



قدم به قدم، همراه دانشجو...

WWW.GhadamYar.Com

جامع ترین و به روز ترین پرتال دانشجویی کشور (پرتال دانش)
با ارائه خدمات رایگان، تحصیلی، آموزشی، رفاهی، شغلی و...
برای دانشجویان

- (۱) راهنمای ارتقاء تحصیلی. (کاردانی به کارشناسی، کارشناسی به ارشد و ارشد به دکتری)
- (۲) ارائه سوالات کنکور مقاطع مختلف سالهای گذشته، همراه پاسخ، به صورت رایگان
- (۳) معرفی روش‌های مقاله و پایان نامه نویسی و ارائه پکیج‌های آموزشی مربوطه
- (۴) معرفی منابع و کتب مرتبط با کنکورهای تحصیلی (کاردانی تا دکتری)
- (۵) معرفی آموزشگاه‌ها و مراکز مشاوره تحصیلی معتبر
- (۶) ارائه جزوایت و منابع رایگان مرتبط با رشته‌های تحصیلی
- (۷) راهنمای آزمون‌های حقوقی به همراه دفترچه سوالات سالهای گذشته (رایگان)
- (۸) راهنمای آزمون‌های نظام مهندسی به همراه دفترچه سوالات سالهای گذشته (رایگان)
- (۹) آخرین اخبار دانشجویی، در همه مقاطع، از خبرگزاری‌های پربازدید
- (۱۰) معرفی مراکز ورزشی، تفریحی و فروشگاه‌های دارای تخفیف دانشجویی
- (۱۱) معرفی همایش‌ها، کنفرانس‌ها و نمایشگاه‌های ویژه دانشجویی
- (۱۲) ارائه اطلاعات مربوط به بورسیه و تحصیل در خارج و معرفی شرکت‌های معتبر مربوطه
- (۱۳) معرفی مسائل و قوانین مربوط به سرگذری، معافیت تحصیلی و امریه
- (۱۴) ارائه خدمات خاص ویژه دانشجویان خارجی
- (۱۵) معرفی انواع بیمه‌های دانشجویی دارای تخفیف
- (۱۶) صفحه ویژه نقل و انتقالات دانشجویی
- (۱۷) صفحه ویژه ارائه شغل‌های پاره وقت، اخبار استخدامی
- (۱۸) معرفی خوابگاه‌های دانشجویی معتبر
- (۱۹) دانلود رایگان نرم افزار و اپلیکیشن‌های تحصیلی و...
- (۲۰) ارائه راهکارهای کارآفرینی، استارت آپ و...
- (۲۱) معرفی مراکز تایپ، ترجمه، پرینت، صحافی و ... به صورت آنلاین
- (۲۲) راهنمای خرید آنلاین ارزی و معرفی شرکت‌های مطرح (۲۳)



WWW.GhadamYar.Ir

WWW.PortaleDanesh.com

WWW.GhadamYar.Org

۰۹۱۲ ۳۰ ۹۰ ۱۰۸

باما همراه باشید...

۰۹۱۲ ۰۹ ۰۲ ۸۰۱

www.GhadamYar.com

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 (نیمه‌تمركز) داخل – سال ۱۳۹۶

روش امتحانی 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 (ساختمن داده‌ها – طراحی الگوریتم‌ها – شناسایی الگو – یادگیری ماشین)	۴۵	۱	۴۵

این آزمون نمره منفی دارد.

استفاده از ماشین حساب مجاز نیست.

اسقندماه – سال ۱۳۹۵

حق چاپ، تکثیر و انتشار سوالات به هر روش (الکترونیکی و ...) پس از برگزاری آزمون، برای تعامل اشخاص حلبی و حقوقی تنها با مجوز این سازمان مجاز می‌باشد و با متخلفین برابر مقررات رفتار می‌شود.

ساختمان داده‌ها:

- ۱ کدام مورد، جواب رابطه بازگشته $T(n) = T(\sqrt{n}) + O(\log n)$ 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 امین 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 جستجو شامل n عنصر داده شده است. با فرض داشتن 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 $i \leq j \leq n$ ، مقدار $\sum_{k=i}^j$ را برابر مجموع مقادیر قرارگرفته در بازه $[i, j]$ از این آرایه تعریف می‌کنیم. می‌خواهیم میانگین تمام $\binom{n}{2}$ های ممکن در این آرایه را پیدا کنیم. با چه مرتبه زمانی این کار امکان‌پذیر است؟ (فرض کنید چهار عمل اصلی در (۱) $O(1)$ 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 $O(\frac{m}{n})$ 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 درهم‌سازی k است.)

$$h(x) = kx \bmod (k - 1) \quad (1)$$

$$h(x) = (k - 1)x \bmod k \quad (2)$$

$$h(x) = x \bmod (k - 1) \quad (3)$$

$$h(x) = x^T \bmod k \quad (4)$$

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) (۴) ۲) (۲) ۳) (۳) ۴) (۱)

-۸ 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) $O(\log n)$

-۹ 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.)

$O(n \log C)$ (۱)

$O(\log C)$ (۲)

$O(n)$ (۳)

$O(1)$ (۴)

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؟

$$O(\log_2 N), O(\log_2 N), O(1) \quad (2) \quad O(1), O(\log_2 N), O(1) \quad (1)$$

$$O(N \log_2 N), O(N \log_2 N), O(1) \quad (4) \quad O(N \log_2 N), O(N \log_2 N), O(N^2) \quad (3)$$

-۱۲ وزارت ارشاد قصد دارد یک کتاب داستان آموزنده را از زبان انگلیسی به زبان‌های رایج در ایران ترجمه و منتشر نماید. هزینه ترجمه 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	۵	۷	۸	۵	۸
فارسی	۰	۱	۱	۲	۱
ترکی		۰	۲	۲	۲
کردی			۰	۵	۲
عربی				۰	۱۰

(۲) سه میلیون و سیصدهزار (۱) یک میلیون

(۴) هشت میلیون (۳) پنج میلیون

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

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

۲- متغیر A را برابر ۲ قرار دهید.

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

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

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

۳-۳- متغیر A را دو برابر کنید.

هزینه الگوریتم در بدترین حالت، کدام است؟

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

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

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

Preorder : fgbceda

Postorder : gedcabf

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

-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

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

- ۱۷ در صورتی که یک آرایه مرتب شده (صعودی) داشته باشیم، کدام الگوریتم مرتب سازی بهترین عملکرد را دارد؟

(۱) ادغامی (۲) درجی (۳) سریع (۴) هیپ

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 (۱) 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 = 5;
    while(j <= n)
    {
        j = j * r;
        i = o;
        b = 5;
        while(b == 1 and i < j)
        {
            if(i + j) / r == o
                b = o;
            i++;
        }
    }
}

```

```

int f(int n , int m)
{
    if(m < n)
        return  r×m;
    else
        return(f(min(m,n),max(m,n)-1)+(r×n));
}

```

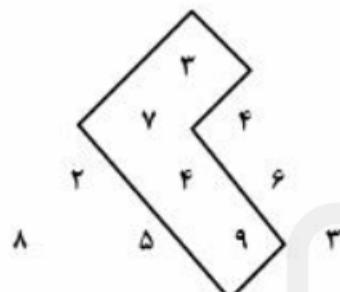
۴) خاتمه نمی‌باید

۷۴

۲۶ (۲)

10 (1)

- 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؟



$$O(n^r \log n) \quad (1)$$

$$O(n \log n) \quad (2)$$

$$O(n^r) \quad (3)$$

$$O(2^n) \quad (4)$$

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 تابانده شود، بیشترین جداسازی ممکن را داشته باشند؟ (project)

$$\begin{bmatrix} 8 \\ 12 \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} 37 \\ 6 \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} 8 \\ 37 \end{bmatrix} \quad (4)$$

قدم به قدم، همراه دانشجو...

WWW.GhadamYar.Ir

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

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

اگر ماتریس هزینه‌ها برابر $\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$ باشد، (i,j -نیان دهنده هزینه تصمیم اشتباہ 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) چند ویژگی را برای ساخت مرز تصمیم استخراج می‌کند؟

(۱) حداقل ۲۵۰

(۲) بین ۱۰۰ تا ۲۵۰

(۳) حداقل ۱۰۰

(۴) حداقل ۹۹

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

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

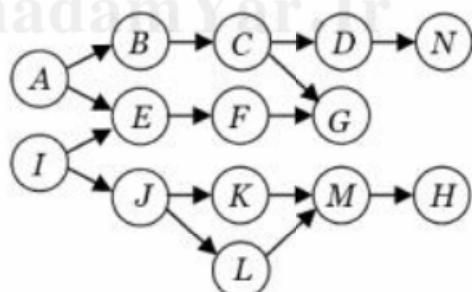
$$P(A, I) = P(A)P(I)$$

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

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

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

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



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 گاوی 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 مختلف $K_1(X_i, X_j)$ 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؟

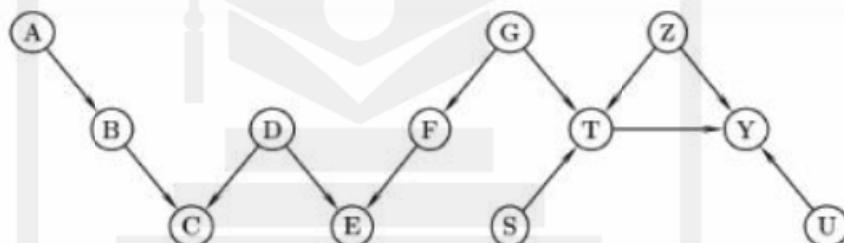
$A \perp Y | CE$

$A \perp U$

$G \perp Y$

$G \perp Y | T$

$S \perp U | TY$



- 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 μ_2 به ترتیب میانگین و پراکندگی داده‌های کلاس اول و کلاس دوم 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 گاووسی با میانگین‌های غیر یکسان باشند، در چه صورت مرز حاصل از جداساز با کمترین نرخ 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 گاووسی با هم برابر باشند.

۳) تنها در صورتی که ماتریس‌های کوواریانس هر دو مدل گاووسی، ماتریس همانی (I) باشد.

۴) تنها در صورتی که ماتریس‌های کوواریانس هر دو مدل گاووسی 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 وقتی k به سمت بینهایت میل کند، برابر کدام مورد است؟

۱)

۲)

$\frac{1}{2}$

$\frac{1}{4}$

- 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:

WWW.GhadamYar.Ir

- 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 درست نمایی (Maximum Likelihood) برای پارامتر θ کدام است؟

شیر	خط	خط	شیر	نتیجه	سکه
۲	۲	۲	۲	۱	۲

۱) ۰/۴

۲) ۰/۳

۳) ۰/۲

۴) ۰/۱

- 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، حداقل دارای چند برگ است؟

۱) 2^m

۲) m

۳) 2^d

۴) $\min(m, 2^d)$

- 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؟

$$(0,666, 0,666) \quad (2)$$

$$(0,707, 0,707) \quad (1)$$

$$(0,5, 0,5) \quad (4)$$

$$(1,1) \quad (3)$$

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

$$k^r(x,y) < (x,x)k(y,y) \quad (2)$$

$$k^r(x,y) \leq (x,x)k(y,y) \quad (1)$$

$$k^r(x,y) > (x,x)k(y,y) \quad (4)$$

$$k^r(x,y) \geq (x,x)k(y,y) \quad (3)$$

- 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 \circ

(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 $p(x) = e^{-x}/\lambda$ 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؟

$$\ln p(X) = \ln \lambda - \lambda \prod_{i=1}^m x_i \quad (2)$$

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

$$\ln p(X) = m \ln \lambda - \lambda \sum_{i=1}^m x_i \quad (4)$$

$$\ln p(X) = m \ln \lambda - \lambda \prod_{i=1}^m x_i \quad (3)$$

- ۳۷ کدام مورد، در خصوص انتخاب یک زیرمجموعه d -بعدی با m نمونه درست است؟

(۱) نمونه‌برداری p ویژگی با جایگذاری

(۲) نمونه‌برداری m نمونه با جایگذاری

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

(۴) نمونه‌برداری $k < m$ ویژگی 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؟

K-Means (1)

(۲) خوش‌بندی طیفی

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

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

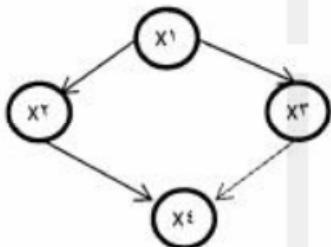
-۴۰ در چه شرایطی SVD و PCA نتایج یکسانی تولید می‌کنند؟

- (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 داده‌ها صفر باشد.

-۴۱ در رگرسیون خطی برای کدام توزیع پیشین وزن‌ها، تابع هدف روش منظم ساز L2 را نتیجه می‌دهد؟

- (۱) دو جمله‌ای
- (۲) پواسن
- (۳) بتا
- (۴) گوسی

-۴۲ شبکه بیز زیر از چهار متغیر دودویی تشکیل شده است. برای محاسبه $P(X_1, X_2, X_3, X_4)$ به چند پارامتر مستقل نیاز است؟



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

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

-۴۴ می‌خواهیم براساس مجموعه داده $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ که x یک عدد حقیقی است. تابع $y = w_0 + w_1 x$ براساس تابع هدف $J(w_0, w_1) = \sum_{i=1}^m (y_i - (w_0 + w_1 x_i))^2$ آموزش دهیم. متاسفانه داده‌ها را نداریم اما اطلاعات زیر ذخیره شده است. کمترین اطلاعاتی که می‌توان براساس آنها w_1 را تخمین زد، کدام است؟

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

(۱) 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}

-۴۵ کدام مورد، درست نیست؟

- (۱) هنگامی که ابعاد ویژگی بزرگتر می‌شود، احتمال بیش برآش افزایش می‌یابد.
- (۲) هنگامی که فضای فرضیه‌ها غنی‌تر باشد، احتمال بیش برآش افزایش می‌یابد.
- (۳) می‌توانیم از روش نزول در امتداد گرادیان برای یادگیری مدل مخلوط گوسی استفاده کنیم.
- (۴) اگر در مساله رگرسیون خطی با تابع هدف مجموع مرربع خطأ از روش نزول در امتداد گرادیان استفاده کنیم ممکن است در کمینه‌های محلی تابع هدف گرفتار شویم.



به اطلاع کاظلیان شرکت گشته در آزمون دکتری سال 1396 می‌رساند: در صورت تغایر می‌توانید حداکثر تاریخ 16/12/95 با مراجعت به سیستم پاسخگویی اینترنتی، نسبت به تکمیل فرم «اعتراف به کلید سوالات از مون» اقدام نمایید. لازم به ذکر است نظرات دوکاظلیان فقط از طریق سامانه پاسخگویی اینترنتی و فرم مذکور دریافت جواهد شد و به هوارد ارسالی از طرف دیگر و پس از تاریخ اعلام شده، به شیخ عنوان رسیدگی نخواهد شد.



گروه امتحان	نحوه پاسخگیری	نوع درجه	عنوان درجه
فندی و فنی مهندسی	1	F	مهندسی کامپیوتر-هوش مخصوصی

نحوه پاسخگیری سوال	گزینه صحیح	نحوه پاسخگیری سوال	گزینه صحیح
1	1	31	2
2	4	32	1
3	3	33	1
4	4	34	1
5	3	35	1
6	3	36	4
7	2	37	2
8	1	38	1
9	2	39	1
10	4	40	4
11	2	41	3
12	1	42	3
13	3	43	2
14	1	44	3
15	4	45	4
16	2		
17	4		
18	2		
19	2		
20	3		
21	2		
22	2		
23	4		
24	3		
25	1		
26	1		
27	4		
28	2		
29	3		
30	4		