Exercise

1. train a NB classifier to predict "Class" using the following training data:

Name	Give Birth	Can Fly	Live in Water	Have Legs	Class
human	yes	no	no	yes	mammal
python	no	no	no	no	non-mammal
salmon	no	no	yes	no	non-mammal
whale	yes	no	yes	no	mammal
frog	no	no	sometimes	yes	non-mammal
komodo	no	no	no	yes	non-mammal
bat	yes	yes	no	yes	mammal
pigeon	no	yes	no	yes	non-mammal
cat	yes	no	no	yes	mammal
leopard shark	yes	no	yes	no	non-mammal
turtle	no	no	sometimes	yes	non-mammal
penguin	no	no	sometimes	yes	non-mammal
porcupine	yes	no	no	yes	mammal
eel	no	no	yes	no	non-mammal
salamander	no	no	sometimes	yes	non-mammal
gila monster	no	no	no	yes	non-mammal
platypus	no	no	no	yes	mammal
owl	no	yes	no	yes	non-mammal
dolphin	yes	no	yes	no	mammal
eagle	no	yes	no	yes	non-mammal

2. use your classifier to predict the class of the following example:

Give Birth	Can Fly	Live in Water	Have Legs	Class
yes	no	yes	no	?

Solution

training data

Name	Give Birth	Can Fly	Live in Water	Have Legs	Class
human	yes	no	no	yes	mammal
python	no	no	no	no	non-mammal
salmon	no	no	yes	no	non-mammal
whale	yes	no	yes	no	mammal
frog	no	no	sometimes	yes	non-mammal
komodo	no	no	no	yes	non-mammal
bat	yes	yes	no	yes	mammal
pigeon	no	yes	no	yes	non-mammal
cat	yes	no	no	yes	mammal
leopard shark	yes	no	yes	no	non-mammal
turtle	no	no	sometimes	yes	non-mammal
penguin	no	no	sometimes	yes	non-mammal
porcupine	yes	no	no	yes	mammal
eel	no	no	yes	no	non-mammal
salamander	no	no	sometimes	yes	non-mammal
gila monster	no	no	no	yes	non-mammal
platypus	no	no	no	yes	mammal
owl	no	yes	no	yes	non-mammal
dolphin	yes	no	yes	no	mammal
eagle	no	yes	no	yes	non-mammal

y/n   y/n   y/s/n   y/n   m/n

values of attributes

20 animals, of which 7 are m(ammal) and 13 n(on-mammal), thus

$$P(C = m) = \frac{7}{20}$$

$$P(C = n) = \frac{13}{20}$$

Solution

training data

	Attributes A$_i$			Class
B	F	W	L	C

Name	Give Birth	Can Fly	Live in Water	Have Legs	Class
human	yes	no	no	yes	mammal
					non-mammal
					non-mammal
whale	yes	no	yes	no	mammal
					non-mammal
					non-mammal
bat	yes	yes	no	yes	mammal
					non-mammal
cat	yes	no	no	yes	mammal
					non-mammal
					non-mammal
					non-mammal
porcupine	yes	no	no	yes	mammal
					non-mammal
					non-mammal
					non-mammal
platypus	no	no	no	yes	mammal
					non-mammal
dolphin	yes	no	yes	no	mammal
					non-mammal

y/n   y/n   y/s/n   y/n   m/n

values of attributes

$$P(L = y|C = m) = \frac{5}{7}$$

$$P(L = n|C = m) = \frac{2}{7}$$

$$P(W = y|C = m) = \frac{2}{7}$$

$$P(W = s|C = m) = \frac{0}{7}$$

$$P(W = n|C = m) = \frac{5}{7}$$

$$P(F = y \,|\, C = m) = \frac{1}{7}$$

$$P(F = n \,|\, C = m) = \frac{6}{7}$$

$$P(B = y \,|\, C = m) = \frac{6}{7}$$

$$P(B = n \,|\, C = m) = \frac{1}{7}$$

Solution

training data

| | Attributes A_i | | | Class |
| | B | F | W | L | C |

Name	Give Birth	Can Fly	Live in Water	Have Legs	Class
					mammal
python	no	no	no	no	non-mammal
salmon	no	no	yes	no	non-mammal
					mammal
frog	no	no	sometimes	yes	non-mammal
komodo	no	no	no	yes	non-mammal
					mammal
pigeon	no	yes	no	yes	non-mammal
					mammal
leopard shark	yes	no	yes	no	non-mammal
turtle	no	no	sometimes	yes	non-mammal
penguin	no	no	sometimes	yes	non-mammal
					mammal
eel	no	no	yes	no	non-mammal
salamander	no	no	sometimes	yes	non-mammal
gila monster	no	no	no	yes	non-mammal
					mammal
owl	no	yes	no	yes	non-mammal
					mammal
eagle	no	yes	no	yes	non-mammal

y/n y/n y/s/n y/n m/n

values of attributes

$$P(L = y | C = n) = \frac{9}{13}$$

$$P(L = n | C = n) = \frac{4}{13}$$

$$P(W = y | C = n) = \frac{3}{13}$$

$$P(W = s | C = n) = \frac{4}{13}$$

$$P(W = n | C = n) = \frac{6}{13}$$

$$P(F = y | C = n) = \frac{3}{13}$$

$$P(F = n | C = n) = \frac{10}{13}$$

$$P(B = y | C = n) = \frac{1}{13}$$

$$P(B = n | C = n) = \frac{12}{13}$$

Solution

Give Birth	Can Fly	Live in Water	Have Legs	Class
yes	no	yes	no	?

mammal?

$$P(C = m)$$
$$\cdot P(B = y | C = m)$$
$$\cdot P(F = n | C = m)$$
$$\cdot P(W = y | C = m)$$
$$\cdot P(L = n | C = m)$$
$$= \frac{7}{20} \cdot \frac{6}{7} \cdot \frac{6}{7} \cdot \frac{2}{7} \cdot \frac{2}{7}$$
$$= 0.021$$

non-mammal?

$$P(C = n)$$
$$\cdot P(B = y | C = n)$$
$$\cdot P(F = n | C = n)$$
$$\cdot P(W = y | C = n)$$
$$\cdot P(L = n | C = n)$$
$$= \frac{13}{20} \cdot \frac{1}{13} \cdot \frac{10}{13} \cdot \frac{3}{13} \cdot \frac{4}{13}$$
$$= 0.0027$$

$>$

mammal!