

کد کنترل

708A



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

«علم و تحقیق، کلید پیشرفت کشور است.»
مقام معظم رهبری

صبح جمعه
۱۴۰۴/۱۱/۱۰
دفترچه شماره ۲ از ۲

آزمون ورودی دوره‌های دکتری (نیمه‌متمرکز) - سال ۱۴۰۵
مهندسی فناوری اطلاعات (کد ۲۳۵۸)

مدت زمان پاسخ‌گویی: ۱۳۵ دقیقه

تعداد سؤال: ۴۵ سؤال

عنوان مواد امتحانی، تعداد و شماره سؤال‌ها

ردیف	مواد امتحانی	تعداد سؤال	از شماره	تا شماره
۱	حل مسئله (ساختمان داده‌ها و طراحی الگوریتم‌ها) - شبکه (شبکه‌های کامپیوتری و امنیت شبکه)	۲۲	۱	۲۲
۲	مهندسی اطلاعات (پایگاه داده پیشرفته، بازیابی پیشرفته اطلاعات و داده‌کاوی) - خدمات فناوری اطلاعات (تجارت الکترونیک و آموزش الکترونیکی)	۲۳	۲۳	۴۵

استفاده از ماشین حساب مجاز نیست.

این آزمون نمره منفی دارد.

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 متخلفین برابر مقررات رفتار می‌شود.

* داوطلب گرامی، عدم درج مشخصات و امضا در مندرجات کادر زیر، به منزله عدم حضور شما در جلسه آزمون است.

اینجانب با شماره داوطلبی با آگاهی کامل، یکسان بودن شماره صندلی خود با شماره داوطلبی مندرج در بالای کارت ورود به جلسه، بالای پاسخنامه و دفترچه سؤالات، نوع و کدکنترل درج شده بر روی جلد دفترچه سؤالات و پایین پاسخنامه را تأیید می‌نمایم.

امضا:

حل مسئله (ساختمان داده‌ها و طراحی الگوریتم‌ها) - شبکه (شبکه‌های کامپیوتری و امنیت شبکه):

۱- توابع $f_1(n) = 2^{(\log n)^5}$ و $f_2(n) = n^{\Delta(\log n)^2}$ و $f_3(n) = \sqrt{2\sqrt{n}}$ و $f_4(n) = \sqrt{2}^{(\log n)^6}$ را در نظر بگیرید.

کدام یک از گزینه‌های زیر درست است؟

(۱) $f_2(n) = O(f_1(n)), f_1(n) = O(f_4(n)), f_4(n) = O(f_3(n))$

(۲) $f_2(n) = O(f_4(n)), f_4(n) = O(f_3(n)), f_3(n) = O(f_1(n))$

(۳) $f_4(n) = O(f_1(n)), f_1(n) = O(f_2(n)), f_2(n) = O(f_3(n))$

(۴) $f_4(n) = O(f_2(n)), f_2(n) = O(f_1(n)), f_1(n) = O(f_3(n))$

۲- رابطه بازگشتی زیر را در نظر بگیرید. مرتبه زمانی $T(n)$ کدام است؟

$$T(n) = 2T\left(\frac{n}{2}\right) + \frac{n}{\log n}$$

(۲) $\Theta(n)$

(۱) $\Theta\left(\frac{n}{\log n}\right)$

(۴) $\Theta(n \log n)$

(۳) $\Theta(n \log \log n)$

۳- فرض کنید $A[1, \dots, n]$ ، $B[1, \dots, n]$ دو آرایه خطی از اعداد دو به دو متمایز و هرکدام به صورت غیرنزولی مرتب باشند. سریع‌ترین الگوریتم برای یافتن میانه‌های تمامی این $2n$ عدد، از چه مرتبه زمانی است؟

(۲) $O(\sqrt{n})$

(۱) $O(\log^2 n)$

(۴) $O(1)$

(۳) $O(\log n)$

۴- یک S-Term ترتیبی از تعدادی نویسه S و پرانتز است که به صورت زیر تعریف می‌شود:
S یک S-Term است.

- اگر M و N دو S-Term باشند، الحاق آنها یعنی MN نیز یک S-Term است.

اگر b_n تعداد S-Term هایی باشد که با n عدد S قابل ساخت است، آنگاه b_n برابر با کدام است؟

(۲) $\sum_{K=1}^{n-1} b_K$

(۱) $\sum_{K=1}^{n-1} b_{2K}$

(۴) $\sum_{K=1}^{n-1} b_{2K} b_{n-2K}$

(۳) $\sum_{K=1}^{n-1} b_K b_{n-K}$

۵- فرض کنید T یک درخت **ranked tree** است که در آن، هر گره x دارای یک مقدار $\text{rank}(x)$ است. در این درخت، خصوصیت زیر برقرار است:

در هر مسیر ساده از یک گره به ریشه، rank ها صعودی هستند. یعنی اگر y والد x باشد، $\text{rank}(y)$ بزرگ‌تر از $\text{rank}(x)$ است. برای هر گره x ، $\text{level}(x)$ را به عنوان تعداد یال‌ها از x تا ریشه تعریف می‌کنیم. کدام گزاره صحیح است؟

(۱) برای تضمین صعود level ها، علاوه بر صعود rank باید مسیرهای درخت به صورت monotone و بدون هرگونه شاخه پیچیده باشند.

(۲) صعود rank ها در مسیر به ریشه، هیچ تضمینی برای صعود یا نزول level ها نمی‌دهد، و rank و level مستقل هستند.

(۳) level ها تنها در صورتی صعودی خواهند بود که درخت علاوه بر صعود rank ، خاصیت heap order نیز داشته باشد.

(۴) در هر مسیر از گره به ریشه، level ها همواره صعودی هستند (با افزایش rank ، سطح‌ها افزایش می‌یابند).

۶- فرض کنید $G = (V, E)$ یک گراف جهت‌دار قویاً همبند (**strongly connected**) است. یک **Euler tour** در G یک چرخه است که از هر یال دقیقاً یک بار عبور می‌کند، ولی ممکن است یک رأس چند بار بازدید شود. کدام یک از گزاره‌های زیر، صحیح و دقیق است؟

(۱) اگر هر رأس v در G دارای $\text{in-degree} \geq \text{out-degree}$ باشد، G حتماً **Euler tour** دارد.

(۲) وجود **Euler tour** در G تنها زمانی ممکن است که G یک گراف کامل (**complete**) جهت‌دار باشد.

(۳) G دارای تور اویلری است اگر و تنها اگر برای هر رأس گراف، درجه‌های ورودی و خروجی برابر باشند.

(۴) وجود **Euler tour** در G تنها وقتی تضمین می‌شود که G **acyclic** نباشد و حداقل یک رأس با in-degree بیشتر از out-degree داشته باشد.

۷- یک لیست پیوندی را در نظر بگیرید که در آن، به جای آنکه دو اشاره‌گر به گره‌های بعدی و قبلی داشته باشیم، تنها یک اشاره‌گر داریم که **XOR** اشاره‌گر قبلی و بعدی را در خود نگه می‌دارد. فرض کنید مقدار nil برابر با صفر است. همچنین فرض کنید ما به آدرس **head** و **tail** لیست دسترسی داریم و می‌خواهیم عملیات‌های زیر را روی لیست انجام دهیم.

الف - حذف یک گره **curr** در وسط فهرست

ب - برعکس کردن کامل فهرست در $O(1)$ زمان

کدام یک از گزینه‌های زیر، صحیح‌ترین روش‌ها برای انجام عملیات بالا هستند؟

(۱) برای برعکس کردن: کافی است اشاره‌گر **head** و **tail** را با هم جابه‌جا کنیم و برای حذف:

$$\text{prev.np} = \text{XOR}(\text{prev.np}, \text{curr}) \text{XOR next}$$

$$\text{next.np} = \text{XOR}(\text{next.np}, \text{curr}) \text{XOR prev}$$

(۲) برای حذف: نمی‌توان بدون پیمایش از ابتدای فهرست، گره میانی را حذف کرد و برای برعکس کردن: باید برای هر گره np را دوباره محاسبه کرد.

(۳) برای حذف: کافی است اشاره‌گر prev را به next تغییر دهیم و برای برعکس کردن: $\text{head} = \text{tail}$ و $\text{tail} = \text{head}$ کافی نیست.

(۴) برای برعکس کردن: پیمایش کامل فهرست و محاسبه دوباره np ها لازم است و برای حذف: $\text{curr.np} = \text{NIL}$

۸- گراف n رأسی و e یالی و بدون جهت G را تک - دوره می‌نامیم هرگاه همبند باشد و فقط یک دور داشته باشد. سریع‌ترین الگوریتم برای تشخیص تک - دوره بودن یک گراف، از چه مرتبه زمانی است؟

$$O(n^2) \quad (۲)$$

$$O(en) \quad (۱)$$

$$O(e+n) \quad (۴)$$

$$O(n) \quad (۳)$$

۹- فرض کنید می‌خواهیم با استفاده از الگوریتم **Backtracking**، تمام راه‌حل‌های مسئله **n-Queens** را پیدا کنیم. در مرحله‌ای از اجرای الگوریتم، در سطر **k** قرار داریم و در حال بررسی ستون‌های ممکن برای قرار دادن ملکه هستیم. فرض کنید تاکنون **k-1** ملکه در سطرها قبل به صورت معتبر قرار گرفته‌اند. در این حالت، کدام یک از موارد زیر تضمین می‌کند که با کمترین تعداد بازگشت (**backtrack**) به سطرها قبل، به اولین جواب ممکن برسیم؟

(۱) بررسی ستون‌ها به ترتیب تصادفی در هر سطر

(۲) بررسی ستون‌ها از چپ به راست بدون هیچ ترتیب خاصی

(۳) بازگشت بلافاصله پس از یافتن اولین تعارض در هر مسیر جستجو، بدون بررسی ادامه ستون‌ها در آن سطر

(۴) استفاده از **forward checking** برای حذف ستون‌هایی که با ملکه‌های قبلی در تعارض هستند، قبل از قرار دادن ملکه در سطر **k**

۱۰- چند مورد از گزاره‌های زیر درست هستند؟

- برای مسئله کوله‌پشتی کسری، یک الگوریتم حریصانه با مرتب‌سازی نسبت ارزش به وزن، همیشه درست‌ترین راه‌حل را ارائه می‌دهد و زمان اجرای آن $O(n \log n)$ است.

- برای مسئله کوله‌پشتی **0-1** الگوریتم حریصانه با انتخاب آیتم‌ها براساس نسبت ارزش به وزن نمی‌تواند تضمین کند که جواب بهینه را پیدا کند. برای یافتن جواب بهینه باید از برنامه‌ریزی پویا با زمان $O(nW)$ استفاده کرد.

- در مسئله **Matrix Chain Multiplication**، استفاده از الگوریتم حریصانه با انتخاب بهترین مکان تقسیم محلی نمی‌تواند تضمین کند که کمترین تعداد ضرب‌ها را پیدا کند. فقط الگوریتم **DP** با پیچیدگی $O(n^3)$ جواب بهینه را می‌دهد.

- الگوریتم **Prim** با **Fibonacci Heap** برای یافتن **MST** روی یک گراف با **n** رأس و **m** یال، همیشه زمان $O(m + n \log n)$ دارد و نسبت به پیاده‌سازی با **Binary Heap** برای هر گرافی سریع‌تر است.

- مسئله **Longest Common Subsequence (LCS)** بین رشته‌های طول **m** و **n** را می‌توان با الگوریتم **DP** در زمان $O(mn)$ و فضای $O(\min(m, n))$ حل کرد، اگر فقط طول **LCS** لازم باشد.

(۱) ۵ (۲) ۴

(۳) ۳ (۴) ۲

۱۱- کدام مورد درست است؟

(۱) اگر مسئله **A** را بتوان به یک مسئله شناخته‌شده **NP**-سخت کاهش داد، مسئله **A** نیز **NP**-سخت است.

(۲) از مرتب‌سازی هرمی می‌توان به‌عنوان روال کمکی مرتب‌سازی در مرتب‌سازی مینایی (**radix**) استفاده کرد.

(۳) زمان اجرای بدترین حالت الگوریتم دیکسترا دوطرفه (**Bidirectional Dijkstra**)، همانند الگوریتم دیکسترا است.

(۴) فرض کنید **P** کوتاه‌ترین مسیر از رأس **S** به رأس **t** در یک گراف وزن‌دار باشد. اگر وزن هر یال در گراف یک واحد افزایش یابد، **P** همچنان کوتاه‌ترین مسیر از **S** به **t** خواهد بود.

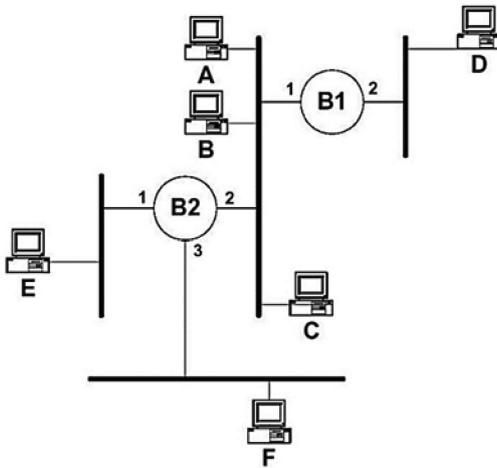
۱۲- فرض کنید در یک مسئله مینیمم‌سازی، تابع کران پایین (**LB**) طوری طراحی شده که در عمق‌های بالاتر، مقدار آن به‌طور تصادفی نوسان می‌کند. در درخت زیر، هر گره عدد **LB** خودش را دارد. اگر الگوریتم از روش **Best-First** استفاده کند و در گام اول Y_1 گسترش یابد ولی منجر به جواب با مقدار ۹ شود، کدام گره در گام بعد انتخاب می‌شود؟



۱۳- در شکل زیر، شبکه‌ای از پل‌های یادگیرنده (learning bridges) نشان داده شده است. فرض کنید در ابتدا جدول هدایت (forwarding table) همه bridgeها خالی است و سپس این ارسال‌ها، به ترتیب زیر، رخ می‌دهند:

- ۱- A به C می‌فرستد.
- ۲- E به F می‌فرستد.
- ۳- F به E می‌فرستد.
- ۴- D به B می‌فرستد.

پس از این چهار ارسال، جدول B2 چگونه خواهد بود؟



destination	port
C	۲
F	۳
E	۱

(۲)

destination	port
C	۲
F	۳
E	۱
D	۲

(۱)

destination	port
A	۲
C	۲
E	۱
F	۳
D	۲

(۴)

destination	port
C	۲
F	۱
E	۳
D	۲

(۳)

۱۴- پروتکل QUIC که اساس HTTP/3 را تشکیل می‌دهد، با هدف حل مشکل مسدود شدن در ابتدای صف (Head-of-Line Blocking) طراحی شده است. کدام گزینه به دقیق‌ترین شکل ممکن، مکانیزم اصلی QUIC

برای رفع این مشکل و یکی از پیامدهای کلیدی آن را توصیف می‌کند؟

(۱) QUIC با استفاده از قابلیت 0-RTT Handshake، زمان برقراری اتصال را به صفر کاهش می‌دهد. این امر باعث می‌شود که صف بسته‌ها در مبدأ و گره‌های میانی شبکه هرگز به اندازه‌ای طولانی نشود که منجر به از دست رفتن بسته و فعال شدن مکانیزم‌های بازبازی TCP مانند Fast Retransmit شود، که این خود، عامل اصلی HOL Blocking در لایه انتقال است.

(۲) QUIC با پیاده‌سازی مکانیزم کنترل ازدحام قابل‌اتصال (Pluggable Congestion Control) در فضای کاربری (User Space)، به سرور اجازه می‌دهد تا به‌صورت هوشمندانه، بسته‌های مربوط به جریان‌های (Streams) کم‌اهمیت‌تر را در زمان ازدحام حذف (drop) کند تا از مسدود شدن جریان‌های حیاتی مانند HTML جلوگیری کرده و بدین ترتیب HOL Blocking را مدیریت نماید.

(۳) QUIC با معرفی شناسه اتصال (Connection ID) مستقل از آدرس IP و پورت، مکانیزم مهاجرت اتصال (Connection Migration) را فراهم می‌کند. این ویژگی به QUIC اجازه می‌دهد تا در صورت تشخیص ازدحام شدید که منجر به از دست رفتن بسته‌ها و HOL Blocking می‌شود، به‌صورت پویا مسیر شبکه را تغییر دهد و بدین ترتیب از مسدود شدن جریان‌ها جلوگیری کند.

(۴) QUIC بر روی UDP اجرا شده و مفهوم جریان‌ها (Streams) را به‌عنوان شهروندان درجه اول (first-class citizens) در لایه انتقال خود بازآفرینی می‌کند. هر جریان دارای توالی و بافرینگ مستقل است. در نتیجه، از دست رفتن یک دیتاگرام UDP که حامل داده‌های یک جریان است، تنها همان جریان را متوقف می‌کند و پردازش دیتاگرام‌های بعدی که متعلق به جریان‌های دیگر هستند، ادامه می‌یابد و بدین ترتیب HOL Blocking لایه انتقال حذف می‌شود.

۱۵- در یک پروتکل Selective-Repeat ARQ، از شماره توالی (Sequence Number) به طول ۵ بیت استفاده می‌شود. برای جلوگیری از خطا در تشخیص بسته‌های جدید و تکراری، به ترتیب، حداکثر اندازه ممکن برای پنجره ارسال (Send Window) و پنجره دریافت (Receive Window) چقدر است؟

(۱) ۱ و ۳۱

(۲) ۸ و ۸

(۳) ۱۶ و ۱۶

(۴) ۳۲ و ۳۲

۱۶- دو گره به‌طور هم‌زمان شروع به ارسال یک بسته به طول L بیت بر روی یک کانال اشتراکی می‌کنند. نرخ ارسال کانال R بیت بر ثانیه و تأخیر انتشار سیگنال (propagation delay) بین دو گره t_{prop} ثانیه است. کدام گزاره در مورد تشخیص برخورد (Collision Detection) درست است؟

(۱) هر گره در صورت تداوم ارسال، در زمانی متناسب با t_{prop} (پس از شروع ارسال بسته) می‌تواند برخورد را تشخیص دهد. در نتیجه، اگر مدت زمان ارسال بسته (L/R) از مدت زمان تشخیص برخورد کمتر باشد، گره ارسال خود را تمام می‌کند قبل از آنکه برخورد را تشخیص دهد.

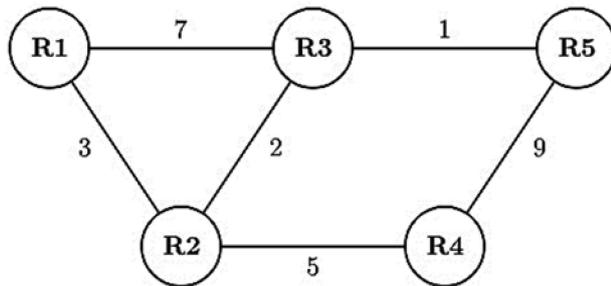
(۲) هر گره پس از زمان $L/R + t_{prop}$ (پس از شروع ارسال بسته) می‌تواند برخورد را تشخیص دهد. در نتیجه، اگر گره زودتر از این زمان ارسال خود را تمام کند، نمی‌تواند برخورد را تشخیص دهد.

(۳) برخورد در لحظه $t = 0$ رخ می‌دهد و هر دو گره بلافاصله آن را تشخیص می‌دهند.

(۴) اگر $L/R \geq 2 * t_{prop}$ باشد، هیچ برخوردی تشخیص داده نمی‌شود.

۱۷- در شکل زیر، شبکه‌ای شامل پنج روتر R1 تا R5 نشان داده شده است. هزینه هر لینک روی آن نوشته شده است (برای مثال، هزینه لینک بین R1 و R2 برابر ۳ است). فرض کنید پروتکل مسیریابی بردار فاصله (Distance Vector) بر پایه الگوریتم Bellman-Ford اجرا شده و شبکه در حالت پایدار است.

اکنون لینک بین R3 و R5 قطع می‌شود. بردار فاصله نهایی قبل از قطع لینک را به‌عنوان تکرار صفر (0th iteration) در نظر بگیرید. بدون استفاده از هیچ بهینه‌سازی (مانند poison reverse, split horizon و ...)، الگوریتم Bellman-Ford چند تکرار دیگر (غیر از تکرار صفر) لازم دارد تا به بردار فاصله جدید همگرا شود؟



- ۸ (۱)
- ۷ (۲)
- ۴ (۳)
- ۳ (۴)

۱۸- در کدام یک از موقعیت‌های زیر، از الگوریتم رمزنگاری نامتقارن (Asymmetric Key) استفاده می‌شود؟

- ۱) مدیر اداری قبل از ذخیره‌سازی فایل‌های محرمانه روی یک حافظه قابل حمل، آن‌ها را رمز می‌کند.
- ۲) دو روتر سیسکو با استفاده از CHAP، یکدیگر را احراز هویت می‌کنند.
- ۳) پس از برقراری یک VPN، داده‌های کاربر روی شبکه ارسال می‌شود.
- ۴) یک مدیر شبکه با استفاده از SSH به روتر سیسکو متصل می‌شود.

۱۹- در حمله SYN spoofing، هدف اصلی مهاجم کدام مورد است؟

- ۱) جدول اتصالات (connections table) پروتکل TCP
- ۲) اجرای حمله انکار سرویس برای سرقت آدرس IP
- ۳) سرویس DNS
- ۴) سرویس ایمیل

۲۰- مراحل زیر، نشان‌دهنده گام‌های مختلف رفتار یک مهاجم در طی یک نفوذ هستند:

- A: حفظ دسترسی با نصب backdoor و بدافزار Maintaining Access
- B: شناسایی بیشتر شبکه، حرکت جانبی و بهره‌برداری بیشتر Information Gathering/ System Exploit
- C: پوشاندن ردپا با دستکاری لاگ‌ها یا نصب rootkit Covering Tracks
- D: ارتقای سطح دسترسی از کاربر عادی به سطح Privilege Escalation Root/Admin
- E: به‌دست آوردن دسترسی اولیه (استفاده از آسیب‌پذیری راه دور) Intitial Access
- F: شناسایی هدف و جمع‌آوری اطلاعات Target acquisition and information gathering

کدام گزینه ترتیب صحیح این مراحل را از ابتدا تا انتهای فرایند نفوذ نشان می‌دهد؟

- ۱) F → E → D → B → A → C
- ۲) F → B → E → D → A → C
- ۳) E → F → D → B → A → C
- ۴) F → E → B → D → A → C

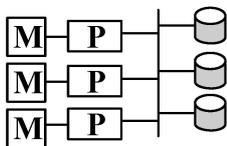
۲۱- چه نوع حمله‌ای براساس ارسال تعداد زیاد درخواست INVITE با آدرس‌های جعلی به سمت یک سرور انجام می‌شود؟

- ۱) SIP flood attack
- ۲) Smurf attack
- ۳) Slashdot attack
- ۴) Reflection attack

- ۲۲- فیلتر کردن نشست‌ها (Session filtering)، براساس متن و وضعیت درون یک نشست انجام می‌شود. برای این کار، فایروال نشست یا اتصال را نگه می‌دارد و چه عملی را انجام می‌دهد؟
- | | |
|-------------------------|-------------------|
| Virtual Switch (۱) | Traffic Block (۲) |
| Stateful inspection (۳) | DMZ re route (۴) |

مهندسی اطلاعات (پایگاه داده پیشرفته، بازیابی پیشرفته اطلاعات و داده‌کاوی) - خدمات فناوری اطلاعات (تجارت الکترونیک و آموزش الکترونیکی):

- ۲۳- نسبت OLAP به سیستم مدیریت انبار داده (data warehouse management system)، مانند کدام مورد است؟
- | | |
|-----------------------------|---------------------------------|
| Transaction Managements (۱) | به سیستم مدیریت پایگاه داده |
| OLTP (۲) | به سیستم مدیریت بازیابی اطلاعات |
| OLAP (۳) | به سیستم مدیریت بازیابی اطلاعات |
| OLTP (۴) | به سیستم مدیریت پایگاه داده |
- ۲۴- کدام مورد از اجزای اصلی مدل رابطه‌ای (Relational model) است؟
- | | |
|-----------|--------------------------|
| View (۱) | SQL (۲) |
| Index (۳) | Basic Set operations (۴) |
- ۲۵- مناسب‌ترین مدل برای ترکیب قابلیت‌های سیستم مدیریت پایگاه داده و زبان‌های برنامه‌سازی کدام است؟
- | | |
|--------------------------------|-------------|
| (۱) تابعی | (۲) شیء‌گرا |
| (۳) چندبعدی (multidimensional) | (۴) منطقی |
- ۲۶- کدام نوع ارتباط نسبت به بقیه، کلی‌تر است؟
- | | |
|-------------------|--------------------|
| Dependence (۱) | Inheritance (۲) |
| Participation (۳) | Generalization (۴) |
- ۲۷- شکل داده‌شده، نمایش کدام نوع معماری پایگاه داده است؟
- | |
|--|
| (۱) پایگاه داده موازی از نوع دیسک اشتراکی |
| (۲) پایگاه داده توزیع‌شده از نوع دیسک اشتراکی |
| (۳) پایگاه داده موازی از نوع غیراشتراکی (Shared nothing) |
| (۴) پایگاه داده توزیع‌شده از نوع غیراشتراکی (Shared nothing) |
- ۲۸- کدام مورد، نقش Smoothing را در مدل‌های زبانی برای بازیابی اطلاعات به‌درستی بیان می‌کند؟
- کاهش سهم واژگان کمتر مرتبط با موضوع غالب و اصلی سند با کاهش احتمال رخداد آنها
 - افزایش حساسیت مدل به فراوانی نسبی واژگان پُرسامد و بهبود رتبه‌بندی براساس آنها
 - جلوگیری از حذف اسناد کوتاه با افزایش وزن کلمات پُر تکرار در کل مجموعه
 - تقویت احتمال رخداد کلمات ناموجود در سند و محاسبه پایدارتر توزیع واژگان



۲۹- در پیکره‌ای شامل ۱۰۰۰ سند، واژه w در ۱۰ سند آمده است. همچنین در سند خاص d به طول ۱۰۰، واژه w پنج مرتبه

آمده است. مقدار تقریبی $TF-IDF(d,w)$ چند است؟ (فرض کنید $TF = \frac{n}{L}$ که در آن، L طول سند و n تعداد تکرار

یک واژه است. $idf = \log\left(\frac{N}{df}\right)$ است که در آن، N اندازه پیکره و df تعداد اسناد حاوی یک واژه خاص است.)

(۱) ۰٫۰۵ (۲) ۰٫۱

(۳) ۰٫۲۵ (۴) ۲

۳۰- در یک سیستم بازیابی اسناد، برای یک پرسش ۲۰ سند مرتبط وجود دارد. سیستم برای این پرسش، ۱۵ سند

برگردانده که از میان آنها ۱۲ سند واقعاً مرتبط هستند. Recall و Precision به ترتیب چه مقادیری دارند؟

(۱) ۰٫۷۵ و ۰٫۸ (۲) ۰٫۶ و ۰٫۷۵

(۳) ۰٫۶ و ۰٫۸ (۴) ۰٫۸ و ۰٫۶

۳۱- فرض کنید در مدل فیدبک روکیو (Rocchio)، پارامترهای $a = 1$ ، $b = 0.75$ ، $c = 0.25$ را داریم. اگر بردار پرسش

اولیه Q_0 ، یک سند مرتبط و D_n یک سند نامربوط باشند، بردار پرسش اصلاح شده Q_m چگونه محاسبه می‌شود؟

(۱) $Q_m = Q_0 + 0.75 D_r - 0.25 D_n$ (۲) $Q_m = Q_0 + 0.75 D_r$

(۳) $Q_m = 0.75 Q_0 + D_r$ (۴) $Q_m = Q_0 - 0.25 D_n$

۳۲- یک پایگاه داده با پنج تراکنش زیر داریم. معیار support برای مجموعه $\{A, C\}$ و معیار confidence برای

قانون $A \rightarrow C$ به ترتیب چه مقداری دارند؟

$T_1 : \{A, B, C\}$, $T_2 : \{A, C\}$, $T_3 : \{B, C\}$, $T_4 : \{A, B\}$, $T_5 : \{A, B, C\}$

(۱) ۱ و ۱

(۲) ۰٫۴ و ۰٫۴

(۳) ۰٫۶ و ۰٫۷۵

(۴) ۰٫۷۵ و ۰٫۲۵

۳۳- کدام مورد، بهترین تعریف داده‌کاوی است؟

(۱) کشف الگوهای پنهان و قابل استفاده از داده‌ها با استفاده از روش‌های محاسباتی

(۲) تحلیل و بازنمایی داده‌ها توسط کارشناس متخصص

(۳) طراحی مدل‌های آماری برای پیش‌بینی

(۴) استخراج داده از پایگاه داده‌های بزرگ

۳۴- عملیات طبقه‌بندی (Classification) و خوشه‌بندی (Clustering)، در کدام مورد زیر مشابه هستند؟

(۱) هر دو نظارت شده هستند.

(۲) هر دو نیاز به داده‌های کامل (غیرگمشده) دارند.

(۳) هر دو مبتنی بر یک معیار شباهت بین داده‌ها عمل می‌کنند.

(۴) هر دو را می‌توان به صورت یادگیری یک تابع تصمیم در فضای ویژگی مدل کرد.

۳۵- اگر برای مجموعه $A = \{10, 20, 3, 15, 1000, 50, 14\}$ ، بازه داده‌های معتبر را به صورت

$Median \pm 2 \text{Inter-quartile range}$ در نظر بگیریم: آنگاه این بازه کدام است؟

(۱) $[3, 50]$ (۲) $[-65, 95]$

(۳) $[10, 50]$ (۴) $[3, 1000]$

- ۳۶- رابطه میان اعتماد (Trust) و امنیت (Security) در تجارت الکترونیکی چیست؟
- (۱) اعتماد در فضای فیزیکی و امنیت در فضای مجازی معنا پیدا می کنند.
 - (۲) امنیت پیش نیاز اعتماد است، ولی تضمین کننده آن نیست.
 - (۳) اعتماد پیش نیاز و تضمین کننده امنیت است.
 - (۴) هر دو مترادف هستند.
- ۳۷- هدف اصلی گواهی نامه های دیجیتال (Digital Certificate) در تجارت الکترونیکی چیست؟
- (۱) تأیید هویت وبسایت ها برای کاربران
 - (۲) ارائه خدمات دولت الکترونیکی به کاربران
 - (۳) حذف نیاز به رمز عبور
 - (۴) فشرده سازی داده ها
- ۳۸- عدم تقارن اطلاعات در تجارت الکترونیکی، به چه مفهومی اشاره دارد؟
- (۱) نابرابری اطلاعات بین خریدار و فروشنده
 - (۲) عدم بهره مندی از فرایندهای استاندارد مدیریت اطلاعات
 - (۳) فقدان قوانین کافی جهت حفظ حریم شخصی و اطلاعات کاربران
 - (۴) شکاف اطلاعاتی و عدم دسترسی یکسان مردم در مناطق مختلف به فضای مجازی
- ۳۹- فراگیری (Ubiquity) در تجارت الکترونیکی به چه معناست؟
- (۱) دسترسی نامحدود به بازارهای بین المللی در هر ساعتی از شبانه روز
 - (۲) استفاده از ارز دیجیتال و سایر روش های پرداخت جدید
 - (۳) برخورداری از فرصت استفاده از واسطه های متعدد
 - (۴) امکان انجام تجارت در هر مکان و هر زمان
- ۴۰- نرخ همگرایی (Conversion Rate) در بازاریابی دیجیتال به چه معناست؟
- (۱) ضریب همگرایی بین بازدیدهای کاربران مختلف براساس ویژگی های دموگرافیک
 - (۲) نسبت بازدیدکنندگان جدید نسبت به بازدیدکنندگان قدیمی
 - (۳) درصد بازدیدکنندگانی که به مشتری تبدیل می شوند.
 - (۴) متوسط زمان حضور کاربران در سایت
- ۴۱- چرا یادگیری اکتشافی (Discovery Learning)، بدون هدایت معمولاً ناموفق است؟
- (۱) نیاز به فناوری های گران قیمت
 - (۲) افزایش بیش از حد بار شناختی
 - (۳) حذف فرصت تعامل
 - (۴) کاهش انگیزه
- ۴۲- طبق اصل پیوستگی زمانی، کدام مورد درست است؟
- (۱) ابتدا صدا سپس تصویر ارائه می شود.
 - (۲) پیوستگی مطالب در ارائه حفظ شود.
 - (۳) فاصله زمانی زیاد بین دو مفهوم، باعث درک بهتر می شود.
 - (۴) متن و تصویر باید همزمان ارائه شوند تا حافظه کاری، کمتر درگیر شود.
- ۴۳- انتقال دور «Far Transfer» در آموزش، به چه معنا است؟
- (۱) انتقال مهارت ها به موقعیت های جدید و متفاوت
 - (۲) انتقال یادگیری در موقعیت های مشابه و یکسان
 - (۳) یادگیری از راه دور و به صورت مجازی با استفاده از LMS
 - (۴) یادگیری تلفیقی حضوری و مجازی و کلاس های درس ترکیبی

۴۴- اصل فیدینگ (Fading) در آموزش با مثال‌های حل شده، به چه معنا است؟

(۱) استفاده از مثال‌های کاربردی و مهارتی در کنار آموزش

(۲) حذف تدریجی مثال‌ها و جایگزینی با تمرین

(۳) تلفیق مثال‌های حل شده در آموزش

(۴) افزودن تدریجی اطلاعات اضافی

۴۵- طبق اصل مدالیتی (Modality Principle)، در هنگام ارائه گرافیک‌های پیچیده باید کدام مورد رعایت شود؟

(۱) در کنار تصویر، هیچ توضیح اضافی قرار داده نشود تا انتقال مطلب به خوبی صورت پذیرد.

(۲) از متن روی صفحه، به جای گفتار استفاده شود.

(۳) از متن و گفتار، به صورت همزمان استفاده شود.

(۴) از گفتار برای توضیح تصویر استفاده شود.

