



تمرین یازدهم
درس هوش مصنوعی

نام استاد: دکتر محمدی
دستیاران آموزشی مرتبط:
علیرضا خیرخواه
مهلت تحویل: 30 دی

حل یک maze (Solving a Maze)

هدف از این تکلیف، آشنایی با مفاهیم اساسی یادگیری تقویتی (RL) از طریق پیاده‌سازی یک الگوریتم یادگیری-Q است. شما یک عامل RL را برای پیمایش یک maze و رسیدن به هدف در حالی که پاداش‌های تجمعی را به حداکثر می‌رساند، آموزش خواهید داد. در پایان این تکلیف:

- اصول اولیه یادگیری تقویتی و یادگیری-Q را درک خواهید کرد.
- نحوه نمایش محیط‌ها و پاداش‌ها در RL را یاد خواهید گرفت.
- تجربه عملی در آموزش یک عامل RL برای حل یک مسئله کسب خواهید کرد.

1. شرح مسئله

در این تکلیف، شما یک عامل یادگیری-Q را برای حل یک maze که به صورت شبکه‌ای نمایش داده شده است، طراحی خواهید کرد. maze شامل اجزای زیر است:

- نقطه شروع (S): موقعیت اولیه عامل.
- هدف (G): موقعیت هدف که عامل باید به آن برسد.
- دیوارها (W): موانع غیرقابل عبور.
- سلول‌های خالی (.) : فضاهای قابل پیمایش.

عامل بر اساس اقدامات خود پاداش دریافت می‌کند و باید یک سیاست بهینه را برای رسیدن به هدف به طور کارآمد یاد بگیرد. شما همچنین باید نحوه مدیریت موارد خاص، مانند برخورد عامل با دیوارها یا حرکت به خارج از شبکه را در نظر بگیرید.

2. راهنمایی:

- تابع پاداش را به دقت تعریف کنید تا تعادل بین اکتشاف و بهره‌برداری حفظ شود.
- با یک maze کوچک و ساده شروع کنید و به تدریج به سمت محیط‌های پیچیده‌تر پیش بروید.
- الگوریتم را با پارامترهای مختلف مانند نرخ یادگیری (α) و ضریب تخفیف (γ) آزمایش کنید.

3. محیط maze

S	.	.	W	G
.	W	.	W	.
.	W	.	.	.
.	.	.	W	.
.	W	W	.	.

maze به صورت یک شبکه 5×5 نمایش داده می‌شود که هر سلول متناظر با یک حالت است.

مثال رو به رو چیدمان را نشان می‌دهد:

توجه: شما در این تمرین باید برای یک شبکه 7×7 مدل آموزش بدهید.



تمرین یازدهم
درس هوش مصنوعی

نام استاد: دکتر محمدی

دستیاران آموزشی مرتبط:

علیرضا خیرخواه

مهلت تحویل: 30 دی

4. قوانین محیط

- عامل از سلول مشخص شده با S شروع می‌کند.
- هدف رسیدن به سلول مشخص شده با G است.
- عامل می‌تواند به بالا، پایین، چپ یا راست حرکت کند.
- دیوارها (W) غیرقابل عبور هستند و اگر عامل سعی کند به آنها حرکت کند، جریمه می‌شود.

❖ پاداش‌ها:

- $10+$ برای رسیدن به هدف
- $1-$ برای هر قدم
- $10-$ برای برخورد با دیوار

5. راهنمای پیاده‌سازی

- نمایش maze به صورت یک آرایه یا لیست دو بعدی در پایتون.
- ایجاد تابعی برای شبیه‌سازی رفتار محیط، شامل پاداش‌ها و انتقال‌ها.
- پیاده‌سازی بررسی‌های مرزی برای جلوگیری از خارج شدن عامل از محدوده.
- تجسم maze و پیشرفت عامل برای اشکال‌زدایی و درک رفتار آن.

6. ایده‌های امتیاز اضافی

- اضافه کردن اهداف چندگانه، مانند جمع‌آوری اقلام در maze قبل از رسیدن به هدف.
- پیاده‌سازی دیوارها یا موانع پویا که در طول آموزش تغییر موقعیت می‌دهند.
- آزمایش با mazeهای بزرگتر یا نامنظم برای افزایش پیچیدگی.

7. وظیفه شما

- I. تعریف محیط maze در پایتون با استفاده از نمایش مبتنی بر شبکه.
- II. پیاده‌سازی الگوریتم یادگیری Q برای آموزش عامل.
- III. آموزش عامل در طول چندین قسمت و تجسم سیاست یادگرفته شده آن.
- IV. تحلیل تأثیرات ابرپارامترهای مختلف بر عملکرد عامل.

8. تحویل‌دادنی‌ها

- کد پایتون برای محیط و عامل یادگیری Q.
- نموداری که پاداش تجمعی عامل را در طول قسمت‌ها نشان می‌دهد.
- تجسمی از مسیر عامل در maze پس از آموزش.
- یک گزارش مختصر (1-2 صفحه) که پیاده‌سازی، نتایج و بینش‌های شما را توضیح می‌دهد.



تمرین یازدهم
درس هوش مصنوعی

نام استاد: دکتر محمدی

دستیاران آموزشی مرتبط:

علیرضا خیرخواه

مهلت تحویل: 30 دی

9. نکات برای نمرات مثبت:

- کد خود را به وضوح مستند کنید و توضیحاتی برای شرح مراحل کلیدی اضافه کنید.
- تجسم‌سازی‌ها را شامل کنید، مانند نقشه‌های حرارتی مقادیر Q یا انیمیشن‌های گام به گام مسیر عامل.
- مسئله را با معرفی چالش‌های اضافی گسترش دهید، مانند اهداف چندگانه یا جریمه برای بازدید مجدد از سلول‌ها.