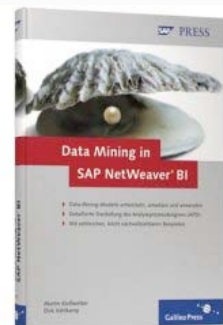




[REVIEWED FOR YOU]

mayato book recommendations in the area of business intelligence and customer relationship management

Title of book: **Data Mining in SAP NetWeaver® BI**
 Author(s): **Martin Kiesswetter, Dirk Vahlkamp**
 Details: **in German / 375 pages / SAP Press 2007/ Price €69.90**
 Reviewed by: **Dr. Marcus Dill**



1 Recommendation

I can recommend this book for BW specialists who have, up until now, not looked in detail at SAP Data Mining and the Analysis Process Designer in SAP BW, but who wish to gain an overview of the functionality. However, readers expecting to find background information, insiders' tips, and hints from experts will be largely disappointed. In particular, the assessment of the strengths and weaknesses of the different methods and features lacks the treatment it deserves. As a result, a usable handbook is still not available to readers looking for a practical approach. This particularly applies to administrators, architects, and project managers, because the book concentrates primarily on superficial software features. The authors have missed the opportunity to provide real added value, beyond SAP's standard documentation.

2 Summary of content

After an introductory chapter, the authors embark on the second section, in which they introduce readers to the fundamentals of data mining, its methods, and its applications in a highly compact but nevertheless well explained manner. They initially orientate themselves toward the methods provided by SAP, and dispense with a more comprehensive description. Chapter 3 summarizes the functions of SAP BW's Analysis Process Designer and Data Mining Workbench. The next two chapters present the procedures available in SAP Data Mining. In chapter 6, the authors introduce readers to two more types of method: genetic algorithms and neural networks. Implementation examples show how these can be integrated into SAP BW.

3 Assessment criteria

Content: The authors mainly take the perspective of an interested BW user and hardly examine the architecture and technology at all. Likewise, they barely touch on algorithms and analysis methods. Essentially, the book regurgitates the content of SAP's online help and accompanying documentation and support notes. Readers looking for background information, insiders' tips, or recommendations about how to use the available tools properly will be disappointed. Data mining experts will, in general, find the description of the methods superficial and will probably learn nothing new from reading these sections. Important interrelations that have a major impact on system performance and usability are not covered. Many users are not aware of the fact that, if an analysis process is poorly designed, they will not only have to wait a long time for results, but they may well be putting a great load on the system and database. Yet there are no references to the point or the possibility of using samples. Project managers will find no recommendations for deployment strategies and procedure models. Architects could integrate data mining or analysis processes more easily into broader data warehouse processes – or even into operations – if only someone would explain the many opportunities available to them. But, for example, exporting models in XML format (in the case of data mining models as PMML) goes without a mention. The authors apparently do not even know that SAP provides a certifiable interface through which BW customers can integrate their own custom-built data mining procedures – because they suggest integrating them using ABAP routines. They fail to mention the official interface, just as they fail to mention the fact that the range of data mining procedures in the Analysis Process Designer can be easily extended using certified supplementary products. For instance, Oracle has enabled the integration of its database-based data mining methods – which are available anyway in most BW databases – using this interface. Finally, administrators and BW managers are left in the dark about, for example, security aspects and authorizations: The Analysis Process Designer allows access to many different data sources and opens up an infinite number of options (including table updates and the chance to change all sorts of system settings). If you don't stick to a few simple rules, the Analysis Process Designer can thus cause security breaches or become a risk to system stability. Even if this book was not designed to be supplementary documentation for end users, the lack of such information must surely be seen as negligent.

Readability: The content is generally comprehensible, very well written, and illustrated using plausible examples. The explanations are supported by helpful graphics and numerous screenshots, and are generally easy to follow. The chapters are structured clearly and take a standardized approach, making it easy for readers to find their way around. In general, the book is not a struggle to read. In fact, its style is very pleasant.

Practical use: The fact that the book is highly readable makes it recommendable as an introduction to the subject. However, it is not really suitable as a practical handbook, because many aspects are not addressed. Due to a dearth of specialist literature, experienced – and expensive – consultants will continue to dominate the field. The great lack of information leads me to suspect that the authors acquired their knowledge purely from official sources and their own experiments in sandbox systems – but are not actually familiar with the Analysis Process Designer and SAP Data Mining in a production system and live operation, and nor do they profit from third-party expert knowledge. In the acknowledgements section of the preface, the authors do not list any SAP employees whose names might lead one to assume that background information from developers or experienced consultants could have found its way into the book. While the book makes a very good impression in terms of editing and production, the index was rather neglected. It is fairly obvious that it was created using automated intelligence – and without subsequent quality assurance. Many important key words are not there, even though the terms are explained in the text. Instead, key words such as “problem 4A” are listed. As a result, the book's usefulness as a work of reference is limited.