

Sieci Samouczące Się

Laboratorium 7, 8 – Sprawozdanie

Celem laboratorium było zaimplementowanie aproksymatora wykorzystującego sieć neuronową oraz zastosować techniki poprawiające stabilność uczenia.

W porównaniu do laboratorium 6 zostały wprowadzone następujące zmiany:

- najpierw została dodana prosta sieć neuronowa, która normalizuje wejścia:

```
self.model = nn.Sequential(  
    nn.Linear(state_dim, hidden_dim),  
    nn.Sigmoid(),  
    nn.ReLU(),  
    nn.Linear(hidden_dim, hidden_dim * 2),  
    nn.ReLU(),  
    nn.Linear(hidden_dim * 2, action_dim),  
    nn.Sigmoid()  
)
```

Zastosowano optymalizator Adam. Funkcja kosztu MSE (mean square error). Na wejściu otrzymywane są stany lub zakodowane stany. Zakodowane stany pogorszały znacznie naukę sieci, ponieważ kodowane stany były w formie indeksów Q dla poprzedniej implementacji (z tego powodu w różnych miejscach na wejściu pokazywały się indeksy w różnych miejscach co nie miało sensu). Pozbycie się kodowania poprawiło naukę.

- bufor zawierający wykonane kroki (stan, akcja, nagroda, stan_następny)

- bufor został wykorzystany do wybierania losowych kroków i aktualizacji wag sieci batchami, co znacznie poprawiło naukę, ale również wydłużyło czas wykonania kroku algorytmu. Domyślnie są wykorzystywane batche w liczbie 100.

- wykorzystano 10 akcji z przedziału -1200, 1200

- ostatnim krokiem było dodanie sieci target (miękkie kopiowanie sieci), co 5 kroków. Dodanie drugiej sieci sprawiło, że algorytm lepiej stosuje się do funkcji nagrody.