



17



ANALYTICS

Weitere Artikel zum Thema:

- Hadoop: Die Big-Data-Plattform erweist sich im Praxiseinsatz noch als kompliziert. S. 114
- Sportlich: Die Analyse großer Datenmengen führt im Sport zu neuen Erkenntnissen. S. 118

DIE ZUKUNFT BERECHNEN

Aufgrund immer umfangreicherer Datenmengen und -quellen wird die Zukunft mit Hilfe von Predictive Analytics, der Königsdisziplin im Bereich Business Intelligence, bis zu einem gewissen Grad vorhersagbar.

Oliver Weiss

© Fotolia / Nikolai Chertyshev



Schon vor Big Data war Business Intelligence (BI) ein heißes Thema und laufend wachsende Datenmengen machen es endgültig zum Muss. Die Anforderungen bei Reporting, Analyse und Prognosen sind in den letzten Jahren dramatisch gestiegen und setzen die bisherigen Datenarchitekturen in Unternehmen unter Druck, so die Beobachtung der Analysten des Business Research Application Center (BARC). Zwar verlieren bestehende Data-Warehouse-Systeme nicht ihren Wert, doch können sie mit traditionellen Reports und einer vorwiegend auf Transaktionsdaten basierenden Datenhaltung nicht alle Informationsbedürfnisse stillen. Unternehmen müssen immer mehr zusätzliche Datenquellen und -typen berücksichtigen, die aus neuen Datencon-

tainern wie Hadoop stammen. Hand in Hand mit dieser Entwicklung nimmt auch die Bedeutung von Datenintegration und Datenqualitätsmanagement weiter zu. Ohne konsistente, vollständige und korrekte Daten fehlt BI-Vorhaben eine solide Basis, um Entscheidungen unterstützen zu können. Zusätzlich sollte das Datenqualitätsmanagement organisatorisch durch die Benennung von »Data Stewards« aus den Reihen der Fachanwender gestärkt werden, rät BARC. Diese sind dafür verantwortlich, dass Qualitätsaspekte wie Datenstandards, Konsistenz oder Korrektheit eingehalten werden. Ein weiteres Thema für die Zukunft bleibt der Aufbau einer Data Governance, die bislang in vielen Unternehmen ein bloßes Lippenbekenntnis ist. Auch wenn es aufwändig er-

scheint, Governance-Strukturen und -Prozesse in der eigenen Organisation aufzubauen: Es ist alternativlos. Umso mehr, da die Datenmengen und neuen Datenquellen rasant zunehmen. Neben der traditionellen OLAP-Analyse haben sich in den letzten Jahren weitere Verfahren und Techniken am Markt etabliert, wie Set-based, Visual, Real-time oder Predictive Analytics. Sie kommen immer häufiger zum Einsatz, weil sie heutigen Anforderungen und Anwendungsmöglichkeiten in der Datenanalyse, beispielsweise in punkto Datenmenge und Geschwindigkeit, besser nachkommen. Entsprechende Werkzeuge bieten oft eine am Fachanwender orientierte Arbeitsumgebung, die mehr Flexibilität bei der Informationsnutzung und -aufbereitung gibt. Speziell in



© Fotolia / Industrieblick

Ein Beispiel für Predictive Analytics ist Predictive Maintenance, die vorausschauende Wartung von Maschinen.

Anwendungsszenarien, in denen es um zeitnahe Abfragen und eine ebenso schnelle Bereitstellung von Datenauswertungen oder Events in den operativen Prozessen geht, sind Werkzeuge für Real-time oder Streaming Analytics im Kommen.

Die größte Nachfrage werden in den kommenden Jahren Werkzeuge für Predictive Analytics erfahren, prophezeit BARC. Ziel ist der Einsatz von Algorithmen zur Vorhersage von Werten. Oft werden solche Produkte mit denen für Data Mining gleichgesetzt, was aber falsch ist, da Data Mining auch die Segmentierung und Assoziierung als Verfahren umfasst, die aber nicht vorhersagend oder vorausschauend sind. Doch genau darum geht es bei Predictive Analytics: Unternehmen aus vielen Branchen werfen auf Basis statistischer Methoden einen Blick in die Zukunft. Dynamische Preisgestaltung, Absatzprognosen oder vorausschauende Wartung – das Spektrum der Anwendungen von Predictive Analytics ist breit gefächert.

Ein Beispiel für den Einsatz von Predictive Analytics ist die dynamische Preisgestaltung. Unternehmen können damit ihre Preise – auch in Echtzeit – an die aktuelle Markt- und Wettbewerbssituation sowie an Verfügbarkeiten und Lieferbedingungen anpassen. Im Modebereich sind Preise beispielsweise im Winter- oder Sommerschlussverkauf entscheidend, um die Lager pünktlich zum Saisonende zu räumen. Auch Wetterdaten oder Feiertage fließen in die Berechnung ein.

DYNAMISCHE PREISE

Ein konkretes Beispiel: Wie viel darf eine Badehose kosten, wenn es in den Pfingstferien regnet? Was darf sie kosten, wenn die Sonne scheint? Ein Modehändler weiß aus seinen historischen Daten, dass der Absatz von Badehosen an den ersten warmen und freien Tagen generell deutlich steigt. Kunden, die das schöne Wetter für einen Tag am See genießen wollen oder über Pfingsten in den Urlaub fahren, sind bereit, genau jetzt Geld für eine Badehose auszugeben. Das heißt: Der Händler kann mehr Geld verlangen. Eine Predictive-Analytics-Lösung rechnet hier mit Hilfe externer Daten (Wetter, Feiertage etc.) und interner Faktoren wie Einkaufspreise, Umsatzziel, historische Preise etc. mehrere Szenarien durch und ermittelt automatisiert den optimalen Preis.

Egal wie der Anwendungsfall aussieht, Ziel von Predictive Analytics ist es, auf Basis von Data Mining, maschinellem Lernen und anderen statistischen Methoden Vorhersagen über die Wahrscheinlichkeit von zukünftigen Ereignissen zu treffen. Das wirft einige Fragen auf: Wie stichhaltig sind diese Prognosen? Worauf müssen Unternehmen bei der Implementierung einer Predictive-Analytics-Lösung achten? Welche Vorge-

hensweise (Strategie, Auswahl der Daten, Methodik) ist zu empfehlen? Wie lassen sich die gewonnenen Erkenntnisse erfolgreich in die Geschäftspraxis übertragen?

Vor jedem Pilotprojekt steht erst eine ausführliche Analyse der Situation des Anwenders. Welche Einflussfaktoren treiben das Geschäftsmodell? Wann sind die Umsätze niedrig, wann hoch? Wann steigt der Gewinn? Welche Ziele hat der Kunde? Um welche externen Informationen müssen die internen Daten angereichert werden, damit Ursache und Wirkung richtig zueinander passen? »Auf Basis der ausgewählten Daten bauen wir ein Modell, das aufzeigt, welcher Mehrwert sich mit besser abgesicherten Prognosen für das Business konkret erzielen lässt. So kommen wir zum Business Case«, erklärt Dirk Böckmann, Vorstand und Partner bei der Avantum Consult AG, einer Tochter des SAP-Dienstleisters All for One Steeb AG. Anhand individueller Kennzahlen werde darin ein konkretes Ziel festgelegt, das der Kunde über Predictive Analytics erreichen will.

AUSWAHL DER DATEN

Nach der Analyse der Geschäftsprozesse und dem Festlegen der Ziele folgt die Auswahl der Daten, die für die konkrete Analyse relevant sind. »Ausgangspunkt ist immer die konkrete Fragestellung. Sie können nicht die komplette Datenbasis verwenden, erläutert Michael Milnik, Team Lead Customer Analysis bei der Blue Yonder GmbH, einem Anbieter von Predictive-Analytics-Lösungen. Dabei sei es meist wirtschaftlich sinnvoller, die Daten zuzuschneiden, die später über Algorithmen und Vorhersagemodelle ausgewertet werden. Die Auswahl der Datensätze hängt immer von der Fragestellung ab, lässt sich aber Schritt für Schritt erweitern«, so



Milnik weiter. Interne Daten werden meist um externe Informationen wie Wetterdaten oder Ferien- und Feiertags-Termine ergänzt, um aussagekräftige Prognosen zu erhalten.

INFRASTRUKTUR

Grundlegende Voraussetzung für die effiziente Datenanalyse ist ein leistungsfähiges Data Warehouse, beziehungsweise eine geeignete Big-Data-Infrastruktur, die auch große Datenmengen schnell auswertet. Im Rahmen von Predictive Analytics ist es unerlässlich, Daten aus unterschiedlichsten Quellen zu aggregieren und zu analysieren. Sämtliche Datenquellen (Maschinen, Prozesse, Produkte, Wetterdaten etc.) und Systeme (ERP, CRM, BI etc.) müssen miteinander integriert werden. Eine weitere Herausforderung ist es, die große Vielfalt von Datentypen und Formaten sowie die Qualität der einzelnen Daten umfassend zu verstehen und daraus entsprechende Einsichten zu gewinnen. Die Daten sollten natürlich alle Kriterien für hohe Qualität erfüllen, wie etwa Korrektheit, Konsistenz,

Vollständigkeit, Aktualität oder Einheitlichkeit. Insbesondere für Predictive Analytics müssen genügend historische Daten vorhanden sein, damit sich diese mit geeigneten Prognose-techniken aussagekräftig in die Zukunft fortschreiben lassen.

Es gibt viele unterschiedliche Methoden und Vorhersagemodelle, um Prognosen aus Daten abzuleiten. Sie reichen von klassischen Data-Mining-Methoden, wie Clustering oder Regressionsanalyse, über Elemente der Spieltheorie bis hin zum maschinellen Lernen. Bei letzterem werden die Algorithmen so trainiert, dass sie aus den vorliegenden Daten lernen, selbstständig ein Datenmodell erzeugen und dieses für Prognosen oder Entscheidungen einsetzen.

SCHRITT FÜR SCHRITT

Angesichts der Komplexität der Predictive-Analytics-Projekte mit vielen Variablen und Einflussfaktoren raten beide Experten Unternehmen, schrittweise vorzugehen und mit einem kleinen Pilotprojekt zu starten, das in einem laufenden Prozess ständig weiterentwickelt wird.

Dazu Milnik: »Es besteht immer das Risiko, dass die Vorhersagemodelle nicht das erwartete Ergebnis bringen. Das kann unterschiedlichste Ursachen haben. Beispielsweise kann es an der Datenqualität mangeln. Daher ist es äußerst wichtig, gemeinsam mit Experten bei den Anwendern die Projekte schrittweise zu erarbeiten und auch die Datenquellen sukzessive zu erweitern.«

Es gibt nicht die eine richtige Predictive-Analytics-Methode für alle Unternehmen. Welche Methode und welcher Algorithmus am besten funktionieren, hängt vom individuellen Ziel und der Fragestellung ab. Predictive Analytics ist zudem als immerwährender Prozess zu sehen mit dem Test und der Kombination verschiedener Datensätze und Vorhersagemodelle. Da sich die Modelle im Laufe der Zeit immer weiter verbessern, werden auch die Vorhersagen immer präziser. Unternehmen sollten daher mit einem kleineren Projekt starten und die Lösung dann Schritt für Schritt erweitern. ♦