

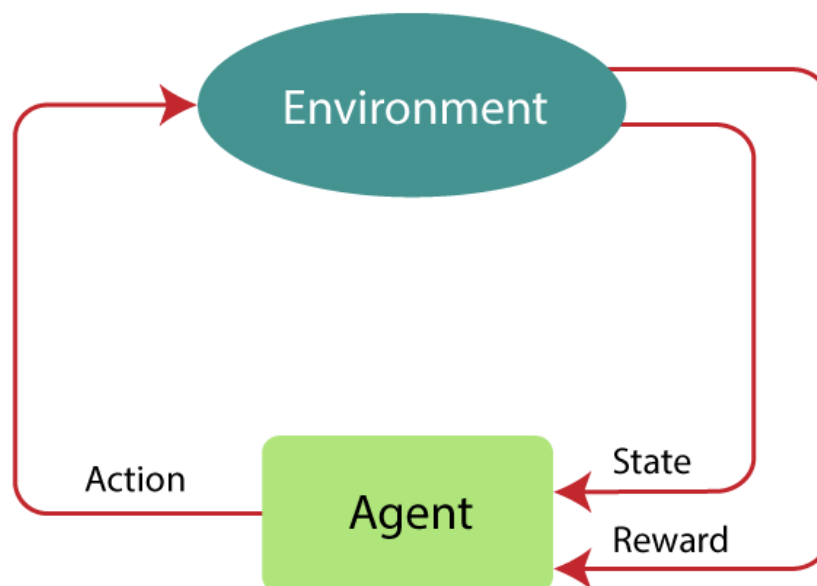


پروژه درس هوش مصنوعی ۴۰۱۲
استاد درس: دکتر سمانه حسینی
دستیاران آموزشی: نوید نصیری، علیرضا حبیبی

۱ مقدمه

همان‌طور که می‌دانید، یادگیری تقویتی یا Reinforcement Learning بخشی از ماشین‌لرنینگ است که در آن عامل از طریق تعامل با محیط برای دستیابی به استراتژی بهینه جهت رسیدن به اهداف خود، یاد می‌گیرد. این روش برخلاف یادگیری با نظارت، به داده و برچست برای یادگیری نیاز ندارد در عوض، از محیط و سیستم پاداش یا جریمه دریافت می‌کند تا بر اساس آن بیاموزد تصمیمات بهتری بگیرد.

برای مثال اگر بازی ویدیویی Pacman را در نظر بگیرید، با حرکت کاراکتر بازی به هر سمت، خود را در یک موقعیت جدیدی قرار می‌دهد و بر اساس آن موقعیت جدید، مشاهدات مربوط به آن موقعیت را دریافت کرده و همچنین امتیازی از محیط دریافت می‌کند که بر اساس آن می‌تواند بفهمد که آیا این حرکت آن درست بوده یا نه. مثلاً اگر این حرکت باعث خورده شدن آن شده باشد می‌توانیم امتیاز -100 را به آن بدهیم و اگر باعث خوردن غذاهای کوچک شده باشد، امتیاز $+10$ را به آن بدهیم، اگر باعث خوردن غذاهای پر انرژی شده باشد، امتیاز $+50$ را به آن بدهیم و در غیر این صورت امتیاز 0 را نصیب آن کنیم. در این صورت با هر دست بازی کردن، می‌تواند بر اساس نوع بازی‌ای که در دست قبل انجام داده بود، بهترین حرکت را در این دست از بازی انجام دهد تا امتیاز بیشتری نسبت به قبل، دریافت کند.



۲ شرح پروژه

شرکت Open AI یک شرکت در حوزه هوش مصنوعی است که در زمینه‌های پردازش زبان طبیعی، بینایی کامپیوتر، رباتیک و یادگیری تقویتی در حال تحقیق و توسعه است. یکی از محصولات جذاب این شرکت [gymlibrary](#) هست که در آن بستر بسیار مناسبی به کمک بازی‌های متنوع قرار داده است تا علاقه‌مندان حوزه یادگیری تقویتی بتوانند الگوریتم‌های مختلفی را بر روی آن‌ها پیاده سازی کنند و خروجی آن را ببینند. در اینجا دو مورد از این بازی‌ها معرفی می‌شود و شما می‌توانید یکی از آن‌ها را برای پروژه انتخاب کرده و انجام دهید. بدیهی است که نمره تعلق گرفته به هر کدام از آنها بر اساس سطحشان، متفاوت خواهد بود. برای آشنایی بیشتر با این کتابخانه می‌توانید از [مستندات](#) سایت استفاده کنید.

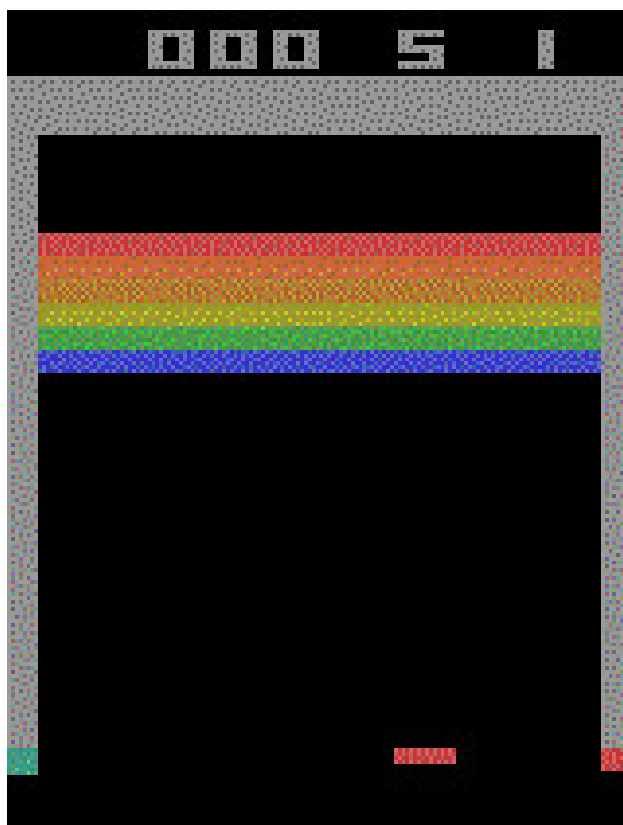
۱.۲ FrozenLake

این بازی شامل عبور از یک دریاچه یخزده از نقطه شروع به نقطه پایان و بدون افتادن به درون گودال است. در این بازی عامل که شناختی نسبت به محیط ندارد، شروع به حرکت می‌کند و به کمک الگوریتمی که شما انتخاب می‌کنید، محیط را شناخته و اقدام‌های خود را بر اساس شناخت خود انجام می‌دهد. در این بازی، اقدامات عامل شامل حرکت به بالا، پایین، چپ و راست است که در هر مرحله می‌تواند یکی از این اعمال را انجام داده و به state جدید برود. مشاهدات عامل در این محیط، موقعیت آن در نقشه است. برای مشاهده اطلاعات کامل در رابطه با این بازی می‌توانید به [مستندات](#) مربوط به آن مراجعه کنید.



۲.۲ Breakout

این بازی یکی از بازی‌های معروف آتاری است که در آن عامل یک چوب است که به چپ یا راست حرکت می‌کند و توپ را به دیوار موجود در بالای صفحه می‌زند. هدف تخریب کامل آن دیوار است. در این بازی اعمالی که عامل مجاز به انجام آن‌ها است، حرکت به چپ، راست، بی‌حرکت ایستادن و یا شروع بازی است. همچنین مشاهده‌ای که عامل در هر state دارد، پیکسل‌های فریم بازی است که نشان دهنده رنگ هر پیکس در آن state است. برای مشاهده اطلاعات کامل در رابطه با این بازی می‌توانید به [مستندات](#) مربوط به آن مراجعه کنید.



۳ نکات قابل توجه و نحوه تحویل

در این مستندات توضیح مفصل در رابطه با بازی‌ها و همچنین روش حل مورد انتظار آنها داده نشده است چون در این رابطه به مقدار زیادی منبع می‌توانید در بستر اینترنت بیابید. برای همین چگونگی ارائه حضوری شما بخش اصلی از نمره را در بر خواهد گرفت و تحویل پروژه بدون ارائه، شامل نمره‌ای نخواهد شد. به نکات زیر در تحویل و انجام پروژه دقت کنید.

- در صورت انتخاب موضوع اول، انتظار می‌رود که از الگوریتم QLearning استفاده کنید و نسبت به پیاده سازی و طرز کار آن آشنایی کامل را داشته باشید.
- از آن جایی که پیچیدگی دو پروژه یکسان نیستند، نمره مربوط به هر کدام نیز یکسان نیستند. لذا معیار دریافت نمره کامل پروژه، پیاده سازی و ارائه کامل پروژه دوم است. اما با این حال نمره دریافتی شما، کاملاً وابسته به نوع الگوریتم و ایده‌های مورد استفاده شما خواهد بود.
- پروژه به صورت تک‌نفره بوده و شامل ارائه حضوری خواهد بود. زمان دقیق ارائه هر نفر به زودی اعلام می‌شود.
- آخرین زمان تحویل پروژه تاریخ ۱۰ تیر بوده و قابل تمدید نمی‌باشد. لذا نسبت به آپلود تمام فایل‌ها تا این تاریخ اقدام فرمایید. فایل‌های پروژه شامل کد مربوط به پروژه، فایل پی‌دی اف شامل توضیحات مربوط به کد و الگوریتم مورد استفاده، یک ویدیوی حداکثر ۵ دقیقه‌ای شامل اجرای یک سناریوی موفق بازی توسط عامل می‌باشد.
- آپلود کد پایتون مربوط به پروژه در کلاس [کویرا](#) الزامی است.