

Nama : Endang Prayoga Hidayatulloh
NIM : 2006189
Mata Kuliah : Sistem Pakar

1. apa saja kelebihan dan kekurangan Naive bayesian klasifikasi

Jawab.

Naive Bayesian classification adalah metode klasifikasi yang menggunakan teori probabilitas untuk memprediksi kelas dari suatu data. Berikut adalah kelebihan dan kekurangan dari metode klasifikasi ini:

Kelebihan:

- Efektif dalam data yang besar: Naive Bayesian classification dapat mengatasi data yang besar dengan cepat dan efisien. Karena model ini hanya memerlukan perhitungan sederhana dan tidak memerlukan iterasi yang rumit, maka waktu komputasi yang dibutuhkan relatif singkat.
- Sederhana dan mudah dipahami: Metode ini relatif mudah dipahami oleh pengguna yang memiliki pengetahuan dasar tentang teori probabilitas. Selain itu, metode ini juga mudah diimplementasikan dalam bahasa pemrograman seperti Python dan R.
- Tidak memerlukan banyak data training: Model ini dapat menghasilkan hasil yang cukup akurat bahkan dengan sedikit data training.
- Baik untuk dataset kategorikal: Metode ini sangat efektif dalam menangani dataset kategorikal, terutama dalam klasifikasi teks atau pengolahan bahasa alami.

Kekurangan:

- Mengabaikan hubungan antara fitur: Model ini mengasumsikan bahwa fitur dalam data adalah independen satu sama lain, sehingga dapat menghasilkan akurasi yang rendah ketika hubungan antar fitur diabaikan.
- Sensitif terhadap data yang tidak seimbang: Jika jumlah data untuk setiap kelas sangat tidak seimbang, model ini dapat menghasilkan bias terhadap kelas mayoritas.
- Tidak cocok untuk data kontinu: Model ini kurang efektif dalam menangani data kontinu, seperti data sensor atau pengukuran yang diambil secara terus-menerus.
- Mengalami kesulitan dalam memodelkan masalah yang kompleks: Model ini dapat mengalami kesulitan dalam memodelkan masalah yang kompleks, terutama ketika terdapat banyak fitur dalam data atau ketika terdapat banyak kelas yang mungkin.

2. Di Coba aja

berikut merupakan pendekatan bayesian untuk estimasi probabilitas yang ada

$$P(a_i | v_j) = \frac{n_c + mp}{n + m} \quad (2.1)$$

Keterangan :

- n : Jumlah data traning dimana $v = v_j$.
- n_c : Jumlah data traning dimana $v = v_j$ dan $a = a_i$.
- P : *Prior* estimate untuk $P(a_i | v_j)$.
- m : Ukuran sampel ekuivalen.

atau

Keterangan :

$P(v_j)$: Probabilitas hipotesis v_j (*prior*).

N : Jumlah data traning dimana $v = v_j$.

Jumlah : Jumlah data training.

Terdapat data seperti di bawah ini

Warna	Tipe	Asal	Tercuri?
a1	a2	a4	vj
merah	Sport	domestik	ya
merah	Sport	domestik	tidak
merah	Sport	domestik	ya
kuning	Sport	domestik	tidak
kuning	Sport	import	ya
kuning	SUV	import	tidak
kuning	SUV	import	ya
kuning	SUV	domestik	tidak
merah	SUV	import	tidak
merah	Sport	import	ya

Dari tabel di atas, data mobil yang tercuri bisa dilihat dari atribut warna, tipe, dan asal. Misalkan kita ingin mengelompokkan *mobil warna merah, tipe SUV, dan asal domestik*. tentukan probabilitas tercuri dan probabilitas tidak tercuri, dan kemudian tentukan berapa persen mobil yang tercuri dan berapa persen

mobil yang tidak tercuri, serta tentukan *mobil dengan warna merah, tipe SUV, dan asal domestik* tersebut tercuri atau tidak?

Soal nya :

1. hitung Mobil yang dapat tercuri YA & TIDAK ($P(v_j)$) (probabilitas)
- 2 hitung probabilitas *ya tercuri* pada mobil warna merah, tipe SUV, dan asal domestik (probabilitas)
3. hitung probabilitas *tidak tercuri* pada mobil warna merah, tipe SUV, dan asal domestik
4. tentukan berapa persen mobil tercuri dan tidak tercuri

Penyelesaian

1. Untuk menghitung probabilitas tercuri menggunakan rumus:
 - a. $P(YA) = 5/10 = 0.5$
 - b. $P(TIDAK) = 5/10 = 0.5$
2. Menghitung probabilitas ya tercuri pada mobil warna merah, tipe SUV, dan asal domestik.
 - a. $P(\text{merah}|ya) = 3/5 = 0.6$
 - b. $P(SUV|ya) = 1/5 = 0.2$
 - c. $P(\text{domestik}|ya) = 2/5 = 0.4$
3. Menghitung probabilitas tidak tercuri pada mobil warna merah, tipe SUV, dan asal domestik.
 - a. $P(\text{merah}|tidak) = 2/5 = 0.4$
 - b. $P(SUV|tidak) = 3/5 = 0.6$
 - c. $P(\text{domestik}|tidak) = 3/5 = 0.6$
4. Menentukan berapa persen mobil tercuri dan tidak tercuri.
 - a.
$$\begin{aligned} \text{tercuri (ya)} &= P(ya) * P(\text{merah}|ya) * P(SUV|ya) * P(\text{domestik}|ya) \\ &= 0.5 * 0.6 * 0.2 * 0.4 \\ &= 0.024 \text{ atau } 2,4 \% \end{aligned}$$
 - b.
$$\begin{aligned} \text{tercuri (tidak)} &= P(tidak) * P(\text{merah}|tidak) * P(SUV|tidak) * P(\text{domestik}|tidak) \\ &= 0.5 * 0.4 * 0.6 * 0.6 \\ &= 0.072 \text{ atau } 7,2\% \end{aligned}$$