

УДК 656.078

DOI: [https://doi.org/10.32515/2664-262X.2024.10\(41\).1.60-67](https://doi.org/10.32515/2664-262X.2024.10(41).1.60-67)

**В.В. Аулін**, проф., д-р техн. наук, **С.Г. Ковальов**, докторант, канд. пед. наук,  
**А.В. Гриньків**, ст. дослідник, канд. техн. наук

*Центральноукраїнський національний технічний університет, м. Кропивницький,  
Україна*

*e-mail: aulinvv@gmail.com, kovalyovserggr@ukr.net*

**В.В. Варваров**, пров. наук. співр., канд. техн. наук

*Харківський національний університет повітряних сил імені Івана Кожедуба, м. Харків,  
Україна*

*e-mail: varvarov\_@ukr.net*

## Алгоритм оптимізації надійності функціонування та ефективності використання виробничого обладнання методами штучного інтелекту

Зміст статті орієнтований на проблематику підвищення надійності та ефективності виробничих процесів за рахунок зниження витрат на обслуговування та мінімізації простоїв. Представлено підхід, який передбачає оптимізацію технічного обслуговування виробничого обладнання на основі алгоритмів штучного інтелекту. Розглядається методика використання теорії марківських процесів в контексті навчання з підкріпленням і її застосування для моделювання виробничого середовища. Проведене дослідження дає змогу говорити про значне покращення надійності та продуктивності виробничого обладнання у виробничій системі завдяки впровадженню ШІ-технологій.

**виробнича система, виробниче обладнання, штучний інтелект, марківські процеси, надійність функціонування, ефективність використання**

**Постановка проблеми.** Важливою складовою забезпечення ефективного функціонування сучасного виробництва є надійність роботи та ефективність використання виробничого обладнання. Підтримування параметрів надійності та ефективності обладнання у необхідному діапазоні забезпечується вчасним проведенням необхідних технічних дій. Традиційно вчасне проведення технічного обслуговування (ТО) та ремонту, виконують у визначені інтервали часу або при виникненні відмов, що не є оптимальним у процесі виробництва. Негативний вплив на виробництво здійснює непередбачувані простої та підвищення витрат на використання обладнання.

Вирішення цієї проблеми потребує залучення інноваційних дослідницьких методів та механізмів, що здатні підвищити інформаційну складову використання виробничого обладнання. Інформаційна складова дозволяє інтегрувати у виробничий процес сучасні інформаційні методи в тому числі і інтелектуальні системи, які здатні комплексно виконувати діагностику, контроль, прогнозування та адаптацію до умов функціонування самої виробничої системи.

На сьогоднішній день серед інтелектуальних виробничих систем найбільш перспективними, досліджуваними та затребуваними є моделі та алгоритми штучного інтелекту (ШІ)[1], а тому дослідження застосуванням ШІ для підвищення надійності функціонування та ефективності використання обладнання у виробничому процесі є безумовно актуальною науковою проблемою.

**Аналіз останніх досліджень і публікацій.** Виробничі процеси та зміни технічного стану виробничого обладнання можливо описати теорією марківських процесів [2,3]. При цьому виробниче обладнання розглядається як підсистема, що входить у виробничу систему. Теорія марківських процесів є потужним інструментом для опису складних систем, які можуть перебувати у деяких станах  $S$ , перехід між якими має стохастичний характер. Такий інструмент добре підходить для моделювання параметрів надійності та ефективності виробничого обладнання, які теж носять стохастичний або відносний характер[4].

Основною властивістю марківських процесів є умова, що майбутній стан системи залежать тільки від теперішнього стану і умовно не залежить від минулого її стану. Математично це можна записати через ймовірність виникнення подій, які відбулись в минулому, теперішньому та будуть відбуватись в майбутньому часі [3]:

$$P(\text{Майбутнє} \mid \text{Минуле} \cap \text{Сучасне}) = P(\text{Майбутнє} \mid \text{Сучасне}). \quad (1)$$

Важливими понятійними елементами у теорії марківських процесів є: матриця ймовірностей переходу між станами системи  $\{p_{ij}\}$ , вектор ймовірностей перебування системи у деякому стані -  $P$  та теорема про матрицю ймовірностей переходу між станами на  $n$ - кроків:

$$P = P^n \cdot \{p_{ij}\}, \quad (2)$$

де  $P$  – вектор ймовірностей знаходження системи у деякому стані для поточної ітерації;

$P'$  – вектор ймовірностей знаходження системи у деякому стані для попередньої ітерації;

$\{p_{ij}\}$  – матриця ймовірності переходу між станами системи.

Матриця ймовірностей  $\{p_{ij}\}$  формується з графа, у вузлах якого відображені усі стани системи, а у ребрах графу містять значення ймовірності переходу від  $i$ -го до  $j$ -го стану. Вектор ймовірності стану системи  $P$  має координати, значення яких відповідають ймовірності перебування системи у відповідному стані.

Враховуючи, що марківські процеси є однорідними із зазначеної вище теореми впливає співвідношення, яке дозволяє знаходження вектора ймовірності перебування системи в деякому стані на основу попереднього його значення та матриці ймовірності переходів між станами, яка формується на основі поточного стану системи [2,3].

Для розв'язання завдань марківських процесів останнім часом намагаються використовувати різні методи ШІ, одним з різновидів яких є машинне навчання [5,6]. Машинне навчання поділяється на супервізоване навчання (Supervised Learning, SL), ненаглядне навчання (Unsupervised Learning, UL) та навчання з підкріпленням (Reinforcement Learning, RL). Навчання з підкріпленням базується на теорії марківських процесів і дозволяє точно моделювати динаміку технічних станів виробничого обладнання. Враховуючи ймовірності переходів між станами, така модель ШІ добре відповідає дослідницьким вимогам у розглядуваному контексті.

Навчання з підкріпленням RL — це метод машинного навчання, де агент навчається приймати рішення шляхом взаємодії з середовищем і отримання зворотного зв'язку у вигляді винагород або штрафів. Основна мета агента — максимізувати сукупну винагороду з часом [7,8].

RL-модель навчання передбачає три реалізації: Q-навчання, SARSA та DQN навчання. Q-навчання і DQN-навчання у своєму алгоритмі використовують Q-функцію, яка повертає довготривалу винагороду агенту за виконану ним дію  $a$  в стані  $S_a$ . При цьому, відмінність між Q- та DQN-навчаннями полягає у тому, що у Q-навчанні Q-функція задається таблицею, а у випадку DQN-навчання її роль виконує штучна нейронна мережа [3,9,10].

Агент у середовищі виробничої системи перебуває в певному стані  $S_a$  у якому він виконує дію  $a$ , після чого переходить у інший стан. Цей процес виконується циклічно. У випадку Q-навчання для виконання однієї дії і переходу у інший стан, агент виконує наступний алгоритм [11,13]:

1. Агент визначає свій стан  $S_a$  і отримує список можливих дій  $a$  у цьому стані.
2. Для кожної дії  $a$  викликається функція  $Q(s, a)$ , яка повертає табличне значення, яке є довготривалою винагородою агента для конкретної дії  $a$  у поточному стані системи  $S$ .
3. Отримавши множину значень довготривалих винагород  $m_{ij}$  для кожної пари  $(s, a_i)$ , політика (компонент агента) відбирає на основі усіх отриманих значень від  $Q$ -функції лише одну оптимальну дію і агент її виконує.
4. Виконавши обрану політикою дію  $a$ , агент переходить у наступний стан, а середовище передає агенту миттєву винагороду за виконану дію (винагорода може бути від'ємною, тобто покарання).
5. На основі отриманої винагороди корегується функція  $Q(s, a)$ , а саме, корегується значення довготривалої винагороди за виконану дію. Формула за якою відбувається корекція називається формулою Беллмана:

$$Q_{\text{корект}}(s_b, a_t) = Q(s_b, a_t) + \alpha (r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_b, a_t)), \quad (3)$$

де  $\alpha$  – швидкість навчання,  
 $r_t$  – негайна винагорода,  
 $\gamma$  – коефіцієнт дискотування,  
 $\max_{a'} Q(s_{t+1}, a')$  – максимально очікувана довгострокова винагорода для наступного стану  $s_{t+1}$  та дії  $a'$ .

6. Зазначений цикл повторюється для наступних станів [13,14].

Виконуючи вище описаний алгоритм, агент оптимізує свої дії так, щоб отримати максимальне значення довготривалої винагороди і цим самим формує стратегію оптимального функціонування виробничого обладнання в даному середовищі виробничої системи.

**Постановка завдання.** Метою даної роботи є побудова алгоритму підвищення надійності та ефективності виробничого обладнання шляхом впровадження технологій штучного інтелекту, зокрема моделей навчання з підкріпленням та теорії марківських процесів, оптимізацією технічного обслуговування, прогнозуванням відмов та плануванням профілактичних заходів, що сприятиме зниженню витрат на обслуговування та мінімізації простоїв, а також покращенню продуктивності виробничих процесів.

Для реалізації мети розв'язувалися наступні завдання:

1. З'ясувати використання моделей навчання з підкріпленням і теорії марківських процесів у виробничих системах та оцінити їх вплив на підвищення надійності функціонування та ефективності використання обладнання.
2. Створити алгоритми для прогнозування можливих відмов виробничого обладнання та розробити стратегії планування профілактичних заходів з метою мінімізації простоїв та зниження витрат на технічне обслуговування.
3. Провести дослідження щодо впровадження ШІ-технологій у виробничі процеси та визначити їх вплив на покращення продуктивності використання та надійності функціонування обладнання у виробничій системі, що сприятиме оптимізації ефективного функціонування сучасного виробництва.

**Викладення основного матеріалу.** Розглянемо моделювання функціонування виробничої системи з оптимізацією параметрів надійності функціонування та ефективності використання її обладнання за допомогою RL-методу машинного

навчання. Основна мета застосування RL-методу передбачає створення оптимальної системи динамічного контролю надійності та ефективності виробничого процесу, що дозволить прогнозувати можливі відмови та виконувати адаптацію графіку ТО обладнання у режимі реального часу. Це дозволить мінімізувати простой та знизити витрати на обслуговування, а також підвищити загальну ефективність виробничих процесів.

Виробниче обладнання умовно розглядаємо як підсистему середовища виробничої системи, що може перебувати у шести станах:

- робочий стан  $S_1$ , в якому обладнання працює належним чином, виконуючи свої функції з оптимальною продуктивністю;
- стан очікування  $S_2$ , коли обладнання готове до роботи, але очікує на завдання або матеріали;
- стан простою  $S_3$ , для якого характерно те, що обладнання не працює, але готове до запуску, ситуація корисна для зниження енергоспоживання під час перерв;
- стан технічного обслуговування  $S_4$ , коли обладнання знаходиться в процесі планового або непланового технічного обслуговування для запобігання несправності та пошкодження обладнання або відновлення його працездатності;
- стан несправності  $S_5$ , в якому обладнання вийшло з ладу і потребує ремонту або заміни.
- стан деградації  $S_6$ , коли обладнання працює, але знижена продуктивність або ефективність через спрацювання.

Сукупність станів середовища виробничої системи наведено на рис.1 у вигляді графу.

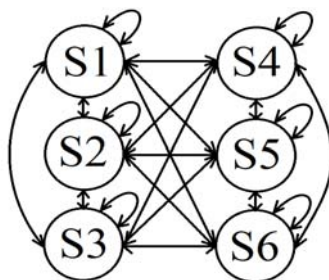


Рисунок 1 – Граф станів середовища виробничої системи

Джерело: розроблено авторами

Можна бачити, що між станами середовища виробничої системи є взаємозв'язки і взаємопереходи, відображені дуговими та прямими двосторонніми векторами. Кожен із векторів характеризують числа, які вказують на ймовірність переходу між відповідними станами. Таке представлення середовища відповідає марківській моделі, де перехід між станами системи має стохастичний характер і описується матрицею ймовірностей переходу між існуючими станами середовища. Дані рис. 1 дають можливість побудувати матрицю ймовірностей станів виробничого середовища (рис. 2).

$$\{P_{ij}\} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} \end{pmatrix}; \quad \{P_{ij}\} = \begin{pmatrix} 0,3 & 0,2 & 0,29 & 0,1 & 0,01 & 0,1 \\ 0,2 & 0,1 & 0,3 & 0,1 & 0,2 & 0,1 \\ 0,1 & 0,1 & 0,3 & 0 & 0,2 & 0,3 \\ 0,1 & 0,1 & 0,1 & 0,1 & 0,1 & 0,2 \\ 0,01 & 0,2 & 0,2 & 0,1 & 0,29 & 0,2 \\ 0,1 & 0,1 & 0,3 & 0 & 0,2 & 0,2 \end{pmatrix}$$

а б

Рисунок 2 – Представлення марковської матриці ймовірностей переходу виробничої системи між доступними станами середовища: а - загальний вигляд; б – конкретизований стан

Джерело: розроблено авторами

Як зазначалось вище, алгоритм RL-навчання працює циклічно і для його реалізації потрібна певна кількість ітерацій [3,14]. На практиці елементи матриці ймовірностей  $\{p_{ij}\}$  формуються для кожної ітерації на основі апаратних, програмно-інформаційних засобів, що інтегровані у виробниче обладнання виробничої системи [15]. Їх можливе визначення на основі вимірювань тривалості експлуатації, рівня завантаження і спрацювання обладнання, загальної інформації про обладнання: кількість відмов різних вузлів, затрати на ТО та ремонт.

Маючи матрицю ймовірностей переходу між станами, можна визначити вектор ймовірностей знаходження виробничої системи у конкретному стані. Визначення вектора ймовірностей для поточної ітерації передбачає операцію множення вектора ймовірностей для попередньої ітерації на матрицю ймовірностей переходу системи між станами, яка отримана для поточної ітерації (2). Такий підхід дозволяє розраховувати на основі стохастичних методів, що описані у теорії марківських процесів, стан виробничої системи з ймовірнісним характером. На початку навчання координати вектора ймовірності перебування у певному стані. Приймають, що всі вони рівні нулю за виключенням однієї координати, що визначає стартовий стан агента, відповідно, ця координата має значення одиниці.

Виходячи з графу станів виробничої системи (рис. 1), в матричному вигляді представлення її є таблиця Q-функції, орієнтована на середовище виробничої системи.

Таблиця 1 – Матричне представлення станів виробничої системи

| Стан виробничої системи  | Продовжуват<br>и поточний | Перейти в<br>стан простою | Перейти в<br>стан ТО | Запустити<br>обладнання | Перейти в<br>стан | Завершити<br>ТО | Ремонт<br>обладнання | Замінити<br>обладнання |
|--------------------------|---------------------------|---------------------------|----------------------|-------------------------|-------------------|-----------------|----------------------|------------------------|
| Робочий стан, $S_1$      | Q11                       | Q12                       | Q13                  | -                       | Q15               | -               | Q17                  | Q18                    |
| Стан очікування, $S_2$   | Q21                       | Q22                       | Q23                  | Q24                     | -                 | -               | Q27                  | Q28                    |
| Стан простою, $S_3$      | Q31                       | -                         | Q33                  | Q34                     | -                 | -               | Q37                  | Q38                    |
| Стан ТО, $S_4$           | Q41                       | Q42                       | -                    | -                       | -                 | Q46             | Q47                  | Q48                    |
| Стан деградації, $S_5$   | Q51                       | Q52                       | Q53                  | -                       | Q55               | -               | Q57                  | Q58                    |
| Стан несправності, $S_6$ | Q61                       | Q62                       | Q63                  | Q64-                    | Q65               | Q66             | Q67                  | Q68                    |

Джерело: розроблено авторами

Коефіцієнти  $Q_{mn}$  в таблиці 1 мають загальний вигляд і на початку навчання вони всі рівні 0,5. Таке значення визначається рівноймовірністю дій, які може виконувати агент на початку вивчення виробничого середовища. У процесі навчання ці коефіцієнти корегуються до стану, у якому виробнича система буде мати оптимальний варіант функціонування.

На основі розглянутого побудуємо алгоритм для кожної ітерації, що реалізує агент у відповідності до Q-навчання при моделюванні виробничого обладнання у відповідності до станів виробничої системи, відображеної на рис.1:

1. Агент обраховує стохастичне значення свого стану на основі матриці ймовірностей та вектора ймовірностей для попередньої ітерації і отримує від середовища виробничої системи список дій, які він може виконати у цьому стані. Стани

та відповідні дії агента відображені у таблиці 1.

2. Агент використовує Q-функцію (Таблиця 1) і визначає для кожної можливої дії значення довготривалої винагороди, що дозволяє йому отримати множину пар: дія, довготривала винагорода.

3. Масив пар (дія, довготривала винагорода) передається для аналізу до об'єкта політика, яка обирає дію для поточного стану. Це може бути: проведення ТО, заміна вузлів, продовження перебування в поточному стані та ін., що на логіку політики призведе до максимального значення довготривалої винагороди. Така політика функціонування агента дозволяє йому вибудовувати оптимальну стратегію пристосування до середовища.

4. Після виконання дії агент переходить умовно до іншого стану і отримує від середовища короткотривалу винагороду на основі якої, за допомогою формули Беллмана здійснюється корекція Q-функції, що відповідає накопиченню досвіду агента. Для прикладу відзначимо, що перебуваючи у стані деградації, агент може прийняти рішення про ТО чи ремонт виробничого обладнання, що дозволить йому отримувати позитивну короткотривалу винагороду, а у випадку прийняття рішення про простій матиме покарання і негативний довгостроковий досвід тощо.

5. Перехід до наступної ітерації описаного алгоритму.

Таким чином, моделювання середовища виробничої системи за допомогою Q-навчання у випадку, коли середовище буде відрізнятись кількістю станів чи набором допустимих дій у цих станах, є подібним до розглядуваного і не порушує логіки побудови та використання запропонованої моделі.

**Висновки:** 1. Показано, що впровадження штучного інтелекту, зокрема моделей навчання з підкріпленням та теорії марківських процесів в середовищі виробничої системи, дозволяє значно підвищити надійність функціонування та ефективність використання виробничого обладнання в ньому.

2. Відзначено, що прогнозуванням можливих відмов та плануванням профілактичних заходів, можливо знизити витрати на обслуговування та мінімізувати простої виробничого обладнання.

3. Результати дослідження свідчать, що використання технологій штучного інтелекту сприяють покращенню продуктивності та надійності виробничих процесів, що є важливим кроком до оптимізації ефективного функціонування сучасного виробництва з належним рівнем надійності виробничого обладнання та ефективності його використання.

## Список літератури

1. Miguel Neves, Miguel Vieira, and Pedro Neto. "A study on a Q-Learning algorithm application to a manufacturing assembly problem". In: *Journal of Manufacturing Systems* 59 (2021), P. 426–440.
2. Аулін В.В., Гриньків А.В., Лисенко С.В., Голуб Д.В. Синергетика підвищення надійності машин використанням моделей марківських процесів. *Перспективи і тенденції розвитку конструкторської та технічного сервісу сх машин і знарядь*: зб. матеріалів доп. учасн. V Всеукраїнської науково-практичної конф. Житомир: Житомирський агротехнічний коледж, 2019. С. 242-245.
3. Аулін В. В., Гриньків А. В., Головатий А. О., Лисенко С. В., Голуб Д. В., Кузик О.В., Тихий А. А. Методологічні основи проектування та функціонування інтелектуальних транспортних і виробничих систем: монографія під заг. ред. д.т.н., проф. Ауліна В.В. Кропивницький: Видавель Лисенко В.Ф., 2020. 428 с.
4. Zhao, M.; Lu, H.; Yang, S.; Guo, F. The Experience-Memory Q-Learning Algorithm for Robot Path Planning in Unknown Environment. *IEEE Access* 2020, 8. P. 47824–47844.
5. Palacio, J.C.; Jiménez, Y.M.; Schietgat, L.; Van Doninck, B.; Nowé, A. A Q-Learning algorithm for flexible job shop scheduling in a real-world manufacturing scenario. *Procedia CIRP* 2022, 106. P. 227–232.

6. Ha D. Reinforcement learning for improving agent design. *Artificial life*. 2019. T. 25, № 4. P. 352–365. URL: [https://doi.org/10.1162/artl\\_a\\_00301](https://doi.org/10.1162/artl_a_00301).
7. Han R., Chen K., Tan C. Curiosity - driven recommendation strategy for adaptive learning via deep reinforcement learning. *British journal of mathematical and statistical psychology*. 2020. T. 73, № 3. P. 522–540. URL: <https://doi.org/10.1111/bmsp.12199>.
8. Inverse reinforcement learning-based time-dependent A\* planner for human-aware robot navigation with local vision / S. Sun et al. *Advanced robotics*. 2020. T. 34, № 13. P. 888–901. URL: <https://doi.org/10.1080/01691864.2020.1753569>.
9. L. A. P., Fu M. C. Risk-Sensitive reinforcement learning via policy gradient search. *Foundations and trends® in machine learning*. 2022. T. 15, № 5. P. 537–693. URL: <https://doi.org/10.1561/22000000091>.
10. Reinforcement learning and adaptive optimization of a class of Markov jump systems with completely unknown dynamic information / S. He et al. *Neural computing and applications*. 2019. T. 32, № 18. P. 14311–14320. URL: <https://doi.org/10.1007/s00521-019-04180-2>.
11. Reinforcement learning / B. L. Moore et al. *Anesthesia & analgesia*. 2011. T. 112, № 2. P. 360–367. URL: <https://doi.org/10.1213/ane.0b013e31820334a7>.
12. Reinforcement learning for logistics and supply chain management: methodologies, state of the art, and future opportunities / Y. Yan et al. *Transportation research part E: logistics and transportation review*. 2022. Vol. 162. 102712. URL: <https://doi.org/10.1016/j.tre.2022.102712>.
13. Dynamic handoff policy for RAN slicing by exploiting deep reinforcement learning / Y. Wu et al. *EURASIP journal on wireless communications and networking*. 2021. T. 2021, № 1. URL: <https://doi.org/10.1186/s13638-021-01939-x>.
14. Deep deterministic policy gradient for navigation of mobile robots in simulated environments / J. C. Jesus et al. 2019 19th international conference on advanced robotics (ICAR), м. Belo Horizonte, Brazil, 2–6 груд. 2019 р. 2019. URL: <https://doi.org/10.1109/icar46387.2019.8981638>.
15. Ковальов С.Г. Ковальов Ю.Г. Особливості реалізація моделі штучної нейронної мережі апаратними засобами. «Наука і техніка сьогодні» (Серія «Педагогіка», Серія «Право», Серія «Економіка», Серія «Фізико-математичні науки», Серія «Техніка»): журнал. 2024. №6(34) 2024. 1131с. URL: DOI: [https://doi.org/10.52058/2786-6025-2024-6\(34\)](https://doi.org/10.52058/2786-6025-2024-6(34))

## References

1. Neves, M., Vieira, M., & Neto, P. (2021). A study on a Q-Learning algorithm application to a manufacturing assembly problem. *Journal of Manufacturing Systems*, P. 426–440.
2. Aulin, V. V., Hrynkyv, A. V., Lysenko, S. V., & Holub, D. V. (2019). Synergetyka pidvyshchennia nadiinosti mashyn vykorystanniam modelei markivskykh protsesiv [Synergetics of increasing machine reliability using Markov process models]. *Perspektyvy i tendentsii rozvytku konstruktstii ta tekhnichnoho servisu shk mashyn i znariad: zb. materialiv dop. uchashnykiv V Vseukrainskoi naukovo-praktychnoi konf [Perspectives and trends in the development of constructions and technical service of machines and tools: collection. additional materials participation V All-Ukrainian Scientific and Practical Conf]* Zhytomyr: Zhytomyrskyi ahrotekhnichniyi koledzh, P. 242-245 (in Ukrainian).
3. Aulin, V. V., Hrynkyv, A. V., Holovaty, A. O., Lysenko, S. V., Holub, D. V., Kuzyk, O. V., & Tykhyi, A. A. (2020). Metodolohichni osnovy proektuvannia ta funktsionuvannia intelektualnykh transportnykh i vyrobnychykh system [Methodological principles of designing and functioning of intelligent transport and production system] (Prof. V.V.Aulin Ed) Kropyvnytskyi: Vydavets Lysenko V.F., 2020. 428 p.
4. Zhao, M., Lu, H., Yang, S., & Guo, F. (2020). The Experience-Memory Q-Learning Algorithm for Robot Path Planning in Unknown Environment. *IEEE Access*, 8. P. 47824–47844.
5. Palacio, J.C., Jiménez, Y.M., Schietgat, L., Van Doninck, B., & Nowé, A. (2022). A Q-Learning algorithm for flexible job shop scheduling in a real-world manufacturing scenario. *Procedia CIRP*, 106, P. 227–232.
6. Ha, D. (2019). Reinforcement learning for improving agent design. *Artificial Life*, 25(4), P. 352–365.
7. Han, R., Chen, K., & Tan, C. (2020). Curiosity-driven recommendation strategy for adaptive learning via deep reinforcement learning. *British Journal of Mathematical and Statistical Psychology*, 73(3), P. 522–540.
8. Sun, S. (2020). Inverse reinforcement learning-based time-dependent A\* planner for human-aware robot navigation with local vision. *Advanced Robotics*, 34(13), P. 888–901
9. L. A. P., & Fu, M. C. (2022). Risk-Sensitive reinforcement learning via policy gradient search. *Foundations and Trends® in Machine Learning*, 15(5), P. 537–693.
10. He, S., et al. (2019). Reinforcement learning and adaptive optimization of a class of Markov jump systems with completely unknown dynamic information. *Neural Computing and Applications*, 32(18), P. 14311–14320.

11. Moore, B. L., et al. (2011). Reinforcement learning. *Anesthesia & Analgesia*, 112(2), P. 360–367.
12. Yan, Y., et al. (2022). Reinforcement learning for logistics and supply chain management: methodologies, state of the art, and future opportunities. *Transportation Research Part E: Logistics and Transportation Review*. 102712.
13. Wu, Y., et al. (2021). Dynamic handoff policy for RAN slicing by exploiting deep reinforcement learning. *EURASIP Journal on Wireless Communications and Networking*, 2021(1).
14. Jesus, J. C., et al. (2019). Deep deterministic policy gradient for navigation of mobile robots in simulated environments. *19th International Conference on Advanced Robotics (ICAR)*, Belo Horizonte, Brazil.
15. Kovalov, S. H., & Kovalov, Yu. H. (2024). Osoblyvosti realizatsiia modeli shtuchnoi neironnoi merezhi aparatnymy zasobamy [Features of implementing the model of artificial neural network with hardware means]. *Nauka i tekhnika sohodni.., No6(34)*, 1131. (in Ukrainian).

**Victor Aulin**, Prof., DSc, **Serhiy Kovalov**, Doctoral student, PhDped. sci., **Andriy Hrynkiv**, Senior Researcher PhD tech. sci.

*Central Ukrainian National Technical University, Kropyvnytskyi, Ukraine*

**Valeriy Varvarov**, Leading researcher, PhDtech. sci..

*Kharkiv National Ivan Kozhedub Air Force University, Kharkiv, Ukraine*

### **Algorithm for Optimizing the Reliability of Operation and Efficiency of Use of Production Equipment Using Artificial Intelligence Methods**

This paper presents an innovative approach that involves the optimization of maintenance for production equipment by leveraging advanced artificial intelligence (AI) algorithms. The study explores the application of Markov process theory within the context of reinforcement learning and its integration into the modeling of production environments. The focus of this research is to address critical issues related to enhancing the reliability and efficiency of production processes. This is achieved by reducing maintenance costs and minimizing equipment downtime.

The proposed model for optimizing the use of the production equipment system is described as an environment with discrete states, where an agent operates with the capability to perform specific actions. This model emphasizes the implementation of the Q-learning algorithm, a form of reinforcement learning that aims to optimize production processes by enabling the agent to learn and make decisions that enhance system performance. Through the iterative process of Q-learning, the agent evaluates the potential benefits of various actions in different states, gradually refining its strategy to maximize long-term rewards.

Q-learning, with its ability to handle environments with unknown dynamics, is particularly effective in this context. It helps the agent to develop an optimal maintenance policy by balancing immediate maintenance costs against the long-term benefits of reduced downtime and extended equipment lifespan. The iterative nature of Q-learning ensures continuous improvement and adaptation, making the system robust against varying operational conditions and unforeseen disruptions.

Through rigorous analysis and experimentation, the findings demonstrate a significant improvement in the reliability and productivity of the production equipment system. The introduction of AI technologies, specifically through the Q-learning framework, not only streamlines maintenance practices but also ensures a more efficient utilization of resources, ultimately leading to a more robust and cost-effective production environment. The results of this study highlight the potential of AI in transforming traditional production maintenance strategies and setting new standards for operational excellence.

**production system, production equipment, artificial intelligence, Markov processes, reliability of functioning, efficiency of use**

*Одержано (Received) 15.10.2024*

*Прорецензовано (Reviewed) 19.10.2024*

*Прийнято до друку (Approved) 28.10.2024*