

Моделювання поведінки автономних роботизованих систем із застосуванням методів підкріплювального навчання

Олексій Мацієвський асистент¹(ORCID: 0009-0008-2341-8166), Владислав Гоц¹ к.т.н. доц. (ORCID: 0000-0003-4384-4011)

¹ Київський національний університет будівництва і архітектури, 03037, м. Київ, проспект Повітряних Сил, 31, Україна

АНОТАЦІЯ

У цьому дослідженні представлено методи підкріплювального навчання (RL) для моделювання поведінки автономних роботизованих систем. RL дозволяє роботам адаптуватися до динамічного середовища через нагороди за правильні дії. Описано ключові компоненти RL та виклики його реалізації у робототехніці. Також розроблено нейромережеву модель для роботизованих агентів і продемонстровано її ефективність у симуляціях. Результати показують, що RL значно підвищує адаптивність роботів у складних умовах.

Ключові слова: автономні роботизовані системи, підкріплювальне навчання, нейромережеві моделі, штучний інтелект, динамічні середовища.

1. ВСТУП

Сучасний розвиток робототехніки та штучного інтелекту створює можливості для застосування підкріплювального навчання (RL) у моделюванні поведінки автономних роботизованих систем [1]. RL дозволяє роботам самостійно вивчати найкращі стратегії через взаємодію з навколишнім середовищем, отримуючи винагороду за правильні дії та штрафи за помилки. Крім того, аналізуються сучасні підходи до інтеграції методів підкріплювального навчання з алгоритмами глибинного навчання для підвищення ефективності роботи роботизованих систем у складних середовищах. Досліджується можливість створення адаптивних моделей, які дозволяють роботам швидко пристосовуватися до змінних умов і вдосконалювати свою поведінку в реальному часі.

2. МЕТА ДОСЛІДЖЕННЯ

Дослідження спрямоване на розробку та тестування нейромережевої моделі [2], що використовує методи RL для автономних роботизованих систем. Метою є створення моделі, здатної адаптувати поведінку роботів у складних умовах, таких як навігація в динамічному середовищі та уникнення перешкод. Додатково, однією з цілей є розробка гібридної нейромережевої моделі, яка зможе поєднувати методи Q-learning та convolutional neural networks (CNN) для підвищення ефективності адаптації роботизованих агентів до складних умов середовища.

3. ПІДКРІПЛЮВАЛЬНЕ НАВЧАННЯ В АВТОНОМНИХ РОБОТИЗОВАНИХ СИСТЕМАХ

Методи підкріплювального навчання [3] є потужним інструментом для створення адаптивних та самонавчальних систем. У випадку автономних роботів, ці методи дозволяють агентам поступово вивчати середовище через експерименти з ним. RL заснований на взаємодії між агентом та навколишнім середовищем [4]. Агент виконує дії, отримує винагороду або покарання за свої дії і використовує цей досвід для вдосконалення своєї поведінки. Підхід комбінованого використання підкріплювального навчання

та методів глибинного навчання дозволяє створювати гнучкі моделі, які ефективно пристосовуються до нових умов завдяки здатності CNN обробляти великі обсяги вхідних даних з високою точністю. В такій гібридній системі нейронна мережа відповідає за аналіз середовища, тоді як RL відповідає за вибір найкращої дії в кожному стані.

4. МАРКОВСЬКИЙ ПРОЦЕС ПРИЙНЯТТЯ РІШЕНЬ

Марковського процесу прийняття рішень (MDP) [5] у робототехніці є його здатність чітко моделювати послідовні дії агента в умовах невизначеності. Це особливо важливо для автономних систем, де алгоритм повинен враховувати ймовірності переходів між станами для забезпечення максимальної ефективності рішень. Додатково, впровадження динамічного налаштування параметрів, таких як коефіцієнт дисконтування, дозволяє коригувати політику прийняття рішень, що робить систему гнучкішою під час виконання складних завдань.

Марковський процес прийняття рішень використовується для моделювання RL-систем. Однією з ключових переваг використання

$$V(s) = \max_a [R(s, a) + \gamma \sum_{s'} P(s'|s, a) V(s')] \quad (1)$$

Де: $V(s)$ - очікувана сума винагороди для стану s ;

a - дія, яку виконує агент;

$R(s, a)$ - винагорода, отримана після виконання дії a у стані s ;

γ - коефіцієнт дисконтування (від 0 до 1), який відображає важливість майбутніх винагород;

$P(s'|s, a)$ - ймовірність переходу до стану s' з стану s при виконанні дії a ;

s' - наступний стан

Метою RL є оптимізація політики агента, щоб максимізувати очікувану винагороду за його дії. Одним з популярних алгоритмів RL є Q-learning, який використовує функцію оцінки стану та дії для визначення найкращих рішень у кожному стані.

5. НЕЙРОННІ МЕРЕЖІ ТА ЇХ ВИКОРИСТАННЯ В RL

Для реалізації RL в робототехніці часто використовують нейронні мережі [6]. Вони дозволяють агентам обробляти складні вхідні дані, такі як зображення чи сигнали з датчиків, і приймати рішення на основі аналізу отриманої інформації. У цьому дослідженні було розроблено нейромережеву модель, що складається з вхідного, прихованого і вихідного шарів.

- **Вхідний шар** відповідає за прийом вхідних даних від середовища, таких як позиція робота чи перешкоди на шляху.
- **Прихований шар** обробляє дані та виявляє взаємозв'язки між параметрами середовища.
- **Вихідний шар** генерує ймовірності вибору кожної можливої дії, що дозволяє агенту приймати рішення в реальному часі.

6. РЕЗУЛЬТАТИ

Проведене дослідження підтвердило ефективність застосування методів підкріплювального навчання для моделювання поведінки автономних роботизованих систем. Розроблена модель нейронної мережі дозволила агенту успішно навчатися через взаємодію із середовищем, демонструючи здатність адаптуватися до змінних умов та покращувати свої стратегії для досягнення поставлених цілей.

Таблиця 1 та рисунок 1 демонструють прогрес навчання нейронної мережі для автономної роботизованої системи. Результати показують, що зі збільшенням кількості епізодів середня винагорода агента поступово зростала, а кількість кроків, необхідних для виконання завдань, зменшувалася. Від епізоду 1 до епізоду 50 спостерігалось значне зниження середньої винагороди, що свідчить про складність початкових етапів навчання. Однак з епізоду 100 середня винагорода починає підвищуватися, і в епізоді 350 досягає максимального значення +100, що є індикатором успішного навчання системи.

Значна частка епізодів завершувалася успішно починаючи з епізоду 100, що підтверджує поступове вдосконалення стратегії поведінки агента. Отримані результати підтверджують ефективність розробленої моделі нейронної мережі та методів підкріплювального навчання для автономних роботизованих систем

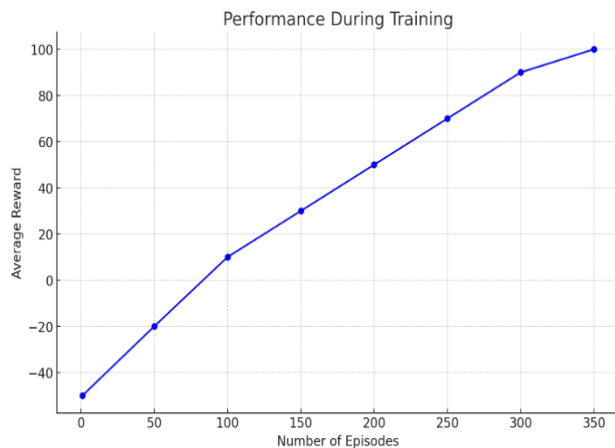


Рисунок 1. Результати навчання нейронної мережі

Таблиця 1: Прогрес навчання нейронної мережі для автономної роботизованої системи

Episode	Average reward	Number of steps	Successful episode (Yes/No)
1	-50	10	No
50	-20	30	No
100	10	50	Yes
150	30	60	Yes
200	50	80	Yes
250	70	90	Yes
300	90	100	Yes
350	100	110	Yes

7. ВИСНОВКИ

Дослідження показало, що методи підкріплювального навчання дозволяють створювати адаптивні автономні роботизовані системи, здатні до навчання на основі взаємодії з середовищем. Використання нейронних мереж для реалізації RL дає можливість обробляти складні вхідні дані та ефективно приймати рішення, що робить цей підхід актуальним для вирішення завдань у різних галузях робототехніки.

Список джерел

- [1] Structural and modular designing of supporting frameworks of autonomous multifunctional robotic platforms for cover combat and special operations / V. Belikov та ін. *Collection of scientific works of Odesa Military Academy*. 2020. Т. 1, № 13. С. 113–121. URL: <https://doi.org/10.37129/2313-7509.2020.13.1.113-121>.
- [2] Lutska N., Vlasenko L., Zaiets N. Prognostication resource efficiency of the sugar factory based on neuron network models. *Scientific Works of National University of Food Technologies*. 2023. Т. 29, № 3. С. 7–18. URL: <https://doi.org/10.24263/2225-2924-2023-29-3-3>.
- [3] Lorenz U. Bestärkendes Lernen als Teilgebiet des Maschinellen Lernens. Reinforcement Learning. Berlin, Heidelberg, 2020. С. 1–11. URL: https://doi.org/10.1007/978-3-662-61651-2_1.
- [4] Fergus P., Chalmers C. Deep Reinforcement Learning. *Computational Intelligence Methods and Applications*. Cham, 2022. P. 255–264. URL: https://doi.org/10.1007/978-3-031-04420-5_11.
- [5] Sivaraman R. MARKOV PROCESS AND DECISION ANALYSIS. *JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES*. 2020. Т. 15, № 7. URL: <https://doi.org/10.26782/jmcmms.2020.07.00002>.
- [6] Bräunl T. *Neural Networks. Embedded Robotics*. Singapore, 2022. С. 421–437. URL: https://doi.org/10.1007/978-981-16-0804-9_19.