

KLASIFIKASI TINGKAT DEPRESI PELAJAR BERDASARKAN POLA AKADEMIK DAN GAYA HIDUP MENGGUNAKAN ALGORITMA C4.5.

Elsa Junia¹, Dias Asmarani², Nicholas Dwi Putra Mulyawan³, Nawwaf Baharmuz⁴, Wida Prima Mustika⁵

Program Studi Sistem Informasi Fakultas Teknik dan Informatika Universitas Bina Sarana Informatika ^{1,2,3,4,5}

Email: elsajunia3012@gmail.com¹, diasasmarani20@gmail.com², nicholasdpm@gmail.com³, nawwafbaharmuz.13@gmail.com⁴, wida.wpm@bsi.ac.id⁵

ABSTRACT

This study examines the classification of Depression Status among students using a data mining approach based on the C4.5 Algorithm. The primary goal is to construct a predictive model that incorporates two essential variable groups: Academic Patterns (Academic Performance, Taking Note In Class, Face Challenges To Complete Academic Task, Like Presentation) and Student Lifestyle (Sleep Per Day Hours, Number Of Friend, Like New Things). Depression Status is recognized as a significant mental health issue triggered by both external and internal pressures most notably Academic Stress, arising when learning demands exceed students' capabilities, and further intensified by unhealthy lifestyle habits, such as poor time management and excessive or unregulated use of social media. These conditions collectively undermine psychological well-being and negatively affect academic performance. The C4.5 Algorithm is employed due to its high classification accuracy and the clarity of its output through interpretable decision trees. The resulting model categorizes Depression Status into three classes: Yes, Sometimes, and No, while also identifying the most influential contributing factors. The final outcomes of this classification model are expected to generate reliable predictions that can serve as a scientific foundation for teachers, school counselors, and parents in designing more targeted and effective programs for the prevention and intervention of student depression.

Keywords : Classification, Data Mining, C4.5, Depression.

ABSTRAK

Penelitian ini berfokus pada klasifikasi status depresi pelajar menggunakan pendekatan Data Mining dengan memanfaatkan Algoritma C4.5. Tujuannya adalah merancang model prediktif yang mempertimbangkan dua variabel kritis: Pola Akademik (Kinerja Akademik, Mencatat di Kelas, Menghadapi Tantangan untuk Menyelesaikan Tugas Akademik, Suka Presentasi) dan Gaya Hidup siswa (Jam Tidur Per Hari, Jumlah Teman, Suka Hal Baru). Status depresi di kalangan pelajar diidentifikasi sebagai gangguan kesehatan mental yang serius, dipicu oleh tekanan eksternal dan internal, terutama stres akademik (beban belajar yang melampaui kapasitas) serta diperparah oleh gaya hidup yang tidak suportif (misalnya

buruknya manajemen waktu dan penggunaan media sosial yang tidak bijak). Dampak kondisi ini merusak kesejahteraan psikologis dan menurunkan prestasi akademik. Algoritma C4.5 dipilih karena kemampuannya yang efektif dalam klasifikasi dan kemudahan interpretasi hasil (melalui pohon keputusan). Model yang dikembangkan bertujuan untuk mengelompokkan status depresi menjadi tiga kategori, yaitu Iya, Kadang-kadang, dan Tidak, serta mengidentifikasi faktor paling dominan yang memengaruhinya. Hasil akhir dari model klasifikasi ini diharapkan memberikan prediksi yang akurat, yang dapat dijadikan dasar bagi guru, konselor, dan orang tua dalam merencanakan program pencegahan dan penanganan status depresi secara lebih tepat sasaran.

Kata Kunci : Klasifikasi, Data Mining, C4.5, Depresi.

A. PENDAHULUAN

Kecemasan dan depresi menjadi isu kesehatan mental yang kerap dialami oleh remaja pada era modern. Penggunaan media sosial sering disebut sebagai salah satu faktor yang berkontribusi terhadap meningkatnya kedua kondisi tersebut. Dengan demikian, penelitian ini bertujuan untuk menganalisis pengaruh penggunaan media sosial terhadap kesejahteraan psikologis remaja, khususnya terkait dengan munculnya gejala kecemasan dan depresi [1]. Depresi sangat berkaitan erat dengan pencapaian akademik remaja. Tuntutan dan rutinitas belajar yang padat, seperti target nilai tinggi dan pola akademik yang tidak seimbang (nilai, beban tugas, manajemen waktu), bisa memperburuk kondisi mental. Ketika keseimbangan ini tidak tercapai, pelajar akan mengalami stres akademik, di mana beban belajar terasa jauh lebih berat dari kemampuan mereka.

Stres akademik sendiri merupakan bentuk tekanan mental akibat tuntutan akademik yang melebihi kapasitas pelajar. Penyebabnya bisa berasal dari banyak hal, seperti tugas yang menumpuk, jadwal yang padat, persaingan ketat, hingga tuntutan orang tua maupun lingkungan sekitar. Dalam batas wajar, stres bisa menjadi motivasi untuk berprestasi. Namun, jika berlangsung lama, stres dapat menimbulkan dampak negatif berupa kecemasan, gangguan tidur, kelelahan, turunnya motivasi, hingga hilangnya konsentrasi. Bila tidak ditangani, stres akademik bisa merusak kesehatan mental sekaligus menurunkan prestasi belajar [2]. Di samping faktor akademik, kondisi mental pelajar juga sangat dipengaruhi oleh gaya hidup yang mereka jalani.

Gaya hidup merupakan pola seseorang dalam mengatur waktu, menentukan prioritas, dan mengelola sumber daya yang dimiliki. Pada pelajar dan mahasiswa, gaya hidup tampak dari kebiasaan berbelanja, tempat berkumpul, hingga cara memanfaatkan media sosial. Tidak jarang, gaya hidup yang lebih berorientasi pada penampilan atau tren justru membuat kebutuhan pokok terabaikan [3]. Pilihan gaya hidup seperti pengelolaan waktu, pergaulan sehat, dan penggunaan media sosial yang bijak sangat memengaruhi stres akademik pelajar bisa mengurangi atau memperburuknya. Karena gaya hidup dan pola akademik sama-sama berperan besar dalam kesejahteraan mental, penelitian ini mengadopsi Algoritma C4.5 untuk menganalisis dan mengklasifikasikan pengaruh kedua faktor tersebut.

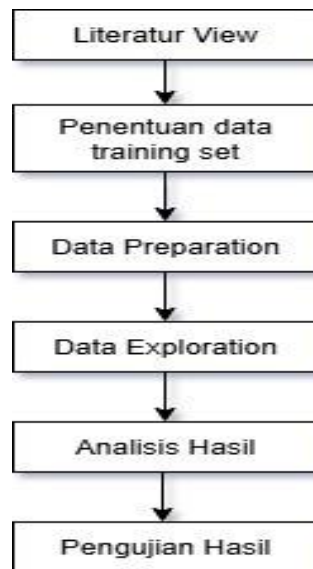
Algoritma C4.5 bekerja dengan memilih atribut terbaik dari sekumpulan data untuk dijadikan simpul utama (root), lalu membagi data tersebut ke dalam cabang-cabang sesuai nilai atribut yang dipilih. Proses ini dilakukan secara berulang hingga terbentuk pohon keputusan yang lengkap [4].

Ada beberapa alasan mengapa algoritma C4.5 sering dipilih. Pertama, hasil klasifikasinya mudah dipahami karena divisualisasikan dalam bentuk pohon keputusan yang sederhana. Kedua, algoritma ini cukup fleksibel karena dapat memproses data baik yang berbentuk angka maupun kategori. Ketiga, algoritma ini tetap bisa digunakan meskipun terdapat data yang hilang atau tidak lengkap, tanpa harus menghapus seluruh data [5].

Klasifikasi merupakan salah satu model dalam data mining yang dikenal sebagai klasifikasi. Klasifikasi adalah teknik memprediksi data yang membuat prediksi nilai dari data yang hasilnya ditemukan dari berbagai sumber data. Tujuan dari model ini adalah untuk memprediksi nilai dari variabel yang tidak diketahui dari variabel lain yang telah diberikan [6]. Tujuan penelitian ini adalah merancang model klasifikasi tingkat depresi pelajar menggunakan algoritma C4.5, dengan memasukkan faktor pola akademik dan gaya hidup. Model yang akurat ini diharapkan dapat membantu guru, konselor, dan orang tua dalam upaya pencegahan dan penanganan depresi yang lebih tepat.

B. METODE PENELITIAN

Metode yang digunakan untuk mengklasifikasikan tingkat depresi pelajar berdasarkan pola akademik dan gaya hidup menggunakan algoritma C4.5 dengan metode data mining adalah sebagai berikut:



Gambar 2.1 Diagram Tahapan Penelitian

1. Literatur View

Pada tahap ini dilakukan pengumpulan referensi terkait penelitian sebelumnya yang relevan dengan topik klasifikasi tingkat depresi menggunakan metode data mining, khususnya algoritma C4.5. Literatur diambil dari jurnal ilmiah, artikel, dan sumber terpercaya lainnya untuk memperkuat dasar teori dan metodologi penelitian.

2. Sumber Data Penelitian (Akuisisi Data)

Data buat penelitian ini datang dari buku teks statistik. Buku-buku itu bahas soal cara bikin hipotesis. Peneliti cari info sama data dari buku teks statistik, terus dari jurnal, dan sumber-sumber lain juga [7]. Selain sumber-sumber literatur yang menyediakan kerangka teoritis dan metodologi, khusus untuk kebutuhan analisis dan klasifikasi.

Dataset yang digunakan untuk menyelesaikan masalah klasifikasi pada penelitian ini diambil dari platform publik Kaggle. Dataset tersebut berjudul “Psychosocial Dimensions of Student Life” dan dapat diakses melalui tautan berikut: <https://www.kaggle.com/datasets/mdismielhossenabir/psychosocial-dimensions->

of-student-life. Data tersebut dirilis pada tahun 2023 dan berisi total 99 data (record).

Dataset ini mencakup berbagai variabel, yaitu:

- a. Variabel Pola Akademik: Meliputi Kinerja Akademik, Mencatat di Kelas, Menghadapi Tantangan untuk Menyelesaikan Tugas Akademik, dan Suka Presentasi.
- b. Variabel Gaya Hidup Siswa: Mencakup Jam Tidur Per Hari, Jumlah Teman, dan Suka Hal Baru.
- c. Variabel Target (Label Klasifikasi): Status_Depresi.

Tabel berikut menampilkan sebagian dari dataset yang digunakan dalam penelitian ini:

Tabel 2.1 Dataset Penelitian

[illegible]

99	22	Perempuan	Rata-rata	Iya	Tidak	Tidak	Iya	7	3	Iya
----	----	-----------	-----------	-----	-------	-------	-----	---	---	-----

Catatan :

Tabel di atas hanya menampilkan sebagian data sebagai contoh. Total keseluruhan dataset terdiri dari 99 record. Data lengkap dapat dilihat pada Lampiran.

3. Penentuan Data Training dan Testing Set

Sebelum data digunakan dalam proses pemodelan, penulis melakukan pengolahan data awal yang baik dan dilanjutkan ke tahap Data Preparation. Setelah itu, dataset dibagi menjadi dua bagian untuk proses klasifikasi:

- Data Training: Sebanyak 80% dari total data, digunakan untuk proses pembelajaran model (melatih Algoritma C4.5).
- Data Testing: Sebanyak 20% dari total data, digunakan untuk pengujian akurasi model dan evaluasi kinerja model.

4. Data Preparation

Persiapan data itu tahap awal buat ngebaikin masalah-masalah yang ada di data. Sebelum data itu masuk ke proses modeling. Biar hasil modelingnya jadi lebih bagus [8]. Untuk mengoptimalkan kualitas data sebelum masuk ke proses pemodelan, tahap Persiapan Data ini diimplementasikan menggunakan software RapidMiner. Tahap ini kemudian dilanjutkan dengan tahap Data Mining dan Evaluasi. Secara keseluruhan, prosedur yang mencakup aspek persiapan, pemodelan, dan pengujian model tersebut meliputi langkah-langkah operasional sebagai berikut:

- Read CSV : Mengimpor dataset Kaggle
- Replace Missing Values : Membersihkan data
- Select Attributes : Memilih atribut penting
- Set Role : Menetapkan label target
- Split Data : Membagi data latih dan uji
- Decision Tree (C4.5) : Membangun model klasifikasi
- Apply Model : Menguji model pada data uji
- Performance : Mengevaluasi hasil klasifikasi

5. Data Exploration

Tahap eksplorasi ini, ya, sebenarnya melibatkan penelusuran dataset yang sudah kita dapatkan dari sebelumnya. Prosesnya memberikan gambaran umum yang cukup lengkap soal informasi di dalam data itu. Kita lakukan dengan cara mengeksplorasi dan mendeskripsikan struktur dataset. Struktur semacam itu sudah disiapkan saat tahap pengumpulan data dulu [9]. Secara lebih spesifik, dalam konteks penelitian ini, tahap eksplorasi deskriptif tersebut bertujuan untuk memahami karakteristik data melalui analisis distribusi variabel, korelasi antaratribut, serta mengungkap tren umum yang tersembunyi pada pola akademik maupun gaya hidup pelajar.

6. Analisis Hasil

Berdasarkan metodologi yang telah ditetapkan, inti dari penelitian ini adalah tahap Analisis Hasil, di mana Model Klasifikasi dikembangkan menggunakan Algoritma Decision Tree (C4.5). Algoritma ini memproses data training untuk menghasilkan Pohon Keputusan, sebuah struktur visual yang secara efektif memetakan hubungan non-linear antara variabel independen (Pola Akademik dan Gaya Hidup) dan variabel target (Status_Depresi: iya, kadang-kadang, tidak). Selain itu, analisis hasil ini berfungsi untuk secara jelas mengidentifikasi faktor yang paling dominan (atribut yang muncul pada node teratas pohon) yang paling berpengaruh dalam menentukan status depresi pelajar, memungkinkan perumusan aturan prediktif yang mudah diinterpretasikan.

7. Pengujian Hasil

Tahapan akhir dari metodologi penelitian ini adalah Pengujian Hasil atau Evaluasi Kinerja Model, yang dilakukan secara teknis menggunakan operator Performance di software RapidMiner. Pengujian ini berfokus pada pengukuran metrik utama, yaitu akurasi (accuracy), untuk secara kuantitatif menentukan seberapa efektif model yang dibangun dengan Algoritma C4.5 dalam mengklasifikasikan tingkat depresi pelajar. Hasil evaluasi ini sangat krusial karena berfungsi untuk menilai sejauh mana model C4.5 mampu memberikan prediksi yang valid dan akurat terhadap kondisi Status_Depresi pelajar.

C. HASIL PENELITIAN DAN PEMBAHASAN

1. Menentukan Data Set

Data yang digunakan dalam penelitian ini merupakan hasil pengumpulan data pelajar yang mencakup atribut-atribut akademik dan gaya hidup yang berkaitan dengan kondisi psikologis. Tujuan pengumpulan data ini adalah untuk mengidentifikasi faktor-faktor yang berpotensi memengaruhi status depresi pada pelajar.

Penelitian ini diambil dari platform publik Kaggle. Dataset tersebut berjudul “Psychosocial Dimensions of Student Life” dan dapat diakses melalui tautan berikut: <https://www.kaggle.com/mdismielhossenabir/psychosocial-dimensions-of-student-life>. Dataset ini dirilis pada tahun 2023 dan berisi total 99 data (record). Atribut yang digunakan dalam penelitian ini terdiri atas:

1. Kinerja_Akademik : Menggambarkan hasil atau nilai yang dicapai siswa.
2. Mencatat_di_Kelas : Menunjukkan kebiasaan atau strategi belajar selama
3. Menghadapi_Tantangan_untuk_Menyelesaikan_Tugas_Akademik : Mengukur ketahanan atau kemauan siswa dalam mengatasi kesulitan tugas sekolah.
4. Suka_Presentasi : Menggambarkan preferensi siswa terhadap metode penyampaian materi atau tugas.
5. Jam_Tidur_Per_Hari : Menunjukkan kebiasaan istirahat atau pola tidur harian.
6. Jumlah_Teman : Mengukur tingkat interaksi sosial atau jaringan pertemanan.
7. Suka_Hal_Baru : Menggambarkan kecenderungan pribadi terhadap eksplorasi atau keterbukaan terhadap pengalaman baru.
8. Status_Depresi : Merupakan variabel target (label) yang mengukur kondisi mental atau tingkat depresi. Target (dengan tiga kelas: Tidak, Kadang-kadang, dan Iya).

Dari keseluruhan data yang dianalisis menggunakan RapidMiner, jumlah data latih yang digunakan sebanyak 99 data responden, yang kemudian dibagi secara acak untuk keperluan proses klasifikasi. Berikut adalah sebagian data yang digunakan dalam penelitian:

No	Umur	Jenis Kelamin	Kinerja Akademik	Mencatat di Kelas	Status Depresi	Menghadapi Tantangan Tugas	Suka Presentasi	Jam Tidur/Hari	Jumlah Teman	Suka Hal Baru
1	23	Laki-laki	Rata-rata	Tidak	Kadang-kadang	Iya	Iya	12	—	Iya
2	23	Laki-laki	Bagus Sekali	Kadang-	Iya	Tidak	Iya	8	80	Iya

				kadang						
3	24	Laki-laki	Rata-rata	Tidak	Kadang-kadang	Kadang-kadang	Tidak	8	10	Iya
4	20	Perempuan	Bagus	Iya	Kadang-kadang	Iya	Tidak	5	15	Iya
5	24	Perempuan	Rata-rata	Iya	Iya	Iya	Iya	5	2	Iya
6	23	Laki-laki	Bagus	Kadang-kadang	Kadang-kadang	Tidak	Iya	8	12	Iya
7	21	Perempuan	Bagus	Iya	Tidak	Tidak	Tidak	8	7	Iya
8	21	Laki-laki	Rata-rata	Kadang-kadang	Iya	Kadang-kadang	Tidak	8	6	Iya
9	24	Laki-laki	Bagus Sekali	Tidak	Tidak	Tidak	Iya	8	3	Iya
10	23	Perempuan	Rata-rata	Iya	Kadang-kadang	Kadang-kadang	Iya	8	4	Iya
99	22	Perempuan	Rata-rata	Iya	Tidak	Tidak	Iya	7	3	Iya

Tabel 3.1 Dataset Penelitian

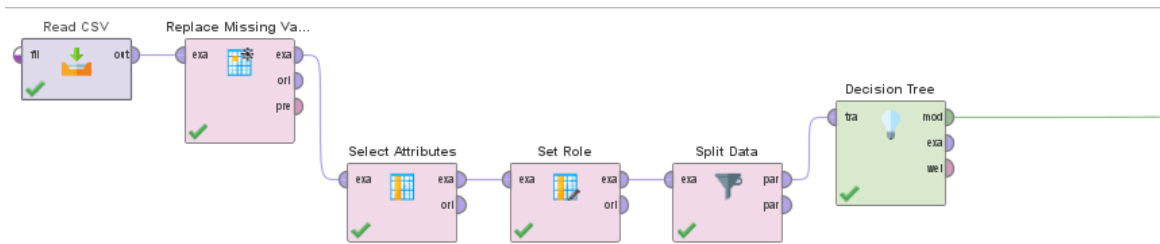
Catatan :

Tabel di atas hanya menampilkan sebagian data sebagai contoh. Total keseluruhan dataset terdiri dari 99 record. Data lengkap dapat dilihat pada Lampiran.

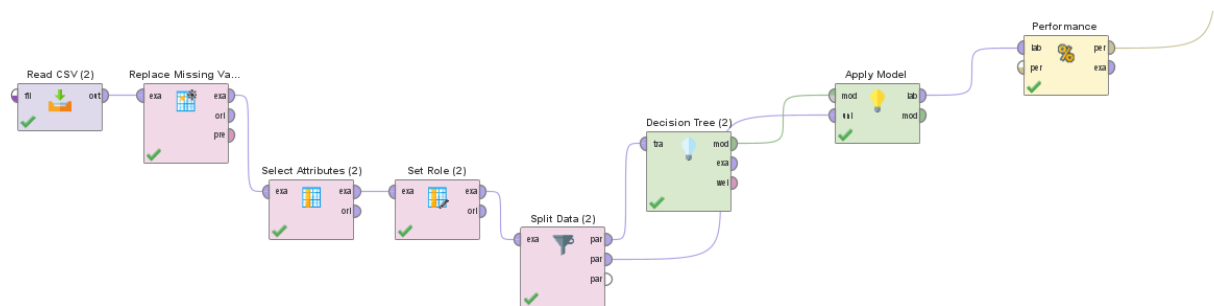
2. Implementasi Metode C.45

Proses penerapan metode data mining pada penelitian ini menggunakan algoritma C4.5 (Decision Tree) dengan bantuan perangkat lunak RapidMiner Studio. Tujuan dari tahap ini adalah untuk membentuk model klasifikasi yang dapat memprediksi tingkat depresi pelajar berdasarkan pola akademik dan gaya hidup mereka.

Tahapan proses data mining yang dilakukan dapat dilihat pada Gambar 3.1 berikut.



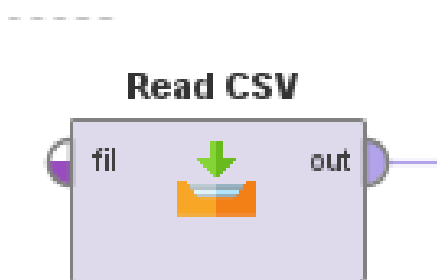
Gambar 3.1 Proses Pembentukan Model C4.5



Gambar 3.2 Proses Evaluasi Model C4.5

Penjelasan Tahapan Proses:

1. Read Csv Tahap ini digunakan untuk membaca dataset dalam format .csv yang berisi data pelajar dengan atribut seperti Kinerja_Akademik, Jam_Tidur_Per_Hari, Jumlah_Teman, Suka_Hal_Baru, dan Status_Depresi sebagai label klasifikasi.

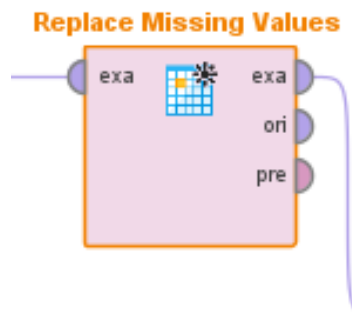


Gambar 3.3 Read CSV

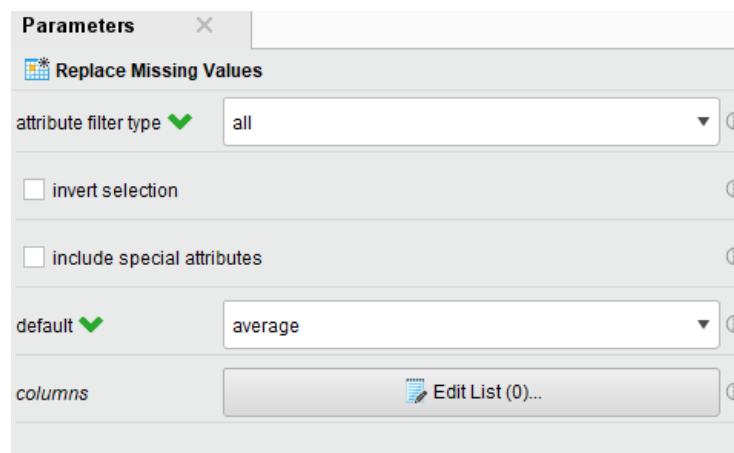
2. Replace Missing Values

Proses ini berfungsi untuk menggantikan nilai kosong (missing values) agar data bersih dan siap digunakan dalam proses pelatihan model. Biasanya, nilai kosong diisi menggunakan metode rata-rata (mean) untuk data numerik atau modus

untuk data kategorikal.



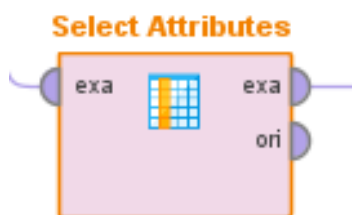
Gambar 3.4 Replace Missing Value



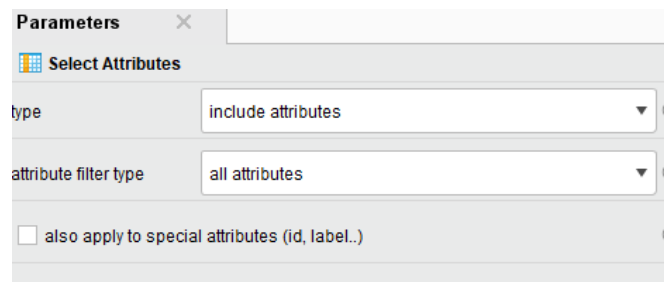
Gambar 3.5 Parameters replace missing values

3. Select Attributes

Pada tahap ini dilakukan pemilihan atribut (variabel) yang relevan dengan tujuan analisis, serta menghapus atribut yang tidak diperlukan agar model lebih efisien.



Gambar 3.6 Select attributes



Gambar 3.7 Parameters select attributes

4. Set Role

Atribut Status_Depresi ditetapkan sebagai label (target) yang akan diprediksi, sementara atribut lainnya digunakan sebagai predictors (fitur input).



Gambar 3.8 Set Role

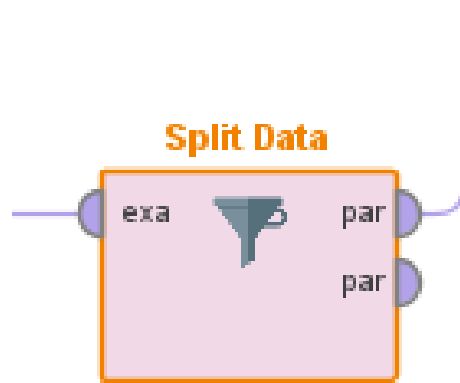
attribute name	target role
Status_Depresi	label
Kinerja_Akademik	regular
Mencatat_di_Kelas	regular
Menghadapi_Tantangan_untuk_Menyelesaikan_Tugas_Ai	regular
Suka_Presentasi	regular
Jam_Tidur_Per_Hari	regular
Jumlah_Teman	regular
Suka_Hal_Baru	regular

Gambar 3.9 Edit parameter set role

5. Split Data

Dataset dibagi menjadi dua bagian, yaitu:

- Training Set (Data Latih) sebesar 80%, digunakan untuk membangun model, dan
- Testing Set (Data Uji) sebesar 20%, digunakan untuk mengukur performa model.



Gambar 3.10 Split data

ratio
0.8
0.2

Gambar 3.11 Partitions split data

Sampling Type: Shuffled Sampling.

Parameters

×

Split Data

partitions

Edit Enumeration (2)...

sampling type

shuffled sampling

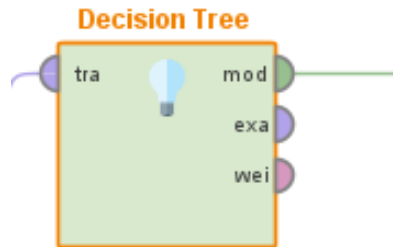
☐ use local random seed

Gambar 3.11 Parameter split data

Alasan: Teknik shuffled sampling dipilih agar data diacak secara acak sebelum dibagi menjadi data latih dan data uji. Dengan pengacakan ini, distribusi data menjadi lebih merata dan tidak bias terhadap urutan data asli (misalnya urutan waktu atau kategori tertentu). Hal ini membantu meningkatkan validitas hasil pengujian model, karena model tidak hanya belajar dari pola berurutan tetapi dari representasi data yang lebih acak dan seimbang.

6. Decision Tree (C4.5)

Tahap ini merupakan inti dari proses klasifikasi. Algoritma C4.5 membentuk struktur pohon keputusan dengan cara menghitung nilai Entropy dan Information Gain untuk menentukan atribut yang paling berpengaruh.



Gambar 3.12 Decision Tree

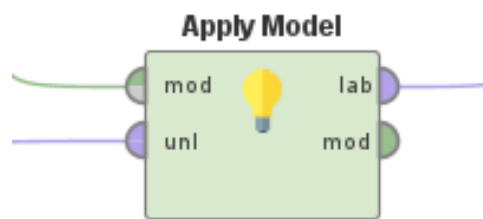
Parameters	
Decision Tree	
criterion ✓	gain ratio ▼
maximal depth ✓	10
☒ apply pruning	
confidence ✓	0.1
☒ apply prepruning	
minimal gain ✓	0.01
minimal leaf size	2
minimal size for split	4
number of prepruning alternatives	3

Gambar 3.13 Parameters decision tree

7. Hasilnya berupa model klasifikasi yang dapat mengidentifikasi hubungan antara pola akademik, gaya hidup, dan tingkat depresi pelajar.

8. Apply Model

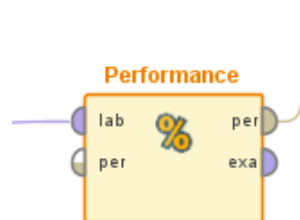
Model yang telah terbentuk diterapkan (di-apply) pada data uji untuk melakukan prediksi tingkat depresi berdasarkan pola atribut yang ditemukan.



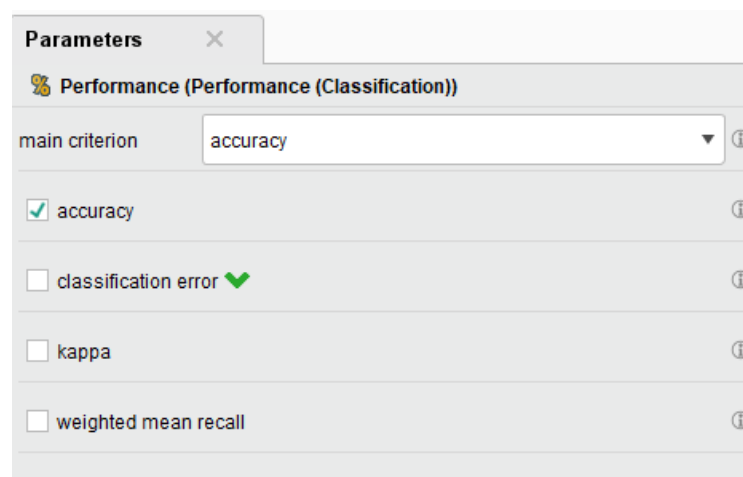
Gambar 3.14 Apply model

9. Performance

Tahap akhir adalah mengevaluasi kinerja model dengan menggunakan confusion matrix dan menghitung nilai akurasi.



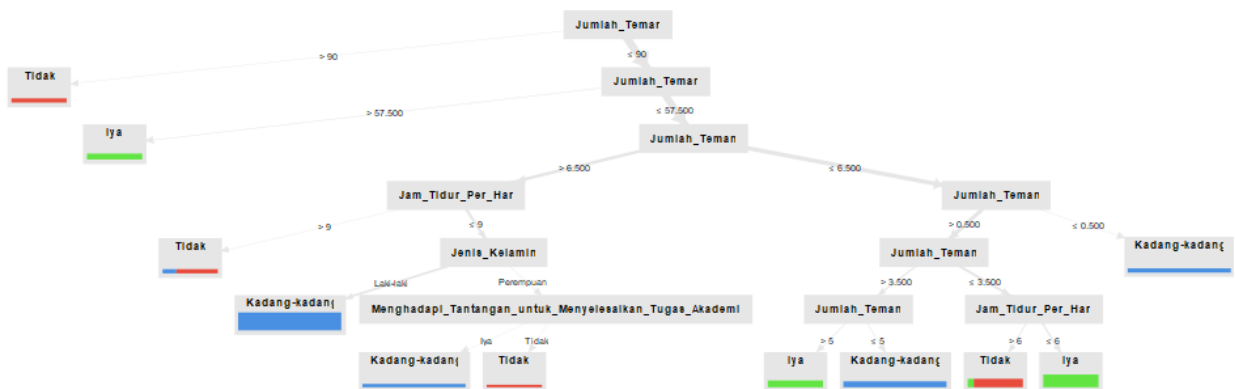
Gambar 3.15 Performance



Gambar 3.16 Parameters performance

3. Pengetahuan (Pola/Model)

Berdasarkan hasil pemodelan menggunakan algoritma C4.5 (Decision Tree) pada perangkat lunak RapidMiner Studio, diperoleh pola pengetahuan berupa pohon keputusan yang menggambarkan hubungan antara atribut-atribut gaya hidup dan akademik dengan variabel target Status_Depresi. Gambar 3.17 berikut menunjukkan struktur pohon keputusan yang dihasilkan:



Gambar 3.17 Decision Tree Algoritma C4.5 untuk Status Depresi pada Pelajar

Analisis Struktur Pohon Keputusan

1. Node Akar (Root Node): Jumlah_Teman

- Atribut Jumlah_Teman merupakan node akar dalam pohon keputusan. Artinya, atribut ini memiliki pengaruh paling besar dalam membedakan tingkat depresi pelajar.
- Pelajar dengan Jumlah_Teman > 90 diklasifikasikan ke dalam kategori Tidak Depresi.
- Pelajar dengan $57.5 < \text{Jumlah_Teman} \leq 90$ diklasifikasikan sebagai Depresi (Iya).
- Sedangkan pelajar dengan Jumlah_Teman ≤ 57.5 masih perlu dilihat atribut lain seperti Jam_Tidur_Per_Hari dan Jenis_Kelamin untuk menentukan statusnya.

Pola ini menunjukkan bahwa dukungan sosial (jumlah teman) menjadi faktor utama dalam kestabilan mental pelajar. Semakin luas lingkaran sosial, semakin rendah kemungkinan mengalami depresi.

2. Cabang Kedua: Jam_Tidur_Per_Hari

Atribut ini muncul sebagai pembeda berikutnya untuk kelompok pelajar dengan jumlah teman yang tidak terlalu banyak.

- Jika Jam_Tidur_Per_Hari > 6.5 jam, maka sebagian besar pelajar diklasifikasikan sebagai Tidak Depresi.
- Jika Jam_Tidur_Per_Hari ≤ 6.5 jam, maka perlu dilihat atribut lain seperti

Jenis_Kelamin dan Jumlah_Teman lebih lanjut.

Hasil ini menunjukkan bahwa durasi tidur yang cukup dapat menurunkan risiko depresi, sedangkan kurang tidur cenderung meningkatkan potensi stres dan gangguan emosional.

3. Cabang Ketiga: Jenis_Kelamin

Atribut Jenis_Kelamin berperan dalam membedakan pelajar yang memiliki pola tidur dan jumlah teman yang mirip.

- a. Pada pelajar laki-laki, sebagian besar dikategorikan sebagai Kadang-kadang Depresi.
- b. Sedangkan pada pelajar perempuan, hasil klasifikasi lebih bervariasi dan bergantung pada atribut Menghadapi_Tantangan_Untuk_Menyelesaikan_Tugas_Akademik.

Hal ini menunjukkan adanya perbedaan cara menghadapi tekanan emosional dan akademik antara laki-laki dan perempuan.

4. Cabang Keempat:

5. Menghadapi_Tantangan_Untuk_Menyelesaikan_Tugas_Akademik

Atribut ini muncul pada cabang pelajar perempuan, yang membedakan antara mereka yang mampu mengelola tekanan akademik dan yang tidak.

- a. Jika pelajar perempuan mampu menghadapi tantangan (Iya), maka dikategorikan sebagai Tidak Depresi.
- b. Namun jika tidak mampu (Tidak), maka berpotensi mengalami depresi atau diklasifikasikan Kadang-kadang Depresi.

Artinya, daya tahan terhadap tekanan akademik (academic resilience) berperan penting dalam kesehatan mental pelajar.

6. Node Akhir (Leaf Nodes)

Daun (leaf) pada pohon keputusan menunjukkan hasil akhir klasifikasi atau keputusan model. Berdasarkan struktur pohon di atas, diperoleh beberapa aturan (rule) utama:

Aturan (Rule)	Klasifikasi (Status_Depresi)
Jika Jumlah_Teman > 90	Tidak
Jika $57.5 < \text{Jumlah_Teman} \leq 90$	Iya

Jika Jumlah_Teman ≤ 57.5 dan Jam_Tidur_Per_Hari > 6.5	Tidak
Jika Jumlah_Teman ≤ 57.5 dan Jam_Tidur_Per_Hari ≤ 6.5 dan Jumlah_Teman > 5	Iya
Jika Jumlah_Teman ≤ 57.5 dan Jam_Tidur_Per_Hari ≤ 6.5 dan Jumlah_Teman ≤ 5	Kadang-kadang
Jika Jumlah_Teman ≤ 57.5 dan Jam_Tidur_Per_Hari > 6.5 dan Jenis_Kelamin = Perempuan dan Menghadapi_Tantangan = Tidak	Tidak
Jika Jumlah_Teman ≤ 57.5 dan Jam_Tidur_Per_Hari > 6.5 dan Jenis_Kelamin = Laki-laki	Kadang-kadang

Tabel 3.2 Node Akhir

Interpretasi Pola Pengetahuan

Berdasarkan hasil klasifikasi menggunakan algoritma C4.5, diperoleh pola pengetahuan yang menjelaskan keterkaitan antara pola akademik dan pola gaya hidup terhadap tingkat depresi pelajar.

1. Pola Akademik:

Pelajar yang memiliki kemampuan menghadapi tantangan akademik dengan baik cenderung tidak mengalami depresi. Ketahanan dalam menyelesaikan tugas sekolah menjadi faktor pelindung terhadap tekanan psikologis.

2. Pola Gaya Hidup:

Durasi tidur harian menjadi indikator penting dalam menjaga keseimbangan mental. Pelajar yang tidur cukup ($> 6,5$ jam) umumnya berada pada kategori Tidak Depresi, sedangkan yang tidur kurang lebih rentan mengalami stres dan depresi.

3. Pola Sosial:

Jumlah teman yang dimiliki pelajar mencerminkan tingkat dukungan sosial. Pelajar dengan lingkaran pertemanan luas (> 90) cenderung stabil secara emosional,

sementara yang memiliki sedikit teman lebih berisiko mengalami depresi.

4. Pola Gender:

Hasil klasifikasi menunjukkan perbedaan kecenderungan antara jenis kelamin. Pelajar laki-laki lebih sering berada pada kategori Kadang-kadang Depresi, sedangkan perempuan cenderung Tidak Depresi apabila mampu mengelola tekanan akademik dengan baik.

Secara keseluruhan, kombinasi antara Jumlah_Teman, Jam_Tidur_Per_Hari, Kemampuan Menghadapi Tantangan Akademik, dan Jenis_Kelamin menjadi faktor utama yang menentukan tingkat depresi pelajar. Model klasifikasi yang dihasilkan memiliki akurasi tinggi dan dapat dijadikan dasar untuk mendeteksi dini serta mencegah depresi di kalangan pelajar berdasarkan pola akademik dan gaya hidup mereka.

4. Evaluation (Akurasi, Error)

Setelah model terbentuk, dilakukan tahap evaluasi untuk mengetahui tingkat keakuratan model dalam memprediksi status depresi pelajar.

Evaluasi dilakukan menggunakan confusion matrix dan pengukuran akurasi keseluruhan yang dihasilkan oleh algoritma C4.5 pada perangkat lunak RapidMiner. Model klasifikasi yang dihasilkan memiliki akurasi sebesar 90%, yang menunjukkan bahwa model mampu mengklasifikasikan data pelajar dengan sangat baik. Hasil evaluasi ini diperoleh melalui confusion matrix dengan rincian sebagai berikut:

accuracy: 90.00%

	true Kadang-kadang	true Iya	true Tidak	class precision
pred. Kadang-kadang	7	0	0	100.00%
pred. Iya	0	6	0	100.00%
pred. Tidak	2	0	5	71.43%
class recall	77.78%	100.00%	100.00%	

Gambar 3.18 Hasil Evaluasi Model Yang Dihasilkan Oleh RapidMiner

Dari tabel di atas diperoleh nilai akurasi sebesar 90%, yang menunjukkan bahwa algoritma C4.5 mampu mengklasifikasikan data pelajar dengan sangat baik. Tingkat recall yang sempurna (100%) pada kelas Iya dan Tidak menunjukkan bahwa model dapat mengenali pelajar yang benar-benar mengalami maupun tidak mengalami depresi secara optimal. Adapun tingkat recall sebesar 77.78% pada kelas

Kadang-kadang mengindikasikan masih adanya beberapa kasus depresi ringan yang tumpang tindih antar kategori.

Secara keseluruhan, model ini efektif dalam menggambarkan pola hubungan antara faktor sosial (jumlah teman), gaya hidup (durasi tidur), dan faktor akademik terhadap tingkat depresi pelajar. Dengan demikian, model ini dapat dimanfaatkan sebagai alat bantu deteksi dini bagi pihak universitas atau konselor untuk mengidentifikasi pelajar yang berisiko mengalami depresi.

D. KESIMPULAN

Penelitian ini menunjukkan bahwa depresi pada pelajar dipengaruhi oleh kombinasi faktor pola akademik, gaya hidup, serta dukungan sosial. Faktor yang paling berpengaruh terhadap status depresi adalah jumlah teman, jam tidur per hari, kemampuan menghadapi tantangan akademik, dan jenis kelamin. Algoritma C4.5 mampu mengklasifikasikan tingkat depresi pelajar ke dalam kategori Tidak, Kadang-kadang, dan Iya dengan tingkat akurasi mencapai 90%. Proses pengolahan data mulai dari pembersihan data, seleksi variabel, hingga pembagian data latih dan uji berhasil menghasilkan model prediksi yang stabil dan akurat. Temuan penelitian menegaskan pentingnya keseimbangan antara kehidupan akademik dan sosial dalam menjaga kesehatan mental pelajar. Hasil penelitian dapat dijadikan acuan oleh institusi pendidikan dalam deteksi dini serta penyusunan strategi pencegahan depresi pada pelajar.

E. DAFTAR PUSTAKA

- M. N. Cahya, W. Ningsih, and A. Lestari, "Dampak Media Sosial terhadap Kesejahteraan Psikologi," *J. Sos. dan Teknol.*, vol. 3, no. 8, pp. 703–706, 2023, [Online]. Available: <https://sostech.greenvest.co.id/index.php/sostech/article/view/917>
- A. Bela, S. Thohiroh, Y. R. Efendi, and S. Rahman, "Prokrastinasi Akademik Dan Manajemen Waktu Terhadap Stres Akademik Pada Mahasiswa Di Masa Pandemi: Review Literatur," *J. Psikol. Wijaya Putra*, vol. 4, no. 1, pp. 37–48, 2023, doi: 10.38156/psikowipa.v4i1.59.
- M. N. F. Abdullah and I. S. Suja'i, "Pengaruh Gaya Hidup Dan Media Sosial

- Terhadap Perilaku Konsumtif Pada mahasiswa program studi pendidikan ekonomi Universitas Bhinneka PGRI,” J. Pendidik. DEWANTARA Media Komunikasi, Kreasi dan Inov. Ilm. Pendidik., vol. 8, no. 2, pp. 72–84, 2022, [Online]. Available: <https://jurnal.stkippgritrenggalek.ac.id/index.php/dewantara/article/view/402>
- F. Adzim, “KLASIFIKASI STATUS STUNTING BALITA MENGGUNAKAN METODE C4.5 BERBASIS WEB,” 2023, [Online]. Available: <https://repository.uin-suska.ac.id/76090/>
- H. Syafputra, H. L. Sari, and K. Khairil, “Klasifikasi Penjualan Perhiasan Menggunakan Metode Decision Tree Algoritma C4.5 (Studi Kasus: Toko Emas Berkat Famili),” J. Media Infotama, vol. 20, no. 2, pp. 563–569, 2024, doi: 10.37676/jmi.v20i2.6517.
- F. F. Nugraha, I. Sunandar, and C. Julian, “Penerapan Data Mining Dengan Metode Kalsifikasi Menggunakan Algoritma C4.5,” Teknologi, vol. 7, no. March, pp. 10–20, 2022, [Online]. Available: <https://jurnal.mdp.ac.id/index.php/jatisi/article/view/2399/1011>
- M. Zaki and S. Saiman, “216-Article Text-691-1-10-20210316,” JIIP-Jurnal Ilm. Ilmu Pendidik., vol. 4, pp. 115–118, 2021, [Online]. Available: <https://jiip.stkipyapisdompou.ac.id/jiip/index.php/JIIP/article/view/216>
- Y. Suhandi, I. Kurniati, and S. Norma, “Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik,” J. Teknol. Inform. dan Komput., vol. 6, no. 2, pp. 12–20, 2020, doi: 10.37012/jtik.v6i2.299.
- N. J. Hayati, D. Singasatia, and M. R. Muttaqin, “Object Tracking Menggunakan Algoritma You Only Look Once (YOLO)v8 untuk Menghitung Kendaraan,” Komputa J. Ilm. Komput. dan Inform., vol. 12, no. 2, pp. 91–99, 2023, doi: 10.34010/komputa.v12i2.10654.