

О. М. Клименко

Харківський національний університет радіоелектроніки, Харків, Україна

КОНЦЕПЦІЯ РОБОТИ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ НА БАЗІ МЕТОДУ НАЙБЛИЖЧИХ СУСІДІВ ДЛЯ БАЛАНСУВАННЯ НАВАНТАЖЕННЯ НА КОНВЕЄРНИХ ЛІНІЯХ СОРТУВАННЯ ПРОДУКЦІЇ

Анотація. Ця стаття представляє дослідження концепції роботи алгоритмів машинного навчання з учителем на основі методу найближчих сусідів для оптимізації процесу балансування навантаження на фармацевтичних лініях сортування. Сфера фармацевтики стикається з навантаженням, що постійно змінюється, і різноманітними характеристиками упаковок, що вимагає ефективних стратегій управління виробничими процесами. Автор пропонує концепцію, що базується на принципі найближчих сусідів, для визначення оптимального навантаження для кожної зони сортування. Використання цього методу дозволяє враховувати контекст і схожість зон сортування, що призводить до точних та адаптивних рішень. Роботи-сортирувальники, що працюють на фармацевтичних конвеєрних лініях, стають більш ефективними у розподілі упаковок по всіх робочих зонах, що призводить до зниження часу складання замовлень та оптимізації ресурсів.

Ключові слова: індустрія 4.0, розумне виробництво, логістика 4.0, складське господарство 4.0, Інтернет речей, бережливе виробництво, балансування навантаження, сортувальна конвеєрна лінія, фармацевтичне виробництво.

Вступ

У сучасному світі, де концепції Індустрія 4.0, Розумне виробництво, Логістика 4.0, Складське господарство 4.0, відіграють ключову роль у перетворенні виробничих процесів, увага до оптимізації та ефективності у сфері фармацевтики стає критично важливою [1,2]. Науковці розглядають цілу низку проблем, пов'язаних із впровадженням принципів концепції Складське господарство 4.0 (Warehousing 4.0). Серед цих проблем можна виділити використання шатлів [3-5], питання логістики на фармацевтичних підприємствах [6, 7], розробки програмного забезпечення для складів [8], розробки та роботи конвеєрної стрічки [9], використання роботів [10, 12] тощо.

На тлі цього контексту дана стаття спрямована на дослідження застосування передових технологій у галузі балансування навантаження на фармацевтичних лініях сортування, що є надзвичайно актуальною задачею [13-17].

Основним фокусом дослідження є використання алгоритмів машинного навчання з учителем для оптимізації процесу сортування упаковок. Роботи-сортирувальники, що працюють на фармацевтичних конвеєрних лініях, стикаються з навантаженням, що постійно змінюється, і різними характеристиками упаковок.

Ми доповнили наше дослідження завдяки інтеграції Інтернету речей (IoT) та принципів Бережливого виробництва, забезпечуючи систему реального часу для збору даних про навантаження, швидкість та інші фактори, що впливають на ефективність сортування. Цей комплексний підхід не тільки покращує виробничі показники, а й сприяє більш стійкій та адаптивній фармацевтичній логістиці. Актуальність досліджень проявляється у необхідності мінімізації трудомісткості формування замовлень та підвищення ефективності роботи конвеєрних ліній. Запропоновані алгоритм машинного навчання надають інтелектуальні інструменти для досягнення фармацевтичної

індустрії, що динамічно розвивається, де ефективність виробництва надзвичайно важливою для забезпечення конкурентоспроможності.

Загальний принцип роботи алгоритму машинного навчання з учителем

Загальний принцип роботи алгоритму машинного навчання з учителем (Supervised Learning) полягає у використанні розмічених даних для навчання моделі, яка здатна робити передбачення на нових, раніше невідомих даних. Цей тип машинного навчання має на увазі наявність "вчителя" або "експерта", який надає моделі дані та відповідні їм правильні відповіді. Основні кроки, які включає процес навчання алгоритму машинного навчання з учителем:

- підготовка даних. Завантаження та попередня обробка даних, включаючи очищення, масштабування, перетворення ознак та інші операції. Поділ даних на навчальний та тестовий набори для оцінки продуктивності моделі;

- вибір моделі. Вибір архітектури моделі в залежності від характеру задачі (класифікація, регресія і т.д.). Вибір алгоритму, який найкраще підходить для вирішення конкретної задачі.

- навчання моделі. Подача навчальних даних на вхід моделі. Модель "навчається" на основі розмічених прикладів, коригуючи свої ваги або параметри для мінімізації помилки між передбаченими значеннями та реальними відповідями;

- оцінка моделі. Використання тестового набору даних для оцінки продуктивності моделі. Використання метрик оцінки (наприклад, точність, F1 міра, середньоквадратична помилка), щоб визначити, наскільки добре модель справляється з новими даними;

- налаштування гіперпараметрів. При необхідності проведення налаштувань гіперпараметрів для покращення продуктивності моделі.

- прогнозування нових даних. Після успішного навчання модель може використовуватися для передбачення відповідей на нові, раніше невідомі дані.

Прикладами алгоритмів машинного навчання з учителем є лінійна регресія, метод найближчих сусідів, дерева рішень, метод опорних векторів (SVM), ансамблі (наприклад, випадковий ліс) та глибокі нейронні мережі.

Розробка методу машинного навчання з учителем на основі методу найближчих сусідів

Метод найближчих сусідів (k-Nearest Neighbors, k-NN) – це алгоритм класифікації та регресії, заснований на принципі близькості об'єктів у просторі ознак. Розглянемо застосування методу найближчих сусідів балансування навантаження на фармацевтичних лініях сортування.

Визначимо у межах даного дослідження під поняттям простір ознак будемо розуміти математичне поняття, що визначає всі можливі комбінації значень ознак кожного об'єкта. Позначимо його як X і кожен об'єкт представлений у цьому просторі як вектор ознак. Нехай X є n -простір, де n – кількість ознак. Кожен об'єкт представлений вектором ознак:

$$x_i = (x_{i1}, x_{i2}, \dots, x_{in}), \quad (1)$$

де x_{in} – ознака, яка може бути числовим значенням, категоріальним (фактором) або бінарним.

Внаслідок цього матриця ознак X містить усі об'єкти та їх ознаки. Якщо у нас є m об'єктів та n ознак, то матриця буде розміру $m \times n$.

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix}. \quad (2)$$

У межах цих досліджень уявімо, що є фармацевтична лінія сортування, і необхідно визначити оптимальне навантаження для кожної зони. Ознаками можуть бути такі параметри: поточне навантаження (x_1), швидкість обробки (x_2), ефективність (x_3) та інші параметри. Виходячи з цього вектор ознак для першої робочої зони роботів-сортувальників можна подавати наступним чином:

$$x_1 = (x_{11}, x_{12}, x_{13}). \quad (3)$$

Простір ознак надає абстрактний, математичний фреймворк для подання та аналізу даних, що є ключовим компонентом багатьох методів машинного навчання.

Розглядаючи завдання балансування навантаження на фармацевтичних лініях сортування, цільовою змінною може бути оптимальне навантаження для кожної зони сортування.

Цільова змінна (y) в математичній моделі є величиною, яку ми прагнемо передбачити або пояснити з використанням ознак з простору ознак. У різних завданнях машинного навчання цільова змінна може бути числовою (у задачах регресії) або категоріальною (у задачах класифікації). Якщо ми

розглядаємо завдання балансування навантаження на фармацевтичних лініях сортування, цільовою змінною може бути оптимальне навантаження для кожної зони сортування.

У задачі регресії:

$$y_i \in R. \quad (4)$$

У задачі класифікації:

$$y_i \in \{C_1, C_2, \dots, C_k\}, \quad (5)$$

де y_i – цільова змінна для зони i .

Мета полягає в побудові моделі $f(X)$, яка максимально точно передбачає або класифікує цільову змінну y на основі вхідних ознак з простору X .

Наступною дією необхідно визначити відстані в математичних моделях, в контексті методу найближчих сусідів (k-NN) це є мірою віддаленості між двома об'єктами в просторі ознак. Це поняття необхідно для визначення "близькості" об'єктів i , таким чином, для ухвалення рішення у методі найближчих сусідів. Нехай нас є два об'єкти, представлені векторами ознак x_i і x_j . Відстань між ними позначимо як $d(x_i, x_j)$. Виходячи з цього можна використовувати такі типи відстаней:

– евклідова відстань:

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}; \quad (6)$$

– манхеттенська відстань

$$d(x_i, x_j) = \sum_{k=1}^n |x_{ik} - x_{jk}|; \quad (7)$$

– косинусна відстань

$$d(x_i, x_j) = 1 - \frac{\sum_{k=1}^n x_{ik} \cdot x_{jk}}{\sqrt{\sum_{k=1}^n (x_{ik})^2} \cdot \sqrt{\sum_{k=1}^n (x_{jk})^2}}. \quad (8)$$

Вибір конкретної метрики залежить від природи даних та характеру завдання. Наприклад, евклідова відстань часто використовується, коли важливо враховувати абсолютні значення ознак. При використанні k-NN для кожного об'єкта обчислюються відстані до всіх інших об'єктів у вибірці (6-8). Потім обираються k найближчих сусідів з урахуванням цих відстаней. Тобто для кожної зони сортування x_i знаходимо k найближчих сусідів за обраною метрикою відстані. Зважуємо їх цільові змінні y_i на основі відстані, наприклад через зворотню відстань:

$$y_i = \left(\frac{1}{d(x_i, x_j)} \right). \quad (9)$$

Це дає можливість обчислити зважене середнє для визначення оптимального навантаження для зони x_i .

Коли з'являються нові дані (нова зона сортування), використовуємо алгоритм k-NN для прогнозування оптимального навантаження на основі найближчих сусідів.

Висновки

У процесі дослідження концепції роботи алгоритмів машинного навчання з учителем, заснованих на методі найближчих сусідів, для балансування навантаження на конвеєрних лініях сортування при розподілі упаковок по робочих зонах роботів сортувальників було виявлено значні та перспективні результати: моделі, засновані на даному методі, успішно прогнозують оптимальні навантаження для

різних зон сортування; враховуються їх характеристики та розподіл упаковок.

Використання різних метрик відстані, таких як евклідова, манхеттенська та косинусна відстані, дозволило обрати найкращу метрику в залежності від особливостей конкретного завдання. Це додатково підкреслило гнучкість та адаптивність методу до різноманітних умов роботи фармацевтичних ліній сортування.

Отримані результати мають пряме практичне застосування у фармацевтичній індустрії та можуть бути узагальнені на інші галузі промисловості з аналогічними принципами сортування та розподілу продукції.

СПИСОК ЛІТЕРАТУРИ

1. 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>
2. Nevludov, I., & et al. (2020). Method of Algorithms for Cyber-Physical Production Systems Functioning Synthesis. International Journal of Emerging Trends in Engineering Research, 8(10), 7465-7473.
3. Maksymova, S., & et al. (2023). Shuttle-Based Storage And Retrieval System 3d Model Improvement and Development. Journal of Natural Sciences and Technologies, 2(2), 232-237.
4. Nevludov, I., & et al. (2022). Analysis of Software Products for Simulation Modeling of the Operation of the System of Shuttles for Warehousing. In Manufacturing & Mechatronic Systems 2022: Proceedings of VIst Int. Conf., Kharkiv, 24-26.
5. Igor, N., & et al. (2023). (2023). Using Mecanum Wheels for Radio Shuttle. Multidisciplinary Journal of Science and Technology, 3(3), 182-187.
6. Igor, N., & et al. (2023). Automated Logistics Processes Improvement in Logistics Facilities. Multidisciplinary Journal of Science and Technology, 3(3), 157-170.
7. Nevludov, I., & et al. (2023). Features of Wave Algorithm Application in Warehouse Logistics Transport Systems. Information systems in project and program management: Coll. mon. European University Press. Riga: ISMA, 251-261.
8. Nevludov, I., & et al. (2023). Software development for small details production warehouse automated system. Scientific Collection «InterConf», 320-323.
9. V. V. Yevsieiev, & et al. (2023) Conveyor Belt Object Identification: Mathematical, Algorithmic, and Software Support. Appl. Math. Inf. Sci. 17, No. 6. - P. 1073-1088.
10. Yevsieiev, V., & Gurin, D. (2023) Comparative Analysis of the Characteristics of Mobile Robots and Collaboration Robots Within INDUSTRY 5.0. In the VI International Scientific and Theoretical Conf., September 8, 2023. Chicago, USA. 92-94
11. Yevsieiev V. Some aspects of the development of the BEAM robot control scheme. In IV International Scientific and Theoretical Conference, Singapore, Republic of Singapore. - P. 79-81.
12. Невлюдов І. Ш. та інші (2024) BEAM робототехніка : навч. посіб. Кривий Ріг : Вид. Чернявський Д. О., 2024. – 276 с.
13. Kuo, Yiyo, Ssu-Han Chen, Taho Yang, and Wei-Chen Hsu. 2023. "Optimizing a U-Shaped Conveyor Assembly Line Balancing Problem Considering Walking Times between Assembly Tasks" Applied Sciences 13, no. 6: 3702. <https://doi.org/10.3390/app13063702>
14. Zhang, S. & Xia, X. (2010). Optimal control of operation efficiency of belt conveyor systems. Appl. Energy, 87, 1929–1937.
15. Ponnambalam, S.G. & et al. (1999). A comparative evaluation of assembly line balancing heuristics. Int. J. Adv. Manuf. Technol. 1999, 15, 577–586.
16. Oksuz, M.K.; Buyukozkan, K.; Satoglu, S.I. U-shaped assembly line worker assignment and balancing problem: A mathematical model and two meta-heuristics. Comput. Ind. Eng. 2017, 112, 246–263.
17. Khalid, M. S., & et al. (2022). HMI Development Automation with GUI Elements for Object-Oriented Programming Languages Implementation. International Journal of Engineering Trends and Technology, 70.1, 139-145.

Received (Надійшла) 23.11.2023

Accepted for publication (Прийнята до друку) 17.01.2024

Concept of supervised machine learning algorithms based on the k-nearest neighbors method for load balancing on pharmaceutical sorting lines

Oleksandr Klymenko

Abstract. This article presents a study on the concept of supervised machine learning algorithms based on the k-nearest neighbors method to optimize the load balancing process on pharmaceutical sorting lines. The pharmaceutical industry faces constantly changing loads and diverse packaging characteristics, requiring efficient strategies for managing production processes. The author proposes a concept based on the nearest neighbors principle to determine the optimal load for each sorting zone. The use of this method allows considering the context and similarity of sorting zones, leading to accurate and adaptive solutions. Sorting robots operating on pharmaceutical conveyor lines become more efficient in distributing packages across all working zones, resulting in reduced order assembly time and resource optimization.

Keywords: Industry 4.0, Smart Manufacturing, Logistics 4.0, Warehousing 4.0, Internet of Things, Lean Production, Load Balancing, Sorting Conveyor Line, Pharmaceutical Production.