

330

F

نام:

نام خانوادگی:

محل امضا:



330F

صبح جمعه

۱۳۹۵/۱۲/۶

دفترچه شماره (۱)



«اگر دانشگاه اصلاح شود مملکت اصلاح می‌شود.»

امام خمینی (ره)

جمهوری اسلامی ایران

وزارت علوم، تحقیقات و فناوری

سازمان سنجش آموزش کشور

آزمون ورودی

دوره دکتری (نیمه‌متمرکز) داخل - سال ۱۳۹۶

رشته امتحانی مهندسی کامپیوتر - هوش مصنوعی (کد ۲۳۵۶)

مدت پاسخگویی: ۱۵۰ 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 داده‌ها - طراحی الگوریتم‌ها - شناسایی الگو - یادگیری ماشین)	۴۵	۱	۴۵

این آزمون نمره منفی دارد.

استفاده از ماشین حساب مجاز نیست.

اسفندماه - سال ۱۳۹۵

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 می‌باشد و با متغلبین برابر مقررات رفتار می‌شود.

ساختمان داده‌ها:

- ۱- کدام مورد، جواب رابطه بازگشتی $T(n) = T(\sqrt{n}) + O(\log n)$ است؟
- (۱) $O(\log n)$
 - (۲) $O(\log^2 n)$
 - (۳) $O(\sqrt{n})$
 - (۴) $O(n)$
- ۲- یک هرم کمینه با n عنصر متمایز داده شده است. می‌خواهیم به ازای عدد صحیح داده شده k ($k \leq \sqrt{n}$)، k آمین کوچک‌ترین عنصر را در این هرم پیدا کنیم (یعنی عددی که دقیقاً $k-1$ عنصر از آن کوچک‌تر هستند)، با چه مرتبه زمانی این کار امکان‌پذیر است؟ (بهترین گزینه را انتخاب کنید).
- (۱) $O(n)$
 - (۲) $O(\sqrt{n})$
 - (۳) $O(k \log n)$
 - (۴) $O(k \log k)$
- ۳- یک درخت دودویی جست‌وجو شامل n عنصر داده شده است. با فرض دانستن محل عنصر x در این درخت، کوچک‌ترین عنصر بزرگ‌تر از x را در چه زمانی می‌توان در درخت به دست آورد؟
- (فرض کنید تمام عناصر درخت متمایزند و درخت به صورت استاندارد و بدون هیچ گونه اطلاعات کمکی ذخیره شده است.)
- (۱) $O(\log n)$
 - (۲) $O(\log^2 n)$
 - (۳) $O(n)$
 - (۴) $O(1)$
- ۴- آرایه‌ای شامل n عدد صحیح داده شده است. به ازای $1 \leq i \leq j \leq n$ ، مقدار c_{ij} را برابر مجموع مقادیر قرارگرفته در بازه i تا j از این آرایه تعریف می‌کنیم. می‌خواهیم میانگین تمام c_{ij} ‌های ممکن در این آرایه را پیدا کنیم. با چه مرتبه زمانی این کار امکان‌پذیر است؟ (فرض کنید چهار عمل اصلی در $O(1)$ قابل انجام‌اند.)
- (۱) $O(n \log n)$
 - (۲) $O(n \log^2 n)$
 - (۳) $O(n^2)$
 - (۴) $O(n)$
- ۵- فرض کنید یک کاهش چندجمله‌ای از مسئله ۱ به مسئله ۲ داریم. کدام مورد، درست است؟
- (۱) اگر مسئله ۲ ان پی-سخت باشد، آن گاه مسئله ۱ ان پی-تمام است.
 - (۲) اگر مسئله ۱ ان پی-تمام باشد، آن گاه مسئله ۲ نیز ان پی-تمام است.
 - (۳) اگر مسئله ۱ ان پی-تمام باشد، آن گاه مسئله ۲ ان پی-سخت است.
 - (۴) اگر مسئله ۲ ان پی-سخت باشد، آن گاه مسئله ۱ نیز ان پی-سخت است.

- ۶- کدام مورد در خصوص الگوریتم دایکسترا درست است؟
- ۱) هزینه سرشکن به روزرسانی هر رأس $O(1)$ است.
 - ۲) هزینه سرشکن به روزرسانی هر رأس $O(n)$ است.
 - ۳) هزینه سرشکن به روزرسانی هر رأس $O(m/n)$ است.
 - ۴) فاصله هر رأس تا مبدأ در طول الگوریتم دقیقاً یک بار به روز می‌شود.
- ۷- کدام یک از توابع درهم‌سازی زیر یکنوا (uniform) است؟ (فرض کنید اندازه جدول درهم سازی k است.)
- ۱) $h(x) = kx \bmod (k-1)$
 - ۲) $h(x) = (k-1)x \bmod k$
 - ۳) $h(x) = x \bmod (k-1)$
 - ۴) $h(x) = x^2 \bmod k$
- ۸- چه تعداد از گزاره‌های زیر، درست است؟
- اگر وزن تمام یال‌های یک گراف با مقدار مثبت C جمع شود، کوتاه‌ترین مسیرها تغییر نمی‌کند.
 - اگر وزن تمام یال‌های یک گراف در مقدار مثبت C ضرب شود، کوتاه‌ترین مسیرها تغییر نمی‌کند.
 - اگر وزن تمام یال‌های یک گراف با مقدار منفی C جمع شود، کوتاه‌ترین مسیرها تغییر نمی‌کند.
 - اگر وزن تمام یال‌های یک گراف در مقدار منفی C ضرب شود، کوتاه‌ترین مسیرها تغییر نمی‌کند.
- ۱ (۱) ۲ (۲) ۳ (۳) ۴ (۴)
- ۹- میانگین ارتفاع درخت DFS بر روی یک گراف کامل با فرض آنکه راس شروع تصادفی انتخاب شده است از چه مرتبه‌ای است؟
- ۱) $O(1)$
 - ۲) $O(n)$
 - ۳) $O(\sqrt{n})$
 - ۴) $O(\log n)$
- ۱۰- شبکه‌ای متشکل از n رأس، دو رأس معین s و t داده شده است. فرض کنید ظرفیت تمام یال‌های شبکه نامتناهی است. به ازای یک شار f از s به t ، یالی که بیش‌ترین شار از آن عبور می‌کند را یال تنگنا و مقدار شار عبوری از آن یال را «تنگنای» شار f می‌نامیم. می‌خواهیم به ازای یک مقدار صحیح C داده شده، شاری با مقدار C را با کم‌ترین تنگنا از s به t منتقل کنیم. با چند بار استفاده از الگوریتم فورد-فالکرسن می‌توان این شار را به دست آورد؟ (بهترین گزینه را انتخاب کنید.)
- ۱) $O(n \log C)$
 - ۲) $O(\log C)$
 - ۳) $O(n)$
 - ۴) $O(1)$

طراحی الگوریتم‌ها:

۱۱- در مرتب‌سازی آرایه‌ای به طول N با الگوریتم‌های MergeSort, Randomized quicksort, Insertion sort

میزان استفاده از پشته فراخوانی (Callstack) به ترتیب از چه مرتبه‌ای است؟

$$(1) O(1), O(\log_p N), O(1) \quad (2) O(\log_p N), O(\log_p N), O(1)$$

$$(3) O(N \log_p N), O(N \log_p N), O(N^2) \quad (4) O(N \log_p N), O(N \log_p N), O(1)$$

۱۲- وزارت ارشاد قصد دارد یک کتاب داستان آموزنده را از زبان انگلیسی به زبان‌های رایج در ایران ترجمه و منتشر

نماید. هزینه ترجمه یک صفحه بین هر دو زبان به هزار تومان در جدول زیر داده شده است. اگر این کتاب صد

صفحه داشته باشد، کمترین هزینه ترجمه آن به همه زبان‌ها چند تومان است؟

	لری	عربی	کردی	ترکی	فارسی
انگلیسی	۸	۵	۸	۷	۵
فارسی	۱	۲	۱	۱	۰
ترکی	۳	۲	۲	۰	
کردی	۲	۵	۰		
عربی	۱۰	۰			

(۱) یک میلیون (۲) سه میلیون و سیصد هزار

(۳) پنج میلیون (۴) هشت میلیون

۱۳- آرایه‌ای به طول n داده شده که n توان درست ۲ است. الگوریتم زیر را در نظر بگیرید:

۱- لیست را به n/k زیر لیست k تایی تقسیم کنید. هر زیر لیست را با Insertion sort مرتب کنید.

۲- متغیر i را برابر ۲ قرار دهید.

۳- تا زمانی که $k \times i$ کوچکتر یا مساوی n است، عملیات زیر را تکرار کنید:

۱-۳- آرایه را به صورت قسمت‌های $i \times k$ در نظر بگیرید.

۲-۳- هر قسمت را از وسط به دو زیر لیست تقسیم کرده و آن‌ها را با هم ادغام merge کنید.

۳-۳- متغیر i را دو برابر کنید.

هزینه الگوریتم در بدترین حالت، کدام است؟

$$(1) \theta(n \cdot \log n) \quad (2) \theta\left(\frac{n}{k} \cdot \log\left(\frac{n}{k}\right)\right)$$

$$(3) \theta\left(nk + n \cdot \log\left(\frac{n}{k}\right)\right) \quad (4) \theta\left(nk + \left(\frac{n}{k}\right) \cdot \log\left(\frac{n}{k}\right)\right)$$

۱۴- پیمایش Preorder و Postorder یک درخت دودویی داده شده است. پیمایش inorder آن، کدام است؟

Preoder : fgbcda

Postorder : gedcabf

(۱) gfcdba (۲) gfecabd (۳) gfcdba (۴) نمی‌توان به دست آورد.

۱۵- استفاده از کدام داده ساختار، در مرتب‌سازی ادغامی (mergesort) به پیچیدگی $O(n \log n)$ منجر می‌شود؟

(i) لیست پیوندی یک طرفه، (ii) لیست پیوندی دو طرفه، (iii) آرایه

(۱) فقط ii (۲) فقط iii (۳) i و ii (۴) هر سه مورد

- ۱۶- در صورتی که یک آرایه مرتب‌شده (صعودی) داشته باشیم، کدام الگوریتم مرتب‌سازی بهترین عملکرد را دارد؟
 (۱) ادغامی (۲) درجی (۳) سریع (۴) هیپ
- ۱۷- فرض کنید که $2n+1$ عدد داریم و می‌دانیم که هر کدام از این اعداد دقیقاً دو بار آمده است به جز یک عدد. پیچیدگی زمانی الگوریتمی که عدد یکتا را تعیین کند چقدر است؟ فرض کنید اعمال رایج روی دو عدد در $O(1)$ انجام می‌شود.
 (۱) $O(n \log n)$
 (۲) $O(\log n)$
 (۳) $O(n^2)$
 (۴) $O(n)$
- ۱۸- مرتبه زمانی قطعه کد زیر، کدام است؟

```
for k = n Down to n-1000
{
  j = 1;
  while(j <= n)
  {
    j = j * 2;
    i = 0;
    b = 1;
    while(b == 1 and i < j)
    {
      if(i + j) % 2 == 0
        b = 0;
      i++;
    }
  }
}
```

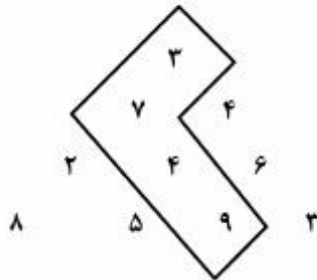
- (۱) $O(n^2)$ (۲) $O(\log n)$
 (۳) $O(n \log n)$ (۴) $O(n(\log n)^2)$

- ۱۹- کدام مورد، خروجی تابع زیر برای ورودی $f(2,5)$ است؟

```
int f(int n , int m)
{
  if(m < n)
    return 2 * m;
  else
    return (f(min(m,n), max(m,n) - 1) + (2 * n));
}
```

- (۱) ۲۰ (۲) ۲۶ (۳) ۴۴ (۴) خاتمه نمی‌یابد

۲۰- مثلثی از اعداد صحیح و مثبت در n ردیف به صورت زیر داده شده است. از رأس مثلث شروع کرده و در هر قدم به عدد مجاور در سطر پایین حرکت می‌کنیم. هدف پیدا کردن مسیری حداکثری از مجموع اعداد هم‌جوار است. (برای مثال در شکل زیر مسیر حداکثری به طول ۲۳ نشان داده شده است). هزینه زمانی بهترین الگوریتمی که می‌توان برای یافتن این مسیر حداکثری نوشت، کدام است؟



$$O(n^2 \log n) \quad (۱)$$

$$O(n \log n) \quad (۲)$$

$$O(n^2) \quad (۳)$$

$$O(2^n) \quad (۴)$$

شناسایی الگو:

۲۱- داده‌های زیر را در یک مساله دو کلاسه در نظر بگیرید:

$$C_1 = \{(1,1)^T, (1,2)^T, (2,1)^T, (2,4)^T, (3,1)^T, (3,3)^T\}$$

$$C_2 = \{(2,2)^T, (3,4)^T, (4,2)^T, (5,1)^T, (5,4)^T, (5,5)^T\}$$

با استفاده از جداساز خطی فیشر (FLD) جهت خطی را پیدا کنید، تا در صورتی که داده‌ها در آن جهت تابنده (project) شود، بیشترین جداسازی ممکن را داشته باشند؟

$$\begin{bmatrix} 8 \\ 12 \end{bmatrix} \quad (۱)$$

$$\begin{bmatrix} 37 \\ 6 \end{bmatrix} \quad (۲)$$

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix} \quad (۳)$$

$$\begin{bmatrix} 8 \\ 37 \end{bmatrix} \quad (۴)$$

۲۲- فرضیات زیر را در یک مساله کلاسه‌بندی دو کلاسه در نظر بگیرید:

$$x \in [0,1], P(x|C_1) = x + \frac{1}{2}, P(x|C_2) = \frac{3x^2}{4} + \frac{3}{4}, P(C_2) = 0.25$$

اگر ماتریس هزینه‌ها برابر $\lambda = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$ باشد، λ_{ij} نشان دهنده هزینه تصمیم اشتباه تعلق به کلاس C_i است، در صورتی که کلاس واقعی داده C_j باشد، کلاسه‌بندی را پیدا می‌کنیم که ریسک شرطی را کمینه کند. کدام مورد، ناحیه کلاس اول را نشان می‌دهد؟

(۱) $[0, 0.5]$

(۲) $[0, 1]$

(۳) $\left[\frac{8 - \sqrt{76}}{2}, 0 \right]$

(۴) $\left[1, \frac{8 + \sqrt{76}}{2} \right]$

۲۳- در یک مسأله طبقه‌بندی دو کلاسه ۱۰۰ نمونه آموزشی با ۲۵۰ ویژگی داده شده است. الگوریتم (LDA)

Linear Discrimination Analysis چند ویژگی را برای ساخت مرز تصمیم استخراج می‌کند؟

(۱) حداکثر ۲۵۰

(۲) بین ۱۰۰ تا ۲۵۰

(۳) حداقل ۱۰۰

(۴) حداکثر ۹۹

۲۴- مدل گرافیکی احتمالی زیر را در نظر بگیرید. چه تعداد از عبارتهای زیر درست است؟

$$P(D, H) = P(D)P(H)$$

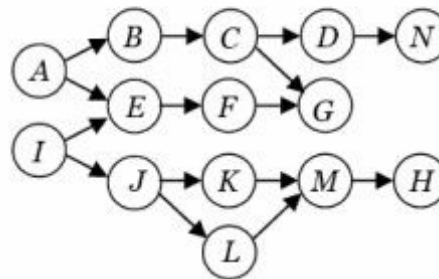
$$P(A, I) = P(A)P(I)$$

$$P(A, I|G) = P(A|G)P(I|G)$$

$$P(J, G|F) = P(J|F)P(G|F)$$

$$P(J, M|K, L) = P(J|K, L)P(M|K, L)$$

$$P(E, C|A, G) = P(E|A, G)P(C|A, G)$$



۱ (۱)

۲ (۲)

۳ (۳)

۴ (۴)

۲۵- کدام مورد، درست است؟

- (۱) روش ۱- نزدیک‌ترین همسایگی (1-NN) با استفاده از فاصله اقلیدسی در یک فضای ویژگی با ابعاد بالاتر که بوسیله نگاشت با یک کرنل گاوسی بدست آمده است، ممکن نیست نتایج بهتری را به نسبت فضای اولیه به دست دهد.
- (۲) مقدار حاشیه بدست آمده در دو ماشین بردار پشتیبان با کرنل‌های متفاوت $K_1(x_i, x_j)$ و $K_2(x_i, x_j)$ بر روی یک مجموعه داده آموزشی یکسان نشان دهنده میزان کارایی دو ماشین فوق بر روی مجموعه داده تست است.
- (۳) در مسائل کلاسه‌بندی، مرزهای با بالاترین حاشیه‌ای که ماشین‌های جداساز خطی می‌سازند، دارای کمترین خطای تعمیم (generalization) در بین تمامی کلاسه‌بندهای خطی دیگر هستند.
- (۴) حذف یکی از بردارهای پشتیبان از مجموعه داده‌ها ممکن است باعث شود که حاشیه ماشین بردار پشتیبان کاهش پیدا کند.

۲۶- با توجه به گراف زیر، چند مورد از عبارات درست است؟

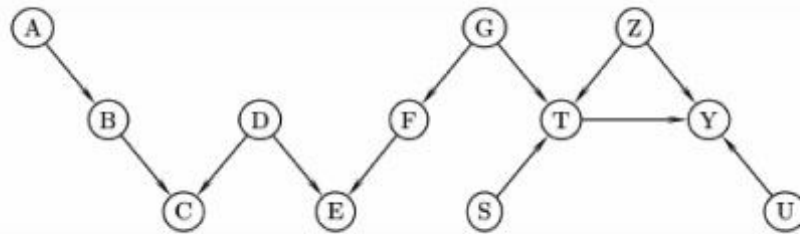
$$A \perp Y | CE$$

$$A \perp U$$

$$G \perp Y$$

$$G \perp Y | T$$

$$S \perp U | TY$$



(۱) ۱

(۲) ۲

(۳) ۳

(۴) ۴

- ۲۷- فرض کنید می‌خواهیم جهت بردار v را به قسمی پیدا کنیم، تا در صورتی که داده‌ها در آن جهت تصویر شوند $(y = v^T x)$ ، دارای بیشترین جداسازی ممکن باشد (نظیر روش جداسازی فیشر). در صورتی که μ_1 و σ_1^2 به ترتیب میانگین و پراکندگی داده‌های کلاس اول پس از تصویر شدن در جهت v و μ_2 و σ_2^2 به ترتیب میانگین و پراکندگی داده‌های کلاس دوم پس از تصویر شدن در همان جهت باشند، کدام مورد یک هدف مناسب برای این منظور می‌باشد، به قسمی که بیشینه‌سازی آن نسبت به بردار v ، جهت آن بردار را به ما بدهد؟ (دقت داشته باشید که در توابع زیر v به صورت ضمنی در پارامترهای میانگین و پراکندگی حضور دارد.)

$$J(v) = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 - \sigma_2^2} \quad (1)$$

$$J(v) = \frac{\sigma_1^2 + \sigma_2^2}{(\mu_1 - \mu_2)^2} \quad (2)$$

$$J(v) = \frac{\sigma_2^2 (\mu_1 - \mu_2)^2}{\sigma_1^2} \quad (3)$$

$$J(v) = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2} + \frac{(\mu_1 - \mu_2)^2}{\sigma_2^2} \quad (4)$$

۲۸- در یک مساله کلاسه‌بندی دو کلاسه در فضای دو بعدی، اگر داده‌های دو کلاس دارای احتمال‌های پیشین یکسان بوده و احتمال شرطی آنها دو مدل گاوسی با میانگین‌های غیر یکسان باشند، در چه صورت مرز حاصل از جداساز با کمترین نرخ خطا خطی خواهد بود؟

- (۱) تنها در صورتی که ماتریس‌های کوواریانس با هم برابر بوده و قطری باشند.
- (۲) تنها در صورتی که ماتریس کوواریانس‌های هر دو مدل گاوسی با هم برابر باشند.
- (۳) تنها در صورتی که ماتریس‌های کوواریانس هر دو مدل گاوسی، ماتریس همانی (I) باشد.
- (۴) تنها در صورتی که ماتریس‌های کوواریانس هر دو مدل گاوسی قطری (و نه لزوماً برابر) باشند.

۲۹- یک مساله k کلاسه را با مدل‌های احتمال شرطی یکنواخت $P(x|C_i) \sim U(-2^{i-2}, 2^{i-2})$ در نظر بگیرید. در صورت استفاده از کلاسه‌بند آماری بر پایه بیشینه شباهت (maximum likelihood ratio test)، مقدار $P(\text{error})$ وقتی k به سمت بی‌نهایت میل کند، برابر کدام مورد است؟

- (۱) ۰
- (۲) ۱
- (۳) $\frac{1}{2}$
- (۴) $\frac{1}{4}$

۳۰- در کدام مورد، هر دو روش discriminative هستند؟

- (۱) CRF و HMM
- (۲) Naïve Bayes و SVM
- (۳) Naïve Bayes و HMM
- (۴) logistic regression و SVM

یادگیری ماشین:

۳۱- دو سکه داریم که سکه یک با احتمال θ و سکه دوم با احتمال 2θ شیر می‌آید. این دو سکه را چندین بار پرت می‌کنیم و نتیجه به صورت زیر است. بیشینه درست‌نمایی (Maximum Likelihood) برای پارامتر θ کدام است؟

سکه	۱	۲	۲	۲	۲
نتیجه	شیر	خط	خط	خط	شیر

- (۱) ۰/۱
- (۲) ۰/۲
- (۳) ۰/۳
- (۴) ۰/۴

۳۲- یک مجموعه داده با اندازه m داریم که هر نمونه از d ویژگی دودویی تشکیل شده است. درخت تصمیمی که با استفاده از این مجموعه داده آموزش داده می‌شود، حداکثر دارای چند برگ است؟

- (۱) $\min(m, 2^d)$
- (۲) 2^d
- (۳) m
- (۴) 2^m

۳۳- سه نقطه $(۳,۳)$ ، $(۲,۲)$ و $(۱,۱)$ در فضای دو بعدی داریم. نخستین مولفه اصلی (Principal component) در این داده کدام است؟

$$(۱) (۰,۷۰۷, ۰,۷۰۷) \quad (۲) (۰,۶۶۶, ۰,۶۶۶)$$

$$(۳) (۱, ۱) \quad (۴) (۰,۵, ۰,۵)$$

۳۴- اگر $k(x,y)$ یک هسته معتبر باشد، کدام مورد درست است؟

$$(۱) k^T(x,y) \leq (x,x)k(y,y) \quad (۲) k^T(x,y) < (x,x)k(y,y)$$

$$(۳) k^T(x,y) \geq (x,x)k(y,y) \quad (۴) k^T(x,y) > (x,x)k(y,y)$$

۳۵- کدام مورد، در خصوص بعد VC درست نیست؟

(۱) برای مجموعه نقاط در فضای دو بعدی و فرضیه h که یک چند جمله‌ای درجه d است داریم $VC(h) > 0$

(۲) برای هر دو کلاس فرضیه H_1 و H_2 که شرط $H_1 \subseteq H_2$ برقرار باشد داریم $VC(H_1) \leq VC(H_2)$

(۳) برای هر دو کلاس فرضیه H_1 و H_2 که شرط $H_1 = H_2 \cup \{h_1, h_2, \dots, h_k\}$ برقرار باشد داریم

$$VC(H_1) \leq VC(H_2) + k$$

(۴) برای هر سه کلاس فرضیه H_1 و H_2 و H_3 که شرط $H_1 = H_2 \cup H_3$ برقرار باشد داریم

$$VC(H_1) \leq VC(H_2) + VC(H_3)$$

۳۶- یک متغیر تصادفی $t \in [0, \infty]$ دارای توزیع نمایی $p(x) = e^{-x}$ با پارامتر λ است. از این توزیع m نمونه مستقل

$X = (x_1, \dots, x_m)$ نمونه‌برداری می‌شود. $\ln p(X)$ به شرط داده شدن λ کدام است؟

$$(۱) \ln p(X) = \ln \lambda - \lambda \sum_{i=1}^m x_i \quad (۲) \ln p(X) = \ln \lambda - \lambda \prod_{i=1}^m x_i$$

$$(۳) \ln p(X) = m \ln \lambda - \lambda \prod_{i=1}^m x_i \quad (۴) \ln p(X) = m \ln \lambda - \lambda \sum_{i=1}^m x_i$$

۳۷- کدام مورد، در خصوص انتخاب یک زیر مجموعه bootstrap از یک مجموعه d - بعدی با m نمونه درست است؟

(۱) نمونه‌برداری p ویژگی با جایگذاری

(۲) نمونه‌برداری m نمونه با جایگذاری

(۳) نمونه‌برداری $k < p$ ویژگی بدون جایگذاری

(۴) نمونه‌برداری $k < m$ نمونه بدون جایگذاری

۳۸- کدام مورد، درست است؟

(۱) در روش K -NN هرچه تعداد نمونه‌ها بیشتر باشد دقت پیش‌بینی دسته‌بند K -NN بیشتر می‌شود.

(۲) هزینه محاسباتی آموزش روش K -NN بیشتر از هزینه محاسباتی آزمون این روش است.

(۳) روش K -NN برای رگرسیون مناسب نیست.

(۴) روش K -NN به داده پرت حساس نیست.

۳۹- کدام روش، بیشترین بایاس (bias) را دارد؟

(۱) K -Means

(۲) خوشه بندی طیفی

(۳) مخلوط گوسی (GMM) با هر ماتریس کوواریانس

(۴) مخلوط گوسی (GMM) با ماتریس کوواریانس همانی

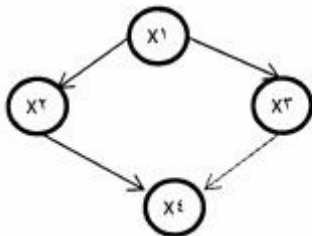
۴۰- در چه شرایطی SVD و PCA نتایج یکسانی تولید می کنند؟

- (۱) برای ماتریس کوواریانس همانی و هر بردار میانگینی پاسخ هر دو روش یکسان است.
- (۲) بزرگترین مقدار ماتریس کوواریانس یک باشد.
- (۳) ماتریس کوواریانس همانی باشد.
- (۴) میانگین داده ها صفر باشد.

۴۱- در رگرسیون خطی برای کدام توزیع پیشین وزن ها، تابع هدف روش منظم ساز L2 را نتیجه می دهد؟

- (۱) دو جمله ای
- (۲) پواسن
- (۳) گوسی
- (۴) بتا

۴۲- شبکه بیز زیر از چهار متغیر دودویی تشکیل شده است. برای محاسبه $P(X1, X2, X3, X4)$ به چند پارامتر مستقل نیاز است؟



- (۱) ۶
- (۲) ۸
- (۳) ۹
- (۴) ۱۵

۴۳- اگر $K1$ و $K2$ دو هسته معتبر باشند، کدامیک از هسته های زیر معتبر نیست؟

- (۱) $K1(x,y) + aK2(x,y)$ برای همه a های نامنفی
- (۲) $K1(x,y) + aK2(x,y)$ برای همه a ها
- (۳) $K1(x,y) + K2(x,y)$
- (۴) $K1(x,y) K2(x,y)$

۴۴- می خواهیم براساس مجموعه داده $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ که x یک عدد حقیقی است. تابع

$$J(w_0, w_1) = \sum_{i=1}^m (y_i - (w_0 + w_1 x_i))^2$$

هدف $y = w_0 + w_1 x$ براساس تابع هدف آموزش دهیم. متاسفانه داده ها

را نداریم اما اطلاعات زیر ذخیره شده است. کمترین اطلاعاتی که می توان براساس آنها w_1 را تخمین زد، کدام است؟

$C_{xx} = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})^2$	$C_{yy} = \frac{1}{m} \sum_{i=1}^m (y_i - \bar{y})^2$	$C_{xy} = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})$
---	---	--

(\bar{x} میانگین x_i ها و \bar{y} میانگین y_i ها است.)

(۱) C_{xy} و C_{yy}

(۲) C_{yy} و C_{xx}

(۳) C_{xy} و C_{xx}

(۴) C_{yy} و C_{xy} و C_{xx}

۴۵- کدام مورد، درست نیست؟

- (۱) هنگامی که ابعاد ویژگی بزرگتر می شود، احتمال بیش برآزش افزایش می یابد.
- (۲) هنگامی که فضای فرضیه ها غنی تر باشد، احتمال بیش برآزش افزایش می یابد.
- (۳) می توانیم از روش نزول در امتداد گرادیان برای یادگیری مدل مخلوط گوسی استفاده کنیم.
- (۴) اگر در مساله رگرسیون خطی با تابع هدف مجموع مربع خطا از روش نزول در امتداد گرادیان استفاده کنیم ممکن است در کمینه های محلی تابع هدف گرفتار شویم.

