Knowledge-Based Systems 80 (2015) 14–23

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Transfer learning using computational intelligence: A survey

Jie Lu *, Vahid Behbood, Peng Hao, Hua Zuo, Shan Xue, Guangquan Zhang

Decision Systems & e-Service Intelligence Lab, Centre for Quantum Computation & Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology Sydney, PO Box 123, Broadway, NSW 2007, Australia

ARTICLE INFO

Article history:
Received 3 December 2014
Received in revised form 7 January 2015
Accepted 17 January 2015
Available online 22 January 2015

Keywords:
Transfer learning
Computational intelligence
Neural network
Bayes
Fuzzy sets and systems
Genetic algorithm

ABSTRACT

Transfer learning aims to provide a framework to utilize previously-acquired knowledge to solve new but similar problems much more quickly and effectively. In contrast to classical machine learning methods, transfer learning methods exploit the knowledge accumulated from data in auxiliary domains to facilitate predictive modeling consisting of different data patterns in the current domain. To improve the performance of existing transfer learning methods and handle the knowledge transfer process in real-world systems, computational intelligence has recently been applied in transfer learning. This paper systematically examines computational intelligence-based transfer learning techniques and clusters related technique developments into four main categories: (a) neural network-based transfer learning; (b) Bayes-based transfer learning; (c) fuzzy transfer learning, and (d) applications of computational intelligence-based transfer learning. By providing state-of-the-art knowledge, this survey will directly support researchers and practice-based professionals to understand the developments in computational intelligence-based transfer learning research and applications.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Although machine learning technologies have attracted a remarkable level of attention from researchers in different computational fields, most of these technologies work under the common assumption that the training data (source domain) and the test data (target domain) have identical feature spaces with underlying distribution. As a result, once the feature space or the feature distribution of the test data changes, the prediction models cannot be used and must be rebuilt and retrained from scratch using newly-collected training data, which is very expensive and sometimes not practically possible. Similarly, since learning-based models need adequate labeled data for training, it is nearly impossible to establish a learning-based model for a target domain which has very few labeled data available for supervised learning. If we can transfer and exploit the knowledge from an existing similar but not identical source domain with plenty of labeled data, however, we can pave the way for construction of the learning-based model for the target domain. In real world scenarios, there are many situations in which very few labeled data are available, and collecting new labeled training data and forming a particular model are practically impossible.

Transfer learning has emerged in the computer science literature as a means of transferring knowledge from a source domain to a target domain. Unlike traditional machine learning and semi-supervised algorithms [1–4], transfer learning considers that the domains of the training data and the test data may be different [5]. Traditional machine learning algorithms make predictions on the future data using mathematical models that are trained on previously collected labeled or unlabeled training data which is the same as future data [6–8]. Transfer learning, in contrast, allows the domains, tasks, and distributions used in training and testing to be different. In the real world, we observe many examples of transfer learning. We may find that learning to recognize apples might help us to recognize pears, or learning to play the electronic organ may facilitate learning the piano. The study of transfer learning has been inspired by the fact that human beings can utilize previously-acquired knowledge to solve new but similar problems much more quickly and effectively. The fundamental motivation for transfer learning in the field of machine learning focuses on the need for lifelong machine learning methods that retain and reuse previously learned knowledge. Research on transfer learning has been undertaken since 1995 under a variety of names: learning to learn; life-long learning; knowledge transfer; meta learning; inductive transfer; knowledge consolidation; context sensitive learning and multi-task learning [9]. In 2005, the

* Corresponding author.
E-mail addresses: jie.lu@uts.edu.au (J. Lu), vahid.behbood@uts.edu.au (V. Behbood), peng.hao@student.uts.edu.au (P. Hao), hua.zuo@student.uts.edu.au (H. Zuo), shan.xue@student.uts.edu.au (S. Xue), guangquan.zhang@uts.edu.au (G. Zhang).

http://dx.doi.org/10.1016/j.knosys.2015.01.010
0950-7051/© 2015 Elsevier B.V. All rights reserved.

Broad Agency Announcement of the Defense Advanced Research Projects Agency's Information Processing Technology Office gave a new mission to transfer learning: the ability of a system to recognize and apply knowledge and skills learned in previous tasks to novel tasks. In this definition, transfer learning aims to extract the knowledge from one or more source tasks and then apply the knowledge to a target task. Traditional machine learning techniques only try to learn each task from scratch, while transfer learning techniques try to transfer the knowledge from other tasks and/or domains to a target task when the latter has few high-quality training data.

Several survey papers on transfer learning have been published in the last few years. For example, the paper by [9] presented an extensive overview of transfer learning and different categories. However, these papers focus on transfer learning techniques and approaches only; none of them discusses how the computational intelligence approach can be used in transfer learning. Since the computational intelligence approach has been applied in transfer learning more recently and has already demonstrated its advantage, this survey is timely.

There are three main types of articles being reviewed in this survey: Type 1 – articles on transfer learning techniques (including related methods and approaches) and Type 2 – articles on transfer learning using computational intelligence techniques. Type 3 – articles on related computational intelligence techniques. The search and selection of these articles were performed according to the following five steps:

Step 1. Publication database identification and determination: The eminent publication databases such as Science Direct, ACM Digital Library, IEEE Xplore and SpringerLink, were searched to provide a comprehensive bibliography of research papers on transfer learning and transfer learning using computational intelligence.
Step 2. Type 1 article selection: These papers were selected according to the two criteria: (1) novelty; and (2) impact- published in high quality (high impact factor) journals, or in conference proceedings or book chapters but with high citations[1]. These types of article are mainly used in Section 2.
Step 3. Preliminary screening of Type 2 articles: The search was first performed based on related keywords of computational intelligence in transfer learning.
Step 4. Result filtering for Type 2 articles: The keywords of the preliminary references were extracted and clustered manually. Based on the keywords related to application domain, these papers were divided, using "topic clustering", into four groups: (a) Neural Network in transfer learning; (b) Bayes in transfer learning; (c) fuzzy and genetic algorithm in transfer learning and (d) application of transfer learning. This article selection process was based on the following criteria: (1) novelty – published within the last few years; (2) impact – see Step 2; (3) coverage – reported a new or particular application domain; and (4) typicality – only the most typical methodology and applications were retained.
Step 5. Type 3 article selection: These papers were selected according to the requirement of Step 4, aiming to introduce related concepts of computational intelligence techniques.

The main contributions of this paper are: (1) it comprehensively and perceptively summarizes research achievements on transfer learning from the point of view of applications of computational intelligence, and strategically clusters the transfer learning into four computational intelligence application domains; (2) for each computational intelligence technique it carefully analyses typical transfer learning frameworks and effectively identifies the specific requirements of computational intelligence techniques in transfer learning. This will directly support researchers and practitioners to promote the popularization and application of computational intelligence in transfer learning in different domains; and (3) it also covers several very new transfer learning techniques with computational intelligence, and reveals their successful applications.

The remainder of this paper is structured as follows. In Section 2, the transfer learning techniques are reviewed and analyzed. Sections 3–5 respectively present the 4 main application domains of transfer learning. Section 6 discusses the applications of computational intelligence-based transfer learning methods. Section 7 presents our analysis and main findings.

2. Basic transfer learning techniques

To understand and analyze the application developments of transfer learning by using computational intelligence, this section first reviews the main transfer learning techniques. The notations and definitions that will be used throughout the section are introduced. According to the definitions, we then categorize the various settings of transfer learning methods that exist in the literature of machine learning.

Definition 2.1 (*Domain [9]*). A domain, which is denoted by $D = \{\chi, P(X)\}$, consists of two components:

(1) Feature space χ; and
(2) Marginal probability distribution $P(X)$, where $X = \{x_1, \ldots, x_n\} \in \chi$.

Definition 2.2 (*Task [9]*). A task, which is denoted by $T = \{Y, f(\cdot)\}$, consists of two components:

(1) A label space $Y = \{y_1, \ldots, y_m\}$; and
(2) An objective predictive function $f(\cdot)$ which is not observed and is to be learned by pairs $\{x_i, y_i\}$.

The function $f(\cdot)$ can be used to predict the corresponding label, $f(x_i)$, of a new instance x_i. From a probabilistic viewpoint, $f(x_i)$ can be written as $P(y_i|x_i)$. In the bank failure prediction example, which is a binary prediction task, y_i can be the label of failed or survived. More specifically, the source domain can be denoted as $D_s = \{(x_{s_1}, y_{s_1}), \ldots, (x_{s_n}, y_{s_n})\}$ where $x_{s_i} \in \chi_s$ is the source instance or bank in the bank failure prediction example and $y_{s_i} \in Y_s$ is the corresponding class label which can be failed or survived for bank failure prediction. Similarly, the target domain can be denoted as $D_t = \{(x_{t_1}, y_{t_1}), \ldots, (x_{t_n}, y_{t_n})\}$ where $x_t \in \chi_t$ is the target instance and $y_{t_i} \in Y_t$ is the corresponding class label and in most scenarios $t_n \ll s_n$.

Definition 2.3 (*Transfer learning [9]*). 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 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 above definition, the condition $D_s \neq D_t$ implies that either $\chi_s \neq \chi_t$ or $P_s(X) \neq P_t(X)$. Similarly, the condition $T_s \neq T_t$ implies that either $Y_s \neq Y_t$ or $f_s(\cdot) \neq f_t(\cdot)$. In addition, there are some explicit or implicit relationships between the feature spaces of two domains such that we imply that the source domain and target domain are related. It should be mentioned that when the target and

[1] "high citation" means that the citation of the paper is greater than the average citation rates listed in the "ISI Web of Knowledge – Essential Science Indicators", and the citation per year of the paper is larger than 1.

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات