

ПЕРСПЕКТИВИ МОДЕЛЕЙ НАВЧАННЯ З ПІДКРІПЛЕННЯМ У ЗАДАЧАХ ПІДТРИМКИ ПРИЙНЯТТЯ ІНВЕСТИЦІЙНИХ РІШЕНЬ

Горенський Г. Г.

Науковий керівник – проф. каф. ЕОМ Рубан І. В.

Харківський національний університет радіоелектроніки, каф. ЕОМ,
м. Харків, Україна

e-mail: hennadii.horenskyi@nure.ua

The investigation aims to evaluate the effectiveness of machine learning algorithms in investment decision-making by examining reinforcement learning models within the investment decision support system. The research includes an analysis of the PPO, IMPALA, and A2C models. The Gym environment was used to simulate automated trading, and it was found that PPO was the most effective in maximizing portfolio profitability. The findings from 30 experiments were compared using Welch's t-test to ensure objectivity and robustness.

Системи підтримки прийняття інвестиційних рішень відносяться до сучасного інструментарію суб'єктів фінансового сектора, що обумовлює необхідність адаптації традиційних методів машинного навчання до особливостей великої кількості складно пов'язаних фінансових процесів.

Альтернативою використання традиційного аналізу при прийнятті інвестиційних рішень [1], є автоматизовані системи, які використовують алгоритми машинного навчання для прийняття рішень на основі результатів попередніх фінансових реалізацій.

Такий підхід дає змогу здійснювати ефективні транзакції з загально низьким рівнем ризику. Сучасні системи підтримки прийняття інвестиційних рішень можуть працювати безперервно, аналізуючи великі обсяги даних і враховуючи ринкові тенденції в режимі реального часу.

Моделі навчання з підкріпленням (Reinforced Learning) представляють собою підходи до машинного навчання, які спрямовані на навчання агента взаємодіяти з оточенням і приймати оптимальні рішення для досягнення поставленої мети. Такі моделі стали популярними в останні роки завдяки їх здатності ефективно вирішувати завдання, де необхідно управляти послідовними діями з метою максимізації нагороди [2].

Мета дослідження полягає у адаптації традиційних моделей з підкріпленням до особливостей процесу прийняття рішень на фінансових ринках, що забезпечить раціональне розподілення капіталу та максимізацію прибутковості інвестиційного портфеля [3].

У дослідженні порівнювалися такі моделі навчання з підкріпленням PPO (Proximal Policy Optimization), IMPALA (Importance Weighted Actor-Learner Architecture) та A2C (Advantage Actor-Critic).

Процес навчання у цих моделях полягає у тому, що агент взаємодіє з оточенням, приймає дії, отримує нагороди і здійснює навчання на основі отриманого досвіду. У випадку PPO, агент користується проксимальною оптимізацією політики, щоб ефективно навчати політику, яка максимізує нагороду. IMPALA використовує архітектуру актор-учень з ваговими коефіцієнтами важливості, що дозволяє ефективно використовувати паралельне навчання для швидкого навчання. А A2C базується на архітектурі актор-критик, де критик оцінює переваги дій, а актор вчиться вибирати дії, які максимізують ці переваги.

Для моделювання автоматичної торгівлі використовувалось середовище Gym.

Для порівняльного аналізу моделей використовувалась метрика – кумулятивна нагорода за епізод. Яка у досліджуваній задачі представляє собою сумарний щоденний прибуток:

$$\sum_{t=1}^n \left(\frac{NW_t}{NW_{t-1}} - 1 \right),$$

де NW – net worth, t – часовий крок.

Задля отримання статистично достовірних результатів, усі моделі навчались на однакових даних протягом однакової кількості епох. Усього було проведено 30 експериментів зі зміною початкового числа (seed), після чого результати порівнювались (Рис. 1) між собою за допомогою t-критерія Уелча.

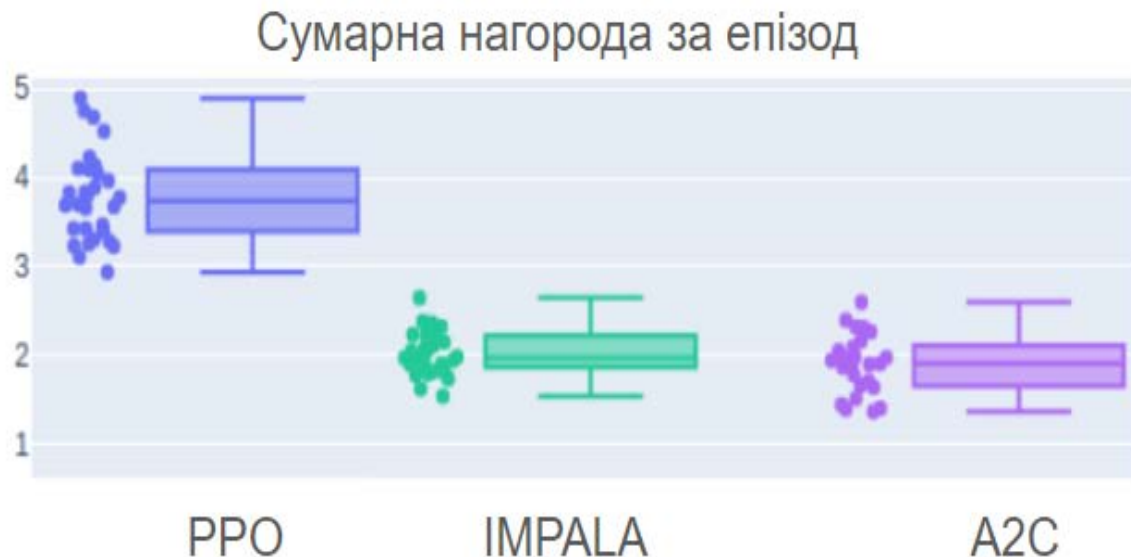


Рисунок 1 – Порівняння моделей навчання з підкріпленням PPO, IMPALA та A2C

Цей метод дозволяє забезпечити об'єктивність та порівняти ефективність кожної моделі в різних умовах, враховуючи випадковість у початковому стані навчання [4]. За результатами порівняння – усі групи статистично відрізняються одна від одної, значну перевагу показує модель РРО.

Таким чином, адаптація традиційних моделей навчання з підкріпленням, для використання з фінансовими даними, що характеризується величезним обсягом та зашумленістю, виглядає перспективним напрямком розвитку спеціалізованих інформаційних фінансових технологій.

Список використаних джерел:

1. Романенков Ю. О., Зейнієв Т. Г. Завдання контуру стратегічного управління ефективністю бізнес-процесів в організації // Системні дослідження та інформаційні технології. 2015. №. 3. С. 43-47.

2. Martovytskyi V., Argunov V., Ruban I., Romanenkov Y. Developing a Risk Management Approach based on reinforcement Training in the Formation of an investment Portfolio. Eastern-European Journal of Enterprise Technologies. 2023. Vol. 2, No. 3(122). PP.106-116.

3. Горенський Г.Г., Рубан І.В. Аналіз методів навчання з підкріпленням у стратегіях фінансової торгівлі // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: збірник тез доповідей 13-ої МНТК (26-27 квітня 2023 р.). Баку – Харків – Жиліна, 2023. Т. 2, С.72-73.

4. Colas, Cédric, Olivier Sigaud, and Pierre-Yves Oudeyer. How many random seeds? statistical power analysis in deep reinforcement learning experiments. arXiv preprint arXiv:1806.08295. 2018. DOI: <https://doi.org/10.48550/arXiv.1806.08295>.