

Використання машинного навчання для професійного розвитку персоналу

Любомудрова Надія Петрівна¹, Ямщиков Ігор Ігорович²

Опубліковано	Секція	УДК
24.05.2024	Економіка	005.95:004.8

DOI: <https://doi.org/10.5281/zenodo.11275894>

Ліцензовано за умовами Creative Commons BY 4.0 International license

Анотація. У сучасному світі, де швидкість розвитку технологій невідомо зростає, особливої уваги набуває питання застосування машинного навчання для ефективного професійного розвитку персоналу. У статті, присвяченій цій тематиці, запропоновано детальний аналіз використання алгоритмів машинного навчання, які можуть значно оптимізувати процеси навчання та розвитку працівників у різних організаціях. Автори фокусують увагу на тренуванні моделей машинного навчання, зокрема, на таких методах, як дерева рішень та кластеризація, та їх використанні для автоматизації процесів навчання. Дерева рішень представляють собою ефективний інструмент для класифікації та прогнозування. Використання дерев рішень дозволяє компаніям розробляти персоналізовані навчальні плани, які враховують особливості кожного працівника, їхній досвід, кваліфікацію та поточні потреби в навчанні.

Ключові слова: машинне навчання, професійний розвиток, дерева рішень, кластеризація, автоматизація навчання, навчання персоналу, розвиток персоналу, управління персоналом.

Using machine learning for the professional staff development

Annotation. In today's world, where the pace of technological development is relentlessly increasing, the application of machine learning for effective professional development of personnel is gaining particular attention. The article dedicated to this topic offers a detailed analysis of the use of machine learning algorithms that can significantly optimize the processes of employee training and development in various organizations. The authors focus on training machine learning models, particularly on methods such as decision trees and clustering, and their use in automating training processes. Decision trees represent an effective tool for classification and prediction. The use of decision trees allows companies to develop personalized training plans that consider each employee's individual characteristics, experience, qualifications, and current training needs.

Keywords: machine learning, professional development, decision trees, clustering, training automation, staff training, staff development, personnel management.

¹ к.е.н, доцент, доцент кафедри менеджменту персоналу та адміністрування, НУ «Львівська політехніка», ORCID: <https://orcid.org/0000-0002-1982-0442>

² аспіранти кафедри менеджменту персоналу та адміністрування, НУ «Львівська політехніка» ORCID: <https://orcid.org/0009-0005-7366-8129>

Вступ

Застосування штучного інтелекту, машинного навчання та інших сучасних технологій у навчанні персоналу набуває все більшої популярності протягом останніх років. Ця популярність пов'язана з успіхом машинного навчання – підходом до розробки систем штучного інтелекту, де моделі, які використовуються в HR-менеджменті, тренуються на базі даних, зібраних з прикладів реального світу. Важливим моментом є розуміння принципів роботи моделей машинного навчання та штучного інтелекту для покращення ефективності навчання персоналу.

Проблематикою управління персоналом займається багато дослідників; зокрема серед українських науковців варто виділити: Балановську Т.І., Михайліченко М.В., Троян А.В., які у своїх працях розкривають сутність та види оцінювання персоналу, приділяють увагу основним завданням оцінки роботи персоналу організації, розглядають елементи процесу оцінювання персоналу. [1] Т. П. Збрицька, Г. О. Савченко, М. С. Татаревська працювали над проблемами розвитку персоналу на основі його оцінювання, навчання та підвищенні кваліфікації персоналу організації та інших питаннях. [2] Серед зарубіжних учених варто відзначити Неупане Р. та Ханал Т. Р, які у своїх працях розглядають, як навчання персоналу впливає на продуктивність працівників, проводять аналіз на основі дослідження даних щодо навчання персоналу працівників конкретних компаній, вивчають вплив різних типів навчання, таких як, навчання на робочому місці. У публікаціях цих науковців проаналізовано вплив онлайн-навчання на продуктивність працівників, на основі проведеного дослідження сформовано рекомендації для компаній, щоб допомогти покращити навчальні програми. У цих працях вчені здебільшого намагаються встановити кореляційний зв'язок між різними показниками щодо навчання персоналу та продуктивності працівників. [3]

Метою статті є здійснити огляд та аналітичне дослідження наявних алгоритмів машинного навчання з акцентом на вивченні їхніх архітектурних особливостей та механізмів функціонування. Крім цього, одним із завдань даної статті є визначення доцільності застосування проаналізованих моделей для професійного корпоративного навчання персоналу.

Результати

У сучасному науковому дискурсі штучний інтелект стає об'єктом зростаючої уваги та інтересу. Популярність штучного інтелекту безпосередньо пов'язана з успіхами машинного навчання, методології розробки систем на основі аналізу реальних прикладів. Застосування машинного навчання демонструє вражаючу універсальність, внаслідок чого ми спостерігаємо поширення інструментів на основі штучного інтелекту у всіх сферах економіки та життєдіяльності. Сфера управління людськими ресурсами не залишається осторонь від цієї тенденції. Ці інструменти надають значні перспективи та викликають інтерес завдяки своїй здатності швидко обробляти інформацію, а також потенціалу покращення процесів управління персоналом, що може призвести до прийняття більш обґрунтованих рішень та покращення результатів. Різноманіття таких інструментів відображає креативність та інноваційність, які стимулюються останніми досягненнями в області штучного інтелекту, оскільки розробники прагнуть вирішити існуючі проблеми в сфері управління ресурсами та розширити можливості в нові сфери діяльності.

Розробники створюють інструменти на основі штучного інтелекту майже для кожної ланки в ланцюжку процесів управління персоналом.

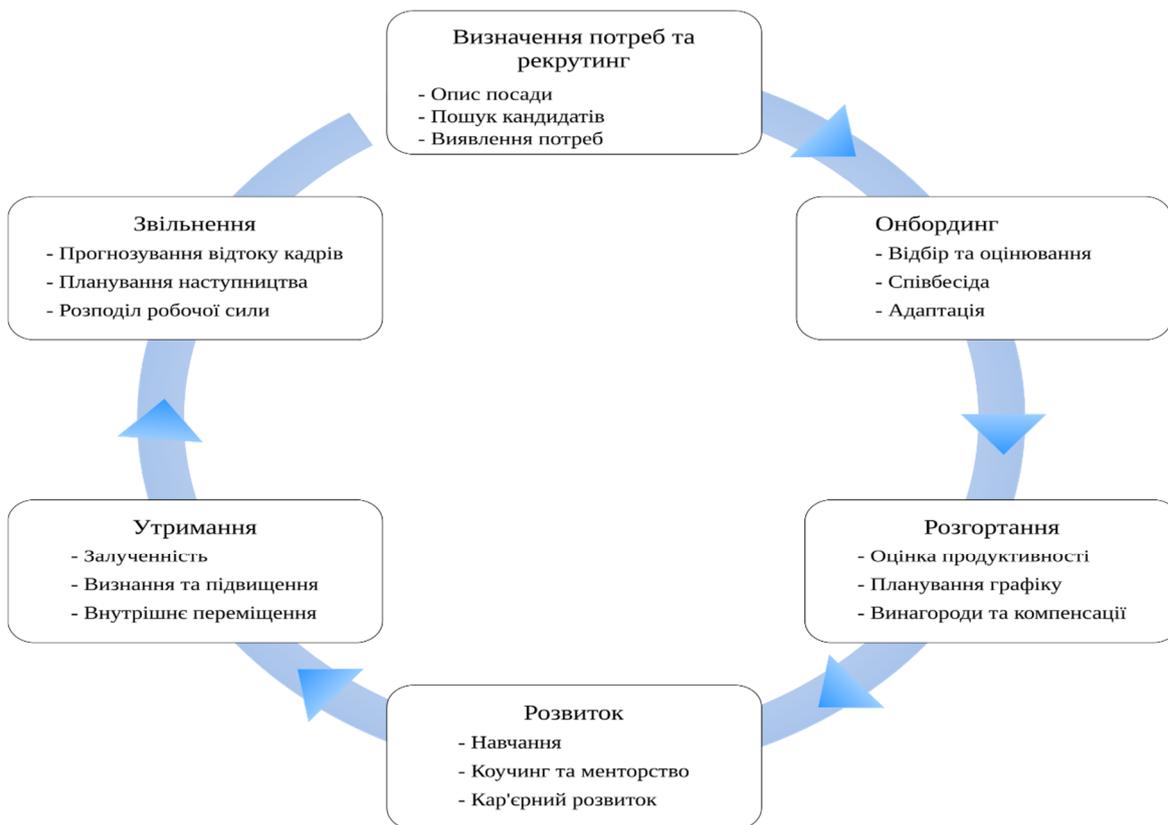


Рис. 1. Послідовність функціонування окремих стадій в процесі управління персоналом

Джерело: розроблено авторами на основі джерела [4]

На рис. 1 відображено послідовність функціонування окремих стадій в процесі управління персоналом, а також і різноманіття інструментів, які стимулюються останніми досягненнями в області штучного інтелекту, прагнуть вирішити існуючі проблеми в сфері управління ресурсами та розширити можливості використання у нових сферах діяльності. Ці інструменти можуть бути застосовані у кожній з відображених стадій у процесі управління персоналом. [4] Розробники створюють інструменти на основі штучного інтелекту майже для кожної ланки в ланцюжку процесів управління персоналом. Проте, в межах нашої статті ми акцентуватимемо увагу на стадії розвитку персоналу, зокрема навчання, визначення потреби у менторстві чи наставництві. Також ми розглянемо, які моделі машинного навчання доцільно застосовувати для створення рекомендацій щодо кар'єрного розвитку працівників. Для того, щоб зрозуміти, які саме підходи та моделі штучного інтелекту доцільно використовувати для навчання персоналу, ми розглянемо декілька основних типів моделей, а також принципи реалізації їхніх алгоритмів. У рамках даної статті ми не розглядатимемо такі концепції, як PAC навчання, теорема No-Free-Lunch, VC-Dimension. Хоча вони надають важливе теоретичне підґрунтя, детальне вивчення цих концепцій не вважаємо необхідним для практичного застосування штучного інтелекту в контексті навчання та оцінювання персоналу. Ці концепції вимагають глибокого розуміння статистики та теорії машинного навчання, що може бути занадто складним і часозатратним для фахівців HR-менеджменту, які прагнуть швидко застосувати штучний інтелект для конкретних завдань. Натомість, ми розглянемо дерева рішень та алгоритми кластеризації.

Дерево рішень є таким інструментом прогнозування, який можна визначити за формулою (1):

$$h: X \rightarrow Y \quad (1)$$

де, h визначає мітку, асоційовану з конкретним випадком X , шляхом проходження від кореневого вузла дерева до листка. Для спрощення ми зосередимось на бінарній класифікації, тобто, $Y = \{0,1\}$, хоча дерева рішень також можуть бути застосовані для інших задач прогнозування. На кожному вузлі шляху від кореня до листка вибір наступного дочірнього вузла здійснюється на основі розбиття вхідного простору. Зазвичай, розбиття базується на одній з ознак x або на попередньо зазначеному наборі правил розбиття. Листки містять специфічну мітку. [5]

Загальна схема побудови дерева рішень є такою: ми починаємо з дерева з одним листком (коренем) і присвоюємо цьому листку мітку згідно з більшістю голосів серед усіх міток навчального набору. Потім ми виконуємо серію ітерацій. На кожній ітерації ми аналізуємо ефект від розбиття одного листка. Ми визначаємо деяку міру «приросту», яка кількісно оцінює покращення внаслідок цього розбиття. Потім, серед усіх можливих розбиттів, ми або вибираємо той, який максимізує приріст і виконуємо його, або вирішуємо взагалі не розбивати листок. Існує багато можливих варіантів визначення вище згаданої міри «приросту». Кожна така реалізація містить свої недоліки та позитивні сторони, одним з можливих варіантів яких є приріст інформації – різниця між ентропією мітки до і після розбиття, і досягається шляхом використання функції ентропії, що визначається формулою (2):

$$C(a) = a \log(a) - (1 - a) \log(1 - a) \quad (2)$$

Важливо зазначити, що дерево рішень не є результатом програмування, тобто не визначається логічними операторами певної мови програмування, а є результатом тренування моделі машинного навчання на певному наборі даних. [5]

Наприклад, якщо модель тренувати на основі даних про стаж роботи працівників, враховувати їхні посади, а у даних для верифікації результатів тренування використовувати оцінку ефективності працівників та її кореляцію з курсами підвищення кваліфікації, що були призначені цим працівникам, то в результаті тренування ми отримаємо певну модель дерева рішень. Ця модель матиме змогу розглядати стаж роботи працівника та його посаду для визначення траєкторії професійного розвитку працівника шляхом надання рекомендацій щодо призначення курсів відповідно до стажу та посади. Наприклад, якщо стаж роботи в компанії менше 1 року, то в такому випадку модель негайно пропонує пройти базовий курс, розроблений для нових працівників. Якщо ж працівник має стаж роботи від 1 року і більше, то, зважаючи на посаду працівника, буде обрано один з двох інших курсів: управлінський курс для менеджерів, або сертифікаційний курс для інженерів, відповідно до рівня кваліфікації. Іншими словами, під час навчання на тренувальних даних для побудови дерева рішень алгоритм випробовує різні можливі варіанти розбиття на гілки, а також маркування відповідних розбиттів, оцінюючи після цього ефективність такого розбиття. Цей процес повторюється допоки не знаходиться найбільше ефективно розбиття, що врешті і стане натренованою моделлю, яка дозволить приймати рішення, використовуючи вже нові дані. Проблемою алгоритмів такого типу є те, що дерево, яке вони розробляють, зазвичай буде дуже великим. Такі дерева можуть мати низький емпіричний ризик, але їхній реальний ризик, як показує теоретичний аналіз та практика, схильний бути високим. Одним із рішень є обмеження кількості ітерацій алгоритму, що призводить до створення дерева з обмеженою кількістю вузлів. Візуалізацію такого дерева рішень для визначення траєкторії розвитку працівника зображено на рис.2:



Рис.2 Візуалізація дерева рішень для визначення напрямку розвитку працівника

Джерело: розроблено авторами на основі джерела [5]

Кластеризація є ще одним із найбільш поширених методів для початкового аналізу даних, що застосовується у великому діапазоні областей від соціальних наук до біології та інформатики. Вона дозволяє дослідникам отримати перші враження про набір даних шляхом групування точок даних у значимі категорії. По суті, кластеризація – це процес розподілу об'єктів на групи за принципом схожості, де об'єкти в одній групі мають високий рівень подібності, а об'єкти різних груп – значні відмінності. Однією з проблем кластеризації є можливість існування довгого ланцюжка об'єктів, x_1, \dots, x_m , де кожне x_i дуже схоже на своїх двох сусідів, x_{i-1} та x_{i+1} , але x_1 і x_m дуже відрізняються між собою. Якщо ми хочемо, щоб схожі елементи завжди були у одному кластері, тоді ми повинні помістити всі елементи ланцюжка в один кластер. У такому випадку ми отримуємо несхожі елементи (x_1 і x_m), які опиняються в одному кластері, тим самим порушуючи другу вимогу. [6]

У контексті навчання персоналу кластеризація може бути корисним елементом для визначення груп людей, які потребують додаткового навчання, і розділення працівників за результатами навчання.

Розглянемо алгоритми кластеризації на основі зв'язків. Алгоритми кластеризації, засновані на принципі зв'язності, є найбільш базовими та інтуїтивно зрозумілими підходами до групування даних. Ці методи реалізують послідовність кроків, починаючи зі стадії, на якій кожен елемент даних утворює власний кластер. Далі, у процесі виконання, алгоритми поетапно зливають між собою найменш віддалені один від одного кластери згідно з попередньою кластеризацією. Це призводить до зменшення кількості кластерів на кожному етапі. Продовжуючи цей процес до кінця, можна досягти ситуації, коли всі елементи опиняться в єдиному загальному кластері. Таким чином, для точного визначення алгоритму необхідно встановити два основні параметри: метод визначення відстані між кластерами та критерій для припинення процесу об'єднання.[6]

Кластеризацію на основі зав'язків можна розділити на 3 типи:

1. Кластеризація за методом найближчого сусіда – визначає відстань між кластерами як мінімальну відстань між елементами цих двох кластерів.
2. Кластеризація за методом середнього зв'язку визначає відстань між двома кластерами як середню відстань між точкою в одному кластері та точкою в іншому.
3. Кластеризація за методом найвіддаленішого сусіда визначає відстань між двома кластерами як максимальну відстань між їх елементами.

Алгоритми кластеризації, що базуються на принципах зв'язності, працюють за агломеративним принципом: вони починають із даних, які є цілковито розділеними, та крок за кроком об'єднують їх у все більш об'ємні групи. За відсутності критерію для припинення процесу, кінцевий результат такої кластеризації можна зобразити за допомогою дендрограми, деревоподібної структури, що відображає ієрархію підмножин даних, де на найнижчому рівні знаходяться індивідуальні елементи, а в корені – весь домен даних. Щоб перетворити дендрограму на поділ простору (кластеризацію), необхідно використати критерій зупинки. Можна використовувати такі критерії:

- 1) фіксована кількість кластерів – встановити обмежуючий параметр k і припинити об'єднання кластерів, як тільки їх кількість досягне k
- 2) верхня межа відстані – встановити деяке значення $r \in R$. Припинити об'єднання, як тільки всі відстані між кластерами стануть більші за r . Можна також встановити r за формулою (3):

$$a \max\{d(x, y) : x, y \in X\} \quad (3)$$

для деякого $a < 1$. У цьому випадку критерій зупинки називають «масштабована верхня межа відстані». [6]

Складність використання алгоритмів кластеризації для навчання персоналу швидше за все полягатиме у коректному виборі критерію зупинки, а також в отриманні необхідної вибірки даних для тренувань.

Висновки

У ході написання цієї роботи було проведено огляд сучасних підходів до реалізації моделей штучного інтелекту, а саме – використання кластеризації та дерев рішень. З урахуванням стрімкого розвитку технологій та зростаючого інтересу до автоматизації процесів професійного навчання ми зосередили нашу увагу на ключових аспектах інтеграції штучного інтелекту у навчання та розвиток персоналу. Аналізуючи потенціал дерев рішень та методів кластеризації в контексті розробки методик навчання персоналу та автоматизації процесів навчання, можна зробити висновок про високу ефективність цих підходів. Попри їхню відносну простоту, порівняно з іншими методами машинного навчання, дерева рішень та кластеризація дозволяють ефективно структурувати вхідні дані, що може слугувати основою для подальшого покращення процесів навчання та розвитку персоналу. Застосування дерев рішень може допомогти у виборі оптимальних навчальних курсів для різних категорій працівників на основі їхніх поточних компетентностей та потреб у розвитку. Кластеризація, в свою чергу, дозволяє ідентифікувати групи співробітників зі схожими потребами у навчанні, сприяючи більш цілеспрямованому підходу до розроблення навчальних програм та організування навчальних заходів. Таким чином, обидва ці методи разом забезпечують можливість не тільки ефективно організувати навчальний процес, але й оцінити його результати, що є критично важливим для постійного покращення якості навчання. Як зазначалось вище, у практичній роботі складністю використання алгоритмів кластеризації може бути вибір коректного критерію зупинки, а для дерев рішень – це обмеження кількості ітерацій. Визначення такого критерію зупинки та оптимальної

кількості операцій для моделей, націлених на навчання персоналу, може стати перспективним елементом у подальших дослідженнях.

Список використаних джерел

1. Збрицька Т. П., Савченко Г. О., Татаревська М. С. Управління розвитком персоналу. *Атлант*, 2013.
2. Балановська Т. І., Михайліченко М. В., Троян А. В. Сучасні технології управління персоналом. *ФОП Ямчинський О.В.*, 2020.
3. Neupane R., Khanal T. R. Evaluation of the Impacts of Training on Employee Performance in The Uk Hotel Industry: An Analysis of Employees' Perspectives at Marriott International. *International Journal of Research*, 2022
4. Human-Centred Artificial Intelligence for Human Resources – офіційний сайт World Economic Forum. URL: <https://www.weforum.org/publications/human-centred-ai-for-hr-state-of-play-and-the-path-ahead/toolkit/>
5. Ben-David S., Shalev-Shwartz S. Understanding Machine Learning: From Theory to Algorithms From Theory to Algorithms. *Cambridge University Press*, 2014
6. Bishop C. M. Pattern Recognition and Machine Learning. *Springer Science+Business Media, LLC*, 2006