

WSI - ćwiczenie 8.

Uczenie się ze wzmocnieniem

grupa 101

27 maja 2021

1 Sprawy organizacyjne

1. Ćwiczenie realizowane jest samodzielnie.
2. Ćwiczenie wykonywane jest w języku R lub Python.
3. Ćwiczenie powinno zostać wykonane do 16.06.2021 23:59. Do tego czasu na adres mailowy `jakub.lyskawa.dokt@pw.edu.pl` należy przesłać plik `.zip` albo `.tar.gz` zawierający kod, dokumentację oraz skan lub zdjęcie podpisanego oświadczenia o pracy zdalnej.
4. Dokumentacja powinna być w postaci pliku `.pdf`, `.html` albo notebooka jupyterowego. Szczegółowe informacje co dokumentacja powinna zawierać oraz na co będzie zwracana uwaga podczas oceniania znajdują się na stronie <http://staff.elka.pw.edu.pl/~rbiedrzy/WSI/index.html>
5. Wzór oświadczenia o pracy zdalnej jest załącznikiem do zarządzenia <https://www.bip.pw.edu.pl/var/pw/storage/original/application/9bfa38aad48ba019ab4cd5449ef209b6.pdf>
6. W przypadku pytań lub wątpliwości zachęcam do pisania na adres mailowy `jakub.lyskawa.dokt@pw.edu.pl` albo na platformie MS Teams (konto powiązane z powyższym adresem email).

2 Zadanie

W ramach ósmego ćwiczenia należy zaimplementować algorytm Q-learning.

Zaimplementowany algorytm powinien działać na dowolnym środowisku *OpenAI gym* o dyskretnych przestrzeniach akcji oraz stanów.

Następnie należy przetestować działanie algorytmu na środowisku *FrozenLake8x8-v0* (z limitem czasu 200 kroków czasowych, jest to domyślna konfiguracja jeżeli się utworzy środowisko z użyciem metody *make*).

3 Uwagi

Pakiet *OpenAI gym* jest pakietem pythonowym, ale istnieje jego port do języka R: <https://github.com/paulhendricks/gym-R>.

Do pomiaru jakości polecam użyć deterministycznej polityki - gdzie wybierana jest najbardziej prawdopodobna akcja z dostępnych. Ponadto, ze względu na losowość środowiska, otrzymany wynik dobrze jest uśrednić na wielu przebiegach.

4 Wskazówki

- W implementacji nie powinno być magicznych stałych, parametry algorytmu powinny być przekazywane np. jako parametry funkcji która ten algorytm implementuje, nie powinny być również przekazywane jako zmienne globalne
- Implementacje powinny być ogólne. Należy unikać pisania osobnej implementacji algorytmu dla każdego problemu.
- Przełączanie wariantów implementacji poprzez komentowanie fragmentów kodu nie jest dobrą praktyką.
- W miarę możliwości warto korzystać z gotowych implementacji np. operacji macierzowych i wektorowych (oczywiście wskazane w poleceniach algorytmy należy zaimplementować samodzielnie).
- Dokumentacja powinna zawierać opis przeprowadzonych eksperymentów, prezentować w jakiejś formie ich wyniki oraz zawierać komentarz do tych wyników.