

io-port 05486049

Karunaratne, Thashmee; Boström, Henrik

The effect of background knowledge in graph-based learning in the chemoinformatics domain.

Castillo, Oscar (ed.) et al., Trends in intelligent systems and computer engineering (IMECS 2007). Selected papers based on the presentations at the international conference, Hong Kong, China, March 21–23, 2007. New York, NY: Springer (ISBN 978-0-387-74934-1/hbk). Lecture Notes in Electrical Engineering 6, 141-153 (2008).

Summary: Typical machine learning systems often use a set of previous experiences (examples) to learn concepts, patterns, or relations hidden within the data [1]. Current machine learning approaches are challenged by the growing size of the data repositories and the growing complexity of those data [1, 2]. In order to accommodate the requirement of being able to learn from complex data, several methods have been introduced in the field of machine learning [2]. Based on the way the input and resulting hypotheses are represented, two main categories of such methods exist, namely, logic-based and graph-based methods [3]. The demarcation line between logic- and graph-based methods lies in the differences of their data representation methods, hypothesis formation, and testing as well as the form of the output produced. The main purpose of our study is to investigate the effect of incorporating background knowledge into graph learning methods. The ability of graph learning methods to obtain accurate theories with a minimum of background knowledge is of course a desirable property, but not being able to effectively utilize additional knowledge that is available and has been proven important is clearly a disadvantage. Therefore we examine how far additional, already available, background knowledge can be effectively used for increasing the performance of a graph learner. Another contribution of our study is that it establishes a neutral ground to compare classification accuracies of the two closely related approaches, making it possible to study whether graph learning methods actually would outperform ILP methods if the same background knowledge were utilized [9].

doi:10.1007/978-0-387-74935-8_10