

Optimization

①

$$h_1 = \frac{1}{1 + e^{-((0,05 \times 0,1) + (0,1 \times 0,3) + 0,25)}} = \frac{1}{1 + e^{-0,3}} \approx 0,57 \text{ (الف)}$$

$$h_2 = \frac{1}{1 + e^{-((0,05 \times 0,2) + (0,1 \times 0,4) + 0,35)}} = \frac{1}{1 + e^{-0,42}} \approx 0,6$$

$$Q_1 = \frac{1}{1 + e^{-((0,57 \times 0,5) + (0,6 \times 0,7) + 0,45)}} = \frac{1}{1 + e^{-1,155}} \approx 0,76$$

$$Q_2 = \frac{1}{1 + e^{-((0,57 \times 0,6) + (0,6 \times 0,1) + 0,55)}} = \frac{1}{1 + e^{-1,322}} \approx 0,79$$

$$E = \frac{1}{2} \left((0,1 - 0,76)^2 + (0,9 - 0,79)^2 \right) \approx 0,22 \quad (ب)$$

$$L = \frac{1}{2} (0,1 - 0,76)^2 = 0,2178$$

$$\text{Sigmoid}' = 0,76(1 - 0,76) =$$

$$0,1824$$

$$\frac{d(E)}{d(f)} = -(0,1 - 0,76) = 0,66 \Rightarrow$$

$$\text{میزان تغییر E نسبت به } w: 0,12 \Rightarrow \frac{dE}{dz} \times \frac{dz}{dw} = \frac{dE}{dw} = 0,12 \times 0,57 =$$

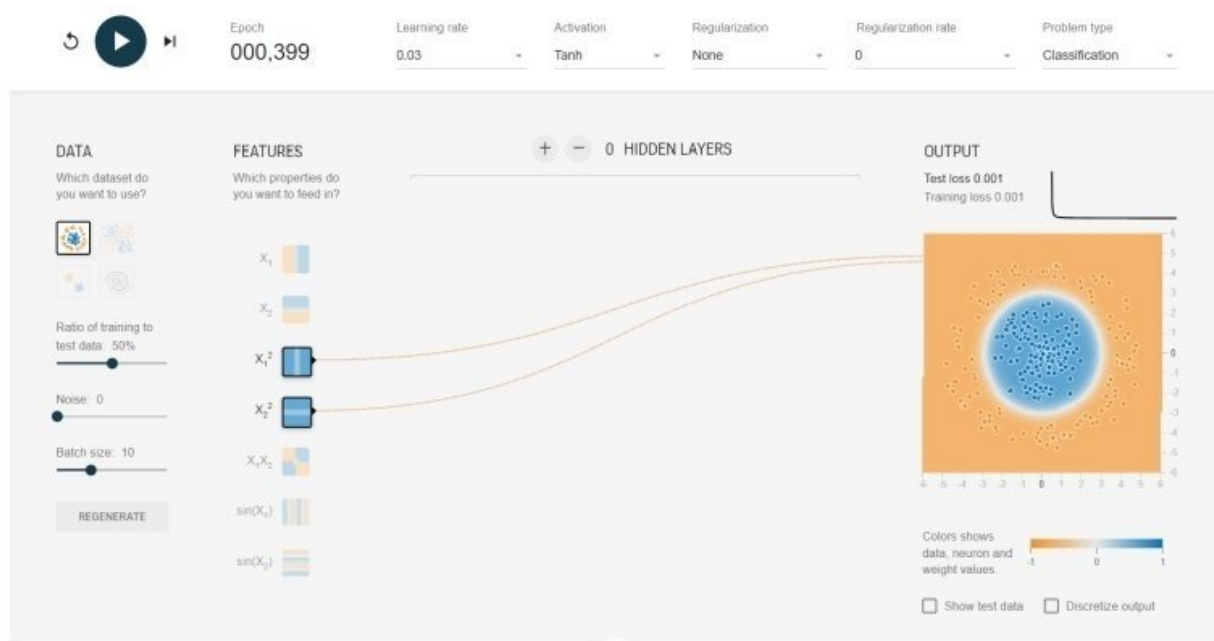
$$0,06$$

$$\Rightarrow w_5 = \frac{1}{2} - \left(\frac{1}{2} \times 0,06 \right) = 0,47$$

شبکه های عصبی

سوال اول:

(الف)



از آنجایی که دو فیچر x_1^2 و x_2^2 را داریم و معادله دایره نیز به شکل $x^2 + y^2 = 1$ است بنابراین این مدل به خوبی کار خواهد کرد.

(ب)



وزن های ورودی نرون اول (از بالا به پایین): -0.28 و -1.4

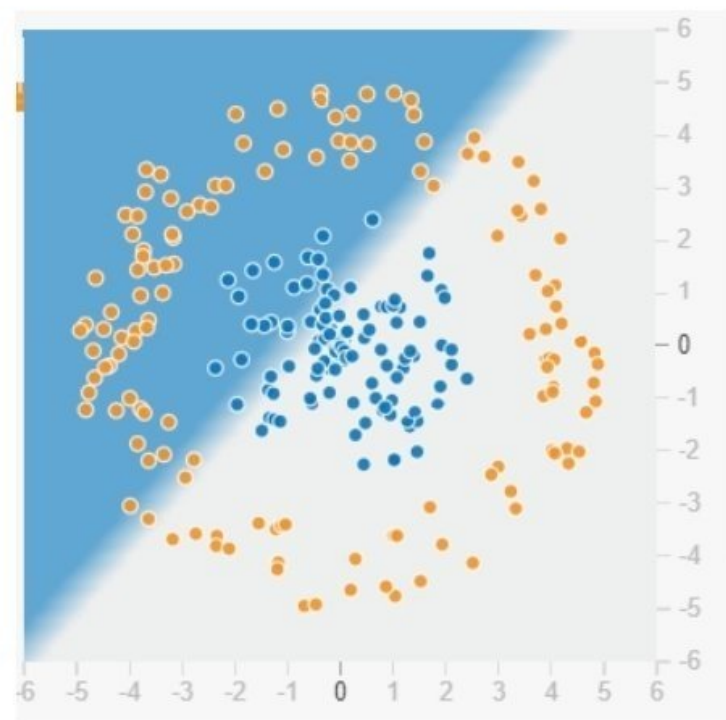
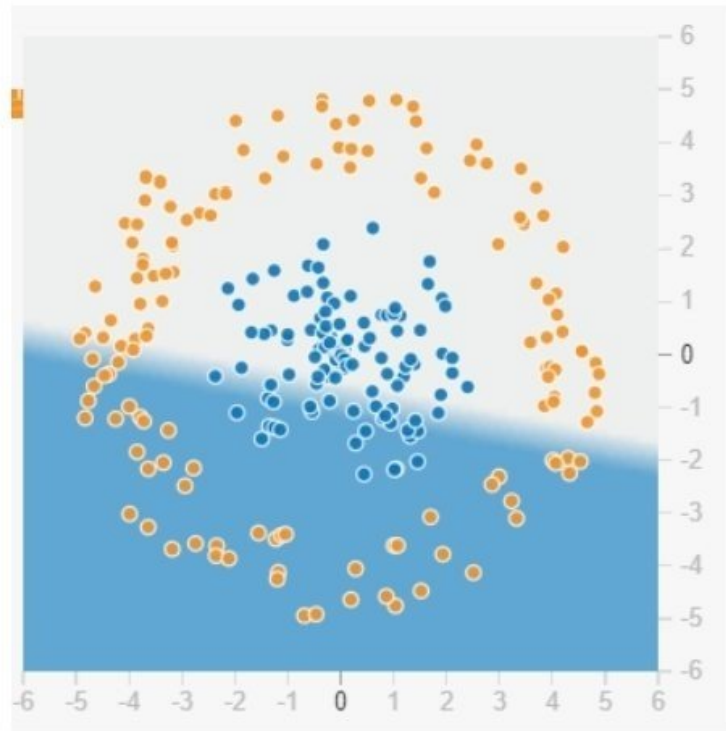
وزن های ورودی نرون دوم (از بالا به پایین): -1.1 و 0.91

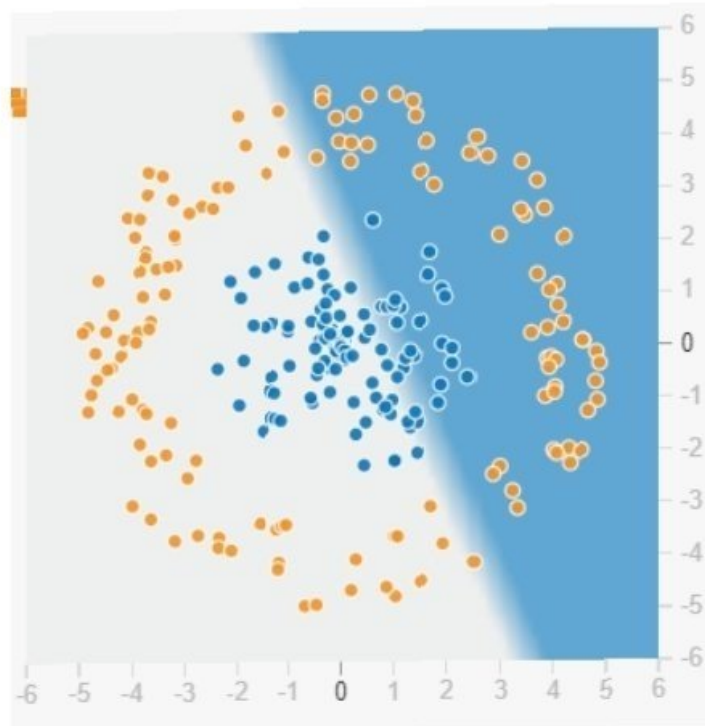
وزن های ورودی نرون سوم (از بالا به پایین): 1.3 و 0.55

وزن خروجی نرون اول: -1.6

وزن خروجی نرون دوم: -1.5

وزن خروجی نرون سوم: -1.6





1. شکل‌گیری مثلث در شبکه‌های عصبی:

- وقتی که در یک شبکه‌ی عصبی، نورون‌ها به صورت خطی با یکدیگر ارتباط دارند، می‌توانیم آن‌ها را به عنوان یک مثلث در نظر بگیریم. این مثلث توسط وزن‌های نورون‌ها شکل می‌گیرد.
- نورون‌ها به صورت خطی با یکدیگر ارتباط دارند و وزن‌های آن‌ها یک مثلث ایجاد می‌کنند. راس‌های این مثلث در وسط خطوط کوتاه 6 ضلعی قرار دارند.
- نقطه متراکم (نقطه وسط مثلث) با مقدار predict صفر شده است. این مقدار توسط bias تعیین می‌شود و در نهایت گروهی که به آن تعلق می‌گیرد 1 است.
- راس‌های مثلث از دو سمت اثر می‌گیرند و باعث می‌شوند که در نقاط اطراف نیز اثر گذاشته شود و یک 6 ضلعی به وجود بیاید.

2. شکستگی تند در گوشه‌ها:

- دلیلی که گوشه‌های مثلث تیز هستند، استفاده از تابع فعال‌سازی ReLU است. تابع ReLU شکستگی تندی دارد و به طور معمول در شبکه‌های عصبی استفاده می‌شود.

(پ)



همانند مورد قبل اینجا هم میتوان درمورد این مثلث به وجود آمده نتیجه گیری و تحلیل کرد و همچنین توابع Tanh و Sigmoid توابع smooth هستند و دیگر همانند مورد قبل، این شکل گوشه های تیز ندارد.

شبکه های عصبی پیچشی

سوال اول:

1. الیه های پیچشی (Convolutional Layers):

- الیه های پیچشی به تصویر ورودی فیلترها (کرنل ها) اعمال می کنند. این فیلترها به طور خودکار ویژگی های مختلف مانند لبه ها، نقاط روشن و تیرگی ها را تشخیص می دهند.
- با استفاده از پیچش ها، ما می توانیم ویژگی های سطح پایین مانند لبه ها را استخراج کنیم. این ویژگی ها به شبکه های عصبی کمک می کنند تا الگوهای پیچیده تر را در تصاویر تشخیص دهند.

2. الیه های ادغام (Pooling Layers):

- الیه های ادغام با کوچک کردن اندازهی تصویر ورودی و خلاصه سازی ویژگی های اصلی و مهم موجود در عکس، به کاهش محاسبات شبکه و مشکل Overfitting کمک می کنند.
- این الیه ها معمولاً با استفاده از میانگین گیری یا حداکثرگیری از منطقه های کوچک تصویر، اطلاعات را خلاصه می کنند.

سوال دوم :

$$[x_1, x_2, x_3] \Rightarrow \left. \begin{array}{l} x_1 = 1 \\ x_2 = 4 \\ x_3 = 0 \end{array} \right\} \Rightarrow x_1 + 4x_2 = -2$$

$$\left. \begin{array}{l} x_1 = 4 \\ x_2 = 0 \\ x_3 = -2 \end{array} \right\} \Rightarrow 4x_1 - 2x_3 = -2, \quad \left. \begin{array}{l} x_1 = 0 \\ x_2 = -2 \\ x_3 = 3 \end{array} \right\} \Rightarrow -2x_2 + 3x_3 = 11$$

$$\Rightarrow x_1 = \frac{14}{13}, \quad x_2 = \frac{-10}{13}, \quad x_3 = \frac{123}{39} = \frac{41}{13}$$

$$\Rightarrow \text{filter} = \frac{1}{13} [14, -10, 41]$$