

Klasifikasi Kemampuan Mahasiswa Berdasarkan *Automatic Essay Scoring* terhadap Jawaban *Essay* Ujian Kompetensi dengan Metode *Machine Learning*

Muhammad Hakiki ^{1*}, Chastine Fatichah ²

^{1,2} Program Studi Teknik Informatika, Fakultas Teknologi Elektro dan Informatika Cerdas, Institut Teknologi Sepuluh Nopember, Kota Surabaya, Provinsi Jawa Timur, Indonesia.

Corresponding Email: 6025221035@student.its.ac.id ^{1*}

Histori Artikel:

Dikirim 17 Januari 2025; *Diterima dalam bentuk revisi* 1 Juli 2025; *Diterima* 10 Juli 2025; *Diterbitkan* 10 September 2025. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Penilaian jawaban mahasiswa dan pengelompokan kemampuan mahasiswa yang dilakukan secara manual sangat memakan waktu yang cukup banyak. Oleh karena itu, dibutuhkan sistem yang dapat melakukan penilaian terhadap jawaban esai mahasiswa secara otomatis dan pengelompokan kemampuan mahasiswa. Penelitian ini mengusulkan metode klasifikasi kemampuan mahasiswa berdasarkan nilai *Automatic Essay Scoring* menggunakan metode *LSTM* dan beberapa metode klasifikasi. Jumlah dataset yang digunakan dalam penelitian ini sebanyak 98 mahasiswa, sedangkan soal yang diujikan dalam ujian kompetensi ini sebanyak 200 soal. Parameter yang digunakan untuk *LSTM* adalah jawaban dari mahasiswa. Manfaat penelitian ini adalah untuk mengetahui mahasiswa yang menguasai perkuliahan dan mahasiswa yang belum menguasai perkuliahan. Hasil penelitian ini menunjukkan bahwa metode *LSTM* berhasil memberikan penilaian esai secara otomatis dengan nilai akurasi sebesar 0,9, sedangkan metode klasifikasi yang paling unggul adalah metode *Decision Tree* dengan metode *oversampling* ROS yaitu sebesar 0,654.

Kata Kunci: *Automatic Essay Scoring*; Klasifikasi; *LSTM*; *Decision Tree*.

Abstract

Manually assessing student answers and grouping student abilities is very time-consuming. Therefore, a system is needed that can automatically assess student essay answers and group student abilities. This study proposes a method for classifying student abilities based on the Automatic Essay Scoring value using the LSTM method and several classification methods. The number of datasets used in this study was 98 students, while the questions tested in this competency exam were 200 questions. The parameters used for LSTM are student answers. The benefit of this study is to find out which students have mastered the lecture and which students have not mastered the lecture. The results of this study indicate that the LSTM method successfully provides automatic essay assessment with an accuracy value of 0.9, while the most superior classification method is the Decision Tree method with the ROS oversampling method, which is 0.654.

Keyword: Automatic Essay Scoring; Classification; LSTM; Naive Bayes; SVM; Decision Tree.

1. Pendahuluan

Pada awal tahun 2020, seluruh dunia dihadapkan pada bencana besar *COVID-19* yang berdampak signifikan, khususnya di Indonesia. Pandemi tersebut menyebabkan berbagai kerusakan terhadap kehidupan manusia, salah satunya dalam aspek pendidikan. Pendidikan merupakan salah satu *human resource* yang dapat membantu perubahan dalam pembangunan negara. Pemerintah Indonesia melalui Kementerian Pendidikan dan Kebudayaan menerbitkan Surat Edaran Nomor 4 Tahun 2020 tentang Penyelenggaraan Pendidikan selama Darurat *COVID-19*. Surat edaran tersebut berisi kebijakan-kebijakan dalam penyelenggaraan pembelajaran, baik Ujian Nasional, Ujian Sekolah, maupun sistem pembelajaran dari rumah. Kebijakan tersebut mengharuskan proses belajar mengajar dilaksanakan dari rumah akibat dampak *COVID-19*, sehingga pembelajaran yang biasanya dilakukan secara *offline* kini harus dilakukan secara *online* dengan menggunakan *E-Learning* (Susanti *et al.*, 2022).

Sistem *E-Learning* telah semakin dikenal di tingkat pendidikan tinggi. Dalam sistem *E-Learning* terdapat bagian yang cukup penting, yaitu esai. Esai memberikan kesempatan kepada siapa pun untuk mengeluarkan seluruh pemahaman yang dimiliki tentang suatu materi dalam bentuk jawaban esai. Namun, penilaian jawaban esai memakan waktu dan tenaga yang cukup banyak (Ifenthaler, 2022). *Essay scoring* merupakan salah satu bagian yang sangat penting dalam bidang pendidikan. Penilaian esai yang dilakukan secara manual sangat memakan waktu yang cukup banyak. Oleh karena itu, dibutuhkan sistem untuk melakukan penilaian terhadap jawaban esai secara otomatis yang memungkinkan proses penilaian esai dengan bantuan program komputer (Susanti *et al.*, 2022).

Automatic Essay Scoring merupakan salah satu bagian dari *machine learning* yang dapat digunakan untuk melakukan penilaian terhadap jawaban esai secara otomatis. Kini *Automatic Essay Scoring* (AES) semakin mendapatkan perhatian di sektor pendidikan karena dapat mengurangi beban penilaian secara manual (Ludwig *et al.*, 2021). *Automatic Essay Scoring* merupakan metode yang cukup rumit yang memanfaatkan teknologi komputer, sedangkan penilaian esai secara manual hanya berfokus pada fitur linguistik (Li *et al.*, 2023). Pada penelitian lain, dilakukan *scoring* terhadap jawaban esai bahasa Indonesia dengan menggunakan metode *Cosine Similarity*. Metode ini digunakan dengan prinsip *similarity* di mana hanya teks yang relevan dari dokumen asli yang akan diproses untuk *scoring*. Sistem penilaian esai secara otomatis dilakukan menggunakan metode pembobotan TF-IDF, kemudian metode TF-IDF dikombinasikan dengan metode *Cosine Similarity* untuk mengukur kemiripan teks. Metode *Cosine Similarity* merupakan metode yang digunakan untuk menentukan kemiripan antara dua teks, yaitu antara kunci jawaban dan hasil jawaban esai siswa (Lahitani, 2022). Metode *automatic essay scoring* juga digunakan untuk menilai jawaban esai ujian akhir semester mata kuliah Bahasa Indonesia. Metode yang digunakan adalah *Rabin-Karp*, *Dice Coefficient Similarity*, dan *Synonym Recognition*. Metode yang digunakan dapat menghasilkan persentase tingkat akurasi hasil nilai jawaban mahasiswa dengan tingkat akurasi yang tinggi, yaitu sebesar 93,75% (Saputro & Budiyanto, 2024).

Pada penelitian terhadap penilaian hasil belajar yang dilakukan seorang guru bertujuan untuk mengevaluasi, memantau pembelajaran, dan meningkatkan hasil belajar. Arsitektur *deep learning* yang banyak digunakan adalah *Convolutional Neural Network* (CNN) dan *Recursive Neural Network* (RNN). Metode RNN yang paling umum digunakan adalah *Long Short Term Memory* (LSTM). Metode LSTM digunakan untuk memberikan penilaian dengan cara menghitung tingkat kemiripan tiap kata. Metode ini bekerja dengan cara membandingkan dua kalimat dalam setiap jaringan yang memiliki distribusi relasional kata berurutan dari kalimat tersebut (Wiratmo & Fatichah, 2020).

Penerapan metode LSTM juga digunakan untuk mengidentifikasi toksisitas yang ada dalam komentar yang berasal dari forum publik, media sosial, dan situs web. Penelitian ini bertujuan untuk mengklasifikasikan tingkat toksisitas di forum obrolan *online*. Dalam penelitian ini, beberapa klasifikasi teks digunakan untuk mengidentifikasi toksisitas yang ada dalam komentar. Penelitian ini berfokus pada efek dari *word embeddings* pada tingkat akurasi dari klasifikasi LSTM (Neuroscience, 2023). Metode LSTM juga digunakan untuk melakukan prediksi terhadap aliran sungai. Penelitian ini menggabungkan dua metode, yaitu sistem permodelan hidrologi *Weather Research and Forecasting* atau *WRF-Hydro* dan metode *Long Short-Term Memory* atau LSTM. Metode gabungan dari kedua metode

tersebut bernama WRF-*Hydro*-LSTM yang digunakan untuk meningkatkan simulasi aliran sungai. Penelitian ini menunjukkan bahwa metode WRF-*Hydro*-LSTM sangat berguna untuk merepresentasikan kondisi fisik di dunia nyata, sehingga meningkatkan hasil prediksi aliran sungai (Cho & Kim, 2022). Selain proses penilaian esai secara otomatis, penelitian ini juga berencana melakukan proses klasifikasi terhadap hasil *score* jawaban esai. Salah satu penelitian yang pernah dilakukan sebelumnya adalah dengan membandingkan beberapa metode klasifikasi sehingga dapat diperoleh nilai terbaik. Metode klasifikasi yang dibandingkan adalah *Naive Bayes* (NB), *Neural Network* (NN), *Logistic Regression* (LR), dan *Support Vector Machine* (SVM). Hasil menunjukkan bahwa SVM merupakan metode yang sangat efektif dan efisien. Namun, hasil ini hanya berhasil dibuktikan pada mata kuliah Pengantar Teknologi Informasi (Firdaus, 2022). Metode klasifikasi juga dilakukan terhadap penelitian yang membahas pengembangan metode CNN yang digunakan untuk mengidentifikasi model-model *drone* dalam kehidupan nyata. Hasil pengujian menunjukkan bahwa metode CNN menghasilkan akurasi sebesar 92,4%, presisi sebesar 88,8%, *recall* sebesar 88,6%, dan skor F1 sebesar 88,7% (Wisniewski *et al.*, 2022).

Metode CNN juga digunakan dalam penelitian untuk meningkatkan kinerja klasifikasi citra medis. Penelitian ini menggunakan metode momentum adaptif yang baru untuk melakukan klasifikasi citra medis dengan menggunakan metode CNN (Aytaç *et al.*, 2022). Penelitian terdahulu pernah melakukan perbandingan dua metode klasifikasi, yaitu *Decision Tree* dan SVM. Hasil menunjukkan bahwa metode SVM memiliki nilai akurasi yang lebih tinggi dan stabil, yaitu sebesar 96,62%, sedangkan metode *Decision Tree* menghasilkan nilai interpretabilitas yang lebih tinggi (Kusumawati *et al.*, 2024). Metode klasifikasi *Naive Bayes* juga digunakan untuk melakukan klasifikasi dan telah digunakan untuk menghasilkan prediksi mahasiswa berprestasi. Metode *Naive Bayes* yang digunakan untuk memprediksi mahasiswa berprestasi berhasil menunjukkan nilai akurasi yang sangat tinggi, yaitu 98%, nilai *recall* sebesar 97%, nilai *precision* sebesar 100%, dan nilai *F1-Score* sebesar 99% (Nurul Jariah Sintam *et al.*, 2024). Metode klasifikasi *Naive Bayes* pernah digunakan dalam penelitian dengan tujuan untuk mengelompokkan bidang kemampuan mahasiswa sehingga dapat dengan mudah memilih fokus studi dan judul skripsi yang sesuai dengan bidang studi mahasiswa. Hasil penelitian ini menunjukkan bahwa metode klasifikasi *Naive Bayes* menghasilkan nilai akurasi rata-rata sebesar 100% dan nilai *recall* rata-rata sebesar 100% (Astari Uumbu Zaza *et al.*, 2024). Metode klasifikasi *Naive Bayes* juga digunakan dalam penelitian untuk memberikan prediksi apakah mahasiswa jurusan matematika tersebut lulus tepat waktu atau terlambat. Hasil penelitian ini menunjukkan bahwa metode klasifikasi *Naive Bayes* berhasil mendapatkan nilai akurasi sebesar 70%, nilai *error* sebesar 30%, dan nilai AUC sebesar 72% (Ilham Rizki Maziz Tomu & Malon Stivo Noya Van Delsen, 2024).

Metode SVM juga pernah digunakan dengan membandingkannya dengan metode *Random Forest* untuk memprediksi capaian studi mahasiswa. Penelitian ini melakukan perbandingan antara metode klasifikasi SVM dan metode *Random Forest*. Hasil penelitian ini menunjukkan bahwa metode *Random Forest* dan SVM setelah dilakukan *feature selection* menghasilkan nilai akurasi yang sama, yaitu sebesar 97,67% (Novianto *et al.*, 2024). Penelitian yang dilakukan dengan tujuan untuk memprediksi tingkat kelulusan tepat waktu mahasiswa dengan membandingkan beberapa metode, yaitu metode *Naive Bayes*, *Random Forest*, SVM, dan ANN. Hasil penelitian ini menunjukkan bahwa metode SVM memiliki nilai akurasi tertinggi dibandingkan metode *Naive Bayes*, *Random Forest*, dan ANN, yaitu sebesar 0,94 (Satrio Junaidi *et al.*, 2024). Terdapat penelitian yang menganalisis pengaruh musik terhadap penurunan tingkat stres mahasiswi semester 7 saat mengerjakan skripsi. Penelitian ini melakukan perbandingan dua metode, yaitu metode SVM dan metode *Random Forest*. Hasil penelitian ini menunjukkan bahwa metode *Random Forest* memiliki nilai akurasi yang lebih tinggi, yaitu sebesar 85% (Muriyatmoko *et al.*, 2024). Metode klasifikasi lain yang digunakan untuk melakukan klasifikasi adalah *Decision Tree*. Metode ini pernah digunakan untuk menentukan kelulusan mahasiswa. Metode *Decision Tree* menghasilkan nilai akurasi yang sangat tinggi, sebesar 99%. Namun, penggunaan metode *Decision Tree* dapat ditingkatkan dengan menambah variabel yang lebih banyak (Jurnal *et al.*, 2024).

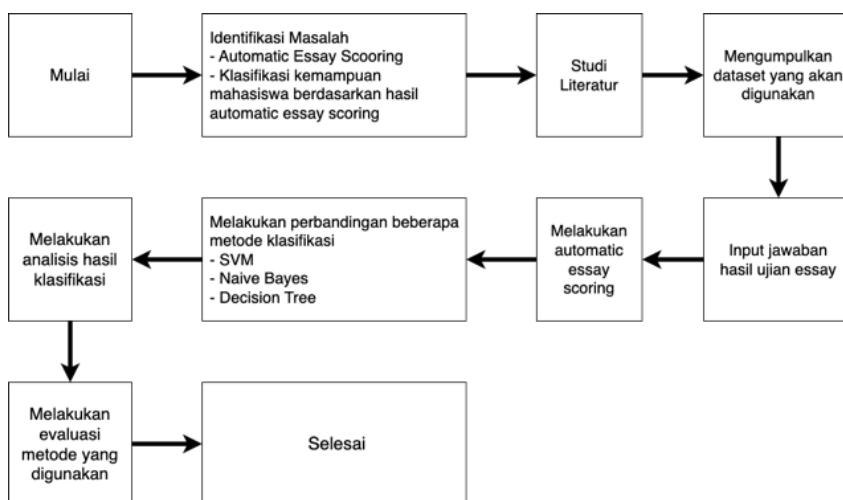
Penelitian dengan melakukan klasifikasi mahasiswa berpotensi putus studi menggunakan *Decision Tree* menunjukkan bahwa metode *Decision Tree* berhasil digunakan untuk mengklasifikasikan

mahasiswa yang memiliki kemungkinan untuk berhenti studi. Hasil penelitian ini menunjukkan bahwa metode *Decision Tree* menghasilkan nilai akurasi sebesar 100%, nilai *precision* sebesar 100%, nilai *recall* sebesar 100%, dan nilai *F1-score* sebesar 100% (Penulis *et al.*, 2024). Metode *Decision Tree* juga digunakan dalam penelitian untuk memprediksi kelulusan mahasiswa tepat waktu menggunakan metode *Naïve Bayes* dan *Decision Tree*. Penelitian ini dilakukan dengan cara membandingkan dua metode klasifikasi, yaitu metode *Naïve Bayes* dan metode *Decision Tree*. Hasil penelitian ini menunjukkan bahwa metode *Naïve Bayes* lebih unggul dengan nilai akurasi sebesar 71,24% (Inda S, 2024).

Berdasarkan beberapa metode yang telah dilakukan sebelumnya untuk melakukan *Automatic Essay Scoring*, dapat disimpulkan bahwa kekurangan yang dihadapi adalah keterbatasan dalam penerapannya yang hanya dilakukan terhadap satu objek saja. Sedangkan untuk metode klasifikasi, kekurangan yang dimiliki dari metode yang telah dilakukan adalah penerapannya yang hanya diterapkan pada satu objek saja, sehingga dibutuhkan penerapan lain dengan objek yang berbeda. Pada penelitian ini, diusulkan metode LSTM untuk melakukan *Automatic Essay Scoring* terhadap hasil jawaban esai ujian mahasiswa dan mengklasifikasikan kemampuan mahasiswa berdasarkan hasil dari *Automatic Essay Scoring*. *Score* esai didapat dari hasil pencocokkan jawaban esai mahasiswa dengan kunci jawaban ujian menggunakan metode LSTM. Kemudian dari hasil LSTM tersebut dilakukan klasifikasi terhadap *score* esai sehingga dapat diketahui mahasiswa yang menguasai materi dan yang belum. Diharapkan dengan adanya penelitian ini dapat diterapkan di bidang kedokteran, kemudian penerapannya akan diterapkan bersamaan dengan penerapan beberapa metode klasifikasi seperti SVM, *Naïve Bayes*, dan *Decision Tree*.

2. Metode Penelitian

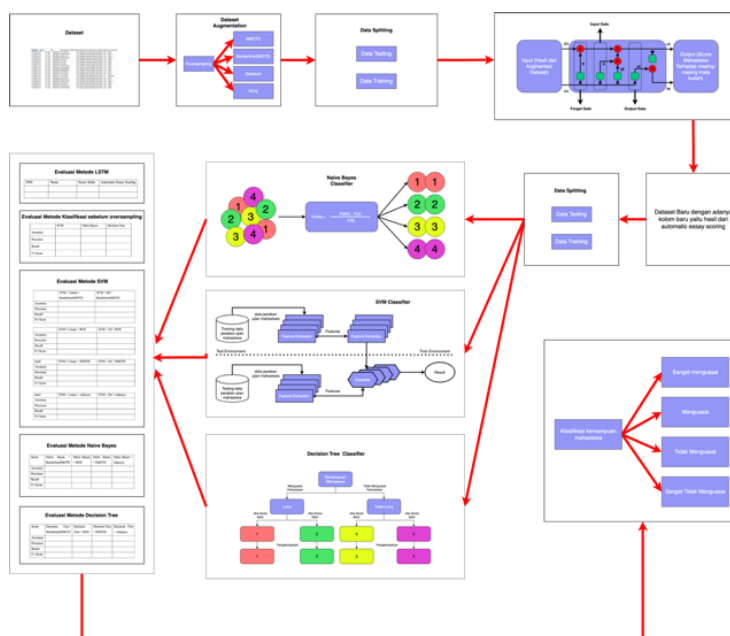
Fokus dari penelitian ini adalah untuk mengetahui kemampuan setiap mahasiswa dari fakultas kedokteran ketika menjalani ujian kompetensi sehingga dapat diketahui mahasiswa mana yang menguasai perkuliahan dan mahasiswa yang belum menguasai perkuliahan melalui proses klasifikasi berdasarkan hasil *score* jawaban esai yang didapat dari perhitungan *automatic essay scoring*. Dalam upaya untuk mencapai tujuan tersebut, penelitian ini akan dilakukan dengan melibatkan beberapa serangkaian tahapan yang dilakukan secara sistematis.



Gambar 1. Model Proses Penelitian

Berdasarkan Gambar 1, terdapat beberapa proses yang akan dilakukan dalam penelitian ini. Langkah pertama yang harus dilakukan adalah menentukan objek penelitian. Langkah selanjutnya adalah melakukan beberapa studi literatur, baik yang berkaitan dengan beberapa metode klasifikasi

maupun studi literatur yang berkaitan dengan metode *Long Short-Term Memory* atau LSTM. Langkah selanjutnya adalah mengumpulkan semua *dataset*. Kemudian melakukan *Automatic Essay Scoring*. Hasil dari *Automatic Essay Scoring* akan digunakan untuk melakukan klasifikasi kemampuan mahasiswa.



Gambar 2. Diagram Alir Metode yang Diusulkan

Gambar 2 merupakan *flowchart* alur kerja metode yang akan dilakukan. Dimulai dari pengumpulan *dataset*, melakukan augmentasi data dengan menggunakan metode *oversampling*, yaitu metode SMOTE, *BorderlineSMOTE*, ROS, dan *Adasyn*. Kemudian setelah melakukan augmentasi data, dilakukan proses *Automatic Essay Scoring* dengan menggunakan metode LSTM. Hasil dari metode LSTM tersebut akan diteruskan dengan melakukan klasifikasi menggunakan beberapa metode klasifikasi, yaitu metode SVM, *Naïve Bayes*, dan *Decision Tree*. Hasil akhir dari proses ini adalah klasifikasi kemampuan mahasiswa dimulai dari kelompok tidak menguasai, menguasai, sampai pada sangat menguasai.

2.1 Menentukan Objek Penelitian

Langkah pertama yang akan dilakukan dalam penelitian ini adalah menentukan objek penelitian. Objek yang akan digunakan dalam penelitian ini adalah jawaban esai mahasiswa dari hasil ujian kompetensi mahasiswa kedokteran.

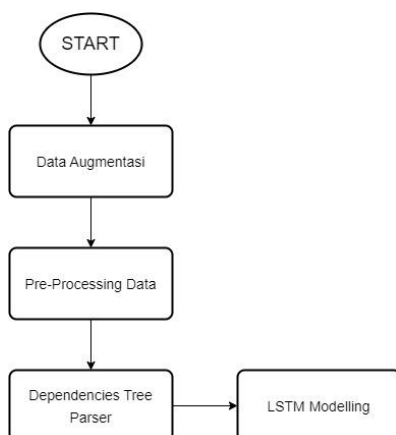
2.2 Dataset yang Digunakan

Dataset yang digunakan dalam penelitian ini diambil dari data hasil jawaban esai ujian kompetensi mahasiswa kedokteran di Perguruan Tinggi Swasta di Surabaya. *Dataset* ini berisi jawaban-jawaban mahasiswa. Jumlah mahasiswa yang mengikuti ujian kompetensi ini sebanyak 98 mahasiswa, sedangkan soal yang diujikan dalam ujian kompetensi ini sebanyak 200 soal.

2.3 Studi Literatur

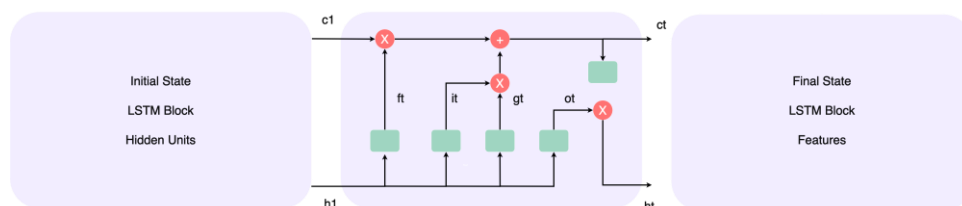
Langkah berikutnya yang dilakukan adalah melakukan studi literatur yang bertujuan untuk meringkas, menyusun, mendeskripsikan, dan menjelaskan teori untuk menyelesaikan permasalahan sebelumnya. Studi literatur dilakukan terhadap beberapa metode klasifikasi seperti metode SVM, *Naïve Bayes*, dan *Decision Tree*, serta metode *Automatic Essay Scoring*, khususnya metode *Long Short-Term Memory* (LSTM).

2.4 Automatic Essay Scoring



Gambar 3. Diagram Alur *Automatic Essay Scoring*

Gambar 3 merupakan alur kerja dari metode *Long Short-Term Memory* atau LSTM. Dari gambar di atas dapat dilihat bahwa terdapat beberapa langkah yang harus dilakukan. Langkah tersebut dimulai dengan melakukan *Augmentation Data*. Tahapan selanjutnya adalah melakukan *Pre-Processing* data. Tahapan selanjutnya adalah melakukan *Dependencies Tree Parser*, dan yang terakhir adalah melakukan *Long Short-Term Memory* atau LSTM Modelling.



Gambar 4. Arsitektur Metode LSTM

Gambar 4 merupakan arsitektur metode LSTM. Terdapat beberapa komponen yang ada dalam metode LSTM. f_t atau *Forget Gate* berfungsi untuk menentukan *input* data mana yang akan diolah dan dipilih serta yang akan disimpan atau dibuang dalam *memory cell*.

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f)$$

Dimana:

f_t = *forget gate*

σ = fungsi aktivasi *sigmoid*

W_f = bobot *forget gate*

h_{t-1} = nilai *hidden state cell* sebelum

X_t = nilai *input*

b_f = bias *forget gate*

i_t atau *Input Gate* berfungsi untuk mengambil nilai *input* sebelumnya dan *input* baru. Persamaan pada *gate* ini menggunakan fungsi aktivasi *sigmoid* dan aktivasi *tanh*.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$C_t = \tanh (W_c \cdot [h_{t-1}, x_t] + b_c)$$

Dimana:

i_t = *input gate*

W_i = bobot *input gate*

b_i = bias *input gate*

C_t = *candidate gate*

Tanh = fungsi aktivasi tanh

C_t atau *Cell State* digunakan untuk menggantikan nilai pada *memory cell* sebelumnya dengan *memory cell* yang baru dengan melakukan operasi hasil pada *gate unit* sebelumnya.

$$c_t = (i_t * C_t + f_t * c_{t-1})$$

Dimana:

c_t = *cell gate*

i_t = *input gate*

C_t = *candidate gate*

f_t = *forget gate*

c_{t-1} = nilai *cell state* sebelum

O_t atau *Output Gate* digunakan untuk menentukan nilai pada bagian *memory cell* mana yang akan dikeluarkan dengan menggunakan aktivasi *sigmoid*.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh (c_t)$$

Dimana:

o_t = *output gate*

W_o = bobot *output gate*

b_o = bias *output gate*

h_t = *hidden state*

c_t = *cell gate*

Dalam prosesnya, sebelum menggunakan metode LSTM, terdapat proses *pre-processing* yang harus dilakukan terlebih dahulu melalui beberapa tahapan penting. Langkah pertama dalam melakukan *Dependencies Tree Parser* adalah mengumpulkan *list* atau daftar kata yang didapatkan dari proses *tokenizing* yang dilakukan dalam proses *pre-processing* data. Langkah selanjutnya yang harus dilakukan adalah memeriksa relasi dari setiap *token* yang ada dalam *stack* dengan menggunakan model yang sudah dibuat sebelumnya. Relasi dari setiap *token* yang ada dalam *stack* terdiri dari *arc* kiri, *arc* kanan, dan *shift*. Untuk *arc* kiri merupakan ketergantungan relasi dari *stack* ke *stack* sebelumnya, sedangkan untuk *arc* kanan sendiri merupakan ketergantungan relasi dari *stack* ke *stack* selanjutnya. Namun, jika tidak ada hubungan relasi maka disebut dengan *shift*. Langkah selanjutnya yang harus dilakukan adalah mengubah *stack* yang bergantung pada relasi yang dimilikinya. Jika relasinya merupakan *arc* kiri maka *stack* terakhir yang sebelumnya akan dihapus. Jika relasinya memiliki *arc* kanan maka *stack* terakhir akan dihapus, namun jika memiliki *shift* maka *token* pertama yang ada dalam daftar kata ditambahkan ke dalam *stack*. Proses ini dilakukan secara berulang kali hingga memiliki [ROOT] dan [] dalam daftar kata. Dalam proses *modelling* yang pertama ini, *input* data yang dimasukkan berupa vektor dari setiap

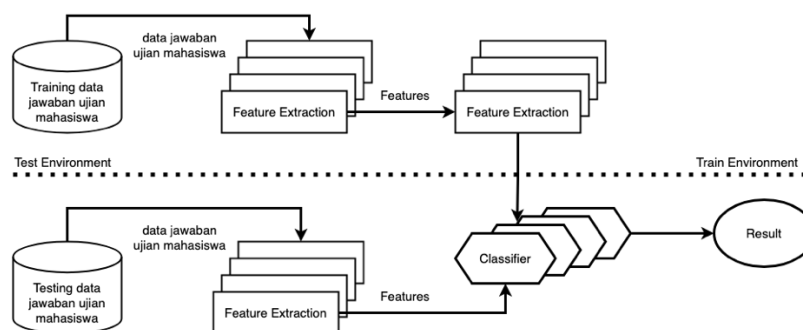
kata yang ada di setiap kalimat di setiap *prompt* dalam kumpulan data *source* dan *dependency parser* dari kalimat ini. Vektor data yang digunakan dibuat dengan menggunakan dimensi *GloVe* 250. *Output* dari proses klasifikasi yang dilakukan ini adalah untuk mendapatkan nilai skor.

2.5 Augmentasi Data

Langkah selanjutnya sebelum melakukan proses klasifikasi adalah melakukan proses augmentasi data. Terdapat dua metode yang dapat digunakan untuk melakukan augmentasi data ini, yaitu metode *oversampling* dan *undersampling*. Dalam penelitian ini dilakukan *oversampling* karena *dataset* yang digunakan memiliki jumlah yang sedikit. *Oversampling* dilakukan dengan menggunakan metode *SMOTE*, *BorderlineSMOTE*, *ROS*, dan *Adasyn*.

2.6 Klasifikasi Kemampuan Mahasiswa

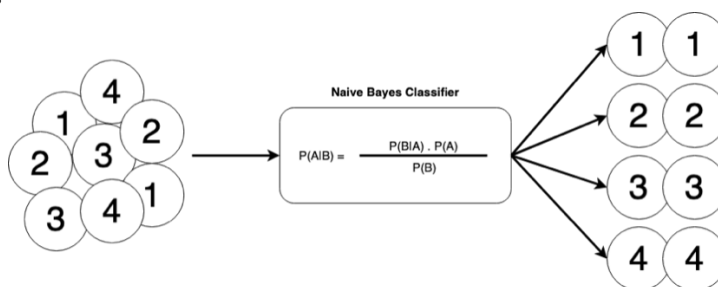
2.6.1 SVM



Gambar 5. Diagram Alur Klasifikasi Kemampuan Mahasiswa

Gambar 5 merupakan diagram alur klasifikasi kemampuan mahasiswa yang digunakan dalam penelitian ini. Data dipecah menjadi data *testing* dan data *training*, kemudian dilakukan *feature extraction*, setelah itu akan dilanjutkan dengan proses *classifier* dengan hasil akhir adalah pengelompokan data.

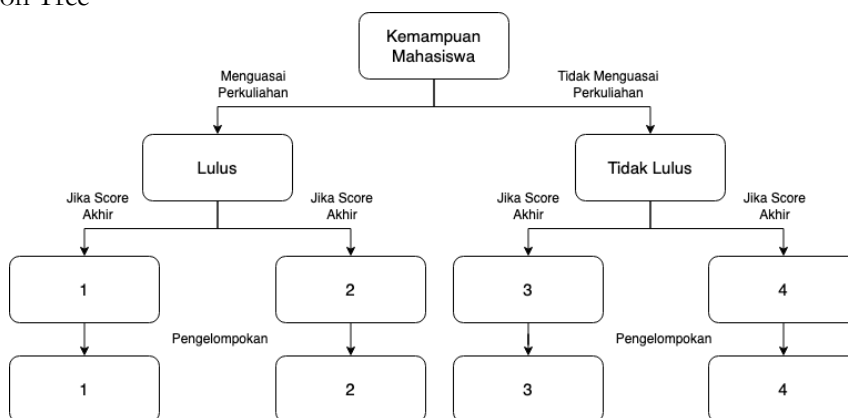
2.6.2 Naïve Bayes



Gambar 6. Diagram Kerja Metode *Naïve Bayes*

Gambar 6 merupakan struktur dari penggunaan metode klasifikasi *Naïve Bayes*. Dari gambar di atas dapat dilihat bahwa sekelompok data "1,2,3,4" sebelum dilakukan klasifikasi *Naïve Bayes* masih berkumpul menjadi satu dan belum terkelompokkan. Setelah dilakukan klasifikasi *Naïve Bayes* dengan rumus *Naïve Bayes*, sekumpulan data "1,2,3,4" berhasil dikelompokkan sesuai dengan *dataset*-nya. Untuk data berisi "1" akan dikelompokkan dengan data berisi "1", begitu juga data yang lainnya sampai ke data berisi "4". Tahapan-tahapan dalam menggunakan metode klasifikasi *Naïve Bayes* akan dijelaskan sebagai berikut:

2.6.3 Decision Tree



Gambar 7. Diagram Kerja Metode *Decision Tree*

Gambar 7 merupakan struktur dari metode klasifikasi *Decision Tree*. Dari gambar di atas didapat bahwa mahasiswa akan lulus jika masuk ke dalam kelompok dengan skor total 1 dan 2. Sedangkan mahasiswa tidak akan lulus jika masuk ke dalam kelompok dengan skor total 3 dan 4.

3. Hasil dan Pembahasan

3.1 Hasil

Pada bagian ini akan dijelaskan mengenai hasil dari *Automatic Essay Scoring* dengan menggunakan metode *Long Short-Term Memory (LSTM)* dan klasifikasi kemampuan mahasiswa dengan menggunakan beberapa metode klasifikasi seperti metode *SVM*, *Naïve Bayes*, dan metode *Decision Tree*.

3.1.1 Hasil Metode LSTM

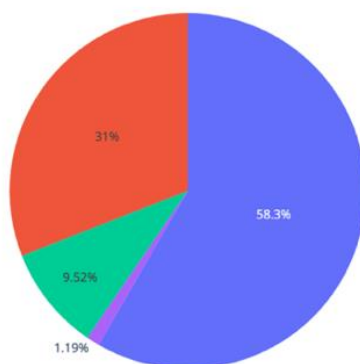
Tabel 1. Hasil *Automatic Essay Scoring* dengan Metode LSTM

NIM	Nama	Score Akhir	<i>Automatic Essay Scoring</i>
001	A	0,441	1
002	B	0,5	0,626
003	C	0,744	0,005
004	D	0,662	0,486

Tabel 1 merupakan hasil akhir dari jawaban esai yang didapatkan dari penggunaan metode *Long Short-Term Memory (LSTM)* untuk melakukan *Automatic Essay Scoring*. Dari tabel di atas dapat dilihat semua hasil *Automatic Essay Scoring* dari masing-masing mahasiswa. Hasil dari *Automatic Essay Scoring* menunjukkan nilai *loss* sebesar 0,0051 dan nilai akurasi sebesar 99%.

3.1.2 Hasil Klasifikasi Kemampuan Mahasiswa

Gambar 8 merupakan grafik dari hasil kemampuan mahasiswa. Grafik yang berwarna biru menunjukkan mahasiswa yang tidak menguasai perkuliahan, grafik yang berwarna merah menunjukkan mahasiswa yang menguasai perkuliahan, grafik yang berwarna hijau menunjukkan mahasiswa yang sangat menguasai pembelajaran, dan grafik yang berwarna ungu menunjukkan mahasiswa yang sangat menguasai pembelajaran.



Gambar 8. Hasil Klasifikasi

Tabel 2 merupakan hasil akhir dari perbandingan beberapa metode klasifikasi, yaitu metode SVM, *Naïve Bayes*, dan *Decision Tree* dengan beberapa metode *oversampling*, yaitu metode SMOTE, *BorderlineSMOTE*, ROS, dan *Adasyn*. Dari tabel di atas dapat disimpulkan bahwa metode dengan nilai tertinggi adalah metode *Decision Tree* + ROS dengan nilai akurasi sebesar 0,654.

Tabel 2. Hasil Perbandingan Metode Klasifikasi

Score	SVM + Linear + <i>Adasyn</i>	<i>Naïve Bayes</i> + <i>Adasyn</i>	<i>Decision Tree</i> + ROS
<i>Accuracy</i>	0,615	0,5	0,654
<i>Precision</i> 1 (kelas 1)	0	0	0
<i>Precision</i> 2 (kelas 2)	0,71	0,69	0,75
<i>Precision</i> 3 (kelas 3)	0,62	0,75	1
<i>Precision</i> 4 (kelas 4)	0,25	0,11	0,17
<i>Recall</i> 1 (kelas 1)	0	0	0
<i>Recall</i> 2 (kelas 2)	0,77	0,69	0,92
<i>Recall</i> 3 (kelas 3)	0,45	0,27	0,36
<i>Recall</i> 4 (kelas 4)	1	1	1
<i>F1-Score</i> 1 (kelas 1)	0	0	0
<i>F1-Score</i> 2 (kelas 2)	0,74	0,69	0,83
<i>F1-Score</i> 3 (kelas 3)	0,53	0,40	0,53
<i>F1-Score</i> 4 (kelas 4)	0,40	0,20	0,29

3.2 Pembahasan

Berdasarkan hasil yang diperoleh dari implementasi metode LSTM untuk *Automatic Essay Scoring*, terlihat bahwa metode ini mampu menghasilkan nilai akurasi yang sangat tinggi, yaitu sebesar 99% dengan nilai *loss* yang sangat rendah sebesar 0,0051. Hal ini menunjukkan bahwa metode LSTM efektif dalam melakukan penilaian otomatis terhadap jawaban esai mahasiswa kedokteran. Tingginya akurasi ini dapat disebabkan oleh kemampuan LSTM dalam memahami konteks dan struktur bahasa yang kompleks dalam jawaban esai mahasiswa. Metode LSTM mampu menangkap dependensi jangka panjang dalam teks, sehingga dapat memahami hubungan antar kata dan kalimat yang berkontribusi terhadap kualitas jawaban secara keseluruhan. Hasil *scoring* yang diperoleh menunjukkan variasi yang cukup signifikan antar mahasiswa, dengan rentang skor dari 0,005 hingga 1. Variasi ini mengindikasikan bahwa metode LSTM mampu membedakan tingkat pemahaman mahasiswa terhadap materi yang diujikan dengan baik. Perbedaan skor yang mencolok ini mencerminkan kemampuan model dalam mengidentifikasi nuansa kualitas jawaban, mulai dari jawaban yang sangat lemah hingga jawaban yang sangat baik. Hal ini menunjukkan bahwa model LSTM tidak hanya

memberikan penilaian secara kasar, tetapi mampu memberikan gradasi penilaian yang detail sesuai dengan tingkat pemahaman masing-masing mahasiswa.

Hasil perbandingan ketiga metode klasifikasi menunjukkan bahwa metode *Decision Tree* dengan teknik *oversampling* ROS memberikan performa terbaik dengan nilai akurasi sebesar 0,654 atau 65,4%. Meskipun nilai akurasi ini tidak terlalu tinggi, namun masih dalam kategori yang dapat diterima untuk klasifikasi dengan empat kelas. Performa ini dapat dijelaskan oleh karakteristik metode *Decision Tree* yang mampu menangani data numerik dengan baik dan tidak memerlukan asumsi distribusi tertentu seperti pada metode *Naïve Bayes*. Selain itu, metode *Decision Tree* juga memiliki kemampuan interpretabilitas yang tinggi, sehingga hasil klasifikasi dapat dijelaskan dengan mudah kepada pengguna. Metode *Decision Tree* menunjukkan keunggulan dalam beberapa aspek yang signifikan. Pertama, dari segi akurasi tertinggi dengan nilai 0,654, metode ini unggul dibandingkan SVM yang mencapai 0,615 dan *Naïve Bayes* yang hanya mencapai 0,5. Keunggulan ini menunjukkan bahwa struktur pohon keputusan lebih sesuai untuk mengklasifikasikan data skor esai yang memiliki karakteristik numerik kontinu. Kedua, dari segi performa per kelas, metode *Decision Tree* menunjukkan konsistensi yang baik. Untuk kelas 2, metode ini memiliki nilai *recall* tertinggi sebesar 0,92 dan *F1-Score* tertinggi sebesar 0,83, yang menunjukkan kemampuan model dalam mengidentifikasi mahasiswa yang menguasai perkuliahan dengan baik. Untuk kelas 3, metode ini bahkan mencapai *precision* sempurna sebesar 1,0, yang berarti semua prediksi untuk kelas ini benar. Konsistensi yang baik juga terlihat pada kelas 4 dengan *recall* sempurna sebesar 1,0, menunjukkan bahwa model mampu mengidentifikasi semua mahasiswa yang sangat menguasai pembelajaran. Kesesuaian metode *Decision Tree* dengan *dataset* yang digunakan dapat dijelaskan melalui beberapa alasan teknis. Pertama, metode ini sangat efektif untuk data numerik seperti skor hasil *Automatic Essay Scoring* karena dapat membuat pemisahan (*split*) yang optimal berdasarkan nilai numerik tertentu. Kedua, metode *Decision Tree* tidak sensitif terhadap ukuran *dataset* yang kecil, dalam hal ini 98 mahasiswa, karena tidak memerlukan estimasi parameter yang kompleks seperti pada metode SVM atau *neural network*. Ketiga, metode ini mampu menangani data yang tidak seimbang dengan baik ketika dikombinasikan dengan teknik ROS (*Random Over Sampling*), yang membantu mengatasi masalah *class imbalance* yang sering terjadi dalam data klasifikasi kemampuan mahasiswa.

Hasil penelitian ini menunjukkan bahwa kombinasi metode LSTM untuk *Automatic Essay Scoring* dan *Decision Tree* untuk klasifikasi kemampuan mahasiswa dapat menjadi solusi yang efektif untuk evaluasi otomatis dalam pendidikan kedokteran. Meskipun masih terdapat ruang untuk perbaikan, terutama dalam hal akurasi klasifikasi, sistem ini dapat memberikan kontribusi yang signifikan dalam proses pembelajaran. Sistem ini dapat membantu dosen dalam mengurangi beban kerja penilaian manual yang selama ini memakan waktu dan tenaga yang besar. Dengan otomatisasi penilaian, dosen dapat mengalokasikan waktu mereka untuk aktivitas yang lebih strategis seperti pengembangan kurikulum dan bimbingan mahasiswa. Selain itu, sistem ini juga dapat memberikan evaluasi yang lebih objektif dan konsisten dibandingkan penilaian manual yang dapat dipengaruhi oleh faktor subjektif seperti kelelahan atau *bias* personal. Konsistensi penilaian ini sangat penting dalam konteks pendidikan kedokteran di mana standar kompetensi harus dijaga dengan ketat. Sistem ini dapat membantu mengidentifikasi mahasiswa yang memerlukan perhatian khusus dalam pembelajaran dengan lebih cepat dan akurat. Dengan klasifikasi otomatis, dosen dapat segera mengetahui mahasiswa mana yang berada dalam kategori "tidak menguasai" atau "menguasai" sehingga dapat memberikan intervensi yang tepat waktu. Hal ini sangat penting dalam pendidikan kedokteran di mana setiap mahasiswa harus mencapai standar kompetensi minimum sebelum dapat melanjutkan ke tahap praktik klinis.

Beberapa keterbatasan yang perlu diperhatikan dalam penelitian ini mencakup aspek teknis dan metodologis yang dapat mempengaruhi generalisasi hasil. Pertama, ukuran *dataset* yang relatif kecil dengan hanya 98 mahasiswa dapat mempengaruhi generalisasi model ke populasi yang lebih besar. *Dataset* yang kecil dapat menyebabkan *overfitting* pada model LSTM dan mengurangi kemampuan model untuk beradaptasi dengan data baru yang memiliki karakteristik berbeda. Untuk penelitian selanjutnya, diperlukan *dataset* yang lebih besar dan beragam untuk meningkatkan robustitas model.

Kedua, nilai akurasi klasifikasi tertinggi yang hanya mencapai 65,4% menunjukkan bahwa masih terdapat ruang yang signifikan untuk perbaikan. Akurasi ini mungkin dapat ditingkatkan melalui *hyperparameter tuning* yang lebih optimal, penggunaan *ensemble methods*, atau eksplorasi metode klasifikasi lain seperti *Random Forest* atau *Gradient Boosting*. Ketiga, ketidakseimbangan kelas yang terlihat dari nilai *precision* dan *recall* yang bervariasi antar kelas menunjukkan bahwa distribusi data tidak merata. Hal ini dapat mempengaruhi performa model, terutama untuk kelas minoritas. Meskipun telah dilakukan teknik *oversampling*, masih diperlukan strategi yang lebih sophisticated seperti *cost-sensitive learning* atau *ensemble methods* yang khusus dirancang untuk menangani *class imbalance*.

4. Kesimpulan dan Rekomendasi

Penelitian ini telah berhasil mengimplementasikan metode *Automatic Essay Scoring* untuk melakukan penilaian jawaban esai dengan menggunakan *machine learning*. Metode *Automatic Essay Scoring* yang digunakan dalam penelitian ini adalah metode *Long Short-Term Memory* (LSTM) yang terbukti efektif dalam memahami konteks dan struktur bahasa dalam jawaban esai mahasiswa kedokteran. Implementasi metode LSTM menghasilkan performa yang sangat baik dengan tingkat akurasi mencapai 99% dan nilai *loss* yang sangat rendah sebesar 0,0051, menunjukkan bahwa metode ini mampu melakukan penilaian otomatis dengan tingkat presisi yang tinggi. Mengingat jumlah *dataset* yang digunakan dalam penelitian ini tergolong cukup sedikit, yaitu 98 mahasiswa, maka diperlukan teknik *oversampling* untuk mengatasi ketidakseimbangan data dan meningkatkan performa model klasifikasi. Dalam penelitian ini, empat metode *oversampling* yang berbeda telah diimplementasikan dan dibandingkan, yaitu metode SMOTE (*Synthetic Minority Oversampling Technique*), metode *BorderlineSMOTE*, metode ROS (*Random Over Sampling*), dan metode *Adasyn* (*Adaptive Synthetic Sampling*). Setiap metode memiliki karakteristik dan keunggulan tersendiri dalam menangani masalah ketidakseimbangan kelas yang sering terjadi dalam data klasifikasi kemampuan mahasiswa.

Berdasarkan evaluasi yang telah dilakukan, penelitian ini melakukan perbandingan sistematis terhadap berbagai kombinasi metode klasifikasi dan teknik *oversampling*. Perbandingan dimulai dengan evaluasi tiga metode klasifikasi utama, yaitu metode SVM (*Support Vector Machine*), metode *Naïve Bayes*, dan metode *Decision Tree* tanpa menggunakan teknik *oversampling*. Selanjutnya, setiap metode klasifikasi dikombinasikan dengan keempat metode *oversampling* untuk mengevaluasi peningkatan performa yang dapat dicapai. Evaluasi dilakukan berdasarkan beberapa metrik performa yang komprehensif, meliputi akurasi, *precision*, *recall*, dan *F1-score* untuk memberikan gambaran yang menyeluruh tentang kemampuan setiap kombinasi metode.

Hasil evaluasi menunjukkan bahwa kombinasi metode klasifikasi *Decision Tree* dengan teknik *oversampling* ROS memberikan performa terbaik di antara semua kombinasi yang diuji, dengan nilai akurasi sebesar 0,654 atau 65,4%. Keunggulan kombinasi ini dapat dijelaskan melalui beberapa faktor. Pertama, metode *Decision Tree* memiliki kemampuan yang baik dalam menangani data numerik seperti skor hasil *Automatic Essay Scoring*. Kedua, metode ini tidak memerlukan asumsi distribusi data tertentu, sehingga lebih fleksibel dalam menangani variasi data yang ada. Ketiga, teknik ROS terbukti efektif dalam mengatasi ketidakseimbangan kelas tanpa mengubah distribusi data asli secara signifikan, sehingga mempertahankan karakteristik data yang penting untuk proses klasifikasi.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan yang perlu diakui dan menjadi perhatian untuk pengembangan selanjutnya. Keterbatasan utama adalah ukuran *dataset* yang digunakan cenderung memiliki jumlah yang sedikit, yaitu hanya 98 mahasiswa. Ukuran *dataset* yang kecil ini dapat mempengaruhi generalisasi model dan mengurangi kemampuan sistem untuk beradaptasi dengan data baru yang memiliki karakteristik berbeda. Selain itu, ketidakseimbangan kelas yang masih terlihat dari variasi nilai *precision* dan *recall* antar kelas menunjukkan bahwa masih diperlukan strategi yang lebih canggih untuk menangani masalah ini. Untuk penelitian selanjutnya, diharapkan implementasi sistem *Automatic Essay Scoring* ini dapat diperluas tidak hanya untuk ujian kompetensi Fakultas Kedokteran, tetapi juga dapat diterapkan di semua fakultas untuk membantu

penilaian kinerja mahasiswa secara lebih luas. Ekspansi ini akan memberikan manfaat yang lebih besar dalam dunia pendidikan tinggi, membantu dosen dari berbagai disiplin ilmu dalam melakukan evaluasi yang lebih objektif dan efisien. Selain itu, penelitian selanjutnya juga diharapkan dapat menggunakan *dataset* yang lebih besar dan beragam untuk meningkatkan robustitas dan generalisasi model, serta mengeksplorasi metode-metode *machine learning* yang lebih canggih seperti *ensemble methods* atau *deep learning* yang lebih kompleks untuk mencapai akurasi yang lebih tinggi.

5. Daftar Pustaka

- Ameliatus, Q., & Fatah, Z. (2024). Perbandingan metode klasifikasi Random Forest dan Support Vector Machine dalam memprediksi capaian studi mahasiswa. <https://doi.org/10.69714/em8qnw54>
- Arifin, A. R., Purnamasari, P. D., & Ratna, A. A. P. (2021, June). Automatic essay scoring for indonesian short answers using siamese manhattan long short-term memory. In *2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)* (pp. 1-6). IEEE. <https://doi.org/10.1109/ICECCE52056.2021.9514223>
- Astari Uumbu Zaza, D., Umar, E., Ema Ose Sanga, F., Maris Sumba Jl Karya kasih No, S., Tambolaka, K., Barat Daya, S., & Abstrak, N. (2024). Penerapan text mining dalam klasifikasi judul skripsi mahasiswa pada Universitas Stella Maris Sumba menggunakan metode Naïve Bayes. *Journal of Informatics and Business*, 2(3).
- Aytaç, U. C., Güneş, A., & Ajlouni, N. (2022). A novel adaptive momentum method for medical image classification using convolutional neural network. *BMC Medical Imaging*, 22(1). <https://doi.org/10.1186/s12880-022-00755-z>
- Bakti, R. Y. (2024). Klasifikasi Mahasiswa Berpotensi Putus Studi Menggunakan Algoritma Decision Tree pada Fakultas Teknik Unismuh Makassar. *Arus Jurnal Sains dan Teknologi*, 2(2), 532-542. <https://doi.org/10.57250/ajst.v2i2.674>
- Chimingyang, H. (2020). An automatic system for essay questions scoring based on LSTM and word embedding. *Proceedings - 2020 5th International Conference on Information Science, Computer Technology and Transportation, ISCTT 2020*, 355–364. <https://doi.org/10.1109/ISCTT51595.2020.00068>
- Cho, K., & Kim, Y. (2022). Improving streamflow prediction in the WRF-Hydro model with LSTM networks. *Journal of Hydrology*, 605, 127297. <https://doi.org/10.1016/j.jhydrol.2021.127297>
- Ifenthaler, D. (2022). Automated essay scoring systems. In *Handbook of open, distance and digital education* (pp. 1-15). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-19-0351-9_59-1
- Intelligence, C. (2023). Retracted: An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding. *Computational Intelligence and Neuroscience*, 2023, 9850820. <https://doi.org/10.1155/2023/9850820>
- Kusumawati, R., Ainul Yaqin, M., & Artikel, S. (2024). Klasifikasi data mahasiswa lampau menggunakan metode Decision Tree dan Support Vector Machine. *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, 6(3), 328–340. <https://doi.org/10.28926/ilkomnika.v6i3.687>

- Lagakis, P., & Demetriadis, S. (2021, November). Automated essay scoring: A review of the field. In *2021 International Conference on Computer, Information and Telecommunication Systems (CITS)* (pp. 1-6). IEEE. <https://doi.org/10.1109/CITS52676.2021.9618476>
- Lahitani, A. R. (2022). Automated Essay Scoring menggunakan Cosine Similarity pada Penilaian Esai Multi Soal. *Jurnal Kajian Ilmiah*, *22*(2), 107-118. <https://doi.org/10.31599/m8n1x493>
- Li, F., Xi, X., Cui, Z., Li, D., & Zeng, W. (2023). Automatic essay scoring method based on multi-scale features. *Applied Sciences*, *13*(11), 6775. <https://doi.org/10.3390/app13116775>
- Ludwig, S., Mayer, C., Hansen, C., Eilers, K., & Brandt, S. (2021). Automated essay scoring using transformer models. *Psych*, *3*(4), 897–915. <https://doi.org/10.3390/psych3040056>
- Muriyatmoko, D., Musthafa, A., & Fa-Idzaa, M. (2024, December). Perbandingan Metode Support Vector Machine dan Random Forest dalam Menganalisis Pengaruh Musik Terhadap Penurunan Tingkat Stress Mahasiswa Semester 7 saat Skripsi (Studi Kasus: Universitas Darussalam Gontor). In *Prosiding Seminar Nasional Amikom Surakarta* (Vol. 2, pp. 128-135).
- Novianto, E., Suhirman, S., & Prasetyo, D. (2024). Perbandingan metode klasifikasi Random Forest dan Support Vector Machine dalam memprediksi capaian studi mahasiswa. *9*(4), 1821–1833. <https://doi.org/10.29100/jipi.v9i4.5423>
- Nurhidayat, A., Asmunin, A., & Suyatno, D. F. (2021). Prediksi Kinerja Akademik Mahasiswa Menggunakan Machine Learning dengan Sequential Minimal Optimization untuk Pengelola Program Studi. *JIEET (Journal of Information Engineering and Educational Technology)*, *5*(2), 84-91. <https://doi.org/10.26740/jieet.v5n2.p84-91>
- Nurul Jariah Sintam, R., Akram Nur Risal, A., Fatmarani Suriyanto, D., Komputer, T., Negeri Makassar Jl Mallengkeri Raya, U., Tambung, P., Tamalate, K., Makassar, K., & Selatan, S. (2024). Klasifikasi Mahasiswa Berprestasi Menggunakan Fuzzy C-Means Dan Naive Bayes. *Jurnal Ilmiah Informatika Global*, *15*(1), 9-16.
- Saputro, H., & Budiyanto, U. (2024). Model Penilaian Esai Otomatis Menggunakan Algoritma Rabin-Karp, Dice Coefficient Similarity dan Synonym Recognition Studi Kasus Pada Universitas Budi Luhur. *SKANIKA: Sistem Komputer dan Teknik Informatika*, *7*(1), 1-11. <https://doi.org/10.36080/skanika.v7i1.3146>
- Sari, J. S. I., Umar, E., & Momo, L. L. (2024). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Naive Bayes Dan Decision Tree Pada Universitas Stella Maris Sumba. *Journal Of Informatics And Busines*, *2*(3), 362-368.
- Satrio Junaidi, Valicia Anggela, R., & Kariman, D. (2024). Klasifikasi metode data mining untuk prediksi kelulusan tepat waktu mahasiswa dengan algoritma Naive Bayes, Random Forest, Support Vector Machine (SVM) dan Artificial Neural Network (ANN). *Journal of Applied Computer Science and Technology*, *5*(1), 109–119. <https://doi.org/10.52158/jacost.v5i1.489>
- Shedriko, S., & Firdaus, M. (2022, January). Penentuan Klasifikasi Dengan Crisp-Dm Dalam Memprediksi Kelulusan Mahasiswa Pada Suatu Mata Kuliah. In *Seminar Nasional Riset dan Inovasi Teknologi (SEMNAS RISTEK)* (Vol. 6, No. 1). <https://doi.org/10.30998/semnasristek.v6i1.5814>

- Susanti, M. N. I., Ramadhan, A., & Warnars, H. L. H. S. (2022). Automatic essay exam scoring system: A systematic literature review. *Procedia Computer Science*, 216, 531–538. <https://doi.org/10.1016/j.procs.2022.12.166>
- Tomu, I. R. M., Noya Van Delsen, M. S., & Lewaherilla, N. (2024). Klasifikasi Ketepatan Waktu Lulus Mahasiswa Jurusan Matematika Universitas Pattimura Menggunakan Metode Naïve Bayes Classifier. *Proximal: Jurnal Penelitian Matematika Dan Pendidikan Matematika*, 7(2), 864–872. <https://doi.org/10.30605/proximal.v7i2.4153>
- Wiratmo, A., & Fatichah, C. (2020). Indonesian short essay scoring using transfer learning dependency tree LSTM. *International Journal of Intelligent Engineering and Systems*, 13(2), 278–285. <https://doi.org/10.22266/ijies2020.0430.27>
- Wisniewski, M., Rana, Z. A., & Petrunin, I. (2022). Drone model classification using convolutional neural network trained on synthetic data. *Journal of Imaging*, 8(8), 218. <https://doi.org/10.3390/jimaging8080218>.