



University of
Nottingham
UK | CHINA | MALAYSIA

COMP3055

Machine Learning

Bayesian Learning Exercise Solution

Ying Weng
2024 Autumn

Exercise

- **Car theft**
 - **Question: A Red Domestic SUV will be stolen or not?**
 - Note there is no example of a Red Domestic SUV in our data set.

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Prompts

- **Car theft**
 - Attributes are Color , Type , Origin, and the subject, stolen can be either yes or no.

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Solution to Exercise

- **Car theft**
 - We want to classify a Red Domestic SUV.
 - Note there is no example of a Red Domestic SUV in our data set.

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Solution to Exercise

- We need to estimate

$$P(x_i | d_j) = \frac{N_c + mP}{N + m}$$

N = the number of training examples for which $d = d_j$

N_c = number of examples for which $d = d_j$ and $x = x_i$

P = a priori estimate for $P(x_i/d_j)$

m = the equivalent sample size

Solution to Exercise

- To classify a Red, Domestic, SUV, we need to estimate

$$Y = \arg \max_{d_i \in \{yes, no\}} P(d_i)P(x_1 = RED | d_i)P(x_2 = SUV | d_i)P(x_3 = Domestic | d_i)$$

Solution to Exercise

Yes:

Red:

$N = 5$

$N_c = 3$

$P = .5$

$m = 3$

SUV:

$N = 5$

$N_c = 1$

$P = .5$

$m = 3$

Domestic:

$N = 5$

$N_c = 2$

$P = .5$

$m = 3$

No:

Red:

$N = 5$

$N_c = 2$

$P = .5$

$m = 3$

SUV:

$N = 5$

$N_c = 3$

$P = .5$

$m = 3$

Domestic:

$N = 5$

$N_c = 3$

$P = .5$

$m = 3$

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Solution to Exercise

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

$$P(Red|Yes) = \frac{3 + 3 * .5}{5 + 3} = .56$$

$$P(SUV|Yes) = \frac{1 + 3 * .5}{5 + 3} = .31$$

$$P(Domestic|Yes) = \frac{2 + 3 * .5}{5 + 3} = .43$$

$$P(Red|No) = \frac{2 + 3 * .5}{5 + 3} = .43$$

$$P(SUV|No) = \frac{3 + 3 * .5}{5 + 3} = .56$$

$$P(Domestic|No) = \frac{3 + 3 * .5}{5 + 3} = .56$$

Solution to Exercise

- To classify a Red, Domestic, SUV, we need to estimate

$$Y = \arg \max_{d_i \in \{yes, no\}} P(d_i)P(x_1 = RED | d_i)P(x_2 = SUV | d_i)P(x_3 = Domestic | d_i)$$

$$\begin{aligned} &P(yes)P(x_1 = RED | yes)P(x_2 = SUV | yes)P(x_3 = Domestic | yes) \\ &= 0.5 * 0.56 * 0.31 * 0.43 = 0.037 \end{aligned}$$

$$\begin{aligned} &P(no)P(x_1 = RED | no)P(x_2 = SUV | no)P(x_3 = Domestic | no) \\ &= 0.5 * 0.43 * 0.56 * 0.56 = 0.069 \end{aligned}$$

$$Y = no$$