



به نام خدا

دانشگاه تهران

پردیس دانشکده های فنی

دانشکده مهندسی برق و کامپیوتر

تمرین شماره 5

بخش کتبی

مدرسین:

دکتر فدائی

دکتر یعقوبزاده

نگارش:

فاطمه محمدی 810199489

بخش کتبی

3..... بهینه سازی

3..... سوال اول

3..... الف)

4..... ب)

4..... ج)

5..... شبکه های عصبی

5..... سوال اول

5..... الف)

6..... ب)

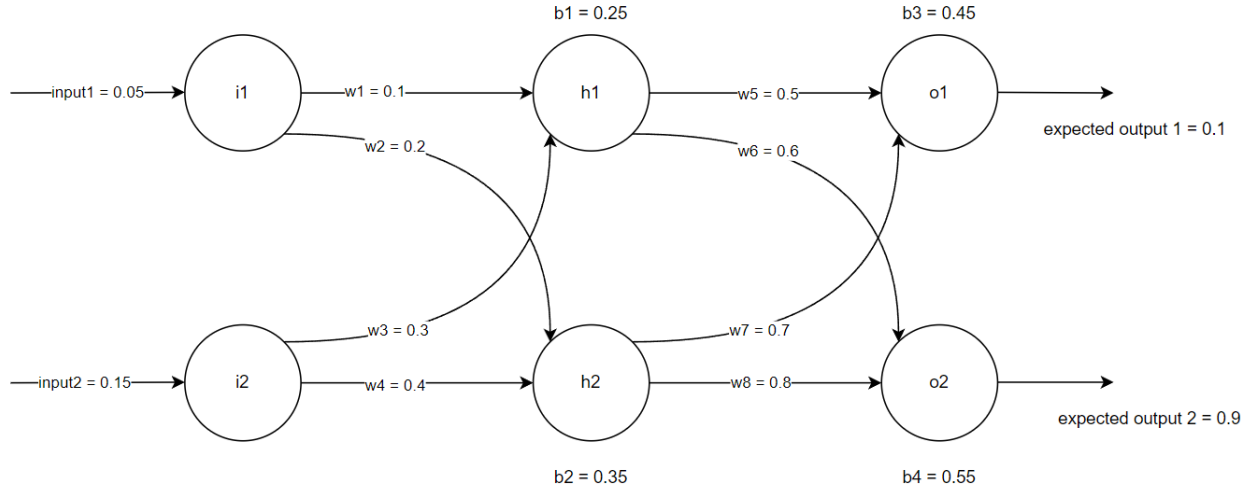
6..... پ)

9..... شبکه های عصبی پیچشی

9..... سوال اول

بهینه سازی

سوال اول



(الف)

$$\sigma(x) = \frac{1}{1+e^{-x}} \text{ تابع فعال ساز:}$$

$$In_{h1} = (w1 \cdot i1) + (w3 \cdot i2) + b1 \rightarrow In_{h1} = 0.1 \cdot 0.05 + 0.3 \cdot 0.15 + 0.25 = 0.3$$

$$In_{h2} = (w2 \cdot i1) + (w4 \cdot i2) + b2 = In_{h2} = 0.2 \cdot 0.05 + 0.4 \cdot 0.15 + 0.35 = 0.42$$

$$Out_{h1} = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-0.3}} = 0.5744$$

$$Out_{h2} = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-0.42}} = 0.6035$$

$$In_{o1} = (w5 \cdot Out_{h1}) + (w7 \cdot Out_{h2}) + b3 \rightarrow In_{o1} = 0.5 \cdot 0.5744 + 0.7 \cdot 0.6035 + 0.45 = 1.1597$$

$$In_{o2} = (w6 \cdot Out_{h1}) + (w8 \cdot Out_{h2}) + b4 \rightarrow In_{o2} = 0.6 \cdot 0.5744 + 0.8 \cdot 0.6035 + 0.55 = 1.3774$$

$$Out_{o1} = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-1.1597}} = 0.7613$$

$$Out_{o2} = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-1.3774}} = 0.7985$$

(ب)

$$E = \frac{1}{2} \sum_i (target_i - output_i)^2$$

| | i = 1 | i = 2 |
|------------|--------|--------|
| $target_i$ | 0.1 | 0.9 |
| $output_i$ | 0.7613 | 0.7985 |

$$\rightarrow E = \frac{1}{2} ((0.1 - 0.7613)^2 + (0.9 - 0.7985)^2) = \frac{1}{2} (0.4373 + 0.0103) = 0.2238$$

(ج)

$$\sigma'(x) = \sigma(x)(1 - \sigma(x))$$

O1:

$$\delta_{o1} = (out_1 - target_1) \cdot \sigma'(in_{o1})$$

$$\sigma'(in_{o1}) = \sigma(in_{o1}) \cdot (1 - \sigma(in_{o1})) = 0.7613 \cdot (1 - 0.7613) = 0.1816$$

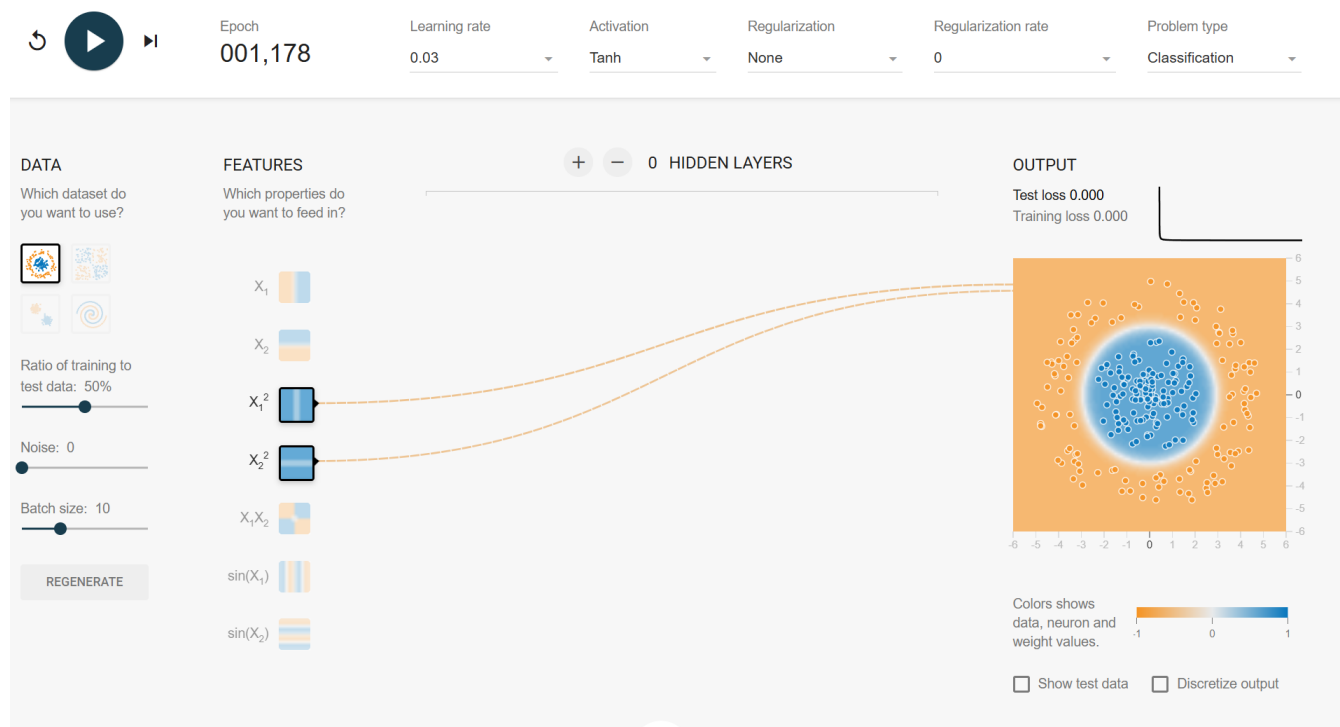
$$\delta_{o1} = (0.7613 - 0.1) \cdot 0.1816 = 0.1201$$

$$w5_{new} = w5_{old} - \alpha \cdot \delta_{o1} \cdot h1 = 0.5 - 0.5 \cdot 0.1201 \cdot 0.5744 = 0.4656$$

شبکه های عصبی

سوال اول

(الف)



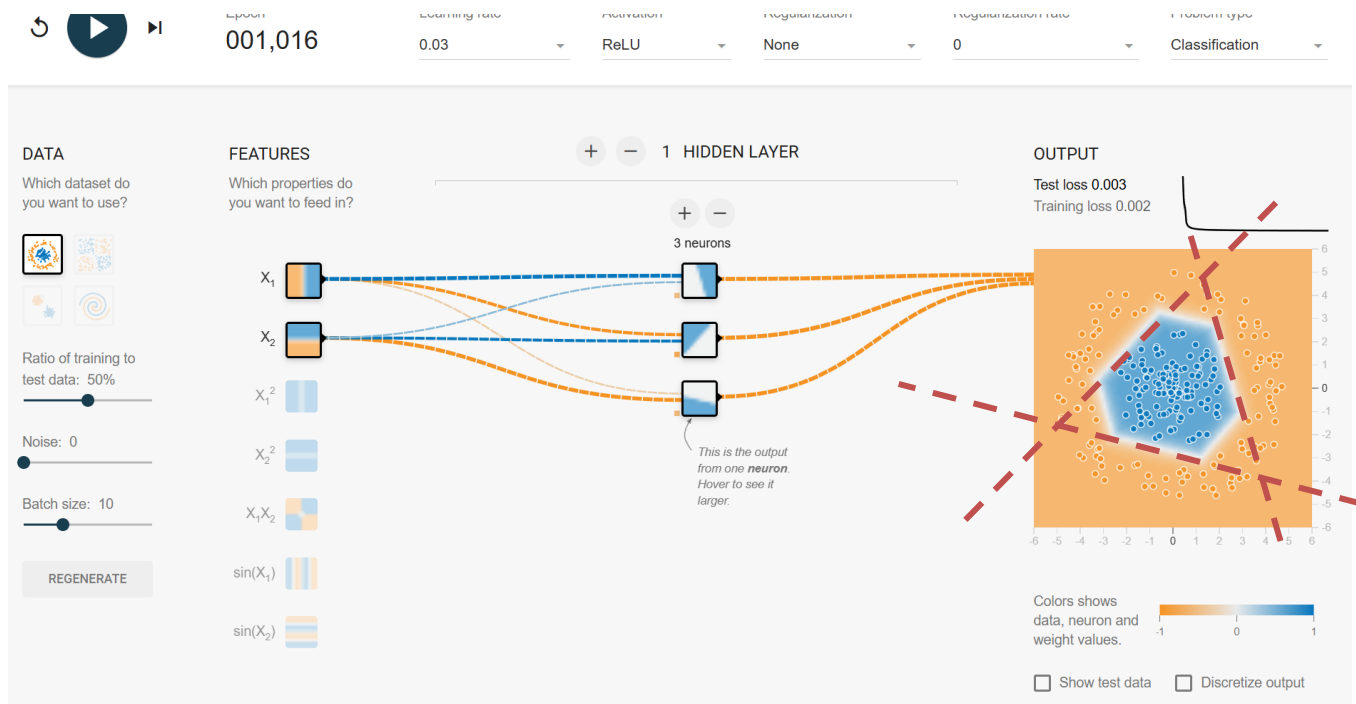
$$w_1 = -0.45, w_2 = -0.43$$

داد های موجود به صورت دایره ای شکل میباشند و مرز جداساز آنها پس دایره میباشد با توجه به اینکه در فرمول دایره نیز تنها از توان دو، x_1 و x_2 استفاده میشود و اینکه ما از توان 2 های feature ها استفاده میکنیم (که یعنی تنها feature های مورد نیاز برای طبقه بندی دایره ای را داریم) مدل به درستی کار میکند.

برای شفاف سازی بیشتر مرز تصمیم به صورت زیر است:

$$w_1 \cdot x_1^2 + w_2 \cdot x_2^2 = R^2$$

(ب)



$$w_1 = -1.6, w_2 = -1.6, w_3 = -1.5$$

در واقع معادله ای که داریم به صورت زیر است:

$$-1.6 \times n_1 - 1.6 \times n_2 - 1.5 \times n_3 + b$$

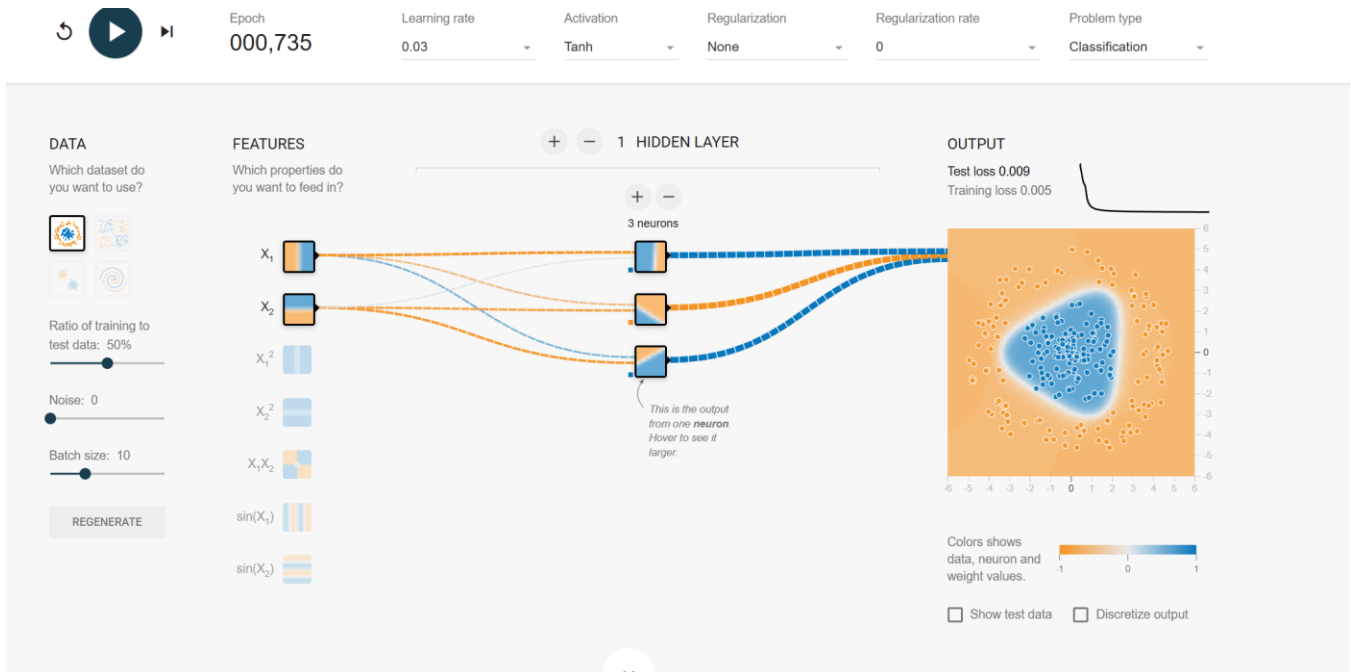
که هرکدام از n_i ها به صورت زیر است:

$$a_i X + b_i Y + c_i$$

که (X, Y) نقاطی هستند که برای هر نقطه باید بدست آورد، برای نقاط مرکزی مقدار $-1.6 \times n_1 - 1.6 \times n_2 - 1.5 \times n_3$ ، منفی میشود که میتوان نتیجه گرفت بایاس (b) مقدار مثبتی دارد که در نهایت خروجی مثبت شود، اما نقاط گوشه ای (که توقع داشتیم به عنوان کلاس ایی شناخته شوند، اما نشده اند، این مقدار منفی شده است)

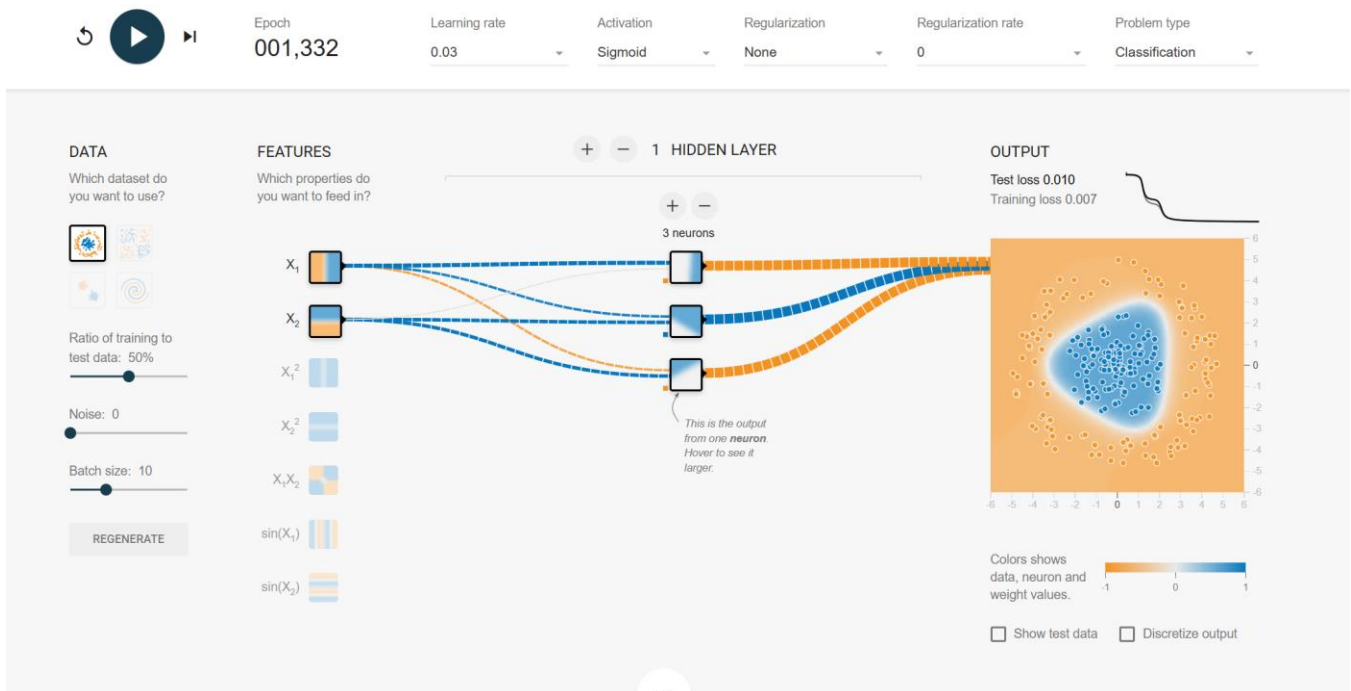
(پ)

1. Tanh:



$$w_1 = 3, w_2 = -2.9, w_3 = 2.9$$

:Sigmoid .2



$$w_1 = -5.5, w_2 = 5.3, w_3 = -5.5$$

اصلی ترین ویژگی و تفاوت این می باشد که Sigmoid و Tanh غیرخطی هستند و در نتیجه میتوانند مرز های غیر خطی (منحنی) داشته باشند (که به وضوح تاثیر آن ها دیده میشود).

شبکه های عصبی پیچشی

سوال اول

1. عملیات convolution قابلیت موازی سازی دارد که میتواند موجب افزایش سرعت محاسبات بشود.
2. در لایه های پیچیدگی وزن های مشترک و میدان های دریافتی محلی در لایه های Convolution به عنوان شکلی از منظم سازی عمل می کنند و با محدود کردن مدل برای یادگیری الگوهای محلی فضایی، خطر **overfitting** را کاهش می دهند اما در لایه های کاملاً متصل، پارامترهای زیادی داریم، احتمال **overfitting** افزایش پیدا میکند.
3. در لایه های کاملاً متصل، با تمام پیکسل های ورودی یکسان برخورد میشود ولی در لایه های Convolution از فیلتر استفاده میشود (فیلترهای $n \times n$ بعدی) که تلاش میکنند روی تصویر ورودی اشکال ساده، الگوها و لبه ها را پیدا کنند و در عمق های بالاتر با ترکیب این موارد، ساختارهای پیچیده تری شناسایی کنند و به این صورت سلسه مراتبی از ویژگی ها ایجاد میشود که برای درک محتوای تصویر هم بسیار کاربردی است.

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

$$Filter = [f_1, f_2, f_3]$$

$$Output = [-2, -2, 11]$$

$$\begin{pmatrix} 1 & 4 & 0 \\ 4 & 0 & -2 \\ 0 & -2 & 3 \end{pmatrix} * \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \\ 11 \end{pmatrix}$$

$$1 f_1 + 4 f_2 = -2 \quad (1)$$

$$4 f_1 - 2 f_3 = -2 \quad (2)$$

$$-2 f_2 + 3 f_3 = 11 \quad (3)$$

$$(1) \rightarrow f_1 = -2 - 4 f_2 \quad (4)$$

$$(1, 4) \rightarrow 4(-2 - 4 f_2) - 2 f_3 = -2 \rightarrow -8 f_2 - f_3 = 3 \quad (5)$$

$$(5) \rightarrow f_3 = -3 - 8 f_2 \quad (6)$$

$$(6, 3) \rightarrow -2 f_2 + 3(-8 f_2 - 3) = 11 \rightarrow f_2 = -\frac{10}{13} \quad (7)$$

$$(4, 7) \rightarrow f_1 = -2 - 4 * \left(-\frac{10}{13}\right) \rightarrow f_1 = \frac{14}{13} \quad (8)$$

$$(8, 2) \rightarrow 4 * \left(\frac{14}{13}\right) - 2 f_3 = -2 \rightarrow f_3 = \frac{41}{13} \quad (9)$$

$$(7, 8, 9) \rightarrow \text{Filter} = \left[\frac{14}{13}, -\frac{10}{13}, \frac{41}{13}\right] \quad (10)$$