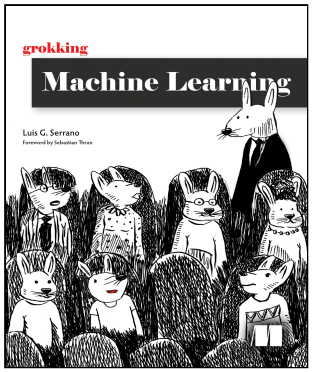
MANNING PUBLICATIONS



Errata: July 17, 2023

Thank you for purchasing *Grokking Machine Learning*. Please post errata not listed below in this book's LiveBook Errata thread. We'll update this list as necessary. Thank you!

The following corrections need to be made to all formats.

Chapter 3, Section "How to get the computer to draw this line: The linear regression algorithm", Subsection "Crash course on slope and y-intercept", page 47

In the last paragraph: "This line cuts the x-axis at height 2, and that is the y-intercept." should be "This line cuts the y-axis at height 2, and that is the y-intercept."

Chapter 3, Section "The general linear regression algorithm (optional)", Subsection "Pseudocode for the general square trick", page 60

Under "Procedure", third line: "ηx" should be "ηx_i" and "ηr" should be "ηx_i"

Chapter 4, page 79

In Paragraph 1: "Paramaters" should be "Parameters"

Chapter 4, Section "How do we get the computer to pick the right model? By testing", page 82

In Figure 4.4 description, third sentence: "The columns represent the training and the testing error." should be "The rows represent the training and the testing error."

Chapter 4, Section "Another alternative to avoiding overfitting: Regularization", Subsection "Another example of overfitting: Movie recommendations", page 89

In Paragraph 7: "But unfortunately, if model 2 produces a smaller error than model 2," should be "But unfortunately, if model 2 produces a smaller error than model 1,"

Chapter 6, Section "Coding the logistic regression algorithm", Subsection "Coding the logistic regression algorithm by hand", page 170

After Figure 6.9, Paragraph 1: "On the plot of the intermediate classifiers, the final one corresponds to the dark line." should be "On the plot of the intermediate

classifiers (Figure 6.10, left), the final one corresponds to the dark line."

Chapter 7, Section "A useful tool to evaluate our model: The receiver operating characteristic (ROC) curve", Subsection "Sensitivity and specificity: Two new ways to evaluate our model", page 191

The titles of the paragraphs are reversed. They should be "Calculating the sensitivity" and then "Calculating the specificity"

Chapter 7, Section "A useful tool to evaluate our model: The receiver operating characteristic (ROC) curve", Subsection "The receiver operating characteristic (ROC) curve: A way to optimize sensitivity and specificity in a model", page 192

In the last Paragraph: "there are no true positives," should be "there are no true negatives,"

Chapter 7, Section "A useful tool to evaluate our model: The receiver operating characteristic (ROC) curve", Subsection "Recall is sensitivity, but precision and specificity are different", page 199

In Paragraph 7: "If we focus on the bottom row (the negatively labeled examples), we can calculate specificity by dividing the number on the left column" should be "If we focus on the bottom row (the negatively labeled examples), we can calculate specificity by dividing the number on the right column"

Chapter 8, Section “Sick or healthy? A story with Bayes' theorem as the hero”, page 209

After Figure 8.1, Paragraph 1: The equation “99/9,999=0.0089” should be “99/(99+9,999)=0.0098”

Chapter 8, Section “Use case: Spam-detection model”, Subsection “What about two words? The naive Bayes algorithm”, page 222

In Paragraph 3: “0.6” should be “0.3”

Chapter 8, Section “Use case: Spam-detection model”, Subsection “What the math just happened? Turning ratios into probabilities”, page 220

At the top of the page, before the second equation: “$F|E$ and $F|E^C$” should be “$F \cap E$ and $F \cap E^C$”, please see the correction *here*

Chapter 8, Section “Use case: Spam-detection model”, Subsection “What about two words? The naive Bayes algorithm”, page 222

In the last bullet point on the page: “a spam email contains both words is 0.45,” should be “a spam email contains both words is 0.225,”

Chapter 10, Section “Neural networks with an example: A more complicated alien planet”, Subsection “Combining the outputs of perceptrons into another perceptron”, page 284

In the paragraph before Figure 10.5: “and a third table in which the first two columns are the inputs and the outputs of the career and family classifier, and the last column is the output of the family classifier.” should be “and a third table in which the first two columns are the outputs of the career and family classifier, and the last column is the output of the happiness classifier.”

Chapter 10, Section “Neural networks with an example: A more complicated alien planet”, Subsection “Combining the outputs of perceptrons into another perceptron”, page 284

One of the values in Figure 10.5 is incorrect, in third Table “Happiness Classifier”: In Column 3, Row 4, “-0.5” should be “-1.5”

Chapter 10, Section “A graphical example in two dimensions”, Subsection “The architecture of the neural network”, page 302

Figure 10.21 is incorrect, please see the correct Figure *here*

Chapter 11, Section “Using polynomial equations to our benefit: The polynomial kernel”, Subsection “Going beyond quadratic equations: The polynomial kernel”, page 334

In Table 11.5 description: “We have added three more rowcolumns” should be “We have added three more columns”

Chapter 11, Section "Using polynomial equations to our benefit: The polynomial kernel", Subsection "Going beyond quadratic equations: The polynomial kernel", page 334

After Table 11.5, in first sentence: "x_4" should be "x_5"

Chapter 13, Section "Turning categorical data into numerical data: One-hot encoding", Subsection "Can we one-hot encode numerical features? If so, why would we want to?", page 397

In Paragraph 1, bullet point 2: "40.38%" should be "47.28%"

In Paragraph 1, bullet point 3: "55%" should be "24.24%"

Chapter 13, Section "Which model is better? Evaluating the models", page 402

In Paragraph 1: "chapter 4" should be "chapter 7"

Appendix A, Section "Chapter 6: A continuous approach to splitting points: Logistic classifiers", Subsection "Exercise 6.2, Solution", page 423

In Solution, part c, first sentence: "w_1x_1" should be "w_2x_2"

Appendix A, Section "Chapter 8: Using probability to its maximum: The naive Bayes model", Subsection "Exercise 8.3, Solution", page 433

In Solution, part b, in the third equation: "$P(T^C \mid S) = 2/4$" should be "$P(T^C \mid H) = 2/4$"

Appendix A, Section "Chapter 9: Splitting data by asking questions: Decision trees", Subsection "Exercise 9.3, Solution, Splitting based on the T feature:", page 440

In Paragraph 1: "(based only on the F feature)" should be "(based only on the T feature)"

Appendix A, Section "Chapter 11: Finding boundaries with style: Support vector machines and the kernel method", pages 445-446

The 11.1 Solution is incorrect, please find the correct Solution *here*

The 11.1 Solution graph on page 446 is incorrect, please find the correct image posted *here*

Chapter Appendix B: The math behind gradient descent, Section "Using gradient descent to train linear regression models", Subsection "Training a linear regression model using gradient descent to reduce the mean absolute error", page 455

In the fourth equation on the page, on the Left side: "w_i" should be "w_j"

Chapter Appendix B: The math behind gradient descent, Section "Using gradient descent to train classification models", Subsection "Training a logistic regression

model using gradient descent to reduce the log loss", page 462

In first equation, on Right side, two times: "xj(i)" should be "$x_j^{(i)}$"

In first equation, on Right side, two times: "yi" should be "y_i"

The following corrections have been made to all formats for the book's second printing Feb 2022.

Front Matter, Section "contents", page xvi

In table of contents, the Appendices should be:

"Appendix A Solutions to the exercises"

"Appendix B The math behind gradient descent: Coming down a mountain using derivatives and slopes"

"Appendix C References"

Author refers to Appendix A, B, and C on p. xvi.

Chapter 4, Section "Another alternative to avoiding overfitting: Regularization", Subsection "Measuring how complex a model is: L1 and L2 norm", page 90

In Paragraph 12:

"**Model 1:** $2^2 = 2$"

should be

"**Model 1:** $2^2 = 4$"

Chapter 6, Section "Logistic classifiers: A continuous version of perceptron classifiers", Subsection "Comparing classifiers using the log loss", page 160

In Paragraph 7, Bullet point 2:

"0.73" should be "0.731"

AND

"ln(0.721)" should be "ln(0.731)"

In Paragraph 7, Bullet point 3:

"0.73" should be "0.731"

AND

"ln(731)" should be "ln(0.731)"
