Searching for Credible Relations in Machine Learning

Vedrana Vidulin
Department of Intelligent Systems, Jožef Stefan Institute, Jamova cesta 39, Ljubljana, Slovenia
E-mail: vedrana.vidulin@ijs.si; Web: http://dis.ijs.si/Vedrana

Thesis Summary

This paper presents a summary of the doctoral dissertation of the author on the topic of searching for credible relations in machine learning.
Povzetek: Članek predstavlja povzetek doktorske disertacije avtorja, ki obravnava temo iskanja verodostojnih relacij v strojnem učenju.

1   Introduction

When machine learning (ML) and data mining (DM) methods construct models in complex domains, models can contain less-credible parts [2], which are statistically significant, but meaningless to the human analyst. For example, let us consider a decision tree model presented in Figure 1. The tree is constructed with the J48 algorithm in Weka [8] for a complex domain indicating which segments of research and development (R&D) sector have the highest impact on economic welfare of a country. Nodes in the tree represent segments of the R&D sector. Leaves in the tree represent economic welfare of the majority of countries that reached the specific leaf. Economic welfare can be: low, middle or high. In each leaf, the first number in brackets represents the number of countries that reached that leaf. The second number represents the number of countries in that leaf with the level of welfare different than the one represented by the leaf. The quantities are expressed in decimals to account for those countries with missing values for segments appearing in the tree. Note that the left subtree is omitted to simplify the example.
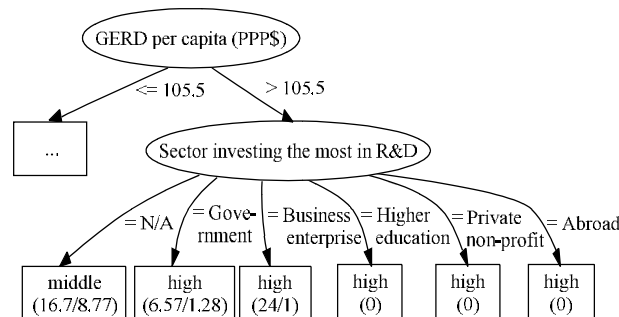

Figure 1: Decision tree constructed with J48 algorithm.

The tree represents two relations, one defining how the level of investment in R&D ("GERD per capita (PPP$)") is related to economic welfare and another defining how sector that invests the most in R&D is related to the welfare. The relation including sector is an example of a less-credible relation. It is statistically significant because the ML method included it in the tree. However, it is meaningless, since the single "middle" leaf represents the countries for which the sector is unknown ("N/A" value), while all other countries have high level of welfare. In other words, the subtree does not bring any additional information to the human analyst.

To eliminate less-credible relations from the models, both automatic and interactive approaches were suggested. Examples of the former are the pruning of decision trees [4] and the correction of a quality estimate to eliminate the random classification rules with optimistically high values of quality [3]. Typical examples of the latter provide improvements in the form of new training examples [1] or a list of attributes that would better describe the class [5]. The presented approaches aim to improve the model's predictive performance, while meaningless relations can remain a part of the model, as long as they positively influence the quality. The resulting models are not acceptable when the task is domain analysis.

Therefore, we proposed a novel method that constructs multiple models in an algorithmic way, enabling the human analyst to examine interesting relations from different angles and in different contexts, and based on additional evidence to conclude which relations are indeed credible.

2   Human-machine data mining

The interactive method for the construction of credible relations in complex domains, proposed in the thesis [6], is named Human-Machine Data Mining (HMDM). The basic idea of our method is to construct a large number of models to extract the credible relations, i.e., relations that are both meaningful and of high quality. The task is computationally very demanding, and for other than simple cases there is no possibility for humans to analyse

a meaningful share of all the hypothesized models on their own. However, the introduced combination of human understanding and raw computer power enables a smart examination of the parts of the huge search space with most credible relations. While ML and DM methods perform the search, humans examine and evaluate the results, make conclusions and redo the search in a way that seems to be the most promising based on the previous attempts. In this way, the humans guide the DM to search the subspaces with the most credible relations and finally the humans construct the overall conclusions from the various, most interesting solutions.

The HMDM defines a toolbox composed of semi-automated DM procedures and a set of scenarios for the human to guide the analysis towards credible relations. Furthermore, it defines a scheme for the extraction of credible relations from multiple models, which provides support to the human analyst in the process of constructing correct conclusions about the domain.

3   Evaluation

The proposed method was demonstrated in two complex domains that show how the higher education and the R&D sectors are related to economic welfare [7]. With the HMDM method we were able to extract relations well-established in the literature, which shows the capability of the method to find important relations in the domain. We also extracted new relations that were not previously addressed in the literature. In addition, we showed in a domain of automatic web genre identification that HMDM can be successfully used for learning predictive models in another domain.

A user study justified the HMDM method by showing that the users are frequently not able to detect meaningless relations by observing a single model constructed by a ML algorithm. However, by observing interesting variations, i.e., candidate solutions suggested by the HMDM method, the participants realized the weaknesses of the default model and created better domain models.

4   Conclusions

The thesis made several contributions to the area of ML and application areas of macroeconomics and automatic web genre identification. First, we proposed a new method Human-Machine Data Mining for extracting credible relations from data, based on interactive and iterative processes exploiting advantages of both human and artificial intelligence. Second, the quality measure corrected class probability estimate was adjusted to decision tree models. Third, interactive explanations of DM results were designed, conceived to facilitate the extraction of credible relations. Fourth, a computer program was developed to support the HMDM method. Fifth, for two real-life domains, showing how the higher education and R&D sector influence the economic welfare of a country, we extracted credible relations with the new method, confirming some well-known relations and providing some new ones. Finally, for the real-life

domain of automatic web genre identification, we constructed credible models with the new method, which provide an insight into the role of content words in recognizing web genres.

References

[1] J.A. Fails, D.R. Olsen Jr. (2003) Interactive Machine Learning, In: *Proceedings of the 8th International Conference on Intelligent User Interfaces*, Miami, FL, pp. 39-45.

[2] D. Jensen, P. Cohen (2000) Multiple Comparisons in Induction Algorithms. *Machine Learning* 38(3):309-338.

[3] M. Možina, J. Demšar, J. Žabkar, I. Bratko (2006), Why is Rule Learning Optimistic and How to Correct It, In: *Machine Learning: ECML 2006*, Springer, Berlin, pp. 330-340.

[4] J.R. Quinlan (1993) C4.5: Programs for Machine Learning, Morgan Kaufmann, San Mateo, CA.

[5] S. Stumpf, V. Rajaram, L. Li, W. Wong, M. Burnett, T. Dietterich, E. Sullivan, J. Herlocker (2009) Interacting Meaningfully with Machine Learning Systems: Three Experiments, *International Journal of Human–Computer Studies* 67:639–662.

[6] V. Vidulin (2012) *Searching for Credible Relations in Machine Learning*, PhD Thesis, IPS Jožef Stefan, Ljubljana, Slovenia.

[7] V. Vidulin, M. Gams (2011) Impact of High-Level Knowledge on Economic Welfare through Interactive Data Mining. *Applied Artificial Intelligence* 25(4):267-291.

[8] I. Witten, E. Frank (2005) *Data Mining. Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, San Francisco, CA.