

Politechnika Wrocławska

Wydział Matematyki

KIERUNEK:

Matematyka Stosowana

**PRACA DYPLOMOWA
INŻYNIERSKA**

TYTUŁ PRACY:

**Analiza efektywności metod uczenia przez wzmacnianie w grach
komputerowych**

AUTOR:

Adrian Galik

PROMOTOR:

dr hab. Janusz Szwabiński

WROCŁAW 2024

1 Wstęp

Rozwój technologii w tempie przekraczającym wszelkie oczekiwania oraz zwiększająca się dostępność mocy obliczeniowej doprowadziły do tego że algorytmy uczenia maszynowego stanowią nieoderwalną część życia codziennego każdego z nas. Zastosowanie ich można znaleźć w dziedzinach robotyki, rozpoznawania obrazów, przetwarzania języka naturalnego, klasyfikacja spamu, systemy nawigacyjne, diagnostyka chorób, sztuczna inteligencja w grach oraz wiele innych gałęzi technologii które oddziałują na nas w sposób pośredni lub bezpośredni. Jedną z najbardziej fascynujących, a zarazem najstarszych dziedzin uczenia maszynowego jest uczenie przez wzmacnianie. Znana już od lat 50 ubiegłego wieku będzie ona kluczowym działem z którego algorytmy będą stanowiły fundament mojej pracy.

Celem niniejszej pracy inżynierskiej jest analiza efektywności wybranych metod uczenia przez wzmacnianie w grach komputerowych. Przede wszystkim badania oraz porównania algorytmów zarówno jeśli chodzi o czas uczenia oraz efektywność zostały przeprowadzone na przykładzie gry Pong, która jest bardzo często wykorzystywana jako dobry przykład środowiska testowego do badań nad algorytmami sztucznej inteligencji. W ramach pracy zaimplementowałem trzy popularne metody uczenia przez wzmacnianie: Deep Q-Learning (DQN), Advantage Actor-Critic (A2C) oraz Asynchronous Advantage Actor-Critic (A3C), a w następnym kroku zbadałem ich efektywność na zasadzie różnych parametrów m. in. prędkość uczenia oraz skuteczność gry.

2 Wprowadzenie do uczenia maszynowego

Uczenie maszynowe jest jedną z kluczowych gałęzi sztucznej inteligencji, której celem jest tworzenie algorytmów zdolnych do uczenia się na podstawie danych i podejmowania decyzji bez konieczności programowania reguł działania. Oto nieco ogólniejsza definicja: Uczenie maszynowe to "dziedzina nauki dająca komputerom możliwość uczenia się bez konieczności ich jawnego programowania". - Arthur Samuel, 1959. A tu bardziej techniczna: "Mówimy, że program komputerowy uczy się na podstawie doświadczenia E w odniesieniu do jakiegoś zadania T i pewnej miary wydajności P , jeśli jego wydajność (mierzona przez P) wobec zadania T wzrasta wraz z nabywaniem doświadczenia E ". - Tom Mitchell, 1997. Przykładowe dane używane do trenowania systemu noszą nazwę **zbioru/zestawu uczącego** (ang. training set). Każdy taki element uczący jest nazywany **przykładem uczącym (próbką uczącą)**. Część systemu uczenia maszynowego odpowiedzialna za uczenie się i uzyskiwanie przewidywań nazywana jest modelem. Przykładowymi modelami są sieci neuronowe i lasy losowe. Dla przykładu klasyfikacji spamu to zgodnie z definicją Toma Mitchella: naszym zadaniem T jest oznaczenie spamu, doświadczeniem E - dane uczące a do wyznaczenia pozostaje miara wydajności P . Może być nią na przykład stosunek prawidłowo oznaczonych wiadomości do przykładów nieprawidłowo zaklasyfikowanych. (książka uczenie maszynowe z użyciem Scikit-Learn, Keras i TensorFlow (5 zdań ostatnich))

2.1 Podział uczenia maszynowego

(Można dodać do każdego jakieś wykresy) Algorytmy uczenia maszynowego można podzielić na cztery ogólne kategorie:

2.1.1 Uczenie nadzorowane

To najczęstszy przypadek uczenia maszynowego. W tym przypadku algorytm uczy się na podstawie oznaczonych danych wejściowych które są opisane przez człowieka oraz odpowiadających im wyników. Głównymi zastosowaniami algorytmów uczenia nadzorowanego to klasyfikacja i regresja. Klasycznym przykładem jest klasyfikacja spamu, polega ona na analizie przez algorytm e-maila i przypisanie do niego kategorii "spam" lub "nie spam". Przykład algorytmów: regresja liniowa, drzewa decyzyjne, SVM

2.1.2 Uczenie nienadzorowane

Algorytm analizuje dane bez użycia jakichkolwiek oznaczeń w celu znalezienie grup lub ukrytych wzorców. Kluczowymi zadaniami uczenia nienadzorowanego są między innymi: wizualizacja danych, redukcja wymiarowości, analiza skupień, wykrywanie anomalii, wykrywanie nowości, usuwanie szumu oraz uczenie przy użyciu reguł asocjacyjnych. Przykład algorytmów: K-Means DBSCAN

2.1.3 Uczenie częściowo nadzorowane

Jest to specyficzny przypadek uczenia nadzorowanego, lecz ma ono na tyle odmienne zasady działania że tworzy oddzielną kategorię. W uczeniu częściowo nadzorowanym algorytm nie używa oznaczeń nadanych przez człowieka, lecz są one wygenerowane na podstawie danych wejściowych (zazwyczaj stosowane są do tego algorytmy heurystyczne). Jest to szczególnie przydatne w sytuacjach, gdy oznaczanie danych jest kosztowne lub czasochłonne jak przykładowo w diagnostyce medycznej.

2.1.4 Uczenie przez wzmacnianie

Dziedzina która była zaniedbywana do momentu w którym autorzy projektu Google DeepMind wykorzystali ją w celu nauki komputerów gier Atari. Jest to specyficzna forma uczenia maszynowego gdyż w zasadniczy sposób różni się od wszystkich poprzednich metod gdyż algorytm nie uczy się za pomocą danych lecz na podstawie interakcji z dynamicznym środowiskiem stąd nazwa "wzmacnianie". Agentem nazywamy element który jest odpowiedzialny za interakcję ze środowiskiem, a same interakcje nazywamy akcjami. Algorytm za wykonanie każdej akcji definiowanej przez autora otrzymuje adekwatnie do oczekiwań nagrodę i karę. Na podstawie tej metody algorytm uczy się strategii która pozwala mu maksymalizować nagrodę na podstawie konkretnego stanu środowiska.

3 Teoretyczne podstawy uczenia przez wzmacnianie

3.1 Podstawowe pojęcia i definicje

- Agent -