

DAFTAR ISI

HALAMAN PENGESAHAN	iii
PERNYATAAN BEBAS PLAGIASI	iv
PRAKATA	v
DAFTAR ISI	vii
DAFTAR GAMBAR	x
DAFTAR TABEL	xiii
INTISARI	xiv
ABSTRACT	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan	3
1.5 Manfaat	3
1.6 Keaslian Penelitian	4
1.7 Metodologi Penelitian	4
BAB II TINJAUAN PUSTAKA	6
BAB III LANDASAN TEORI	12
3.1 Automated Essay Scoring	12
3.2 Pembelajaran Unsupervised	12
3.2.1 HDBSCAN	13
3.3 Jaringan Syaraf Tiruan	16
3.3.1 Self organizing map	19
3.4 Word Embedding	23
3.4.1 Word2Vec	23
3.4.2 Doc2Vec	26
3.5 Readability Index	28
3.6 Automatic Text Summarization	31
3.7 Peringkasan Ekstraktif	31
3.7.1 Peringkasan TF-IDF	32

3.7.2	Peringkasan TextRank	35
3.8	Ukuran Evaluasi	37
3.8.1	ROUGE	37
3.8.2	Quantization error	38
3.8.3	Silhouette score	39
BAB IV	ANALISIS DAN PERANCANGAN SISTEM	42
4.1	Analisis Permasalahan	42
4.1.1	Definisi permasalahan	42
4.1.2	Analisis kebutuhan	42
4.1.3	Pendekatan penyelesaian masalah	44
4.2	Analisis Data	47
4.3	Pra-pemrosesan Data	49
4.3.1	Casefolding	50
4.3.2	Filtering – menghapus stopword	50
4.3.3	Tokenisasi – lemmatisasi	51
4.4	Representasi Data	51
4.4.1	Word embedding	52
4.4.2	Readability index	53
4.4.3	Handcrafted features	54
4.5	Perancangan Arsitektur	55
4.5.1	Peringkasan TF-IDF	56
4.5.2	Peringkasan TextRank	58
4.5.3	Model klastering	61
4.6	Rancangan Evaluasi	66
BAB V	IMPLEMENTASI SISTEM	70
5.1	Spesifikasi Hardware dan Software	70
5.2	Pengambilan Data	70
5.3	Implementasi Pra-Pemrosesan Data	70
5.3.1	Implementasi pembersihan data	71
5.3.2	Implementasi lemmatisasi	72
5.4	Implementasi Representasi Data	73
5.4.1	Implementasi word embedding	74
5.4.2	Implementasi readability index	75
5.4.3	Implementasi handcrafted features	76

5.4.4	Implementasi pembentukan data	78
5.5	Implementasi Pemodelan TF-IDF	79
5.5.1	Implementasi pos-tagging	79
5.5.2	Implementasi menghitung frekuensi kata	80
5.5.3	Implementasi perhitungan TF-IDF	81
5.5.4	Implementasi nilai kepentingan kalimat	82
5.5.5	Implementasi pembentukan ringkasan TF-IDF	83
5.6	Implementasi Pemodelan TextRank	86
5.6.1	Implementasi sentence embedding	86
5.6.2	Implementasi matriks persamaan	87
5.6.3	Implementasi pembentukan ringkasan TextRank	88
5.7	Implementasi Pemodelan Klastering	90
5.8	Implementasi Evaluasi Pemodelan	92
BAB VI	HASIL DAN PEMBAHASAN	95
6.1	Deskripsi Umum Pengujian	95
6.2	Hyperparameter Tuning	98
6.2.1	Hyperparameter tuning model ATS	98
6.2.2	Hyperparameter tuning model SOM	106
6.2.3	Hyperparameter tuning model HDBSCAN	110
6.3	Evaluasi Klastering dengan ATS	112
6.3.1	Analisis kluster hasil pemodelan	113
6.4	Analisis hasil klastering dengan data ringkasan TextRank	116
6.4.1	Analisis I: Kriteria Cohesion of Ideas	119
6.4.2	Analisis II: Kriteria Completeness	121
BAB VII	KESIMPULAN DAN SARAN	125
7.1	Kesimpulan	125
7.2	Saran	125
	DAFTAR PUSTAKA	127
	LAMPIRAN	132