



دفترچه سوال رسمی آزمون
واحد سنتش و ارزیابی باشگاه دانش‌پژوهان جوان

بسمه تعالی
جمهوری اسلامی ایران
وزارت آموزش و پرورش
باشگاه دانش‌پژوهان جوان



مبارزه علمی برای جوانان، زنده کردن روح جست‌وجو و کشف واقعیت‌هاست. «امام خمینی (ره)»

دفترچه سؤالات مرحله اول سال ۱۴۰۴

دومین دوره المپیاد هوش مصنوعی

تعداد سؤالات	مدت آزمون
۲۰	۱۵۰ دقیقه

نام:	نام خانوادگی:	شماره صندلی:
------	---------------	--------------

استفاده از هر نوع ماشین حساب ممنوع است.

توضیحات مهم

- ۱- بلافاصله پس از آغاز آزمون، تعداد سؤالات داخل دفترچه و همه برگه‌های دفترچه سؤالات را بررسی نمایید، در صورت هرگونه نقصی در دفترچه، در اسرع وقت مسئول جلسه را مطلع کنید.
- ۲- یک برگ پاسخ‌برگ در اختیار شما قرار گرفته که مشخصات شما بر روی آن نوشته شده است، در صورت نادرست بودن آن، در اسرع وقت مسئول جلسه را مطلع کنید. ضمناً مشخصات خواسته شده در پایین پاسخ‌برگ را با مداد مشکی بنویسید.
- ۳- برگه پاسخ‌برگ را دستگاه تصحیح می‌کند، پس آن را تا نکنید و تمیز نگه دارید و به علاوه، پاسخ هر پرسش را با مداد مشکی نرم در محل مربوط علامت بزنید. لطفاً خانه مورد نظر را کاملاً سیاه کنید.
- ۴- دفترچه سوال باید همراه پاسخ‌برگ تحویل داده شود.
- ۵- هر پاسخ درست ۳ نمره مثبت و هر پاسخ نادرست ۱ نمره منفی دارد.

کلیه حقوق این سؤالات برای باشگاه دانش‌پژوهان جوان محفوظ است.
آدرس سایت اینترنتی: ysc.medu.gov.ir



این صفحه جهت استفاده به عنوان چرک نویس در نظر گرفته شده است

۱ یک کتابدار می‌خواهد سیستم طبقه‌بندی خود را ارتقا دهد. برای این کار او سه ویژگی داشتن منبع، حجم تصاویر و حجم متن کتاب‌ها را بررسی می‌کند. کتاب‌های زیر به کتابدار داده شده است:

نام کتاب	حجم تصاویر	حجم متن	داشتن منابع	دسته‌ی واقعی
کتاب «الف»	زیاد	کم	خیر	داستانی
کتاب «ب»	کم	زیاد	بله	علمی
کتاب «ج»	زیاد	زیاد	بله	علمی
کتاب «د»	کم	کم	خیر	داستانی
کتاب «ه»	کم	زیاد	خیر	علمی

کتابدار می‌خواهد تنها با پرسیدن یک سوال (بر اساس یکی از ویژگی‌ها)، کتاب‌ها را به دو دسته‌ی داستانی و علمی تقسیم کند. پرسیدن کدام سوال سبب می‌شود کتابدار بیشترین دقت را در دسته‌بندی کتاب‌ها داشته باشد؟

- ۱) آیا کتاب تصاویر زیادی دارد؟
- ۲) آیا کتاب متن زیادی دارد؟
- ۳) آیا کتاب دارای منابع است؟
- ۴) هیچکدام از سوالات نمی‌توانند دسته‌بندی بدون خطا ایجاد کنند.

۲ یک سامانه هوشمند، ایمیل‌های ورودی را به سه دسته تقسیم می‌کند: خبرنامه، شخصی و هرزنامه. بنابر آمارهای موجود، ۵۰ درصد ایمیل‌ها خبرنامه، ۳۰ درصد شخصی و ۲۰ درصد هرزنامه هستند. این سامانه بر اساس محتوای ایمیل‌ها، به آن‌ها برچسب «مشکوک» می‌زند. نرخ برچسب‌گذاری برای هر دسته به شرح زیر است:

- خبرنامه‌ها با احتمال ۲۰ درصد مشکوک شناسایی می‌شوند.
 - ایمیل‌های شخصی با احتمال ۱۰ درصد مشکوک شناسایی می‌شوند.
 - هرزنامه‌ها با احتمال ۶۰ درصد مشکوک شناسایی می‌شوند.
- یک ایمیل به صورت تصادفی انتخاب شده و مشاهده می‌شود که دارای برچسب «مشکوک» است.
۱. احتمال اینکه این ایمیل خبرنامه باشد چقدر است؟ (کسر چپ)
 ۲. با فرض اینکه بدانیم این ایمیل مشکوک، خبرنامه نیست، احتمال اینکه ایمیل شخصی باشد چقدر است؟ (کسر راست)

$$(۱) \frac{2}{5}, \frac{1}{4} \quad (۲) \frac{2}{5}, \frac{1}{5} \quad (۳) \frac{3}{5}, \frac{1}{4} \quad (۴) \frac{3}{5}, \frac{1}{5}$$

۳ در یک مرکز امنیت سایبری، سامانه‌ای به نام «شکارچی الگو» فعالیت می‌کند. این سامانه باید حملات را تشخیص دهد و در هنگام مواجهه با آن‌ها هشدار دهد. بدیهی است که هشدارهای نابجا در زمانی که حمله‌ای رخ نداده کارایی سامانه امنیت را پایین می‌آورد. آمار یک سال گذشته دو واقعیت را نشان می‌دهد:

۱. واقعیت اول (کیفیت): هشدارهای سامانه دقت بالایی دارند (احتمال اینکه هشدار درست باشد زیاد است).
۲. واقعیت دوم (کمیت): سامانه «محتاط» است؛ یعنی تعداد کل «هشدارها» کمتر از تعداد کل «حملات» است.

$$N(\text{هشدار}) < N(\text{حمله})$$

سه معیار ارزیابی تعریف شده است:

- معیار الف (درستی): نسبت هشدارهای درست به کل هشدارها.
- معیار ب (پوشش): نسبت هشدارهای درست به کل حملات.

• معیار ج (کارایی کل): میانگین همساز دو معیار «درستی» و «پوشش».

کدام گزاره لزوماً صحیح است؟

راهنمایی: میانگین همساز اعداد x_1, x_2 به صورت زیر تعریف می‌شود.

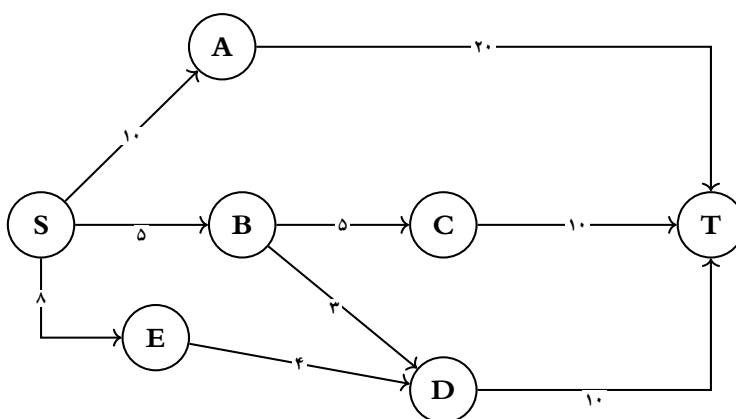
$$H = \frac{2}{\frac{1}{x_1} + \frac{1}{x_2}}$$

- گزاره ۱: مقدار «پوشش» از «درستی» بیشتر است.
- گزاره ۲: مقدار «کارایی کل» از «درستی» کمتر است.
- گزاره ۳: سامانه دارای نرخ بالایی از هشدارهای اشتباه است.

(۱) هیچ‌کدام (۲) فقط گزاره ۱ (۳) فقط گزاره ۲ (۴) گزاره‌های ۲ و ۳

۴ یک شبکه انتقال داده شامل ۴ مسیر مختلف بین مبدأ (S) و مقصد (T) است. ساختار شبکه به گونه‌ای است که برخی مسیرها دارای «یال‌های مشترک» هستند. اعداد نوشته شده روی هر یال، زمان عبور سیگنال (طول یال) را نشان می‌دهند. شرایط مسئله:

۱. برای عملکرد صحیح، سیگنال‌های هر ۴ مسیر باید دقیقاً در یک زمان واحد به مقصد (T) برسند.
 ۲. طولانی‌ترین مسیر شبکه (مسیر $S \rightarrow A \rightarrow T$) به عنوان «زمان مرجع» در نظر گرفته می‌شود و قابل تغییر نیست. سایر مسیرها باید با اضافه کردن «تأخیر» به یال‌ها، با این زمان هماهنگ شوند.
 ۳. شما می‌توانید به هر یالی در شبکه، مقدار عدد صحیح و غیرمنفی ($w \geq 0$) اضافه کنید.
 ۴. نکته کلیدی: اضافه کردن ۱ واحد تأخیر به یک «یال مشترک»، باعث می‌شود تمام مسیرهایی که از آن یال عبور می‌کنند، ۱ واحد به زمان مرجع نزدیک‌تر شوند (یک تیر و چند نشان).
- با فرض اینکه هدف ما کمینه کردن مجموع کل تأخیرهای اضافه شده به شبکه باشد، این مقدار حداقل چقدر خواهد بود؟



(۱) ۳۰ واحد (۲) ۲۲ واحد (۳) ۱۸ واحد (۴) ۱۲ واحد

۵ یک ربات نگهبان برای تشخیص وضعیت «خطر» از دو حسگر «صوتی» و «حرارتی» استفاده می‌کند. مهندسان با تحلیل داده‌های آماری، دو واقعیت زیر را درباره عملکرد این حسگرها کشف کرده‌اند:

۱. **واقعیت اول:** مشاهده ناهنجاری توسط حسگر صوتی، احتمال وجود خطر را نسبت به حالت عادی افزایش می‌دهد.

۲. **واقعیت دوم:** در مواقعی که خطر وجود دارد و حسگر صوتی فعال شده است، احتمال فعال شدن حسگر حرارتی، از «احتمال همزمانی فعال شدن دو حسگر (در حالت کلی)» بیشتر است.

با توجه به این دو واقعیت، وضعیت صحت گزاره‌های زیر را بررسی کنید:

• **گزاره ۱:** در صورت وقوع خطر، احتمال فعال شدن حسگر صوتی کمتر از احتمال فعال شدن آن در حالت کلی است.

• **گزاره ۲:** اگر هر دو حسگر فعال شوند، احتمال وقوع خطر بیشتر از «احتمال وقوع همزمان خطر و فعال شدن حسگر صوتی» است.

• **گزاره ۳:** اگر هر دو حسگر فعال شوند، احتمال وقوع خطر بیشتر از حاصل ضرب «احتمال وقوع خطر» در «احتمال فعال شدن حسگر صوتی» است.

چند مورد از گزاره‌های بالا لزوماً صحیح هستند؟

(۱) صفر (۲) یک (۳) دو (۴) سه

۶ یک پروتکل امنیتی جدید، «دنباله‌های سیگنال» با طول مشخص تولید می‌کند. این پروتکل برای ساخت دنباله‌ها، از سه نوع «بلوک پایه» استفاده می‌کند:

۱. **بلوک نوع A (نقطه):** سیگنالی که دقیقاً ۱ واحد زمان می‌برد.

۲. **بلوک نوع B (خط تیره):** سیگنالی که دقیقاً ۲ واحد زمان می‌برد.

۳. **بلوک نوع C (خط ممتد):** سیگنال دیگری که آن هم ۲ واحد زمان می‌برد.

توجه کنید که بلوک‌های B و C اگرچه طول یکسانی دارند، اما دو سیگنال متمایز محسوب می‌شوند. با استفاده از این سه نوع بلوک، به چند طریق مختلف می‌توان یک «دنباله سیگنال» با طول کل دقیقاً ۷ واحد ساخت؟

(۱) ۴۳ (۲) ۶۴ (۳) ۸۵ (۴) ۲۱

۷ در یک دشت کویری بزرگ، ۱۱ شهرک مسکونی روی یک خط مستقیم (جاده‌ی ترانزیت) واقع شده‌اند. موقعیت کیلومتری این شهرک‌ها نسبت به ابتدای جاده عبارت است از:

$$S = \{4, 5, 7, 9, 10, 12, 13, 14, 15, 18, 91\}$$

دولت قصد دارد یک «مرکز امداد رسانی» در یک نقطه‌ی بهینه P روی این جاده احداث کند. دو گروه کارشناسی، دو فرمول متفاوت برای محاسبه‌ی «هزینه‌ی عدم رضایت» پیشنهاد داده‌اند. هدف نهایی، پیدا کردن مکان P است به طوری که این هزینه کمینه شود.

• **پیشنهاد گروه اول (مدل خطی):** این گروه معتقد است که هزینه، مستقیماً با مسافت طی شده مرتبط است. بنابراین آن‌ها می‌خواهند مقدار زیر را کمینه کنند:

$$L_1 = \sum_{i=1}^{11} |x_i - P|$$

فرض کنید P_1 نقطه‌ای است که این هزینه را حداقل می‌کند.

- پیشنهاد گروه دوم (مدل توانی): این گروه معتقد است که هزینه با «توان دوم فاصله» مرتبط است و فواصل دورتر باید جریمه‌ی سنگین‌تری شوند. آن‌ها می‌خواهند مقدار زیر را کمینه کنند:

$$L_2 = \sum_{i=1}^{11} (x_i - P)^2$$

فرض کنید P_2 نقطه‌ای است که این هزینه را حداقل می‌کند.

فاصله‌ی بین دو مکان پیشنهادی بهینه $(|P_1 - P_2|)$ چند کیلومتر است؟

۵ (۱) ۶ (۲) ۱۲ (۳) ۱۸ (۴)

۸ یک محقق، الگوهای سفر گردشگران در تعطیلات نوروز را بین سه شهر بزرگ «تهران»، «اصفهان» و «شیراز» تحلیل می‌کند. فرض بر این است که این سامانه «بسته» است؛ یعنی هیچ گردشگری از خارج از این سه شهر وارد نمی‌شود و کسی هم از این مجموعه خارج نمی‌گردد. مدل محقق نشان می‌دهد که جابه‌جایی روزانه مسافران از الگوی زیر پیروی می‌کند:

مقصد ↓ / مبدأ ←	از تهران	از اصفهان	از شیراز
به تهران	۶۰٪	۱۰٪	۱۰٪
به اصفهان	۲۰٪	۷۰٪	۱۰٪
به شیراز	۲۰٪	۲۰٪	۸۰٪

توضیح: برای مثال، ستون اول می‌گوید که در پایان هر روز، ۲۰ درصد از کل جمعیتی که در تهران هستند، به شیراز می‌روند.

پس از گذشت مدتی، سیستم به یک «تعادل پویا» می‌رسد. در این وضعیت، مردم همچنان هر روز سفر می‌کنند و جابه‌جا می‌شوند، اما چون تعداد افرادی که وارد هر شهر می‌شوند دقیقاً با تعداد افرادی که از آن خارج می‌شوند برابر است، درصد کل جمعیت در هر شهر ثابت می‌ماند. در این وضعیت پایدار، چه درصدی از کل جمعیت توریست‌ها در شهر شیراز مستقر خواهند بود؟

۲۰٪ (۱) ۳۰٪ (۲) ۴۰٪ (۳) ۵۰٪ (۴)

۹ ما قصد داریم یک «سلول حافظه» با استفاده از یک سلول بازگشتی بسازیم. این سلول در هر لحظه، علاوه بر ورودی یا فرمان جدید (x) ، خروجی لحظه‌ی قبل خود (y_{old}) را نیز دریافت می‌کند. برای محاسبه‌ی خروجی جدید (y_{new}) ، ابتدا مقدار «سیگنال کل» (S) را محاسبه می‌کنیم:

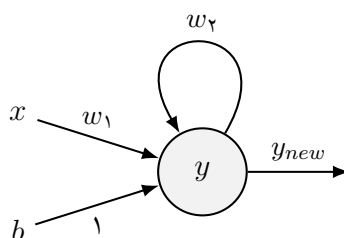
$$S = w_1 \cdot x + w_2 \cdot y_{old} + b$$

سپس خروجی نهایی بر اساس علامت S تعیین می‌شود:

$$y_{new} = 1 \iff S \geq 0$$

$$y_{new} = 0 \iff S < 0$$

که در آن w_1 وزن ورودی، w_2 وزن حافظه و b یک عدد ثابت است که روی تمایل این سلول به دادن خروجی ۰ و ۱ اثر می‌گذارد.



می‌خواهیم این سلول رفتار یک «کلید حافظه‌دار» را داشته باشد، به این معنا که:

۱. فرمان روشن ($x = +1$): خروجی حتماً ۱ شود (مستقل از حالت قبلی).
 ۲. فرمان خاموش ($x = -1$): خروجی حتماً ۰ شود (مستقل از حالت قبلی).
 ۳. فرمان حفظ حالت ($x = 0$): خروجی تغییر نکند (اگر قبلاً ۱ بوده، ۱ بماند و اگر ۰ بوده، ۰ بماند).
- اگر فرض کنیم پارامترهای w_1, w_2, b همگی اعداد صحیحی از مجموعه‌ی $\{-2, -1, 0, 1, 2\}$ باشند، چند دسته‌ی سه‌تایی متمایز (w_1, w_2, b) وجود دارد که این رفتار را دقیقاً پیاده‌سازی کند؟

(۱) ۳ حالت (۲) ۴ حالت (۳) ۵ حالت (۴) ۶ حالت

۱۰ ما در خانه‌ی $(0, 0)$ یک مربع شطرنجی 6×6 قرار داریم. محور افقی و عمودی از ۰ تا ۵ شماره‌گذاری شده‌اند. در هر گام می‌توانیم بوسیله‌ی یکی از دو بردار $(2, 3)$ یا $(4, 0)$ بین خانه‌های جدول حرکت کنیم. در صورتی که بر اثر یک حرکت از جدول بیرون بزنیم به اندازه‌ی ۶ واحد به مولفه‌ی اول یا دوم مختصات اضافه یا کم می‌کنیم تا درون جدول قرار بگیریم.

مثال: فرض کنید در خانه‌ی $(5, 0)$ قرار داشته باشیم. در این صورت بوسیله‌ی بردارها می‌توانیم به خانه‌های $(3, -3), (7, 3), (1, 0), (9, 0)$ برویم. از بین این خانه‌ها $(9, 0), (3, -3), (7, 3)$ خانه‌هایی خارج از جدول هستند. برای این که درون جدول قرار بگیرند:

$$\begin{aligned}(7, 3) - (6, 0) &= (1, 3) \\ (3, -3) + (0, 6) &= (3, 3) \\ (9, 0) - (6, 0) &= (3, 0)\end{aligned}$$

بنابراین می‌توانیم از $(5, 0)$ به یکی از چهارخانه‌ی $\{(1, 0), (1, 3), (3, 3), (3, 0)\}$ برویم. حال با فرض شروع کردن از خانه‌ی پایین چپ جدول $(0, 0)$ تعداد خانه‌هایی که می‌توانیم در آن باشیم چندتا است؟

(۱) ۴ (۲) ۶ (۳) ۸ (۴) ۹

۱۱ واحد کنترل یک بازوی رباتیک، وضعیت پایداری خود را با سه پارامتر طبیعی x, y, z در قالب یک «ماتریس وضعیت» متقارن به صورت $M = \begin{bmatrix} x & z \\ z & y \end{bmatrix}$ تنظیم می‌کند. برای اینکه سیستم در حالت «تعادل پایدار» قرار گیرد، ماتریس M باید دو ویژگی ساختاری زیر را (که وابسته به ثابت‌های طبیعی a و b هستند) داشته باشد:

۱. ثابت بودن اثر ماتریس: مجموع عناصر قطر اصلی برابر با مقدار ثابت a باشد. $(x + y = a)$.
۲. ثابت بودن دترمینان: دترمینان ماتریس برابر با مقدار ثابت b باشد. $(xy - z^2 = b)$.

می‌دانیم مقادیر a و b به گونه‌ای تنظیم شده‌اند که دقیقاً یک ماتریس وضعیت پایدار با درایه‌های طبیعی وجود دارد. در این صورت، کدام گزاره در مورد عدد a لزوماً صحیح است؟

(۱) a فرد است. (۲) a زوج است. (۳) a مربع کامل است. (۴) $a > b$ است.

۱۲ سیستم هوش مصنوعی «سینما- مایند» (Cinemind) سلیقه هر کاربر را به صورت یک بردار در فضای سه بعدی اکشن، درام و علمی-تخیلی مدل می‌کند. در این سیستم برای پیشنهاد فیلم، از معیاری به نام «شاخص هم‌راستایی» استفاده می‌شود که اگر بردار سلیقه کاربر U و بردار ویژگی‌های فیلم V باشد، سیستم فیلمی را پیشنهاد بدهد که زاویه (θ) بین بردار U و V کمتر باشد. کاربری با بردار سلیقه $U = (3, 4, k)$ داریم که سلیقه‌اش در دو بُعد «اکشن» و «درام» مشخص است، اما بُعد «علمی-تخیلی» او مجهول است (چون تا به حال فیلم علمی-تخیلی ندیده است). اکنون دو فیلم جدید اکران شده‌اند که باید با هم رقابت کنند:

- **سلاطین سرعت:** فیلمی کاملاً اکشن که برای آن داریم $V = (2, 0, 0)$
 - **میان ستاره‌ای:** شاهکاری از نولان که ترکیبی از درام و علمی-تخیلی است: $V = (0, 3, 4)$
- میزان علاقه کاربر به ژانر علمی-تخیلی حداقل چقدر باشد تا سیستم، میان ستاره‌ای را به او پیشنهاد کند؟

$$k > 4/5 \quad (4) \quad k > 3 \quad (3) \quad k > 1/25 \quad (2) \quad k > 0/75 \quad (1)$$

۱۳ یک مدل هوش مصنوعی خراب داریم که تصاویر را به صورت تصادفی به یکی از ۵ دسته‌ی مختلف (مثلاً گربه، سگ، ماشین، درخت یا قایق) دسته‌بندی کند. فرض کنید انتخاب دسته برای هر تصویر کاملاً شانسی و مستقل از تصاویر ورودی است. اگر این مدل ۳ تصویر را پردازش کند، احتمال اینکه حداقل دو تصویر برچسب یکسانی دریافت کنند چقدر است؟

$$60\% \quad (4) \quad 52\% \quad (3) \quad 20\% \quad (2) \quad 4\% \quad (1)$$

۱۴ مریخ‌نورد «کنجکاوی» برای تخمین احتمال «وجود حیات میکروبی» روی یک سنگ جدید، از الگوریتم هوشمندی استفاده می‌کند که ورودی‌های آن «میزان تشعشع (R)» و «غلظت آب (W)» هستند. این الگوریتم می‌تواند از تجربیات قبلی که روی بعضی سنگ‌ها آزمایش انجام شده و نتیجه وجود یا عدم وجود حیات برای آن‌ها مشخص شده و در حافظه نگهداری شده‌اند نیز استفاده کند. استراتژی تصمیم‌گیری این ربات به شرح زیر است:

- **حالت اول (تشعشع کم):** اگر میزان تشعشع کمتر از ۵۰۰ باشد ($R < 500$)، سیستم فرض می‌کند شرایط دشوار است و احتمال وجود حیات را ثابت و برابر با ۲۰ درصد در نظر می‌گیرد.
- **حالت دوم (تشعشع زیاد):** اگر میزان تشعشع ۵۰۰ یا بیشتر باشد ($R \geq 500$)، سیستم به سراغ فاکتور «آب» می‌رود. در این حالت، ۳ نمونه سنگ از حافظه که نزدیکترین غلظت آب را به سنگ فعلی دارند انتخاب کرده و «کسری از آن‌ها که دارای حیات هستند» را به عنوان احتمال حیات گزارش می‌دهد.

سنسور تشعشع برای سنگ فعلی عدد ۴۹۸ را نشان می‌دهد. اما طبق کاتالوگ، مقدار واقعی تشعشع یک متغیر تصادفی با توزیع یکنواخت در بازه‌ی ± 5 واحدی نسبت به عدد نمایش داده شده است ($[493, 503]$). اگر بدانیم غلظت آب در سنگ فعلی دقیقاً $W = 12$ است، با توجه به جدول داده‌های زیر و عدم قطعیت در سنسور تشعشع، سیستم نهایتاً چه احتمالی را برای وجود حیات اعلام می‌کند؟

کد سنگ	غلظت آب (W)	نتیجه آزمایش قبلی
A	۱۰	دارای حیات
B	۱۱	بدون حیات
C	۱۴	دارای حیات
D	۱۵	بدون حیات
E	۱۸	دارای حیات

$$0/68 \quad (4) \quad 0/52 \quad (3) \quad 0/46 \quad (2) \quad 0/34 \quad (1)$$

۱۵ محمدجواد سامانه‌ای طراحی کرده است که ایمیل‌های ورودی را به دو دسته «مزاحم» و «غیر مزاحم» تقسیم‌بندی می‌کند. برای این منظور، او دستگاه معادلات خطی $Xw = y$ را تشکیل داده است:

- هر سطر ماتریس X نمایانگر یک «ایمیل» و هر ستون آن نمایانگر «تعداد تکرار یک کلمه‌ی کلیدی در ایمیل‌ها» (مانند «سلام»، «جایزه» و ...) است.
- بردار w شامل وزن‌ها (میزان اهمیت) هر کلمه است که باید محاسبه شود.
- بردار y برچسب واقعی ایمیل‌ها (مزاحم یا غیر مزاحم) است.

فرض کنید ماتریس X مربعی است. سیستم دو ویژگی جداگانه برای کلمات «رایگان» و «مجان» در نظر گرفته است، اما در تمام ایمیل‌های ورودی، هر جا «رایگان» آمده، کلمه «مجان» هم بلافاصله آمده است. در این شرایط (وابستگی خطی دو ستون)، چه اتفاقی برای محاسبه بردار وزن‌ها (w) می‌افتد؟

- (۱) دستگاه دارای یک جواب منحصر به فرد است و سیستم دقیقاً بهترین وزن‌ها را پیدا می‌کند.
- (۲) دستگاه یا «هیچ جوابی» ندارد و یا «بی‌نهایت جواب» دارد؛ سیستم نمی‌تواند یک ترکیب وزن یکتا برای این کلمات انتخاب کند.
- (۳) این مشکل صرفاً با ترانپوز (Transpose) کردن ماتریس X حل می‌شود و جواب یکتا به دست می‌آید.
- (۴) درمیان ماتریس X مخالف صفر است، اما جواب‌های به دست آمده اعداد بسیار بزرگی خواهند بود.

۱۶ شما در حال آموزش دو مدل هوش مصنوعی (A و B) روی دو سرور جداگانه هستید. فرآیند آموزش به این صورت است که در پایان هر دور دیدن کل داده‌های آموزش (ایپاک)، یک بررسی تصادفی انجام می‌شود. سپس با احتمال ثابت p ، آموزش متوقف می‌شود (Crash) و با احتمال $1 - p$ مدل دور دیگری روی کل داده‌ها آموزش می‌بیند. این احتمالات در هر مرحله مستقل هستند.

- مدل A : تاکنون موفق شده ۱۰۰ دور را بدون توقف پشت سر بگذارد و هنوز در حال اجراست.
- مدل B : همین لحظه آموزش را از صفر شروع کرده است.

اگر R_A «امید ریاضی تعداد دورهای باقی‌مانده برای مدل A » باشد و T_B «امید ریاضی تعداد دورهای کل برای مدل B » باشد، نسبت $\frac{R_A}{T_B}$ کدام است؟

- (۱) این نسبت بزرگتر از ۱ است (مدل A پایدارتر است).
- (۲) این نسبت کوچکتر از ۱ است (مدل A پیر شده و احتمال توقفش بیشتر است).
- (۳) این نسبت دقیقاً ۱ است.
- (۴) بستگی به مقدار p دارد.

۱۷ فرض کنید می‌خواهیم یک مدل ریاضی بسازیم که بتواند بر اساس داده‌های ورودی (ماتریس X)، خروجی (برداری y) را پیش‌بینی کند (مثلاً تخمین قیمت خانه بر اساس مترائ).

$$Xw = y$$

به فرآیند پیدا کردن بهترین خط یا نمودار برای این کار (پیدا کردن بردار وزن w که معادله بالا را تقریباً برآورده کند) «رگرسیون» (Regression) می‌گوییم.

- حالت اول (X_1):

$$X_1 = \begin{bmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & 0 & -3 \end{bmatrix}$$

- حالت دوم (X_2):

$$X_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1/0.001 \end{bmatrix}$$

برای داشتن یک مدل پایدار (که نویز بردار y باعث تغییرات انفجاری در وزن‌های w نشود)، کدام حالت ترجیح داده می‌شود؟

- (۱) ماتریس X_2 پایدارتر است، زیرا دترمینان آن فاصله بیشتری از صفر دارد و معکوس پذیرتر است.
- (۲) هر دو ماتریس به شدت ناپایدار هستند چون دترمینان هر دو بسیار کوچک و نزدیک به صفر است.
- (۳) ماتریس X_1 پایدارتر است، زیرا ستون‌های آن کاملاً مستقل (عمود برهم) هستند، اما X_2 دارای «هم‌خطی» شدید (Near Linear Dependence) است.
- (۴) پایداری هر دو یکسان است زیرا هر دو ماتریس 2×2 هستند و رتبه کامل دارند.

۱۸ فرض کنید $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ یک ماتریس تبدیل باشد. بردار تصادفی $v = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ را در نظر بگیرید که در آن X_1 و X_2 متغیرهای تصادفی مستقل هستند که فقط مقادیر $+1$ یا -1 را با احتمال برابر (0.5) می‌گیرند. مقدار امید ریاضی توان دوم طول بردار خروجی، یعنی $E[\|Av\|^2]$ چقدر است؟

- (۱) ۰ (۲) ۲ (۳) ۴ (۴) ۸

۱۹ ما دو جعبه سیاه داریم که هر کدام یک ضرب ماتریسی انجام می‌دهند: دو جعبه‌ی سیاه داریم که هر کدام یک بردار ورودی را در یک ماتریس مشخص ضرب می‌کنند:

- جعبه‌ی A: بردار ورودی را در ماتریس M_A ضرب می‌کند که معکوس پذیر است ($\det(M_A) \neq 0$).
- جعبه‌ی B: بردار ورودی را در ماتریس M_B ضرب می‌کند که یک ماتریس تصویر روی محور افقی است (y را صفر می‌کند).

ما یکی از جعبه‌ها را با احتمال برابر انتخاب می‌کنیم و بردار ورودی $x = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ را به آن می‌دهیم. مشاهده می‌کنیم که خروجی بردار $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ است احتمال اینکه جعبه انتخاب شده، جعبه A باشد چقدر است؟

$$P(A \text{ جعبه} \mid \text{خروجی} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}) = ?$$

- (۱) ۰ (۲) ۰/۲۵ (۳) ۰/۵ (۴) ۱

۲۰ سهراب از زال خواست تا به عنوان هدیه تولدش یک پر سیمرخ نفیس به او هدیه دهد. زال اما گفت که "اگر با من و اسفندیار رویین‌تن مسابقه کشتی دهی و در حداقل ۴ مسابقه متوالی از پنج مسابقه کشتی برنده شوی یک پر سیمرخ به تو هدیه می‌دهم." برای این کار سهراب باید یکی در میان با زال و اسفندیار کشتی بگیرد اما می‌تواند انتخاب کند که اولین مسابقه را با کدام یک از این دو برگزار کند. احتمال برنده شدن سهراب در مسابقه با اسفندیار رویین‌تن فقط ۰/۳ است اما از آنجا که زال پیر شده احتمال برنده شده او در مسابقه با زال ۰/۸ است. کدام گزینه برای سهراب درست است؟

- (۱) سهراب باید اولین مسابقه را با زال انجام دهد تا احتمال بردن جایزه تقریباً ۰/۰۳ بیشتر شود
- (۲) سهراب باید اولین مسابقه را با اسفندیار انجام دهد تا احتمال بردن جایزه تقریباً ۰/۰۳ بیشتر شود
- (۳) سهراب باید اولین مسابقه را با زال انجام دهد تا احتمال بردن جایزه تقریباً ۰/۳ بیشتر شود
- (۴) سهراب باید اولین مسابقه را با اسفندیار انجام دهد تا احتمال بردن جایزه تقریباً ۰/۳ بیشتر شود



این صفحه جهت استفاده به عنوان چرک نویس در نظر گرفته شده است



این صفحه جهت استفاده به عنوان چرک نویس در نظر گرفته شده است



این صفحه جهت استفاده به عنوان چرک نویس در نظر گرفته شده است



این صفحه جهت استفاده به عنوان چرک نویس در نظر گرفته شده است



این صفحه جهت استفاده به عنوان چرک نویس در نظر گرفته شده است



این صفحه جهت استفاده به عنوان چرک نویس در نظر گرفته شده است