<u>Chapter 3 Section 2</u>

#: 2.1, 2.11, 2.12, 3.1

---

**Exercise 2.1a.** *Set $\delta = 0.03$ and let $\epsilon(M, N, \delta) = \sqrt{\frac{1}{2N} ln \frac{2M}{\delta}}$. For $M = 1$, how many examples do we need to make $\epsilon \leq 0.05$*

*Proof.*

$$E_{out}(g) \Leftarrow E_{in}(g) + \sqrt{\left( \frac{1}{2N} ln \frac{2M}{\delta} \right)}$$

$$\epsilon(M, N, \delta) = \sqrt{\left( \frac{1}{2N} ln \frac{2M}{\delta} \right)} \Leftarrow 00.5$$

$$= \sqrt{\left( \frac{1}{2N} ln \frac{2}{0.03} \right)} \Leftarrow 0.05$$

$$= \sqrt{\left( \frac{1}{2N} \cdot 4.2 \right)} \Leftarrow 0.05$$

$$= \left( \sqrt{\left( \frac{1}{2N} \cdot 4.2 \right)} \right)^2 \Leftarrow 0.05^2$$

$$= \left( \frac{1}{2N} \cdot 4.2 \right) \Leftarrow 0.0025$$

$$= 4.2 \Leftarrow 0.005N$$

$$= N \geq 840 \quad samples$$

$\square$

**Exercise 2.1b.** *Set $\delta = 0.03$ and let $\epsilon(M, N, \delta) = \sqrt{\frac{1}{2N} ln \frac{2M}{\delta}}$. For $M = 100$, how many examples do we need to make $\epsilon \leq 0.05$*

*Proof.*

$$E_{out}(g) \Leftarrow E_{in}(g) + \sqrt{\left(\frac{1}{2N} ln \frac{2M}{\delta}\right)}$$

$$\epsilon(M, N, \delta) = \sqrt{\left(\frac{1}{2N} ln \frac{2M}{\delta}\right)} \Leftarrow 00.5$$

$$= \sqrt{\left(\frac{1}{2N} ln \frac{200}{0.03}\right)} \Leftarrow 0.05$$

$$= \sqrt{\left(\frac{1}{2N} \cdot 8.8\right)} \Leftarrow 0.05$$

$$= \left(\sqrt{\left(\frac{1}{2N} \cdot 8.8\right)}\right)^2 \Leftarrow 0.05^2$$

$$= \left(\frac{1}{2N} \cdot 8.8\right) \Leftarrow 0.0025$$

$$= 8.8 \Leftarrow 0.005N$$

$$= N \geq 1760 \ \ samples$$

$\square$

**Exercise 2.1c.** *Set $\delta = 0.03$ and let $\epsilon(M, N, \delta) = \sqrt{\frac{1}{2N} ln \frac{2M}{\delta}}$. For $M = 10,000$, how many examples do we need to make $\epsilon \leq 0.05$*

*Proof.*

$$E_{out}(g) \Leftarrow E_{in}(g) + \sqrt{\left(\frac{1}{2N} ln \frac{2M}{\delta}\right)}$$

$$\epsilon(M, N, \delta) = \sqrt{\left(\frac{1}{2N} ln \frac{2M}{\delta}\right)} \Leftarrow 00.5$$

$$= \sqrt{\left(\frac{1}{2N} ln \frac{20000}{0.03}\right)} \Leftarrow 0.05$$

$$= \sqrt{\left(\frac{1}{2N} \cdot 13.4\right)} \Leftarrow 0.05$$

$$= \left(\sqrt{\left(\frac{1}{2N} \cdot 13.4\right)}\right)^2 \Leftarrow 0.05^2$$

$$= \left(\frac{1}{2N} \cdot 13.4\right) \Leftarrow 0.0025$$

$$= 13.4 \Leftarrow 0.005N$$

$$= N \geq 2680 \quad samples$$

$\square$

**Exercise 2.11a.** *Suppose $m_{\mathcal{H}}(N) = N+1$, so $d_{vc} = 1$. You have 100 training examples. Use the generalization bound to give a bound for $E_{out}$ with confidence 90%.*

*Proof.*

$$E_{out}(g) \leq E_{in}(g) + \sqrt{\left(\frac{8}{N}ln\frac{4_{\mathcal{H}}(2N)}{8}\right)}$$

$$= E_{in}(g) + \sqrt{\frac{8}{100}ln\frac{4(2N+1)}{0.1}}$$

$$= E_{in}(g) + \sqrt{0.08 \cdot \frac{8N+4}{0.1}}$$

$$= E_{in}(g) + \sqrt{0.08 \cdot \frac{804}{0.1}}$$

$$= E_{in}(g) + \sqrt{0.08 \cdot ln8040}$$

$$= E_{in}(g) + \sqrt{0.08 \cdot 9}$$

$$E_{out}(g) \leq E_{in}(g) + \sqrt{0.72}$$

$$E_{out}(g) \leq E_{in}(g) + 0.85$$

$\square$

**Exercise 2.11b.** *Suppose $m_{\mathcal{H}}(N) = N + 1$, so $d_{vc} = 1$. You have $10,000$ training examples. Use the generalization bound to give a bound for $E_{out}$ with confidence 90%.*

*Proof.*

$$E_{in}(g) + \sqrt{\frac{8}{10000} \cdot \frac{ln80004}{0.1}}$$

$$E_{in}(g) + \sqrt{0.0008 \cdot 13.9}$$

$$E_{out}(g) \leq E_{in}(g) + \sqrt{0.012}$$

$$E_{out}(g) \leq E_{in}(g) + 0.1$$

$\square$

**Exercise 2.12.** *For an $\mathcal{H}$ with a $d_{vc} = 10$, what sample size do you need (as prescribed by the generalization bound) to have 95% confidence that your generalization error is at most 0.05?*

*Proof.*

$$N \geq \frac{8}{\varepsilon^2} ln(\frac{4((2N)^{10} + 1)}{\delta})$$

$$\geq \frac{8}{0.05^2} ln(\frac{4((2N)^{10} + 1)}{0.05})$$

$$\geq \frac{8}{0.025} (ln(4(1024N^{10} + 1)) - ln(0.05))$$

$$\geq 3200(ln(4096) + 10ln(N) + ln(4) - ln(0.05))$$

$$\geq 3200(ln(\frac{4096(4)}{0.05}) + 10ln(N))$$

$$\geq 3200(12.7 + 10ln(N))$$

$$\geq 40639 + 32000ln(N)$$

The rest of this equation was solved using a graphing calculator to graph the following equations $f(x) = 12.7 + ln(x)$ and $g(x) = \frac{x}{32000}$. The result of which equal to 457729.046     □
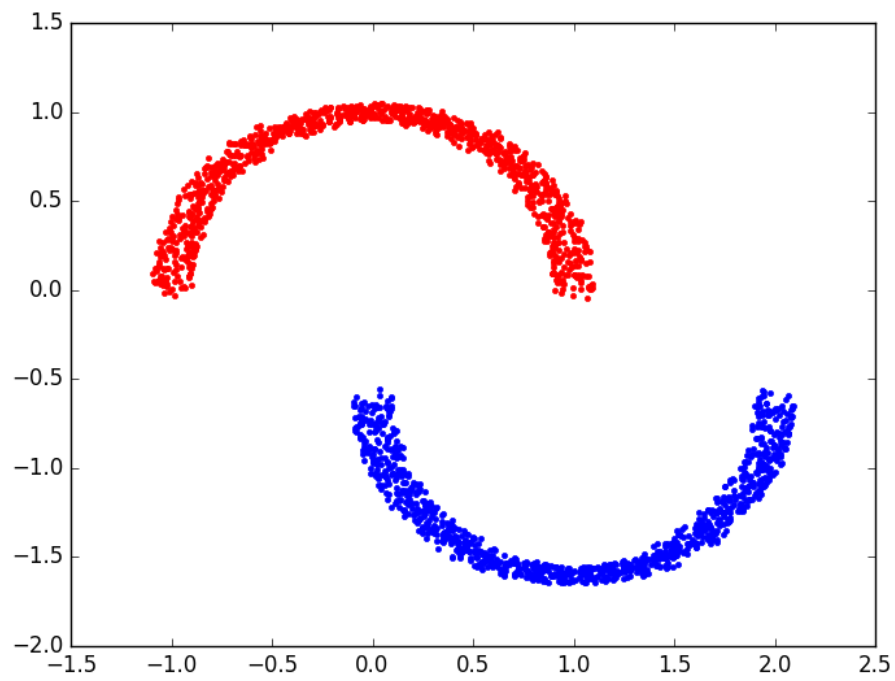
---

**Exercise 3.1.** *Run the PLA starting from $w = 0$ until it converges. Plot the data and the final hypothesis.*

*Proof.* (See next page)     □

---

3.1a



3.1b
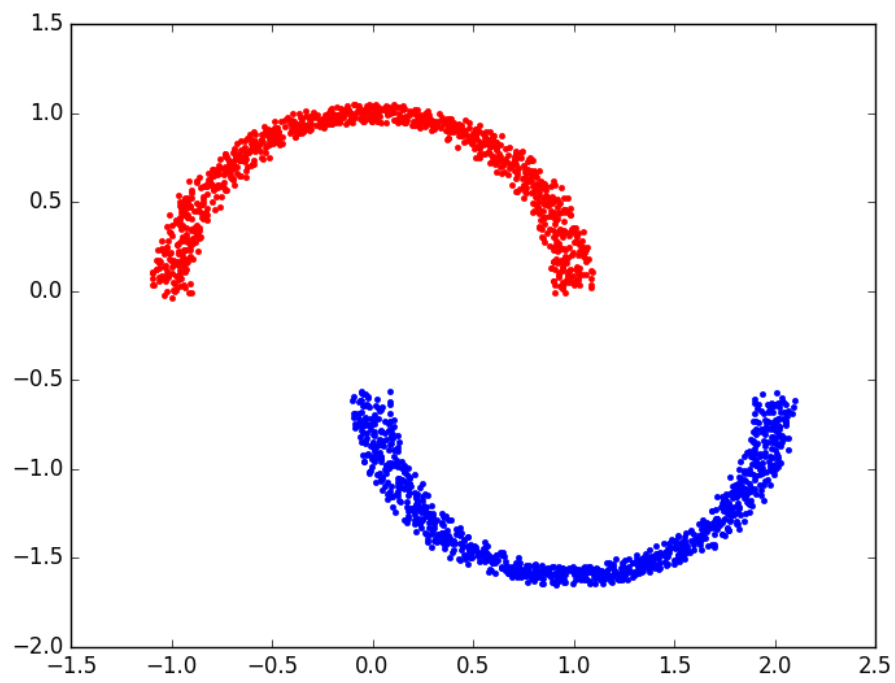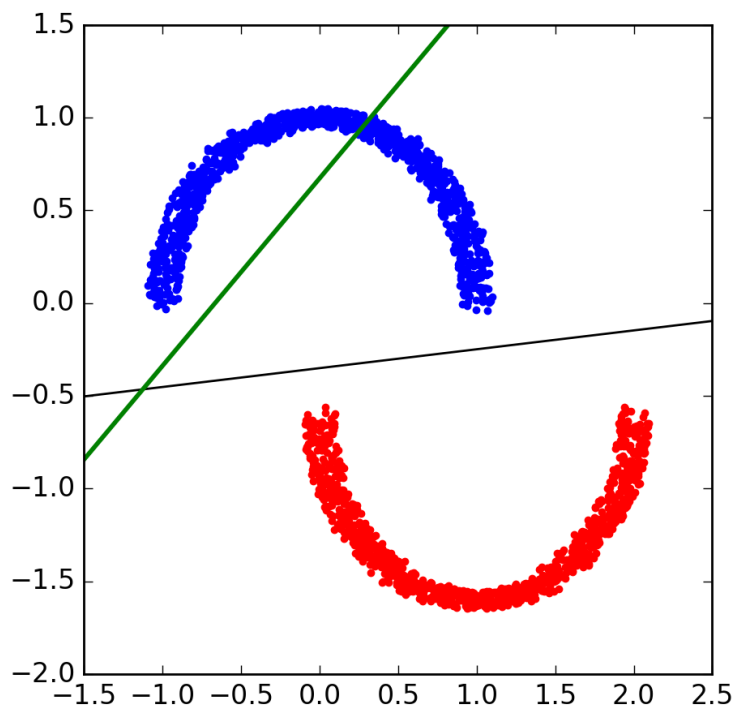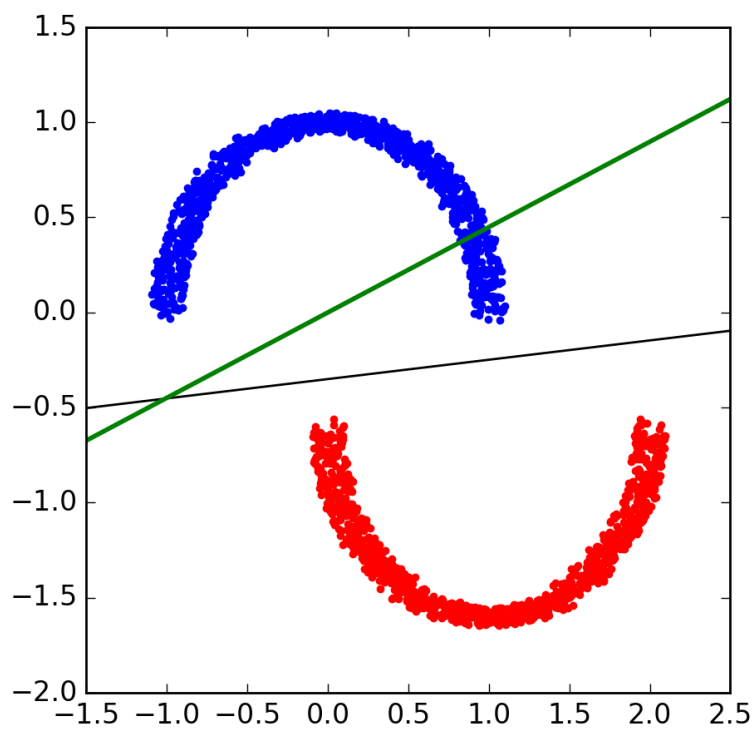
N = 2000, Iteration 1



N = 2000, Iteration 2

N = 2000, Iteration 3