



سوال ۱. {هدف این سوال بررسی تاثیر ریسک و ابهام در تصمیم گیری است.}



امروزه علاوه بر فروشگاه‌های سنتی (مغازه‌ها)، فروشگاه‌های اینترنتی نیز بسیار فراگیر شده‌اند. با توجه به تفاوت شرایط و خدماتی که هریک از این دو ارائه می‌کنند، هر دو هنوز به حیات خود ادامه می‌دهند و مشتری‌های خود را دارند. فرض کنید یک لپ‌تاپ وجود دارد که در هر دو این فروشگاه‌ها به فروش می‌رسد. قیمت این لپ‌تاپ در مغازه ۱۰ میلیون تومان و در فروشگاه اینترنتی یوتی‌کالا ۱۲ میلیون تومان است. احتمال خرابی جنس خریده شده از مغازه،

با احتمال خرابی جنس خریداری شده از یوتی‌کالا مساوی و برابر p است. اگر از مغازه خرید کرده باشید و کالا خراب باشد، خرابی شامل گارانتی می‌شود و باید لپ‌تاپ را به تعمیرگاه ببرید. هزینه‌ای بابت تعمیر لپ‌تاپ پرداخت نمی‌کنید ولی همان دستگاه تعمیر شده و به شما برمی‌گردد. اگر از یوتی‌کالا خرید کنید و لپ‌تاپ خراب از آب در بیاید، تا ۷ روز بعد از خرید، دستگاه با یک نمونه‌ی نو تعویض می‌شود (در بازه‌ی ۷ روزه هرچندبار دستگاه خراب به شما تحویل داده شود، با نمونه‌ی نو تعویض می‌شود).

الف) شما با بودجه ۱۵ میلیون تومانی، قصد خرید همین لپ‌تاپ را دارید. تابع سودمندی^۱ شما به صورت زیر است. B نشان‌دهنده‌ی بودجه‌ی شما و C نشان‌دهنده‌ی قیمت لپ‌تاپ است. ابتدا بیان کنید بازه‌ی منطقی α ، و بتا چیست و سپس بر اساس مقادیر این پارامترها و P توضیح دهید که طبق این تابع سودمندی انتخاب بهینه کدام است؟

$$U(C) = \begin{cases} (B - C)^\alpha & \text{در صورتی که لپ‌تاپ خراب نشود} \\ -\beta(B - C)^\alpha & \text{در صورتی که لپ‌تاپ به تعمیرگاه برود} \end{cases}$$

ب) در ماه‌های اخیر به علت افزایش چشم‌گیر قیمت لپ‌تاپ، فروشگاه‌هایی مانند قوطی‌کالا هم به وجود آمده‌اند که لپ‌تاپ‌های دست دوم و تعمیرشده را با قیمت بسیار پایین‌تر از بازار عرضه می‌کنند. برای مثال

^۱ Utility Function

لپتایی که در بالا در مورد آن صحبت کردیم، در قوطی کالا با قیمت ۴ میلیون تومان به فروش می‌رسد. با این تفاوت که چون این دستگاه‌ها گارانتی ندارند، در صورت خرابی همه‌ی هزینه تعمیر بر عهده‌ی خریدار است (دقت کنید که در قسمت قبلی، کالایی که از مغازه خریداری شده بود، در صورت خرابی، شامل گارانتی می‌شد و خریدار هزینه‌ای برای تعمیر پرداخت نمی‌کرد). احتمال خرابی دستگاهی که از مغازه خریداری شده با احتمال خرابی دستگاهی که از قوطی کالا خریداری شده مساوی و برابر P است.

همچنین به علت قیمت‌های بالای یوتی کالا، این فروشگاه ورشکست شده و دیگر وجود ندارد. فردی قصد خرید یک لپتاپ از یکی از این دو فروشگاه موجود را دارد. تابع سودمندی این فرد به صورت زیر است:

$$U(r) = \begin{cases} \log_W \beta \times r^\alpha & W > 0 \\ r^\alpha & W = 0 \end{cases}$$

که در آن W هزینه تعمیر دستگاه و مقداری مثبت است. مقدار α مثبت و برابر پاداش ذهنی‌ای است که فرد از صرفه‌جویی خود بدست می‌آورد. مشخص کردن نحوه‌ی محاسبه مقدار α بر عهده شماست. محاسبه این مقدار باید سازگار با مفروضات مسئله و فرضیات شما باشد. تمام فرضیات خود در مورد نحوه محاسبه‌ی این مقدار را شرح دهید.

بر اساس مقادیر مختلف منطقی آلفا، بتا و مقدار P توضیح دهید فرد کدام فروشگاه را برای خرید انتخاب می‌کند؟

سوال ۲. {هدف از این سوال آشنایی با مفهوم تابع ارزش در انسان‌هاست.}



کبری باز هم می‌خواهد تصمیم بگیرد. او هر روز برای رفتن از منزل به سمت محل کارش ۳ انتخاب دارد و می‌تواند از ۳ مسیر مختلف به مقصد خود برسد. هزینه‌هایی که کبری برای رفتن از هر یک از مسیرها بابت وسایل نقلیه می‌پردازد با هم متفاوت است و هریک از توزیعی نرمال می‌آید:

$$\text{مسیر اول : } P(r) \sim N(15,8)$$

$$\text{مسیر دوم : } P(r) \sim N(25,8)$$

$$\text{مسیر سوم : } P(r) \sim N(30,8)$$

مسیر اول در ۳۰ درصد مواقع به دلیل ترافیک شدیدی که وجود دارد احتمال دیر رسیدن کبری را به دنبال دارد. مسیر دوم در ۲۰ درصد مواقع به دلیل تصادفاتی که در آن رخ می‌دهد، باعث می‌شود کبری دیرتر به محل کارش برسد. وجود ترافیک و یا تصادف تاثیری در زمان رسیدن از مسیر سوم نخواهد داشت.

کبری اگر سر موقع به محل کارش نرسد و تاخیر داشته باشد توسط مدیرش جریمه می‌شود. جریمه‌ی او به این صورت است که مبلغی از حقوق ماهیانه‌اش کسر می‌شود که از توزیع نرمال زیر می‌آید.

$$\text{توزیع جریمه : } P(r) \sim N(40,8)$$

هدف کبری این است که مسیری را برای رفتن به محل کارش انتخاب کند که در طولانی‌مدت بتواند بیشتر پولش را پس انداز کند و کمتر جریمه شود.

اگر بدانیم که تابع سودمندی کبری به صورت زیر است و فرض کنیم پارامتر a برابر ۰.۳ و w نمایان‌گر هزینه باشد:

$$U(w) = \begin{cases} \frac{w^{1-a} - 1}{1-a}, & w \geq 30 \\ \frac{w - e^{-aw}}{1+a}, & w < 30 \end{cases}$$

الف) شما کدام مسیر را به وی پیشنهاد می‌دهید؟ تغییر پارامتر a چه تغییری در تابع ارزش کبری ایجاد می‌کند؟

ب) می‌دانیم که احتمال ذهنی گرفتن جایزه برای افراد (Subjective Probability) لزوماً با احتمال واقعی گرفتن جایزه (Objective Probability) یکسان نیست. اگر روزی کبری صبح زود از خواب بیدار شود و زمان اضافی داشته باشد، احتمال دیر رسیدن خود را با وجود احتمال ترافیک و تصادفات (Subjective Probability) از تابع زیر محاسبه می‌کند. با به دست آوردن احتمالات تاخیر جدید برای دو مسیر اول این تغییر را توجیه کنید. در این صورت بهتر است از کدام مسیر به محل کارش برود؟ (p احتمال دیر رسیدن است و پارامتر b را در این قسمت برابر ۰.۳ فرض کنید.)

$$f(p) = \frac{b \frac{p}{1-p}}{1 + b \frac{p}{1-p}}$$

سوال ۳. {هدف از این سوال ارائه‌ی مثالی واقعی از تصمیم‌گیری‌های غیرمنطقی در انسان‌ها و استفاده از آن در یک مسئله‌ی کاربردی است.}

دو رقم آخر شماره دانشجویی شما چیست؟ (برای مثال ۲۰) آیا حاضرید به اندازه ۲۰ هزار تومان برای یک خوراکی جدید (که تا به حال نخورده‌اید و اطلاعاتی راجع به آن ندارید) هزینه کنید؟ پس از پاسخ دادن به این سوال، فکر می‌کنید هزینه‌ی داوطلبانه‌ای که برای این خوراکی خواهید پرداخت چقدر متأثر از دو رقم آخر شماره دانشجویی شما خواهد بود؟

به گروهی از دانشجویان مدیریت دانشگاه MIT لیستی از کالاها داده شد. از آن‌ها خواستند دو رقم آخر شماره دانشجویی خود را بالای لیست بنویسند. همچنین به آن‌ها گفته شد این عدد را در فرم قیمت، کنار نام هر کالا بنویسند و در کنار هر کالا با 'بله' یا 'خیر' مشخص کنند که آیا حاضرند به اندازه قیمت مشخص شده توسط دو رقم آخر کارت دانشجویی آن را خریداری کنند یا خیر. در مرحله آخر آزمایش از آن‌ها خواسته شد بیشینه قیمتی که برای خرید هر یک از کالاها حاضرند بپردازند را کنار هر کالا بنویسند. بررسی داده‌های حاصل از پاسخ دانشجویان نشان داد دانشجویان با شماره دانشجویی بالاتر قیمت بیشتری پیشنهاد داده‌اند. در واقع سوال اولیه (تمایل به خرید کالا به قیمت مشخص شده توسط دو رقم آخر شماره دانشجویی) باعث بوجود آمدن یک قیمت پایه در ذهن این دانشجویان شده است. نکته جالب دیگر این آزمایش این بود که پس از شکل‌گیری این قیمت پایه، کالاهای هم‌دسته متناسب با این قیمت پایه ارزش‌گذاری شدند. برای مثال، اگر موس در لیست داده شده به دانشجویان وجود داشت و پس از آن از آن‌ها قیمت صفحه‌کلید پرسیده می‌شد، آن‌ها قیمت صفحه‌کلید را بیشتر از موس و متناسب با قیمت موس پیشنهاد می‌دادند.

«زمانی که به لحاظ ذهنی ارزش یک کالا را پذیرفتیم، هزینه تمامی کالاهای هم‌دسته را با این ارزش پایه مقایسه می‌کنیم.» این پدیده اصطلاحاً مهر زدن^۲ نامیده می‌شود. قیمت اولیه می‌تواند بدون هیچ منطقی و کاملاً تصادفی انتخاب شده باشد، اما به محض پذیرفته شدن آن در ذهن نه تنها قیمت فعلی آن کالا، بلکه قیمت تمامی کالاهای هم‌خانواده با این کالا که در آینده با آن‌ها روبرو می‌شویم تحت تاثیر قرار می‌گیرد.

پس از این مقدمه (!) فرض کنید می‌خواهیم از این پدیده غیرمنطقی برای بیشینه‌کردن سود در پیک موتوری اینترنتی یوتی‌باکس استفاده کنیم. در این شرکت تعدادی موتور سوار استخدام شده است که هر کدام به صورت ماهانه حقوق ثابتی دریافت می‌کنند. معمولاً مشتری‌های این شرکت مسیرهای تکراری زیادی دارند. یعنی

^۲ Imprinting

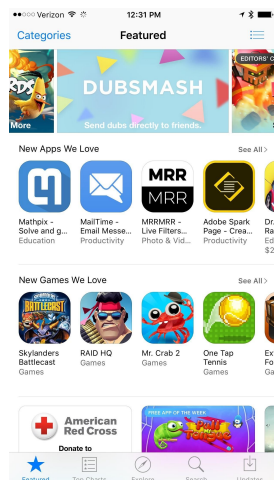
مسیرهایی که چندین بار در ماه بین آن‌ها کالا جابه‌جا می‌کنند. طبق آزمایش گفته شده در سوال می‌توان نتیجه گرفت در ذهن مشتریان به ازای هر کدام از این مسیرهای متداول یک قیمت پایه مشخص شده است. اگر بتوانیم این قیمت پایه را به دست آوریم، می‌توانیم از آن برای قیمت‌گذاری مسیرهای غیر متداول‌تر استفاده کنیم زیرا مشتری‌ها هزینه یک مسیر غیر متداول را نسبت به هزینه‌ای که برای مسیر پایه پذیرفته‌اند می‌سنجند و سرویس ارائه‌شده توسط شرکت را قبول یا رد می‌کنند. با استفاده از این ایده می‌توانیم میزان پذیرش سرویس توسط مشتری و در نتیجه سود شرکت را بالا ببریم.

چه راه حلی پیشنهاد می‌کنید که:

الف) سود شرکت به ازای مشتری‌های جدید بیشینه شود.

ب) سود شرکت به ازای مشتری‌های قدیمی بیشینه شود.

سوال ۴. {هدف از این سؤال طراحی حالت‌ها، اعمال و پاداش برای یادگیری تقویتی در یک مسئله‌ی کاربردی است.}



شرکت یوتی‌مارکت که یک فروشگاه آنلاین خرید و دانلود برنامه‌های موبایل است در نظر دارد جهت بهبود سیاست‌های بازاریابی خود از روش‌های یادگیری تقویتی استفاده کند.

یکی از چالش‌های اصلی که این شرکت با آن روبه‌روست، تشخیص برنامه‌های پرمخاطب و درآمدزا در ابتدای معرفی آن‌هاست. به عبارت دیگر اگر این شرکت بتواند در هفته‌های ابتدایی که یک برنامه به بازار عرضه می‌شود تشخیص دهد که مورد اقبال مخاطبین قرار خواهد گرفت، ۱-می‌تواند آن را بیشتر تبلیغ کند و ۲-با توجه به نیازی که برای آن برنامه وجود دارد، سیاست قیمت‌گذاری خود را تغییر دهد تا به سود بیشتری دست پیدا کند (یا تخفیف بگذارد و با تعداد بیشتر فروش سود کند یا قیمت‌ها را بالا ببرد و از طریق آن درآمد را افزایش دهد).

متأسفانه این شرکت داده‌های رفتارهای قبلی مشتریان را ذخیره نکرده است و قصد دارد در این پروژه به صورت تعاملی و با استفاده از یادگیری تقویتی آن را به انجام رساند. همچنین اطلاعات فروش، نظرات کاربران و نرخ کلیک روی کلیه لینک‌ها نیز در دسترس است. اقدام اصلی شرکت در ابتدا این است که چینش مناسب برنامه‌ها در صفحه‌ی اصلی را کشف کند. به این منظور شبانه‌روز به ۶ بازه‌ی ۴ ساعته و دو دسته‌ی روزهای تعطیل و غیرتعطیل تقسیم‌بندی شده است. با توجه به این شرایط:

(الف) تیم بازاریابی تعدادی از برنامه‌های مستعد را به شما معرفی کرده است و در بخش اول این پروژه الگوریتم باید تصمیم بگیرید که کدامیک از این برنامه‌ها را و در چه ساعتی از روز روی صفحه‌ی اصلی تبلیغ کند. به نظر شما تابع ارزش الگوریتم یادگیری چگونه می‌تواند باشد؟ (با استفاده از مدل ریاضی تابع ارزش را بیان کنید).

(ب) فضای حالت و اعمال در این مسئله چگونه تعریف می‌شوند؟

(پ) برای اینکه سرعت یادگیری مسئله افزایش یابد چه اقداماتی می‌توان انجام داد؟

امتیازی

(ت) همواره یکی از نگرانی‌های مدیران برای استفاده از الگوریتم‌های یادگیری ماشین، پیش‌بینی‌ناپذیر بودن تصمیم‌های آن‌ها و امکان اشتباه کردن است. پیشنهاد شما برای حل این موضوع چیست؟ یکی از ایده‌های پیشنهادی در [این لینک](#) قابل مشاهده است.

سوال ۵. {هدف از این سؤال طراحی حالت‌ها، اعمال و پاداش برای یادگیری تقویتی است.}



یکی از مسائل مهم در صنعت بازی‌سازی، تعیین سطح کاربر است. در صورتی که بازی بسیار ساده باشد، باعث خستگی او می‌شود و در صورتی که بازی بسیار سخت باشد، ناامیدی را برای او به همراه خواهد داشت. این مسأله، با نام DGDB³ شناخته می‌شود. در این سوال، می‌خواهیم با حداقل سعی و خطای ممکن، سطح بازی کاربر را بدست آورده و متناسب با همان سطح، با او بازی کنیم. (سطح بازی یک کاربر به این معناست که اگر بازی کمی سخت‌تر شود، کاربر با احتمال زیادی همیشه بازنده است.)

بازی فیفا شامل ۷۰۰ تیم باشگاهی و ملی است. این تیم‌ها قدرت‌های متفاوت دارند که با ستاره نمایش داده می‌شوند. برای مثال، تیم فرانسه، ۵ ستاره و تیم ژاپن ۳ ستاره است (تعداد ستاره‌ها مقداری از ۱ تا ۶ دارد). کاربران بازی فیفا، می‌توانند با انتخاب تیم مورد علاقه‌شان، در برابر سیستم به رقابت بپردازند. هر تیم سیستم می‌تواند در پنج سطح⁴ در مقابل کاربر بازی کند. پس از پایان بازی، گزارشی از تعداد گل‌ها، درصد مالکیت، تعداد شوت‌های در چارچوب، تعداد پاس‌های موفق، کرنرها، خطاها و تعداد کارت‌های زرد و قرمز است نمایش داده می‌شود.

الف) فرض کنید که کاربر و سیستم، همیشه تیم‌های ۳ ستاره را انتخاب می‌کنند. مدلی از یادگیری تقویتی ارائه دهید که با چند بار بازی کردن در مقابل کاربر، بتواند سطح بازی او را ارزیابی کند (حالت‌ها⁵، اعمال⁶ و تابع ارزش را مشخص کنید).

³ Dynamic Game Difficulty Balancing

⁴ 1-Amateur 2-Semi-Pro 3-Professional 4-World Class 5-Legendary

⁵ States

⁶ Actions

(ب) چه فرض هایی نیاز است تا بتوانیم توزیع تابع ارزش را ایستا⁷ در نظر بگیریم؟

(پ) با گذشت زمان و با انجام n تا بازی، با احتمال p سطح بازی کاربر تغییر می‌کند. با فرض اینکه p و n را ندانیم، چگونه می‌توانیم همواره تخمین درستی از سطح بازی کاربر داشته باشیم؟

(ت) فرض کنید که در بازی اول، سطح professional انتخاب می‌شود و کاربر با نتیجه ۱۰ بر ۱ بازی را واگذار می‌کند. در بازی دوم، سیاست تعیین سطح بازی چیست؟ با تعمیم مسأله به سایر سطح‌ها و تابع سیاست را به شکل ریاضی بنویسید.

(ث) در قسمت الف برای تیم‌های ۳ ستاره، سطح بازی کاربر بدست آمد. از اطلاعات بدست آمده در قسمت الف، چگونه می‌توان برای تسریع تخمین سطح بازی کاربر در دیگر شرایط استفاده کرد؟ (برای مثال سیستم تیم ۲ ستاره انتخاب کند و کاربر ۴ ستاره انتخاب کند.)

⁷ Stationary

لطفا به نکات زیر توجه کنید:

- ✓ حجم گزارش شما به هیچ وجه معیار نمره دهی نیست، پس لطفا در حد نیاز توضیح دهید.
- ✓ سعی کنید از پاسخ های روشن در گزارش خود استفاده کنید و اگر پیش فرضی در حل سوال در ذهن خود دارید، حتما در گزارش خود آن را ذکر کنید.
- ✓ از نمودارهای واضح در گزارش خود استفاده کنید، نمودارهایی که دارای لیبل گذاری روشن روی هر محور و همینطور توضیح مناسب باشد.
- ✓ کدهایی که به همراه گزارش تحویل می دهید باید قابل اجرا باشد. همچنین توجه کنید که به تمرین بدون گزارش نمره ای تعلق نمی گیرد.
- ✓ لطفا در گزارش و کدهای خود از تمرین دیگران استفاده نکنید، مشورت و همفکری در مورد سوال ها اشکالی ندارد اما اگر شباهت بیش از اندازه در تمرین ها دیده شود منجر به صفر شدن نمره خواهد شد.
- ✓ تمام فایل ها را در قالب یک فایل zip یا rar در سایت درس بارگذاری کنید.
- ✓ برای پیاده سازی تمرین فقط از زبان های MATLAB و یا Python می توانید استفاده کنید.

موفق و سلامت باشید. (: