

## Penentuan Kelayakan Mahasiswa untuk Mendapat Beasiswa Menggunakan Algoritma Naive Bayes

### 1. Latar Belakang

Beasiswa adalah salah satu bentuk dukungan pendidikan yang diberikan kepada mahasiswa untuk meringankan beban finansial dan mendorong prestasi akademik maupun non-akademik. Dalam menentukan kelayakan mahasiswa untuk menerima beasiswa, pihak penyelenggara sering kali menggunakan kriteria tertentu, seperti IPK, penghasilan orang tua, dan prestasi. Namun, proses ini dapat menjadi rumit dan membutuhkan waktu jika dilakukan secara manual, terutama jika jumlah pelamar beasiswa sangat banyak.

Untuk mengatasi masalah ini, algoritma Naive Bayes dapat digunakan sebagai solusi dalam melakukan klasifikasi secara otomatis. Naive Bayes cocok untuk menyelesaikan masalah klasifikasi seperti ini karena algoritmanya sederhana, cepat, dan efektif dalam menangani data yang memiliki banyak fitur.

### 2. Dataset

Dataset dalam studi kasus ini berfokus pada penentuan kelayakan mahasiswa untuk mendapatkan beasiswa berdasarkan beberapa faktor:

- IPK (Indeks Prestasi Kumulatif): Representasi kemampuan akademik mahasiswa, dikategorikan menjadi Tinggi, Sedang, dan Rendah.
- Penghasilan Orang Tua: Nominal pendapatan keluarga, dikategorikan menjadi Rendah, Menengah, dan Tinggi.
- Keaktifan Organisasi: Partisipasi mahasiswa dalam organisasi atau kegiatan kemahasiswaan, terdiri dari Sangat Aktif, Cukup Aktif, dan Tidak Aktif.
- Tanggungan Orang Tua: Jumlah anggota keluarga yang menjadi tanggungan, dikategorikan menjadi Sedikit, Sedang, dan Banyak.
- Target (Label): Apakah mahasiswa layak (Layak) atau tidak layak (Tidak Layak) untuk mendapatkan beasiswa.

IPK	Penghasilan	Organisasi	Tanggungan	Target
Tinggi	Rendah	Sangat Aktif	Sedikit	Layak
Sedang	Menengah	Cukup Aktif	Sedang	Layak
Rendah	Tinggi	Tidak Aktif	Banyak	Tidak Layak
Tinggi	Menengah	Sangat Aktif	Sedang	Layak
Rendah	Rendah	Tidak Aktif	Banyak	Tidak Layak

## Perhitungan Manual Algoritma Naive Bayes

Studi Kasus: Seorang mahasiswa memiliki atribut:

- IPK: 3.2 (kategori: Sedang)
- Penghasilan Orang Tua: Rp4.000.000 (kategori: Menengah)
- Keaktifan Organisasi: Cukup Aktif
- Tanggungan Orang Tua: Sedang

Langkah 1: Hitung Probabilitas Prior Probabilitas prior dihitung berdasarkan jumlah data yang memiliki label Layak dan Tidak Layak:

$$P(\text{Layak}) = \frac{\text{Jumlah Layak}}{\text{Total Data}} = \frac{3}{5} = 0.6$$

$$P(\text{Tidak Layak}) = \frac{\text{Jumlah Tidak Layak}}{\text{Total data}} = \frac{2}{5} = 0.4$$

Langkah 2: Hitung Probabilitas Likelihood

Kita menghitung probabilitas dari setiap atribut terhadap label tertentu:

- Untuk Target Layak:

$$P(\text{IPK} = \text{Sedang} | \text{Layak}) = \frac{\text{Jumlah IPK Sedang Pada Layak}}{\text{Jumlah Layak}} = \frac{1}{3} = 0.33$$

$$P(\text{Penghasilan} = \text{Menengah} | \text{Layak}) = \frac{\text{Jumlah Penghasilan Menengah Pada Layak}}{\text{Jumlah Layak}} = \frac{1}{3} = 0.33$$

$$P(\text{Organisasi} = \text{Cukup Aktif} | \text{Layak}) = \frac{\text{Jumlah cukup Aktif Pada Layak}}{\text{Jumlah Layak}} = \frac{1}{3} = 0.33$$

$$P(\text{Tanggungan} = \text{Sedang} | \text{Layak}) = \frac{\text{Jumlah Sedang Pada Layak}}{\text{Jumlah Layak}} = \frac{1}{3} = 0.33$$

- Untuk Target Tidak Layak:

$$P(\text{IPK} = \text{Sedang} | \text{Tidak Layak}) = \frac{\text{Jumlah IPK Sedang Pada Tidak Layak}}{\text{Jumlah Tidak Layak}} = \frac{0}{2} = 0.000001 \text{ (Laplace Smoothing)}$$

$$P(\text{Penghasilan} = \text{Menengah} | \text{Tidak Layak}) = \frac{\text{Jumlah Penghasilan Menengah Pada Tidak Layak}}{\text{Jumlah Tidak Layak}} = \frac{0}{2} = 0.000001$$

$$P(\text{Organisasi} = \text{Cukup Aktif} | \text{Tidak Layak}) = \frac{\text{Jumlah Organisasi Cukup Aktif Pada Tidak Layak}}{\text{Jumlah Tidak Layak}} = \frac{0}{2} = 0.000001$$

$$P(\text{Tanggungan} = \text{Sedang} | \text{Layak}) = \frac{\text{Jumlah Tanggungan Sedang Pada Layak}}{\text{Jumlah Tidak Layak}} = \frac{1}{2} = 0.5$$

Langkah 3: Hitung Probabilitas Posterior

Gunakan rumus Naive Bayes:

$$P(\text{Target} | \text{Fitur}) = P(\text{Target}) \cdot P(\text{Fitur1} | \text{Target}) \cdot P(\text{Fitur2} | \text{Target}) \cdot \dots$$

- Untuk Target Layak

$$P(\text{Layak} | \text{Fitur}) = 0.6 \cdot 0.33 \cdot 0.33 \cdot 0.33 \cdot 0.33 = 0.0071$$

- Untuk Target Tidak Layak

$$P(\text{Tidak Layak} | \text{Fitur}) = 0.4 \cdot 0.000001 \cdot 0.000001 \cdot 0.000001 \cdot 0.5 = 0.00000000000002$$

Langkah 4: Prediksi Akhir Prediksi akhir adalah kelas dengan probabilitas tertinggi:

Prediksi = Layak