



پروژه مبحث RL

درس هوش مصنوعی و سیستم های خبره

استاد درس: دکتر آرش عبدی

طراح: محمد امین رضاپور

نیم سال دوم

سال تحصیلی ۱۴۰۲-۱۴۰۳

فهرست مطالب

۱	مقدمه	۱.۰
۱	توضیح روند بازی	۲.۰
۲	نحوه پیاده سازی	۳.۰
۲	قوانین	۴.۰

۱.۰ مقدمه

در این سند روند انجام پروژه Reinforcement Learning بررسی خواهد شد .
صفحه [GitHub](#) برای انجام این پروژه کمک فراوانی به شما می کند .
در صفحه کدی گذاشته است که می توانید محیط بازی را با استفاده از آن مشاهده کنید.
طبق توضیحات صفحه گیت هاب برای نصب از python ۳.۸ استفاده شده است اگر با بقیه ورژن ها به مشکل
خوردید از anaconda استفاده کنید برای نصب دستور زیر را بزنید :
`pip install flappy-bird-gym`

۲.۰ توضیح روند بازی

در flappy bird بازیکن باید با کلیک کردن بر روی صفحه یا فشار دادن دکمه ای که پرواز پرنده را انجام می دهد، از میان لوله هایی که به طور پیوسته در حال حرکت به بالا و پایین هستند عبور کند . هدف این بازی افزایش امتیاز بازیکن با عبور از بیشترین تعداد لوله هاست.
Space Observation پارامتر مهمی است که برای انتخاب پروژه باید به آن توجه کرد . این پارامتر در واقع بیان کننده روش درک Agent از محیط اطراف خود است که می تواند به صورت لیست یا ماتریسی از پارامتر

های گسسته یا پیوسته باشد.

در محیط FlappyBird-v۰ فضای مشاهده اطلاعات عددی ساده ای درباره وضعیت بازی را برمی گرداند. این اطلاعات شامل فاصله افقی تا لوله بعدی و تفاوت بین موقعیت Y بازیکن و موقعیت Y سوراخ بعدی است. با استفاده از این اطلاعات عامل می تواند تصمیم گیری های خود را بر اساس موقعیت دو شرایط بازی بگیرد.

پارامتر مهم دیگر که باید به آن توجه کنید Space Action می باشد که مانند پارامتر قبل در روش آموزش شما و انتخاب آن تاثیر گذار خواهد بود.

۳.۰ نحوه پیاده سازی

فایل پروژه شامل دو فایل `utils.py`، `main.py` است و فایل `main.py` شامل الگوریتم و `policy` هست توابع خواسته شده را در این فایل باید پیاده سازی کنید. در فایل `utils.py` شامل یکسری توابع کمکی هست که می توانید از یک `qtable` استفاده کنید که همان `class counter` در فایل `utils` هست. یک مشکلی هست که `reward` همیشه ۱ برگردانده می شود به همین علت باید یک تابع مخصوص به `reward` پیاده سازی کنید و این `reward` باید از طریق `info` محاسبه کنید:

```
observation, reward, done, info = env.step(action)
```

۴.۰ قوانین

حتما توجه داشته باشید که الگوریتم پیاده سازی شده برای آموزش باید `qlearning` باشد. کدهای زیادی از انواع الگوریتم های برای بازی ها در اینترنت موجود است بنابراین در ارائه پروژه دقت بیشتری خواهد شد و طبیعتا به استفاده از کدهای آماده و نفهمیدن پروژه نمره ای تعلق نخواهد گرفت. تک تک اقدامات شما در راه رسیدن به الگوریتم نهایی در افزایش نمره شما موثر خواهد بود. تمام مراحل رسیدن به کد نهایی پروژه و راه حل های خلاقانه برای بهبود نتیجه به همراه گزارش نتیجه و مقایسه آن با حالت قبل را در ارائه خود داشته باشید. و حتما یک ویدیو از رفتار `Agent` در محیط بازی داشته باشید. در زمان ارائه هم حتما `Agent` آموزش دیده باشد و بتواند به خوبی عمل کند.