

$$E = \frac{1}{2} \sum_i (\text{target}_i - \text{output}_i)^2$$

(۱)

۱ نرون = $(0/0 \times 0) + (0/0 \times 0/2) + (0/1 \times 0) = 0/0$

۲ نرون : $(0/0 \times 0/4) + (0/1 \times 0/0) + (0/0 \times 0) = 0/0$

hidden layer (الف)

Sigmoid function:

۱: $\text{sigmoid}(0/0) \approx 0/0$

۲: $\text{sigmoid}(0/0) \approx 0/0$

output layer:

۱ نرون : $(0/0 \times 0/0) + (0/0 \times 0/0) + 0/0 = 0/0$

۲ " : $(0/0 \times 0/0) + (0/0 \times 0/0) + 0/0 = 0/0$

Sigmoid \Rightarrow $\begin{cases} o_1 = \text{sigmoid}(0/0) \approx 0/0 \\ o_2 = \text{sigmoid}(0/0) \approx 0/0 \end{cases}$

۱/۱ $o_1 = (0/0 - 0/0)^2 = 0/0$
۱/۱ $o_2 = (0/0 - 0/0)^2 = 0/0$
 $\Rightarrow 0/0 + 0/0 = 0/0$ (ب) هزینه کلی

Total loss = $\frac{1}{2} \times 0/0 = 0/0$

(ج) از بخش قبل (اربع) ۱، ۲ و ۳، ۴ = ۰/۰ و ۰/۰ = ۰/۰

derivation

$\text{sigmoid} \Rightarrow \sigma'(x) = \sigma(x) \times (1 - \sigma(x))$

$\sigma_j = \sum_k (\sigma_k \times w_{kj})$
error hidden j

$\begin{cases} h_1 = \sigma h_1 = (\sigma_{o1} \times w_{h1}) + (\sigma_{o2} \times w_{h2}) \\ h_2 = \sigma h_2 = (\sigma_{o1} \times w_{h3}) + (\sigma_{o2} \times w_{h4}) \end{cases}$

$\Delta w_{h1} = -\eta \times \sigma_{h1} \times \sigma'(\text{net}_{h1})$

$\sigma_{o1} = 0/0, \sigma_{o2} = 0/0$

$\sigma h_1 = (\sigma_{o1} \times w_{h1}) + (\sigma_{o2} \times w_{h2})$

$\Rightarrow \Delta w_{h1} = -\eta \times \sigma h_1 \times \sigma'(\text{net}_{h1})$

$= -0/0 \times 0/0 \times \sigma'(0/0)$



Epoch
000,047

Learning rate
0.03

Activation
Tanh

Regularization
None

Regularization rate
0

Problem type
Classification

DATA

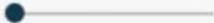
Which dataset do you want to use?



Ratio of training to test data: 50%



Noise: 0



Batch size: 10



REGENERATE

FEATURES

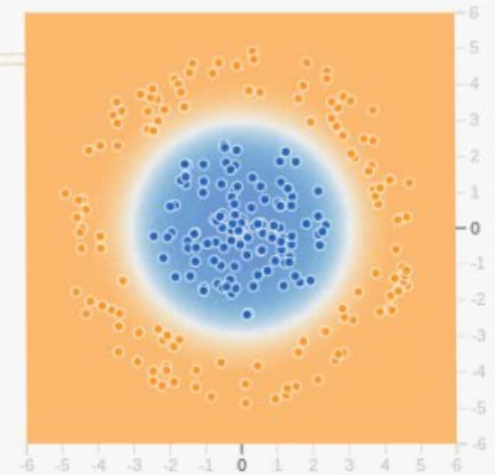
Which properties do you want to feed in?



+ - 0 HIDDEN LAYERS

OUTPUT

Test loss 0.012
Training loss 0.009



Colors shows data, neuron and weight values.



☐ Show test data

☐ Discretize output



Epoch
000,099

Learning rate
0.03

Activation
ReLU

Regularization
None

Regularization rate
0

Problem type
Classification

DATA

Which dataset do you want to use?



Ratio of training to test data: 50%



Noise: 0



Batch size: 10



REGENERATE

FEATURES

Which properties do you want to feed in?

- X_1
- X_2
- X_1^2
- X_2^2
- $X_1 X_2$
- $\sin(X_1)$
- $\sin(X_2)$

+ - 1 HIDDEN LAYER

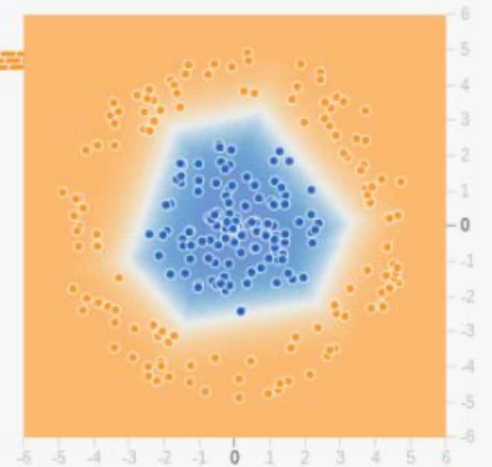
+ -

3 neurons

This is the output from one neuron. Hover to see it larger.

OUTPUT

Test loss 0.023
Training loss 0.018



Colors shows data, neuron and weight values.



☐ Show test data

☐ Discretize output



Epoch
000,161

Learning rate

0.03

Activation

Tanh

Regularization

None

Regularization rate

0

Problem type

Classification

DATA

Which dataset do you want to use?



Ratio of training to test data: 50%



Noise: 0



Batch size: 10



REGENERATE

FEATURES

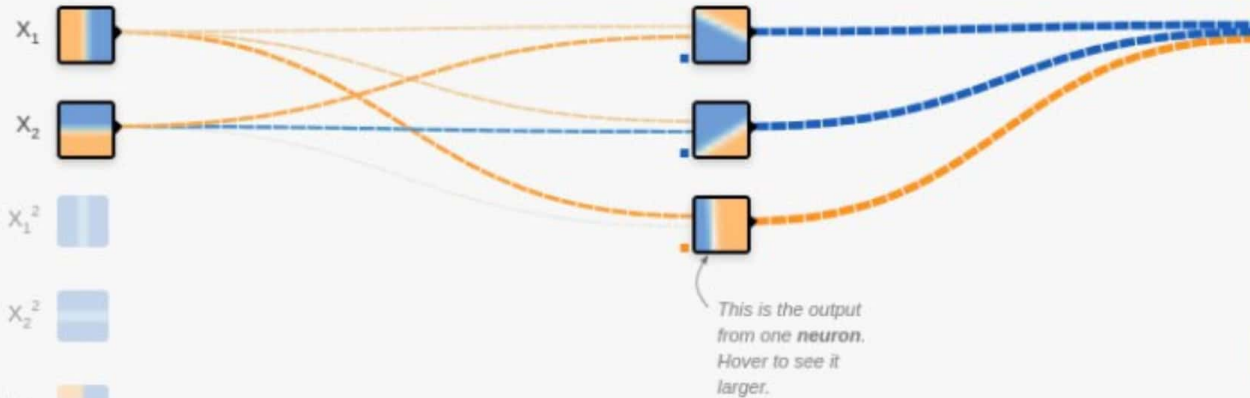
Which properties do you want to feed in?

- X_1
- X_2
- X_1^2
- X_2^2
- $X_1 X_2$
- $\sin(X_1)$
- $\sin(X_2)$

+ - 1 HIDDEN LAYER

+ -

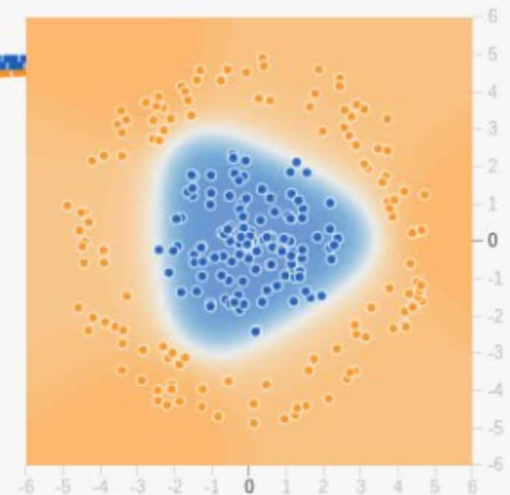
3 neurons



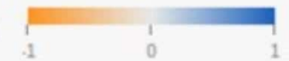
OUTPUT

Test loss 0.027

Training loss 0.022



Colors shows data, neuron and weight values.



☐ Show test data

☐ Discretize output



Epoch
000,279

Learning rate
0.03

Activation
Sigmoid

Regularization
None

Regularization rate
0

Problem type
Classification

DATA

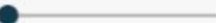
Which dataset do you want to use?



Ratio of training to test data: 50%



Noise: 0



Batch size: 10



REGENERATE

FEATURES

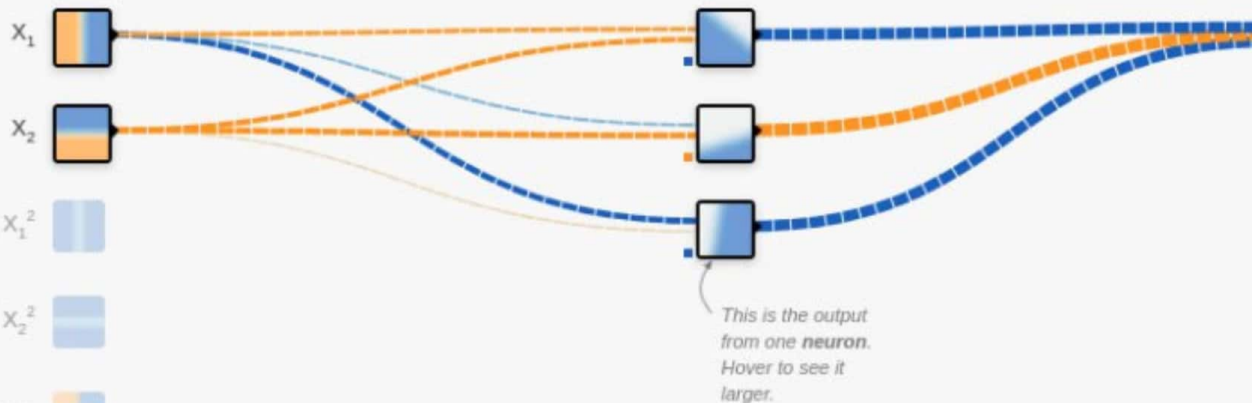
Which properties do you want to feed in?

- X_1
- X_2
- X_1^2
- X_2^2
- $X_1 X_2$
- $\sin(X_1)$
- $\sin(X_2)$

+ - 1 HIDDEN LAYER

+ -

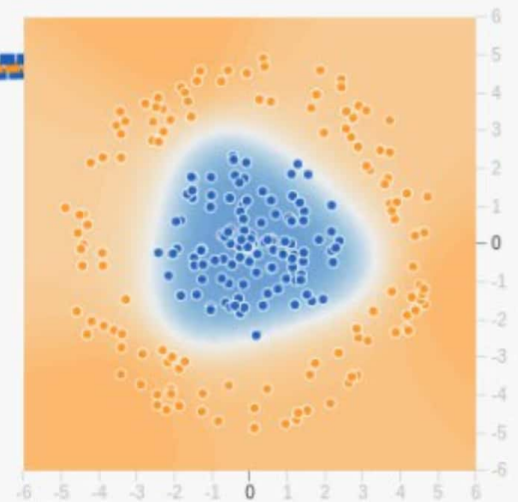
3 neurons



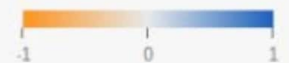
This is the output from one neuron. Hover to see it larger.

OUTPUT

Test loss 0.050
Training loss 0.043



Colors shows data, neuron and weight values.



☐ Show test data

☐ Discretize output

شبکه های عصبی:

سوال اول:

(الف) از آنجایی که داده ها از قبل به صورت خطی قابل تفکیک هستند، برای یادگیری الگوهای غیرخطی نیاز به پیچیدگی ۰ یا به بیان سبک تر، در چند صدم ثانیه همگرا می شوند؛ وقتی مرتبه ۲ ماشین (اسان) رسم بندی غیر خطی می شود و معیار مومن داخلش فاصله از مرکز می باشد و به این علت شبکه به درسی کاری کند. د این ویژگی نباید فاصله شغلی از مبدأ است)

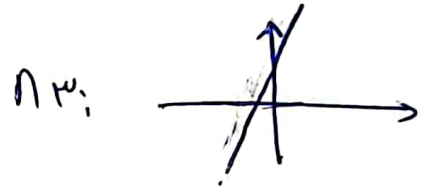
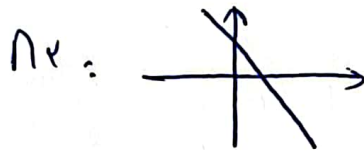
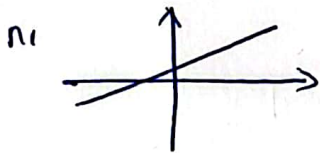
(ب)
$$X_1 w_1 + X_2 w_2 + b = 0$$
 معادله نورون ها

$$X_2 = \frac{-w_1}{w_2} X_1 - \frac{b}{w_2}$$

نارنجی ها منفی
آبی مثبت

که تا نورون داریم

ها نگونده که روی شکل مهره د است شکل عددی به نورون به این صورت است:



$$n_1 \Rightarrow X_2 = 0.94 X_1 + B$$

$$n_2 \Rightarrow X_2 = -0.97 X_1 + B$$

$$n_3 \Rightarrow 0.94 X_1 + B$$

جمع هراس ها با هم = bias final

(ب) sigmoid, tanh برای موارد پیچیده تر هستند ولی در Relu ساده سازی شده و برای موارد ساده تر به شکل خطی می sigmoid و tanh به صورت صحنی بوده و به اکسپوننسیال نرم تری دارند و به طبعی پیچیده تر هم به نتایج می باشند

سوال اول:

اشتبک پارامترها و تعداد پارامترهای گفته: در لایه های کاملاً متصل، هر نرون ارتباطات خود را با تمام نرون های لایه قبل دارد، که تعداد زیادی از پارامترها را به دنبال دارد. این سبب شد که مدل ها پیچیده تر و تعداد پارامترها زیاد شود که در آموزش و مراحل آن مشکلاتی پدید می آورد. ولی در لایه های کانفولوشن و فیلترها به ازای هر بخش از تصویر اعمال شده و پارامترهای مربوط به فیلترها به اشتباه گفته می شوند که سبب کاهش تعداد پارامترها و در نتیجه کاهش اندازه مدل و افزایش کارایی و سرعت آموزش می شود.

حفظ مشخصات مکانی: لایه های کانفولوشن و ادغام قادرند اطلاعات مکانی مهم تصویر را حفظ کنند. با اعمال فیلترها و ادغام در تصویرها، ویژگی های مهم مکانی مانند لبه ها و گوشه ها و الگوهای مختلف به خوبی استخراج می شوند و مشخصات مکانی تصویر حفظ می گردد. در لایه های کاملاً متصل، این اطلاعات مکانی از دست رفته و مدل ها قادر به تحلیل دقیق تر تب و مکان و ویژگی ها نیستند که تاثیر منفی بر عملکرد دارد.

سوال دوم:

$$\begin{bmatrix} 1 & 4 & 6 & 5 & 9 & 2 & 3 \end{bmatrix}$$

ماتریس:

$$\begin{bmatrix} -2 & 11 & -2 & 11 \end{bmatrix}$$

$$\Rightarrow \begin{cases} f_1 \times 1 + f_2 \times 4 + f_3 \times 0 = -2 \\ f_1 \times 4 + f_2 \times 0 + f_3 \times (-2) = -2 \\ f_1 \times 0 + f_2 \times (-2) + f_3 \times 3 = 11 \end{cases}$$

دومین
سومین

$$\Rightarrow \begin{cases} f_1 + 4f_2 = -2 \\ 4f_1 - 2f_3 = -2 \\ -2f_2 + 3f_3 = 11 \end{cases}$$

$$\begin{cases} f_1 = \frac{14}{13} \\ f_2 = \frac{-10}{13} \\ f_3 = \frac{41}{13} \end{cases}$$

$$\Rightarrow \left[\frac{14}{13}, \frac{-10}{13}, \frac{41}{13} \right]$$