List of corrections

1. page 50, Figure 4.5: Remove arrows from the ends of the two lines.

2. page 64, In the first line after the first mathematical display the word "by" is duplicated. Remove one.

3. page 70, line -4: write

$$= \mathbf{E}\left\{\sum_{j=0}^{k}\cdots\right.$$

   (i.e., replace "1" by "0").

4. page 94, line -2: Put: "By conditioning on the random variable $N(x)$, it is easy to see ..."

5. page 166, add the following sentence to the end of Problem 10.22:
   "Prove that the rule is weakly consistent for all $r \le d$ whenever X has a density."

6. page 180, Figure 11.2: Remove arrows from the ends of all lines.

7. page 191, in the fourth line of the second paragraph replace "points from $X_1, \ldots, X_{2d}$" by "points from $X_1, \ldots, X_n$"

8. page 206, line 10: replace expression by the following:

$$\mathbf{P}\left\{L(\phi_n^*) - \inf_{\phi\in\mathcal{C}} L(\phi) > \epsilon\right\} \le 8\mathcal{S}(\mathcal{C}, n)e^{-n\epsilon^2/8(L+\epsilon)}.$$

9. page 210, line 10: replace "(Lugosi (1995))" by "(Shawe-Taylor, Anthony, and Biggs (1993))".

10. page 211, line -15, and page 212, line 1: replace "Buescher and Kumar (1994a)" by "Buescher and Kumar (1996a)".

11. page 236, lines -4, -3:
    If on the other hand $n \ge 15$ and $n \le (V-1)/(12\epsilon)$, then

$$\sup_{(X,Y)\in\mathcal{X}} \mathbf{P}\{L_n \ge \epsilon\} \ge \frac{1}{10}.$$

12. page 237, line 15: replace expression by

$$= \frac{1}{2}\mathbf{P}\{X \neq X_1, \ldots, X \neq X_n, X \neq x_V | X_1, \ldots, X_n\}$$

13. page 237, line -6: Replace the first two sentences of this paragraph by:
CHOICE A. Take $p = 1/n$, and assume $12n\epsilon \leq V-1$, $\epsilon < 1/2$. Note that for $n \geq 15$, $\mathbf{E}|J| = (V-1)(1-p)^n \geq (V-1)/3$. Also ...

14. pages 237–238: replace the chain of inequalities starting on page 237 and ending on page 238 by the following:

$$\begin{aligned}
\mathbf{P}\{|J| \geq 2n\epsilon\} &= 1 - \mathbf{P}\{|J| < 2n\epsilon\} \\
&\geq 1 - \mathbf{P}\{|J| < (V-1)/6\} \\
&= 1 - \mathbf{P}\{|J| - \mathbf{E}|J| \leq -(V-1)/6\} \\
&\geq 1 - \frac{\mathbf{Var}|J|}{\mathbf{Var}|J| + (V-1)^2/36} \\
&\geq 1 - \frac{(V-1)^2/4}{(V-1)^2/4 + (V-1)^2/36} \\
&= \frac{1}{10}.
\end{aligned}$$

15. page 246, lines 6,7: Replace these two lines by
Finally, for $\delta \leq 1/10$ *and* $\epsilon < 1/2$,

$$N(\epsilon, \delta) \geq \frac{V-1}{12\epsilon}.$$

16. page 292, lines 13 and 19, and page 301, line 3: replace "Lugosi and Zeger (1994)" by "Lugosi and Zeger (1996)".

17. page 298, line 1: replace "Buescher and Kumar (1994b)" by "Buescher and Kumar (1996b)".

18. page 320, Figure 20.9: remove arrow from the left end of horizontal line.

19. page 324, Figure 20.12: remove the three empty-headed arrows.

20. page 329, at the beginning of the first and sixth lines after the first mathematical display, replace $\mathbf{P}\{E_2 \cup E_5 \cup E_6\} \to \infty$ by $\mathbf{P}\{E_2 \cup E_5 \cup E_6\} \to 1$.

21. page 339, Figure 20.19: remove arrow from the upper right end of line.

22. page 348, Figure 20.22: remove all arrows from the ends of the lines in the right half of the picture.

23. page 364, lines 8 and -5; page 365, line -4; page 369, line 1; page 378, line 1 (under the figure); and page 383, line 9: replace "Lugosi and Nobel (1993)" by "Lugosi and Nobel (1996)".

24. page 393, line 1: Put "Theorem 21.5" instead of "Theorem 21.7".

25. page 485, line 21: replace "Buescher and Kumar (1994a)" by "Buescher and Kumar (1996a)".

26. page 538, line -4: replace the first $-$ sign in the expression by $+$.

27. page 589, line 16: replace the last mathematical display in Section A.9 by

$$\mathbf{P}\left\{\left|X - \frac{bn}{N}\right| \geq n\epsilon\right\} \leq 2e^{-2n\epsilon^2}.$$

28. page 596, line 19: update reference to:
Buescher, K. and Kumar, P. (1996a). Learning by canonical smooth estimation, Part I: Simultaneous estimation. *IEEE Transactions on Automatic Control*, **41**:545–556.

29. page 596, line 21: update reference to:
Buescher, K. and Kumar, P. (1996b). Learning by canonical smooth estimation, Part II: Learning and choice of model complexity. *IEEE Transactions on Automatic Control*, **41**:557–569.

30. page 599, line -6: update reference to:
Devroye, L. and Lugosi, G. (1995). Lower bounds in pattern recognition and learning. *Pattern Recognition*, **28**:1011–1018.

31. page 609, line 18: update reference to:
Lugosi, G. and Nobel, A. (1996). Consistency of data-driven histogram methods for density estimation and classification. *Annals of Statistics*, **24**:687–706.

32. page 609, line 24: update reference to:
Lugosi, G. and Zeger, K. (1996). Concept learning using complexity regularization. *IEEE Transactions on Information Theory*, **42**:48–54

33. page 614, line -3: add the following reference:
Shawe-Taylor, J., Anthony, M. and Biggs, N.L. (1993). Bounding sample size with the Vapnik-Chervonenkis dimension. *Discrete Applied Mathematics*, **42**:65–73.

34. page 615, line -16: At the end of reference Stone (1977), replace **8**:1348–1360 by **5**:595–645