

LEMBAR KERJA MAHASISWA (LKM)

LK.8 Perancangan Project Data Science

Nama	: ALDI SAPUTRA
Tanggal	: 6-12-2025
Kelas	: 5AI-A
Judul Project	: Sistem Peringatan Dini ( <i>Early Warning System</i> ) Mahasiswa Berisiko Gagal Studi Menggunakan Algoritma Random Forest

A. Format Perancangan

Tahapan CRISP-DM	Instruksi (Dari LKM)	Rancangan Implementasi
<b>1. Business Understanding</b> (Pemahaman Bisnis)	Identifikasi permasalahan & rumuskan tujuan bisnis.	<p><b>Konteks:</b> Perguruan Tinggi (Universitas Teknologi).</p> <p><b>Permasalahan:</b> Pihak program studi sering terlambat menyadari adanya mahasiswa yang mengalami penurunan performa akademik. Evaluasi biasanya baru dilakukan setelah hasil ujian keluar dengan status <i>Fail</i> atau <i>Pass with Grace</i> (Lulus Bersyarat), sehingga intervensi dosen wali menjadi tidak efektif.</p> <p><b>Tujuan Bisnis:</b> Membangun model prediksi otomatis untuk mengelompokkan mahasiswa ke dalam kategori "Aman" atau "Berisiko" berdasarkan tren nilai indeks prestasi (SGPA &amp; CGPA), agar Bimbingan Konseling dapat dilakukan lebih dini.</p>
<b>2. Data Understanding</b> (Pemahaman Data)	Jelaskan sumber, jenis data, fitur, dan target.	<p><b>Sumber Data:</b> <i>College Exam Result Dataset 2025</i> (Data Akademik Mahasiswa B.Tech).</p> <p><b>Jenis Data:</b> Numerik (Nilai IP) dan Kategorikal (Status Kelulusan).</p> <p><b>Fitur (Input - X):</b></p> <ol style="list-style-type: none"><li><b>SGPA</b> (<i>Semester Grade Point Average</i>): Nilai rata-rata semester berjalan.</li><li><b>CGPA</b> (<i>Cumulative Grade Point Average</i>): Nilai rata-rata kumulatif dari semester awal.</li></ol> <p><b>Target (Output - Y):</b></p>

Tahapan CRISP-DM	Instruksi (Dari LKM)	Rancangan Implementasi
		<p><b>Status Risiko.</b> Target ini diturunkan dari kolom Result Description.</p> <ul style="list-style-type: none"> <li>- Label 0 (Aman): Status <i>PASS</i>.</li> <li>- Label 1 (Berisiko): Status <i>FAIL</i> atau <i>PASS WITH GRACE</i>.</li> </ul>
<b>3. Data Preparation</b> (Persiapan Data)	Langkah pembersihan dan transformasi data.	<p><b>Pembersihan Data (<i>Cleaning</i>):</b></p> <p>Menghapus baris data mahasiswa yang memiliki nilai kosong (<i>missing values</i>) pada kolom SGPA atau CGPA untuk menjaga kualitas model.</p> <p><b>Transformasi Data:</b></p> <p>Melakukan <i>Feature Engineering</i> pada kolom target (Result Description). Kita mengubah teks menjadi label biner:</p> <ul style="list-style-type: none"> <li>- Jika status = "PASS" <math>\rightarrow</math> diubah menjadi <b>0</b>.</li> <li>- Jika status = "FAIL" atau "PASS WITH GRACE" <math>\rightarrow</math> diubah menjadi <b>1</b>.</li> </ul> <p>Hal ini agar algoritma <i>Machine Learning</i> dapat memproses klasifikasi risiko.</p>
<b>4. Modeling</b> (Pemodelan)	Pilih algoritma dan jelaskan alasannya.	<p><b>Algoritma:</b> Random Forest Classifier.</p> <p><b>Alasan Pemilihan:</b></p> <ol style="list-style-type: none"> <li><b>Akurasi Tinggi:</b> Random Forest bekerja dengan menggabungkan banyak <i>Decision Tree</i>, sehingga hasil prediksinya lebih akurat dan stabil dibandingkan pohon keputusan tunggal.</li> <li><b>Non-Linear:</b> Hubungan antara SGPA, CGPA, dan risiko kelulusan seringkali tidak linear (misal: IPK tinggi tapi IP semester anjlok drastis). Random Forest sangat baik menangani pola data seperti ini.</li> </ol>
<b>5. Evaluation</b> (Evaluasi)	Pilih metode evaluasi.	<p><b>Metode:</b> Confusion Matrix dan Akurasi.</p>

Tahapan CRISP-DM	Instruksi (Dari LKM)	Rancangan Implementasi
		<p><b>Fokus Evaluasi:</b></p> <p>Selain melihat akurasi total, kami memantau <b>Recall</b> untuk kelas "Berisiko".</p> <p>Tujuannya adalah memastikan sistem <b>sedikit mungkin melewatkan</b> mahasiswa yang sebenarnya berisiko gagal. Lebih baik sistem memberi peringatan palsu (waspada) daripada gagal mendeteksi mahasiswa yang akan <i>Drop Out</i>.</p>
6. Deployment (Penerapan)	Rancangan deployment & interface.	<p><b>Rancangan:</b> Dashboard Monitoring Akademik untuk Dosen Wali (Kaprodi).</p> <p><b>Tampilan Interface:</b></p> <p>Halaman web sederhana di mana Dosen dapat mengunggah file Excel nilai sementara. Sistem akan menampilkan tabel:</p> <ul style="list-style-type: none"> <li>- <b>Nama Mahasiswa</b></li> </ul>