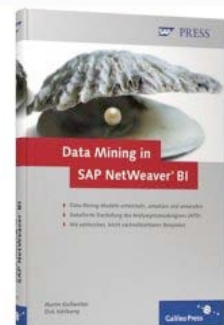




[FÜR SIE GELESEN]

mayato Buchempfehlungen im Umfeld Business Intelligence und Customer Relationship Management

Buchtitel: Data Mining in SAP NetWeaver® BI
Autor(en): Martin Kießwetter, Dirk Vahlkamp
Details: Sprache Deutsch / 375 Seiten / SAP Press, 2007 / Preis 69,90 €
Rezensent: Dr. Marcus Dill



1 Empfehlung

Das Buch ist zu empfehlen für BW-Kundige, die sich mit SAP Data Mining und dem Analyseprozessdesigner im SAP BW bisher nicht beschäftigt haben, sich nun aber einen Überblick über die vorhandene Funktionalität verschaffen wollen. Wer Hintergründe, Geheimtipps und Expertenhinweise erwartet, wird jedoch über weite Strecken enttäuscht. Zu kurz kommt insbesondere die Bewertung der Stärken und Schwächen der einzelnen Verfahren und Features. Praxisorientierten Lesern fehlt somit nach wie vor ein brauchbares Handbuch. Dies gilt insbesondere für Administratoren, Architekten und Projektleiter, da sich das Buch weitgehend auf oberflächliche Softwarefeatures konzentriert. Die Autoren verpassen die Chance, einen echten Mehrwert über die Standarddokumentation der SAP hinaus zu liefern.

2 Inhaltzusammenfassung

Nach einem einleitenden Kapitel führen die Autoren im zweiten Abschnitt in stark verkürzter, aber doch gut verständlicher Weise in die Grundlagen des Data Mining, seine Verfahren und seine Anwendung ein. Sie orientieren sich dabei zunächst an den von SAP bereitgestellten Methoden und verzichten auf eine umfassendere Darstellung. Kapitel 3 fasst die Funktionen des Analyseprozessdesigners und der Data Mining Workbench in SAP BW zusammen. Die beiden anschließenden Kapitel stellen die verfügbaren Verfahren des SAP Data Mining im Detail vor. Im Kapitel 6 führen die Autoren mit Genetischen Algorithmen und Neuronalen Netzen in zwei zusätzliche Verfahrensklassen ein. Implementierungsbeispiele zeigen, wie diese Verfahren in SAP BW integriert werden können.

3 Bewertungskriterien

Inhalt: Die Autoren beschränken sich weitgehend auf die Sicht des interessierten BW-Anwenders und verzichten fast vollständig auf eine tiefer gehende Betrachtung von Architektur und Technologie, aber auch von Algorithmen und Analyseverfahren. Im Wesentlichen gibt das Buch die Inhalte der SAP Onlinehilfe sowie begleitender Dokumentationen und Support-Hinweise wieder. Hintergrundinformationen, Insidertipps oder Empfehlungen über den richtigen Umgang mit den verfügbaren Werkzeugen sucht der Leser vergeblich. Data-Mining-Experten werden generell die Darstellung der Verfahren als oberflächlich empfinden und von der Lektüre dieser Abschnitte kaum neue Erkenntnisse gewinnen. Wichtige Zusammenhänge, die maßgeblich die Performance und die Nutzbarkeit des Systems beeinflussen, werden nicht dargestellt. Vielen Anwendern ist nicht klar, dass sie bei schlechtem Design eines Analyseprozesses nicht nur lange auf Ergebnisse warten müssen, sondern u.U. auch das System und die Datenbank stark belasten. Hinweise auf den Sinn und die Möglichkeit des Einsatzes von Stichproben findet man jedoch nicht. Als Projektleiter vermisst man Empfehlungen zu Einsatzstrategien und Vorgehensmodellen. Architekten könnten Data-Mining- oder Analyseprozesse leichter in größere Abläufe im Data Warehouse oder sogar in operative Prozesse integrieren, wenn ihnen die Vielfalt der Möglichkeiten erläutert würde. Es bleibt jedoch z.B. der Modellexport im XML-Format (im Falle der Data-Mining-Modelle als PMML) unerwähnt. Dass SAP eine zertifizierbare Schnittstelle anbietet, über die BW-Kunden auch selbstentwickelte Data-Mining-Verfahren integrieren können, scheinen die Autoren selbst nichts zu wissen, da sie stattdessen eine Integration über ABAP-Routinen vorschlagen. Sie erwähnen die offizielle Schnittstelle genauso wenig wie die Tatsache, dass die Palette an Data-Mining-Verfahren im APD über diese Schnittstelle leicht durch entsprechend zertifizierte Zusatzprodukte ergänzt werden kann. So bietet Oracle die Integration seiner datenbankbasierten – und in den meisten BW-Datenbanken sowieso vorhandenen - Data-Mining-Methoden über dieses Interface an. Administratoren und BW-Verantwortliche schließlich werden z.B. über Sicherheitsaspekte und Berechtigungen im Dunkeln gelassen: Der APD ermöglicht Zugriffe auf verschiedenste Datenquellen und eröffnet über ABAP-Routinen eine unendliche Vielzahl von Optionen (bis hin zu Tabellenupdates und Eingriffen in jegliche Systemeinstellungen). Bei Nichtbeachtung einiger einfacher Regeln kann der APD so zur Sicherheitslücke oder zum Risiko für die Systemstabilität werden. Auch wenn das Buch sich nur als ergänzende Dokumentation für Endanwender versteht, muss der Verzicht auf einen diesbezüglichen Hinweis als fahrlässig gelten.

Lesbarkeit: Die Inhalte sind generell verständlich, sprachlich ausgezeichnet und mit eingängigen Beispielen beschrieben. Die Erklärungen sind mit hilfreichen Grafiken und zahlreichen Screenshots unterlegt und insgesamt gut nachzuvollziehen. Die Kapitel sind sauber und einheitlich strukturiert und erlauben eine gute Orientierung. Das Buch liest sich so über weite Strecken sehr angenehm und ohne Anstrengung.

Praxistauglichkeit: Die gute Lesbarkeit macht das Buch zu einer empfehlenswerten Einstiegsliteratur. Als Praxishandbuch ist das Buch allerdings nur bedingt geeignet, da zahlreiche zentrale Aspekte nicht behandelt werden. Da es an Spezialliteratur weitgehend fehlt, wird das Feld in der Praxis erfahrenen – und teuren - Beratern überlassen bleiben. Die Vielzahl der fehlenden Informationen lässt vermuten, dass die Autoren ihr Wissen ausschließlich über offiziell verfügbare Informationen und durch eigenes Herumprobieren in Sandboxsystemen erworben haben, den APD und SAP Data Mining aber weder im produktiven und Projektbetrieb kennen noch vom Expertenwissen Dritter profitierten. Im Vorwort findet sich unter den Danksagungen keine Nennung von SAP-Mitarbeitern, die darauf schießen ließe, dass Hintergrundinformationen von Entwicklern oder erfahrenen Beratern eingeflossen sein könnten. Während das Buch hinsichtlich Lektorat und Herstellung einen sehr guten Eindruck macht, wurde dem Index keine ausreichende Aufmerksamkeit geschenkt. Er wurde ziemlich offensichtlich ausschließlich mit maschineller Intelligenz erstellt - ohne anschließende Qualitätsprüfung. Viele wichtige Stichworte fehlen, auch wenn die jeweiligen Begriffe im Text erläutert werden. Stattdessen stehen dort Stichworte wie "Problem 4A". Eine Verwendung des Buches als Nachschlagewerk ist somit stark eingeschränkt.