

Nama : Bismar Wasykuru
Nim : 2210018
Prodi : Informatika
Mata Kuliah : Kecerdasan Buatan

Implementasi Sederhana dari Naive Bayes Classifier : Program Prediksi Kinerja Mahasiswa

Dataset yang digunakan dalam program ini adalah dataset yang berisi informasi mengenai beberapa faktor yang dapat mempengaruhi kinerja mahasiswa, yaitu jam belajar, kehadiran, dan partisipasi. Berdasarkan informasi ini, program akan memprediksi apakah kinerja mahasiswa tersebut tergolong baik atau buruk.

Berikut adalah **struktur dataset** yang digunakan:

Jam Belajar	Kehadiran	Partisipasi	Kinerja
rendah	kurang	tidak aktif	buruk
sedang	baik	aktif	Baik
tinggi	baik	aktif	baik
rendah	kurang	tidak aktif	buruk
sedang	kurang	aktif	buruk

Penjelasan Atribut dalam Dataset (Sederhana)

- Jam Belajar**
Menunjukkan seberapa banyak waktu mahasiswa belajar. : Rendah/ Sedang/ Tinggi
- Kehadiran**
Menunjukkan seberapa sering mahasiswa hadir di kelas. : Kurang/ Baik
- Partisipasi**
Menunjukkan seberapa aktif mahasiswa dalam kelas. : Aktif/ Tidak Aktif
- Kinerja**
Menunjukkan hasil akhir mahasiswa berdasarkan jam belajar, kehadiran, dan partisipasi. : Baik/ Buruk

Penjelasan Dataset (Sederhana)

Dataset ini berisi informasi tentang tiga faktor yang mempengaruhi kinerja mahasiswa:

- Jam Belajar
- Kehadiran
- Partisipasi

Setiap baris data menunjukkan kombinasi dari ketiga faktor tersebut dan kinerja mahasiswa (baik atau buruk).

Contohnya:

- Mahasiswa yang belajar rendah, kurang hadir, dan tidak aktif memiliki kinerja buruk.
- Mahasiswa yang belajar tinggi, selalu hadir, dan aktif memiliki kinerja baik.

Dataset ini digunakan untuk memprediksi apakah seorang mahasiswa akan memiliki kinerja baik atau buruk berdasarkan faktor-faktor tersebut.

Menghitung Prior Probabilitas

Dari data di atas ada 5 data: 2 data dengan kinerja baik dan 3 data dengan kinerja buruk

Rumus Prior Probabilitas:

$$P(\text{baik}) = \frac{\text{Jumlah data dengan kinerja baik}}{\text{Jumlah Total Data}} = \frac{2}{5} = 0,4$$

$$P(\text{buruk}) = \frac{\text{Jumlah data dengan kinerja buruk}}{\text{Jumlah Total Data}} = \frac{3}{5} = 0,6$$

Jadi, prior probabilitasnya adalah:

- $P(\text{baik})=0.4$
- $P(\text{buruk})=0.6$

Menghitung Probabilitas Likelihood untuk Kinerja Baik:

1. Keterangan : Di data dengan kinerja baik (baris 2 dan 3), ada **1 data** yang memiliki jam belajar "sedang" (baris 2).

$$\text{Jam belajar : } P(\text{jam_belajar} = \text{sedang} \mid \text{baik}) = \frac{1}{2} = 0.5$$

2. Keterangan : Di data dengan kinerja baik (baris 2 dan 3), ada **2 data** yang memiliki kehadiran "baik" (baris 2 dan 3).

$$\text{Kehadiran : } P(\text{kehadiran} = \text{baik} \mid \text{baik}) = \frac{2}{2} = 1.0$$

3. Keterangan : Di data dengan kinerja baik (baris 2 dan 3), ada **2 data** yang memiliki partisipasi "aktif" (baris 2 dan 3).

$$\text{Partisipasi : } P(\text{kehadiran} = \text{baik} \mid \text{baik}) = \frac{2}{2} = 1.0$$

Jadi Perhitungan Probabilitas untuk Kinerja Baik:

$$P(\text{baik} \mid \text{input}) = P(\text{baik}) \times P(\text{jam_belajar} = \text{sedang} \mid \text{baik}) \times P(\text{kehadiran} = \text{baik} \mid \text{baik}) \times P(\text{partisipasi} = \text{tidak aktif} \mid \text{baik})$$

Menghitung Probabilitas Likelihood untuk Kinerja Buruk:

1. Keterangan : Di data dengan kinerja buruk (baris 1, 4, dan 5), ada **1 data** yang memiliki jam belajar "sedang" (baris 5).

$$\text{Jam belajar : } P(\text{jam_belajar} = \text{sedang} \mid \text{buruk}) = \frac{1}{3} \approx 0.333$$

2. Keterangan : Di data dengan kinerja buruk (baris 1, 4, dan 5), tidak ada data yang memiliki kehadiran "baik". Semua data memiliki kehadiran "kurang".

$$\text{Kehadiran : } P(\text{kehadiran} = \text{baik} \mid \text{buruk}) = \frac{0}{3} = 0$$

3. Keterangan : Di data dengan kinerja buruk (baris 1, 4, dan 5), ada **1 data** yang memiliki partisipasi "aktif" (baris 5).

$$\text{Partisipasi : } P(\text{Partisipasi} = \text{aktif} \mid \text{buruk}) = \frac{1}{3} \approx 0.333$$

Jadi Perhitungan Probabilitas untuk Kinerja Buruk:

$$P(\text{buruk} \mid \text{input}) = P(\text{buruk}) \times P(\text{jam_belajar} = \text{sedang} \mid \text{buruk}) \times \\ P(\text{kehadiran} = \text{baik} \mid \text{buruk}) \times P(\text{partisipasi} = \text{tidak aktif} \mid \text{buruk})$$
