



هوش مصنوعی

بهار ۱۴۰۰

استاد: محمدحسین رهبان

گردآورندگان: کیمیا یزدانی - نیما فتحی

Reinforcement Learning

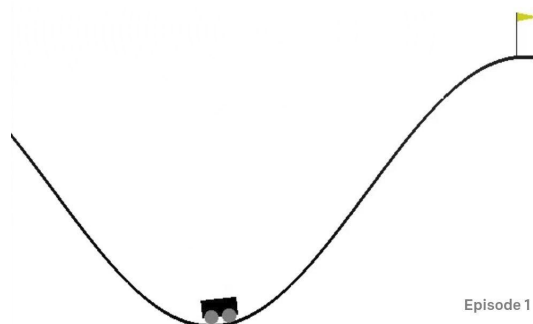
تمرین عملی پنجم

مهلت ارسال: ۱۷ تیر

- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص شده است.
- همکاری و همفکری شما در انجام تمرین مانعی ندارد اما پاسخ ارسالی هر کس حتما باید توسط خود او نوشته شده باشد.
- در صورت همفکری و یا استفاده از هر منابع خارج درسی، نام همفکران و آدرس منابع مورد استفاده برای حل سوال مورد نظر را ذکر کنید.
- لطفا تصویری واضح از پاسخ سوالات نظری بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.

سوالات عملی (۱۲۵ نمره)

۱. (۱۰۰ + ۲۵ نمره) در این سوال شما باید با استفاده از دانشتان در زمینه RL یکی از مسائل Open AI Gym را پیاده سازی کنید. مسئله ای که باید پیاده کنید **MountainCar-V0** است. در این مسئله یک ماشین در دره ای وسط دو کوه قرار دارد. شما باید بتوانید ماشین را از شیب سمت راست بالا ببرید و به پرچم برسائید، اما ماشین نمی تواند به طور مستقیم شتاب لازم برای این صعود را کسب کند. پس برای این کار لازم است ماشین شروع به عقب و جلو رفتن کند تا تکانه ای لازم ایجاد شود و در نهایت بتواند از شیب بالا برود.



توضیحات مسئله:

- (آ) در این محیط برای ماشین دو ویژگی تعریف شده است، ویژگی مکان و ویژگی سرعت. ویژگی سرعت در ابتدا صفر و بر اساس انتخاب های کد شما تغییر می کند و ویژگی دیگر مختصات x ماشین است که در ابتدا بین $-۰/۶$ تا $-۰/۴$ - مقداری را به صورت رندوم توسط کد پایه اتخاذ می کند.
- (ب) کد شما در هر لحظه سه نوع انتخاب دارد که با سه شماره ی صفر، یک و دو تعیین شده اند. به ترتیب انتخاب هر یک به معنی شتاب گرفتن به چپ، شتاب نگرفتن و شتاب گرفتن به راست است.
- (ج) پاداش در کد پایه به این صورت تعریف می شود که در صورت رسیدن به پرچم صفر و در غیر این صورت -۱ است.
- (د) شرط پایان هر اپیزود شبیه سازی رسیدن به پرچم (به معنی بیشتر شدن x از $۰/۵$) یا طولانی شدن زمان به بیش از ۲۰۰ واحد است.

(ه) نمره‌دهی کد شما با استفاده از دو محدوده تعیین می‌شود. امتیاز نهایی، میانگین امتیاز اجرای تست است که پس از حداکثر پانزده دقیقه training در یک اجرای جداگانه اندازه‌گیری می‌شود. اگر امتیاز نهایی شما بیش از ۱۵۰- باشد نمره‌ی کامل کد را از سوال دریافت می‌کنید. اگر بتوانید امتیازی بیشتر از ۱۴۰- هم بگیرید، علاوه بر نمره‌ی کامل کد، ۰/۲۵ نمره امتیازی دریافت می‌کنید. بخشی از نمره‌ی تمرین به نحوه‌ی تحویل اختصاص دارد که در ادامه توضیح داده شده‌است.

(و) نحوه‌ی تحویل تمرین به این صورت است که باید کد خود را در دو قالب فایل تکمیل شده و جویتر نوت‌بوک تحویل دهید. در جویتر نوت‌بوک کد را بخش بخش وارد کرده تا توضیحات لازم و نتیجه‌ی اجرای هر بخش جداگانه قابل مشاهده باشد. همچنین از شبیه‌سازی اسکرین‌شات‌های گویا تهیه کنید و در فلدر نهایی ارسال کنید. زمان training هم در کد چاپ کنید.

(ز) برای استفاده از کتابخانه‌ی gym برای این سوال از این قطعه کد استفاده کنید:

```
import gym
env = gym.make('MountainCar-v0')
env.reset()
```