

Web Intelligence & Representing of Mined Knowledge

Maurice Mulvenna¹, Alex Büchner, Sarabjot Anand, Matthias Baumgarten & Rüdiger Böhm
Faculty of Informatics, University of Ulster, Shore Road, Newtownabbey, Co. Antrim, Northern Ireland, BT37 0QB, UK

Abstract

This paper describes a research project, called NetMODEL, designed to facilitate the representation and manipulation of mined data. The paper draws upon a new European research project called CERENA to describe the first applications of the NetMODEL architecture.

NetMODEL's rationale is founded on the observation that more and more commerce-related transactions are becoming digital. This is happening not just on the supplier side of the value chain, but increasingly on the consumer side. That is, new digital interactive devices are emerging in the mass-market arena. These include the Internet, Digital Interactive Television, Personal Digital Assistants (PDA) and mobile broadband telephony devices. These devices have at least one important facet in common; *all* the marketing channel interactions between digital service provider (information content, product and service provision) and the consumer can be recorded and stored in digital databases.

Thus, for the first time, aspects of consumer behaviour are available for analysis. The major problem — the scale of analysis — can be addressed by using data mining technology, which is designed to discover knowledge in large databases; in this case, consumer behavioural patterns. However, existing data mining algorithms — artificial intelligence, database, or statistics based — have several generic drawbacks. Firstly, the output of data mining algorithms is usually in a form that is not readily understandable to those most in need to the results — marketing professionals. In the context of electronic commerce, this hinders the application of discovered marketing intelligence, which is the basis for personalisation activities. Secondly, the results are not represented in a format that allows the *exchange* of discovered knowledge. Thus, the application or exploitation of data mining results is a tedious, often impossible task. Some existing attempts, mainly the Knowledge Interchange Format (KIF), have proven too generic to be applicable in the area of knowledge discovery. It is expected that much of future research in the data mining community will focus less on algorithmic improvements and more on the generic representation and application of discovered patterns.

NetMODEL investigates and develops new ways of representing information generated from data mining algorithms. Specifically NetMODEL focuses on the development of techniques for representing the output of the MIDAS (Mining Internet Data for Associative Sequences) algorithm into a machine-readable and portable form. The NetMODEL representation is based on the eXtended Markup Language (XML), and acts as a bridge between mining and usage of discovered patterns, such as exchange across data mining systems, visualisation et cetera.

This paper outlines the technical architecture in development for NetMODEL, and how this architecture is being used within a major European bank for provision of value added services in electronic retail banking.

Key Words:

Data Mining, Web Mining, Digital Markets, Electronic Retail Banking, XML, PMML, RDF

¹ Corresponding Author Address & Contact Information: MINEIT Software Limited, c/o Faculty of Informatics, University of Ulster, Shore Road, Newtownabbey, Co. Antrim, Northern Ireland, BT37 0QB, UK
Email: maurice@mineit.com Tel: +44 28 90368875