

Klasifikasi Risiko *Dropout* Mahasiswa Menggunakan LSTM dan *Naïve Bayes*

¹Kristina Annatasia Br Sitepu, ²I Gusti Prahmana

³Novenda Putra Linarta, ⁴Wanda Arfilla Daulay

^{1,4}Sistem Informasi, STMIK Kaputama, Binjai, Indonesia

^{2,3}Teknik Informatika, STMIK Kaputama, Binjai, Indonesia

*Korespondensi: kannatasia88@gmail.com

Submit : 20 Sept 2026 | Diterima : 02 Okt 2026 | Terbit : 08 Okt 2026

ABSTRACT

High Dropout rates significantly undermine the effectiveness of higher education institutions, making early risk detection essential. This research proposes a hybrid classification framework to predict Dropout vulnerability by combining Long Short-Term Memory (LSTM)-based sentiment analysis with a Naïve Bayes classifier. Data were gathered from active undergraduates via comprehensive questionnaires, capturing structured metrics current semester, GPA, academic and economic constraints, and motivation levels alongside three qualitative essay responses. The textual data underwent standard preprocessing phases, including cleaning, case folding, tokenization, stopword filtering, and stemming. Utilizing the Indonesian Sentiment Lexicon (InSet), these texts were categorized into positive, neutral, or negative polarities to train the LSTM network. A majority voting scheme then aggregated the predicted sentiments from the three essays into a single representative sentiment score per student. This derived sentiment, combined with the structured attributes, served as the input for the Naïve Bayes algorithm to categorize students into low, medium, or high Dropout risk profiles. Performance assessments yielded an accuracy of 67% for the LSTM sentiment extraction. The subsequent Naïve Bayes classification recorded an accuracy of 22.22%, precision of 72.22%, recall of 22.22%, and an F1-score of 26.46%. While the current metrics reflect limitations due to severe data imbalance, the proposed architecture demonstrates theoretical viability for early Dropout detection. Future iterations require expanded, well-balanced datasets to optimize predictive reliability.

Keywords: *Dropout_risk; Long_Short-Term_Memory; machine_learning; Naïve_Bayes; sentiment_analysis*

ABSTRAK

Tingginya angka putus kuliah secara signifikan berdampak pada keberhasilan institusi pendidikan tinggi, sehingga deteksi risiko dini menjadi sangat krusial. Penelitian ini mengusulkan kerangka klasifikasi hibrida untuk memprediksi kerentanan putus kuliah dengan menggabungkan analisis sentimen berbasis *Long Short-Term Memory* (LSTM) dan pengklasifikasi *Naïve Bayes*. Data dikumpulkan dari mahasiswa aktif melalui kuesioner yang mencakup atribut terstruktur semester, IPK, faktor akademik dan ekonomi, serta tingkat motivasi ditambah tiga respons esai kualitatif. Data tekstual melalui tahapan pra-pemrosesan standar meliputi pembersihan, *case folding*, tokenisasi, penghapusan *stopword*, dan *stemming*. Menggunakan *Indonesian Sentiment Lexicon* (InSet), teks tersebut dilabeli menjadi polaritas positif, netral, atau negatif untuk melatih jaringan LSTM. Skema *majority voting* kemudian digunakan untuk mengagregasi prediksi sentimen dari ketiga esai menjadi satu sentimen akhir per mahasiswa. Sentimen ini, digabungkan dengan atribut terstruktur lainnya, digunakan sebagai fitur masukan bagi algoritma *Naïve Bayes* untuk mengklasifikasikan mahasiswa ke dalam risiko rendah, sedang, atau tinggi. Pengujian performa menghasilkan akurasi sebesar 67% untuk ekstraksi sentimen LSTM. Sementara itu, klasifikasi *Naïve Bayes* mencatat akurasi 22,22%, presisi 72,22%, *recall* 22,22%, dan *F1-score* 26,46%. Meskipun metrik klasifikasi akhir mencerminkan keterbatasan akibat ketidakseimbangan kelas data, arsitektur yang diusulkan menunjukkan potensi kelayakan sebagai model awal untuk identifikasi risiko. Penelitian mendatang sangat merekomendasikan penggunaan dataset yang lebih besar dan seimbang untuk meningkatkan keandalan prediksi.

Kata Kunci: analisis sentimen; *Long Short-Term Memory*; *machine learning*; *Naïve Bayes*; risiko putus kuliah

PENDAHULUAN

Fenomena putus studi atau *Dropout* mahasiswa masih menjadi salah satu tantangan krusial yang dihadapi oleh berbagai lembaga perguruan tinggi. Tingginya angka *Dropout* tidak hanya berdampak negatif pada kelangsungan akademik individu, tetapi juga memengaruhi indikator kinerja serta efisiensi operasional institusi pendidikan. Secara umum, kerentanan seorang mahasiswa terhadap risiko putus studi tidak dipicu oleh faktor tunggal, melainkan merupakan akumulasi dari serangkaian variabel yang saling berkaitan. Variabel tersebut mencakup capaian prestasi akademik (seperti indeks prestasi kumulatif), stabilitas kondisi ekonomi keluarga, tingkat motivasi internal, hingga dinamika psikososial yang dialami mahasiswa sepanjang proses pembelajaran. (Ulinnuha et al., 2023) menunjukkan bahwa data akademik dapat digunakan untuk mengklasifikasikan status *Dropout* mahasiswa menggunakan algoritma *Naïve Bayes*. Sementara itu, (Tumbilung & Efendi, 2025) memanfaatkan data akademik seperti IPK, beban mata kuliah, kehadiran, dan tingkat kegagalan untuk memprediksi risiko *Dropout* mahasiswa.

Perkembangan machine learning memberikan peluang untuk melakukan deteksi dini berdasarkan pola yang terdapat pada data mahasiswa. (Alban & Mauricio, 2019) menerapkan jaringan saraf untuk memprediksi *Dropout* pada mahasiswa, sedangkan (Sutoyo & Almaarif, 2020) menerapkan algoritma *Naïve Bayes* dalam educational data mining untuk memprediksi kelulusan mahasiswa. Penelitian (Ratniasih, n.d.), (Ramadhany et al., 2026), (Abdah Syakiroh Gustian & Fathoni Mahardika, 2025), (Arora, 2023), dan (Satya, 2026) juga menunjukkan penerapan algoritma klasifikasi untuk menentukan mahasiswa yang berpotensi mengalami *Dropout*. Penelitian-penelitian tersebut menunjukkan bahwa data mahasiswa dapat dimanfaatkan sebagai dasar pembangunan model prediktif dalam bidang pendidikan.

Meskipun demikian, sebagian besar penelitian tersebut masih menggunakan data yang bersifat terstruktur, terutama data akademik. Data tersebut belum sepenuhnya menggambarkan pengalaman, kendala, dan motivasi mahasiswa yang dapat disampaikan melalui jawaban terbuka. Dalam konteks ini, analisis sentimen dapat digunakan untuk memperoleh informasi dari data tekstual. (Koto & Rahmanytyas, 2017) mengembangkan Indonesian Sentiment Lexicon (InSet) yang dapat digunakan untuk memberikan skor sentimen pada kata-kata berbahasa Indonesia. Sementara itu, penelitian (Daulay et al., 2026) dan (Syachdini et al., 2026) menunjukkan penerapan text preprocessing, pelabelan sentimen, dan algoritma machine learning dalam mengklasifikasikan data teks.

Dalam mengolah data berstrukturnya sekuensial seperti respons teks, pendekatan *Natural Language Processing* (NLP) konvensional kerap kali gagal menangkap konteks jangka panjang. Untuk mengatasi keterbatasan tersebut, Hochreiter dan Schmidhuber (1997) mengembangkan arsitektur *Long Short-Term Memory* (LSTM) sebagai evolusi dari *Recurrent Neural Network* (RNN). LSTM dirancang khusus dengan mekanisme gerbang (*gating mechanism*) yang mumpuni untuk menyimpan, memperbarui, serta mengingat informasi penting dalam deret urutan data tanpa mengalami masalah *vanishing gradient*. Kemampuan ini menjadikan LSTM sangat efektif untuk mengekstrak makna implisit dan polaritas sentimen dari ungkapan kualitatif yang dituliskan oleh mahasiswa. (Yunusa Umar et al., 2025) memanfaatkan LSTM untuk mempelajari pola data mahasiswa dalam prediksi *Dropout*. Selain itu, penelitian yang dirangkum dalam laporan penelitian ini menunjukkan penggunaan LSTM dalam klasifikasi sentimen dan pengolahan teks.

Berdasarkan penelitian terdahulu, klasifikasi risiko *Dropout* masih banyak berfokus pada data akademik, sedangkan analisis sentimen lebih banyak digunakan untuk mengklasifikasikan opini atau data teks. (Tumbilung & Efendi, 2025), misalnya, membangun model risiko *Dropout* berdasarkan data akademik, sedangkan penelitian analisis sentimen yang dirangkum dalam (PRAYOGA et al., 2026) berfokus pada perbandingan metode klasifikasi teks seperti *Naïve Bayes*, SVM, Random Forest, LSTM, CNN, dan IndoBERT. Dengan demikian, terdapat ruang penelitian untuk mengintegrasikan informasi sentimen dari jawaban terbuka mahasiswa dengan data akademik dan faktor lainnya dalam klasifikasi risiko *Dropout*.

Berdasarkan kondisi tersebut, penelitian ini mengembangkan model klasifikasi risiko *Dropout* mahasiswa menggunakan LSTM dan *Naïve Bayes*. Jawaban terbuka mahasiswa diproses melalui tahap preprocessing dan pelabelan sentimen menggunakan InSet, kemudian diklasifikasikan dengan LSTM. Hasil klasifikasi dari tiga respons setiap mahasiswa digabungkan

menggunakan majority voting untuk memperoleh sentimen akhir. Selanjutnya, sentimen akhir diintegrasikan dengan data akademik, faktor akademik, faktor ekonomi, dan motivasi belajar sebagai masukan algoritma *Naïve Bayes* untuk mengklasifikasikan risiko *Dropout* ke dalam kategori rendah, sedang, dan tinggi. Penelitian ini bertujuan mengembangkan model tersebut serta mengevaluasi kinerjanya menggunakan accuracy, precision, recall, dan F1-score.

METODE PENELITIAN

Penelitian ini menerapkan pendekatan kuantitatif dengan memanfaatkan metode klasifikasi berbasis *machine learning* untuk memodelkan tingkat risiko *Dropout* mahasiswa. Kebaruan pendekatan ini terletak pada pengintegrasian analisis sentimen dari data teks kualitatif berupa respons esai terbuka mahasiswa menggunakan arsitektur *Long Short-Term Memory* (LSTM), yang kemudian dikombinasikan dengan algoritma *Naïve Bayes* untuk klasifikasi risiko akhir. Data primer diperoleh melalui penyebaran kuesioner yang merekam atribut kuantitatif mencakup performa akademik, indikator faktor pendukung akademik, kondisi sosio-ekonomi, dan tingkat motivasi belajar, serta tiga domain pertanyaan terbuka.

Pengumpulan Data

Data primer dalam penelitian ini dikumpulkan melalui penyebaran kuesioner secara daring memanfaatkan platform *Google Forms* yang ditujukan kepada mahasiswa aktif di STMIK Kaputama. Penentuan sampel penelitian menerapkan teknik *purposive sampling*, dengan kriteria inklusi berupa mahasiswa aktif yang bersedia melengkapi seluruh butir pertanyaan instrumen secara utuh. Instrumen kuesioner dirancang untuk menjangkau informasi multidimensi, meliputi indikator akademis, faktor pendukung proses pembelajaran, kondisi sosio-ekonomi, tingkat motivasi belajar, serta tiga pertanyaan terbuka yang memfasilitasi ekspresi kualitatif responden sebagai sumber data teks analisis sentimen.

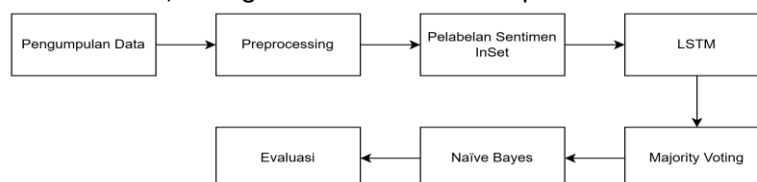
Tabel 1 Variabel Penelitian

Kelompok Data	Variabel
Data akademik	Semester, IPK
Faktor akademik	Faktor/permasalahan akademik mahasiswa
Faktor ekonomi	Kondisi ekonomi mahasiswa
Motivasi	Motivasi belajar
Data tekstual	Tiga jawaban esai mahasiswa
Output sentimen	Positif, netral, negatif
Output akhir	Risiko rendah, sedang, tinggi

Model klasifikasi risiko pada penelitian ini menggunakan semester, IPK, faktor akademik, faktor ekonomi, motivasi belajar, dan sentimen akhir sebagai atribut masukan. Kategori keluaran terdiri atas risiko rendah, risiko sedang, dan risiko tinggi.

Tahapan Penelitian

Secara umum, alur pelaksanaan penelitian ini terbagi ke dalam enam tahapan utama yang terintegrasi secara sistematis, sebagaimana diilustrasikan pada Gambar 2.1.



Gambar 1 Alur Penelitian

Tahapan tersebut dirancang untuk menghubungkan data tekstual mahasiswa dengan data terstruktur dalam proses klasifikasi risiko *Dropout*. Konsep integrasi tersebut berbeda dari penelitian yang hanya menggunakan data akademik sebagai masukan model *Dropout*. Misalnya, (Tumbilung & Efendi, 2025) menggunakan IPK, beban mata kuliah, kehadiran, dan tingkat

kegagalan sebagai data prediksi risiko *Dropout*, sedangkan penelitian ini menambahkan informasi sentimen yang diperoleh dari jawaban terbuka mahasiswa.

Preprocessing Data Teks

Tahap pra-pemrosesan data tekstual dilakukan untuk membersihkan, mereduksi *noise*, serta menyeragamkan struktur kalimat dari respons terbuka mahasiswa agar siap diolah oleh algoritma pembelajaran mesin. Rangkaian tahapan pra-pemrosesan yang diterapkan meliputi:

1. *Cleaning*: Mereduksi karakter yang tidak relevan dengan analisis, seperti simbol, angka, tanda baca, emotikon, serta spasi berlebih.
2. *Case Folding*: Mengubah seluruh huruf dalam dokumen menjadi format huruf kecil (*lowercase*) untuk meminimalkan variasi penulisan kata yang sama.
3. *Tokenizing*: Memotong rangkaian kalimat menjadi unit-unit kata tunggal (token) sebagai dasar analisis.
4. *Stopword Removal*: Mengeliminasi kata-kata umum (*stop words*) yang sering muncul namun tidak memiliki bobot informatif atau kontribusi signifikan dalam penentuan polaritas sentimen (seperti kata hubung dan kata depan).
5. *Stemming*: Memetakan kata-kata berimbuhan kembali ke bentuk kata dasarnya (*root word*) untuk mereduksi kompleksitas kosakata.

Implementasi seluruh rangkaian pra-pemrosesan tersebut menghasilkan kumpulan data teks yang lebih terstruktur, homogen, dan bebas dari informasi redundan, sehingga siap digunakan untuk tahap pelabelan sentimen maupun pelatihan model *Long Short-Term Memory* (LSTM). Laporan penelitian secara eksplisit mencantumkan kelima tahapan tersebut sebagai preprocessing yang digunakan. Penggunaan preprocessing tersebut juga sesuai dengan penelitian klasifikasi sentimen berbahasa Indonesia oleh (PRAYOGA et al., 2026), yang menerapkan tahapan text preprocessing sebelum membandingkan pendekatan berbasis leksikon, *Naïve Bayes*, dan LSTM.

Pelabelan Sentimen Menggunakan InSet

Setelah proses preprocessing, setiap jawaban esai diberikan label sentimen menggunakan Indonesian Sentiment Lexicon (InSet). InSet digunakan sebagai sumber pelabelan awal untuk menentukan kecenderungan sentimen dalam teks berbahasa Indonesia.

Pada tahap ini, setiap dokumen respons esai dikategorikan ke dalam tiga kelas polaritas sentimen, yaitu Positif, Netral, dan Negatif. Proses pelabelan otomatis ini memanfaatkan kamus *Indonesian Sentiment Lexicon* (InSet) dengan menghitung akumulasi bobot valensi dari setiap kata yang terkandung dalam teks. Label polaritas yang dihasilkan oleh InSet selanjutnya dialokasikan sebagai label target (*ground truth*) dalam proses pelatihan dan pengujian model *Long Short-Term Memory* (LSTM). Menerapkan pendekatan berbasis leksikon untuk anotasi data awal memungkinkan pembentukan label target secara objektif dan konsisten sebelum dilakukan klasifikasi mendalam menggunakan arsitektur *deep learning*.

Pendekatan berbasis leksikon juga relevan dengan penelitian (PRAYOGA et al., 2026), yang menggunakan InSet sebagai salah satu pendekatan dalam klasifikasi sentimen bahasa Indonesia sebelum dibandingkan dengan model *Naïve Bayes* dan LSTM.

Klasifikasi Sentimen Menggunakan LSTM

Kumpulan data teks yang telah dianotasi dengan label sentimen selanjutnya dimanfaatkan untuk melatih arsitektur *Long Short-Term Memory* (LSTM). Pemilihan algoritma ini didasari oleh keunggulannya dalam mengolah data berstrukturnya sekuensial, sehingga mampu menangkap dependensi kontekstual dan hubungan antarkata dalam kalimat secara presisi. Arsitektur LSTM yang pertama kali dikembangkan oleh Hochreiter dan Schmidhuber ini memiliki mekanisme pemodelan memori yang secara efektif mempertahankan informasi krusial dalam deret urutan data tanpa kehilangan konteks awal (Nisa et al., 2026).

Sebelum diumpankan ke dalam jaringan LSTM, seluruh teks hasil pra-pemrosesan ditransformasikan ke dalam representasi numerik. Tahapan ini diawali dengan *tokenization* untuk memetakan kata ke dalam indeks angka, diikuti dengan teknik *padding* guna diselaraskan ke dalam panjang sekuens (*sequence length*) yang seragam. Vektor numerik yang dihasilkan kemudian digunakan sebagai masukan (*input layer*) pada proses pelatihan model untuk memprediksi tiga kategori polaritas sentimen, yaitu positif, netral, dan negatif.

Penerapan LSTM dalam ranah prediksi kelangsungan studi juga didukung oleh temuan (Yunusa Umar et al., 2025), yang membuktikan keandalan LSTM dalam memelajari serta

mengekstrak pola sekuensial dari data karakteristik mahasiswa untuk mengidentifikasi risiko *Dropout*. Hal ini memperkuat justifikasi penggunaan LSTM untuk menangkap aspek emosional dan psikososial mahasiswa melalui teks kualitatif secara akurat.

Penentuan Sentimen Akhir dengan Majority Voting

Setiap responden memberikan tiga jawaban esai sehingga satu responden memperoleh tiga hasil prediksi sentimen dari model LSTM. Ketiga hasil tersebut kemudian digabungkan menggunakan metode majority voting untuk memperoleh satu nilai sentimen akhir.

Secara matematis, sentimen akhir dapat dinyatakan sebagai:

$$S_{akhir} = mode(S_1, S_2, S_3)$$

dengan:

S_{akhir} = sentimen akhir responden;

S₁ = hasil prediksi sentimen jawaban pertama;

S₂ = hasil prediksi sentimen jawaban kedua;

S₃ = hasil prediksi sentimen jawaban ketiga.

Dengan pendekatan tersebut, hasil analisis tiga jawaban tidak digunakan secara terpisah pada tahap klasifikasi risiko, tetapi diringkas menjadi satu atribut Sentimen Akhir.

Klasifikasi Risiko *Dropout* Menggunakan *Naïve Bayes*

Sentimen akhir yang diperoleh dari proses majority voting kemudian digabungkan dengan data terstruktur mahasiswa sebagai atribut masukan bagi algoritma *Naïve Bayes*. Atribut yang digunakan meliputi: Semester; IPK; Faktor akademik; Faktor ekonomi; Motivasi belajar; dan Sentimen akhir. Pemodelan ini mengelompokkan tingkat kerentanan putus studi mahasiswa ke dalam tiga kelas target, yang terdiri atas risiko rendah, risiko sedang, serta risiko tinggi.

Prinsip kerja pengklasifikasi *Naïve Bayes* berakar pada Teorema Bayes, yang secara komputasional dirumuskan sebagai berikut:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)}$$

dengan C sebagai kelas risiko *Dropout* dan X sebagai sekumpulan atribut mahasiswa.

Penggunaan *Naïve Bayes* untuk klasifikasi risiko *Dropout* memiliki dasar dari penelitian terdahulu. (Salmawati et al., 2022), misalnya, menggunakan *Naïve Bayes* untuk mengklasifikasikan mahasiswa berpotensi *Dropout* berdasarkan beberapa atribut akademik dan non-akademik. Sementara itu, (Tumbilung & Efendi, 2025) juga menggunakan *Naïve Bayes* sebagai salah satu model untuk prediksi risiko *Dropout* berdasarkan data akademik.

Evaluasi Model

Pengujian performa dilakukan secara terpisah pada dua tahapan pemodelan utama, yakni evaluasi arsitektur LSTM dalam memprediksi sentimen serta pengujian algoritma *Naïve Bayes* dalam mengklasifikasikan tingkat risiko *Dropout*. Evaluasi secara komprehensif mengukur tingkat keandalan sistem memanfaatkan instrumen *confusion matrix* melalui parameter akurasi (*accuracy*), presisi (*precision*), daya ingat (*recall*), dan *F1-score*.

Persamaan yang digunakan adalah:

Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision

$$Precision = \frac{TP}{TP + FP}$$

Recall

$$Recall = \frac{TP}{TP + FN}$$

F1-Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Pada klasifikasi multikelas, nilai metrik dapat dihitung untuk masing-masing kelas dan kemudian diringkas menggunakan metode rata-rata yang diterapkan pada implementasi model.

HASIL DAN PEMBAHASAN

Hasil Preprocessing Data

Tahap pra-pemrosesan menghasilkan kumpulan data teks yang telah dibersihkan melalui tahapan *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Implementasi seluruh rangkaian proses tersebut secara signifikan mereduksi *noise* serta meningkatkan ringkasan dan keterstrukturkan respons esai mahasiswa, sehingga data siap diumpankan ke tahap pelabelan sentimen dan pelatihan model. Ilustrasi perubahan struktur teks sebelum dan sesudah pra-pemrosesan disajikan pada Tabel 2.

Tabel 2 Contoh Hasil Preprocessing Data

Teks Asli	Hasil Preprocessing
Di Pembayaran UKT	bayar ukt
tugas deadline dari dosen	tugas deadline dosen
kendala nya hanya di bagian pengelompokan tim atau tugas berkelompok	kendala nya kelompok tim tugas kelompok
internet wifi untuk belajar	internet wifi ajar
Hambatan utama yang dialami terletak pada manajemen waktu antara kewajiban akademik, penyelesaian tugas, dan aktivitas non-akademik. Penumpukan tenggat waktu tugas yang berdekatan secara signifikan meningkatkan beban kerja dan memicu kondisi <i>academic burnout</i> atau rasa kewalahan.	kendala besar hadap bagi kuliah tugas giat terkadang tugas sama kewalahan

Contoh tersebut menunjukkan bahwa proses preprocessing mengurangi kata atau karakter yang tidak diperlukan serta mengubah kata ke bentuk yang lebih sesuai untuk proses klasifikasi.

Hasil Pelabelan Sentimen

Pelabelan menggunakan InSet menghasilkan 132 dokumen teks dari 44 responden. Distribusi hasil pelabelan ditunjukkan pada Tabel 3.

Tabel 3 Distribusi Hasil Pelabelan Sentimen

Label Sentimen	Jumlah Data	Persentase (%)
Positif	77	58,33
Netral	34	25,76
Negatif	21	15,91
Total	132	100

Berdasarkan Tabel 3, sentimen positif merupakan kategori terbanyak, yaitu 77 data (58,33%), diikuti sentimen netral sebanyak 34 data (25,76%), dan sentimen negatif sebanyak 21 data (15,91%). Distribusi tersebut menunjukkan bahwa data sentimen tidak seimbang, dengan jumlah data positif lebih besar dibandingkan dua kelas lainnya.

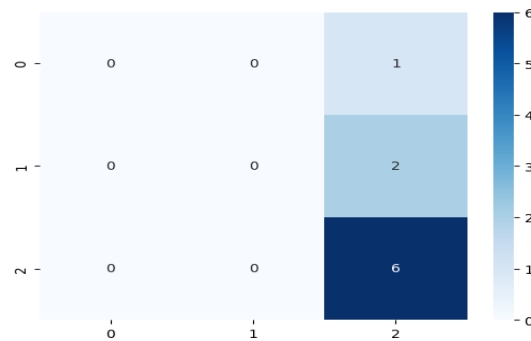
Kondisi distribusi data tersebut perlu diperhatikan ketika menginterpretasikan hasil LSTM karena jumlah data pada masing-masing kelas tidak sama. Penelitian (PRAYOGA et al., 2026) juga menunjukkan pentingnya evaluasi beberapa metrik, bukan hanya accuracy, dalam klasifikasi sentimen multikelas.

Hasil Klasifikasi Sentimen Menggunakan LSTM

Arsitektur *Long Short-Term Memory* (LSTM) diimplementasikan untuk memprediksi tiga kelas polaritas sentimen, yaitu positif, netral, dan negatif. Capaian performa dan metrik evaluasi dari pemodelan LSTM tersebut dirangkum pada Tabel 4.

Tabel 4 Hasil Evaluasi Model LSTM

	Precision	Recall	F1-Score	Support
Negatif	0	0	0	1
Netral	0	0	0	2
Positif	0,67	1	0,8	6
Accuracy	0,67	-	-	9
Macro Average	0,22	0,33	0,27	9
Weighted Average	0,44	0,67	0,53	9



Gambar 2 Confusion Matrix model LSTM

Pengujian model LSTM menghasilkan akurasi sebesar 67%. Berdasarkan evaluasi per kelas, model memperoleh precision sebesar 0,67, recall 1,00, dan F1-score 0,80 pada kelas positif. Sementara itu, kelas negatif dan netral belum dapat diprediksi pada data pengujian.

Hasil tersebut menunjukkan bahwa kinerja klasifikasi dipengaruhi oleh ketidakseimbangan distribusi kelas dan keterbatasan jumlah data. Kelas negatif dan netral memiliki jumlah sampel pengujian yang lebih sedikit sehingga representasinya dalam proses klasifikasi terbatas.

Temuan ini berbeda dengan (PRAYOGA et al., 2026), yang memperoleh performa LSTM lebih tinggi pada dataset komentar YouTube yang jauh lebih besar. Perbedaan tersebut perlu dipahami berdasarkan karakteristik dataset, bukan semata-mata berdasarkan algoritmanya. Penelitian tersebut menggunakan ribuan komentar, sedangkan penelitian ini hanya menggunakan 132 dokumen teks dengan data pengujian yang jauh lebih kecil.

Penelitian (Yunusa Umar et al., 2025) juga memperoleh performa LSTM yang lebih tinggi pada prediksi *Dropout*, tetapi penelitian tersebut menggunakan dataset akademik dengan jumlah dan karakteristik data yang berbeda.

Hasil Majority Voting

Setiap responden memiliki tiga jawaban esai yang menghasilkan tiga prediksi sentimen dari LSTM. Ketiga prediksi tersebut kemudian digabungkan menggunakan majority voting untuk memperoleh satu sentimen akhir pada tingkat responden. Contoh mekanisme penentuan sentimen akhir disajikan pada Tabel 5.

Tabel 5 Contoh Penentuan Sentimen Akhir Menggunakan Majority Voting

Prediksi Esai 1	Prediksi Esai 2	Prediksi Esai 3	Sentimen Akhir
Positif	Positif	Positif	Positif
Positif	Positif	Netral	Positif
Netral	Netral	Positif	Netral
Negatif	Negatif	Netral	Negatif
Negatif	Positif	Negatif	Negatif

Sumber: Diolah penulis, 2026

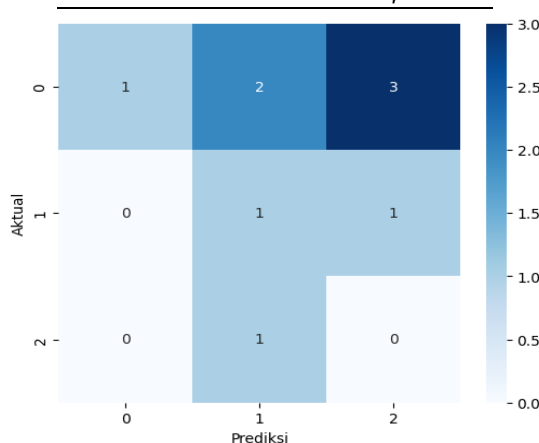
Sentimen akhir tersebut selanjutnya digunakan sebagai salah satu atribut dalam klasifikasi risiko *Dropout* menggunakan *Naïve Bayes*.

Hasil Klasifikasi Risiko *Dropout* Menggunakan *Naïve Bayes*

Algoritma *Naïve Bayes* diimplementasikan untuk memetakan tingkat risiko *Dropout* mahasiswa berdasarkan integrasi fitur yang mencakup semester, IPK, indikator faktor akademik, kondisi faktor ekonomi, tingkat motivasi, serta nilai Sentimen Akhir. Melalui pemodelan ini, kerentanan mahasiswa diklasifikasikan ke dalam tiga kategori risiko, yakni risiko rendah, sedang, dan tinggi.

Tabel 6 Hasil Evaluasi Model *Naïve Bayes*

Metrik Evaluasi	Nilai
Accuracy	22,22%
Precision	72,22%
Recall	22,22%
F1-Score	26,46%



Gambar 3 Confusion Matrix Model *Naïve Bayes*

Hasil pengujian menunjukkan bahwa model *Naïve Bayes* memperoleh akurasi sebesar 22,22%, precision 72,22%, recall 22,22%, dan F1-score 26,46%. Nilai precision yang lebih tinggi dibandingkan recall menunjukkan bahwa prediksi yang dihasilkan cukup tepat pada sebagian prediksi yang dibuat, tetapi model masih terbatas dalam mengenali sampel dari kelas yang sebenarnya.

Rendahnya accuracy dan recall perlu dilihat dalam konteks ukuran dan distribusi data penelitian. Jumlah responden hanya 44 orang dan keluaran risiko terdiri atas tiga kelas. Dengan jumlah data yang terbatas, pembentukan pola probabilistik pada *Naïve Bayes* menjadi lebih rentan terhadap distribusi data yang tidak merata.

Hasil tersebut berbeda dengan (Salmawati et al., 2022), yang memperoleh accuracy *Naïve Bayes* sebesar 97,83% pada 1.155 data mahasiswa. Perbedaan tersebut menunjukkan bahwa performa model tidak dapat dibandingkan secara langsung hanya berdasarkan algoritma, karena ukuran dataset, jumlah kelas, atribut, dan karakteristik data yang digunakan berbeda.

(Tumbilung & Efendi, 2025) juga memperoleh accuracy *Naïve Bayes* sebesar 95% pada penelitian prediksi risiko *Dropout* berbasis data akademik. Penelitian tersebut menggunakan dataset dan karakteristik variabel yang berbeda dari penelitian ini.

Pembahasan

Hasil penelitian menunjukkan bahwa integrasi analisis sentimen menggunakan Long Short-Term Memory (LSTM) dengan algoritma *Naïve Bayes* dapat diterapkan untuk menghubungkan informasi tekstual mahasiswa dengan data akademik dan faktor pendukung lainnya dalam klasifikasi risiko *Dropout*. Model LSTM menghasilkan akurasi sebesar 67% dalam klasifikasi sentimen, sedangkan model *Naïve Bayes* yang menggunakan hasil sentimen akhir bersama atribut terstruktur menghasilkan accuracy 22,22%, precision 72,22%, recall 22,22%, dan F1-score 26,46%. Hasil tersebut menunjukkan bahwa kendala utama penelitian berada pada tahap klasifikasi risiko *Dropout*, bukan pada proses integrasi kedua metode.

Rendahnya kinerja *Naïve Bayes* perlu dipahami berdasarkan karakteristik dataset yang digunakan. Penelitian ini hanya melibatkan 44 responden, dengan tiga kategori keluaran risiko *Dropout*, yaitu rendah, sedang, dan tinggi. Jumlah data tersebut relatif terbatas untuk

membentuk pola probabilistik yang stabil pada beberapa kategori risiko. Pada *Naïve Bayes*, keputusan klasifikasi ditentukan berdasarkan probabilitas kelas dan probabilitas atribut terhadap masing-masing kelas. Dengan jumlah observasi yang sedikit, distribusi setiap atribut pada masing-masing kategori risiko menjadi lebih mudah dipengaruhi oleh karakteristik individual responden. Akibatnya, probabilitas yang terbentuk dapat kurang merepresentasikan pola umum mahasiswa dan menyebabkan model kesulitan membedakan ketiga kategori risiko secara konsisten.

Selain ukuran dataset, distribusi kelas juga menjadi faktor yang perlu diperhatikan. Pada tahap analisis sentimen, dari 132 dokumen teks yang berasal dari 44 responden, kelas positif berjumlah 77 dokumen (58,33%), sedangkan kelas negatif hanya 21 dokumen (15,91%) dan kelas netral 34 dokumen (25,76%). Dominasi kelas tertentu menunjukkan bahwa informasi yang tersedia untuk mempelajari karakteristik kelas minoritas lebih terbatas. Kondisi tersebut juga berpengaruh terhadap fitur Sentimen Akhir yang kemudian digunakan sebagai salah satu atribut pada model *Naïve Bayes*. Dengan demikian, informasi sentimen yang menjadi bagian dari masukan klasifikasi risiko belum sepenuhnya mampu memberikan representasi yang seimbang untuk setiap kategori. Perbedaan kinerja dengan penelitian terdahulu juga memperlihatkan pentingnya karakteristik dataset dalam menentukan performa *Naïve Bayes*. Misalnya, Salmawati et al. (2022) menggunakan 1.155 data mahasiswa, sedangkan Prayoga et al. (2026) menggunakan ribuan data komentar untuk klasifikasi sentimen.

Nilai *precision* sebesar 72,22% yang lebih tinggi dibandingkan *recall* sebesar 22,22% juga memberikan informasi penting mengenai perilaku model. Hasil tersebut menunjukkan bahwa pada prediksi yang dibuat oleh model, terdapat prediksi yang relatif tepat, tetapi model belum mampu mengenali sebagian besar sampel yang sebenarnya termasuk dalam kategori tertentu. Dengan kata lain, model masih memiliki keterbatasan dalam cakupan pengenalan kelas. Kondisi ini sejalan dengan nilai *F1-score* yang hanya mencapai 26,46%, sehingga tingginya *precision* belum dapat diartikan sebagai kemampuan klasifikasi risiko yang baik secara keseluruhan. Temuan tersebut menunjukkan bahwa integrasi LSTM dan *Naïve Bayes* dalam penelitian ini lebih tepat dipandang sebagai model awal untuk menguji integrasi informasi tekstual dan data terstruktur, bukan sebagai model yang telah memiliki kemampuan prediksi yang stabil. Keterbatasan utama berasal dari jumlah responden yang kecil dan distribusi data yang belum seimbang. Oleh karena itu, peningkatan jumlah dan keragaman responden, penerapan teknik class balancing, serta pengujian pada dataset yang lebih besar diperlukan agar pola hubungan antara sentimen, karakteristik akademik, faktor ekonomi, motivasi, dan risiko *Dropout* dapat dipelajari dengan lebih representatif.

Dari sisi pendekatan, penelitian ini memiliki perbedaan dengan penelitian *Dropout* yang umumnya menggunakan data akademik sebagai masukan utama. Tumbilung dan Efendi (2025), Salmawati et al. (2022), dan Saputra (2026) berfokus pada data akademik atau karakteristik mahasiswa dalam membangun model prediksi *Dropout*. Sementara itu, penelitian ini menambahkan sentimen dari jawaban terbuka mahasiswa sebagai salah satu atribut dalam klasifikasi risiko.

KESIMPULAN

Penelitian ini berhasil mengintegrasikan Long Short-Term Memory (LSTM) dan algoritma *Naïve Bayes* untuk mengklasifikasikan risiko *Dropout* mahasiswa berdasarkan kombinasi data akademik, faktor akademik, faktor ekonomi, motivasi belajar, dan respons teks terbuka. Tahapan analisis sentimen yang meliputi preprocessing, pelabelan menggunakan Indonesian Sentiment Lexicon (InSet), klasifikasi LSTM, serta agregasi majority voting menghasilkan akurasi sebesar 67%. Selanjutnya, sentimen akhir digunakan sebagai salah satu atribut pada klasifikasi risiko *Dropout* menggunakan *Naïve Bayes*. Hasil pengujian *Naïve Bayes* memperoleh accuracy 22,22%, precision 72,22%, recall 22,22%, dan F1-score 26,46%. Hasil tersebut menunjukkan bahwa model masih memiliki keterbatasan dalam mengenali seluruh kategori risiko, terutama karena jumlah responden yang terbatas dan distribusi kelas yang belum seimbang. Oleh karena itu, pengembangan penelitian selanjutnya perlu dilakukan dengan menambah jumlah dan keragaman dataset, menerapkan teknik penyeimbangan kelas (class balancing), serta menguji model pada data yang lebih representatif agar kemampuan model dalam mengenali setiap kategori risiko dapat dievaluasi secara lebih stabil dan mendukung pengembangan sistem deteksi dini *Dropout* mahasiswa.

REFERENSI

- Abdah Syakiroh Gustian, & Fathoni Mahardika. (2025). Analisis Klasifikasi Risiko *Dropout* Mahasiswa Menggunakan Algoritma Decision Tree dan Random Forest. *Jupiter: Publikasi Ilmu Keteknikan Industri, Teknik Elektro Dan Informatika*, 3(4), 182–189. <https://doi.org/10.61132/jupiter.v3i4.980>
- Alban, M., & Mauricio, D. (2019). Neural networks to predict *Dropout* at the universities. *International Journal of Machine Learning and Computing*, 9(2), 149–153. <https://doi.org/10.18178/ijmlc.2019.9.2.779>
- Arora, D. (2023). Predicting Students Academic Success and *Dropout* Using Supervised Machine Learning. In *72 72 International Journal of Scientific Study* (Vol. 11, Number 6). <https://archive.ics.uci.edu/dataset/697/predict+students>
- Daulay, W. A., Buaton, R., Annatasia, K., & Sitepu, B. (2026). *Implementation of Naïve Bayes Algorithm for Sentiment Classification of Public Youtube Opinions Related to Nickel Mining Issues in Raja Ampat*. <https://ioinformatic.org/>
- Koto, F., & Rahmaningtyas, G. Y. (2017). *2017 International Conference on Asian Language Processing*. IEEE.
- Nisa, K., Annatasia Br Sitepu, K., & Kaputama, S. (2026). *Prediction of the Number of Motor Vehicle Inspections using the Long Short-Term Memory (LSTM) Method as a Decision-Support Tool at the Transportation Agency of Langkat Regency*. <https://ioinformatic.org/>
- Prayoga, J., Fajri, T. I., & Dristyan, F. (2026). Sentiment Classification Model Based On Comparative Studies Using Machine Learning Technology. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, 11(3), 755–764. <https://doi.org/10.33480/jitk.v11i3.7105>
- Ramadhany, M. Z., Wati, L., Lunak, R. P., Negeri, P., Jalan, B., Alam, S., Bengkalis, K., Bengkalis, K., Riau, I., Prof, J., Hamka, K., Air Tawar, S., & Barat, I. (2026). Klasifikasi Mahasiswa Berpotensi Drop Out (Do) Menggunakan Algoritma Random Forest. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 10, Number 2).
- Ratniasih, N. L. (n.d.). *Penerapan Algoritma Klasifikasi Untuk Penentuan Mahasiswa Berpotensi Drop Out*.
- Salmawati, Yuyun, & Hazriani. (2022). *Klasifikasi Mahasiswa Berpotensi Drop Out Menggunakan Algoritma Naïve Bayes Dan Decision Tree*. 8(2). <http://ejournal.fikom-unasman.ac.id>
- Satya, S. (2026). Analisis Prediktif *Dropout* Mahasiswa Berdasarkan Kinerja Akademik Semester Awal Menggunakan Machine Learning. *Jurnal Riset Dan Aplikasi Mahasiswa Informatika (JRAMI)*, 07.
- Sutoyo, E., & Almaarif, A. (2020). Educational Data Mining untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritme *Naïve Bayes* Classifier. *Masa Berlaku Mulai*, 1(3), 95–101.
- Syachdini, N., Manurung, H., Annatasia, K., & Sitepu, B. (2026). *Journal of Artificial Intelligence and Engineering Applications Sentiment Analysis of Lau Berte Waterfall Using the Naïve Bayes Algorithm*. <https://ioinformatic.org/>
- Tumbilung, C. F., & Efendi, R. (2025). *Prediksi Risiko Drop Out Mahasiswa Menggunakan Model Machine Learning Berbasis Data Akademik (Studi Kasus : Universitas XYZ)*. 16(2). <http://ejurnal.provisi.ac.id/index.php/JTIKP>
- Ulinnuha, N., Fanani, A., & Sunan Ampel Surabaya, U. (2023). *Klasifikasi Status Drop Out Mahasiswa Menggunakan Naïve Bayes dengan Seleksi Fitur Information Gain Classification of Drop Out Status for Students Using Naïve Bayes with Information Gain Feature Selection* (Vol. 22, Number 4).
- Yunusa Umar, I., Musa, K. I., & Tella, M. (2025). Long Short-Term Memory (LSTM) Based Model for Forecasting of Students' *Dropout*. *International Journal of Scientific Research in Multidisciplinary Studies*, 11(7), 124–135. www.isroset.org