

Homework 2

Parsa Eissazadeh 97412364

Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exerci tation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan.

Lorem ipsum

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam,

سوال 2

برای هر داده یک forward propagation داریم که مقدار تابع loss را حساب میکند و یک backward propagation که مشتق را حساب می کند .
برای backward propagation در هر لایه مشتق نسبت به لایه قبل را حساب میکنیم سپس طبق قانون chain rule آن ها را در هم ضرب می کنیم .

داده اول $x_1 = 3$ و $x_0 = 3$:

مشتق y نسبت به w_0 برابر با ۳ و نسبت به w_1 برابر با -1 است .

مقادیر وزن ها $w_1=1$ و $w_0=2$ و $b=2$

Forward propagation :

$$Z = 3 * 2 + 1 * -1 + 2 = 7 \Rightarrow \text{sigmoid}(7) = 0.9990889488055994$$

$$\text{Loss function} = (\text{label} - \text{output})^2 = (1 - 0.9990889488055994)^2 =$$

$$8.300142788187649 \times 10^{-7}$$

Backward propagation :

$$\text{MSE}' = 2 (\text{label} - \text{output}) = 2 * (0.9990889488055994 - 1) = -0.0018221023888012056$$

$$\text{Sigmoid}' = (1 - \text{sigmoid}) * \text{sigmoid} = (1 - 0.9990889488055994) * 0.9990889488055994 = 0.000910221180121784$$

مشتق نسبت به W_0 :

$$-0.0018221023888012056 * 0.000910221180121784 * 3 = -4.975548559912065 \times 10^{-6}$$

مشتق نسبت به W_1 :

$$-0.0018221023888012056 * 0.000910221180121784 * -1 = 1.6585161866373552e \times 10^{-6}$$

آپدیت کردن مقادیر :

$$W_0^{\text{new}} = W_0 - 10^{-3} * -4.975548559912065 \times 10^{-6} = 2 + 4.975548559912065 \times 10^{-9} =$$

$$2.0000000049755484$$

$$W_1^{\text{new}} = W_1 - 10^{-3} * 1.6585161866373552e \times 10^{-6} = 0.9999999983414838$$

=====

داده دوم $x_1 = -2$ و $x_0 = 1$:

مشتق y نسبت به w_0 برابر با 1 و نسبت به W_1 برابر با -2 است .

مقادیر وزن ها $w_0 = 2.0000000049755484$ و $w_1 = 0.9999999983414838$ و $b = 2$.

سوال 4

از آنجایی که کلا دو کلاس داریم ، یک نرون هم کافی است . زیرا احتمالی برای سگ بودن عکس وجود دارد و احتمال گربه بودن برابر است با 1 منهای آن مقدار .

گزینه ها برای activation function لایه آخر :

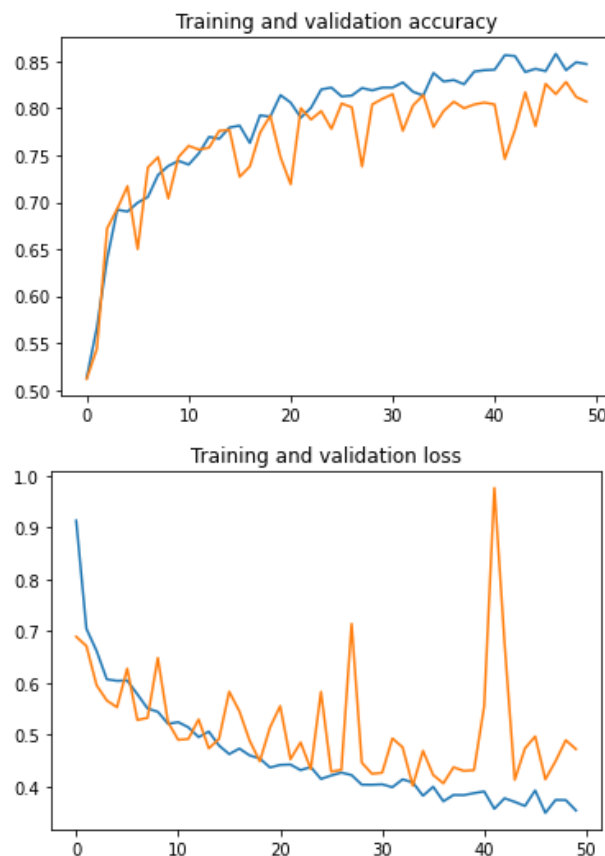
- سیگموید
- تانژانت هایپربولیک

برای loss function از binray_crossentropy استفاده میکنیم .

حال activation function های مختلف را امتحان میکنیم .

سیگموید :

بعد از اجرا کردن مدل به نتایج زیر رسیدیم :

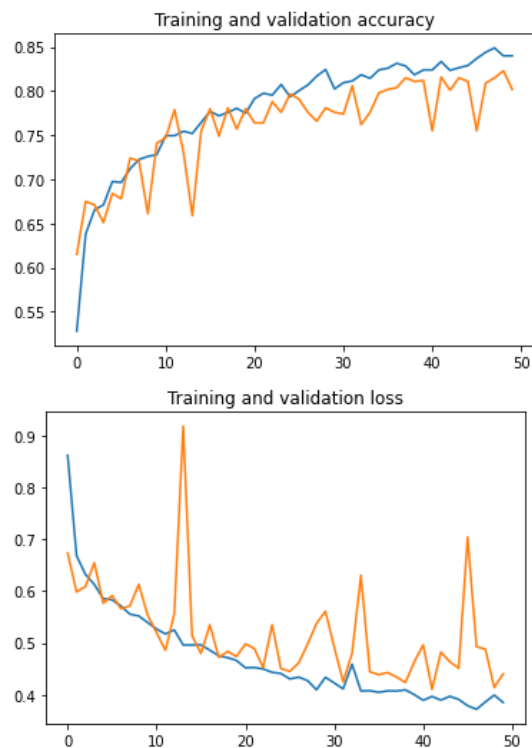


تانژانت هایپربولیک :

وقتی از تانژانت هایپربولیک استفاده کردیم ، نتایج هیچ پیشرفتی نکرد .

```
100/100 [=====] - 28s 175ms/step - loss: 7.6095 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 2/50
100/100 [=====] - 18s 182ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 3/50
100/100 [=====] - 17s 173ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 4/50
100/100 [=====] - 17s 174ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 5/50
100/100 [=====] - 18s 176ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 6/50
100/100 [=====] - 18s 176ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 7/50
100/100 [=====] - 17s 174ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 8/50
100/100 [=====] - 17s 175ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 9/50
100/100 [=====] - 18s 175ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 10/50
100/100 [=====] - 18s 175ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 11/50
100/100 [=====] - 18s 175ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 12/50
100/100 [=====] - 17s 173ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 13/50
100/100 [=====] - 17s 175ms/step - loss: 7.6246 - acc: 0.5000 - val_loss: 7.6246 - val_acc: 0.5000
Epoch 14/50
```

همچنین میتوان از دو نورون در لایه خروجی استفاده کرد . یکی برای سگ یکی برای گربه . در حالت دو نورون از softmax استفاده میکنیم.



نتیجه تقریباً مشابه شد و دقت اول و آخر یکی بود و همچنین سرعت converge شدن هم مشابه بود .