

AUTOMATED ESSAY SCORING UNTUK PENILAIAN JAWABAN
ESAI BAHASA INDONESIA DENGAN INDOBERT *EMBEDDING*
DAN *FEEDFORWARD NEURAL NETWORK*

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh :

Ramadhania Humaira
NIM : 09021282126064

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2025

HALAMAN PENGESAHAN

SKRIPSI

**AUTOMATED ESSAY SCORING UNTUK PENILAIAN
JAWABAN ESAI BAHASA INDONESIA DENGAN INDOBERT
EMBEDDING DAN FEEDFORWARD NEURAL NETWORK**

Sebagai salah satu syarat untuk penyelesaian studi di

Program Studi S1 Teknik Informatika

Oleh:

RAMADHANIA HUMAIRA

09021282126064

**Pembimbing 1 : Dr. Abdiansah, S.Kom., M.Cs.
NIP. 198410012009121005**

Mengetahui

Ketua Jurusan Teknik Informatika



**Hadipurnawan Satria, Ph.D
198004182020121001**

TANDA LULUS UJIAN KOMPREHENSIF

Pada hari Jumat tanggal 28 Februari 2025 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Ramadhania Humaira
NIM : 09021282126064
Judul : *Automated Essay Scoring* untuk Penilaian Jawaban Esai Bahasa Indonesia dengan *IndoBERT Embedding* dan *Feedforward Neural Network*

dan dinyatakan LULUS

1. Ketua Penguji

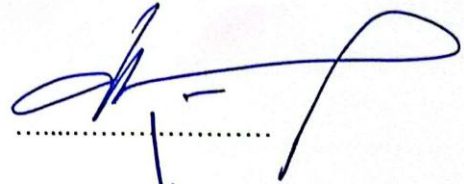
Alvi Svahrini Utami, M.Kom.
NIP. 197812222006042003



.....

2. Penguji

Julian Supardi, M.T., Ph.D.
NIP. 197207102010121001



.....

3. Pembimbing I

Dr. Abdiansah, S.Kom., M.CS.
NIP. 198410012009121005



.....



HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Ramadhania Humaira

NIM : 09021282126064

Program Studi : Teknik Informatika

Judul Skripsi : *Automated Essay Scoring* untuk Penilaian Jawaban Esai Bahasa Indonesia dengan *IndoBERT Embedding* dan *Feedforward Neural Network*

Hasil pengecekan *Software iThenticate/Turnitin*: 2%

Menyatakan bahwa laporan penelitian saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan penelitian ini, maka saya bersedia menerima sanksi akademik Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapa pun.



Palembang, 01 Maret 2025



Ramadhania Humaira
NIM 09021282126064

MOTTO DAN PERSEMBAHAN

“Kembalilah kepada Tuhanmu dengan hati yang ridha dan diridhai-Nya”
(Q.S Al-Fajr: 28)

Kupersembahkan karya tulis ini kepada:

- Allah SWT
- Keluarga
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

ABSTRACT

Improving the quality of education can be supported by a more effective assessment system, one of which is Automated Essay Scoring (AES) for automatic essay evaluation. This study develops an Indonesian-language AES system using IndoBERT Embedding and a Feedforward Neural Network (FNN). The dataset used is the secondary dataset from the UKARA Challenge, developed by the NLP Research Team at Universitas Gadjah Mada, which has a limited number of data and an imbalanced class distribution (labels 1 and 0). Overall, the developed model, after being trained and evaluated on datasets A and B, achieved an F1-score of 0.767. On dataset A, the model trained using the SMOTE technique obtained an F1-score of 0.835 with a batch size of 16, epoch 7, and a learning rate of $1.26e-4$. The best model on dataset B achieved an F1-score of 0.699 with a batch size of 64, epoch 4, and a learning rate of $5.7e-3$. These results indicate that IndoBERT Embedding and FNN provide a reasonably good performance compared to the baseline provided by UKARA for the training set, although challenges remain regarding data imbalance and limited dataset size.

Keywords: AES, IndoBERT, FNN, SMOTE

**AUTOMATED ESSAY SCORING UNTUK PENILAIAN JAWABAN ESAI
BAHASA INDONESIA DENGAN INDOBERT EMBEDDING DAN
FEEDFORWARD NEURAL NETWORK**

Oleh

Ramadhania Humaira

09021282126064

ABSTRAK

Peningkatan kualitas pendidikan dapat didukung dengan sistem penilaian yang lebih baik, salah satunya melalui *Automated Essay Scoring* (AES) untuk penilaian esai secara otomatis. Penelitian ini mengembangkan sistem AES berbahasa Indonesia menggunakan *IndoBERT Embedding* dan *Feedforward Neural Network* (FNN). Dataset yang digunakan adalah dataset sekunder UKARA *Challenge* yang dikembangkan oleh Tim Riset NLP Universitas Gadjah Mada, dengan jumlah data terbatas dan distribusi kelas yang tidak seimbang (label 1 dan 0). Secara keseluruhan, model yang dikembangkan, setelah dilatih dan dievaluasi pada dataset A dan B, mencapai F1-Score sebesar 0,767. Pada dataset A, model yang dilatih menggunakan teknik SMOTE memperoleh F1-Score sebesar 0,835 dengan batch size 16, epoch 7, dan learning rate $1,26e-4$. Model terbaik pada dataset B mencapai F1-Score 0,699 dengan batch size 64, epoch 4, dan learning rate $5,7e-3$. Hasil ini menunjukkan bahwa *IndoBERT Embedding* dan FNN memberikan kinerja yang cukup baik dibandingkan baseline yang diberikan UKARA untuk train set, meskipun masih terdapat tantangan dalam ketidakseimbangan dan jumlah data yang terbatas.

Kata Kunci: AES, *IndoBERT*, FNN, SMOTE

Palembang, 05 Maret 2025

Mengetahui,
Ketua Jurusan Teknik Informatika

Pembimbing I



Hadipurnawan Satria, Ph.D.
NIP. 198004182020121001



Dr. Abdiansah, S.Kom., M.CS.
NIP. 198410012009121005

KATA PENGANTAR

Puji syukur kepada Allah SWT atas rahmat dan berkah-Nya yang telah diberikan kepada penulis sehingga mampu menyelesaikan skripsi ini dengan baik. Skripsi ini disusun sebagai salah satu syarat kelulusan program Strata-1 di Fakultas Ilmu Komputer Universitas Sriwijaya. Berkat dukungan, bimbingan, dan bantuan dari banyak pihak secara langsung maupun tidak langsung, penulis dapat menyelesaikan skripsi ini. Penulis ingin menyampaikan terima kasih kepada:

1. Kedua orang tua (Ayah dan Ibu) serta saudara penulis yang telah mendoakan juga memberi semangat kepada penulis untuk menyelesaikan skripsi ini.
2. Bapak Hadipurnawan Satria, Ph.D. sebagai Ketua Jurusan Teknik Informatika Universitas Sriwijaya.
3. Bapak Dr. Abdiansah, S.Kom., M.Cs. sebagai Dosen Pembimbing yang telah memberikan bimbingan dan motivasi sehingga penulis mampu menyelesaikan tugas akhir dengan tepat waktu.
4. Ibu Anggina Primanita, M.IT., Ph.D. sebagai Pembimbing Akademik penulis sejak awal perkuliahan hingga akhir.
5. Yayasan Karya Salemba Empat sebagai yayasan yang telah memberikan beasiswa pendidikan kepada penulis sejak semester 3 hingga akhir kuliah. Bapak Hengky Poerwowidagdo dan Bapak Satriadi Indarmawan yang telah memotivasi penulis untuk meraih cita-cita yang lebih tinggi.
6. Ibu Parwati Surjaudaja, PT Profesional Telekomunikasi Indonesia (Protelindo), PT Jaringan Kreasi Nusantara (Veyor), dan Give2Asia sebagai

Ibunda dan Ayahanda donatur beasiswa KSE yang telah mendukung penuh beasiswa pendidikan penulis.

7. Teman-teman Paguyuban KSE Nusantara dan Paguyuban KSE Universitas Sriwijaya yang telah mewarnai hari-hari penulis dengan kegiatan positif.
8. Teman-teman dan pihak lainnya yang tidak dapat penulis sebutkan satu-persatu.

Penulis menyadari bahwa penyusunan skripsi ini masih memiliki kekurangan. Penulis menerima kritik dan saran yang membangun untuk menyempurnakan penelitian. Semoga tugas akhir ini mampu memberikan manfaat. Terima kasih.

Indralaya, 03 Maret 2025
Penulis

Ramadhania Humaira

DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
TANDA LULUS UJIAN KOMPREHENSIF	iii
HALAMAN PERNYATAAN	iv
MOTTO DAN PERSEMBAHAN	v
ABSTRACT.....	vi
ABSTRAK	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xv
BAB I PENDAHULUAN.....	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang Masalah.....	I-1
1.3 Rumusan Masalah	I-4
1.4 Tujuan Penelitian.....	I-4
1.5 Manfaat Penelitian	I-4
1.6 Batasan Masalah.....	I-5
1.7 Sistematika Penulisan	I-5
1.8 Kesimpulan	I-6
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori	II-1
2.2.1 Automated Essay Scoring	II-1
2.2.2 Pra-Pemrosesan Teks.....	II-2
2.2.3 SMOTE	II-4
2.2.4 IndoBERT <i>Embedding</i>	II-6
2.2.5 <i>Feedforward Neural Network (FNN)</i>	II-8
2.2.6 <i>Confusion Matrix</i>	II-11
2.2.7 <i>Rational Unified Process</i>	II-13
2.3 Penelitian Lain yang Relevan.....	II-14
2.4 Kesimpulan	II-16

BAB III METODOLOGI PENELITIAN.....	III-1
3.1 Pendahuluan	III-1
3.2 Pengumpulan Data	III-1
3.2.1 Jenis dan Sumber Data	III-1
3.2.2 Metode Pengumpulan Data	III-2
3.3 Tahapan Penelitian	III-2
3.3.1 Mengumpulkan Data	III-3
3.3.2 Menentukan Kerangka Kerja Penelitian	III-5
3.3.3 Menentukan Kriteria Pengujian	III-6
3.3.4 Menentukan Format Data Pengujian.....	III-7
3.3.5 Menentukan Alat Bantu Penelitian.....	III-7
3.3.6 Melakukan Pengujian Penelitian.....	III-8
3.3.7 Melakukan Analisis dan Menarik Kesimpulan Penelitian	III-8
3.4 Kesimpulan	III-10
 BAB IV RANCANGAN PERANGKAT LUNAK	 IV-1
4.1 Pendahuluan	IV-1
4.2 Fase Insepsi	IV-1
4.2.1 Pemodelan Bisnis	IV-1
4.2.2 Kebutuhan Sistem	IV-2
4.2.3 Analisis dan Desain	IV-2
4.3 Fase Elaborasi	IV-12
4.3.1 Pemodelan Bisnis	IV-13
4.3.2 Kebutuhan Sistem	IV-15
4.3.3 Analisis dan Perancangan	IV-15
4.4 Fase Konstruksi.....	IV-17
4.4.1 Kebutuhan Sistem	IV-17
4.4.2 Implementasi	IV-18
4.5 Fase Transisi.....	IV-20
4.5.1 Pemodelan Bisnis	IV-21
4.5.2 Rencana Pengujian	IV-21
4.6 Kesimpulan	IV-22
 BAB V HASIL DAN ANALISIS	 V-1
5.1 Pendahuluan	V-1

5.2	Hasil Percobaan.....	V-1
5.2.1	Konfigurasi Percobaan.....	V-1
5.2.2	Hasil Percobaan dengan <i>Answer Embedding</i>	V-4
5.2.3	Pelatihan Model dengan <i>Distance Feature</i>	V-11
5.2.4	Percobaan SMOTE.....	V-16
5.3	Analisis Hasil Penelitian	V-19
5.3.1	Analisis Hasil Penelitian Skenario 1	V-19
5.3.2	Analisis Hasil Penelitian Skenario 2	V-21
5.3.3	Analisis Hasil Penelitian Skenario 3	V-22
5.3.4	Perbandingan Hasil dengan Penelitian Terdahulu.....	V-24
5.4	Kesimpulan	V-26
BAB VI KESIMPULAN DAN SARAN		VI-1
6.1	Pendahuluan	VI-1
6.2	Kesimpulan	VI-1
6.3	Saran.....	VI-1
DAFTAR PUSTAKA		vii
LAMPIRAN.....		x

DAFTAR TABEL

Tabel II-1. Tabel Confusion Matrix.....	II-11
Tabel III-1. Pertanyaan untuk Dataset A dan Dataset B.....	III-3
Tabel III-2. Kunci Jawaban Esai A dan Esai B	III-4
Tabel III-3. Contoh Jawaban Siswa untuk Esai A dan Esai B	III-5
Tabel III-4. Format Data Pengujian	III-7
Tabel III-5. Rancangan Tabel Hasil Pengujian Pertama dan Kedua	III-9
Tabel III-6. Rancangan Tabel Hasil Pengujian Ketiga.....	III-9
Tabel III-7. Rancangan Tabel Analisis Pengujian Pertama dan Kedua.....	III-9
Tabel III-8. Rancangan Tabel Perbandingan Hasil Model.....	III-9
Tabel IV-1. Kebutuhan Fungsional	IV-2
Tabel IV-2. Kebutuhan Non-Fungsional	IV-2
Tabel IV-3. Contoh Dataset Jawaban Esai A.....	IV-4
Tabel IV-4. Contoh Dataset Jawaban Esai B.....	IV-5
Tabel IV-5. Hasil Tokenisasi Dataset Jawaban Esai A	IV-6
Tabel IV-6. Hasil Tokenisasi Dataset Jawaban Esai B	IV-7
Tabel IV-7. Hasil Embedding Dataset Jawaban Esai A.....	IV-9
Tabel IV-8. Hasil Embedding Dataset Jawaban Esai B.....	IV-9
Tabel IV-9. Defisini Actor	IV-11
Tabel IV-10. Definisi Use Case	IV-11
Tabel IV-11. Skenario Use Case Demonstrasi.....	IV-11
Tabel IV-12. Perancangan Data	IV-13
Tabel IV-13. Keterangan Implementasi Kelas.....	IV-19
Tabel IV-14. Rencana Pengujian Use Case Demonstrasi	IV-21
Tabel IV-15. Pengujian Use Case Demonstrasi.....	IV-22
Tabel V-1. Pembagian Dataset.....	V-1
Tabel V-2. Arsitektur Model	V-3
Tabel V-3. Konfigurasi Hyperparameter	V-3
Tabel V-4. Hasil F1-Score Terbaik Train Set Answer Embedding A	V-5
Tabel V-5. Hasil Parameter Terbaik Train Set Answer Embedding A.....	V-6
Tabel V-6. Hasil F1-Score Terbaik Validation Set Answer Embedding A	V-6
Tabel V-7. Hasil F1-Score Terbaik Test Set Answer Embedding A	V-7
Tabel V-8. Hasil F1-Score Terbaik Train Set Answer Embedding B	V-8
Tabel V-9. Hasil Parameter Terbaik Train Set Answer Embedding B.....	V-9
Tabel V-10. Hasil F1-Score Terbaik Validation Set Answer Embedding B	V-9
Tabel V-11. Hasil F1-Score Terbaik Test Set Answer Embedding B	V-10
Tabel V-12. Hasil F1-Score Terbaik Train Set Distance Feature A.....	V-11
Tabel V-13. Hasil Parameter Terbaik Train Set Distance Feature A.....	V-12
Tabel V-14. Hasil F1-Score Terbaik Validation Set Distance Feature A	V-13
Tabel V-15. Hasil F1-Score Terbaik Train Set Distance Feature B.....	V-14
Tabel V-16. Hasil Parameter Terbaik Train Set Distance Feature B.....	V-14
Tabel V-17. Hasil F1-Score Terbaik Validation Set Distance Feature B	V-15
Tabel V-18. Hasil F1-Score Terbaik Test Set Distance Feature B	V-16
Tabel V-19. F1-Score Terbaik Train Set dengan SMOTE pada Dataset A.....	V-17
Tabel V-20. Parameter Terbaik Train Set dengan SMOTE pada Dataset A	V-18

Tabel V-21. F1-Score Terbaik Validation Set dengan SMOTE Dataset A	V-18
Tabel V-22. Hasil F1-Score Terbaik Set Uji dengan SMOTE pada Dataset A..	V-19
Tabel V-23. Perbandingan Hasil Input Fitur Dataset A	V-19
Tabel V-24. Perbandingan Hasil Input Fitur Dataset B	V-21
Tabel V-25. Perbandingan Hasil Penelitian Terdahulu	V-24

DAFTAR GAMBAR

Gambar II-1. Ilustrasi Cara Kerja SMOTE	II-4
Gambar II-2. Ilustrasi Cara Kerja SMOTE-Tomek.....	II-5
Gambar II-3. Ilustrasi Cara Kerja IndoBERT Embedding dan FNN	II-7
Gambar III-1. Tahapan Kegiatan Penelitian	III-2
Gambar III-2. Kerangka Kerja Penelitian	III-6
Gambar IV-1. Diagram Use Case	IV-10
Gambar IV-2. Desain Antarmuka Landing Page	IV-14
Gambar IV-3. Desain Antarmuka Penilaian Esai	IV-15
Gambar IV-4. Diagram Aktivitas	IV-16
Gambar IV-5. Diagram Sequence.....	IV-17
Gambar IV-6. Diagram Kelas.....	IV-18
Gambar IV-7. Hasil Tampilan Landing Page	IV-19
Gambar IV-8. Hasil Tampilan Penilaian Esai.....	IV-20
Gambar IV-9. Hasil Tampilan Penilaian Esai.....	IV-20
Gambar V-1. Distribusi Label Set Latih Dataset A	V-2
Gambar V-2. Distribusi Label Set Latih Dataset B	V-2
Gambar V-3. Grafik Train dan Validation Loss Answer Embedding A.....	V-5
Gambar V-4. Grafik Train dan Validation Loss Answer Embedding B.....	V-8
Gambar V-5. Grafik Train dan Validation Loss Distance Feature A	V-11
Gambar V-6. Grafik Train dan Validation Loss Distance Feature B	V-13
Gambar V-7. Grafik Train dan Validation Loss dengan SMOTE.....	V-17
Gambar V-8. Hasil Answer Embedding dan Distance Feature A	V-20
Gambar V-9. Hasil Sentence Embedding dan Distance Feature B.....	V-21
Gambar V-10. Perbandingan Hasil SMOTE dan Tidak SMOTE	V-23
Gambar V-11. Perbandingan Hasil dengan Penelitian Terdahulu.....	V-25
Gambar V-12. Perbandingan Hasil Performa Train Set Dataset A.....	V-25

BAB I

PENDAHULUAN

1.1 Pendahuluan

Bab ini akan membahas secara mendalam mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah. Keseluruhan penelitian akan dijelaskan secara garis besar dalam bab ini.

1.2 Latar Belakang Masalah

Meningkatkan kualitas pendidikan dapat diupayakan melalui peningkatan kualitas sistem penilaian dan pembelajaran. Dua hal ini saling berkaitan sebab sistem penilaian yang berkualitas mampu memotivasi pengajar untuk menentukan strategi pengajaran yang lebih baik lagi, sehingga kegiatan pembelajaran pun akan lebih memotivasi murid selama belajar (Warlizasusi et al., 2022). Bentuk sistem penilaian selama Kegiatan Belajar Mengajar (KBM) cukup beragam, salah satu bentuk penilaian yang sering digunakan adalah penilaian esai. Esai adalah metode penilaian yang bersifat subjektif dan dirancang untuk menguji kemampuan murid dalam menyusun diskusi serta deskripsi dengan kata-kata. Setiap murid yang menyelesaikan esai dapat memberikan jawaban yang beragam sehingga banyaknya variasi dalam jawaban tersebut dapat memperpanjang waktu yang dibutuhkan pengajar untuk melakukan koreksi.

Automated Essay Scoring (AES) adalah salah satu implementasi dari ilmu Pemrosesan Bahasa Alami yang digunakan untuk mengevaluasi dan memberi penilaian pada esai. Implementasi AES dilakukan dengan menganalisis teks dan

ekstraksi fitur yang diterapkan untuk melihat tren dan pola pada jawaban siswa (Xu et al., 2024). Fitur ekstraksi dapat diperoleh dari jarak absolut antara *embedding* jawaban siswa dan *embedding* jawaban esai (Liang et al., 2018), atau hanya dengan *embedding* jawaban siswa saja (Rajagede, 2021). AES dapat menganalisis tulisan berdasarkan relevansi antara jawaban siswa dan jawaban esai. Penggunaan AES dapat disarankan sebagai penilaian efektif karena dapat menghindari penilaian manual yang lambat dan kurang akurat (Mizumoto & Eguchi, 2023). Contoh metode yang digunakan untuk mengembangkan AES dapat berupa TF IDF, *Cosine Similarity*, FastText, dan *Pre-trained Transformer* seperti BERT. Perkembangan terbaru dari pembuatan model AES adalah penggunaan *embedding* kata kontekstual yang memanfaatkan *transformers*. Berbeda dengan pendekatan tradisional sebelumnya yang merepresentasikan makna kata sebagai vektor tetap, model *transformers* secara dinamis menyesuaikan representasi kata sesuai konteks (Mayfield et al., 2020). Sebuah *Neural Network* dapat melatih sebuah esai yang direpresentasikan dalam bentuk representasi terdistribusi dan menghasilkan *single dense vector* yang mewakili keseluruhan esai. Sebuah model *nonlinear Neural Network* memiliki keunggulan lebih adaptif dibanding pendekatan model statistik lainnya (Liang et al., 2018).

Penelitian menunjukkan bahwa model *transformer* IndoBERT merupakan model terbaik dibanding TF-IDF, *cosine similarity*, dan *linear regression* (Pradani & Hulliyyatus, 2023) yang menghasilkan nilai MAE sebesar 0,285 dan RMSE sebesar 0,2001. Hasil penelitian lainnya yang dilakukan dengan *Transformer* BERT *Sentence Embedding* dan model sederhana *Neural Network* memiliki fitur lebih

kaya dan ukuran *file* yang lebih kecil dibanding metode FastText dan membuahkan hasil *F1-score* sebesar 0,829 terhadap dataset UKARA (Rajagede, 2021). Penelitian lain yang relevan dalam kasus dan menggunakan dataset yang sama adalah menggunakan FastText dan XGBoost yang menghasilkan *F1-Score* 0,821 (Rajagede et al., 2020). Penelitian lain juga mencoba menggunakan *Bidirectional LSTM* dan *Word2Vec* sebagai *word embedding* dengan hasil *F1-Score* 0,811 (Septiandri et al., 2019). Penggunaan metode TF-IDF, Random Forest, dengan *Logistic Regression* dilakukan oleh (Putra, 2019) dengan hasil akhir *F1-Score* sebesar 0,81.

Meskipun bahasa Indonesia memiliki potensi besar untuk dieksplorasi lebih lanjut di bidang teknologi, bahasa ini masih menghadapi tantangan dalam perkembangan ilmu Pemrosesan Bahasa Alami akibat keterbatasan data beranotasi dan sumber daya bahasa yang memadai (Koto et al., 2020). AES merupakan salah satu penerapan penting dalam Pemrosesan Bahasa Alami yang berperan besar dalam meningkatkan kualitas pendidikan di Indonesia. Berdasarkan hal tersebut, penelitian ini akan mengeksplorasi pemanfaatan sumber daya bahasa untuk mendukung penelitian Pemrosesan Bahasa Alami dalam bahasa Indonesia dengan menggunakan dataset berbahasa Indonesia untuk meneliti pengembangan sistem AES. Penelitian dilakukan dengan metode *IndoBERT embedding* dan menggunakan *Feedforward Neural Network* sebagai model klasifikasi. Pemilihan *IndoBERT embedding* didasarkan pada pertimbangan kinerjanya yang telah disesuaikan dengan karakteristik bahasa Indonesia, serta *Feedforward Neural*

Network sebagai model sederhana yang dapat mengklasifikasikan penilaian yang tidak membebani komputasi pada dataset yang terbatas.

1.3 Rumusan Masalah

Merujuk pada permasalahan yang telah diuraikan dalam latar belakang, penelitian ini merumuskan masalah sebagai berikut:

1. Bagaimana metode IndoBERT *embedding* dan *Feedforward Neural Network* sebagai model klasifikasi dapat digunakan untuk menghasilkan sistem *Automated Essay Scoring* berbahasa Indonesia?
2. Bagaimana kinerja metode IndoBERT *embedding* dan *Feedforward Neural Network* sebagai model klasifikasi dalam sistem *Automated Essay Scoring*?

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Membangun sebuah sistem dalam bentuk situs web untuk penilaian esai bahasa Indonesia dengan IndoBERT *embedding* dan *Feedforward Network*.
2. Mengetahui nilai F1-Score IndoBERT *embedding* dan *Feedforward Network*.
3. Membandingkan hasil penelitian dengan penelitian terdahulu.

1.5 Manfaat Penelitian

Manfaat yang dihasilkan dari penelitian ini adalah sebagai berikut.:

1. Meningkatkan kualitas penilaian secara otomatis untuk esai berbahasa Indonesia.
2. Hasil penelitian ini dapat dijadikan sebagai penelitian selanjutnya di bidang *Automated Essay Scoring*.

1.6 Batasan Masalah

Lingkup permasalahan dalam penelitian ini ditetapkan sebagai berikut:

1. Objek penelitian fokus pada pengembangan *Automated Essay Scoring* berbahasa Indonesia.
2. Penelitian ini menggunakan dataset sekunder berisi jawaban esai berbahasa Indonesia dengan penilaian 1 untuk benar dan 0 untuk salah dari dua pertanyaan berbeda.
3. *Pretrained Language Model* untuk proses *embedding* menggunakan IndoBERT dan model klasifikasi yang digunakan adalah *Feedforward Neural Network*.

1.7 Sistematika Penulisan

Struktur penulisan dalam penelitian ini disusun sebagai berikut:

BAB I. PENDAHULUAN

Bab ini menguraikan latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, serta sistematika penelitian yang akan menjadi landasan utama dalam penelitian ini.

BAB II. KAJIAN LITERATUR

Pada bab ini membahas landasan teori yang digunakan dalam penelitian, yakni *Automated Essay Scoring*, *Text Preprocessing*, SMOTE, IndoBERT *Embedding*, *Feedforward Neural Network*, *Confusion Matrix*, *Rational Unified Process*, penelitian yang relevan, dan kesimpulan.

BAB III. METODOLOGI PENELITIAN

Bab ini menguraikan tahapan yang akan dilakukan dalam penelitian, termasuk pengumpulan data, analisis data, dan perancangan perangkat lunak. Setiap langkah dijelaskan sesuai dengan kerangka kerja yang telah disusun.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Bab ini mengulas metode pengembangan perangkat lunak menggunakan *Rational Unified Process* (RUP) dalam membangun sistem AES. Selain itu, perancangan struktur perangkat lunak serta pengujian untuk memastikan sistem yang dikembangkan sesuai dengan kebutuhan penelitian juga dibahas dalam bab ini.

BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini memaparkan hasil pengujian AES menggunakan *IndoBERT Embedding* dan *Feedforward Neural Network* sesuai dengan tahapan yang telah dirancang. Analisis yang disajikan berfungsi sebagai dasar dalam merumuskan kesimpulan penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Bab ini menyajikan kesimpulan yang diperoleh berdasarkan pembahasan pada bab sebelumnya serta saran yang diberikan sebagai tindak lanjut dari penelitian yang telah dilakukan.

1.8 Kesimpulan

Bab ini menyajikan kesimpulan yang diperoleh berdasarkan pembahasan pada bab sebelumnya serta saran yang diberikan sebagai tindak lanjut dari penelitian yang telah dilakukan.

DAFTAR PUSTAKA

- Albab, M. U., Karuniawati P, Y., & Fawaiq, M. N. (2023). *Optimization of the Stemming Technique on Text preprocessing President 3 Periods Topic*. 20(2), 1–10. <https://doi.org/10.26623/transformatika.v20i2.5374>
- Anjarwani Endang Sri, Afwani Royana, & Perwitasari Ririn. (2020). Penerapan Metode Rational Unified Process (Rup) Dalam Pengembangan Sistem Informasi Medical Check Up Pada Citra Medical Centre. *Jurnal Teknologi Informasi, Komputer, Dan Aplikasinya (JTika)*, 2(1), 76–88. <https://doi.org/https://doi.org/10.29303/jtika.v2i1.85>
- Ayu Pradani, K., & Suada Hulliyyatus, L. (2023). Automated Essay Scoring Menggunakan Semantic Textual Similarity Berbasis Transformer Untuk Penilaian Ujian Esai. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTik)*, 10(6), 1177–1184. <https://doi.org/10.25126/jtiik.2023107338>
- Baharuddin, F., & Naufal, M. F. (2023). Fine-Tuning IndoBERT for Indonesian Exam Question Classification Based on Bloom’s Taxonomy. *Journal of Information Systems Engineering and Business Intelligence*, 9(2), 253–263. <https://doi.org/10.20473/jisebi.9.2.253-263>
- Budiman, I., Faisal, M. R., Faridhah, A., Farmadi, A., Mazdadi, M. I., Saragih, T. H., & Abadi, F. (2023). Classification Performance Comparison of BERT and IndoBERT on SelfReport of COVID-19 Status on Social Media. *Journal of Computer Sciences Institute*, 30, 61–67. <https://doi.org/10.35784/jcsi.5564>
- Chamidah, N., Mega Santoni, M., Nurramdhani Irmanda, H., Astriratma, R., Ilmu Komputer, F., & Pembangunan Nasional Veteran Jakarta, U. (2022). *Penilaian Esai Pendek Otomatis Berdasarkan Similaritas Semantik dengan SBERT Semantic Similarity-Based Short Essay Auto Scoring using SBERT* (Vol. 21, Issue 4).
- Chandra, W., Suprihatin, B., & Resti, Y. (2023). Median-KNN Regressor-SMOTE-Tomek Links for Handling Missing and Imbalanced Data in Air Quality Prediction. *Symmetry*, 15(4). <https://doi.org/10.3390/sym15040887>
- Chassab, R. H., Zakaria, L. Q., & Tiun, S. (2024). An Optimized LSTM-Based Augmented Language Model (FLSTM-ALM) Using Fox Algorithm for Automatic Essay Scoring Prediction. *IEEE Access*, 12, 48713–48724. <https://doi.org/10.1109/ACCESS.2024.3381619>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. <https://arxiv.org/abs/1810.04805>

- Gong, Z., Gao, C., Feng, Z., Dong, P., Qiao, H., Zhang, H., Shi, L., & Guo, W. (2024). Integrating masked generative distillation and network compression to identify the severity of wheat fusarium head blight. *Computers and Electronics in Agriculture*, 227, 109647. <https://doi.org/10.1016/j.compag.2024.109647>
- Hadiprakoso, R. B., Setiawan, H., Yasa, R. N., & Girinoto. (2023, December). Text preprocessing for optimal accuracy in Indonesian sentiment analysis using a deep learning model with word embedding. *AIP Conference Proceedings*. <https://doi.org/10.1063/5.0126116>
- Heydarian, M., Doyle, T. E., & Samavi, R. (2022). MLCM: Multi-Label Confusion Matrix. *IEEE Access*, 10, 19083–19095. <https://doi.org/10.1109/ACCESS.2022.3151048>
- Jayadianti, H., Kaswidjanti, W., Utomo, A. T., Saifullah, S., Dwiyanto, F. A., & Drezewski, R. (2022). Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN. *ILKOM Jurnal Ilmiah*, 14(3), 348–354. <https://doi.org/10.33096/ilkom.v14i3.1505.348-354>
- Kalyan, K. S., Rajasekharan, A., & Sangeetha, S. (2021). *AMMUS: A Survey of Transformer-based Pretrained Models in Natural Language Processing*. <http://arxiv.org/abs/2108.05542>
- Kang, J., Schwartz, R., Flickinger, J., & Beriwal, S. (2015). Machine Learning Approaches for Predicting Radiation Therapy Outcomes: A Clinician's Perspective. *International Journal of Radiation Oncology*Biography*Physics*, 93(5), 1127–1135. <https://doi.org/10.1016/j.ijrobp.2015.07.2286>
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). *IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP*. <http://arxiv.org/abs/2011.00677>
- Krstinić, D., Braović, M., Šerić, L., & Božić-Štulić, D. (2020). *Multi-label Classifier Performance Evaluation with Confusion Matrix*. 01–14. <https://doi.org/10.5121/csit.2020.100801>
- Li, W., & Liu, H. (2024). Applying large language models for automated essay scoring for non-native Japanese. *Humanities and Social Sciences Communications*, 11(1). <https://doi.org/10.1057/s41599-024-03209-9>
- Maldonado, S., Vairetti, C., Fernandez, A., & Herrera, F. (2022). FW-SMOTE: A feature-weighted oversampling approach for imbalanced classification. *Pattern Recognition*, 124, 108511. <https://doi.org/10.1016/j.patcog.2021.108511>
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2). <https://doi.org/10.1016/j.rmal.2023.100050>

- Rai, A., & Borah, S. (2021). Study of Various Methods for Tokenization. *Lecture Notes in Networks and Systems*, 193–200. https://doi.org/10.1007/978-981-15-6198-6_18
- Rajagede, R. A. (2021). Improving Automatic Essay Scoring for Indonesian Language using Simpler Model and Richer Feature. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 11–18. <https://doi.org/10.22219/kinetik.v6i1.1196>
- Rodriguez, P. U., Jafari, A., & Ormerod, C. M. (2019). *Language models and Automated Essay Scoring*. <http://arxiv.org/abs/1909.09482>
- Samsir, Irmayani, D., Edi, F., Harahap, J. M., Jupriaman, Rangkuti, R. K., Ulya, B., & Watrianthos, R. (2021). Naives Bayes Algorithm for Twitter Sentiment Analysis. *Journal of Physics: Conference Series*, 1933(1). <https://doi.org/10.1088/1742-6596/1933/1/012019>
- Sarasuartha Mahajaya, N., Desiana, P., Ayu, W., & Huizen, R. R. (n.d.). *Pengaruh Optimizer Adam, AdamW, SGD, dan LAMB terhadap Model Vision Transformer pada Klasifikasi Penyakit Paru-paru*. <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database>,
- Sarica, S., & Luo, J. (2021). Stopwords in technical language processing. *PLOS ONE*, 16(8). <https://doi.org/10.1371/journal.pone.0254937>
- Warlizasusi, J., Kholilah Harahap, E., Pascasarjana Manajemen Pendidikan Islam, M., Tarbiyah, F., Curup, I., Lebong Bengkulu, R., & Pascasarjana Manajemen Pendidikan Islam, D. (2022). *Supervisi Dalam Peningkatan Mutu Evaluasi Pembelajaran Di Sekolah Dasar Negeri 114 Rejang Lebong*. 1(1).
- Wibawa, A. P., Fithri, H. K., Zaeni, I. A. E., & Nafalski, A. (2020). Generating Javanese Stopwords List using K-means Clustering Algorithm. *Knowledge Engineering and Data Science*, 3(2), 106. <https://doi.org/10.17977/um018v3i22020p106-111>
- Zhang, B., Chang, K., & Li, C. (2024). *Simple Techniques for Enhancing Sentence Embeddings in Generative Language Models* (pp. 52–64). https://doi.org/10.1007/978-981-97-5669-8_5

LAMPIRAN



Lampiran kode terdapat pada bit.ly/SkripsiRamadhaniaHumaira