ABSTRACT

Classification is one of the main tasks in machine learning, data mining, and pattern recognition. Compared with the extensively studied automation approaches, the interactive approaches, centered on human users, are less explored. This chapter studies interactive classification at 3 levels. At the philosophical level, the motivations and a process-based framework of interactive classification are proposed. At the technical level, a granular computing model is suggested for re-examining not only existing classification problems, but also interactive classification problems. At the application level, an interactive classification system (ICS), using a granule network as the search space, is introduced. ICS allows multi-strategies for granule tree construction, and enhances the understanding and interpretation of the classification process. Interactive classification is complementary to the existing classification methods.

INTRODUCTION

Human cognitive activities rely on classification to organize the vast number of known matter, plants, animals and events into categories that can be named, remembered, and discussed. The problems of human-based classification are that, for very large and separate datasets, it is difficult for people to be aware, to extract, to memorize, to search and to retrieve classification patterns, in addition to interpreting and evaluating classification results that are constantly changing, and then making recommendations or predictions in the face of inconsistent and incomplete data.

Computers perform classification by revealing the internal structures of data according to programmed algorithms. They maintain precise operations under a heavy information load and preserve steady performance. A typical automatic classification approach is batch processing, where all the input is prepared before the program runs. The problems of automatic classification are that the systems often do not allow users, or limit users' ability,
Interactive Classification Using a Granule Network

to contact and participate in the discovery process. A fixed algorithm may not satisfy the diverse requirements of users; a user often cannot relate to the answers, and is left wondering about the meaning and value of the so-called discovered knowledge.

In this chapter, we propose a framework of human-machine interactive classification. Although human-machine interaction has been emphasized for many disciplines, such as information retrieval and pattern recognition, it has received some, though not yet enough, attention in the domain of data mining (Ankerst et al., 1999, Brachmann & Anand, 1996, Han, Hu & Cercone, 2003, Zhao & Yao, 2006). The fundamental idea of interactive classification is: on one hand, computers can help users to carry out description, prediction and explanation activities efficiently; on the other hand, human insights, judgements and preferences can effectively interfere with method selection, application and adjustment, thus improving the existing methods and generating new methods. Interactive classification uses the advantages of both a computer system and a human user. A foundation of human-computer interaction in data mining may be provided by cognitive informatics (Wang, 2002, 2007a, 2007b; Wang & Kinsner, 2006; Wang et al., 2006). As Wang suggests that, for cognitive informatics, relations and connections of neurons represent information and knowledge in the human brain might be more important than the neurons. Following the same way of thinking, we believe that interactive data mining is sensitive to capacities and needs of humans and machines. A critical issue is not how intelligent the user is, or how efficient the algorithm is, but how well these two parts can be connected and communicated, stimulated and improved.

More specifically, interactive classification systems allow users to suggest preferred classifiers and knowledge structures, and use machines to assist calculation and analysis during the discovery process. A user can freely explore the dataset according to his/her preference and priority, ensuring that each classification stage and the corresponding result are all understandable and comprehensible. The constructed classifier is not necessarily efficient when compared with most of the automatic classifiers. However, it is close to human thinking by its very nature. The evaluation of an interactive classification, involving the understandability and applicability of the final classification results, relies heavily on the interaction between the computer and the human user, not just on one single factor.

In the rest of this chapter, we discuss the interactive classification at three levels: philosophical level, technical level and application level. At the philosophical level (Section 2), we discuss the motivation of interactive classification and present the process-based framework. At the technical level (Section 3), we apply granular computing as the methodology for examining the search space and complexity issues of interactive classification. At the application level (Section 4), an interactive classification implementation based on a granule network is introduced. The main results demonstrate the usefulness of the proposed approach. The conclusion is in Section 5.

STUDY OF INTERACTIVE CLASSIFICATION AT THE PHILOSOPHICAL LEVEL

For an interactive system, the most critical aspects are not only how intelligent the user is, or how efficient the system is, but also how well these two parts connect and communicate. Through interaction and communication, computers and users can divide labour in order to achieve a good balance of automation and human control.

Motivations for Interactive Classification

Human-computer interaction should cope with users’ either vague or clear requirements, and fixed or changing goals, as well as encouraging the curiosity of exploration, and supporting multiple views and operations for different inquiries. The interaction allows users to engage in all important processes of classification. Typically, a user can design and conduct a specific classifier, think critically and logically in his/her own way, and compare the different sets of classification rules based on his/her own understanding. These interactive processes encourage human learning, improve insight and understanding of the domain, and stimulate users to explore creative possibilities. User feedback can be used to improve the system. The interaction is bi-beneficial.

It is important to note that users process a variety of skills, intelligence, cognitive styles, frustration tolerances and other mental abilities. They come to a classification problem with varying preferences, requirements and background knowledge. Given a set of data, each user will try to make sense of data by seeing it from different
Related Content

Towards Cognitive Machines: Multiscale Measures and Analysis
www.igi-global.com/article/towards-cognitive-machines/1527?camid=4v1a

Communication and Automatic Interpretation of Affect from Facial Expressions
www.igi-global.com/chapter/communication-automatic-interpretation-affect-facial/49534?camid=4v1a

The Theoretical Framework of Cognitive Informatics
www.igi-global.com/article/theoretical-framework-cognitive-informatics/1526?camid=4v1a

Cognitively Informed Systems: Justifications and Foundations
www.igi-global.com/chapter/cognitively-informed-systems/6620?camid=4v1a