

بسمه تعالیٰ

بازشناسی آماری الگو



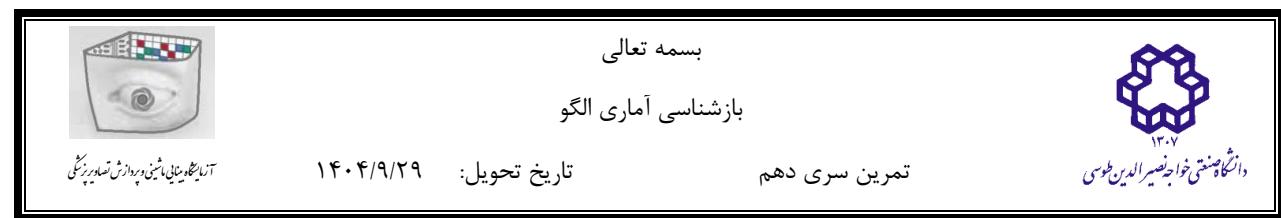
دانشگاه علوم پزشکی اسلامی طوسی

آستانه علمی پژوهشی پرورش پرگز

۱۴۰۴/۹/۲۹

تاریخ تحويل:

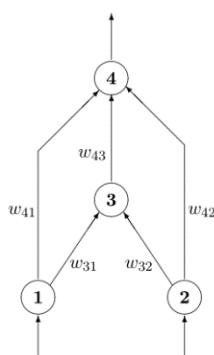
تمرین سری دهم



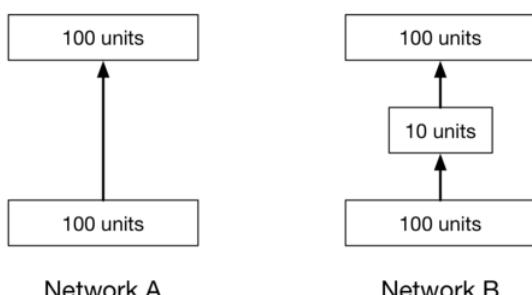
2- (اختیاری) شبکه MLP زیر را در نظر بگیرید.

ورودی و خروجی همه نورون‌ها باینری است (صفر یا یک). وزن‌های شبکه به شکل زیر است:

$$w_{31} = 1, \quad w_{32} = 1, \quad w_{41} = -1, \quad w_{42} = -1, \quad w_{43} = 3$$



آستانه نورون شماره ۳ برابر با ۱.۵ است. آستانه نورون شماره ۴ برابر با ۰.۵ است. آستانه نورون‌های شماره ۱ و ۲ برابر با ۰.۵ است (پس خروجی این دو نورون با ورودی آنها برابر است). این شبکه معادل کدام یک از توابع Boolean است؟



3- (اختیاری) دو شبکه MLP زیر را در نظر بگیرید. همه لایه‌ها از توابع فعالیت خطی استفاده می‌کنند.

الف. مزیت شبکه A را نسبت به شبکه B بیان کنید.

ب. مزیت شبکه B را نسبت به شبکه A بیان کنید.

4- (اختیاری) مساله ۱ را برای تابع فعالیت softmax که بصورت زیر تعریف می‌شود:

$$\hat{y}_k = \frac{e^{v_k^L}}{\sum_{k'} e^{v_{k'}^L}}$$

حل کنید و نشان دهید:

$$\delta_j^L(i) = \hat{y}_j(i) - y_j(i)$$

۵- (اختیاری) چنانچه تابع هزینه زیر در یک شبکه عصبی پرسپترون چند لایه استفاده شود:

$$J = \frac{1}{2} \sum_{i=1}^N \sum_{m=1}^{k_L} (\hat{y}_m(i) - y_m(i))^2$$

عناصر ماتریس هسیان (Hessian Matrix) که توسط رابطه زیر تعریف می شود

$$\frac{\partial^2 J}{\partial w_{kj}^r \partial w_{k'j'}^{r'}}$$

را محاسبه کنید و نشان دهید که در نزدیکی یک نقطه کمینه تابع انرژی مقدار تقریبی آن بصورت زیر قابل محاسبه است:

$$\frac{\partial^2 J}{\partial w_{kj}^r \partial w_{k'j'}^{r'}} = \sum_{i=1}^N \sum_{m=1}^{k_L} \frac{\partial \hat{y}_m(i)}{\partial w_{kj}^r} \frac{\partial \hat{y}_m(i)}{\partial w_{k'j'}^{r'}}$$

یعنی مشتقات مرتبه دوم می تواند با حاصلضرب مشتقات مرتبه اول تقریب زده شود. سپس با استفاده از روش مشابه با الگوریتم پس انتشار خط نشان دهید:

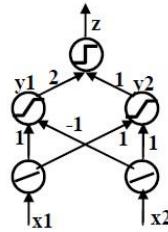
$$\frac{\partial \hat{y}_m(i)}{\partial w_{kj}^r} = \hat{\delta}_{jm}^r y_k^{r-1}$$

که در آن:

$$\hat{\delta}_{jm}^r = \frac{\partial \hat{y}_m(i)}{\partial v_j^r(i)}$$

۶- (اختیاری) مساله امتحان پایان ترم بازشناسی آماری الگو سال ۱۳۸۶

شبکه عصبی سه لایه زیر را در نظر بگیرید که در آن وزن‌های ارتباطی نرونها نیز مشخص شده است.



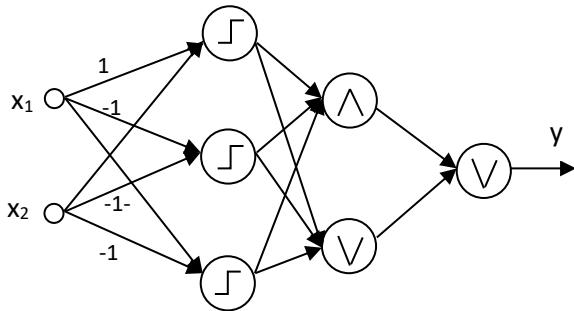
روابط ورودی-خروجی برای نرون‌های لایه ورودی ( $i$ )، مخفی ( $h$ ) و خروجی ( $o$ ) بصورت زیر می‌باشد:

$$o_i = i_i, \quad o_h = \begin{cases} +1 & i_h > 1 \\ i_h & -1 < i_h < 1, \\ -1 & i_h < -1 \end{cases}, \quad o_o = \begin{cases} +1 & i_o > 0 \\ -1 & i_o < 0 \end{cases}$$

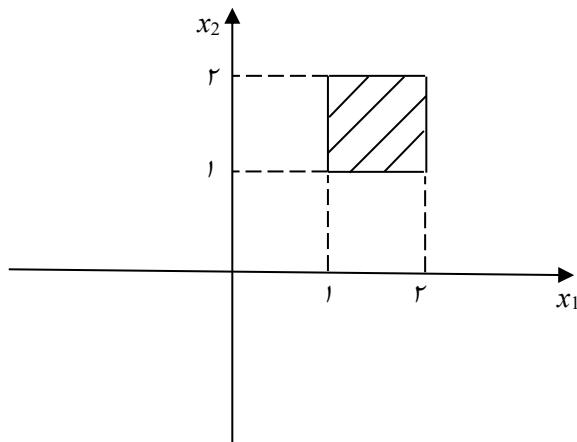
در حالت دو کلاسه (با فرض  $\omega_1 > 0 \Rightarrow d(z) = \omega_1$ ، نواحی تصمیم‌گیری را بازای این دو کلاس در فضای ویژگی دو بعدی با ذکر دلیل رسم نمایید.

۷- (اختیاری) سوال امتحانی پایان ترم ۱۳۹۲

شبکه عصبی پیشخور (feedforward) شکل زیر با دو لایه مخفی را در نظر بگیرید که ورودی آن بردارهای  $X$  در فضای دو بعدی و خروجی آن مشخص کننده یکی از دو کلاس  $\omega_1$  و  $\omega_2$  می‌باشد. نرون‌های لایه میانی اول دارای تابع فعالیت از نوع پله واحد و نرون‌های لایه میانی دوم دارای تابع فعالیت منطقی "و" و "یا" می‌باشند. وزنهای ارتباطی نرون‌ها (به جز دو مورد مشخص شده در شکل) همگی برابر واحد هستند و فرض می‌شود نرون‌های لایه میانی اول دارای مقدار بایاس برابر واحد باشند. نرون‌های لایه مخفی دوم و لایه آخر بدون بایاس هستند. مطلوب است رسم مرز بین دو کلاس در فضای دو بعدی با ذکر دلیل و انجام محاسبات لازم.



۸- اختیاری (سوال امتحانی پایان ترم ۹۶) یک شبکه عصبی از نوع چندلایه پیشرو طراحی نمایید (تعداد لایه‌ها، تعداد نورنها در لایه‌ها، فعالیت هر نuron، ضرایب ارتباطی نورنها در لایه‌های مختلف را مشخص کنید) طوری که فضای ویژگی دو بعدی را بصورت شکل زیر در دو کلاس هاشور خورده (داخل مربع) و بدون هاشور (خارج مربع) دسته‌بندی نماید.



۹- (سوال امتحانی پایان ترم ۱۳۹۷)

می‌خواهیم یک شبکه عصبی را برای رانندگی بک خودرو آموزش دهیم. داده‌های آموزش شامل تصاویر سطوح خاکستری به ابعاد  $64 \times 64$  پیکسل می‌باشد. خروجی شبکه شامل زاویه گردش فرمان (درجه) و سرعت (کیلومتر بر ساعت) است. شبکه عصبی دارای یک لایه ورودی با ۴۰۹۶ نuron، یک لایه پنهان با ۱۰۲۴ نuron و یک لایه خروجی با ۲ نuron (یکی زاویه گردش فرمان و دیگری سرعت) است. تابع فعالیت نورنها لایه‌های ورودی و خروجی از نوع خطی (همانی) و لایه پنهان از نوع ReLU می‌باشد.

$$RELU: f(y) = \max(0, y), \quad f'(y) = \begin{cases} 1 & y > 0 \\ 0 & y \leq 0 \end{cases}$$

(چنانچه آرگومان تابع  $f$  بردار باشد، روابط فوق به تک تک عناصر بردار قابل اعمال است.)

الف- تعداد پارامترها (وزنهای) این شبکه را بدست آورید (مقادیر بایاس نuron‌ها را نیز به عنوان پارامتر در نظر بگیرید). برای پاسخ خود دلیل ذکر کنید.

ب- تابع هزینه شبکه به ازای یک نمونه ورودی بصورت زیر تعریف می‌شود:

$$J(V, W) = \frac{1}{2} \|y - \hat{y}\|^2$$

که در آن  $y$  بردار ستونی خروجی مطلوب،  $\hat{y}$  بردار ستونی خروجی شبکه،  $V$  ماتریس وزن‌های ارتباطی نورنها لایه ورودی به لایه پنهان و  $W$  ماتریس وزن‌های ارتباطی نورنها لایه پنهان به لایه خروجی می‌باشد و داریم:

$$\begin{aligned} g &= Vx \\ h &= f(g) \\ \hat{y} &= Wh \end{aligned}$$

در روابط فوق  $\mathbf{x}$  بردار ستونی معرف نمونه ورودی گسترش یافته (یک عنصر اضافی با مقدار ۱ به انتهای آن اضافه شده)،  $\mathbf{g}$  بردار ستونی معرف آرگومان نزونهای لایه میانی و  $\mathbf{h}$  بردار ستونی معرف خروجی نزونهای لایه میانی (با یک عنصر اضافی با مقدار ۱ که به انتهای آن اضافه شده) می‌باشد. ابعاد  $\mathbf{x}, \mathbf{y}, \hat{\mathbf{y}}, \mathbf{W}, \mathbf{V}, \mathbf{g}$  و  $\mathbf{h}$  را مشخص کنید.

ج- مطلوب است محاسبه مشتق تابع هزینه نسبت به عنصر  $j$ -ام ماتریس  $\mathbf{W}$ :

$$\frac{\partial J}{\partial W_{ij}}$$

برحسب  $\hat{\mathbf{y}}_i, \mathbf{y}_i$  و  $\mathbf{h}_j$  که در آن اندیس‌های  $i$  و  $j$  به عنصر  $i$ -ام و  $j$ -ام بردار اشاره می‌کنند.

د- مطلوب است محاسبه:

$$\frac{\partial J}{\partial W}$$

برحسب  $\hat{\mathbf{y}}$ ,  $\mathbf{y}$  و  $\mathbf{h}$ . (راهنمایی):  $J$  اسکالر است و مشتق آن نسبت به ماتریس  $\mathbf{W}$  ماتریسی با ابعاد برابر با  $\mathbf{W}$  خواهد شد)

ه- مطلوب است محاسبه:

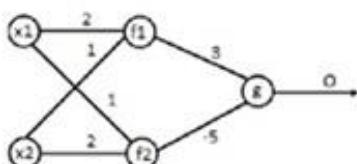
$$\frac{\partial J}{\partial V_{ij}}$$

برحسب  $\hat{\mathbf{y}}, \mathbf{y}_i, \mathbf{W}, \mathbf{y}_j$  و  $\mathbf{g}_i$

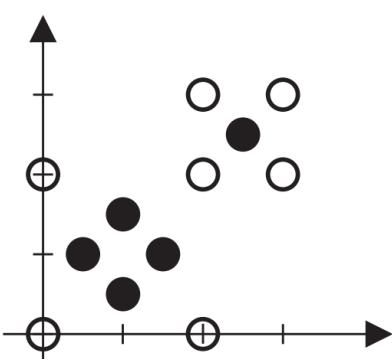
و- با فرض نرخ آموزش برابر  $\eta$  و نتایج بندهای قبل، روابط بازگشتی به روزسازی وزنهای شبکه  $(V_{ij} \text{ و } W_{ij})$  را برحسب متغیرهای شبکه بنویسید

۱۰- (اختیاری) سوال امتحانی پایان ترم ۱۳۹۸

در شبکه عصبی ذیل، که در آن  $g(\text{net}) = \text{net}$  و  $f_t(\text{net}) = \frac{1}{1+e^{-a_t(\text{net})}}$  است. قصد داریم تا ضرب  $\mathbf{a}$  را نیز در هر گام به کمک الگوریتم پس انتشار خطأ آموزش دهیم. اگر در این لحظه  $[1.1] = a_t$  باشد و بردار ویژگی  $X_{t+1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$  با مقدار مطلوب  $0 = y_{t+1}$  از راه بررسد، محلولیست محاسبه  $a_{t+1}$  ؟ (نرخ آموزش را  $0.5$  در نظر بگیرید)



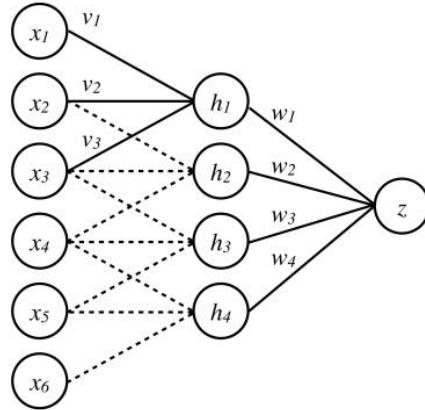
۱۱- (مساله امتحانی پایان ترم ۱۴۰۰) شکل زیر یک مسئله دسته بندی را نشان میدهد که در آن نمونه‌های کلاس ۱ با دایره‌های مشکی و نمونه‌های کلاس ۲ با دایره‌های سفید نمایش داده شده‌اند. همانطور که مشخص است این دو دسته به صورت خطی قابل جداسازی نیستند. شبکه عصبی RBF را طوری طراحی کنید که این دسته بندی را انجام دهید، سپس نواحی تصمیم‌گیری آن را به صورت تقریبی رسم کنید. ( واضح است که شبکه شما در لایه ورودی باید دو نورون و در لایه خروجی یک نورون داشته



باشد).

## ۱۲- مساله امتحانی پایان ترم ۱۳۹۹

Consider this convolutional neural network architecture.



In the first layer, we have a one-dimensional convolution with a single filter of size 3 such that  $h_i = s\left(\sum_{j=1}^3 v_j x_{i+j-1}\right)$ . The second layer is fully connected, such that  $z = \sum_{i=1}^4 w_i h_i$ . The hidden units' activation function  $s(x)$  is the logistic (sigmoid) function with derivative  $s'(x) = s(x)(1 - s(x))$ . The output unit is linear (no activation function). We perform gradient descent on the loss function  $R = (y - z)^2$ , where  $y$  is the training label for  $x$ .

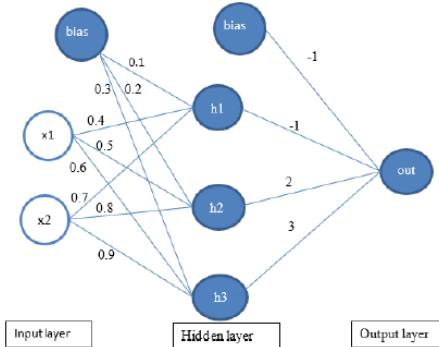
- (a) What is the total number of parameters in this neural network? Recall that convolutional layers share weights.  
There are no bias terms.
- (b) Compute  $\partial R / \partial w_i$ .
- (c) Vectorize the previous expression—that is, write  $\partial R / \partial w$ .
- (d) Compute  $\partial R / \partial v_j$ .

## ۱۳- اختیاری (پایان ترم ۱۴۰۱)

الف- شبکه عصبی شکل مقابل را در نظر بگیرید که در آن در هر دو لایه مقادیر بایاس برابر ۱- هستند. همچنینتابع فعالیت نرون‌های لایه مخفی از نوع سیگموید و نرخ آموزش برابر ۰.۱ است. چنانچه نمونه ورودی بصورت زیر باشد:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

در اینصورت خروجی نرون مخفی  $h_3$  را بدست آورید.



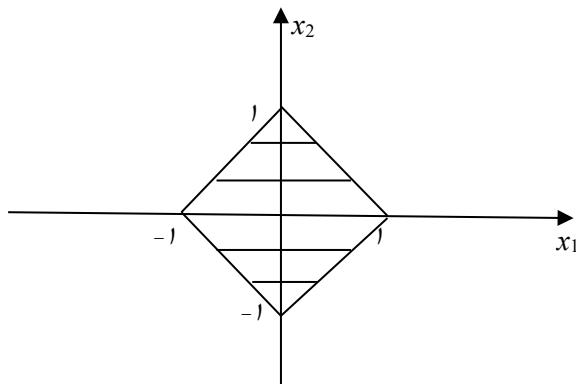
ب- چنانچه نرون لایه خروجی تابع فعالیت همانی داشته باشد خروجی آن به ازای نمونه ورودی زیر چه خواهد بود؟

$$\mathbf{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

ج- با فرض اینکه پیش‌بینی صحیح مقدار خروجی به ازای  $t=5$  باشد در اینصورت وزن ارتباطی نرون  $h_3$  در لایه مخفی با نرون لایه خروجی پس از به روز رسانی بوسیله پس انتشار خطأ چقدر خواهد بود؟

د- در فرض «ج» وزنهای ارتباطی نرون  $x_2$  به نرون  $h_3$  را پس از به روز رسانی بوسیله پس انتشار خطأ بدست آورید.

۱۴- (اختیاری) (پایان ترم ۱۴۰۲) یک شبکه عصبی از نوع چندلایه پیش رو طراحی نمایید (تعداد لایه ها، تعداد نرونهای هر لایه، تابع فعالیت هر نuron، ضرایب ارتباطی نرونها در لایه های مختلف را مشخص کنید) طوری که فضای ویژگی دو بعدی را بصورت شکل زیر در دو کلاس هاشور خورده (داخل لوزی) و بدون هاشور (خارج لوزی) دسته بندی نماید.



۱۵- (پایان ترم ۱۴۰۳) تابع زیر را به عنوان تابع فعالسازی شبکه عصبی در نظر بگیرید:

$$\phi(x) = \begin{cases} 1 & x \geq 0, \\ 0 & O.W. \end{cases}$$

الف) تابع AND با تعداد n ورودی را در نظر بگیرید. شبکه عصبی ای رسم کنید که با حداقل تعداد لایه ممکن، این تابع را پیاده سازی نماید (ورودی شبکه نیز به صورت باینری می باشد).

ب) تابع OR با تعداد n ورودی را در نظر بگیرید. شبکه عصبی ای رسم کنید که با حداقل تعداد لایه ممکن، این تابع را پیاده سازی نماید (ورودی شبکه نیز به صورت باینری می باشد).

ج) نشان دهید که تابع XOR با دو ورودی را نمی توان با شبکه ای با کمتر از دو لایه پیاده سازی نمود.

د) به نظر شما چرا در شبکه های عصبی از تابع فعالسازی به صورت تابع داده شده در این سوال استفاده نمی شود؟ مزیت استفاده از تابع فعالساز sigmoid نسبت به این تابع چیست؟

موفق باشید