دانشکده‌ی مهندسی کامپیوتر

# یادگیری عمیق

نیم‌سال دوم ۰۰ـ۹۹

مدرس: حمید بیگی

تمرین سری پنجم | یادگیری عمیق | زمان تحویل: ۲۵ تیرماه

نکات زیر را رعایت کنید:
فایل گزارش را به همراه تمامی کدها در یک فایل فشرده و با عنوان #HW5_STD در سایت Quera.ir بارگذاری نمایید.
بخش‌های پیاده‌سازی مربوط به هر سوال را در فایل مربوطه با شماره‌ی آن سوال قرار دهید. برای مثال تمامی بخش‌های پیاده‌سازی سوال چهارم را در فایل Q۴ قرار دهید.
سوالات خود را از طریق Piazza مطرح کنید.

## مسئله‌ی ۱. Sum-Product Networks (SPN) (۲۵ نمره)

(آ) چگونه از ساخت شبکه‌ی Max-Product Network در جهت استنتاج MPE میتوان استفاده نمود؟

(ب) مدل‌های SPN و Bayes Net را از لحاظ مرتبه‌ی زمانی استنتاج‌های marginal، conditional و MPE مقایسه نمایید.

(ج) آیا مدل SPN را می‌توان به مدل Bayesian Network تبدیل نمود؟ بالعکس چطور؟ توضیح دهید.

(د) چگونگی استفاده از متغیر پیوسته در SPN را شرح دهید.

(ه) توضیح دهید چگونه می‌توان از شبکه‌ی SPN به منظور یادگیری بازنمایی استفاده نمود.

## مسئله‌ی ۲. Dual Learning (۱۰ نمره)

(آ) تفاوت الگوریتم‌های dual unsupervised learning و dual supervised learning را شرح دهید.

(ب) معماری شبکه‌ی DualGAN و کاربرد آن را توضیح دهید [۱].

## مسئله‌ی ۳. Deep Q-Netwok (DQN) (۲۰ نمره)

به سوالات زیر پیرامون الگوریتم DQN پاسخ دهید.

(آ) دلیل Off Policy بودن الگوریتم Q-Learning را بیان کرده و مزایای این خصوصیت را ذکر کنید.

(ب) تفاوت الگوریتم‌های DQN و Q-Learning را ذکر کرده و مزایا و معایب آنها نسبت به یکدیگر را بیان کنید.

(ج) در الگوریتم DQN از تکنیک‌های Exprience Replay و Target Network جهت بهبود روند آموزش استفاده شده است. دلیل استفاده از هر کدام از این تکنیک‌ها و تاثیر آنها بر روند آموزش را توضیح دهید.

(د) الگوریتم Double DQN یکی از الگوریتم‌های ارائه شده جهت بهبود الگوریتم DQN است. روش استفاده شده در این الگوریتم و تاثیر آن بر روند آموزش را توضیح داده و تفاوت این روش با تکنیک Target Network را بیان کنید.

## مسئله‌ی ۴. پیاده سازی **DQN** (۳۵ نمره)

در این سوال می‌خواهیم مدل DQN را به صورت کامل از پایه پیاده سازی کرده و آن را حهت کنترل پاندول معکوس آموزش دهیم. موارد زیر را هنگام پیاده سازی مدل در نظر بگیرید.

(آ) پیاده سازی مدل را از پایه انجام داده و از تکنیک‌های Exprience Replay و Target Network بهره ببرید.

(ب) از محیط CartPole-v1 در کتابخانه Gym جهت آموزش مدل استفاده کنید.

(ج) مدل را حداقل ۵۰۰ اپیزود آموزش داده و نمودار پاداش کل در هر اپیزود را رسم کنید. این نمودار روند آموزش را نشان داده و باید به صورت کلی حرکت صعودی داشته باشد.

(د) توجه داشته باشید که تنها می‌توانید از کتابخانه Pytorch جهت پیاده سازی و آموزش شبکه‌های عمیق مدل استفاده کنید.

## مسئله‌ی ۵. **Deep Deterministic Policy Gradient (DDPG)** (۱۰ نمره)

الگوریتم DDPG صورت بهبود یافته الگوریتم DQN برای آموزش در محیط‌های با فضای حالت و عمل پیوسته است. به سوالات زیر در رابطه با این الگوریتم پاسخ دهید.

(آ) از چه روشی جهت به روز رسانی شبکه عامل استفاده شده است؟

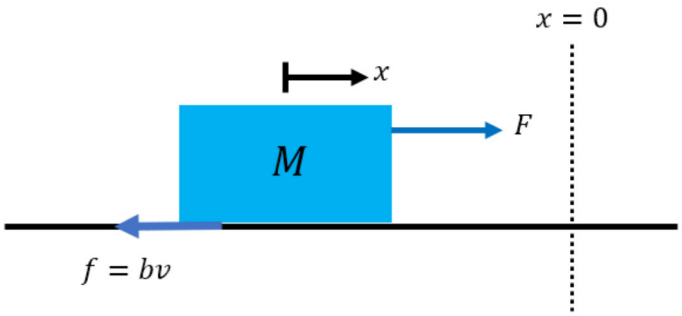
(ب) جهت حل چالش Exploration چه ایده‌ای به کار گرفته شده است؟

## مسئله‌ی ۶. **Custom Gym Environment** (+۱۰ نمره)

هدف این سوال ایجاد یک Custom Environment در قالب کتابخانه Gym است. در این مسئله یک جسم با جرم $M$ بر روی یک سطح با اصطکاک ویسکوز حرکت می‌کند.
که متغیر $x$ مکان جسم و $v$ سرعت جسم است. معادله حرکت جسم به صورت زیر است:

$$M\ddot{x} + b\dot{x} = F$$

هدف مسئله طراحی کنترلی است که بتواند از طریق نیروی $F$ این جرم را در نقطه $x = 0$ با سرعت صفر در کمترین زمان ممکن متوقف کند. برای سادگی مسئله و کوانتیزه شدن، نیروی $F$ را به صورت نرمالیز شده نسبت به جرم جسم به صورت $F = M$ در نظر می‌گیریم. با در نظر گرفتن ضریب $k = \frac{b}{M}$ و ورودی نیرو




$a = \frac{F}{M}$ و در نظر گرفتن مکان جسم به عنوان جالت اول ($s_۱$) و سرعت جسم به عنوان جالت دوم ($s_۲$) معادلات زمان گسسته دینامیک سیستم در زمان $t$ با زمان نمونه برداری $T$ به صورت زیر است:

$$s_۱(t+۱) = s_۱(t) + T(s_۲(t))$$
$$s_۲(t+۱) = s_۲(t) + T(-k.s_۲(t) + a)$$

در واقع Agent یک شبکه عصبی است که با گرفتن دو ورودی مکان و سرعت جسم، خروجی نیروی پیوسته بین $-۱$ تا $+۱$ را می‌دهد. محیط مسئله را در قالب محیط کتابخانه Gym پیاده کنید. برای این کار از محیط‌های آماده Gym مانند CartPole-v1 استفاده کرده و ایده بگیرید.

# References

[۱] Zili Yi، Hao Zhang، Ping Tan، and Minglun Gong. Dualgan: Unsupervised dual learning for image-to-image translation. In Proceedings of the IEEE international conference on computer vision، pages ۲۸۴۹–۲۸۵۷، ۲۰۱۷.