

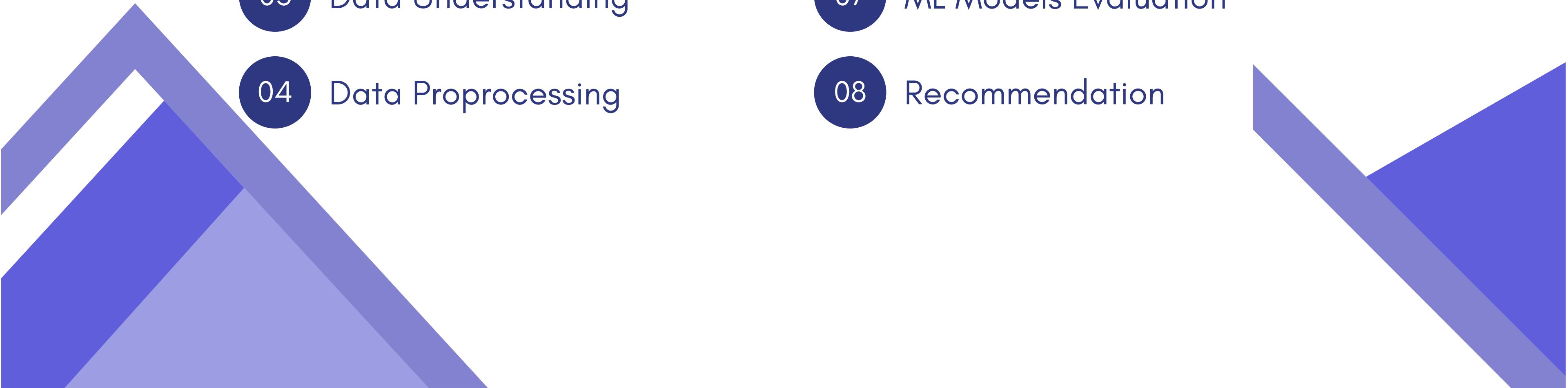
Final Project DS 33

Student Stress Prediction

Hanif Abdi Hakim Hani

13 September 2025

Table of Contents

- 
- 01 About Me
 - 02 About Project
 - 03 Data Understanding
 - 04 Data Proprocessing
 - 05 EDA
 - 06 ML Preparation
 - 07 ML Models Evaluation
 - 08 Recommendation

About Me



I'm an aspiring data scientist with a passion for transforming raw data into actionable insights.

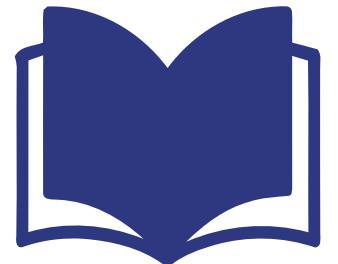
Education

Data Science Bootcamp (Apr 2025 – present)
dibimbing.id

Working Experience

Application Support Engineer (Feb 2024 – Feb 2025)
Nawa Data Solutions

About Project



Project Background

- Stres adalah masalah umum pada mahasiswa & individu muda.
- Tekanan berasal dari faktor akademik, sosial, maupun pribadi.
- Eustress → mendorong motivasi & kinerja.
- Distress → berdampak negatif pada kesehatan mental & prestasi.



Project Goals

- Menganalisis faktor-faktor penyebab stres.
- Mengklasifikasikan jenis stres (eustress vs distress).
- Memberikan wawasan & rekomendasi untuk pengelolaan stres yang lebih efektif.

Project Benefit

- Mendukung upaya peningkatan kesejahteraan mahasiswa.
- Membantu strategi pencegahan & penanganan stres.

Data Understanding

Dataset Overview

- Rows (entries): 843
- Columns (features): 26 Columns
- Memory Usage: 172 KB

Columns

1. Demographics:
 - Gender (int64)
 - Age (Int64)
2. Stress Indicators (Question):
 - 22 questions (0 to 5)
3. Target Variable:
 - Which type of stress do you primarily experience? (object)

Data Pre-Processing

Duplicated Data

Terdapat 27 duplikat data

Missing Value

Tidak ada missing value

Feature Engineering

Encoding get dummies untuk
penambahan kolom baru berdasarkan
stress type

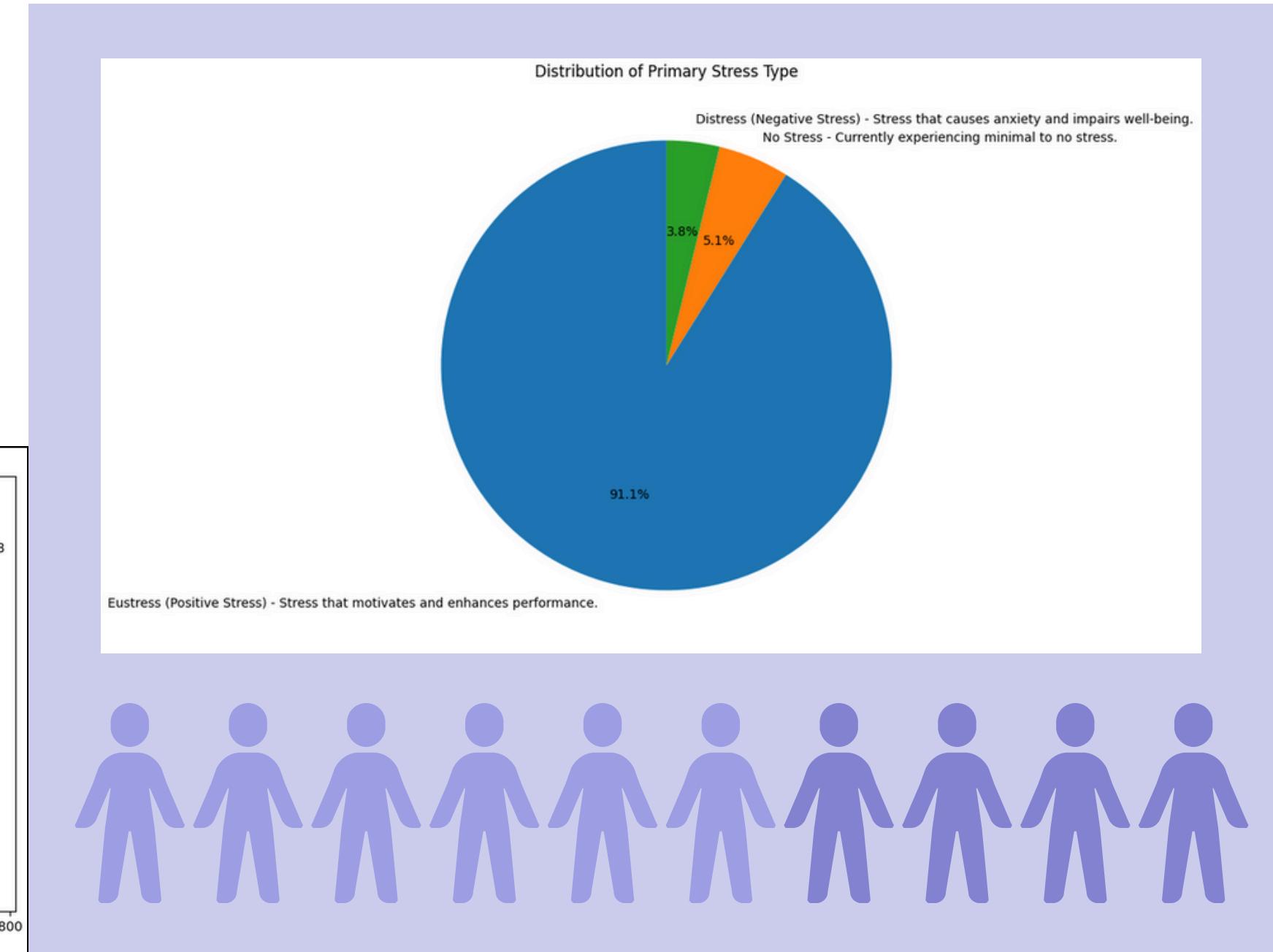
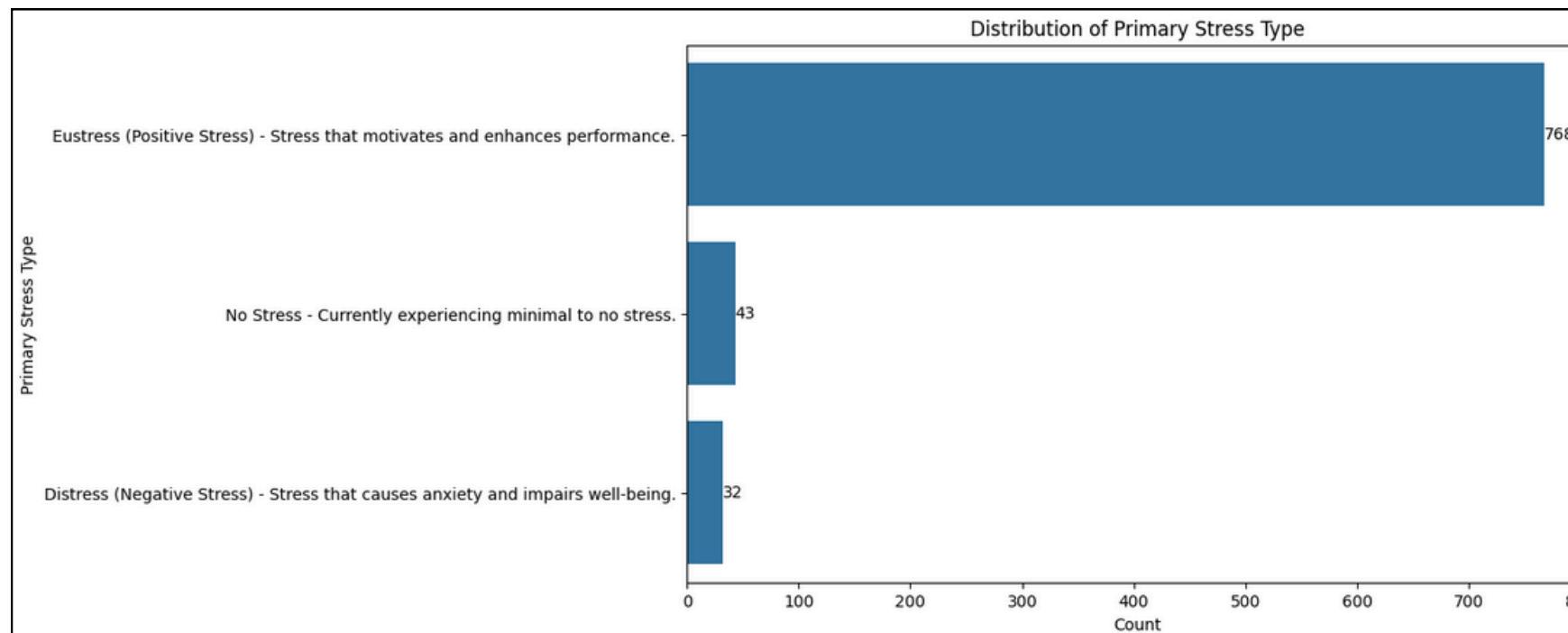


EDA

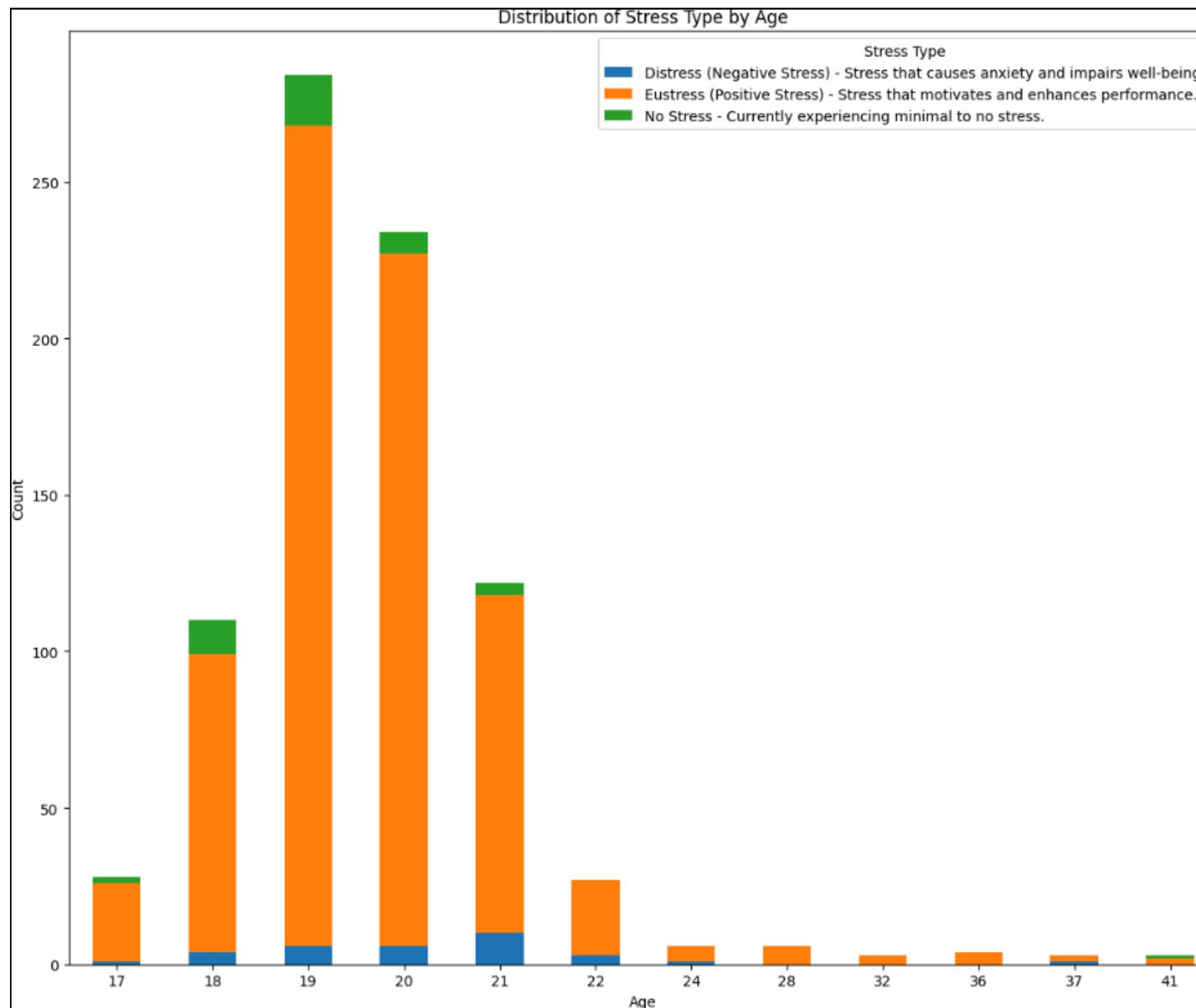
(Exploratory Data Analysis)

Stress Data Distribution

- Mayoritas responden (lebih dari 90%) mengalami Eustress
- Hanya sedikit yang mengalami Distress (3.8%), yang biasanya berdampak negatif pada kesehatan mental maupun produktivitas.
- Responden yang melaporkan tidak ada stres (5.1%) juga jumlahnya relatif kecil.

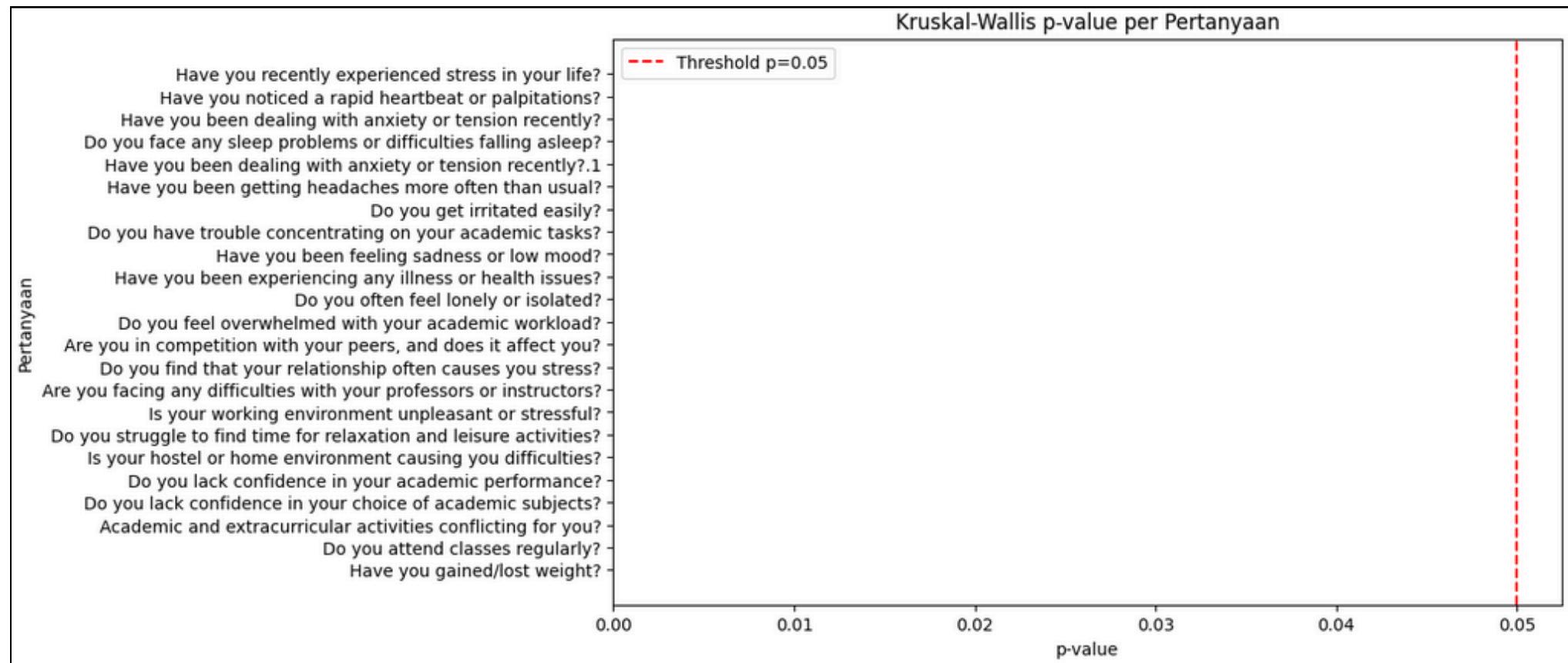


Distribution of Stress Type by Age



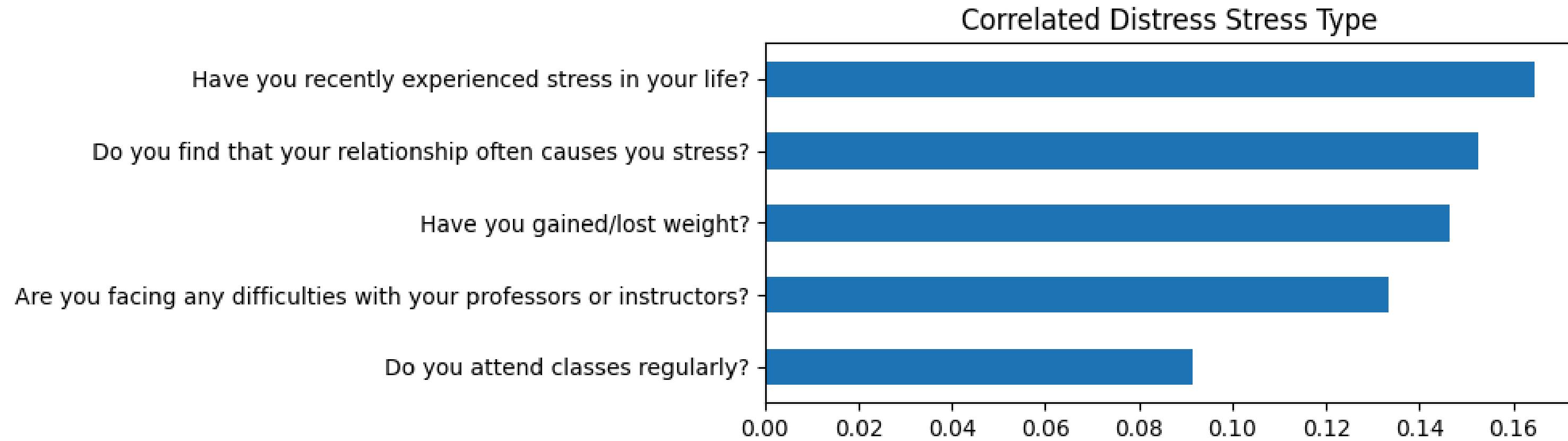
- Mayoritas usia 18–21 tahun.
- Dominasi eustress → stres positif yang mendorong motivasi & kinerja.
- Adanya distress → tetap perlu diwaspadai karena berdampak negatif pada kesejahteraan & prestasi akademik.

Hypothesis Testing



- Seluruh pertanyaan signifikan → p-value < 0,05
- Semua dimensi penting untuk dipertimbangkan dalam memahami faktor penyebab stres mahasiswa

Distress Type

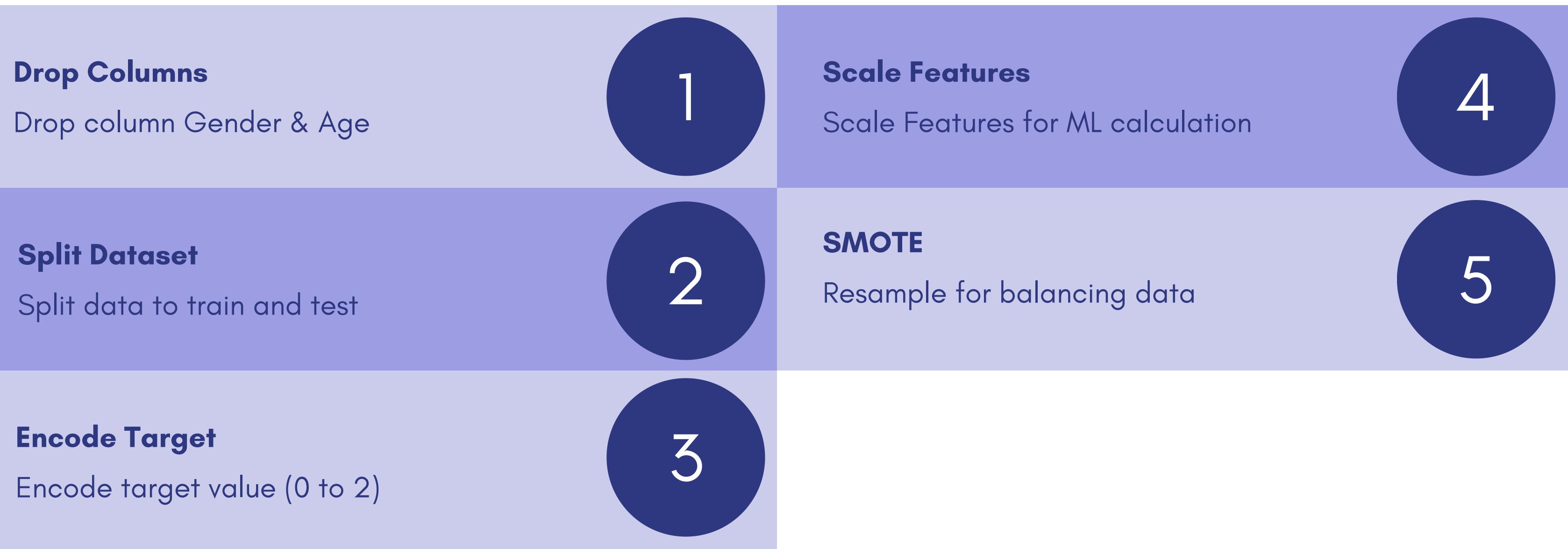


- Pengalaman stres sebelumnya → mahasiswa yang sudah sering mengalami stres lebih rentan terkena distress.
- Masalah hubungan sosial → hubungan pribadi sering menjadi sumber stres signifikan.
- Perubahan berat badan → indikasi adanya dampak fisik dari stres (gain/loss weight).
- Kesulitan dengan dosen/instruktur → interaksi akademik berpengaruh besar terhadap distress.
- Kehadiran kuliah → ketidakteraturan hadir kuliah berkorelasi dengan tingkat distress.

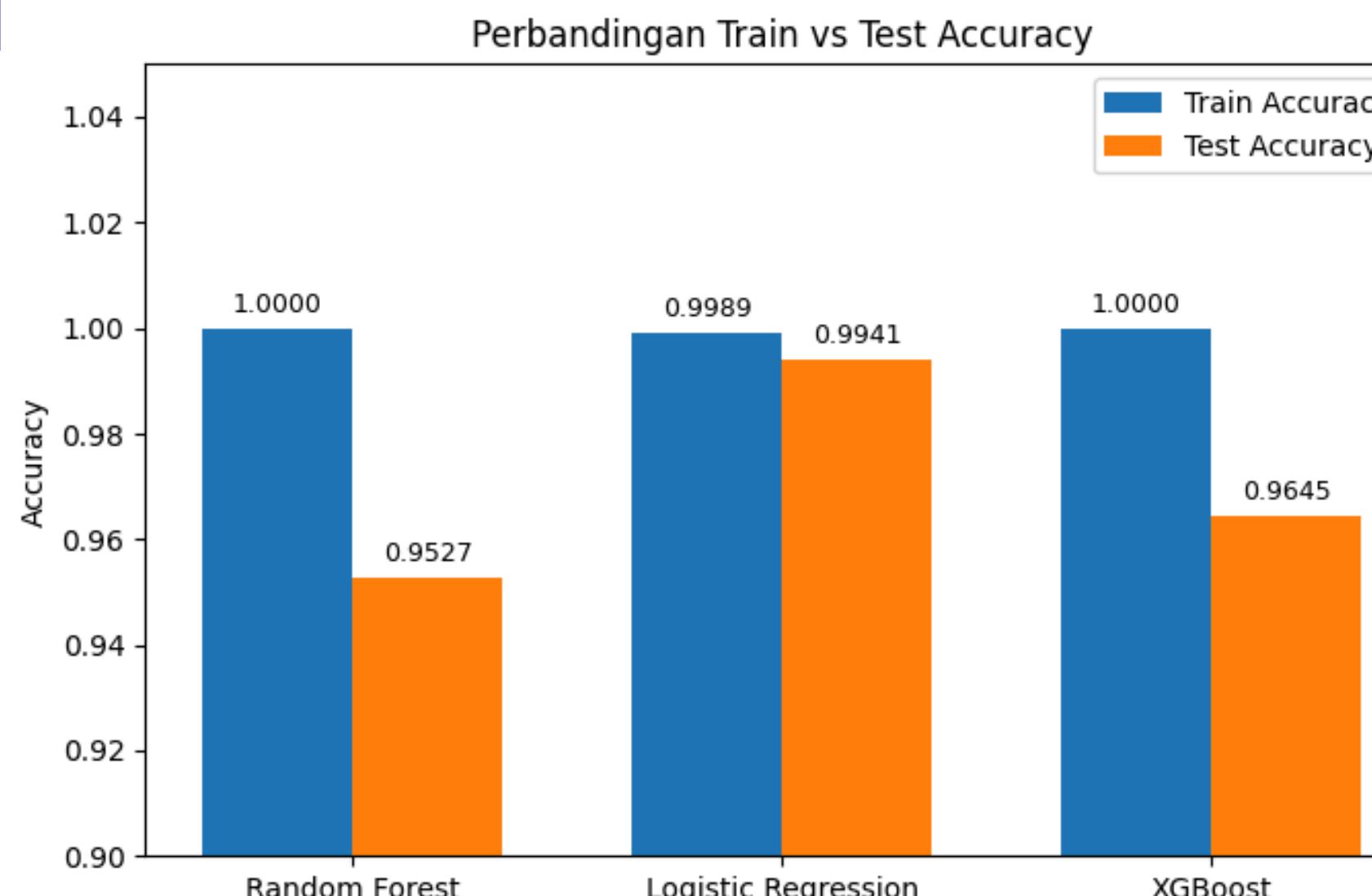


Machine Learning Preparation

ML Data Preparation



ML Models



Best Model: Logistic Regression with recall: 0.9941

Random Forest

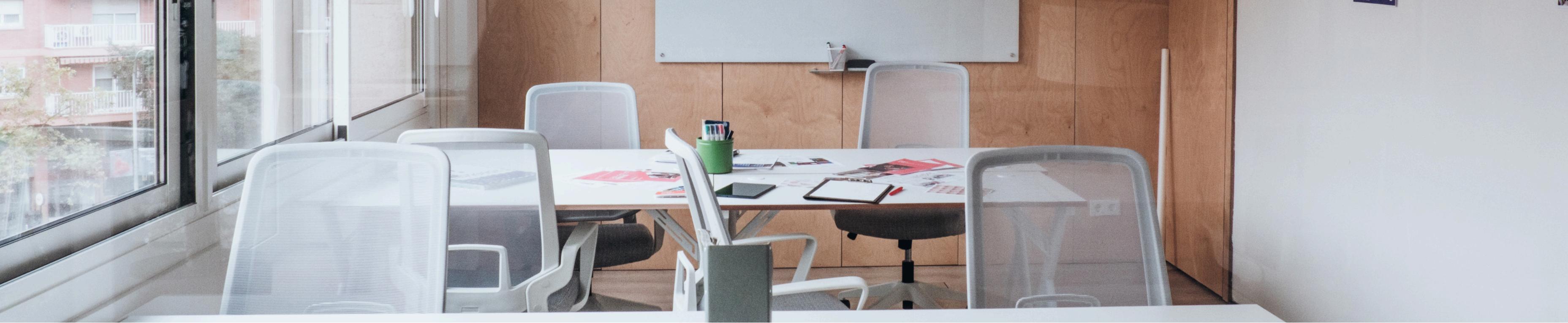
- Train: 100%, Test: 95,27%
- Indikasi overfitting → performa terlalu tinggi di train, menurun di test.

Logistic Regression

- Train: 99,89%, Test: 99,41%
- Paling seimbang → generalisasi kuat dan stabil.

XGBoost

- Train: 100%, Test: 96,45%
- Juga ada tanda overfitting, meski test accuracy cukup tinggi.



Machine Learning Evaluation

Classification Report Logistic Regression

	precision	recall	f1-score	support
<i>Distress (Negative Stress)</i> - Stress that causes anxiety and impairs well-being.	0.86	1.00	0.92	6
<i>Eustress (Positive Stress)</i> - Stress that motivates and enhances performance.	1.00	0.99	1.00	156
<i>No Stress</i> - Currently experiencing minimal to no stress.	1.00	1.00	1.00	7
		accuracy		0.99
		macro avg	0.95	1.00
		weighted avg	0.99	0.99

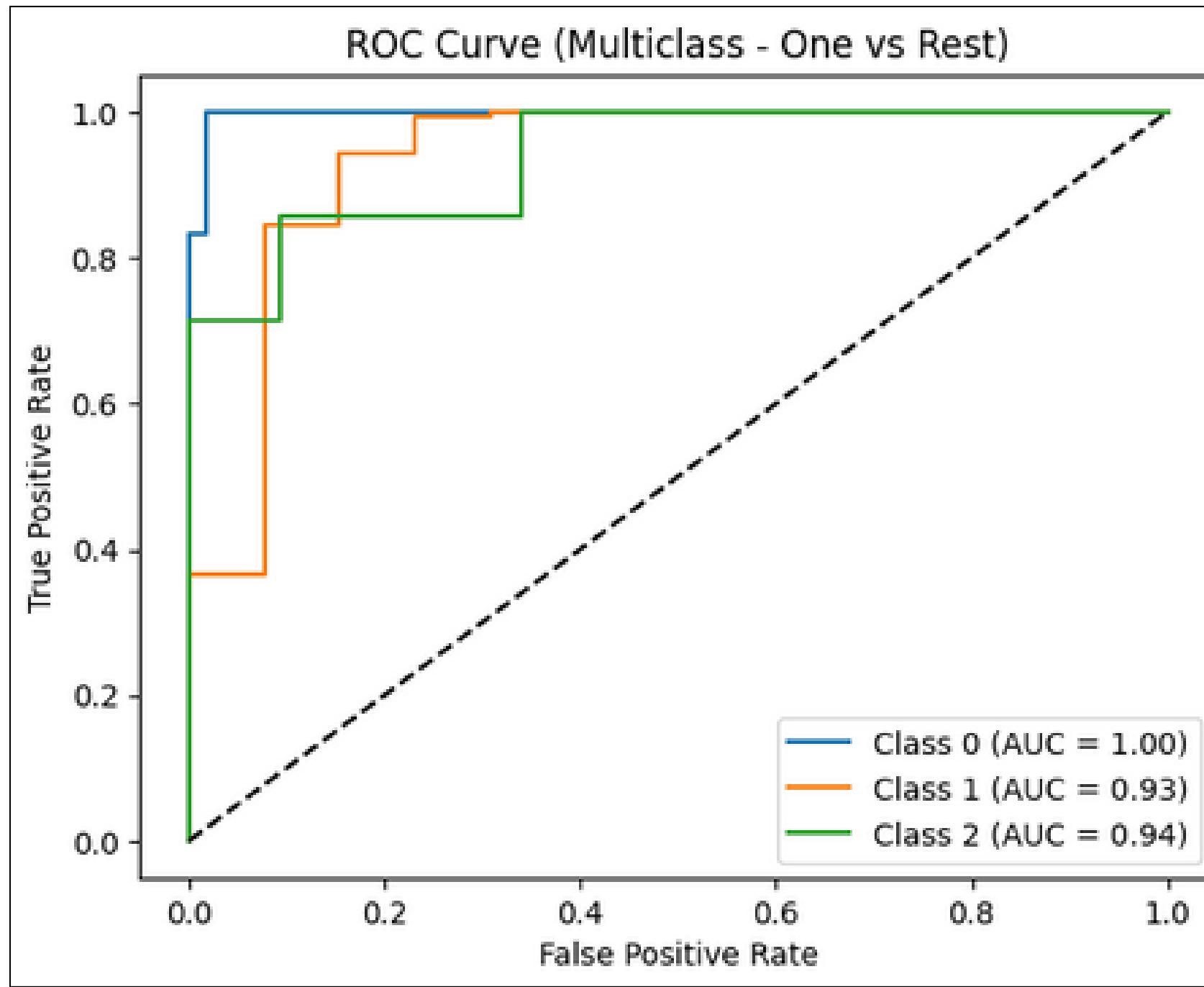
- Akurasi sangat tinggi: 99%.
- Eustress & No Stress → terdeteksi hampir sempurna.
- Distress → tetap teridentifikasi dengan baik meskipun jumlah data kecil.
- Kesalahan sangat rendah → tingkat error minimal.
- Model terbukti andal untuk klasifikasi tipe stres responden.

Confusion Matrix Logistic Regression

- Akurasi tinggi dalam mengklasifikasikan ketiga tipe stres.
- Distress, eustress, dan no stress dapat dibedakan dengan baik.
- Kesalahan kecil hanya terjadi pada kategori eustress.
- Model terbukti andal dan konsisten untuk prediksi tipe stres.

		Confusion Matrix		
		Distress (Negative Stress) - Stress that causes anxiety and impairs well-being. -	Eustress (Positive Stress) - Stress that motivates and enhances performance. -	No Stress - Currently experiencing minimal to no stress. -
Actual	Distress (Negative Stress) - Stress that causes anxiety and impairs well-being. -	6	0	0
	Eustress (Positive Stress) - Stress that motivates and enhances performance. -	1	155	0
	No Stress - Currently experiencing minimal to no stress. -	0	0	7
	Predicted			

ROC AUC Logistic Regression

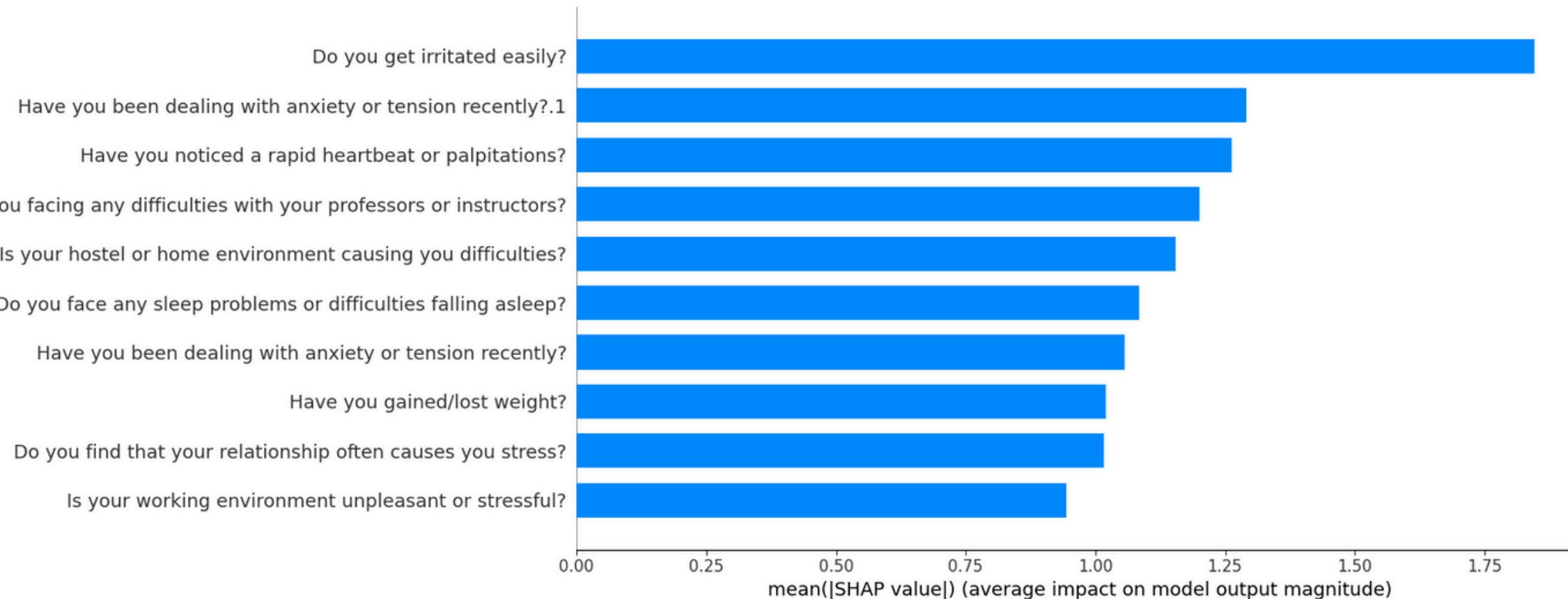


- Performa klasifikasi sangat tinggi → AUC > 0,90 pada semua kelas.
- Kelas 0 (No Stress) → teridentifikasi sempurna (AUC = 1.00).
- Kelas 1 (Eustress) → sangat baik (AUC = 0.93).
- Kelas 2 (Distress) → sangat baik (AUC = 0.94).
- Model memiliki kemampuan diskriminasi yang kuat dalam membedakan tipe stres responden.

Conclusion Model Logistic Regression

- Konsisten dalam klasifikasi stres.
- Akurasi tinggi.
- Kesalahan minimal.
- Generalisasi kuat → tetap akurat pada data uji.
- Kemampuan diskriminasi sangat baik (dibuktikan oleh nilai AUC > 0,90).

Interpretasi Model



Faktor emosional & psikologis

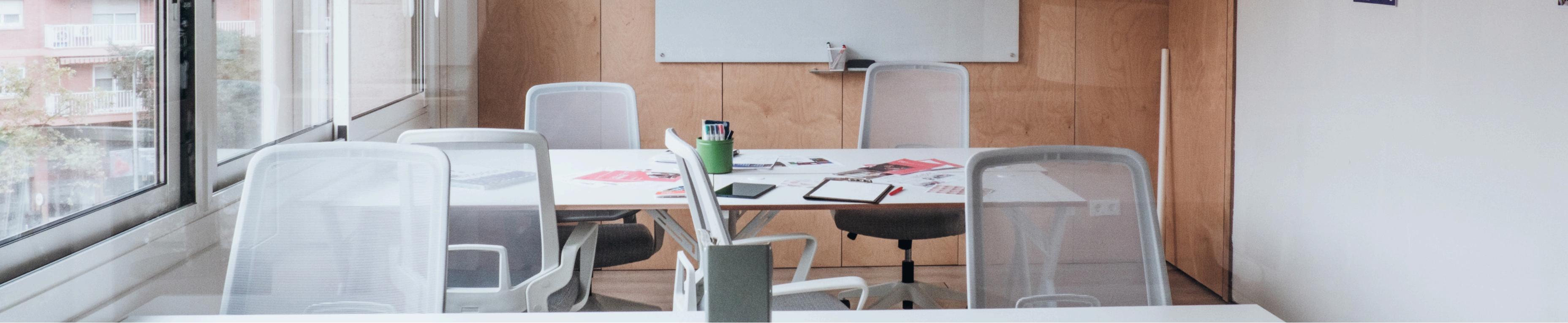
- Mudah marah
- Kecemasan/ketegangan

Faktor fisik

- Jantung berdebar
- Gangguan tidur
- Perubahan berat badan

Faktor lingkungan akademik & sosial

- Kesulitan dengan dosen/instruktur.
- Masalah di rumah/kos.
- Hubungan personal.
- Lingkungan akademik yang tidak nyaman.



Recommendation

Recommendation

1

Pendekatan Psikologis & Emosional

- Adakan counseling rutin untuk mengatasi kecemasan, mudah marah, dan ketegangan.
- Latih keterampilan stress management seperti mindfulness, relaksasi, atau journaling.

2

Kesehatan Fisik

- Fasilitasi pemeriksaan kesehatan berkala terkait gejala fisik (jantung berdebar, gangguan tidur, perubahan berat badan).
- Dorong gaya hidup sehat: olahraga, tidur cukup, pola makan seimbang.

3

Lingkungan Akademik

- Tingkatkan komunikasi antara dosen/instruktur dengan mahasiswa.
- Sediakan ruang diskusi atau academic support center untuk mahasiswa yang mengalami kesulitan.



Thank You