

A European Virtual Enterprise on Collaborative Data Mining and Decision Support

Dunja Mladenić

Jožef Stefan Institute, Slovenia

Nada Lavrač

Jožef Stefan Institute, Slovenia

INTRODUCTION

One of the challenging research problems is to develop data mining and decision support integration techniques and to propose new methods for collaborative data mining. Advances in this area were achieved within the European project *Data Mining and Decision Support for Business Competitiveness: A European Virtual Enterprise (SolEuNet, 2000-2003)*, in which a virtual enterprise model was proposed as a dynamic problem-solving link between a network of experienced data mining and decision support experts on the one hand, and customers in need of specific solutions on the other.

Successful applications of data mining and decision support technologies in a real-world setting show great potential of these technologies and open the need for their further development. Directions for extending the existing technologies include collaborative problem solving and integration of the technologies at different levels, which were addressed in the European 5FP IST project *SolEuNet: Data Mining and Decision Support for Business Competitiveness: A European Virtual Enterprise* (Mladenić & Lavrač, 2003a; Mladenić, Lavrač, Bohanec, & Moyle, 2003b), whose aim was to develop a framework, methods, and tools for the integration of data mining and decision support, as well as their application to business problems in a collaborative setting.

Data mining and decision support are, each on their own, well-developed research areas, but until recently there has been no systematic attempt to integrate them. The main innovation achieved in SolEuNet was the bridging of these two technologies, that had a significant impact on the developments of both fields, by improving approaches for problem solving in

real-world settings, enabling the fusion of knowledge from experts and knowledge extracted from data, and consequently enabling the successful solution of new types of problems.

To enhance competitiveness and find new collaborative business opportunities in the global market, the objective of the project was also to provide access to cutting-edge ICT technologies through a proposed model of a European virtual enterprise composed of companies and research laboratories with highly specialized expertise in data mining and decision support. The SolEuNet virtual enterprise was envisioned as a business structure made of small, cross-organizational, time-focused, task-driven work teams, providing problem solutions to end users in industry, businesses, and public services. The work included the development and evaluation of the virtual enterprise model, enhancement of tools and method for cooperative work, combining problem solutions and consensus building, advances in decision support and data mining techniques enhancing the CRISP data mining methodology, advances in text mining and Web mining methods and tools, as well as new education and training methods, and Web information source maintenance.

The rest of this contribution describes the background by describing the methodology for collaborative data mining projects, the main SolEuNet achievements, the e-collaboration aspects of the project, and the conclusions describing the lessons learned.

BACKGROUND: METHODOLOGY FOR COLLABORATIVE PROBLEM SOLVING

The core of data mining is the extraction of useful patterns or models from data (Hand, Mannila, & Smyth,

2001). However, to reach actionable results from data usually requires a long and non-trivial process (Berry & Linoff, 1997) involving aspects of business and technology (Pyle, 1999), as well as human skill; the human factor is one of the most important success factors, including project management and control. A well-defined process is important for achieving successful data mining results, particularly if the number of participants involved in carrying out the data mining tasks is large, involving teams of individuals with different expertise, skills, habits, and cultural backgrounds.

Many authors have suggested broadly defined process models to perform data mining (Adriaans & Zantinge, 1996; Fayyad, Piatetsky-Shapiro, & Smyth, 1996). The emerging standard data mining process model is the CRoss Industry Standard Process for Data Mining (CRISP-DM) (Chapman et al., 2000). CRISP-DM subdivides a data mining project into six interrelated phases: (1) business understanding, (2) data understanding, (3) data preparation, (4) modeling, (5) evaluation, and (6) deployment. Like in alternative data mining processes, there are numerous feedback loops connecting the phases in CRISP-DM. As data mining is multi-disciplinary it often requires the expertise of numerous individuals. The business understanding phase requires communication skills to work closely with the data mining client (the organization interested in the data mining results). The modeling phase—which requires the use of statistics or machine learning—can be undertaken largely independently of others, making it possible to perform parts of a data mining process in a remote e-collaboration setting. To ensure that collaborative data mining is successful, we propose well-defined collaboration principles and support tools (Jorge et al., 2003a; Mladeníc et al., 2003a, 2003b).

A data mining project that is collaborative involves more complexity than the one that is small and local, but there are benefits to combining expertise. To realize such benefits it is vital that all collaborating parties share their results, either complete or intermediate. For example, in the data preparation phase, any data transformations should be made available; while in the modeling phase, the models should be made available in a standardized format. Information needs to be securely but easily shared using an appropriate e-collaboration knowledge management system. The evaluation phase is important in the data mining process as in this phase the key results—the models—are evaluated against the initial project objectives. When working in a collabora-

tive setting it is important that all models be evaluated fairly and consistently. This is best done by centralizing the model evaluation as much as practical.

E

MAIN SOLEUNET ACHIEVEMENTS

SolEuNet project partners succeeded to perform their research and achieve their business goals through Internet-supported collaboration, as partners of a virtual enterprise of expert teams from academia and business, offering expertise and tools designed to solve end-users' data mining and decision support needs. This was achieved by the innovations in collaborative data mining and decision support, combining solutions, integrating data mining and decision support, and integrating data mining and decision support with information systems. The main project achievements are outlined next.

Business

The project has established a virtual enterprise of researchers and marketing experts in data mining and decision support, formed through the business opportunity of proposing a project to be funded by the European Commission. In order to act as a networked organization, SolEuNet needed to develop a model and procedures for solving client problems by forming virtual enterprises involving business and academic partners for each individual business opportunity occurring at the advent of each data mining and decision support solution developed for a particular client. Specific systems and procedures were developed enabling collaborative data mining and decision support in solving client problems, occasionally with the collaboration of external data mining and decision support experts.

Case Studies and Applications

Several collaborative data mining and decision support end-user projects were conducted including the analysis of media research data for a marketing research company, brand name recognition for a direct marketing campaign, customer quality evaluation, and stock market prediction for a large financial house, predicting the use of resources in health farms, analysis of road traffic accidents, semi-automated ontology construction from education materials on the Web for a large publishing house, analysis of Web page access to

4 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/european-virtual-enterprise-collaborative-data/17655

Related Content

Interaction, Imagination and Community Building at the Math Forum

Wesley Shumar (2009). *Virtual Community Practices and Social Interactive Media: Technology Lifecycle and Workflow Analysis* (pp. 269-282).

www.irma-international.org/chapter/interaction-imagination-community-building-math/30822

GLARE: An Open Source Augmented Reality Platform for Location-Based Content Delivery

Enrico Gandolfi, Richard E. Ferdig, David Carlyn, Annette Kratcoski, Jason Dunfee, David Hassler, James Blank, Chris Lenartand Robert Clements (2021). *International Journal of Virtual and Augmented Reality* (pp. 1-19).

www.irma-international.org/article/glare/290043

Omnichannel Retailing: A Comprehensive Exploration of Integration, Customer Engagement, and Market Share in Today's Retail

Sanjay Tanejaand Rishi Prakash Shukla (2024). *Omnichannel Approach to Co-Creating Customer Experiences Through Metaverse Platforms* (pp. 60-76).

www.irma-international.org/chapter/omnichannel-retailing/341023

The Role of Mechanics in Gamification: An Interdisciplinary Perspective

Miralem Helmeffalk, Siw Lundqvistand Leif Marcusson (2019). *International Journal of Virtual and Augmented Reality* (pp. 18-41).

www.irma-international.org/article/the-role-of-mechanics-in-gamification/228944

Thinking in Virtual Spaces: Impacts of Virtual Reality on the Undergraduate Interior Design Process

Elizabeth Poberand Matt Cook (2019). *International Journal of Virtual and Augmented Reality* (pp. 23-40).

www.irma-international.org/article/thinking-in-virtual-spaces/239896