

# Projekt 5: Uczenie przez wzmocnianie w przestrzeniach ciągłych

Projekt przygotowujemy w grupach dwuosobowych.

## Zadania na 4 punkty:

1. Wybierz środowisko Gymnasium ze stanami ciągłymi oraz jeden z algorytmów z biblioteki stable-baselines3 <https://stable-baselines3.readthedocs.io/en/master/guide/quickstart.html>.
2. Narysuj krzywe uczenia dla minimum trzech zestawów hiperparametrów. Krzywe powinny obrazować średnią nagrodę i odchylenie standardowe obliczone na podstawie 10 uruchomień algorytmu w tych samych warunkach (za wyjątkiem różnej inicjalizacji). W ramach każdego uruchomienia przeprowadź minimum 50 000 kroków czasowych (w przypadku zadań epizodycznych: sumarycznie we wszystkich epizodach). Na osi poziomej krzywej uczenia ma być numer kroku czasowego/epizodu, a na osi pionowej wartość nagrody.
3. W sprawozdaniu napisz po jednym paragrafie o charakterystyce wybranego środowiska i algorytmu. Dalej wymień rozważane hiperparametry i ich wartości, umieść krzywe uczenia, zanotuj czas potrzebny na wykonanie jednego epizodu lub kroku czasowego środowiska. Czy któryś z zestawów hiperparametrów jest wyraźnie lepszy lub gorszy od pozostałych? Uzasadnij swoją opinię.

Każda para środowisko-algorytm może być wybrana przez co najwyżej 1 grupę.

## Zadania na 6 punktów:

1. Wszystkie zadania na 4 punkty.
2. Przetestuj minimum dwie różne architektury sieci wykorzystywanych przez agenta. Narysuj w sprawozdaniu ich schematy (rodzaje warstw, ich wielkości, funkcje aktywacji). Opisz, co znajduje się na wejściu i wyjściu sieci.

## Zadania na 8 punktów:

1. Wszystkie zadania na 6 punktów.
2. Zapisz stan agenta z wcześniejszych dającego najlepsze wyniki. Wykonaj symulację jego działania z wyłączonym trybem eksploracji (wykonuje deterministycznie najlepsze akcje). Porównaj nagrody z krzywą uzyskaną w trakcie uczenia.

Rozwiązania (kod i raporty) należy przesłać na platformie MS TEAMS do dnia 2025-06-09.