

Klasifikasi Risiko Burnout Akademik Siswa Menggunakan Decision Tree dengan Seleksi Fitur Information Gain

Aditya Aji Sagara^{1*}, Herri Setiawan², Muhammad Haviz Irfani^{3*}

^{1,2,3}Teknik Informatika, Universitas Indo Global Mandiri, Palembang, Indonesia

Email: ^{1*}2021110001@students.uigm.ac.id, ²herri@uigm.ac.id,

Abstrak– Burnout akademik merupakan kondisi kelelahan emosional, fisik, dan mental akibat tekanan akademik yang berkepanjangan yang dapat menurunkan motivasi, prestasi belajar, serta kesejahteraan psikologis siswa. Oleh karena itu, diperlukan suatu metode deteksi dini untuk membantu sekolah mengidentifikasi siswa yang berisiko mengalami burnout akademik. Penelitian ini bertujuan mengklasifikasikan tingkat risiko burnout akademik siswa menggunakan algoritma Decision Tree. Penelitian dilakukan terhadap 101 siswa kelas XI SMA LTI IGM dengan menggunakan data hasil kuesioner yang memuat berbagai indikator burnout akademik. Tahapan penelitian meliputi preprocessing data, seleksi fitur menggunakan Information Gain, pembangunan model Decision Tree, penyeimbangan data menggunakan teknik oversampling, serta evaluasi menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model sebelum proses balancing memperoleh akurasi sebesar 86%, sedangkan setelah dilakukan oversampling performa model meningkat menjadi 94%, dengan nilai precision, recall, dan F1-score masing-masing sebesar 94%. Hasil seleksi fitur juga menunjukkan bahwa indikator keraguan terhadap potensi diri, kecemasan berlebihan, dan kekhawatiran terhadap masa depan merupakan faktor yang paling berpengaruh dalam klasifikasi risiko burnout akademik. Hasil penelitian menunjukkan bahwa penerapan Decision Tree yang dikombinasikan dengan Information Gain dan oversampling mampu meningkatkan performa klasifikasi serta berpotensi digunakan sebagai alat bantu deteksi dini burnout akademik untuk mendukung pengambilan keputusan dan penyusunan intervensi yang lebih tepat di lingkungan sekolah.

Kata Kunci: Information Gain, Feature Selection, Oversampling, Educational Data Mining, Class Imbalance

Abstract– Academic burnout is a condition of emotional, physical, and mental exhaustion due to prolonged academic pressure that can reduce students' motivation, academic achievement, and psychological well-being. Therefore, an early detection method is needed to help schools identify students at risk of academic burnout. This study aims to classify students' academic burnout risk levels using the Decision Tree algorithm. The study was conducted on 101 11th-grade students of SMA LTI IGM using questionnaire data containing various indicators of academic burnout. The research stages include data preprocessing, feature selection using Information Gain, Decision Tree model development, data balancing using oversampling techniques, and evaluation using confusion matrix, accuracy, precision, recall, and F1-score. The results showed that the model before the balancing process achieved an accuracy of 86%, while after oversampling the model's performance increased to 94%, with precision, recall, and F1-score values of 94% each. The feature selection results also showed that indicators of self-doubt, excessive anxiety, and worry about the future were the most influential factors in classifying academic burnout risk. The results of the study indicate that the application of Decision Tree combined with Information Gain and oversampling is able to improve classification performance and has the potential to be used as an early detection tool for academic burnout to support decision making and the development of more appropriate interventions in the school environment.

Keywords: Information Gain, Feature Selection, Oversampling, Educational Data Mining, Class Imbalance

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi yang pesat telah berdampak signifikan terhadap dunia pendidikan, mendorong institusi untuk beradaptasi dalam meningkatkan mutu pembelajaran melalui sistem digital dan otomatisasi [1]. Selain itu, penggunaan sistem digital dan otomatisasi juga telah terbukti mampu meningkatkan efisiensi dalam pengelolaan data serta memperkuat proses pengambilan keputusan di berbagai institusi, termasuk institusi pendidikan [2]. Sejalan dengan Undang-Undang No. 20 Tahun 2004 pasal 4 ayat (1), pendidikan ditujukan untuk membentuk peserta didik yang mandiri dan bertanggung jawab, yang kini didukung melalui penerapan Kurikulum Merdeka. Namun, fleksibilitas kurikulum ini belum sepenuhnya diimbangi dengan kesiapan siswa, sehingga menimbulkan stres hingga risiko burnout [3].

Penelitian ini berfokus pada siswa kelas XI SMA LTI IGM Palembang, yang berada di jenjang kritis dengan beban akademik tinggi. Berdasarkan data sekolah, terdapat 325 siswa dan 34 tenaga

pengajar. Indikasi burnout di kalangan siswa mendorong perlunya penerapan model klasifikasi berbasis machine learning untuk mendeteksi risiko lebih awal.

Penelitian sebelumnya telah menunjukkan bahwa algoritma Decision Tree memiliki performa yang baik dalam klasifikasi berbagai permasalahan psikologis. Penelitian oleh [4] berhasil mengklasifikasikan tingkat depresi mahasiswa berdasarkan skor PHQ-9 dengan akurasi yang tinggi, sedangkan penelitian [5] menerapkan Decision Tree untuk klasifikasi tingkat stres siswa menggunakan atribut seperti kecemasan dan kualitas tidur dengan akurasi sebesar 88,18%. Selain itu, penelitian [6] telah menerapkan teknik pengenalan pola untuk memprediksi burnout pada programmer sebagai upaya deteksi dini. Temuan-temuan tersebut menunjukkan bahwa pendekatan machine learning, khususnya Decision Tree, memiliki potensi yang baik dalam mendukung identifikasi kondisi psikologis.

Meskipun demikian, penelitian terdahulu masih memiliki beberapa keterbatasan. Sebagian besar penelitian difokuskan pada depresi, stres, atau burnout pada mahasiswa maupun pekerja, sehingga penelitian mengenai klasifikasi risiko burnout akademik pada siswa sekolah menengah masih relatif terbatas. Selain itu, penelitian sebelumnya umumnya belum mengintegrasikan proses seleksi fitur menggunakan Information Gain dan penanganan ketidakseimbangan kelas (class imbalance) melalui teknik oversampling dalam pembangunan model klasifikasi. Padahal, distribusi data yang tidak seimbang berpotensi menurunkan kemampuan model dalam mengenali kategori minoritas sehingga memengaruhi performa klasifikasi.

Berdasarkan kesenjangan tersebut, penelitian ini mengusulkan model klasifikasi risiko burnout akademik pada siswa kelas XI SMA LTI IGM Palembang menggunakan algoritma Decision Tree yang dikombinasikan dengan seleksi fitur Information Gain dan teknik oversampling untuk meningkatkan representasi kelas minoritas. Kombinasi metode tersebut diharapkan mampu menghasilkan model klasifikasi yang lebih akurat, lebih seimbang dalam mengenali seluruh kategori risiko burnout, serta lebih mudah diinterpretasikan sebagai dasar pengambilan keputusan. Oleh karena itu, tujuan penelitian ini adalah membangun dan mengevaluasi model klasifikasi risiko burnout akademik siswa menggunakan Decision Tree, Information Gain, dan oversampling sebagai pendekatan deteksi dini yang dapat mendukung penyusunan intervensi secara lebih tepat di lingkungan sekolah.

2. METODOLOGI PENELITIAN

Algoritma Decision Tree merupakan metode klasifikasi yang bekerja dengan membentuk struktur pohon, di mana setiap node merepresentasikan atribut, dan cabang mewakili hasil dari pengujian kondisi tertentu hingga mencapai daun sebagai representasi keputusan akhir, sedangkan klasifikasi merupakan proses dalam menentukan objek data ke dalam kelas[7]. Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk membangun model klasifikasi risiko burnout akademik siswa menggunakan algoritma Decision Tree. Tahapan penelitian meliputi pengumpulan data, pra-pemrosesan data, penyeimbangan data, seleksi fitur, pembangunan model klasifikasi, serta evaluasi performa model. Populasi penelitian adalah seluruh siswa kelas XI SMA LTI IGM Palembang yang berjumlah 325 siswa. Sampel penelitian sebanyak 101 siswa diperoleh menggunakan teknik purposive sampling dengan mempertimbangkan ketersediaan responden dan kelengkapan pengisian kuesioner.

Data penelitian dikumpulkan menggunakan kuesioner berbasis Google Form yang terdiri atas 24 butir pernyataan untuk mengukur indikator burnout akademik, meliputi kelelahan emosional, depersonalisasi, serta penurunan pencapaian diri. Jawaban responden kemudian dikonversi ke dalam bentuk numerik sehingga dapat diproses menggunakan algoritma *machine learning*. Selanjutnya dilakukan proses *preprocessing* yang meliputi pemeriksaan data kosong (*missing value*), penghapusan data duplikat, konversi data kategorikal menjadi numerik, dan perhitungan skor total setiap responden. Berdasarkan distribusi skor tersebut, tingkat risiko burnout diklasifikasikan ke dalam tiga kategori, yaitu rendah, sedang, dan tinggi menggunakan pendekatan kuartil.

Untuk mengatasi permasalahan ketidakseimbangan kelas (*class imbalance*), penelitian ini menerapkan teknik *Random Oversampling* pada kelas minoritas hingga jumlah data setiap kategori menjadi seimbang. Setelah itu dilakukan seleksi fitur menggunakan metode *Information Gain* untuk menentukan atribut yang memiliki kontribusi terbesar terhadap proses klasifikasi. Nilai *Information Gain* dihitung berdasarkan pengurangan nilai entropi awal terhadap entropi setelah data dipartisi menggunakan suatu atribut sebagaimana ditunjukkan pada Persamaan (1) Information Gain.

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} \cdot Entropy(S_v) \quad (1)$$

Keterangan:

S adalah himpunan kasus atau data awal,

A adalah atribut yang sedang diuji,

Values(A) adalah himpunan nilai-nilai unik dari atribut A,

S_v adalah subset dari S yang memiliki nilai v untuk atribut A,

$|S_v|$ adalah jumlah data dalam subset S_v ,

$|S|$ adalah jumlah data total dalam S.

Persamaan (2) *Entropy*.

$$Entropy(S) = - \sum_{i=1}^n P_i \cdot \log_2(P_i) \quad (2)$$

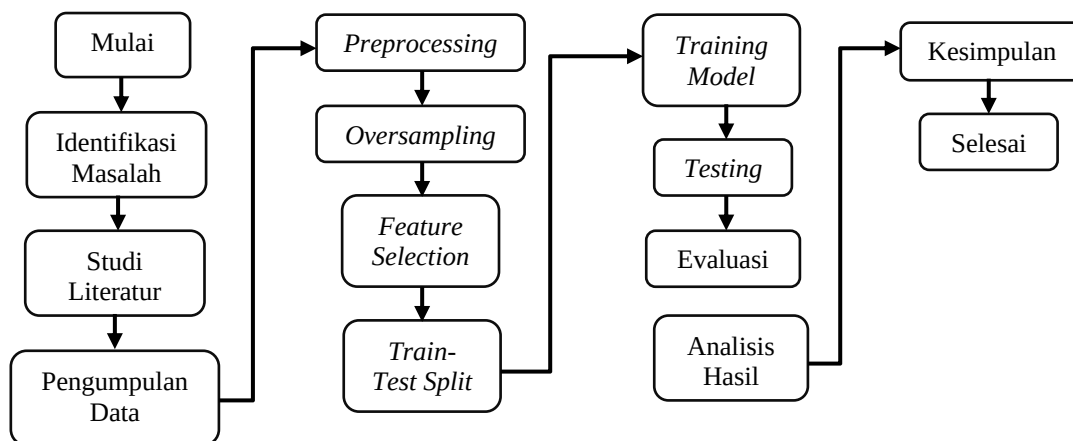
Keterangan:

S adalah himpunan data (kasus) yang dianalisis,

n adalah jumlah kelas yang terdapat dalam himpunan data tersebut,

Pi adalah proporsi dari jumlah data di kelas ke-i terhadap seluruh data dalam S.

1. Pengumpulan data dilakukan melalui [10] kuesioner berbasis Google Form yang disebarakan kepada siswa kelas [11] XI SMA LTI IGM Palembang. Kuesioner berisi 24 butir pernyataan yang mengukur berbagai dimensi burnout akademik, seperti kelelahan emosional, depersonalisasi, dan penurunan pencapaian diri.
2. Tahap preprocessing data mencakup pemeriksaan data kosong dan duplikat, konversi jawaban menjadi format numerik (0 dan 1), dan perhitungan skor total dari seluruh butir pertanyaan untuk masing-masing responden guna dijadikan dasar klasifikasi.
3. Penentuan kategori risiko burnout dilakukan dengan mengelompokkan skor total menjadi tiga kategori, yaitu Rendah, Sedang, dan Tinggi. Batasan tiap kategori ditentukan berdasarkan nilai kuartil, di mana kuartil pertama (Q1) dan kuartil ketiga (Q3) digunakan sebagai ambang klasifikasi.
4. Untuk menangani ketidakseimbangan kelas dalam data, dilakukan teknik random oversampling pada kategori Sedang dan Tinggi agar jumlah data sebanding dengan kategori mayoritas yaitu Rendah. Teknik ini dilakukan dengan menduplikasi data dari kelas minoritas secara acak (sampling with replacement) hingga distribusi antar kategori menjadi proporsional.
5. Seleksi fitur dilakukan dengan metode Information Gain untuk menentukan atribut-atribut yang paling berpengaruh dalam proses klasifikasi. Sebanyak 15 fitur dengan nilai Information Gain tertinggi dipilih sebagai input dalam pembangunan model klasifikasi.
6. Model klasifikasi dibangun menggunakan algoritma Decision Tree dengan memanfaatkan platform Google Colab. Proses pelatihan model dilakukan dengan membagi dataset menjadi 90% data latih dan 10% data uji untuk memperoleh hasil prediksi.
7. Evaluasi model dilakukan dengan menggunakan [12] confusion matrix dan metrik evaluasi lainnya seperti akurasi, precision, recall, dan f1-score. Performa model [13] dievaluasi baik sebelum maupun sesudah proses balancing untuk melihat perbandingan dan efektivitas model dalam mengklasifikasikan risiko burnout secara lebih adil.



Gambar 1. Diagram Tahapan Penelitian

2.1 Pra-pemrosesan Data

Tahap awal setelah data diperoleh adalah proses pembersihan dan normalisasi. Kolom yang tidak relevan seperti nama, email, umur, dan timestamp dihapus, kemudian seluruh jawaban kuesioner yang semula dalam bentuk kategorikal ("Ya" dan "Tidak") dikonversi menjadi format numerik (1 dan 0) agar dapat diproses oleh algoritma klasifikasi. Selanjutnya, dilakukan perhitungan total skor untuk setiap responden berdasarkan jawaban yang diberikan. Distribusi skor kemudian diklasifikasikan ke dalam tiga kategori risiko burnout menggunakan pendekatan kuartil, yaitu: rendah (di bawah Q1), sedang (antara Q1 dan Q3), dan tinggi (di atas Q3). Penggunaan kuartil sebagai dasar klasifikasi bertujuan untuk menghasilkan pembagian data yang proporsional dan objektif sesuai prinsip distribusi statistik.

2.2 Seleksi Fitur

Seleksi fitur dilakukan untuk memilih atribut yang paling berpengaruh dalam menentukan risiko burnout akademik. [14] Metode Information Gain digunakan untuk mengukur kontribusi masing-masing fitur terhadap klasifikasi. Proses ini diawali dengan perhitungan nilai entropy dan Information Gain pada setiap fitur untuk menilai tingkat relevansi terhadap target klasifikasi. Dari hasil perhitungan, dipilih 15 fitur terbaik dengan nilai Information Gain tertinggi. Untuk memudahkan proses visualisasi model klasifikasi, fitur-fitur tersebut kemudian disederhanakan dengan menggunakan penamaan ulang (misalnya: P1, P2, ..., P15). Dataset yang telah terpilih inilah yang kemudian digunakan pada tahap pelatihan model Decision Tree (DT). Dapat dilihat pada Gambar 2. Hasil dari seleksi fitur.

```
15 Fitur Terbaik dengan Kode Singkat dan Nilai Information Gain-nya:
P1 : Apakah kamu ragu dengan potensi atau kemampuanmu sendiri? --> IG = 0.4512
P2 : Apakah kamu takut gagal sampai sulit tidur? --> IG = 0.3827
P3 : Apakah kamu terlalu khawatir tentang masa depan akademimu? --> IG = 0.3768
P4 : Apakah kamu sering merasa tidak sehebat teman-temanmu? --> IG = 0.3640
P5 : Apakah kamu merasa kehilangan semangat untuk melakukan aktivitas sehari-hari? --> IG = 0.3496
P6 : Apakah kamu menghabiskan banyak waktu memikirkan hal-hal yang belum tentu terjadi? --> IG = 0.3426
P7 : Apakah kamu kesulitan mengingat materi yang sudah kamu pelajari? --> IG = 0.2941
P8 : Apakah Anda mengalami kesulitan tidur karena pikiran yang terus-menerus tentang sekolah atau belajar? --> IG = 0.2866
P9 : Apakah Anda merasa bahwa Anda terlalu terbebani oleh tugas atau pelajaran? --> IG = 0.2824
P10 : Apakah kamu enggan mencoba hal baru karena takut tidak berhasil? --> IG = 0.2717
P11 : Apakah Anda merasa bahwa Anda tidak mampu mengatasi tekanan belajar atau sekolah? --> IG = 0.2555
P12 : Apakah kamu merasa cemas berlebihan jika mendapatkan nilai buruk? --> IG = 0.2477
P13 : Apakah kamu menganggap semua usaha belajar sia-sia? --> IG = 0.2358
P14 : Apakah kamu merasa kelelahan sepanjang waktu meskipun sudah tidur cukup? --> IG = 0.2320
P15 : Apakah kamu sulit membuka diri kepada teman karena takut dihakimi? --> IG = 0.2218
```

Gambar 2. Hasil Dari Seleksi fitur

2.3 Implementasi Algoritma Decision Tree

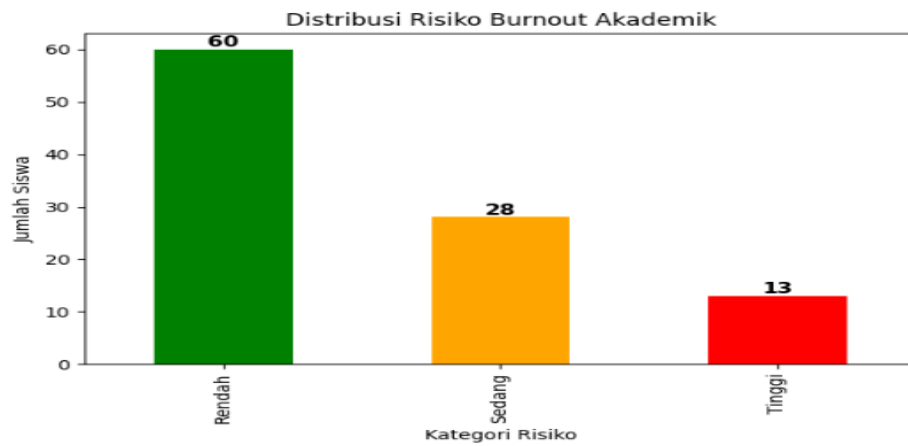
Pada tahap ini, data dibagi menjadi data latih dan data uji dengan rasio awal 90:10. Pengujian juga dilakukan dengan variasi test size mulai dari 10% hingga 50% untuk mengevaluasi performa model pada berbagai skenario pembagian data. Dataset yang digunakan adalah hasil dari tahap seleksi fitur ($x_{selected}$) sebelumnya. Model Decision Tree dibangun dengan menggunakan kriteria entropy untuk pemilihan atribut terbaik berdasarkan nilai Information Gain. Parameter `random_state` diatur agar proses pelatihan konsisten di setiap eksekusi. Setelah pelatihan selesai, model digunakan untuk memprediksi data uji, dan hasilnya dievaluasi berdasarkan metrik akurasi dan performa klasifikasi lainnya.

2.4 Evaluasi Model dan Visualisasi

Evaluasi model dilakukan menggunakan confusion matrix dan metrik klasifikasi seperti akurasi, precision, recall, dan F1-score. Model Decision Tree dievaluasi dengan berbagai rasio test size (10%–50%) untuk melihat pengaruh proporsi data latih terhadap performa model. Hasil terbaik diperoleh pada test size 10%, dengan akurasi 91%, precision dan recall 89%, serta F1-score 87%.

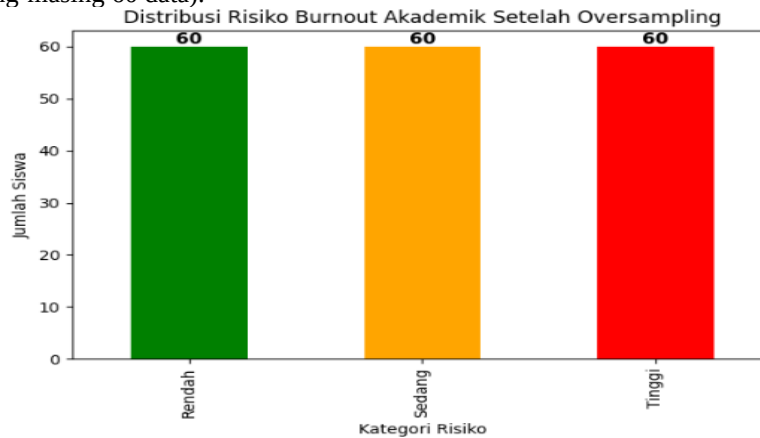
Tabel 1. Hasil Pengujian Model Menggunakan Data Uji

Test Size	Precision	Recall	F1-Score	Akurasi
10%	0.89	0.89	0.87	0.91
20%	0.83	0.87	0.81	0.90
30%	0.77	0.84	0.77	0.87
40%	0.76	0.85	0.79	0.88
50%	0.76	0.83	0.77	0.86

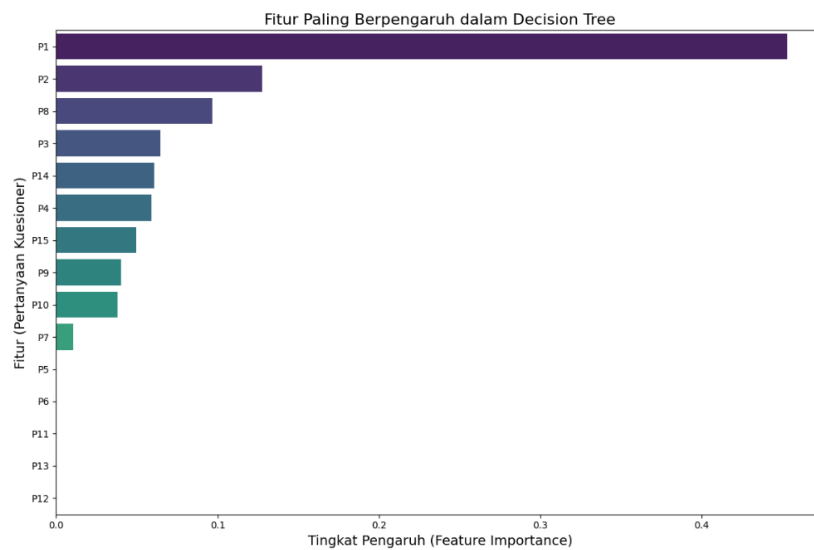


Gambar 3. Visualisasi Distribusi Risiko Burnout

Distribusi awal kategori risiko burnout menunjukkan ketidakseimbangan kelas, di mana kategori “Rendah” mendominasi, disusul “Sedang” dan “Tinggi”. Untuk mengatasi hal ini, dilakukan random oversampling terhadap kategori minoritas. Setelah dilakukan rebalancing, distribusi tiap kategori menjadi seimbang (masing-masing 60 data).

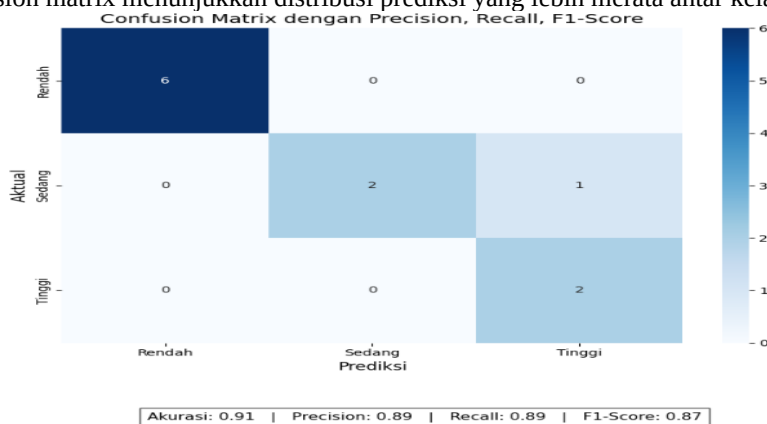


Gambar 4. Visualisasi Distribusi Risiko Setelah Balancing



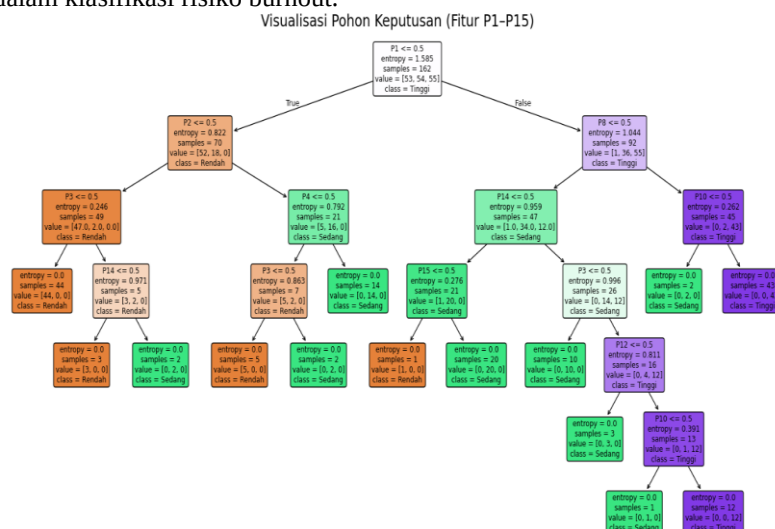
Gambar 5. Fitur Paling Berpengaruh

Model kemudian dilatih kembali dengan data yang telah seimbang. Evaluasi menunjukkan peningkatan performa signifikan: akurasi, precision, recall, dan F1-score masing-masing mencapai 94%, serta hasil confusion matrix menunjukkan distribusi prediksi yang lebih merata antar kelas.



Gambar 6. Visualisasi Confusion Matrix

Visualisasi decision tree menunjukkan alur pemisahan data berdasarkan 15 fitur terpilih (P1–P15), dengan fitur P1, P2, dan P8 sebagai pemisah utama. Selain itu, visualisasi feature importance dan confusion matrix mendukung interpretasi model serta menunjukkan bahwa model hasil rebalancing lebih adil dan akurat dalam klasifikasi risiko burnout.



Gambar 7. Visualisasi DT Setelah Oversampling

3. HASIL DAN PEMBAHASAN

Berdasarkan hasil penelitian yang telah dilakukan, penerapan algoritma Decision Tree (DT) untuk klasifikasi risiko burnout akademik pada siswa kelas XI SMA LTI IGM Palembang memberikan hasil yang cukup baik. Pada tahap awal, model dibangun menggunakan data asli tanpa balancing, dengan jumlah data yang tidak seimbang—kategori “Rendah” memiliki jumlah terbanyak. Dengan skema pembagian data 90% untuk pelatihan dan 10% untuk pengujian, model menghasilkan akurasi sebesar 91%, precision 89%, recall 89%, dan f1-score 87%. Meskipun metrik evaluasi tergolong tinggi, ketidakseimbangan distribusi data dapat menyebabkan model bias terhadap kategori mayoritas dan kurang optimal dalam mengenali kategori minoritas seperti “Sedang” dan “Tinggi”.

Untuk mengatasi hal tersebut, dilakukan proses balancing data dengan metode oversampling pada kategori “Sedang” dan “Tinggi” agar jumlah data di setiap kategori menjadi seimbang dengan kategori “Rendah”. Setelah balancing, proses seleksi fitur menggunakan Information Gain tetap diterapkan untuk memilih 15 fitur terbaik. Model kemudian dilatih kembali menggunakan data seimbang dan skema split yang sama (90:10). Hasil evaluasi menunjukkan bahwa akurasi model tetap tinggi sebesar 94%, dengan

precision dan recall mencapai 94% serta f1-score meningkat menjadi 94%. Hal ini menandakan bahwa model yang dibangun dari data seimbang lebih mampu mengenali semua kategori secara adil tanpa dominasi kelas tertentu. Peningkatan performa model setelah proses oversampling menunjukkan bahwa ketidakseimbangan distribusi data (class imbalance) merupakan salah satu faktor yang memengaruhi kemampuan algoritma Decision Tree dalam mengenali kelas minoritas. Dengan distribusi data yang lebih seimbang, setiap kategori memperoleh representasi yang memadai selama proses pelatihan sehingga model mampu mempelajari karakteristik masing-masing kelas secara lebih optimal. Kondisi ini mengurangi kecenderungan model untuk memprediksi kelas mayoritas dan meningkatkan kemampuan generalisasi terhadap seluruh kategori risiko burnout akademik.

Hasil penelitian ini sejalan dengan penelitian yang dilakukan oleh Sarita dkk. [4] yang menunjukkan bahwa algoritma Decision Tree memiliki kemampuan yang baik dalam klasifikasi kondisi psikologis. Demikian pula penelitian Deena dkk. [5] melaporkan bahwa Decision Tree mampu mengklasifikasikan tingkat stres siswa dengan performa yang tinggi. Perbedaan penelitian ini terletak pada penerapan teknik Information Gain untuk seleksi fitur serta oversampling untuk mengatasi ketidakseimbangan data sehingga model mampu mengenali setiap kategori risiko burnout secara lebih proporsional.

Penelitian ini memberikan kontribusi melalui kombinasi seleksi fitur Information Gain dengan teknik oversampling pada algoritma Decision Tree untuk klasifikasi burnout akademik siswa. Pendekatan tersebut tidak hanya meningkatkan performa klasifikasi tetapi juga menghasilkan model yang lebih mudah diinterpretasikan melalui visualisasi pohon keputusan dan feature importance sehingga dapat dimanfaatkan sebagai dasar pengambilan keputusan oleh guru maupun konselor sekolah.

Hasil feature importance menunjukkan bahwa indikator keraguan terhadap potensi diri, kecemasan berlebihan, dan kekhawatiran terhadap masa depan memiliki kontribusi terbesar terhadap proses klasifikasi. Temuan ini menunjukkan bahwa faktor psikologis memiliki peranan yang dominan dalam menentukan tingkat risiko burnout akademik. Oleh karena itu, strategi pencegahan burnout di sekolah tidak hanya perlu difokuskan pada aspek akademik, tetapi juga pada penguatan kesehatan mental siswa melalui layanan bimbingan dan konseling.

Dari sisi implementasi, model yang dihasilkan berpotensi digunakan sebagai sistem pendukung keputusan (*Decision Support System*) dalam layanan Bimbingan dan Konseling di sekolah. Guru BK dapat memanfaatkan hasil klasifikasi sebagai dasar untuk mengidentifikasi siswa yang memiliki risiko burnout sehingga intervensi psikologis maupun akademik dapat diberikan secara lebih dini dan tepat sasaran.

Visualisasi confusion matrix dari model balanced memperlihatkan distribusi prediksi yang sempurna di semua kategori, sedangkan visualisasi pohon keputusan dan feature importance memperkuat temuan bahwa fitur-fitur seperti perasaan ragu terhadap potensi diri, kecemasan, dan kekhawatiran berlebihan memainkan peran penting dalam klasifikasi. Dengan demikian, model yang telah dioptimalkan ini dinilai lebih layak digunakan sebagai alat bantu deteksi dini risiko burnout akademik yang adil dan akurat, serta dapat mendukung intervensi yang lebih tepat oleh pihak sekolah atau guru BK.

4. KESIMPULAN

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, terdapat beberapa saran yang dapat disampaikan:

Penelitian ini berhasil membangun model klasifikasi risiko burnout akademik siswa menggunakan algoritma Decision Tree dengan dukungan seleksi fitur Information Gain dan teknik oversampling untuk mengatasi ketidakseimbangan kelas. Hasil penelitian menunjukkan bahwa kombinasi metode tersebut mampu meningkatkan kemampuan model dalam mengenali seluruh kategori risiko burnout secara lebih seimbang sehingga dapat dimanfaatkan sebagai alat bantu deteksi dini bagi pihak sekolah dalam mengidentifikasi siswa yang memerlukan pendampingan atau intervensi. Selain menghasilkan model yang mudah diinterpretasikan, penelitian ini juga mengidentifikasi bahwa indikator keraguan terhadap potensi diri, kecemasan berlebihan, dan kekhawatiran terhadap masa depan merupakan faktor yang paling berpengaruh dalam proses klasifikasi.

Penelitian ini masih memiliki beberapa keterbatasan, yaitu jumlah responden yang relatif terbatas dan hanya berasal dari satu sekolah sehingga hasil penelitian belum dapat digeneralisasikan pada populasi yang lebih luas. Selain itu, penelitian hanya menggunakan algoritma Decision Tree sehingga belum diketahui performanya dibandingkan dengan algoritma klasifikasi lainnya.

Penelitian selanjutnya disarankan untuk melibatkan jumlah responden yang lebih besar dan berasal dari berbagai sekolah, menerapkan teknik validasi seperti k-fold cross validation, serta membandingkan performa Decision Tree dengan algoritma lain, seperti Random Forest, Support Vector Machine, Extreme Gradient Boosting (XGBoost), atau metode ensemble lainnya. Selain itu, pengembangan sistem deteksi dini berbasis web atau aplikasi juga dapat dilakukan agar hasil penelitian lebih mudah diimplementasikan dalam layanan bimbingan dan konseling di sekolah.

REFERENCES

- [1] A. Lailan, "Peran Teknologi Pendidikan Dalam Pembelajaran," *SENTRI J. Ris. Ilm.*, vol. 3, no. 7, pp. 3257–3262, 2024, doi: 10.55681/sentri.v3i7.3115.
- [2] A. Destria, E. Purnamasari, and M. H. Irfani, "Pembuatan Aplikasi Pengelolaan Arsip Surat Dinas PU.BMTR Prov. SumSel," *J. Softw. Eng. Comput. Intell.*, vol. 2, no. 01, pp. 26–36, 2024, doi: 10.36982/jsoci.v2i01.4100.
- [3] I. N. Anggraini, K. Kartika, and H. Darmayanti, "Pengaruh Kecerdasan Emosional terhadap School Burnout Siswa di SMA Muhammadiyah 8 Palembang The Impact of Emotional Intelligence on School Burnout among Students at SMA Muhammadiyah 8 Palembang," vol. 3, no. 2, pp. 79–89, 2024, doi: 10.15575/jops.v3i2.37407.
- [4] J. Sarita, I. Lu, I. Maylani, and P. Setyawati, "Pengaruh Kesehatan Mental Terhadap Stress Akademik Pada Mahasiswa Di Universitas X Dengan Menggunakan Metode Algoritma Decision Tree J48," 2024.
- [5] G. Deena, A. Sandhya, and K. Raja, "Machine Learning-Based Classification and Prediction of Student Stress Levels: a Comparative Study of Algorithms," *J. Theor. Appl. Inf. Technol.*, vol. 102, no. 19, pp. 7063–7076, 2024.
- [6] C. H. Saputra, A. Hermawan, and D. Avianto, "Prediksi Burnout Pada Programmer Menggunakan Teknik Pengenalan Pola Untuk Identifikasi Dini Dan Intervensi," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 3, pp. 667–674, 2024, doi: 10.25126/jtiik.1138070.
- [7] D. Sartika, I. Saluza, and M. H. Irfani, "Perbandingan Akurasi Metode Principal Component Analysis (PCA) dan Correlation-Based Feature Selection (CFS) Pada Klasifikasi Perpanjangan Kontrak Karyawan Menggunakan Metode Naïve Bayes," *J. Ilm. Inform. Glob.*, vol. 13, no. 2, pp. 82–87, 2022, doi: 10.36982/jiig.v13i2.2292.
- [8] Z. A. Almaghfiroh and D. Darmawan, "Pengaruh Model Pembelajaran Terhadap Minat Belajar Siswa Tingkat Madrasah Tsanawiyah," vol. 3, 2025.
- [9] M. Penerima and B. Kip, "Implementasi Algoritma Naïve Bayes dalam Prediksi Penerimaan," vol. 15, no. 1, pp. 108–114, 2025.
- [10] N. L. S. H. Seli Marlina Radja Leba, "Pelatihan Model Pembelajaran Jarak Jauh Berbasis Google Form Sebagai Media Pembelajaran," *J. Pengabd. Kpd. Masy. Univ. Pamulang*, vol. 1, pp. 42–46, 2020.
- [11] A. Aseptianova and R. Ajazqi Casanobe, "Efektivitas Asesmen Formatif pada Pembelajaran Biologi di SMA LTI IGM Palembang," *J. Sos. dan sains*, vol. 5, no. 4, pp. 945–961, 2025, doi: 10.59188/jurnalsosains.v5i4.32155.
- [12] B. P. Pratiwi, A. S. Handayani, and S. Sarjana, "Pengukuran Kinerja Sistem Kualitas Udara Dengan Teknologi Wsn Menggunakan Confusion Matrix," *J. Inform. Upgris*, vol. 6, no. 2, pp. 66–75, 2021, doi: 10.26877/jiu.v6i2.6552.
- [13] A. Fahri and Y. Ramdhani, "Visualisasi Data dan Penerapan Machine Learning Menggunakan Decision Tree Untuk Keputusan Layanan Kesehatan COVID-19," *J. Tekno Kompak*, vol. 17, no. 2, p. 50, 2023, doi: 10.33365/jtk.v17i2.2438.
- [14] I. Maulida, A. Suyatno, and H. R. Hatta, "Seleksi Fitur Pada Dokumen Abstrak Teks Bahasa Indonesia Menggunakan Metode Information Gain," *J. SIFO Mikroskil*, vol. 17, no. 2, pp. 249–258, 2016, doi: 10.55601/jsm.v17i2.379.