به نام خدا

دانشگاه تهران

پردیس دانشکده های فنی

دانشکده مهندسی برق و کامپیوتر

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

تمرین شماره‌ی 5

بخش کتبی

**مدرسین:**

دکتر فدائی

دکتر یعقوب‌زاده

**نگارش:**

فاطمه محمدی 810199489

**بخش کتبی**

[بهینه سازی 3](#_Toc168090601)

[سوال اول 3](#_Toc168090602)

[الف) 3](#_Toc168090603)

[ب) 4](#_Toc168090604)

[ج) 4](#_Toc168090605)

[شبکه های عصبی 5](#_Toc168090606)

[سوال اول 5](#_Toc168090607)

[الف) 5](#_Toc168090608)

[ب) 6](#_Toc168090609)

[پ) 6](#_Toc168090610)

[شبکه های عصبی پیچشی 9](#_Toc168090611)

[سوال اول 9](#_Toc168090612)

# بهینه سازی

## سوال اول



### الف)

تابع فعال‌ساز:

### ب)

|  |  |  |
| --- | --- | --- |
| i = 2 | i = 1 |  |
| 0.9 | 0.1 |  |
| 0.7985 | 0.7613 |  |

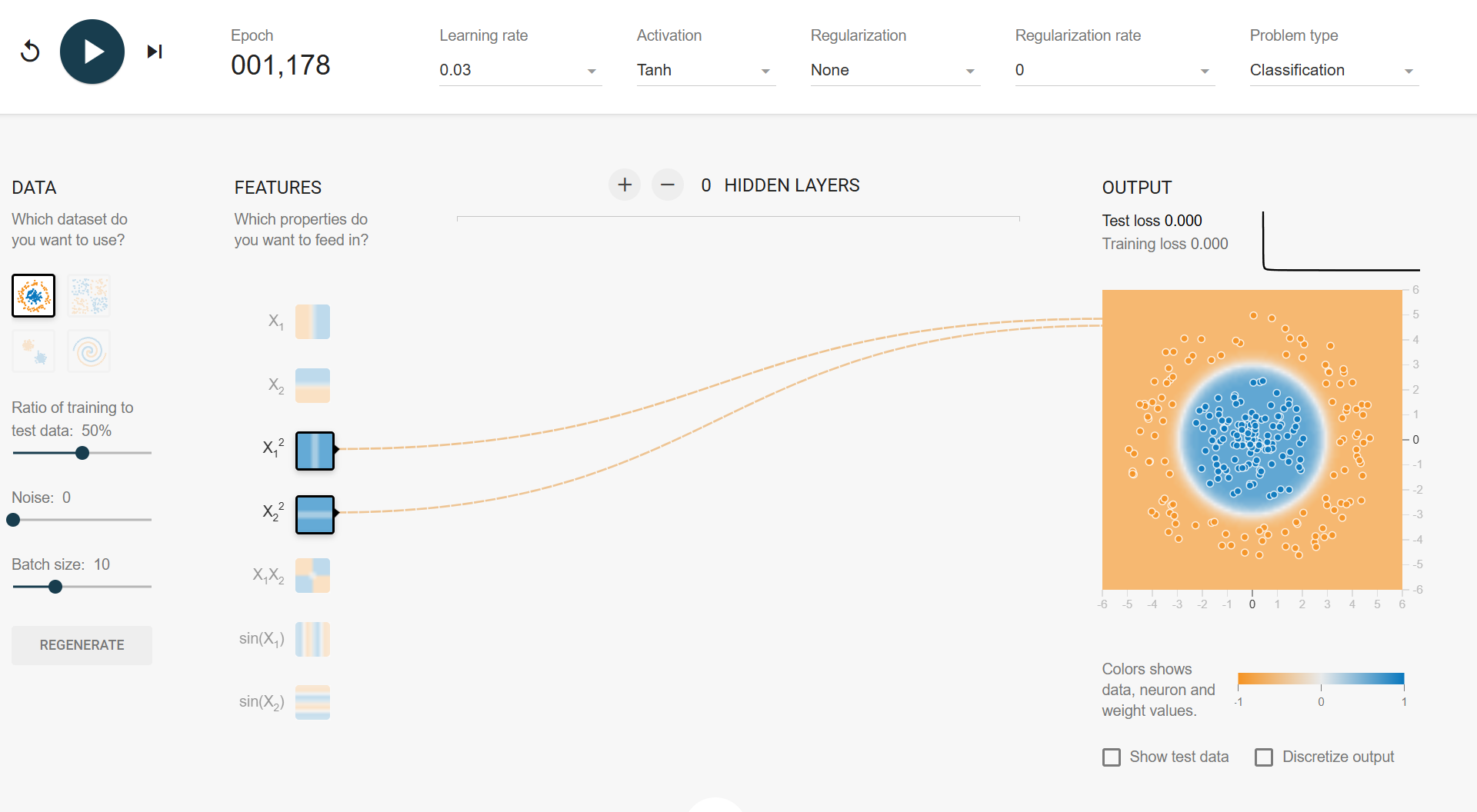
### ج)

O1:

# شبکه های عصبی

## سوال اول

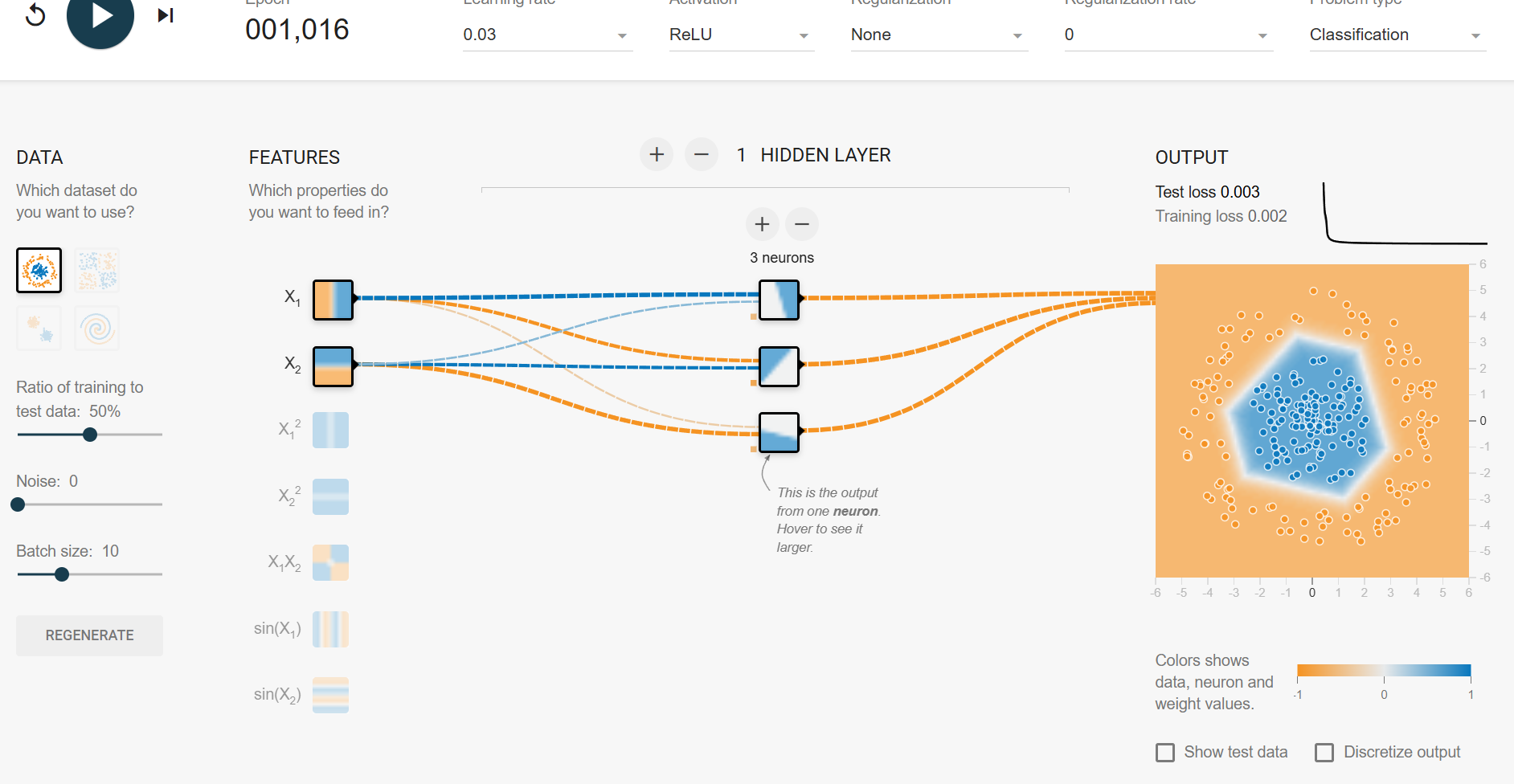
### الف)



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 x1, x2 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 تصیم به صورت زیر است:

### ب)



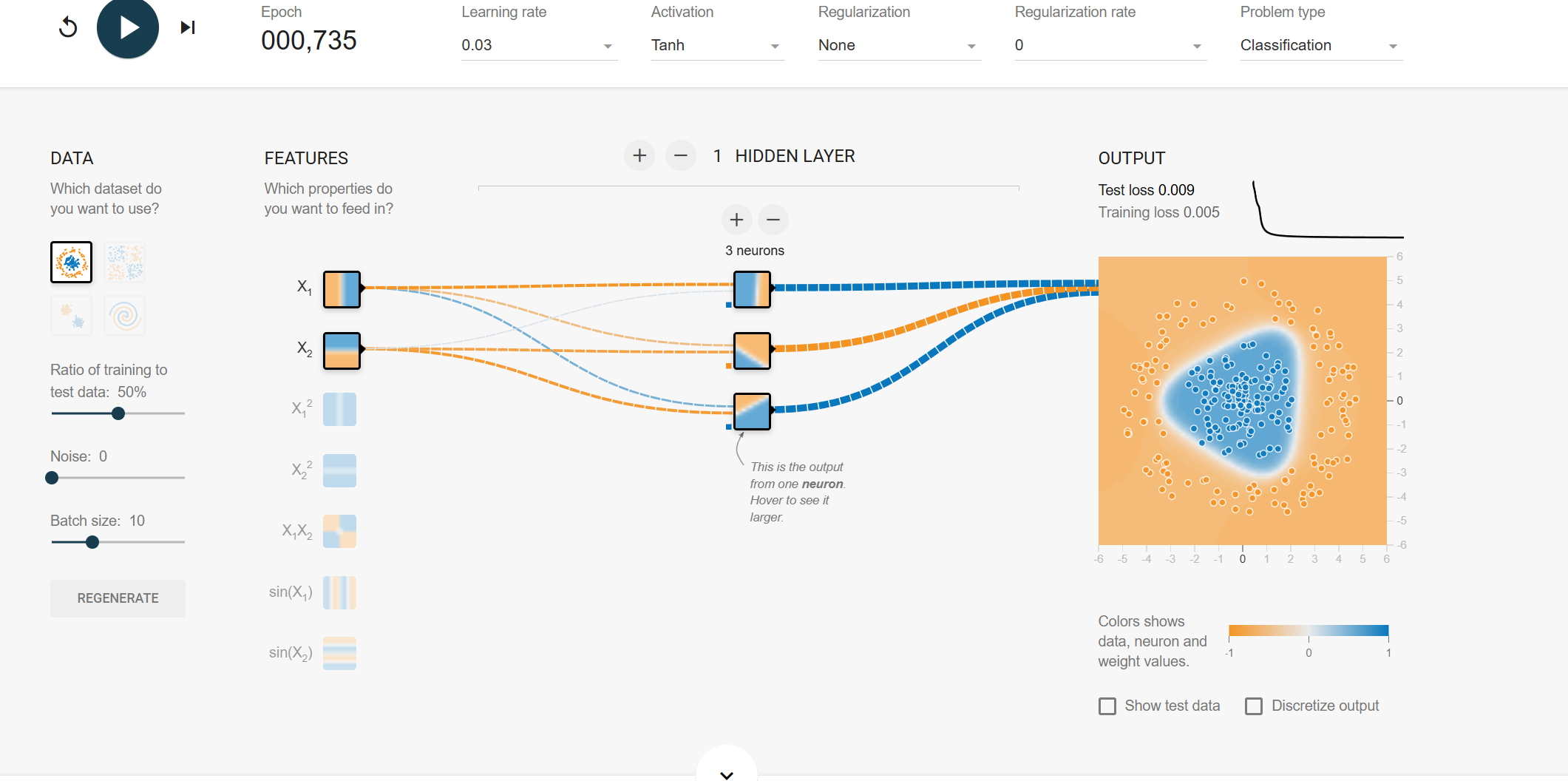
در واقع معادله ای که داریم به صورت زیر است:

که هرکدام از ها به صورت زیر است:

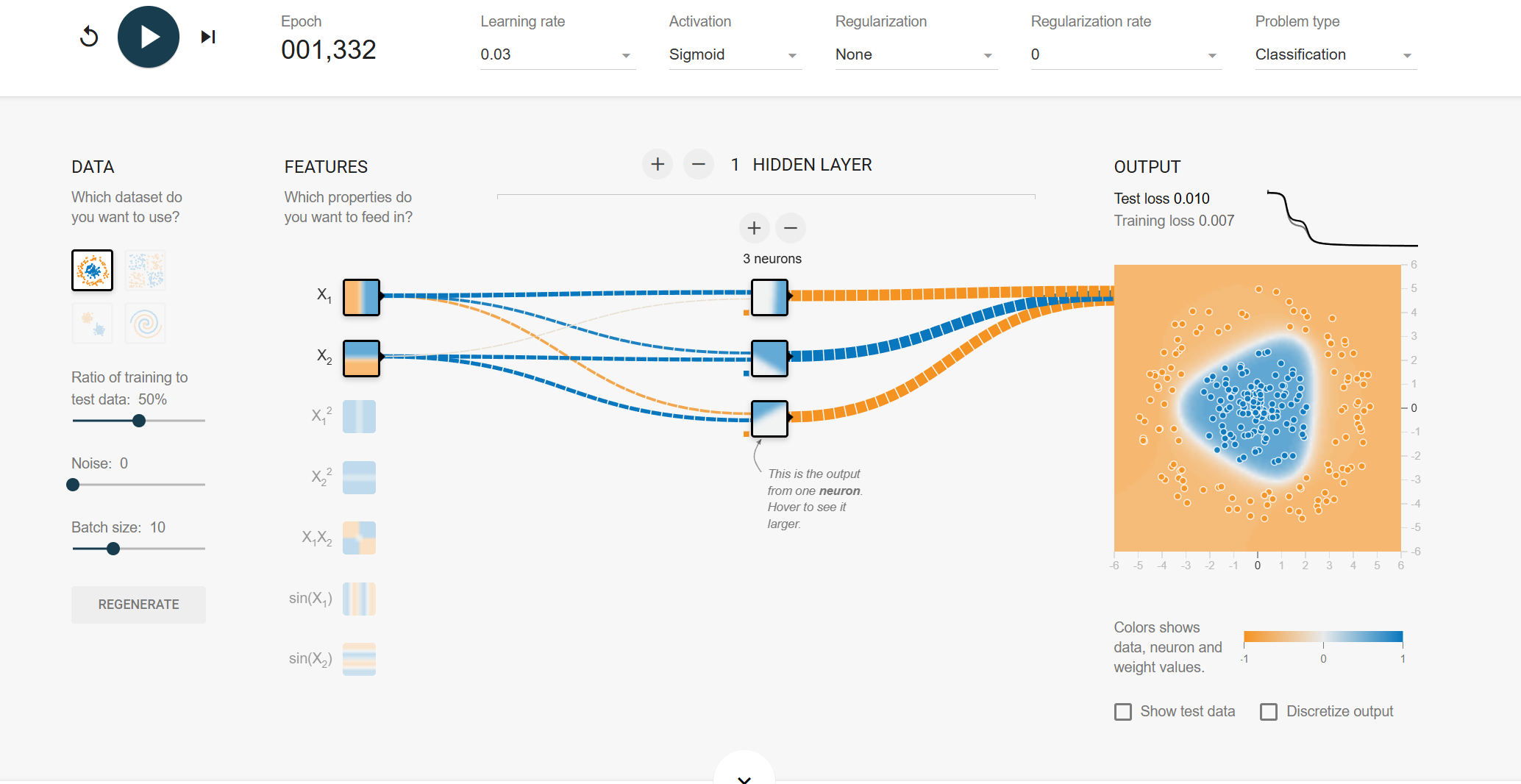
که (X, Y) نقاطی هستن که برای هر نقطه باید بدست اورد، برای نقاط مرکزی مقدار ، منفی میشود که میتوان نتیجه گرفت بایاس (b) مقدار مثبتی دارد که در نهایت خروجی مثبت شود، اما نقاط گوشه ای (که توقع داشتیم به عنوان کلاس ابی شناخته شوند، اما نشده اند، این مقدار منفی شده است)

### پ)

1. Tanh:



2. Sigmoid:



اصلی ترین ویژگی و تفاوت این میباشد که Sigmoid و Tanh غیرخطی هستند و در نتیجه میتواندد مرز های غیر خطی (منحنی) داشته باشند (که به وضوح تاثیر ان 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 کاملا متصل، با تمام پیکسل های ورودی یکسان برخورد میشود ولی در لایه های Convolution از فیلتر استفاده میشود (فیلتر های n\*n بعدی) که تلاش میکنند روی تصویر ورودی اشکال ساده، الگوها و لبه ها را پیدا کند و در عمق های بالاتر با ترکیب این موارد، ساختار های پیچیده تری شناسایی کتتد و به این صورت سلسه مراتبی از ویزگی ها ایجاد میشود که برای درک محتوای تصویر هم بسیار کاربردی است.
4. سوال دوم

(1)

(2)

(3)

(1) (4)

(1, 4) (5)

(5) (6)

(6, 3) (7)

(4, 7) (8)

(8, 2) (9)

(7,8,9) (10)