



SKRIPSI

PENGEMBANGAN MODEL GLOVE-LSTM DAN ALGORITMA PENILAIAN JAWABAN SISWA MENGGUNAKAN ROUGE SCORE, TF-IDF, DAN COSINE SIMILARITY

Ahmad Sofian Aris Saputra
NPM 21081010211

DOSEN PEMBIMBING

Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU
Dr. Ir. Mohammad Idhom, S.P, S.Kom.,MT

**KEMENTERIAN PENDIDIKAN,
KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**



SKRIPSI

PENGEMBANGAN MODEL GLOVE-LSTM DAN ALGORITMA PENILAIAN JAWABAN SISWA MENGGUNAKAN ROUGE SCORE, TF-IDF, DAN COSINE SIMILARITY

Ahmad Sofian Aris Saputra
NPM 21081010211

DOSEN PEMBIMBING

Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU
Dr. Ir. Mohammad Idhom, S.P, S.Kom.,MT

**KEMENTERIAN PENDIDIKAN KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**

Halaman ini sengaja dikosongkan

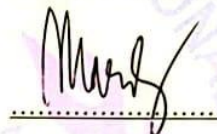
LEMBAR PENGESAHAN

PENGEMBANGAN MODEL GLOVE-LSTM DAN ALGORITMA PENILAIAN JAWABAN SISWA MENGGUNAKAN ROUGE SCORE, TF-IDF, DAN COSINE SIMILARITY

Oleh :
AHMAD SOFIAN ARIS SAPUTRA
NPM. 21081010211

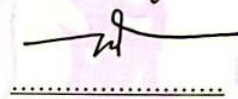
Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 05 Juni 2025

Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU
NIP. 19700619 202121 1 009



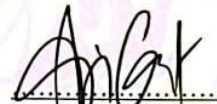
(Pembimbing I)

Dr. Ir. Mohammad Idhom, S.P, S.Kom., MT.
NIP. 19830310 202121 1 006



(Pembimbing II)

Chrystia Aji Putra, S.Kom, MT
NIP. 19861008 202121 1 001



(Ketua Penguji)

Eka Prakarsa Mandyartha, ST, M.Kom
NIP. 19880525 2018031 001



(Penguji I)

Mengetahui,
Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

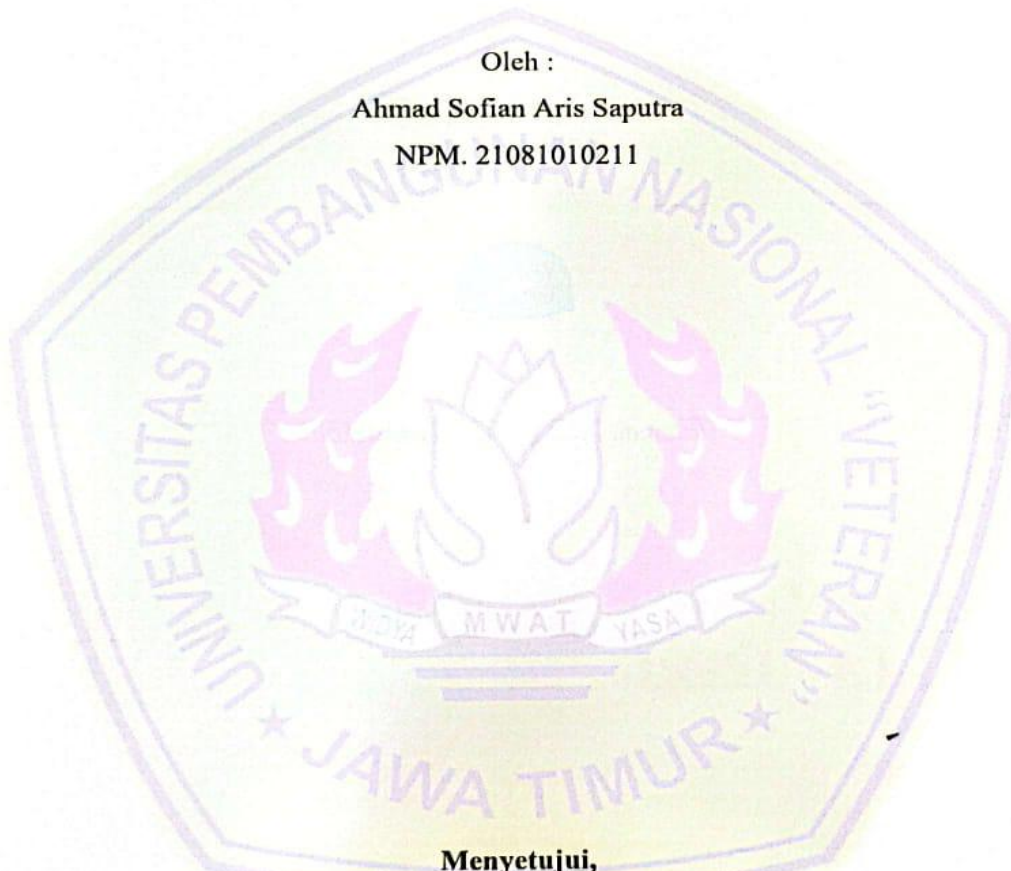
LEMBAR PERSETUJUAN

PENGEMBANGAN MODEL GLOVE-LSTM DAN ALGORITMA PENILAIAN JAWABAN SISWA MENGGUNAKAN ROUGE SCORE, TF-IDF, DAN COSINE SIMILARITY

Oleh :

Ahmad Sofian Aris Saputra

NPM. 21081010211



**Menyetujui,
Koordinator Program Studi Informatika
Fakultas Ilmu Komputer**

Fetty Tri Anggraeny, S.Kom. M. Kom
NIP. 19820211 2021212 005

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan dibawah ini :

Nama : Ahmad Sofian Aris Saputra
NPM : 21081010211
Program : Sarjana (S1)
Program Studi : Informatika
Fakultas : Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Tugas Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila di kemudian hari ditemukan indikasi plagiat pada Skripsi/Tesis/Desertasi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 05 Juni 2025
Yang Membuat Pernyataan



AHMAD SOFIAN ARIS SAPUTRA
21081010211

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa /NPM : Ahmad Sofian Aris Saputra / 21081010211
Judul Skripsi : Pengembangan Model GloVe-LSTM Dan
Algoritma Penilaian Jawaban Siswa Menggunakan
Rouge Score, TF-IDF, dan *Cosine Similarity*

Dosen Pembimbing : 1. Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU
2. Dr. Ir. Mohammad Idhom, S.P, S.Kom.,MT

Penilaian jawaban siswa secara manual menghadapi berbagai tantangan seperti subjektivitas, inkonsistensi, dan waktu koreksi yang lama, terutama dalam menghadapi variasi linguistik jawaban siswa. Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model penilaian jawaban siswa otomatis berbahasa Indonesia yang lebih objektif, akurat, dan adaptif. Model utama yang diusulkan mengintegrasikan representasi kata GloVe dengan *Long Short-Term Memory* (LSTM) serta didukung oleh algoritma penilaian teks ROUGE Score, TF-IDF, dan *Cosine Similarity*. Metodologi penelitian mencakup perancangan arsitektur model, pengumpulan dataset mandiri berisi 3420 item jawaban siswa yang kemudian diproses menjadi 3152 sampel untuk pelatihan dan evaluasi, pra-pemrosesan data, pelatihan model menggunakan *TensorFlow* dan Keras, serta pengujian melalui berbagai skenario. Selain itu, dilakukan studi eksploratif terhadap Model Pembanding dengan implementasi LSTM manual menggunakan NumPy. Hasil evaluasi pada data uji menunjukkan Model Utama mencapai Mean Absolute Error (MAE) sebesar 0.0761, koefisien korelasi *Pearson* 0.8429, dan *Quadratic Weighted Kappa* (QWK) 0.7332, mengindikasikan potensi dalam memberikan peringkat jawaban yang relatif sesuai dengan penilaian manual. Meskipun demikian, analisis skenario mengungkapkan bahwa Model Utama cenderung memberikan skor lebih tinggi dari seharusnya (*overestimate*) untuk jawaban berkualitas rendah, tidak relevan, atau salah, serta menunjukkan inkonsistensi dalam menangani variasi panjang jawaban, penggunaan sinonim, kesalahan ejaan, dan jawaban tanpa kata kunci eksplisit. Model Pembanding menunjukkan kemampuan belajar dasar namun dengan kinerja evaluasi yang masih terbatas. Penelitian ini menyimpulkan bahwa model hibrida yang dikembangkan memiliki potensi, namun memerlukan penyempurnaan signifikan dalam akurasi skor *absolut* dan kemampuan diskriminasi, khususnya untuk jawaban berkualitas rendah dan tidak relevan dalam konteks Bahasa Indonesia.

Kata kunci : Penilaian Otomatis, GloVe, LSTM, ROUGE Score, TF-IDF, *Cosine Similarity*, Pemrosesan Bahasa Alami, Pembelajaran Mendalam, Skoring Esai Otomatis, Bahasa Indonesia.

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Ahmad Sofian Aris Saputra / 21081010211
Thesis Title : *Development Of A GloVe-LSTM Based Hybrid Model For Automated Scoring Of Student Answers Using Rouge, TF-IDF, And Cosine Similarity*
Advisor : 1. Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU
2. Dr. Ir. Mohammad Idhom, S.P, S.Kom.,MT

Manual assessment of student answers faces various challenges such as subjectivity, inconsistency, and prolonged grading time, particularly when dealing with the linguistic variations in student responses. This research aims to develop and evaluate an automated scoring model for Indonesian student answers that is more objective, accurate, and adaptive. The primary proposed model integrates GloVe word representations with Long Short-Term Memory (LSTM), supported by text assessment algorithms such as ROUGE Score, TF-IDF, and Cosine Similarity. The research methodology encompasses model architecture design, the collection of a proprietary dataset comprising 3420 student answer items (subsequently processed into 3152 samples for training and evaluation), data pre-processing, model training using TensorFlow and Keras, and testing across various scenarios. Furthermore, an exploratory study was conducted on a Comparative Model featuring a manual LSTM implementation using NumPy. Evaluation results on the test data indicate that the Main Model achieved a Mean Absolute Error (MAE) of 0.0761, a Pearson correlation coefficient of 0.8429, and a Quadratic Weighted Kappa (QWK) of 0.7332, suggesting its potential in ranking answers relatively consistently with manual assessments. Nevertheless, scenario analysis revealed that the Main Model tends to overestimate Scores for low-quality, irrelevant, or incorrect answers. It also demonstrated inconsistencies in handling variations in answer length, synonym usage, spelling errors, and responses lacking explicit keywords. The Comparative Model demonstrated basic learning capabilities but exhibited limited evaluation performance. This research concludes that the developed hybrid model shows potential but requires significant refinement in absolute Score accuracy and discriminatory power, particularly for low-quality and irrelevant answers within the Indonesian language context.

Keywords: *Automated Scoring, GloVe, LSTM, ROUGE Score, TF-IDF, Cosine Similarity, Natural Language Processing, Deep Learning, Automated Essay Scoring, Indonesian Language.*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Pertama, saya panjatkan puji syukur kehadiran Allah SWT, atas rahmat dan karunia-Nya, sehingga skripsi dengan judul **“Pengembangan Model GloVe-LSTM Dan Algoritma Penilaian Jawaban Siswa Menggunakan Rouge Score, TF-IDF, dan Cosine Similarity”** dapat diselesaikan dengan baik dan tepat pada waktunya. Penulisan skripsi ini merupakan salah satu persyaratan akademis guna memperoleh gelar Ilmu Komputer Program S-1 dari Prodi Informatika Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Penyusunan skripsi ini tidak akan dapat terwujud sebagaimana mestinya tanpa bimbingan, bantuan, serta fasilitas yang diberikan oleh berbagai pihak. Oleh karena itu, pada kesempatan ini saya menyampaikan rasa terima kasih dan penghormatan yang tulus kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT., selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom. selaku Koordinator Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. bapak Dr. Ir. I Gede Susrama Mas Diyasa, S.T., M.T., IPU, selaku Dosen Pembimbing I, yang telah membimbing saya dengan penuh kesabaran, dedikasi, serta memberikan arahan yang sangat berarti hingga skripsi ini dapat diselesaikan tepat pada waktunya.
5. Bapak Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T., selaku Dosen Pembimbing II, yang telah berkenan meluangkan waktu, serta memberikan pemikiran, wawasan, dan masukan yang sangat membantu dalam pengembangan sistem serta penyempurnaan isi skripsi ini.
6. Bapak Chrystia Aji Putra, S.Kom., M.T., selaku Dosen Penguji I, yang telah memberikan kritik, saran, dan masukan yang membangun dalam proses penulisan serta penyempurnaan skripsi ini.
7. Bapak Eka Prakarsa Mandyartha, S.T., M.Kom., selaku Dosen Penguji II, yang turut memberikan arahan, evaluasi, serta pandangan kritis yang sangat bermanfaat dalam penyusunan dan pengembangan isi skripsi ini.

8. Bapak Andreas Nugroho Sihananto, S.Kom., M.Kom., selaku Dosen pengurus skripsi, yang telah membantu dan memfasilitasi proses administrasi serta kelancaran tahapan penyusunan skripsi ini dengan penuh tanggung jawab.
9. Seluruh dosen Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur, yang telah memberikan ilmu, pengalaman, serta motivasi selama masa studi saya. Setiap pembelajaran yang diberikan menjadi bekal berharga dalam penyusunan skripsi ini maupun dalam perjalanan saya ke depannya.
10. Kepada kedua orang tua dan seluruh keluarga besar penulis yang selalu memberikan doa, semangat, dan dukungan penuh tanpa henti, baik secara moral maupun emosional, selama proses pengerjaan skripsi ini. Tanpa mereka, perjalanan ini tentu akan jauh lebih berat.
11. Kepada pacarku yang menjadi penyemangat, motivasi, dan memberikan dukungan yang amat besar. Terima kasih telah bersedia menjadi tempat curhat dan keluh kesah selama pengerjaan skripsi ini, sehingga skripsi ini selesai tepat waktu. *I love you.*
12. Terima kasih yang sebesar-besarnya kepada teman-teman Banteng UwU, serta seluruh teman-teman Angkatan 2021, yang telah menjadi sumber semangat, informasi, motivasi, dan kadang juga tempat sambat. Dukungan kalian sangat berarti hingga akhirnya skripsi ini dapat diselesaikan dengan lancar.
13. Terinspirasi dari perjalanan para kru Topi Jerami dalam anime One Piece, saya menyadari bahwa perjalanan panjang penuh tantangan ini tidak bisa dilalui sendirian. Layaknya Luffy yang dikelilingi oleh nakama setia, saya pun bersyukur memiliki orang-orang hebat di sekitar saya yang selalu mendukung, mengingatkan tujuan, dan tidak pernah menyerah meskipun badai datang silih berganti. Semoga semangat pantang menyerah, keberanian untuk bermimpi besar, dan tekad untuk terus maju seperti yang ditunjukkan dalam One Piece dapat terus menjadi inspirasi dalam kehidupan saya ke depannya.

Saya menyadari bahwa skripsi ini masih memiliki kekurangan yang perlu disempurnakan. Oleh karena itu, saya memohon maaf yang sebesar-besarnya dan sangat terbuka terhadap segala kritikan serta masukan yang konstruktif guna meningkatkan kualitas skripsi ini. Semoga skripsi ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan di masa yang akan datang.

Surabaya, 05 Juni 2025
Penulis,

Ahmad Sofian Aris Saputra

Halaman ini sengaja dikosongkan

DAFTAR ISI

LEMBAR PENGESAHAN	i
LEMBAR PERSETUJUAN.....	iii
SURAT PERNYATAAN BEBAS PLAGIASI.....	v
ABSTRAK	vii
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xv
DAFTAR GAMBAR.....	xviii
DAFTAR TABEL	xix
BAB I.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Fokus Penelitian	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
1.6 Batasan Penelitian	7
BAB II	9
2.1 Penelitian Terdahulu.....	9
2.2 Dasar Teori	13
2.2.1 Machine Learning	13
2.2.2 Deep Learning.....	14
2.2.3 NLP	15
2.2.4 GloVe	15
2.2.5 LSTM	17
2.2.6 ROUGE	20
2.2.7 TF-IDF	21

2.2.8	<i>Cosine Similarity</i>	23
2.3	Alasan Pemilihan Teori	24
BAB III		27
3.1	Perancangan Model	28
3.2.1	Arsitektur Model	29
3.2.2	Penerapan Model GloVe-LSTM	30
3.2.3	Penerapan Algoritma Evaluasi	32
3.2	Pengumpulan Dataset	35
3.3	<i>Pre-processing</i> Data	37
3.4	<i>Training</i> Model	38
3.5	Uji Coba Model	40
3.6	Evaluasi dan Validasi Model	42
3.7	Skenario Pengujian	45
3.8	Kesimpulan dan Dokumentasi	55
BAB IV		57
4.1	Deskripsi Lingkungan dan Skenario Eksperimen	57
4.1.1	Lingkungan Eksperimen	57
4.1.2	Dataset Eksperimen.....	58
4.1.3	Konfigurasi Parameter Model	59
4.1.4	Metrik Evaluasi	62
4.1.5	Skenario Pengujian yang Ditetapkan	64
4.2	Hasil Eksperimen Model Utama	65
4.2.1	Hasil Pelatihan Model Utama	65
4.2.2	Hasil Evaluasi Kinerja Model Utama pada Data Uji	68
4.2.3	Hasil Pengujian Model Utama Berdasarkan Skenario	70
4.2.4	Analisis dan Pembahasan Hasil Model Utama	84
4.3	Observasi Pengembangan Model Pembanding	85

4.3.1	Observasi Pelatihan <i>Word embedding</i> (GloVe)	86
4.3.2	Observasi Pengembangan dan Pelatihan Awal LSTM Manual	88
4.3.3	Refleksi Awal dan Potensi Lanjutan	92
4.4	Pembahasan Umum Hasil Penelitian	94
4.4.1	Interpretasi Keseluruhan Hasil	94
4.4.2	Keterbatasan Penelitian	97
4.4.3	Implikasi Hasil Penelitian	99
BAB V		103
5.1	Kesimpulan	103
5.2	Saran.....	104
DAFTAR PUSTAKA		107
LAMPIRAN		110
Lampiran I – Kode Model		110
Lampiran II – Dataset yang Digunakan (Format CSV)		117
Lampiran III - Dataset yang Digunakan (Format JSON)		123
Lampiran IV – Gambar Hasil Pelatihan Model		128
Lampiran V – Soal Skenario Pengujian		130

DAFTAR GAMBAR

Gambar 2. 1 GloVe Sematik Antar Kata	16
Gambar 2. 2 Relasi Sematik dan Sintaksis	17
Gambar 3. 1 <i>Flowchart</i> Pengerjaan.....	28
Gambar 3. 2 Arsitektur Model	30
Gambar 3. 3 Penerapan GloVe	31
Gambar 3. 4 Penerapan LSTM	32
Gambar 3. 5 Penerapan Rouge <i>Score</i>	33
Gambar 3. 6 Penerapan TF-IDF.....	33
Gambar 3. 7 Penerapan Cosine Similarity	34
Gambar 3. 8 Penerapan Gabungan Model dan Algoritma	35
Gambar 4. 1 <i>Mean Squared Error</i>	66
Gambar 4. 2 <i>Mean Absolute Error</i>	67
Gambar 4. 3 Grafik Ringkasan Perbandingan Skor	70
Gambar 4. 4 Grafik <i>Average Loss</i> (MSE) Manual	90

DAFTAR TABEL

Tabel 2. 1 <i>Reseach Gap</i>	11
Tabel 2. 2 Arsitektur LSTM.....	17
Tabel 3. 1 Contoh Kasus Evaluasi	44
Tabel 3. 2 Pengujian Fungsionalitas Dasar	46
Tabel 3. 3 Pengujian Variasi Panjang Jawaban	48
Tabel 3. 4 Pengujian Penggunaan Sinonim	49
Tabel 3. 5 Pengujian Kesalahan Tata Bahasa	51
Tabel 3. 6 Pengujian Jawaban Tanpa Kata Kunci	52
Tabel 3. 7 Pengujian Jawaban Tidak Relevan	54
Tabel 4. 1 Pelatihan Model Utama	60
Tabel 4. 2 Model Pembanding	61
Tabel 4. 3 Metrik Evaluasi.....	68
Tabel 4. 4 Kinerja Model Utama	68
Tabel 4. 5 Skenario 1	71
Tabel 4. 6 Skenario 2	73
Tabel 4. 7 Skenario 3	75
Tabel 4. 8 Skenario 4	77
Tabel 4. 9 Skenario 5	79
Tabel 4. 10 Skenario 6	82
Tabel 4. 11 Hasil <i>Epoch</i> Manual GloVe	86
Tabel 4. 12 Hasil <i>Epoch</i> Manual LSTM.....	89
Tabel 4. 13 Hasil Metrik Evaluasi Manual	90