

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN TIM PEMBIMBING	ii
HALAMAN PENGESAHAN TIM PENGUJI.....	iv
PERNYATAAN BEBAS PLAGIASI	vi
PRAKATA.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR ARTI LAMBANG DAN SINGKATAN	xiv
ABSTRAK.....	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Penelitian.....	4
1.4 Tujuan Penelitian.....	4
1.5 Keaslian dan Kontribusi Penelitian.....	4
1.6 Manfaat Penelitian.....	10
1.7 Sistematika Penulisan	10
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	11
2.1 Tinjauan Pustaka.....	11
2.1.1 Penilaian Otomatis Menggunakan <i>Machine Learning</i>	11
2.1.2 Penilaian Otomatis Menggunakan <i>Deep Learning</i>	13
2.1.3 Penilaian Otomatis Menggunakan <i>Large Language Model</i>	14
2.1.4 Penilaian Otomatis Menggunakan Multimodal LLMs	17
2.2 Landasan Teori	18
2.2.1 Esai.....	18
2.2.1.1 Esai Monomodal.....	18
2.2.1.2 Esai Multimodal.....	19
2.2.2 Esai Panjang: <i>Systematic Literature Review</i>	19
2.2.3 Penilaian Esai Berbasis Rubrik.....	20
2.2.4 <i>Human Rater</i>	21
2.2.5 Penilaian Esai Otomatis.....	21
2.2.6 <i>Large Language Model</i>	22
2.2.6.1 <i>Multimodal Large Language Model</i>	23
2.2.6.2 <i>Generative Pre-trained Transformer (GPT)</i>	24

2.2.7	<i>Prompt Engineering</i>	27
2.2.10.1	<i>Zero-shot prompting</i>	28
2.2.10.2	<i>Few-shot prompting</i>	28
2.2.10.3	<i>Role-based prompting</i>	28
2.2.10.4	<i>Rubric-guided prompting</i>	29
2.2.10.5	<i>Chain-of thought prompting</i>	29
2.2.10.6	<i>Chain-of-evidence prompting</i>	29
2.2.8	<i>Optical Character Recognition (OCR)</i>	30
2.2.9	Metrik Evaluasi Kinerja Penilaian Esai Otomatis	30
2.2.9.1	Quadratic Weighted Kappa.....	30
2.2.9.2	Pearson Correlation Coefficient.....	31
2.2.9.3	Mean Absolute Error	32
2.2.9.4	Root Mean Squared Error.....	32
2.2.9.5	Akurasi.....	33
2.3	Pertanyaan Penelitian.....	33
BAB III METODE PENELITIAN		34
3.1	Alat dan Bahan Penelitian.....	34
3.1.1	Alat.....	34
3.1.1.1	Perangkat Keras.....	34
3.1.1.2	Perangkat Lunak	34
3.1.2	Bahan	36
3.1.2.1	Data Paper SLR	36
3.1.2.2	Dataset Daftar Nilai dengan Rubrik Lengkap.....	37
3.1.2.1	Rubrik Penilaian Lengkap	37
3.2	Tahapan Penelitian.....	39
3.2.1	Identifikasi Masalah dan Analisis Kebutuhan	40
3.2.2	Peninjauan Konsep dan Teori.....	40
3.2.3	Penentuan Variabel Penelitian.....	40
3.2.4	Pengumpulan dan Pengolahan Data Penelitian	41
3.2.5	Perancangan Konseptual Sistem Penilaian Otomatis	41
3.2.6	Implementasi Sistem dan Analisis Performa.....	42
3.2.7	Evaluasi Hasil	43
3.2.8	Penarikan Kesimpulan dan Saran	43
3.3	Perancangan Sistem.....	43
3.3.1	Gambaran Umum Alur Sistem	43
3.3.2	Preprocessing.....	45
3.3.2.1	<i>Data Cleaning</i>	45
3.3.2.2	<i>Data Cropping</i>	45
3.3.3	Identifikasi Format Dokumen dan Proses OCR	46
3.3.4	Konfigurasi Model Multimodal LLM dan Eksplorasi <i>Prompt</i>	47
3.3.4.1.	<i>Zero-shot prompting</i> dengan <i>role</i> dan rubrik.....	49
3.3.4.2.	<i>Chain-of-thought prompting</i> dengan <i>role</i> dan rubrik	50
3.3.4.3.	<i>Multi-trait prompting</i> dengan <i>role</i> , <i>chain-of-evidence</i> , rubrik, dan <i>guidance</i>	50
3.3.4.4.	<i>Multi-chain prompting</i> dengan <i>role</i> , <i>chain-of-evidence</i> , rubrik, <i>guidance</i> , dan <i>multi-chain reasoning</i>	51
3.3.5	Penilaian seluruh Paper menggunakan <i>Prompt</i> Terbaik yang Terpilih..	53

3.4	Pendekatan Analisis dan Evaluasi Hasil.....	53
BAB IV HASIL DAN PEMBAHASAN		57
4.1	Hasil Penelitian.....	57
4.1.1	Hasil Pengolahan Data Penelitian.....	57
4.1.2	Hasil Perbandingan Teknik <i>Prompt</i> Terhadap Pengujian 30 Sampel.....	58
4.1.2.1.	Analisis Hasil <i>Zero-shot prompting</i> dengan <i>Role</i> dan Rubrik.....	58
4.1.2.2.	Analisis Hasil <i>Chain-of-Thought Prompting</i>	59
4.1.2.3.	Analisis Hasil <i>Multi-trait Prompting</i> Menggunakan GPT-5.0.....	59
4.1.2.4.	Analisis Hasil <i>Multi-trait Prompting</i> Menggunakan GPT-4o.....	59
4.1.2.5.	Analisis Hasil <i>Multi-chain Prompting</i>	59
4.1.3	Hasil Penilaian Esai Otomatis Terhadap Seluruh Data	60
4.1.3.1.	Hasil Penilaian Paper Menggunakan Rubrik Lengkap.....	60
4.1.3.2.	Hasil Penilaian Paper Menggunakan Rubrik yang Berfokus pada Elemen Visual.....	66
4.1.4	Fenomena Interpretasi QWK dan Akurasi untuk Distribusi Tidak Seimbang ⁶⁹	
4.1.5	Umpan Balik yang Dihasilkan oleh multimodal LLM	71
4.1.4.1.	Analisis Koherensi Skor dan Umpan Balik	71
4.1.4.2.	Analisis Kesesuaian Umpan Balik dengan Indikator Rubrik	72
4.1.4.3.	Analisis Konsistensi Umpan Balik terhadap Logika Penilaian	72
4.2	Pembahasan Hasil.....	73
4.3	Penggunaan ChatGPT.....	77
BAB V KESIMPULAN DAN SARAN		80
5.1	Kesimpulan.....	80
5.2	Saran	80
DAFTAR PUSTAKA		82
LAMPIRAN.....		L-1
Lampiran 1. <i>Zero-shot Prompt : role + rubric</i>		L-1
Lampiran 2. <i>Chain-of-thought Prompt (zero-shot): role + chain-of-thought + rubric</i>		L-2
Lampiran 3. <i>Multi-trait Prompt (zero-shot): role + chain-of-evidence + rubric + guidance</i>		L-3
Lampiran 4. <i>Multi-chain (zero-shot): role + chain-of-evidence + rubric + guidance + multi-chain-reasoning</i>		L-7
Lampiran 5. <i>Prompt</i> Berfokus pada Penilaian Tabel.....		L-11
Lampiran 6. <i>Prompt</i> Berfokus pada Penilaian Gambar.....		L-13