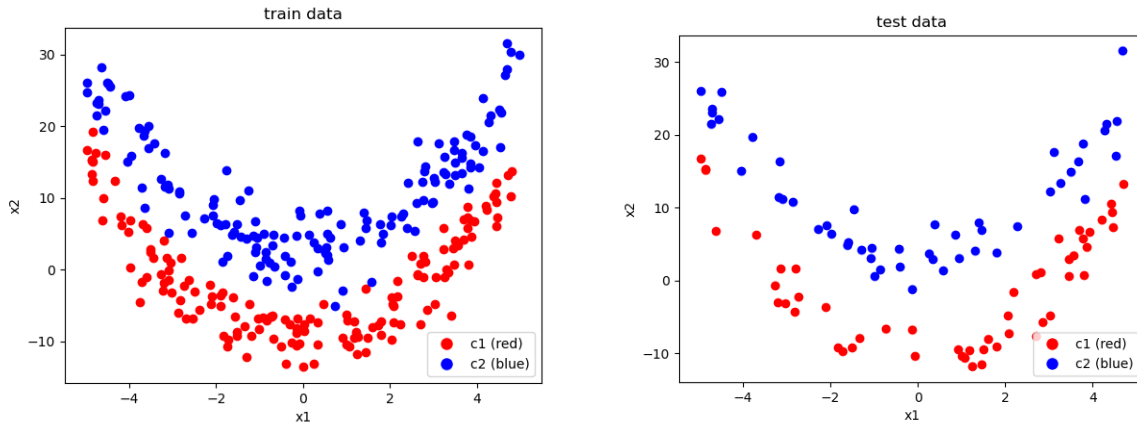


## گزارش تمرین پیاده سازی سوم درس هوش محاسباتی

1.

نمودارهای مربوط به داده‌های آموزشی و داده‌های تست در ادامه آورده شده است:



2.

b.

داریم:

$$\frac{\partial J}{\partial b} = \frac{\partial J}{\partial a} * \frac{\partial a}{\partial (x_0 w_0 + x_1 w_1 + b)} * \frac{\partial (x_0 w_0 + x_1 w_1 + b)}{\partial b}$$

$$\Rightarrow \frac{\partial J}{\partial b} = 2(y - a) * \sigma'(x_0 w_0 + x_1 w_1 + b) * 1$$

$$\frac{\partial J}{\partial w_0} = \frac{\partial J}{\partial a} * \frac{\partial a}{\partial (x_0 w_0 + x_1 w_1 + b)} * \frac{\partial (x_0 w_0 + x_1 w_1 + b)}{\partial w_0}$$

$$\Rightarrow \frac{\partial J}{\partial w_0} = 2(y - a) * \sigma'(x_0 w_0 + x_1 w_1 + b) * x_0$$

$$\frac{\partial J}{\partial w_1} = \frac{\partial J}{\partial a} * \frac{\partial a}{\partial (x_0 w_0 + x_1 w_1 + b)} * \frac{\partial (x_0 w_0 + x_1 w_1 + b)}{\partial w_1}$$

$$\Rightarrow \frac{\partial J}{\partial w_1} = 2(y - a) * \sigma'(x_0 w_0 + x_1 w_1 + b) * x_1$$

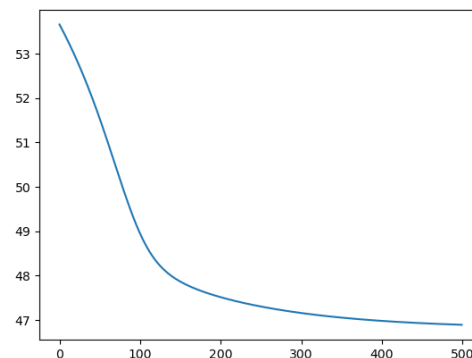
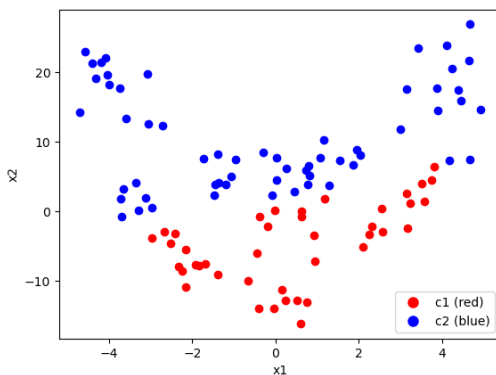
عبارات حاصل با در نظر گرفتن  $\sigma'(x) = \sigma(x) * (1 - \sigma(x))$ ، ساده تر نیز خواهند شد.

c.

تعداد گام‌ها و نرخ آموزش مناسب برای این مدل شبکه، به ترتیب 500 گام و نرخ آموزش 0.2 می‌باشد.

d.

نمونه‌ای از خروجی شبکه بر روی داده‌های تست به همراه نمودار میزان خطا برحسب تعداد گام در ادامه آورده شده است:



دقت این مدل شبکه، 83% برآورد شده است.

با توجه به اینکه شبکه دارای لایه پنهان نمی‌باشد، مدلی ساده است و به نوعی رگرسیون خطی (linear regression) انجام می‌دهد و بهترین خط ممکن را به عنوان مرز کلاس‌ها مشخص می‌کند.

3.

b.

داریم:

$$\begin{aligned}\frac{\partial J}{\partial b_2} &= \frac{\partial J}{\partial a} * \frac{\partial a}{\partial (a_0 v_0 + a_1 v_1 + b_2)} * \frac{\partial (a_0 v_0 + a_1 v_1 + b_2)}{\partial b_2} \\ \Rightarrow \frac{\partial J}{\partial b} &= 2(y - a) * \sigma'(a_0 v_0 + a_1 v_1 + b_2) * 1\end{aligned}$$

$$\begin{aligned}\frac{\partial J}{\partial v_1} &= \frac{\partial J}{\partial a_2} * \frac{\partial a_2}{\partial (a_0 v_0 + a_1 v_1 + b_2)} * \frac{\partial (a_0 v_0 + a_1 v_1 + b_2)}{\partial v_1} \\ \Rightarrow \frac{\partial J}{\partial v_1} &= 2(y - a_2) * \sigma'(a_0 v_0 + a_1 v_1 + b_2) * a_1\end{aligned}$$

$$\begin{aligned}\frac{\partial J}{\partial v_0} &= \frac{\partial J}{\partial a_2} * \frac{\partial a_2}{\partial (a_0 v_0 + a_1 v_1 + b_2)} * \frac{\partial (a_0 v_0 + a_1 v_1 + b_2)}{\partial v_0} \\ \Rightarrow \frac{\partial J}{\partial v_0} &= 2(y - a_2) * \sigma'(a_0 v_0 + a_1 v_1 + b_2) * a_0\end{aligned}$$

$$\begin{aligned}\frac{\partial J}{\partial u_0} &= \frac{\partial J}{\partial a_2} * \frac{\partial a_2}{\partial(a_0 v_0 + a_1 v_1 + b_2)} * \frac{\partial(a_0 v_0 + a_1 v_1 + b_2)}{\partial a_1} * \frac{\partial a_1}{\partial(x_0 u_0 + x_1 u_1 + b_1)} * \frac{\partial(x_0 u_0 + x_1 u_1 + b_1)}{\partial u_0} \\ &= 2(y - a_2) * \sigma'(a_0 v_0 + a_1 v_1 + b_2) * v_1 * \partial'(x_0 u_0 + x_1 u_1 + b_1) * x_0\end{aligned}$$

$$\begin{aligned}\frac{\partial J}{\partial u_1} &= \frac{\partial J}{\partial a_2} * \frac{\partial a_2}{\partial(a_0 v_0 + a_1 v_1 + b_2)} * \frac{\partial(a_0 v_0 + a_1 v_1 + b_2)}{\partial a_1} * \frac{\partial a_1}{\partial(x_0 u_0 + x_1 u_1 + b_1)} * \frac{\partial(x_0 u_0 + x_1 u_1 + b_1)}{\partial u_1} \\ &= 2(y - a_2) * \sigma'(a_0 v_0 + a_1 v_1 + b_2) * v_1 * \partial'(x_0 u_0 + x_1 u_1 + b_1) * x_1\end{aligned}$$

$$\begin{aligned}\frac{\partial J}{\partial b_1} &= \frac{\partial J}{\partial a_2} * \frac{\partial a_2}{\partial(a_0 v_0 + a_1 v_1 + b_2)} * \frac{\partial(a_0 v_0 + a_1 v_1 + b_2)}{\partial a_1} * \frac{\partial a_1}{\partial(x_0 u_0 + x_1 u_1 + b_1)} * \frac{\partial(x_0 u_0 + x_1 u_1 + b_1)}{\partial b_1} \\ &= 2(y - a_2) * \sigma'(a_0 v_0 + a_1 v_1 + b_2) * v_1 * \partial'(x_0 u_0 + x_1 u_1 + b_1) * 1\end{aligned}$$

$$\begin{aligned}\frac{\partial J}{\partial w_0} &= \frac{\partial J}{\partial a_2} * \frac{\partial a_2}{\partial(a_0 v_0 + a_1 v_1 + b_2)} * \frac{\partial(a_0 v_0 + a_1 v_1 + b_2)}{\partial a_0} * \frac{\partial a_0}{\partial(x_0 w_0 + x_1 w_1 + b_0)} * \frac{\partial(x_0 w_0 + x_1 w_1 + b_0)}{\partial w_0} \\ &= 2(y - a_2) * \sigma'(a_0 v_0 + a_1 v_1 + b_2) * v_0 * \partial'(x_0 w_0 + x_1 w_1 + b_0) * x_0\end{aligned}$$

$$\begin{aligned}\frac{\partial J}{\partial w_1} &= \frac{\partial J}{\partial a_2} * \frac{\partial a_2}{\partial(a_0 v_0 + a_1 v_1 + b_2)} * \frac{\partial(a_0 v_0 + a_1 v_1 + b_2)}{\partial a_0} * \frac{\partial a_0}{\partial(x_0 w_0 + x_1 w_1 + b_0)} * \frac{\partial(x_0 w_0 + x_1 w_1 + b_0)}{\partial w_1} \\ &= 2(y - a_2) * \sigma'(a_0 v_0 + a_1 v_1 + b_2) * v_0 * \partial'(x_0 w_0 + x_1 w_1 + b_0) * x_1\end{aligned}$$

$$\begin{aligned}\frac{\partial J}{\partial b_0} &= \frac{\partial J}{\partial a_2} * \frac{\partial a_2}{\partial(a_0 v_0 + a_1 v_1 + b_2)} * \frac{\partial(a_0 v_0 + a_1 v_1 + b_2)}{\partial a_0} * \frac{\partial a_0}{\partial(x_0 w_0 + x_1 w_1 + b_0)} * \frac{\partial(x_0 w_0 + x_1 w_1 + b_0)}{\partial b_0} \\ &= 2(y - a_2) * \sigma'(a_0 v_0 + a_1 v_1 + b_2) * v_0 * \partial'(x_0 w_0 + x_1 w_1 + b_0) * 1\end{aligned}$$

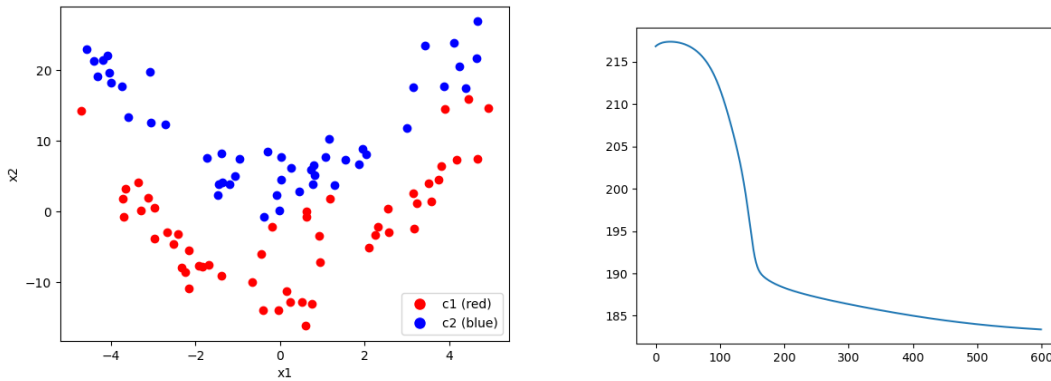
عبارات حاصل با در نظر گرفتن  $\sigma'(x) = \sigma(x) * (1 - \sigma(x))$  ساده تر نیز خواهند شد.

c.

تعداد گام ها و نرخ آموزش مناسب برای این مدل شبکه، به ترتیب 600 گام و نرخ آموزش 0.43 می باشد.

d.

نمونه‌ای از خروجی شبکه بر روی داده‌های تست به همراه نمودار میزان خطا برحسب تعداد گام در ادامه آورده شده است:



دقت این مدل شبکه، 96% برآورد شده است.

با توجه به اینکه شبکه دارای یک لایه پنهان می‌باشد، قابلیت بیشتری برای مرزبندی کلاس‌ها به نسبت مدل شبکه قبلی دارد.

همانطور که از روی شکل مشخص است، این مرزبندی به صورت یک سهمی می‌باشد؛ که با توجه به اینکه داده‌ها (چه آموزشی و چه تست) خود نیز به صورت سهمی توزیع شده‌اند، باعث شده تا این مدل شبکه به داده‌ها به خوبی fit شود.