



Probabilitas Bersyarat

Achmad Basuki

Politeknik Elektronika Negeri Surabaya

2004



Definisi

- Probabilitas bersyarat dituliskan dengan $p(A|B)$ yang menyatakan probabilitas A bila diketahui B, dimana A dan B menyatakan kejadian acak.
- Probabilitas bersyarat dapat dihitung menggunakan:

$$p(A | B) = \frac{p(A, B)}{p(A)}$$

Dimana: $p(A,B)$ adalah probabilitas A dan B
 $p(A)$ adalah probabilitas A



Contoh 1⁽¹⁾

Dari hasil survey terhadap 100 orang mengenai hobby olah raga sepak bola dan bola volley. Dari 100 responden, 20 orang menyatakan menyukai keduanya, 50 orang menyatakan hanya menyukai sepak bola dan 20 hanya menyukai bola volley dan 10 orang tidak menyukai keduanya.

Bila data tersebut dibuat tabel distribusi, maka diperoleh:

Sepak bola Bola Volley	Ya	Tidak
Ya	20	20
Tidak	50	10

Contoh 1(2)

Sepak bola Bola Volley	Ya	Tidak
Ya	20	20
Tidak	50	10



Sepak bola Bola Volley	Ya	Tidak
Ya	0.2	0.2
Tidak	0.5	0.1

Bila A menyatakan menyukai sepak bola
B menyatakan menyukai bola volley

Sepak bola Bola Volley	Ya	Tidak	p(B)
Ya	0.2	0.2	0.4
Tidak	0.5	0.1	0.6
p(A)	0.7	0.3	1

$$A = \{\text{suka, tidak}\}$$

$$p(A) = \{ 0.7 \quad 0.3 \}$$

$$B = \{\text{suka, tidak}\}$$

$$p(B) = \{ 0.4 \quad 0.6 \}$$

p(A) dan p(B) dinamakan dengan probabilitas marginal



Contoh 1⁽³⁾

Sepak bola Bola Volley	Ya	Tidak	p(B)
Ya	0.2	0.2	0.4
Tidak	0.5	0.1	0.6
p(A)	0.7	0.3	1

$p(A,B)$ menyatakan probabilitas bersama A dan B, dan dibaca probabilitas A dan B.

Kejadian $(A,B) = \{(ya,ya), (ya,tidak), (tidak,ya), (tidak,tidak)\}$

$$p(A=ya,B=ya) = 0.2$$

$$p(A=ya,B=tidak) = 0.5$$

$$p(A=tidak,B=ya) = 0.2$$

$$p(A=tidak,B=tidak) = 0.1$$



Contoh 1⁽⁴⁾

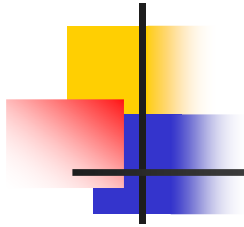
Sepak bola Bola Volley	Ya	Tidak	p(B)
Ya	0.2	0.2	0.4
Tidak	0.5	0.1	0.6
p(A)	0.7	0.3	1

Probabilitas seseorang menyukai sepak bola bila diketahui dia menyukai bola volley adalah:

$$p(A = ya | B = ya) = \frac{p(A = ya, B = ya)}{p(B = ya)} = \frac{0.2}{0.4} = 0.5$$

Probabilitas seseorang tidak akan menyukai bola volley bila diketahui dia menyukai sepak bola adalah:

$$p(B = tidak | A = ya) = \frac{p(A = ya, B = tidak)}{p(A = ya)} = \frac{0.5}{0.7} = 0.714$$



Sepak bola Bola Volley	Ya	Tidak	p(B)
Ya	0.2	0.2	0.4
Tidak	0.5	0.1	0.6
p(A)	0.7	0.3	1

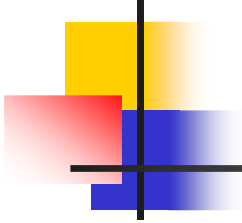
$$p(A = ya | B = ya) = \frac{p(A = ya, B = ya)}{p(B = ya)} = \frac{0.2}{0.4} = 0.5$$

$$p(B = ya | A = ya) = \frac{p(A = ya, B = ya)}{p(A = ya)} = \frac{0.2}{0.7} \approx 0.28$$

Bayes Theory

$$P(A|B)P(B)=P(B|A)P(A)$$

$$0.5*0.4=(0.2/0.7)*0.7$$
$$0.2 = 0.2$$



Sepak bola Bola Volley	Ya	Tidak	p(B)
Ya	0.2	0.2	0.4
Tidak	0.5	0.1	0.6
p(A)	0.7	0.3	1

$$P(A) = 0.7$$

$$P(\bar{A}) = 0.3$$

$$P(+|A) = 0.2/0.7$$

$$P(-|A) = 0.5/0.7$$

$$P(+|\bar{A}) = 0.2/0.3$$

$$P(-|\bar{A}) = 0.1/0.3$$

$$P(+|A)P(A) = (0.2/0.7) * 0.7 = 0.2$$

$$P(+|\bar{A})P(\bar{A}) = (0.2/0.3) * 0.3 = 0.2$$



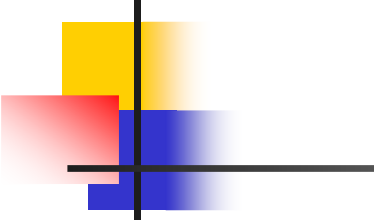
Contoh 2⁽¹⁾

Survey dilakukan untuk mengetahui hubungan antara merokok dan potensi menderita penyakit paru-paru. Survey ini dilakukan pada 20 orang, yang hasilnya sebagai berikut:

Ubah tabel di atas menjadi tabel distribusi bersama sebagai berikut:

paru-paru	merokok	
	ya	tidak
ya	6	2
tidak	5	7

no.	merokok	paru-paru
1	ya	tidak
2	ya	ya
3	tidak	tidak
4	tidak	tidak
5	ya	tidak
6	ya	ya
7	ya	tidak
8	tidak	tidak
9	ya	ya
10	tidak	ya
11	tidak	tidak
12	tidak	tidak
13	ya	tidak
14	ya	ya
15	tidak	tidak
16	ya	ya
17	tidak	ya
18	tidak	tidak
19	ya	tidak
20	ya	ya



paru-paru	merokok		p(B)
	ya	tidak	
ya	0.3	0.1	0.4
tidak	0.25	0.35	0.6
p(A)	0.55	0.45	1

$$P(A) = 0.55$$

$$P(\bar{A}) = 0.45$$

$$P(+|A) = 0.3/0.55$$

$$P(-|A) = 0.25/0.55$$

$$P(+|\bar{A}) = 0.1/0.45$$

$$P(-|\bar{A}) = 0.35/0.45$$

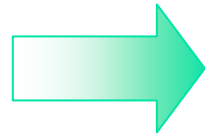
$$P(+|A)P(A) = (0.3/0.55) * 0.55 = 0.3$$

$$P(+|\bar{A})P(\bar{A}) = (0.1/0.45) * 0.45 = 0.1$$

$$P(A|+) = 0.3 / (0.3 + 0.1) = 0.75$$

Contoh 2⁽²⁾

paru-paru	merokok	
	ya	tidak
ya	6	2
tidak	5	7



paru-paru	merokok		p(B)
	ya	tidak	
ya	0.3	0.1	0.4
tidak	0.25	0.35	0.6
p(A)	0.55	0.45	1

$P(A)$ = Probabilitas seseorang menderita penyakit paru-paru
 = {0.4 0.6}

$P(B)$ = Probabilitas seseorang merokok
 = {0.55 0.45}

Probabilitas seseorang menderita penyakit paru-paru bila diketahui dia merokok adalah:

$$p(A = ya | B = ya) = \frac{P(A = ya, B = ya)}{p(B = ya)} = \frac{0.3}{0.55} = 0.545$$



Contoh 3(1)

Masalah multivariate
pada probabilitas bersyarat

Dari hasil survey terhadap 20 orang mengenai pengaruh usia, berat badan terhadap kemungkinan terkena hipertensi terlihat pada tabel berikut:

A = usia = {muda,tua}

B = berat badan = {over,rata-rata,kurang}

C = Hipertensi = {ya, tidak}

no.	usia	berat badan	hipertensi
1	muda	over	tidak
2	muda	rata-rata	tidak
3	tua	kurang	ya
4	muda	kurang	tidak
5	tua	rata-rata	ya
6	tua	kurang	tidak
7	tua	over	ya
8	muda	rata-rata	tidak
9	tua	over	ya
10	muda	kurang	ya
11	muda	rata-rata	tidak
12	muda	rata-rata	tidak
13	tua	over	ya
14	muda	over	tidak
15	tua	kurang	tidak
16	tua	rata-rata	tidak
17	muda	rata-rata	tidak
18	muda	rata-rata	ya
19	muda	over	ya
20	tua	kurang	tidak

Contoh 3(2)

Masalah multivariate
pada probabilitas bersyarat

A = usia = {muda,tua}

B = berat badan = {over,rata-rata,kurang}

C = Hipertensi = {ya, tidak}

Usia dan berat badan	Hipertensi		p(A,B)
	ya	tidak	
(muda,over)	1	2	3
(muda,rata-rata)	0	5	5
(muda,kurang)	1	1	2
(tua,over)	3	0	3
(tua,rata-rata)	1	1	2
(tua,kurang)	1	4	5
p(C)	7	13	20

no.	usia	berat badan	hipertensi
1	muda	over	tidak
2	muda	rata-rata	tidak
3	tua	kurang	ya
4	muda	kurang	tidak
5	tua	rata-rata	ya
6	tua	kurang	tidak
7	tua	over	ya
8	muda	rata-rata	tidak
9	tua	over	ya
10	muda	kurang	ya
11	muda	rata-rata	tidak
12	muda	rata-rata	tidak
13	tua	over	ya
14	muda	over	tidak
15	tua	kurang	tidak
16	tua	rata-rata	tidak
17	muda	rata-rata	tidak
18	muda	rata-rata	ya
19	muda	over	ya
20	tua	kurang	tidak

Contoh 3(3)

Masalah multivariate
pada probabilitas bersyarat

A = usia = {muda,tua}

B = berat badan = {over,rata-rata,kurang}

C = Hipertensi = {ya, tidak}

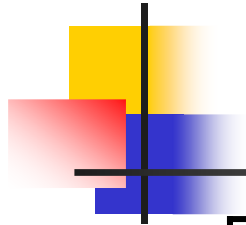
Fungsi distribusi bersama

Usia dan berat badan	Hipertensi		p(A,B)
	ya	tidak	
(muda,over)	1	2	3
(muda,rata-rata)	0	5	5
(muda,kurang)	1	1	2
(tua,over)	3	0	3
(tua,rata-rata)	1	1	2
(tua,kurang)	1	4	5
p(C)	7	13	20

Usia dan berat badan	Hipertensi		p(A,B)
	ya	tidak	
(muda,over)	0.05	0.1	0.15
(muda,rata-rata)	0	0.25	0.25
(muda,kurang)	0.05	0.05	0.1
(tua,over)	0.15	0	0.15
(tua,rata-rata)	0.05	0.05	0.1
(tua,kurang)	0.05	0.2	0.25
p(C)	0.35	0.65	1

Probabilitas seseorang akan menderita hipertensi bila diketahui dia tua dan over adalah:

$$p(C = ya | A = tua, B = over) = \frac{p(A = tua, B = over, C = ya)}{p(A = tua, B = over)} = \frac{0.15}{0.15} = 1$$



Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	No
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Naïve Bayes Classifier

Case: P.59 Table 3.2

$$P(\text{PlayTennis}=\text{yes}) = 9/14$$

$$P(\text{PlayTennis}=\text{no}) = 5/14$$

If (Outlook=sunny) ^ (Temperature=mild) ^
(Humidity=high) ^ (Wind=strong) → PlayTennis or not

Answer:

$$P(\text{yes}) P(\text{sunny}|\text{yes}) P(\text{mild}|\text{yes}) P(\text{high}|\text{yes}) P(\text{strong}|\text{yes}) \\ 9/14 \times 2/9 \times 4/9 \times 3/9 \times 3/9 = (2 \times 4 \times 3 \times 3) \times 1 / (14 \times 9^3)$$

$$P(\text{no}) P(\text{sunny}|\text{no}) P(\text{mild}|\text{no}) P(\text{high}|\text{no}) P(\text{strong}|\text{no}) \\ 5/14 \times 3/5 \times 2/5 \times 3/5 \times 4/5 = (2 \times 4 \times 3 \times 3) \times 1 / (14 \times 5^3)$$

→ Not play tennis



Learner

If a new data is:

$(\text{Outlook}=\text{sunny}) \wedge (\text{Temperature}=\text{mild}) \wedge$
 $(\text{Humidity}=\text{high}) \wedge (\text{Wind}=\text{strong}) \rightarrow \text{PlayTennis}$

The result (using Naïve Bayes Classifier) \rightarrow no

When the new data is added to the table
the results will change to "yes"