

### “Mental Stress Classification on Student Dataset”



# Universitas Telkom

Dzaky Alfari	103012300391
Rafiq Labib	103012330353

**Bandung 2024/2025**

# **BAB I**

## **Pendahuluan**

### **1.1. Latar Belakang**

Kasus bunuh diri merupakan salah satu masalah kesehatan masyarakat yang semakin menjadi perhatian serius di Indonesia. Dilansir dari Goodstats Pada tahun 2020 tercatat ada lebih dari 640 kasus bunuh diri yang ditangani oleh Polri. Di tahun selanjutnya pada 2021 jumlah kasus bunuh diri turun tipis menjadi 629 kasus. Namun, angka tersebut menukik tajam setahun setelahnya. Tahun 2022 sebanyak 887 jiwa melayang akibat bunuh diri. Jumlah kasus ini terus naik di tahun 2023 bahkan mencapai 1.288 kasus.

Salah satu faktor utama yang berkontribusi signifikan adalah kondisi kesehatan mental, khususnya tingkat stres yang dialami oleh kelompok rentan seperti mahasiswa. Kesehatan mental mahasiswa merupakan isu yang sangat penting dan semakin mendapat perhatian seiring dengan peningkatan kasus stres serta depresi di lingkungan akademik. Mahasiswa menghadapi berbagai tekanan, mulai dari tuntutan akademik, adaptasi sosial, hingga permasalahan pribadi, yang dapat memicu timbulnya stres mental. Apabila tidak ditangani dengan baik, kondisi tersebut dapat berdampak serius terhadap kualitas hidup maupun prestasi akademik mahasiswa. Sebagaimana dijelaskan oleh Sugihartono et al. (2020), gangguan kesehatan mental cenderung lebih sulit dideteksi dibandingkan gangguan fisik karena gejalanya yang tidak langsung terlihat, sehingga mengakibatkan rendahnya kesadaran terhadap permasalahan tersebut di kalangan mahasiswa.

Oleh karena itu, pemanfaatan pendekatan berbasis data dengan metode klasifikasi menjadi solusi yang sangat potensial untuk mengatasi permasalahan ini. Dengan memanfaatkan dataset “Student Mental Stress and Coping Mechanisms”, sistem yang dikembangkan dapat secara efektif mengidentifikasi dan mengklasifikasikan tingkat stres mental mahasiswa secara akurat. Pendekatan ini tidak hanya memungkinkan deteksi dini terhadap individu yang berisiko tinggi mengalami stres berat, tetapi juga membuka peluang untuk intervensi yang lebih tepat sasaran. Dengan demikian, sistem klasifikasi ini dapat menjadi alat bantu penting dalam mendukung kesejahteraan mental mahasiswa dan mencegah konsekuensi serius, termasuk risiko bunuh diri.

### **1.2. Rumusan Masalah**

1. Bagaimana merancang Fuzzy Inference System untuk masalah klasifikasi mental stress pada mahasiswa?
2. Bagaimana hasil klasifikasi yang dihasilkan dari Fuzzy Inference System?
3. Bagaimana hasil perbandingan algoritma Fuzzy Inference System dan algoritma Random Forest?

### **1.3. Tujuan**

1. Melakukan implementasi algoritma Fuzzy Inference System pada dataset yang sudah di preprocessing.
2. Menunjukkan hasil evaluation metrics dari klasifikasi mental stress pada mahasiswa dengan algoritma Fuzzy.

3. Membandingkan hasil klasifikasi mental stress (evaluation metrics) mahasiswa dengan algoritma Fuzzy dengan algoritma Random Forest.

#### 1.4.Solusi

- Melakukan eksplorasi dataset dan PreProcessing pada dataset.
- Menggunakan model Fuzzy dan Random Forest, untuk mengevaluasi
- Membandingkan hasil dari evaluation metrics dari kedua model yang digunakan.
- Visualisasi evaluation metrics yang dihasilkan oleh kedua model yang digunakan dan melakukan analisis secara kuantitatif.

## BAB II METODE

### 2.1. Dataset

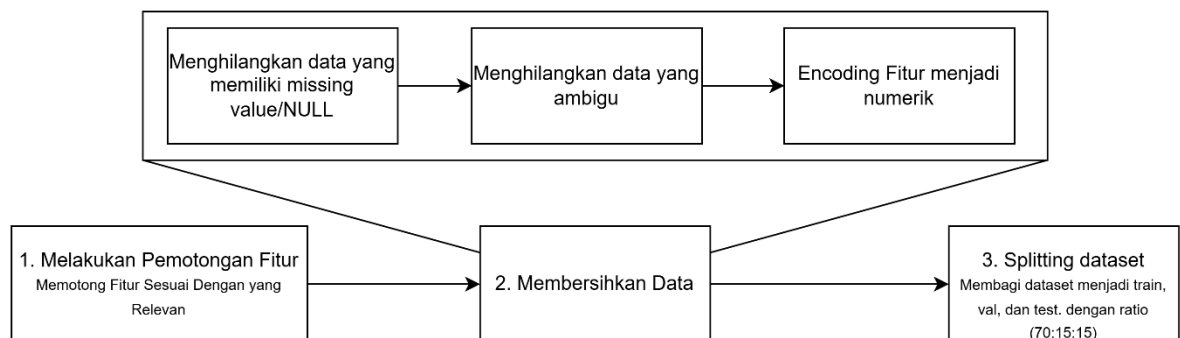
Dataset yang digunakan adalah “Student Depression Dataset” dataset terdiri dari beberapa fitur diantaranya Gender, Age, City, Profession, Academic Pressure, dan sebagainya. Dataset bersifat open public dan dapat di download via *Kaggle*. Dataset berguna untuk pengembangan model klasifikasi dengan tipe binary classification dengan hasil stress (1) dan tidak stress(0).

Adapun fitur lengkap dari dataset “Student Depression Dataset” , berikut merupakan penjabaran tiap fitur yang tersedia dari dataset “Student Depression Dataset”:

Fitur	Deskripsi
ID	Identifier unik untuk setiap mahasiswa
Age	Usia mahasiswa
Gender	Jenis kelamin mahasiswa
City	Kota tempat mahasiswa tinggal
CGPA	Nilai IPK atau nilai rata-rata akademik mahasiswa
Academic Pressure	Tingkat tekanan yang dirasakan mahasiswa terkait akademik
Study Satisfaction	Tingkat kepuasan mahasiswa terhadap proses dan hasil belajar
Sleep Duration	Durasi tidur harian (kategori seperti '5-6 hours', 'Less than 5 hours', dll)
Dietary Habits	Kebiasaan makan mahasiswa (Healthy, Moderate, dll)
Work Pressure	Tingkat tekanan yang dirasakan mahasiswa terkait pekerjaan
Job Satisfaction	Tingkat kepuasan mahasiswa terhadap pekerjaan
Work/Study Hours	Jumlah jam kerja atau belajar per hari (numerik skala 0.0 - 5.0)
Profession	Pekerjaan atau profesi mahasiswa
Degree	Tingkat pendidikan yang sedang ditempuh oleh mahasiswa

Financial Stress	Tingkat stres yang dirasakan mahasiswa terkait masalah keuangan
Family History of Mental Illness	Riwayat penyakit mental dalam keluarga (Yes/No)
Have you ever had suicidal thoughts ?	Apakah mahasiswa pernah memiliki pikiran untuk bunuh diri (Yes/No)
Depression_Status	Variabel target, indikator biner (0/1 atau Ya/Tidak) yang menunjukkan apakah mahasiswa mengalami depresi

## 2.2. Pra Pemrosesan



Gambar 1. WorkFlow dari Pra Pemrosesan

- **Melakukan pemotongan fitur**

Pemotongan fitur ini didasarkan dengan dukungan jurnal penelitian yang sudah dilakukan oleh para peneliti sebelumnya.

Fitur yang dipilih	Deskripsi
Have you ever had suicidal thoughts ?	Apakah mahasiswa pernah berpikir untuk bunuh diri,
Academic Pressure	Tingkat tekanan yang dirasakan mahasiswa terkait akademik
Financial Stress	Tingkat tekanan yang dirasakan mahasiswa terkait keuangan

Pemilihan fitur "Financial Stress" didasarkan pada temuan Kendall A. Lawley dkk. (2023) yang menyatakan bahwa mahasiswa yang mengalami tekanan finansial cenderung memiliki tingkat stres lebih tinggi dan kesehatan mental yang lebih buruk.

Hal ini diperkuat oleh analisis korelasi Pearson dan Bayesian yang menunjukkan hubungan signifikan antara financial strain dengan gejala kesehatan mental.

Sementara itu, fitur "Academic Pressure" dipilih berdasarkan hasil meta-analisis M. Agung Rahmadi dkk. (2024) yang mengonfirmasi hubungan kuat dan konsisten antara stres akademik dengan berbagai indikator kesehatan mental.

Adapun fitur "Have you ever had suicidal thoughts?" didukung oleh penelitian Xueqing Wang & Huafang Li (2025) yang menyimpulkan bahwa distress psikologis dapat mempengaruhi munculnya ideasi bunuh diri, terutama pada individu yang rentan. Dengan demikian, ketiga fitur ini saling terkait dan relevan dalam memprediksi risiko gangguan kesehatan mental.

- **Menghilangkan data yang missing value**

dengan menggunakan module pandas. Menghilangkan data yang missing value berguna untuk menghindari error dan bias dalam membangun model klasifikasi.

- **Menghilangkan nilai ambigu**

Menghilangkan nilai ambigu pada fitur yang telah dipilih. Hal ini bertujuan untuk membuang data yang nilainya ambigu, dan sulit diinterpretasikan secara konsisten. Pada kasus dari fitur yang kami pilih ada pada kolom "Financial Stress" yaitu nilai "?", dan pada kolom "Academic Pressure" yaitu nilai "0.0".

- **Encoding fitur kategorikal**

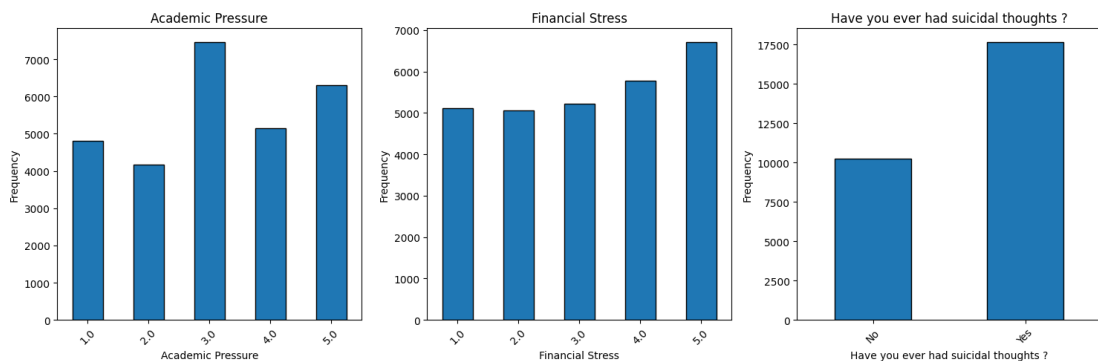
Pada tahap ini bertujuan untuk mengubah nilai kategorikal menjadi numerik, pada fitur yang sudah kami pilih yaitu mengubah nilai "yes" menjadi 1, dan nilai "no" menjadi nilai 0.

- **Splitting Dataset**

Pada tahap ini dataset dibagi menjadi train, val, dan test. dengan ratio (70:15:15) yang bertujuan agar data yang digunakan untuk melatih model dan hasil klasifikasi dari model tidak bias.

### 2.3. Exploratory Data Analysis

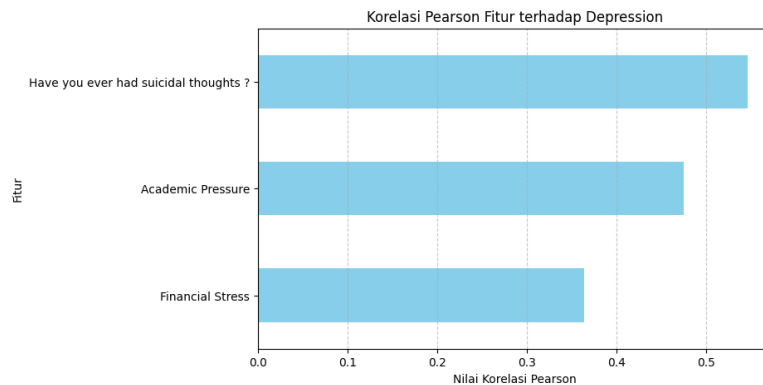
Eksplorasi analisis dataset ini dilakukan untuk memahami distribusi dan hubungan antar fitur dalam dataset terhadap label. Fokus utama diberikan pada tiga fitur yang sudah dipilih menjadi fitur utama, yaitu "Have you ever had suicidal thoughts ?", "Academic Pressure", "Financial Stress".



Gambar 2. Distribusi Analisis Fitur

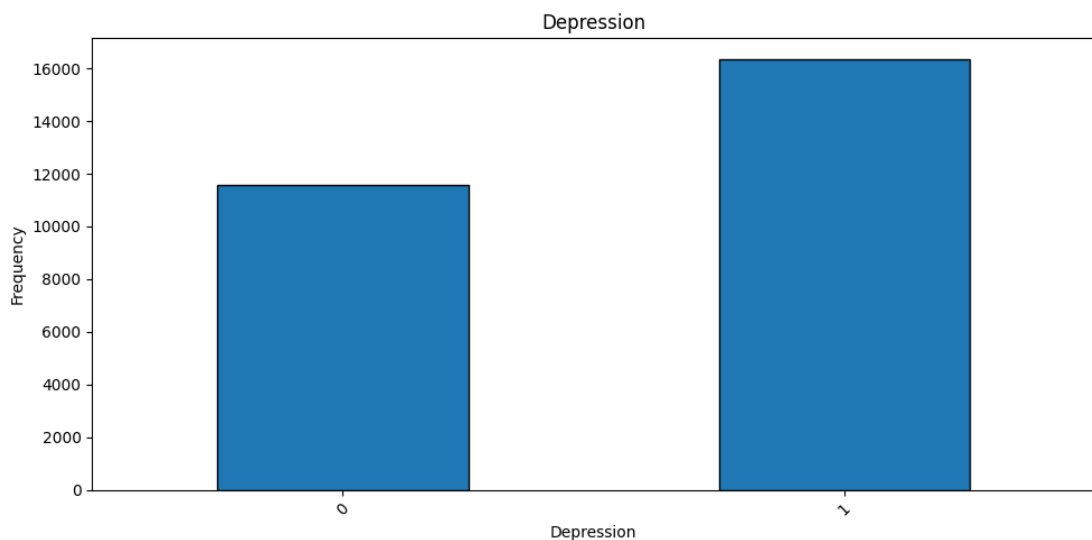
Gambar 2 menunjukkan visualisasi Distribusi fitur pada dataset Student Depression Dataset dimana sebanyak 6000 lebih mahasiswa memilih Academic Pressure pada tingkat 5. Sedangkan Financial Stress terdapat sebanyak 6000 lebih mahasiswa memilih di Financial Stress di tingkat 5. Sedangkan pada fitur mahasiswa memiliki pemikiran bunuh diri sebanyak 17500+ menjawab ya.

Korelasi ketiga fitur yang dipilih terhadap target/label:



Gambar 3. Korelasi Setiap Fitur dengan Target

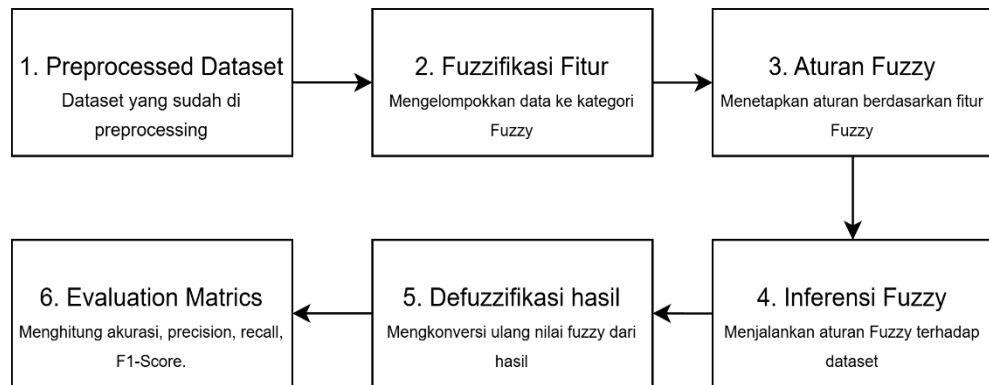
Gambar 3 menunjukkan bahwa ketiga fitur yang dipilih memiliki korelasi diatas 0,3. dengan suicidal thoughts memiliki korelasi setinggi 0,54. dan Financial stress memiliki korelasi 0,36 terhadap target. dan Academic Pressure setinggi 0,475 terhadap label/target.



Gambar 3. Distribusi nilai target pada dataset

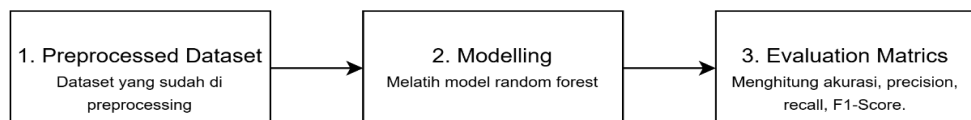
Gambar 3 menunjukkan bahwa pada dataset sebanyak 16314 mahasiswa masuk kedalam kategori "1" yaitu depresi. sedangkan terdapat 11545 mahasiswa masuk kedalam kategori "0" yaitu tidak depresi.

## 2.4.WorkFlow Rancangan Model



Gambar 4. Workflow Fuzzy Mamdani

Metode klasifikasi Fuzzy, melakukan klasifikasi Fuzzy dimulai dengan melakukan fuzzifikasi fitur, di mana data yang ada dikelompokkan menjadi kategori fuzzy, seperti Tinggi, Sedang dan Rendah. Setelah melakukan fuzzifikasi dilakukan pula menerapkan aturan kombinasi dari fuzzy, yang bertujuan untuk menentukan kategori dari klasifikasi Mental Stress Level yaitu Tinggi, Sedang, dan Rendah. Dilanjutkan dengan melakukan Inferensi Fuzzy System, yaitu menjalankan fuzzy sesuai dengan aturan fuzzy yang sudah dibuat. Selanjutnya adalah defuzzifikasi yaitu mengkonversi output fuzzy menjadi nilai yang pasti.



Gambar 5. WorkFlow Random Forest

Metode klasifikasi Random Forest, melakukan klasifikasi menggunakan Random Forest dimulai dengan melakukan preprocessing dataset, di mana data yang ada dibersihkan, di-encode, dan dinormalisasi agar siap digunakan untuk pemodelan. Selanjutnya, dilakukan tahap modelling dengan membangun model Random Forest menggunakan data pelatihan. Model ini terdiri dari kumpulan pohon keputusan (decision trees) yang bekerja secara ensemble untuk meningkatkan akurasi dan mengurangi risiko overfitting. Setelah model selesai dilatih, dilakukan prediksi pada data pengujian untuk menghasilkan klasifikasi tingkat Mental Stress Level menjadi kategori Tinggi, Sedang, dan Rendah.



## 2.5.Rancangan Model

### 2.5.1. Sistem Fuzzy

Berdasarkan EDA didapatkan Fungsi Keanggotaan sebagai berikut:

- Fungsi keanggotaan

Have you ever had suicidal thoughts ?

- Yes : (0, 0, 0.4, 0.5)
- No : (0.5, 0.6, 1, 1)

Academic Pressure

- Low = (0, 0, 2.5, 3.5)
- Medium = (2.3, 3.5, 4.3)
- High = (3.5, 4.5, 5, 5)

Financial Stress

- Low = (0, 0, 1.5, 2.5)
- High = (2.5, 3.5, 5, 5)

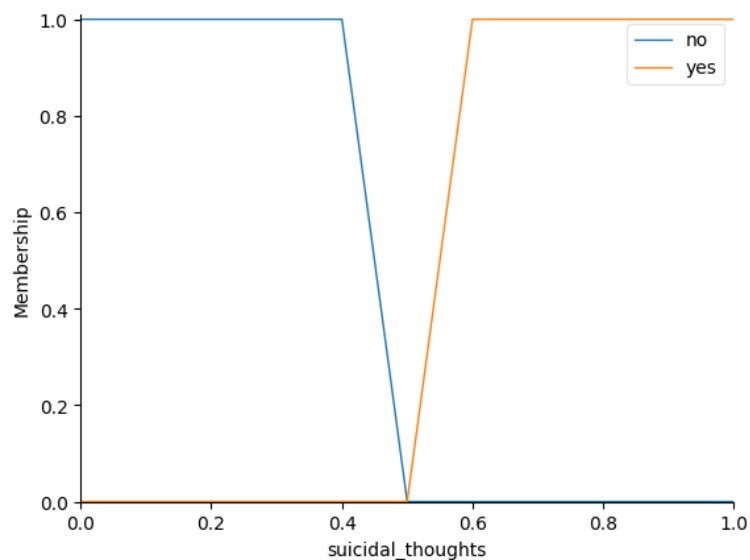
Depression

- No: (0,0,0.3,0.5)
- Yes: (0.5,0.7,1,1)

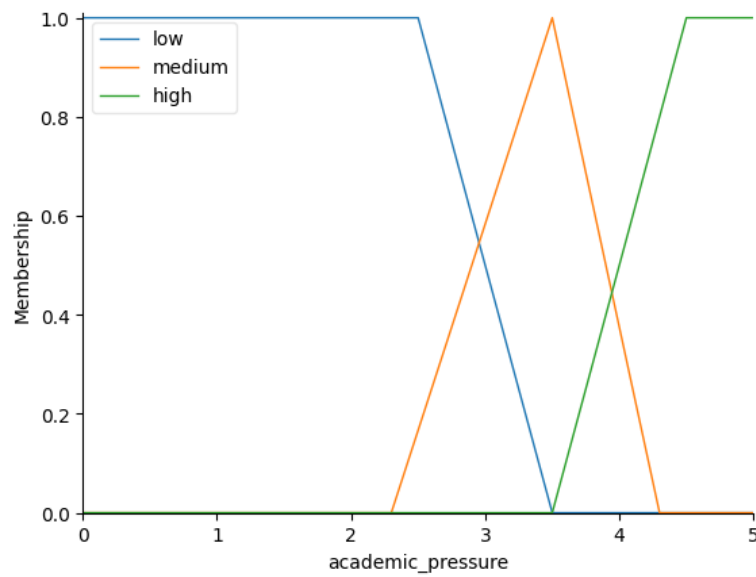
Berdasarkan korelasi data dan label maka ditetapkan aturan fuzzy yang dimana ketika suicidal thoughtnya “no” maka mayoritas outputnya adalah “no” pula. Sedangkan suicidal thoughtnya “yes” maka mayoritas outputnya adalah “yes”.

- Aturan Fuzzy :
  - IF suicidal\_thoughts == 'no' AND academic\_pressure == 'low' AND financial\_stress == 'low' THEN depression\_status = 'no'
  - IF suicidal\_thoughts == 'no' AND academic\_pressure == 'low' AND financial\_stress == 'high': depression\_status = 'no'
  - IF suicidal\_thoughts == 'no' AND academic\_pressure == 'medium' AND financial\_stress == 'low' THEN depression\_status = 'no'
  - IF suicidal\_thoughts == 'no' AND academic\_pressure == 'medium' AND financial\_stress == 'high' THEN depression\_status = 'yes'
  - IF suicidal\_thoughts == 'no' AND academic\_pressure == 'high' AND financial\_stress == 'low' THEN depression\_status = 'no'
  - IF suicidal\_thoughts == 'no' AND academic\_pressure == 'high' AND financial\_stress == 'high' THEN depression\_status = 'yes'

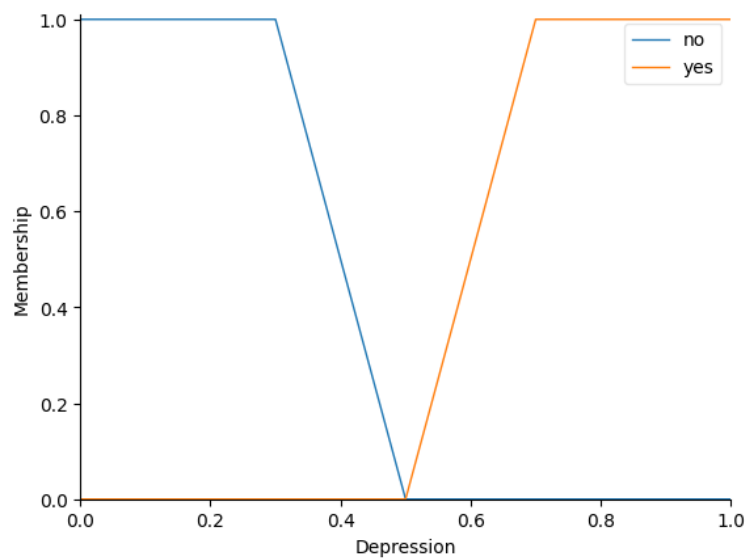
- IF suicidal\_thoughts == 'yes' AND academic\_pressure == 'low' AND financial\_stress == 'low' THEN depression\_status = 'no'
- IF suicidal\_thoughts == 'yes' AND academic\_pressure == 'low' AND financial\_stress == 'high' THEN depression\_status = 'yes'
- IF suicidal\_thoughts == 'yes' AND academic\_pressure == 'medium' AND financial\_stress == 'low' THEN depression\_status = 'yes'
- IF suicidal\_thoughts == 'yes' AND academic\_pressure == 'medium' AND financial\_stress == 'high' THEN depression\_status = 'yes'
- IF suicidal\_thoughts == 'yes' AND academic\_pressure == 'high' AND financial\_stress == 'low' THEN depression\_status = 'yes'
- IF suicidal\_thoughts == 'yes' AND academic\_pressure == 'high' AND financial\_stress == 'high' THEN depression\_status = 'yes'



Gambar 6. Graph Visualisasi Suicidal Thoughts



Gambar 7. Graph Visualisasi Academic Pressure



Gambar 8. Graph Visualisasi Financial Stress

### 2.5.2. Random Forest

- Membuat objek model Random Forest, dengan pengaturan seed random = SEED.
- Melatih model Random Forest dengan dataset train.
- Melakukan testing prediksi dengan dataset test.

## 2.6. Hasil dan Analisis

### 2.6.1. Sistem Fuzzy

Berdasarkan fuzzy sistem yang kami kembangkan didapatkan hasil :

Confusion Matrix	Jumlah
True Positive	1217
True Negative	2180
False Positive	517
False Negative	270

Evaluation Metrik	Hasil
Akurasi	0.8119
Recall	0.8897
Precision	0.8083
F1-Score	0.8470

### 2.6.2. Random Forest

Berdasarkan model random forest yang kami kembangkan didapatkan hasil :

Confusion Matrix	Jumlah
True Positive	1329
True Negative	2112
False Positive	405
False Negative	338

Evaluation Metrik	Hasil
Akurasi	0.8224
Recall	0.8390
Precision	0.8620
F1-Score	0.8504

### 2.6.3. Analisis

Berikut adalah ringkasan hasil didapatkan oleh kedua model

Metrik	Fuzzy System	Random Forest	Selisih
Akurasi	0.8119	0.8224	+0.105
Recall	0.8897	0.8390	-0.050
Precision	0.8083	0.8620	+0.015
F1-Score	0.8470	0.8504	+0.003

Perbandingan metrik evaluasi antara Sistem Fuzzy dan Random Forest menunjukkan bahwa Random Forest memiliki keunggulan pada akurasi, precision, dan F1-score, sementara Sistem Fuzzy unggul pada metrik recall. Perbedaan ini dapat dijelaskan oleh karakteristik dan pendekatan masing-masing model dalam memproses data. Random Forest secara otomatis belajar pola dari data melalui ensemble banyak pohon keputusan, sehingga mampu menghasilkan prediksi yang lebih tepat secara keseluruhan. Dengan mekanisme ini, Random Forest cenderung mengurangi kesalahan prediksi dan lebih fleksibel menghadapi variasi data dibandingkan Sistem Fuzzy yang mengandalkan aturan tetap yang telah didefinisikan secara manual. Sistem Fuzzy unggul dalam recall karena aturan-aturan yang dirancang berhasil meminimalkan risiko melewatkan kasus positif (False Negative). Dengan kata lain, sistem ini cenderung lebih konservatif dalam mengenali kasus depresi, sehingga menghasilkan lebih banyak True Positive, meskipun hal ini dapat meningkatkan jumlah False Positive.

Precision yang lebih tinggi pada Random Forest menunjukkan bahwa model ini lebih akurat saat memprediksi kelas positif, sehingga menghasilkan lebih sedikit False Positive dibanding Sistem Fuzzy. Hal ini berkaitan dengan kemampuan Random Forest dalam mengidentifikasi pola yang benar-benar relevan secara statistik tanpa terlalu banyak memicu prediksi positif palsu, yang kadang terjadi pada Sistem Fuzzy karena aturan yang lebih luas atau general. F1-score, yang merupakan hasil gabungan dari perhitungan precision dan recall, menunjukkan bahwa meskipun perbedaan antara kedua model cukup kecil, Random Forest berhasil mempertahankan keseimbangan yang baik antara kemampuan

menangkap kasus positif dan menghindari prediksi positif palsu. Sistem Fuzzy, meskipun sangat baik dalam recall, sedikit tertinggal dalam menjaga keseimbangan ini.

### **BAB III**

### **KESIMPULAN**

Berdasarkan serangkaian eksperimen dan evaluasi menggunakan metrik evaluasi , akurasi, precision, recall, dan F1-score, Random Forest berhasil menunjukkan performa tinggi dibandingkan Fuzzy Logic Mamdani. Random Forest mencapai mendapatkan score akurasi sebesar 82,24%, lebih unggul dibanding Fuzzy Logic Mamdani dengan akurasi 81,19%. Selain akurasi, Random Forest juga menunjukkan keunggulan pada metrik evaluasi pada precision dan F1-score dengan nilai masing-masing 0.8083 dan 0.8620, yang sedikit lebih tinggi dibandingkan sistem fuzzy yang masing-masing mencapai 0.8470 dan 0.8504. Hal ini mengindikasikan bahwa Random Forest lebih efektif dalam memprediksi kasus positif secara tepat, sehingga mampu mengurangi jumlah false positive. Namun, sistem fuzzy memiliki nilai recall yang lebih tinggi sebesar 0.8897 dibandingkan dengan Random Forest yang 0.8390, menunjukkan bahwa sistem fuzzy lebih sensitif dalam menangkap semua kasus positif atau depresi, meskipun dengan risiko meningkatnya false positive.

Dengan demikian, pemilihan metode terbaik dapat disesuaikan dengan kebutuhan aplikasi. Jika prioritas utama adalah mengurangi kemungkinan melewatkan kasus depresi, maka sistem fuzzy bisa menjadi pilihan yang lebih sesuai. Sebaliknya, jika diinginkan keseimbangan antara sensitivitas dan spesifisitas dengan akurasi yang lebih tinggi, maka Random Forest merupakan solusi yang lebih optimal.

Secara keseluruhan, penggunaan model Random Forest memberikan pendekatan yang lebih adaptif dan mampu menangani kompleksitas data dengan lebih baik melalui proses pembelajaran otomatis dari data, sehingga menghasilkan performa yang lebih konsisten dan andal dalam klasifikasi mental stress mahasiswa. Oleh karena itu, model Random Forest sangat direkomendasikan untuk digunakan sebagai alat bantu dalam deteksi dini stres mental mahasiswa guna mendukung intervensi yang lebih tepat dan efektif.



## BAB IV

### Referensi

1. E. H. Mamdani, "Application of fuzzy algorithms for control of simple dynamic plant," *Proceedings of the Institution of Electrical Engineers*, vol. 121, no. 12, pp. 1585–1588, 1974.
2. M. Sugeno and G. Kang, "Fuzzy modeling and control of multivariable systems," in *Proceedings of the IEEE International Symposium on Fuzzy Information Processing*, 1985, pp. 262–268.
3. Priscillia Pravina Putri Sugihartono, Nurul Hidayat, Tibyani, "Implementasi Metode Fuzzy Tsukamoto Untuk Deteksi Dini Tingkat Depresi Mahasiswa Yang Sedang Menempuh Skripsi (Studi Kasus: Fakultas Ilmu Komputer Universitas Brawijaya)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 4, no. 10, pp. 3432–3438, 2020.
4. Adil Shamim, "Student Depression Dataset," [Kaggle Dataset], Available at: <https://www.kaggle.com/datasets/adilshamim8/student-depression-dataset>
5. Novitayani, S., & Rekan. (2022). *Risiko bunuh diri pada mahasiswa tingkat akhir*. JIM Fkep, 6(2), 145-150. Universitas Syiah Kuala.
6. Goodstats. (2024). *Angka Kasus Bunuh Diri di Indonesia Meningkat 60% dalam 5 Tahun Terakhir*. Diakses dari <https://data.goodstats.id/statistic/angka-kasus-bunuh-diri-di-indonesia-meningkat-60-dalam-5-tahun-terakhir-2FzH6>.
7. K. A. Lawley, T. C. S. Caley, dan B. J. Lehman, "Financial strain and the health and well-being of college students during the COVID-19 pandemic," *Journal of American College Health*, vol. 73, no. 4, pp. 1320–1327, 2023, doi: 10.1080/07448481.2023.2253919.
8. X. Wang and H. Li, "Psychological distress and suicidal ideation in patients with depressive disorders: the chain mediation of psychological resilience and neuroticism," *\*International Journal of Nursing Studies Advances\**, vol. 8, p. 100325, 2025, doi: 10.1016/j.ijnsa.2025.100325.
9. M. A. Rahmadi, H. Nasution, L. Mawar, dan M. Sari, "Analisis Keterkaitan Kesehatan Mental dan Stres Akademik Mahasiswa," *Counseling For All: Jurnal Bimbingan dan Konseling*, vol. 4, no. 2, pp. 37–54, Sep. 2024. [Online]. Tersedia: <https://jurnal.uniraya.ac.id/index.php/Counseling>.