

DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan dan Manfaat	4
1.4 Ruang Lingkup.....	4
1.5 Sistematika Penulisan	5
BAB II LANDASAN TEORI.....	7
2.1 <i>State of the Art</i>	7
2.2 <i>Automatic Short Answer Grading</i>	11
2.3 <i>Regular Expressions</i>	11
2.4 Pra-Pemrosesan Data	12
2.4.1 <i>Data Formatting</i>	12
2.4.2 Normalisasi Data.....	12
2.4.3 Pembersihan Data.....	13
2.4.4 <i>Case Folding</i>	13
2.5 <i>K-Means Clustering</i>	13
2.6 <i>Silhouette Score</i>	15
2.7 <i>Fine-Tuning Triplet Loss</i>	16
2.8 <i>Forward dan Backward Pass</i>	17
2.8.1 <i>Feed-Forward</i>	17
2.8.2 <i>Back-Propagation</i>	18
2.9 <i>Sentence Embedding</i>	18

2.10	<i>Mean Pooling</i>	19
2.11	<i>Model Pre-Trained</i>	20
2.11.1	BERT.....	20
2.11.2	IndoSBERT	22
2.11.3	XLM-RoBERTa	24
2.11.4	<i>Paraphrase-Multilingual-Mpnet-V2</i>	24
2.12	<i>Cosine Similarity</i>	25
2.13	<i>Few-Shot Learning</i>	26
2.14	Evaluasi SMAPE.....	27
BAB III METODOLOGI PENELITIAN		28
3.1	Gambaran Umum Penelitian	28
3.2	Pengumpulan Data	30
3.2.1	Data MPI 1.....	30
3.2.2	Data MPI 2.....	32
3.2.3	Data Siscer A1	33
3.2.4	Data Siscer B1	34
3.3	Pra-Pemrosesan Data	36
3.3.1	<i>Data Formatting</i>	37
3.3.2	Normalisasi Data	39
3.3.3	Pembersihan Data	40
3.3.4	<i>Case Folding</i>	41
3.4	<i>Clustering</i>	42
3.4.1	<i>Load Model SBERT</i>	43
3.4.2	<i>Encode Jawaban Mahasiswa</i>	43
3.4.3	<i>K-Means Clustering</i>	50
3.4.4	Menghitung <i>Silhouette Score</i>	51
3.4.5	Menghitung <i>Cosine Similarity</i> Antara Referensi dan <i>Centroid Cluster</i>	53
3.5	<i>Fine-Tuning</i>	55
3.6	<i>Sentence Embedding</i>	61
3.7	Perhitungan <i>Cosine Similarity</i>	62
3.8	Pembagian Data	64
3.9	Pembangunan Model Assesmen	65
3.10	Evaluasi.....	67

3.11	<i>Experimental Setup</i>	68
3.11.1	Lingkungan Implementasi.....	68
3.11.2	Skenario Eksperimen	69
BAB IV HASIL DAN PEMBAHASAN.....		70
4.1	Hasil Penelitian	70
4.1.1	Pengumpulan Data.....	70
4.1.1.1	Data MPI 1.....	70
4.1.1.2	Data MPI 2.....	71
4.1.1.3	Data Siscer A1	73
4.1.1.4	Data Siscer B1	74
4.1.2	Pra-Pemrosesan Data.....	75
4.1.2.1	Data MPI 1.....	75
4.1.2.2	Data MPI 2.....	82
4.1.2.3	Data Siscer A1	93
4.1.2.4	Data Siscer B1	96
4.1.3	<i>Clustering</i>	102
4.1.3.1	Data MPI 1.....	102
4.1.3.2	Data MPI 2.....	107
4.1.3.3	Data Siscer A1	115
4.1.3.4	Data Siscer B1	119
4.1.4	<i>Fine-Tuning</i>	123
4.1.4.1	Data MPI 1.....	123
4.1.4.2	Data MPI 2.....	126
4.1.4.3	Data Siscer A1	129
4.1.4.4	Data Siscer B1	130
4.1.5	<i>Sentence Embedding</i>	131
4.1.5.1	Data MPI 1.....	132
4.1.5.2	Data MPI 2.....	133
4.1.5.3	Data Siscer A1	137
4.1.5.4	Data Siscer B1	139
4.1.6	Perhitungan <i>Cosine Similarity</i>	141
4.1.6.1	Data MPI 1.....	141
4.1.6.2	Data MPI 2.....	143

4.1.6.3 Data Siscer A1	147
4.1.6.4 Data Siscer B1	148
4.1.7 Pembagian Data	151
4.1.7.1 Data MPI 1	151
4.1.7.2 Data MPI 2	154
4.1.7.3 Data Siscer A1	160
4.1.7.4 Data Siscer B1	162
4.1.8 Model Assesmen	165
4.1.8.1 Data MPI 1	166
4.1.8.2 Data MPI 2	166
4.1.8.3 Data Siscer A1	167
4.1.8.4 Data Siscer B1	168
4.1.9 Evaluasi SMAPE	169
4.1.9.1 Data MPI 1	169
4.1.9.2 Data MPI 2	171
4.1.9.3 Data Siscer A1	174
4.1.9.4 Data Siscer B1	175
4.2 Hasil Pelatihan dan Analisis	176
4.3 Hasil Pengujian dan Analisis	177
4.3.1 Skenario 1: Pengaruh <i>Fine-Tuning</i> pada Setiap <i>Dataset</i>	179
4.3.2 Skenario 2: Nilai α dan β Optimal pada Setiap <i>Dataset</i>	181
4.3.3 Skenario 3: Model Terbaik pada Setiap <i>Dataset</i> Berdasarkan Skenario 1 dan 2	183
BAB V KESIMPULAN DAN SARAN	190
5.1 Kesimpulan	190
5.2 Saran	190
DAFTAR PUSTAKA	191
LAMPIRAN 1. Contoh Perhitungan Representasi Vektor Model IndoSBERT	197
LAMPIRAN 2. Contoh Perhitungan Representasi Vektor Model <i>Paraphrase-</i> <i>Multilingual-Mpnet-V2</i>	199
LAMPIRAN 3. Contoh Perhitungan <i>Silhouette Score</i>	201