368

F

368F

نام

نام خانوادگی

محل امضاء

صبح جمعه
۹۱/۱/۲۵

اگر دانشگاه اصلاح شود مملکت اصلاح می‌شود.
امام خمینی (ره)

جمهوری اسلامی ایران
وزارت علوم، تحقیقات و فنّاوری
سازمان سنجش آموزش کشور

آزمون ورودی
دوره‌های دکتری (نیمه متمرکز) داخل
در سال ۱۳۹۱

رشته‌ی
مهندسی کامپیوتر ـ هوش مصنوعی (کد ۲۳۵۶)

نام و نام خانوادگی داوطلب:	شماره داوطلبی:
تعداد سؤال: ۴۵	مدت پاسخگویی: ۱۵۰ دقیقه

عنوان مواد امتحانی، تعداد و شماره سؤالات

ردیف	مواد امتحانی	تعداد سؤال	از شماره	تا شماره
۱	مجموعه دروس تخصصی (ساختمان داده‌ها، شناسایی آماری الگو، یادگیری ماشین)	۴۵	۱	۴۵

فروردین سال ۱۳۹۱

استفاده از ماشین حساب مجاز نمی‌باشد.

حق چاپ و تکثیر سؤالات پس از برگزاری آزمون برای تمامی اشخاص حقیقی و حقوقی تنها با مجوز این سازمان مجاز می‌باشد و با متخلفین برابر مقررات رفتار می‌شود.

۱) در ورودی یک پشته اعداد ۱ تا n به ترتیب قرار دارند (۱ در ابتدای ورودی است.) عمل PUSH اولین عدد ورودی را برداشته و در بالای پشته قرار می‌دهد. عمل POP عدد بالای پشته را برداشته و در انتهای دنباله‌ی خروجی می‌نویسد. با ترکیبی مناسب از n عدد PUSH و n عدد POP، می‌توان جای‌گشتی از اعداد ۱ تا n را در خروجی تولید کرد که به آن «جای‌گشت قابل تولید» می‌گوییم. مثلاً برای $n = 4$، جای‌گشت $\langle 2,3,1,4 \rangle$ (که عناصر آن از چپ به راست POP شده‌اند) قابل تولید است. برای $n = 8$، کدام‌یک از جای‌گشت‌های زیر قابل تولید نیست؟

۱) $\langle 4,3,7,8,6,2,5,1 \rangle$
۲) $\langle 8,7,6,5,4,3,2,1 \rangle$
۳) $\langle 4,3,2,1,8,7,6,5 \rangle$
۴) $\langle 5,4,7,8,6,3,2,1 \rangle$

۲) چه تعداد درخت دودویی با n گره و با برچسب‌های ۱ تا n دارای ترتیب‌های یکسان در دو روش پس‌ترتیب و میان‌ترتیب هستند؟

۱) ۰ ۲) ۱ ۳) $n!$ ۴) عدد n ام کاتالان

۳) با n عنصر چند درخت دودویی متوازن با ارتفاع $h = \lfloor \lg n \rfloor$ می‌توان ساخت؟

۱) ۱ ۲) $\binom{2^h}{n-2^h+1}$ ۳) $\binom{n}{n-2^h}$ ۴) $\binom{n}{n-2^{h-1}}$

۴) ماتریس A_0، A_1 و ... را با این تعریف در نظر بگیرید. A_0 یک ماتریس 1×1 با مقدار ۱ و A_k برای $k > 0$ یک ماتریس $2^k \times 2^k$ است که به‌صورت تعریف می‌شود:

$$A_k = \begin{bmatrix} A_{k-1} & A_{k-1} \\ A_{k-1} & -A_{k-1} \end{bmatrix}$$

فرض کنید V یک بردار ستونی به اندازه‌ی $n = 2^k$ است. هزینه‌ی ضرب $A \times V$ کدام است؟

۱) $\Theta(n)$ ۲) $\Theta(n \lg n)$ ۳) $\Theta(k \lg n)$ ۴) $\Theta(n \lg k)$

۵) یک درخت ۲- کامل درختی است که هر گره‌ی آن صفر یا ۲ فرزند دارد. اگر $n(h)$ و $N(h)$ به‌ترتیب بیشینه و کمینه‌ی تعداد گره‌های یک درخت ۲- کامل به ارتفاع h باشد، برای $h > 0$ مقادیر $n(h)$ و $N(h)$ به‌ترتیب کدامند؟

۱) $h + 1$ و 2^{h-1}
۲) $h + 1$ و 2^{h+1}
۳) $2h + 1$ و $2^h - 1$
۴) $2h + 1$ و $2^{h+1} - 1$

۶) پیمایش سطح‌ترتیب گره‌های یک درخت را به‌ترتیب سطح آن‌ها و در هر سطح آن‌ها را از چپ‌به‌راست ملاقات می‌کند.
ترتیب ملاقات برگ‌های هر درخت در پیمایش سطح‌ترتیب برابر کدام روش زیر است؟

۱) میان‌ترتیب ۲) پس‌ترتیب ۳) پیش‌ترتیب ۴) هیچ‌کدام

۷) با ۲۵ عنصر چند درخت دودریی باارتفاع کمینه می‌توان ساخت؟

۱) $\binom{16}{10}$
۲) $56 + \binom{14}{4} + \binom{16}{6}$
۳) $\binom{16}{6} + \binom{14}{3} + 8 + \binom{8}{2}$
۴) $\binom{16}{10} + \binom{14}{3} + \binom{8}{2} + \binom{8}{2}$

۸) یک درخت دودویی جست‌وجو با n گره را در نظر بگیرید که عناصر آن متمایز و خاصیت هرم بیشینه هم داشته باشد (مقدار هر رأس از فرزندانش کم‌تر نباشد.) اگر h ارتفاع این درخت باشد، داریم:

۱) $h = O(n)$
۲) میانگین h برابر $O(\lg n)$ است.
۳) $h = O(\lg n)$
۴) چنین درختی وجود ندارد.

۹) عبارت $a - b * c + d - e/g/h$ را، با فرض اِعمال اولویت عمل‌گرها، به‌چند طریق می‌توان به‌درستی پرانتزگذاری کرد به‌طوری‌که مقدار عبارت حاصل به‌ازای همه‌ی مقادیر a تا h با مقدار عبارت اصلی برابر شود؟
دقت کنید که دور یک متغیر تنها در عبارت پرانتز گذاشته نمی‌شود و یک عبارت پرانتزگذاری شده باید صفر یا تعداد زوجی پرانتز داشته باشد. مثلاً $((a + b) + c)$، $a + (b + c)$ و $a + b + c$ سه‌عبارت پرانتزگذاری‌شده‌ی درست برای $a + b + c$ است ولی $(a) + b + c$ نیست.

۱) ۳۲
۲) ۶۴
۳) ۹۶
۴) ۱۲۸

۱۰) کدام‌یک از گزینه‌های زیر ترتیب درجه‌ی رشد تابع‌های زیر را نشان می‌دهد؟ می‌نویسیم $g_i < g_j$ اگر $g_j = \Omega(g_i)$

g_1	g_2	g_3	g_4	g_5	g_6
$\lg^*(n^n)$	$n^{2+\lg n}$	$n \lg n$	2^n	n	$2^{2\sqrt{2 \lg n}}$

۱) $g_2 > g_3 > g_6 > g_5 > g_1 > g_4$
۲) $g_4 > g_3 > g_5 > g_6 > g_2 > g_1$
۳) $g_4 > g_2 > g_1 > g_3 > g_5 > g_6$
۴) $g_4 > g_2 > g_3 > g_5 > g_5 > g_6$

۱۱) زمان اجرای برنامه‌ای برای اندازه‌ی ورودی ۱۰۰، برابر ۶ ثانیه و برای اندازه‌ی ورودی ۱۰۰۰، برابر ۱۰ دقیقه است. کدام‌یک از گزینه‌های زیر به‌احتمال زیاد پیچیدگی این برنامه است؟

۱) $O(1)$
۲) $O(N)$
۳) $O(N^2)$
۴) $O(N^3)$

۱۲) بزرگ‌ترین عنصر یک هرم کمینه با عناصر متفاوت را به‌صورت کارا با چند مقایسه بین عناصر آن می‌توان به‌دست آورد؟

۱) $\lfloor \lg n \rfloor - 1$
۲) $n - 1$
۳) $\lceil n/2 \rceil - 1$
۴) $\lfloor (n+1)/2 \rfloor$

۱۳) داده‌ساختاری بر روی مجموعه‌ی A از اعداد صحیح در نظر بگیرید که اعمال درج، حذف و «یافتن نزدیک‌ترین» (FindClosest) را فراهم می‌آورد. منظور از FindClosest(y) یافتن عدد $x \in A$ در مجموعه است به‌طوری‌که برای هر $x_i \in A$ داشته باشیم: $|x - y| \leq |x_i - y|$.
فرض کنید T بیش‌ترین زمان اجرای این اعمال است. یعنی:

$$T = \max\{T_{Insert}, T_{Delete}, T_{FindClosest}\}$$

کدام‌یک از داده‌ساختارهای زیر کم‌ترین مقدار T را خواهد داشت؟

۱) لیست مرتب ۲) درخت دودویی جست‌وجو ۳) هرم ۴) درخت قرمز- سیاه

۱۴) احتمال آن‌که الگوریتم مرتب‌سازی سریع تصادفی n عنصر را در زمان $\Omega(n^2)$ مرتب کند دست‌کم برابر است با:

۱) $1/n$ ۲) $1/(n!)$ ۳) $n/(n!)$ ۴) $n^2/(n!)$

۱۵) در گونه‌ای جدید مرتب‌سازی سریع آرایه‌ی n عضوی A، $2\sqrt{n}+1$ عنصر اول را انتخاب کرده با یک الگوریتم ساده مثل مرتب‌سازی درجی، این عناصر را مرتب و سپس عنصر میانه‌ی آن‌ها را به‌دست می‌آوریم. این عنصر را محور قرار می‌دهیم و بقیه‌ی الگوریتم مرتب‌سازی را بر روی A را اجرا می‌کنیم. کدام‌یک از رابطه‌های زیر زمان اجرای این الگوریتم جدید را در بدترین حالت نشان می‌دهد؟

۱) $T(n) \leq T(n - \sqrt{n}) + O(\sqrt{n})$
۲) $T(n) \leq T(2\sqrt{n}) + T(n - 2\sqrt{n}) + O(n)$
۳) $T(n) \leq T(\sqrt{n}) + T(n - \sqrt{n}) + O(n)$
۴) $T(n) \leq 2T(n - \sqrt{n}) + O(n)$

۱۶) در یک داده‌ساختار «مجموعه‌های مجزا (disjoint sets)» که با درخت و بدون فشرده‌سازی مسیر پیاده‌سازی شده‌است و در عمل ادغام درخت با ارتفاع کم‌تر فرزند ریشه‌ی درخت با ارتفاع بیش‌تر می‌شود، اگر تعداد عناصر n باشد، هزینه‌ی دنباله‌ای از m عمل «ادغام» و f عمل «یافتن» حداکثر چه‌قدر است؟

۱) $O(f + m)$ ۲) $O(f + m \lg n)$
۳) $O(m + f \lg n)$ ۴) $O((m + f) \lg n)$

۱۷) کمینه‌ی تعداد گره‌های داخلی یک درخت قرمز- سیاه که سیاه- ارتفاع (Black-Height) ریشه‌ی آن h باشد برابر است با:

۱) $2^h - 1$ ۲) $2^{h+1} - 1$ ۳) $2^{2h} - 1$ ۴) $2^{2h+1} - 1$

۱۸) یک شمارنده‌ی k بیتی در ابتدا مقدار صفر دارد. این شمارنده را $n = 2^k$ بار و هر بار یک واحد افزایش می‌دهیم. مجموع تعداد تغییرات بیت‌های این شمارنده را از ۱ به صفر یا از صفر به یک چند تاست؟ این تعداد بخش‌بر 2^k هزینه‌ی سرشکن‌شده‌ی هر عمل افزایش است.

۱) $\sum_{i=1}^{n} \frac{n}{2^i}$ ۲) $\Theta(n^2)$ ۳) $\sum_{i=1}^{n} \frac{n}{2^{i-2}}$ ۴) $\Theta(n \lg n)$

۱۹) فرض کنید که $(a_1, a_2, \ldots, a_n)$ دنباله‌ی میان‌ترتیب یک درخت دودویی جست‌وجو باشد. کدام‌یک از گزینه‌های زیر در مورد این درخت نادرست است؟

۱) ترتیب حذف دو عنصر a_i و a_{i+1} تأثیری در درخت حاصل ندارد.
۲) اگر a_i برگ نباشد، یا a_{i-1} در زیر درخت چپ a_i است یا a_{i+1} در زیر درخت راست آن (یا هردو).
۳) اگر a_i $(i < n)$ دو فرزند داشته باشد، a_{i+1} فرزند چپ ندارد.
۴) اگر a_i $(i > 1)$ دو فرزند داشته باشد، a_{i-1} فرزند راست ندارد.

۲۰) می‌خواهیم داده‌ساختاری برای نگه‌داری n نقطه بر روی محور xها طراحی کنیم تا اعمال $\textsc{Insert}(x)$ (درج نقطه‌ی x)، $\textsc{Delete}(x)$ (حذف نقطه‌ی x) و $\textsc{Find}(a, b)$ (یافتن تعداد نقاط موجود در بازه‌ی $[a, b]$) را در زمان کارا انجام دهد. کدام‌یک از زمان‌های زیر ممکن است. فرض کنید که تعداد نقاط پاسخ آخرین عمل k است.

۱) $\textsc{Find}$ در زمان $k \lg n$ و بقیه در زمان $O(\lg n)$
۲) $\textsc{Find}$ در زمان $\lg n + k$ و بقیه در زمان $O(\lg n)$
۳) همه در زمان $O(\lg n)$
۴) همه در بدترین حالت در زمان $O(n)$

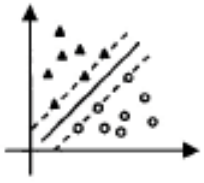
۲۱- **در رابطه با مشکل بیش برازش (overfitting) کدام یک از گزینه‌های زیر نادرست است؟**

۱) اعتبارسنجی متقابل (cross-validation) می‌تواند بروز بیش برازش را تشخیص دهد.
۲) بیش‌برازش در هنگام طبقه‌بندی خطرساز می‌شود، ولی در هنگام خوشه‌بندی رخ نمی‌دهد.
۳) کاهش ابعاد با استفاده از روش‌های انتخاب ویژگی، باعث کاهش احتمال بیش برازش می‌شود.
٤) احتمال رخ دادن بیش برازش در طبقه‌بندی که با داده‌های بیشتری آموزش داده شده است، کمتر است.

۲۲- **کدام یک از ادعاهای زیر منطقا قابل پذیرش است؟**

۱) الگوریتم A از الگوریتم B بهتر است، زیرا نرخ خطای آن بر روی داده‌های آموزشی کمتر است.
۲) الگوریتم A از الگوریتم B بهتر است، زیرا نرخ خطای آن بر روی داده‌های تست کمتر است (نتایج به ازای پارامتر L=1.7894898783).
۳) الگوریتم A از الگوریتم B بهتر است، زیرا نرخ خطای آن بر روی داده‌های تست کمتر است (نتایج به ازای بهترین مقدار پارامتر L برای مجموعه داده‌های تست).
٤) الگوریتم A از الگوریتم B بهتر است، زیرا نرخ خطای آن بر روی داده‌های تست کمتر است (نتایج به ازای بهترین مقدار پارامتر L بر اساس اعتبارسنجی متقابل 10-fold).

۲۳- **تخمین خطای leave-one-out برای جداساز با بیشترین حاشیه در شکل زیر چقدر است؟ (مجموع تعداد نمونه‌های غلط تشخیص داده شده در تمامی مراحل مد نظر است).**



۱) صفر
۲) ۱
۳) ٦
٤) ۱۵

۲۴- فرض کنید می‌خواهیم جهت بردار $\mathbf{v}$ را به قسمی پیدا کنیم که در صورتی که داده‌ها در آن جهت تصویر شوند ($\mathbf{y = v^t x}$)، دارای بیشترین جداسازی ممکن باشند (نظیر روش جداسازی فیشر). در صورتی که μ_1 و σ_1^2 به ترتیب میانگین و پراکندگی داده‌های کلاس اول، پس از تصویر شدن در جهت $\mathbf{v}$ و μ_2 و σ_2^2 به ترتیب میانگین و پراکندگی داده‌های کلاس دوم، پس از تصویر شدن در همان جهت باشند، کدام یک از توابع زیر یک تابع هدف مناسب برای این منظور می‌باشد، به قسمی که بیشینه‌سازی آن نسبت به بردار $\mathbf{v}$، جهت آن بردار را به ما بدهد؟ دقت داشته باشید که در توابع زیر $\mathbf{v}$ به صورت ضمنی در پارامترهای میانگین و پراکندگی حضور دارد.

۱) $J(v)=\frac{(\mu_1-\mu_2)^2}{\sigma_1^2-\sigma_2^2}$

۲) $J(v)=\frac{\sigma_1^2+\sigma_2^2}{(\mu_1-\mu_2)^2}$

۳) $J(v)=\frac{\sigma_2^2(\mu_1-\mu_2)^2}{\sigma_1^2}$

۴) $J(v)=\frac{(\mu_1-\mu_2)^2}{\sigma_1^2}+\frac{(\mu_1-\mu_2)^2}{\sigma_2^2}$

۲۵- می‌خواهیم $\mathbf{n}$ داده $\mathbf{d}$ بعدی، که $\mathbf{n<<d}$ است، را با روش تحلیل مولفه اصلی **(PCA)** به قسمی کاهش ابعاد دهیم که خطای تبدیل و نمایش در فضای جدید صفر باشد (تمام انرژی فضای قبلی حفظ شود). کمترین مقداری که تعداد ابعاد فضای جدید قطعا از آن بیشتر نخواهد بود، کدام است؟

۱) $d/2$ بعد

۲) $d-1$ بعد

۳) $d-n$ بعد

۴) $n-1$ بعد

۲۶- در یک مساله کلاسه‌بندی دو کلاسه در فضای دو بعدی، اگر داده‌های دو کلاس دارای احتمال‌های پیشین یکسان بوده و احتمال شرطی آنها دو مدل گاوسی با میانگین‌های غیر یکسان باشند، در چه صورت مرز حاصل از جداساز با کمترین نرخ خطا خطی خواهد بود؟

۱) تنها در صورتی که ماتریس‌های کوواریانس با هم برابر بوده و قطری باشند.

۲) تنها در صورتی که ماتریس کوواریانس‌های هر دو مدل گاوسی با هم برابر باشند.

۳) تنها در صورتی که ماتریس‌های کوواریانس هر دو مدل گاوسی، ماتریس همانی $(\mathbf{I})$ باشد.

۴) تنها در صورتی که ماتریس‌های کوواریانس هر دو مدل گاوسی قطری (و نه لزوما برابر) باشند.

۲۷- یک مساله $\mathbf{k}$ کلاسه را با مدل‌های احتمال شرطی یکنواخت $P(x|C_i)\sim U(-(2^{i-z}),\ 2^{i-z})$ در نظر بگیرید. در صورت استفاده از کلاسه‌بند آماری بر پایه بیشینه شباهت **(maximum likelihood ratio test)** مقدار $\mathbf{P(error)}$ وقتی $\mathbf{k}$ به سمت بی‌نهایت میل کند، برابر چه مقداری می‌شود؟

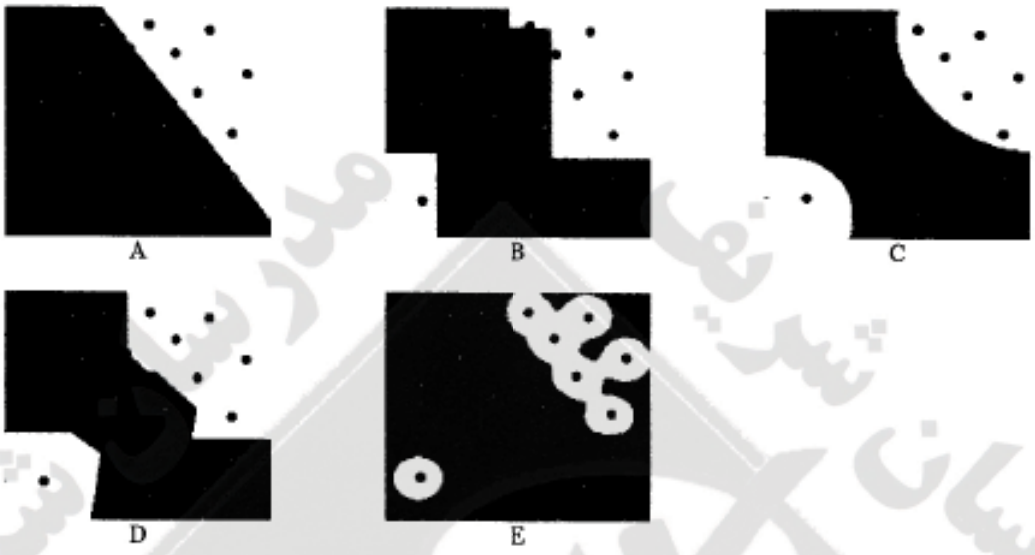
۱) ¼

۲) ½

۳) 0

۴) 1

۲۸- برای داده‌های نشان داده شده در شکل زیر، کدام خروجی‌ها می‌توانند به ترتیب از راست به چپ حاصل ماشین بردار پشتیبان خطی، ماشین بردار پشتیبان با کرنل گاوسی و $\sigma = 1$ و ماشین بردار پشتیبان با کرنل گاوسی و $\sigma = 0.25$ باشند؟



۱) D و C ،A ۲) E و C ،A ۳) D و C ،B ٤) E و C ،B

۲۹- کدام یک از گزینه‌های زیر صحیح است؟

۱) حذف یکی از بردارهای پشتیبان از مجموعه داده‌ها ممکن است باعث شود که حاشیه ماشین بردار پشتیبان کاهش پیدا کند.

۲) در مسائل کلاسه‌بندی، مرزهای با بالاترین حاشیه‌ای که ماشین‌های جداساز خطی می‌سازند، دارای کمترین خطای تعمیم (generalization) در بین تمامی کلاسه‌بندهای خطی دیگر هستند.

۳) مقدار حاشیه بدست آمده در دو ماشین بردار پشتیبان با کرنل‌های متفاوت $K_1(x_i,x_j)$ و $k_2(x_i,x_j)$ بر روی یک مجموعه داده آموزشی یکسان نشان دهنده میزان کارایی دو ماشین فوق بر روی مجموعه داده تست است.

٤) روش ۱-نزدیک‌ترین همسایگی (1-NN) با استفاده از فاصله اقلیدسی در یک فضای ویژگی با ابعاد بالاتر که بوسیله نگاشت با یک کرنل گاوسی بدست آمده است، ممکن نیست نتایج بهتری را به نسبت فضای اولیه به دست دهد.

۳۰- در صورتی که در الگوریتم خوشه‌بندی **k-means** k برابر تعداد نمونه‌های مجموعه داده آموزشی در نظر گرفته شود، خوشه‌بندی مجموعه تست معادل کلاسه‌بندی داده‌های تست با کدام یک از روش‌های زیر خواهد بود؟

۱) k-NN و k برابر ۱

۲) k-NN و k برابر تعداد نمونه‌های مجموعه داده آموزشی

۳) کلاسه‌بندی آماری تنها با استفاده از احتمال‌های پیشین کلاس‌ها

٤) بستگی به مجموعه داده‌های آموزش و تست دارد.

۳۱- H مجموعه تمام مستطیل هایی است که اضلاع آنها موازی محورهای مختصات است. بعد VC مجموعه H چند است؟

۱) 2

۲) 3

۳) 4

۴) 5

۳۲-پس از کامل شدن یادگیری یک عامل یادگیر Q، تابع Q چه چیزی را نشان می دهد؟

۱) کنش بهینه برای هر حالت

۲) میانگین پاداش دریافتی برای رسیدن به هر حالت

۳) میانگین پاداش دریافتی آتی

۴) تفاضل میانگین پاداش دریافتی و پاداش لحظه ای

۳۳-الگوریتم AdaBoost را روی مجموعه ای با m نمونه آموزشی اجرا می کنیم. فرض کنید خطای دسته بند پایه k ام که با e_k نشان داده می شود حداکثر برابر 0.5-b باشد که b یک عدد مثبت می باشد. پس از آموزش دست کم چند دسته بند پایه، خطای آموزشی برابر صفر می شود؟

۱) $\ln\left(\frac{m}{2b^2}\right)$

۲) $\ln\left(\frac{m}{b^2}\right)$

۳) $\ln\left(\frac{m}{2b}\right)$

۴) $\ln\left(\frac{m}{b}\right)$

۳۴- در الگوریتم های یادگیری درخت تصمیم، فضای فرضیه ها شامل کدامیک از گزینه های زیر است؟

۱) همه گره ها در درخت تصمیم

۲) همه درخت های تصمیم

۳) همه ترکیبات ویژگی ها

۴) همه برگ های درخت تصمیم

۳۵- کدامیک از گزینه های زیر در باره فرضیه های MAP و ML درست است؟

۱) $h_{MAP} = \underset{h}{argmax}\, P(D|h)$ $\quad h_{ML} = \underset{h}{argmax}\, P(h|D)$

۲) $h_{MAP} = \underset{h}{argmax}\, P(h|D)$ $\quad h_{ML} = \underset{h}{argmax}\, P(D,h)$

۳) $h_{MAP} = \underset{h}{argmax}\, P(h|D)$ $\quad h_{ML} = \underset{h}{argmax}\, P(D|h)$

۴) $h_{MAP} = \underset{h}{argmax}\, P(h,D)$ $\quad h_{ML} = \underset{h}{argmax}\, P(D|h)$

۳۶- هنگامی که یک الگوریتم بایاس از نوع محدودیت (Restriction bias) دارد این الگوریتم کدامیک از محدودیت های زیر را دارد؟

۱) فضای فرضیه ها توانایی نمایش همه فرضیه ها جهت دسته بندی داده ها را ندارد.

۲) قدرت تعمیم این الگوریتم ها کم است و برای افزایش قدرت تعمیم باید بایاس کمینه شود.

۳) ترتیب ارائه نمونه های آموزشی به الگوریتم روی کارایی الگوریتم یادگیری تاثیر می گذارد.

۴) برای یادگیری، الگوریتم نیاز به نمونه های آموزشی زیادی دارد.

۳۷ در یادگیری تقویتی، کدام گزینه در باره تابع سیاست درست است؟

۱) این تابع مشخص می کند که در هر حالت چه کنشی باید انجام شود.

۲) این تابع مشخص می کند که برای دریافت پاداش مشخص در هر حالت چه کنشی باید انجام شود.

۳) این تابع ارزش هر کنش در هر حالت را مشخص می کند.

٤) این تابع حالت بعدی را پس از انتخاب هر کنش مشخص می کند.

۳۸- در یک مسئله یادگیری، هر نمونه دارای دو ویژگی دودویی می باشد. اگر فضای فرضیه شامل تمامی درختان تصمیم با عمق 2 باشد این فضا دارای چند درخت است؟

۱) 8

۲) 12

۳) 14

٤) 16

۳۹- کدامیک از گزینه های زیر در باره الگوریتم های یادگیری درخت تصمیم و بیز ساده (Naïve Bayes) درست است؟

۱) در هر دو الگوریتم زوج ویژگی ها باید مستقل از هم باشند.

۲) در هر دو الگوریتم نیازی نیست زوج ویژگی ها مستقل یا وابسته باشند.

۳) در هر دو الگوریتم زوج ویژگی ها به شرط برچسب باید به هم وابسته باشند.

٤) در هر دو الگوریتم زوج ویژگی ها به شرط برچسب باید مستقل از هم باشند.

۴۰-یک دسته بند روی مجموعه آموزشی با 100 نمونه آموزش می بینید و پس از آموزش 90 نمونه آموزشی را درست دسته بندی می کند. سپس این دسته بند روی یک مجموعه آزمایشی با 50 نمونه آزمایش می شود که 43 نمونه آزمایشی را درست دسته بندی می کند. با 90٪ اطمینان مشخص نمایید که خطای واقعی این دسته بند در چه بازه ای است؟ (توجه : $Z_{90}=1.28$)

۱) $0.14 \pm \frac{1.28}{10}\sqrt{0.28(1-0.14)}$

۲) $0.1 \pm \frac{1.28}{10}\sqrt{0.1(1-0.1)}$

۳) $0.14 \pm \frac{1.28}{10}\sqrt{0.1(1-0.1)}$

٤) $0.1 \pm \frac{1.28}{10}\sqrt{0.28(1-0.14)}$

۴۱ - کدامیک از گزینه های زیر درست است؟

۱) اگر یک دسته بند روی مجموعه آموزشی با تعداد نمونه های کم آموزش داده شود دسته بند بیش برازش (overfit) می شود.

۲) اگر یک دسته بند روی مجموعه آموزشی با m نمونه آموزش داده شود اگر $m \to \infty$ آنگاه خطای آموزش به خطای واقعی همگرا می شود.

۳) الگوریتم AdaBoost نمونه های آموزشی را به ترتیب از نمونه های سخت به آسان یاد می گیرد.

٤) اگر توزیع یکنواخت برای داده ها فرض شود تخمین MAP به تخمین ML تبدیل می شود.

۴۲ - کدامیک از گزینه های زیر درست است؟

۱) بعد VC فضای فرضیه $[0,a)$ (برای a حقیقی مثبت) برابر 2 است.

۲) بعد VC فضای فرضیه (a,b) (برای a و b حقیقی و $a < b$) برابر 3 است.

۳) بعد VC فضای فرضیه $(-\infty,a)$ (برای a حقیقی) برابر 2 است.

٤) برای هر فضای فرضیه H، شرط $VC(H) \leq \log_2(|H|)$ برقرار است.

۴۳- کدامیک از الگوریتم های دسته بندی زیر با داده های آموزشی سازگار نیستند؟

۱) 1-NN

۲) K-NN

۳) Candidate Elimination

۴) ID3

۴۴- برای تولید دسته بندهای پایه متنوع کدامیک از روش های زیر موثر نمی باشد؟

۱) استفاده از دسته بندهای پایدار

۲) استفاده از دسته بندهای ناپایدار

۳) استفاده از مجموعه های آموزشی متفاوت

۴) استفاده از مجموعه ویژگی های متفاوت

۴۵- فضای فرضیه H را در نظر بگیرید. احتمال اینکه یک فرضیه از H که با m نمونه آموزشی سازگار باشد و دارای حداقل خطای ε باشد چقدر است؟

۱) حداقل $\frac{(13/\varepsilon)^4}{2^{\frac{\varepsilon m}{4}-1}}$

۲) حداقل $\frac{(13/\varepsilon)}{2^{\frac{\varepsilon m}{4}-1}}$

۳) حداقل $\frac{(13/\varepsilon)^4}{2^{\varepsilon m}}$

۴) حداقل $\frac{(13/\varepsilon)^4}{2^{\frac{\varepsilon m}{4}}}$