

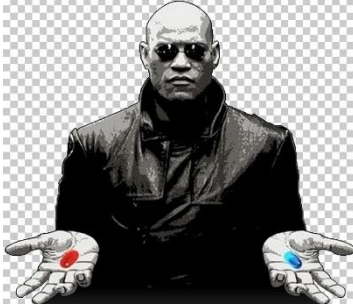


به نام او، برای او



تمرین اول درس یادگیری ماشین
پاییز ۱۳۹۸

سوال ۱



فرض کنید شما یک بیماری کشنده دارید و به ناچار برای معالجه به دکتر مراجعه می‌کنید. دکتر به شما می‌گوید که برای این بیماری عجیب، فقط دو نوع درمان وجود دارد که در هرکدام از آنها یک قرص برای شما تجویز می‌شود.

- قرص اول: حتماً برای ۱۲ سال دیگر زنده می‌مانید.
- قرص دوم: با احتمال ۸۹ درصد برای ۱۲ سال دیگر زنده می‌مانید. با احتمال ۱۰ درصد برای ۱۸ سال دیگر زنده می‌مانید و با احتمال ۱ درصد بدن شما به قرص حساسیت نشان می‌دهد و پس از مصرف آن به خواب ابدی می‌روید.

(الف)

کدام قرص را انتخاب می‌کنید؟

(ب)

تحلیل کنید که چرا این قرص را انتخاب کردید.

سوال ۲

۴ تابع ارزش (Utility Function) زیر را در نظر بگیرید.

$$U(w) = w^4$$

$$U(w) = \log(w)$$

$$U(w) = 2^w$$

$$U(w) = \tanh(w)$$

(الف)

به نظر شما، کدام توابع بالا ریسک پذیر و کدام ریسک گریز هستند؟ علت دسته بندی خود را توضیح دهید.

(ب)

پس از اینکه توابع بالا را به دو دسته ریسک پذیر و ریسک گریز تقسیم کردید، به نظر شما در هر دسته کدام تابع نسبت به دیگر تابع های مشابه در دسته خود، ریسک پذیرتر یا ریسک گریز تر است و نسبت به تغییرات واکنش شدیدتری دارد؟

راهنمایی: رسم این توابع می تواند به شما دید بهتری برای قضاوت در این مسئله بدهد.



کبرا و اکبر، اخیراً روی گوشی‌های خود بازی اکشنی به نام Bandit Country نصب کرده‌اند که در حال رشد بسیار سریعی است. یکی از مراحل اصلی این بازی، بسیار به شانس بازیکن بستگی دارد. در این مرحله، مقابل بازیکن ۴ صندوق قرار داده می‌شود که بازیکن باید از میان این ۴ صندوق، یک صندوق را انتخاب کند. هر صندوق یا امتیازی مثبت به شما می‌دهد، یا از بدشانسی شما، امتیازی منفی. در صورتی که صندوق را باز کنید و امتیاز شما بیشتر شود، این مرحله را برده‌اید و در غیر این صورت این مرحله را باخته‌اید.

یکی از برنامه‌نویس‌های این بازی اخیراً افشا کرده است که این صندوق‌ها، با توزیع‌های زیر به کاربران امتیاز می‌دهند.

Chest #1 : $N(2.5, 3)$

Chest #2: $Uniform(-1, 4)$

Chest #3: $N(3, 2)$

Chest #4: $N(-2, 1.5)$

اکبر علاقه‌مند است که در این مرحله، تعداد برد خود را افزایش دهد تا در رده بندی صعود کند (یعنی تعداد دفعاتی که امتیاز مثبت کسب می‌کند) اما کبرا می‌خواهد تعداد امتیاز خود را بیشتر کند و علاقه چندانی به رده بندی ندارد.

(الف) یعنی براساس اندازه پاداش مثبت بزرگ‌تر می‌شود. *punishment averse*

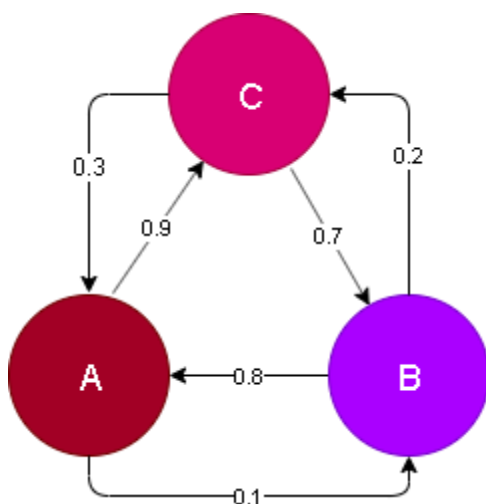
به نظر شما اکبر که می‌خواهد تعداد برد را افزایش دهد، باید چه تابع ارزشی داشته باشد که بتواند به هدف خود برسد؟ چرا احساس می‌کنید که این تابع، تابع مناسبی است؟

(ب) این براساس هم‌ع که پاداش زیاد برده یعنی باطن هم‌نیت جانور می‌شود

به نظر شما کبرا که می‌خواهد امتیاز کل خود را افزایش دهد، باید چه تابع ارزشی داشته باشد که بتواند به هدف خود برسد؟ چرا احساس می‌کنید که این تابع، تابع مناسبی است؟

سوال ۴

فرض کنیم که در یک زندان سه سلول A و B و C وجود دارد و زندانبان بر مبنای گراف زیر به سلولها سر می‌زند. یک زندانی در سلول A می‌خواهد با آپدیت بیزی باورهایش احتمال آمدن زندانبان را محاسبه کند. آن زندانی می‌داند که سه سلول وجود دارد و نگهبان به آنها سر می‌زند ولی در ابتدا احتمال آنها را نمی‌داند. یک روش برای محاسبه بهینه احتمال آمدن زندانبان در دو حالت زیر ارائه کنید. (در شروع فرایند زندانبان در یکی از سلولهاست).



(الف)

زندان فقط می‌تواند مشاهده کند که زندانبان وارد سلولش شده یا نه و از دیگر سلولها بی‌خبر است.

(ب)

زندان می‌تواند بیرون سلول را مشاهده کند که زندانبان در هر لحظه به کدام سلول سر می‌زند.

(ج)

برای دو حالت قبلی مساله را شبیه سازی کنید و عملکرد زندانی را با error bar های مناسب، نمایش دهید و دو حالت را با هم مقایسه کنید.

سوال ۵

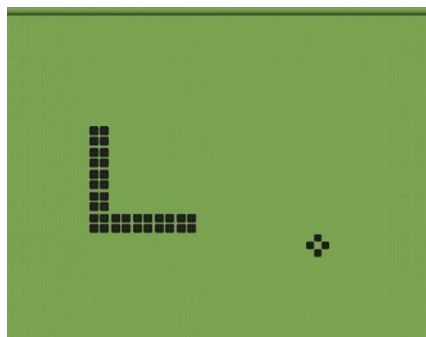


یک تحلیل آماری مبتنی بر داده نشان می‌دهد که احتمال برد یک تیم فوتبال برابر ۱۰ درصد است. یکی از هواداران دو آتشه و تعصبی این تیم، اگر احساس کند که احتمال برد تیمش بیش از ۵۰ درصد است به ورزشگاه می‌رود تا این پیروزی شیرین را از نزدیک ببیند. اگر بدانیم که برای این هوادار، رابطه بین subjective probability و objective probability به صورت زیر است:

$$P_s(o) = \frac{e^{\beta P_o(o)}}{\sum_{o'} e^{\beta P_o(o')}}$$

o ، در این رابطه، یک outcome دلخواه است. به نظر شما به ازای چه مقدار β این هوادار به ورزشگاه می‌رود و مسابقه را از نزدیک می‌بیند؟

سوال ۶



بازی Snake، یکی از خاطره انگیز و نوستالژیک ترین بازی های تاریخ است. سر همین علاقه، یکی از دانشجویان ترم اولی دانشگاه تهران می خواهد با استفاده از یادگیری تقویتی، عاملی را تربیت کند که بتواند این بازی را انجام دهد و با اسکرین شات گرفتن از High Score این عامل، به دوستان خود که یادگیری تقویتی بلد نیستند فخر بفروشد. وی برای سادگی کار خود، ابتدا بازی را در حالت عادی تست می کند. در این حالت همه میوه ها در بازی یکسان هستند (میوه بزرگ نداریم) و تا زمانی که میوه خورده نشده باشد، میوه جدیدی در نقشه ظاهر نمی شود. در صورت برخورد سر مار با دیوار ها و یا بدنه مار، بازی تمام می شود و با خوردن هر میوه، طول مار یک واحد افزایش می یابد.

(الف)

به نظر شما فضای حالات و فضای اعمال را چگونه طراحی کنیم که بتوانیم عامل خود را train کنیم؟ دلیل انتخاب فضای حالات و فضای اعمال خود را توجیه کنید.

(ب)

به نظر شما نحوه پاداش دادن به عامل هنگام یادگیری باید چگونه باشد؟ چرا فکر می کنید که این عملکرد در پاداش دادن مناسب است؟

(ج)

فرض کنید با بزرگتر شدن محیط بازی، بازی حالت رقابتی پیدا می کند و یک مار دیگر نیز به بازی اضافه می شود. تمامی قوانین مانند حالت معمولی است، تنها قانون جدید این است که اگر هر مار به بدن مار دیگر برخورد می بازد (در صورتی که هر دو مار به سر هم برخورد کنند بازی مساوی می شود). آیا فضای حالات و اعمالی که در قسمت الف تعریف کرده اید، نیاز به تغییر دارد؟ آیا می توانید این تغییرات در حداقل ترین شکل ممکن بیان کنید؟

(د)

به نظر شما اگر عامل روی هر دوی این مار ها نصب شود و در محیط رقابتی با خودش رقابت کند، چه عواملی ممکن است باعث برتری یک مار به مار دیگر شود؟ آیا فکر می کنید در این صورت همواره به تساوی خواهیم رسید؟ پیشبینی خود را از این شرایط توضیح دهید.

سوال ۷ {امتیازی}

هدف یک بانک تعیین حداقل میزان سپرده قرض الحسنه یکساله لازم برای کسانی است که می‌خواهند در قرعه‌کشی بانک شرکت کنند. احتمال برد افرادی که این سپرده و بیشتر از آن را داشته باشند، یکسان است. برای تعیین حداقل سپرده، بانک نیاز به اجرای یک آزمون کامپیوتری دارد. آزمون مناسب برای این امر را طراحی کنید.

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 پیشفرضی 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 میدهید باید قابل اجرا باشد. همچنین توجه کنید که به تمرین بدون گزارش نمرهای تعلق نمیگیرد.
- ✓ لطفا در گزارش و کدهای خود از تمرین دیگران استفاده نکنید، مشورت و همفکری در مورد سوالها اشکالی ندارد اما اگر شباهت بیش از اندازه در تمرینها دیده شود منجر به صفر شدن نمره خواهد شد.
- ✓ تمام فایلها را در قالب یک فایل zip در سایت درس بارگذاری کنید.
- ✓ برای پیاده سازی تمرین فقط از زبانهای MATLAB و یا Python میتوانید استفاده کنید

موفق و سلامت باشید :)