Errata and additions in the text

Below are errata and minor alterations that improve the style or clarify the book. To see which printing of the book you have, look on the third page of the text itself, which faces the dedication page. At the bottom you will see:

Printed in the United States of America
10 9 8 7 6 5 4 3 2

The last number at the right gives the number of the printing; thus, as illustrated here, “2” means that this is the second printing.

First and second printings

Below, “line +7” means the seventh line from the top of the text body (not including figure or table captions, or other headlines unless otherwise indicated), and “line -5” means the fifth line from the bottom of the text body. In Algorithms, the numbering refers to the line numbers within the algorithm itself. Thus Algorithm 4, line 5” means line 5 in Algorithm 4, not the fifth line from the top of the page.

Front matter

page x line +6: Change “4.8 Reduced Coulomb” to “*4.8 Reduced Coulomb”

page xv line -13: Change “A.4.7 The Law of Total Probability and Bayes’ Rule” to “A.4.7 The Law of Total Probability and Bayes Rule”

page xviii Take the last sentence under Examples, “In addition, in response to popular demand, a Solutions Manual has been prepared to help instructors who adopt this book for courses.” and move it to be the final sentence under Problems, lower on the same page.

page xviii lines -10– -11: Change “and they are generally” to “and they are typically”

page xix line -15: Change “Ricoh Silicon Valley” to “Ricoh Innovations”
Chapter 1

page 1 line +4: Change “data and taking” to “data and making”

page 4 the only equation on the page: Change “$\mathbf{x} = [x_1 \ldots x_2]$” to “$\mathbf{x} = (x_1 \ldots x_2)$”

page 11 line +21: Change “us for practical, rather than” to “us for practical rather than”

page 14 line -4 in the caption to Figure 1.8: Change “of the data impact both” to “of the data affect both”

page 19 line +15: Change “is achieved in humans” to “is performed by humans”

Chapter 2

page 21 line -6 in the footnote: Change “should be written as $p_X(x|\omega)$” to “should be written as $p_x(x|\omega)$”

page 21 line +4 in the footnote: Change “clear that $p_X(\cdot)$ and $p_Y(\cdot)$” to “clear that $p_x(\cdot)$ and $p_y(\cdot)$”

page 22 second line after Eq. 3: Change “probability (or posterior) probability” to “probability (or posterior)”

page 23 second line after Eq. 7: Change “By using Eq. 1, we can” to “By using Eq. 1 we can”

page 26 first line after Eq. 17: Change “and $\omega_2$ otherwise.” to “and otherwise decide $\omega_2$."

page 28 Equation 23: Change “$(\lambda_{11} - \lambda_{22}) - (\lambda_{21} - \lambda_{11})$” to “$(\lambda_{11} - \lambda_{22}) + (\lambda_{21} - \lambda_{11})$”

page 28 second line after Eq. 24: Change “decision boundary gives” to “decision boundary then gives”

page 32 second line after Eq. 33: Change “expected values — by these” to “expected values by these”

page 36 first equation after Eq. 49: Change “Let us examine the discriminant” to “Let us examine this discriminant”

page 41 Figure 2.13, caption, line +2: Change “unequal variance.” to “unequal variance, as shown in this case with $P(\omega_1) = P(\omega_2)$.”

page 47 Equation 73: Change “for0” to “for 0” (i.e., add space)

page 47 line - 10: Change “substituting the results in Eq. 73” to “substituting this $\beta$ into Eq. 73”

page 47 line -2: Change “This result is the so-called” to “This gives the so-called”

page 48 Example 2, line +3: Change “4.11,” to “4.06,”

page 48 Example 2, line +4: Change “0.016382.” to “0.0087.”
At the time of the release of the first printing, the problem of inference in Bayes belief nets with loops was not fully solved. Since that time, however, such a solution has emerged and for this reason Section 2.11 has been rewritten accordingly. This revision is posted on the Wiley site.

Page 66 Problem 2, part (b), line +2: Change “for arbitrary $a_i$ and $b_i$.” to “for arbitrary $a_i$ and positive $b_i$."

Page 66 Problem 3, part (a) equation: End the equation with a period (full stop).

Page 67 Problem 5, part (d): Change “What is the minimax risk?” to “What is the minimax risk for part (c)?”

Page 67 Problem 6, part (2), line +2: Change “Determine the decision boundary” to “Determine the single-point decision boundary”

Page 69 Move the title “Section 2.4” to the top of the page so that Problem 13 is now under Section 2.4.

Page 71 Problem 20, part (a), line +1: Change “we know only that a distribution is nonzero in” to “we know solely that a distribution is nonzero only in”

Page 71 Problem 20, part (b), line +1: Change “we know only that a distribution is nonzero for” to “we know solely that a distribution is nonzero only for”

Page 71 Problem 23, at the center of the typeset equation, change “and $\Sigma$” to “and $\Sigma$” (i.e., add space)

Page 72 Problem 24, line +1: Change “normal density for which $\sigma_{ij} = 0$” to “normal density with mean $\mu, \sigma_{ij} = 0$”

Page 75 Problem 37, first equation: Change “and $P(\omega_1)$” to “and $P(\omega_1)$” (i.e., add space)

Page 75 Problem 37, part (c) equation: Change “and $p(x|\omega_2)$” to “and $p(x|\omega_2)$” (i.e., add space)

Page 75 Problem 39, line +1: Change “Use the signal detection” to “Use signal detection”

Page 75 Problem 39, part (a), line +1: Change “and $P(x < x^*|x \in \omega_2)$, taken” to “and $P(x > x^*|x \in \omega_1)$, taken”

Page 75 Problem 39, part (b): Replace the last two sentences with “Estimate $d'$ if $P(x > x^*|x \in \omega_1) = 0.8$ and $P(x > x^*|x \in \omega_2) = 0.3$. Repeat for $P(x > x^*|x \in \omega_1) = 0.7$ and $P(x > x^*|x \in \omega_1) = 0.4$."

Page 75 Problem 39, part (d): Replace the two equation lines with “Case A: $P(x > x^*|x \in \omega_1) = 0.8, P(x > x^*|x \in \omega_2) = 0.3$ or Case B: $P(x > x^*|x \in \omega_1), P(x > x^*|x \in \omega_2) = 0.7.$”

Page 76 Problem 41, first line after the equation: Change “$((\mu_2 - \mu_1)/\delta, \ldots, (\mu_2 - \mu_1)/\delta$” to “$((\mu_2 - \mu_1)/\delta$”

Page 76 Problem 41, part (b): Change “$d'_r = 1.0$” to “$d'_r = 1.0$ and 2.0.”
ERRATA IN THE TEXT

page 76  Problem 41, part (c): Change “P(x > x^*|x \in \omega_1) = .2.” to “P(x > x^*|x \in \omega_1) = .7.”

page 76  Problem 76, part (e): Change “measure P(x > x^*|x \in \omega_2) = .9 and (x > x^*|x \in \omega_1) = .3.” to “measure P(x > x^*|x \in \omega_2) = .3 and P(x > x^*|x \in \omega_1) = .9.”

page 81  Computer exercise 6, part (b), line +1: Change “Consider” to “Consider the normal distributions”

page 81  Computer exercise 6, part (b), equation: Change “and p(x|\omega_2)” to “and p(x|\omega_2)” (i.e., add space)

page 81  Computer exercise 6, part (b), equation: Move “with P(\omega_1) = P(\omega_2) = 1/2.” out of the centered equation, and into the following line of text.


Chapter 3

page 88  Equation 9: Change “\nabla_{\theta \mu}” to “\nabla_{\mu}”

page 91  Ninth line after Eq. 22: Change “shall consider), the samples” to “shall consider) the samples”

page 100  line -5: Change “are equivalent to” to “are more similar to”

page 100  line -5 – -4: Change “If there are much data” to “If there is much data”

page 102  line -5: Change “(Computer exercise 22)” to “(Problem 22)”

page 103  line -2: Change “choice of an prior” to “choice of a prior”

page 104  line +6: Change “if” to “only if”

page 104  line +16: Change “only if” to “if”

page 104  Equation 62: Make the usage and style of the summation sign (\sum) uniform in this equation. Specifically, in two places put the arguments beneath the summation sign, that is, change “\sum_{D \in D}” to “\sum_{D \in D}”

page 105  first line after Eq 63: Change “to this kind of scaling.” to “to such scaling.”

page 111  lines +9 – 10: Change “constants c_0 and x_0 such that |f(x)| \leq c_0|h(x)| for all” to “constants c and x_0 such that |f(x)| \leq c|h(x)| for all”

page 111  line +14: Change “proper choice of c_0 and x_0.” to “proper choice of c and x_0.”

page 116  Equation 86: Change “\lambda e^e” to “\lambda(e^e - 1)”

page 125  Algorithm 1, line 1: Change “i = 0” to “i \leftarrow 0”
Chapter 3

Page 126 First line of the equation at the middle of the page: Change
\[ Q(\theta; \theta^0) = \mathcal{E}_{x_{41}}[\ln p(x_g, x_b; \theta|\theta^0; D_g)] \]

to
\[ Q(\theta; \theta^0) = \mathcal{E}_{x_{41}}[\ln p(x_g, x_b; \theta|\theta^0; D_g)] \]

Page 128 Line +6, (second line after the Example): Change “the EM algorithm, and they” to “the EM algorithm as they”

Page 129 Second line above Section 3.10.3: Change “while the \( \omega_i \) are unobservable” to “while the \( \omega_j \) are unobservable”

Page 132 Line +3: Change “\( b_{kj} \), and thus” to “\( b_{jk} \), and thus”

Page 132 Equation 136: Replace by:
\[
\alpha_j(t) = \begin{cases} 
0 & \text{if } t = 0 \text{ and } j \neq \text{initial state} \\
1 & \text{if } t = 0 \text{ and } j = \text{initial state} \\
\sum_{c} \alpha_i(t-1)a_{ij}b_{kj}v(t) & \text{otherwise.}
\end{cases}
\]

Page 132 Third line after Eq. 136: Change “Consequently, \( \alpha_i(t) \) represents” to “Consequently, \( \alpha_j(t) \) represents”

Page 132 Fourth line after Eq. 136: Change “hidden state \( \omega_i \)” to “hidden state \( \omega_j \)”

Page 132 Algorithm 2, Line 1: Delete “\( \omega(1), \)”

Page 132 Algorithm 2, Line 1: Change “\( t = 0 \)” to “\( t \leftarrow 0 \)”

Page 132 Algorithm 2, Line 1: Change “\( \alpha(0) = 1 \)” to “\( \alpha_j(0) \)”

Page 132 Algorithm 2, Line 3: Replace entire line by “\( \alpha_j(t) \leftarrow b_{jk}v(t)\sum_{c} \alpha_i(t-1)a_{ij} \)”

Page 132 Algorithm 3: Somehow the line numbering became incorrect. Change the line number to be sequential, 1, 2, \ldots 6.

Page 132 Algorithm 3, Line 1: Change “\( \omega(t), t = T \)” to “\( \beta_j(T), t \leftarrow T \)”

Page 132 Algorithm 3, old Line 4, renumbered to be Line 3: Replace entire line by “\( \beta_j(t) \leftarrow \sum_{c} \beta_j(t+1)a_{ij}b_{jk}v(t+1) \)”

Page 132 Algorithm 3, old Line 7, renumbered Line 5: Change “\( P(V^T) \)” to “\( P(V^T) \)” (i.e., make the “\( V \)” bold)

Page 133 Figure 3.10, caption, Line +5: Change “was in state \( \omega_j(t = 2) \)” to “was in state \( \omega_i(t = 2) \)”

Page 133 Figure 3.10, caption, Line +6: Change “is \( \alpha_j(2) \)” for \( j = 1, 2 \)” to “is \( \alpha_i(t) \)” for \( i = 1, 2 \)”

Page 133 Figure 3.10, caption, Line -1: Change “\( \alpha_2(3) = b_{2k}\sum_{c} \alpha_j(2)a_{j2} \)” to “\( \alpha_2(3) = b_{2k}\sum_{c} \alpha_i(2)a_{i2} \)”
ERRATA IN THE TEXT

page 133 line -6: Change \( V^5 = \{v_3, v_1, v_3, v_2, v_0\} \) to \( V^4 = \{v_1, v_3, v_2, v_0\} \)

page 133 line -44: Change “is shown above,” to “is shown at the top of the figure.”

page 134 Figure in Example 3, caption, line +4: Change “\( \alpha_i(t) \) — the probability” to “\( \alpha_j(t) \) — the probability.”

page 134 Figure in Example 3, caption, line +6: Change “calculation of \( \alpha_i(1) \)” to “calculation of \( \alpha_j(1) \).”

page 134 Algorithm 4 old line 5 (now renumbered 4): Change “\( k \leftarrow k + 1 \)” to “\( j \leftarrow j + 1 \)”

page 134 Algorithm 4 old line 7 (now renumbered 5): Change “\( \alpha_k(t) \)” to “\( \alpha_j(t) \)”

page 137 Equation 138: Replace equation by

\[
\beta_i(t) = \begin{cases} 
0 & \omega_i(t) \neq \omega_0 \text{ and } t = T \\
1 & \omega_i(t) = \omega_0 \text{ and } t = T \\
\sum_j \beta_j(t+1)a_{ij}b_{jk}v(t+1) & \text{otherwise.}
\end{cases}
\]

page 137 seventh line after Eq. 138: Change “\( \beta_i(T - 1) = \sum_j a_{ij}b_{ij}v(T)\beta_j(T) \)” to “\( \beta_i(T - 1) = \sum_j a_{ij}b_{jk}v(T)\beta_j(T) \)”.

page 137 fourth line before Eq. 139: Change “probabilities \( a_{ij} \) and \( b_{ij} \)” to “probabilities \( a_{ij} \) and \( b_{jk} \)”

page 137 Equation 139: Change “\( \hat{b}_{ij} \)” to “\( \hat{b}_{jk} \)”

page 138 line +3: Change “whereas at step \( t \) it is” to “whereas the total expected number of any transitions from \( \omega_i \) is”

page 138 first line after Eq. 140: Change “\( \hat{b}_{ij} \)” to “\( \hat{b}_{jk} \)”
Equation 141: Replace equation by:

\[ \hat{b}_{jk} = \frac{\sum_{t=1}^{T} \sum_{l} \gamma_{jl}(t)}{\sum_{t=1}^{T} \sum_{l} \gamma_{jl}(t)} \]

Algorithm 5, line 1: Change “criterion \( \theta \)” to “criterion \( \theta, z \leftarrow 0 \)”

Problem 22, line between the two equations: Change “has a uniform” to “has a uniform distribution”

Problem 27, part (a), line after the equation: Change “as given in Table 3.1.” to “as in Table 3.1.1.”

Problem 31, line +1: Change “suppose \( a \) and \( b \) are constants” to “suppose \( a \) and \( b \) are positive constants”

Problem 32, line +1: Change “where the \( n \) coefficients” to “at a point \( x \), where the \( n \) coefficients”

Problem 34, line +4: Change “the number of operations \( n \)” to “the maximum size \( n \)”

Problem 38, line +1: Change “\( p_x(x|\omega_i) \)” to “\( p_x(x|\omega_i) \)”

Algorithm 5, line +1: Change “and” to “is maximized by”

Problem 43, line +4: Change “and the \( d \) mean vectors.” to “and the \( c \) mean vectors.”

Problem 46, (top equation): Change “\( 0 \) otherwise.” to “\( \epsilon \) otherwise.”

Problem 46, line +1 (after the equation): Change “missing feature values.” to “missing feature values and \( \epsilon \) is a very small positive constant that can be neglected when normalizing the density within the above bounds.”

Problem 46, part (b): Add “You may make some simplifying assumptions.”

Problem 47, first equation: Change “\( e^{-\theta_1 x_1} \)” to “\( e^{-x_1/\theta_1} \)”

Computer exercise 4, line -3: Change “apply it to the \( x_1 - x_2 \) components” to “apply it to the \( x_1 - x_2 \) components” (i.e., eliminate spaces and note that the dash is not subtraction sign, but an n-dash)

Computer exercise 6, part (a), line +2: Change “in the Table above.” to “in the table above.”
Chapter 4

page 172 Figure 4.9, caption, line -1: Change “where \( I \) is the \( d \times d \) identity matrix.” to “where \( I \) is the \( d \)-by-\( d \) identity matrix.”

page 173 Algorithm 1, line 1: Change “\( j = 0 \)” to “\( j \leftarrow 0 \)”

page 173 Algorithm 2, line 1: Change “test pattern,” to “test pattern”

page 179 Equation 41: Change “\( \int_{x' \in S} \)” to “\( \int_{x' \in S} \)” (i.e., place the limits underneath the integral sign)

page 184 lines +2 – 3: Keep “\( k - i > i \)” on the same line (i.e., do not put a line break in this equation)

page 186 Algorithm 3, line 1: Change “\( j = 0, D = \) data set, \( n = \) # prototypes” to “\( j \leftarrow 0, D \leftarrow \) data set, \( n \leftarrow \) # prototypes”

page 187 Figure 4.18, caption, line -4: Change “by a factor 1/3” to “by a factor \( \alpha = 1/3 \)”

page 188 First margin note: Change “MINKOWSKI METRIC” to “MINKOWSKI METRIC” (i.e., capitalize the M in MINKOWSKI)

page 188 Second margin note: Change “MANHATTAN DISTANCE” to “MANHATTAN DISTANCE” (i.e., capitalize the M in MANHATTAN)

page 188 Third margin note: Change “TANIMOTO METRIC” to “TANIMOTO METRIC” (i.e., capitalize the T in TANIMOTO)

page 189 line +5: Change “the relative shift is a mere” to “the relative shift \( s \) is a mere”

page 192 Figure 4.23, caption, lines -4 – -2: Eliminate the second-to-last sentence.

page 193 Equation 61: Replace the full equation with “\( \mu_x(x) \cdot \mu_y(y) \)”

page 194 line +19: Change “belief about memberships” to “belief about memberships”

page 195 line +9: Change in the title “4.8 REDUCED COULOMB” to “*4.8 REDUCED COULOMB”

page 196 Algorithm 4, line 1: Change “\( j = 0, n = \) # patterns, \( \epsilon = \) small parameter, \( \lambda_m = \) max radius” to “\( j \leftarrow 0, n \leftarrow \) # patterns, \( \epsilon \leftarrow \) small parameter, \( \lambda_m \leftarrow \) max radius”

page 197 Algorithm 5, line 1: Change “\( j = 0, k = 0, x = \) test pattern, \( D_t = \{ \} \)” to “\( j \leftarrow 0, k \leftarrow 0, x \leftarrow \) test pattern, \( D_t \leftarrow \{ \} \)”

page 199 line -7: Change “prior knowledge.” to “prior beliefs.”

page 203 Problem 11, part (d), line +1: Change “close to an edge” to “close to a face”

page 203 Problem 11, part (d) line +4: Change “closer to an edge” to “closer to a face”
10.0. CHAPTER 5

page 203  Problem 11, part (d), line +5: Change “even though it is easier to calculate here” to “and happens to be easier to calculate”

page 204  Problem 13 line +1: Change “from the distributions” to “with priors $P(\omega_1) = P(\omega_2) = 0.5$ and the distributions”

page 205  Problem 21, part (a), line +1: Change “As given in the text, take” to “Follow the treatment in the text and take”

page 206  Problem 23 part (d) line +2: Change “space, then the $b$” to “space, then in the $b$”

page 207  Problem 26, line +4: Change “and find its nearest” to “and seek its nearest”

page 207  Problem 27, part (a) equation: Replace current equation with “$D_{\text{Tanimoto}}(S_1, S_2) = \frac{n_1+n_2-2n_{12}}{n_1+n_2-n_{12}}$”

page 207  Problem 29, first equation: Change “$\delta_i$” to “$\delta_1$” in all three places

page 207  Problem 29, second equation: Change “$\hat{C}(x, \mu_i)$” to “$\hat{C}(x; \mu_i)$”

page 207  Problem 29 second equation: Change “$\delta_i$” to “$\delta_2$” in all three places

page 207  Problem 29 line -5: Change “we have for the length $\delta_i = 5$” to “we have for the length $\delta_1 = 5$”

page 207  Problem 29 line -4: Change “and for lightness $\delta_j = 30$” to “and for the lightness $\delta_2 = 30$”

Chapter 5

page 218  line +6: Change “the problem to $c - 1$” to “the problem to $c$”

page 218  Equation 2: Change “$w^T x_i$” to “$w_i^T x$”

page 219  second line after the second (unnumbered) equation: Change “is given by $(g_i - g_j)/\|w_i - w_j\|$” to “is given by $(g_i(x) - g_j(x))/\|w_i - w_j\|$”

pages 221–222 (across the page break): Change “mul-ti-layer” to “multi-layer” (i.e., hyphenate as “multi-layer”)

page 225  Algorithm 1, line 1: Change “$k = 0$” to “$k = 0$”

page 228  Algorithm 3, line 1: Change “$k = 0$” to “$k = 0$”

page 229  line +5: Change “We shall begin our examination” to “We begin our examination”

page 229  line -3: Change “Thus we shall denote” to “Thus we denote”

page 230  Algorithm 4, line 1: Change “$k = 0$” to “$k = 0$”

page 230  fourth line before Theorem 5.1: Change “correction is clearly moving” to “correction is hence moving”
ERRATA IN THE TEXT

page 230  line -2: Change “From Eq. 20,” to “From Eq. 20 we have”

page 232  first line after Eq. 24: Change “Because the squared distance” to “Because this squared distance”

page 233  Algorithm 5, line 1: Change “k = 0” to “k ← 0”

page 233  Algorithm 6, line 1: Change “k = 0” to “k ← 0”

page 234  line +22 (counting from the end of the Algorithm): Change “that it will have little effect at all.” to “that it will have little if any effect.”

page 235  lines +3 –4 (counting from the end of the Algorithm): Change “and this means the “gap,” determined by these two vectors, can never increase in size for separable data.” to “and this means that for separable data the “gap,” determined by these two vectors, can never increase in size.”

page 235  second and third lines after the section title 5.6 RELAXATION PROCEDURES: Change “in so-called “relaxation procedures” to include” to “in so-called “relaxation procedures,” to include”

page 235  first and second lines after Eq. 32: Change “misclassified by a, as do Jp, Jq focus attention” to “misclassified by a. Both Jp and Jq focus attention”

page 235  second line after Eq. 32: Change “Its chief difference is that its gradient” to “The chief difference is that the gradient of Jq is”

page 236  Algorithm 8, line 1: Change “k = 0” to “k ← 0”

page 236  Algorithm 8, line 4: Change “j = 0” to “j ← 0”

page 236  Algorithm 9, line 1: Change “k = 0” to “k ← 0”

page 238  line -5: Change “procedures, because” to “procedures because”

page 242  line -2: Change “We begin by writing Eq. 47” to “We begin by writing Eq. 45”

page 246  lines +1 – 2: Don’t split the equation by the line break

page 246  first (unnumbered) equation, second line: Change “(Ya_k - b)” to “(Ya(k) - b)”.

page 246  margin note: Change “LMS RULE” to “LMS RULE” (i.e., capitalize “LMS”)

page 246  Equation 61, second line: Change “(b_k - a(k)y_k)” to “(b(k) - a(k)y_k)”

page 246  Algorithm 10, line 1: Change “k = 0” to “k ← 0”

page 248  second line after Eq. 66: Change “obtain the MSE optimal” to “obtain the MSE-optimal”

page 250  Equation 79: Co-align vertically the “>” in the top equation with the “=” in the bottom equation

page 252  top equation: Change “= 0 =” to “= 0 =” (i.e., de-bold the “0”)
page 253  line -7: Change “requires that $e^+(k) = 0$ for” to “requires that $e^+(k) = 0$ for”

page 257  Figure 5.18, caption, line +2: Change “form $Au\beta$” to “form $Au = \beta$”

page 262  fourth line after Eq. 105: Change “with the largest margin” to “with the largest margin” (i.e., italicize “largest”)

page 263  fourth line after Eq. 107: Change “equation in Chapter 9,” to “topic in Chapter 9,”

page 266  Equation 114: Change “$a_i^I(k)y^k \leq a_j^I(k)y^k.$” to “$a_i^I(k)y^k \leq a_j^I(k)y^k.$”


page 270  line -13: Change “support vector machines” to “Support Vector Machines”

page 270  line -8: Change “support vector machines” to “Support Vector Machines”

page 271  Problem 2, line +3: Change “if $0 \leq \lambda \leq 1.$” to “for $0 \leq \lambda \leq 1.$”

page 272  Problem 8, line +2: Change “if $a_i^I y_i \geq 0$” to “if $a_i^I y_i \geq b$”

page 277  lines +2 – 3: Change “satisfies $z_k a_i^I y_k = 0$” to “satisfies $z_k a_i^I y_k = 1$”

page 277  Problem 38, line -2: Change “procedures Perceptron” to “procedures. Generalize the Perceptron”

page 278  Computer exercise 2, part (a): Change “Starting with $a = 0,$” to “Starting with $a = 0,$” (i.e., make bold the “0”)

page 279  Computer exercise 4, part (a), line +2: Change “and $\mu_1 = 0,$” to “and $\mu_1 = 0,$” (i.e., make bold the “0”)

Chapter 6

page 286  Equation 5: Change “$z_k = f(\text{net}_k).$” to “$z_k = f(\text{net}_k).$”

page 287  line +2: Change “all identical.” to “all the same.”

page 288  line -3: Change “depend on the” to “depends on the”

page 292  Equation 19, first line (inside brackets): Change “$1/2$” to “$\frac{1}{2}$” (i.e., typeset as a full fraction)

page 299  line -1: Change “and can be linearly separable” to “and are linearly separable”

page 306  sixth and seventh line after Eq. 33: Change “in a sum squared error sense” to “in a sum-squared-error sense”

page 295  line -5: Change “In addition to the use of the training set, here are” to “In addition to the use of the training set, there are”
ERRATA IN THE TEXT

page 307 lines +2 – 3: Change “been found useful” to “found to be useful”

page 307 line +6: Change “as continuity of f and its derivative” to “as continuity of f(·) and its derivative”

page 308 line +1b: Change “that is,” to “or is an “odd” function, that is,”

page 308 line +19: Change “values that ensure \( f'(0) \simeq 1 \)” to “values that ensure \( f'(0) \simeq 0.5 \)”

page 315 first line after Eq. 38: Change “reducing the criterion” to “reducing the error”

page 316 Figure 6.19 caption, line -1: Change “trained network.” to “trained network (red).”

page 318 line +8: Change “to compute is nonnegative” to “to compute, is nonnegative”

page 318 margin note: Change “MINKOWSKI ERROR” to “MINKOWSKI ERROR” (i.e., capitalize “M” in “MINKOWSKI”)

page 320 line between Eqs. 50 and 51: Change “The optimum change” to “Therefore, the optimum change”

page 319 Equation 48: Change last entry from \( f'(\text{net})y^nx^nH^x^d \) to \( f'(\text{net})f'(\text{net})y^nx^nH^x^d \)

page 322 Equation 56: Replace the current equation by

\[
\beta_m = \frac{\nabla J'(w(m))\nabla J(w(m))}{\nabla J'(w(m - 1))\nabla J(w(m - 1))}
\] (56)

page 322 Equation 57: Replace the current equation by

\[
\beta_m = \frac{\nabla J'(w(m))|\nabla J(w(m)) - \nabla J(w(m - 1))|}{\nabla J'(w(m - 1))\nabla J(w(m - 1))}
\] (57)

page 323 Fourth (unnumbered) equation on the page: Replace the left portion by

\[
\beta_1 = \frac{\nabla J'(w(1))\nabla J(w(1))}{\nabla J'(w(0))\nabla J(w(0))}
\]

page 323 Caption to bottom figure, line +2: Change “shown in the contour plot,” to “shown in the density plot,”

page 325 last line of the section Special Bases: Add “This is very closely related to model-dependent maximum-likelihood techniques we saw in Chapter 3.”

page 326 first line after Eq. 63: Change “of the filter in analogy” to “of the filter, in analogy”

page 326 line -5: Add a red margin note “TIME DELAY NEURAL NETWORK”

page 330 three lines above Eq. 67: Change “write the error as the sum” to “write the new error as the sum”
Chapter 7

page 332 Figure 6.28, caption, line +1: Change “a function of weights, \(J(w)\)” to “a function of weights, \(J(w)\),”

page 337 Problem 14, part (c), line +2: Change “the 2 \(\times\) 2 identity” to “the 2-by-2 identity”

page 339 Problem 22, line +4: Add “Are the discriminant functions independent?”

page 341 Problem 31, lines +1 – 2: Change “for a sum squared error criterion” to “for a sum-square-error criterion”

page 344 Computer exercise 2, line +2: Change “backpropagation to (Algorithm 1)” to “backpropagation (Algorithm 1)”

page 345 Computer exercise 7, line +2: Change “on a random problem.” to “on a two-dimensional two-category problem with \(2^k\) patterns chosen randomly from the unit square. Estimate \(k\) such that the expected error is \(25\%\). Discuss your results.

page 346 Computer exercise 10, part (c), line +1: Change “Use your network” to “Use your trained network”


page 348 column 2, entry for [43], line +4: Change “2000.” to “2001.”

Chapter 7

page 351 fourteenth line after Eq. 1: Change “of the magnets with the most stable configuration” to “of the magnets that is the most stable”

page 351 footnote, line -1: Change “in a range of problem domains.” to “in many problem domains.”

page 352 Figure 7.1, caption, line +7: Change “While our convention” to “While for neural nets our convention”

page 352 Figure 7.1, caption, line +8: Change “Boltzmann networks is” to “Boltzmann networks here is”

page 353 Figure 7.2, caption, line +3: Change “or “temperature” \(T\) to avoid” to “or “temperature” \(T\), to avoid”

page 357 Figure 7.4, caption, line +3: Change “values \(e^{-E_i/T}\)” to “values \(e^{E_i/T}\)”

page 360 first line in Section 7.3: Change “will use modify the” to “will use the”

page 360 second line in Section 7.3: Change “to specifically identify” to “and specifically identify”

page 360 second line in Section 7.3: Change “and other units as outputs” to “and others as outputs”
ERRATA IN THE TEXT

page 361 fourth line after Eq. 6: Change “possible hidden states.” to “possible hidden states consistent with $\alpha$.”

page 364 at the end of the body of the text: insert “One benefit of such stochastic learning is that if the final error seems to be unacceptably high, we can merely increase the temperature and anneal — we do not need to re-initialize the weights and re-start the full anneal.”

page 365 seventh line after the subsection Pattern Completion: Change “components of a partial pattern” to “components of the partial pattern”

page 367 line +1: Change “Recall, at the end” to “As mentioned, at the end”

page 373 fourteenth line in Section 7.5: Change “repeated for subsequent” to “repeated for the subsequent”

page 373 fifth line above the subsection Genetic Algorithms: Change “In both cases, a key” to “In both cases a key”

page 374 line +2: Change “used in the algorithm. Below” to “used in the algorithm; below”

page 374 line +3: Change “$P_{co}$ and $P_{mut}$, respectively, but first we present the general algorithm:” to “$P_{co}$ and $P_{mut}$.”

page 379 third line in the subsection Representation: Change “Here the syntactic” to “Whereas the syntactic”

page 381 third line under BIBLIOGRAPHICAL AND HISTORICAL REMARKS: Change “branch-and-bound, $A^*$” to “branch-and-bound and $A^*$”

page 382 line +28: Change “been fabricated as described” to “been fabricated, as described”

page 383 Problem 3, part (b), line +1: Change “The figure shows” to “That figure shows”

page 384 Problem 7, part (c), line +1: Change “magnets, total” to “magnets, the total”

page 384 Problem 8, part (b), equation: Change “= $P(s = +1)(+1) + P(s = -1)(-1).$” to “= Pr[$s = +1$](+1) + Pr[$s = -1$](–1).”

page 385 Problem 12, line +1: Change “Train a six-unit Hopfield network with the following three patterns using the” to “Determine the weights in a six-unit Hopfield network trained with the following three patterns. Use the”

page 385 Problem 15 line +1: Change “not be in a set of” to “not be in a subset of”

page 387 Problem 24: Change “crossover operator” to “crossover operator, and the multiplication operator, $\times$, and the addition operator, $\dagger$, can take two or more operands.”

page 393 column 1, entry for [54], line +3: Change “Evolutions-” to “Evolutions-”

page 393 column 2, entry for [62], line +2: Change “neurobiological system.” to “neurobiological systems.”
Chapter 8

page 403 line above Section 8.3.4 Pruning: Change “splitting is stopped.” to “splitting should be stopped.”

page 405 caption to figure, line -2: Change “marked * were instead slightly lower (marked †),” to “marked * were instead slightly lower (marked †),” (i.e., change the color of the special symbols to red)

page 414 line +16: Change “GACTG” to “GACTG” (i.e., eliminate space)

page 417 line -9: Change “The good-suffix function, \( \mathcal{G}(x) \), creates a table that for each suffix gives the location of its other occurrences in \( x \)” to “The good-suffix function, \( \mathcal{G}(x) \), creates a table that for each suffix gives the location of its second right-most occurrence in \( x \).”

page 418 fifth line before Algorithm 3: Change “consider interchanges.” to “consider the interchange operation.”

page 418 fourth line before Algorithm 3: Change “be an \( m \times n \) matrix” to “be an \( m \)-by-\( n \) matrix”

page 422 line -3: Change “specify how to transform” to “specifies how to transform”

page 426 line +1: Change “The grammar takes digit6 and” to “This grammar takes digit6 and”

page 428 line -1: Change “subsequent characters.” to “subsequent characters, or instead starting and the last (right) character in the sentence.”

page 429 caption to Figure 8.16, add after the last line: “Such finite-state machines are sometimes favored because of their clear interpretation and learning methods based on addition of nodes and links. In Section 8.7, though, we shall see general methods for grammatical learning that apply to a broader range of grammatical models.”

page 445 Problem 36, part (d), line +1: Change “Give a derivation” to “Attempt a derivation”

page 445 Problem 40, part (d), line +2: Change “either grammar as” to “either grammar, or can be parsed in both grammars, as”

Chapter 9

page 455 line -9 - 10: Change “a zero-one loss function, or, more generally, the cost for a general loss function \( L(\cdot, \cdot) \)” to “a zero-one or other loss function.”

page 462 line +12: Change “upon a specification method \( L \),” by “upon a specification method,”

page 462 line +13: Change “transmitted as \( y \), denoted \( L(y) = x \)” to “transmitted as \( y \) and decoded given some fixed method \( L \), denoted \( L(y) = x \).”
ERRATA IN THE TEXT

page 462 line +15–16: Change “denoted min L(y) = x; this minimal...[to end of paragraph]” to “denoted \( \min_{y: L(y) = x} \frac{|y|}{|y|} \)”.

page 462 line +17–18: Change “by analogy to entropy, where instead of a specification method \( L \) we consider” to “by analogy to communication, where instead of a fixed decoding method \( L \), we consider”.

page 462 line +25: Change “and so on. A universal description would” to “and so on. Such a description would”.

page 462 line +26: Change “different binary strings.” to “different binary strings \( x_1 \) and \( x_2 \).”

page 462 third and second line above Eq. 7: Change “the shortest program \( y \) (where the length” to “the shortest program string \( y \) (where \( y \)’s length”.

page 462 Equation 7: Replace the entire equation by:

\[
K(x) = \min_{y: \hat{U}(y) = x} |y|,
\]

page 463 line -10: Change “No Free Lunch Theorems.” to “No Free Lunch Theorem.”

page 472 Equation 23: Change “\( \frac{1}{(n-1)} \)” to “\( \frac{1}{n(n-1)} \)”.

page 472 Equation 24: Change “\( \sum_{j \neq i} \)” to “\( \frac{1}{n} \sum_{j \neq i} \)” (i.e., place the upper limit \( n \) over the summation sign).

page 472 line -4: Change “jackknife estimate of the variance” to “variance of the jackknife estimate”.

page 479 line +1: Change “in line 4” to “in line 5”.

page 485 line -10: Change “good estimates, because” to “good estimates because”.

page 488 Caption to Fig. 9.13, line +6: Change “approximated as a \( k \)-dimensional” to “approximated as a \( p \)-dimensional”.

page 488 line +10 (i.e., second line of the second paragraph of text): Change “is \( k \)-dimensional and the” to “is \( p \)-dimensional and the”.

page 496 Equation 54: Change “\( P(r|x, \eta_0) \)” to “\( P(r|x, \theta_0^0) \)”.

page 497 Equation 58: Change \( \mu_r \) to \( \theta_r \) in two places.

page 501 line -15: Change “and learning algorithm was first described” to “and learning algorithm were first described.”

page 502 Problem 6, last line: Change “this way” to “this sense”.

page 504 Problem 20, line -2: Change “denoted \( p(g(x; D)) \) is a” to “denoted \( p(g(x; D)) \), is a”.

page 508 Problem 45, line +1: Change “mixture of experts classifier” to “mixture-of-experts classifier”.

page 512 line -3: Add “Now what is the training error measured using \( \omega_A \) and \( \omega_B \)”.

Chapter 10

page 524 The second to last equation: Increase the size of the final bracket, “]” to match its mate

page 525 line following Eq. 23: Change “results in Eq. 12” to “results into Eq. 12”

page 529 line +5: Change “\( \hat{P}(w_j) \)” to “\( \hat{P}(\omega_j) \)”

page 529 Algorithm 2: Change “(Fuzzy k-Means Clustering)” to “(Fuzzy \( k \)-Means Clustering)”

page 534 line +2: Change “overlap, thus” to “overlap; thus”

page 535 First equation: move the second, third, and fourth lines so as to co-align vertically the corresponding terms

page 536 line +19: Change “classification analog of Chapter 3” to “classification case in Chapter 3”

page 537 line -20: Change “distributed, these statistics” to “distributed these statistics”

page 571 line +1 - +2: Change “microphones” to “micro-phones” (i.e., re-hyphenate)

page 578 Figure 10.31, caption, line +2: Change “space that leads maximally” to “space that maximally”

page 579 Figure 10.32, caption, line -1: Change “to this center region” to “to this central region”

page 580 line +15: Change “cluster centers being used to” to “cluster centers being used to” (i.e., italicize “cluster centers”)

page 580 line +16: Change “with combined features being” to “with combined features being” (i.e., italicize “combined features”)

page 582 four lines above BIBLIOGRAPHICAL and HISTORICAL REMARKS: Change “between points that, too, seeks to preserve neighborhoods” to “between points that preserves neighborhoods”

page 583 lines +9 - 10: Change “The classificatory foundations of biology, cladistics (from the Greek \( klados \), branch) provide useful” to “Cladistics, the classificatory foundation of biology (from the Greek \( klados \), branch), provides useful”

page 583 line +18: Change “analysis, and explained the very close” to “analysis, and in reference [36] explained the very close”

page 583 line +19: Change “information maximization in reference [36].” to “information maximization.”

page 584 Problem 2, equation: Change “\( (1 - |x - \mu_1|)/(2w_j) \)” to “\( (w_i - |x - \mu_i|)/w_i^2 \)”

page 585 Problem 6, line +1: Change “Consider a \( c \) component” to “Consider a \( c \)-component”
ERRATA IN THE TEXT

page 587 Problem 13, line +3: Change “that for any observed $x$, all but one” to “that for any observed $x$ all but one”

page 591 Problem 41, line +1: Change “null hypothesis in associated” to “null hypothesis associated”

page 592 Problem 44, part (c), line +3: Change “$(\delta e)^T \Sigma (\delta e) e = 0$” to “$(\delta e)^T \Sigma (\delta e) e = 0$”

page 592 Problem 46, line +1: Change “principal componet analysis” to “principal component analysis”

page 593 Problem 48, line +3: Change “linearity is given by” to “linearity is the one given by”

page 596 Problem 11, lines +3 – 4: Change “to the date in the table above using the distance measure indicated” to “to the data in the table above using the distance measures indicated”

page 597 Problem 13, line +1: Change “a basic competitive learning” to “a basic Competitive Learning”

page 597 Problem 13, lines +4 – 5: Change “hy-persphere” to “hyper-sphere” (i.e., re-hyphenate “hypersphere”)

Appendix

page 609 first line after Eq. 26: Change “the $i, j$ cofactor or” to “the $i, j$ cofactor or” (i.e., eliminate the space in “$i, j$”) 

page 609 second line after Eq. 26: Change “is the $(d - 1) \times (d - 1)$ matrix” to “is the $(d - 1)$-by-$(d - 1)$ matrix”

page 609 fourth line after Eq. 26: Change “whose $i, j$ entry is the $j, i$ cofactor” to “whose $i, j$ entry is the $j, i$ cofactor” (i.e., eliminate the space in “$i, j$” and “$j, i$”)

page 609 line -2: Add “The inverse of the product of two square matrices obeys $[MN]^{-1} = N^{-1}M^{-1}$, as can be verified by multiplying on the right or the left by $MN$."

page 615 line -12 (i.e., just before Section A.4.7): Change “and $\frac{n_{11}/n}{n_{01} + n_{11}/n}$ is approximately” to “and $(n_{01} + n_{11})/n$ is approximately”

page 624 Figure A.3, caption, line +2: Change “between $-\sqrt{2}u$ and $\sqrt{2}u$, that is” to “between $-\sqrt{2}u$ and $\sqrt{2}u$; that is”

page 629 line +9: Change “from a distribution” to “from a standarized Gaussian distribution”

page 631 line +9: Change “equally likely is” to “equally likely, is”

page 631 line +12: Change “outcome and $H = \log_2 2^3 = 3$” to “outcome and $H = -\sum_{i=0}^{7} \frac{1}{2} \log_2 2^3 = \log_2 2^3 = 3$”
page 631  Equation 118: Change “ln p(x)” to “ln p(x)” (i.e., reduce the space between “ln” and “p(x)”)
page 631  first line after Eq. 119: Change “For this Dirac function” to “For this Dirac density”
page 632  Red margin note: Change “KULLBACK-LEIBLER distance” to “KULLBACK-LEIBLER DISTANCE” (i.e., capitalize the “L” in “LEIBLER”)
page 632  line -2: Change is always larger than” to “is never smaller than”
page 634  line +3: Change “f(x) ≤ c_0 g(x) for all” to “f(x) ≤ c g(x) for all”
page 634  line +7: Change “proper choice of c_0 and x_0” to “proper choice of c and x_0”

Index

page 644  column 1, line -14: Insert “Gini impurity, 399, 401”
page 653  column 1, line -3: Change “multi-variate” to “multivariate”