Information Sciences 293 (2015) 59–79

Contents lists available at ScienceDirect

Information Sciences

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

Fuzzy Transfer Learning: Methodology and application

Jethro Shell *, Simon Coupland [1]

The Centre for Computational Intelligence, De Montfort University, The Gateway, Leicester LE1 9BH, United Kingdom

A R T I C L E I N F O

Article history:
Received 13 November 2013
Received in revised form 27 August 2014
Accepted 7 September 2014
Available online 18 September 2014

MSC classification:
94D05
68T05

Keywords:
Fuzzy logic
Transfer learning
Context
Online adaptation
Sensor network
Intelligent environment

A B S T R A C T

Producing a methodology that is able to predict output using a model is a well studied area in Computational Intelligence (CI). However, a number of real-world applications require a model but have little or no data available of the specific environment. Predominantly, standard machine learning approaches focus on a need for training data for such models to come from the same domain as the target task. Such restrictions can severely reduce the data acquisition making it extremely costly, or in certain situations, impossible. This impedes the ability of these approaches to model such environments. It is on this particular problem that this paper is focussed.

In this paper two concepts, Transfer Learning (TL) and Fuzzy Logic (FL) are combined in a framework, Fuzzy Transfer Learning (FuzzyTL), to address the problem of learning tasks that have no prior direct contextual knowledge. Through the use of a FL based learning method, uncertainty that is evident in dynamic environments is represented. By applying a TL approach through the combining of labelled data from a contextually related source task, and little or no unlabelled data from a target task, the framework is shown to be able to accomplish predictive tasks using models learned from contextually different data.

1. Introduction

The availability of information impacts on the understanding of a problem. Further to this, a lack of information reduces the ability to understand a problem. Differences in information and understanding about one problem domain and a second similar domain can be defined as being contained within a *knowledge gap*. This paper presents a novel composition of methods that use a combination of inductive and deductive learning mechanisms that draw inspiration from human approaches to problem solving, to bridge the knowledge gap. Two concepts, TL, a methodology that allows information gained in different contextual situations to assist new learning tasks,[2] and FL, an approach to capture imprecision and uncertainty, are brought together in a novel framework to address the problem of learning tasks that have no prior direct contextual knowledge.

Real-world applications often consist of many unknowns increasing the knowledge gap. To predict or classify based on the information gathered from these applications can be difficult. Standard machine learning approaches require that there is a form of training data. Predominantly such training data has to come from the same domain. Some applications make the procurement of *a priori* labelled training data demanding, or in some cases, not possible at all. For example, to measure

* Corresponding author. Tel.: +44 (0)116 255 1551.
E-mail addresses: jethros@dmu.ac.uk (J. Shell), simonc@dmu.ac.uk (S. Coupland).
[1] Tel.: +44 (0)116 207 8419.

[2] In this research the term *task* is referred to as any action the learning method is required to accomplish.

http://dx.doi.org/10.1016/j.ins.2014.09.004

certain physical areas such as remote forest locations, impromptu set ups such as disaster zones, or small user groups that have very defined requirements such as disabled users.

The procurement of training data produces an interesting problem. If there is a requirement to classify or predict the output from such environments, *how can a model be produced?* The examples given previously present situations where labelled data from the same distribution is costly. The lack of information within the same problem domain can cause standard supervised to be ineffective. The process of labelling data can also be costly. For example, there may be few images that are labelled from a given feature space [9]. Recognising that supervised methods require that training data is supplied from the same domain, Semi-Supervised Learning (SSL) was introduced to approach this issue [25]. SSL sits in between supervised and unsupervised learning. Unlike supervised learning, where the goal is to learn a mapping from x to y, given a training set of pairs (x_i, y_i), SSL is supplied with unlabelled data. $y_i \in Y$ is referred to as the labels of x_i. Typically the focus of unsupervised learning is to find structure in the unlabelled data, $X = (x_1, \ldots, x_n)$ where $x_i \in X$ for all of $i \in n$ [7]. Additionally, large quantities of unlabelled data may not be available. The initial reduced quantity of unlabelled domain data within the implementations we are focussing upon also renders the use of SSL and unsupervised less effective.

The contexts discussed, however, can be related to other implementations which may contain previously discovered knowledge. The transferral of knowledge from one context to another is in keeping with the concept of a more humanist style of learning, to reuse and repurpose information. Within the study of human learning, ordinary learning is viewed as being ordinary when it is within the same context (a student may solve similar problems that are at the end of a chapter that have appeared previously), whereas TL occurs outside of a single context (problems are solved when they occur mixed with others at the end of the course, for instance) [41]. Studies have shown that humans often draw upon more than just training data for generalisation [52]. What can make humans successful is not learning from scratch, but adapting to the changing environment [40]. In recent years there has been significant quantities of research in the area of TL and its application to real-world problems in the area of CI [59,16,21,2,24,4]. TL can be broadly defined as a learning technique that uses knowledge from a source domain to increase the performance of learning within the target task domain. The methodology allows the domains, tasks and distributions used within the training and testing to be different. The research within this paper presents a novel use of a TL method to model scenarios where little or no information is initially known.

There is a strong relationship between context and uncertainty. As individuals endeavour to learn a new task they often afford uncertainty to it. There is a clear codependency on the level of certainty in any learning activity and the amount of information that is available. Problems with little information can have a high degree of uncertainty [31]. Dynamic applications such as Intelligent Environments (IEs) can exhibit this uncertainty in the sensors that are used and the decision structures that are applied. The incorporation of a FL system is proposed to assist in the modelling of environments in the presence of uncertainty and vagueness. The use of FL allows for the incorporation of approximation and a greater expressiveness of the uncertainty within the data [62].

This paper introduces a novel framework for the learning of target tasks from limited unlabelled target data and related, differing source labelled data using a Fuzzy Inference System (FIS). The primary goal of the framework is to model environments where no or little unlabelled data is available by using contextually different but related information. The framework uses an adaptive online learning methodology to enhance the transfer of a FIS between the contextually differing learning tasks. Within this paper the adaptive, predictive framework was demonstrated on real-world data from intelligent environments. In comparison to current state of the art methods, the FuzzyTL framework was shown to be able to outperform two mature regression algorithms, K-Nearest Neighbour (KNN) and Support Vector Machine (SVM). When compared to the Support Vector Regression (SVR) approach, the FuzzyTL framework produced a lower Root Mean Squared Error (RMSE) in 67.8996% contexts. When compared to a KNN implementation, the FuzzyTL framework produced a lower RMSE in 54.5493% of the tested contexts.

The structure of the paper is as follows. The next section, Section 2 gives an overview of the current research in this area. Section 3 describes the methodology used within the FuzzyTL framework. The experiments and results carried out to analyse the proposed methodology are discussed in Sections 4–6. The final two sections conclude the paper and outline future work.

2. Background

2.1. Context

The defining of context is broad and far reaching. The study of context is a multi-disciplinary pursuit, ranging from psychology and linguistics, to computing (especially within Computational Intelligence).

The concept of context plays an important role in both FL and TL. There is, however, no single consensus of how context should be defined. The structure of the FuzzyTL framework has foundations in the notion of context. TL has the ability to use information from one domain to close the information gap in a learning process from a differing but similar domain. The domains can be defined as contexts. To analyse the contexts, a valid definition of a context must be put forward.

There has been considerable work surrounding context and computing. Dey [11] puts forward a definition of context:

> "Context is any information that can be used to characterise the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.".