

Editorial

Special issue on soft computing for dynamic data mining

An ancient Greek saying “Change is the only constant” emerges as a universal wisdom that cuts across almost all facets of our life. Likewise, it becomes equally relevant in science and engineering when we commonly encounter phenomena of inherent dynamic behavior. This is also the case in Data Mining which over the last two decades has enjoyed a rapid development of its applications along with a plethora of advanced architectures and techniques.

There are numerous successful applications that have spurred more interest in the area and fostered the need for further research on methodological issues, algorithms and specialized development platforms. What still remains an open issue, however, is the question how to update/revise/optimize the respective systems when new data becomes available. Here, dynamic data mining (DDM) comes into play by offering a coherent suite of technologies for “updating” and enhancing functionality of the existing data mining systems.

Such *dynamic* aspects of systems have become highly relevant to all algorithmic categories of data mining such as, e.g., clustering, classification, association building, deviation detection, and visualization, to name the most representative examples.

Soft computing, neural networks, fuzzy logic, and evolutionary computing, in particular forming a coherent conceptual and algorithmic platform of design and analysis of intelligent systems, offer evident advantages by supporting understanding, modeling, and experimenting in the presence of dynamic aspects of data mining. For instance, neural networks offer superb learning abilities and in this sense become of evident interest when dealing with various facets of pattern updating. Some other aspects of dynamic data mining are related to concept drift and in this case fuzzy logic and fuzzy sets offer an interesting insight into the vagueness of the underlying phenomenon. Genetic optimization helps address various aspects of structural optimization. Evolutionary computing with its inherent adaptation capabilities brings powerful optimization mechanisms for dynamic data mining.

This special issue documents the state-of-the-art picture of this emerging area and serves as a testimony to a genuine wealth of concepts, methodologies, algorithms, applications, and case studies we encounter in dynamic data mining.

The following summaries of the contributions to this special issue highlight the diversity of techniques and application areas where dynamic data mining comes into play.

The paper by Last et al. presents a series of incremental learning algorithms that are shown to produce more accurate classification models in the presence of a concept drift than the batch algorithms. Furthermore they are computationally cheaper than the incre-

mental methods existing in the literature. These algorithms are based on an advanced methodology of developing decision-tree called “info-fuzzy network” (IFN) developed by the same authors. Using a sliding window of examples the existing models are updated unless a major concept drift is detected which triggers a construction of new models. The applications to real-world streams of traffic and intrusion detection data highlight the potential of the presented approach.

Lee presents a fuzzy rule-based system for adaptive scheduling in a dynamic manufacturing environment. The proposed method aims at the development of a robust scheduling strategy rather than finding a good solution to some isolated problems. The presented approach thus combines the human ability to gather expertise in a manufacturing environment and quantify it via vague (linguistic) terms with the power of nowadays soft- and hardware to analyze huge data sets. An application to a hypothetical Flexible Manufacturing System (FMS) shows the strength of the proposed approach to generate robust and interpretable rules that outperform some existing methods as well as neural networks.

Klawonn and Georgieva suggest the use of modified clustering algorithms to discover pattern in streaming data, such as, e.g., weather records. The same authors had developed previously a prototype-based clustering algorithm called Dynamic Data Assigning Assessment (DDAA) being able to identify single clusters in a step-by-step manner. Hence, a subsequent development was to apply the underlying idea to continuously arriving data as is the case in streaming data. The proposed clustering method assigns new input data to the already existing clusters or discovers new interesting groups that currently had been created by the underlying phenomenon. An application to weather data shows the usefulness of the proposed approaches.

Awad and Motai analyze streaming data however in their case they show the applicability of Support Vector Machines (SVM) to dynamic video classification. A system to classify motion into multiple classes using SVM is considered. The contribution of this paper consists in a simple yet effective procedure to update an initial SVM-model when new data (images) become available.

Nasraoui and Zhang have developed a method inspired by Fuzzy Associative Memories (FAMs) in order to mine search engine query logs. Analyzing recent queries the respective search engine could provide better recommendations based on search behavior of previous visitors. Using a limited active sliding window, for storing and processing recent query logs, a updating strategy have

been developed in order to adapt to dynamically changing search behaviour. The proposed approach has been applied successfully in a real search engine. Recent developments regarding, e.g. Web 2.0 provide a rich field of potential applications of the presented approach.

We hope that this special issue will become instrumental in better understanding of the challenges of various facets of dynamic data mining and will stimulate further research pursuits and applications.

We would like to take this opportunity and express our thanks to the authors who enthusiastically contributed to this special issue and shared their recent research findings. Our thanks go to reviewers whose critical yet highly constructive input was very instrumental. We are grateful to the Editor-in-Chief, Professor Rajkumar Roy for providing us with the opportunity to present this special issue under the aegis of Applied Soft Computing.

Acknowledgement

Support from the Chilean Fondecyt project 1040926 and the Millennium Science Institute "Complex Engineering Systems" (www.sistemasdeingenieria.cl) is greatly acknowledged.

Witold Pedrycz*
*Department of Electrical and Computer Engineering,
University of Alberta, Canada
Systems Research Institute,
Polish Academy of Sciences,
Warsaw, Poland*

Richard Weber
*Department of Industrial Engineering,
University of Chile, Chile*

*Corresponding author
E-mail address: pedrycz@ee.ualberta.ca
(W. Pedrycz)