



LAPORAN PROYEK AKHIR

**KLASIFIKASI RISIKO STRESS
MAHASISWA DI INDONESIA
BERDASARKAN POLA AKADEMIK DAN
AKTIVITAS HARIAN MENGGUNAKAN
ALGORITMA RANDOM FOREST**

Disusun oleh: Kelompok 1

Latar Belakang

Kesehatan mental mahasiswa berperan penting dalam menunjang motivasi belajar dan prestasi akademik, namun tekanan tugas, tuntutan prestasi, dan perubahan gaya hidup sering memicu stres. Tingginya persentase mahasiswa yang mengalami gangguan mental menunjukkan perlunya deteksi dini berbasis data agar institusi dapat melakukan intervensi secara tepat.

Penelitian ini menerapkan algoritma Random Forest untuk mengklasifikasikan risiko stres mahasiswa berdasarkan faktor akademik, gaya hidup, dan sosial-ekonomi. Model dievaluasi menggunakan metrik klasifikasi serta diimplementasikan dalam aplikasi web berbasis Streamlit sebagai alat skrining awal yang praktis bagi lingkungan kampus.



Tujuan Penelitian



1 Menerapkan algoritma Random Forest untuk mengklasifikasikan tingkat risiko stres mahasiswa berdasarkan data akademik, gaya hidup, dan sosial-ekonomi.



2 Mengevaluasi kinerja model Random Forest dalam memprediksi risiko stres mahasiswa menggunakan metrik evaluasi seperti akurasi, precision, recall, dan confusion matrix.



3 Mengidentifikasi faktor-faktor yang paling berpengaruh terhadap risiko stres mahasiswa pada modelling.



Rumusan Masalah

1

Bagaimana menerapkan algoritma Random Forest untuk mengklasifikasikan risiko stres mahasiswa berdasarkan data akademik dan aktivitas harian?

2

Bagaimana performa model Random Forest dalam memprediksi risiko stres mahasiswa berdasarkan metrik evaluasi seperti akurasi, precision, recall, dan confusion matrix?

3

Faktor apa saja yang paling berpengaruh terhadap klasifikasi risiko stres mahasiswa berdasarkan analisis feature importance pada model Random Forest?

Tinjauan Studi

Penelitian terdahulu menunjukkan bahwa machine learning efektif untuk klasifikasi stres mahasiswa. Terdapat salah satu publikasi yang menggunakan algoritma Support Vector Machine (SVM).

Namun, Beberapa studi komparatif menunjukkan bahwa Random Forest memberikan performa yang lebih optimal dibandingkan Decision Tree dan Logistic Regression karena kemampuannya menangani hubungan nonlinier serta mengurangi overfitting melalui ensemble learning

Oleh karena itu, Random Forest dipilih sebagai metode utama dalam penelitian ini.

Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif menggunakan machine learning.

Penelitian ini termasuk dalam penelitian terapan, diimplementasikan dalam bentuk sistem prediksi berbasis web. Metode yang digunakan adalah klasifikasi supervised learning dengan algoritma Random Forest



Variabel Penelitian & Dependen

Penelitian ini menggunakan dua jenis fitur :

Numerikal

No.	Nama Fitur
1	Umur
2	Jam Belajar per Hari
3	Jam Tidur per Hari
4	IPK
5	Jumlah Tugas Besar per Minggu

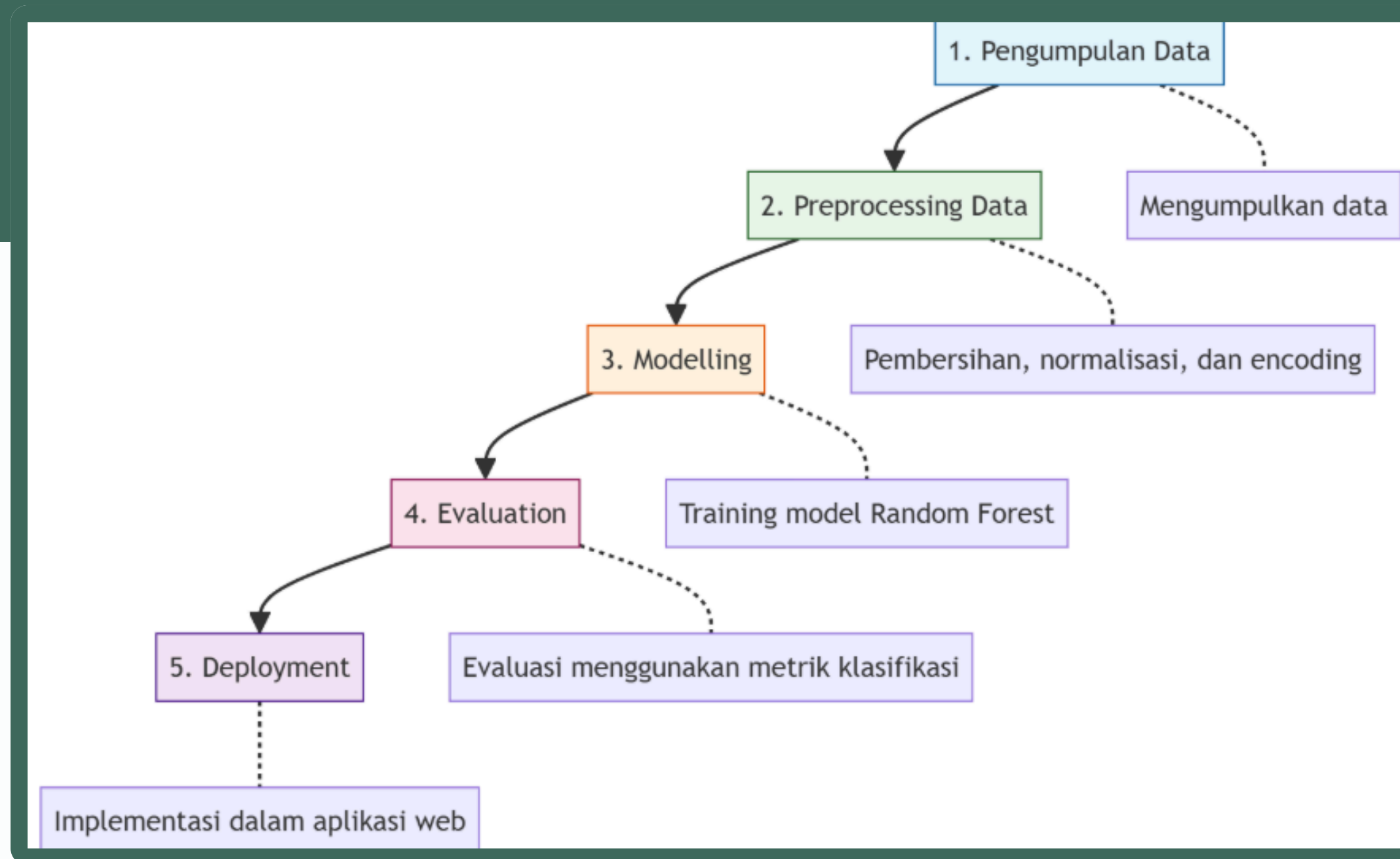
Kategorikal

No.	Nama Fitur
1	Gender
2	Jurusan/Program Studi
3	Frekuensi Olahraga
4	Pemasukan Keluarga
5	Status Hubungan

V. Dependen

Label	Keterangan
Sehat	Mahasiswa dengan risiko stres rendah
Risiko Stres	Mahasiswa dengan potensi risiko stres tinggi

Tahapan Penelitian



Penelitian ini mengikuti tahapan CRISP-DM (Cross-Industry Standard Process for Data Mining):

Preprocessing data & Metrik Evaluasi

Data dilakukan preprocessing dengan langkah-langkah sebagai berikut :

1. Pembersihan Data
2. Normalisasi Data
3. Encoding Data
4. Splitting Data

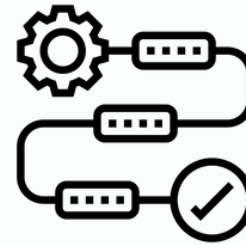
Berikut adalah Matriks Evaluasi yang digunakan untuk mengukur performa model :

Metrik	Fungsi
Accuracy	Mengukur proporsi prediksi yang benar
F1-Score	Harmonic mean dari precision dan recall
Confusion Matrix	Menunjukkan distribusi prediksi (TP, TN, FP, FN)
Classification Report	Laporan precision, recall, F1 per kelas

Kerangka Berpikir



**Input Data
Mahasiswa**



**Preprocessing,
Training, dan
Evaluasi
performa
model**



**Prediksi Risiko
Stres dengan
probabilitas
ketepatan
prediksi**




Kesimpulan

Berdasarkan hasil penelitian, algoritma Random Forest berhasil diterapkan untuk mengklasifikasikan kondisi kesehatan mental mahasiswa melalui tahapan pengumpulan data, preprocessing, pembangunan, dan evaluasi model. Dataset diproses dengan normalisasi dan encoding variabel agar dapat digunakan secara optimal oleh model.

Evaluasi menggunakan metrik klasifikasi menunjukkan bahwa model memiliki performa yang cukup baik serta mampu mengidentifikasi faktor-faktor utama yang memengaruhi kondisi kesehatan mental mahasiswa.

Model kemudian diimplementasikan dalam aplikasi berbasis Streamlit sebagai sistem peringatan dini yang interaktif dan mudah digunakan untuk mendukung deteksi risiko stres mahasiswa.



TERIMA KASIH

