



INNOVATIVE: Journal Of Social Science Research

Volume 5 Nomor 3 Tahun 2025 Page 5915-5934

E-ISSN 2807-4238 and P-ISSN 2807-4246

Website: <https://j-innovative.org/index.php/Innovative>

## Implementasi Algoritma Decision Tree untuk Deteksi Depresi Pada Pelajar

Sofyh Zacksavira<sup>1✉</sup>, Cindy Anggraeni Br Sebayang<sup>2</sup>, Jonathan Imanuel<sup>3</sup>, Grestia Br Sibarani<sup>4</sup>,

Universitas Satya Terra Bhinneka

Email: [2303311258@students.satyaterrabhinneka.ac.id](mailto:2303311258@students.satyaterrabhinneka.ac.id)<sup>1✉</sup>

### Abstrak

Depresi merupakan masalah kesehatan mental yang berdampak signifikan terhadap kualitas hidup dan produktivitas individu. Penelitian ini bertujuan mengembangkan model klasifikasi depresi menggunakan algoritma Decision Tree yang diimplementasikan melalui platform Google Colab. Dataset yang digunakan terdiri dari 503 entri dengan 11 variabel yang mencakup aspek demografis, kondisi akademik, dan kesehatan mental. Proses preprocessing mencakup konversi format data, transformasi variabel kategorikal dengan metode Nominal to Binomial, serta normalisasi data numerik menggunakan Z-transformation. Model Decision Tree dioptimalkan dengan parameter gain ratio, kedalaman maksimum 10, confidence factor 0,1, serta aktivasi pruning dan pre-pruning. Hasil pengujian menunjukkan performa model yang tinggi dengan akurasi sebesar 89%, weighted mean recall 89%, dan weighted mean precision 89%. Confusion matrix mengindikasikan keberhasilan model dalam mengidentifikasi 43 kasus non-depresi dan 3 kasus depresi secara akurat, dengan tingkat kesalahan klasifikasi yang rendah. Temuan ini menunjukkan potensi Decision Tree sebagai alat bantu dalam deteksi awal kondisi depresi berbasis data.

Kata Kunci: Decision Tree, Kesehatan Mental, Depresi Machine Learning

### Abstract

Depression is a mental health problem that has a significant impact on the quality of life and productivity of individuals. This research aims to develop a depression classification model using the Decision Tree algorithm implemented through the Google Colab platform. The dataset used consists of 503 entries with 11 variables covering demographic aspects, academic conditions, and mental health. The preprocessing process includes data format conversion, categorical variable transformation using Nominal to Binomial method, and numerical data normalization using Z-transformation. The Decision Tree model was optimized with gain ratio parameters, maximum depth of 10, confidence factor of 0.1, as well as pruning and pre-pruning activation. The test results showed high model performance with 89% accuracy, 89% weighted mean recall, and 89% weighted mean precision. The confusion matrix indicated the success of the model in accurately identifying 43 non-depression cases and 3 depression cases, with a low misclassification rate. These findings demonstrate the potential of Decision Tree as a tool in data-driven early detection of depression.

Keywords: Decision Tree, Mental Health, Depression Machine Learning

### PENDAHULUAN

Depresi kini telah berkembang menjadi salah satu isu krusial dalam ranah kesehatan mental di era modern. Berdasarkan data yang dirilis oleh World Health Organization (WHO), lebih dari 264 juta individu secara global mengalami gangguan depresi, dengan tren peningkatan kasus yang konsisten setiap tahunnya. Survei mengenai kesehatan mental remaja di Indonesia pada tahun 2022 menunjukkan bahwa sekitar 5,5% remaja berusia 10–17 tahun mengalami gangguan mental. Dari jumlah tersebut, 1% teridentifikasi mengalami depresi, 3,7% mengalami gangguan kecemasan, 0,9% menderita post-traumatic stress disorder (PTSD), dan 0,5% mengalami attention-deficit/hyperactivity disorder (ADHD). Sementara itu, berdasarkan Riset Kesehatan Dasar (Riskesdas) tahun 2023, tercatat bahwa 6,2% penduduk berusia 15–24 tahun mengalami depresi. Fakta yang mengkhawatirkan adalah bahwa sekitar 61% anak muda yang mengalami depresi dalam satu bulan terakhir pernah memiliki pemikiran untuk mengakhiri hidup. Namun, dari jumlah tersebut, hanya 10,4% yang aktif mencari pertolongan atau pengobatan profesional (Kemenkes, 2023). Dampak depresi tidak hanya terbatas pada aspek psikologis individu, melainkan juga berdampak luas terhadap penurunan produktivitas kerja, terganggunya relasi sosial, serta penurunan kualitas hidup secara keseluruhan. Ketika dialami oleh pelajar, depresi tidak hanya memengaruhi kondisi psikologis mereka, tetapi juga berdampak luas terhadap perkembangan akademik, sosial, serta kesehatan fisik secara keseluruhan.

Proses diagnosis gangguan depresi tidaklah sederhana, mengingat keterlibatan berbagai faktor yang saling memengaruhi. Beberapa determinan utama yang berkontribusi terhadap kondisi ini antara lain tekanan dalam lingkungan kerja, tingkat kepuasan kerja, kualitas tidur, pola konsumsi makanan, serta riwayat gangguan mental dalam keluarga. Selain itu, variabel demografis seperti usia, jenis kelamin, dan durasi jam kerja turut memainkan peran penting dalam menentukan tingkat kerentanan seseorang terhadap depresi (Arirahmayanti et al., 2024). Sementara itu, metode diagnosis konvensional yang mengandalkan konsultasi tatap muka dengan profesional kesehatan mental kerap memerlukan waktu yang panjang dan biaya yang tidak sedikit. Hambatan ini sering kali membuat individu enggan atau tertunda dalam mengakses layanan yang dibutuhkan. Ditambah lagi, keberadaan stigma sosial terhadap isu kesehatan mental memperkuat resistensi masyarakat untuk mencari pertolongan secara terbuka (Rismayani et al, 2024).

Seiring dengan kemajuan teknologi informasi, khususnya dalam pengembangan kecerdasan buatan dan machine learning, muncul peluang inovatif dalam mendeteksi dan mengklasifikasikan kondisi depresi secara lebih cepat dan efisien. Salah satu metode yang menonjol adalah Decision Tree, yang terbukti efektif dalam menyusun model klasifikasi berdasarkan berbagai variabel yang relevan. Keunggulan metode ini terletak pada kemampuannya menghasilkan model yang tidak hanya akurat tetapi juga interpretatif dan mudah dipahami oleh pengguna non-ahli (Aini et al., 2023). Decision Tree bekerja dengan membentuk rangkaian keputusan yang terstruktur, berdasarkan karakteristik data yang dianalisis, menyerupai pendekatan klinis seorang profesional dalam menilai kondisi pasien namun dengan kecepatan dan konsistensi yang lebih tinggi.

Algoritma Decision Tree merupakan salah satu metode dalam data mining yang mampu mengidentifikasi pola dan tren dalam data berdasarkan atribut-atribut yang tersedia. Metode ini telah terbukti efektif dalam menganalisis data dan menghasilkan keputusan yang bersifat interpretable, sehingga memudahkan proses pemahaman serta penerapan hasil analisis (Nasrullah, 2021). Dalam berbagai studi sebelumnya, algoritma Decision Tree telah diaplikasikan secara luas untuk meningkatkan akurasi prediksi, termasuk dalam konteks deteksi dini terhadap tingkat depresi pada pelajar. Penerapan algoritma ini diharapkan dapat menjadi salah satu pendekatan strategis dalam membantu proses identifikasi dan penanganan depresi pada pelajar secara lebih dini dan tepat sasaran. Hal ini menjadi sangat relevan mengingat bahwa, meskipun terdapat berbagai inisiatif integrasi teknologi dalam sektor pendidikan di Indonesia, pemanfaatan teknologi prediktif khususnya

dalam mendeteksi masalah kesehatan mental pelajar masih tergolong terbatas (Ramadhan et al., 2025).

Melalui pendekatan Decision Tree, diharapkan penelitian ini dapat memberikan kontribusi signifikan terhadap pengembangan sistem pendeteksi dini yang lebih efisien dan responsif, serta mendorong pengambilan keputusan berbasis data dalam upaya peningkatan kesejahteraan mental di lingkungan pendidikan. Penelitian ini menggunakan dataset komprehensif yang melibatkan 503 responden dan mencakup 11 variabel yang merepresentasikan aspek kehidupan profesional maupun personal. Implementasi metode Decision Tree dilakukan melalui platform Google Colab, yang dikenal luas karena kapabilitasnya dalam pengolahan data berskala besar serta antarmuka yang user-friendly. Pemilihan Google Colab didasarkan pada efektivitasnya dalam mendukung pemodelan data secara optimal, termasuk fitur-fitur yang mendukung proses validasi dan evaluasi model. Dengan memanfaatkan kombinasi antara dataset yang representatif, metode analisis yang kuat, serta dukungan teknologi yang tepat, penelitian ini bertujuan untuk mengembangkan sistem klasifikasi depresi yang dapat digunakan sebagai alat bantu dalam proses skrining awal dan deteksi dini gangguan depresi, khususnya di kalangan pelajar (Nasution et al., 2024).

## METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif, yang merupakan salah satu metode utama dalam penelitian ilmiah selain pendekatan kuantitatif. Pendekatan ini dipilih karena memungkinkan eksplorasi mendalam terhadap fenomena yang kompleks dalam konteks alami, khususnya untuk memahami dinamika psikososial yang berkaitan dengan kondisi depresi. Metode kualitatif berlandaskan pada paradigma postpositivisme, yang menekankan pada pemahaman makna subjektif dan kontekstual dari suatu fenomena, berbeda dengan paradigma positivistik yang lebih berfokus pada pengukuran dan generalisasi. Oleh karena itu, pendekatan ini kerap disebut sebagai metode kontemporer, seiring dengan meningkatnya popularitasnya dalam riset-riset ilmu sosial dan kesehatan mental. Dalam pendekatan ini, peneliti berperan sebagai instrumen utama yang secara langsung terlibat dalam proses pengumpulan dan interpretasi data. Teknik pengumpulan data dilakukan secara triangulatif, yakni dengan mengombinasikan berbagai metode seperti observasi, wawancara mendalam, dan studi dokumentasi, guna memperoleh gambaran yang komprehensif dan meningkatkan validitas hasil penelitian. Proses analisis dilakukan secara induktif, yaitu dengan membangun pola dan temuan dari data empiris yang

dikumpulkan, bukan berdasarkan hipotesis yang telah ditentukan sebelumnya. Hasil dari penelitian kualitatif lebih menekankan pada pemaknaan dan pemahaman yang mendalam terhadap suatu fenomena, daripada sekadar menghasilkan generalisasi statistik. Sebagaimana ditegaskan oleh Sari (2018), pendekatan kualitatif bertujuan untuk mengungkap esensi makna dari suatu gejala sosial melalui perspektif partisipan secara holistik (Samsul, 2023).

## 1. Data Mining

Data mining merupakan suatu proses sistematis untuk mengidentifikasi pola dan informasi bermakna dari sejumlah besar data. Sumber data yang dianalisis dapat berasal dari basis data, gudang data, data berbasis internet, serta berbagai sistem penyimpanan informasi lainnya yang bersifat dinamis. Secara umum, data mining melibatkan eksplorasi basis data dengan menggunakan metode-metode analitis yang inovatif dan dapat dipahami oleh pemilik data, dengan tujuan untuk mengekstraksi informasi yang berguna. Proses implementasi data mining mencakup sejumlah tahapan, antara lain prediksi, deskripsi, klasifikasi, estimasi, pengelompokan (clustering), dan analisis asosiasi (Farisi dan Zaehol, 2024).

Data mining merujuk pada proses penggalian informasi bermakna dari sekumpulan data dalam jumlah besar dengan tujuan mengidentifikasi pola, tren, atau hubungan tersembunyi yang dapat dimanfaatkan dalam proses pengambilan keputusan. Prosedur ini mencakup beberapa tahapan penting, seperti pemahaman awal terhadap data, pembersihan dan transformasi data (data preprocessing), pengembangan model analitik, serta evaluasi terhadap hasil model yang diperoleh. Penerapan data mining telah meluas ke berbagai disiplin, salah satunya di bidang kesehatan, yang memungkinkan analisis terhadap variabel-variabel kompleks untuk mengungkap indikasi klinis secara lebih akurat. Dalam studi terkait kesehatan mental, pendekatan ini dapat digunakan untuk mengidentifikasi keterkaitan antara aspek-aspek seperti pola hidup, tekanan kerja, dan faktor sosial terhadap tingkat kerentanan seseorang terhadap depresi (Primadona dan Rahmat., 2023).

## 2. Decision Tree

Pohon keputusan (decision tree) merupakan salah satu metode pendukung keputusan yang menggunakan representasi model berbentuk struktur pohon. Metode ini digunakan untuk memecahkan permasalahan dengan memetakan berbagai alternatif keputusan yang memungkinkan, termasuk faktor-faktor yang dipengaruhi oleh masing-masing alternatif beserta estimasi konsekuensinya. Decision tree bekerja dengan menampilkan algoritma

yang tersusun atas pernyataan-pernyataan kontrol bersyarat. Proses pengambilan keputusan dilakukan secara bertahap, dimulai dari bagian atas (root node) hingga mencapai bagian bawah (leaf node), sehingga diperoleh keputusan akhir secara sistematis dan akurat (Ramadhan et al., 2025).

Decision Tree ini didasarkan pada prinsip yang menyerupai proses pengambilan keputusan manusia, di mana setiap simpul (node) dalam struktur pohon merepresentasikan suatu kondisi atau pertanyaan, sedangkan setiap cabang menggambarkan kemungkinan respons terhadap kondisi tersebut, misalnya "ya" atau "tidak". Dalam konteks deteksi depresi, algoritma Decision Tree dapat digunakan untuk mengklasifikasikan data gejala yang dialami siswa ke dalam beberapa kategori, seperti "tidak mengalami depresi", "depresi ringan", "depresi sedang", atau "depresi berat". Proses klasifikasi ini diawali dengan pembelajaran dari data historis yang telah diketahui kelasnya, untuk kemudian diterapkan dalam mengklasifikasikan data baru.

Algoritma Decision Tree memiliki sejumlah keunggulan, antara lain:

- Kemampuan untuk menangani baik data kategorikal maupun numerik;
- Struktur hasil yang bersifat interpretatif dan mudah dipahami oleh manusia;
- Tidak mensyaratkan asumsi tertentu terkait distribusi data.

Decision Tree yang dihasilkan dari proses klasifikasi dapat berfungsi sebagai panduan logis yang bermanfaat bagi pihak non-teknis, seperti guru bimbingan konseling atau konselor, dalam memahami dasar pengambilan keputusan sistem deteksi, sehingga memungkinkan intervensi yang lebih tepat dan informatif. Proses pembangunan pohon keputusan diawali dengan penentuan root node atau simpul akar, yaitu simpul paling atas yang merepresentasikan atribut atau variabel utama dalam proses klasifikasi. Selanjutnya, dibentuk cabang (branch) yang mewakili hasil pengujian berdasarkan data latih (training dataset). Tiap pengujian pada atribut ditunjukkan oleh simpul internal (internal node), yang berfungsi sebagai simpul penghubung dalam struktur pohon. Proses ini diakhiri dengan pembentukan leaf node atau simpul daun, yang merepresentasikan hasil akhir berupa prediksi atau label klasifikasi.

Decision Tree adalah metode pemodelan prediktif berbasis pohon yang digunakan untuk tujuan klasifikasi maupun prediksi. Metode ini menyusun struktur keputusan secara hierarkis, di mana setiap simpul mewakili suatu atribut yang dipilih berdasarkan nilai terbaik dalam memisahkan data, dan setiap cabang menunjukkan hasil dari keputusan tersebut. Atribut dipilih menggunakan metrik seperti Information Gain dan Gain Ratio untuk mengukur seberapa baik atribut tersebut membagi data. Nilai entropy digunakan untuk

menggambarkan tingkat ketidakpastian dalam dataset, yang dihitung menggunakan rumus (Setiawan et al., 2023).

$$\text{Entropy}(S) = -\sum (p_i \log_2 p_i)$$

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum \left( \frac{|S_v|}{|S|} \times \text{Entropy}(S_v) \right)$$

$$\text{SplitInfo}(S,A) = -\sum \left( \frac{|S_v|}{|S|} \log_2 \frac{|S_v|}{|S|} \right)$$

$$\text{GainRatio}(S,A) = \frac{\text{Gain}(S,A)}{\text{SplitInfo}(S,A)}$$

Keunggulan utama dari metode Decision Tree adalah interpretabilitasnya yang tinggi, sehingga model yang dihasilkan dapat dengan mudah dipahami oleh pengguna, termasuk mereka yang tidak memiliki latar belakang teknis, sekaligus memberikan hasil yang andal dalam klasifikasi data secara sistematis. Model yang dihasilkan oleh pohon keputusan bersifat intuitif dan representatif, karena menyajikan proses pengambilan keputusan dalam bentuk struktur hierarkis yang menyerupai pohon dengan simpul-simpul logis. Hal ini memungkinkan pengguna untuk mengikuti alur logika klasifikasi secara visual dan sistematis, dari akar hingga ke daun, tanpa memerlukan pemahaman mendalam mengenai algoritma komputasional yang kompleks.

Tingkat keterbacaan yang tinggi ini menjadikan Decision Tree sebagai metode yang sangat ramah bagi pengguna non-teknis, seperti praktisi di bidang bisnis, pendidikan, atau kebijakan publik, yang memerlukan hasil analisis data yang transparan dan mudah ditafsirkan. Selain itu, pohon keputusan mampu mengidentifikasi dan mengelompokkan pola-pola tersembunyi dalam data secara efisien, sehingga memberikan hasil klasifikasi yang tidak hanya akurat tetapi juga konsisten dalam mendukung proses pengambilan keputusan berbasis data (data-driven decision making). Dengan kemampuan untuk menggabungkan transparansi, efektivitas klasifikasi, serta fleksibilitas dalam menangani berbagai tipe data dan variabel, Decision Tree menjadi salah satu metode analisis data yang banyak digunakan dalam berbagai disiplin ilmu, termasuk ilmu komputer, statistik, ekonomi, hingga ilmu sosia (Bahri dan Akhyar, 2020).

### 3. Depresi

Menurut The Diagnostic and Statistical Manual of Mental Disorders edisi kelima (DSM-5), depresi diklasifikasikan sebagai gangguan suasana hati (mood disorder) yang ditandai oleh perasaan sedih yang menetap, kekosongan emosional, atau mudah tersinggung, yang disertai dengan gejala kognitif dan somatik tertentu. Suatu diagnosis depresi ditegakkan apabila individu menunjukkan sedikitnya lima dari sejumlah gejala berikut secara simultan:

perubahan signifikan dalam pola makan, baik berupa penurunan maupun peningkatan berat badan, gangguan tidur seperti insomnia atau hipersomnia, rasa lelah yang persisten, perasaan tidak berharga atau rendah diri, gangguan konsentrasi dan memori, serta munculnya pikiran untuk mengakhiri hidup. Gejala-gejala ini harus berlangsung secara konsisten selama minimal dua minggu dan menyebabkan penderita merasa sedih hampir sepanjang waktu. Dalam kasus major depressive disorder atau depresi berat, kekambuhan sering terjadi dan gangguan ini dapat berkembang menjadi episode depresif yang berulang seumur hidup. Berdasarkan temuan epidemiologis, depresi lebih umum terjadi pada dewasa muda berusia antara 20 hingga 30 tahun, dengan prevalensi pada perempuan tercatat tiga kali lebih tinggi dibandingkan laki-laki (Safinah, 2024).

Menurut Maina (dalam Ramadani et al., 2024), gejala depresi mencakup berbagai manifestasi emosional, kognitif, dan fisik, antara lain: (a) perasaan sedih mendalam atau ketidakbahagiaan yang terus-menerus; (b) iritabilitas yang berlebihan bahkan terhadap persoalan sepele; (c) hilangnya minat dan kesenangan dalam aktivitas sehari-hari; (d) gangguan tidur berupa insomnia atau hipersomnia; (e) perubahan nafsu makan yang dapat menyebabkan penurunan atau peningkatan berat badan secara signifikan; (f) kegelisahan motorik, seperti ketidakmampuan untuk diam atau dorongan untuk terus bergerak; (g) perlambatan dalam proses berpikir, berbicara, atau aktivitas motorik; (h) penurunan konsentrasi dan daya ingat; (i) kelelahan ekstrem meskipun hanya melakukan aktivitas ringan; (j) kecenderungan menyalahkan diri sendiri atas kejadian negatif; (k) keluhan somatik tanpa penyebab medis yang jelas, seperti nyeri kepala atau gangguan pencernaan; serta (l) pikiran untuk mengakhiri hidup atau melakukan tindakan bunuh diri.

Gejala-gejala depresi tidak selalu muncul secara seragam pada setiap individu, karena depresi memengaruhi setiap orang dengan cara yang unik. Faktor-faktor seperti usia, jenis kelamin, tingkat pendidikan, dan latar belakang budaya memainkan peran penting dalam memengaruhi bagaimana seseorang mengalami dan mengekspresikan depresi. Pada anak-anak, gejala depresi dapat berupa rasa sedih, lekas marah, perasaan putus asa, dan kekhawatiran yang berlebihan, sedangkan pada remaja, gejalanya sering kali mencakup kecemasan, kemarahan, dan kecenderungan untuk menghindari interaksi sosial (Ramadani et al., 2024).

#### 4. Tahap Penelitian

Tahap awal dalam proses penelitian dimulai dengan pengumpulan data, di mana dataset yang digunakan diperoleh dari platform Kaggle dan terdiri atas 503 entri data yang



mencakup 11 variabel. Variabel-variabel tersebut meliputi jenis kelamin, usia, tingkat tekanan kerja, kepuasan terhadap pekerjaan, durasi tidur, pola konsumsi makanan, riwayat pikiran untuk bunuh diri, jumlah jam kerja, tekanan finansial, riwayat gangguan kesehatan mental dalam keluarga, serta status depresi yang berfungsi sebagai variabel target. Dataset yang semula tersedia dalam format CSV kemudian dikonversi ke dalam format Excel guna mempermudah proses pemrosesan dan analisis data menggunakan perangkat lunak Google Colab (Hakim dan Faldy, 2024).



Gambar 1. Metodologi Penelitian

Tahap kedua dalam proses penelitian adalah data preprocessing, yang merupakan langkah krusial untuk memastikan kualitas dan kelayakan data sebelum dianalisis lebih lanjut. Pada tahap ini, sejumlah prosedur dilakukan, dimulai dengan penetapan variabel target melalui operator Set Role, di mana kolom "Depression" ditentukan sebagai label klasifikasi utama. Selanjutnya, dilakukan transformasi data menggunakan operator Nominal to Binominal untuk mengubah data kategorikal menjadi format biner, serta Nominal to Numerical untuk mengkonversi variabel kategorikal lain ke dalam bentuk numerik. Selain itu, proses normalisasi diterapkan menggunakan operator Normalize dengan pendekatan Z-transformation, yang bertujuan untuk menyeragamkan skala antar variabel numerik agar tidak menimbulkan bias dalam proses pemodelan (Farisi dan Zaehol, 2024).

Tahap ketiga adalah implementasi algoritma Decision Tree. Pada fase ini, data yang telah dipreproses dibagi ke dalam dua subset, yakni 80% sebagai data latih (training set) dan 20% sebagai data uji (testing set), dengan memanfaatkan operator Split Data. Konfigurasi model Decision Tree dilakukan secara terstruktur, di antaranya dengan menetapkan criterion menggunakan metode gain\_ratio, batas maksimum kedalaman pohon (maximal depth) pada nilai 10, serta confidence factor sebesar 0.1. Untuk meningkatkan performa dan menghindari overfitting, fitur pruning dan prepruning turut diaktifkan selama proses pelatihan model (Latifah et al., 2019).

Tahap terakhir adalah evaluasi model, yang bertujuan untuk menilai kinerja sistem klasifikasi yang telah dikembangkan. Evaluasi ini menggunakan berbagai metrik, antara lain accuracy, confusion matrix, weighted mean recall, dan weighted mean precision. Penggunaan metrik-metrik ini memungkinkan penilaian performa model secara

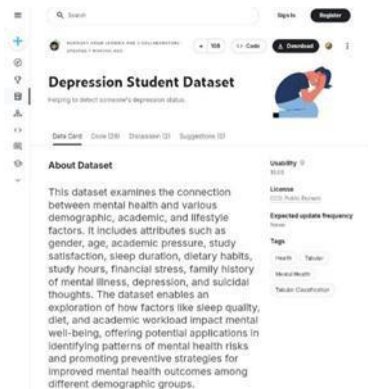
menyeluruh, baik dari sisi ketepatan klasifikasi maupun keseimbangan dalam mendeteksi kategori depresi dan non-depresi. Analisis terhadap masing-masing metrik dilakukan untuk memastikan bahwa model yang dibangun tidak hanya akurat, tetapi juga andal dalam mengidentifikasi kedua kelas secara proporsional (Abrori dan Zaehol, 2024).

## HASIL DAN PEMBAHASAN

### Hasil penelitian

#### Indikator Decision Tree

Dataset yang dijadikan landasan empiris dalam penelitian ini merupakan himpunan data yang mengkaji prevalensi depresi di kalangan pelajar, diperoleh dari repositori terbuka Kaggle pada pranala berikut: <https://www.kaggle.com/datasets/ikynahidwin/depression-student-dataset>. Dataset tersebut terdiri atas 503 entri observasional dengan 11 variabel independen yang merefleksikan dimensi-dimensi kehidupan pelajar serta kondisi psikososial individu yang berkaitan dengan kesehatan mental. Berdasarkan karakteristik struktural dan tipologisnya, variabel-variabel dalam dataset ini dapat diklasifikasikan ke dalam tiga kelompok utama sesuai dengan bentuk data dan fungsinya dalam analisis statistik selanjutnya.



Gambar 2. Halaman Dataset Kaggle

Kategori pertama mencakup variabel bertipe kategorikal, yang terdiri atas Gender (jenis kelamin, dengan nilai Female atau Male), Sleep Duration (durasi tidur, diklasifikasikan ke dalam kategori Less than 5 hours, 5–6 hours, 7–8 hours, dan More than 8 hours), serta Dietary Habits (kebiasaan makan, dengan klasifikasi Unhealthy, Moderate, dan Healthy). Kategori kedua meliputi variabel bertipe numerik, antara lain Age (usia responden dalam satuan tahun), Academic Pressure (tingkat tekanan akademik yang dinyatakan dalam skala ordinal 1 hingga 5), Study Satisfaction (tingkat kepuasan belajar dalam rentang skala yang sama), Study Hours (jumlah jam belajar per hari), serta Financial Stress (tekanan finansial,

juga diukur dengan skala ordinal 1 sampai 5). Selain itu, variabel Suicidal Thoughts (riwayat pemikiran untuk melakukan bunuh diri, dengan nilai Yes atau No) dan Family History (riwayat gangguan kesehatan mental dalam keluarga, juga dengan nilai Yes atau No) turut dianalisis sebagai bagian dari faktor-faktor relevan yang memengaruhi status depresi individu.

Variabel target atau label klasifikasi dalam dataset ini adalah Depression, yang dibagi ke dalam dua kategori, yaitu Yes untuk mengindikasikan adanya gejala depresi, dan No untuk menunjukkan tidak adanya indikasi depresi. Keberadaan variabel ini memungkinkan proses supervised learning dalam membangun model klasifikasi. Secara substansial, dataset ini merepresentasikan kompleksitas multidimensional dari faktor-faktor yang dapat memengaruhi kondisi depresi individu mulai dari aspek personal seperti pola tidur dan kebiasaan konsumsi makanan, hingga dimensi profesional seperti intensitas tekanan kerja dan tingkat kepuasan kerja. Ragam dan heterogenitas variabel yang termuat dalam dataset memberikan landasan analitis yang kuat untuk menghasilkan pemahaman yang lebih menyeluruh mengenai determinan depresi di kalangan profesional.

### Preprocessing Data

Tahapan preprocessing data dalam penelitian ini diawali dengan konversi format berkas dari Comma Separated Values (CSV) ke format Microsoft Excel (.xlsx). Proses transformasi ini dilakukan dengan mempertimbangkan kemudahan dalam pembacaan serta fleksibilitas pengolahan data menggunakan platform Google Colaboratory (Google Colab). Format Excel dipilih karena kompatibilitasnya yang tinggi terhadap berbagai pustaka pemrosesan data seperti pandas, serta kemampuannya dalam memfasilitasi manipulasi data multivariabel secara lebih terstruktur. Dataset yang digunakan dalam penelitian ini terdiri atas 503 baris observasi dan 11 atribut atau variabel yang memiliki relevansi terhadap konstruksi kondisi psikologis individu, khususnya dalam mengidentifikasi tingkat depresi.

Setelah konversi format selesai, langkah berikutnya adalah penerapan tahapan data preprocessing melalui serangkaian prosedur yang difasilitasi oleh operator dan fungsi pada Google Colab. Langkah awal dalam proses ini adalah penetapan peran masing-masing atribut melalui operator Set Role, yang digunakan untuk mendefinisikan fungsi setiap variabel dalam kerangka analisis. Dalam konteks ini, atribut "Depression" secara eksplisit ditetapkan sebagai label atau target variable, yakni variabel dependen yang menjadi fokus utama dalam proses klasifikasi. Variabel ini bersifat kategorikal dan memiliki dua nilai diskret,

yaitu Yes, yang menunjukkan keberadaan gejala depresi pada individu, serta No, yang merepresentasikan tidak adanya gejala depresi.

Penetapan atribut "Depression" sebagai variabel target merupakan langkah fundamental dalam membangun kerangka supervised learning, karena secara langsung menentukan struktur pembelajaran model dan arah estimasi prediktif. Dengan mendefinisikan variabel target secara jelas sejak awal, model dapat dilatih untuk mengidentifikasi pola dan hubungan antara atribut-atribut independen (predictors) terhadap status depresi sebagai keluaran yang diharapkan. Selain itu, tahap ini juga memastikan bahwa seluruh proses selanjutnya mulai dari pemilihan algoritma, evaluasi kinerja, hingga validasi hasil dilandaskan pada relasi yang sah antara fitur input dan label output.

Proses preprocessing tidak hanya terbatas pada penetapan peran variabel, tetapi juga mencakup kegiatan lain seperti pembersihan data (data cleaning), transformasi nilai variabel (termasuk pengkodean variabel kategorikal jika diperlukan), dan normalisasi atau standarisasi data numerik. Semua prosedur ini dilakukan untuk memastikan kualitas data yang optimal sebelum digunakan dalam tahap pemodelan, sehingga dapat meminimalkan potensi bias atau kesalahan sistematis yang dapat mempengaruhi performa algoritma klasifikasi yang diterapkan.

	A	B	C	D	E	F	G	H	I	J	K
	Gender	Age	Academic Phase	Study Satisfaction	Sleep Duration	Dietary Habits	Have you ever been to a doctor	Study Hours	Financial Stress	Family History	Depression
1	Male	20	2	4	7-8 hours	Healthy	Yes	6	2	No	No
2	Male	25	4	5	5-6 hours	Healthy	Yes	7	1	No	No
3	Male	20	1	3	5-6 hours	Unhealthy	Yes	10	4	No	No
4	Male	25	1	4	More than 8 hours	Unhealthy	Yes	7	2	No	No
5	Female	31	1	4	More than 8 hours	Healthy	Yes	4	2	No	No
6	Female	19	4	4	5-6 hours	Unhealthy	Yes	1	4	No	No
7	Female	24	4	2	More than 8 hours	Moderate	Yes	6	2	No	No
8	Female	20	4	1	More than 8 hours	Healthy	Yes	3	4	No	No
9	Female	33	1	4	More than 8 hours	Moderate	No	10	3	No	No
10	Male	31	4	3	Less than 5 hours	Unhealthy	Yes	10	3	No	No
11	Female	31	5	4	5-6 hours	Healthy	Yes	6	4	No	No
12	Male	24	2	1	7-8 hours	Unhealthy	Yes	11	3	No	No
13	Female	23	5	3	Less than 5 hours	Unhealthy	Yes	2	1	No	No
14	Male	25	1	1	5-6 hours	Moderate	Yes	12	3	No	No
15	Male	21	5	1	More than 8 hours	Unhealthy	Yes	3	3	No	No
16	Male	28	5	3	5-6 hours	Healthy	Yes	8	3	No	No
17	Male	23	5	2	More than 8 hours	Moderate	No	10	4	No	No
18	Female	23	1	3	Less than 5 hours	Healthy	Yes	5	3	No	No
19	Female	20	5	4	More than 8 hours	Unhealthy	Yes	2	4	No	No
20	Male	29	4	3	More than 8 hours	Unhealthy	Yes	1	3	No	No
21	Male	31	2	3	More than 8 hours	Unhealthy	No	3	3	No	No
22	Male	31	2	4	More than 8 hours	Healthy	No	1	3	No	No
23	Male	31	2	4	More than 8 hours	Unhealthy	No	10	1	No	No
24	Female	31	2	2	7-8 hours	Moderate	No	11	3	No	No
25	Female	31	2	2	7-8 hours	Moderate	No	12	3	No	No
26	Male	31	2	2	7-8 hours	Healthy	No	2	4	No	No
27	Male	30	2	2	7-8 hours	Moderate	Yes	6	3	No	No
28	Male	21	5	3	7-8 hours	Unhealthy	No	6	4	No	No
29	Female	20	3	3	Less than 5 hours	Moderate	Yes	4	3	No	No
30	Female	34	3	4	Less than 5 hours	Unhealthy	Yes	12	1	No	No
31	Female	29	3	2	More than 8 hours	Healthy	No	2	3	No	No

Gambar 3. Dataset yang Telah Dilakukan Convert Format Excel

Tahap berikutnya dalam proses preprocessing adalah transformasi data, yang dilakukan menggunakan operator Nominal to Binominal untuk mengonversi variabel-variabel kategorikal ke dalam format biner. Transformasi ini esensial, terutama untuk atribut seperti Gender (Male/Female), Sleep Duration (Less than 5 hours, 5–6 hours, 7–8 hours, More than 8 hours), Dietary Habits (Healthy, Moderate, Unhealthy), serta variabel kategorikal lainnya. Prosedur ini diperlukan karena algoritma Decision Tree meskipun fleksibel terhadap data kategorikal akan beroperasi lebih efisien dan akurat apabila seluruh input berada dalam format numerik atau biner yang dapat ditafsirkan secara eksplisit oleh

model. Sementara itu, untuk variabel bertipe numerik seperti Age, Academic Pressure, Study Satisfaction, Study Hours, dan Financial Stress, dilakukan proses normalisasi menggunakan operator Normalize dengan metode Z-transformation. Tujuan utama dari normalisasi ini adalah untuk menyamakan skala nilai antarvariabel numerik, sehingga tidak terdapat variabel yang secara tidak proporsional memengaruhi pembentukan model akibat perbedaan rentang nilai. Melalui Z-transformation, data diubah sedemikian rupa agar memiliki nilai rata-rata (mean) sebesar 0 dan deviasi standar (standard deviation) sebesar 1, yang secara signifikan dapat meningkatkan kestabilan dan kinerja algoritma klasifikasi yang digunakan.

### Implementasi Decision Tree

Dalam rangka mengoptimalkan performa algoritma Decision Tree dalam proses klasifikasi tingkat depresi, penelitian ini menerapkan serangkaian parameter utama yang telah disesuaikan secara cermat guna meningkatkan akurasi serta kemampuan generalisasi model. Salah satu parameter kunci yang digunakan adalah criterion, yaitu metode evaluasi untuk menentukan atribut pemisah terbaik pada setiap node. Dalam konteks ini, metode yang dipilih adalah gain ratio, yang merupakan penyempurnaan dari pendekatan information gain konvensional. Gain ratio mengombinasikan nilai information gain dengan split information, yaitu ukuran yang mempertimbangkan jumlah dan sebaran nilai dari atribut yang diuji, sehingga memberikan informasi tambahan mengenai signifikansi atribut tersebut dalam proses pemisahan data.

Penerapan gain ratio bertujuan untuk mengatasi bias inheren dalam metode information gain, yang cenderung memilih atribut dengan jumlah kategori terbanyak sebagai pemisah utama, meskipun kontribusi informatifnya terhadap target klasifikasi relatif rendah. Bias semacam ini sering kali berujung pada overfitting, yaitu kondisi di mana model terlalu menyesuaikan diri terhadap karakteristik khusus data pelatihan, namun kehilangan kemampuan untuk menggeneralisasi pola-pola tersebut pada data baru. Dengan demikian, penggunaan gain ratio tidak hanya memberikan keunggulan dalam menghasilkan pohon keputusan yang lebih seimbang secara struktural, tetapi juga memperkuat representasi distribusi data secara keseluruhan, sehingga meningkatkan validitas dan reliabilitas hasil klasifikasi terhadap variabel target, yakni identifikasi kondisi depresi.

Untuk menjaga kompleksitas model tetap terkendali dan menghindari pembentukan struktur pohon yang terlalu dalam, parameter maximum depth ditetapkan pada nilai 10. Penetapan batas kedalaman ini bertujuan untuk membatasi jumlah level pemisahan yang

dapat terbentuk dalam hierarki pohon keputusan. Pembatasan ini memiliki implikasi penting dalam menghindari overfitting, karena pohon dengan kedalaman yang tidak terkontrol cenderung menangkap noise atau anomali dalam data pelatihan yang tidak relevan terhadap pola sebenarnya. Sebaliknya, pohon dengan kedalaman yang optimal mampu mengidentifikasi fitur-fitur utama yang relevan secara sistematis, tanpa mengorbankan akurasi pada saat pengujian terhadap data baru.

### Evaluasi Model

Berdasarkan hasil evaluasi yang diperoleh dari proses pengujian, model klasifikasi menunjukkan performa yang sangat baik dengan tingkat akurasi mencapai 89%. Nilai akurasi tersebut mengindikasikan bahwa model mampu mengklasifikasikan status depresi secara tepat pada 89% dari keseluruhan data uji yang digunakan. Artinya, proporsi prediksi yang benar baik dalam mengidentifikasi individu yang mengalami depresi (true positive) maupun yang tidak mengalami depresi (true negative) sangat dominan dibandingkan dengan kesalahan klasifikasi. Pencapaian ini mencerminkan tingkat presisi dan efektivitas model yang tinggi dalam membedakan antara kategori kelas yang berbeda, sehingga meningkatkan keandalan model dalam konteks diagnosis awal atau proses skrining gangguan depresi. Dalam aplikasi praktis, performa seperti ini sangat penting, terutama dalam konteks klinis, karena membantu meminimalkan risiko false classification yang dapat berdampak pada kualitas intervensi atau pengambilan keputusan terapeutik.

```
[503 rows x 11 columns]

Evaluasi Model

Accuracy: 0.89
Precision: 0.89
Recall: 0.89
F1-Score: 0.89
Confusion Matrix:
[[43  3]
 [ 8 47]]
```

Gambar 4. Evaluasi Model

Analisis lanjutan yang dilakukan melalui confusion matrix memberikan pemahaman yang lebih komprehensif terkait performa model klasifikasi dalam mengidentifikasi status depresi. Berdasarkan hasil evaluasi, model menunjukkan kemampuan yang cukup baik dalam membedakan antara kasus positif dan negatif. Tercatat sebanyak 43 kasus true negative, yaitu individu yang tidak mengalami depresi dan berhasil diprediksi secara tepat oleh model, serta 3 kasus true positive, yakni individu yang mengalami depresi dan

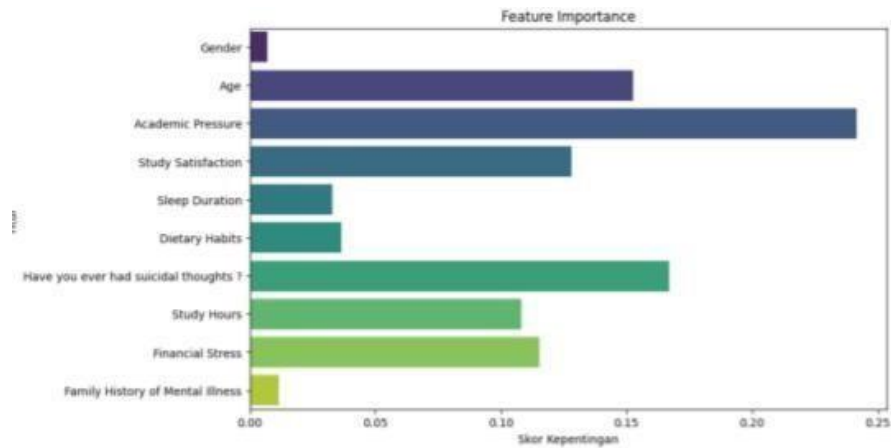
diklasifikasikan secara akurat. Namun demikian, model juga menghasilkan 8 kasus false positive, di mana individu yang sebenarnya tidak mengalami depresi secara keliru diidentifikasi sebagai penderita depresi. Selain itu, terdapat 47 kasus false negative, yaitu kondisi ketika model gagal mendeteksi individu yang sebetulnya mengalami depresi.

Tingkat kesalahan klasifikasi yang muncul terutama pada kasus false negatif perlu menjadi perhatian dalam pengembangan lanjutan model, mengingat konsekuensi dari kegagalan mendeteksi individu yang mengalami depresi dapat berdampak serius terhadap upaya penanganan awal. Meskipun demikian, secara umum, hasil ini menunjukkan bahwa model memiliki reliabilitas yang cukup tinggi dalam proses klasifikasi status depresi, khususnya dalam menghindari kesalahan positif. Dengan demikian, model ini berpotensi dimanfaatkan sebagai instrumen skrining awal dalam konteks klinis maupun non-klinis, terutama oleh tenaga profesional dalam mendeteksi gejala gangguan depresi pada tahap awal intervensi.

#### Analisis Fitur yang Berpengaruh

Setelah model Decision Tree berhasil dikonstruksi dan divalidasi, langkah selanjutnya adalah melakukan analisis mendalam terhadap tingkat kontribusi masing-masing fitur atau variabel masukan (input features) terhadap hasil klasifikasi. Tujuan dari tahap ini adalah untuk mengidentifikasi gejala-gejala atau faktor-faktor yang memiliki pengaruh paling signifikan dalam menentukan status kondisi psikologis siswa, khususnya dalam konteks deteksi dini depresi. Analisis ini dilakukan dengan mengukur tingkat kepentingan (feature importance score) dari setiap atribut yang digunakan dalam pelatihan model.

Hasil analisis tersebut divisualisasikan dalam bentuk diagram batang horizontal (lihat Gambar 5), yang menampilkan urutan fitur berdasarkan besarnya kontribusi terhadap model prediksi. Berdasarkan hasil visualisasi tersebut, ditemukan bahwa beberapa fitur memiliki skor pengaruh yang menonjol dibandingkan fitur lainnya. Lima fitur teratas yang menunjukkan tingkat kontribusi paling tinggi antara lain:



Gambar 5. Tingkat Kontribusi Tertinggi

- Academic Pressure (Tekanan Akademik)
- Have you ever had suicidal thoughts? (Pernahkah Anda memiliki pikiran untuk bunuh diri?)
- Age (Usia)
- Study Satisfaction (Tingkat Kepuasan dalam Belajar)
- Financial Stress (Stres Finansial)

Kelima fitur ini mencerminkan dimensi yang kompleks dalam kehidupan siswa, yang saling berinteraksi dan berdampak langsung terhadap kesejahteraan psikologis mereka. Tekanan akademik, misalnya, secara konsisten ditemukan dalam berbagai literatur psikologis sebagai salah satu faktor stres utama di kalangan pelajar, terutama ketika ekspektasi akademik yang tinggi tidak diimbangi dengan dukungan emosional atau kapasitas manajemen stres yang memadai. Demikian pula, riwayat pikiran bunuh diri merupakan indikator yang sangat penting dalam asesmen risiko depresi, karena sering kali menjadi manifestasi dari tekanan psikologis yang mendalam dan tidak tertangani.

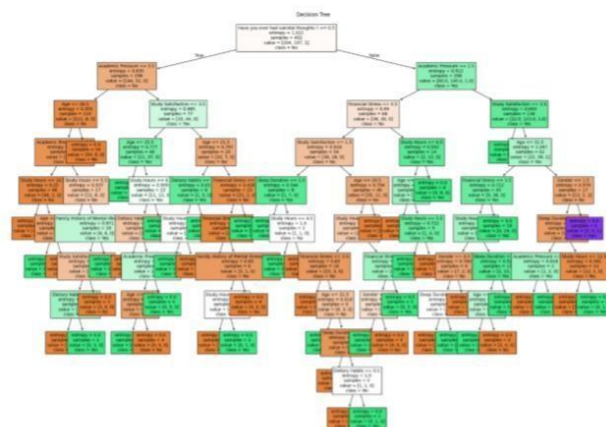
Usia juga memainkan peran penting dalam klasifikasi, mengingat rentang usia remaja dan dewasa awal merupakan fase perkembangan yang rentan terhadap gangguan emosional akibat transisi kehidupan, tekanan sosial, dan pencarian jati diri. Sementara itu, kepuasan terhadap proses belajar dapat merefleksikan sejauh mana siswa merasa terlibat secara positif dengan aktivitas akademik mereka; tingkat kepuasan yang rendah sering kali berkorelasi dengan perasaan tidak berdaya atau kehilangan motivasi, yang merupakan gejala umum dari depresi. Faktor terakhir, yakni stres finansial, menjadi aspek kontekstual yang turut memperburuk kondisi mental siswa, terutama bagi mereka yang harus menghadapi tekanan ekonomi keluarga atau tanggung jawab keuangan pribadi di usia muda.



Secara keseluruhan, temuan ini menunjukkan bahwa kombinasi antara faktor internal (seperti usia dan kepuasan belajar) dan eksternal (seperti tekanan akademik dan finansial) berkontribusi secara signifikan terhadap kemungkinan seorang siswa mengalami depresi. Hasil ini juga sejalan dengan temuan-temuan empiris dalam bidang psikologi pendidikan dan kesehatan mental remaja, yang menekankan pentingnya pendekatan multidimensional dalam memahami determinan kesejahteraan psikologis pelajar. Dengan demikian, hasil analisis ini tidak hanya memperkuat validitas model klasifikasi yang dibangun, tetapi juga memberikan kontribusi penting bagi upaya intervensi berbasis data yang dapat dilakukan oleh tenaga pendidik, konselor, dan pembuat kebijakan di sektor pendidikan.

### Struktur Pohon Keputusan

Untuk memahami cara kerja model dalam mengambil keputusan secara sistematis dan transparan, dilakukan visualisasi terhadap struktur pohon keputusan (decision tree) yang dihasilkan dari proses pelatihan data. Visualisasi ini bertujuan untuk memperlihatkan alur logika yang diikuti model dalam mengklasifikasikan kondisi psikologis siswa berdasarkan atribut-atribut yang diberikan.



Gambar 6. Struktur Pohon Keputusan

Pohon keputusan dimulai dari simpul akar (root node), yang merupakan fitur dengan tingkat pengaruh paling tinggi terhadap hasil klasifikasi, yaitu pertanyaan: "Have you ever had suicidal thoughts?" (Apakah Anda pernah memiliki pikiran untuk bunuh diri?). Pemilihan fitur ini sebagai titik awal mengindikasikan bahwa variabel tersebut memiliki nilai prediktif yang sangat tinggi dalam membedakan antara kategori depresi dan non-depresi. Dari simpul akar ini, pohon kemudian bercabang ke simpul-simpul berikutnya berdasarkan variabel lain seperti tekanan akademik (academic pressure), tingkat kepuasan belajar (study satisfaction), stres finansial (financial stress), dan usia (age), hingga akhirnya mencapai

simpul terminal (leaf nodes) yang mewakili kelas akhir dari prediksi, seperti "tidak mengalami depresi", "depresi ringan", "depresi sedang", atau "depresi berat".

Struktur hierarkis pohon ini mencerminkan alur pengambilan keputusan yang logis dan mudah diikuti, menjadikan Decision Tree sebagai salah satu metode klasifikasi yang sangat dapat diinterpretasikan (interpretable) dibandingkan dengan model-model black box seperti neural networks atau support vector machines. Kelebihan ini memberikan nilai tambah yang signifikan dalam konteks implementasi sistem di lingkungan non-teknis, seperti di sekolah, pusat layanan konseling, atau komunitas pendidikan, di mana pengguna akhir seperti guru Bimbingan Konseling (BK), konselor, maupun orang tua memerlukan penjelasan yang jelas dan masuk akal terkait alasan di balik keputusan yang diberikan oleh sistem.

## SIMPULAN

Penelitian ini berhasil mengimplementasikan algoritma Decision Tree secara efektif dalam mengklasifikasikan status depresi pada kalangan pelajar, berdasarkan pemanfaatan dataset yang dikurasi dari platform terbuka Kaggle. Dataset yang terdiri dari 503 entri dan 11 variabel independen merepresentasikan keragaman faktor psikososial yang relevan dengan kesehatan mental pelajar, seperti tekanan akademik, pola tidur, kebiasaan makan, dan tekanan finansial. Kompleksitas multidimensional data ini menjadi fondasi yang kokoh bagi pengembangan model prediktif yang akurat.

Tahapan preprocessing data dilakukan secara sistematis, dimulai dari konversi format file ke Excel untuk kompatibilitas yang lebih baik di Google Colab, dilanjutkan dengan penetapan peran atribut, transformasi variabel kategorikal menjadi biner, serta normalisasi variabel numerik dengan metode Z-transformation. Rangkaian proses ini memastikan bahwa data yang digunakan memenuhi standar kesiapan analitik untuk pemodelan berbasis pembelajaran mesin. Dalam implementasi model, algoritma Decision Tree dioptimalkan menggunakan kriteria gain\_ratio untuk pemilihan atribut, yang terbukti lebih efektif dibandingkan information gain konvensional karena mengurangi bias terhadap atribut dengan jumlah kategori yang besar. Selain itu, pembatasan kedalaman pohon pada tingkat maksimum 10 dilakukan untuk menghindari kompleksitas berlebihan dan meningkatkan generalisasi model terhadap data baru.

Hasil evaluasi menunjukkan bahwa model mencapai tingkat akurasi yang sangat tinggi, yaitu 89%, dengan distribusi prediksi yang menunjukkan performa sangat baik dalam membedakan kasus depresi dan non-depresi. Nilai true positive dan true negative

yang dominan, serta jumlah kesalahan klasifikasi (false positive dan false negative) yang minimal, mengindikasikan bahwa model ini sangat andal dan aplikatif sebagai alat bantu deteksi awal gejala depresi.

#### DAFTAR PUSTAKA

- Primadona., Rahmat, F. (2023). Penerapan Data Mining Pada Penjualan Produk Elektronik. *Jurnal Comasie*, 9(4): 463-472.
- Nasution, A.N., Ardilla, S., Zulfahmi, I. (2024). Implementasi Algoritma Kriptografi Modern melalui Google Colab: Studi Kasus AES dan RSA. *MOTEKAR: Jurnal Multidisiplin Teknologi dan Arsitekt*, 2(2): 841-845.
- Ramadhan, P., Yuhandri., Jhon, V. (2025). Eksplorasi Algoritma Decision Tree untuk Penentuan Siswa Berprestasi. *Bit-Tech (Binary Digital - Technology)*, 7(3): 823-833.
- Aini, Y.N., Ahmad, F., Gifthera, D. (2025). Penerapan Metode Decision Tree dalam Penentuan Jurusan Siswa. *Jurnal Ilmiah Informatika*, 13(1): 8-12.
- Rismayani., Samsul, A., Andi, H.E., Hasyrif, S., Noor, E. (2024). Identifikasi Gangguan Kesehatan Mental Pada Remaja Generasi Z Menggunakan Artificial Neural Network. *JUSTIN (Jurnal Sistem dan Teknologi Informasi)* , 12(4): 776-783.
- Arirahmayanti, I.G.A., Ardani, I.G.A., Aji, I.P.D.K. (2024). Tatalaksana Non-farmakologi Pada Depresi Remaja : Laporan Kasus. *HEALTHY : Jurnal Inovasi Riset Ilmu Kesehatan*, 3(4): 202-210.
- Abrori, S., Zaehol, F. (2024). Implementasi Metode Decission Tree Dalam Mengklasifikasi Depresi Menggunakan Rapidminer. *Jurnal JSRCS*, 5(2): 123-132.
- Latifah, Retnani., Emi, S.W., Priadhana, E.K. (2019). Model Decision Tree untuk Prediksi Jadwal Kerja menggunakan Scikit-Learn. *Seminar Nasional Sains dan Teknologi Fakultas Teknik Universitas Muhammadiyah Jakarta*.
- Farisi, A., Zaehol, F. (2024). Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Menggunakan Metode K-Means Clustering. *Gudang Jurnal Multidisiplin Ilmu*, 2(10): 205-212.
- Hakim, M.G.A., Faldy, I. (2024). Analisis Sentimen Terhadap Ulasan Pengguna Pada Aplikasi BCA Mobile Menggunakan Metode Naïve Bayes. *Journal of Information System Research (JOSH)*, 5(4): 911-921.
- Setiawan, I., Renata, F.A.C., Irfan, S. (2023). Exploring Complex Decision Trees: Unveiling Data Patterns And Optimal Predictive Power. *Jurnal Innovation and Future Technology (IFTECH)*, 5(2): 112-123.

- Bahri, S., Akhyar, L. (2020). Metode Klasifikasi Decision Tree Untuk Memprediksi Juara English Premier League. *Jurnal Sintaksis: Pendidikan Guru Sekolah Dasar, IPA, IPS dan Bahasa Inggris*, 2(1): 63-70.
- Samsul, M. (2023). *Modul Model Penelitian Kualitatif*. Jakarta: Universitas Muhammadiyah Prof. Dr. Hamka.
- Nasrullah, A.H. (2021). Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris. *Jurnal Ilmiah Ilmu Komputer*, 7(2): 45-51.
- Gisela, E.S., Emmanuella, A.K., Auliana, S. Untung, S. (2025). Pengaruh Stres Akademik terhadap Kesejahteraan Psikologis Mahasiswa Semester Akhir yang Terlambat Lulus. *AKADEMIK Jurnal Mahasiswa Humanis*, 5(1): 331-341.
- Ramadani, I.R., Tryana, F., Bima, K.R. (2024). Depresi, Penyebab Dan Gejala Depresi. *BERSATU: Jurnal Pendidikan Bhinneka Tunggal Ika*, 2(4): 88-99.