



به نام خدا
دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر

درس سیستم‌های هوشمند

پروژه پایانی

یادگیری تعاملی عمیق

بهمن ماه 1401

فهرست مطالب

3	محیط بازی Lunar Lander :
3	الف: بررسی دقیق الگوریتم های مفروض:
3	ب: آشنایی با محیط بازی:
3	ج: پیاده سازی الگوریتم DQN:
3	د: پیاده سازی الگوریتم DDQN:
4	ث: معیار Regret vs Convergence Rate:
4	ح: تفسیر و مقایسه نتایج:
4	ز: بررسی گرافیکی عملکرد عامل:
6	نکات تحویل:

محیط بازی Lunar Lander :

در این سوال قصد داریم در قالب حل محیط یک بازی، با دو الگوریتم پرکاربرد DQN و DDQN آشنا شویم و عاملی را طراحی نماییم که قادر باشد مسئله مفروض را بخوبی حل نماید.

الف: بررسی دقیق الگوریتم های مفروض:

در این قسمت شما میبایست در رابطه با الگوریتم های Deep Q Network و Double Deep Q Network تحقیق کنید و بصورت دقیق تئوری این الگوریتم ها و نحوه کارکرد آنها را توضیح دهید.

ب: آشنایی با محیط بازی:

محیط LunarLander را مطالعه کرده و ویژگی های آن را شرح دهید. این ویژگی ها میبایست شامل مشخصات فضای حالت، مشخصات فضای عمل و سیستم پاداش باشد.

ج: پیاده سازی الگوریتم DQN:

عملکرد عامل را با رسم پاداش تجمعی در هر episode برای batch size های 128,64,32 بررسی نمایید و برای بهترین حالت به ازای episode های 50,100,150,200,250 فیلمی از عملکرد عامل تهیه کنید.

د: پیاده سازی الگوریتم DDQN:

بعد از پیاده سازی فوق، قصد داریم که از دو جدول Q به عنوان حافظه عامل استفاده نماییم پس در این قسمت نیز عملکرد عامل را با رسم پاداش تجمعی در هر episode برای batch size های 128,64,32 بررسی نمایید و برای بهترین حالت به ازای episode های 50,100,150,200,250 فیلمی از عملکرد عامل تهیه کنید.

ث: معیار Regret vs Convergence Rate :

در ابتدا در رابطه با معیار Regret توضیح دهید و بصورت دقیق توضیح دهید این معیار چه چیزی را مورد بررسی قرار میدهد. سپس علاوه بر بررسی معیار سرعت همگرایی، پاداش بهینه معیار Regret را نیز تحلیل نمایید و مشخص کنید در انتخاب batch-size بهینه چگونه میتوانیم این معیارها را بکار بگیریم.

ج: تفسیر و مقایسه نتایج:

عملکرد مدل DQN و DDQN را با رسم پاداش تجمعی در هر episode و به ازای batch-size های برابر مقایسه نمایید.

د: بررسی گرافیکی عملکرد عامل:

برای هر دو مدل به ازای episode های 50,150,250 به صورت گرافیکی فیلم هایی از عملکرد مدل تهیه نمایید.

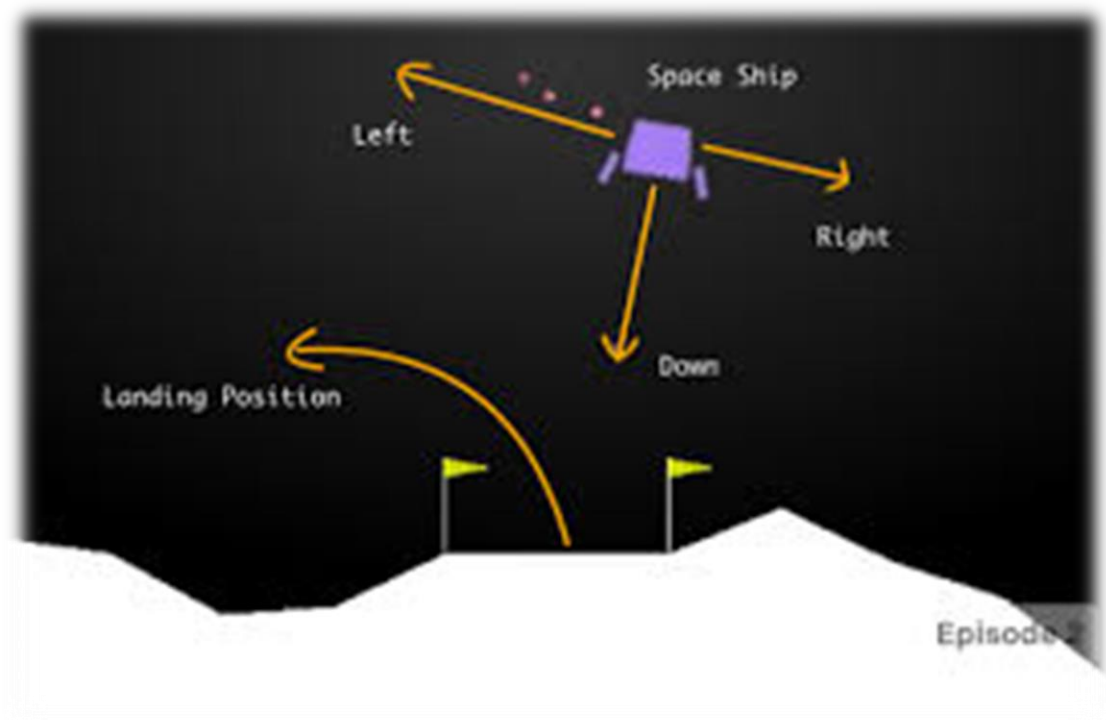


Figure1: Game Environment

نکات پروژه:

- برای تهیه خروجی نمی توانید از سرویس google colab استفاده نمایید و میبایست checkpoint های مدل را دانلود کرده و روی سیستم خود فیلم ها را تهیه کنید.
- برای پیاده سازی الگوریتم های گفته شده یک notebook ضمیمه شده است که در وهله اول شما میبایست آن را با دقت بررسی نمایید و سپس بخش های گفته شده را مطابق الگوریتم تکمیل نمایید.
- شما میبایست در طول آموزش عامل، با استفاده از تحلیل دقیق مقدار پارامترهای آموزش مانند decay epsilon و تعداد اپیزودها و هر پارامتر مورد استفاده دیگری در طول الگوریتم را بدست آورید و سپس از آنها در طول الگوریتم استفاده نمایید.

نکات تحویل:

- مهلت تحویل این تمرین 16 بهمن می باشد.
- برای انجام این تمرین تنها مجاز به استفاده از زبان برنامه نویسی پایتون هستید. در سولاتی که از شما خواسته شده یک الگوریتم را پیاده سازی کنید مجاز به استفاده از توابع آماده نمی باشید مگر اینکه در صورت سوال مجاز بودن استفاده از این توابع یا کتابخانه ها صریح ذکر شده باشد.
- کدهای مربوط به هربخش می بایست در پوشه ای با نام Codes در کنار گزارش کار شما موجود باشد. این کدها باید خوانا و به صورت مرتبط نام گذاری شده باشند، لذا توضیحات لازم را به صورت یادداشت¹ در کدهای خود قرار دهید.
- لطفا تمامی نکات و مفروضاتی که برای پیاده سازی ها و محاسبات خود در نظر می گیرید را در گزارش ذکر کنید. همچنین رعایت موازین نگارشی در گزارش توجه ویژه ای داشته باشید (بطور مثال استفاده از زیرنویس برای تصاویر و بالانویس برای جداول).
- برای پروژه هر گروه علاوه بر گزارش کتبی ملزم به ارائه گزارشی در قالب ارائه علمی خواهد بود، این ارائه از اهمیت ویژه ای در مراحل نمره دهی برخوردار است و تمام افراد گروه باید به تمام مباحث پروژه اشراف داشته باشند و به سولات مطرح شده در ارائه پاسخ دهند.
- لطفا گزارش، فایل کدها و سایر ضمائم مورد نیاز را با ترتیب نام گذاری زیر در صفحه درس در سامانه بارگذاری کنید.

FinalProject_[StudentNumber(s)].zip

- در صورت وجود هر گونه ابهام یا مشکل لطفا به مسئولان پروژه ایمیل بنزید.

[Mohammad Heydari](#)

[Erfan Hajihashemi](#)

[Oveys Delafrooz](#)

¹ Comment