

سوال اول

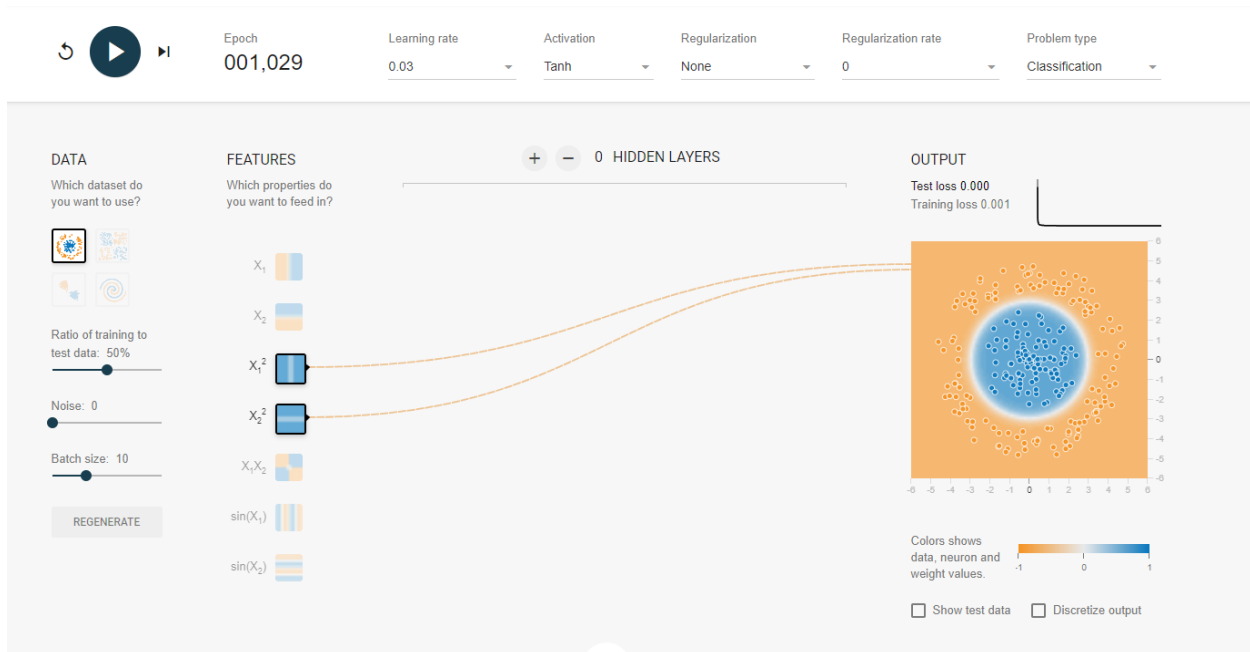
$E = \frac{1}{2} \sum_i (target_i - output_i)^2$   
 $E = \frac{1}{2} [(0.1 - 0.1)^2 + (0.9 - 0.9)^2] = 0$

$h_1 = 0.1 \times 0.1 + 0.2 \times 0.2 + 0.1 \times 0.3 = 0.13$   
 $h_2 = 0.1 \times 0.2 + 0.2 \times 0.1 + 0.1 \times 0.3 = 0.09$   
 $o_1 = h_1 \times 0.3 + h_2 \times 0.4 + 0.1 \times 0.5 = 0.13 \times 0.3 + 0.09 \times 0.4 + 0.1 \times 0.5 = 0.13$   
 $o_2 = h_1 \times 0.4 + h_2 \times 0.5 + 0.1 \times 0.6 = 0.13 \times 0.4 + 0.09 \times 0.5 + 0.1 \times 0.6 = 0.13$

$E = \frac{1}{2} [(0.1 - 0.13)^2 + (0.9 - 0.13)^2] = 0.123$

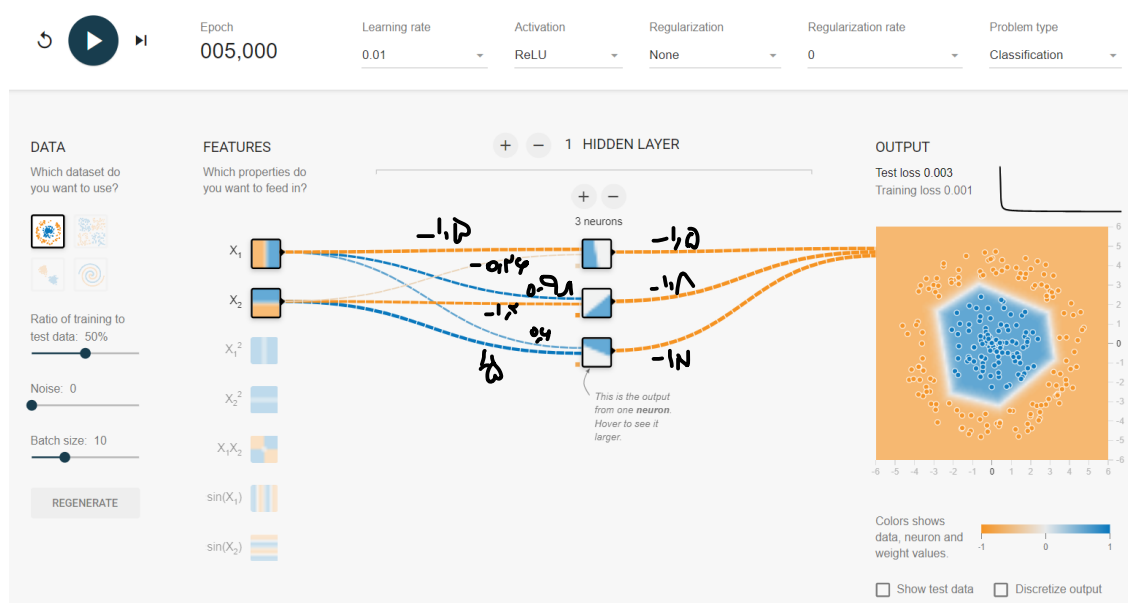
$w_{\omega}^{t+1} = w_{\omega}^t - \eta \frac{d \text{loss}}{d w_{\omega}}$   
 $\text{loss} = \frac{1}{2} (o_{final} - d)^2 \rightarrow \frac{d \text{loss}}{d y} = o_{final} - d$   
 $\frac{d \text{loss}}{d w_{\omega}} = \frac{d \text{loss}}{d o_{final}} \times \frac{d o_{final}}{d t} \times \frac{d t}{d w_{\omega}} = (0.13 - 0.1) \times (0.13 - 0.1) \times (0.13 - 0.1) = 0.0009$   
 $w_{\omega}^{t+1} = 0.1 - \frac{1}{2} \times 0.0009 = 0.0995$

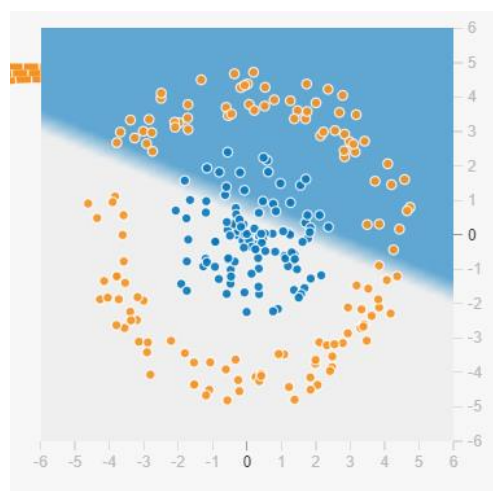
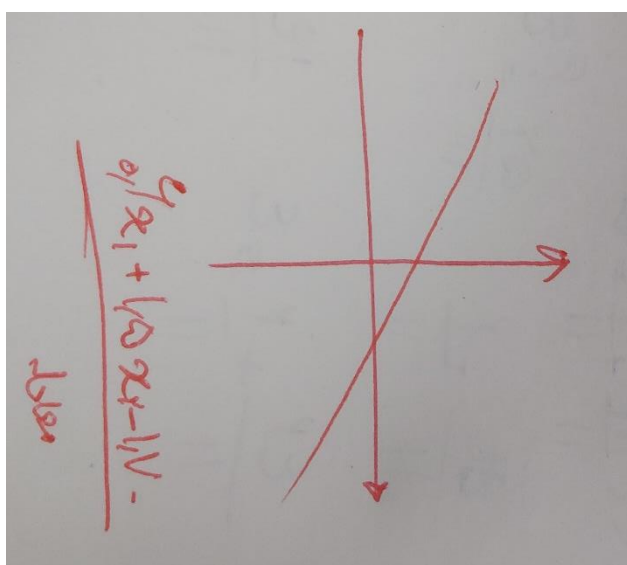
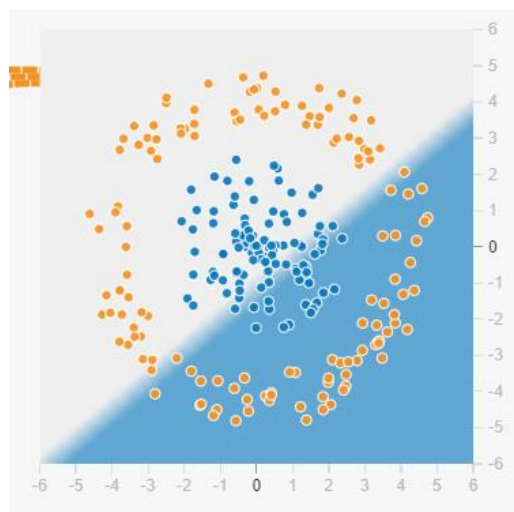
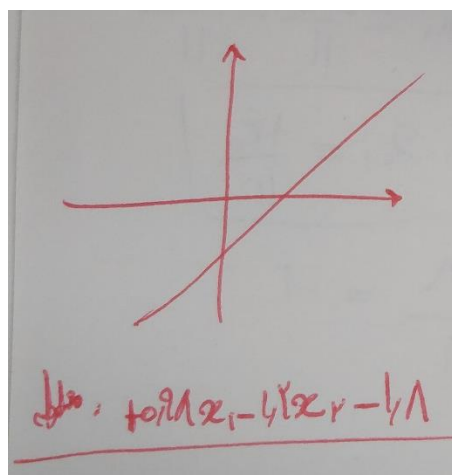
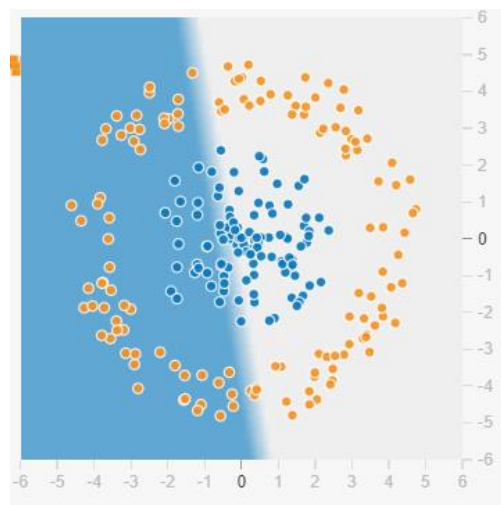
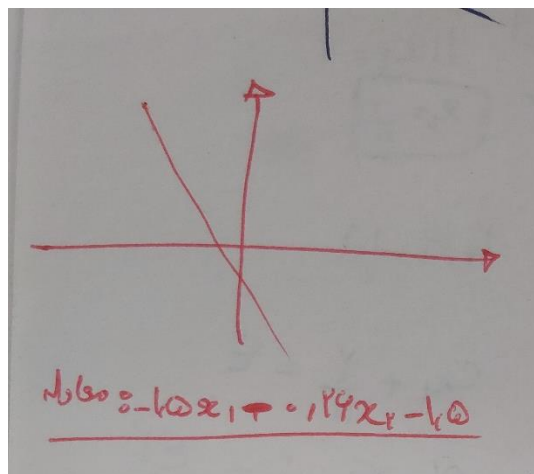
$$w_{\omega}^{t+1} = w_{\omega}^t - \eta \frac{d \text{loss}}{d w_{\omega}} = 0.1 - \frac{1}{2} \times (0.0009) = 0.0995$$



مرز جداکننده همانطور که مشخص است، باید یک دایره باشد. همانطور که در شکل مشاهده می شود، شبکه به خوبی توانسته داده ها را دسته بندی کند. در واقع با perceptron تونست جدا بکنه. ما یکی از ویژگی ها را  $X_1^2$  و دیگری را  $X_2^2$  گذاشتیم و میدانیم معادله آن دایره هم که در تصویر آمده در واقع  $X_1^2 + X_2^2$  که این جمع از یک threshold ای کوچکتر است. شبکه ما هم در واقع همین کار را کرده است و مرز تصمیم را به شکل دایره در آورده است. نسبت به ورودی ها خطی است اما ما چون در ورودی ها توان 2 اعمال کرده ایم، تونستیم دایره را بدست آوریم.

(ب)

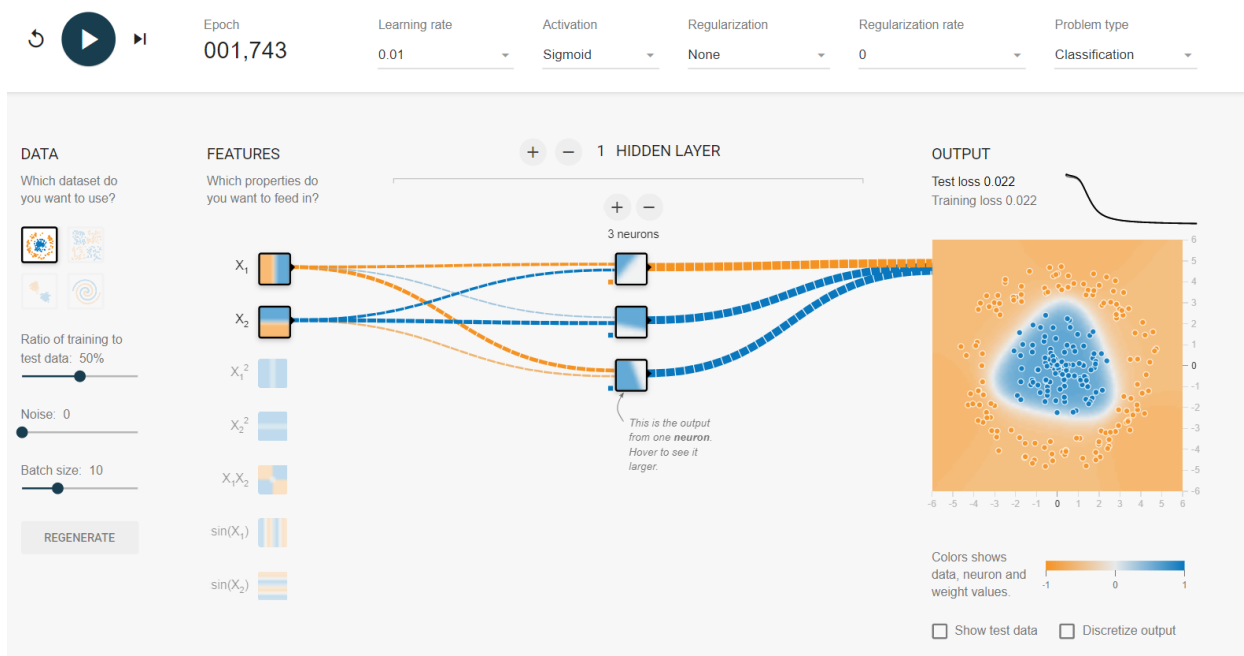
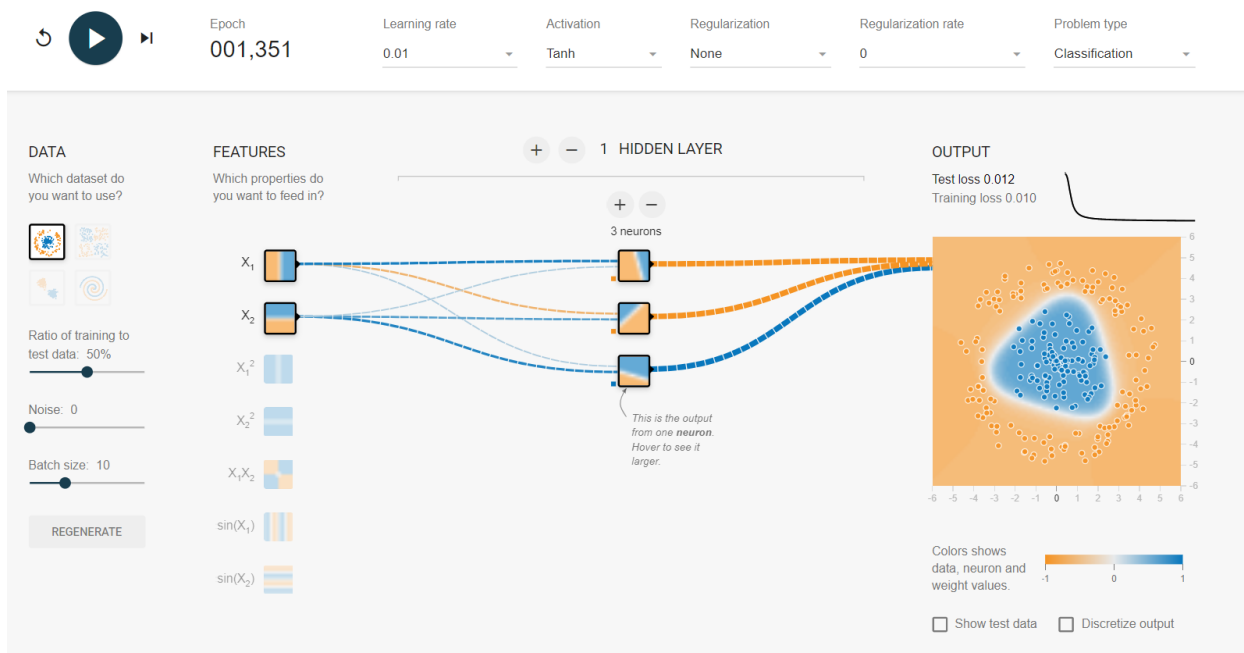




از آنجایی که مقادیر وزن های لایه خروجی برای تصمیم گیری منفی شده است، برای اینکه بتوان مقداری مثبت در خروجی گرفت، بایستی که بایاس مقداری مثبت در لایه خروجی باشد.

رفتار تابع فعالساز Relu بعد از صفر باعث میشود فقط وابستگی های خطی را پس از فعال شدن در نظر میگیرد که منجر به تولید خط به عنوان مرز شده است.

(پ)



به نظر میرسد تفاوت در خطی یا غیرخطی بودن توابع فعالساز است. Relu وابستگی های خطی را شبیه سازی میکند، و این منجر به تولید مرزهای خطی میشود. اما Tanh و sigmoid توابع غیرخطی هستند و مرزهای خطی تولید نمیکنند و میتوانند که مرزهای

غیرخطی و منحنی طور ایجاد کنند. این توابع این امکان را به ما میدهند تا فضا های پیچیده تر را به کمک این توابع فعالساز غیرخطی بهتر طبقه بندی بکنیم.

### شبکه های عصبی پیچشی:

سوال اول:

حفظ اطلاعات مکانی و کاهش پارامترها:

حفظ اطلاعات مکانی: لایه های پیچشی ویژگی های محلی یک تصویر را با حفظ ساختار فضایی آن استخراج می کنند. این بدان معناست که روابط فضایی بین پیکسل ها و الگو های محلی در تصویر حفظ می شود. این قابلیت بسیار مهم است، زیرا بسیاری از ویژگی های مهم در تصاویر به صورت محلی ظاهر می شوند و حفظ این اطلاعات مکانی کمک می کند که شبکه بتواند به درستی این ویژگی ها را شناسایی کند.

کاهش تعداد پارامترها: در لایه های کاملاً متصل، هر نورون به تمامی نورون های لایه قبلی متصل است، که منجر به افزایش سریع تعداد پارامترها به خصوص برای تصاویر بزرگ می شود. اما در لایه های پیچشی، هر فیلتر تنها با یک ناحیه کوچک از ورودی در هر زمان تعامل دارد، که به طور قابل توجهی تعداد پارامترها را کاهش می دهد. این کاهش پارامترها نه تنها به کاهش نیاز به حافظه و زمان پردازش کمک می کند، بلکه احتمال بیش پرازش (Overfitting) را نیز کاهش می دهد.

کاهش حساسیت به جابجایی و تغییر مقیاس:

کاهش حساسیت به جابجایی: لایه های ادغام (Pooling) با کاهش ابعاد ویژگی های استخراج شده و تجمیع اطلاعات، می توانند به کاهش حساسیت مدل به جابجایی و تغییرات کوچک در ورودی کمک کنند. این ویژگی به شبکه کمک می کند تا پایداری بیشتری در تشخیص الگوها در تصاویر داشته باشد، حتی اگر این الگوها به میزان کمی جابجا شده باشند.

کاهش حساسیت به تغییر مقیاس: با استفاده از تکنیک های مانند ماکس-پولینگ (Max-Pooling)، شبکه می تواند ویژگی های با اهمیت را حتی در صورت تغییرات مقیاس حفظ کند. این امر باعث می شود شبکه بتواند به خوبی با تصاویر ورودی که اندازه و مقیاس آنها متفاوت است، سازگار شود.

سوال دوم

Signal:  $[1, 4, 0, -2, 3]$

output:  $[-2, -2, 11]$

Stride = 1

$[1, 4, 0, -2, 3]$

$F = [\omega_1, \omega_r, \omega_p]$

$$\omega_1 + 4\omega_r = -2$$

$$4\omega_1 - 2\omega_r = -2$$

$$[\omega_1, \omega_r, \omega_p] \rightarrow -2\omega_r + 3\omega_p = 11$$

$$\textcircled{1} \omega_1 + 4\omega_r = -2$$

$$\textcircled{2} 4\omega_1 - 2\omega_r = -2$$

$$\textcircled{3} -2\omega_r + 3\omega_p = 11$$

$$-2 \times \textcircled{1} + \textcircled{2} = -8\omega_1 - 14\omega_r = 11$$

$$\textcircled{+} \quad 8\omega_1 - 2\omega_r = -2$$

$$-2\omega_p - 14\omega_r = +9$$

$$\textcircled{2} \rightarrow \begin{cases} 2\omega_p + 14\omega_r = -9 \\ 2\omega_p - 2\omega_r = 11 \end{cases}$$

$$2\omega_p + 14\omega_r = -9$$

$$\times 1 \rightarrow 2\omega_p - 2\omega_r = 11$$

$$24\omega_r = 12$$

$$\omega_r = \frac{12}{24} = \frac{1}{2}$$

$$2 \times \frac{41}{13} - 2\omega_r = 11$$

$$\rightarrow \frac{82}{13} - \frac{11 \times 13}{13} = 2\omega_r$$

$$\frac{82 - 143}{13} = 2\omega_r \rightarrow$$

$$\omega_r = \frac{-61}{26}$$

$$\omega_1 + 4\left(\frac{-61}{26}\right) = -2$$

$$\omega_1 = \frac{61}{13} - 2 = \frac{61 - 26}{13} = \frac{35}{13}$$

$$F = \left[ \frac{35}{13}, \frac{-61}{26}, \frac{14}{13} \right]$$

النتيجة  $F = \left[ \frac{35}{13}, \frac{-61}{26}, \frac{14}{13} \right]$  درست مع (تفاوت نمیکند) به نوع مسئله و محل کاتالوگ (طرح)