



## آزمون پایان ترم

- زمان در نظر گرفته شده برای آزمون ۱۸۰ دقیقه است.
- لطفاً پاسخ‌های خود را خوانا و خوش خط بنویسید.

### سوالات (۱۰۰ نمره)

۱. (۱۵ نمره) به سوالات زیر به طور مختصر پاسخ دهید:

- (آ) درستی یا نادرستی عبارت رو به رو را با ذکر دلیل مشخص کنید: در یک شبکه بیزی که در آن  $X$  به شرط  $Z$  از  $Y$  مستقل است؛ ممکن است فرض استقلال این دو متغیر با شرطی کردن شواهد اضافی برای متغیرهای دیگر در شبکه، برقرار نباشد.
- (ب) دو مورد از روش‌های جلوگیری از بیش‌برازش (overfitting) در درخت تصمیم را ذکر کرده و مختصراً توضیح دهید.
- (ج) آیا همواره با استفاده از مدل درخت تصمیم بدون محدودیت عمق می‌توان به دقت ۱۰۰٪ رسید؟ توضیح دهید.
- (د) برای طبقه‌بندی تصاویر به دو کلاس سگ و گربه، از یک شبکه عصبی عمیق استفاده شده است. در انتهای این شبکه، از یک تابع فعال‌ساز  $ReLU$  و سپس از تابع  $sigmoid$  استفاده شده است. در صورتی که خروجی نهایی شبکه  $\hat{y} \geq 0.5$  باشد آن را به عنوان گربه و در غیر این صورت به عنوان سگ در نظر می‌گیریم. آیا این شبکه ایرادی در برچسب‌گذاری تصاویر دارد؟ مختصراً توضیح دهید.
- (ه) فرض کنید از رابطه‌ی

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha \left[ R(s, a, s') + \gamma \max_{a'} f(Q(s', a'), N(s', a')) \right]$$

برای بروزرسانی  $Q$ -value‌ها استفاده کنیم که  $N(s', a')$  تعداد دفعاتی است که از حالت  $s'$  کنش  $a'$  را انجام داده‌ایم و  $f(x, y) = x + \frac{\beta}{y}$  که  $\beta$  عدد ثابت مثبتی است. توضیح دهید که این تغییر چه تاثیری بر جست‌وجوی فضای حالت خواهد گذاشت؟

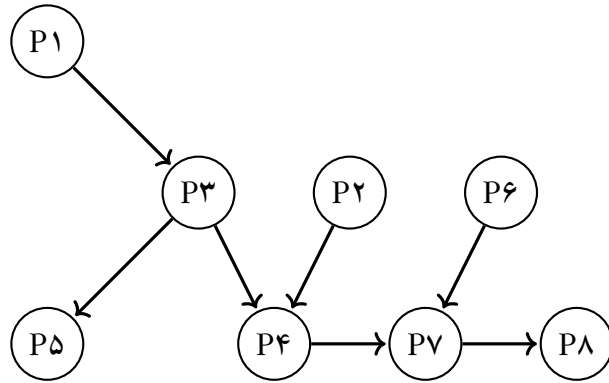
۲. (۸ نمره) مدل مارکوف زیر را در نظر بگیرید که سه استیت ۱ و ۲ و ۳ دارد. این مدل دارای Transition Matrix به صورت زیر است:

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

(آ) برای این مدل یک نمودار حالت (state diagram) رسم کنید.

(ب) اگر بدانیم که  $P(X_1 = 1) = P(X_1 = 2) = \frac{1}{4}$  مقدار  $P(X_1 = 3, X_2 = 2, X_3 = 1)$  را محاسبه کنید.

۳. (۱۲ نمره) فرض کنید گراف زیر را به عنوان یک شبکه‌ی بیزین داریم:



درستی یا نادرستی هر کدام از عبارتهای زیر را با ذکر دلیل مشخص کنید.

- (آ)  $P_5 \perp\!\!\!\perp P_6 | P_7$
- (ب)  $P_1 \perp\!\!\!\perp P_7 | P_5$
- (ج)  $P_1 \perp\!\!\!\perp P_7 | P_8$
- (د)  $P_5 \perp\!\!\!\perp P_8 | P_7$

۴. (۱۴ نمره) یک مدل خطی به فرم زیر را در نظر بگیرید:

$$y(x_n, \mathbf{w}) = w_0 + \sum_{i=1}^D w_i x_{ni}$$

خطای آن را نیز به صورت زیر در نظر می‌گیریم:

$$E_D(\mathbf{w}) = \frac{1}{N} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2$$

حال فرض کنید که یک نویز گوسی  $\epsilon_i \sim \mathcal{N}(0, \sigma^2 I)$  به هر ورودی  $x_i$  اضافه شده است.  $\epsilon_i$ ها به صورت i.i.d تولید شده‌اند. اگر  $\tilde{E}_D(\mathbf{w})$  خطای مدل وقتی از  $x_i + \epsilon$  استفاده می‌کنیم باشد، رابطه‌ی زیر را اثبات کنید.

$$\mathbb{E}[\tilde{E}_D(\mathbf{w})] = E_D(\mathbf{w}) + \frac{N}{N} \sum_{i=1}^D w_i^2 \sigma^2$$

۵. (۱۰ نمره) با توجه به جدول زیر قصد داریم یک درخت تصمیم بسازیم که مقدار  $Y$  را براساس  $X_1, X_2$  و  $X_3$  تعیین کند.

$X_1$	$X_2$	$X_3$	$Y$
No	No	No	No
Yes	No	Yes	Yes
No	Yes	Yes	Yes
Yes	Yes	Yes	No

(آ) درخت تصمیم را رسم کنید. در هر مرحله ریشه را براساس Information Gain انتخاب کنید. (لزومی به محاسبه‌ی عددی IG نیست اما در صورت نیاز از آن استفاده کنید.)

(ب) آیا درخت به‌دست آمده (با انتخاب ریشه‌ها از طریق IG) بهینه است؟ اگر بهینه است علت آن را بیان کنید؛ در غیر این صورت درخت بهینه را رسم کنید. (منظور از بهینه، درختی با کوتاه‌ترین ارتفاع ممکن است که نمونه‌های سازگار را جداسازی کند.)

۶. (۱۵ نمره) می‌دانیم در شبکه‌های عصبی از توابع مختلفی برای فعال‌سازی استفاده می‌شود. در این سوال می‌خواهیم به کمک تابع پله، شبکه‌های عصبی ساده‌ای برای پیش‌بینی مقادیر توابع منطقی بدست آوریم. توجه داشته باشید در عبارات زیر،  $\wedge$  نماد عملگر منطقی AND،  $\vee$  نماد عملگر منطقی OR و  $\neg$  نماد عملگر منطقی NOT است.

در این سوال ورودی‌ها ( $x_i$ ) صفر یا یک هستند و تابع فعال‌ساز هم به صورت زیر است:

$$f(x) = \begin{cases} 0 & - (w_0 + \sum_i w_i x_i) < 0 \\ 1 & - (w_0 + \sum_i w_i x_i) \geq 0 \end{cases}$$

در این فرمول  $w_0$  نمایانگر بایاس بوده و بقیه  $w_i$  ها وزن‌های نسبت داده شده به یال‌های درونی شبکه عصبی هستند.

(آ) یک شبکه عصبی با یک پرسپترون تشکیل دهید که بتواند تابع منطقی  $y = x_1 \vee x_2$  را به درستی محاسبه کند.

(ب) یک شبکه عصبی با یک پرسپترون تشکیل دهید که بتواند تابع منطقی  $y = x_1 \wedge x_2$  را به درستی محاسبه کند.

(ج) یک شبکه عصبی متشکل از یک لایه‌ی نهان تشکیل بدهید که تابع منطقی زیر را به درستی محاسبه کند.

$$y = (x_1 \wedge x_2 \wedge x_3) \vee (\neg x_1 \wedge x_3)$$

۷. (۱۴ نمره) می‌خواهیم مسئله‌ی پک‌من را با استفاده از مدل MDP حل کنیم. در مسئله‌ی پک‌من، یک موجود به نام پک‌من داریم که در یک جدول قرار دارد، می‌تواند به چهار جهت حرکت کند و هدف آن خوردن غذاهای موجود در صفحه‌ی بازی است. در مسئله‌ی مدل شده، حالت‌ها همان خانه‌های جدول و کنش‌ها حرکت به چهار جهت مختلف هستند. هر حرکتی که به سمت دیوار انجام شود، چه دیوارهای داخل محیط بازی و چه دیوارهای مرزی محیط بازی، بی‌تاثیر بوده و پک‌من در همان خانه‌ی قبلی خود باقی می‌ماند. هم‌چنین فرض کنید در کل محیط بازی فقط یک غذا وجود دارد (که با  $\circ$  در نقشه‌ی بازی نشان داده شده است) و پک‌من به محض رسیدن به آن یک امتیاز دریافت کرده و بازی تمام می‌شود.

• در این قسمت محیط بازی را مطابق شکل زیر در نظر بگیرید. (حروف موجود در خانه‌ها نام هر خانه هستند.)

A	B	C
D	E	F, $\circ$

ضریب تخفیف برابر  $0.5$  بوده و پاداش زنده‌بودن صفر است. با در نظر گرفتن این موارد به سوال‌های زیر پاسخ دهید:

(آ) سیاست بهینه را برای هر حالت مشخص کنید.

(ب) ارزش بهینه‌ی خانه‌ی A یا همان  $V^*(A)$  چه قدر است؟

توجه کنید امتیاز موجود در یک خانه هنگام ورود به خانه‌ی دارای امتیاز محاسبه می‌شود. یعنی به طور مثال در شکل بالا،  $R(E, \text{راست}, F) = 1$  می‌باشد.

- حال محیط دیگری مانند شکل زیر را برای بازی در نظر بگیرید:

A	B, ○		
C	D, *	E	F, *

در این محیط علاوه بر غذای اصلی (○) دو ماده‌ی غذایی دیگر (\*) نیز در نقشه وجود دارند که خوردن آن‌ها ۵ امتیاز مثبت خواهد داشت. این امتیازها نیز هنگام ورود به خانه‌ی دارای غذا کسب می‌شوند. در این حالت هم‌چنان غذای اصلی همان ○ است که خوردن آن باعث تمام شدن بازی می‌شود و پاداش ۱ دارد. در این حالت به سوال‌های زیر پاسخ دهید:

(ج) اگر ضریب تخفیف برابر ۱ باشد و هزینه‌ی زنده‌ماندن برابر ۱-، سیاست بهینه را برای حالت‌های مختلف پیدا کنید.

(د) هم‌چنان فرض کنید ضریب تخفیف برابر ۱ است، هزینه‌ی زنده‌ماندن در چه بازه‌ای می‌تواند باشد تا پک‌من با شروع از خانه‌ی A دقیقاً یک \* را بخورد؟

۸. (۱۲ نمره) فرض کنید بازی‌ای فقط شامل دو حالت A و B بوده و کنش‌های قابل انجام از هر کدام از این حالت‌ها نیز حرکت به سمت بالا یا پایین است. عاملی در این محیط با استفاده از سیاست  $\pi$  به انجام بازی پرداخته و دنباله‌ی حالت‌ها و پاداش‌های زیر را مشاهده کرده است:

$t$	$s_t$	$a_t$	$s_{t+1}$	$r_t$
۰	A	پایین	B	۲
۱	B	پایین	B	-۴
۲	B	بالا	B	۰
۳	B	بالا	A	۳
۴	A	بالا	A	-۱

با در نظر گرفتن ضریب تخفیف برابر ۰/۵ و  $\alpha = ۰/۵$  به سوال‌های زیر پاسخ دهید:

(آ) می‌خواهیم از روش Q-learning برای بروزرسانی Q-value ها استفاده کنیم. ابتدا رابطه‌ی بروزرسانی را برای Q-value ها بنویسید. سپس با در نظر گرفتن مقدار صفر برای مقدار اولیه‌ی Q-value ها مقدار  $Q(A, \text{پایین})$  و  $Q(B, \text{بالا})$  را بدست آورید.

(ب) در روش model-based ابتدا تابع گذار  $T(s, a, s')$  و تابع پاداش  $R(s, a, s')$  را تخمین می‌زنیم. برای این منظور با استفاده از دنباله‌ی حالت‌های مشاهده شده این توابع را به‌ازای مقادیر مختلف  $s, a$  و  $s'$  محاسبه کنید (در مجموع مقادیر هشت حالت را باید تخمین بزنید). اگر داده‌های لازم برای تخمین یکی از ورودی‌های هر کدام از توابع وجود ندارد به این موضوع اشاره کنید.

(ج) فرض کنید تجربه‌ی جدیدی از بازی کسب کرده‌ایم و از روی آن توابع  $\hat{T}$  و  $\hat{R}$  را تخمین زده‌ایم. نتیجه جدول زیر شده است:

$s$	$a$	$s'$	$\hat{T}(s, a, s')$	$\hat{R}(s, a, s')$
A	بالا	A	۱	۱۰
A	پایین	A	۰/۵	۲
A	پایین	B	۰/۵	۲
B	بالا	A	۱	-۵
B	پایین	B	۱	۸

با استفاده از این جدول سیاست بهینه  $\hat{\pi}^*(s)$  و تابع ارزش بهینه  $\hat{V}^*(s)$  را بدست آورید. (راهنمایی: برای هر  $x$  حقیقی که  $|x| < ۱$  داریم:  $۱ + x + x^۲ + x^۳ + \dots = \frac{1}{1-x}$ )