

به نام خدا



درس هوش مصنوعی و سیستم‌های خبره

---

## تمرین عملی پنج

---

مدرس درس:  
جناب آقای دکتر محمدی

طراحان:  
احمد رضا طهماسبی  
فاطمه عسکری

مهلت ارسال: ۱۴۰۲/۱۰/۲۲

## مقدمه

به عنوان یکی از تمرین های عملی این درس شما موضوعی که در ادامه می آید را باید با استفاده از رویکرد یادگیری تقویتی حل کنید.

## ۱ آماده سازی محیط پروژه

در این پروژه قرار هست از یادگیری تقویتی در بازی flappy-bird استفاده کنید. flappy bird gym یک پروژه نرم افزاری است که برای توسعه و آموزش الگوریتم های یادگیری تقویتی در حوزه بازی flappy bird طراحی شده است.

● [صفحه GitHub](#)

طبق توضیحات صفحه گیت هاب برای نصب از python 3.8 استفاده شده است اگر با بقیه ورژن ها به مشکل خوردید از anacond استفاده کنید برای نصب دستور زیر را بزنید:

- pip install flappy-bird-gym

با استفاده از کد زیر می توانید محیط بازی را مشاهده کنید:

```
import time
import flappy_bird_gym
env = flappy_bird_gym.make("FlappyBird-v0")

obs = env.reset()
while True:
    # Next action:
    # (feed the observation to your agent here)
    action = ... # env.action_space.sample() for a random action

    # Processing:
    obs, reward, done, info = env.step(action)

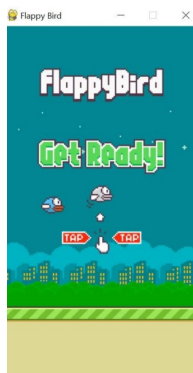
    # Rendering the game:
    # (remove this two lines during training)
    env.render()
    time.sleep(1 / 30) # FPS

    # Checking if the player is still alive
    if done:
        break

env.close()
```

## ۲ 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 بالا و پایین هستند، عبور کند. هدف این بازی افزایش امتیاز بازیکن با عبور از بیشترین تعداد لوله ها است.



در این بازی دو نوع فضای مشاهده (observation space) برای عامل فراهم میکند:

● FlappyBird-v۰

● FlappyBird-rgb-v۰

در محیط FlappyBird-rgb-v0 فضای مشاهده تصاویر RGB هستند که صفحه بازی را نمایش می دهند. در این فضا مشاهده عامل میتواند با تحلیل تصویر و استفاده از الگوریتم های یادگیری تقویتی تصمیم گیری های مربوط به بازی را انجام دهد  
**اگر کسی با این فضا مشاهده کار کند نمره امتیازی تعلق نمیگیرد.**

در محیط FlappyBird-v0 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 و شرایط بازی بگیرد.

### ۳ نحوه پیاده سازی

فایل پروژه شامل دو فایل `utils.py` و `main.py` است و فایل `main.py` شامل الگوریتم و `policy` هست توابع خواسته شده را در این فایل باید پیاده سازی کنید و عملکرد `agent` را بسنجید. و در فایل `utils.py` شامل یک سری توابع کمکی هست که می توانید از یک `qtable` آماده استفاده کنید که همان `class counter` در فایل `utils` هست.

نکته: یک مشکلی که هست **reward** همیشه ۱ برگردانده میشود به همین علت باید یک تابع مخصوص به **reward** پیاده سازی کنید باید **reward** را از طریق **info** محاسبه کنید در **info** امتیاز پرنده موجود است.

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

می توانید برای تست درستی `qtable` خود از کتابخانه `gym` پروژه `Mountain Car` استفاده کنید و اگر روی آن پروژه به درستی کار میکرد سعی کنید الگوریتم خود را بر روی `flappy bird` بهتر کنید دقت کنید تست `qtable` بر پروژه `Mountain Car` نمره ای ندارد و صرفاً جهت اطمینان درستی `qtable` خودتان هست.

#### • صفحه پروژه `MoutainCar`

می توانید برای پیاده سازی اولیه `qtable` خود از لینک های گیت هاب موجود در اینترنت استفاده کنید به عنوان مثال می توانید از لینک زیر ایده بگیرید

#### • پیاده سازی `qtable`

### قوانین:

۱. تمرین به صورت انفرادی انجام شود.
۲. نمره شما بر اساس گزارش راه طی شده برای حل مسئله و پاسخ صحیح خواهد بود لذا از هرگونه اطناب در گزارش پرهیز و به موارد خواسته شده به صورت کامل پاسخ دهید.
۳. برای تمرین از شما ارائه شفاهی گرفته خواهد شد بنابراین تسلط لازم را بر کدی که پیاده میکنید داشته باشید.
۴. در صورتی مشاهده شباهت غیرعادی بین پیاده سازی ها نمره طرفین طبق قوانین درس محاسبه خواهد شد.
۵. برای تحویل تمرین یک فایل zip شامل گزارش حل سوالات، با نام [HW\_ID\_NAME1] در سامانه LMS بارگذاری کنید.