



## PRESSEMITTEILUNG

mayato Data Mining Studie 2011

### **Data-Mining-Werkzeuge: Praktischer Einsatz im Customer Relationship Analytics**

Berlin, 16. März 2011 — Die diesjährige Data Mining Studie des BI-Analysten- und Beraterhauses mayato legt den Schwerpunkt auf Kundensegmentierung als wesentlicher Bestandteil von Customer Relationship Analytics. Dazu traten anhand eines 270.000 Zeilen umfassenden Testdatensatzes eines Finanzdienstleisters vier Data-Mining-Tools und -Suiten gegeneinander an: Der SAS Enterprise Miner 6.2, der STATISTICA Data Miner 10 von StatSoft, das Open-Source-Werkzeug WEKA 3.6.3 und das SAP BW 7 (Data Mining Workbench). Die Studie ist ab sofort zu einem Preis von 299 Euro für Unternehmen und 99 Euro für Studenten und Hochschulen unter [www.mayato.com](http://www.mayato.com) erhältlich.

Für den Vergleichstest schuf mayato ein komplexes, praxisnahes Testszenario, das von der Datenvorverarbeitung über die eigentliche Erstellung der Kundensegmentierung und Interpretation der Ergebnisse bis hin zur praktischen Anwendung der Segmentierung zur Neukundenklassifikation alle Aspekte des Analyseprozesses abdeckt.

Darüber hinaus wurde für jedes der Werkzeuge eine detaillierte Funktionsübersicht zur Clusteranalyse erstellt. Die Bewertungskriterien umfassen unter anderem die Bedienung der Tools, den Funktionsumfang, das Systemverhalten bei großen Datenmengen und die Stabilität. Weiterhin wurde die Ausführungsgeschwindigkeit mit zahlreichen Parametereinstellungen gemessen und dokumentiert.

Im Endergebnis zeigen sich starke Unterschiede insbesondere im Bedienkomfort, bei der Ausführungsgeschwindigkeit sowie bei den Möglichkeiten der grafischen Ergebnisauswertung.

#### **Das Testfeld**

Data-Mining-Funktionen werden inzwischen in gänzlich unterschiedlichen Tool- und Preiskategorien angeboten. Entsprechend breit fiel das Testfeld aus: Die klassische,

funktionsmächtige Data-Mining-Suite (SAS, StatSoft) findet sich ebenso darunter wie das Business-Intelligence-Werkzeug, das Data-Mining-Verfahren eher als Zusatz zu mächtigen Datenverwaltungsfunktionen offeriert (SAP). Um die Frage zu klären, ob sich die Analyseaufgabe auch mit einer kostenlosen Data-Mining-Suite zufriedenstellend lösen lässt, wurde zusätzlich eine Open-Source-Variante in den Test aufgenommen (WEKA).

„Die Kundensegmentierung ist eine stark explorativ geprägte und im CRM-Kontext wichtige analytische Fragestellung“, äußert Peter Neckel, Analyst und Leiter der Studie beim BI-Analysten- und Beraterhaus mayato. „Man durchläuft dabei typischerweise mehrere Analysezyklen; daher sind die Analyseeffizienz und eine gute Bedienbarkeit der Tools in diesem Einsatzbereich besonders wichtig“.

### **Der Praxistest**

Im Praxistest wurden rasch die Stärken und Schwächen der jeweiligen Tools sichtbar: Den höchsten Reifegrad haben deutlich erkennbar der SAS Enterprise Miner und der STATISTICA Data Miner erreicht. Sie sind flüssig zu bedienen und leisten sich im gesamten Test keine echte Schwäche. Sie punkten beide besonders mit hohen Ausführungsgeschwindigkeiten und dem sicherem Umgang mit großen Datenmengen.

Mit bereits deutlichem Abstand folgt das SAP BW mit der Data Mining Workbench: Hier muss der Anwender vor allem Abstriche hinsichtlich des Funktionsumfangs und des Bedienkomforts in Kauf nehmen. Zudem ist die Dokumentation stark ergänzungsbedürftig, was die Einarbeitung erschwert. Überzeugen kann hingegen die grafische Darstellung der Ergebnisse; sie unterstützt die Interpretation durch einen guten optischen Eindruck und durch praxisgerechte Voreinstellungen.

WEKA wird von dem großen Testdatensatz sichtlich am meisten gefordert. Die gemessenen Ausführungszeiten sind wenig konkurrenzfähig, die allgemeine Systemstabilität ist für den professionellen Einsatz noch nicht ausreichend. Die nur rudimentär vorhandene Ergebnispräsentation erschwert die Interpretation der Clustereinteilungen zusätzlich.

### **Einfachere Bedienkonzepte erleichtern den Einstieg**

Ein dominierendes Bedienkonzept lässt sich – wie schon in den letzten Jahren – weiterhin noch nicht erkennen. Allerdings sind die Hersteller zunehmend um Einsteigerfreundlichkeit bemüht: SAS bietet neuerdings mit dem Rapid Predictive Modeler eine Data-Mining-Umgebung mit sinnvoll begrenzten Parametrisierungsmöglichkeiten an; im Data Miner von StatSoft kann man bereits

seit der letzten Version wahlweise in den Assistentenmodus umschalten, der vorgefertigte „Data Mining Rezepte“ für Standardanalyseaufgaben bereitstellt.

### **Flexible Preismodelle**

Der Investitionsaufwand für Data-Mining-Suiten liegt weiterhin – abhängig von der Nutzeranzahl – meist im sechsstelligen Euro-Bereich. Hier geht der Trend jedoch zu günstigeren Konditionen, da immer mehr Hersteller einzelne Funktionen in Pakete schnüren und separat anbieten. So kann der Nutzer ein an seine individuellen Bedürfnisse angepasstes Produkt erwerben, das zudem bei Bedarf einfach zum Differenzpreis erweiterbar ist.

Aber auch komplette Suiten kosten nicht mehr zwangsläufig sechsstellige Summen: Die Einzelplatzlizenz des STATISTICA Data Miners ist bei vollem Funktionsumfang beispielsweise ab 20.000 Euro erhältlich. Für Open-Source-Lösungen entfällt der Anschaffungspreis; bei professioneller Nutzung sollten jährliche Supportgebühren im vierstelligen Eurobereich einkalkuliert werden.

Insgesamt betrachtet lässt die diesjährige Studie erkennen, dass Data-Mining-Software inzwischen einen beachtlichen Reifegrad erreicht hat. Vor allem das vielfältige Produktangebot, zunehmend flexible Preismodelle sowie leichter erlernbare Bedienkonzepte haben die Einstiegshürden spürbar verringert.

Zeichenzahl: 5.530

### **Über mayato**

[mayato](#) ist als Analysten- und Beraterhaus spezialisiert auf Business Intelligence. Von Niederlassungen in Berlin, Bielefeld und Heidelberg aus arbeitet ein [Team](#) von erfahrenen IT- und BI-Architekten, Statistikern, Analysten sowie fachlichen Experten für spezielle Themen wie [Betrugserkennung](#), [Data Mining](#) und [Analytisches CRM](#). Zu den [Kunden](#) von mayato zählen namhafte Unternehmen aus unterschiedlichen Branchen. Als Partner mehrerer [Softwareanbieter](#) ist mayato grundsätzlich der Neutralität und in erster Linie der Qualität seiner eigenen Dienstleistungen verpflichtet. Nähere Infos unter [www.mayato.com](http://www.mayato.com).

### **Pressekontakt:**

mayato GmbH  
Am Borsigturm 9  
D-13507 Berlin  
Dr. Marcus Dill  
Tel. +49 178 490 6727  
marcus.dill@mayato.com

Stemmermann – Text & PR  
Hinter dem Rathaus 1  
D-23966 Wismar  
Tanja Stemmermann  
Tel. +49 3841 22 43 14  
info@stemmermann-pr.de