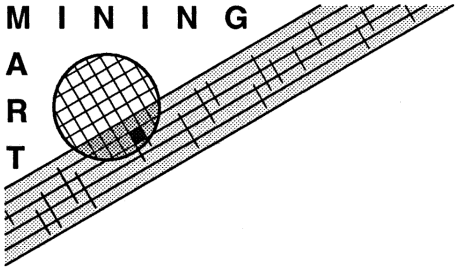

Enabling End-User Datawarehouse Mining
Contract No. IST-1999-11993
Deliverable No. D11.1

One Day Seminar – Data Mining in Practice
February 18th, 2003

Martin Scholz

University of Dortmund, Computer Science
D-44227 Dortmund, Germany
scholz@ls8.cs.uni-dortmund.de

February 28, 2003

1 Introduction

The "*One Day Seminar – Data Mining in Practice*" was held in the scope of the MiningMart project. It was carried out by and held at the University of Dortmund as part of work package 11 ("Exploitation and Transfer of Results") on February 18th, 2003.

The objective of the seminar was to present the results of the project and to attract contributions to the case-base. In preparation to this workshop it was announced at conferences and on several mailing lists for KDD, Machine Learning and neighbouring disciplines. There were about 70 registrations from the industrial and scientific sector.

The programme consisted of a variety of presentations, illustrating the benefit of using the MiningMart system for real world applications in KDD. A system demo and several examples of exploiting the system for rapid application development in various domains gave a good picture of the applicability and the advantages compared to manual or SAS-like preprocessing. The mixed contributions of project partners from the industrial and academic sector made it possible to point out the results of the project at different levels of abstraction, and to address different audiences using their preferred terminology.

Two industrial partners showed serious interest in using the system and cooperating with some of the project partners. A participant of the European IST project cInQ suggested a meeting to share experience, and to evaluate possible ways of collaboration in the near future.

2 The Presentations

All presentations and a list of participants are accessible on the Internet pages of the MiningMart project under the following address:
http://www-ai.cs.uni-dortmund.de/MMWEB/content/oneDaySeminar.html
The following sections give short summaries of the main contributions.

2.1 The MiningMart Approach

Presented by: Katharina Morik (UniDo)

The first presentation illustrated the philosophy guiding the MiningMart project. The general ideas underlying the M4 model, the role of re-using best-practice cases and the realization were shown in detail. The introduction was followed by a demonstration of the system, showing how easy powerful operators can be applied with MiningMart. Finally the Internet case-base of best-practice cases was announced.

2.2 Spatial Data Mining for Customer Segmentation

Presented by: Michael May (AiS)

Along some practical applications from various domains the notion of Spatial Data Mining was introduced. Possible synergies between the projects SPIN! and MiningMart became obvious, and the importance of the correct representation for the Data Mining tasks was stressed.

2.3 Analyzing Churn of Customers

Presented by: Marco Richeldi (TILab)

The Churn Prediction Case provided by Telecom Italia Lab in the scope of the project was presented. After introducing the important problem of customer churn the talk focussed on using MiningMart to aggregate data from different sources and to build models for the prediction task. Several screen shots illustrated how easy a complex chain of preprocessing can be set up with the MiningMart system.

2.4 Knowledge Discovery Services and Applications

Presenter: Jörg-Uwe Kietz (kdlabs)

The presentation stressed the high relevance of MiningMart's basic features for KDD, namely clever preprocessing as the key for successful knowledge discovery, re-use as the key for providing knowledge discovery services, and DB-based processing for handling large amounts of data. The Swiss kdlabs company provides KDD services in the fields of Marketing/CRM, Website analysis, Credit risk analysis, and Fraud detection. Examples of practical applications were given, to reformulate the abstract advantages of MiningMart in practical terms, and to illustrate the potential of KDD.

2.5 Telecommunication Case Modelling - Call Center

Presented by: Janusz Granat (NIT)

The Call Center Case was introduced by the National Institute of Telecommunication, Poland. Apart from introducing a case re-usable in similar form for other domains, the main contribution of this presentation was a direct comparison of MiningMart to formerly applied systems like SAS. The documentation facility and the opportunity to set up re-usable preprocessing chains without having to make use of a programming language were pointed out as real valuable benefits of MiningMart. Additionally, by the means of a conceptual data model cases can much easier be adapted to structural changes in the data than if all attributes are hard coded like in many low-level preprocessing procedures.

2.6 Prediction from Quasi-Random Time-Series

Presented by: Lorenza Saitta (DISTA)

A case study on a sales prediction task from time series was presented. The learning task could not be addressed with several intuitive representations of the sales data. Another representation based on cumulative sales turned the random-like data into a format, which finally allowed for building an appropriate model. This example proved, how fundamentally clever preprocessing influences the quality of data mining results.

2.7 Integrating Knowledge Discovery into Knowledge Management

Presented by: Katharina Morik, (UniDo)

The presentation showed how easy Knowledge Discovery technologies can be integrated into an overall framework of Knowledge Management. The prototype of a system was presented, which was capable of answering high-level queries by a combination of KDD and queries to the Internet or the companies Intranet.