The methods for solving the problems of transfer learning are reviewed from the point of view of solving the metalearning tasks. The concepts of metalearning, transfer learning, multi-task learning, inductive transfer and lifelong learning are given. A scheme for solving the metalearning problem using the inductive approach is proposed. Metalearning is a generalization of all previously solved problems. It gives a possibility to save resources and to use the existing experience in solving learning problems in an optimal way.

**Keywords**: machine learning, transfer learning, modelling, inductive approach, generalization.

**Introduction**

Data mining and machine learning technologies have made great strides in many fields of knowledge development. However, the traditional methods of machine learning learn every task from scratch, while the methods of transfer learning try to transfer the knowledge of some previously solved problems to a new problem in the process of learning.

The concept of transfer learning was introduced in 1993 [1], and it was considered more particularly in 1997 in [2]. The knowledge transfer providing methods aim to make the machine learning process as effective as when a person is taught.

The head of artificial intelligence at Google Jeff Dean recommended his colleagues to pay attention to such algorithms as multi-task and transfer learning, as he considers them to be promising and effective algorithms that give a possibility to achieve the same goal, but with the significantly lower learning costs [3].

Transfer training as one of the tools of machine learning serves to train and improve the productivity of the resulting model by transferring the knowledge from the solved problem to a new one that will need to be solved.

The study of transfer learning is motivated by the fact that people can optimally apply the earlier obtained knowledge in order to quickly solve new problems or to get the better solutions. The need for transfer training arises when the data are outdated and the distribution in them at the cur-
rent moment of time does not correspond to the
distribution in the previous period of time.

Thus, the transfer learning is the optimization
and a quick way to save time or to improve pro-
ductivity. It is not known whether the use of trans-
fer learning in this area will be beneficial until the
model is developed and evaluated on new data.

In practice, machine learning models are mainly
designed to perform only one task. However, when
configuring computers, they use past experiences
not only to repeat the same task in the future, but
also to learn completely new tasks, like humans do.
That is, if a new problem that we are trying to solve
is similar to some of our past experiences, it will be
easier for us to learn. Thus, when using the same
approach in machine learning, transfer learning in-
cludes the methods of past experience transfer on
one or more original problems and uses it to acce-
lerate learning in the current problem.

The aim of this work is to study the principles
of transfer learning in order to learn how to op-
timally use the source data, as well as to com-
pare such concepts as transfer learning, inductive
transfer, metalearning. And as a result, to develop
an approach to generalizing the results of solving
previous problems based on the principles of me-
talearning.

Machine learning concepts
for generalization of learning results

A number of machine learning methods have been
developed by now. They give a possibility to gene-
ralize the previously obtained results of similar
problems’ solving. One of the methods for the gen-
eralizing of the previous outcomes of learning is
called metalearning. The main principle of meta-
learning is learning to learn.

Metalearning is a department of machine
learning [4], where the automatic learning algo-
rithms based on meta-data about all earlier per-
formed computer experiments are used. Today,
this term has not found an unambiguous defini-
tion yet, but the main purpose of its application
is to understand how automatic learning can help to solve learning problems. It will improve
the efficiency of existing learning algorithms or
to train the computer to invoke the learning al-
gorithm itself automatically.

Metalearning (or learning-to-learn) is notable
for the fact that previously solved learning prob-
lems are the precedents [5]. It is required to deter-
mine which of the heuristics used there work more
efficiently. The ultimate goal is to provide continu-
ous automatic improvement of the learning algo-
rithm over time.

Metalearning includes the following learning
methods:

- **Multi-task learning** [6–8]. A set of interrelated
or similar learning problems is solved simultane-
ously using different learning algorithms that have
a similar internal representation. Information on
the similarity of tasks between each other gives a
possibility to improve the learning algorithm and
the quality of the main problem’s solving more ef-
fectively.

One of the particular cases of multi-task learn-
ing when two tasks are considered is transfer
learning [1].

- **Inductive transfer** [9, 10]. The experience of
solving of individual particular learning problems
by precedents is transferred to the solution of sub-
sequent particular learning problems. Relational
or hierarchical structures of knowledge represen-
tation are used in order to formalize and preserve
this experience.

In [11], it is argued that the main idea of meta-
learning is to reduce the problem of choosing an
algorithm to a problem of learning with a teacher,
when the problem is described by the target-at-
tributes. A meta-attribute is a property of a task,
for example, the number of variables in the data,
the number of possible labels, the dataset size,
and many other parameters. This approach gives
a possibility to choose from a variety of algorithms
the most suitable algorithm for a particular task
(sometimes immediately choosing its parameters
as well).

Research in the field of transfer learning is at-
tracting more and more attention under different
names: learning for learning, lifelong learning,
knowledge transfer, inductive transfer, multi-tas-
kling learning, knowledge consolidation, context-
sensitive learning, knowledge-based inductive bias,
metalearning [12 –14]. A closely related technique of transfer learning is a multi-task learning structure that attempts to teach several tasks at the same time, even if they are different [15].

In [11], such concepts as metalearning, multi-task learning, inductive transfer are compared.

Multi-task learning [4, 6] is one of the machine learning approaches based on the principles of metalearning, where a simultaneous learning of a group of interrelated tasks takes place, each of which is given its own pair “situation – solution” [8]. Multi-task learning solves multiple problems at the same time, exploiting common properties and differences between them, which can improve learning efficiency and forecast accuracy. Parallelization of operations can be attributed to multi-task learning methods, but only if the tasks do not intersect with each other [11].

Multi-task learning is a form of inductive transfer. It expresses an approach to the simultaneous study of several interrelated tasks. In this way, the core task can be better learned through experience gained from other tasks. This approach is effective when the tasks have some similarity. This approach can be useful when there is a data scarcity problem.

The difference between multi-task and transfer learning is that in multi-task learning, different models can exchange information in any direction when solving a problem, and in transfer learning, it is possible to transfer information strictly in the direction from the original task to the current one [16].

Inductive transfer is one of the methods of transfer learning, the difference of which lies in the purpose of this transfer. The purpose of inductive transfer is to increase the performance of the model for solving a target task, and the purpose of transfer learning is to teach input and target tasks simultaneously by using information about a previously solved similar learning task [11].

Inductive transfer is usually used in the same sense as learning transfer.

Learning to learn can be used interchangeably with the term inductive transfer. It is mainly aimed at improving the learning process over time. Meta-data, including past learning experiences, is used for future learning, even for learning in different fields. Thus, learning to learn or inductive transfer can be considered as an approach to metalearning.

Let’s consider the principles of transfer learning in more detail.

The Transfer Learning Task
Transfer learning is the ability to combine a pre-trained model with the customer’s learning data. It means that you can use the functionality of the model and add your own data without having to create everything from scratch [17].

For example, if an algorithm has been trained by the thousands of images to create a classification model, then instead of building another model, transfer learning gives you a possibility to combine customer’s image data with a pretrained model to create a new image classifier. This feature makes it quick and easy to have a custom classifier.

In transfer learning, the algorithm is split into two stages. The first stage consists of retraining, when the algorithm is trained on a reference data-set representing a variety of data. Then fine tuning follows, where the algorithm learns to perform a specific target task. The preliminary preparation stage helps to tune the model for general characteristics that can be used again in the target problem, increasing its accuracy [18].

In [19], the problem of using transfer learning for solving the problems of Neuro-linguistic programming was investigated.

Here the transfer learning tasks are roughly classified into three areas:

• based on whether the initial and target settings refer to the same task or not;
• on the nature of the initial and target tasks;
• based on the sequence in which the tasks are studied.

This classification of transfer learning tasks in various directions is shown in Fig. 1.

Figure 1 shows that transfer learning methods are divided into two classes: transductive and inductive transfer learning. They differ in that if they study the same or different learning tasks.
Fig. 1. Classification of methods of transfer learning in various directions [19]

Fig. 2. An overview of different settings of transfer learning [20]
In [20], the tasks of transfer learning are divided into three classes: transductive, inductive, and unsupervised learning (Fig. 2).

Based on the Fig. 2 in [20], transfer learning methods are classified based on the type of traditional ML algorithms used, such as:

- **Inductive Learning Transfer**. In this scenario, the source and target domains are the same, but the source and target tasks are different from each other. The algorithms try to use inductive bias of the source domain in order to help improve the target task. Depending on whether the source domain contains labelled data or not, they can be further divided into two subcategories, similar to multi-tasking learning and self-learning, respectively.

- **Unsupervised transfer learning**. This parameter is similar to the inductive transfer itself, with an emphasis on unsupervised tasks in the target area. The source and target domains resemble, but the tasks are different. In this case, the labelled data is not available in any of the domains.

- **Transductive learning**. In this scenario, there is a similarity between the source and target tasks, but the corresponding domains are different. In this parameter, the source domain contains a lot of labelled data, but the target domain does not.

### When to use transfer learning?

Transfer learning is an optimization, a fast way to save time or to increase productivity.

Authors of [17] and [21] describe three potential advantages to look for when using transfer learning:

These advantages can be described as follows:

- **Higher start**. The best starting model: in other types of learning you need to build a model without any knowledge. Transfer learning offers a better starting point and can perform the tasks at some level even without learning. Transfer learning offers a higher learning rate during learning as the problem is already prepared for a similar task.

- **Higher slope**. Higher accuracy after learning: with a better starting point and higher learning speed, transfer learning provides a machine learning model for convergence with a higher level of productivity, providing a more accurate result.

- **Higher asymptote**. Faster learning: Learning can achieve desired performance faster than traditional learning methods because it uses a pre-trained model.

However, the effectiveness of transfer learning may be not much higher than that of traditional learning models. The impact of transfer learning cannot be determined until a target model is developed.

Ideally, one would reap all three benefits from the successful application of transfer learning. This is a way to try if you can identify a related task with abundant data, and you have the resources to develop a model for this task and reuse it for your own problem, or there is a pre-prepared model that you can use as a starting point for your own model.

The choice of input data or source model is an open issue and may require expert knowledge and/or intuition developed from experience.

### Problem statement of transfer learning

Traditional machine learning methods try to learn every task from scratch, while transfer learning methods try to transfer knowledge from some previous tasks to the target task when the latter has less high-quality learning data [8].

In [1], the following statement of the transfer learning problem is considered.

The definition of transfer learning is given in terms of domain and task [23]. The domain \( D \) consists of: a feature space \( X \) and a marginal probability distribution \( P(X) \), where \( X = \{x_1, x_2, \ldots, x_n\} \in X \). Given a specific domain, \( D = \{X, P(X)\} \), a task
consists of two components: a label space $y$ and an objective predictive function $f(\cdot)$, which is learned from the learning data consisting of pairs, which consist of pairs $\{x_i, y_i\}$, where $x_i \in X$ and $y_i \in Y$. The function $f(\cdot)$ can be used to predict the corresponding label, $f(x)$, of a new instance $x$.

Given a source domain $D_s$ and learning task $T_s$, a target domain $D_T$ and learning task $T_T$, transfer learning aims to help improve the learning of the target predictive function $f_T(\cdot)$ in $D_T$ using the knowledge in $D_s$ and $T_s$, where $D_s \neq D_T$, or $T_s \neq T_T$.

In the process of distance learning, it is necessary to answer the following three important questions [23].

**What to carry?**

It involves the allocation of the part of the task that must be transferred. For this purpose, general knowledge is allocated for different tasks, which are true for many problems at the same time [17]. This is the first and the most important step in the whole process. Here you need to find answers about what part of knowledge can be transferred from the source to the target in order to improve the performance of the target task, i.e. how much of knowledge depends on the source, and what is common between the source and the target.

**How to transfer?**

If there is information about what to transfer, the question is how to transfer? is solved in a natural way. Here it is necessary to begin to determine how to actually transfer knowledge between the tasks. This includes changes in existing algorithms and various methods.

**When to transfer?**

The answer gives a possibility to assess the situations where the application of learning transfer is justified. If the tasks are not related to each other, then the learning transfer may be ineffective (the concept of “negative transfer” is known). We should strive to use transfer learning so as to improve the results of the target task, rather than to worsen them. We must be careful with the situation when it is useful to do so and when not to.

### Applications of transfer learning method

Transfer learning is successfully applied in various real-life tasks.

One example of the application of the learning transfer methods is its use for neuro-linguistic programming [19]. Textual data present all kinds of problems when it comes to deep learning. They are usually transformed or vectorized using various techniques and are used in various tasks such as sentiment analysis and document classification by transferring knowledge from the original tasks.

In [24], the properties of transfer learning for the visualization of medical images are studied. Transferring learning from natural image data sets using standard large models and corresponding pre-trained weights has become a method for applying deep learning to medical visualization.

**Transfer learning for audio/speech.** Similar to areas such as NLP and Computer Vision, deep learning has been successfully used for tasks based on audio data [25]. For example, automatic speech recognition models developed for English have been successfully used to improve speech recognition performance for other languages, such as German. In addition, automatic speaker identification is another example where learning is very helpful.

**Learning Transfer for Computer Vision.** Deep learning has been used quite successfully for various computer vision tasks such as object recognition and identification using various neural network architectures. In [25], the authors present the findings on how the lower layers of the neural network act as conventional extractors of computer vision functions such as edge detectors, while the latter layers work towards task-specific functions [26].

In [16], it is proposed to use transfer learning to build artificial intelligence systems for the diagnosis of diabetic retinopathy and five other diseases based on chest X-ray. For this purpose, artificial neural networks are being built that use the transfer of learning using the approach of behavioral genetics proposed in [27] and demonstrate its effectiveness on financial data.
The Transfer Learning Task as the Means of Metalearning Tasks Solution

**Inductive transfer and inductive bias**

Inductive transfer is one of the names for transfer because it uses the principle of induction, i.e. the experience of solving individual particular problems of learning by precedents is transferred to the solution of subsequent particular problems of learning.

Two inductive approaches to knowledge transfer are Multi-task and Feature Net. In [28], they were used to build predictive neural networks. Unlike traditional modeling, learning with one task, focused only on one target property without any relation to other properties, and in the framework of inductive transfer, individual models are considered as nodes in a network of interconnected models built in parallel (Multi-task Learning) or sequentially (Feature Net). Both of these methods proved to be extremely useful when modeling structure properties on small and structurally diverse datasets where traditional modeling cannot generate any predictive model.

Inductive transfer of learning is divided into two classes [23]:

- If there are marked learning data in the original problem, the inductive transfer of learning corresponds to the task of multitask learning. The difference lies in the transfer goal: for transfer learning the goal is to increase the productivity of the model when solving the target problem and the goal of multi-task learning is to train the source and target tasks at the same time.

- If the labeled learning data are absent in the original problem, inductive transfer of learning corresponds to self-taught learning. Label spaces in the domains of the source and target tasks can be different, which implies the inability to directly use side information of the source task.

Inductive bias is some a priori knowledge that modifies the learning mechanism in the process of solving of the learning problem [29]. In other words, while learning, the solution of the problem shifts toward the opening of certain hypotheses or patterns, and other temporal correlations between events are ignored.

The author [29] argues that the inclusion of a priori knowledge that displaces the learning process is a prerequisite for the success of learning algorithms. Developing tools to express domain knowledge, translating it into a learning algorithm bias, and quantifying the impact of this bias on learning success is a central theme in machine learning theory. The stronger the a priori knowledge or hypotheses with which we begin the learning process, the easier the learning is on further examples, however the less flexible the learning will be because it is limited by the need to comply with these assumptions.

The main problem of inductive bias is how to choose the learning hypothesis space in such a way that it will be large enough to find a solution to the problem under study, but at the same time it will be small enough to provide a reliable generalization from learning sets of reasonable size. As a rule, it is done manually using expert knowledge [30].

Inductive transfer methods use inductive bias of the original problem in order to help the target problem [25]. This can be done in various ways, for example, by adjusting the inductive bias of the target task by limiting the model space, narrowing the space of hypotheses, or making adjustments to the search process itself using knowledge from the original problem. This process is visually depicted in the following Figure 4.

Thus, having conducted a study of machine learning methods, we can conclude that the algorithms of the inductive approach or self-organization can be attributed to the learning methods without a teacher, when no hypotheses about the data on which the learning is performed are known.

**The Solving of the Problem of Metalearning by Inductive Methods**

The goal of metalearning is to study the common properties that are easily adaptable to new challenges. It studies them based on learning a lot of different tasks. In a certain sense, metalearning can be regarded as almost a historical multi-task learning, since it uses many different tasks to find the ideal set of properties [31]. Recently, metalearning...
Ye.A. Savchenko, M.Yu. Savchenko

has focused on a search for model-independent solutions where multi-task learning remains closely related to model architecture.

Metalearning gives a possibility to use the experience of previously solved problems in a specific area of knowledge in order to solve the current problem. Meta-data describing the solved tasks are used for this purpose. Some machine learning algorithms are applied to these meta-data.

Meta-data include the properties of the algorithm used, the learning task itself, and so on. Using meta-data, you can better make a decision

Fig. 4. Inductive transfer illustration [25]

Fig. 5. Scheme of application of inductive methods for solving metalearning problems
about the chosen learning algorithm to solve the current problem more effectively.

After reviewing the tasks of transfer learning, we can come to conclusion that these algorithms can be considered as one of the metalearning units.

Inductive methods or methods of self-organization [32, 33] can be considered as one of the units of machine learning, i.e. case studies. The experience gained in the field of inductive modeling can be used to develop new approaches and methods in the field of metalearning and metamodelling.

Many approaches to improving the accuracy and speed of algorithms’ performance in the field of inductive modeling have been developed by now. Databases, tasks, as well as a database of inductive modeling algorithms are accumulated.

Fig. 5 shows a diagram of the application of inductive methods for solving meta-learning problems.

In the future, it is planned to develop methods and tools based on the inductive approach using meta-data in order to generalize all the accumulated knowledge in this area in new approaches and to use them in order to solve the problem of metalearning and metamodelling.

Conclusions

Methods and tools of machine learning give a possibility to teach a computer to solve problems, like a human. The methods for solving the problems of transfer learning are reviewed from the point of view of solving of the metalearning task. The principles of metalearning allow summarizing a human experience in the knowledge and database to formulate rules for decision-making in the modeling process. Using the experience of solving of many problems gives a possibility to build a system based on the principles of metalearning and metamodelling.

Metalearning is a generalization of all previously solved problems. It allows you to save resources and use the available experience in solving learning tasks optimally. In a certain sense, meta-learning can be regarded as almost a “historical” multi-task learning, since it uses many different tasks to find the ideal set of properties. Recently, metalearning has focused on a search for “model-independent” solutions while multi-task learning remains closely related to model architecture.

A scheme for solving the metalearning problem using the experience of transfer learning is proposed. Some problems may not contain a lot of initial data. That is why transfer learning can give a possibility to develop such a models, what simply could not be developed without using transfer learning.

In the future, it is planned to develop a computer system using the inductive approach and experience in solving of many applied problems in the field of inductive modeling.

REFERENCES


The Transfer Learning Task as the Means of Metalearning Tasks Solution


Received 10.04.2020

ЛІТЕРАТУРА


17. Нежное введение в трансферное обучение для глубокого обучения. [online] Available at: <www.machinelearningmastery.ru> (дата звернення: 20.02.2020).


ЗАДАЧА ТРАНСФЕРНОГО НАВЧАННЯ ЯК ЗАСІБ РОЗВ’ЯЗАННЯ ЗАДАЧІ МЕТАНАВЧАННЯ

Вступ. Метанавчання — це узагальнення всіх раніше розв’язаних задач, яке дозволяє економити ресурси та оптимально використовувати наявний досвід такого розв’язання для нових завдань. Завдання метанавчання — автоматизувати процес навчання, використовуючи методи машинного навчання та індуктивного підходу.

Методи метанавчання засновані на багаторазовому використанні алгоритмів навчання з різними можливими компонентами та безпосередньому порівнянні результатів навчання. У разі розв’язання реальних завдань це вимагає високих обчислювальних потужностей та великих затрат часу. Можлива альтернатива — навчити модель, яка передбачає оптимальні компоненти на основі мета-властивостей вхідних даних.

Мета. У статті поставлено завдання вивчити принципи трансферного навчання, щоб навчитися оптимально використовувати вхідні дані, а також порівняти такі поняття як трансферне навчання, трансферне перенесення, багатозадачне навчання, індуктивне переносу, індуктивне зміщення та метанавчання. В результаті планується розробити підхід до узагальнення результатів розв’язання попередніх завдань на основі принципів метанавчання та метамоделювання.

Методи. Вивчаються методи машинного моделювання: метанавчання, трансферного навчання, індуктивного перенесення та індуктивного підходу.

Результати. Проведено огляд методів розв’язання задачі трансферного навчання з точки зору розв’язання задачі метанавчання. Запропоновано схему розв’язання задачі метанавчання з використанням досвіду трансферного навчання. Метанавчання це узагальнення всіх розв’язаних раніше задач, яке дозволяє економити ресурси та оптимально використовувати наявний досвід розв’язання цих задач.

Висновки. Таким чином, досліджено задачу трансферного навчання як одну з задач метанавчання. Наведено постановку задачі трансферного навчання, а також класифікацію завдань трансферного навчання з точки зору різних напрямків. Порівнюються поняття метанавчання, навчання протягом усього життя, передачі знань, індуктивної передачі, багатозадачного навчання, індуктивного перенесення та індуктивного зміщення. Запропоновано схему застосування індуктивного підходу для розв’язання задач метанавчання та метамоделювання.

Ключові слова: машинне навчання, трансферне навчання, моделювання, індуктивний підхід, узагальнення.
ЗАДАЧА ТРАНСФЕРНОГО ОБУЧЕНИЯ КАК СРЕДСТВО РЕШЕНИЯ ЗАДАЧИ МЕТАОБУЧЕНИЯ

Введение. Метаобучение — это обобщение всех ранее решенных задач, позволяющее экономить ресурсы и оптимально использовать имеющийся опыт решения всех предыдущих задач обучения.

Задача метаобучения — автоматизировать процесс обучения, используя методы машинного обучения и индуктивного моделирования.

Методы метаобучения основаны на многократном использовании алгоритмов обучения с разными возможными компонентами и непосредственном сравнении результатов обучения. В случае решения реальных задач это требует высоких вычислительных мощностей и больших временных затрат. Возможная альтернатива этого — обучить модель, которая предсказывает оптимальные компоненты на основе мета-свойств всех входящих данных.

Цель. В статье поставлена задача изучить принципы трансферного обучения, чтобы научиться оптимально использовать исходные данные, а также сравнить такие понятия как трансферное обучение, трансферный перенос, индуктивный перенос, индуктивное смещение и метаобучение. В результате планируется разработать подход к обобщению результатов решения предыдущих задач на основе принципов метаобучения и метамоделирования.

Методы. Изучаются методы машинного моделирования: метаобучение, трансферное обучение, индуктивный перенос и индуктивное моделирование.

Результаты. Проведен обзор методов решения задач трансферного обучения с точки зрения решения задачи метаобучения. Метаобучение — это обобщение всех решенных ранее задач, которое позволяет экономить ресурсы и оптимально использовать имеющийся опыт решения задач обучения. Предложена схема решения задачи метаобучения с использованием индуктивного подхода.

Выводы. Исследована задача трансферного обучения как одной из задач метаобучения. Приведена постановка задачи трансферного обучения, а также классификация задач трансферного обучения с точки зрения различных направлений. Сравниваются понятия метаобучения, многозадачного обучения, обучения в течение всей жизни, передачи знаний, индуктивной передачи, многозадачного обучения, индуктивного переноса и индуктивного смещения. Предложена схема применения индуктивного подхода для решения задач метаобучения.

Ключевые слова: машинное обучение, трансферное обучение, моделирование, индуктивный подход, обобщение.