

آزمایشگاه ملی ماشین بردارشی تصویر پزشکی

بسمه تعالی

بازشناسی آماری الگو



دانشگاه صنعتی خواجه نصیرالدین طوسی

۱۴۰۴/۹/۲۹

تاریخ تحویل:

تمرین سری دهم

۱- (اختیاری) اگر برای تابع هزینه یک شبکه پرسپترون چندلایه از رابطه آنتروپی عرضی (cross entropy) به فرم زیر استفاده کنیم:

$$J = - \sum_{i=1}^N \sum_{k=1}^{k_L} y_k(i) \ln \frac{\hat{y}_k(i)}{y_k(i)}$$

و تابع فعالیت تابع سیگموئید باشد، نشان دهید که مقدار گرادیان در لایه آخر بصورت زیر بدست می آید:

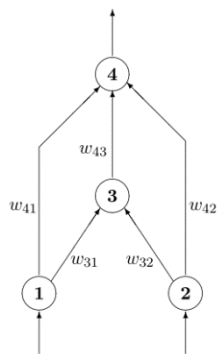
$$\delta_j^L(i) = a(1 - \hat{y}_j(i)) y_j(i)$$

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

ورودی و خروجی همه ی نورون ها باینری است (صفر یا یک). وزن های شبکه به شکل زیر است:

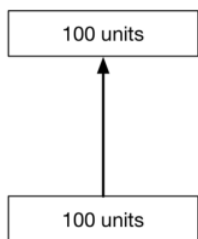
$$w_{31} = 1, \quad w_{32} = 1, \quad w_{41} = -1, \quad w_{42} = -1, \quad w_{43} = 3$$

آستانه نورون شماره ۳ برابر با ۱.۵ است. آستانه نورون شماره ۴ برابر با ۰.۵ است. آستانه نورون های 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؟

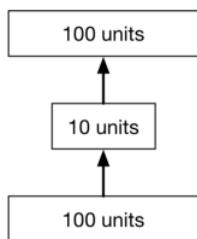


۳- (اختیاری) دو شبکه MLP زیر را در نظر بگیرید. همه لایه ها از توابع

فعالیت خطی استفاده می کنند.



Network A



Network B

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

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

۴- (اختیاری) مساله ۱ را برای تابع فعالیت softmax که بصورت زیر تعریف می شود:

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

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 خطا نشان دهید:

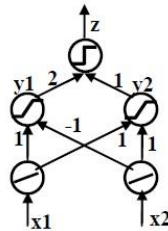
$$\frac{\partial \hat{y}_m(i)}{\partial w_{kj}^r} = \delta_{jm}^r y_k^{r-1}$$

که در آن:

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

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

شبکه عصبی سه لایه شکل زیر را در نظر بگیرید که در آن وزن های ارتباطی نرون ها نیز مشخص شده است.



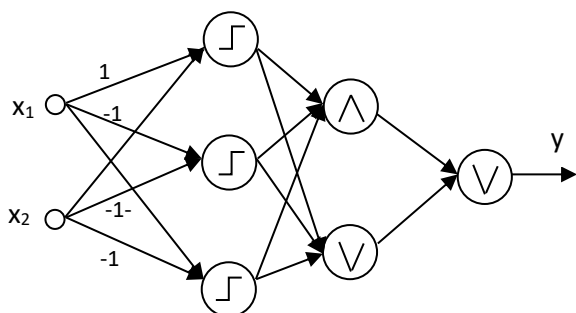
روابط ورودی-خروجی برای نرون های لایه ورودی (i)، مخفی (h) و خروجی (o) بصورت زیر می باشد:

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

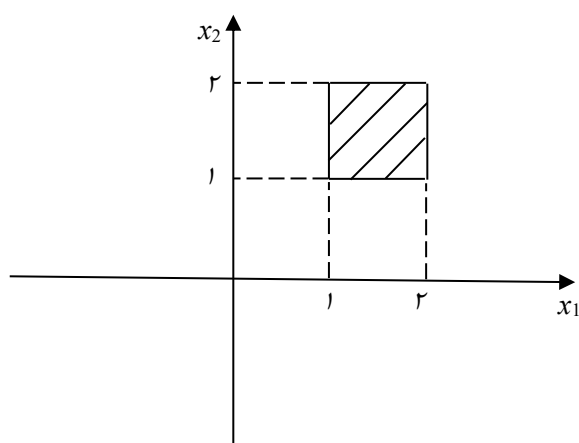
در حالت دو کلاسه ($d(z) = \omega_1$ با فرض $z > 0$)، نواحی تصمیم گیری را برای این دو کلاس در فضای ویژگی دو بعدی با ذکر دلیل رسم نمایید.

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 دوبعدی و خروجی آن مشخص کننده یکی از دو کلاس ω_1 و ω_2 می باشد. نرون های لایه میانی اول دارای تابع فعالیت از نوع پله واحد و نرون های لایه میانی دوم دارای تابع فعالیت منطقی "و" و "یا" می باشند. وزن های ارتباطی نرون ها (به جز دو مورد مشخص شده در شکل) همگی برابر واحد هستند و فرض می شود نرون های لایه میانی اول دارای مقدار بایاس برابر واحد باشند. نرون های لایه مخفی دوم و لایه آخر بدون بایاس هستند. مطلوب است رسم مرز بین دو کلاس در فضای دوبعدی 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 نرونهای هر لایه، تابع فعالیت هر نرون، ضرایب ارتباطی نرونها در لایه‌های مختلف را مشخص کنید) طوری که فضای ویژگی دوبعدی 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 یک خودرو آموزش دهیم. داده‌های آموزش شامل تصاویر سطوح خاکستری به ابعاد 64×64 پیکسل می‌باشد. خروجی شبکه شامل زاویه گردش فرمان (درجه) و سرعت (کیلومتر بر ساعت) است. شبکه عصبی دارای یک لایه ورودی با $64 \times 64 \times 96 = 40$ نرون، یک لایه پنهان با 1024 نرون و یک لایه خروجی با 2 نرون (یکی زاویه گردش فرمان و دیگری سرعت) است. تابع فعالیت نورنهای لایه‌های ورودی و خروجی از نوع خطی (همانی) و لایه پنهان از نوع ReLU می‌باشد.

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

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

الف- تعداد پارامترها (وزنهای) این شبکه را بدست آورید (مقادیر بایاس نرون‌ها 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:

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

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

$$\begin{aligned} \mathbf{g} &= V\mathbf{x} \\ \mathbf{h} &= f(\mathbf{g}) \\ \hat{\mathbf{y}} &= W\mathbf{h} \end{aligned}$$

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 نرونهای لایه میانی و \mathbf{h} بردار ستونی معرف خروجی نرونهای لایه میانی (با یک عنصر اضافی با مقدار ۱ که به انتهای آن اضافه شده) می‌باشد. ابعاد $\mathbf{x}, \mathbf{y}, \hat{\mathbf{y}}, \mathbf{V}, \mathbf{W}, \mathbf{g}, \mathbf{h}$ را مشخص کنید.

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

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

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

د- مطلوب است محاسبه:

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

برحسب $\hat{\mathbf{y}}, \mathbf{y}, \mathbf{h}$. (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 $\hat{\mathbf{y}}, \mathbf{y}, \mathbf{W}, \mathbf{g}$ و \mathbf{x}_j

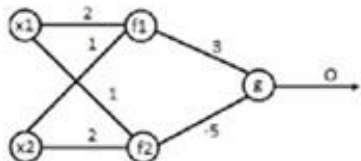
و- با فرض نرخ آموزش برابر η و نتایج بندهای قبل، روابط بازگشتی به‌روزرسانی وزنه‌های شبکه (V_{ij} و W_{ij}) را برحسب متغیرهای شبکه بنویسید

۱۰- (اختیاری) سوال امتحانی پایان ترم ۱۳۹۸

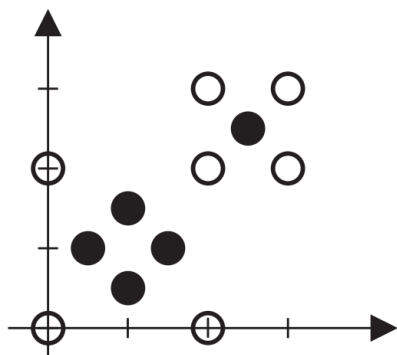
در شبکه عصبی ذیل، که در آن $f_i(\text{net}) = \frac{1}{1+e^{-a_i(\text{net})}}$ و $g(\text{net}) = \text{net}$ است، قصد داریم تا ضرب \mathbf{a} را نیز در هر گام به کمک

الگوریتم پس انتشار خطا آموزش دهیم. اگر در این لحظه $\mathbf{a}_t = [1.1]$ باشد و بردار ویژگی $X_{t+1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ با مقدار مطلوب $y_{t+1} = 0$ از راه

برسد، مطلوبست محاسبه \mathbf{a}_{t+1} ؟ (نرخ آموزش را ۰.۵ در نظر بگیرید)



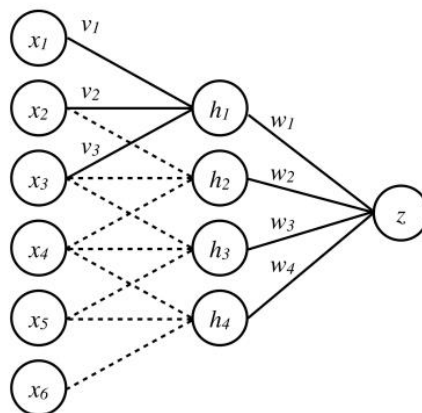
۱۱- (مساله امتحانی پایان ترم ۱۴۰۰) شکل زیر یک مسئله دسته بندی را نشان می‌دهد که در آن نمونه های کلاس ۱ با دایره های مشکی و نمونه های کلاس ۲ با دایره های سفید نمایش داده شده اند. همانطور که مشخص است این دو دسته به صورت خطی قابل جداسازی نیستند. شبکه عصبی RBF را طوری طراحی کنید که این دسته بندی را انجام دهید، سپس نواحی تصمیم گیری 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 .

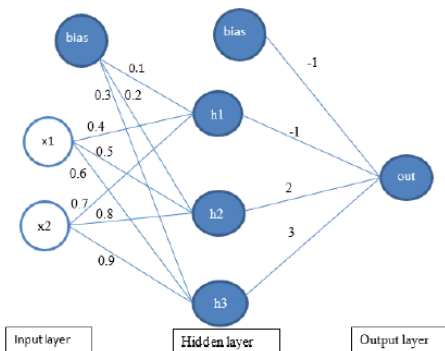
- What is the total number of parameters in this neural network? Recall that convolutional layers share weights. There are no bias terms.
- Compute $\partial R / \partial w_i$.
- Vectorize the previous expression—that is, write $\partial R / \partial w$.
- Compute $\partial R / \partial v_j$.

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. همچنین تابع فعالیت نرون‌های لایه مخفی از نوع سیگموئید 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.



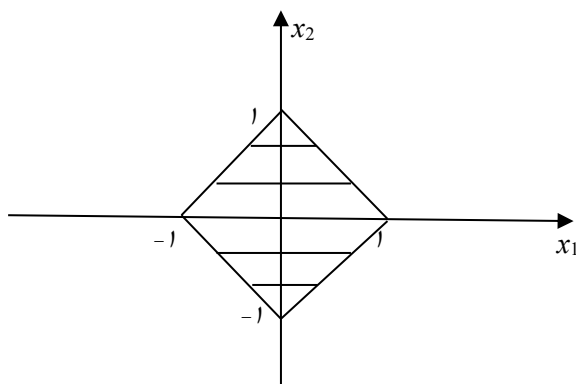
ب- چنانچه نرون لایه خروجی تابع فعالیت همانی باشد خروجی آن به ازای نمونه ورودی زیر چه خواهد بود؟

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

ج- با فرض اینکه پیش‌بینی صحیح مقدار خروجی به ازای $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ بصورت $t=5$ باشد در اینصورت وزن ارتباطی نرون h_3 در لایه مخفی با نرون لایه خروجی پس از به روز رسانی بوسیله پس انتشار خطا چقدر خواهد بود؟

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

۱۴- (اختیاری) (پایان ترم ۱۴۰۲) یک شبکه عصبی از نوع چندلایه پیشرو طراحی نمایید (تعداد لایه‌ها، تعداد نرونهای هر لایه، تابع فعالیت هر نرون، ضرایب ارتباطی نرونها در لایه‌های مختلف را مشخص کنید) طوری که فضای ویژگی دوبعدی را بصورت شکل زیر در دو کلاس هاشور خورده (داخل لوزی) و بدون هاشور (خارج لوزی) دسته‌بندی 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. شبکه‌ی عصبی‌ای رسم کنید که با حداقل تعداد لایه ممکن، این تابع را پیاده‌سازی نماید (ورودی شبکه نیز به صورت باینری می‌باشد).

ب) تابع OR با تعداد n ورودی را در نظر بگیرید. شبکه‌ی عصبی‌ای رسم کنید که با حداقل تعداد لایه ممکن، این تابع را پیاده‌سازی نماید (ورودی شبکه نیز به صورت باینری می‌باشد).

ج) نشان دهید که تابع XOR با دو ورودی را نمی‌توان با شبکه‌ای با کمتر از دو لایه پیاده‌سازی نمود.

د) به نظر شما چرا در شبکه‌های عصبی از تابع فعالسازی به صورت تابع داده شده در این سوال استفاده نمی‌شود؟ مزیت استفاده از تابع فعالساز sigmoid نسبت به این تابع چیست؟

موفق باشید