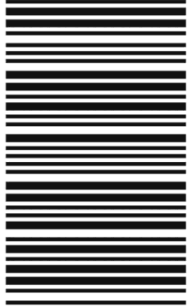


کد کنترل

532

F



532F

آزمون (نیمه متمرکز) ورود به دوره های دکتری - سال ۱۴۰۲

دفترچه شماره (۱)

صبح پنجشنبه

۱۴۰۱/۱۲/۱۱



«اگر دانشگاه اصلاح شود مملکت اصلاح می شود.»
امام خمینی (ره)

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

مهندسی کامپیوتر - هوش مصنوعی (کد ۲۳۵۶)

زمان پاسخ گویی: ۱۳۵ دقیقه

تعداد سؤال: ۴۵

عنوان مواد امتحانی، تعداد و شماره سؤالات

ردیف	مواد امتحانی	تعداد سؤال	از شماره	تا شماره
۱	مجموعه دروس تخصصی: - ساختمان داده ها و طراحی الگوریتمها - شناسایی الگو - یادگیری ماشین	۴۵	۱	۴۵

این آزمون نمره منفی دارد.

استفاده از ماشین حساب مجاز نیست.

حق چاپ، تکثیر و انتشار سؤالات به هر روش (الکترونیکی و...) پس از برگزاری آزمون، برای تمامی اشخاص حقیقی و حقوقی تنها با مجوز این سازمان مجاز می باشد و با متخلفین برابر مقررات رفتار می شود.

* داوطلب گرامی، عدم درج مشخصات و امضا در مندرجات جدول زیر، به منزله عدم حضور شما در جلسه آزمون است.

اینجانب با شماره داوطلبی با آگاهی کامل، یکسان بودن شماره صندلی خود با شماره داوطلبی مندرج در بالای کارت ورود به جلسه، بالای پاسخنامه و دفترچه سؤالات، نوع و کد کنترل درج شده بر روی جلد دفترچه سؤالات و پایین پاسخنامه ام را تأیید می نمایم.

امضا:

مجموعه دروس تخصصی (ساختمان داده‌ها و طراحی الگوریتم‌ها - شناسایی الگو - یادگیری ماشین):

۱- الگوریتم فلوید - وارشال از یک الگوریتم برای حل مسئله کوتاه‌ترین مسیرهای تمام جفت رؤوس در یک گراف جهت‌دار $G = (V, E)$ در زمان استفاده می‌کند.

(۱) حریمانه، $\Theta(v^3)$ (۲) حریمانه، $\Theta(V^2 \log E)$

(۳) برنامه‌نویسی پویا، $\Theta(v^3)$ (۴) برنامه‌نویسی پویا، $\Theta(V^2 \log E)$

۲- با فرض اینکه $P \neq NP$ باشد، کدام مورد درست است؟

(۱) $NP - hard = NP$ (۲) $NP - complete = P$

(۳) $NP - complete = NP$ (۴) $NP - complete \cap P = \emptyset$

۳- کم‌ترین تعداد مقایسه مورد نیاز برای تعیین اینکه یک عدد صحیح بیش از $\frac{n}{4}$ مرتبه در یک آرایه مرتب از اعداد صحیح به طول n ظاهر می‌شود، از کدام مرتبه است؟

(۱) $\Theta(1)$ (۲) $\Theta(\log n)$

(۳) $\Theta(n)$ (۴) $\Theta(n \log n)$

۴- n آرایه نامرتب A_1, \dots, A_n را در نظر بگیرید (n عددی فرد است). هر کدام از این آرایه‌ها دارای n عنصر متمایز است. هیچ عنصر مشترکی میان هیچ دو آرایه‌ای وجود ندارد. کمترین پیچیدگی زمانی الگوریتمی برای محاسبه میانه این آرایه‌ها از چه مرتبه‌ای است؟

(۱) $\Theta(n)$ (۲) $\Theta(n \log n)$

(۳) $\Theta(n^2)$ (۴) $\Omega(n^2 \log n)$

۵- فرض کنید $W(n)$ و $A(n)$ ، به ترتیب، نشان‌دهنده بدترین حالت و میانگین زمان اجرای الگوریتم اجراشده بر روی ورودی با اندازه n باشند. کدام مورد همواره درست است؟

(۱) $A(n) = O(W(n))$ (۲) $A(n) = \Theta(W(n))$

(۳) $A(n) = \Omega(W(n))$ (۴) $A(n) = o(W(n))$

۶- یک آرایه مرتب‌شده از اعداد داریم. می‌خواهیم دو عدد در این آرایه پیدا کنیم که جمع آن دو عدد مساوی یک عدد داده‌شده x باشد. کمترین پیچیدگی زمانی حل این مسئله کدام است؟

(۱) $\Theta(n)$ (۲) $\Theta(n^2)$

(۳) $\Theta(\log n)$ (۴) $\Theta(n \log n)$

۷- فرض کنید آرایه‌ای از اعداد صحیح $A = [a_1; a_2; \dots; a_n]$ داده شود. فرض کنید یک اندیس (ناشناخته) k وجود دارد به طوری که زیر آرایه $A = [a_1; a_2; \dots; a_k]$ به ترتیب اکیداً افزایشی مرتب شده است و زیر آرایه $A = [a_k; a_{k+1}; \dots; a_n]$ به ترتیب اکیداً نزولی مرتب شده است (یعنی اگر $1 \leq i < j \leq k$ ، آنگاه $a_i < a_j$ ، و اگر $k \leq i < j \leq n$ ، آنگاه $a_i > a_j$ هدف شما تعیین k است. یک الگوریتم بهینه برای حل این مسئله چه زمان اجرایی دارد؟

$$\begin{array}{ll} \Theta(n^2 \log n) & (۱) \\ \Theta(n \log n) & (۲) \\ \Theta(\log n) & (۴) \\ \Theta(n) & (۳) \end{array}$$

۸- کدام یک از موارد زیر درست است؟

(۱) آرایه $A = [۱۰; ۳; ۵; ۱; ۴; ۲]$ یک max heap است.

(۲) هر مسئله محاسباتی با اندازه ورودی n را می‌توان با یک الگوریتمی با زمان چندجمله‌ای بر حسب n حل کرد.

(۳) برای تمام توابع مثبت $f(n)$ ، $g(n)$ و $h(n)$ ، اگر $f(n) = O(g(n))$ و $f(n) = \Omega(h(n))$ باشد، آنگاه $g(n) + h(n) = \Omega(f(n))$ است.

(۴) اگر هر رقم جداگانه در RADIX SORT را با استفاده از INSERTION SORT به جای COUNTING SORT مرتب کنیم، آنگاه RADIX SORT به درستی کار نمی‌کند (یعنی خروجی صحیح را تولید نمی‌کند).

۹- کدام مورد زیر مطمئناً عبارت $f(n) = \Omega(g(n))$ را پشتیبانی می‌کند؟

$$(۱) \quad f(n) \leq 4 \times g(n) \text{ برای تمام } n \geq 1$$

$$(۲) \quad f(n) \geq 4 \times g(n) \text{ برای تمام } n \geq 136$$

$$(۳) \quad f(n) \leq 4 \times g(n) \text{ برای تمام } n \geq 100$$

$$(۴) \quad \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$$

۱۰- کدام یک از گزاره‌های زیر درست است؟

الف- اگر مسئله P_1 بتواند به مسئله P_2 در زمان خطی کاهش (reduce) یابد، آنگاه اگر P_2 یک مسئله NP-hard باشد، می‌توان نتیجه گرفت P_1 نیز NP-hard است.

ب- یک Clique در یک گراف بدون جهت لزوماً یک vertex cover در گراف مکمل نیست.

(۱) فقط گزاره «الف» درست است.

(۳) هر دو گزاره «الف» و «ب» درست است.

(۴) هر دو گزاره «الف» و «ب» نادرست است.

۱۱- کدام یک از گزاره‌های زیر درست است؟

الف- اینکه زمان حل یک مسئله P حد پایین $\Omega(n^2)$ دارد به این معنی است که برای هر الگوریتم A که P را حل می‌کند فقط برخی از نمونه‌های P وقتی به عنوان ورودی به A داده شوند، باعث می‌شود A زمان $\Omega(n^2)$ صرف کند.

ب- اینکه زمان حل یک مسئله P حد پایین $\Omega(n^2)$ دارد به این معنی است که برای هر الگوریتم A که P را حل می‌کند هر نمونه از P که به عنوان ورودی به A داده شود، باعث می‌شود A زمان $\Omega(n^2)$ صرف کند.

(۱) فقط گزاره «الف» درست است.

(۳) هر دو گزاره «الف» و «ب» درست هستند.

(۲) فقط گزاره «ب» درست است.

(۴) هر دو گزاره «الف» و «ب» نادرست هستند.

- ۱۲- کدام یک از گزاره‌های زیر درست است؟
 الف- اگر یک الگوریتم زمان چندجمله‌ای برای یک مسئله که NP-hard است ارائه شود، آنگاه می‌توان نتیجه گرفت $P = NP$ است.
 ب- اگر یک مسئله NP-complete است، آنگاه می‌توان نتیجه گرفت که آن مسئله هیچ راه‌حلی ندارد.
 (۱) فقط گزاره «الف» درست است.
 (۲) فقط گزاره «ب» درست است.
 (۳) هر دو گزاره «الف» و «ب» درست هستند.
 (۴) هر دو گزاره «الف» و «ب» نادرست هستند.
- ۱۳- کدام یک از گزاره‌های زیر درست است؟
 الف- اگر یک مسئله که به کلاس NP تعلق دارد یک راه‌حل زمان چند جمله‌ای داشته باشد، آنگاه $P = NP$ است.
 ب- اگر کسی یک حد پایین زمان نمایی برای یک مسئله که NP-complete است بدهد، آنگاه $P \neq NP$ است.
 (۱) فقط گزاره «الف» درست است.
 (۲) فقط گزاره «ب» درست است.
 (۳) هر دو گزاره «الف» و «ب» درست هستند.
 (۴) هر دو گزاره «الف» و «ب» نادرست هستند.
- ۱۴- مسئله کوله‌پشتی ۱-۰ را در نظر بگیرید که n شیء با وزن صحیح داریم و گنجایش کوله‌پشتی عدد صحیح M است. این مسئله دارای یک الگوریتم مبتنی بر روش برنامه‌ریزی پویا با زمان $O(M.n)$ است. این مرتبه زمانی بر حسب اندازه مسئله چگونه است؟
 (۱) خطی
 (۲) درجه دو
 (۳) نمایی
 (۴) شبه چندجمله‌ای
- ۱۵- برای پیدا کردن k امین عدد در میان n عدد که به‌عنوان کلید در گره‌های یک درخت جستجوی دودویی متوازن ذخیره شده‌اند، کمترین پیچیدگی زمانی ممکن کدام است؟ (هر گره درخت فقط شامل کلید و اشاره‌گر به پدر و فرزند چپ و راست است.)
 (۱) $\Theta(n)$
 (۲) $\Theta(\log n)$
 (۳) $\Theta(n \log n)$
 (۴) $\Theta(\log \log n)$
- ۱۶- n عدد به‌عنوان کلید در گره‌های یک درخت جستجوی دودویی متوازن ذخیره شده‌اند. هر گره علاوه بر کلید و اشاره‌گر به پدر و فرزند چپ و راست، تعداد گره‌های زیر درخت خود را هم نگهداری می‌کند. برای پیدا کردن rank کلید یک گره (یعنی اینکه کلید گره چندمین عدد در بین n عدد است) کمترین پیچیدگی زمانی ممکن کدام است؟
 (۱) $\Theta(\log \log n)$
 (۲) $\Theta(\log n)$
 (۳) $\Theta(n)$
 (۴) $\Theta(n \log n)$
- ۱۷- در یک درخت قرمز - سیاه، طول طولانی‌ترین مسیر ساده از یک گره x به یک برگ در زیر درخت خودش حداکثر چند برابر طول کوتاه‌ترین مسیر از گره x به یک برگ در زیر درخت خودش است؟
 (۱) ۴
 (۲) ۳
 (۳) ۲
 (۴) ۱
- ۱۸- کدام یک از گزاره‌های زیر درست است؟
 الف- هر درخت جستجوی دودویی دلخواه با n گره می‌تواند به یک درخت جستجوی دودویی دلخواه دیگر با n گره با انجام $O(n)$ عمل rotation تبدیل شود.
 ب- برای هر دو تابع $f(n)$ و $g(n)$ یکی از سه حالت (۱) $f(n) \in o(g(n))$ ، (۲) $f(n) \in w(g(n))$ و (۳) $f(n) \in \theta(g(n))$ برقرار است.
 (۱) فقط گزاره «الف» درست است.
 (۲) فقط گزاره «ب» درست است.
 (۳) هر دو گزاره «الف» و «ب» درست است.
 (۴) هر دو گزاره «الف» و «ب» نادرست است.

۱۹- کمترین پیچیدگی زمانی ممکن برای مرتب‌سازی n عدد طبیعی که مقادیر کمتر از n^2 دارند، کدام است؟

(۱) $\Theta(n \log^2 n)$ (۲) $\Theta(n^2)$

(۳) $\Theta(n \log n)$ (۴) $\Theta(n)$

۲۰- اگر عدد ۳۶۳ را در یک درخت جستجوی دودویی، جستجو کنیم، کدام دنباله زیر نمی‌تواند دنباله‌ای از کلید گره‌هایی

باشد که بررسی می‌شوند؟ (ترتیب از راست به چپ است.)

(۱) ۳۶۳، ۳۹۷، ۳۴۴، ۳۳۰، ۳۹۸، ۴۰۱، ۲۵۲، ۲

(۲) ۳۶۳، ۳۶۲، ۲۵۸، ۸۹۸، ۲۴۴، ۹۱۱، ۲۲۰، ۹۲۴

(۳) ۳۶۳، ۲۷۸، ۳۸۱، ۳۸۲، ۲۶۶، ۲۱۹، ۳۸۷، ۳۹۹، ۲

(۴) ۳۶۳، ۲۴۵، ۹۱۲، ۲۴۰، ۹۱۱، ۲۰۲، ۹۲۵

۲۱- نزدیک‌ترین فاصله نقطه z از ابرصفحه $W^T x = 0$ کدام است؟

(۱) $W^T z$

(۲) $\frac{W^T z}{\|W\|}$

(۳) $\|W\| \cdot \|z\|$

(۴) $\frac{W^T z}{\|W\|^2}$

۲۲- در دسته‌بندی یک مسئله M دسته‌ای به کمک روش بیزین، کدام رابطه در مورد احتمال خطای دسته‌بندی

دسته‌بند بهینه یا همان P_e صادق است؟

(۱) $\frac{1}{M} \leq P_e \leq 1 - \frac{1}{M}$

(۲) $P_e \leq 1 - \frac{M}{1+M}$

(۳) $P_e \leq 1 - \frac{1}{M}$

(۴) $P_e \leq \frac{1}{M}$

۲۳- در ساخت یک جنگل تصادفی هر درخت را براساس نمونه‌گیری Bootstrap از داده‌های آموزشی می‌سازیم. اگر n

تعداد کل نمونه‌های آموزشی و d تعداد ویژگی باشد، به‌طور متوسط هر درخت براساس چه تعداد نمونه آموزشی

ساخته می‌شود؟

(۱) $(1 - \frac{1}{n})^n$

(۲) $(1 - \frac{1}{n})^d$

(۳) $1 - (1 - \frac{1}{n})^d$

(۴) $1 - (1 - \frac{1}{n})^n$

۲۴- فرض کنید در یک مسئله دسته‌بندی، داده‌های دارای دو ویژگی x_1 و x_2 هستند که هر کدام مقادیر -1 ، 0 ، یا $+1$ را می‌گیرند و برچسب طبقه هم مقدار صفر یا یک را می‌گیرد. سه فرضیه h_1 ، h_2 ، h_3 به صورت زیر تعریف شده‌اند:

$$h_1(x) = \begin{cases} 1 & \text{if } x_1 = 1 \\ 0 & \text{else} \end{cases} \quad h_2(x) = \begin{cases} 1 & \text{if } x_1 \neq x_2 \\ 0 & \text{else} \end{cases} \quad h_3(x) = \begin{cases} 1 & \text{if } x_1 \cdot x_2 = 0 \\ 0 & \text{else} \end{cases}$$

کدام یک از فرضیه‌های فوق را می‌توان از طریق یک پرسپترون (با گرفتن وزن‌های مناسب) پیاده‌سازی کرد؟

(۱) h_1

(۲) h_2

(۳) h_3

(۴) h_2 و h_3

۲۵- در دسته‌بندی یک مسئله دو دسته‌ای با داده‌های آموزش به شرح جدول زیر، دو دسته‌بند **Naïve Bayes** و **Bayes** داده آزمون ($\text{Weight} = 80$ ، $\text{Height} = 180$) را در کدام دسته، دسته‌بندی می‌کنند؟

#	Height	Weight	Class
۱	۸۰	۱۸۰	+
۲	۷۰	۱۸۰	+
۳	۸۰	۱۸۰	-
۴	۸۰	۱۸۰	+
۵	۷۰	۱۷۰	+
۶	۵۰	۱۷۰	-
۷	۴۰	۱۶۰	-
۸	۳۰	۱۴۰	-
۹	۱۰۰	۱۷۰	+
۱۰	۱۱۰	۱۶۰	+
۱۱	۵۰	۱۶۰	-
۱۲	۱۱۰	۱۶۰	+

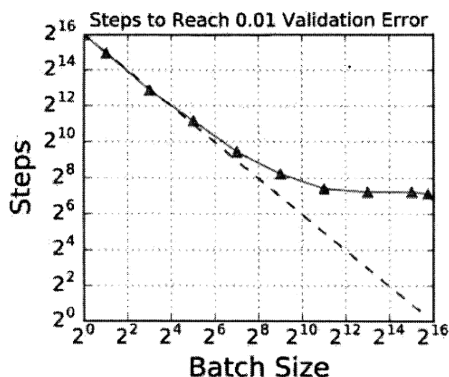
(۱) Naïve Bayes $\rightarrow -$ ، Bayes $\rightarrow +$

(۲) Naïve Bayes $\rightarrow +$ ، Bayes $\rightarrow -$

(۳) هر دو مثبت

(۴) هر دو منفی

۲۶- نمودار زیر در رابطه با تعداد تکرارهای مورد نیاز برای دستیابی به کمینه خطا با استفاده از الگوریتم نزول در امتداد گرادیان تصادفی (SGD) را در نظر بگیرید. گزینه درست کدام است؟



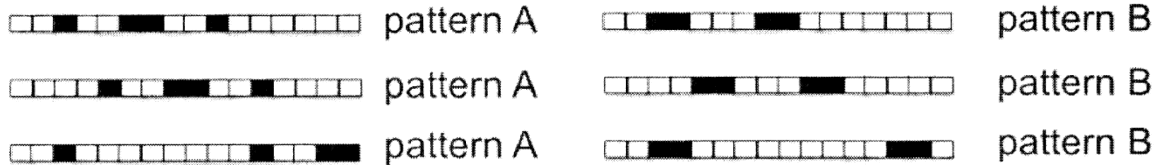
(۱) اندازه دسته یا Batch-size های بزرگتر، واریانس را در تخمین گرادیان در الگوریتم SGD افزایش می‌دهند.

(۲) با تعداد بروزرسانی یکسان، اندازه دسته‌های بزرگتر نسبت به اندازه دسته‌های کوچکتر سریع‌تر همگرا می‌شوند.

(۳) براساس این شکل، اندازه دسته‌های نسبتاً کوچک هم می‌تواند همگرایی سریع را تضمین نماید.

(۴) همه موارد

۲۷- مسئله دسته‌بندی دودویی زیر را در نظر بگیرید. فرض کنید مجموعه آموزش حاوی نمونه‌های متفاوتی از الگوی A و B باشند. آیا یک شبکه عصبی حاوی یک لایه کانولوشنی یک‌بعدی با یک تابع فعال‌سازی (Activation function) خطی که به دنبال آن لایه خطی با خروجی لجستیک قرار دارد، می‌تواند به‌طور کامل همه نمونه‌های آموزش را دسته‌بندی نماید؟



(۱) بله، فقط باید خروجی لجستیک تغییر یابد.

(۲) خیر، اما افزودن یک لایه کانولوشنی یک‌بعدی دیگر می‌تواند مسئله را حل کند.

(۳) بله، لایه کانولوشنی به‌خوبی می‌تواند شرایط جداپذیری را در این الگوها قابل شناسایی نماید.

(۴) خیر، این مسئله ذاتاً قابل جدایی‌پذیری خطی نیست و نیاز است شبکه پیچیده‌تری جایگزین گردد.

۲۸- در یک مسئله دودسته‌ای بر روی داده‌های یک‌بعدی، اگر تابع توزیع دسته‌های w_1 و w_2 به ترتیب توابع گوسی $N(0, \sigma^2)$ و $N(1, \sigma^2)$ باشد، مقدار آستانه x_0 که مقدار Average risk را کمینه می‌کند، کدام است؟ λ_{k1} را ریسک آن در نظر بگیرید که داده‌ای از دسته k را در دسته i دسته‌بندی نماییم. همچنین فرض کنید $(\lambda_{11} = \lambda_{22} = 0)$

$$x_0 = \frac{1}{2} - \sigma^2 \ln \frac{\lambda_{12}P(w_1)}{\lambda_{21}P(w_2)} \quad (2) \quad x_0 = \frac{1}{2} - \sigma^2 \ln \frac{\lambda_{21}P(w_2)}{\lambda_{12}P(w_1)} \quad (1)$$

$$x_0 = -\frac{1}{2} \sigma^2 \ln \frac{\lambda_{12}P(w_1)}{\lambda_{21}P(w_2)} \quad (4) \quad x_0 = -\frac{1}{2} \sigma^2 \ln \frac{\lambda_{21}P(w_2)}{\lambda_{12}P(w_1)} \quad (3)$$

۲۹- فرض کنید x_1, x_2, \dots, x_n نمونه‌هایی باشند که به‌صورت i.i.d از تابع $P(X=x) = \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}}$ انتخاب شده‌باشند. تخمین MLE برای پارامتر σ کدام است؟

$$\hat{\sigma} = \frac{n}{\sum_{i=1}^n |x_i|} \quad (2) \quad \hat{\sigma} = \frac{\sum_{i=1}^n |x_i|}{n} \quad (1)$$

$$\hat{\sigma} = \frac{2n}{\sum_{i=1}^n |x_i|} \quad (4) \quad \hat{\sigma} = \frac{\sum_{i=1}^n |x_i|}{2n} \quad (3)$$

۳۰- زنجیره مارکوف با ماتریس گذار زیر را در نظر بگیرید. امید ریاضی تعداد روزهای ماندن در وضعیت Sunny با شروع از وضعیت Sunny چند است؟

	rain	cloudy	sunny	
rain	0/4	0/3	0/3	1/25 (۱)
cloudy	0/2	0/6	0/2	2/5 (۲)
sunny	0/1	0/1	0/8	4 (۳)

$$A = \{a_{ij}\} = \begin{matrix} & \begin{matrix} \text{rain} \\ \text{cloudy} \\ \text{sunny} \end{matrix} \\ \begin{matrix} \text{rain} \\ \text{cloudy} \\ \text{sunny} \end{matrix} & \begin{bmatrix} 0/4 & 0/3 & 0/3 \\ 0/2 & 0/6 & 0/2 \\ 0/1 & 0/1 & 0/8 \end{bmatrix} \end{matrix}$$

۳۱- در مبحث کاهش ابعاد و انتخاب ویژگی‌ها کدام یک از گزاره‌های زیر نادرست است؟
 (۱) روش آنالیز تشخیص خطی (LDA) یک ترکیب خطی از ویژگی‌ها برای تفکیک بهتر فضای داده به دست می‌دهد.
 (۲) تبدیل موجک روشی برای انتقال فضای داده اصلی به فضای جدید با استفاده از ترکیب توابع موجک پایه دلخواه است.
 (۳) شبکه‌های عصبی خودکدگذار روشی برای کاهش بُعد داده به شمار می‌آیند که قادرند یک بردار ویژگی غنی از اطلاعات را به دست دهند.
 (۴) الگوریتم PCA و ICA فضای داده را به یک فضای جدید با همان تعداد بُعد یا کمتر انتقال می‌دهند به گونه‌ای که فاصله بین هر دو نمونه داده با فاصله در فضای قبلی یکسان باقی می‌ماند.

۳۲- در نگاشت داده $x \in \mathbb{R}^k$ از فضای داده به فضای با بعد بالاتر به صورت $y = \phi(x) \in \mathbb{R}^{2k+1}$ شکل تابع هسته متناظر با ضرب داخلی داده‌ها، در فضای هسته به کدام صورت است، اگر تابع $\phi(x)$ به صورت رابطه زیر باشد؟

$$\phi(x) = \left[\frac{1}{\sqrt{2}}, \cos x, \cos 2x, \dots, \cos kx, \sin x, \sin 2x, \dots, \sin kx \right]$$

$K(x_i, x_j) = \frac{2 \sin \left((k + \frac{1}{2}) \left(\frac{x_i - x_j}{2} \right) \right)}{\cos(x_i - x_j)} \quad (۲)$	$K(x_i, x_j) = \frac{2 \cos \left((k + \frac{1}{2}) (x_i - x_j) \right)}{\sin \left(\frac{x_i - x_j}{2} \right)} \quad (۱)$
$K(x_i, x_j) = \frac{\sin \left((k + \frac{1}{2}) (x_i - x_j) \right)}{2 \sin \left(\frac{x_i - x_j}{2} \right)} \quad (۴)$	$K(x_i, x_j) = \frac{\cos \left(\frac{x_i - x_j}{2} \right)}{2 \sin \left((k + \frac{1}{2}) (x_i - x_j) \right)} \quad (۳)$

۳۳- یک تابع PDF به صورت زیر داده شده است:

$$P(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{2}{\omega} - \frac{2x}{\omega^2} & \text{if } 0 \leq x \leq \omega \\ 0 & \text{if } x > \omega \end{cases}$$

E[X] کدام است؟

$$\int_0^{\omega} x \left(\frac{2}{\omega} - \frac{2x}{\omega^2} \right) dx \quad (۲)$$

$$\int_{-\infty}^{\infty} \omega \left(\frac{2}{\omega} - \frac{2x}{\omega^2} \right) dx \quad (۱)$$

$$\int_0^x x \left(\frac{2}{\omega} - \frac{2x}{\omega^2} \right) dx \quad (۴)$$

$$\int_0^{\omega} \left(\frac{2}{\omega} - \frac{2x}{\omega^2} \right) dx \quad (۳)$$

۳۴- اگر S ماتریس کواریانس یک مجموعه داده و \bar{x} میانگین آن باشد و داشته باشیم $S = QDQ^T$ (ستون Q بردارهای ویژه و قطر D مقادیر ویژه S هستند). در کدام حالت داده‌ها ناهمبسته (uncorrelated) خواهند بود؟

(۱) $(\bar{x} - \bar{\mu})^T QD^{-\frac{1}{2}}$ (۲) $(\bar{x} - \bar{\mu}) S^{-1}$

(۳) $(\bar{x} - \bar{\mu})^T S^{-1}$ (۴) $\bar{x}^T Q$

۳۵- کدام مورد نتیجه میان‌گیری خروجی تعدادی درخت تصمیم است؟

- (۱) کاهش بایاس (۲) افزایش بایاس
(۳) کاهش واریانس (۴) افزایش واریانس

۳۶- در هنگام انتخاب یک زیرمجموعه از ویژگی‌ها، تلاش می‌کنیم که ویژگی‌های نامناسب را تشخیص دهیم و آنها را حذف کنیم، کدام مورد، دلیل حذف ویژگی‌ها نیست؟

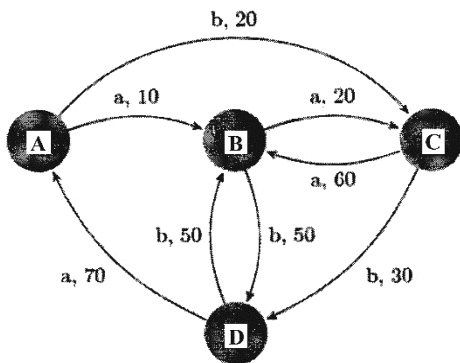
- (۱) کاهش بایاس مدل (۲) کاهش واریانس مدل

- (۳) افزایش تفسیرپذیری مدل (۴) افزایش سرعت پیش‌بینی روی نمونه‌های آزمون

۳۷- یک دسته‌بند **Hard margin linear SVM** که بر روی یک مجموعه داده n تایی آموزش دیده است با $k = 2$ بردار پشتیبان (Support vector) را در نظر بگیرید. اگر یک داده برچسب‌دار جدید به مجموعه داده موجود اضافه کنیم، حداکثر تعداد Support vectorها چند خواهد بود؟

- (۱) k (۲) n
(۳) $n + 1$ (۴) $k + 1$

۳۸- قصد داریم تا به روش یادگیری تقویتی، سیستمی با چهار حالت (A,B,C,D) را که در هر حالت تنها دو کنش a و b امکان‌پذیر است، آموزش دهیم. شکل زیر امکان انتقال بین حالات و همچنین پاداش دریافتی به ازای هر کنش در هر حالت را نشان می‌دهد. در صورتی که از الگوریتم **Q-Learning** برای آموزش استفاده نماییم و $\gamma = 0.9$ باشد، مقادیر تابع Q چه خواهد بود در صورتی که از حالت A شروع کرده و کنش‌های $\{a, a, b, a, b, a\}$ را به ترتیب از چپ به راست انجام دهیم؟



$Q(s,a)$	A	B	C	D
a	۱۰	۲۰	۷۸	۷۹
b	۴۷	۰	۳۰	۰

(۲)

$Q(s,a)$	A	B	C	D
a	۱۰	۲۰	۳۰	۷۰
b	۲۰	۰	۶۰	۰

(۱)

$Q(s,a)$	A	B	C	D
a	۷۸	۱۰	۴۷	۳۰
b	۷۹	۰	۲۰	۰

(۴)

$Q(s,a)$	A	B	C	D
a	۱۰	۲۰	۳۰	۷۰
b	۲۰	۵۰	۶۰	۵۰

(۳)

۳۹- اگر در دسته‌بندی مسئله‌ای با ویژگی‌های عددی به کمک درخت تصمیم، یک ویژگی خاص مانند x_1 را نرمال کنیم، تأثیر آن چگونه است؟

(۱) ویژگی x_1 در ارتفاع بالاتری از درخت قرار می‌گیرد.

(۲) ویژگی x_1 در ارتفاع پایین‌تری از درخت قرار می‌گیرد.

(۳) هیچ تأثیری بر شکل درخت و نتیجه دسته‌بندی ندارد.

(۴) با نرمال‌سازی ویژگی x_1 نمی‌توان در مورد شکل درخت و نتیجه دسته‌بندی نظر قطعی داد.

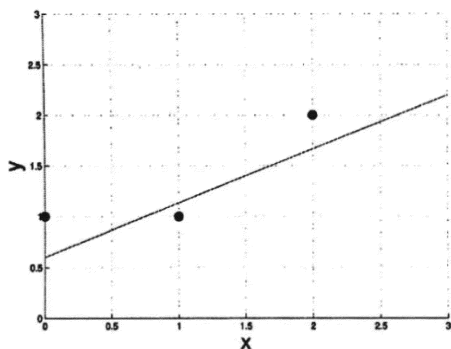
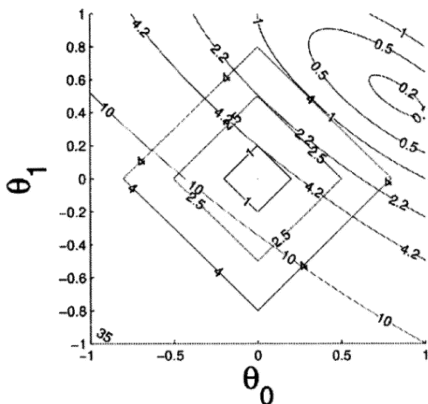
۴۰- در رگرسیون خطی سه داده $(0, 1)$ ، $(1, 1)$ و $(2, 2)$ در فضای دو بعدی، L_1 -regularization با مقدار

$\lambda = 5$ به صورت زیر استفاده شده است. در صورتی که خطای رگرسیون یا همان $\sum_{i=1}^3 (y_i - \theta_1 x_i - \theta_0)^2$

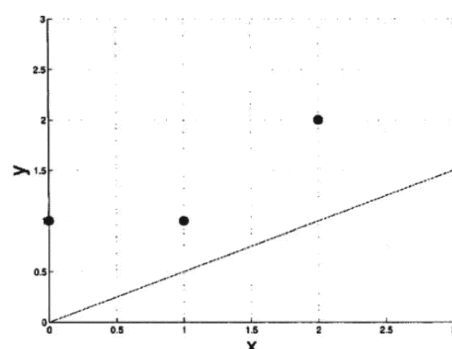
به صورت اعداد درج شده بر روی کانتورهای بیضی‌شکل و جریمه ناشی از Regularization یا همان $\lambda(|\theta_1| + |\theta_0|)$ به صورت اعداد درج شده بر روی کانتورهای مربع‌شکل رسم شده در شکل زیر باشد:

$$\min_{\theta_1, \theta_0} \sum_{i=1}^3 (y_i - \theta_1 x_i - \theta_0)^2 + \lambda(|\theta_1| + |\theta_0|) \text{ where } \lambda = 5$$

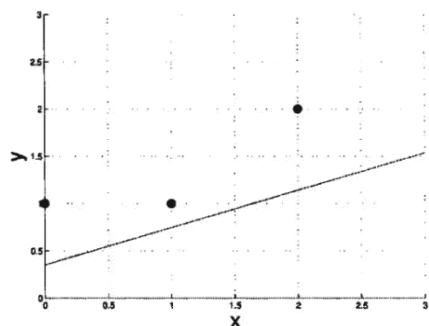
کدام یک از شکل‌های زیر نشان‌دهنده نتیجه رگرسیون است؟



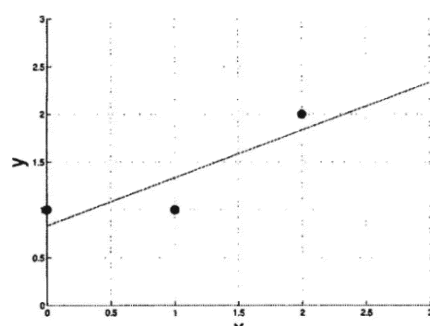
(۲)



(۱)



(۴)

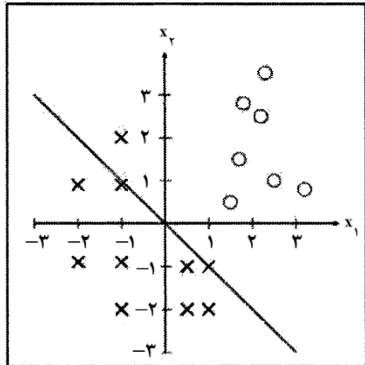


(۳)

۴۱- اگر بخواهیم مسئله دسته‌بندی دو دسته‌ای شکل زیر را به کمک مدل Regularized logistic regression بیان شده زیر حل کنیم، مرز تصمیم رسم شده در شکل برای کدام حالت می‌تواند درست باشد؟

$$\text{Max} \sum_{i=1}^n \log(P(y^{(i)} | x_1^{(i)}, x_2^{(i)})) - (\lambda_0 w_0^2 + \lambda_1 w_1^2 + \lambda_2 w_2^2)$$

Where $P(y | x_1, x_2) = \frac{1}{1 + e^{-(w_0 + w_1 x_1 + w_2 x_2)}}$



(۱) $\lambda_0 = 0, \lambda_1 = 0, \lambda_2 = 0$

(۲) $\lambda_0 = \infty, \lambda_1 = 0, \lambda_2 = 0$

(۳) $\lambda_0 = 0, \lambda_1 = \infty, \lambda_2 = 0$

(۴) $\lambda_0 = 0, \lambda_1 = 0, \lambda_2 = \infty$

۴۲- کدام یک از دسته‌بندی‌های زیر نسبت به چرخش داده‌ها (Rotation) بسیار حساس است؟ (فرض کنید همه ویژگی‌ها عددی هستند.)

(۱) SVM رگرسیون لجستیک

(۲) K نزدیک‌ترین همسایه

۴۳- اگر بخواهیم روش رگرسیون لجستیک را برای یک مسئله K دسته‌ای بسط دهیم و اینگونه عمل کنیم که دسته K را به عنوان Pivot انتخاب نماییم و برای هر دسته $i < K$ به صورت جداگانه رگرسیون لجستیک را نسبت به دسته Pivot انجام دهیم و β_i را به صورت زیر محاسبه نماییم، آنگاه کدام رابطه برای $\Pr(Y_i = K)$ درست است؟

$$\ln \frac{\Pr(Y_i = 1)}{\Pr(Y_i = K)} = \beta_1 \cdot X_i$$

$$\ln \frac{\Pr(Y_i = 2)}{\Pr(Y_i = K)} = \beta_2 \cdot X_i$$

.....

$$\ln \frac{\Pr(Y_i = K-1)}{\Pr(Y_i = K)} = \beta_{K-1} \cdot X_i$$

$$\Pr(Y_i = K) = \frac{e^{\beta_K \cdot X}}{1 + \sum_{K=1}^K e^{\beta_K \cdot X}} \quad (۲)$$

$$\Pr(Y_i = K) = \frac{1}{1 + \sum_{K=1}^{K-1} e^{\beta_K \cdot X}} \quad (۱)$$

$$\Pr(Y_i = K) = 1 - \frac{1}{1 + \sum_{K=1}^{K-1} e^{\beta_K \cdot X}} \quad (۴)$$

$$\Pr(Y_i = K) = 1 - \frac{1}{\sum_{K=1}^{K-1} e^{\beta_K \cdot X}} \quad (۳)$$

۴۴- در یک دسته‌بند که همیشه با احتمال $P = 0.7$ ، برچسب مثبت به داده‌ها انتساب می‌دهد، مقدار AUC (Area under ROC curve) کدام است؟

(۱) ۰.۷

(۲) ۰.۵

(۳) ۰.۳

(۴) ۰.۱۵

۴۵- در محاسبه حد بالای Sample complexity به کمک روش‌های مبتنی بر یادگیری PAC، در صورتی که Learner هیچ فرضی در مورد اینکه target concept به کمک فضای فرضیه H قابل نمایش است نداشته باشد و فقط فرضیه با کمترین خطای آموزش را برگرداند، از کدام رابطه می‌توان بهره برد؟

$$m \geq \frac{1}{\epsilon} (\ln |H| + \ln(\frac{1}{\delta})) \quad (۱)$$

$$m \geq \frac{1}{\epsilon} (\ln |H| + \ln(\frac{1}{\delta})) \quad (۲)$$

$$m \geq \max(\frac{1}{\epsilon} (\log_2(\frac{1}{\delta})), \frac{VC(H) - 1}{32\epsilon}) \quad (۳)$$

(۴) با توجه به اینکه Learner نمی‌تواند یک فرضیه سازگار (Consistent) بیاید امکان محاسبه حد بالای Sample complexity وجود ندارد.