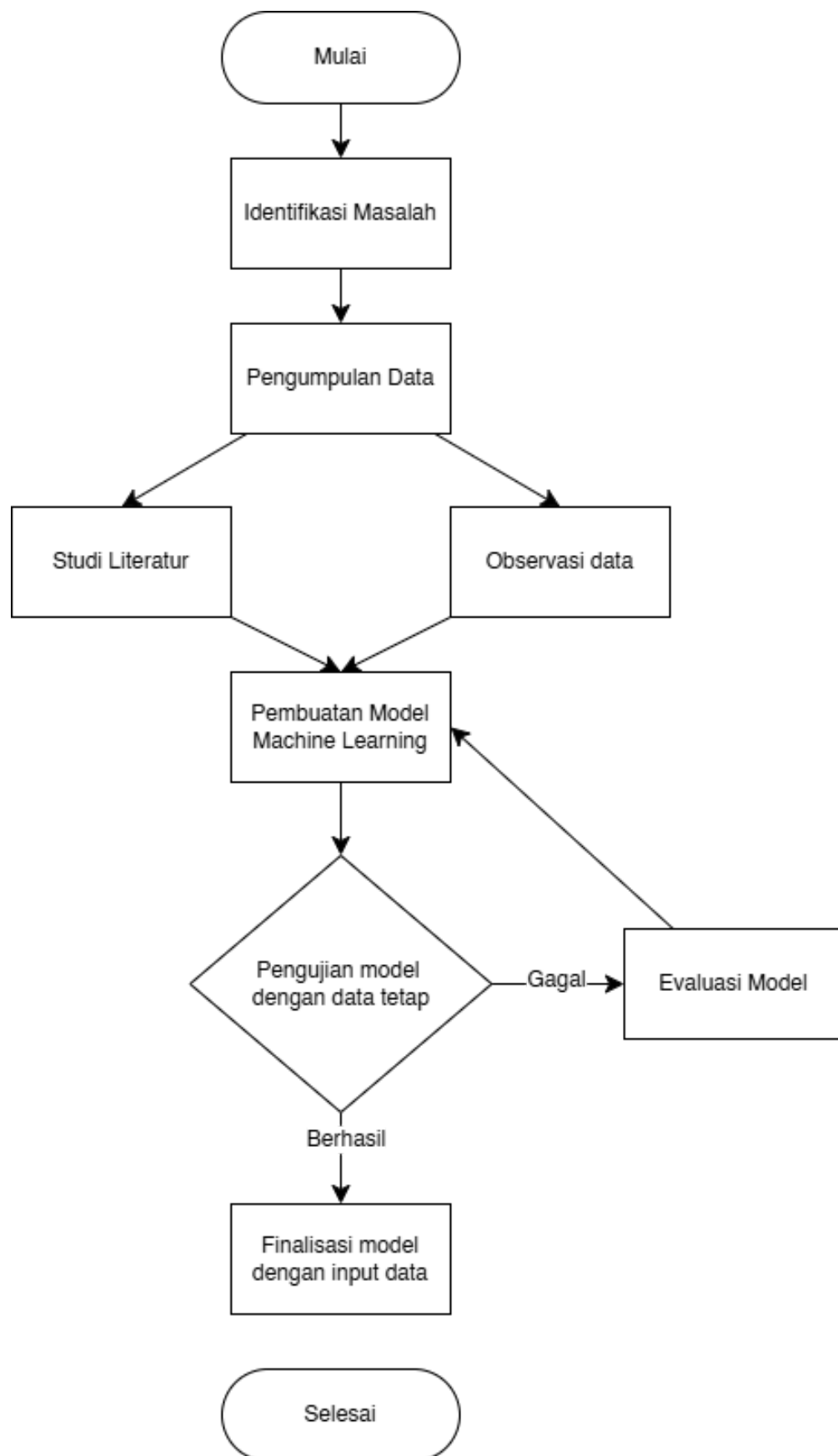


TAHAPAN PERANCANGAN



1. Identifikasi Masalah

Tahap awal ini bertujuan untuk mengenali permasalahan utama yang menjadi fokus penelitian, yaitu meningkatnya risiko *burnout* di kalangan mahasiswa. Pada tahap ini dilakukan analisis terhadap latar belakang fenomena, faktor penyebab, dan urgensi permasalahan di dunia pendidikan tinggi. Hasil dari tahap ini berupa rumusan masalah yang akan dijawab melalui pendekatan *data science* dan *machine learning*.

2. Pengumpulan Data

Setelah masalah teridentifikasi, dilakukan proses pengumpulan data yang relevan. Data diperoleh dari berbagai sumber daring yang berisi informasi mahasiswa seperti:

- jenis kelamin, umur, tahun kuliah,
- status kerja, status kesehatan, jam belajar per minggu, status hubungan
- serta hasil survei burnout (MBI-EX, MBI-CY, MBI-EA).

Data ini menjadi dasar untuk membangun model prediksi.

3. Studi Literatur dan Observasi Data

Tahapan ini dilakukan secara paralel:

- **Studi Literatur:** mengkaji teori dan hasil penelitian sebelumnya terkait *burnout*, faktor-faktor yang mempengaruhi, serta metode *machine learning* yang relevan.
- **Observasi Data:** melakukan eksplorasi awal terhadap dataset untuk memahami struktur, karakteristik variabel, serta mendeteksi data kosong (*missing values*) atau anomali.

Kedua tahap ini memberikan pemahaman teoretis dan empiris sebelum melangkah ke tahap pemodelan.

4. Pembuatan Model Machine Learning

Tahap ini adalah inti dari penelitian. Model prediksi dibangun menggunakan algoritma **CatBoostRegressor**, yang mampu menangani data numerik dan kategorikal tanpa perlu proses *encoding* manual.

Library Python yang digunakan meliputi:

- pandas dan numpy untuk pengolahan data,
- sklearn untuk pembagian data dan evaluasi,
- catboost untuk pemodelan,
- joblib untuk penyimpanan model.

Proses ini juga mencakup tahap *preprocessing* (normalisasi data dengan StandardScaler) dan pembagian dataset menjadi data pelatihan dan data uji (*train_test_split*).

5. Pengujian Model dengan Data Tetap

Setelah model dibuat, dilakukan pengujian menggunakan data uji yang telah dipisahkan sebelumnya. Tujuannya untuk menilai kemampuan generalisasi model

terhadap data baru. Apabila hasil pengujian menunjukkan performa yang rendah, model akan diperbaiki kembali melalui tahap tuning parameter dan pelatihan ulang.

6. Evaluasi Model

Jika hasil pengujian belum sesuai harapan, dilakukan evaluasi mendalam menggunakan metrik R^2 Score dari `sklearn.metrics`. Evaluasi ini digunakan untuk mengetahui sejauh mana model mampu menjelaskan variabilitas data burnout mahasiswa. Model akan diperbaiki hingga diperoleh performa terbaik sesuai target yang diharapkan.

7. Finalisasi Model dengan Input Data

Setelah model dinyatakan berhasil dan akurat, dilakukan tahap finalisasi dengan cara menyimpan model menggunakan `joblib`. Model yang telah disimpan kemudian dapat digunakan untuk melakukan prediksi burnout pada data mahasiswa baru atau digunakan sebagai bagian dari sistem pendeteksi dini tingkat *burnout*.