

Analisis Klasifikasi Tingkat Kesehatan Mental pada Mahasiswi Akhir dalam Menempuh Tugas Akhir Menggunakan Algoritma *Support Vector Machine* Studi Kasus: Universitas Darussalam Gontor

Widya Kurniawan, Aziz Musthafa, & Anisa Kirani*

Universitas Darussalam Gontor, Jl. Raya Siman, Ponorogo, Jawa Timur, Indonesia.

Email: rektorat@unida.gontor.ac.id, anisakirani42013@mhs.unida.gontor.ac.id

Corresponding author: anisakirani42013@mhs.unida.gontor.ac.id

Abstrak

Kesehatan secara keseluruhan bergantung pada kesehatan mental atau kesehatan jiwa. Kesehatan mental juga penting sama halnya dengan Kesehatan fisik. Satu dari dua anak muda di bawah usia 25 tahun akan mengalami gangguan kesehatan mental di beberapa titik, dan 75% penyakit mental dimulai sebelum usia 25 tahun. Mahasiswi termasuk dalam kategori tertinggi dalam gangguan kesehatan mental. Umumnya, mahasiswa semester awal dan semester akhir banyak mengalami kecemasan akademik. Salah satu faktor utama yang menyebabkan tekanan psikologis pada mahasiswa yaitu tugas akhir atau biasa disebut skripsi. Penelitian ini bertujuan untuk mengklasifikasi tingkat kesehatan mental yaitu stress dan kecemasan melalui proses pelatihan model klasifikasi menggunakan algoritma *Support Vector Machine*. Penelitian ini menggunakan data dari hasil dari kuesioner yang telah disebarkan kepada mahasiswi akhir yang sedang mengerjakan skripsi. Dataset yang digunakan sebanyak 249 record data dan dibagi menjadi 2 dataset untuk klasifikasi stres dan kecemasan. Hasil penelitian ini menunjukkan ketepatan akurasi tertinggi pada dataset klasifikasi stres dengan menggunakan kernel RBF dan polynomial mencapai 68% yaitu pada kernel RBF dengan gamma 1 dan C 100. Sedangkan pada dataset klasifikasi kecemasan ketepatan akurasi tertinggi mencapai 50% yaitu pada kernel polynomial dengan degree 3 dan C 100. Penerapan model terbaik menunjukkan fitur yang paling berpengaruh dalam dataset Stress dan Anxiety adalah Literature Review, Support System, dan Analysis Method. Akurasi yang diperoleh dapat dijadikan sebagai tolak ukur untuk penelitian selanjutnya dengan menggunakan data yang lebih kompleks. Penelitian mendatang diharapkan tidak hanya mempertimbangkan fitur berupa faktor eksternal yang menentukan tingkat stres dan kecemasan mahasiswa, tetapi juga menambahkan faktor internal.

Kata Kunci: Kesehatan Mental, Mahasiswi, Tugas Akhir, Klasifikasi, *Support Vector Machine*

Abstract

Overall health depends on mental health or psychological well-being. Mental and physical well being are equally vital. One in two young people under the age of 25 will experience mental health issues at some point, and 75% of mental illnesses begin before the age of 25. Female university students are among those at the highest risk for mental health issues. Typically, students in their early and final semesters experience a high level of academic anxiety. One of the main factors causing psychological distress in students is the final project or thesis. This study aims to classify mental health levels, namely stress and anxiety, by training a classification model that makes use of the *Support Vector Machine* algorithm. The dataset used in this study is derived from a questionnaire distributed to final-year female students working on their thesis. The dataset consists of 249 records and is divided into two datasets for stress and anxiety classification. The results of this study show the highest accuracy in the stress classification dataset using the RBF and polynomial kernels, reaching 68% with the RBF kernel at gamma 1 and C 100. Meanwhile, the highest accuracy in the anxiety classification dataset reached 50% achieved with the polynomial kernel at degree 3 and C 100. The application of the best model indicates that the most influential features in the Stress and Anxiety datasets are Literature Review, Support System, and Analysis Method. The obtained accuracy can be used as a standard for upcoming studies using more complex data.

Future studies are expected not only to consider external factors as features determining students levels of stress and anxiety but also to add internal factors.

Keywords: *Mental health, Female Student, Final Thesis, Classification, Support Vector Machine*

1. PENDAHULUAN

Kesehatan mental adalah suatu hal penting untuk mencapai kesehatan yang holistik. Sama halnya dengan kesehatan fisik, perhatian terhadap kesehatan mental juga sangat penting. Menurut definisi kesehatan yang diajukan oleh *World Health Organization (WHO)* [1], yaitu “*health as a state of complete physical, mental and social well-being and not merely the absence of diases or infirmity*” Kesehatan merupakan kondisi kesejahteraan yang menyeluruh, mencakup fisik, mental, dan sosial, bukan sekadar hilangnya penyakit atau kelemahan. Sebagian besar anak muda (berusia lanjut 12-25 tahun) di seluruh dunia sedang mengalami kesehatan mental yang buruk. Satu dari dua anak muda di bawah usia 25 tahun akan mengalami gangguan kesehatan mental di beberapa titik, dan 75% penyakit mental dimulai sebelum usia 25 tahun [2]. Dalam *International Statistical Classification of Diseases and Related Health Problems, Tenth revision (ICD-10)* menyebutkan bahwa penyakit dengan jumlah terbanyak salah satunya adalah depresi dan gangguan kecemasan.

Berdasarkan informasi dari Pusat Informasi Kriminal Nasional (Pusiknas) Kepolisian RI (Polri), kasus bunuh diri di Indonesia mencapai angka 971 kasus antara Januari hingga 18 Oktober 2023. Angka ini telah melebihi total kasus bunuh diri di seluruh tahun 2022, yang tercatat sebanyak 900 kasus. Salah satu kelompok yang sering melakukan tindakan bunuh diri adalah mahasiswa [3]. Tingkat gangguan kesehatan mental di kalangan mahasiswa tergolong tinggi. Kecemasan akademik sering dialami oleh mahasiswa yang berada di semester awal maupun semester akhir. Beberapa faktor utama yang menyebabkan tekanan psikologis pada mahasiswa antara lain adalah tugas akhir, yang lebih dikenal sebagai skripsi [4]. Skripsi dirancang untuk mengasah kemampuan mahasiswa dalam melakukan analisis sesuai dengan prinsip-prinsip ilmiah, yang ditulis dalam bentuk laporan tugas akhir [5]. Selain itu, tantangan dan tantangan yang muncul dari

sumber eksternal seperti dosen pembimbing dan lainnya selama proses menyelesaikan tugas akhir atau skripsi, dapat menjadi sumber stress bagi mahasiswa tingkat akhir. Adapun 5 faktor yang menyebabkan stres mahasiswa tingkat akhir, meliputi *role of supervisor, literature review, analysis method, support system, research gap* [6]. Salah satu skala untuk mengukur tingkat depresi, kecemasan, dan stres seseorang adalah DASS-42 (*Depression Anxiety Stress Scale*). Skala ini terdiri dari 42 pertanyaan yang membagi individu ke dalam tingkat instrument yang berbeda yaitu normal, ringan, sedang, berat, dan sangat berat [7].

Berdasarkan jenis kelamin, jumlah responden mahasiswi lebih banyak daripada mahasiswa yaitu 64,7%. Dalam menghadapi stres, wanita memiliki kemampuan mengatasi lebih rendah daripada pria, yang menyebabkan wanita lebih beresiko mengalami gangguan psikologis. Selain itu, kebanyakan wanita memiliki sifat karakter yang sensitif [8]. Banyaknya perguruan tinggi di Indonesia, Universitas Darussalam Gontor (UNIDA) merupakan salah satu universitas yang bersistem pesantren yang terbagi menjadi 2 kampus yaitu Kampus Putra yang berada di Ponorogo dan Kampus Putri yang berada di Mantingan, Ngawi. Universitas dengan jumlah mahasiswa sebanyak 5.076 orang ini menaungi 7 fakultas dan 18 program studi S1.

Beberapa peneliti telah melakukan penelitian dalam jurnal Analisis Data Hasil Diagnosa untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5. Data yang digunakan untuk proses klasifikasi merupakan hasil yang didapat dari angket masyarakat. Nilai akurasi pada penerapan algoritma C4.5 mencapai nilai 72,67% [9]. Dalam jurnal Implementasi Algoritma *Naïve Bayes* untuk Analisis Klasifikasi Survei Kesehatan Mental dengan menggunakan data dari *Open Source Mental Illness (OSMI)* yang terdiri dari 1254 record data, menghasilkan hasil akurasi sebesar 72% [10]. Pada salah satu artikel dengan judul Perbandingan Algoritma *Naïve Bayes Classifier* dan *Support Vector*

Machine pada Klasifikasi Data Kesehatan Mental Mahasiswa dengan menggunakan data *Kaggle* didapati hasil dari akurasi pada algoritma *Support Vector Machine* mencapai 94,37% dan pada algoritma *Naïve Bayes Classifier* mencapai 86,87%. Hal ini menunjukkan bahwa kinerja dan performa algoritma *Support Vector Machine* lebih tinggi dibandingkan algoritma *Naïve Bayes Classifier* [11].

Penulis ingin melakukan penelitian tentang klasifikasi pada kesehatan mental mahasiswi berdasarkan beberapa temuan penelitian sebelumnya dengan menggunakan algoritma *Support Vector Machine (SVM)* karena *SVM* memiliki kemampuan untuk menangani beberapa kontinu dan variabel kategoris [12]. Dibandingkan dengan metode lain, metode *Support Vector Machine (SVM)* adalah yang paling baru. Sebab fungsi kernel menentukan area fitur untuk mencari fungsi klasifier, pemilihan fungsi kernel yang tepat sangat penting. Berdasarkan permasalahan dan penjelasan yang telah dipaparkan maka dari itu penulis ingin menyelesaikan kasus kesehatan mental pada mahasiswa dengan penelitian yang berjudul Analisis Klasifikasi Kesehatan Mental pada Mahasiswi dalam Menempuh Tugas Akhir dengan menggunakan Algoritma *Support Vector Machine*. Dengan menggunakan metode *Support Vector Machine (SVM)*, penelitian ini diharapkan dapat mengembangkan sistem cerdas yang dapat menentukan apakah seseorang mungkin mengalami gangguan kesehatan mental.

2. DASAR TEORI

2.1 Kesehatan Mental

Kesehatan adalah aspek penting yang perlu dijaga dan diperharikan, mencakup kesehatan fisik, mental, serta sosial untuk keseimbangan yang harmonis. Berdasarkan definisi *WHO (World Health Organization)*, kesehatan adalah keadaan yang lengkap, meliputi kesejahteraan fisik, mental, dan sosial, serta bebas dari penyakit atau kelemahan. Dengan demikian, kesehatan tidak hanya terkait dengan kondisi fisik saja, tetapi juga mencakup kesehatan mental dan kesejahteraan sosial [13]. *WHO* mendefinisikan Kesehatan mental didefinisikan sebagai keadaan Sejahtera di mana seseorang memiliki pemahaman tentang

potensi dirinya, kemampuan untuk mengatasi tekanan hidup sehari-hari, kemampuan untuk bekerja dengan produktif dan bermanfaat, dan kemampuan untuk berkontribusi positif kepada komunitasnya [14].

2.2 Kecemasan

Kecemasan adalah kondisi psikologis dimana seseorang merasakan kecemasan berlebihan dan sulit untuk dikendalikan, yang pada akhirnya mengganggu kehidupan sehari-hari. Bagi kebanyakan orang, kecemasan sering muncul dalam situasi tertentu, seperti ketika akan melewati ujian atau wawancara kerja. Namun, pada penderita gangguan kecemasan, rasa cemas sering kali muncul di berbagai situasi. Akibatnya, mereka mengalami kesulitan untuk merasa tenang. Gejala psikologis lain yang dapat dialami oleh penderita gangguan kecemasan termasuk kurangnya *self-confident*, mudah tersinggung, stress, kesulitan untuk fokus, dan kecenderungan untuk menyendiri [15].

2.3 Stres

Stres adalah respon non-spesifik tubuh terhadap situasi atau kondisi yang terjadi di sekitar atau di lingkungannya. Pada suatu waktu dalam hidup, setiap orang akan mengalami stress. Gejala stress bisa berupa tanda-tanda fisik maupun psikologis. Secara fisik, individu yang mengalami stress dapat mengalami peningkatan tekanan darah, mudah lelah, jantung berdebar, mual, gangguan pada lambung, menstruasi tidak teratur, tangan dan kaki dingin, gemetar, ketegangan otot, dan napas pendek. Sedangkan gejala psikologis stress meliputi mudah marah, cemas, rendah diri, hipersensitif, kehilangan kreativitas, menarik diri dari interaksi sosial, menangis secara tiba-tiba, kesulitan fokus, tidak tertarik pada hal-hal yang biasa disenangi, dan mudah tersinggung [16].

2.4 Tugas Akhir

Penyusunan skripsi adalah tugas akhir yang harus dilakukan oleh mahasiswa tingkat akhir untuk memenuhi persyaratan selesai studi. Skripsi adalah karya ilmiah yang harus ditulis secara individual oleh setiap mahasiswa sebagai bukti kemampuan akademisnya. Selain menjadi persyaratan kelulusan, skripsi juga merupakan syarat untuk memperoleh gelar sarjana. Dalam proses penulisannya, mahasiswa memerlukan banyak sumber

referensi. Dengan membaca berbagai referensi, mahasiswa dapat memperdalam pengetahuan mereka dan mendapatkan pemahaman lebih tentang topik yang akan dikerjakan [17].

2.5 Data Mining

Data Mining atau Penambangan Data adalah proses mengekstraksi serta mengidentifikasi informasi bermanfaat dari kumpulan database besar dengan menggunakan statistic, matematika, kecerdasan buatan, dan pembelajaran mesin. Penambangan data, yang juga disebut sebagai pengenalan pola adalah metode pengolahan data yang bertujuan mengidentifikasi pola tersembunyi pada data yang diolah. Penambangan data memiliki peran penting di berbagai bidang seperti industry, keuangan, meteorologi, ilmu pengetahuan, dan teknologi [18].

2.6 Klasifikasi

Klasifikasi adalah proses untuk mengidentifikasi model, pola, atau fungsi yang mampu memberikan gambar dan memisahkan objek data sehingga dapat dikelompokkan ke dalam kelas-kelas tertentu dari beberapa kelas yang tersedia. Akurasi klasifikasi diukur berdasarkan persentase *instance* yang berhasil diberikan label kelas yang benar. Tujuan dari klasifikasi adalah untuk melakukan prediksi kelas objek yang belum diketahui. Klasifikasi juga merupakan proses analisis data yang menghasilkan model untuk menunjukkan berbagai kelas yang ada dalam data tersebut.

2.7 Algoritma Support Vector Machine

Support Vector Machine (SVM) adalah metode klasifikasi yang didasarkan pada prinsip dasar linear classifier, yaitu klasifikasi yang dapat dipisahkan secara linear. Namun, dengan menerapkan konsep kernel pada ruang kerja berukuran besar, SVM juga mampu menangani masalah non-linier. Kernel berfungsi sebagai pemetaan ulang data dari ruang dimensi asal ke ruang dimensi yang lebih tinggi untuk sementara [19]. Pemilihan SVM didasarkan pada kemampuannya untuk menyelesaikan masalah klasifikasi dengan dataset yang kompleks. Pengaturan hyperparameter digunakan untuk menyelesaikan masalah dengan menetapkan rentang nilai untuk gamma dan C yang akan diuji. Dataset yang digunakan bersifat kategorikal, *Radial Basis Function (RBF)*

kernel dijadikan acuan dalam penggunaan SVM ini [12].

Formula untuk menghitung jarak dengan kernel RBF adalah sebagai berikut

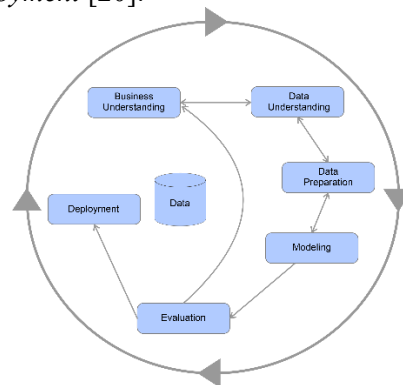
$$K(x, xl) = \exp(-\text{gamma} * ||x - xl||^2) \quad (1)$$

Sedangkan formula untuk menghitung jarak dengan kernel polynomial adalah :

$$K(w, x) = (\text{gamma} \cdot w^T x + b)^N \quad (2)$$

3. METODOLOGI

Metode dalam penelitian ini menggunakan metode *CRISP-DM (Cross-Industry Standard Process for Data Mining)*. CRISP-DM adalah model proses industry-independen untuk penambangan data yang terdiri dari 6 fase berulang dari *Business Understanding* hingga *Deployment* [20].



Gambar 1. Diagram Alur CRISP-DM

3.1 Business Understanding

Pada tahapan pertama dari CRISP-DM adalah *Business Understanding (Pemahaman Bisnis)*. Tahap ini adalah fase pemahaman tentang pentingnya proses penambangan data yang akan dilakukan, serta kebutuhan dari perpektif penelitian. Aktivitas pada tahap ini meliputi penetapan sasaran atau tujuan penelitian, pemahaman terhadap situasi penelitian, serta penerjemahan tujuan penelitian ke dalam tujuan penambangan data.

3.2 Data Understanding

Dataset yang akan digunakan untuk penelitian ini adalah data dari kuisisioner yang disebarakan kepada 249 mahasiswi akhir dari program studi yang berbeda-beda. Kuisisioner terdiri dari tiga komponen, (1) Identitas responden, (2) Pertanyaan-pertanyaan yang merujuk dari *Depression Anxiety Stress Scale (DASS-42)* diambil hanya pertanyaan-pertanyaan terkait kecemasan dan stress kemudian diparafrase, (3) Pertanyaan mengenai faktor-faktor stres mahasiswi ketika

mengerjakan skripsi. Pengalaman responden dalam menyelesaikan skripsi, cara mereka menangani masalah, dan perasaan mereka adalah topik pertanyaan dalam kuesioner.

3.3 Business Preparation

Setelah memiliki dataset, selanjutnya adalah melakukan *Data Preparation*. *Data Preparation* adalah mempelajari dan memberi pemahaman pada data yang akan diolah hingga data tersebut layak untuk melaju ke tahap pemodelan.

3.4 Modelling

Pada tahap ini, langkah yang dilakukan yaitu mengolah dataset yang digunakan, yaitu dataset kesehatan mental mahasiswa yang sedang menempuh tugas akhir menggunakan tools *Google Colabs* dengan format *.xlsx. Tahap pemodelan pada penelitian ini adalah menggunakan algoritma *Support Vector Machine (SVM)* untuk melakukan proses klasifikasi.

3.5 Evaluation

Metode validasi yang digunakan untuk mengetahui performa model adalah *Cross Validation*. Karena SVM sangat sensitive terhadap *hyperparameter*, maka sangat dianjurkan untuk menggunakan *Cross Validation*. Metode ini bekerja dengan membagi data menjadi sejumlah k. Untuk mengetahui nilai akurasi, jumlah data uji yang diprediksi benar harus dibagi dengan jumlah total data uji. Kemudian, hasil rata-rata dari setiap k dibandingkan.

3.6 Deployment

Pada tahapan ini, penulis menemukan hasil klasifikasi tingkat kesehatan mental mahasiswa yang sedang menempuh tugas akhir. Hasil klasifikasi ini akan menjadi insight baru mengenai tingkat kesehatan mahasiswa akhir dan faktor-faktor penyebabnya dengan identifikasi yang komprehensif.

4. HASIL DAN PEMBAHASAN

4.1 Business Understanding

Mahasiswa sering mengalami gangguan Kesehatan mental pada semester awal dan akhir, dan gangguan ini dikategorikan sebagai gangguan Kesehatan mental yang tinggi. Tugas akhir adalah komponen utama yang menyebabkan tekanan psikologis pada mahasiswa. Terbentuk 5 faktor yang memengaruhi stres mahasiswa tingkat akhir,

meliputi *role of supervisor, literature review, analysis method, support system, research gap*.

4.1.1 Problem Statement

- Faktor-faktor apa yang paling berpengaruh terhadap tingkat stress dan kecemasan mahasiswa saat mengerjakan tugas akhir?
- Dapatkah kita membangun model prediktif untuk mengidentifikasi tingkat stress mahasiswa saat mengerjakan tugas akhir?

4.1.2 Goals

- Mengidentifikasi faktor-faktor yang berpengaruh terhadap tingkat stress dan kecemasan. Tujuannya adalah menentukan fitur-fitur mana yang paling berpengaruh terhadap tingkat stress mahasiswa. Dengan memahami faktor-faktor tersebut, dapat mengidentifikasi karakteristik mahasiswa yang rentan mengalami stress.
- Membangun model prediktif yang akurat. Tujuannya adalah untuk mengembangkan model *machine learning* yang dapat secara akurat mengidentifikasi apakah seorang mahasiswa mengalami stress atau tidak.

4.2 Data Understanding

Dataset yang akan digunakan untuk penelitian ini adalah data dari kuisisioner yang disebarkan kepada 249 mahasiswa akhir dari program studi yang berbeda-beda. Kuisisioner terdiri dari tiga bagian, (1) Identitas responden, (2) Pertanyaan-pertanyaan yang merujuk dari *Depression Anxiety Stress Scale (DASS-42)* diambil hanya pertanyaan-pertanyaan terkait kecemasan dan stress kemudian diparafrase, (3) Pertanyaan mengenai faktor-faktor stress mahasiswa ketika mengerjakan skripsi. Pengalaman responden dalam menyelesaikan skripsi, cara mereka menangani masalah, dan perasaan mereka adalah topik pertanyaan dalam kuisisioner.

4.2.1 Dataset Features Information

Variabel-variabel pada Data Kesehatan Mental Mahasiswa Akhir adalah sebagai berikut:

Tabel 1. Dataset Features Information

Kolom	Deskripsi	Sub Variabel
ID	ID unik untuk setiap mahasiswa	Q1, Q2, etc
Umur	Usia mahasiswa	19,20,21,22, 23
Program Studi	Program studi yang diambil oleh mahasiswa dalam	PBA, PAI, TBI, HES, PM, EI, MB,

	S1 di Universitas Darussalam Gontor	FARMASI, GIZI, TI, AGRO, IQT, HI, ILKOM
Fakultas	Fakultas dari program studi yang diambil oleh mahasiswa	Tarbiyah, Syari'ah, Ekonomi, FEM, Ilkes, Saintek, Ushuluddin, Humaniora
Metode Penelitian	Metode yang digunakan mahasiswa untuk meneliti dalam penelitian tugas akhir.	Kuantitatif, Kualitatif
Role of Supervisor	Menunjukkan bagaimana peran seorang pembimbing dalam pengerjaan tugas akhir. Apakah waktu yang dimiliki untuk bimbingan teratur. Apakah dosen pembimbing mudah untuk ditemui. Dan apakah arahan dari dosen pembimbing mampu untuk memberi motivasi mahasiswa	<i>Very Bad, Bad, Good, Very Good</i>
Literature Review	Tingkatan menemukan buku dan jurnal melalui internet maupun perpustakaan dan kemudahan mencari penelitian yang sejenis sebagai bahan perbandingan	<i>Very Bad, Bad, Good, Very Good</i>
Analysis Method	Tingkatan kesulitan metode analisis yang digunakan oleh mahasiswa dan kemampuan mahasiswa untuk mengaplikasikannya ke dalam penelitian	<i>Very Bad, Bad, Good, Very Good</i>

<i>Support System</i>	Apakah mahasiswa merasa termotivasi ketika menghadiri seminar proposal teman dan apakah mahasiswa merasakan tekanan dari orang tua agar segera menyelesaikan studi	<i>Very Bad, Bad, Good, Very Good</i>
<i>Self Management</i>	Kemampuan untuk mengelola waktu dan menemukan cara untuk tetap termotivasi bahkan ketika merasa lelah dan stres	<i>Very Bad, Bad, Good, Very Good</i>
<i>Environment</i>	Pengaruh lingkungan dan disiplin kampus terhadap pengerjaan tugas akhir	<i>Very Bad, Bad, Good, Very Good</i>
<i>Stress</i>	Menunjukkan tingkat stress yang dialami mahasiswa	<i>Normal, Mild, Moderate, Severe, Extremely Severe</i>
<i>Anxiety</i>	Menunjukkan tingkat kecemasan yang dialami mahasiswa	<i>Normal, Mild, Moderate, Severe, Extremely Severe</i>

4.2.2 Tipe Data dan Info Dataset

Pada bagian ini, akan dijelaskan mengenai tipe data yang digunakan dalam dataset serta informasi umum terkait dataset tersebut.

ID	User	Program Studi	Fakultas	Metode Penelitian	Basis of Supervision	Literature Review	Analysis Method	Support System	Self Management	Environment	Stress	Anxiety
count	249	249	249	249	249	249	249	249	249	249	249	249
unique	249	NAN	14	7	2	3	4	4	4	4	3	5
top	Q1	NAN	TI	Tarbiyah	Kualitatif	Very Good	Good	Good	Good	Good	Normal	Extremely Severe
freq	1	NAN	20	60	140	130	111	170	133	154	177	92
mean	NAN	21.625000	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN
std	NAN	9.850642	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN
min	NAN	19.000000	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN
25%	NAN	21.000000	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN
50%	NAN	22.000000	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN
75%	NAN	22.000000	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN
max	NAN	24.000000	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN	NAN

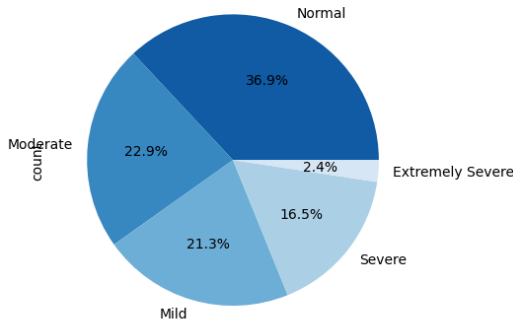
Gambar 2. Info Dataset

4.2.3 Exploratory Data Analysis

Setelah memiliki dataset, selanjutnya adalah melakukan Data Preparation. Data Preparation adalah melakukan eksplorasi dan

pemahaman pada data yang akan diolah hingga data tersebut layak untuk melaju ke tahap pemodelan.

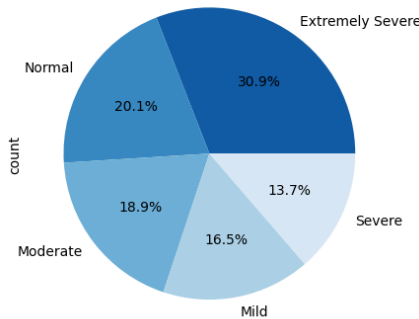
Presentase Tingkat Stres Mahasiswi



Gambar 3. Presentase Tingkat Stres Mahasiswi

Pada gambar 3 menjelaskan bahwa mayoritas mahasiswi berada dalam kondisi normal atau mengalami stres ringan hingga sedang, sementara sebagian kecil mahasiswi mengalami stres berat atau sangat berat. Hal ini menunjukkan adanya variasi yang signifikan dalam tingkat stres di kalangan mahasiswi, dengan proporsi yang lebih besar berada pada kategori normal.

Presentase Tingkat Kecemasan Mahasiswi



Gambar 4. Presentase Tingkat Kecemasan Mahasiswi

Pada gambar 4 dapat disimpulkan bahwa kecemasan pada mahasiswi merupakan masalah yang signifikan, dengan lebih dari setengah dari populasi mahasiswi mengalami kecemasan pada tingkat berat hingga sangat berat. Persentase mahasiswi yang mengalami kecemasan normal dan ringan berada pada tingkat yang lebih rendah dibandingkan dengan kategori yang lebih serius.

4.3 Perancangan Component Diagram

4.3.1 Pengumpulan Data

Sebelum memulai data preparation, dataset dibagi 2 menjadi Dataset Stress, dan Dataset Anxiety. Hal ini dilakukan penulis untuk mempermudah dalam proses pengolahan

data. Data yang digunakan masing-masing sebanyak 249 record data dengan 12 kolom.

ID	Umur	Program Studi	Fakultas	Metode Penelitian	Role of Supervisor	Literature Review	Analysis Method	Support System	Self Management	Environment	Stress
0	01	21	TI	Sainik	Kuantitatif	Good	Good	Bad	Very Good	Bad	Good Severe
1	02	22	PAI	Talqiah	Kuantitatif	Very Good	Good	Very Good	Bad	Good	Good Severe
2	03	23	TI	Sainik	Kuantitatif	Good	Bad	Good	Bad	Very Bad	Very Good Severe
3	04	20	ICT	Ushuludin	Kuantitatif	Very Good	Good	Good	Bad	Good	Very Good Mild
4	05	22	HI	Humaira	Kuantitatif	Very Good	Good	Good	Good	Good	Bad Severe

Gambar 5. Dataset Stress

ID	Umur	Program Studi	Fakultas	Metode Penelitian	Role of Supervisor	Literature Review	Analysis Method	Support System	Self Management	Environment	Anxiety
0	01	21	TI	Sainik	Kuantitatif	Good	Good	Bad	Very Good	Bad	Good Extremely Severe
1	02	22	PAI	Talqiah	Kuantitatif	Very Good	Good	Very Good	Bad	Good	Good Extremely Severe
2	03	23	TI	Sainik	Kuantitatif	Good	Bad	Good	Bad	Very Bad	Very Good Extremely Severe

Gambar 6. Dataset Anxiety

Pembagian variabel X dan y pada kedua dataset adalah sebagai berikut.

Tabel 2. Pembagian variabel X dan y pada Dataset Stress

X	Umur, Program Studi, Fakultas, Metode Penelitian, Role of Supervisor, Literature Review, Analysis Method, Support System, Self Management, Environment
y	Stress

Tabel 3. Pembagian variabel X dan y pada Dataset Anxiety

X	Umur, Program Studi, Fakultas, Metode Penelitian, Role of Supervisor, Literature Review, Analysis Method, Support System, Self Management, Environment
y	Anxiety

4.3.2 Pembersihan Data

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 249 entries, 0 to 248
Data columns (total 12 columns):
# Column Non-Null Count Dtype
---
0 ID 249 non-null object
1 Umur 249 non-null int64
2 Program Studi 249 non-null object
3 Fakultas 249 non-null object
4 Metode Penelitian 249 non-null object
5 Role of Supervisor 249 non-null object
6 Literature Review 249 non-null object
7 Analysis Method 249 non-null object
8 Support System 249 non-null object
9 Self Management 249 non-null object
10 Environment 249 non-null object
11 Stress 249 non-null object
dtypes: int64(1), object(11)
memory usage: 23.5+ KB
```

Gambar 7. Info Dataset Stress

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 249 entries, 0 to 248
Data columns (total 12 columns):
# Column Non-Null Count Dtype
---
0 ID 249 non-null object
1 Umur 249 non-null int64
2 Program Studi 249 non-null object
3 Fakultas 249 non-null object
4 Metode Penelitian 249 non-null object
5 Role of Supervisor 249 non-null object
6 Literature Review 249 non-null object
7 Analysis Method 249 non-null object
8 Support System 249 non-null object
9 Self Management 249 non-null object
10 Environment 249 non-null object
11 Anxiety 249 non-null object
dtypes: int64(1), object(11)
memory usage: 23.5+ KB
```

Gambar 8. Info Dataset Anxiety

Dari gambar 7 dan 8 menunjukkan bahwa tidak ada missing value dari kedua dataset. Kemudian menghapus kolom yang tidak digunakan ketika pengolahan data yaitu kolom 'ID'.

4.3.3 Transformasi Data

Data yang digunakan sebagian besar berupa data kategorikal, sehingga perlu dilakukan transformasi agar dapat diolah menggunakan teknik analisis yang sesuai. Oleh karena itu, dilakukan konversi manual dari kategori ke nilai numerik. Proses ini melibatkan pemberian kode angka pada setiap kategori, sehingga data dapat diinterpretasikan dengan lebih baik oleh model yang digunakan dalam tahap analisis selanjutnya.

Umur	Program Studi	Fakultas	Metode Penelitian	Role of Supervisor	Literature Review	Analysis Method	Support System	Self Management	Environment	Stress	
0	21	12	6	0	2	2	1	3	1	2	3
1	22	0	0	1	3	2	3	1	2	2	3
2	23	12	6	1	2	1	2	1	0	3	3
3	20	7	3	0	3	2	2	1	2	3	1
4	22	0	0	0	3	2	2	2	2	1	3

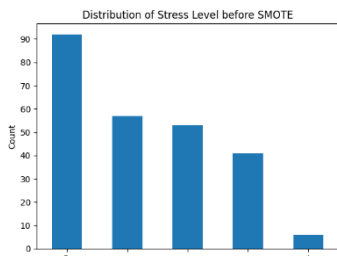
Gambar 9. Dataset Stres setelah Transformasi Data

Umur	Program Studi	Fakultas	Metode Penelitian	Role of Supervisor	Literature Review	Analysis Method	Support System	Self Management	Environment	Anxiety	
0	21	12	6	0	2	2	1	3	1	2	4
1	22	0	0	1	3	2	3	1	2	2	4
2	23	12	6	1	2	1	2	1	0	3	4
3	20	7	3	0	3	2	2	1	2	3	4
4	22	0	0	0	3	2	2	2	2	1	4

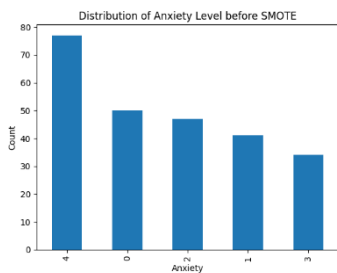
Gambar 10. Dataset Anxiety setelah Transformasi Data

4.3.4 Balancing Data

Salah satu tantangan yang dihadapi dalam penelitian ini adalah ketidakseimbangan kelas pada dataset, terutama untuk variabel stress dan anxiety. Ketidakseimbangan ini dapat mempengaruhi performa model prediktif, karena model cenderung bias terhadap kelas yang lebih dominan. Untuk mengatasi hal ini, digunakan metode SMOTE (*Synthetic Minority Over-Sampling Technique*), yang bertujuan untuk menyeimbangkan jumlah sampel di setiap kelas dengan cara membuat sampel sintetis dari kelas minoritas.

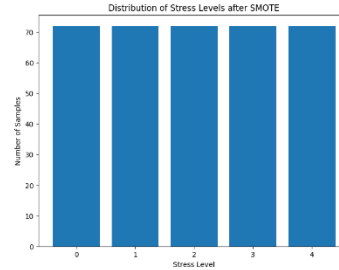


Gambar 11. Distribusi Variabel Stres sebelum SMOTE

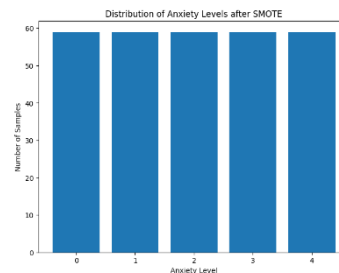


Gambar 12. Distribusi Variabel Anxiety sebelum SMOTE

Gambar 11 dan 12 menunjukkan bahwa ketidakseimbangan data pada variabel Stress dan Anxiety. Setelah menerapkan SMOTE, jumlah sampel di setiap kelas untuk dataset stress berhasil diseimbangkan menjadi 72 sampel, sedangkan untuk dataset anxiety, masing-masing kelas berjumlah 59 sampel. Dengan dataset yang lebih seimbang ini, model diharapkan dapat memiliki performa yang lebih baik dalam memprediksi kedua label tersebut



Gambar 13. Distribusi Variabel Stres setelah SMOTE



Gambar 14. Distribusi Variabel Anxiety setelah SMOTE

4.3.5 Standarisasi Data

Setelah proses *oversampling* menggunakan metode SMOTE untuk menyeimbangkan jumlah sampel di setiap kelas, langkah selanjutnya adalah melakukan standarisasi data. Dalam analisis ini, standarisasi dilakukan menggunakan *Standard Scaler*, yang bekerja dengan mengubah data sehingga setiap fitur memiliki rata-rata nol dan standar deviasi satu.

0	-0.579570	-0.436271	-0.648485	1.239036	-0.481759
1	-0.579570	-0.694165	-0.648485	-0.887079	-0.481759
2	-0.579570	-0.952059	-1.124746	-0.887079	1.124104
3	0.775269	-1.467847	-1.124746	1.239036	1.124104
4	-0.579570	-0.436271	-0.648485	-0.887079	1.124104
Literature Review Analysis Method Support System Self Management \					
0	0.380946	0.175142	-0.311473	-0.667861	
1	0.380946	0.175142	-0.311473	-0.667861	
2	0.380946	1.576281	1.362113	0.564437	
3	0.380946	0.175142	-0.311473	1.795935	
4	-1.271351	0.175142	-3.658645	0.564437	
Environment					
0	2.385425				
1	2.385425				
2	0.275467				
3	-1.834492				
4	0.275467				

Gambar 15. Dataset Stres setelah Standarisasi Data


```

Umur Program Studi Fakultas Metode Penelitian Role of Supervisor \
0 -0.651153 -0.417431 -0.630048 1.296849 -0.818526
1 -0.651153 -0.683397 -0.630048 -0.771100 -0.818526
2 -0.651153 -0.949363 -1.137874 -0.771100 1.024720
3 0.596185 -1.481295 -1.137874 1.296849 1.024720
4 -0.651153 -0.417431 -0.630048 -0.771100 1.024720

Literature Review Analysis Method Support System Self Management \
0 0.089290 -0.237926 -0.465123 -1.267972
1 0.089290 -0.237926 -0.465123 -1.267972
2 0.089290 1.711744 1.316842 0.277696
3 0.089290 -0.237926 -0.465123 1.823365
4 -1.792173 -0.237926 -4.029053 0.277696

Environment
0 2.318520
1 2.318520
2 0.324458
3 -1.669605
4 0.324458
    
```

Gambar 16. Dataset Anxiety setelah Standarisasi Data

4.4 Perancangan Deployment Diagram

Analisis *Support Vector Machine* pada tingkat Kesehatan mental mahasiswi menggunakan 2 fungsi kernel yaitu *Polynomial* dan *Radial Basis Function*. Fungsi *polynomial* menggunakan 3 jenis tingkat yaitu $p=1$, $p=2$, dan $p=3$. Sedangkan pada *radial basis function* dengan $\gamma = 1, 2$, dan 3 . Nilai C yang dibandingkan yaitu $C=10$, $C=100$, dan $C = 1000$. Pada analisis tingkat kesehatan mental ini menggunakan nilai C , degree, dan γ yang berbeda, dapat dibandingkan ketepatan klasifikasi terbaik pada analisis tingkat kesehatan.

4.5 Evaluation

Evaluation (evaluasi) adalah proses untuk mengukur kinerja model klasifikasi atau prediksi dengan menggunakan metrik yang sesuai. Salah satu metode validasi yang digunakan dalam penelitian ini adalah *Cross Validation* sebanyak 5 fold. Pada metode ini, data dibagi menjadi 5 bagian yang sama besar. Untuk setiap fold :

- 4 bagian dari data akan digunakan sebagai data latih (80% dari data)
- 1 bagian sisanya akan digunakan sebagai data uji (20% dari data)

Proses ini akan diulang sebanyak 5 kali, sehingga setiap bagian data menjadi data uji tepat satu kali dan data latih empat kali. Hasil dari setiap *fold* kemudian dirata-rata untuk memperoleh performa keseluruhan model.

Tabel 4. Hasil Tingkat Akurasi Klasifikasi Stres

Mean Accuracy Cross Validation	Polynomial			RBF		
	P=1	P=2	P=3	G=1	G=2	G=3
C = 10	0.5278	0.5750	0.6722	0.6750	0.6472	0.6250
C = 100	0.5194	0.6139	0.6750	0.6833	0.6444	0.6250
C = 1000	0.5194	0.6167	0.6583	0.6833	0.6444	0.6250

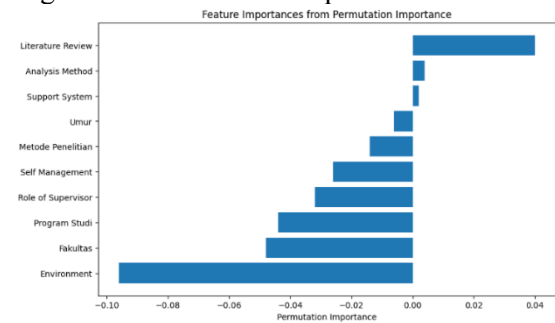
Berdasarkan Tabel 2 menunjukkan bahwa nilai akurasi tertinggi klasifikasi dataset *stress*

terdapat pada *Support Vector Machine*, dengan fungsi kernel RBF dengan nilai γ 1 dan C 100 mencapai 68%.

Tabel 5. Hasil Tingkat Akurasi Klasifikasi Anxiety

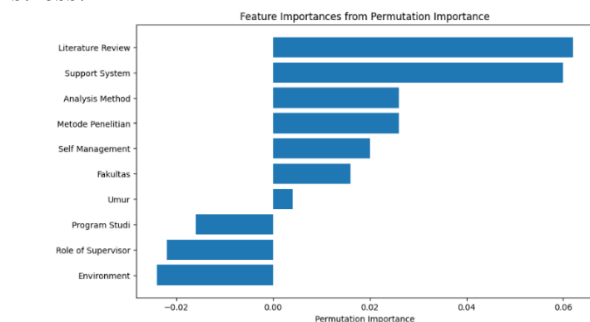
Mean Accuracy Cross Validation	Polynomial			RBF		
	P=1	P=2	P=3	G=1	G=2	G=3
C = 10	0.2915	0.4407	0.4746	0.4542	0.4407	0.4339
C = 100	0.2814	0.4644	0.5017	0.4542	0.4407	0.4339
C = 1000	0.2847	0.4678	0.4847	0.4542	0.4407	0.4339

Berdasarkan Tabel 3 menunjukkan bahwa nilai akurasi tertinggi klasifikasi dataset *anxiety* terdapat pada *Support Vector Machine*, dengan fungsi kernel *Polynomial* dengan nilai degree 3 dan C 100 mencapai 50%.



Gambar 17. Feature Importances Dataset Stress

Hasil analisis *Permutation Importance* terhadap faktor-faktor yang mempengaruhi stres. Diagram pada gambar 17 menunjukkan bahwa *Literature Review*, *Analysis Method*, dan *Support System* memiliki nilai paling signifikan positif berpengaruh pada variabel *stress*.



Gambar 18. Feature Importances Dataset Anxiety

Berdasarkan pada gambar 18 menunjukkan bahwa *Literature Review*, *Support System*, dan *Analysis Method* sangat signifikan dalam mempengaruhi tingkat kecemasan.

5. SIMPULAN

Kesimpulan dalam penelitian ini adalah fitur yang paling berpengaruh dalam dataset *Stress* dan *Anxiety* adalah *Literature Review*, *Support System*, dan *Analysis Method*. Dan klasifikasi pada dataset *Stress* menghasilkan akurasi tertinggi mencapai 68% dengan fungsi kernel RBF γ 1 dan C 100. Sedangkan klasifikasi pada dataset *Anxiety* menghasilkan akurasi tertinggi mencapai 50% dengan fungsi kernel polynomial degree 3 dan C 100. Akurasi yang diperoleh dapat dijadikan sebagai tolak ukur untuk penelitian selanjutnya dengan menggunakan data yang lebih kompleks. Penelitian mendatang diharapkan tidak hanya mempertimbangkan fitur berupa faktor eksternal yang menentukan tingkat stres dan kecemasan mahasiswa, tetapi juga menambahkan faktor internal. Dengan demikian, diharapkan akurasi yang dihasilkan dapat lebih meningkat.

DAFTAR REFERENSI

- [1] World Health Organization, "Comprehensive Mental Health Action Plan 2013-2030," *Geneva*, 2021.
- [2] W. Economic Forum, "A Global Youth Mental Health Advocacy Toolkit," *Orygen*, 2020.
- [3] B. Media.com, "Tren Bunuh Diri di Kalangan Mahasiswa Sepanjang 2023, Apa Solusinya _," Dec. 2023, Accessed: Feb. 07, 2024. [Online]. Available: <https://bekasimedia.com/2023/12/23/tren-bunuh-diri-di-kalangan-mahasiswa-sepanjang-2023-apa-solusinya/>
- [4] A. T. Setyanto, "Deteksi Dini Prevalensi Gangguan Kesehatan Mental Mahasiswa di Perguruan Tinggi," *Wacana*, vol. 15, no. 1, p. 66, Feb. 2023, doi: 10.20961/wacana.v15i1.69548.
- [5] S. Suhandiah, A. Ayuningtyas, and P. Sudarmaningtyas, "Tugas Akhir dan Faktor Stres Mahasiswa," *JAS-PT (Jurnal Analisis Sistem Pendidikan Tinggi Indonesia)*, vol. 5, no. 1, p. 65, Jul. 2021, doi: 10.36339/jaspt.v5i1.424.
- [6] Nabila and A. Sayekti, "Manajemen Stres pada Mahasiswa dalam Penyusunan Skripsi di Institut Pertanian Bogor Stress Management for Undergraduate Students in Writing Thesis at IPB University," *Jurnal Manajemen dan Organisasi (JMO)*, vol. 12, no. 2, pp. 156–165, 2021.
- [7] A. M. Ali, A. A. Alkhamees, H. Hori, Y. Kim, and H. Kunugi, "The depression anxiety stress scale 21: Development and validation of the depression anxiety stress scale 8-item in psychiatric patients and the general public for easier mental health measurement in a post COVID-19 world," *Int J Environ Res Public Health*, vol. 18, no. 19, Oct. 2021, doi: 10.3390/ijerph181910142.
- [8] N. Romadhona, S. Fitriyana, R. G. Ibnusantosa, and T. Respati, "Level of Depression, Anxiety, and Stress of College Students in Indonesia during the Pandemic COVID-19," *Global Medical and Health Communication (GMHC)*, vol. 9, no. 3, Dec. 2021, doi: 10.29313/gmhc.v9i3.8337.
- [9] S. Febriani and H. Sulistiani, "Analisis Data Hasil Diagnosa untuk Klasifikasi Gangguan Kepribadian menggunakan Algoritma C4.5," *Jurnal Teknologi dan Sistem Informasi (JTSI)*, 2021.
- [10] R. Alfaresy Chaerudin *et al.*, "Implementasi Algoritma Naïve Bayes Untuk Analisis Klasifikasi Survei Kesehatan Mental (Studi Kasus: Open Sourcing Mental Illness)," 2022.
- [11] H. Dwi Putra, L. Khairani, D. Hastari, P. Studi Sistem Informasi, F. Sains dan Teknologi, and C. Author, "SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat Comparison of Naive Bayes Classifier and Support Vector Machine Algorithms for Classifying Student Mental Health Data Perbandingan Algoritma Naive Bayes Classifier dan Support Vector Machine untuk Klasifikasi Data Kesehatan Mental Mahasiswa," 2023. [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas>
- [12] S. Ghosh, A. Dasgupta, and A. Swetapadma, *A Study on Support Vector Machine based Linear and Non-*

- Linear Pattern Classification*. International Conference on Intelligent Sustainable Systems (ICISS), 2019.
- [13] D. V. Fakhriyani, “Kesehatan Mental,” 2019. [Online]. Available: <https://www.researchgate.net/publication/348819060>
- [14] World Health Organization, “Mental Health WHO,” Jun. 2022, Accessed: Feb. 06, 2024. [Online]. Available: <https://www.who.int/news-room/factsheets/detail/mental-health-strengthening-our-response>
- [15] N. Izzati Septia, N. Kamal, and P. Kesehatan Banjarmasin, “KESEHATAN MENTAL DAN KETENANGAN JIWA KAJIAN PSIKOLOGI AGAMA,” 2023.
- [16] J. Andriana and N. N. Prihantini, “HUBUNGAN TINGKAT STRES DENGAN INDEKS MASSA TUBUH PADA MAHASISWA FAKULTAS KEDOKTERAN UNIVERSITAS KRISTEN INDONESIA,” 2021.
- [17] S. B. Seto, M. T. S. Wondo, and M. F. Mei, “Hubungan Motivasi Terhadap Tingkat Stress Mahasiswa Dalam Menulis Tugas Akhir (Skripsi),” *Jurnal Basicedu*, vol. 4, no. 3, pp. 733–739, Jun. 2020, doi: 10.31004/basicedu.v4i3.431.
- [18] A. Trifani, S. Tunas Bangsa, A. Perdana, W. Stikom, T. Bangsa, and H. Qurniawan, “Penerapan Data Mining Klasifikasi C4.5 dalam Menentukan Tingkat Stress Mahasiswa Akhir,” *Jurnal Riset Rumpun Ilmu Teknik (JURRITEK)*, vol. 1, no. 2, 2022.
- [19] M. Wibowo, M. Rizieq, and F. Djafar, “Perbandingan Metode Klasifikasi Untuk Deteksi Stress Pada Mahasiswa di Perguruan Tinggi,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 2023, doi: 10.30865/mib.v7i1.5182.
- [20] C. Schröer, F. Kruse, and J. M. Gómez, “A systematic literature review on applying CRISP-DM process model,” in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 526–534. doi: 10.1016/j.procs.2021.01.199.