کد کنترل

528

A

528A

آزمون ورودی دورهٔ دکتری (نیمه‌متمرکز) ــ سال ۱۴۰۰

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

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

دفترچهٔ شمارهٔ (۱)
صبح جمعه
۹۹/۱۲/۱۵

رشتهٔ مهندسی کامپیوتر ــ هوش مصنوعی ــ (کد ۲۳۵۶)

تعداد سؤال: ۴۵ — مدت پاسخ‌گویی: ۱۵۰ دقیقه

عنوان مواد امتحانی، تعداد و شمارهٔ سؤالات

ردیف	مواد امتحانی	تعداد سؤال	از شمارهٔ	تا شمارهٔ
۱	مجموعه دروس تخصصی: ــ ساختمان داده‌ها و طراحی الگوریتم‌ها ــ شناسایی الگو ــ یادگیری ماشین	۴۵	۱	۴۵

این آزمون نمرهٔ منفی دارد.

استفاده از ماشین‌حساب مجاز نیست.

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

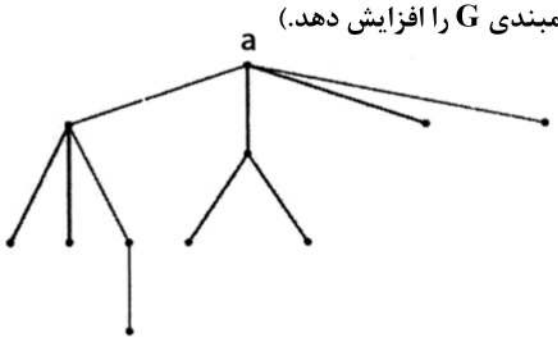
اینجانب ..... با شمارهٔ داوطلبی ..... با آگاهی کامل، یکسان‌بودن شمارهٔ صندلی خود را با شمارهٔ داوطلبی مندرج در بالای کارت ورود به جلسه، بالای پاسخ‌نامه و دفترچهٔ سؤالات، نوع و کد کنترل درج شده بر روی دفترچهٔ سؤالات و پائین پاسخ‌نامه‌ام را تأیید می‌نمایم.

امضا:

۱- **اگر به یک گراف جهت‌دار یک یال اضافه کنیم، تعداد اجزای قویا همبند در گراف چه مقدار تغییر می‌کند؟**

۱) حداکثر یک واحد کمتر می‌شود.
۲) حداکثر دو واحد کمتر می‌شود.
۳) ممکن است بیش از دو واحد کم شود.
۴) تغییر نمی‌کند یا ممکن است افزایش پیدا کند.

۲- **فرض کنید شکل زیر، درخت حاصل از اجرای DFS روی یک گراف همبند و بدون جهت G است. با فرض این‌که می‌دانیم گراف G دقیقاً یک رأس برشی دارد، کدام گزاره در مورد درجهٔ رأس a در این گراف صحیح است؟ (یک رأس گراف G برشی است، اگر حذف آن تعداد مؤلفه‌های همبندی G را افزایش دهد.)**



۱) درجهٔ رأس a در G هر عددی بین ۴ تا ۷ می‌تواند باشد.
۲) درجهٔ رأس a در G هر عددی بین ۸ و ۱۰ است.
۳) درجهٔ رأس a در G هر عددی بین ۹ و ۱۰ است.
۴) درجهٔ رأس a در G دقیقاً برابر ۴ است.

۳- **آرایه‌ای شامل n عدد متمایز داده شده است. می‌خواهیم از روی این اعداد یک درخت دودویی بسازیم، با این خاصیت که به ازای هر رأس درخت ساخته شده این خصوصیت را نیز داشته باشد که پیمایش میان ترتیب آن دقیقاً معادل ترتیب عناصر در آرایه شود. کدام گزاره درست است؟**

۱) چنین درختی لزوماً به ازای هر آرایه وجود دارد، اما یکتا نیست.
۲) چنین درختی لزوماً به ازای هر آرایه شامل n عدد متمایز وجود ندارد.
۳) بهترین زمان برای ساخت چنین درختی از روی یک آرایه $O(n)$ است.
۴) بهترین زمان برای ساخت چنین درختی از روی یک آرایه $O(n \log n)$ است.

۴- **فرض کنید می‌خواهیم n تومان را با استفاده از سکه‌های a و b و c تومانی خرد کنیم. به ازای چه تعداد از سه‌تایی‌های (a,b,c) زیر، الگوریتم حریصانه، n تومان را با کمترین تعداد سکه خرد می‌کند؟**

- **(۵,۲,۱)**
- **(۵,۴,۱)**
- **(۶,۳,۱)**

۱) ۰ ۲) ۱ ۳) ۲ ۴) ۳

۵- **آرایهٔ A حاصل ترکیب دو زیر آرایهٔ B و C است که B یک آرایهٔ صعودی و C یک آرایهٔ نزولی است. به‌عنوان نمونه، A می‌تواند به‌صورت [۲,۹ ,۷,۴,۶,۵,۸ ,۱] باشد، که در واقع ترکیب آرایهٔ صعودی [۲,۴,۶,۸] و آرایهٔ نزولی [۹,۷,۵,۱] است. این آرایه را در چه زمانی می‌توان مرتب کرد؟ (بهترین گزینه را انتخاب کنید.)**

۱) $O(n)$ ۲) $O(n \log n)$
۳) $O(n/\log n)$ ۴) $O(n \log\log n)$

۶- فرض کنید یک آرایهٔ مرتب از اعداد طبیعی به طول n داریم، که در آن هر عدد به غیر از یکی دقیقاً دو بار ظاهر شده است. عضو غیرتکراری را در چه زمانی می‌توان پیدا کرد؟

۱) $O(n\log n)$ ۲) $O(\log n)$

۳) $O(n)$ ۴) $O(1)$

۷- یک هرم کمینه شامل n عنصر داده شده است. سومین کوچک‌ترین عنصر این آرایه را در چه زمانی می‌توان پیدا کرد؟

۱) $O(1)$ ۲) $O(n)$

۳) $O(\log n)$ ۴) $O(n\log n)$

۸- اگر در الگوریتم هافمن نویسه‌ای بیش از $\frac{۲}{۵}$ کل متن تکرار شود، در آن صورت کد این نویسه چند بیت می‌تواند باشد؟

۱) هر عددی بین ۱ تا ۳ ۲) فقط ۱

۳) فقط ۲ ۴) ۱ یا ۲

۹- در یک درخت T با n گره، فرض کنید تعداد برگ‌ها B و تعداد فرزندان هر گره غیربرگ ۲ باشد. همچنین فرض کنید $E[T]$ و $I[T]$ به‌ترتیب مجموع عمق برگ‌ها و مجموع عمق عناصر غیربرگ T باشند. اگر $n = ۹۹۹۹$ باشد، کدام گزینه همیشه درست است؟

۱) $B = ۵۰۰۰۰$ ۲) $E[T] = ۹۹۹۹۹$

۳) $I[T] = ۹۹۹۹۸$ ۴) $E[T] - I[T] = ۱۰۰۰۰۰$

۱۰- چند تا از گزاره‌های زیر صحیح است؟

- اگر پس از اتمام الگوریتم بلمن فورد، به‌روزرسانی فاصله‌ها را ادامه دهیم و فاصلهٔ مربوط به یک رأس v باز هم به روز شود، v در یک دور منفی قرار دارد.
- اگر در جست‌وجوی عمق اول گراف جهت‌دار G تنها یک یال بازگشتی (back edge) e وجود داشته باشد، آنگاه یال e' به جز یال e وجود دارد که $G - e'$ بدون دور است.
- اگر در الگوریتم دایکسترا که روی یک DAG و از مبدأ اجرا شده، فقط برخی از یال‌های خروجی s وزن منفی داشته باشند، الگوریتم ممکن است به درستی فاصله‌ها را محاسبه نکند.

۱) ۰ ۲) ۱

۳) ۲ ۴) ۳

۱۱- در یک مجموعه از اعداد صحیح به اندازهٔ n، دنبال یک ۴ تایی‌هایی مثل (a,b,c,d) هستیم که $d = \sqrt{a^۲ + b^۲ + c^۲}$. این کار در چه زمانی قابل انجام است؟ (بهترین گزینه را انتخاب کنید.)

۱) $O(n^۳)$ ۲) $O(n^۴)$

۳) $O(n\log n)$ ۴) $O(n^۲\log n)$

۱۲- با فرض درهم سازی یکنواخت ساده، احتمال این‌که سه عضو متفاوت a و b و c به یک خانهٔ جدول درهم‌سازی نگاشت شوند برابر کدام گزینه است؟ (فرض کنید اندازهٔ جدول m است.)

۱) $۱/m$ ۲) $۱/m^۲$

۳) $۱/m^۳$ ۴) بستگی به مقادیر a و b و c دارد.

۱۳- خانوادهٔ $H = \{h_1, \ldots, h_k\}$ از توابع درهم‌ساز را در نظر بگیرید که $h_i : \{a,b,c,d\} \to \{0,1\}$، برای آن‌که این خانواده یک خانوادهٔ درهم‌ساز سراسری باشد، k حداقل چقدر باید باشد؟ (خانوادهٔ توابع H سراسری است اگر و فقط اگر به ازای هر دو مقدار u و v داشته باشیم. $Pr_{h \in H}[h(u) = h(v)] \leq \frac{1}{m}$ که m اندازهٔ جدول درهم‌سازی است.)

۱) ۱۶ ۲) ۴

۳) ۲ ۴) ۱

۱۴- آرایهٔ $A[1..n]$ از اعداد صحیح داده شده است. زیر دنبالهٔ متوالی $A[i..j]$ یک «بازهٔ مثبت» نامیده می‌شود، اگر جمع اعضای $A[i]$ تا $A[j]$ مثبت (بزرگ‌تر از ۰) باشد. می‌خواهیم کمترین تعداد بازه‌های مثبت که تمام اعداد مثبت آرایه را پوشش می‌دهد پیدا کنیم. اگر ورودی آرایهٔ زیر باشد، جواب کدام است؟

$A[1..15] = \langle 3, -5, 4, 1, -9, -8, 2, 3, 4, -10, 1, -2, -3, 6, -1 \rangle$

۱) ۵ ۲) ۴

۳) ۳ ۴) ۲

۱۵- یک گراف ۵ رأسی همبند و بدون جهت داریم که رأس‌های آن با شماره‌های ۱ تا ۵ شماره‌گذاری شده‌اند. فرض کنید از رأس ۱ الگوریتم BFS را اجرا می‌کنیم و تمام حالت‌هایی که BFS می‌تواند رئوس را ملاقات کند عبارتند از $\langle 1,2,3,4,5 \rangle$، $\langle 1,3,2,5,4 \rangle$، $\langle 1,3,2,4,5 \rangle$ و $\langle 1,2,3,5,4 \rangle$. حال اگر از رأس ۵ الگوریتم DFS را اجرا کنیم، کدام گزینه نمی‌تواند ترتیب ملاقات‌ها رئوس گراف باشد؟

۱) ۵,۴,۳,۱,۲ ۲) ۵,۴,۳,۲,۱

۳) ۵,۳,۴,۱,۲ ۴) ۵,۴,۱,۲,۳

۱۶- برنامه زیر چه کاری می‌کند و زمان اجرای آن کدام است؟

```
SS(A[0 . . n-1])
  If  n = 2 and A[0] > A[1] then
     Swap (A[0] , A[1])
else if n > 2
     m= ⌈2n/3⌉
     SS(A[0. . m-1])
     SS(A[n-m. .n-1])
     SS(A[0. . m-1])
```

۱) آرایهٔ A را مرتب می‌کند و زمان اجرای آن $\theta\left(n^{\log_{3/2} 3}\right)$ است.

۲) آرایهٔ A را مرتب می‌کند و زمان اجرای آن $\theta\left(n^{\log_{2/3} 3}\right)$ است.

۳) آرایهٔ A را لزوماً مرتب نمی‌کند اما زمان اجرای آن $\theta\left(n^{\log_{3/2} 3}\right)$ است.

۴) آرایهٔ A را لزوماً مرتب نمی‌کند اما زمان اجرای آن $\theta\left(n^{\log_{2/3} 3}\right)$ است.

۱۷- دنبالهٔ $\langle ۳,۱,۴,۱,۵,۹,۲,۶,۵,۳,۸,۹,۷,۹,۳,۲,۳, ۸,۴,۶,۲,۷,۹ \rangle$ را در نظر بگیرید. چند عضو متوالی این دنباله را می‌توان به‌صورت یک عدد تصور کرد. مثلاً سه عنصر متوالی ۵ و ۳ و ۸ را عدد ۵۳۸ تصور کرد. دو عدد به این شکل را مجزا گوییم، اگر هیچ‌یک از عناصر دنباله در ساخت هر دوی آن‌ها نقش نداشته باشند. حداکثر چند عدد مجزا به این شکل می‌توان ساخت که به‌ترتیب از چپ به راست صعودی باشند؟

۱) ۱۱ ۲) ۱۰
۳) ۹ ۴) ۸

۱۸- فرض کنید تابعی داریم که به‌عنوان ورودی دو دنباله گرفته و به‌عنوان خروجی طول بزرگ‌ترین زیر دنبالهٔ مشترک آن‌ها را برمی‌گرداند. با حداکثر یک بار فراخوانی این تابع به علاوهٔ هزینهٔ $O(n)$ چند مورد زیر را می‌توان محاسبه کرد؟

- محاسبهٔ طول بزرگ‌ترین زیر دنبالهٔ آینه‌ای یک دنباله
- تشخیص این‌که آیا یک دنباله زیر دنبالهٔ یک دنبالهٔ دیگر است.
- تشخیص این‌که آیا یک دنباله آینه‌ای است.

۱) ۰ ۲) ۱ ۳) ۲ ۴) ۳

۱۹- درخت فراگیر T از گراف وزن‌دار G یک درخت گلوگاهی است، اگر سنگین‌ترین یال آن در بین تمامی درخت‌های فراگیر G سبک‌ترین باشد، کدام گزینه درخصوص گزاره‌های زیر درست است؟

(الف) هر درخت گلوگاهی یک درخت فراگیر کمینه است.

(ب) هر درخت فراگیر کمینه یک درخت گلوگاهی است.

۱) (الف) درست ـ (ب) درست ۲) (الف) درست ـ (ب) نادرست
۳) (الف) نادرست ـ (ب) درست ۴) (الف) نادرست ـ (ب) نادرست

۲۰- گراف جهت‌دار G با وزن یال‌های صحیح و دو رأس خاص s و t از گراف داده شده است. فرض کنید شار بیشینه از s به t در گراف داده شده است. کدام گزینه درخصوص گزاره‌های زیر درست است؟

(الف) اگر ظرفیت یکی از یال‌های G یک واحد افزایش داده شود، شار بیشینه در گراف جدید در زمان خطی قابل محاسبه است.

(ب) اگر ظرفیت یکی از یال‌های G یک واحد کاهش داده شود، شار بیشینه در گراف جدید در زمان خطی قابل محاسبه است.

۱) (الف) درست ـ (ب) درست ۲) (الف) درست ـ (ب) نادرست
۳) (الف) نادرست ـ (ب) درست ۴) (الف) نادرست ـ (ب) نادرست

۲۱- کدام‌یک از روش‌های کاهش ابعاد زیر از تکنیک «کشف ساختار درونی با بعد کمتر» یا «intrinsic low-dimensional structure detection» برای کاهش ابعاد داده‌ها استفاده می‌نماید؟

۱) ISOMAP
۲) Locally Linear Embedding (LLE)
۳) Multidimensional Scaling (MDS)
۴) Principal Components Analysis (PCA)

۲۲- از معیار واگرایی (Divergence) برای اندازه‌گیری میزان جدایی‌پذیری دو کلاس می‌توان استفاده کرد. برای حالتی که توابع توزیع ویژگی x در دو کلاس w_i و w_j به‌صورت توابع گوسی یک متغیره $p(x|w_i) \sim N(\mu_i, \sigma_i^2)$ و $p(x|w_j) \sim N(\mu_j, \sigma_j^2)$ باشد، معیار واگرایی d_{ij} از کدام رابطه به‌دست می‌آید؟
نکته: معیار واگرایی d_{ij} توسط رابطه زیر محاسبه می‌شود.

$$d_{ij} = \int_{-\infty}^{\infty} (p(x|w_i) - p(x|w_j)) \ln\left(\frac{p(x+w_i)}{p(x|w_j)}\right) dx$$

۱) $$d_{ij} = \frac{1}{2}\left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_j^2} - 2\right) + \frac{1}{2}(\mu_i - \mu_j)^2\left(\frac{\sigma_j^2}{\sigma_i^2} + \frac{\sigma_i^2}{\sigma_j^2}\right)$$

۲) $$d_{ij} = \frac{1}{2}(\mu_i - \mu_j)^2\left(\frac{\sigma_j^2}{\sigma_i^2} + \frac{\sigma_i^2}{\sigma_j^2} - 2\right) + \frac{1}{2}\left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_j^2}\right)$$

۳) $$d_{ij} = \frac{1}{2}(\mu_i - \mu_j)^2\left(\frac{\sigma_j^2}{\sigma_i^2} + \frac{\sigma_i^2}{\sigma_j^2}\right) + \frac{1}{2}\left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_j^2} - 2\right)$$

۴) $$d_{ij} = \frac{1}{2}\left(\frac{\sigma_j^2}{\sigma_i^2} + \frac{\sigma_i^2}{\sigma_j^2} - 2\right) + \frac{1}{2}(\mu_i - \mu_j)^2\left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_j^2}\right)$$

۲۳- چنانچه مجموعه داده‌های زیر برای دسته‌بندی دودسته‌ای در اختیار باشند، نمونه آزمایشی $x = 5$ با استفاده از دسته‌بند بیز و با فرض توزیع گاوس برای هر دسته و دسته‌بند $NN-1$، به‌ترتیب از راست به چپ به کدام دسته تعلق خواهد داشت؟

$C_1 = \{1, 3\}$
$C_2 = \{6, 7, 8, 9, 10\}$

۱) یک و یک ۲) یک و دو ۳) دو و یک ۴) دو و دو

۲۴- اگر (x_1, x_2) به‌طور توأم دارای توزیع نرمال با میانگین $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ و ماتریس کوواریانس $\sum = \begin{bmatrix} \sigma_1^2 & \sigma \\ \sigma & \sigma_2^2 \end{bmatrix}$ باشد، کدام گزینه میانگین توزیع شرطی $P(x_1 | x_2 = a)$ را نشان می‌دهد؟

۱) $\mu_1 + \frac{\sigma^2}{\sigma_2}(a - \mu_2)$

۲) $\mu_1 + \frac{\sigma}{\sigma_2}(a - \mu_2)$

۳) $\mu_1 + \frac{\sigma}{\sigma_2^2}(a - \mu_2)$

۴) $\mu_1 + \frac{\sigma_1}{\sigma_2}(a - \mu_2)$

۲۵- در جدول زیر با فرض اینکه متغیرهای باینری X و Y مستقل باشند، مقادیر p_1 و p_2 مربوط به احتمال توأم $P(X,Y)$ کدام است؟

X	Y	P(X,Y)
+	+	$\frac{3}{5}$
+	−	$\frac{1}{5}$
−	+	$p_1 = ?$
−	−	$p_2 = ?$

۱) $p_1 = \frac{3}{20}$ و $p_2 = \frac{1}{20}$

۲) $p_1 = \frac{1}{15}$ و $p_2 = \frac{2}{15}$

۳) $p_1 = \frac{1}{20}$ و $p_2 = \frac{3}{20}$

۴) $p_1 = \frac{2}{15}$ و $p_2 = \frac{1}{15}$

۲۶- چهار نمونه که هر کدام دارای ۲ ویژگی هستند و به‌صورت ماتریس x نشان داده شده است.

$$x = \begin{bmatrix} 6 & -4 \\ -3 & 5 \\ -2 & 6 \\ 7 & -3 \end{bmatrix}$$

می‌خواهیم این داده‌ها را توسط الگوریتم PCA به یک بعد کاهش بدهیم، کدام بردار تبدیل می‌تواند برای این منظور به‌کار رود؟

۱) $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$

۲) $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^T$

۳) $\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$

۴) $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^T$

۲۷- یک مجموعه آموزشی داریم که نمونه‌های آن برچسب‌های ۲+ و ۱+ دارند و هر نمونه دارای m ویژگی با مقدار حقیقی است. برای هر مقدار توزیع گاوسی برای ویژگی در نظر می‌گیریم، پس $P(x_k \mid c_j) = N(\mu_{jk}, \sigma^2_{jk})$ که μ_{jk} میانگین ویژگی k برای دسته j و σ^2_{jk} واریانس ویژگی k برای دسته j است. به این دسته‌بند، دسته‌بند بیز ساده گاوسی می‌گویند، کدام گزینه در مورد این دسته‌بند نادرست است؟

۱) اگر $\sigma_{1k} = \sigma_{2k}$ باشد، دسته‌بند خطی است.

۲) قدرت تفکیک دسته‌بند خطی کمتر از قدرت تفکیک دسته‌بند بیز ساده گاوسی است.

۳) بدون هیچ شرطی روی σ_{1k} و σ_{2k}، دسته‌بند بیز ساده گاوسی، یک دسته‌بند خطی است.

۴) دسته‌بند بیز ساده گاوسی و logistic regression در حالت مجانبی وقتی‌که شرط $\sigma_{1k} = \sigma_{2k}$ برقرار باشد، یک مدل را می‌سازند.

۲۸- در رگرسیون چند جمله‌ای همراه با منظم‌سازی، با افزایش λ کدام‌یک از گزاره‌های زیر درست است؟

$$E = (\omega \mid x) = \frac{1}{2}\sum_{i=1}^{N}[y^t - g(x^t \mid \omega)]^2 + \lambda \sum_i \omega_i^2$$

۱) واریانس افزایش می‌یابد.

۲) بایاس افزایش می‌یابد.

۳) مدل دچار بیش برازش می‌شود.

۴) به داده‌های پرت اهمیت بیشتری داده می‌شود.

۲۹- آزمونی داده شده است که پیش‌بینی می‌کند یک شخص بیمار است $(+c)$ یا بیمار نیست $(-c)$. این آزمون توسط تابع زیر نشان داده می‌شود.

$$p(y=c+\mid x)=\begin{cases}0 & \text{if} \quad x<0\\ x & \text{if} \quad 0\le x<1\\ 1 & \text{if} \quad x\ge 1\end{cases}$$

همچنین می‌دانیم که هزینهٔ خطای FN سه برابر هزینه خطای FP است. ما می‌خواهیم مقدار x^* را پیدا کنیم که احتمال دسته‌بندی اشتباه و ریسک را کمینه نماید. در این خصوص کدام گزینه درست است؟ (مقدار سمت راست احتمال دسته‌بندی اشتباه و مقدار سمت چپ ریسک را نشان می‌دهد.)

۱) $\frac{1}{2}$ و $\frac{1}{2}$ ۲) $\frac{1}{2}$ و $\frac{1}{4}$

۳) $\frac{1}{2}$ و $\frac{1}{4}$ ۴) $\frac{1}{4}$ و $\frac{1}{4}$

۳۰- فرض کنید که x یک بردار تصادفی d-بعدی با میانگین μ و ماتریس کوواریانس Σ است. اگر $Y=Ax+b$، که A یک ماتریس $n\times d$ و b یک بردار n بعدی باشد، ماتریس کوواریانس Y کدام است؟

۱) $A\sum A^T$ ۲) $A\sum^{-1}A^T$

۳) $A^T\sum A$ ۴) $A^T\sum^{-1}A$

۳۱- مسألهٔ دسته‌بندی برای داده‌های دو دستهٔ $S=1$ و $S=2$ با احتمال پیشین برابر $p(S=1)=p(S=2)$ را در نظر بگیرید. ورودی دسته‌بند بردار $X=(x_1,x_2)^T$ با دو ویژگی نامنفی و مستقل از هم x_1 و x_2 است. در هر دسته ویژگی‌های x_1 و x_2 دارای تابع توزیع نمایی به‌صورت زیر هستند.

$$f_{x_k|S}(x_k\mid i)=\begin{cases}\lambda_{ik}e^{-\lambda_{ik}x_k} & , \quad x_k\ge 0\\ 0 & , \quad x_k<0\end{cases}$$

پارامترهای توزیع برای این دو دسته مشخص و به‌صورت زیر است.

$$S=1:\begin{cases}\lambda_{11}=1\\ \lambda_{12}=2\end{cases} \qquad S=2:\begin{cases}\lambda_{21}=2\\ \lambda_{22}=1\end{cases}$$

فرض کنید برای این مسأله، دسته‌بند بهینه خطی طراحی کرده‌ایم. احتمال دسته‌بندی صحیح به کمک دسته‌بند بهینه کدام است؟

۱) $\frac{1}{3}$ ۲) $\frac{1}{2}$ ۳) $\frac{2}{3}$ ۴) $\frac{1}{1+e}$

۳۲- کدام گزینه نادرست است؟

۱) بعد VC پرسپترون کوچک‌تر از بعد VC در SVM خطی است.

۲) احتمال رویداد بیش برازش (Overfitting)، برای مجموعه داده‌های کوچک، بیشتر است.

۳) اگر عمق یک درخت تصمیم به اندازهٔ طول بردار ویژگی باشد، نویز موجود در داده‌ها بیشتر تأثیر خواهد گذاشت.

۴) اگر داده آموزشی شامل نویز در برچسب داده‌ها باشد، روش ادغام Boosting صدمه بیشتری نسبت به روش Bagging خواهد داشت.

۳۳- تعداد همسایه‌ها (یعنی k) در روش نزدیک‌ترین همسایه (KNN)، چه اثری دارد؟

۱) افزایش k باعث کاهش هزینهٔ محاسباتی می‌شود.

۲) افزایش k باعث افزایش حساسیت به نویز ویژگی‌ها می‌شود.

۳) افزایش k باعث کاهش حساسیت به نویز دسته‌بندی می‌شود.

۴) کاهش k باعث کاهش تعداد ویژگی‌های مؤثر در دسته‌بندی می‌شود.

۳۴- کدام عبارت صحیح است؟

۱) Naïve Bayes همیشه یک دسته‌بند خطی است.

۲) یکی از مشکلات روش خوشه‌بندی سلسله مراتبی با معیار فاصله Single-link، مشکل Chain effect است.

۳) دسته بند A دقت ۹۰٪ روی دادهٔ آموزشی و ۷۵٪ روی دادهٔ آزمون و دسته‌بند B دقت ۷۸٪ روی هر دو دادهٔ آموزشی و آزمون دارد. می‌توان نتیجه گرفت که دسته‌بند A از B بهتر است، چون میانگین دقت آن بالاتر است.

۴) یک مجموعه دادهٔ خطی جدایی‌پذیر داریم و دو پرسپترون که یکی با روش نزول در امتداد گرادیان خطا و دیگری با قانون پرسپترون آموزش داده شده است. می‌توان گفت که هر دو پرسپترون دقت یکسانی بر روی مجموعه دادهٔ آموزشی و آزمون دارند.

۳۵- در یک دسته‌بند ماشین بردار پشتیبان با حاشیه سخت (Hard Margin SVM) در فضای دوبعدی $\mathbf{x}=(x_1,x_2)^T$ مرز جداساز آن به‌وسیلهٔ قاعده $4z_1+9z_2+4z_3+16z_4=0$ مشخص شده است و از تابع زیر برای نگاشت بردار ویژگی به فضای با ابعاد بالاتر استفاده می‌کند.

$$\mathbf{z}=(z_1,z_2,z_3,z_4)^T=\phi(\mathbf{x})=(x_1^2,x_2^2,x_1x_2,-x_1)^T$$

آیا نمونه‌های $\mathbf{x}=[+1,-1]^T$ و $\mathbf{y}=[+1,-1]^T$ به‌ترتیب می‌توانند بردارهای پشتیبان این دسته‌بند باشند؟

۱) خیر ـ خیر ۲) خیر ـ بلی

۳) بلی ـ خیر ۴) بلی ـ بلی

۳۶- فرض کنید در یک مجموعه داده، ستون TID نشان‌دهندهٔ کلید مربوط به نمونه‌ها باشد. (برای هر نمونه یک TID منحصر بفرد وجود داشته باشد.) اگر یک درخت تصمیم غیرباینری روی این داده، آموزش دهیم و متغیر TID را هم به‌عنوان یک ویژگی وارد الگوریتم یادگیری کنیم، در صورت استفاده از معیار بهره اطلاعاتی (Information Gain) برای انتخاب ویژگی، کدام گزینه صحیح است؟

۱) درخت حاصل دارای عمق پایین و مشکل کم برازش (Underfitting) خواهد بود.

۲) درخت حاصل دارای عمق بالا و مشکل بیش برازش (Overfitting) خواهد بود.

۳) متغیر TID دارای بهره اطلاعاتی بالا بوده و به‌عنوان ریشهٔ درخت انتخاب خواهد شد.

۴) در صورت مناسب بودن مابقی ویژگی‌ها، الگوریتم قادر به یادگیری درختی با قدرت تعمیم بالا خواهد بود.

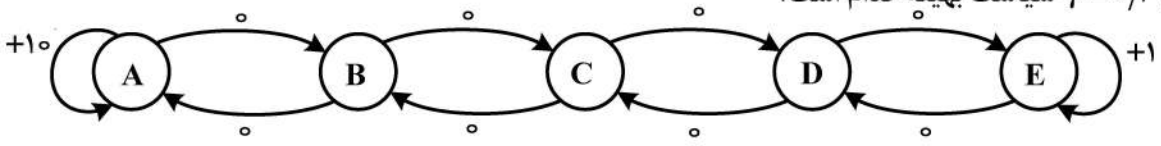
۳۷- بعد VC یک درخت تصمیم با سه گره دودویی (گره‌ها دارای دو فرزند هستند) در $\mathbf{R}^1$ کدام است؟

۱) ۵ ۲) ۴

۳) ۳ ۴) ۲

۳۸- یک عامل یادگیری تقویتی در محیط زیر که شامل ۵ حالت است، فعالیت می‌کند. اعداد روی کمان‌ها، پاداش را نشان می‌دهند. عامل تنها دو کنش چپ و راست که به‌صورت قطعی هستند را دارد.

برای $\gamma = ۰/۱$ سیاست بهینه کدام است؟



۱) $\pi^* = \{B : left, C : left\ or\ right, D : right\}$
۲) $\pi^* = \{B : left, C : right, D : right\}$
۳) $\pi^* = \{B : left, C : left, D : left\}$
۴) $\pi^* = \{B : left, C : left, D : right\}$

۳۹- یک کارخانه محصول **A** را تولید می‌کند. می‌خواهیم احتمال خراب بودن این محصول را تخمین بزنیم. برای این کار مجموعه‌ای از آزمایش‌ها را به‌صورت زیر انجام می‌دهیم. در هر آزمایش چندین محصول **A** را برداشته و بررسی می‌کنیم که آیا این محصول **A** سالم است یا خراب، تا به یک محصول **A** سالم برسیم. در هر آزمایش تعداد محصول **A** آزمایش شده خراب که با **k** نشان می‌دهیم را یادداشت می‌کنیم. فرض کنید که احتمال خرابی محصول **A** برابر **P** باشد و **m** آزمایش مستقل انجام دهیم و تعداد محصولات خراب $k_۱, k_۲, ..., k_m$ را یادداشت کنیم. تخمین **P** با استفاده از تخمین درست‌نمایی بیشینه کدام است؟

۱) $m + ۱ + \sum_{i=۱}^{m} K_i$
۲) $\frac{۱}{m + \sum_{i=۱}^{m} K_i}$
۳) $\frac{\sum_{l=۱}^{m} K_i}{m + ۱ + \sum_{i=۱}^{m} K_i}$
۴) $\frac{\sum_{l=۱}^{m} K_i}{m + \sum_{i=۱}^{m} K_i}$

۴۰- مجموعه داده‌های زیر را داریم.

$S = \{(۱,۰),(۲,۱)\}$

مؤلفهٔ اول هر زوج، بردار ویژگی **x** و مؤلفهٔ دوم برچسب آن است. اگر از یک دسته‌بند **logistic regression** با پارامترهای $\omega_۰ = -۴$ و $\omega_۱ = ۳$ استفاده کنیم، مقدار **likelihood** کدام است؟

۱) ۰/۲۷
۲) ۰/۶۴
۳) ۰/۷۳
۴) ۰/۸۸

۴۱- تابع توزیعی به‌صورت زیر داریم، اما مقدار ω را نداریم و می‌خواهیم با یک نمونه از توزیع مقدار آن (ω) را تخمین بزنیم. یک نمونه از توزیع گرفته می‌شود، که مقدار آن $x = ۳$ است. تخمین درست‌نمایی بیشینه ω برابر کدام گزینه است؟

$$P(x) = \begin{cases} ۰ & \text{if } x < ۰ \\ \frac{۲}{\omega} - \frac{۲x}{\omega_۲} & \text{if } ۰ \le x \le \omega \\ ۰ & \text{if } x > \omega \end{cases}$$

۱) ۶
۲) ۵
۳) ۴
۴) ۳

۴۲- فرض کنید فضای فرضیه (Hypothesis space) شامل درخت‌های تصمیم با عمق ۲ بر روی ۲ ویژگی است. همچنین فرض کنید هر داده شامل $n = ۱۰$ ویژگی دودویی باشد. چه تعداد داده (m) کافی است تا مطمئن باشیم که یک درخت تصمیم با مشخصه ذکر شده می‌یابیم که به احتمال حداقل ۰/۹۹ خطای آن کمتر از ۰/۰۵ است؟

۱) $m \geq ۷$ ۲) $m \geq ۱۰۰$

۳) $m \geq ۱۶۰$ ۴) $m \geq ۲۲۴$

۴۳- یک ژن به شکل دو الل A و a ظاهر می‌شود، به‌گونه‌ای که احتمال رخداد الل A در جمعیت برابر با θ است. بدین معنی که یک کپی تصادفی از این ژن به احتمال θ الل A و به احتمال $۱-\theta$ الل a است. یک ژنوتیپ از دو ژن تشکیل شده است و احتمال هر ژنوتیپ به‌صورت زیر است.

ژنوتیپ	AA	Aa	aa
احتمال	$\theta^۲$	$۲\theta(۱-\theta)$	$(۱-\theta)^۲$

فرض کنید که یک نمونه برداری تصادفی از جمعیت شامل $k_۱$ ژنوتیپ AA، $k_۲$ ژنوتیپ Aa و $k_۳$ ژنوتیپ aa است. تخمین بیشینه درست‌نمایی (MLE) برای θ کدام است؟

۱) $\hat{\theta} = \frac{۲k_۱ + k_۲}{۲k_۱ + ۲k_۲ + ۲k_۳}$ ۲) $\hat{\theta} = \frac{۲k_۱ + k_۲}{k_۱ + k_۲ + k_۳}$

۳) $\hat{\theta} = \frac{k_۱ + k_۲}{k_۱ + k_۲ + k_۳}$ ۴) $\hat{\theta} = \frac{k_۱}{k_۱ + k_۲ + k_۳}$

۴۴- در مورد روش‌های خوشه‌بندی کدام مورد درست است؟

۱) تابع هزینه روش k-means ممکن است در برخی از گام‌های الگوریتم افزایش یابد.

۲) الگوریتم EM برای پیدا کردن پارامترهای GMM، پارامترهایی را پیدا می‌کند که متناظر با بیشینه سراسری تابع (Likelihood) هستند.

۳) در روش EM ممکن است حالتی از خوشه‌بندی پیدا شود که مقدار درست‌نمایی (Likelihood) را بی‌نهایت کند، ولی خوشه‌بندی مطلوبی نباشد.

۴) اگر خوشه‌ها توزیع گاوسی متقارن (با ماتریس کوواریانس αI) باشند، جواب $k-means$ و EM روی GMM لزوماً یکی می‌شود.

۴۵- کدام‌یک از موارد زیر درست است؟

۱) در تشخیص داده‌های پرت، روش‌های پارامتری نسبت به روش‌های ناپارامتری حساسیت کمتری در برابر نویز دارند.

۲) تخمین‌زن $k-NN$ توانایی محاسبه تابع چگالی احتمال را دارد.

۳) در $k-NN$ با افزایش k، واریانس مدل افزایش می‌یابد.

۴) در $k-NN$ با کاهش k، بعد VC افزایش می‌یابد.