

PREDIKSI KESEHATAN MENTAL MAHASISWA UNIVERSITAS YUDHARTA PASURUAN DENGAN DECISION TREE

Choirun Nisa^{1*}, Muhammad Imron Rosyadi²,

¹Universitas Yudharta Pasuruan No.7, Kembangkuning, Songonagung, Kec. Purwosari, Pasuruan

Keywords:

Decision Tree;
Kesehatan Mental;
Mahasiswa;
Naïve Bayes;
RapidMiner

Correspondent Email:

Choirunnisak253@gmail.com.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstrak. Kesehatan mental mahasiswa menjadi perhatian penting di lingkungan pendidikan tinggi karena berdampak langsung terhadap prestasi akademik dan kualitas hidup mereka. Penelitian ini bertujuan untuk membangun sistem prediksi tingkat kesehatan mental mahasiswa Universitas Yudharta Pasuruan menggunakan algoritma klasifikasi *Decision Tree* dengan 3 kriteria yaitu *gain ratio*, *information gain* dan *gini indeks* serta peneliti membandingkan dengan algoritma *naive bayes*. Data diperoleh dari 215 responden melalui kuesioner DASS-21 yang mengukur tiga aspek utama: depresi, kecemasan, dan stres. Data tersebut diproses dan diklasifikasikan ke dalam lima kategori kondisi mental: Normal, Ringan, Sedang, Berat, dan Sangat Berat. Proses pengolahan dilakukan melalui tahapan *preprocessing data*, penentuan label kelas, hingga implementasi algoritma *Decision Tree* 3 kriteria dan algoritma *Naive Bayes* menggunakan *RapidMiner*. Evaluasi performa model dilakukan dengan *confusion matrix*, akurasi prediksi, serta nilai koefisien Kappa. Hasil penelitian menunjukkan bahwa mayoritas dari 215 responden berada dalam kategori Sangat Berat. Di antara ketiga kriteria tersebut, *Information Gain* menghasilkan performa terbaik dengan akurasi 94,41%, *precision* 93,7%, *recall* 92,8%, dan *F1-score* 93,2%. Sistem ini diharapkan dapat membantu deteksi dini masalah kesehatan mental mahasiswa dan mendukung layanan konseling kampus secara lebih efektif.

Abstract. Students' mental health is an important concern in the higher education environment because it has a direct impact on their academic achievement and quality of life. This study aims to build a system for predicting the level of mental health of students of Yudharta University Pasuruan using the *Decision Tree* classification algorithm with 3 criteria, namely *gain ratio*, *information gain* and *Gini index* and the researcher compares with the *naive Bayes* algorithm. Data were obtained from 215 respondents through the DASS-21 questionnaire that measured three main aspects: depression, anxiety, and stress. The data is processed and classified into five categories of mental conditions: Normal, Mild, Moderate, Severe, and Very Severe. The processing process is carried out through the stages of data preprocessing, determination of class labels, to the implementation of the *Decision Tree* 3 criteria algorithm and the *Naive Bayes* algorithm using *RapidMiner*. The evaluation of the model performance was carried out with a *confusion matrix*, prediction accuracy, and Kappa coefficient values. The results showed that the majority of the 215 respondents were in the Very Heavy category. Among these three criteria, *Information Gain* produces the best performance with an accuracy of 94.41%, precision of 93.7%, recall of 92.8%, and an *F1-score* of 93.2%. This system is expected to help early detection of student mental health problems and support campus counseling services more effectively..

1. PENDAHULUAN

Kesehatan mental mahasiswa merupakan isu krusial di lingkungan perguruan tinggi karena berdampak langsung terhadap pencapaian akademik, kualitas hidup, serta produktivitas mahasiswa. Tekanan akademik, persoalan ekonomi, kesulitan dalam penyesuaian sosial, dan ketidakpastian masa depan menjadi faktor utama yang berkontribusi terhadap gangguan psikologis seperti stres, kecemasan, dan depresi. Studi oleh [1] menunjukkan bahwa lebih dari 40% mahasiswa mengalami gejala stres dan depresi yang signifikan, sehingga menekankan pentingnya deteksi dini terhadap kondisi tersebut.

Seiring berkembangnya teknologi, pendekatan berbasis machine learning semakin banyak digunakan untuk mendukung sistem prediksi kesehatan mental yang bersifat preventif. Salah satu instrumen yang telah teruji validitas dan reliabilitasnya dalam mengukur tingkat depresi, kecemasan, dan stres adalah DASS-21 (*Depression Anxiety Stress Scales*) yang dikembangkan oleh Lovibond & Lovibond (1995). Instrumen ini terdiri dari 21 item yang dibagi secara merata untuk mengukur ketiga aspek tersebut dan telah digunakan secara luas dalam berbagai penelitian lintas negara [2].

Dalam bidang klasifikasi data, algoritma *Decision Tree* dan *Naïve Bayes* merupakan salah satu metode yang paling banyak digunakan karena kemudahannya dalam interpretasi dan efektivitasnya dalam menangani data kompleks. Penelitian sebelumnya menunjukkan bahwa algoritma ini mampu memberikan akurasi tinggi dalam klasifikasi kondisi mental mahasiswa. Selain itu, pemilihan atribut yang relevan seperti kebiasaan tidur, aktivitas sosial, dan kelelahan terbukti dapat meningkatkan performa model prediksi. Penelitian lain membandingkan performa beberapa algoritma klasifikasi dalam mendeteksi tingkat depresi mahasiswa, termasuk *Decision Tree* dan *Naïve Bayes*. Hasil penelitian tersebut menunjukkan bahwa *Decision Tree* memberikan hasil yang kompetitif dalam klasifikasi tingkat depresi berdasarkan atribut-atribut psikologis mahasiswa [3].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membangun sistem klasifikasi tingkat kesehatan mental

mahasiswa Universitas Yudharta Pasuruan dengan menggunakan algoritma *Decision Tree* dan *Naïve Bayes* serta instrumen DASS-21. Sistem ini diharapkan dapat menjadi alat bantu dalam melakukan deteksi dini terhadap mahasiswa yang berisiko mengalami gangguan kesehatan mental, sehingga memungkinkan intervensi psikologis yang lebih cepat dan tepat.

2. TINJAUAN PUSTAKA

2.1. Kesehatan Mental Mahasiswa

Kesehatan mental mencakup aspek emosional, psikologis, dan sosial yang memengaruhi pola pikir, perilaku, serta interaksi seseorang dalam kehidupan sehari-hari [4]. Mahasiswa merupakan kelompok yang rentan mengalami gangguan mental seperti stres, kecemasan, dan depresi karena berada dalam fase transisi dewasa muda yang penuh tekanan akademik dan sosial. UNICEF (2024) mencatat bahwa gangguan mental seperti kecemasan, stres, depresi, dan panik adalah kondisi umum yang dapat mengganggu fungsi kehidupan sehari-hari bila tidak ditangani dengan tepat [5].

2.2. *Depression Anxiety Stress Scales* (DASS-21)

DASS-21 adalah instrumen psikologis yang dikembangkan untuk mengukur tiga kondisi emosional negatif: depresi, kecemasan, dan stres. Alat ini terdiri dari 21 item (7 item per skala) dan merupakan versi singkat dari DASS-42 [6]. Keunggulan DASS-21 terletak pada efisiensi waktu dan validitasnya yang telah teruji di berbagai negara [7].

2.3. *Decision Tree* C4.5

Decision Tree merupakan algoritma klasifikasi yang banyak digunakan dalam data mining karena memiliki kelebihan dalam hal interpretasi dan efisiensi proses. Struktur pohon dengan simpul dan cabang memudahkan pemahaman pola klasifikasi dari data yang dianalisis. Algoritma ini efektif untuk mengidentifikasi aturan klasifikasi dari dataset dan dapat digunakan untuk prediksi berbasis data historis [8].

2.4. RapidMiner

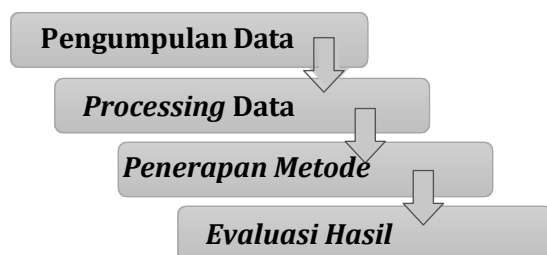
RapidMiner adalah perangkat lunak *open-source* yang mendukung analisis data dan pemodelan *machine learning* tanpa kebutuhan *coding*. Platform ini sangat berguna dalam implementasi algoritma klasifikasi seperti *Decision Tree* karena menyediakan antarmuka visual yang interaktif [9].

2.5. Kuesioner

Kuesioner merupakan metode pengumpulan data yang umum digunakan dalam penelitian kuantitatif. Instrumen ini memudahkan pengumpulan data dalam jumlah besar secara efisien, terutama dengan dukungan teknologi seperti *Google Form* yang dapat disebar melalui media sosial [10].

3. METODE PENELITIAN

Metode penelitian ini menggunakan algoritma *Decision Tree* dan *Naïve Bayes* yang diimplementasikan dalam *RapidMiner*. Data dikumpulkan dari 215 responden mahasiswa Universitas Yudharta Pasuruan menggunakan kuesioner DASS-21. Kuesioner merupakan salah satu instrumen pengumpulan data primer dengan metode survei yang fleksibel dan mudah digunakan untuk memperoleh opini responden. Data diklasifikasikan ke dalam lima kategori tingkat kesehatan mental [11]. Tahapan penelitian untuk model klasifikasi kesehatan mental mahasiswa Universitas Yudharta ditunjukkan pada gambar 3.1.



Gambar 3.1 Tahapan Penelitian

3.1 Pengumpulan Data

Data dikumpulkan dari 215 mahasiswa aktif Universitas Yudharta Pasuruan menggunakan kuesioner DASS-21. DASS-21 yaitu sebuah alat ukur yang digunakan untuk mengukur kondisi emosi dalam bentuk negatif dari depresi, kecemasan dan stress seseorang. Skala alat ukur ini dapat digunakan baik dalam

pengaturan klinis seperti rumah sakit atau non klinis seperti penelitian atau populasi Kuesioner terdiri dari 21 pertanyaan yang terbagi dalam tiga kategori: depresi, kecemasan, dan stres, masing-masing terdiri dari 7 butir soal. Setiap butir diukur menggunakan skala Likert 4 poin (0–3) [11].

3.2 Processing Data

Dalam tahap pengolahan data (*preprocessing*), dilakukan serangkaian proses untuk mempersiapkan data sebelum masuk ke tahap klasifikasi. Pertama, dilakukan seleksi atribut dengan hanya menggunakan tiga variabel utama, yaitu depresi, kecemasan, dan stres, yang dianggap paling relevan dalam menentukan kondisi kesehatan mental. Atribut lain seperti nama, jenis kelamin, dan jurusan tidak disertakan karena tidak memberikan kontribusi signifikan terhadap klasifikasi

Selanjutnya, dilakukan pemeriksaan data kosong (*missing values*), dan data yang tidak lengkap dihapus untuk menjaga integritas dataset. Setelah itu, dilakukan penentuan label kelas kesehatan mental dengan menjumlahkan skor dari ketiga variabel (depresi, kecemasan, stres), kemudian mengelompokkannya ke dalam lima kategori: Normal, Ringan, Sedang, Berat, dan Sangat Berat, berdasarkan tabel acuan skor DASS-21. Terakhir, normalisasi data dilakukan apabila diperlukan untuk memastikan skala nilai antar atribut seimbang dan mendukung performa algoritma klasifikasi secara optimal [12].

3.3 Penerapan Metode

Setelah data melalui tahap pra-pemrosesan, langkah selanjutnya adalah melakukan proses klasifikasi menggunakan algoritma *Decision Tree* dan *Naïve Bayes* pada platform *RapidMiner*. *RapidMiner* merupakan perangkat lunak yang dibuat oleh Dr. Markus Hofman dari *Institute of Teknologi Blanchardstown* dan Ralf Klinkenberg Software ini menyediakan banyak tools untuk persiapan data, pemodelan, evaluasi dan penerapannya [13]. Algoritma ini dipilih karena kemampuannya dalam membentuk model pohon keputusan yang mudah dipahami dan memiliki tingkat akurasi yang baik dalam klasifikasi data dan *naïve bayes* dalam pengklasifikasian data Kesehatan mental mahasiswa dapat digunakan secara terukur dan efisien sekalipun data yang diuji berjumlah

ribuan dan dengan variasi yang beragam. Proses klasifikasi dilakukan dengan memanfaatkan operator *Cross Validation* untuk membagi data ke dalam data latih dan data uji secara otomatis, sehingga memungkinkan evaluasi model secara lebih obyektif. Selain itu, digunakan juga operator *Performance* untuk menghitung dan menampilkan metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-score*, guna mengetahui seberapa baik model dalam mengklasifikasikan tingkat kesehatan mental mahasiswa.

3.4 Evaluasi Hasil

Setelah hasil klasifikasi diperoleh, selanjutnya hasil di evaluasi menggunakan *Confusion Matrix* untuk melihat *Accuracy*, *Precision* dan *Recall* yang dihasilkan oleh model yang diusulkan serta nilai *kappa coefficient*.

1. *Accuracy* mengukur seberapa sering model membuat prediksi yang benar, baik untuk kelas positif maupun negatif.

Rumus:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

2. *Precision* menunjukkan proporsi prediksi positif yang benar-benar positif.

Rumus:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

3. *Recall* (atau sensitivitas) mengukur seberapa baik model mendeteksi sampel positif dari seluruh sampel yang sebenarnya positif.

Rumus:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

4. *F1-score* adalah rata-rata harmonik dari *precision* dan *recall*. Metrik ini memberikan keseimbangan antara *precision* dan *recall*, terutama berguna ketika data tidak seimbang.

Rumus:

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

4. HASIL DAN PEMBAHASAN

4.1 Pengolahan data

1. Data Selection

Pada penelitian ini data yang digunakan berasal dari kuesioner mengenai kesehatan mental usia remaja. Data yang dipilih dan digunakan penulis yaitu atribut depresi, kecemasan, dan stress. Atribut *timestamp*, nama, jenis kelamin, semester dan jurusan dihapus oleh penulis dan tidak dipilih karena tidak berkaitan dengan tingkat kesehatan mental [14], hasil tersebut ditunjukkan pada gambar 4.1.

No.	Timestamp	Nama	Jenis Kelamin	Semester	Depresi	Kecemasan	Stres	Kategori Depresi	Kategori Kecemasan	Kategori Stres	Tingkat Kesehatan Mental
1	2021-06-18 17:27:35	Indah Laila	Perempuan	4	Subnormal Rendah	16	20	22	Ringan	Sangat Berat	Sangat Berat
2	2021-06-17 17:33:39	Shan Rizkiyah	Laki-Laki	2	Tinggi Infeksius	18	14	22	Sangat Berat	Ringan	Sangat Berat
3	2021-06-17 17:34:19	Yeni Nurhidayati	Perempuan	6	Subnormal Rendah	12	14	12	Sangat Berat	Sangat Berat	Sangat Berat
4	2021-06-17 17:34:39	Kusman Azzahra Purnama	Perempuan	6	Subnormal Rendah	8	8	14	Normal	Ringan	Indung
5	2021-06-17 17:34:59	Yuli Indrawati	Perempuan	4	Subnormal Rendah	20	18	20	Sangat Berat	Sangat Berat	Sangat Berat
6	2021-06-17 17:35:19	Yuliana	Perempuan	6	Subnormal	12	8	8	Ringan	Ringan	Normal
7	2021-06-17 17:35:39	Indah Nurhidayati	Perempuan	2	Subnormal Rendah	12	18	18	Ringan	Sangat Berat	Indung
8	2021-06-17 17:35:59	Shan Azzahra dan Yuli	Perempuan	4	Subnormal Rendah	14	14	18	Ringan	Sangat Berat	Sangat Berat
9	2021-06-17 17:36:19	Shan Azzahra dan Yuli	Perempuan	4	Subnormal Rendah	14	14	18	Ringan	Sangat Berat	Sangat Berat
10	2021-06-17 17:36:39	Yeni Nurhidayati	Perempuan	2	Subnormal Rendah	18	12	14	Ringan	Indung	Indung

Gambar 4. 1 Data Sebelum Seleksi

No.	Depresi	Kecemasan	Stres
1	16	26	22
2	18	14	22
3	32	16	32
4	6	8	14
5	20	18	20
6	12	8	6
7	12	18	16
8	14	14	10
9	8	12	20
...	...	...	...
215	10	12	14

Gambar 4. 2 Data Setelah Seleksi

Atribut tingkat kesehatan mental dibagi ke dalam lima kategori: Normal, Ringan, Sedang, Berat, dan Sangat Berat. Nilai kategori ini ditentukan berdasarkan penjumlahan skor depresi, kecemasan, dan stres, sesuai dengan acuan skor DASS-21. Hasil penjumlahan tersebut kemudian digunakan untuk membuat kolom baru yang merepresentasikan kelas target dalam proses klasifikasi [15].

Tabel 1 Score Nilai DASS-21

	Depresi	Kecemasan	Stress	Total
Normal	0-6	0-5	0-11	0-23
Ringan	7-8	6-7	12-13	24-29
Sedang	9-13	8-12	14-16	30-39
Berat	14-16	13-15	17-18	40-46
Sangat Berat	17+	16+	19+	47+

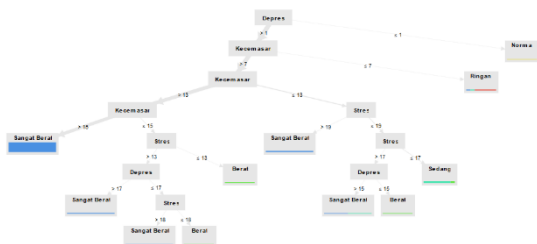
Setelah seleksi data, dilakukan integrasi data dengan menambahkan secara manual atribut tingkat kesehatan mental, yang ditentukan berdasarkan hasil penjumlahan skor depresi, kecemasan, dan stres sesuai kriteria pada tabel DASS-21.

No.	Depresi	Kecemasan	Stres	Tingkatan Kesehatan Mental
1	16	26	22	Sangat Berat
2	18	14	22	Sangat Berat
3	32	16	32	Sangat Berat
4	6	8	14	Sedang
5	20	18	20	Sangat Berat
6	12	8	6	Sedang
7	12	18	16	Sangat Berat
8	14	14	10	Berat
9	8	12	20	Sangat Berat
...	...	...	...	...
215	10	12	14	Sedang

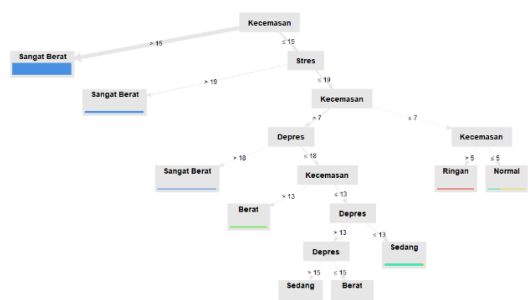
Gambar 4. 3 Dataset Setelah Proses Integrasi Data

2. Data Processing

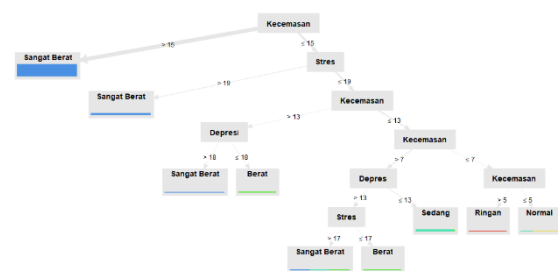
Setelah melalui tahapan *preprocessing* dan penerapan algoritma *Decision Tree* pada *RapidMiner*, diperoleh model pohon keputusan seperti ditunjukkan pada gambar 4.4 yang menyajikan serangkaian aturan untuk memprediksi kategori kesehatan mental mahasiswa, antara lain:



Gambar 4. 4 Tampilan Grafik *Decision Tree* Gain Ratio



Gambar 4. 5 Tampilan Grafik *Decision Tree* Information Gain



Gambar 4. 6 Tampilan Grafik *Decision Tree* Gini Index

3. Evaluasi Model

Pada tahap ini dilakukan evaluasi model menggunakan confusion matrix guna mengetahui *performance* dalam melakukan klasifikasi menggunakan algoritma *Decision Tree* dan *Naïve Bayes*. *Confusion matrix* ini merupakan ringkasan tabel jumlah prediksi benar yang sesuai dengan actual dan yang tidak sesuai. Tabel ini terdiri dari baris yang berisi nilai prediksi dan kolom yang berisi nilai actual. Model dapat dikategorikan baik jika yang dihasilkan mempunyai *True Positive* (TP) dan *True Negative* (TN) bernilai tinggi sedangkan *False Positive* (FP) dan *False Negative* (FN) bernilai rendah.

accuracy: 91.62% +/- 6.89% (micro average: 91.63%)

	true Sangat Berat	true Sedang	true Berat	true Normal	true Ringan	class precision
pred. Sangat B...	171	5	3	2	0	94.48%
pred. Sedang	2	15	3	0	0	75.00%
pred. Berat	1	1	6	0	0	75.00%
pred. Normal	0	0	0	0	0	0.00%
pred. Ringan	0	1	0	0	5	83.33%
class recall	98.28%	68.18%	50.00%	0.00%	100.00%	

Gambar 4. 7 Tampilan Hasil *Confusion Matrix* *Decision Tree* Gain Ratio

Berdasarkan *confusion matrix* pada gambar 4.7, model klasifikasi menghasilkan akurasi sebesar 91,62% dengan deviasi $\pm 6,89\%$. Hasil ini menunjukkan performa yang baik dalam memprediksi lima kategori tingkat kesehatan mental mahasiswa, yaitu Sangat Berat, Sedang, Berat, Normal, dan Ringan. Nilai class precision tertinggi dicapai oleh kategori Sangat Berat, yaitu 94,48%, sedangkan precision terendah ada pada kategori Normal sebesar 0,00%. Untuk nilai *recall*, kategori Ringan mencapai 100%, diikuti oleh Sangat Berat sebesar 98,28%. Kategori Normal memiliki recall terendah, yaitu 0,00%. Meskipun demikian, masih terdapat kesalahan prediksi, misalnya, 5 data yang seharusnya Sedang diprediksi sebagai Sangat Berat, dan 3 data yang seharusnya Berat diprediksi sebagai

Sangat Berat. Selain itu, ada 1 data yang seharusnya Ringan diprediksi sebagai Sedang. Secara keseluruhan, model ini menunjukkan potensi yang baik dalam klasifikasi kesehatan mental mahasiswa, meskipun masih ada ruang untuk peningkatan, terutama pada kategori Normal.

accuracy: 96.77% +/- 3.76% (micro average: 96.74%)

	true Sangat Berat	true Sedang	true Berat	true Normal	true Ringan	class precision
pred. Sangat B...	173	0	0	0	0	100.00%
pred. Sedang	1	20	2	2	0	80.00%
pred. Berat	0	1	10	0	0	90.91%
pred. Normal	0	0	0	0	0	0.00%
pred. Ringan	0	1	0	0	5	83.33%
class recall	99.43%	90.91%	83.33%	0.00%	100.00%	

Gambar 4. 8 Tampilan Hasil *Confusion Matrix Decision Tree Information Gain*

Berdasarkan *confusion matrix* pada Gambar 4.8, model klasifikasi menghasilkan akurasi sebesar 96,77% dengan deviasi $\pm 3,76\%$. Hasil ini menunjukkan performa yang sangat baik dalam memprediksi lima kategori tingkat kesehatan mental mahasiswa, yaitu Sangat Berat, Sedang, Berat, Normal, dan Ringan. Nilai *class precision* tertinggi dicapai oleh kategori Sangat Berat sebesar 100,00%, diikuti oleh kategori Berat sebesar 90,91%, dan Ringan sebesar 83,33%. Sementara itu, precision terendah terdapat pada kategori Normal sebesar 0,00%, yang menunjukkan bahwa tidak ada data yang berhasil diprediksi benar dalam kategori ini. Untuk nilai *recall*, kategori Sangat Berat mencapai 99,43%, diikuti oleh Ringan sebesar 100,00%, dan Sedang sebesar 90,91%. Sama seperti *precision*, *recall* terendah juga terdapat pada kategori Normal dengan nilai 0,00%. Secara keseluruhan, model ini menunjukkan potensi yang sangat baik dalam klasifikasi kesehatan mental mahasiswa, khususnya pada kategori dominan seperti Sangat Berat. Namun demikian, masih ada ruang perbaikan, terutama pada kategori Normal yang belum berhasil diprediksi dengan benar.

accuracy: 95.84% +/- 4.05% (micro average: 95.81%)

	true Sangat Berat	true Sedang	true Berat	true Normal	true Ringan	class precision
pred. Sangat B...	173	0	3	0	0	98.30%
pred. Sedang	1	20	1	2	0	83.33%
pred. Berat	0	1	8	0	0	88.89%
pred. Normal	0	0	0	0	0	0.00%
pred. Ringan	0	1	0	0	5	83.33%
class recall	99.43%	90.91%	66.67%	0.00%	100.00%	

Gambar 4. 9 Tampilan Hasil *Confusion Matrix Decision Tree Gini Index*

Berdasarkan *confusion matrix* pada Gambar 4.9, model klasifikasi menghasilkan

akurasi sebesar 95,84% dengan deviasi $\pm 4,05\%$. Hasil ini menunjukkan performa yang sangat baik dalam memprediksi lima kategori tingkat kesehatan mental mahasiswa, yaitu Sangat Berat, Sedang, Berat, Normal, dan Ringan. Nilai *class precision* tertinggi dicapai oleh kategori Sangat Berat sebesar 98,30%, diikuti oleh Berat sebesar 88,89%, dan Sedang sebesar 83,33%. *Precision* terendah terdapat pada kategori Normal sebesar 0,00%, karena tidak ada data yang berhasil diprediksi sebagai kategori tersebut. Sementara itu, nilai *recall* tertinggi dicapai oleh kategori Ringan sebesar 100,00%, diikuti oleh Sangat Berat dengan 99,43%, dan Sedang sebesar 90,91%. Kategori Normal kembali mencatatkan nilai *recall* terendah sebesar 0,00%, yang menunjukkan bahwa tidak ada data Normal yang berhasil diklasifikasikan dengan benar. Secara keseluruhan, model menunjukkan kinerja yang sangat baik dan cukup stabil, terutama pada kategori dominan seperti Sangat Berat. Namun demikian, prediksi pada kategori Normal masih menjadi kelemahan utama yang perlu ditingkatkan.

accuracy: 92.58% +/- 5.42% (micro average: 92.56%)

	true Sangat Berat	true Sedang	true Berat	true Normal	true Ringan	class precision
pred. Sangat B...	170	0	3	0	0	96.27%
pred. Sedang	2	17	2	2	0	73.91%
pred. Berat	2	5	7	0	0	50.00%
pred. Normal	0	0	0	0	0	0.00%
pred. Ringan	0	0	0	0	5	100.00%
class recall	97.70%	77.27%	58.33%	0.00%	100.00%	

Gambar 4. 10 Tampilan Hasil *Confusion Matrix Algoritma Naïve Bayes*

Berdasarkan *confusion matrix* pada Gambar 10, model klasifikasi menggunakan *Naive Bayes* menghasilkan akurasi sebesar 92,58% dengan deviasi $\pm 5,42\%$. Hasil ini menunjukkan performa yang cukup baik dalam memprediksi lima kategori tingkat kesehatan mental mahasiswa, yaitu Sangat Berat, Sedang, Berat, Normal, dan Ringan. Nilai *class precision* tertinggi dicapai oleh kategori Ringan sebesar 100,00%, diikuti oleh Sangat Berat sebesar 98,27%, dan Sedang sebesar 73,91%. *Precision* terendah terdapat pada kategori Normal sebesar 0,00%, karena tidak ada data yang berhasil diprediksi sebagai kategori tersebut. Sementara itu, nilai *class recall* tertinggi dicapai oleh kategori Ringan sebesar 100,00%, diikuti oleh Sangat Berat dengan 97,70%, dan Sedang sebesar 77,27%. Kategori Normal kembali mencatatkan nilai *recall* terendah sebesar 0,00%, yang menunjukkan

bahwa tidak ada data Normal yang berhasil diklasifikasikan dengan benar. Secara keseluruhan, model Naive Bayes menunjukkan kinerja yang cukup baik dengan akurasi di atas 90% dan performa stabil pada kategori dominan seperti Sangat Berat. Namun demikian, prediksi pada kategori Normal serta Berat masih menjadi kelemahan utama yang perlu ditingkatkan, karena kedua kategori ini belum terklasifikasi dengan optimal.

Pada tabel 2 menunjukkan hasil *accuracy*, *precision* dan *recall* yang didapat dari penerapan *decision tree* dan *naïve bayes* menggunakan *rapidminer*.

Tabel 2 Perbandingan Nilai *Decision Tree* dan *Naïve Bayes*

Matrix	Gain Ratio	Information Gain	Gini Indeks	Naive Bayes
Accuracy	91.63%	96.74%	95.81%	92.56%
Precision	92.06%	97.65%	95.85%	92.56%
Recall	95.81%	98.10%	96.26%	92.56%

Keempat model yang diuji *Gain Ratio*, *Information Gain*, *Gini Indeks*, dan *Naive bayes* menunjukkan kemampuan yang bervariasi dalam memprediksi data, seperti yang tercermin dari metrik *Accuracy*, *Precision*, dan *Recall*. Yang mana *Gain Ratio* menunjukkan performa yang solid di ketiga metrik. Model ini berhasil mencapai tingkat akurasi sebesar 91.63%, presisi 92.06%, dan *recall* 95.81%. Angka-angka ini mengindikasikan bahwa *Gain Ratio* adalah model yang cukup andal, mampu membuat prediksi yang relatif tepat dan komprehensif.

Selanjutnya, *Information Gain* muncul sebagai pemimpin dalam perbandingan ini, menunjukkan performa terbaik secara keseluruhan. Dengan presisi mencapai 97.65% dan *recall* 98.10%, *Information Gain* membuktikan keunggulannya dalam mengidentifikasi kelas positif dengan sangat akurat dan meliputi sebagian besar kasus positif yang sebenarnya. Akurasinya yang mencapai 96.74% juga menegaskan kemampuan model ini dalam membuat prediksi yang benar secara keseluruhan.

Selanjutnya, *Gini Index* juga menunjukkan performa yang sangat baik, bahkan mendekati keunggulan *Information Gain*. Model ini mencapai akurasi 95.82%, presisi 95.85%, dan *recall* 96.26%. Angka-angka ini menempatkan

Gini Index sebagai pilihan yang kuat, menunjukkan keseimbangan yang baik antara akurasi, presisi, dan *recall*.

Terakhir *Naïve Bayes* juga memperlihatkan performa yang kompetitif meskipun berada sedikit di bawah metode berbasis pohon keputusan. Model ini mencapai akurasi sebesar 92.56%, dengan presisi dan *recall* yang sama tinggi yaitu 92.56%. Konsistensi nilai presisi dan *recall* tersebut menunjukkan bahwa *Naive Bayes* mampu menjaga keseimbangan antara ketepatan dan kelengkapan prediksi. Meskipun tidak menjadi yang terbaik, algoritma ini tetap dapat dijadikan alternatif yang andal dalam melakukan klasifikasi.

Berdasarkan tabel 2, *Information Gain* menunjukkan performa terbaik dalam memprediksi data, dengan nilai *Accuracy*, *Precision*, dan *Recall* tertinggi dibandingkan dengan *Gain Ratio* dan *Gini Indeks*, dan *Naive Bayes*. Meskipun *Gini Index* juga memberikan hasil yang sangat baik, *Information Gain* sedikit unggul dalam semua aspek. *Gain Ratio*, meskipun memiliki performa yang baik, masih di bawah dua model lainnya.

5. KESIMPULAN

Penelitian ini dilakukan untuk merancang dan membangun sistem prediksi kondisi kesehatan mental mahasiswa Universitas Yudharta Pasuruan menggunakan algoritma *Decision Tree*. Berdasarkan hasil penelitian yang telah dilakukan dapat ditarik kesimpulan bahwa sistem prediksi kesehatan mental berbasis algoritma *Decision Tree* yang dibangun mampu mengklasifikasikan tingkat kesehatan mental mahasiswa ke dalam lima kategori, yaitu Normal, Ringan, Sedang, Berat, dan Sangat Berat, berdasarkan skor *Depression Anxiety Stress Scales (DASS-21)*.

Hasil pengolahan data menunjukkan bahwa dari 215 responden, mayoritas mahasiswa berada pada kategori Sangat Berat sebanyak 174 orang, diikuti kategori Sedang sebanyak 22 orang, kategori Ringan sebanyak 5 orang, dan kategori Normal sebanyak 2 orang. Dari tiga kriteria *Decision Tree* dan yang diuji, yaitu *Gain Ratio*, *Information Gain*, dan *Gini Index*, kriteria *Information Gain* memberikan performa terbaik dengan tingkat akurasi sebesar 98,14%, presisi tertinggi sebesar 100% pada kategori Sangat Berat dan Sedang, serta nilai

recall yang tinggi dengan tingkat kesalahan klasifikasi (false negative) yang minimal.

Analisis dari hasil wawancara memperkuat temuan tersebut, di mana faktor-faktor dominan yang memengaruhi kesehatan mental mahasiswa meliputi tekanan akademik, kurangnya dukungan sosial, kebiasaan hidup tidak sehat, masalah keluarga/ekonomi, serta penggunaan media sosial secara berlebihan. Mahasiswa dengan tingkat kecemasan dan stres yang tinggi cenderung memiliki kondisi mental yang lebih buruk, sejalan dengan hasil klasifikasi model *Decision Tree* yang menempatkan variabel kecemasan dan stres sebagai atribut paling berpengaruh.

Kombinasi analisis kuantitatif dan kualitatif ini memberikan gambaran yang lebih komprehensif mengenai kondisi kesehatan mental mahasiswa. Hasil ini juga menjadi dasar untuk merekomendasikan penguatan layanan bimbingan konseling, pelatihan manajemen stres dan waktu, serta kampanye kesadaran kesehatan mental secara berkelanjutan di lingkungan kampus.

Secara keseluruhan, penelitian ini menunjukkan bahwa algoritma Decision Tree, khususnya dengan pendekatan Information Gain, efektif digunakan untuk membangun sistem klasifikasi tingkat kesehatan mental mahasiswa. Sistem ini tidak hanya memberikan klasifikasi yang akurat, tetapi juga memiliki potensi untuk diintegrasikan dalam layanan bimbingan konseling universitas sebagai sarana deteksi dini terhadap potensi gangguan mental di kalangan mahasiswa. Meskipun demikian, sistem ini tidak dimaksudkan sebagai alat diagnosis medis, melainkan sebagai supporting tool yang dapat melengkapi proses asesmen psikologis secara profesional.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] K. Battista, K. A. Patte, L. Diao, J. A. Dubin, and S. T. Leatherdale, "Using Decision Trees to Examine Environmental and Behavioural Factors Associated with Youth Anxiety, Depression, and Flourishing," *Int. J. Environ. Res. Public Health*, vol. 19, no. 17, 2022, doi: 10.3390/ijerph191710873.
- [2] O. Ahmed, R. A. Faisal, S. M. A. H. M. Alim, T. Sharker, and F. A. Hiramoni, "The psychometric properties of the Depression Anxiety Stress Scale-21 (DASS-21) Bangla version," *Acta Psychol. (Amst.)*, vol. 223, p. 103509, 2022, doi: 10.1016/j.actpsy.2022.103509.
- [3] Y. Huang, S. Li, B. Lin, S. Ma, J. Guo, and C. Wang, "Early Detection of College Students' Psychological Problems Based on Decision Tree Model," *Front. Psychol.*, vol. 13, no. August, pp. 1–10, 2022, doi: 10.3389/fpsyg.2022.946998.
- [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".
- [5] K. JASMINE, "濟無No Title No Title No Title," *Penambahan Natrium Benzoat Dan Kalium Sorbat Dan Kecepatan Pengadukan Sebagai Upaya Penghambatan Reaksi Inversi Pada Nira Tebu*, vol. 05, no. 01, pp. 25–36, 2024.
- [6] H. Hamka and M. Yusuf, "Spiritualitas, Kecemasan, dan Stres selama Pandemi COVID-19," *J. Psikol. Islam dan Budaya*, vol. 6, no. 1, pp. 47–54, 2023, doi: 10.15575/jpib.v6i1.14145.
- [7] M. A. Hakim and N. V. Aristawati, "Mengukur depresi, kecemasan, dan stres pada kelompok dewasa awal di Indonesia: Uji validitas dan reliabilitas konstruk DASS-21," *J. Psikol. Ulayat*, vol. 10, no. 2, pp. 232–250, 2023, doi: 10.24854/jpu553.
- [8] Q. Ameliatus and Z. Fatah, "Klasifikasi Kelulusan Mahasiswa Menggunakan Metode Decision Tree Menggunakan Aplikasi Rapidminer," *J. Ris. Tek. Komput.*, vol. 1, no. 4, pp. 58–64, 2024.
- [9] M. Sholeh, E. K. Nurnawati, and U. Lestari, "Penerapan Data Mining dengan Metode Regresi Linear untuk Memprediksi Data Nilai Hasil Ujian Menggunakan RapidMiner," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 8, no. 1, pp. 10–21, 2023, doi: 10.14421/jiska.2023.8.1.10-21.
- [10] A. H. Suasapha, "Skala Likert Untuk Penelitian Pariwisata; Beberapa Catatan Untuk Menyusunnya Dengan Baik," *J. Kepariwisata*, vol. 19, no. 1, pp. 26–37, 2020, doi: 10.52352/jpar.v19i1.407.
- [11] A. Priya, S. Garg, and N. P. Tigga, "ScienceDirect ScienceDirect Predicting Anxiety , Depression and Stress in Modern Life using Predicting Anxiety , Depression and Stress in Modern Life using Machine Learning Algorithms Machine Learning

- Algorithms,” *PROCS*, vol. 167, no. 2019, pp. 1258–1267, 2020, doi: 10.1016/j.procs.2020.03.442.
- [12] K. Duangchaemkarn, “Machine Learning-Based Classification of Mental Health State Using the DASS-21 Profile,” no. November, 2024, doi: 10.1109/BMEiCON64021.2024.10896316.
- [13] A. Naik and L. Samant, “Correlation Review of Classification Algorithm Using Data Mining Tool: WEKA, Rapidminer, Tanagra, Orange and Knime,” *Procedia Comput. Sci.*, vol. 85, no. Cms, pp. 662–668, 2016, doi: 10.1016/j.procs.2016.05.251.
- [14] S. Widaningsih, “Penerapan Data Mining untuk Memprediksi Siswa Berprestasi dengan Menggunakan Algoritma K Nearest Neighbor,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 3, pp. 2598–2611, 2022, doi: 10.35957/jatisi.v9i3.859.
- [15] M. Makara-Studzińska, E. Tyburski, M. Załuski, K. Adamczyk, J. Mesterhazy, and A. Mesterhazy, “Confirmatory Factor Analysis of Three Versions of the Depression Anxiety Stress Scale (DASS-42, DASS-21, and DASS-12) in Polish Adults,” *Front. Psychiatry*, vol. 12, no. January, pp. 1–9, 2022, doi: 10.3389/fpsyt.2021.770532.