

УДК 681.513

ЗАДАЧА БАГАТОЗАДАЧНОГО НАВЧАННЯ ЯК ОДНА З ВАЖЛИВИХ ЗАДАЧ МАШИННОГО НАВЧАННЯ

Є.А. Савченко

*Міжнародній науково-навчальній центрі інформаційних технологій
та систем НАН та МОН України*

savchenko@irtc.org.ua

Проведено огляд проблеми багатозадачного навчання. Описано такі близькі поняття, як метанавчання, багатозадачне навчання, трансферне навчання та індуктивний перенос. Представлено огляд різних завдань багатозадачного навчання. Показано схожість індуктивного підходу та індуктивного переносу як одного із способів розв'язання задач багатозадачного навчання.

Ключові слова: багатозадачне навчання, машинне навчання, індуктивний підхід, узагальнення.

Проведен обзор проблемы многозадачного обучения. Описаны такие близкие понятия, как метаобучение, многозадачное обучение, трансферное обучение и индуктивный перенос. Представлен обзор разных задач многозадачного обучения. Показано сходство индуктивного подхода и индуктивного переноса как одного из способов решения задачи многозадачного обучения.

Ключевые слова: многозадачное обучение, машинное обучение, индуктивный подход, обобщение,

A review of the problem of multitasking learning. Close concepts such as metalearning, multitasking, transfer learning, and inductive transfer are described. An overview of the different tasks of multitasking learning is presented. The similarity of the inductive approach and inductive transfer as one of the ways to solve the problem of multitask learning is shown.

Keywords: multi-task learning, inductive approach, generalization, machine learning.

Вступ

Галузь машинного навчання є однією з підгалузей штучного інтелекту, за допомогою якої можна перекласти розв'язання низки завдань з людини на комп'ютер та інші пристрої.

Навчання відбувається за рахунок машинної обробки величезних масивів вхідних даних та знаходження в них закономірностей. Застосування машинного навчання дозволяє розв'язувати як рутинні завдання, так і більш складні завдання, що містять багато входів та виходів й часто пов'язані між собою.

Багатозадачне навчання є однією з галузей машинного навчання, яка дає можливість покращити ефективність виконання декількох завдань навчання, що виконуються одночасно, узагальнюючи та максимально використовуючи з них корисну інформацію.

За допомогою цього підходу можна підвищити продуктивність розв'язання задач навчання у багатьох сферах, таких як комп'ютерний зір, біоінформатика, інформатика здоров'я, мовлення, природна обробка мови, веб-додатки та різні комп'ютерні обчислення.

Голова підрозділу штучного інтелекту компанії Google Джефф Дін рекомендував колегам звернути увагу на такі алгоритми як багатозадачне та трансферне навчання, оскільки він бачить їх перспективними та ефективними алгоритмами, що дозволять досягти однієї тієї ж мети, але зі значно меншими затратами [1].

Метою цього дослідження є проаналізувати проблему багатозначного навчання як одну з перспективних задач машинного навчання, визначити особливості її розв'язання та продемонструвати можливість застосування та місце індуктивного підходу в розв'язанні цієї задачі.

1. Визначення понять *metalearning*, *multitask learning* та *transfer learning*

Однією з основних властивостей систем, які побудовані за принципами штучного інтелекту, є здатність накопичувати досвід усіх попередньо розв'язаних завдань для використання цього досвіду з метою ефективного розв'язання нових завдань. Методи та засоби машинного навчання є інструментом створення таких систем. Метою машинного навчання є часткова або повна автоматизація знаходження розв'язків складних професійних завдань в різних галузях людської діяльності.

Машинне навчання (англ. *Machine learning*, *ML*) – це клас методів штучного інтелекту, характерною рисою яких є не пряме розв'язання задачі, а навчання в процесі застосування рішень багатьох раніше розв'язаних подібних завдань. Для побудови таких методів використовуються засоби математичної статистики, чисельних методів, методів оптимізації, теорії ймовірностей, теорії графів, різні техніки роботи з даними в цифровій формі [2].

У найзагальнішому випадку розрізняють два типи машинного навчання: навчання за прецедентами, або його називають індуктивним навчанням, та дедуктивне навчання [3]. Оскільки дедуктивне навчання прийнято відносити до галузі експертних систем, то терміни «машинне навчання» та «навчання за прецедентами» або індуктивне навчання можна вважати синонімами. Цей метод навчання зараз у тренді, а експертні системи переживають кризу. Бази знань, що лежать в їх основі, важко узгоджувати з реляційною моделлю даних, тому промислові СУБД неможливо ефективно використовувати для наповнення баз знань експертних систем.

Навчання за прецедентами, у свою чергу, поділяють на три основні типи: контрольоване навчання або навчання з учителем (*supervised learning*), неконтрольоване навчання, або навчання без учителя (*unsupervised learning*), і навчання з підкріпленням (*reinforcement learning*).

Крім названих, розробляють й інші методи навчання: активне, багатозадачне, трансферне та ін. Особливо успішно розвивається в останні роки «глибинне навчання», при використанні якого успішно поєднують алгоритми навчання з учителем та без учителя.

Досвід розв'язання окремих завдань навчання за прецедентами переноситься на розв'язання майбутніх завдань навчання. Для формалізації та збереження цього досвіду застосовують реляційні або ієрархічні структури представлення знань.

Серед методів машинного навчання виділяють також методи метанавчання (англ. *metalearning* або *learning-to-learn*), що визначають як один з підходів навчання за прецедентами [4]. Цей підхід дозволяє вибрати з множини алгоритмів найбільш відповідний алгоритм для конкретного завдання (іноді одразу вибираючи також і його параметри). Основна ідея метанавчання – звести задачу вибору алгоритму до задачі навчання з учителем, коли завдання описуються мета-ознаками. Мета-ознака є властивістю завдання, наприклад, число змінних в даних, число можливих міток, розмір набору даних та багато інших параметрів.

Багатоваріантне навчання (англ. *Multiple-instance learning*) – навчання, коли прецеденти (або сукупність вхідних та вихідних змінних) можуть бути об'єднані в групи, в кожній з яких для всіх прецедентів задано поняття «ситуація», але тільки для одного з них (причому, невідомо для якого) є пара «ситуація-розв'язок» [6].

Трансферне навчання (англ. *Transfer learning*) – це проблема машинного навчання, яка фокусується на збереженні знань, отриманих під час вирішення однієї проблеми з подальшим застосуванням її результатів до іншої, але близької до неї проблеми [7]. Наприклад, знання, отримані під час навчання розпізнаванню легкових автомобілів, можуть бути застосовані при спробі розпізнати вантажівки. З практичної точки зору, повторне використання або передача інформації з раніше засвоєних завдань для вивчення нових може значно підвищити ефективність їх розв'язання [8].

Багатозадачне навчання або *multitask learning (MTL)* – це один з підходів машинного навчання побудованого на принципах метанавчання, в якому відбувається одночасне навчання групи взаємозв'язаних завдань, для кожного з яких задається своя пара «ситуація – розв'язок» [5]. З допомогою *MTL* одночасно розв'язують декілька завдань, використовуючи спільні властивості та відмінності між ними, що може підвищити ефективність навчання та точність прогнозування. Розпаралелення операцій можна також віднести до методів *MTL*, але тільки у випадку, коли задачі між собою не перетинаються.

Індуктивний перенос (англ. *Inductive transfer*) – це один зі способів трансферного навчання, відмінність якого полягає в меті цього переносу. Метою індуктивного переносу є підвищити продуктивність моделі розв'язання цільового завдання, а метою трансферного навчання – навчити вхідні та цільове завдання одночасно за рахунок використання інформації про раніше розв'язану схожу задачу навчання [9].

2. Порівняння понять *metalearning*, *multitask learning* та *transfer learning*

В [10] висловлено гіпотезу, що результат трансферного навчання буде залежати від поділу вибірки на навчальну та перевірну, причому розподіл навчальної та перевіркової вибірки є різним на відміну від методів тонкого налаштування, тобто налаштування параметрів моделі (англ. *FineTuning*), де розподіл є однаковим.

Автори [11] стверджують, що трансферне навчання фактично є випадком *MTL*, коли кількість завдань, між якими передаються знання, дорівнює двом. Метою метанавчання є вивчення загальних властивостей (гіперпараметрів або ваг), які дуже пристосовані до нових завдань, спираючись на велику кількість розв'язаних раніше різних завдань. У такий спосіб метанавчання можна розглядати як "історичне" багатозадачне навчання, оскільки воно використовує кілька різних завдань, щоб знайти ідеальний набір властивостей. Останнім часом метанавчання, як правило, зосереджується на пошуку «модельних агностичних» рішень, тобто воно, як і багатозадачне навчання, глибоко пов'язане з архітектурою моделі.

Метанавчання також називають безперервним навчанням або позитивним навчанням (англ. *Lifelong Learning*) [11]. Це парадигма машинного навчання, за якою машина постійно навчається, накопичує знання, отримані в минулому, і використовує/адаптує їх для подальшого навчання та вирішення проблем. Це безперервна здатність до навчання є однією з ознак людського інтелекту.

В [12] порівнюють поняття метанавчання, трансферного навчання та індуктивного переносу.

Метанавчання визначається як частина теорії машинного навчання, в якій застосовують деякі алгоритми метаданих для конкретного випадку з метою вдосконалення процесу машинного навчання. Метадані містять властивості використовуваного алгоритму, а також самого навчального завдання. Використовуючи метадані, можна краще вибрати алгоритм навчання для більш ефективного вирішення проблеми.

Індуктивний перенос, як правило, використовують в тому ж значенні, що і трансферне навчання.

Багатозадачне навчання є формою індуктивної передачі. Воно виражає підхід до вивчення кількох суміжних завдань разом. Таким чином, основне

завдання можна краще засвоїти, скориставшись досвідом, набутим за іншими завданнями. Цей підхід ефективний, коли завдання мають певну спільність.

Метанавчання може використовуватися як взаємозамінний термін "індуктивної передачі". В основному вона спрямована на вдосконалення навчального процесу в часі. Метадані, включаючи досвід минулих періодів навчання, використовують для подальшого навчання, навіть для навчання в різних галузях. Таким чином, індуктивний перенос може розглядатися як підхід до метанавчання.

В [13] розглядають поняття безперервного, трансферного та багатозадачного навчання. Автори відзначають, що описи цих понять у літературі досить заплутані, оскільки вони тісно пов'язані та перетинаються.

В [14] також стверджують, що безперервне навчання має механізм метанавчання. Щоб проілюструвати, варто порівняти навчання людини з традиційним машинним навчанням. Зіткнувшись зі складним завданням, люди не лише набувають нових понять чи навичок, вони також вчаться узагальнювати. Також людина може дуже добре помітити відповідні риси та досить точно відрізнити їх від невідповідних. Іншими словами, вони «метанавчаються».

Іноді до метанавчання помилково відносять побудову алгоритмічних композицій, зокрема, бустінг. Проте коли в композиціях декілька алгоритмів розв'язують одну й ту ж задачу, метанавчання передбачає, що розв'язуються багато різних завдань одночасно.

Розглянемо детальніше задачу багатозадачного навчання.

3. Постановка задачі багатозадачного навчання

Типова постановка звичайного алгоритму *MTL* наведена в [16] у такому вигляді:

$$\min_{w=[w^1 w^2 \dots w^M]} \sum_{m=1}^M L(X^m, y^m, w^m) + \lambda \text{Reg}(w), \quad (1)$$

де $X^m \in \mathbb{R}^{N_m \times D}$ позначає вхідну матрицю m -ї задачі, $y^m \in \mathbb{R}^{N_m \times 1}$ є відповідним вихідним вектором m -го завдання, а $w^m \in \mathbb{R}^{D \times 1}$ - ваговий вектор (або параметри регресії) для m завдань, який ставить у відповідність X_m та y_m , наприклад, $y^m \approx X^m w^m$ (для побудови регресії). Скаляри N_m , D та M позначають відповідно кількість вхідних вибірок для m -ї задачі, кількість властивостей кожної вхідної матриці та кількість задач відповідно. Зауважимо, що на цьому етапі ми припускаємо, що всі вхідні матриці в $\{X^m, m=1, 2, \dots, M\}$ мають однакову розмірність властивостей (але можуть мати різну кількість спостережень для кожної задачі), та властивості всіх ми можемо поєднати, зібравши всі вагові

вектори в $\{w^m\}$ разом, щоб отримати $W = [w^1 w^2 \dots w^M]$, тобто характеристики у кожному рядку W також відповідають задачі.

Виходячи з цих припущень щодо співвідношення між завданнями, розроблено різні обмеження для W , що зазвичай реалізують в регуляризаторі W , позначеному $\text{Reg}(W)$. Крім того, λ - параметр регуляризації, який керує балансом між функцією втрат (перший доданок) та регуляризатором (другий доданок) у рівнянні (1).

Цей критерій дещо подібний до ідеального критерію МГУА [17], який складається зі структурної та шумової складових.

Якщо значення λ буде дорівнювати нулю, то матимемо рішення W , яке не використовує жодних припущень чи попередніх знань про зв'язаність між завданнями, яке, швидше за все, буде добре працювати лише на навчальних даних, але не на контрольних. Якщо ми встановимо значення λ занадто високим, ми можемо мати загальне рішення W , яке задовольняє припущення про відношення між задачами, відображеному в $\text{Reg}(W)$, але не справляється з усіма завданнями прогнозування. Таким чином, параметр регуляризації (та будь-які інші гіперпараметри) зазвичай визначають за допомогою внутрішньої перехресної оцінки за допомогою навчальних зразків.

На рис. 1 наведено різні варіанти алгоритмів *MTL*.

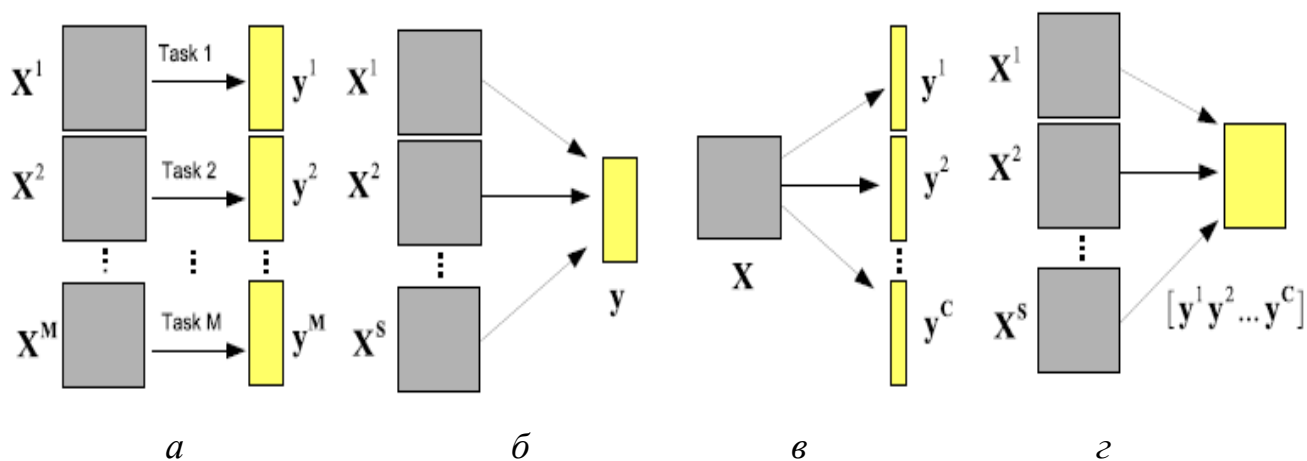


Рис. 1. Варіанти алгоритмів *MTL*

На рис. 1 наведено: (а) – загальна форма *MTL* навчання, коли задана множина входів X_i та множина виходів y_i ; (б) – множина входів та один вихід, де декілька наборів входних даних відповідають одному виходу; (в) – один вхід відповідає багатьом виходам та (г) – багато входів відповідають багатьом виходам.

4. Приклади застосування багатозадачного навчання

Будучи перспективною сферою машинного навчання, *MTL* має на меті покращити ефективність виконання декількох пов'язаних навчальних завдань, максимально використовуючи з них корисну інформацію. В [18] надано огляд задачі та вводяться кілька різних налаштувань *MTL*, включаючи багатозадачне кероване навчання, багатозадачне непідконтрольне навчання, багатозадачне напівконтрольне навчання, багатозадачне активне навчання, багатозадачне навчання з підкріпленням, багатозадачне онлайн-навчання та багатозадачне завдання багатопрофільного навчання. Для кожної настройки представлено репрезентативні моделі *MTL*. Для прискорення процесу навчання вводяться паралельні та розподілені моделі *MTL*. У багатьох сферах, включаючи комп'ютерний зір, біоінформатику, інформатику здоров'я, мовлення, природну обробку мови, веб-додатки, використовується *MTL* для покращення продуктивності застосовуваних додатків, а деякі репрезентативні – переглядаються.

В [19] розглянуто *MTL* як задачу багатокритеріальної оптимізації. В *MTL* кілька завдань розв'язуються спільно, розділяючи індуктивне зміщення між ними. *MTL* за своєю суттю є багатокритеріальною проблемою, тому що різні завдання можуть конфліктувати, і це вимагає компромісу. Поширеним компромісом є оптимізація мети, яка мінімізує зважену лінійну комбінацію втрат для кожного завдання. Однак цей шлях працює тільки тоді, коли завдання не конкурують один з одним, що буває дуже рідко. Автори [19] розглядають *MTL* як багатокритеріальну оптимізацію із загальною метою знайти оптимальний за Парето розв'язок. Для цього вони використовують градієнтні алгоритми з багатокритеріальною оптимізацією. Ці алгоритми погано масштабуються в залежності від розмірності градієнтів та кількості завдань. Тому запропоновано верхню межу для багатоцільових втрат і показано, що вона може бути ефективно оптимізована та застосована до багатьох задач глибиного навчання, включаючи класифікацію цифр, розуміння сцени (спільна семантична сегментація, сегментація зразків та оцінка глибини) і класифікація за кількома мітками. Метод дає більш ефективні моделі, ніж навчання для кожного окремого завдання.

В [20] *MTL* використано для аналізу семантичних послідовностей. Запропоновано нову схему поділу на композиції в декількох завданнях, а саме використано загальну мета-мережу для отримання мета-знань про семантику та згенеровано параметри моделі семантичної композиції. Показано, що загальні мета-знання, засвоєні за запропонованою авторами моделлю, можна розглядати як зовнішнє знання, яке легко переноситься на нові завдання. Автори запропонували мета-мережу, де знання поділяють на кілька суміжних завдань. За допомогою мета-мережі отримано кращу пропозицію для конкретного завдання з використанням отриманих знань інших, пов'язаних з ними завдань. Експериментальні результати показали, що так можна покращити

продуктивність декількох пов'язаних завдань шляхом вивчення спільних рис. Отримані знання за допомогою мета-мережі можуть бути перенесені на інші нові завдання.

5. Індуктивний підхід у задачі багатозадачного навчання

Таким чином, з виконаного огляду можна зробити висновок, що алгоритми *MTL* можна розглядати як один з підрозділів метанавчання.

Індуктивні методи або методи самоорганізації [21-23] можна розглядати як один з підрозділів машинного навчання – навчання за прецедентами.

Враховуючи накопичений досвід розвитку цього напрямку починаючи з кінця 60-х років і донині, він може бути використаний при розробці нових підходів в галузі метанавчання та метамодельовання. Алгоритми індуктивного підходу або самоорганізації можуть бути віднесені до методів навчання без учителя.

На сьогодні в галузі індуктивного моделювання розроблено багато підходів до підвищення точності та швидкодії алгоритмів. Цей досвід може бути використаний для розроблення нових ефективних алгоритмів на принципах метанавчання.

Планується розроблення алгоритмів і засобів на основі індуктивного підходу до метанавчання та метамодельовання [24, 25], для того, щоб усі накопиченні знання в цій галузі узагальнити в нових засобах.

Висновки

Розглянуто задачу багатозадачного навчання як одну з задач метанавчання. Наведено поняття метанавчання, багатозадачного, трансферного, безперервного навчання та індуктивного переносу. Визначено перспективний напрямок розвитку алгоритмів метанавчання на основі індуктивного підходу з урахуванням багатозадачного підходу.

Література

1. https://ai-news.ru/2019/12/glava_google_ai_nazval_glavnye_trendy_mashinnogo_obucheniya_2020.html
2. https://en.wikipedia.org/wiki/Multi-task_learning.
3. Machine Learning - Машинне навчання. <https://www.it.ua/knowledge-base/technology-innovation/machine-learning>.
4. https://ru.wikipedia.org/wiki/Машинное_обучение.
5. Caruana, R., 1997. "Multi-task learning". Machine Learning. 28: 41–75. doi:10.1023/A:1007379606734.
6. <http://www.machinelearning.ru>.
7. https://en.wikipedia.org/wiki/Transfer_learning.

8. Ricardo Vilalta R., Giraud-Carrier C., Brazdil P., Soares C. Inductive Transfer. In book: Encyclopedia of Machine Learning. DOI: 10.1007/978-0-387-30164-8_401.
9. Inductive Transfer. https://link.springer.com/referenceworkentry/10.1007%2F978-0-387-30164-8_401.
10. What is the different between Fine-tuning and Transfer-learning? <https://datascience.stackexchange.com/questions/22302/what-is-the-different-between-fine-tuning-and-transfer-learning>.
11. <https://towardsdatascience.com/icml-2018-advances-in-transfer-multitask-and-semi-supervised-learning-2a15ef7208ec>.
12. In machine learning, what is the difference between the terms transfer learning, multitask learning, inductive transfer, meta-learning, and learning to learn? <https://www.quora.com/In-machine-learning-what-is-the-difference-between-the-terms-transfer-learning-multitask-learning-inductive-transfer-meta-learning-and-learning-to-learn>.
13. Zhiyuan Chen and Bing Liu. Lifelong Machine Learning in the Big Data Era - Tutorial at IJCAI 2015. <https://www.cs.uic.edu/~liub/IJCAI15-tutorial.html>]
14. In machine learning, what's the difference between lifelong learning and meta learning? <https://www.quora.com/In-machine-learning-whats-the-difference-between-lifelong-learning-and-meta-learning>.
15. Кустикова В.Н. Перенос обучения глубоких нейронных сетей http://hpc-education.unn.ru/files/courses/intel-neon-course/Rus/Lectures/Presentations/5_Transfer Learning.pdf
16. Thung, K.-H., Wee, C-Y. A brief review on multi-task learning. *Multimed Tools Appl* 77, 2018, pp. 29705–29725. <https://doi.org/10.1007/s11042-018-6463-x>.
17. Stepashko V.S. Method of Critical Variances as Analytical Tool of Theory of Inductive Modeling. *Journal of Automation and Information Sciences*. 2008. 40(2). P. 4–22.
18. Zhang Yu., Yang Q. An overview of multi-task learning. *National Science Review*, Volume 5, Issue 1, January 2018, pp. 30–43, <https://doi.org/10.1093/nsr/nwx105>.
19. Sener O., Koltun V. Multi-task learning as multi-objective optimization. *NIPS'18 Proceedings of the 32nd International Conference on Neural Information Processing Systems*, Montréal, Canada — December 03 - 08, 2018, pp. 525-536. <https://dl.acm.org/citation.cfm?id=3326992>.
20. Chen J., Qiu X., Liu P., Huang X. Meta Multi-Task Learning for Sequence Modeling. arXiv:1802.08969v1 [cs.AI] 25 Feb 2018.
21. Savchenko Ye., Stepashko V. Metamodeling and Metalearning Approaches in Inductive Modelling Tools. Preprint. <https://easychair.org/publications/preprint/6L1W>.
22. Ivakhnenko A.G. Group Method of Data Handling as a Rival of Stochastic Approximation Method. *Soviet Automatic Control*. 1968. No. 3. P. 43–55.
23. Ivakhnenko A.G. Heuristic Self-Organization in Problems of Automatic Control. *Automatica (IFAC)*. 1970. No. 6. P. 207–219.
24. Savchenko Ye., Stepashko V. Metamodeling as a Way to Universalization of Inductive Modeling Tools. *Proceedings of the 13th IEEE International Conference CSIT-2018 & International Workshop on Inductive Modeling*, September 11-14, 2018, Lviv, Ukraine. Lviv: Publisher “Vezha&Co”, 2018. P. 444-447.
25. Savchenko Ye., Stepashko V. Metamodeling and Metalearning Approaches in Inductive Modelling Tools. Preprint. <https://easychair.org/publications/preprint/6L1W>.