293F

293

F

@RayanAzmoon

نام:

نام خانوادگی:

محل امضا:

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

صبح جمعه
۹۳/۱۲/۱۵
دفترچه شماره ۱ از ۲

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

آزمـــون ورودی
دوره‌های دکتـــری (نیمه متمرکز) داخل ــ ســـال ۱۳۹۴

مهندسی کامپیوتر ــ هوش مصنوعی
(کد ۲۳۵۶)

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 و الگوریتم - شناسایی آماری الگو، یادگیری ماشین)	۴۵	۱	۴۵

این آزمون نمرهٔ منفی دارد.

استفاده از ماشین حساب مجاز نیست.

اسفند ماه ــ سال ۱۳۹۳

حق چاپ، تکثیر و انتشار سوالات به هر روش (الکترونیکی و ...) 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) یک گراف جهت‌دار و بدون وزن $G = (V, E)$ با دو رأس s و t داده شده است که درجه‌ی ورودی s و درجه‌ی خروجی t برابر صفراند. می‌خواهیم بیشینه‌ی تعداد مسیرهای یال‌مجزا از s به t را به‌دست آوریم. با استفاده از کدام‌یک از الگوریتم‌های زیر می‌توان این عدد را به‌دست آورد؟

۱) بلمن-فورد ۲) دایکسترا
۳) هافمن ۴) شبکه‌ی شار

۲) با اضافه کردن یک یال جهت‌دار جدید در یک گراف جهت‌دار، چند تا از موارد زیر در مورد تعداد اجزای هم‌بندقوی گراف، ممکن است درست باشد؟

- حداکثر یک واحد کم می‌شود.
- حداکثر یک واحد زیاد می‌شود.
- شاید تغییر نکند.
- شاید بیش از یک واحد کم شود.

۱) ۰ ۲) ۱ ۳) ۲ ۴) ۳

۳) چندتا از مسئله‌های زیر ان‌پی هستند؟

- نسخه تصمیم‌گیری فروشنده‌ی دوره‌گرد
- مرتب بودن یک آرایه
- یافتن بیشینه‌ی شبکه‌ی شار
- نسخه تصمیم‌گیری کوله‌پشتی صفر و یک

۱) ۱ ۲) ۲ ۳) ۳ ۴) ۴

۴) چندتا از گزاره‌های زیر در مورد طول کوتاه‌ترین مسیر بین دو رأس (که با $\delta(u, v)$ نمایش داده می‌شود) در یک گراف جهت‌دار و وزن‌دار G درست است؟ وزن یال‌ها ممکن است منفی باشد.

۱. با فرض عدم وجود دور منفی همیشه داریم: $\delta(u, t) \leq \delta(u, v) + \delta(v, t)$

۲. اگر دور منفی داشته باشیم، آن‌گاه برای هر دو راس u و v، عدد $\delta(u, v)$ برابر $-\infty$ خواهد بود.

۳. اگر یال با وزن منفی داشته باشیم اما دور منفی نداشته باشیم، $\sum \delta(u, v)$ که مجموع روی همه‌ی زوج راس‌ها گرفته شده نمی‌تواند منفی باشد.

۱) ۰ ۲) ۱ ۳) ۲ ۴) ۳

۵) الگوریتم بلمن-فورد در هر مرحله همه‌ی یال‌ها را بررسی می‌کند و به ازای هر یال (u, v) اگر

$$d(v) > d(u) + w(u, v)$$

آن‌گاه $d(v)$ به‌روز می‌شود که $w(u, v)$ وزن یال (u, v) و $d(u)$ طول بهترین مسیر یافت شده تاکنون برای راس u است. اگر در مرحله‌ای برای هیچ راسی به‌روزرسانی انجام نشود الگوریتم بلمن-فورد خاتمه می‌یابد. با فرض آنکه الگوریتم بلمن-فورد برای یافتن همه‌ی کوتاه‌ترین مسیرها از رأس s در گراف G با n راس پس از $k < n$ بار تکرار به پایان برسد، چندتا از گزاره‌های زیر درست‌اند؟

۱. تعداد یال‌های همه‌ی کوتاه‌ترین مسیرها از s حداکثر $k - ۱$ است.

۲. وزن همه‌ی کوتاه‌ترین مسیرها از s حداکثر $k - ۱$ است.

۳. گراف دور منفی ندارد.

۱) ۰ ۲) ۱ ۳) ۲ ۴) ۳

۶) فرض کنید درخت AVL به این صورت تعریف شود که به‌ازای هر گره v از درخت، اختلاف ارتفاع زیردرخت چپ و زیردرخت راست v حداکثر ۲ باشد. با این تعریف، اگر $F(h)$ کمینه‌ی تعداد عناصر یک درخت AVL با ارتفاع h باشد، کدام‌یک از رابطه‌های بازگشتی درست است؟ فرض کنید $F(\mathcal{O}(۱)) = \mathcal{O}(۱)$.

۱) $F(h) \leq F(h-۱) + F(h-۲) + ۱$
۲) $F(h) \leq F(h-۱) + F(h-۳)$
۳) $F(h) \leq F(h-۱) + F(h-۲)$
۴) $F(h) \leq F(h-۱) + F(h-۳) + ۱$

۷) یک درخت دودویی با ۶ گره داده شده است که هر گره فقط فرزند چپ دارد. با چند عمل «دوران راست» (بدون دوران چپ) می‌توان این درخت را به درختی تبدیل کرد که هر گره فقط فرزند راست داشته باشد. کم‌ترین مقدار ممکن را انتخاب کنید.

۱) ۴ ۲) ۵ ۳) ۶ ۴) ۷

۸) گراف بدون جهت G با تعداد n رأس و m یال و دو رأس s و t داده شده است. می‌خواهیم بدانیم که آیا مسیری از s به t در G وجود دارد یا خیر. می‌دانیم که اگر G با لیست مجاورت پیاده‌سازی شود با استفاده از DFS یا BFS می‌توان این مسئله را در زمان $O(n+m)$ حل کرد. اگر به جای لیست مجاورت، G با ماتریس مجاورت پیاده‌سازی شده باشد، این مسئله را در چه زمانی می‌توان حل کرد؟

۱) $\mathcal{O}(n^۲)$ ۲) $\mathcal{O}(nm)$ ۳) $\mathcal{O}(n+m)$ ۴) $\mathcal{O}(m + n\log n)$

۹) در گراف بدون جهت G به تعریف‌های زیر توجه کنید:

- قطر: بیشینه‌ی کوتاه‌ترین مسیر بین دو رأس در گراف
- $l(s)$: بیشینه‌ی طول کوتاه‌ترین مسیرها از s به دیگر رأس‌ها.
- شعاع گراف: کمینه‌ی مقدار $l(s)$ بین همه‌ی رأس‌های گراف

اگر d و r به‌ترتیب قطر و شعاع یک گراف باشند، کدام‌یک از رابطه‌های زیر همیشه برقرار است؟ بهترین گزینه را انتخاب کنید.

۱) $r \leq d/۲$ ۲) $r \leq d$ ۳) $r \geq d/۲$ ۴) $r \geq d$

۱۰) الگوریتم مرتب‌سازی ادغامی بر روی آرایه‌ی n عضوی را اینطور تغییر می‌دهیم که آرایه را به سه قسمت مساوی تقسیم می‌کنیم. هر قسمت را به‌همین روش و به‌صورت بازگشتی مرتب می‌کنیم و سپس سه بخش مرتب‌شده را در هم ادغام می‌کنیم تا آرایه‌ی مرتب به‌دست آید. زمان اجرای این الگوریتم چیست؟

۱) $\mathcal{O}(n^۲)$ ۲) $\mathcal{O}(n\log n)$ ۳) $\mathcal{O}(n\log^۲ n)$ ۴) $\mathcal{O}(n^{\log_۲ ۳})$

۱۱) تعداد برش‌های کمینه یک درخت بدون وزن با n گره و $n-۱$ یال چند تاست؟

۱) $n-۱$ ۲) n ۳) $\binom{n}{۲}$ ۴) $۲^n - ۲$

۱۲) حل رابطه‌ی بازگشتی $T(n) = ۲^{\log_۲ ۳}T(n/۲) + n\sqrt{n}$ کدام است؟

۱) $\mathcal{O}(n\sqrt{n})$ ۲) $\mathcal{O}(n\sqrt{n}\log n)$ ۳) $\mathcal{O}(n^{\log_۲ ۳})$ ۴) $\mathcal{O}(n^{\log_۲ ۳}\log n)$

۱۳) در مسئله‌ی ضرب ماتریس‌ها فرض کنید که می‌خواهیم $A_۱ \times A_۲ \times \cdots \times A_n$ را محاسبه کنیم که ماتریس A_i به ابعاد $d_{i-۱} \times d_i$ است. می‌خواهیم ترتیبی از ضرب ماتریس‌ها را به‌دست آوریم که تعداد عناصر ماتریس‌های میانی (که در مراحل مختلف ضرب به‌دست می‌آید) کمینه شود. اگر این تعداد کمینه را هزینه‌ی ضرب بنامیم، یعنی C_{ij} هزینه ضرب زیرمسئله‌ی $A_i \times A_۲ \times \cdots \times A_j$ باشد، داریم $C_{i,i} = ۰$. کدام‌یک از رابطه‌های زیر برای این مسئله درست است؟

۱) $C_{ij} = \min_{i \leq k < j} \max\{C_{ik}, C_{k+۱,j} + d_{i-۱}d_j\}$
۲) $C_{ij} = \min_{i \leq k < j} \max\{C_{ik}, C_{k+۱,j}, d_{i-۱}d_j\}$
۳) $C_{ij} = \min_{i \leq k < j} \max\{C_{ik} + C_{k+۱,j} + d_{i-۱}d_j\}$
۴) $C_{ij} = \min_{i \leq k < j} \max\{C_{ik}, C_{k+۱,j}\} + d_{i-۱}d_j$

۱۴) n نفر به‌ترتیب تصادفی وارد یک اتاق می‌شوند. می‌خواهیم با دیدن فرد iام بلندقدترین فرد بین i نفر دیده‌شده را به دست آوریم. ما تنها یک حافظه برای ذخیره‌سازی اندیس بلندقدترین فرد داریم که هر بار با دیدن فرد بلندتر روزآمد می‌شود. می‌خواهیم میانگین تعداد روزآمد کردن این حافظه را به‌دست آوریم. این میانگین به کدام گزینه نزدیک‌تر است؟

۱) $\ln n$ ۲) n ۳) $n/۲$ ۴) $\log^۲ n$

۱۵) بر روی یک هرم کمینه تهی n عمل درج و حذف (با فرض داشتن محل حذف در هرم کمینه) با ترتیب دل‌خواه انجام می‌دهیم. هزینه‌ی سرشکنی هر یک از این دو عمل چه‌قدر است؟

۱) درج: $\mathcal{O}(\log n)$ و حذف: $\mathcal{O}(\log n)$
۲) درج: $\mathcal{O}(۱)$ و حذف: $\mathcal{O}(\log n)$
۳) درج: $\mathcal{O}(\log n)$ و حذف: $\mathcal{O}(۱)$
۴) درج: $\mathcal{O}(۱)$ و حذف: $\mathcal{O}(۱)$

۱۶) الگوریتمی را در نظر بگیرید که یک آرایه نامرتب با $۳n$ عنصر با مقادیر مجزا را بگیرد و دو مقدار $x < y$ را محاسبه کند به‌طوری که n عنصر هر کدام کم‌تر از x، n عنصر بیش‌تر از x، و n عنصر هر کدام بین x و y باشند.
چندتا از گزاره‌های زیر درست‌اند؟

- هر الگوریتمی که این کار را انجام دهد از $\Omega(n \lg n)$ است.
- الگوریتمی از مرتبه‌ی میانگین $O(n)$ برای این مسئله وجود دارد.
- الگوریتمی از مرتبه‌ی $O(n)$ برای این مسئله وجود دارد.
- با فرض داشتن $O(۱)$ حافظه اضافی علاوه بر آرایه ورودی، نمی‌توان الگوریتمی با زمان چندجمله‌ای برای این مسئله ارائه کرد.

۱) ۱ ۲) ۲ ۳) ۳ ۴) ۴

۱۷) فرض کنید در متنی تعداد تکرار iامین حرف انگلیسی برابر ۲^i است. ارتفاع درخت هافمن برای این حرف‌ها در الگوریتم فشرده‌سازی هافمن کدام است؟

۱) ۵ ۲) ۶ ۳) ۱۲ ۴) ۲۵

۱۸) در یک گراف بدون‌جهت، همبند و وزن‌دار (با وزن غیرمنفی و متمایز) G می‌دانیم کوتاهترین مسیر بین هر دو راس بر روی درخت پوشای کمینه قرار دارد. چندتا از گزاره‌های زیر درست‌اند؟

- گراف G یک درخت است.
- وزن هر یال $\{u, v\}$ حداقل به اندازه سنگین‌ترین یال در کوتاه‌ترین مسیر از u به v است.
- کوتاه‌ترین مسیر بین هر دو راس u و v یکتاست.
- با فرض شروع از راس s هر دو الگوریتم پریم (برای محاسبه درخت کمینه) و دایکسترا (برای محاسبه کوتاه‌ترین مسیرها) با ترتیب یکسان راس‌ها را پردازش و به درخت اضافه می‌کنند.

۱) ۰ ۲) ۱ ۳) ۲ ۴) ۳

۱۹) نسخه‌ی تصمیم‌گیری چندتا از مسائل زیر برای یک گراف همبند بدون‌جهت و وزن‌دار G اِن‌پی-تمام است؟

- پیدا کردن تطابق بیشینه
- پیدا کردن دور همیلتنی بیشینه
- پیدا کردن دور اویلری بیشینه
- پیدا کردن برش بیشینه

۱) ۱ ۲) ۲ ۳) ۳ ۴) ۴

۲۰) فرض کنید ۱۳۹۳ نقطه روی محور اعداد حقیقی داده شده است. می‌خواهیم این نقاط را رنگ‌آمیزی کنیم به گونه‌ای که به ازای هر بازه‌ی دل‌خواه $[a, b]$ اگر حداقل یک نقطه از ۱۳۹۳ نقطه‌ی داده‌شده داخل بازه بود، رنگ حداقل یکی از نقاط قرارگرفته داخل این بازه با دیگر نقاط قرارگرفته داخل بازه فرق کند. با چند رنگ این کار امکان‌پذیر است؟ بهترین گزینه را انتخاب کنید.

۱) ۱۱ ۲) ۱۳۹۲ ۳) ۱۳۹۳ ۴) ۶۷

۲۱- داده‌های دو کلاس C_1 و C_2 دارای یک ویژگی x و به ترتیب دارای توزیع های$N(0,0.5)$ و $N(1,0.5)$ نمونه برداری شده اند و احتمال پیشین آنها برابر $P(C_1)=P(C_2)=0.5$ است. برای ماتریس هزینه زیر کدام گزینه برای مقدار آستانه x_0 برای دسته بند کمترین ریسک (هزینه) درست است؟

$$L = \begin{bmatrix} 0 & 0.5 \\ 1 & 0 \end{bmatrix}$$

۱) $x_0= 0.5$

۲) $x_0<0.5$

۳) $x_0=0.55$

۴) $x_0=0.65$

۲۲- اگر $K_1(x,z)$ و $K_2(x,z)$ دو تابع هسته در فضای $R^n \times R^n$ باشند، در چه صورت $aK_1(x,z)+bK_2(x,z)$ = $K(x,z)$یک تابع هسته معتبر نمی‌باشد؟

۱) $a>0$ و $b>0$

۲) $a<0$ و $b>0$

۳) $a>0$ و $b<0$

۴) $a<0$ و $b<0$

۲۳- در یک مجموعه داده با**N** داده، نمونه **i**ام (x_i) دارای یک عدد حقیقی است. اگر نمونه x_i از توزیع زیر نمونه برداری شده باشد (a یک پارامتر است) کدام گزینه زیر تخمین **Maximum likelihood** برای پارامتر a است؟

$$p(x) = \begin{cases} ae^{-ax} & if\ x \geq 0 \\ 0 & otherwise \end{cases}$$

۱) $\frac{\sum_{i=1}^{N} x_i}{N}$

۲) $\frac{N}{\sum_{i=1}^{N} x_i}$

۳) $\frac{\sum_{i=1}^{N} \log(x_i)}{N}$

۴) $\frac{N}{\sum_{i=1}^{N} \log(x_i)}$

۲۴- در کاربردی هر نمونه دارای ۱۰۰۰ ویژگی است که ۵۰ تا از این ویژگی ها واقعا دارای اطلاعات جهت دسته بندی و ۵۰ ویژگی دیگر کپی ۵۰ ویژگی نخست هستند. ۹۰۰ ویژگی دیگر دارای اطلاعات مفیدی برای دسته بندی نیستند. فرض کنید که به اندازه کافی داده در اختیار است و مقادیر پارامتر برای روش های مختلف درست تنظیم شده اند. کدام گزینه زیر صحیح است؟

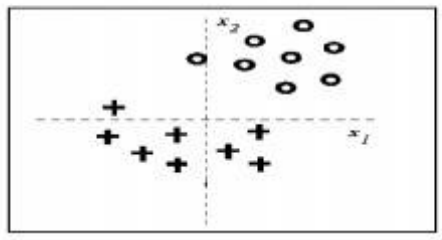
۱) روش پالاینده اطلاعات متقابل (mutual information filter) تقریبا ۵۰ ویژگی و روش لفافه (wrapper) تقریبا ۵۰ ویژگی انتخاب می کنند.

۲) روش پالاینده اطلاعات متقابل (mutual information filter) تقریبا ۵۰ ویژگی و روش لفافه (wrapper) تقریبا ۱۰۰ ویژگی انتخاب می کنند.

۳) روش پالاینده اطلاعات متقابل (mutual information filter) تقریبا ۱۰۰ ویژگی و روش لفافه (wrapper) تقریبا ۵۰ ویژگی انتخاب می کنند.

۴) روش پالاینده اطلاعات متقابل (mutual information filter) تقریبا ۱۰۰ ویژگی و روش لفافه (wrapper) تقریبا ۱۰۰ ویژگی انتخاب می کنند.

۲۵- برای مجموعه داده زیر از دسته بند **logistic regression** استفاده شده است.



تابع تصمیم به صورت زیر است.

$$g(x) = p(t = 1|w, x) = \frac{1}{1 + exp\{-w_0 - w_1x_1 - w_2x_2\}}$$

اگر از منظم ساز L1 استفاده کنیم و تابع هدف به صورت $[\sum_{i=1}^{N} \log(g(x_i))] - C(|w_1| + |w_2|)$ **باشد و پارامتر C به تدریج افزایش یابد کدام گزینه زیر درست است؟**

۱) مقدار w_1 و w_2 همزمان به سمت صفر میل می کنند.

۲) نخست w_2 صفر می شود و سپس w_1 به سمت صفر میل می کند.

۳) نخست w_1 صفر می شود و سپس w_2 به سمت صفر میل می کند.

۴) هیچ کدام از مقدارهای w_1 و w_2 صفر نمی شوند و همزمان با افزایش C مقدار آنها کوچکتر می شود.

۲۶- **در یک مسئله دسته بندی دو کلاسه با احتمال پیشین برابر کلاس ها** $(p(C_1) = p(C_2) = 0.5)$**، اگر توزیع نمونه ها در کلاس ها به صورت زیر باشد**

$$p(x|C_1) = N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}a & c\\c & b\end{bmatrix}\right)$$
$$p(x|C_2) = N\left(\begin{bmatrix}d\\e\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right)$$

که $a \times b - c \times c = 1$**. اگر بخواهیم سطح جداساز خطی باشد مقادیر مناسب برای** a, b, c, d, e **چقدر است؟**

۱) a = ۱, b = ۱, c = ۰, d = ۵, e = ۵

۲) a = ۱, b = ۱, c = ۰, d = ۰, e = ۰

۳) a = ۱, b = ۱, c = ۱, d = ۵, e = ۵

۴) a = ۱, b = ۰, c = ۰, d = ۵, e = ۵

۲۷- در یک مسئله دسته بندی دو کلاسه با احتمال پیشین برابر کلاس ها ($p(C_1) = p(C_2) = 0.5$)، اگر توزیع نمونه ها در کلاس ها به صورت زیر باشد مقدار وزن ها در معیار فیشر چقدر است؟

$$\mathbf{p(x \mid c_1) = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\right)}$$

$$\mathbf{p(x \mid c_2) = N\left(\begin{bmatrix} 4 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)}$$

۱) $w = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$

۲) $w = \begin{bmatrix} -\frac{2}{3} \\ -\frac{2}{3} \end{bmatrix}$

۳) $w = \begin{bmatrix} -\frac{4}{3} \\ -\frac{2}{3} \end{bmatrix}$

۴) $w = \begin{bmatrix} -\frac{4}{3} \\ -\frac{2}{9} \end{bmatrix}$

۲۸- کدامیک از تطبیق های الگوریتم های مدل مخلوط گوسی (Gaussian Mixture model) آن را شبیه تر به الگوریتم خوشه بندی k-means می نماید؟

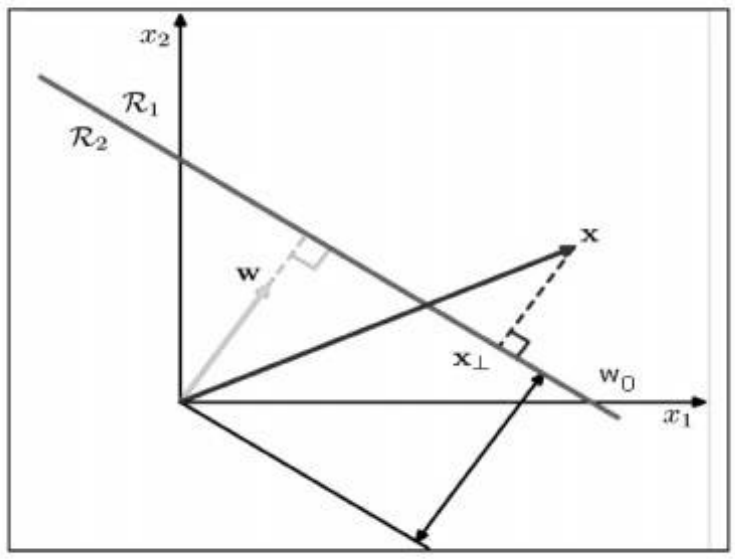
۱) محدود سازی ماتریس کوواریانس (Σ_i) به یک ماتریس قطری.

۲) ماتریس کوواریانس (Σ_i) هر ماتریسی می تواند باشد.

۳) محدود سازی ماتریس کوواریانس (Σ_i) به یک ماتریس قطری به فرم rI که I ماتریس همانی است.

۴) محدود سازی ماتریس کوواریانس (Σ_i) به یک ماتریس قطری به فرم r_iI که I ماتریس همانی است.

۲۹- در یک دسته بند خطی $y=w_0+w_1x_1+w_2x_2$ (که در شکل زیر نشان داده شده است)، فاصله مبدا مختصات تا خط جداساز برابر است با:



۱) $\frac{|w_0|}{|w|^2}$

۲) $\frac{|w_0|}{|w|}$

۳) $\frac{w_0}{|w|}$

۴) $|w_0|$

۳۰- کدام گزینه زیر درست نیست؟

۱) روش تک پیوندی (single link) برای اندازه گیری فاصله بین دو خوشه عبارت است از فاصله بین دورترین نقاط.

۲) یکی از مهمترین ایراد های خوشه بندی سلسله مراتبی تجمیعی این است که اگر اشتباهی صورت پذیرفت قابل تصحیح نمی باشد.

۳) k-means خوشه های با اشکال محدب تولید می کند.

۴) k-means به نویز و داده های پرت حساس است.

۳۱- یک مجموعه داده با N داده که نمونه iام دارای یک عدد حقیقی x_i و مقدار t_i است. اگر مقدار خروجی دارای توزیع $t_i = N(\log(wx_i), 1)$ باشد که w متغیری است که می بایست یاد گرفته شود، در صورتی‌که از تخمین Maximum likelihood برای تخمین w استفاده شود کدام گزینه زیر صحیح است؟

۱) $\sum_i t_i = \sum_i \log(wx_i)$

۲) $\sum_i x_i t_i = \sum_i x_i \log(wx_i)$

۳) $\sum_i x_i \log(wx_i) = \sum_i t_i \log(wx_i)$

۴) $\sum_i x_i \log(wx_i) = \sum_i x_i t_i \log(wx_i)$

۳۲- یک دسته بند SVM را در نظر بگیرید که دوتابع نگاشت $\Phi_1 = [x \quad x^2]$ و $\Phi_2 = [2x \quad 2x^2]$ برای آن در نظر گرفته شده است. کدام گزینه زیر درست است؟

۱) نمی توان در باره حاشیه بدست آمده توسط دو نگاشت قضاوت نمود.

۲) حاشیه بدست آمده توسط نگاشت Φ_2 مساوی حاشیه بدست آمده توسط نگاشت Φ_1 است.

۳) حاشیه بدست آمده توسط نگاشت Φ_2 بزرگتر از حاشیه بدست آمده توسط نگاشت Φ_1 است.

۴) حاشیه بدست آمده توسط نگاشت Φ_2 کوچکتر از حاشیه بدست آمده توسط نگاشت Φ_1 است.

۳۳- بعد VC دسته بند نزدکترین همسایه کدام است؟

۱) ۱

۲) ۲

۳) ۳

۴) ∞

۳۴- کدامیک از گزینه های زیر بیش برازش (overfitting) را کاهش می دهد؟

۱) هرس درخت تصمیم

۲) کاهش k از مقدار ۵ به یک در دسته بند k-NN

۳) افزایش ضریب c در SVM با حاشیه نرم با تابع هدف $\min_w |w|^2 + c \sum_i \zeta_i$

۴) کاهش ضریب منظم ساز (λ) در رگرسیون خطی منظم شده $\frac{1}{2}\sum_i (t_i - w^T x_i) + \lambda |w|^2$

۳۵- بعد VC فضای فرضیه های شامل همه مثلث ها کدام است؟

۱) ۵ ۲) ۶ ۳) ۷ ۴) ۸

۳۶- در یک دسته بند SVM از تابع هسته RBF استفاده شده است. سه نقطه z_1, z_2, x داریم که از نظر هندسی x به z_1 خیلی نزدیک و از z_2 خیلی دور است. کدام گزینه در باره مقدار $k(z_1,x)$ و $k(z_2,x)$ درست است؟

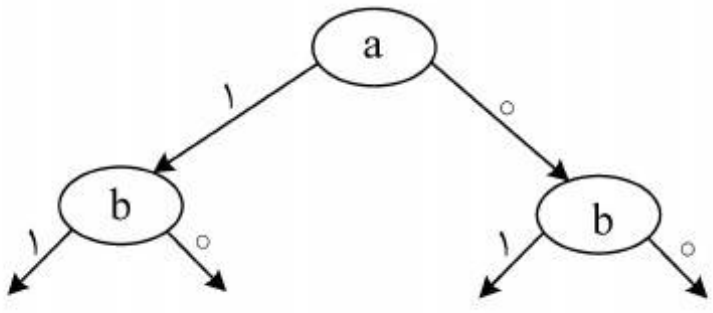
۱) $k(z_2,x)$ تقریبا مساوی یک و $k(z_1,x)$ تقریبا مساوی صفر است

۲) $k(z_1,x)$ تقریبا مساوی یک و $k(z_2,x)$ تقریبا مساوی صفر است

۳) $k(z_1,x)$ تقریبا مساوی c_1 $(c_1 \gg 1)$ و $k(z_2,x)$ تقریبا مساوی c_2 $(c_2 \ll 0)$ است

۴) $k(z_1,x)$ تقریبا مساوی c_1 $(c_1 \ll 0)$ و $k(z_2,x)$ تقریبا مساوی c_2 $(c_2 \gg 0)$ است

۳۷- درخت تصمیم با عمق ۲ زیر برای دو ویژگی دودویی ساخته شده است. فضای فرضیه ای که این درخت می تواند نمایش دهد دارای چند ویژگی است؟



۱) ۴ ۲) ۸ ۳) ۱۶ ۴) ۳۲

۳۸-کدام یک از گزینه های زیر در دو دسته بند درخت تصمیم و بیز ساده درست است؟

۱) در هر دو دسته بند فرض می شود زوج ویژگی ها مستقل باشند

۲) در هر دو دسته بند فرض می شود زوج ویژگی ها وابسته باشند

۳) در هر دو دسته بند فرض می شود زوج ویژگی ها برای یک دسته مستقل باشند

۴) در هر دو دسته بند فرض می شود زوج ویژگی ها برای یک دسته وابسته باشند

۳۹- مجموعه داده آموزشی $D=\{(x_1,t_1), (x_2,t_2), \dots, (x_N,t_N)\}$ داده شده است که x_i و t_i اعداد حقیقی هستند و به اندازه کافی داده آموزشی موجود است. کدامیک از دو مدل $y_1(x)=w_1^2x+w_2x$ و $y_2(x)=wx$ روی این مجموعه داده نتایج بهتری را تولید می کنند؟

۱) مدل $y_1(x)$ روی این مجموعه داده بهتر از مدل $y_2(x)$ عمل می کند.

۲) مدل $y_2(x)$ روی این مجموعه داده بهتر از مدل $y_1(x)$ عمل می کند.

۳) نمی‌توان در مورد عملکرد این دو مدل اظهار نظر کرد.

۴) هر دو مدل یکسانند.

۴۰- مجموعه داده آموزشی $D=\{(x_1,t_1), (x_2,t_2), \dots, (x_N,t_N)\}$ داده شده است که x_i و t_i اعداد حقیقی هستند. t به صورت زیر تولید شده است

$$t = aX + \epsilon$$
$$\epsilon \sim N(0, \sigma^2)$$

که ϵ یک نویز تصادفی مستقل است. اگر σ را بدانیم و دانش پیشین ما درباره a به صورت $a \sim N(0, \lambda^2)$ باشد، کدام یک از گزینه های زیر نادرست است؟

۱) اگر σ و مقدار ثابت باشد $|a_{MAP}$ – کم می شود

۲) اگر σ و مقدار ثابت باشد $|a_{MAP}$ – زیاد می شود

۳) اگر σ و مقدار ثابت باشد $|a_{MAP}$ – کم می شود

۴) اگر σ و مقدار ثابت باشد $(a_{MAP}$ تغییر نمی کند.

۴۱- کدامیک از گزینه های زیر درباره دسته بند های K-NN درست است؟

۱) سطح تصمیم خطی است

۲) سطح تصمیم با کاهش k هموارتر می شود.

۳) دقت دسته بندی با افزایش k افزایش می یابد.

۴) این دسته بند نیاز به فاز آموزش ندارد.

۴۲- برای اینکه نشان دهیم بعد VC یک دسته بند برابر N است کدام یک از گزینه های زیر درباره یک دسته بند درست است؟

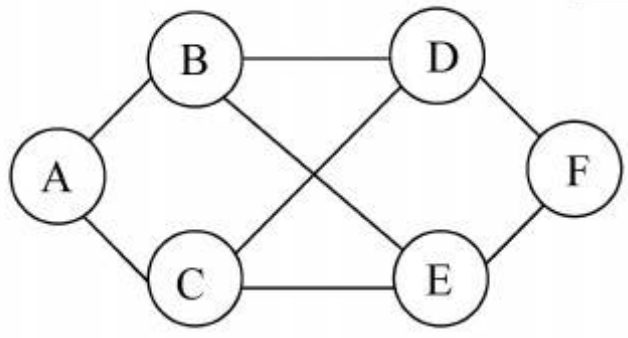
۱) نشان دهیم که یک دسته بند می تواند تمام دسته بندی های ممکن N داده را انجام دهد.

۲) نشان دهیم که یک دسته بند می تواند تمام دسته بندی های ممکن زیرمجموعه ای از N داده را انجام دهد اما نمی تواند یکی از دسته بندی های ممکن $N+1$ داده را انجام دهد

۳) نشان دهیم که یک دسته بند می تواند تمام دسته بندی های ممکن N داده را انجام دهد اما نمی تواند هیچ یک از دسته بندی های ممکن $N+1$ داده را انجام دهد.

۴) نشان دهیم که یک دسته بند می تواند تمام دسته بندی های ممکن N داده را انجام دهد اما نمی تواند یکی از دسته بندی های ممکن $N+1$ داده را انجام دهد.

۴۳- یک مکان پرپیچ و خم را در نظر بگیرید که تغییر مکان ها در شکل زیر نشان داده شده است. در این شکل هر تغییر مکان با یک گره نشان داده شده است. مکان A درب خروجی از مکان پرپیچ و خم را نشان می دهد. هر تغییر مکان جریمه ۱- را نتیجه دارد و عامل در درب خروج پاداش ۰ را دریافت می نماید. اگر ضریب کاهش (γ) برابر با ۰/۹ باشد کدام گزینه در باره ارزش مکان ها نادرست است؟



۱) V(C)= -1 ۲) V(D)=-2 ۳) V(E)=-2 ۴) V(F)=-3

۴۴- کدام گزینه نادرست است؟

۱) احتمال رویداد بیش برازش (Overfitting) برای فضای فرضیه کوچک بیشتر است

۲) احتمال رویداد بیش برازش (Overfitting) برای مجموعه داده های کوچک بیشتر است.

۳) اگر عمق یک درخت تصمیم به اندازه طول بردار ویژگی باشد ، نویز موجود در داده ها بیشتر تاثیر خواهد گذاشت.

۴) اگر نیمی از یک مجموعه داده با N نمونه، برای آموزش و نیمی برای آزمایش بکار رود، اختلاف خطای آموزش و آزمایش با افزایش N کاهش می یابد.

۴۵- کدام گزینه نادرست است؟

۱) واریانس در دقت آزمون با افزایش اندازه مجموعه آزمون کاهش می یابد.

۲) از روش نزول در امتداد گرادیان می توانیم برای یادگیری یک مدل مخلوط گوسی (Gaussian Mixture model) استفاده کنیم.

۳) هنگام ساختن یک درخت تصمیم برای داده های بدون نویز، اگربرای ریشه یک ویژگی درست انتخاب نشود هنوز هم می توانیم یک فرضیه سازگار تولید نماییم.

۴) در مسئله رگرسیون خطی، اگر از روش نزول در امتداد گرادیان برای کمینه نمودن مجموع مربع خطا استفاده شود ممکن است الگوریتم در کمینه های محلی گرفتار شود.