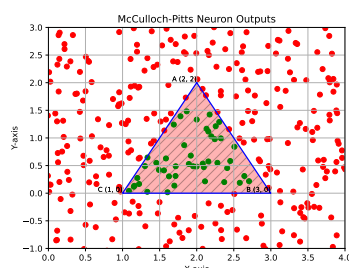




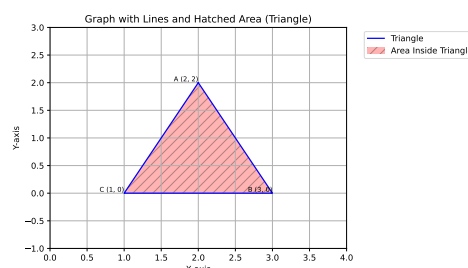
مینی پروژه شماره چهار

۱ پرسش یک

به کمک یک نورون ساده یا پرسپترون یا نورون McCulloch-Pitts^۱ شبکه‌ای طراحی کنید که بتواند ناحیه هاشورزده داخل مثلثی که در نمودار **شکل ۱ (آ)** نشان داده شده را از سایر نواحی تفکیک کند. پس از انجام مرحله طراحی شبکه (که می‌تواند به صورت دستی انجام شود)، برنامه‌ای که در **این دفترچه‌کد** و در کلاس برای نورون McCulloch-Pitts آموخته‌اید را به گونه‌ای توسعه دهید که ۲۰۰۰ نقطه رندوم تولید کند و آن‌ها را به عنوان ورودی به شبکه طراحی شده توسط شما دهد و نقاطی که خروجی «۱» تولید می‌کنند را با رنگ سبز و نقاطی که خروجی «۰» تولید می‌کنند را با رنگ قرمز نشان دهد. خروجی تولید شده توسط برنامه شما باید به صورتی که در **شکل ۱ (ب)** نشان داده شده است باشد (به محدوده عددی محورهای x و y هم دقت کنید). اثر اضافه کردن دو تابع فعال‌ساز مختلف به فرآیند تصمیم‌گیری را هم بررسی کنید.



(ب) خروجی مطلوب برنامه



(آ) نمودار هاشورزده مورد سوال

شکل ۱: نمودارهای مربوط به سوال اول و خروجی برنامه.

۲ پرسش دو

در این سوال به شبیه‌سازی مقاله weather forecasting based on a real-time collaborative machine learning system with multiple predictor locations^۱ نواحی مختلف برای پیش‌بینی آب‌وهوای یک ناحیه استفاده می‌کند. در این تحقیق الگوریتم‌های مختلفی مورد بررسی قرار گرفته است. هر یک از الگوریتم‌ها نیز برای چند حالت مختلف آموزش داده شده‌اند؛ یعنی مدل‌های مختلف برای داده هدف متفاوت. در این سؤال قصد داریم علاوه بر تمرین پیاده‌سازی شبکه عصبی برای مسئله رگرسیون^۲، با collaborative machine learning نیز آشنا شویم.

^۱ تشخیص اینکه با کدام روش می‌توانید این کار را انجام دهید با شماست.
^۲ regression

۱.۲ دادگان

۱.۱.۲

درباره دیتاست جمع‌آوری شده توسط مقاله به صورت مختصر توضیح دهید.

۲.۱.۲

داده این تمرین در فایل `weather_prediction_dataset.csv` قرار دارد. در این پیاده‌سازی از دادگانی غیر از دادگان مقاله مورد مطالعه استفاده می‌شود. در مقاله از نواحی نزدیک به هم برای پیش‌بینی آب‌وهوا استفاده شده است. برای شبیه بودن این تمرین به مقاله از دادگان شهرهای فرانسه که در فایل داده وجود دارد استفاده می‌کنیم. چه شهرهایی از فرانسه در این داده موجود هستند؟ آن داده‌ها را ذخیره کرده و داده مربوط به شهرهای دیگر را حذف کنید. (لینک دادگان)

۳.۱.۲

چند نمونه^۳ در این داده وجود دارد؟ این دادگان چه بازه زمانی را شامل می‌شوند؟ مقاله چه پیش‌پردازش‌هایی را روی داده خود اعمال کرده است؟ آن‌ها را روی داده خود اعمال کنید.

۴.۱.۲

مقاله برای محاسبه متغیر مورد نظر در لحظه t از مقادیر موجود در روزهای $t-1, t-2, \dots$ استفاده می‌کند. یعنی ورودی مدل مربوط به چندین روز بوده و خروجی آن مربوط به تنها یک روز است. در ادامه برای آموزش مدل خود باید دادگان را به این فرمت در آورید. یعنی y شما تنها مربوط به یک روز است ولی x مربوط به چند روز قبل از y است. شما برای آموزش مدل خود باید داده ای به صورت بیان شده ایجاد کنید.

داده‌های مربوط به سال ۲۰۰۹ را جدا کرده و به عنوان دادگان آزمون استفاده نماید. دادگان آموزش و آزمون را با پنجره‌های انتخاب شده بسازید؛ مثلاً اگر اندازه پنجره انتخاب شده ۵ با همپوشانی ۴ باشد، داده آزمون باید از شکل $365 \times n$ به شکل $361 \times 5 \times n$ درآید.

۲.۲ آموزش مدل

در این بخش باید یک شهر را انتخاب کرده و تمامی مدل‌های خواسته شده را برای آن شهر پیاده‌سازی کنید.

۱.۲.۲

مفهوم collaborative machine learning را توضیح دهید. collaborative machine learning در چهارچوب این مقاله به چه صورت استفاده شده است؟ شرح دهید.

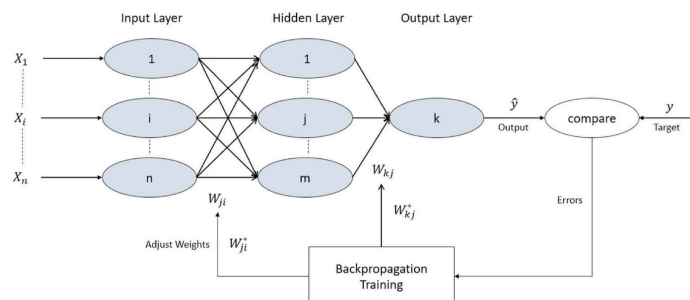
۳.۲

در این مرحله باید یک شبکه عصبی طراحی از صفر طراحی کرده و مدل‌هایی با استفاده از داده‌هایی که ایجاد کردید آموزش دهید.

- معماری شبکه: شبکه عصبی طراحی کنید که ساختار شکل ۲ زیر را داشته باشد.
یعنی شبکه دارای یک لایه پنهان و یک لایه تصمیم‌گیری نهایی است.
نکته: برای محاسبه ویژگی‌های مختلف می‌توانید تنها از یک مدل با چندین خروجی یا از چند مدل با یک خروجی استفاده کنید.
- طراحی شبکه: برای طراحی شبکه خود می‌توانید از کتابخانه tensorflow یا pytorch استفاده کنید.
- پارامترهای مختلف شبکه عصبی باید شبکه خود را برای ۳ حالت مختلف نرخ یادگیری آموزش دهید.

۱. 10^0

^۳sample



شکل ۲: ساختار شبکه عصبی

۲. 10^{-3}

۳. 10^{-8}

از SGD به عنوان تابع الگوریتم بهینه‌ساز و از MSE به عنوان تابع هزینه استفاده نمایید. از ReLU برای تابع فعال‌سازی لایه پنهان و از تابع Linear برای لایه نهایی استفاده کنید. مدل خود را برای ۲۰۰ epoch آموزش دهید.

• دیگر نیازمندی‌ها: باید موارد زیر در حین آموزش مدل نمایش داده شوند:

- progress bar
- خطا آموزش و آزمون
- امتیازی: نمودار خطا آموزش و آزمون

۴.۲ نمایش و تحلیل

۱.۴.۲

نمودار خطا برای داده آموزش و آزمون یک مدل را برای تمامی حالات نرخ یادگیری نمایش دهید.

۲.۴.۲

با استفاده از نمودارهای رسم شده در بخش قبل، تاثیر اندازه نرخ یادگیری را بر عملکرد مدل تحلیل کنید.

۵.۲ مدل جدید

شبکه عصبی را عمیق‌تر کرده و بار دیگری تمامی مراحل آموزش مدل و نمایش و تحلیل را دوباره انجام دهید. تحلیل کنید که مدل عمیق‌تر چه تاثیری بر عملکرد دارد.

۶.۲ بررسی نهایی

تمامی وزن‌ها را در حین آموزش مدل ذخیره کرده و تغییر مقدار وزن‌ها را برای یک یا چند وزن به صورت جداگانه نمایش دهید. تاثیر نرخ یادگیری بر تغییرات وزن مدل حین آموزش را تحلیل کنید. در مورد تاثیر استفاده از مفاهیم زیر نیز با نیز با استدلال و نتیجه مشخص بحث کنید:

- Data Preprocessing
- Weight Initialization – (مطابق پیشنهادهای دوره استنفورد)
- Batch Normalization
- Regularization
- Dropout

به انتخاب خود به یکی از پرسش‌های سه یا چهار پاسخ دهید.

۳ پرسش سه: حل دنیای Wumpus

Wumpus World یک مسئله کلاسیک در هوش مصنوعی و یادگیری تقویتی است که شامل یک محیط مبتنی بر شبکه است که در آن یک عامل باید برای یافتن طلا حرکت کند و در عین حال از خطراتی مانند چاله ها و Wumpus اجتناب کند.

- اهداف پیمایش در شبکه Grid: عامل باید یاد بگیرد که به طور موثر در شبکه حرکت کند.
- اجتناب از خطرات: عامل باید یاد بگیرد که از چاله ها و Wumpus اجتناب کند.
- جمع آوری طلا: عامل باید طلا را پیدا کرده و جمع آوری کند.
- کشتن Wumpus: عامل می تواند برای کشتن Wumpus تیری شلیک کند و آن را به عنوان تهدید از بین ببرد.
- راه اندازی محیط شبکه: یک شبکه 4×4 که در آن هر سلول می تواند خالی باشد، حاوی یک گودال، Wumpus یا طلا باشد.
- فضای اکشن ها: حرکت به بالا، پایین، چپ، راست.
- یک فلش را در هر یک از چهار جهت (بالا، پایین، چپ، راست) شلیک کنید (امتیازی).
- تصورات: Wumpus در شبکه با هر تغییر اکشن به اندازه یک خانه در راستای چپ، راست، بالا یا پایین حرکت می کند (امتیازی).
- فضای Reward:

- ۱۰۰+ برای گرفتن طلا

- ۱۰۰۰- برای افتادن در گودال یا خورده شدن توسط Wumpus

- ۵۰+ برای کشتن Wumpus (امتیازی)

- ۱- برای هر حرکت

- تعریف محیط: یک شبکه 4×4 با موقعیت های دلخواه برای چاله ها، Wumpus و طلا ایجاد کنید. حالت اولیه و حالت های ممکن را بعد از هر عمل تعریف کنید.

• تنظیم پارامترها:

- نرخ یادگیری: ۰.۱

- ضریب تخفیف: ۰.۹

- نرخ اکتشاف: از ۱.۰ شروع می شود و در طول زمان کوچک میشود.

با توجه به موارد کلی گفته شده راجع به مسئله، موارد زیر را پاسخ دهید.

آ. برای این مسئله یک بار با روش Q-learning و یک بار با روش Deep Q-learning عاملی را طراحی کرده و آموزش دهید.

ب. عملکرد Policy:

- پاداش تجمعی را در اپیزودها برای هر دو عامل Q-learning و DQN ترسیم کنید. چگونه عملکرد عامل در طول زمان بهبود می یابد؟
- میانگین پاداش در هر اپیزود را برای هر دو عامل پس از ۱۰۰۰ اپیزود مقایسه کنید. کدام الگوریتم عملکرد بهتری داشت؟

ج. بحث کنید که چگونه نرخ اکتشاف اپسیلون بر فرآیند یادگیری تأثیر می گذارد. وقتی اپسیلون بالا بود در مقابل وقتی کم بود چه چیزی را مشاهده کردید؟

د. کارایی یادگیری:

- چند اپیزود طول کشید تا عامل Q-learning به طور مداوم طلا را بدون افتادن در گودال یا خورده شدن توسط Wumpus پیدا کند؟
- کارایی یادگیری Q-learning و DQN را مقایسه کنید. کدام یک Policy بهینه را سریعتر یاد گرفت؟
- معماری شبکه عصبی مورد استفاده برای عامل DQN را شرح دهید. چرا این معماری را انتخاب کردید؟

۴ پرسش چهار: Deep Q-Learning برای محیط Lunar Lander

در این سوال از شما می‌خواهیم با استفاده از روش Deep Q-Learning برای محیط Lunar Lander یک عامل طراحی کرده و آموزش دهید که بتواند مسأله را حل کند. برای حل این تمرین می‌توانید از این [دفترچه‌کد](#) استفاده کرده و در صورت نیاز آن را تکمیل کرده و تغییر دهید. اگر از دفترچه‌کد ارائه‌شده و کدهای آماده استفاده می‌کنید، لازم است که تمام اجزای کد را با دقت توضیح دهید. دقت داشته باشید که هدف، توضیح روند عملکردی کد است، نه تحلیل خشک و لغتی برآمده از ChatGPT!

آ. در مورد محیط Lunar Lander مطالعه کرده و به‌صورت خلاصه ویژگی‌های آن را شرح دهید. ویژگی‌های مدنظر عبارتند از مشخصات فضای حالت، مشخصات فضای عمل و سیستم پاداش.

ب. عملکرد عامل را با رسم پاداش تجمعی در هر episode و برای batch size های ۳۲، ۶۴ و ۱۲۸ بررسی کنید. تنها برای بهترین حالت به ازای episode های ۵۰، ۱۰۰، ۱۵۰، ۲۰۰ و ۲۵۰ فیلمی از عملکرد عامل تهیه کنید. در صورتی که عملکرد عامل به ازای هر سه مقدار batch size مشابه یکدیگر شد، یکی از آن‌ها را به دلخواه به عنوان بهترین حالت انتخاب کنید. در رابطه با انتخاب بهترین حالت علاوه بر معیار سرعت همگرایی به پاداش بهینه معیار regret را نیز به صورت شهودی بررسی کنید.

ج. عملکرد مدل DQN و DDQN را با رسم پاداش تجمعی در هر episode و به ازای batch size برابر مقایسه کنید. برای هر دو مدل به ازای episode های ۱۰۰ و ۲۵۰، فیلمی از عملکرد مدل تهیه کنید. هر بار آموزش عامل با استفاده از GPU های رایگان گوگل کولب حدوداً بین ۱۰ تا ۱۵ دقیقه زمان لازم خواهد داشت. برای تهیه خروجی می‌توانید checkpoint های مدل را دانلود کرده و روی سیستم خود فیلم‌ها را تهیه کنید.

Algorithm 1 Deep Q-learning with Experience Replay

```

Initialize replay memory  $\mathcal{D}$  to capacity  $N$ 
Initialize action-value function  $Q$  with random weights
for episode = 1,  $M$  do
    Initialize sequence  $s_1 = \{x_1\}$  and preprocessed sequenced  $\phi_1 = \phi(s_1)$ 
    for  $t = 1, T$  do
        With probability  $\epsilon$  select a random action  $a_t$ 
        otherwise select  $a_t = \max_a Q^*(\phi(s_t), a; \theta)$ 
        Execute action  $a_t$  in emulator and observe reward  $r_t$  and image  $x_{t+1}$ 
        Set  $s_{t+1} = s_t, a_t, x_{t+1}$  and preprocess  $\phi_{t+1} = \phi(s_{t+1})$ 
        Store transition  $(\phi_t, a_t, r_t, \phi_{t+1})$  in  $\mathcal{D}$ 
        Sample random minibatch of transitions  $(\phi_j, a_j, r_j, \phi_{j+1})$  from  $\mathcal{D}$ 
        Set  $y_j = \begin{cases} r_j & \text{for terminal } \phi_{j+1} \\ r_j + \gamma \max_{a'} Q(\phi_{j+1}, a'; \theta) & \text{for non-terminal } \phi_{j+1} \end{cases}$ 
        Perform a gradient descent step on  $(y_j - Q(\phi_j, a_j; \theta))^2$ 
    end for
end for

```

شکل ۳: شبکه‌کد الگوریتم DQN

Basic Q-Learning

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(R_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t))$$

Double Q-Learning

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(R_{t+1} + \gamma \overbrace{Q(s_{t+1}, a_t)}^{\text{estimated/expected Q-value}} - Q(s_t, a_t))$$

$$\overbrace{a}^{\text{estimated/expected Q-value}} = \max_a Q(s_{t+1}, a)$$

$$Q_{\text{estimated}} = \overbrace{Q(s_{t+1}, a)}^{\text{estimated/expected Q-value}}$$

شکل ۴: تفاوت مدل DDQN با DQN

در انجام این مینی‌پروژه حتماً به نکات زیر توجه کنید:

- موعد تحویل این تمرین، ساعت ۱۸:۰۰ روز ۲۵ خردادماه ۱۴۰۴ است.
- برای گزارش لازم است که پاسخ هر سوال و زیربخش‌هایش به‌ترتیب و به‌صورت مشخص نوشته شده باشند. بخش زیادی از نمره به توضیحات دقیق و تحلیل‌های کافی شما روی نتایج بستگی خواهد داشت.
- لازم است که در صفحه اول گزارش خود لینک مخزن گیت‌هاب را و گوگل‌کولب مربوط به مینی‌پروژه خود را درج کنید. درخصوص گیت‌هاب، یک مخزن خصوصی درست کنید و آی‌دی‌های MJAHMADEE و AliBagheriNejad را به‌عنوان Collaborator به مخزن اضافه کنید. پروژه‌های گیت‌هاب می‌بایست در انتهای ترم پابلیک شوند. درمقابل، لینک گوگل‌کولب را در حالتی که دسترسی عمومی دارد به اشتراک بگذارید. دفترچه‌کد گوگل‌کولب باید به‌صورت منظم و با بخش‌بندی مشخص تنظیم شده باشد و خروجی سلول‌های اجراشده قابل مشاهده باشد. در گیت‌هاب نیز یک مخزن برای درس و یک پوشه مجزا برای هر مینی‌پروژه ایجاد کنید.
- (آموزش پرایوت‌کردن مخزن گیت‌هاب و آموزش افزودن Collaborator به مخزن گیت‌هاب)
- هرچا از دفترچه‌کد گوگل‌کولب شما نیاز به فراخوانی فایلی خارج از محیط داشت، مطابق آموزش‌های ارائه‌شده ملزم هستید از دستور `gdown` استفاده کنید و مسیرهای فایل‌ها را طوری تنظیم کنید که صرفاً با اجرای سلول‌های کد، امکان فراخوانی و خواندن فایل‌ها توسط هر کاربری وجود داشته باشد.
- در تمامی مراحل تعریف داده و مدل و هرچای دیگری که مطابق آموزش‌های ویدیویی و به لحاظ منطقی نیاز است، Random State را برابر با دو رقم آخر شماره دانشجویی خود در نظر بگیرید.
- استفاده از ابزارهای هوشمند (مانند ChatGPT) در کمک‌گرفتن برای بهبود کدها مجاز است؛ اما لازم است تمام جزئیات مواردی که در خروجی‌های مختلف گزارش خود عنوان می‌کنید را به خوبی خوانده، درک و تحلیل کرده باشید. استفاده از این ابزارهای هوشمند در نوشتن گزارش و تحلیل‌ها ممنوع است.
- در جاهایی که با توجه به دو رقم آخر شماره دانشجویی خود محدود به انتخاب عدد، متغیر و یا داده‌ای خاص شده‌اید، برای تست‌های اضافه‌تر و نمایش بهبود در نتایج خود، مجاز هستید از مقادیر دیگر هم استفاده کنید.
- رعایت نکات بالا به حرفه‌ای‌تر شدن شما کمک خواهد کرد و اهمیتی معادل مطالب درسی فراگرفته‌شده دارد؛ بنابراین، در صورت عدم رعایت هریک از این نکات، از نمره تمرین شما کسر خواهد شد.
- آی‌دی پرسش هرگونه سوال

منابع

[1] <https://github.com/MJAHMADEE/MachineLearning2024W>