

بناام خدا


الگوشناسی آماری (CE-725)

دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف

تمرینات سری سوم (توابع جداساز خطی، مدل سازی و شبکه های عصبی مصنوعی)

بهار ۱۳۹۱

به نکات زیر توجه فرمائید:

۱. زمان تحویل تمرینات در سایت درس مشخص شده است. دقت نمائید که زمانبندی های تعیین شده قابل تغییر نیستند.
۲. تمرینات را با عنوان SPR-HWx-8xxxxxxx (مثلا SPR-HW1-88300785) و در یک فایل فشرده با همین نام به آدرس Muhammadi@dml.ir ایمیل زده و در اولین جلسه بعد از زمان تحویل، بصورت پرینت شده تحویل استاد درس دهید.
۳. گزارش شما باید مختصر و مفید باشد. برای تمرینات پیاده سازی که با لوگوی  مشخص شده اند باید کد متلب نوشته شده ضمیمه گزارش شده و تمامی خروجی های برنامه ها در گزارش شما ذکر شوند.

**سوال ۱)** در بسیاری از مسایل کلاسه بندی یک نمونه جدید یا به یکی از c کلاس نسبت داده می شود یا بدلیل عدم تشخیص رد (reject) می شود. اگر هزینه (cost) رد کردن یک نمونه زیاد نباشد، ممکن است رد کردن، یک عمل دلخواه کلاسه بند شود. با توجه به این توضیحات تابع loss زیر را در نظر بگیرید:

$$\lambda(\alpha_i | w_j) = \begin{cases} 0 & i = j \\ \lambda_r & i = c + 1 \\ \lambda_s & \text{otherwise} \end{cases}$$

الف) نشان دهید که کمترین ریسک زمانی رخ می دهد که داشته باشیم:

«یک نمونه جدید را به کلاس  $w_i$  نسبت بده اگر به ازای تمامی  $j$  های دیگر  $P(w_i | x) \geq P(w_j | x)$  بوده

و همچنین  $P(w_i | x) \geq 1 - \lambda_r / \lambda_s$  باشد، در غیر اینصورت بدلیل عدم تشخیص نمونه را رد کن.»

ب) اگر  $\lambda_r$  صفر باشد چه اتفاقی می افتد؟

پ) اگر  $\lambda_r > \lambda_s$  باشد چه اتفاقی می افتد؟

ت) با افزایش  $\lambda_r / \lambda_s$  از صفر به یک چه اتفاقی می افتد؟

ث) با توجه به نتایج قسمت های قبلی نشان دهید که تابع های جداساز زیر برای این مساله بهینه هستند:

$$g_i(x) = \begin{cases} p(x | w_i) P(w_i) & i = 1, \dots, c \\ \frac{\lambda_s - \lambda_r}{\lambda_s} \sum_{j=1}^c p(x | w_j) P(w_j) & i = c + 1 \end{cases}$$

**سوال ۲)** نشان دهید اگر  $f_i(x)$ ,  $i=1, 2, \dots, M$  توابع جداساز یک مساله  $M$  کلاسه باشند، می‌توانیم  $M-1$  تابع جدید با توجه به آنها بسازیم، که همان عملکرد قبلی را داشته باشند (نکته: از تفاضل توابع جداساز می‌توانید کمک بگیرید).

**سوال ۳)** شش نمونه زیر را در نظر بگیرید:

$$w_1: (1, 2), (2, -4), (-3, -1)$$

$$w_2: (2, 4), (-1, 3), (5, 0)$$

آیا این نمونه‌ها بصورت خطی جدا پذیر هستند؟ در صورت مثبت بودن جواب یک جداساز خطی برای این مساله با روش Minimum Squared Error ارائه دهید.

**سوال ۴)** الف) آیا نمونه‌های جدا پذیر خطی کامل (Totally Linearly separable)، جداپذیر خطی (Linearly Separable) نیز هستند؟ عکس این حالت چطور؟

ب) آیا نمونه‌های جداپذیر خطی دو به دو (Pairwise Linearly Separable)، جداپذیر خطی نیز هستند؟ عکس این حالت چطور؟

**سوال ۵)** مجموعه داده‌ای شامل  $N$  عدد را در نظر بگیرید که هر کدام از این اعداد از یک توزیع گاوسی مجزا تولید شده‌اند. تمامی این توزیع‌های گاوسی دارای میانگین یکسان  $\mu$  و واریانس متفاوت  $\sigma_i$  هستند. واریانس‌ها معلوم‌اند ولی میانگین ثابت، مجهول است.

الف) تخمین maximum-likelihood را برای  $\mu$  بیابید.

ب) چه تعبیری می‌توانید از شکل تابع تخمین بخش قبل داشته باشید؟

**سوال ۶)** نمونه داده‌های یک بعدی  $X = \{-7, -5, -4, -3, -2, 0, 2, 3, 4, 5, 7\}$  داده شده‌اند.

الف) تخمین  $P_j(x)$  را با استفاده از روش پنجره پارزن با پنجره مستطیلی شکل بدست آورید. از  $h_j = 1/\sqrt{j}$  استفاده کنید.

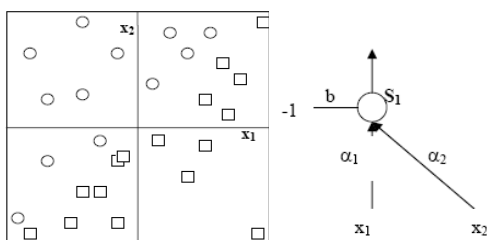
ب) نتایج را برای  $j = 1, 4, 11$  نمایش دهید.

پ) فرض کنید  $h_j = h/\sqrt{j}$  که در آن  $h$  عدد ثابتی است که در ابتدا مشخص می‌شود. در مورد شکل  $P_j(x)$  به ازای مقادیر مختلف  $h$  بحث کنید.

**سوال ۷)** نمونه‌های یک بعدی زیر داده شده‌اند:

$$X = \{x_1, x_2, \dots, x_{16}\} = \{0, 1, 3, 4, 5, 5.5, 6.0, 6.5, 7.0, 7.2, 7.5, 8.0, 8.8, 9.2, 9.3, 11, 13\}$$

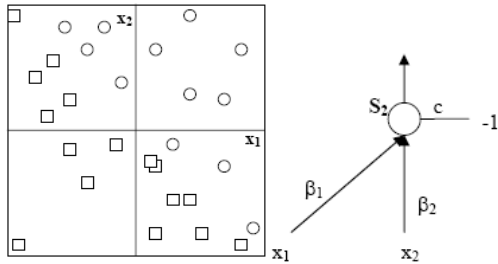
برای  $j=16$  و  $k_j = \sqrt{j}$  مقدار تخمین K-NN برای  $P_j(x)$  را در  $x=2, x=4, x=6, x=8$  و  $x=10$  بدست آورید.



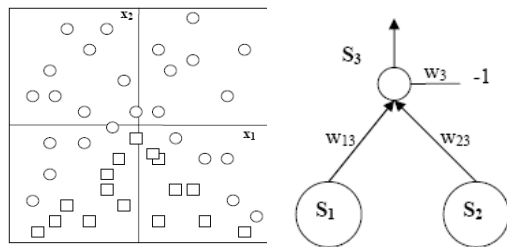
**سوال ۸)** نرون زیر را با دو ورودی، بایاس، یک خروجی و تابع آستانه پله در نظر بگیرید. خروجی این نرون می‌تواند به فرم زیر نوشته شود که در آن  $S_1$  تابع پله می‌باشد:

$$y_1 = S_1(\alpha_1 x_1 + \alpha_2 x_2 - b), S_1(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

در داده‌های داده شده نیز مربع‌ها، کلاس  $y=0$  و دایره‌ها کلاس  $y=1$  را نشان می‌دهند.

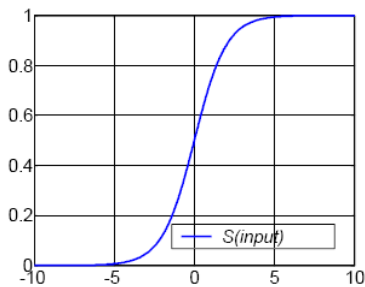


الف) مرز تصمیم‌گیری را رسم کرده و معادله آن را بنویسید. همچنین با متناظر قرار دادن معادله نوشته شده با خروجی نرون فوق، مقادیر ممکن وزن‌های نرون ( $\alpha_1$ ,  $\alpha_2$  و  $b$ ) را تعیین نمایید. ب) مراحل فوق را برای نرون و داده‌های داده شده روبرو تکرار کنید.



پ) برای پیدا کردن مرز بین داده‌های زیر (که خیلی شبیه داده‌های دو مرحله قبل هستند)، می‌خواهیم نرون دیگری را طراحی کنیم که ورودی‌های آن، خروجی‌های دو نرون قبلی باشند. شکل این نرون در زیر نشان داده شده است. خروجی این نرون را بر حسب  $S_1$ ,  $S_2$  و  $S_3$  بیان کنید.

ت) سه داده  $(0, -2)$ ,  $(3, 0)$  و  $(-2, 1)$  را در نظر بگیرید. با استفاده از رابطه‌ای که در مرحله قبل بدست آورده اید و با جایگذاری این سه داده، سه رابطه ساده‌تر بدست آورده و وزن‌های  $w_{13}$ ,  $w_{23}$  و  $w_3$  را بدست آورید.



ث) فرض کنید به جای تابع پله از تابع سیگموئید با شکل روبرو و رابطه زیر استفاده کنیم:

$$S(x) = \frac{1}{1 + e^{-x}}$$

از آنجایی که تابع سیگموئید پیوسته است، از رابطه زیر برای کلاسه‌بندی استفاده می‌کنیم:

$$Class = \begin{cases} 0 & \text{if } S(x) < 1/2 \\ 1 & \text{if } S(x) \geq 1/2 \end{cases}$$

خروجی  $S_3$  را بر حسب  $x_1$  و  $x_2$  با استفاده از این تابع، بازنویسی کنید.

ج) مرزهای تصمیم‌گیری حاصله در دو حالت استفاده از پله و استفاده از سیگموئید را با هم مقایسه کنید.

راهنمایی: مرزهای تصمیم‌گیری را در هر دو حالت، برای دو سری وزن زیر، بصورت تقریبی رسم نمایید:

- سری اول:  $w_{13}=2$ ,  $w_{23}=2$ ,  $w_3=2$  و آلفا و بتاهای با اندازه واحد ولی علامت مناسب (۱- یا ۱+).
- سری دوم:  $w_{13}=2$ ,  $w_{23}=2$ ,  $w_3=1$  و آلفا و بتاهای با اندازه واحد ولی علامت مناسب (۱- یا ۱+).

**سوال ۹)** یک شبکه عصبی feed forward دو لایه برای جداسازی نمونه‌های زیر طراحی کنید (وزن‌ها را دستی تعیین کنید):

$$S_1: (-3, -1), (-1, -1), (-1, 1), (1, 0), (2, -1), (3, -1), (4, -1)$$

$$S_2: (-3, -4), (-1, -4), (1, -3), (1, -3), (2, 4), (3, -4), (4, 2), (5, 1), (6, -3)$$

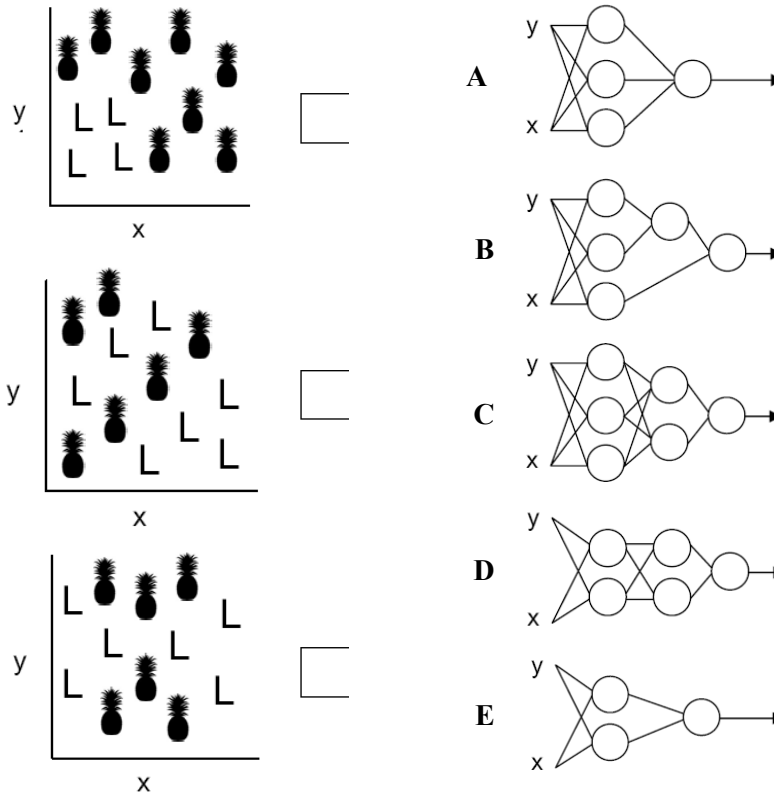
**سوال ۱۰** آیا یک نرون با دو ورودی و یک خروجی قادر به پیاده‌سازی تابع NOR می‌باشد؟ چرا؟

**سوال ۱۱** نمونه داده‌های زیر را در نظر بگیرید:

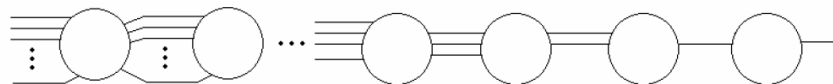
$$S_1: (-1, 0), (2, 0), (2, 3), (6, 4) \quad S_2: (-1, -2), (2, -3), (5, 0), (8, 2)$$

یک شبکه عصبی با کمترین تعداد گره‌های ممکن ارائه دهید که قادر به کلاسه‌بندی این داده‌ها باشد.

**سوال ۱۲** برای هر کدام از مجموعه داده‌های داده شده سمت چپ، مشخص کنید که کدام یک از شبکه‌های سمت راست، مناسب‌ترین گزینه (از لحاظ تعداد نورون‌های کمتر) برای کلاسه‌بندی آن می‌باشد؟



**سوال ۱۳** شبکه زیر را با  $n$  نرون و تابع نورون‌های سیگموئیدی در نظر بگیرید. تعداد ورودی‌های سمت راست‌ترین نود، یک ورودی بوده و به ترتیب، هر نود که به سمت چپ می‌رویم، یک ورودی به ورودی‌ها افزوده می‌شود و تمامی ورودی‌های سمت چپ‌ترین نود برابر ۱ می‌باشند. برای هر نرون وزن تمامی ورودی‌ها برابر می‌باشند و هیچ کدام از نرون‌ها نیز دارای ورودی بایاس نیستند. خروجی سمت راست‌ترین نود را به ازای  $n$ ‌های بزرگ بدست آورید.



**سوال ۱۴** هدف از این تمرین پیاده‌سازی یک عبارت منطقی با یک شبکه عصبی است. شبکه عصبی متناظر با یک عبارت منطقی در صورتی که آن عبارت درست باشد، مقدار ۱ و در غیر این صورت مقدار صفر را برگشت می‌دهد. شبکه عصبی متناظر با هر کدام از عبارت‌های منطقی زیر را طراحی کنید (تمام نورون‌ها را با تابع انتقال پله در نظر بگیرید).

الف) با یک لایه پنهان و دو نورون ورودی عبارت  $A \oplus B$  (A xor B) را طراحی کنید.

ب) با دو لایه پنهان عبارت  $(A \vee \neg B) \oplus (\neg C \vee D)$  را طراحی کنید.

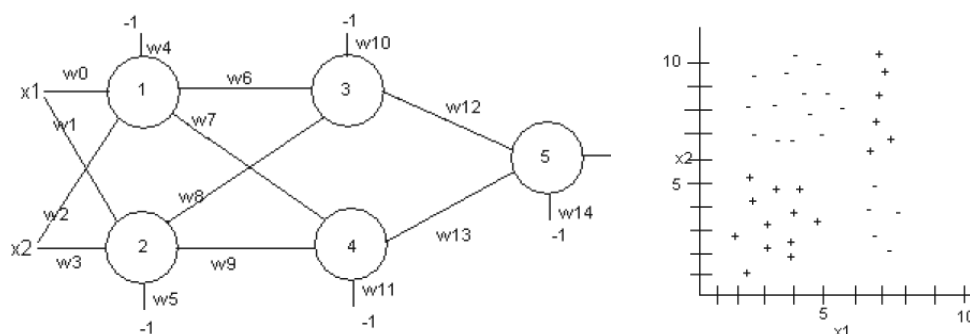
ج) عبارت قسمت (ب) را با یک لایه پنهان طراحی کنید.

د) آیا هر عبارت منطقی را می توان با n ورودی، یک خروجی و یک لایه پنهان طراحی کرد؟ توضیح دهید.

**سوال ۱۵)** شبکه و داده های زیر را در نظر بگیرید. خروجی هر کدام از پرسپترون ها، بر خلاف پرسپترون استاندارد ۱ یا -۱ می باشد. جدول وزن های زیر را بدون استفاده از الگوریتم های یادگیری و بصورت دستی، بصورت مناسب پر کنید، به قسمی که عملکرد جداسازی شبکه داده شده قابل قبول باشد.

راهنمایی: ابتدا فضا را با دو خط به چهار ناحیه تقسیم کرده و سپس ناحیه های مربوط به هر کلاس را مشخص کنید. دو نرون سطح اول نمایانگر دو خط رسم شده، دو نرون مخفی نمایانگر نواحی تشکیل شده و نرون خروجی ترکیب کننده نواحی می باشد. توجه کنید که ساختار چهار ناحیه تشکیل شده مشابه ساختار XOR می باشد.

Weight	W0	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14
Value															



**سوال ۱۶)** به نظر شما اگر فردی در شبکه عصبی از اعتبارسنجی چندگانه K-Fold و در k-نزدیک ترین همسایگی از اعتبارسنجی چندگانه Leave one out استفاده کرده باشد، کار درستی انجام داده است یا خیر/ برای پاسخ خود دلیل بیاورید.

**سوال ۱۷)** دو مثال توزیع گاوسی و ترکیب توزیع یکنواخت و مثلثی ارائه شده در اسلایدهای هشتم درس (اسلایدهای مدل سازی آماری) را برای هر دو حالت Parzen Window و K-NN شبیه سازی کنید.

**سوال ۱۸)** در مجموعه داده Iris ۷۰ درصد داده ها را به عنوان مجموعه داده آموزشی و بقیه را به عنوان مجموعه داده تست در نظر بگیرید.

الف) شبکه MLP را با الگوریتم Back propagation پیاده سازی کرده و آن را آموزش دهید تا مساله دسته بندی Iris را حل کند. تعداد نورون های لایه پنهان را از ۱ تا ۱۰ تغییر داده و خطای MSE را برای داده های آموزش و آزمایش رسم کنید. تغییرات خطای آموزش را برای بهترین تعداد نورون های لایه پنهان، در طول زمان رسم کنید.

ج) برای تشخیص سرطان در مجموعه داده تشخیص سرطان از الگوریتم k-NN استفاده کرده و k بهینه را بیابید (روش خود را توضیح دهید). برای داده های Iris نیز این کار را انجام دهید. از چه روش اعتبارسنجی چندگانه ای استفاده کرده اید و چرا؟