

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
1. Business Understanding (Pemahaman Bisnis)	Identifikasi permasalahan & rumuskan tujuan bisnis.	<p>Konteks: Perguruan Tinggi (Universitas Teknologi).</p> <p>Permasalahan: 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>Tujuan Bisnis: Membangun model prediksi otomatis untuk mengelompokkan mahasiswa ke dalam kategori "Aman" atau "Berisiko" berdasarkan tren nilai indeks prestasi (SGPA & CGPA), agar Bimbingan Konseling dapat dilakukan lebih dini.</p>
2. Data Understanding (Pemahaman Data)	Jelaskan sumber, jenis data, fitur, dan target.	<p>Sumber Data: <i>College Exam Result Dataset 2025</i> (Data Akademik Mahasiswa B.Tech).</p> <p>Jenis Data: Numerik (Nilai IP) dan Kategorikal (Status Kelulusan).</p> <p>Fitur (Input - X):</p> <ol style="list-style-type: none"> SGPA (Semester Grade Point Average): Nilai rata-rata semester berjalan. CGPA (Cumulative Grade Point Average): Nilai rata-rata kumulatif dari semester awal. <p>Target (Output - Y):</p>

Tahapan CRISP-DM	Instruksi (Dari LKM)	Rancangan Implementasi
		<p>Status Risiko. Target ini diturunkan dari kolom Result Description.</p> <ul style="list-style-type: none"> - Label 0 (Aman): Status <i>PASS</i>. - Label 1 (Berasiko): Status <i>FAIL</i> atau <i>PASS WITH GRACE</i>.
3. Data Preparation (Persiapan Data)	Langkah pembersihan dan transformasi data.	<p>Pembersihan Data (<i>Cleaning</i>): Menghapus baris data mahasiswa yang memiliki nilai kosong (<i>missing values</i>) pada kolom SGPA atau CGPA untuk menjaga kualitas model.</p> <p>Transformasi Data: Melakukan <i>Feature Engineering</i> pada kolom target (Result Description). Kita mengubah teks menjadi label biner:</p> <ul style="list-style-type: none"> - Jika status = "PASS" \rightarrow diubah menjadi 0. - Jika status = "FAIL" atau "PASS WITH GRACE" \rightarrow diubah menjadi 1. <p>Hal ini agar algoritma <i>Machine Learning</i> dapat memproses klasifikasi risiko.</p>
4. Modeling (Pemodelan)	Pilih algoritma dan jelaskan alasannya.	<p>Algoritma: Random Forest Classifier.</p> <p>Alasan Pemilihan:</p> <ol style="list-style-type: none"> 1. Akurasi Tinggi: Random Forest bekerja dengan menggabungkan banyak <i>Decision Tree</i>, sehingga hasil prediksinya lebih akurat dan stabil dibandingkan pohon keputusan tunggal. 2. Non-Linear: 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.
5. Evaluation (Evaluasi)	Pilih metode evaluasi.	Metode: Confusion Matrix dan Akurasi.

Tahapan CRISP-DM	Instruksi (Dari LKM)	Rancangan Implementasi
		<p>Fokus Evaluasi:</p> <p>Selain melihat akurasi total, kami memantau Recall untuk kelas "Berisiko". Tujuannya adalah memastikan sistem sedikit mungkin melewatkannya 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>Rancangan: Dashboard Monitoring Akademik untuk Dosen Wali (Kaprodi).</p> <p>Tampilan Interface:</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"> - Nama Mahasiswa