Below are errata for the first and second printings of *Machine Learning*, Tom M. Mitchell, McGraw Hill, 1997.

I am grateful to Ethem Alpaydin, Jonathan Betz, Horacio Carvajal, Pedro Domingos, Olac Fuentes, Haym Hirsh, Ray Mooney, James Reggia, Roni Rosenfeld, Stephen Scott, Nikunj Oza, Tonguc Unluyurt, Ben Wegbreit and Onno Zoeter for pointing out these corrections.

Page 10, line 13

looses the game ...

should be

loses the game ...

Page 18, exercise 1.4

... Experiment Generator module of Figure 1.2

should be

... Experiment Generator module of Figure 1.1

Page 61, Figure 3.4:

$$Gain(S_{sunny}, Humidity) = .970 - (3/5)0.0 - (2/5)0.0 + .970$$

should be

$$Gain(S_{sunny}, Humidity) = .970 - (3/5)0.0 - (2/5)0.0 = .970$$

Page 66, eighth line from bottom:

... handing attributes with differing costs ...

should be

... handling attributes with differing costs ...

Page 68, last sentence of first full paragraph:

... we expect h to outperform h' over subsequent data ...

should be

... we expect h' to outperform h over subsequent data ...

Page 92, third line of text:

... in terms of the linear unit inputs x_{id}, outputs t_d, and ...

should be

... in terms of the linear unit inputs x_{id}, outputs o_d, and ...

Page 108, final paragraph of section 4.6.4

This plots displays ...

should be

This plot displays ...

Page 123, exercise 4.11

http://www.cs.cmu.edu/~tom/book.html

should be

http://www.cs.cmu.edu/~tom/mlbook.html

Page 139, Table 5.4
change "gaussian" to "Gaussian" (in first line of figure, and in caption)

Page 188, in left side of equation 6.25

$$\frac{\partial \ln P(D|h)}{\partial w_{ij}}$$

should be

$$\frac{\partial \ln P(D|h)}{\partial w_{ijk}}$$

Page 193, bottommost equation.

$$\mu_j \leftarrow \frac{1}{m} \sum_{i=1}^{m} E[z_{ij}] \quad x_i$$

should be

$$\mu_j \leftarrow \frac{\sum_{i=1}^{m} E[z_{ij}] \quad x_i}{\sum_{i=1}^{m} E[z_{ij}]}$$

Page 196. Equation (6.31) should be

$$\mu_j \leftarrow \frac{\sum_{i=1}^{m} E[z_{ij}] \quad x_i}{\sum_{i=1}^{m} E[z_{ij}]}$$

Page 199, add in the reference:

Haussler, D., Kearns, M., and Schapire, R.E. (1994). Bounds on the sample complexity of Bayesian learning using information theory and the VC dimension. *Machine Learning*, 14, pg. 83.

Page 200, reference to Opper and Haussler:

... learning a perception. *Physics Review Letters* ...

should be

... learning a perceptron. *Physical Review Letters*

Page 213, first paragraph of section 7.3.3.2. Begining at line 6, should read:

it is easy to show that $|H|$ is at most 3^{nk} (because there are k terms, each of which may take on 3^n possible values). Note 3^{nk} is an overestimate of $|H|$, because ..."

and equation (7.6) on page 213 should read:

$$m \geq \frac{1}{\epsilon}(nk \ln 3 + \ln(1/\delta))$$

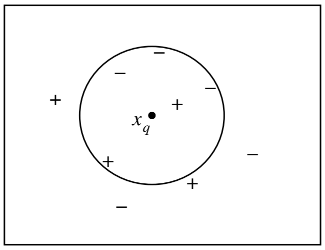
Page 213, tenth line from bottom:

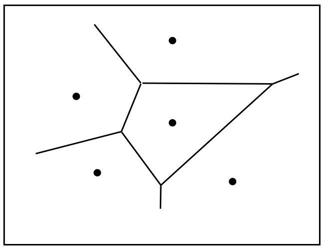
... (unless $RP \neq NP$)

should be

... (unless $RP = NP$)

Page 233, figure 8.1. The figure should be modified as follows:





Page 239, equation at top of page should have minus sign in exponent, as follows:

$$K_u(d(x_u, x)) = e^{-\frac{1}{2\sigma_u^2} d^2(x_u, x)}$$

Page 247, exercise 8.1

... given by Equation (8.1)

should be

... given by Equation (8.7)

Page 255, third line from bottom:

This method is sometimes often called ...

should be

This method is sometimes called ...

Page 259,

Crowding is a phenomena ...

should be

Crowding is a phenomenon ...

Page 318, Table 11.3:

Volume(x,vs)

should be (in both places where it appears)

Volume(x,vx)

Page 361, caption of Figure 12.9

... at each step, FOIL also considers ...

should be

... at each step, FOCL also considers ...

Page 373, sixteenth line from bottom:

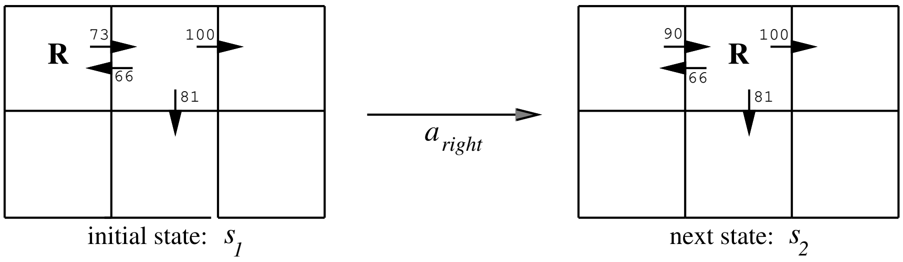
... to its actions, then it can then use ...

should be

... to its actions, it can then use ...

Page 376, Figure 13.3

This figure has incorrect numerical values, and should be changed to:



$$\begin{aligned}\hat{Q}(s_1, a_{right}) &\leftarrow r + \gamma \max_{a'} \hat{Q}(s_2, a') \\ &\leftarrow 0 + 0.9\ \max\{66, 81, 100\} \\ &\leftarrow 90\end{aligned}$$

Page 376, caption of Figure 13.3

For example, the value $\hat{Q}(s_1, a_{right}) = 72$, where

should be

For example, the value $\hat{Q}(s_1, a_{right}) = 72.9$, where

---

Page 382, equation (13.10) should have a γ added. That is,

$$\hat{Q}_n(s,a) \leftarrow (1-\alpha_n)\hat{Q}_{n-1}(s,a) + \alpha_n[r + \max_{a'} \hat{Q}_{n-1}(s',a')]$$

should be

$$\hat{Q}_n(s,a) \leftarrow (1-\alpha_n)\hat{Q}_{n-1}(s,a) + \alpha_n[r + \gamma \max_{a'} \hat{Q}_{n-1}(s',a')]$$

Page 397, author index

Murthy, P.M.

should be

Murthy, S.K.

Page 402, subject index

CIGOL, 302

should be

CIGOL, 293, 302

Back cover of book:

... example date sets, ...

should be

... example data sets, ...
