



کد کنترل

332

F

## آزمون (نیمه متهم گز) ورود به دوره های دکتری - سال ۱۴۰۱

دفترچه شماره (۱)

صبح جمعه ۱۴۰۰/۱۲/۶



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

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

رشته مهندسی کامپیوتر - هوش مصنوعی  
(کد ۲۳۵۶)

جدول مواد امتحانی، تعداد، شماره سؤال ها و زمان پاسخ گویی

زمان پاسخ گویی	تا شماره	از شماره	تعداد سؤال	مواد امتحانی
۱۵۰ دقیقه	۴۵	۱	۴۵	مجموعه دروس تخصصی: - ساختمان داده ها و طراحی الگوریتم ها - شناسایی الگو - یادگیری ماشین

استفاده از ماشین حساب مجاز نیست.

این آزمون نمره منفی دارد.

حق چاپ، تکثیر و انتشار سؤال ها 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 متخلفان برابر مقررات رفتار می شود.

\* متقاضی گرامی، وارد نکردن مشخصات و امضا در کادر زیر، به منزله غیبت و حضور نداشتن در جلسه آزمون است.

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

امضا:

- ۱- اگر زمان اجرای مرتب سازی خیابی  $c_1 n^2$  و زمان اجرای مرتب سازی ادغامی  $c_2 n \log_2 n$  (با فرض  $c_1 \leq c_2$ ) باشد، در کدام یک از حالات، مرتب سازی ادغامی به ازای هر  $n > 2$  سریع تر از مرتب سازی خیابی است؟

$$(1) \quad c_1 > \frac{c_2}{n}$$

$$(2) \quad c_2 < 16c_1$$

$$(3) \quad \frac{n}{2} > \left(\frac{c_2}{c_1}\right)^2$$

(4) مرتب سازی ادغامی همیشه سریع تر از مرتب سازی خیابی است.

- ۲- دو آرایه مرتب A و B از اعداد صحیح با طول های n و m (با فرض  $n \leq m$ ) و عدد صحیح  $k \geq \log_2 n$  داده شده است. با چه تعداد مقایسه می توان k آمین کوچکترین عضو را در اجتماع این دو آرایه پیدا کرد؟ (بهترین گزینه را انتخاب کنید).

$$(1) \quad O(\log(nm)) \quad (2) \quad O(k \log(n+m)) \quad (3) \quad O(n \log m) \quad (4) \quad O(\log n \times \log m)$$

- ۳- اگر بخواهیم داده ساختار صف را با استفاده از پشته پیاده سازی کنیم، طوری که عملیات های پایه ای صف در زمان سرشکن  $O(1)$  قابل انجام باشد، کدام مورد درست است؟

(1) با سه پشته می توان این کار را انجام داد و سه پشته برای این کار لازم است.

(2) با دو پشته می توان این کار را انجام داد و دو پشته برای این کار لازم است.

(3) با یک پشته می توان این کار را انجام داد.

(4) این کار امکان پذیر نیست.

- ۴- فرض کنید m آرایه مرتب داریم که در مجموع آرایه ها شامل n عدد هستند. می خواهیم از هر کدام از آرایه ها یک عدد را انتخاب کنیم، به طوری که اختلاف بیشینه و کمینه اعداد انتخاب شده کمترین مقدار ممکن شود. این کار در چه زمانی ممکن است؟ (بهترین گزینه را انتخاب کنید).

$$(1) \quad O(m + \log n) \quad (2) \quad O(n \log m) \quad (3) \quad O(n + m) \quad (4) \quad O(nm)$$

- ۵- نمایش پیش ترتیب (Preorder) و میان ترتیب (Inorder) یک درخت دودویی مرتب با n رأس داده شده است. کدام مورد درست است؟

(1) درخت متناظر با این لیست رئوس، منحصر به فرد است.

(2) دقیقاً دو درخت با این ترتیب ملاقات رئوس وجود دارد.

(3) دقیقاً  $\frac{n}{2}$  درخت با این ترتیب ملاقات رئوس وجود دارد.

(4) هیچ یک از موارد درست نیست.

- ۶- فرض کنید یک هرم بیشینه شامل حداکثر  $n$  عدد داده شده است. جستجوی یک مقدار در این هرم بیشینه در کدام کلاس پیچیدگی است؟ (بهترین گزینه را انتخاب کنید).

(۱)  $O(\frac{n}{\log(n)})$  (۲)  $O(\log(n))$  (۳)  $O(n)$  (۴)  $O(\sqrt{n})$

- ۷- کدام مورد درخصوص دو گزاره (الف) و (ب) به ترتیب درست است؟

الف - اگر  $f(n) \in O(2^{2n})$ ، آنگاه  $f(n) \in \Omega(2^n)$

ب - اگر  $T(n) = 2T(\frac{n}{2}) + O(n^2)$ ، آنگاه  $T(n) = \Theta(n^2)$

- (۱) درست - درست (۲) درست - نادرست (۳) نادرست - درست (۴) نادرست - نادرست

- ۸- هرم دودویی را ..... از روی یک درخت قرمز - سیاه در زمان  $O(n)$  ساخت و درخت قرمز - سیاه را ..... از روی هرم دودویی در زمان  $O(n)$  ساخت.

- (۱) نمی‌توان - می‌توان (۲) می‌توان - نمی‌توان (۳) نمی‌توان - نمی‌توان (۴) می‌توان - می‌توان

- ۹- رشته‌ای به طول اولیه  $m \geq 1$  در یک صفحه از نرم‌افزار حروف چین داریم. دو عمل زیر را می‌توانیم به ترتیب دلخواه روی این رشته انجام دهیم.

• کپی: کل رشته موجود در صفحه را در حافظه ذخیره کن.

• پیست: رشته ذخیره شده در حافظه را به انتهای رشته موجود در صفحه اضافه کن.

- به‌طور مثال، اگر رشته اولیه  $ab$  باشد، پس از انجام یک عمل کپی و یک عمل پیست، رشته موجود در صفحه به‌صورت  $abab$  درخواهد آمد. اگر  $dp[n]$  نشان‌دهنده طول بزرگ‌ترین رشته قابل ایجاد با  $n$  بار استفاده از عمل کپی یا پیست (به ترتیب دلخواه) باشد، کدام رابطه بازگشتی زیر برای محاسبه  $dp[n]$  به‌ازای  $n \geq 5$  درست است؟ (فرض کنید در ابتدا حافظه خالی است).

(۱)  $dp[n] = 2dp[n-2]$

(۲)  $dp[n] = 2dp[n-1]$

(۳)  $dp[n] = \max(2dp[n-2], 2dp[n-3])$

(۴)  $dp[n] = \max(dp[n-1], 2dp[n-2])$

- ۱۰- محاسبه عنصر با بیشترین تکرار در یک آرایه  $n$  عضوی دلخواه در چه زمانی ممکن است؟ (بهترین گزینه را انتخاب کنید).

(۱)  $O(n \log(n))$  (۲)  $O(n\sqrt{n})$  (۳)  $O(n^2)$  (۴)  $O(n)$

- ۱۱- در مسئله «پوشش رأسی»، گراف ساده  $G = (V, E)$  و عدد صحیح  $k$  داده شده است و می‌خواهیم ببینیم «آیا زیرمجموعه  $C \subseteq V$  با اندازه حداکثر  $k$  وجود دارد طوری که حداقل یک سر هر یال گراف  $G$  در  $C$  باشد؟» چند مورد از گزاره‌های زیر درخصوص این مسئله درست هستند؟

• مسئله پوشش رأسی ان‌پی - کامل است.

• مسئله پوشش رأسی ان‌پی - سخت است.

• مسئله پوشش رأسی قابل کاهش به مسئله SAT است.

• مسئله SAT قابل کاهش به مسئله پوشش رأسی است.

(۱) ۱ (۲) ۲ (۳) ۳ (۴) ۴

(۱) ۱ (۲) ۲ (۳) ۳ (۴) ۴

- ۱۲- کدام یک از دو گزاره (الف) و (ب) در خصوص الگوریتم هافمن به ترتیب درست است؟  
الف - اگر فراوانی یک حرف در یک متن بیشتر از  $5/4$  باشد، آنگاه درخت هافمن متن لزوماً شامل کد یک بیتی است.  
ب - اگر فراوانی هر حرف در یک متن کمتر از  $5/3$  باشد، آنگاه درخت هافمن متن لزوماً شامل کد یک بیتی نیست.  
(۱) نادرست - درست (۲) درست - درست (۳) درست - نادرست (۴) نادرست - نادرست
- ۱۳- ماتریس  $A$  به اندازه  $n \times m$  داده شده است. سطرهای آن از چپ به راست و ستون های آن از بالا به پایین مرتب هستند. هزینه یافتن عدد داده شده  $x$  در این ماتریس چقدر است؟ (بهترین گزینه را انتخاب کنید).  
(۱)  $O((n+m)(\log n + \log m))$  (۲)  $O(m \log n + n \log m)$   
(۳)  $O(n+m)$  (۴)  $O(nm)$
- ۱۴- فرض کنید  $G$  یک گراف کامل وزن دار با  $n$  رأس است. درخت پوشای کمینه  $G$  در چه زمانی قابل محاسبه است؟ (بهترین گزینه ها را انتخاب کنید).  
(۱)  $\Theta(n)$  (۲)  $\Theta(n \log n)$   
(۳)  $\Theta(n^2 \log n)$  (۴)  $\Theta(n^2)$
- ۱۵- مسئله برنامه ریزی خطی با محدودیت های زیر را در نظر بگیرید:  
 $x_1 - x_2 \leq -1$   
 $x_2 - x_1 \leq -4$   
 $x_2 - x_3 = -9$   
 $x_3 - x_1 \leq 5$   
 $x_2 - x_3 \leq -3$   
کدام مورد یک جواب برای مسئله برنامه ریزی فوق است؟  
(۱) مسئله جواب ندارد. (۲) مسئله بی نهایت جواب دارد.  
(۳) حتماً  $x_1 = -9$  (۴) حتماً  $x_3 = -3$
- ۱۶- آرایه مرتب  $A$  شامل  $n$  عدد به صورت اکیداً صعودی داده شده است. یک نفر این آرایه را به اندازه  $k$  واحد شیفت دوری داده و نتیجه را به صورت یک آرایه  $B$  به ما داده است. هدف پیدا کردن مقدار  $k$  است در چه زمانی می توان مقدار  $k$  را با داشتن آرایه  $B$  محاسبه کرد؟ (بهترین گزینه را انتخاب کنید).  
(۱)  $O(n)$  (۲)  $O(\sqrt{n})$   
(۳)  $O(\log n)$  (۴)  $O(\log^2 n)$
- ۱۷- به ازای اعداد صحیح مثبت  $n$  و  $m$  و  $a$ ، در چه مرتبه زمانی می توان  $a^n \bmod m$  را محاسبه کرد؟ فرض کنید عملیات های ضرب و جمع به پیمانه  $m$  در زمان  $O(1)$  قابل انجام هستند. (بهترین گزینه را انتخاب کنید).  
(۱)  $O(\log n)$  (۲)  $O(\sqrt{n})$   
(۳)  $O(n)$  (۴)  $O(m)$
- ۱۸- می خواهیم در آرایه ای به طول  $n$  از اعداد صحیح، بیشترین تعداد درایه های صفر پشت سر هم را پیدا کنیم. این کار در چه مرتبه زمانی قابل انجام است؟ (بهترین گزینه را انتخاب کنید).  
(۱)  $O(n \log(n))$  (۲)  $O(n^2)$   
(۳)  $O(\log n)$  (۴)  $O(n)$

- ۱۹- کدام گزینه در خصوص دو گزاره (الف) و (ب) به ترتیب درست است؟  
الف - به ازای هر گراف  $G$ ، ترتیبی از یال های گراف وجود دارد که با تنها یک دور ریلکس کردن یال ها به آن ترتیب در اجرای الگوریتم بلمن - فورد، کوتاه ترین مسیر از رأس  $s$  به تمام رأس های دیگر محاسبه می شوند.  
ب - در مرتب سازی ادغامی هر عنصر با  $O(\log n)$  عنصر دیگر مقایسه می شود.

(۱) درست - درست

(۲) درست - نادرست

(۳) نادرست - درست

(۴) نادرست - نادرست

ب - اگر ظرفیت تمام یال های شبکه عددی گنگ باشد، آنگاه مقدار شار بیشینه شبکه حتماً عددی گنگ است.

(۱) درست - درست

(۲) درست - نادرست

(۳) نادرست - درست

(۴) نادرست - نادرست

تعداد نمونه	بازه
۹	$(0, 10]$
۶	$(10, 25]$
۵	$(25, \infty)$

$$F(x) = 1 - \frac{\theta}{x} \quad \text{for } \theta < x < \infty$$

(۱) ۵

(۲) ۵/۵

(۳) ۵/۷۵

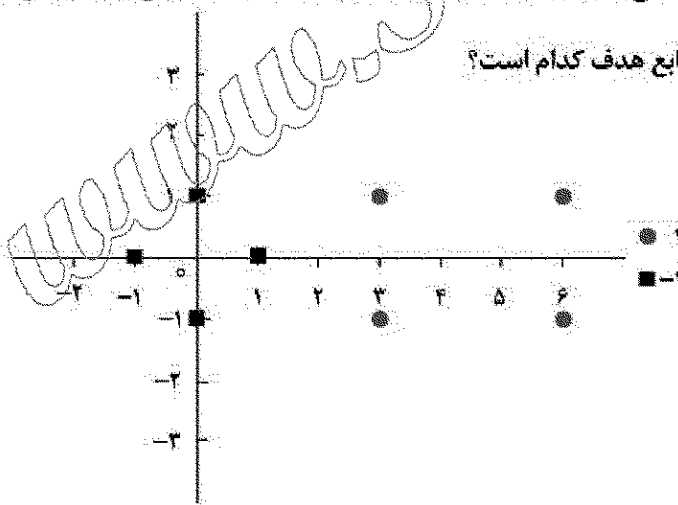
(۴) ۶

(۱) ۰/۳۳

(۲) ۱/۵

(۳) ۲/۵

(۴) ۳/۵



۲۳- اگر بخواهیم از آزمون فرض‌ها برای انتخاب یا عدم انتخاب ویژگی در یک مسئله از دسته‌بندی دودسته‌ای استفاده کنیم و بخواهیم پذیرش یا رد فرض صفر  $H_0: \Delta\mu = \mu_1 - \mu_2 = 0$  را به کمک  $N$  نمونه  $x_1, x_2, \dots, x_N$  از دسته ۱ و  $N$  نمونه  $y_1, y_2, \dots, y_N$  از دسته ۲ را بررسی کنیم، در صورتی که واریانس ویژگی‌های دو دسته برابر باشند،  $(\sigma_1^2 = \sigma_2^2)$  و مقدار واریانس دسته‌ها را ندانیم، مقدار Test Statistic مورد نظر کدام است؟  $\bar{x} = \frac{1}{n} \sum_{i=1}^N x_i$  و  $\bar{y} = \frac{1}{n} \sum_{i=1}^N y_i$

$$q = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\binom{r}{N} \binom{1}{rN-r} \left( \sum_{i=1}^N (x_i - \bar{x})^r + \sum_{i=1}^N (y_i - \bar{y})^r \right)}} \quad (1)$$

$$q = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\frac{1}{rN} \left( \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^r} + \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^r} \right)} \quad (2)$$

$$q = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\frac{1}{rN-r} \sqrt{\frac{r}{N} \left( \sum_{i=1}^N (x_i - \bar{x})^r + \sum_{i=1}^N (y_i - \bar{y})^r \right)}} \quad (3)$$

$$q = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\frac{1}{rN-r} \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^r + \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^r}} \quad (4)$$

۲۴- کدام یک از دو ماتریس داده شده زیر، ماتریس کوواریانس معتبر است؟

$$A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 4 & -1 & 1 \\ -1 & 4 & -1 \\ 1 & -1 & 4 \end{bmatrix}$$

(۲) فقط A

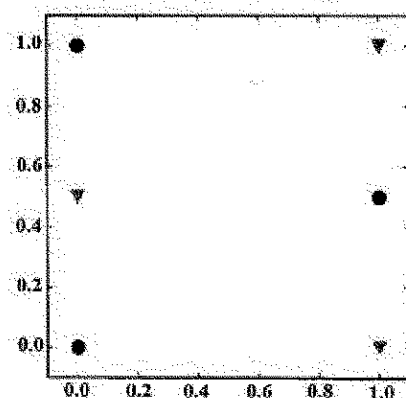
(۱) فقط B

(۴) هیچ کدام

(۳) هر دو ماتریس A و B

۲۵- کدام یک از دسته‌بندی‌های زیر قادر به جداسازی صحیح داده‌های آموزشی شکل زیر (دسته‌های دایره و مثلث) است؟

- 1 Logistic regression
- 2 SVM with linear kernel
- 3 SVM with RBF kernel
- 4 Decision tree
- 5 3-nearest-neighbor classifier (with Euclidean distance)



مجموعه داده‌ها

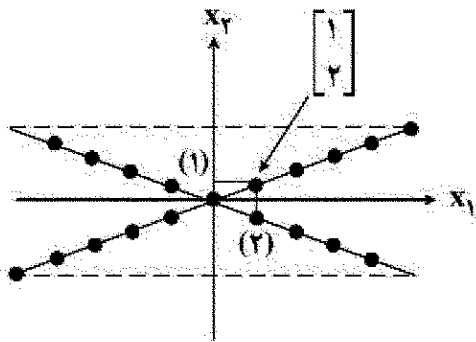
(۲) ۱: no, ۲: no, ۳: yes, ۴: no, ۵: no

(۱) ۱: no, ۲: no, ۳: yes, ۴: yes, ۵: no

(۴) ۱: no, ۲: no, ۳: no, ۴: yes, ۵: no

(۳) ۱: yes, ۲: no, ۳: no, ۴: yes, ۵: no

۲۶- داده های شکل زیر را در نظر بگیرید. اولین مؤلفه اساسی کدام است؟



(۱)  $[0, 1]^T$

(۲)  $[1, 0]^T$

(۳)  $[2, 1]^T$

(۴)  $[2, -1]^T$

۲۷- مجموعه نقاط دو بعدی در دو کلاس مفروض به شرح زیر می باشند:

$$C_1 = \{(0, -1)^T, (1, 0)^T, (2, 1)^T\}$$

$$C_2 = \{(1, 1)^T, (-1, 1)^T, (-1, -1)^T, (-2, -1)^T\}$$

T به معنای ترانزپوز بر دار است. مؤلفه اصلی این نقاط کدام است؟

(۲)  $\frac{1}{\sqrt{34}} \begin{bmatrix} 15 \\ 3 \end{bmatrix}$

(۴)  $\frac{1}{34} \begin{bmatrix} 15 \\ 3 \end{bmatrix}$

(۱)  $\frac{1}{\sqrt{34}} \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

(۳)  $\frac{1}{34} \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

۲۸- یک دسته بند از نوع بیزسالو (Naive Bayes) با سه متغیر ورودی دودویی ( $x_1$  Boolean) و  $x_2$  و  $x_3$  بر حسب

دودویی Y را در نظر بگیرید. برای آموزش این طبقه چه تعداد پارامتر باید تخمین زده شود؟

(۱) ۵

(۲) ۶

(۳) ۷

(۴) ۸

۲۹- در صورتی می گوئیم فرضیه (یا مدل) h بر روی داده های آموزشی overfit شده است که فرضیه دیگری مانند  $h'$

وجود داشته باشد و .....

(۱) خطای  $h'$  بر روی داده های آموزشی کمتر از h و خطای h بر روی کل توزیع داده ها کمتر از  $h'$  باشد.

(۲) خطای h بر روی داده های آموزشی کمتر از  $h'$  و خطای  $h'$  بر روی کل توزیع داده ها کمتر از h باشد.

(۳) خطای h بر روی داده های آموزشی بیشتر از خطای  $h'$  روی داده های تست باشد.

(۴) خطای  $h'$  بر روی داده های آموزشی بیشتر از خطای h روی داده های تست باشد.

۳۰- کدام یک از موارد زیر می تواند موجب کاهش بیش برآزش در SVM شود؟

(۱) کاهش واریانس کرنل RBF

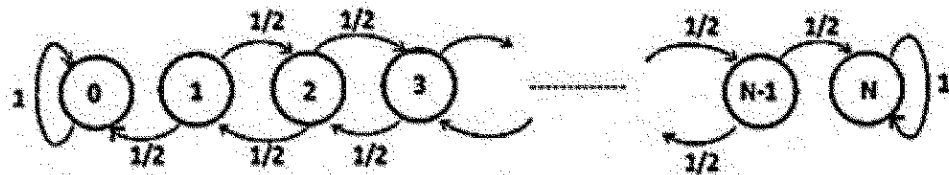
(۲) در نظر گرفتن Soft margin

(۳) افزایش درجه کرنل چند جمله ای

(۴) افزایش ضریب بخش مربوط به جریمه دسته بندی نادرست (C)

۳۱- زنجیره مارکوف مربوط به یک بازی با دیاگرام گذر زیر را در نظر بگیرید.

فرض کنید با شروع از حالت  $i$ ، آنقدر بازی را ادامه می‌دهید تا زمانی که به یکی از حالت‌های  $0$  یا  $N$  برسید که در این صورت بازی تمام است. با شروع از حالت  $i$ ، به طور متوسط چقدر طول می‌کشد تا بازی تمام شود؟



- (۱)  $i$   
(۲)  $N-i$   
(۳)  $i(N-i)$   
(۴)  $\frac{1}{2}(N-i)$

۳۲- یک مجموعه داده در فضای دوبعدی به صورت  $(X_1, X_2, Y)$  را در نظر بگیرید که در آن  $X_1$  و  $X_2$  ویژگی‌های داده  $Y$  بر حسب داده است. پس از یادگیری پارامترهای دسته‌بند بیز ساده (Naive Bayes)، جدول زیر به دست آمده است. اگر بدانیم درست‌نمایی (likelihood) دو نمونه  $(1, 0, 1)$  و  $(0, 1, 0)$  در مدل داده شده برابر با  $\frac{1}{180}$  است،

	$Y=0$	$Y=1$
$X_1$	$P(X_1=1 Y=0)=\frac{1}{5}$	$P(X_1=1 Y=1)=\frac{3}{8}$
$X_2$	$P(X_2=1 Y=0)=\frac{1}{3}$	$P(X_2=1 Y=1)=\frac{3}{4}$

آنگاه  $P(Y=1)$  کدام است؟

- (۱)  $\frac{3}{4}$   
(۲)  $\frac{2}{3}$   
(۳)  $\frac{1}{2}$   
(۴)  $\frac{1}{4}$

۳۳- کدام یک از روابط زیر، رابطه حد بالای Sample complexity با بعد VC را در چارچوب یادگیری PAC به درستی نشان می‌دهد؟

- (۱)  $m \geq \max\left(\frac{1}{\epsilon} \log_2\left(\frac{1}{\delta}\right), \frac{VC(H)-1}{\epsilon^2}\right)$   
(۲)  $m \geq \frac{1}{\epsilon^2} \left(4 \ln\left(\frac{1}{\epsilon}\right) + 8 VC(H) \ln\left(\frac{1}{\delta}\right)\right)$   
(۳)  $m \geq \frac{1}{\epsilon^2} \left(4 \log_2\left(\frac{2}{\delta}\right) + VC(H) \log_2\left(\frac{13}{\epsilon}\right)\right)$   
(۴)  $m \geq \frac{1}{\epsilon} \left(4 \log_2\left(\frac{2}{\delta}\right) + 8 VC(H) \log_2\left(\frac{13}{\epsilon}\right)\right)$



۳۴- فرض کنید یک مجموعه داده برای رگرسیون در اختیار دارید که توسط یک چند جمله ای درجه ۳ تولید شده است. در مورد بایاس و واریانس مدل های تخمین گر زیر بر روی مجموعه داده ذکر شده نسبت به مدل واقعی تولیدکننده داده، کدام نتیجه گیری درست است؟

مدلهای تخمین گر	نتیجه گیری ۱		نتیجه گیری ۲		نتیجه گیری ۳	
	Bias	Variance	Bias	Variance	Bias	Variance
Linear regression	low/high	low/high	low/high	low/high	low/high	low/high
Polynomial regression with degree ۲	low/high	low/high	low/high	low/high	low/high	low/high
Polynomial regression with degree ۱۰	low/high	low/high	low/high	low/high	low/high	low/high

(۴) هیچ کدام

(۳) ۲

(۲) ۲

(۱) ۲

۳۵- اگر ماتریس کوواریانس یک مجموعه داده به صورت ماتریس C باشد، نسبت پراکندگی داده در جهت مؤلفه اول (First component) به جهت مؤلفه دوم (Second component) (یا همان  $\frac{\lambda_1}{\lambda_2}$ ) در روش کاهش ابعاد PCA به

$$C = \frac{1}{5} \begin{bmatrix} ۱۸ & ۵ \\ ۵ & ۲ \end{bmatrix}$$

کدام عدد نزدیکتر است؟

(۱) ۳

(۲) ۹

(۳) ۲۵

(۴) ۲۴

۳۶- به عنوان یک ایده، اگر بخواهیم از بسط تیلور برای تخمین بهتر مقدار تابع هزینه C حول نقطه w در شبکه های عصبی استفاده کنیم و فرایند نزول در امتداد گرادیان (Gradient descent) را برای یافتن وزن های بهینه به صورت موثرتری انجام دهیم، در صورتی که فقط از سه جمله اول بسط استفاده کنیم، رابطه  $\Delta w$  به چه صورت خواهد بود؟ ( $\eta$  نرخ یادگیری، VC بردار گرادیان تابع هزینه و H ماتریس هسیان (Hessian matrix) تابع هزینه است که در آن

$$H[i, j] = \partial^2 C / \partial w_i \partial w_j \text{ است.}$$

$$\Delta w = -\eta HVC \quad (۲)$$

$$\Delta w = -\eta H^{-1}VC \quad (۱)$$

$$\Delta w = -\eta HH^{-1}VC \quad (۳)$$

(۴) استفاده از ایده بسط تیلور بدین منظور مقید نیست.

۳۷- کدام مورد در خصوص الگوریتم خوشه بندی k-means نادرست است؟

(۱) نسبت به داده های نویزی حساسیت بالایی دارد.

(۲) با رسیدن به بهینه سراسری خاتمه پیدا می کند.

(۳) از نظر زمان کارا است و مرتبه زمانی مطلوبی دارد.

(۴) نسبت به داده های پرت (outlier) حساسیت بالایی دارد.

۳۸- کدام روش برای مواجهه با بیش برازش شدن (overfitting) درخت تصمیم مناسب نیست؟

- (۱) اجازه دهیم که درخت به داده های آموزشی overfit شود و سپس، با استفاده از داده های اعتبارسنجی، برخی انشعاب های درخت را هرس کنیم.
- (۲) در حین آموزش درخت تصمیم، پیش از اینکه درخت، همه داده های آموزشی را به درستی دسته بندی کند، به صورت زودهنگام به ساخت درخت خاتمه دهیم.
- (۳) اجازه دهیم که درخت به داده های آموزشی overfit شود، و سپس درخت را به مجموعه قوانین تبدیل می کنیم و با استفاده از داده های اعتبارسنجی، برخی از قوانین را حذف کنیم.
- (۴) داده های آموزشی را با استفاده از نمونه برداری، نامتوازن (imbalance) کنیم و سپس با استفاده از داده های نامتوازن شده، آموزش درخت را انجام دهیم.

۳۹- کدام یک از روش های زیر از روش های یادگیری تنبل (lazy) محسوب نمی شود؟

- (۱) رگرسیون محلی وزن دار (Locally weighted regression)
- (۲) استدلال مبتنی بر مورد (Case based reasoning)
- (۳) بردار ماشین پشتیبان (SVM)
- (۴) K نزدیک ترین همسایه (KNN)

۴۰- داده های زیر را در نظر بگیرید که مؤلفه دوم برچسب داده و مؤلفه اول بردار ویژگی داده است.

$$x_1 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, x_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, x_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, x_4 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

در صورتی که برای جداسازی دو دسته از ماشین بردار پشتیبان (SVM) استفاده شود، در تابع هزینه

$$\frac{1}{2} \omega^T \omega - \sum_{i=1}^n \alpha_i [y_i (\omega^T x_i + b) - 1]$$

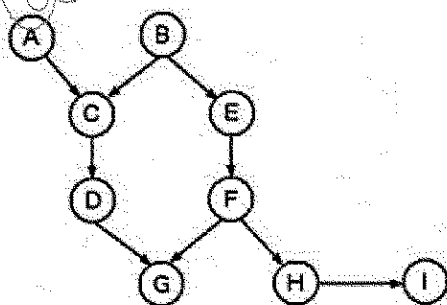
(۱) ۰٫۵

(۲) ۰٫۲۵

(۳) صفر

(۴) -۰٫۵

۴۱- شبکه بیزین با ۹ متغیر دودویی به شکل زیر را در نظر بگیرید. کدام یک از روابط زیر همیشه در مورد این شبکه درست است؟



$$P(A, I) = P(A)P(I) \quad (1)$$

$$P(C|B, F) = P(C|F) \quad (2)$$

$$P(A, B|G) = P(A|G)P(B|G) \quad (3)$$

$$P(B, H|E, G) = P(B|E, G)P(H|E, G) \quad (4)$$

۴۲- یک کرنل گاوسی با رابطه  $k(x_1, x_2) = e^{-\frac{\|x_1 - x_2\|^2}{2}}$  را در نظر بگیرید. فاصله بین دو بردار ویژگی  $x_1 = [1, 1]^T$

و  $x_2 = [2, 2]^T$  در فضای جدید برابر با کدام مورد است؟

(۱)  $\sqrt{2}$

(۲)  $e^{-1}$

(۳)  $e^{-\frac{1}{2}}$

(۴)  $\sqrt{2 - \frac{2}{e^2}}$

۴۳- فرض کنید در حال انجام یک بازی با دو سکه هستید. سکه اول با احتمال  $\theta$  و سکه دوم با احتمال  $2\theta$  شیر می‌آید. پس از چندین پرتاب این سکه‌ها نتایج به شکل جدول زیر به دست آمده است:

پیشامد	سکه
شیر	۱
خط	۲
خط	۲
خط	۲
شیر	۲

مقدار Log-Likelihood این داده‌ها براساس  $\theta$  چقدر است؟

(۱)  $\text{Log } 2 + 2\text{Log}(1 - 2\theta)$

(۲)  $\text{Log } 2 + 2\text{Log } \theta + 2\text{Log}(1 - 2\theta)$

(۳)  $2\text{Log } \theta + 2\text{Log}(1 - 2\theta)$

(۴)  $\text{Log } 2 + 2\text{Log } \theta$

۴۴- کدام گزینه بعد VC یک دسته‌بند  $1 - NN$  را نشان می‌دهد که  $N$  تعداد نمونه‌های آموزشی است؟

(۱) ۱

(۲) ۲

(۳)  $\infty$

(۴)  $N$

۴۵- در رابطه با روش خوشه‌بندی مقید (constraint clustering) کدام گزینه نادرست است؟

(۱) روش خوشه‌بندی مقید، یک روش بدون نظارت (unsupervised) است.

(۲) برخی الگوریتم‌های خوشه‌بندی مقید، قیدها را به صورت قید سخت و برخی دیگر به صورت قید نرم در نظر می‌گیرند.

(۳) در روش خوشه‌بندی مقید، در مورد نمونه‌های داده که باید در یک خوشه قرار بگیرند، قیدها یا راهنمایی‌هایی ارائه می‌شود.

(۴) در روش خوشه‌بندی مقید، در مورد نمونه‌های داده که باید در خوشه‌های متمایز قرار بگیرند، قیدها یا راهنمایی‌هایی

ارائه می‌شود.