

الف) نرمال کردن داده ها وقتی خوب است که توزیع داده ها از توزیع گوسی پیروی نکند و استاندارد کردن داده ها می تواند راه گشای باشد و وقتی که داده اولیه ما دارای توزیع گوسی باشد.

ب) برای جلوگیری از فرارزش برای بخش یادگیری داده ها می توان از Dropout استفاده کرد. این به گونه ای عمل می کند مثلاً بین ۲ لایه مذکور یک لایه Dropout تعیین می کنیم و باید توزیع احتمالی خاصی داشته باشیم Bernoulli برض از اتصال ها حذف می کنیم.

ج) به طور معمول باید مقدار Loss داده های train, validation به هم شبیه باشد و مقدار Loss داده Validation می تواند مقدار بزرگتر باشد. اگر این اتفاق نیفتد یا مدل کم یادگیری کرده است یا فرارزش داریم.

$$\text{training loss} \gg \text{validation loss} \Rightarrow \text{underfitting}$$

$$\text{validation loss} \gg \text{training loss} \Rightarrow \text{overfitting}$$

برای جبران این کار می توانیم از: ۱) مدل کم یاد گرفته ات تعداد لایه ها و یا پیچیدگی آن را افزایش دهیم یا مقدار epoch را افزایش دهیم ۲) اگر فرارزش داشته باشیم از Dropout می توانیم استفاده کنیم.

$$a=1, b=4 \quad (>)$$

$$\begin{aligned} w_{1,1,1} &= 0,11 & H_1 &= in_1 \times w_{1,1,1} + in_2 \times w_{1,2,1} + b_{1,1} = 0,11 - 0,45 + 0,1 = -0,44 \\ w_{1,1,2} &= 1,1 & H_2 &= in_1 \times w_{1,1,2} + in_2 \times w_{1,2,2} + b_{1,2} = 1,1 - 1,4 - 0,2 = -0,5 \\ w_{1,2,2} &= 1,4 & H_3 &= in_1 \times w_{1,1,3} + in_2 \times w_{1,2,3} + b_{1,3} = 0,14 - 0,41 + 0 = -0,27 \\ w_{1,1,3} &= 0,14 & H_{1out} &= \tanh(H_1) = -0,4136 \quad H_{2out} = \tanh(H_2) = -0,4621 \quad H_{3out} = \tanh(H_3) = -0,2636 \\ w_{1,2,3} &= 0,41 \end{aligned}$$

$$\begin{aligned} w_{2,1,1} &= 1,1 & y_{out} &= H_{1out} \times w_{2,1,1} + H_{2out} \times w_{2,2,1} + H_{3out} \times w_{2,3,1} + b_2 = -2,0729 \\ w_{2,2,1} &= 4,3 & & \\ w_{2,3,1} &= -1,4 & & \end{aligned}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y_{out,i})^2 = (y - y_{out})^2 = 16,5889$$

$$b_{20} \frac{dMSE}{db_2} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{db_2} = -8,1458 \rightarrow b_2 = 0,3 - (0,1) \times (-8,1458) = 1,11458$$

$$w_{2,3,1} \frac{dMSE}{dw_{2,3,1}} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dw_{2,3,1}} = 2,1472 \rightarrow w_{2,3,1} = -1,4 - (0,1) \times (2,1472) = -1,6147$$

$$w_2[2,1] : \frac{dMSE}{dw_2[2,1]} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dw_2[2,1]} = 3,7641 \rightarrow w_2[2,1] = f, x - 0,3764 = 3,9236$$

$$w_2[1,1] : \frac{dMSE}{dw_2[1,1]} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dw_2[1,1]} = 3,3671 \rightarrow w_2[1,1] = 1,1 - 0,3367 = 0,7631$$

$$b_1[3] : \frac{dMSE}{db_1[3]} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dH_{3out}} \times \frac{dH_{3out}}{dH_3} \times \frac{dH_3}{b_1[3]} = 1,9979 \rightarrow b_1[3] = 0 - 0,1997 = -0,1997$$

$$b_1[2] : \frac{dMSE}{db_1[2]} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dH_{2out}} \times \frac{dH_{2out}}{dH_2} \times \frac{dH_2}{b_1[2]} = 3,5006 \rightarrow -0,2 - 0,3500 = -0,55$$

$$b_1[1] : \frac{dMSE}{db_1[1]} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dH_{1out}} \times \frac{dH_{1out}}{dH_1} \times \frac{dH_1}{b_1[1]} = 2,7926 \rightarrow 0,1 - 0,2792 = -0,1792$$

$$w_1[2,3] : \frac{dMSE}{dw_1[2,3]} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dH_{3out}} \times \frac{dH_{3out}}{dH_2} \times \frac{dH_2}{w_1[2,3]} = -1,9979 \Rightarrow 0,41 + 0,1997 = 0,6097$$

$$w_1[2,2] : \frac{dMSE}{dw_1[2,2]} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dH_{2out}} \times \frac{dH_{2out}}{dH_2} \times \frac{dH_2}{w_1[2,2]} = -3,5006 \Rightarrow 1,1 - 0,3500 = 0,75$$

$$w_1[2,1] : \frac{dMSE}{dw_1[2,1]} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dH_{1out}} \times \frac{dH_{1out}}{dH_1} \times \frac{dH_1}{w_1[2,1]} = -2,7926 \rightarrow 1,1 + 0,27 = 1,37$$

$$w_1[1,3] : \frac{dMSE}{dw_1[1,3]} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dH_{3out}} \times \frac{dH_{3out}}{dH_1} = 1,9979 \rightarrow 0,14 - 0,1997 = -0,0597$$

$$w_1[1,2] : \frac{dMSE}{dw_1[1,2]} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dH_{2out}} \times \frac{dH_{2out}}{dH_3} \times \frac{dH_3}{w_1[1,2]} = 3,5006 \rightarrow 1,2 - 0,3500 = 0,85$$

$$w_{1,[1,1]} = \frac{dMSE}{dw_{1,[1,1]}} = \frac{dMSE}{dy_{out}} \times \frac{dy_{out}}{dH_{2out}} \times \frac{dH_{2out}}{dH_1} \times \frac{dH_1}{dw_{1,[1,1]}}$$

$$= 2,7926 \rightarrow 0,11 - 0,2792 = -0,1273$$
