

Accepted Manuscript

Conformity-based Source Subset Selection for Instance Transfer

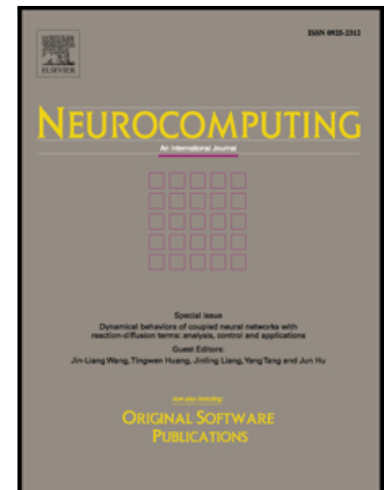
Shuang Zhou, Evgueni Smirnov, Gijs Schoenmakers, Ralf Peeters

PII: S0925-2312(17)30420-4
DOI: [10.1016/j.neucom.2016.11.071](https://doi.org/10.1016/j.neucom.2016.11.071)
Reference: NEUCOM 18170

To appear in: *Neurocomputing*

Received date: 31 May 2016
Revised date: 10 October 2016
Accepted date: 5 November 2016

Please cite this article as: Shuang Zhou, Evgueni Smirnov, Gijs Schoenmakers, Ralf Peeters, Conformity-based Source Subset Selection for Instance Transfer, *Neurocomputing* (2017), doi: [10.1016/j.neucom.2016.11.071](https://doi.org/10.1016/j.neucom.2016.11.071)



This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conformity-based Source Subset Selection for Instance Transfer

Shuang Zhou^{a,*}, Evgueni Smirnov^a, Gijs Schoenmakers^a, Ralf Peeters^a

^a Department of Data Science and Knowledge Engineering,
Maastricht University,
P.O. Box 616, 6200 MD, Maastricht, The Netherlands

Abstract

Instance transfer aims at improving prediction models for a target domain by transferring data from related source domains. The effectiveness of instance transfer depends on the relevance of source data to the target domain. When the relevance of source data is limited, the only option is to select a subset of source data of which the relevance is acceptable. In this paper, we introduce three algorithms that perform source-subset selection prior to model training. The algorithms employ a conformity-based test that estimates the source-subset relevance based on individual instances or on subsets as a whole. Experiments conducted on four real-world data sets demonstrated the effectiveness of the proposed algorithms. Especially, it was shown that pre-training subset-selection based on set relevance is capable of outperforming the existing instance-transfer techniques.

Keywords: instance-transfer learning, source-subset selection, conformal test

1. Introduction

Instance-transfer learning has gained increasing attention over the last decade [1]. It aims at improving prediction models for a *target* domain by exploiting data from (closely) related *source* domains. This implies that the effectiveness of instance-transfer

*Corresponding author

Email addresses: shuang.zhou@maastrichtuniversity.nl (Shuang Zhou),
smirnov@maastrichtuniversity.nl (Evgueni Smirnov),
gm.schoenmakers@maastrichtuniversity.nl (Gijs Schoenmakers),
ralf.peeters@maastrichtuniversity.nl (Ralf Peeters)

Download English Version:

<https://daneshyari.com/en/article/4947225>

Download Persian Version:

<https://daneshyari.com/article/4947225>

[Daneshyari.com](https://daneshyari.com)