

УДК 004.4:[004.738.5:004.722]

## ДОСЛІДЖЕННЯ НАВЧАННЯ РОБОТІВ НА ОСНОВІ ДОСВІДУ

Пара І.І.

Науковий керівник – Старший викладач каф. КІТАР Гурін Д.В.

Харківський національний університет радіоелектроніки

м. Харків, Україна

тел. +38(095) 54-34-566, e-mail: [illia.para@nure.ua](mailto:illia.para@nure.ua).

This article provides a review of the literature on reinforcement learning from demonstrations (RLfD), a subfield of machine learning that allows robots to learn from demonstrations by humans or other robots. The goal is to develop a comprehensive understanding of RLfD methods and their applications in robotics.

Робототехніка – це динамічна галузь, яка постійно розвивається. Одним із ключових напрямків досліджень у цій сфері є розробка методів навчання роботів на основі досвіду. Цей підхід ґрунтується на ідеї, що роботи можуть вчитися на власних помилках та успіхах, а також на досвіді інших роботів.

Метою даної роботи є аналіз предметної області "Дослідження навчання роботів на основі досвіду". В рамках роботи буде проведено огляд існуючих методів навчання роботів на основі досвіду.

Підкріплення навчання (англ. Reinforcement Learning) - це метод машинного навчання, який використовується для навчання агентів (роботів) діяти в середовищі таким чином, щоб максимізувати очікуване значення винагороди.

Основні принципи:

– Агент: Це автономна сутність, яка може сприймати середовище та діяти в ньому.

– Середовище: Це все, що оточує агента, включаючи інші об'єкти, фізичні закони та правила.

– Дія: Це те, що агент може/повинен робити відносно стану середовища.

– Стан: Це опис середовища в певний момент часу.

– Винагорода: Це скалярне значення, що відображає бажаність стану/дії.

– Функція цінності: Це функція, яка оцінює очікуване значення винагороди за певний стан.

– Політика: Це функція, яка визначає, яку дію агент буде виконувати в певному стані.

Алгоритми:

– Метод Монте-Карло: Цей метод оцінює цінність стану, використовуючи середнє значення винагород, отриманих з епізодів, які починаються з цього стану.

– TD-навчання: Цей метод оцінює цінність стану, використовуючи різницю між поточною оцінкою та очікуваною винагородою за наступний стан.

– Q-навчання: Цей метод використовує таблицю Q-значень для зберігання очікуваної винагороди за виконання певної дії в певному стані.

Переваги:

– Підкріплення навчання може використовуватися для навчання агентів в складних середовищах, де складно або неможливо заздалегідь визначити оптимальну політику.

– Підкріплення навчання може використовуватися для навчання агентів адаптувати свою поведінку до мінливих умов середовища.

– Підкріплення навчання може використовуватися для навчання агентів виконувати складні завдання, які неможливо запрограмувати вручну.

Недоліки:

– Підкріплення навчання може потребувати багато даних для навчання.

– Підкріплення навчання може призвести до небажаної поведінки агента, якщо винагорода не правильно визначена.

Навчання з демонстрацією (англ. Learning from Demonstration) - це метод машинного навчання, який використовується для навчання агентів (роботів) на основі прикладів поведінки, продемонстрованої людиною або іншим роботом.

Основні принципи:

– Демонстрація: Це послідовність дій, які виконує людина або інший робот.

– Імітація: Це процес, за допомогою якого агент намагається відтворити дії, продемонстровані в демонстрації.

– Узагальнення: Це процес, за допомогою якого агент використовує знання, отримані з демонстрації, для виконання нових завдань.

Методи:

– Навчання за траєкторією: Цей метод вчить агента слідувати за траєкторією, продемонстрованою в демонстрації.

– Навчання за інваріантними характеристиками: Цей метод вчить агента використовувати інваріантні характеристики для узагальнення з демонстрації.

– Навчання за допомогою зворотного зв'язку: Цей метод використовує зворотний зв'язок від людини або іншого робота для покращення поведінки агента.

Переваги:

– Навчання з демонстрацією може використовуватися для навчання агентів завданням, які важко або неможливо запрограмувати вручну.

– Навчання з демонстрацією може використовуватися для навчання агентів виконувати завдання з високою точністю.

Недоліки:

– Навчання з демонстрацією може призвести до перенавчання, коли агент вчиться відтворювати поведінку демонстратора, а не узагальнювати її на нові завдання.

Імітаційне навчання (англ. Imitation Learning) - це метод машинного навчання, який використовується для навчання агентів (роботів) на основі симуляцій.

Основні принципи:

– Симулятор: Це комп'ютерна програма, яка моделює поведінку реального світу.

– Віртуальне середовище: Це середовище, створене симулятором.

– Агент: Це автономна сутність, яка може сприймати віртуальне середовище та діяти в ньому.

Методи:

– Навчання за траєкторією: Цей метод вчить агента слідувати за траєкторією, згенерованою в симуляторі.

– Навчання за допомогою зворотного зв'язку: Цей метод використовує зворотний зв'язок з симулятора для покращення поведінки агента.

– Навчання з підкріпленням: Цей метод використовує метод підкріплення навчання для навчання агента в симуляторі.

Переваги:

– Імітаційне навчання може використовуватися для навчання агентів в небезпечних або складних середовищах.

– Імітаційне навчання може використовуватися для навчання агентів завданням, які важко або неможливо демонструвати в реальному світі.

Недоліки:

– Імітаційне навчання може призвести до перенавчання, коли агент вчиться відтворювати поведінку в симуляторі, а не узагальнювати її на реальний світ.

Дані методи машинного навчання допоможуть з реалізацією проекту, в якому планується розробити повноцінне віртуальне середовище з можливостями навчання роботів, що, в свою чергу, допоможе розробити: «розумних» роботів, які краще виконують складні завдання, більш автономних роботів, які працюють без постійного втручання. В майбутньому це дослідження буде розвиватися та вдосконалюватися, а також планується розробити макет для роботи із цим віртуальним середовищем.

Список використаної літератури

1. Argall B. D., Chernova S., Veloso M. A survey on robot learning from demonstration. *Robotics and Autonomous Systems*, 2009. Ст.469-483. URL: <https://www.sciencedirect.com/science/article/pii/S0921889008001772>

2. Baxter J., Yew J. Learning In Simulated Environments: A survey. *IEEE Transactions on systems, man and cybernetics Part C (Applications and Reviews)*, 2010. Ст.303-313. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4454909/>

3. Yevsieiev, V., & Gurin, D. (2023). Comparative Analysis of the Basic Methods Used in Industry 4.0 and Industry 5.0. *Collection of Scientific Papers «ΛΟΓΟΣ»*, (Bologna, Italy), 113–115. <https://doi.org/10.36074/logos-29.09.2023.31>