.6.2. Evaluasi Anomali oleh Ahli Bahasa.....	103
<b>BAB 5 PENUTUP .....</b>	<b>109</b>
5.1. Kesimpulan.....	109
5.2. Saran.....	110
<b>DAFTAR PUSTAKA .....</b>	<b>111</b>
<b>LAMPIRAN.....</b>	<b>119</b>

