Tutorial 03

(Version 1.1)

1. Using the data in Table 1, compute:
   (a) The prior probability of playing Tennis tomorrow, $P(Tennis = Yes)$.
   (b) The conditional probability of playing Tennis tomorrow given that the Outlook is Rain, $p(Tennis = Yes|Outlook = Rain)$.

2. The forecast for tomorrow is for:

$$\{Outlook = Rain, Temp = Hot, Humidity = Normal, Wind = Weak\}$$

   Using the data in Table 1, compute the posterior probability of playing Tennis tomorrow using the Naive Bayes model.

3. What is v_{NB} for the data in in Table 1 given the evidence:

$$\{Outlook = Rain, Temp = Hot, Humidity = Normal, Wind = Weak\}$$

4. Using the Naive Bayes model, predict if I would play tennis tomorrow be if the forecast was:

$$\{Outlook = Cloud, Temp = Hot, Humidity = Normal, Wind = Weak\}$$

   according to the data in Table 1?

5. The probability of variable x is a mixture of two univariate Gaussians: $\mathcal{N}_1$, with mean 3 and variance 5; and $\mathcal{N}_2$ with mean 4 and variance 2. $\mathcal{N}_1$ has weight 0.4, $\mathcal{N}_2$ has weight 0.6.

   What is the probability that x takes the value 6?

   Recall that the probability density of a Gaussian distribution is given by:

$$p(x) = e^{-\frac{1}{2\sigma^2}(x-\mu_j)^2}$$

   where μ is the mean of the distribution and σ^2 is the variance.

6. Use k-means to cluster the following dataset:

Instance	Attributes	
	x_1	x_2
X_1	5	8
X_2	6	7
X_3	6	4
X_4	5	7
X_5	5	5
X_6	6	5
X_7	1	7
X_8	7	5
X_9	6	5
X_{10}	6	7

Day	Outlook	Temp	Humidity	Wind	Tennis
D1	Sun	Hot	High	Weak	No
D2	Sun	Hot	High	Strong	No
D3	Cloud	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Cloud	Cool	Normal	Strong	Yes
D8	Sun	Mild	High	Weak	No
D9	Sun	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sun	Mild	Normal	Strong	Yes
D12	Cloud	Mild	High	Strong	Yes
D13	Cloud	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Table 1: Some data on tennis playing — from the textbook and Lecture 3.

Take $(7, 5)$, $(9, 7)$ and $(9, 1)$ to be the initial cluster centres, and for simplicity, use Manhattan distance as the metric.

As a reminder, the Manhattan distance between two examples i with attributes (x_1^i, x_2^i) and j with attributes (x_1^i, x_2^j) is:

$$distance = |x_1^i - x_1^j| + |x_2^j - x_2^j|$$

where $|a|$ is the absolute value of a.

7. With reference to your answer to Question 6, Explain how the the result of the K-means algorithm can be considered to be a simple mixture model.

Version list

- Version 1.0, January 18th 2020.
- Version 1.1, January 11th 2021.