

SKRIPSI

**PENERAPAN METODE *NAIVE BAYES* DAN *RANDOM
OVERSAMPLING* UNTUK DETEKSI TINGKAT GANGGUAN
KECEMASAN**

***APLICATION OF NAIVE BAYES METHOD AND RANDOM
OVERSAMPLING FOR ANXIETY DISORDER DETECTION***



DISUSUN OLEH:

WARDIMAN

D0221345

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS SULAWESI BARAT
MAJENE**

2025

SKRIPSI

**PENERAPAN METODE *NAIVE BAYES* DAN *RANDOM
OVERSAMPLING* UNTUK DETEKSI TINGKAT GANGGUAN
KECEMASAN**

*Aplication Of Naive Bayes Method And Random Oversampling For
Anxiety Disorder Detection*

Diajukan untuk memenuhi sebagai persyaratan
Memeperoleh Sarjana Komputer



Disusun oleh:

**Wardiman
D0221345**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS SULAWESI BARAT
MAJENE**

2025

PERNYATAAN ORISINALITAS

SKRIPSI

**PENERAPAN METODE *NAIVE BAYES* DAN *RANDOM
OVERSAMPLING* UNTUK DETEKSI TINGKAT GANGGUAN
KECEMASAN**

Telah dipersiapkan dan disusun oleh:

WARDIMAN

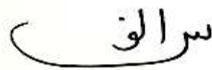
D0221345

Telah dipertahankan di depan Tim Penguji

Pada tanggal 15 Mei 2025

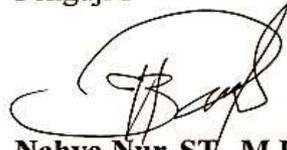
Susunan Tim Penguji

Pembimbing I



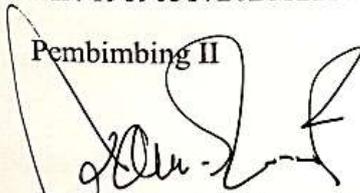
Dr. Eng. Sulfayanti, S.Si., M.T
NIP. 198903172020122011

Penguji I



Nahya Nur, ST., M.Kom
NIP. 199111052019032024

Pembimbing II



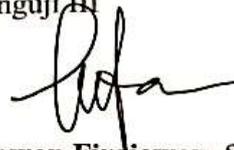
Farid Wajidi, S.Kom., M.T
NIP. 198904182019031018

Penguji II



Arnita Irianti, S.Si., M.Si
NIP.198708062018032001

Penguji III



Wawan Firgiawan, S.T., M.Kom
NIDK. 8948080023

PERNYATAAN ORISINALITAS

**PENERAPAN METODE *NAIVE BAYES* DAN *RANDOM OVERSAMPLING*
UNTUK DETEKSI TINGKAT GANGGUAN KECEMASAN**

SKRIPSI

Untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun Oleh

WARDIMAN

D0221007

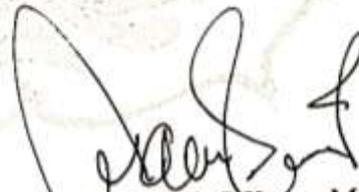
Telah diperiksa dan disetujui oleh:

Pembimbing I

Pembimbing II



Dr. Eng. Sulfayanti, S.Si., M.T
NIP. 198903172020122011



Parid Wajidi, S.Kom., M.T
NIP. 198904182019031018

Dekan Fakultas Teknik,

Ketua Program Studi Informatika



Prof. Dr. Ir. Hafsah Nirwana, M.T
NIP. 1964040519900322002



M. Nur Rasyid, S.Kom, M.T
NIP. 198808182022031006

PERNYATAAN ORISINALITAS

Menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, didalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah diajukan oleh oranglain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya, pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naska ini dan disebutkan dalam daftar referensi.

Apabila ternyata didalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur plagiasi, saya bersedia skripsi ini dugugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang telah berlaku (**UUD NO. 20 Tahun 2003, pasal 25 ayat 2 dan pasal 70**).

Majene, 15 Mei 2025



WARDIMAN
D0221345

ABSTRAK

Wardiman. Penerapan *Metode Naive Bayes* Dan *Random Oversampling* Untuk Deteksi Tingkat Gangguan Kecemasan. (di bimbing oleh Dr. Eng. Sulfayanti, S.Si.,M.T, dan Farid Wajidi, S.Kom.,MT)

Kecemasan merupakan respons alami yang melekat dalam kehidupan manusia, sehingga setiap individu pasti pernah mengalaminya. Dalam aktivitas sehari-hari, keberadaan kecemasan tergolong normal; namun, apabila tingkat kecemasan meningkat secara berlebihan, hal tersebut dapat mengganggu kestabilan psikologis individu dan dikategorikan sebagai gangguan kecemasan. Berdasarkan fenomena tersebut, penelitian ini dilaksanakan dengan tujuan untuk mengidentifikasi tahap awal gangguan kecemasan melalui penerapan Algoritma *Naive Bayes* dan algoritma *Random Oversampling* (ROS). Metodologi yang digunakan dalam penelitian ini adalah *Knowledge Discovery in Database* (KDD) untuk pemrosesan data, dengan teknik pengumpulan data menggunakan instrumen kuesioner *Generalized Anxiety Disorder-7* (GAD-7) yang melibatkan 670 data responden. Hasil akhir dari penelitian ini menunjukkan bahwa hasil prediksi tingkat kecemasan dengan algoritma *Naive Bayes* dan *Random Oversampling* (ROS) memiliki tingkat akurasi sebesar 95.45%. Dari total responden, 12,99% dikategorikan normal atau tidak mengalami gangguan kecemasan, 28,06% mengalami gangguan kecemasan ringan, 32,84% mengalami kecemasan sedang, dan 26,12% mengalami kecemasan berat.

Kata kunci : *Artificial Integence*, Gangguan Kecemasan, *Algotima Naive Bayes*, *Algoritma Random Oversampling* ROS, GAD-7, Deteksi.

ABSTRAK

Wardiman. *Application Of Naive Bayes Method And Random Oversampling For Anxiety Disorder Detection. ((Supervised by Dr. Eng. Sulfayanti, S.Si.,M.T, and Farid Wajidi, S.Kom.,MT)*

Anxiety is a natural response inherent in human life, so every individual must have experienced it. In daily activities, the existence of anxiety is considered normal; however, if the level of anxiety increases excessively, it can disrupt the psychological stability of the individual and is categorized as an anxiety disorder. Based on this phenomenon, this study was conducted with the aim of identifying the early stages of anxiety disorders through the application of the Naive Bayes Algorithm and the Random Oversampling (ROS) algorithm. The methodology used in this study is Knowledge Discovery in Database (KDD) for data processing, with data collection techniques using the Generalized Anxiety Disorder-7 (GAD-7) questionnaire instrument involving 670 respondent data. The final results of this study indicate that the results of predicting anxiety levels with the Naive Bayes and Random Oversampling (ROS) algorithms have an accuracy rate of 95.45%. Of the total respondents, 12.99% were categorized as normal or did not experience anxiety disorders, 28.06% experienced mild anxiety disorders, 32.84% experienced moderate anxiety, and 26.12% experienced severe anxiety.

Keywords : *Artificial Intelligence, Anxiety Disorders, Algorithm Naive Bayes, Algorithm Random Oversampling ROS, GAD-7, Detection.*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Gangguan kecemasan adalah kondisi cemas yang berlebihan disertai dengan respon perilaku yang tidak nyaman dan tidak terkendali terhadap sesuatu yang tidak pasti. Orang dengan gangguan kecemasan akan menunjukkan perilaku yang tidak biasa, seperti panik tak terkendali atau bahkan bertindak tidak rasional dalam kehidupan sehari-hari. *World Health Organization* (WHO) mendefinisikan gangguan kecemasan adalah gangguan kesehatan mental dengan prevalensi yang tinggi dan memberikan ancaman terhadap kesehatan (Zahra et al., 2024). Kecemasan adalah perasaan tidak nyaman atau kekhawatiran yang tidak jelas disertai dengan *respons* emosi yang disebabkan oleh antisipasi bahaya. Kecemasan berbeda dengan ketakutan karena kecemasan membutuhkan *respons* emosi, sedangkan ketakutan menggunakan strategi intelektual untuk mempersiapkan ancaman (Nur Meiji Hasisah, 2023). Kecemasan adalah reaksi emosional terhadap penilaian sesuatu yang dianggap membahayakan, dan cemas sangat terkait dengan perasaan ketidakpastian dan ketidakberdayaan (Mangundap, 2020).

Kecemasan adalah perubahan kondisi yang dirasakan seseorang yang dapat menyebabkan rasa khawatir, gelisah, dan takut karena ancaman dari dalam maupun dari luar diri (Tumanggor et al., 2022). Kecemasan adalah keadaan tidak jelas yang muncul tanpa alasan di masa depan. Dalam keluarga di mana salah satu anggotanya sakit, kecemasan dapat muncul. Sakit seorang anggota keluarga dapat menyebabkan krisis keluarga. Ketakutan adalah *respons* yang tepat terhadap ancaman, tetapi ketakutannya bisa tidak normal jika intensitasnya tidak sebanding dengan ancaman (Eriviana et al., 2023). *Disharmoni* tubuh dapat disebabkan oleh kecemasan yang tidak diatasi.

Perilaku *patologis* seperti kecemasan berlebihan hingga syok disebabkan oleh ketidakmampuan untuk mengatasi kecemasan secara *konstruktif*. Kecemasan pada

anak usia sekolah menyebabkan perubahan perilaku seperti menarik diri dari lingkungan, kesulitan fokus dalam melakukan aktivitas, kehilangan selera makan, mudah tersinggung, kehilangan pengendalian emosi amarah, tidak logis, dan kesulitan tidur (Indria & Intarti, 2023). Gejala pada kecemasan pada anak usia sekolah terdiri dari 2 bentuk yaitu gejala fisiologi dan gejala psikologi, gejala fisiologi dapat dilihat pada gangguan pencernaan yang kurang baik, ujung kaki dengan ujung jari tangan terasa dingin, berkeringat, detak jantung cepat, tidur tidak nyenyak, kepala pusing, tidak nafsu makan dan pernafasan terganggu. Sedangkan gejala psikologi dapat ditandai dengan merasa tertekan, sangat waspada terhadap takut dengan bahaya yang akan datang, sulit untuk rileks dan sulit merasa dalam segala situasi (Oktamarin et al., 2022).

Menurut data yang dimuat laman *Our Better World* dari data Kementerian Kesehatan 2013, sekitar 9 juta penduduk Indonesia mengalami depresi. Ada pun sebagai akibatnya ditemukan 34 kasus bunuh diri per 100.000 orang di Indonesia. Sekitar 16 juta orang berusia 15 tahun ke atas, ditemukan kasus bunuh diri yang diawali gejala kecemasan dan depresi oleh pelakunya. Sementara itu, pada 2019 WHO mencatat sekitar 300 juta orang di seluruh dunia telah mengalami depresi. Sebanyak 15,6 juta di antaranya berasal dari Indonesia. Pada tahun 2020, menurut *World Health Organization* (WHO) kecemasan menjadi penyebab utama ketidakberdayaan orang di seluruh dunia dan menyumbang 15% dari angka kesakitan global

Angka kecemasan pada anak usia sekolah (5-12 tahun) di Indonesia mencapai 6,1%, atau 706.689 orang, menurut Riskesdas 2018. Data Riskesdas tahun 2018, Prevalensi kecemasan di Indonesia usia 15 tahun ke atas mencapai 9,8%, peningkatan dari 6% pada tahun 2016 (Kemenkes RI, 2021). Menurut hasil Riset Kesehatan Dasar (Riskesdas) pada tahun 2018, 9,8% orang Indonesia di usia sekolah mengalami gangguan mental emosional, ditandai dengan gejala depresi dan kecemasan (Kemenkes, 2018).

Pada tingkat kecemasan yang sedang, asumsi individu justru akan lebih condong untuk memfokuskan pikirannya pada hal-hal yang penting saat itu saja dan menyingkirkan hal lainnya yang dirasa kurang penting. Sedangkan pada tingkat

kecemasan tinggi/berat, kognisi individu menurun, dan ia hanya dapat untuk memikirkan hal-hal kecil saja dan tentunya akan menghiraukan yang lainnya, sehingga pada kondisi ini akan lebih sulit untuk berpikir dengan tenang (Faozi et al., 2023).

Klasifikasi pada metode *Naïve Bayes* Salah satu metode yang sering digunakan untuk menerapkan teknik klasifikasi. *Naïve Bayes* merupakan sebuah metode untuk teknik pengklasifikasian dengan konsep probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Adapun keuntungan dalam penggunaan metode *Naïve Bayes* sebagai metode untuk menerapkan teknik klasifikasi yaitu metode ini hanya membutuhkan jumlah data pelatihan (*Training Data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian (Doni et al., 2021).

Penelitian ini menggunakan algoritma *Machine Learning* yang cukup baik dalam menentukan hasil klasifikasi pada algoritma *Naïve Bayes*. *Machine Learning* merupakan sebuah cabang ilmu dari *Artificial Intelligence* dengan menggunakan pemrograman komputer dalam menjadikan proses secara otomatis dengan berperilaku seperti manusia dan dapat meningkatkan pemahamannya melalui pembelajaran pengalaman secara otomatis. *Machine Learning* bekerja dengan menggunakan data sebagai masukan untuk dianalisis terhadap kumpulan data sehingga menemukan pola yang tersembunyi. Didalam *Machine Learning* dikenal data *training* dan data *testing*, data *training* merupakan data untuk melatih algoritma *Machine Learning* dan data *testing* adalah data yang digunakan untuk mengetahui performa dari algoritma dalam *Machine Learning* yang telah dilakukan pelatihan (Atika et al., 2022).

Penelitian terdahulu telah melakukan penelitian terkait tingkat stres mahasiswa menggunakan metode *K-Nearest Neighbor* (KNN) dalam prediksi tingkat stres mahasiswa. Pada penelitian ini dilakukan oleh (Wafiqi et al., 2024), Hasil penelitian menunjukkan bahwa algoritma KNN dapat digunakan untuk mengklasifikasikan mahasiswa kedalam kategori tingkat stres tertentu dengan akurasi sebesar 83,33% dengan data uji sebanyak 6 mahasiswa dan data *training*

sebanyak 84 mahasiswa. Adapun penelitian yang dilakukan oleh (Zahra et al., 2024) dengan judul “Deteksi Tingkat Gangguan Kecemasan Menggunakan Metode *Random Forest*” Penelitian ini menerapkan metode *Knowledge Discovery in Database* (KDD) untuk pemrosesan data dan menggunakan *kuesioner GAD-7* dalam pengumpulan data dengan banyak *responden 670 record*. Penelitian ini juga menerapkan *10-fold cross validation* untuk evaluasi model dengan paramater *spesifisitas, sensitivitas, dan akurasi*. Hasil akhir dari penelitian ini menunjukkan bahwa hasil prediksi tingkat kecemasan dengan algoritma *Random Forest* memiliki tingkat akurasi 89,55%. Dari total responden, 12,99% dikategorikan normal atau tidak mengalami gangguan kecemasan, 28,06% mengalami gangguan kecemasan ringan, 32,84% mengalami kecemasan sedang, dan 26,12% mengalami kecemasan berat.

Berdasarkan uraian diatas maka penulis akan melakukan penelitian dengan judul “**Penerapan Metode *Naïve Bayes* Dan *Random Oversampling* Untuk Deteksi Tingkat Gangguan Kecemasan.**” Metode *Naïve Bayes* dipilih karena kecepatan dan efisiensinya dalam pengklasifikasian, serta kemampuannya untuk bekerja dengan baik meskipun asumsi *independensi* antar fitur seringkali tidak terpenuhi Dengan menerapkan *Random Oversampling*, diharapkan jumlah data pada kelas *minoritas* dapat ditingkatkan sehingga model dapat belajar lebih baik dari data tersebut.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas rumusan penelitian ini adalah sebagai berikut:

1. Bagaimana mengimplementasikan metode *Naïve Bayes* dan algoritma *Random Oversampling* untuk deteksi tingkat gangguan kecemasan?.
2. Seberapa akurat metode *Naïve Bayes* dan algoritma *Random Oversampling* dalam deteksi dan mengklasifikasikan kecemasan?.

1.3 Tujuan Penelitian

Berdasarkan latar belakang diatas rumusan penelitian ini adalah sebagai berikut:

1. Mengimplementasikan metode Metode *Naïve Bayes* dan algoritma *Random Oversampling* dalam deteksi dan mengklasifikasikan kecemasan.
2. Mengetahui tingkat akurasi Metode *Naïve Bayes* dan algoritma *Random Oversampling* dalam deteksi dan mengklasifikasikan kecemasan.

1.4 Manfaat Penelitian.

Berdasarkan dengan permasalahan dan tujuan penelitian, maka penulis mengharapkan penelitian ini dapat memberikan beberapa manfaat dibawah ini

1. Hasil penelitian dapat digunakan digunakan untuk memprediksi deteksi dan mengklasifikasikan kecemasan.
2. Hasil penelitian dapat digunakan dalam perkembangan ilmu pengetahuan.

1.5 Batasan Masalah.

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data yang digunakan pada penelitian ini bersumber dari data *website kaggle* (<https://www.kaggle.com/>).
2. Algoritma yang digunakan pada penelitian yaitu metode *Naïve Bayes* dan *Random Oversampling*.
3. Pembagian kelas pada penelitian ini adalah 4 kelas yaitu normal, ringan, sedang dan berat.

BAB II

KAJIAN PUSTAKA

2.1 Landasan Teori

2.1.1 Kecemasan

Kecemasan adalah reaksi emosional terhadap penilaian sesuatu yang dianggap membahayakan, dan cemas sangat terkait dengan perasaan ketidakpastian dan ketidakberdayaan (Mangundap, 2020). Kecemasan adalah perasaan tidak nyaman atau kekhawatiran yang tidak jelas disertai dengan respons emosi yang disebabkan oleh antisipasi bahaya. Kecemasan berbeda dengan ketakutan karena kecemasan membutuhkan respons emosi, sedangkan ketakutan menggunakan strategi intelektual untuk mempersiapkan ancaman (Nur Meiji Hasisah, 2023).

(Faozi et al., 2023) mengkategorikan kecemasan menjadi beberapa kelompok yaitu respons dari perilaku, respon *kognitif*, dan emosional.

1. Perilaku: mengalami gelisah, adanya dari ketegangan dalam tubuh, respons kaget, bicarayang cenderung lebih cepat, kurang adanya koordinasi, rentan mengalami cedera, menutup diri secara sosial, inhibisi, menghindar dari sebuah masalah, *hiperventilasi*, menghindar yang terlalu waspada atau waspada berlebihan.
2. *Kognitif*: Perhatian terganggu, konsentrasi memburuk, mudah lupa, kurang bisa memberikan kemampuan dalam menilai, *preokupasi*, berpikir lebih terhambat, rentang asumsi berkurang, menurunnya kreativitas, menurunnya produktivitas, kebingungan, kewaspadaan berlebihan, menurunnya kesadaran diri, hilangnya penilaian objektivitas, takut akan adanya kehilangan kendali, ketakutan pada gambar, takut akan adanya kematian, kilas balik, adanya mimpi buruk.

3. Emosi/Afektif: mudah merasa terganggu, tidak bersabar, adanya rasa kegelisahan, tegang, ada perasaan canggung, takut, lebih berhati-hari, takut, khawatir, takut, mati rasa, rasa bersalah, perasaan malu.

Tingkat kecemasan ditentukan dari total skor jawaban seluruh pertanyaan. Tingkat kecemasan yang ada dalam skala ini terbagi menjadi 4, (Zahra, 2024) yaitu, normal, ringan, sedang dan berat:

2.1.2 *Machine Learning*

Machine learning ialah teknologi berbasis komputer dan matematika yang menggunakan data sebagai media pembelajaran oleh mesin komputer dan menghasilkan prediksi di masa depan. *Machine learning* hadir sebagai solusi dalam upaya prediksi dini sebuah penyakit. Prediksi diagnostik pada *machine learning* telah terbukti menjadi metode yang menjanjikan dalam berbagai permasalahan kehidupan. Metode *machine learning* dibuat dengan berbagai rangkaian alat dan teknik yang dapat mengeksplorasi kumpulan data dalam jumlah besar hingga menghasilkan pola-pola tertentu yang bermanfaat bagi ilmu pengetahuan (Raharja et al., 2024).

2.1.3 *Data Preprocessing*

Pada tahap ini dilakukan pembersihan atau merapikan data agar sesuai dengan kebutuhan analisis data yang akan dilakukan. Tahap ini melibatkan penghapusan atribut-atribut yang dianggap tidak relevan atau tidak memberikan kontribusi signifikan terhadap tujuan analisis (Zahra et al., 2024).

2.1.4 *Data Transformation*

Pada Tahap ini dilakukan proses transformasi agar sesuai dengan proses data mining. Transformasi ini mencakup perubahan tipe data dari *nominal* menjadi *numerik* pada 7 atribut (*GAD-7*) *Generalized Anxiety Disorder 7-item*

scale. Lalu selanjutnya akan diterapkan skor 0-3 pada data sesuai dengan petunjuk yang terdapat dalam kuesioner (*GAD-7*) yang ada pada Tabel 1. Skor keseluruhan dari setiap responden dihitung setelah perubahan tipe data untuk memastikan tingkat kecemasan responden. Hasil dari tingkat kecemasan ini yang akan digunakan dalam proses data mining sebagai label (Zahra et al., 2024).

2.1.5 Data Splitting

Subset pelatihan (*training set*) dan subset pengujian (*test set*). Proses ini dikenal sebagai data *splitting*. Subset pelatihan digunakan untuk melatih model agar dapat memahami pola dan relasi dalam data. Di sisi lain, subset pengujian digunakan untuk menguji kinerja model pada data yang tidak pernah dilihat selama pelatihan (Raharja et al., 2024).

2.1.6 Evaluation

ahap evaluasi merupakan langkah krusial dalam penelitian deteksi gangguan kecemasan dengan algoritma *Naive Bayes* dan *Random Oversampling*. Melalui pembagian data yang tepat, pengukuran kinerja menggunakan metrik yang relevan, serta analisis hasil melalui *confusion matrix*, peneliti dapat menilai efektivitas model dalam mendeteksi berbagai tingkat kecemasan secara akurat.

Confusion matrix adalah suatu matriks yang digunakan untuk menguji dan mengestimasi objek yang benar dan yang salah (Zahra, 2024).

Tabel 2. 1 *Confusion Matrix*

Prediksi	Aktual	
	Positif	Negatif
Positif	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>
Negatif	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>

1. *True Positive* (TP) Informasi yang teridentifikasi positif dan diprediksinya positif.
2. *False Positive* (FP) Informasi yang negatif tetapi diprediksi sebagai positif.
3. *True Negative* (TN) Informasi yang teridentifikasi negatif dan diprediksi dengan negatif.
4. *False Negative* (FN) Informasi yang positif tetapi diprediksi sebagai negatif.

Berdasarkan hasil dari *confusion matrix*, dapat pula dilakukan perhitungan untuk mengukur *performa* model, yaitu *accuracy*, *precision*, *recall* (*sensitivity/true positive rate*), dan *f1-score* Masing-masing perhitungannya didefinisikan sebagai berikut:

- a. *Accuracy*, efektivitas keseluruhan dari hasil klasifikasi yang telah dilakukan.

$$Accuracy = \frac{TP+TN}{Total} \quad (2.1)$$

- b. *Precision*, yaitu melihat seberapa sering model memprediksi positif dan secara aktual prediksi itu benar dengan perumusan sebagai berikut:

$$Precision = \frac{TP}{FP+TP} \quad (2.2)$$

- c. *Recall*, yaitu seberapa sering model memprediksi positif pada data yang memiliki klasifikasi aktual yang positif dengan perumusan sebagai berikut:

$$Recall = \frac{TP}{FN+TP} \quad (2.3)$$

- d. *F1-score*, yaitu merupakan hubungan antara data berlabel positif dari hasil klasifikasi yang menunjukkan keseimbangan antara *precision* dan *recall*.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2.4)$$

Cara yang paling umum untuk menunjukkan hasil klasifikasi terutama pada data multiclass ialah menyajikan dalam bentuk *confusion matrix* atau

juga dikenal dengan tabel kontingensi. Dalam permasalahan klasifikasi seringkali digunakan matrik konfusi (*confusion matrix*) untuk memetakan kelas-kelas yang diprediksi (*predicted label*) terhadap kelas-kelas yang sebenarnya (*true label*). *Confusion matrix* untuk klasifikasi label dengan empat kelas ditunjukkan dalam (Adriyan & Istana, 2024).

		Nilai sebenarnya (<i>true value</i>)			
		C ₁	C ₂	C ₃	C ₄
Nilai prediksi (<i>predicted value</i>)	C ₁	TP ₁	K _{2,1}	K _{3,1}	K _{4,1}
	C ₂	K _{1,2}	TP ₂	K _{3,2}	K _{4,2}
	C ₃	K _{1,3}	K _{2,3}	TP ₃	K _{4,3}
	C ₄	K _{1,4}	K _{2,4}	K _{3,4}	TP ₄

Gambar 2. 1 *Multi Class*

2.1.7 Algoritma *Naïve Bayes*

Algoritma *Naïve Bayes* berakar pada *teorema Bayes*. *Teorema Bayes* merupakan teorema yang mengacu pada konsep probabilitas bersyarat. Metode ini merupakan pendekatan statistik untuk melakukan inferensi induksi pada persoalan klasifikasi. Misalkan *A* dan *B* adalah kejadian dalam ruang sampel. *Teorema Bayes* secara matematis, *teorema* ini dapat diekspresikan sebagai berikut.

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (2.5)$$

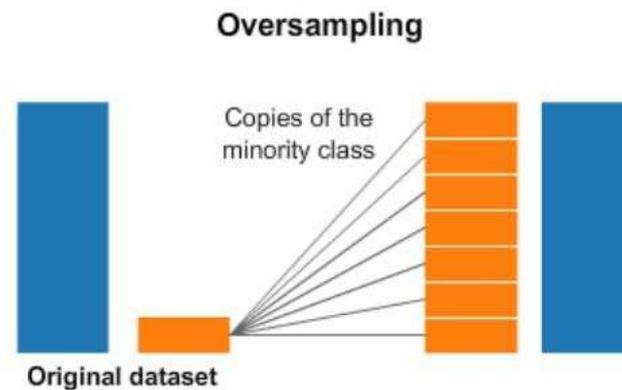
Dimana :

- *X* : Data dengan kelas yang belum di ketahui
- *H* : Hipotesis data merupakan class spesifik
- *P(H|X)*: Probabilitas Hipotesis *H* berdasarkan kondisi *X* (posterior probabilitas)
- *P(H)* : Probabilitas hipotesis (prior probabilitas)

- $P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H
- $P(X)$: Probabilitas X

2.1.8 Algoritma *Oversampling*

Hal ini bertujuan untuk menyeimbangkan distribusi kelas pada dataset yang *imbalanced* sehingga dapat meningkatkan kinerja dari algoritma klasifikasi. Hasil dari penelitian kombinasi ini akan dibandingkan dengan hasil klasifikasi tanpa resampling dan kombinasi antara teknik *resampling* + *clustering* dengan algoritma klasifikasi.



Gambar 2. 2 *Oversampling*

Random oversampling merupakan langkah penting dalam penelitian ini untuk mengatasi ketidakseimbangan data dan meningkatkan akurasi prediksi. Teknik ini membantu model prediktif belajar lebih efisien dan membuat klasifikasi lebih akurat (Ananda et al., 2024).

- Pilih secara acak instance dari kelas minoritas dan gandakan hingga jumlahnya mencapai $N_{\text{mayoritas}}$.
- Proses ini dapat dinyatakan dengan rumus berikut:

$$\text{Selisih} = N_{\text{mayoritas}} - N_{\text{minoritas}} \quad (2.6)$$

2.1.9 Python

Python adalah salah satu bahasa pemrograman yang baru di masa sekarang, pada bahasa pemrograman ini lebih simpel dalam dan singkat dalam membuat sebuah program setiap program yang dibuat pasti dan pasti akan membutuhkan inputan dan hasil outputan. Dalam metode penginputannya pun bahasa ini sedikit berbeda, memang terlihat mudah tapi bukan berarti untuk di hiraukan karena pada kenyataan masih banyak yang kesulitan dalam membuat program pada *python* ini, dengan dibuat nya pembahasan ini semoga dapat membantu bagi pemula yang sedang belajar bahasa pemograman (Raharja et al., 2024).

2.2 Penelitian Terkait

Dalam persiapan penelitian ini, penting untuk merujuk pada beberapa penelitian sebelumnya sebagai landasan dan wawasan. Hal ini bertujuan untuk memperoleh informasi dan ide-ide yang dapat dikembangkan dalam penelitian ini. Untuk itu, pada penelitian ini telah menghimpun sejumlah referensi dari penelitian-penelitian terdahulu yang relevan dengan permasalahan yang dibahas dalam penelitian ini.

Tabel 2. 2 Penelitian Terkait

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
1.	Ghefira Zahra Nur Fadhilah, Rizal Adi Saputra, Asa Hari Wibowo, 2024	Deteksi Tingkat Gangguan Kecemasan Menggunakan Metode <i>Random Forest</i>	Permasalahan yang diangkat yaitu dalam kehidupan sehari- hari memiliki kecemasan adalah hal yang	Hasil akhir dari penelitian ini menunjukkan bahwa hasil prediksi tingkat kecemasan dengan algoritma Random

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
			<p>normal, tetapi kecemasan yang berlebihan dapat mengakibatkan gangguan kestabilan diri sehingga diklasifikasikan sebagai gangguan kecemasan.</p>	<p>Forest memiliki tingkat akurasi 89,55%. Dari total responden, 12,99% dikategorikan normal atau tidak mengalami gangguan kecemasan, 28,06% mengalami gangguan kecemasan ringan, 32,84% mengalami kecemasan sedang, dan 26,12% mengalami kecemasan berat.</p>
2.	<p>Achmad Ulul Azmi Wafiqi ,Tundo, Bobby Arvian James, Abhirama Huga Ramadhan, Amin Nizar. 2024</p>	<p>“Prediksi Tingkat Stres Pada Mahasiswa UNUGHA Cilacap Menggunakan Algoritma <i>K-Nearest Neighbor</i>”</p>	<p>Permasalahan yang diangkat dalam penelitian ini yaitu untuk mengetahui prediksi tingkat kecemasan pada mahasiswa.</p>	<p>Hasil penelitian menunjukkan bahwa Algoritma KNN dapat digunakan untuk mengklasifikasikan mahasiswa kedalam kategori tingkat stress tertentu dengan akurasi sebesar 83,33%,</p>

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
				<p>dengan data uji sebanyak 6 mahasiswa dan data training sebanyak 84 mahasiswa. .</p> <p>Penelitian ini menunjukkan bahwa Algoritma KNN dapat digunakan</p> <p>Hasil penelitian menunjukkan bahwa Algoritma KNN dapat digunakan untuk mengklasifikasikan mahasiswa kedalam kategori tingkat stress tertentu dengan akurasi sebesar 83,33%, dengan data uji sebanyak 6 mahasiswa dan data training sebanyak 84 mahasiswa. .</p> <p>Penelitian ini menunjukkan bahwa Algoritma KNN dapat digunakan</p>

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
3.	Yusmi Nur Aini1, Eva Yulia Puspaningrum, Wahyu Syaifullah Jauharis Saputra, 2021.	“Implementasi <i>Decision Tree</i> Untuk Diagnosis Gangguan Kecemasan Umum”	Gangguan Kecemasan umum atau yang biasa disebut GAD (<i>generalized anxiety disorder</i>) merupakan salah satu satu gangguan kecemasan yang dapat ditandai dengan adanya perasaan cemas yang berlebihan, kecemasan ini biasanya terjadi karena penderita terlalu mengkhawatirkan segala hal yang terjadi diwaktu yang akan datang. Kecemasan ini sangat sulit bahkan tidak dapat dikendalikan	Hasil penelitian ini menggunakan Algoritma C5.0. Algoritma C5.0 dapat mengatasi permasalahan secara terstruktur dengan membentuk sebuah pohon keputusan atau <i>decision tree</i> dapat mengatasi permasalahan secara terstruktur dengan membentuk sebuah pohon keputusan atau <i>decision tree</i> . Pohon keputusan yang terbangun dari sistem ini memiliki nilai akurasi sebesar 95%. Nilai akurasi tersebut dapat berubah sesuai dengan pemilihan data uji dan data latih karena ada setiap data akan

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
				menghasilkan aturan dan ohon keutusan yang berbeda-beda.
4.	Sanilisa Putri Utami, 2024.	Klasifikasi kesehatan mental usia remaja menggunakan algoritma <i>Decision Tree</i> dan <i>Naïve Bayes</i>	Permasalahan yang diangkat dari penelitian ini adalah masa remaja merupakan periode kritis dalam perkembangan kesehatan mental pada remaja.	Hasil penelitian menunjukkan bahwa algoritma <i>Naïve Bayes</i> menghasilkan kinerja yang lebih baik dalam mengklasifikasikan tingkat kesehatan mental remaja menjadi lima kategori (normal, ringan, sedang, berat dan sangat berat). Kinerja algoritma dinilai berdasarkan matriks evaluasi, untuk algoritma <i>naïve bayes</i> didapatkan nilai kappa tertinggi sebesar 0.856, akurasi 89,3%, <i>precision</i> 88,47%

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
				<p>dan <i>recall</i> 89,3%. Sedangkan untuk algoritma <i>Decision Tree</i> mendapatkan nilai kappa sebesar 0.807, akurasi 85:55%, precision 82,4% dan <i>recall</i> 81,35%. Hasil mengindikasikan penggunaan algoritma <i>naïve bayes</i> sebagai alat bantu dalam klasifikasi kesehatan mental pada remaja lebih baik dibandingkan algoritma <i>decision tree</i>.</p>
5.	Nadilla Ayudia Pasa, Yani Maulita, I Gusti Prahmana, 2024.	Penggunaan Metode <i>Rough Set</i> Pada Tingkat Kecemasan (<i>Anxietas</i>) Mahasiswa dalam	Masalah dari penelitian ini adalah penelitian ini menyelidiki tingkat kecemasan mahasiswa dalam menyelesaikan	Hasil penelitian adalah data dikumpulkan melalui kuesioner dari mahasiswa yang saat ini sedang menyelesaikan proyek akhir

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
		Menyusun Tugas Akhir	<p>proyek akhir mereka, yang merupakan persyaratan penting untuk kelulusan. Kecemasan merupakan masalah umum yang dihadapi oleh mahasiswa, yang seringkali di akibatkan oleh proses yang panjang dan rumit dalam persiapan proyek akhir.</p>	<p>mereka. Dengan menerapkan metode rough set, penelitian ini berhasil mengidentifikasi faktor-faktor signifikan yang mempengaruhi tingkat kecemasan, seperti respon psikologis, fisik, dan positif. Temuan ini memberikan wawasan berharga bagi para pendidik dan konseler untuk lebih memahami dan mengatasi kecemasan mahasiswa selama mahasiswa akhir.</p>
6	Rizqulloh Rifqi Edwanto 2024	“Penerapan Metode <i>Support Vector Machine</i> Untuk Prediksi Depresi,	Permasalahn dari penelitian ini adalah adalah Peningkatan prevalensi depresi, kecemasan, dan	Hasil dari penelitian ini adalah Untuk prediksi, digunakan model <i>Support Vector Machine</i> (SVM) dengan <i>kernel Radial Basis</i>

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
		Kecemasan, Dan Stres”	stres di masyarakat modern menyoroti pentingnya pengembangan sistem yang efektif untuk deteksi dini dan klasifikasi kondisi kesehatan mental.	<p><i>Function</i> (RBF) yang dikenal efektif dalam menangani data <i>non-linear</i>. Model SVM yang dikembangkan mencapai akurasi 99.3%, menunjukkan kinerjanya yang sangat baik dalam deteksi dini kondisi kesehatan mental. Penelitian ini tidak hanya meningkatkan akurasi diagnosis mental tetapi juga membuka peluang untuk sistem prediksi berbasis teknologi yang dapat diakses lebih luas, mendukung upaya pencegahan kondisi mental yang lebih serius.</p>

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
7	Arizka Rifanul Mudin, 2024	Sistem pendeteksi dini depresi menggunakan metode kecerdasan buatan dengan analisis sentimen dan <i>natural language processing</i> (NLP)	Masala penelitian ini adalah Depresi, yaang mempengaruhi sekitar 300 juta orang diseluruh dunia setiap tahun, sering kali tidak terdiagnosis pada tahap awal karena banyak pasien yang enggan berkonsultasi dengan dokter.	Hasil penelitian ini adalah algoritma naïve bayes classifer dan <i>support vector machine</i> (SVM) diimplementasikan untuk meningkatkan akurasi dan efesiensi deteksi. Hasil penelitian ini menunjukkan bahwa kombinasi kedua algoritma ini dapat secara signifikan meningkatkan akurasi dan mempercepat proses pendeteksian depresi, sehingga berkontribusi pada upaya deteksu dini.
8	Yusmi Nur Aini, 2021.	Implementasi Algoritma C5.0 Untuk diagnosis	Masalah penelitian ini adalah Kecemasan	Hasil penelitian ini adalah Hasil pengujian menggunakan

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
		gangguan kecemasan umum berdasarkan rumusan GAD-7	merupakan respon normal dalam menghadapi stres, namun sebagian orang dapat mengalami kecemasan yang berlebihan sehingga mengalami kecemasan yang berlebihan sehingga mengalami kesulitan dalam mengatasinya.	<i>Confusion Matrix</i> nilai <i>recall</i> memperoleh keakuratan sebesar 100%, dimana nilai ini menggambarkan tingkat keberhasilan model atau pohon keputusan dalam mengklasifikasikan. Nilai <i>precision</i> sebesar 93.33%, dimana nilai ini menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang diberikan oleh model. Nilai <i>accuracy</i> sebesar 95.23%, dimana menggambarkan seberapa akurat model atau pohon klasifikasi dalam mengklasifikasikan dengan benar Nilai <i>error rate</i> sebesar

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
				4.76%, dimana menunjukkan tingkat <i>error</i> model atau pohon keputusan.
9.	Reza Dwi Fitriani1, Hasbi Yasin, Tarno, 2021	Penyelesaian masalah ketidak seimbangan kelas data melalui teknik <i>Oversampling</i> dan <i>Undersampling</i> pada data klasifikasi siswa tidak naik kelas.	Permasalahan yang diangkat Sebagian besar algoritma data mining bekerja paling baik ketika jumlah sampel di setiap kelas berada pada jumlah yang hampir sama.	Berdasarkan hasil uji yang telah dilakukan dengan tiga dataset yang berbeda, algoritma <i>C5.0</i> dengan <i>k-fold cross validation</i> dapat bekerja lebih baik pada dataset yang diolah melalui teknik <i>random oversampling</i> dibandingkan dataset asli ataupun dataset yang dibentuk dari teknik <i>random undersampling</i> . Ini ditunjukkan dengan akurasi pada setiap lipatan cenderung stabil dan konsisten

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
				pada kisaran 93% sampai 97,6%.
10	Siska Febriani, Heni Sulistiani 2021	Analisis data hasil diagnosa untuk klasifikasi gangguan Kepribadian menggunakan algoritma c4.5	Masalah yang diangkat , penulis hanya memfokuskan pada pemeriksaan gangguan kecemasan (GAD), gangguan mood (gangguan depresi mayor) dan gangguan somatoform (gangguan konversi). Karena gangguan tersebut paling sering dialami oleh masyarakat pada umumnya.	Hasil pada penelitian. Metode yang digunakan dalam penelitian ini adalah klasifikasi data C4.5 dengan menggunakan data primer yang diperoleh dari angket masyarakat untuk proses klasifikasi. Penetapan tersebut diklasifikasikan sebagai proses klasifikasi menggunakan Rapid Miner untuk klasifikasi dan menggunakan validasi silang sebagai validasi keakuratan data. Hasil dari penelitian ini adalah prediksi gangguan

No	Nama Penulis	Judul Penelitian	Masalah Penelitian	Hasil
				kepribadian untuk menghasilkan informasi yang dibutuhkan oleh masyarakat

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kecemasan adalah reaksi yang alami dalam kehidupan, tetapi jika kecemasan tersebut berlebihan, hal itu dapat mengganggu keseimbangan diri dan dikategorikan sebagai gangguan kecemasan. Berdasarkan hasil uji coba dan analisis yang telah dilakukan, model prediksi tingkat kecemasan menggunakan algoritma *Naive Bayes* mencapai tingkat akurasi pada *confusion matrix* 95,45% dengan pembagian data training dan testing 80:20. Pada tingkat akurasi pada *5-fold Cross Validation* yaitu sebesar 92,82%, data *training* pada *5-fold cross validation* yaitu 80% dari total data pada setiap iterasi, karena 4 dari 5 *fold* digunakan untuk pelatihan dan data testing yaitu 20% dari total data pada setiap iterasi, karena 1 *fold* digunakan untuk pengujian. Dalam mendeteksi tingkat kecemasan sejak dini, model prediksi dengan *Naive Bayes* terbukti sangat efektif. Selain itu, untuk menyeimbangkan data, dilakukan teknik *random oversampling* yang berhasil menyamaratakan data hingga 220 perkelas dimana data normal, data kecemasan ringan, data kecemasan sedang dan data kecemasan berat semua di samaratakan.

5.2 Saran

Saran yang diperoleh dari hasil penelitian terkait deteksi tingkat kecemasan dengan menggunakan algoritma *Naive Bayes* adalah perlunya pengumpulan data dalam jumlah yang lebih besar agar mencakup variasi karakteristik responden yang lebih luas, sehingga model menjadi lebih representatif terhadap populasi secara keseluruhan. Selain itu, disarankan untuk melakukan optimasi parameter pada algoritma *Naive Bayes* guna meningkatkan performa prediksi. Penanganan ketidakseimbangan kelas juga penting untuk diperhatikan, salah satunya dengan menerapkan teknik *random oversampling* agar distribusi kelas menjadi seimbang

dan model mampu mengenali setiap kategori dengan baik, akan tetapi pada teknik *Random Oversampling* rentan terhadap *overfitting*, karena data yang ditambahkan hanyalah duplikasi dari data yang sudah ada, maka dari itu lakukan teknik *SMOTE* (*Synthetic Minority Over-sampling Technique*). *SMOTE* yang menghasilkan data sintetik baru secara lebih bervariasi, sehingga mengurangi risiko *overfitting* dan meningkatkan kemampuan model dalam mengenali setiap kategori. Terakhir, disarankan untuk melakukan pengembangan dan perbaikan aplikasi deteksi yang telah dibangun, khususnya dengan menggunakan *framework Streamlit*. Dengan perbaikan ini, antarmuka aplikasi dapat dibuat lebih interaktif, mudah digunakan, dan dapat diakses secara luas oleh pengguna umum. Penggunaan *Streamlit* memungkinkan integrasi model prediksi dengan tampilan berbasis *web* secara efisien, sehingga mendukung penerapan sistem deteksi kecemasan dalam skala yang lebih besar dan praktis.

DAFTAR PUSTAKA

- Adriyan, & Istana, B. (2024). Penerapan Pendekatan Sains Data untuk Pemodelan Klasifikasi Desain Keamanan Rangka Sepeda Melalui Pembelajaran Mesin Otomatis. *Jurnal Surya Teknik*, 11(1), 1–8. <https://doi.org/10.37859/jst.v11i1.7218>
- Ananda, I. K., Fanani, A. Z., Setiawan, D., & Wicaksono, D. F. (2024). Penerapan Random Oversampling dan Algoritma Boosting untuk Memprediksi Kualitas Buah Jeruk. *Edumatic: Jurnal Pendidikan Informatika*, 8(1), 282–289. <https://doi.org/10.29408/edumatic.v8i1.25836>
- Aninda Cahya Savitri, P., & Luh Indah Desira Swandi, N. (2023). Intervensi Kecemasan Pada Mahasiswa: Literature Review. *Psikobuletin: Buletin Ilmiah Psikologi*, 4(1), 43. <https://doi.org/10.24014/pib.v4i1.20628>
- Arsy, G. R., Listyarini, A. D., & Nyumirah, S. (2021). Pendampingan Psikologis Orang Tua Pada Anak Usia Sekolah Selama Pembelajaran Daring. *Jurnal Keperawatan Dan Kesehatan Masyarakat Cendekia Utama*, 10(2), 161. <https://doi.org/10.31596/jcu.v10i2.762>
- Atika, D., Ari Aldino, A., Informasi, S., Pagar Alam No, J., Ratu, L., & Kedaton, K. (2022). Term Frequency-Inverse Document Frequency Support Vector Machine Untuk Analisis Sentimen Opini Masyarakat Terhadap Tekanan Mental Pada Media Sosial Twitter. *Jurnal Teknologi Dan Sistem Informasi (JTISI)*, 3(4), page-page. <http://jim.teknokrat.ac.id/index.php/JTISI>
- Dewi, A. R. (2022). PR (Peer Group) Anak Sekolah Sebagai Upaya Meningkatkan Kreativitas Pada Anak Usia Sekolah Di Gedong Tataan Pesawaran. *Jurnal Pengabdian Masyarakat Jajama (JPMJ)*, 1(2), 86. <https://doi.org/10.47218/jpmj.v1i2.236>

- Doni, B. T. R., Susanti, S., & Mubarak, A. (2021). Penerapan Data Mining Untuk Klasifikasi Penyakit Hepatocellular Carcinoma Menggunakan Algoritma Naïve Bayes. *Jurnal Responsif: Riset Sains Dan Informatika*, 3(1), 12–19. <https://doi.org/10.51977/jti.v3i1.403>
- Erviana, Ismar, & Masniati. (2023). Tingkat Kecemasan Anak Usia Sekolah Pasca Bencana Gempa Bumi. *Jurnal Multidisiplin Ilmu*, 2(24), 501–507.
- Faozi, A., Adzani, A. A., Izza, D. S. N., & Kibtiyah, M. (2023). Dampak Kecemasan Masyarakat Terhadap Kesehatan Mental Di Masa Pandemi Covid 19. *Jurnal Mercusuar: Bimbingan, Penyuluhan, Dan Konseling Islam*, 3(1), 1. <https://doi.org/10.31332/mercusuar.v3i1.6808>
- Gumelar, Gagah, and Hanif Al-Fatta. (2023). “Kombinasi Algoritma Klasifikasi Dengan Algoritma Oversampling Untuk Menangani Ketidakseimbangan Kelas Pada Level Data.” *Jurnal Teknik Informatika dan Sistem Informasi* 10(2): 29–39. <http://jurnal.mdp.ac.id>.
- Indria, G. A., & Intarti, W. D. (2023). Intervensi Kecemasan Akademik pada Mahasiswa Peminatan Kesehatan: Literature Review. *Jurnal Publikasi Kebidanan*, 14(2), 50–59
- Nadilla Ayudia Pasa, Yani Maulita, I Gusti Prahmana (2024), Penggunaan Metode Rough Set Pada Tingkat Kecemasan (Anxietas) Mahasiswa Dalam Menyusun Tugas Akhir.Pdf.”
- Oktamarin, L., Kurniati, F., Sholekhah, M., Nurjanah, S., Oktaria, S. W., Sukmawati, S., & Apriyani, T. (2022). Gangguan Kecemasan (Anxiety Disorder) Pada Anak Usia Dini. *Jurnal Multidisipliner Bharasumba*, 1(02), 119–134. <https://doi.org/10.62668/bharasumba.v1i02.192>
- Pradana, A. I., Studi, P., Rekeyasa, T., Lunak, P., Ilmu, F., Duta, U., Surakarta, B., Studi, P., Komputer, T., Ilmu, F., Duta, U., Surakarta, B., Studi, P., Informasi, S., Ilmu, F., Duta, U., Surakarta, B., & Tengah, J. (2024). 8(2), 221–228.

- Raharja, A. R., Jayadi, Pramudianto, A., & Muchsam, Y. (2024). Penerapan Algoritma Decision Tree dalam Klasifikasi Data “Framingham” Untuk Menunjukkan Risiko Seseorang Terkena Penyakit Jantung dalam 10 Tahun Mendatang. *Technologia Journal*, 1(1). <https://doi.org/10.62872/cwgzpz962>
- Riyanti, N., Setiawan, D., & Rondli, W. S. (2023). Pola Asuh Single Parent Berpendidikan Rendah Dalam Pendidikan Anak. *Jurnal Educatio FKIP UNMA*, 9(2), 507–514. <https://doi.org/10.31949/educatio.v9i2.4703>
- Moch. Anjas Aprihartha, Dicky Zulhan, Fatma Ahardika Nurfaizal, Taufik., Program Studi Informatika, Fakultas Teknologi, Komunikasi Dan, and Universitas Nasional. 2024. “Penyelesaian Masalah Ketidakseimbangan Data Melalui Teknik Oversampling dan Undersampling pada Klasifikasi Siswa Tidak Naik Kelas.”
- Sanilisia Putri Utami., Program studi, Teknik Informatika, Fakultas Sains, D A N Teknologi, Universitas Islam, and Negeri Syarif. 2024. “KLASIFIKASI Kesehatan Mental Usia Remaja Menggunakan Algoritma Decision.”
- Tumanggor, R. D., Siregar, Y. S., Keperawatan, D., & Keperawatan, F. (2022). Hubungan Tingkat Kecemasan Dengan Pola Tidur Mahasiswa Selama Pandemi Covid–19. *Jurnal Endurance*, 6(2), 402–409. <https://doi.org/10.22216/jen.v6i2.337>
- YUSMI NUR AINI, Eva Yulia Puspaningrum, and Wahyu Syaifullah JS. 2021. “Implementasi Decision Tree Untuk Diagnosis Gangguan Kecemasan Umum.” *Jurnal Informatika dan Sistem Informasi* 2(2): 395–402. doi:10.33005/jifosi.v2i2.318.
- Wafiqi, A. U. A., Tundo, T., James, B. A., Ramadhan, A. H., & Nizar, A. (2024). Prediksi Tingkat Stres Pada Mahasiswa UNUGHA Cilacap Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Tekno Kompak*, 18(2), 331. <https://doi.org/10.33365/jtk.v18i2.3933>

Zahra, G., Fadhilah, N., Saputra, R. A., & Wibowo, A. H. (2024). Deteksi Tingkat Gangguan Kecemasan Menggunakan Metode Random Forest. *Jurnal Fakultas Teknik UMT*, 13(1), 38–47.
<http://jurnal.umt.ac.id/index.php/jt/index>

Zhafirah, D. (2023). Penanganan Imbalance Data Dengan Random Oversampling (Ros) Pada Klasifikasi Penderita Diabetes Menggunakan Support Vector Machine (SVM). Universitas Lampung