

۱- در الگوریتم پرسپترون دو لایه:

الف) اگر تعداد نرون‌های لایه پنهان ۲ باشد ( $p = 2$ )، حداکثر تعداد رئوس استفاده شده مربع چقدر باشد تا با همین دو لایه بتوان برای هر مسئله‌ای داده‌های آموزشی را درست طبقه‌بندی کرد؟

ب) اگر تعداد نرون‌های لایه پنهان ۳ باشد ( $p = 3$ )، حداکثر تعداد رئوس استفاده شده مکعب چقدر باشد تا با همین دو لایه بتوان برای هر مسئله‌ای داده‌های آموزشی را درست طبقه‌بندی کرد؟

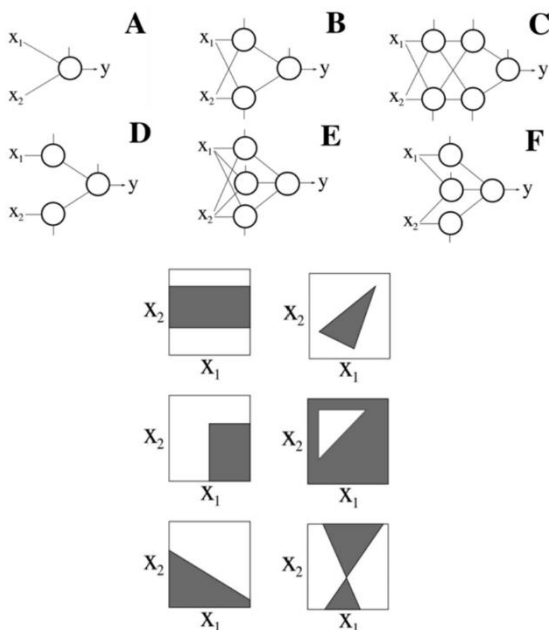
پ) آیا ممکن است فقط یک نرون در لایه پنهان داشته باشیم و داده‌های آموزشی داده‌شده درست طبقه‌بندی شوند؟

ت) در مسئله  $XOR$  با چهار داده، فرض کنید  $y_2 = u(x_1 + x_2 - \frac{1}{4})$ ،  $y_1 = u(x_1 + x_2 - \frac{7}{4})$ . یک دسته ضرایب برای شبکه پیشنهاد کنید.  $u$  تابع پله است.

۲- مسئله ۳ از فصل چهارم کتاب را حل کنید.

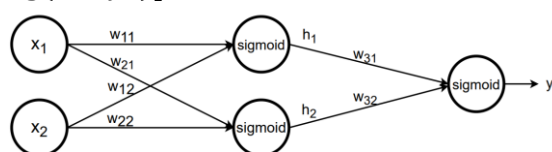
۳- یک شبکه عصبی با دو گره در لایه ورودی و حداقل تعداد لایه و تابع فعال‌سازی علامت را به گونه‌ای تعیین کنید که به ربع اول و سوم برچسب ۱ و به ربع دوم و چهارم برچسب -۱ بدهد.

۴- شبکه‌های عصبی پرسپترون چند لایه با تابع فعال‌سازی پله داده شده در قسمت بالا از شکل روبرو را در نظر بگیرید. ابتدا مشخص کنید هر کدام از پاسخ‌های داده شده در قسمت پایین از شکل روبرو توسط کدام شبکه تولید شده است و سپس علت را توضیح دهید. فرض کنید که هر کدام از شبکه‌ها فقط یک پاسخ را ایجاد کرده است.



۵- شبکه عصبی زیر را در نظر بگیرید. با استفاده از روش پس انتشارخطا، مشتق عبارت زیر را نسبت به وزن  $w_{12}$  بدست آورید (در جواب نهایی می‌بایست از متغیرهایی که در شکل نامگذاری شده است استفاده نمایید). بایاس نرون‌ها در شکل نشان داده نشده است. متغیر  $y$  برچسب واقعی است.

$$H(y, y^*) = -[y \log y^* + (1 - y) \log(1 - y^*)]$$



۶- در طبقه‌بندی کننده خطی تعمیم یافته در مسئله XOR با چهار داده:

الف) در کدامیک از حالات زیر داده‌های آموزش در ست طبقه‌بندی می شوند؟ در این صورت یک معادله مرز در فضای اولیه و تبدیل یافته را رسم کنید.

$$\underline{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 \cdot x_2 \end{pmatrix}, \quad \underline{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1^2 \\ x_2^2 \end{pmatrix}, \quad \underline{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 \\ x_2 \end{pmatrix}$$

ب) برای  $y_1 = f_1(\underline{x}) = |x_1 \cdot x_2 + x_1 - x_2|$ ,  $y_2 = f_2(\underline{x}) = x_1 \cdot x_2$  یک معادله مرز در فضای اولیه و تبدیل یافته را رسم کنید.

۷- در مثال حل شده در صفحه ۱۹۲ کتاب برای مسئله XOR با طبقه‌بندی کننده RBF، مسئله را برای دو مرکز  $\underline{c}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ,  $\underline{c}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  دوباره حل کنید.

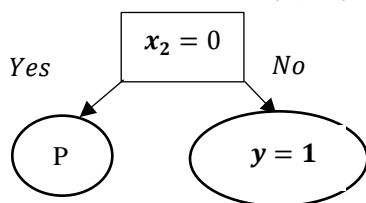
۸- در الگوریتم SVM غیرخطی:

الف) با کرنل  $K(\underline{x}, \underline{z}) = (\underline{x}^T \cdot \underline{z})^3$  تابع  $\phi$  را پیدا کنید فضای اولیه ویژگی‌ها دوبعدی است.

ب) با کرنل RBF  $K(\underline{x}, \underline{z}) = \exp\left(-\frac{1}{2}\|\underline{x} - \underline{z}\|^2\right)$  تابع  $\phi$  را پیدا کنید.

۹- یک مساله یادگیری به روش درخت تصمیم گیری برای  $X = \{0,1\}^4$  در  $Y = \{0,1\}$  در نظر بگیرید. داده آموزشی S شامل ۵ سطر مطابق جدول زیر است. فرض کنید بخشی از درخت تصمیم گیری مطابق شکل ساخته ایم و اکنون می‌خواهیم وضعیت درخت را در نقطه P تعیین کنیم.

مشخص کنید کدام یک از گزینه‌های زیر برای نقطه P وضعیت مطلوب درخت است:



۱) یک برگ با برچسب ۰

۲) یک برگ با برچسب ۱

۳) یک گره تصمیم گیری بر اساس  $x_1$

۴) یک گره تصمیم گیری بر اساس  $x_2$

۵) یک گره تصمیم گیری بر اساس  $x_3$

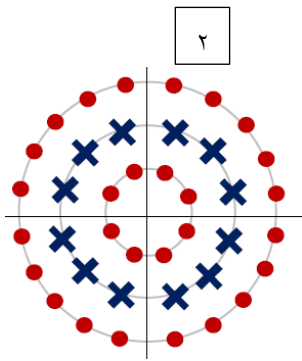
۶) یک گره تصمیم گیری بر اساس  $x_4$

$x_1$	$x_2$	$x_3$	$x_4$	$y$
0	1	0	1	0
1	0	1	0	1
1	0	0	0	0
0	0	1	0	1
0	1	1	0	0

۱۰- با توجه به داده‌های موجود در جدول زیر، کدام ویژگی در درخت تصمیم با روش گفته شده در کلاس ابتدا انتخاب خواهد شد؟

$x_1$	$x_2$	$x_3$	$x_4$	class
0	0	0	0	-
0	0	1	1	+
0	1	0	1	-
0	1	1	0	+
1	0	0	0	+
1	0	1	1	+
1	1	1	0	+
1	1	1	1	+

۱۱- الف) یک طبقه‌بندی کننده خطی تعمیم یافته (متشکل از حداقل تعداد تابع غیرخطی) پیشنهاد کنید که داده‌های شکل زیر خطی جدایی پذیر شوند.  
 ب) کرنلی برای Kernel SVM پیشنهاد کنید که داده‌های شکل زیر به خوبی طبقه‌بندی شوند.



۱۲- در این مساله می‌خواهیم تعیین کنیم که آیا نقطه  $x \in \mathbb{R}^2$  در ناحیه میان یک وتر و دایره‌ای به مرکز مبدا مختصات و شعاع ۲ (مطابق ناحیه هاشور خورده شکل زیر) قرار دارد یا نه. می‌دانیم که این مساله بر اساس یک منحنی درجه ۲ در  $\mathbb{R}^2$  به خوبی قابل جداسازی است. برای این مساله، نخست  $x$  را با نگاشت  $\psi$  به بردار  $\psi(x)$  مناسب تبدیل می‌کنیم و سپس  $\psi(x)$  را به عنوان ورودی به یک شبکه عصبی وارد می‌کنیم.

الف) با انتخاب  $\psi(x)$  مناسب (حداکثر از ترکیبات درجه ۲ در  $\mathbb{R}^2$ ) مساله را به یک طبقه‌بندی خطی تبدیل کنید.

ب) شبکه عصبی با کمترین تعداد لایه‌های ممکن و تابع فعال سازی sign به نحوی طراحی کنید که ورودی آن  $\psi(x)$  و خروجی آن وقتی برابر ۱ باشد که  $x$  در ناحیه هاشور خورده قرار داشته باشد، در غیر این صورت خروجی ۱- باشد.

