Errata Sheet

Last Updated: 9/13/2023

Chapter 1

- On page 20, in the paragraph preceding Section 1.2.3, the word "statistics" should be replaced with "statistic."

Chapter 3

- On page 35, in the beginning of Section 3.1.3, all λ's should be replaced with η.
- On page 42, in the sidenote labelled 8, the video link should now point to `https://www.youtube.com/watch?v=QjOILAQOEFg`.

Errata Sheet for Previous Versions

Applicable to versions before 2/4/2023

Chapter 3

- On page 42 and on page 43, the following Python code

```
ax = plt.figure(figsize=(12, 10)).gca(projection='3d')
```

should be rewritten as

```
fig = plt.figure(figsize=(12,10))
ax = fig.add_subplot(projection='3d')
```

The *gca*() function is now deprecated.

Chapter 4

- On page 53, the gradient of the SVM loss near Equation (4.9) should be written as

$$\nabla \ell = \begin{cases} -y\vec{\mathbf{x}} & y\vec{\mathbf{w}} \cdot \vec{\mathbf{x}} < 1 \\ 0 & y\vec{\mathbf{w}} \cdot \vec{\mathbf{x}} > 1 \end{cases}$$

- On page 54, in Figure 4.4, the lines of boundary should be labeled $\vec{\mathbf{w}} \cdot \vec{\mathbf{x}} = -1, 0, +1$ (left to right).

Chapter 8

- On page 101, the phrase "an loss" should be written as "a loss."

Chapter 11

- On page 137, in Definition 11.4.4, Jacobian matrix should be of size $J(\vec{\mathbf{z}}, \vec{\mathbf{x}}) \in \mathbb{R}^{\ell \times n}$
- On page 137, the equation in Example 11.4.5 should be written as

$$J(\vec{\mathbf{z}}, \vec{\mathbf{x}}) = J(\vec{\mathbf{z}}, \vec{\mathbf{y}}) J(\vec{\mathbf{y}}, \vec{\mathbf{x}}) = diag(\mathbb{1}(\mathbf{A}\vec{\mathbf{x}} > 0))\mathbf{A}$$

Chapter 15

- On page 197, the equation after (15.3) should be updated as

$$Q'_t = r_t + \gamma \max_b Q(s_{t+1}, b)$$

and similarly in page 198, the equation in the Example 15.3.3 should be updated as

$$Q'_t = r_t + \gamma \max_b Q(s_{t+1}, b) = 1$$

Chapter 19

- On page 242, the formula in Definition 19.2.5 should be written as

$$\left.\frac{\partial y_j}{\partial x_i}\right|_{\vec{\mathbf{x}}} = \lim_{h \to 0} \frac{f_j(x_1, \ldots, x_i + h, \ldots, x_n) - f_j(x_1, \ldots, x_i, \ldots, x_n)}{h}$$