

Errors and Corrections in the Book

Chapter 2:

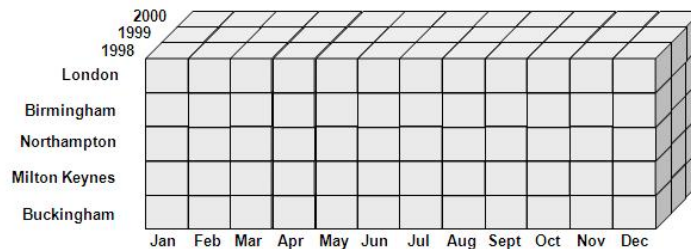
- On page 19, in the first paragraph, “(see figure 2.4)” should be “(see figure 2.4(a))”
- On page 34, in table 2.3, the expression for the sample median if n is even should be

$$\frac{1}{2} \left(\left(\frac{n}{2} \right)^{\text{th}} \text{value} + \left(\frac{n}{2} + 1 \right)^{\text{th}} \text{value} \right) \quad \text{if } n \text{ is even}$$

- On page 36, second paragraph third bullet point, “the standard deviation of the distribution is $\sigma/n^{(1/2)}$ ” should be “the standard deviation of the distribution is $\sigma_x/n^{(1/2)}$ ”

Chapter 3:

- On page 43, at the end of the third paragraph, add the following to complete the sentence: “, but we cannot say John’s year of birth is any number of *times* earlier than Susan’s.”
- On page 63, the 3D hypercube is redrawn as follows:



Chapter 4:

- On page 89, in the paragraph before section heading 4.2.8, “suppose that Habit is asymmetric binary attribute and...” should be changed to “suppose that Habit is asymmetric binary attribute, where “smoker” is considered as 1 and “nonsmoker” as 0, and...”
- On page 92, in the first paragraph below Figure 4.4, “The parameter f represents a distance function.” should be “The parameter **df** represents a distance function.”
- On page 92, in the pseudo code for the algorithm Kmeans, on the line above “S:= \emptyset ”, the reserved word “**for**” should be “**endfor**”.
- On page 96, in Table 4.6, the letter “E” should be “...”.

Chapter 5:

- On page 128, the heading for section 5.5.3 should be “The Expectation Maximisation (EM) Algorithm”.
- On page 138, figure 5.16 shows a scatter plot of points of different shaded colours. All data points should have had the same colour.

Chapter 6:

- On page 153, in the pseudo code for the algorithm ConstructTreeID3, the 6th line “return(Tree)” should be removed.
- On page 155, in Table 6.1, the character “E” should be replaced by “...”.
- On page 161, a divisor 14 has been missed out from the calculations of the chi-square statistic for the example before section 6.4. The correct results

should be $\chi^2(\text{Outlook}) \approx 2.1181$, $\chi^2(\text{Temperature}) \approx 0.5704$, $\chi^2(\text{Humidity}) = 2.8$, and $\chi^2(\text{Windy}) \approx 0.9333$.

- On page 163, in the fourth paragraph, all references to *testing* examples should be to *validation* examples. The last sentence should be changed to “Otherwise, if we have equal number of training examples of positive and negative classes, the label can be either positive or negative”.
- On page 177, in section 6.8, in the third paragraph, the sentences “Based on the relevant ... was also investigated” should be removed.

Chapter 7

- On page 185, in figure 7.2(a) and (b), letter “E” should be “...”.
- On page 215, in the last sentence of the last paragraph, “The understanding should also help readers to understand what how Weka selects a model to present and how ...” should be reworded as “The understanding should help readers to know what model Weka selects as the final model and how ...”

Chapter 8

- On page 223, in the pseudo code for the Apriori algorithm, the last sentence “return ($F_1 \cup F_2 \cup \dots \cup F_k$)” should be “return ($F_1 \cup F_2 \cup \dots \cup F_{k-1}$)”. In the last sentence of the second paragraph after the algorithm pseudo code on the same page, “ F_1, F_2, \dots, F_k ” should therefore be “ F_1, F_2, \dots, F_{k-1} ”.

Chapter 9

- On page 247, in the pseudo code for the AprioriAdapted algorithm, the last sentence “return ($F_1 \cup F_2 \cup \dots \cup F_k$)” should be “return ($F_1 \cup F_2 \cup \dots \cup F_{k-1}$)”. The same error also exists on page 257 in the AprioriAdaptedSeq algorithm.
- On page 266, in the fifth paragraph, the second sentence “Candidates AB, AC, AD,BC and BD are already in CF” should be “Candidates AC, BC and BD are already in CF and hence their support is increased according to the support within the partition p_4 ”.
- On page 267, in the fifth paragraph, the sentence “BD has insufficient support...” should be removed. The following sentence should be changed to “BD and BE have sufficient support and stay in CF”.

Chapter 10

- On page 298, in the last sentence of paragraph 4, “Figure 10.11 shows ...” should be “Figure 10.12 shows ...”.