

training data

id	veg?	type?	busy?	price?	class
1	yes	italian	yes	low	pos
2	no	french	no	med	neg
3	yes	greek	yes	high	neg
4	no	french	yes	low	neg
5	no	french	yes	high	neg
6	yes	italian	yes	med	neg
7	no	french	no	high	pos
8	yes	greek	yes	low	pos

new data

id	veg?	type?	busy?	price?	class
A	yes	italian	no	low	?
B	no	greek	yes	high	?
C	no	greek	no	low	?
D	yes	greek	no	med	?
E	no	italian	yes	med	?
F	yes	greek	no	high	?

Bayes' rule for classification

$$P(C|A_1, \dots, A_n) = \frac{P(C)P(A_1|C) \cdots P(A_n|C)}{P(A_1) \cdots P(A_n)}$$

here:

$$P(C|V, T, B, P) = \frac{P(C)P(V|C) \cdot P(T|C) \cdot P(B|C) \cdot P(P|C)}{\cancel{P(V) \cdot P(T) \cdot P(B) \cdot P(P)}}$$

constant for record, ignore

we abbreviate attributes and values by initials

training data $P(C|V, T, B, P) = \frac{P(C)P(V|C) \cdot P(T|C) \cdot P(B|C) \cdot P(P|C)}{\cancel{P(V) \cdot P(T) \cdot P(B) \cdot P(P)}}$

constant for record, ignore

estimate parts of right-hand side by counting:

8 records in total, 3 pos, 5 neg

$$P(C = pos) = \frac{3}{8}$$

$$P(C = neg) = \frac{5}{8}$$

id	veg?	type?	busy?	price?	class
1	yes	italian	yes	low	pos
2	no	french	no	med	neg
3	yes	greek	yes	high	neg
4	no	french	yes	low	neg
5	no	french	yes	high	neg
6	yes	italian	yes	med	neg
7	no	french	no	high	pos
8	yes	greek	yes	low	pos

training data $P(C|V, T, B, P) = \frac{P(C)P(V|C) \cdot P(T|C) \cdot P(B|C) \cdot P(P|C)}{\cancel{P(V) \cdot P(T) \cdot P(B) \cdot P(P)}}$

constant for record, ignore

estimate parts of right-hand
side by counting:

use pos. examples for probabilities
conditioned on “pos”

id	veg?	type?	busy?	price?	class
1	yes	italian	yes	low	pos
2					
3					
4					
5					
6					
7	no	french	no	high	pos
8	yes	greek	yes	low	pos

$$P(V = y|C = pos) = \frac{2}{3}$$

$$P(T = i|C = pos) = \frac{1}{3}$$

$$P(P = l|C = pos) = \frac{2}{3}$$

$$P(V = n|C = pos) = \frac{1}{3}$$

$$P(T = f|C = pos) = \frac{1}{3}$$

$$P(P = m|C = pos) = \frac{0}{3}$$

$$P(T = g|C = pos) = \frac{1}{3}$$

$$P(P = h|C = pos) = \frac{1}{3}$$

$$P(B = y|C = pos) = \frac{2}{3}$$

$$P(B = n|C = pos) = \frac{1}{3}$$

training data $P(C|V, T, B, P) = \frac{P(C)P(V|C) \cdot P(T|C) \cdot P(B|C) \cdot P(P|C)}{\cancel{P(V) \cdot P(T) \cdot P(B) \cdot P(P)}}$

constant for record, ignore

estimate parts of right-hand
side by counting:

use neg. examples for probabilities
conditioned on “neg”

id	veg?	type?	busy?	price?	class
1					
2	no	french	no	med	neg
3	yes	greek	yes	high	neg
4	no	french	yes	low	neg
5	no	french	yes	high	neg
6	yes	italian	yes	med	neg
7					
8					

$$P(V = y|C = neg) = \frac{2}{5}$$

$$P(V = n|C = neg) = \frac{3}{5}$$

$$P(T = i|C = neg) = \frac{1}{5}$$

$$P(T = f|C = neg) = \frac{3}{5}$$

$$P(T = g|C = neg) = \frac{1}{5}$$

$$P(P = l|C = neg) = \frac{1}{5}$$

$$P(P = m|C = neg) = \frac{2}{5}$$

$$P(P = h|C = neg) = \frac{2}{5}$$

$$P(B = y|C = neg) = \frac{4}{5}$$

$$P(B = n|C = neg) = \frac{1}{5}$$

id	veg?	type?	busy?	price?	class
A	yes	italian	no	low	?
B	no	greek	yes	high	?
C	no	greek	no	low	?
D	yes	greek	no	med	?
E	no	italian	yes	med	?
F	yes	greek	no	high	?

to classify A, compare

pos

$$P(\text{pos}) * P(V=y|\text{pos}) * P(T=i|\text{pos}) * P(B=n|\text{pos}) * P(P=l|\text{pos})$$

$$= 3/8 * 2/3 * 1/3 * 1/3 * 2/3 = 0.0185185$$

neg

$$P(\text{neg}) * P(V=y|\text{neg}) * P(T=i|\text{neg}) * P(B=n|\text{neg}) * P(P=l|\text{neg})$$

$$= 5/8 * 2/5 * 1/5 * 1/5 * 1/5 = 0.002$$

0.0185185 > 0.002 thus A is classified pos

id	veg?	type?	busy?	price?	class
A	yes	italian	no	low	?
B	no	greek	yes	high	?
C	no	greek	no	low	?
D	yes	greek	no	med	?
E	no	italian	yes	med	?
F	yes	greek	no	high	?

to classify B, compare

pos

$$P(\text{pos}) * P(V=n|\text{pos}) * P(T=g|\text{pos}) * P(B=y|\text{pos}) * P(P=h|\text{pos}) \\ = 3/8 * 1/3 * 1/3 * 2/3 * 1/3 = 0.009259$$

neg

$$P(\text{neg}) * P(V=n|\text{neg}) * P(T=g|\text{neg}) * P(B=y|\text{neg}) * P(P=h|\text{neg}) \\ = 5/8 * 3/5 * 1/5 * 4/5 * 2/5 = 0.024$$

0.024 > 0.009259 thus B is classified neg

id	veg?	type?	busy?	price?	class
A	yes	italian	no	low	?
B	no	greek	yes	high	?
C	no	greek	no	low	?
D	yes	greek	no	med	?
E	no	italian	yes	med	?
F	yes	greek	no	high	?

filling in remaining cases in the same way gives:

id	pos	neg	classified as
C	0.009259	0.003	pos
D	0	0.004	neg
E	0	0.024	neg
F	0.009259	0.004	pos

note the 0 for pos for D and E: this is because we haven't seen a training example with $P=m$ labeled pos