Errata for "Notes on probability theory and probabilistic Machine Learning"

Alessio Benavoli

September 30, 2020

Chapter 1

- pag.9,

$$p(x \in \{2,4,6\}) = \sum_{\mathsf{x} \in \{2,4,6\}} p(x = \mathsf{x}) \\ = p(x=2) + p(x=4) + p(x=6) = 0.51,$$

- pag.10,

$$p(x \in \{1,3,5\}) = 1 - p(x \in \{2,4,6\}) = 0.49$$

- pag. 14, there is an extra
at the end of formula (1.4).

- pag. 15, What is $p(y=1)$?

$$p(y=1) = \sum_{\mathsf{x} \in \{H,T\}} p(x = \mathsf{x}, y = 1) \\ = p(x=H, y=1) + p(x=T, y=1) = 0.2.$$

- pag. 16, equation (1.7) should be

$$p(x = \mathsf{x} | y = \mathsf{y}) = p(x = \mathsf{x})$$

- pag. 16, "The three conditions for independence in Equations (1.5), (1.6) and (1.7)".

- pag. 17, the same typo above occurs in

$$p(y=i) = \sum_{\mathsf{x} \in \{H,T\}} p(x = \mathsf{x}, y = i) \\ = p(x=H, y=i) + p(x=T, y=i).$$

- pag. 22, "In practice, we sum all the rows in the table that include the instance $\mathsf{s} = 1$ (colored rows in the following table):"

- pag. 26,

$$p(y = \mathsf{y} \mid x = \mathsf{x}, z = \mathsf{z}) = p(y = \mathsf{y} \mid z = \mathsf{z})$$

- pag. 23, "we aim to answer"

- pag. 29, "ourcome" should be "outcome".

- pag. 30, Exercise 8, "Consider the last exercise. We have now a thermometer whose rate of false negative reading is 5% and false positive reading is 15%"

- pag. 31, Exercise 4.b, "We can also solve it by applying Bayes' rule:

$$p(x=i, y=i | x+y \leq 4) = \frac{p(x+y \leq 4 | x=i, y=i) p(x=i, y=i)}{p(x+y \leq 4)}$$

the numbers at the end should be removed. This is the generic formula.

- pag. 31, Exercise 4.b,

$$p(x=2,y=2|x+y\leq 4)=\frac{p(x+y\leq 4|x=2,y=2)p(x=2,y=2)}{p(x+y\leq 4)}=\frac{\frac{1}{36}}{\frac{1}{6}}=\frac{1}{6}$$

- pag. 32, Exercise 6, in the formulas one should use capital B instead of b (just a notation issue).
- pag. 33, exercise 7 answer =0.192. it is 0.19149 and so it should be 0.191 (when rounded).
- pag. 33, "HighTmp" should be "HighTemp"

Chapter 2

- pag. 38 code at the top of the page, the code can returns 0,1,2,3,4,5 instead 1,2,3,4,5,6 due to the fact that Python counts from zero.
- pag. 39, "and it is also true if we roll the dice more than two times." This sentence is a bit misleading, I mean that the fact that the probability does not depend on the order of the outcomes even fore more than two rolls.
- pag 48, there is typo in the code:
```
def predict(self, X):
return np.argmax(self.predict_proba(X), axis=1)
```
- pag. 49, "the number of possible sentences"
- pag 52, $p(bye)\approx 3/48$.
- pag. 56, Exercise 1, "For instance, the first email (first row) includes the word "money", does not include "win" and it is not spam."
- pag. 59, "Therefore, for the first row in the dataset we have: so" (this extra should be removed)

Chapter 3

- pag. 68, Figure 3.6 caption "Gaussian CDF for different values of"
- pag. 72, in the green box, the Gaussian PDF is wrong, it should be

$$P(x\in B)=\int_{0.2}^{0.3}\frac{1}{\sqrt{2\pi 0.05^2}}\exp\left(-\frac{(x-0.25)^2}{2\cdot 0.05^2}\right)=0.68.$$

  The second equation in the green block has the same typo twice.

Chapter 4

- pag. 101, bullet 1, "Sample 200 prior regression lines from the above model, that is the regression lines before seeing the data (prior regression lines).."
- pag. 103, "After two and, respectively,"