

کد کنترل



332

E

محل امضای:

نام:

نام خانوادگی:

صبح جمعه
۱۳۹۶/۱۲/۴

دفترچه شماره (۱)



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

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

آزمون ورودی دوره دکتری (نیمه‌تمركز) - سال ۱۳۹۷

رشته مهندسی کامپیوتر - هوش مصنوعی (کد ۲۳۵۶)

مدت پاسخگویی: ۱۵۰ دقیقه

تعداد سوال: ۴۵

عنوان مواد امتحانی، تعداد و شماره سوالات

ردیف	مواد امتحانی	تعداد سوال	از شماره	تا شماره
۱	مجموعه دروس تخصصی: ساختمنداده ها و طراحی الگوریتم ها - شناسایی الگو - یادگیری ماشین	۴۵	۱	۴۵

استفاده از ماشین حساب مجاز نیست.

این آزمون نمره منفی دارد.

حق جای، تکیر و انتشار سوالات به هر روش (الکترونیکی و...)، پس از برگزاری آزمون، برای تمامی انتخاب حرفی و حرفی تهاجاً مجوز این سازمان مجاز نباشد و با متخلفین تراویح ممنوع است.

* داوطلب گرامی، عدم درج مشخصات و امضا در مندرجات جدول ذیل، بهمنزله عدم حضور شما در جلسه آزمون است.

اینجانب با شماره داوطلبی در جلسه این آزمون شرکت می‌نمایم.

امضا:

-۱ میزان رشد توابع زیر به ترتیب صعودی (از چپ به راست) کدام است؟

$$n \log^*(n), \log(n)^{\log(n)}, \log(n!), \log(\log(n^n))$$

$$n \log^*(n), \log(n!), \log(n)^{\log(n)}, \log(\log(n^n)) \quad (1)$$

$$\log(\log(n^n)), \log(n!), n \log^*(n), \log(n)^{\log(n)} \quad (2)$$

$$\log(\log(n^n)), \log(n!), \log(n)^{\log(n)}, n \log^*(n) \quad (3)$$

$$\log(\log(n^n)), n \log^*(n), \log(n!), \log(n)^{\log(n)} \quad (4)$$

جواب دو رابطه بازگشته زیر کدام است؟ -۲

$$T(n) = T(3/7n) + T(4/7n) + n, T(1) = 1$$

$$T'(n) = T'(2/7n) + T'(4/7n) + n, T'(1) = 1$$

$$T(n) = \Theta(n), T'(n) = \Theta(n) \quad (1)$$

$$T(n) = \Theta(n), T'(n) = \Theta(n \log n) \quad (2)$$

$$T(n) = \Theta(n \log n), T'(n) = \Theta(n) \quad (3)$$

$$T(n) = \Theta(n \log n), T'(n) = \Theta(n \log n) \quad (4)$$

-۳ فرض کنید یک زبان از حروف الفبای $\{a, b, c, d, e, f, g, h, i\}$ تشکیل شده است و احتمال وقوع a برابر

۱۸ درصد، b برابر ۴ درصد، c برابر ۸ درصد، d برابر ۱۰ درصد، e برابر ۲۰ درصد، f برابر ۵ درصد، g برابر ۵

درصد، h برابر ۱۵ درصد و i برابر ۱۵ درصد است. درخت هافمن این زبان چند گره دارد؟

۱۸ (۱)

۱۷ (۲)

۱۶ (۳)

۱۵ (۴)

-۴ فرض کنید $H : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ یک تابع درهم‌ساز یکنواخت باشد. برای ورودی x عدد Z را برابر تعداد

صفرهای سمت راست ($H(x) \geq c \log n$) قرار می‌دهیم. برای عدد $0 \leq c \leq 1$ ، احتمال $z \geq c \log n$ از چه مرتبه‌ای است؟

(فرض کنید c ثابت است).

$$O(1/n) \quad (1)$$

$$O(1/n^c) \quad (2)$$

$$O(1/\log n) \quad (3)$$

$$O(1/\log^c n) \quad (4)$$

-۵ چه تعداد از تبدیل‌های زیر در زمان $O(n)$ قابل انجام است؟

- تبدیل پیمایش پس‌ترتیب عناصر یک درخت دودویی کامل به پیمایش پس‌ترتیب آن
- تبدیل پیمایش پس‌ترتیب یک درخت دودویی کامل به پیمایش پیش‌ترتیب آن
- تبدیل پیمایش میان‌ترتیب عناصر یک درخت دودویی کامل به یک درخت دودویی جست‌و‌جو

۱ (۱)
۲ (۲)
۳ (۳)
۴ (۴)

-۶ در یک داده ساختار هرم با n عنصر، عدد بعدی یک رأس (عددی که در دنباله‌ی مرتب‌شده بعد از عدد این رأس می‌آید) را در چه زمانی می‌توان به دست آورد؟

$O(1)$ (۱)
 $O(n)$ (۲)
 $O(\sqrt{n})$ (۳)
 $O(\log n)$ (۴)

-۷ اعداد ۱ تا ۱۵ درون آرایه A به‌گونه‌ای ذخیره شده‌اند که تشکیل یک هرم کمینه متوازن می‌دهند. حداقل تعداد نایه‌جایی‌های A چه تعداد است؟

(دو درایه $A[i]$ و $A[j]$ تشكیل یک نایه‌جایی می‌دهند اگر $j < i$ و $A[i] > A[j]$)

۱۱۰ (۱)
۹۴ (۲)
۷۱ (۳)
۵۹ (۴)

-۸ آرایه A از n عدد دلخواه داده شده است. فرض کنید عملیات reverse(i, j) برای $n \leq j < i \leq 1$, زیرآرایه $A[i..j]$ را معکوس می‌کند، یعنی به ازای هر $i, 0 \leq k \leq j - i$ ، $A[i+k]$ را با $A[j-k]$ تعویض می‌کند. با حداقل چندبار استفاده از این عملیات می‌توان آرایه A را مرتب کرد؟

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

-۹ آرایه A شامل n عدد مختلف است. حال می‌خواهیم آرایه B را به این صورت پر کنیم که به ازای هر i , $A[1]$ تا $A[i]$ برابر با میانه اعداد $A[1]$ تا $A[i]$ باشد. بهترین الگوریتم برای این کار از چه مرتبه‌ای است؟

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

- ۱۰ فرض کنید گراف G یک گراف جهت دار و وزن دار است که دور منفی ندارد. رئوس این گراف را با اعداد 1 تا n بر جسب گذاری می‌کنیم و وزن یال از i به j را با $w(i, j)$ نشان می‌دهیم. اگر گراف G' همان گراف G باشد، که فقط وزن یال‌های آن که با w' نشان می‌دهیم، طبق قاعده‌های زیر تغییر کرده است، به ازای چندتا از این قاعده‌ها، کوتاه‌ترین مسیر (خود مسیر نه طول مسیر) بین هر دو رأس داده شده در دو گراف G و G' یکسان است؟

$$w'(i, j) = w(i, j) + i - j$$

$$w'(i, j) = w(i, j) + j - i$$

$$w'(i, j) = w(i, j) + i + j$$

۰ (۱)

۱ (۲)

۲ (۳)

۳ (۴)

- ۱۱ الگوریتمی را در نظر بگیرید که ورودی a_1, \dots, a_n شامل n عدد مجزا را به ترتیب داده شده می‌خواند و هنگام خواندن a_i مقدار متغیر x را به احتمال $1/n$ برابر a_i قرار می‌دهد. الگوریتم در پایان مقدار x را به عنوان خروجی گزارش می‌کند. یا چه احتمالی خروجی الگوریتم برابر a_j است؟

(۱) می‌تواند هر مقداری کوچکتر یا مساوی $1/n$ باشد.

(۲) می‌تواند هر مقداری در بازه $[1/n, 1]$ باشد.

$1/n$ (۳)

$1/n$ (۴)

- ۱۲ در گراف جهت دار G وزن یال‌ها را یک قرار می‌دهیم و شارپیشینه از رأس s به t را محاسبه می‌کنیم. مقدار جربان پیشینه برابر کدام مورد است؟

(۱) تعداد مسیرهای بین s و t

(۲) تعداد کوتاه‌ترین مسیرهای بین s و t

(۳) تعداد مسیرهای مجازی بین s و t

(۴) تعداد مسیرهای مجازی رأسی بین s و t

- ۱۳ چندتا از گزاره‌های زیر درست است؟

* اگر مسئله تصمیم‌گیری X در ان بی باشد، مسئله تصمیم‌گیری $\text{not } X$ نیز در ان بی است.

* هر مسئله ان بی – سخت به یک مسئله ان بی – کامل قابل کاهش است.

* تمام مسائل ان بی – کامل به تمام مسائل ان بی – سخت قابل کاهش‌اند.

۰ (۱)

۱ (۲)

۲ (۳)

۳ (۴)

۱۴- فرض کنید گراف جهتدار G شامل n رأس و m یال، روابط دوستی بین n فرد را مدل می‌کند. به عبارت دقیق‌تر از u به v یال وجود دارد، اگر شخص u شخص v را بشناسد. در این شبکه هرگاه شخص از یک خبر مطلع شود آن را به اطلاع همه دوستان خود می‌رساند. می‌خواهیم یک خبر مشخص را به اطلاع همه برسانیم. می‌خواهیم کمترین تعداد افرادی را پیدا کنیم که با مطلع شدن آن‌ها، همه از خبر فوق مطلع شوند. بهترین الگوریتم برای محاسبه این مجموعه افراد از چه مرتباً است؟

(۱) $O(m+n)$

(۲) $O(\min(n,m))$

(۳) $O(n \log n + m)$

(۴) این مسئله انپی - سخت است.

۱۵- در جست و جوی سطح اول (BFS)، به هر رأس یک بازه زمانی نسبت می‌دهیم، طوری که زمان قراردادن رأس در صفحه شروع بازه و زمان برداشتن آن از صفحه انتهای بازه باشد. کدام مورد درست است؟

(۱) ترتیب طول بازه‌های بین رؤوس، معادل با ترتیب فاصله‌ی آنان از رأس شروع است.

(۲) اگر یک رأس از نوادگان رأس دیگر در درخت BFS باشد، آنگاه بازه‌ی آن‌ها اشتراک ندارد.

(۳) اگر بازه‌ی دو رأس اشتراک نداشته باشد، آن‌گاه یکی از آن‌ها نوی دیگری در درخت BFS است.

(۴) اگر شروع بازه‌ی b بعد از شروع بازه‌ی a باشد، فاصله رأس شروع تا b بیشتر از فاصله رأس شروع تا a است.

۱۶- فرض کنید $T = (V, E)$ یک درخت شامل n رأس باشد. می‌خواهیم $C \subseteq V$ با حداقل تعداد رأس پیدا کنیم، طوری که برای هر یال در E ، حداقل یکی از دو سر آن یال در C باشد. الگوریتم حریصانه زیر را در نظر بگیرید.

C را در ابتدا تهی قرار می‌دهیم. در هر مرحله و اس با درجه بیشینه در T را در C قرار می‌دهیم (درصورتی که بیش از یک رأس با درجه بیشینه بود، یکی از آن‌ها را به دلخواه انتخاب می‌کنیم) و تمام یال‌های مجاور آن را از T حذف می‌کنیم. این کار را تا زمانی که T تهی از یال شود ادامه می‌دهیم. کمترین مقدار n که برای آن، الگوریتم حریصانه فوق درست کار نمی‌کند، کدام است؟

(۱) ۳

(۲) ۴

(۳) ۵

(۴) برای هر n الگوریتم حریصانه جواب بهینه را برمی‌گرداند.

۱۷- گراف $G = (V, E)$ با وزن‌های مثبت را در نظر بگیرید. $d(u, v, k)$ را برابر طول کوتاه‌ترین مسیر از u به v در

نظر بگیرید که حداقل k یال داشته باشد. چندتا از رابطه‌های بازگشتی زیر برای $k > 1$ درست می‌باشند؟

(۱) وزن یال $w = (u, v)$ را نشان می‌دهد و وزن یالی که وجود نداشته باشد بی‌نهایت در نظر گرفته می‌شود.

$$d(u, v, k) = \min_{x \in V} (d(u, v, k-1), d(u, x, k-1) + w(x, v))$$

$$d(u, v, k) = \min_{x \in V} (d(u, x, k-1) + w(x, v))$$

$$d(u, v, k) = \min_{x \in V} (d(u, v, k-1), d(u, x, \lceil k/2 \rceil) + d(x, v, \lfloor k/2 \rfloor))$$

(۱) ۳

(۲) ۲

(۳) ۱

(۴) ۰

- ۱۸ در یک گراف همیند با n رأس و m یال، وزن یال‌ها ۱ یا ۲ هستند. بهترین الگوریتم برای یافتن درخت پوشای کمینه این گراف دارای چه زمان اجرایی است؟

$O(n \log n)$ (۲)

$O(n \log n + m)$ (۱)

$O(m)$ (۴)

$O(n)$ (۳)

- ۱۹ سعید و وحید بازی زیر را انجام می‌دهند. ابتدا دو بازیکن روی یک عدد n توافق می‌کنند. بازی با $x = 2$ آغاز می‌شود. بازیکن اول با انداختن سکه انتخاب شده و بازی را آغاز می‌کند. از این به بعد هر کدام از بازیکنان به دلخواه x را به توان ۲ یا ۳ می‌رساند. مثلاً اگر سعید بازی را آغاز کرده و ۳ را انتخاب کند، x برابر با ۸ خواهد شد. بعد از حرکت هر کدام از بازیکنان که $n > X$ شود، بازی متوقف شده و بازیکن آخر به عنوان برنده بازی انتخاب می‌شود. بازی‌ها به طور متوسط در چند مرحله به پایان می‌رسند؟

$\Theta(\log^2 n)$ (۲)

$\Theta(\log \log n)$ (۱)

(۴) قابل محاسبه نیست.

$\Theta(\log n)$ (۳)

- ۲۰ می‌خواهیم ماتریس‌های (5×10) و M_1 و (10×3) و M_2 و (5×12) و M_3 و (12×5) و M_4 را با همین ترتیب در هم ضرب کنیم. این کار با حداقل چند ضرب عددی قابل انجام است؟

۳۶۰ (۱)

۴۰۵ (۲)

۵۸۰ (۳)

۶۳۰ (۴)

- ۲۱ در یک دسته‌بندی M کلاسه، کدام مورد در تخصوص کران بالای خطای دسته‌بند (P_e) درست است؟

$$P_e \leq \frac{1}{M} \quad (۱)$$

$$P_e \leq \frac{2}{M} \quad (۲)$$

$$P_e \leq \frac{1}{M-1} \quad (۳)$$

$$P_e \leq \frac{M-1}{M} \quad (۴)$$

- ۲۲ در یک کاربرد داده‌ها از توزیع $P(X=k) = \theta(1-\theta)^{k-1}$ نمونه‌برداری می‌شوند. فرض کنید که در مجموعه آموزشی متغیر X دارای مقادیر ۴، ۲، ۷ و ۹ است، چه مقداری از θ تخمین درستنمایی (Maximum Likelihood) را بیشینه می‌کند؟

$\frac{4}{22}$ (۱)

$\frac{3}{22}$ (۲)

$\frac{2}{22}$ (۳)

$\frac{1}{22}$ (۴)

- ۲۳- اگر X متغیری تصادفی با $E[X] = \mu$ و $Var(X) = \sigma^2$ باشد، مقدار $E[(X - c)^2]$ کدام است؟
 (فرض کنید c یک ثابت دلخواه باشد).

$$\mu^2 + c^2 + \sigma^2 \quad (1)$$

$$(\mu - c)^2 + \sigma^2 \quad (2)$$

$$(\mu - c)^2 + 2\sigma^2 \quad (3)$$

$$(\mu + c)^2 + 2\sigma^2 \quad (4)$$

- ۲۴- در یک مسئله دسته‌بندی دو دسته‌ای در فضای دوبعدی، اگر توزیع نمونه‌های داخل هر دسته را گوسی در نظر بگیریم، کدام جمله درست است؟

- (۱) اگر ماتریس کوواریانس توزیع دسته‌ها یکسان فرض شود، دسته‌بند بیز ساده مرز تصمیم خطی را پیدا می‌کند.
 (۲) اگر ماتریس کوواریانس توزیع دسته‌ها یکسان فرض شود، دسته‌بند بیز ممکن است مرز را به صورت یک چندجمله‌ای با درجه ۲ پیدا کند.

- (۳) اگر ماتریس کوواریانس توزیع دسته‌ها ماتریس همانی فرض شود، مرز حاصل توسط دسته‌بند بیز لزوماً عمود منصف خط متصلاً کننده میانگین نمونه‌های دو دسته خواهد بود.

- (۴) اگر برای مدل‌سازی توزیع نمونه‌های دسته‌ها، فرض استقلال شرطی دسته‌بند بیز ساده (Naïve Bayes) را هم اضافه نماییم حتماً مرز خطی توسط دسته‌بند بیز ساده به دست می‌آید.

- ۲۵- یک مسئله دسته‌بندی دو دسته‌ای در فضای دو بعدی دودویی را در نظر بگیرید. اگر سه نمونه آموزشی مطابق شکل زیر داشته باشیم که نمونه‌های مربع دارای برچسب مثبت و نمونه‌های دائره دارای برچسب منفی باشند، در صورتی که نمونه دائرة به دسته‌بند داده شود، احتمال این که دسته‌بند به این نمونه برچسب مثبت دهد، کدام است؟



(۱) ۰

(۲) $\frac{1}{6}$

(۳) $\frac{1}{3}$

(۴) ۱

- ۲۶- کدام مورد درخصوص logistic regression ، درست است؟

- (فرض کنید که فضای ورودی با x و برچسب با y نشان داده شوند).

- (۱) دسته‌بند logistic regression احتمال $p(x)$ را تخمین می‌زنند.

- (۲) دسته‌بند logistic regression احتمال $p(y|x)$ را تخمین می‌زنند.

- (۳) دسته‌بند logistic regression در جداولی داده‌های جداولی بذیر خطی تنها تفکیک داده‌ها را در نظر نمی‌گیرد بلکه فاصله آن‌ها را مرز تصمیم را هم به نوعی لحاظ می‌کند.

- (۴) در دسته‌بند logistic regression مقدار هزینه لحاظ شده برای نمونه x با برچسب مثبت را به صورت

$$f(x) = \frac{1}{1 + \exp(w^T x + w_0)} - \text{در نظر می‌گیرد که}$$

- ۲۷ - کدام مورد، درست است؟

- (۱) برای یک ویژگی که از برچسب مستقل است، ممکن است میزان کوواریانس با برچسب بزرگ باشد.
- (۲) کم بودن میزان همبستگی (correlation) یک ویژگی با برچسب آن به معنای نامفید بودن آن ویژگی است.
- (۳) معیار اطلاعات متقابل (Mutual information) بین یک ویژگی و یک برچسب تنها می‌تواند ارتباط خطی بین برچسب و ویژگی را نشان دهد.
- (۴) ممکن است دو ویژگی هر کدام به صورت تک تک نتوانند به دسته‌بندی داده‌ها کمکی کنند اما در کنار هم بتوانند داده‌ها را به خوبی دسته بندی نمایند.

- ۲۸ - مسأله دسته‌بندی برای دو کلاس با احتمال‌های بیشین برابر ($p(w_1) = p(w_2)$) و توابع توزیع گوسی به صورت داده شده را در نظر بگیرید. در صورتی که $a \times b - c \times e = 1$ ، شرایطی که باید بر مقادیر a, b, c, d, e و حاکم باشد تا مرز تصمیم به صورت خط باشد، کدام است؟

$$p(x|w_1) = N\left(\begin{bmatrix} \circ \\ \circ \end{bmatrix}, \begin{bmatrix} a & c \\ c & b \end{bmatrix}\right) \quad p(x|w_2) = N\left(\begin{bmatrix} d \\ e \end{bmatrix}, \begin{bmatrix} 1 & \circ \\ \circ & 1 \end{bmatrix}\right)$$

$$a=1, b=1, c=\circ, d \neq \circ, e \neq \circ \quad (1)$$

$$a=1, b=1, c \neq \circ, d \neq \circ, e \neq \circ \quad (2)$$

$$a \neq 1, b \neq 1, c=\circ, d \neq \circ, e \neq \circ \quad (3)$$

$$a \neq 1, b \neq 1, c=\circ, d \in \mathbb{R}, e \in \mathbb{R} \quad (4)$$

- ۲۹ - تابع چگالی احتمال $f(x; a) = \begin{cases} \frac{2x}{a^2} & \text{if } 0 < x \leq a \\ 0 & \text{otherwise} \end{cases}$ را در نظر بگیرید. با فرض این که x_1, x_2, \dots, x_n نمونه‌هایی باشند که به صورت i.i.d از این توزیع انتخاب شده باشند، تخمین درست‌نمایی بیشینه برای پارامتر a از این تابع توزیع کدام است؟

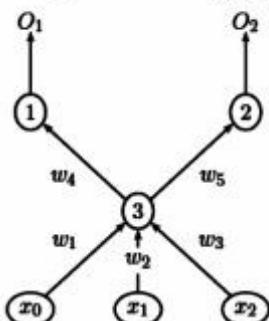
$$\frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

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

$$\frac{\sum_{i=1}^n x_i}{n} \quad (3)$$

$$\max(x_1, x_2, \dots, x_n) \quad (4)$$

- ۳۰ شبکه عصبی زیر را که در آن نرون‌های $i = 1, 2, 3$ از تابع سیگموید $\sigma(x) = \frac{1}{1+e^{-x}}$ به عنوان تابع فعال ساز استفاده می‌کنند در نظر بگیرید. خروجی هر نرون i را O_i می‌نامیم. فرض کنید در لحظه t به ازای مقادیر ورودی $x_0 = 1$ ، $x_1 = 0$ ، $x_2 = 1$ و مقادیر وزن $w_1 = 2$ ، $w_2 = 2$ ، $w_3 = 3$ ، $w_4 = 2$ و $w_5 = 2$ مقادیر خروجی برابر با $y_1 = 0$ ، $y_2 = 1$ و $y_3 = 1$ است. معادله وزن w_5 پس از انتشار (Backpropagation) و در لحظه $t+1$ از کدام رابطه بدست می‌آید؟ (ترجیح بگیرید را ۱ در نظر بگیرید).



$$w_5^{t+1} = 2 + \eta O_2^T (1 - O_2)^T O_2 \quad (1)$$

$$w_5^{t+1} = 2 + \eta O_2 (1 - O_2)^T O_2 \quad (2)$$

$$w_5^{t+1} = 2 + \eta O_2^T (1 - O_2) O_2 \quad (3)$$

$$w_5^{t+1} = 2 + \eta O_2 (1 - O_2) O_2 \quad (4)$$

- ۳۱ در دسته‌بند ماشین بردار پشتیبان برای دسته‌بندی داده‌های دو دسته‌ای، کمترین فاصله دو بردار پشتیبان که از دو دسته مختلف انتخاب شده باشند، کدام است؟ (بردار w ضرایب ابرصفحه جداکننده را مشخص می‌کند).

$$\frac{2}{\|w\|} \quad (1)$$

$$\frac{2}{\|w\|^2} \quad (2)$$

$$\frac{1}{\|w\|^2} \quad (3)$$

$$\frac{1}{\|w\|} \quad (4)$$

- ۳۲ در روش **K-fold cross validation** اگر K افزایش یابد، کدام مورد درخصوص خطای تخمین زده شده درست است؟

(۱) خطای تخمین زده شده افزایش می‌یابد.
(۲) خطای تخمین زده شده کاهش می‌یابد.

(۳) خطای تخمین زده شده تغییر نمی‌کند.
(۴) نمی‌توان درباره تغییر خطای صحبت نمود.

- ۳۳ بعد VC یک مدل مخلوط گوسی با دو مؤلفه که دارای ماتریس کوواریانس یکسان هستند، کدام است؟

۱ (۱)

۲ (۲)

۳ (۳)

۴ (۴)

- ۳۴ درخصوص انتخاب مجموعه **Bootstrap** از یک مجموعه P عضوی که هر عضو آن P بعد دارد، کدام مورد درست است؟

(۱) نمونه‌برداری P نمونه بدون جایگزینی
(۲) نمونه‌برداری \sqrt{P} نمونه با جایگزینی

(۳) نمونه‌برداری P نمونه با جایگزینی
(۴) نمونه‌برداری P نمونه با جایگزینی

- ۳۵ فرض کنید H مجموعه تمام دسته‌بندهایی باشد که $VC(H) = 5$ است. یک مجموعه آموزشی $S = \{(x_1, t_1), (x_2, t_2), (x_3, t_3), (x_4, t_4)\}$ در اختیار داریم، که x_i نمونه و t_i برچسب آن است و الگوریتم یادگیری یک دسته‌بند $h \in H$ را کمینه نمودن خطای آموزشی انتخاب می‌کند. اگر هیچ‌گونه اطلاعات دیگری درباره مجموعه H نداشته باشیم، کدام مورد در خصوص $h(x_5)$ درست است؟

(توجه شود $h(x_5)$ برچسب نمونه x_5 است که در مجموعه آموزشی وجود ندارد.)

(۱) با توجه به اینکه خطای آموزشی صفر است، دسته‌بند بیش برازش شده و x_5 درست برچسب زده نمی‌شود.

(۲) با توجه به اینکه خطای آموزشی صفر است، هیچ سخنی درباره برچسب x_5 نمی‌توان گفت.

(۳) با توجه به اینکه خطای آموزشی صفر است، لذا همیشه برچسب x_5 درست انتخاب می‌شود.

(۴) با توجه به اینکه خطای آموزشی صفر است، برچسب x_5 درست انتخاب می‌شود.

- ۳۶ کدام مورد در خصوص بعد VC برای دسته‌بند در فضای ویژگی دو بعدی، درست نیست؟

(۱) بعد VC دسته‌بند پرسپترون خطی برابر ۳ است.

(۲) بعد VC دسته‌بندی که داخل دایره را مثبت اعلام می‌کند برابر ۳ است.

(۳) بعد VC دسته‌بندی که داخل مثلث را مثبت اعلام می‌کند برابر ۷ است.

(۴) بعد VC دسته‌بندی که مرز تصمیم‌گیری آن یک خط موازی یکی از محورهای مختصات است برابر ۲ است.

- ۳۷ اگر K_1 یک هسته معتبر باشد، در کدام مورد $K(x, y) = K_1(x, y)$ لزوماً هسته معتبر نیست؟

$$K(x, y) = \exp(K_1(x, y)) \quad (1)$$

$$a > 0 \text{ برای } K(x, y) = a K_1(x, y) \quad (2)$$

$$K(x, y) = x^T A y \quad (3)$$

$$f(x) K(x, y) = f(x) K_1(x, y) f(x) \quad (4)$$

- ۳۸ درصد خطای (Leave one out cross validation) LOOCV ۳NN روی مجموعه داده زیر چند درصد است؟



۰ (۱)

۲۵ (۲)

۳۰ (۳)

۶۰ (۴)

- ۳۹ در صورتی که از الگوریتم سلسله مراتبی پایین به بالا برای خوشبندی $n=2^N$ داده $2^1, 2^2, \dots, 2^N$ استفاده کنیم، ارتفاع درخت خوشبندی بر اساس معیار Complete link با استفاده از معیار فاصله اقلیدسی برابر با h خواهد بود.

در صورتی که به جای معیار فاصله اقلیدسی از معیار فاصله اقلیدسی برای محاسبه فاصله دو نقطه A و B استفاده گردد، ارتفاع درخت خوشبندی با استفاده از معیار Complete link برابر با k می‌باشد.

رابطه h و k به چه صورت است؟

(در صورت تساوی فاصله‌ها، اولویت با ادغام خوشبندی از چپ به راست است.)

$$k = \log_2(h+1) \quad (1)$$

$$k = h \quad (2)$$

$$k = 2^h \quad (3)$$

$$k = 2h \quad (4)$$

- ۴۰ فرض کنید X و Y دو بردار در فضای N بعدی هستند. تابع هسته $K(X, Y) = (X^T Y + 1)^2$ ضرب داخلی $X^T Y$ است. بردارهای X و Y را در فضای چند بعدی هسته محاسبه می‌نماید؟

(۱) بیانگر ضرب داخلی دو بردار X و Y است.

$$\frac{(N+1)(N+2)}{2} \quad (1)$$

$$\frac{N(N+1)}{2} \quad (2)$$

$$2N+1 \quad (3)$$

$$N^2 \quad (4)$$

- ۴۱ فرض کنید n نمونه x_1, x_2, \dots, x_n به طور مستقل از تابع توزیع زیر نمونه‌برداری شده‌اند. تخمین درست‌نمایی بیشینه را برای پارامتر θ بدست آورید؟

$$p(x; \theta) = \theta x^{(-\theta-1)} \text{ where } \theta > 1, x \geq 1$$

$$\max \left(1, \frac{\ln(\sum_{i=1}^n x_i)}{n} \right) \quad (1)$$

$$\max \left(1, \frac{n}{\ln(\sum_{i=1}^n x_i)} \right) \quad (2)$$

$$\max \left(1, \frac{n}{\sum_{i=1}^n \ln(x_i)} \right) \quad (3)$$

$$\max \left(1, \frac{\sum_{i=1}^n \ln(x_i)}{n} \right) \quad (4)$$

- ۴۲ فرض کنید $(N, x_i, i=1, \dots, N)$ مقادیر ویژگی مربوط به N نمونه از دسته w_1 با میانگین μ_1 و مقادیر ویژگی مربوط به N نمونه از دسته w_2 با میانگین μ_2 باشند. در انتخاب ویژگی به کمک آزمون فرض‌ها، فرض صفر H_0 چگونه تعریف می‌شود؟

$$\mu_1 - \mu_2 = 0 \quad (1)$$

$$\mu_1 - \mu_2 > 0 \quad (2)$$

$$\mu_1 - \mu_2 < 0 \quad (3)$$

(4) میانگین دسته‌ها تأثیری ندارد.

۴۳- در یک مسئله دسته‌بندی دو دسته‌ای، داده‌های دسته‌های ۱ و ۲ از توزیع گوسی با میانگین‌های به ترتیب

$$\sum^{-1} = \begin{bmatrix} 0,95 & -0,15 \\ -0,15 & 0,95 \end{bmatrix} \text{ نمونه‌برداری شده‌اند. حال } \mu_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ و } \mu_2 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

$$\text{نمونه } X = \begin{bmatrix} 1 \\ 2/2 \end{bmatrix} \text{ داده شده است و می‌خواهیم توسط دسته‌بند بیزی که با داده‌های بالا آموزش دیده‌اند}$$

دسته‌بندی کنیم. کدام مورد درست است؟

۱) فاصله نمونه X تا دسته C1 برابر ۲/۱۵ است و کمتر از فاصله این نمونه تا دسته C2 که برابر است با ۲/۴۱ است. بنابراین نمونه X برچسب دسته C1 را می‌گیرد.

۲) فاصله نمونه X تا دسته C1 برابر ۲/۴۱ است و بیشتر از فاصله این نمونه تا دسته C2 که برابر است با ۲/۱۵ است. بنابراین نمونه X برچسب دسته C2 را می‌گیرد.

۳) فاصله نمونه X تا دسته C1 برابر ۲/۹۵ است و کمتر از فاصله این نمونه تا دسته C2 که برابر است با ۳/۶۷ است. بنابراین نمونه X برچسب دسته C1 را می‌گیرد.

۴) فاصله نمونه X تا دسته C1 برابر ۳/۶۷ است و بیشتر از فاصله این نمونه تا دسته C2 که برابر است با ۲/۹۵ است. بنابراین نمونه X برچسب دسته C2 را می‌گیرد.

۴۴- ماتریس کوواریانس داده‌های ورودی برابر است با $\sum = \begin{bmatrix} 52 & 36 \\ 36 & 73 \end{bmatrix}$ ، می‌خواهیم با روش PCA این داده‌ها را به یک بعد کاهش دهیم. کدام مورد مولفه اصلی این داده‌ها را نشان می‌دهد؟

$$u_1 = \begin{bmatrix} -0/6 \\ -0/8 \end{bmatrix} \quad (1)$$

$$u_1 = \begin{bmatrix} -0/6 \\ 0/8 \end{bmatrix} \quad (2)$$

$$u_1 = \begin{bmatrix} 0/6 \\ -0/8 \end{bmatrix} \quad (3)$$

$$u_1 = \begin{bmatrix} 0/6 \\ 0/8 \end{bmatrix} \quad (4)$$

۴۵- محیطی به صورت زیر داریم، که دارای پنج حالت A, B, C, D, E است. محیط هنگامی که در حالت‌های A و E قرار می‌گیرد به ترتیب پاداش +۱ و +۱ را به عامل داده و فرایند تعام می‌شود. در بقیه حالت‌های محیط، عامل پاداش صفر را دریافت می‌کند. عامل در هر یک از حالت‌های محیط تنها یکی از دو کنش سمت چپ و سمت راست را انتخاب می‌کند. با فرض اینکه محیط قطعی است مقدار $\gamma = ۱$ است. کدام مورد در خصوص سیاست بهینه درست است؟

+1۰	۰	۰	۰	+1
A	B	C	D	E

۱) هنگامی که محیط در حالت B است، کنش بهینه سمت چپ است.

۲) هنگامی که محیط در حالت B است، کنش بهینه سمت راست است.

۳) هنگامی که محیط در حالت D است، کنش بهینه سمت راست است.

۴) هنگامی که محیط در حالت C است، کنش‌های سمت چپ و سمت راست دارای احتمال انتخاب یکسان هستند.