DAFTAR ISI

DAFTAR ISI 3
DAFTAR TABEL 5
DAFTAR GAMBAR 6
BAB I PENDAHULUAN 7
1.1 Latar Belakang Masalah 7
1.2 Rumusan Masalah 8
1.3 Batasan Masalah 8
1.4 Tujuan Penelitian 9
1.5 Manfaat Penelitian 9
1.6 Tahapan Penelitian 9
1.7 Sistematika Penulisan 11
BAB II TINJAUAN LITERATUR 12
2.1 Ulasan Palsu 12
2.2 *Natural Language Processing* 12
2.3 Generasi Teks Sintetis dan Teknik Sampling 13
2.3.1 Model Bahasa untuk Generasi Teks (GPT-2) 13
2.3.2 *Top-k Sampling* 15
2.3.3 *Top-p Sampling* 17
2.4 *Preprocessing Text* 20
2.5 Tokenisasi dan Representasi Teks pada RoBERTa 20
2.6 Model RoBERTa 21
2.7 Evaluasi Performa dan *Robustness* 25
2.7.1 Evaluasi Performa Dengan *Confusion Matrix* 26
2.7.2 Evaluasi *Robustness* Dengan *Cross-Dataset* 26
2.8 Penelitian Sebelumnya 27
BAB III METODOLOGI PENELITIAN 31
3.1 *Data Collection & Generation* 32
3.1.1 *Data Collection* 32
3.1.2 Generasi Teks Ulasan Palsu 33
3.2 *Data Preprocessing* 40
3.3 *Data Splitting* 41
3.4 Tokenisasi dan Representasi Teks 42
3.5 Pemodelan dan Fine-Tuning RoBERTa 44
3.5.1 Arsitektur Dasar RoBERTa 44
3.5.2 Mekanisme *Self-Attention* dan *Multi-Head Attention* 44
3.5.3 *Layer Transformer* dan *Feed-Forward Network* 46
3.5.4 *Fine-Tuning* untuk Deteksi Ulasan Palsu 49
3.6 Skenario Pengujian dan Evaluasi 50
3.6.1 Rencana dan Lingkungan Pengujian 50
3.6.2 *Intra-Dataset Evaluation* 51

3.6.3 *Cross-Dataset Evaluation* 52
BAB IV HASIL PENGUJIAN DAN PEMBAHASAN 54
4.1 Implementasi 54
4.1.1 Pengumpulan Data 54
4.1.2 Implementasi Generasi Teks Sintetis 54
4.1.3 Implementasi *Preprocessing* dan *Label Encoding* 58
4.1.4 Implementasi *Data Splitting* 59
4.1.5 Implementasi Tokenisasi RoBERTa 60
4.1.6 Implementasi *Fine-Tuning* Model RoBERTa 61
4.1.7 Implementasi Evaluasi Intra-Dataset 62
4.1.8 Implementasi Evaluasi Cross-Dataset 63
4.1.9 Implementasi Aplikasi 64
4.2 Hasil 65
4.2.1 Hasil Pengumpulan Data 66
4.2.2 Hasil Generasi Teks Sintetis 66
4.2.3 Hasil Preprocessing dan Label Encoding 68
4.2.4 Hasil *Data Splitting* 70
4.2.5 Hasil Tokenisasi RoBERTa 71
4.2.6 Hasil *Fine-Tuning* Model RoBERTa 72
4.2.7 Hasil Evaluasi *Intra-dataset* 74
4.2.8 Hasil Evaluasi *Cross-Dataset* 77
4.2.9 Hasil Pengujian Aplikasi 81
4.3 Pembahasan 87
4.3.1 Analisis Pengaruh Teknik Sampling terhadap Performa *Intra-Dataset* 88
4.3.2 Analisis Pengaruh Teknik Sampling terhadap Pengujian Cross-Dataset 88
4.3.3 Analisis Kesalahan Prediksi 89
4.3.4 Pembahasan terhadap Rumusan Masalah 90
4.3.5 Keterbatasan Penelitian 91
BAB V PENUTUP 93
5.1 Kesimpulan 93
5.2 Saran 93