



تمرین اول درس یادگیری تقویتی

مهلت تحویل تمرین: دوشنبه ۳۱ اردیبهشت ماه ۱۴۰۳

ایمیل کلاس: rliust2024@gmail.com

در این تمرین شما باید بتوانید اطلاعات مربوط به یک محیط را استخراج کرده و بر اساس آن سیاست بهینه را برای عامل داخل آن محیط بیابید. محیطی که برای این منظور استفاده می شود از صفحه گیت هاب زیر قابل دسترسی است:

<https://github.com/markub3327/flappy-bird-gymnasium/tree/main>

تمامی اطلاعات مربوط به این محیط در این صفحه فوق موجود است.

برای استفاده از این محیط نیازمند کتابخانه gymnasium هستید. این کتابخانه مشابه کتابخانه gym عمل میکند. در ادامه شما باید بتوانید با استفاده از الگوریتم یادگیری Q-Learning سیاست بهینه را بدست آورید.

برای ارسال پاسخ تمرین، فایل ipynb مربوط به google colab و فیلم رندر شده از عملکرد عامل بر اساس سیاست بهینه را در یک فایل rar با نام HW01_sudent-number.rar به ایمیل کلاس ارسال نمایید.

تذکر ۱: لطفا فیلم عملکرد عامل با سیاست بهینه را از طریق نرم افزارهای screen recorder ذخیره کنید و در آن حتما محیط google colab را اسکرول کنید تا کدهای شما نمایش داده شده و در نهایت فیلم عملکرد نمایش داده شود.

تذکر ۲: در صورتی که برای پاسخ تمرین تشخیص کپی داده شود، نمره فقط به شخصی که تمرین را سریعتر ارسال کرده داده می شود.

تذکر ۳: در صورت ارسال تمرین پس از موعد مقرر، به ازای روز اول ۵۰٪ نمره کسر شده و از روز دوم به بعد تمام نمره آن کسر خواهد شد.

موفق باشید