

# Die Arena betreten



## *In diesem Kapitel*

- ▶ Die Grundbausteine erklären
  - ▶ Die Möglichkeiten erkunden
  - ▶ Den Markt überblicken
- 

**P**

redictive Analytics ist ein von Ihren Daten erzeugter heller Lichtblick.

Sie können niemals zu viele Erkenntnisse gewinnen. Je mehr Sie wissen, desto besser sind die von Ihnen getroffenen Entscheidungen – und Sie wollen doch niemals im Dunkeln stehen. Sie wollen am liebsten vor allen anderen wissen, was kommt. Es ist genauso wie bei der Show »Geh aufs Ganze!«, bei der Sie die Tür mit dem versteckten Preis auswählen müssen. Welche Tür wählen Sie? Tür 1, Tür 2 oder Tür 3? Alle sehen gleich aus, daher hängt alles von Ihrer Einschätzung – und von Ihrem Glück ab. Doch was ist, wenn Sie im Vorteil sind – wenn Sie beispielsweise die Möglichkeit haben, durch das Schlüsselloch zu schauen? Predictive Analytics kann Ihnen diese Möglichkeit bieten.

## *Predictive Analytics erklären*

Was würden Sie in einer Welt machen, in der Sie wissen, wie wahrscheinlich es ist, dass Sie irgendwann einmal Ihren Mitbewohner aus Studienzeiten heiraten werden? In der Sie vorher-sagen können, welcher Beruf zu Ihnen passt? In der Sie vorhersagen können, welche Stadt und welches Land für Ihr Leben am besten geeignet ist?

Um es kurz zu machen, stellen Sie sich eine Welt vor, in der Sie die Möglichkeiten in jedem Moment Ihres Lebens maximieren können. Ein derartiges Leben wäre produktiv, effizient und energiegeladen. Sie hätten (in der Tat) Superkräfte – und viel mehr Freizeit. Natürlich kann eine solche Welt Leuten, die unkalkulierbare Risiken lieben, etwas langweilig erscheinen. Aber ein Unternehmen, das Gewinn erwirtschaften muss, würde das sicherlich nicht so sehen. Unternehmen geben Millionen von Euro für das Risikomanagement aus. Und wenn es etwas gibt, das Ihnen hilft, das Risiko zu lenken, Ihre Aufgaben zu optimieren und Ihren Gewinn zu maximieren, dann sollten Sie das unbedingt lernen. Das ist die Welt von Predictive Analytics.

## *Data-Mining*

Big Data ist die neue Wirklichkeit. In der Tat werden die Daten immer umfangreicher, schneller und ergiebiger. Das ist die Zukunft, und Sie sollten Nutzen daraus ziehen.

Daten sind eins der wertvollsten Besitztümer Ihres Unternehmens. Sie besitzen einen großen, verborgenen Wert, den Sie erst erkennen müssen. *Data-Mining* bezeichnet das Erkennen von

in Daten verborgenen Mustern durch maschinelles Lernen – und ausgeklügelte Algorithmen sind die dazu notwendigen Werkzeuge. *Predictive Analytics* ist der Prozess, bei dem mithilfe von Geschäftskennnissen der verborgene Wert aus diesen neu entdeckten Mustern in den Quelldaten extrahiert wird.

*Data-Mining + Geschäftskennnisse = Predictive Analytics → Nutzen*

Führende Unternehmen betrachten heutzutage ihre Daten, untersuchen und verarbeiten sie, um Wege zu finden, ihre Kundschaft besser zu verstehen, ihre Tätigkeiten zu optimieren, ihre Konkurrenten zu übertreffen und sich besser auf dem Markt zu positionieren. Sie untersuchen, wie sie diese Informationen zur Steigerung ihres Marktanteils und ihrer Wettbewerbsfähigkeit nutzen können. Wie können sie den Verkauf steuern und Werbefeldzüge zielgerichteter führen? Wie können sie ihre Kundschaft besser bedienen und deren Bedürfnisse besser befriedigen? Was können sie zur Steigerung des Nettoprofits unternehmen?

Aber diese Instrumente sind auch in Bereichen jenseits des Geschäftslebens nützlich. Ein wichtiges Beispiel ist die Arbeit der Justizbehörden in Bezug auf das Aufdecken und Verhindern von Verbrechen. Welche Person verhält sich auffällig? Bereitet sich diese Person darauf vor, ein schreckliches Verbrechen zu begehen? Wird dieser Kriminelle ein Wiederholungs-täter werden? Wo wird das nächste Verbrechen stattfinden?

Andere Industrien, insbesondere im Bereich des Finanzwesens, könnten einen zuverlässigen Blick in die Zukunft werfen. Unternehmen versuchen alles daran zu setzen herauszufinden, ob der gerade bearbeitete Geschäftsvorgang betrügerisch ist, ob ein Versicherungsanspruch rechtmäßig ist, ob ein Einkauf mit Kreditkarte gedeckt ist, ob ein Kreditantragssteller kreditwürdig ist und so weiter ... Diese Liste kann beliebig fortgesetzt werden.

Behörden, Unternehmen und Einzelpersonen versuchen (je nach Zielsetzung), Trends in sozialen Bewegungen zu entdecken, aufkommende Gesundheitsfragen und Ausbrüche von Krankheiten zu erkennen, neue Modetrends aufzuspüren oder den perfekten Lebenspartner zu finden.

Diese – und viele weitere – Geschäfts- und Forschungsfragen sind Themen, die Sie weiter untersuchen können, um durch das Auswerten der verfügbaren Daten und das Erstellen von auf die Zukunft gerichteten Modellen Antworten zu finden, die zukünftige Entscheidungen stützen.

*Daten + Predictive Analysis = Licht*

### ***Das Modell in den Mittelpunkt stellen***

Ein Modell ist eine mathematische Darstellung eines Objekts oder Prozesses. Wir erstellen Modelle, um als weiteren Untersuchungsschritt Phänomene aus der realen Welt zu simulieren, in der Hoffnung, besser zu verstehen, was wirklich vorgeht. Um beispielsweise das Verhalten unserer Kunden zu modellieren, versuchen wir herauszufinden, wie sich unsere Kunden auf unseren Webseiten bewegt haben:

- ✓ Welche Produkte haben sie angeschaut, bevor sie einen Einkauf getätigt haben?
- ✓ Welche Seiten haben sie angeschaut, bevor sie diesen Einkauf machten?

- ✓ Haben sie sich die Beschreibungen der Produkte angesehen?
- ✓ Haben sie die Besprechungen anderer Nutzer gelesen?
- ✓ Wie viele Besprechungen haben sie gelesen?
- ✓ Haben sie sowohl positive als auch negative Besprechungen gelesen?
- ✓ Haben sie noch etwas anderes als das Produkt gekauft, über das sie sich informiert haben?

Wir sammeln all diese Daten über vergangene Ereignisse. Wir betrachten diese historischen Geschäftsvorgänge zwischen unserem Unternehmen und unseren Kunden – und versuchen, ihre Bedeutung zu verstehen. Wir überprüfen die Daten und schauen, ob sie Antworten auf unsere Fragen enthalten. Das Sammeln dieser Daten – unter besonderer Berücksichtigung ihres Umfangs und ihrer Tiefe, ihrer Qualität und ihrem vorhersagendem Wert – hilft, die Grenzen unseres Modells und seiner Ergebnisse zu bestimmen.

Dieser Prozess ist nicht mit der Angabe der Daten zu verwechseln; er unterscheidet sich auch von der reinen Visualisierung der Daten. Obwohl diese Schritte unverzichtbar sind, sind sie bei der Erforschung der Daten und dem Gewinn von nutzbringenden Erkenntnissen nur der Anfang.

Wir gehen viel weiter, wenn wir über die Entwicklung der Predictive Analytics reden. Im Grunde benötigen wir einen dreifachen Ansatz:

- ✓ Ein gründliches Verständnis der Geschäftsprobleme, die wir versuchen zu lösen
- ✓ Die Gewinnung und Aufbereitung der Daten, mit denen unser Modell arbeiten soll
- ✓ Der Einsatz statistischer Analyseverfahren, des Data-Mining und von Algorithmen des maschinellen Lernens

Bei diesem Prozess müssen wir verschiedene *Attribute* betrachten – Datenpunkte, von denen wir denken, dass sie für unsere Analyse relevant sind. Wir werden etliche *Algorithmen* einsetzen, das heißt, Sätze von mathematischen Anweisungen, die Maschinen zur Lösung von Aufgaben veranlassen.

Wir wenden sie auf verschiedene mögliche Kombinationen von Daten an und untersuchen Was-wäre-wenn-Szenarios. Letztendlich finden wir unsere Antworten, erstellen unser Modell, bereiten seine Anwendung vor und fahren den Gewinn ein.

Wie sieht ein Modell aus? Im Programmierstil ausgedrückt, kann ein Predictive Analytics-Modell aus einigen einfachen Anweisungen `if . . . then` bestehen, die der Maschine sagen: »Wenn diese Bedingung erfüllt ist, dann führe das aus.«

Hier folgen zunächst einige einfache regelbasierte Handelsmodelle:

- ✓ Wenn es später als 10 Uhr ist und die Kurse steigen, dann kaufe 100 Wertpapiere der Sorte XYZ.
- ✓ Wenn meine Aktie um 10 Prozent steigt, dann nimm den Gewinn mit.
- ✓ Wenn mein Depot um 10 Prozent fällt, dann verkaufe meinen Bestand.

Anschließend folgen einige einfache regelbasierte Empfehlungssysteme (in Kapitel 2 finden Sie weitere Erläuterungen zu Empfehlungsdiensten):

- ✓ Wenn jemand ein Buch von einem Autor kauft, dann empfehle andere Bücher vom gleichen Autor.
- ✓ Wenn jemand ein Buch zu diesem Thema kauft, dann empfehle andere Bücher zum gleichen und zu verwandten Themen.
- ✓ Wenn jemand ein Buch zu diesem Thema kauft, dann empfehle Bücher, die andere Kunden beim Erwerb dieses Buchs gekauft haben.

## ***Den Geschäftswert erhöhen***

In einer zunehmend konkurrierenden Welt müssen Unternehmen Wege finden, um wettbewerbsfähiger zu werden. Predictive Analytics wird in diesen Unternehmen als eines der dazu erforderlichen Instrumente eingesetzt. Durch den Einsatz von Technologien wie Algorithmen des maschinellen Lernens, der Statistik und Data-Mining-Verfahren können sie versteckte Muster und Trends in ihren Daten erkennen, die Arbeitsabläufe und Strategien unterstützen sowie helfen können, entscheidende Geschäftsbedürfnisse zu erfüllen.

Die Einbettung der Predictive Analytics in Entscheidungsfindungsprozesse verbessert die Kapitalrendite, da die Unternehmen weniger Zeit für weniger wichtige, nur mit geringen Risiken verbundene Entscheidungen aufwenden müssen. Die Mitarbeiter können mehr Zeit für schwierige, mit hohen Risiken verbundene Entscheidungen aufbringen. So können beispielsweise die meisten üblichen Versicherungsansprüche automatisch ausgezahlt werden. Wenn das voraussagende Modell jedoch auf einen ungewöhnlichen Antrag (einen Sonderfall) stößt oder wenn der Antrag das gleiche Muster wie ein betrügerischer Antrag aufweist, kann das System den Antrag automatisch kennzeichnen und an die zuständige Person liefern, die den Fall dann bearbeitet.

Durch den Einsatz von Predictive Analytics zur Vorhersage von zukünftigen Ereignissen oder Trends kann das Unternehmen eine Strategie entwickeln, um aus diesen Erkenntnissen Nutzen zu ziehen. Wenn Ihr Vorhersagemodell Ihnen (zum Beispiel) sagt, dass der Modetrend in Richtung schwarzer Rollkragenpullover geht, können Sie geeignete Maßnahmen zum Design von mehr schwarzen Rollkragenpullovern oder mehr passenden Accessoires ergreifen.

## ***Unendliche Möglichkeiten***

Unternehmen in der ganzen Welt streben Verbesserung, Wettbewerbsfähigkeit und Verschlan-  
kung an. Sie versuchen, den Prozess der Planung beweglicher zu machen. Sie untersuchen, wie man den Warenbestand verwaltet, und wie man die Aufteilung der menschlichen Ressourcen vorteilhaft optimiert bekommt. Sie überprüfen, wie sie in bestimmten Fällen handeln müssen, wenn diese wirklich eintreten.

Predictive Analytics kann helfen, all diese Ziele zu erreichen. Die Bereiche, in denen Predictive Analytics eingesetzt werden kann, sind unbegrenzt. Das Feld ist weit geöffnet und das Spiel freigegeben. Beginnen Sie das Graben. Starten Sie die Untersuchung.

Gehen Sie zu Ihrem Analyseteam und lassen Sie die von Ihnen gesammelten oder gewonnenen Daten daraufhin untersuchen, eine vorteilhafte Marktlücke für Ihr Produkt zu finden. Führen Sie mithilfe von Daten Neuerungen ein. Bitten Sie Ihr Team, Ihnen zu helfen, Vertrauen zu Ihrer Entscheidungsfindung und Ihrem Risikomanagement zu gewinnen.

Albert Einstein sagte einst: »Zu wissen, wo man Informationen erhält und wie man sie nutzt, ist das Geheimnis des Erfolgs.« Wenn das das Geheimnis des Erfolgs ist, dann werden Sie mithilfe von Predictive Analytics erfolgreich sein: Die Information steckt in Ihren Daten, und mithilfe des Data-Mining werden Sie sie finden. Der Rest der Gleichung beruht auf Ihren Geschäftskennnissen und der Interpretation der Informationen – und letztendlich ihrer Ausnutzung, um Erfolg zu haben.

Den Wert der Daten zu erkennen, bringt den Erfolg. Demzufolge können wir unsere Gleichung mit Predictive Analytics wie folgt umschreiben:

*Data-Mining + Geschäftskennnisse = Predictive Analytics → Erfolg*

### ***Ihr Unternehmen stärken***

Predictive Analytics stärkt Ihr Unternehmen, da es drei Vorteile bietet:

- ✓ Vision
- ✓ Entscheidung
- ✓ Präzision

### ***Vision***

Mithilfe von Predictive Analytics werden Sie entdecken können, was für andere unsichtbar ist – insbesondere nützliche Muster in Ihren Daten.

Predictive Analytics kann Ihnen äußerst wichtige Hinweise geben wie Sie Kunden behalten, neue Kunden gewinnen und Profite maximieren. Predictive Analytics kann große Mengen an früheren Kundendaten bearbeiten, diese mit anderen Daten in Verbindung bringen und am Ende alle Teile in der richtigen Reihenfolge zusammenfügen, um das Puzzle zu lösen. Dazu zählen:

- ✓ Die Kunden einteilen und über ihre Bedürfnisse Vermutungen anstellen
- ✓ Die Wunschlisten Ihrer Kunden kennen
- ✓ Die nächsten Handlungen Ihrer Kunden vorhersagen
- ✓ Ihre Kunden in Stammkunden, saisonabhängige und gelegentliche Kunden einteilen

Diese Art von Information vorab zu kennen, beeinflusst Ihre strategische Planung und hilft Ihnen, Ressourcen sinnvoll einzusetzen, die Zufriedenheit der Kunden zu steigern und den Gewinn zu maximieren.

### **Entscheidung**

Ein gut erstelltes Vorhersagemodell liefert analytische Ergebnisse, die frei von Gefühlen und Vorurteilen sind. Das Modell nutzt mathematische Funktionen, um Erkenntnisse aus Zahlen und Texten zu gewinnen, die ältere Daten und aktuelle Informationen enthalten. Das Modell bietet Ihnen konsistente und vorurteilslose Erkenntnisse zur Unterstützung Ihrer Entscheidungen.

Betrachten Sie das Szenario einer typischen Anwendung bei einem Kredit: Der Prozess dauert ein paar Minuten. Die Bank oder die Agentur trifft eine schnelle, auf Fakten beruhende Entscheidung, ob sie den Kredit verlängern sollte, und steht hinterher zu dieser Entscheidung. Die Geschwindigkeit des Geschäftsvorgangs wird durch Predictive Analytics ermöglicht, da es die Kreditwürdigkeit des Antragstellers vorhersagt.

### **Präzision**

Stellen Sie sich vor, Sie müssen eine Vielzahl von Berichten lesen, Schlussfolgerungen aus den darin verborgenen Fakten ziehen, sich zum Vergleichen von Ergebnissen durch etliche Zeilen von Excel-Tabellen arbeiten oder Informationen aus großen Anordnungen von Zahlen extrahieren. Für diese zeitraubenden Aufgaben benötigt man viele Mitarbeiter. Mithilfe von Predictive Analytics benutzen Sie automatische Hilfsmittel für diese Aufgaben, sodass Zeit und Ressourcen gespart, menschliche Fehler verringert und die Präzision verbessert werden.

Sie können beispielsweise zielgerichtete Marketingkampagnen entwickeln, indem Sie die Daten Ihrer Kunden, ihre Demografien und Ihre Einkäufe auswerten. Wenn Sie genau wissen, an welche Kunden Sie sich wenden sollten, können Sie sich auf die konzentrieren, die am wahrscheinlichsten kaufen werden.

## **Ein Predictive Analytics-Projekt starten**

Lassen Sie uns für den Augenblick alles über Algorithmen und höhere Mathematik vergessen. Vorhersagen spielen in allen Lebensbereichen eine Rolle. Überlegen Sie, wie oft Sie schon gesagt haben (oder von anderen gehört haben): »Ich habe Dir doch gesagt, dass das passieren wird.«

Wenn Sie ein zukünftiges Ereignis mit beliebiger Genauigkeit vorhersagen wollen, dann müssen Sie die Vergangenheit und die derzeitige Situation kennen. Dieser Vorgang beinhaltet verschiedene Prozesse:

- ✓ Bestimmen Sie die augenblickliche Situation.
- ✓ Ermitteln Sie den Unterschied zwischen der jetzigen und der direkt vorangegangenen Situation.
- ✓ Schließen Sie auf mögliche Szenarien, die eintreten könnten.
- ✓ Ordnen Sie die Szenarien nach der Wahrscheinlichkeit, mit der sie eintreten können.

Predictive Analytics hilft Ihnen bei jedem dieser Prozesse, sodass Sie so viel wie möglich über das Geschehen wissen und damit kenntnisreichere Entscheidungen in Bezug auf die Zukunft treffen können.

Firmen entwickeln mithilfe von Predictive Analytics Lösungen, die typischerweise aus einer Kombination von drei Bestandteilen bestehen:

- ✓ Unternehmenskenntnisse
- ✓ Data-Science-Team und -Technologie
- ✓ Die Daten

Obwohl das Verhältnis dieser drei Bestandteile je nach Unternehmen variieren kann, sind alle drei für eine erfolgreiche Lösung mithilfe von Predictive Analytics erforderlich, die aufschlussreiche Erkenntnisse liefert.

### ***Unternehmenskenntnisse***

Da jedes Predictive Analytics-Projekt gestartet wird, um ein Bedürfnis eines Unternehmens zu erfüllen, sind unternehmensspezifische Kenntnisse und ein eindeutiges Geschäftsziel ausschlaggebend für den Erfolg. Ideen für ein Projekt können von jedem aus dem Unternehmen stammen, doch es ist die Aufgabe der Führungsmannschaft, die Geschäftsziele festzulegen und die Zustimmung der im Unternehmen zuständigen Abteilungen zu erhalten.

Sorgen Sie dafür, dass die Entscheidungsträger in Ihrem Team bereit sind zu handeln. Wenn Sie einen Prototyp Ihres Projekts vorstellen, benötigt es einen betriebsinternen Verfechter – jemand, der bereit ist, dessen Annahme voranzutreiben.

Die Führungsmannschaft oder entsprechende Fachleute müssen außerdem klare *Metriken* festlegen – Wege, um die Ergebnisse des Projekts zu quantifizieren und zu messen. Die beteiligten Abteilungen müssen geeignete Metriken so festlegen, dass eindeutig ist, was zu tun ist, wie viel zu tun ist, und ob das, was getan wird, dem Unternehmen hilft, seine Geschäftsziele zu erreichen.

Die *Projektbeteiligten im Unternehmen* sind die, die am besten mit dem Geschäftsbereich des Unternehmens vertraut sind. Sie wissen, welche Verknüpfungen – Beziehungen zwischen Merkmalen – von Daten sinnvoll sind und welche nicht, welche Variablen für das Modell wichtig sind, und ob Sie neue Variablen erzeugen sollten – wie bei abgeleiteten Variablen oder Attributen –, um das Modell zu verbessern.

Wirtschaftsanalytiker und andere Experten können die von Maschinen entdeckten Muster analysieren und interpretieren, eine nützliche Bedeutung in den Mustern erkennen und verfolgbare Erkenntnisse gewinnen.

Dies ist ein *iterativer* Prozess (Aufstellen eines Modells und Interpretation der Ergebnisse), an dem das Unternehmen und die Wissenschaften beteiligt sind. Im Verlauf der Modellbildung müssen Sie zur Verbesserung aufeinanderfolgende Versionen des Modells erstellen (das bezeichnen Datenexperten als *Iteration des Modells über seine Lebensdauer*). Möglicherweise müssen Sie viele Berichtigungen und Überarbeitungen Ihres Modells vornehmen, bevor Sie nachweisen können, dass Ihr Modell für das Unternehmen von großem Nutzen ist. Auch nach

dem Einsatz der Modelle muss das Unternehmen die Ergebnisse kontrollieren, die Genauigkeit der Modelle bewerten und sie verbessern, wenn weitere Daten gesammelt werden.

### ***Data-Science-Team und -Technologie***

Die in Predictive Analytics eingesetzte Technologie umfasst mindestens einige (wenn nicht alle) dieser Fertigkeiten:

- ✓ Data-Mining
- ✓ Statistik
- ✓ Algorithmen des maschinellen Lernens
- ✓ Software-Tools zum Erstellen des Modells

Die Leute im Unternehmen müssen die Einzelheiten der eingesetzten Technologie und der verwendeten Mathematik nicht verstehen – doch sie sollten den durch das Modell beschriebenen Prozess gut kennen und wissen, in welcher Beziehung er zur Infrastruktur des Unternehmens steht. Vergessen Sie nicht, dass es ein gemeinsam durchgeführter Prozess ist; die Datenwissenschaftler und die Mitarbeiter des Unternehmens müssen beim Erstellen des Modells eng zusammenarbeiten.

Wenn man andererseits den Datenwissenschaftlern ein gutes Verständnis der Unternehmenskenntnisse ermöglicht, dann gibt man ihnen die Chance, ein genaues Vorhersagemodell zu erstellen, und hilft ihnen, das Modell schneller anzuwenden. Anschließend kann das Unternehmen beginnen, die Ergebnisse sofort auszuwerten – und die Teams können beginnen, an der Verbesserung des Modells zu arbeiten. Durch Tests lernen die Teams, was geht und was nicht.

Die Verknüpfung von Unternehmenskenntnissen mit der Untersuchung der Daten und der Technologie führt zu einem erfolgreichen Einsatz des Vorhersagemodells. Somit besteht die allgemeine Vorgehensweise darin, das Modell durch sukzessive Versionen zu entwickeln und sicherzustellen, dass die Teammitglieder genug Wissen sowohl im Unternehmensbereich als auch in den Datenwissenschaften besitzen, sodass alle auf dem gleichen Wissenslevel sind.

Einige analytische Tools – spezielle Softwareprodukte – sind so fortgeschritten, dass für deren Anwendung Fachleute mit einem naturwissenschaftlichen Hintergrund benötigt werden. Andere sind so einfach, dass sie von jedem Mitarbeiter des Unternehmens benutzt werden können. Die Auswahl der richtigen Tools ist ebenfalls eine sehr sorgfältig zu treffende Entscheidung. Jedes Unternehmen hat andere Bedürfnisse, und kein Tool kann sie alle erfüllen. Doch eine Sache ist sicher; jedes Unternehmen wird gewisse Tools für Predictive Analytics einsetzen müssen.

Bei der Auswahl der richtigen Softwareprodukte für eine bestimmte Aufgabe spielen folgende Faktoren eine Rolle:

- ✓ Die Kosten des Produkts
- ✓ Die Komplexität des Geschäftsproblems
- ✓ Die Komplexität der Daten



- ✓ Die Quelle(n) der Daten
- ✓ Die Geschwindigkeit der Daten (die Geschwindigkeit, mit der sich die Daten ändern)
- ✓ Die Mitarbeiter innerhalb des Unternehmens, die das Produkt verwenden

## ***Die Daten***

Unter sonst gleichen Umständen erwarten Sie, dass eine Person mit mehr Erfahrung eine Tätigkeit besser ausübt, als eine weniger erfahrene Person. Das Gleiche gilt auch für ein Unternehmen. Wenn Sie sich ein Unternehmen als eine Person vorstellen, können Sie die Daten der Firma als ihr Äquivalent an Erfahrung betrachten. Indem Sie diese Erfahrung nutzen, können Sie kenntnisreichere Geschäftsentscheidungen treffen und mit größerer Effizienz agieren. Genau dies ist der Prozess der Verwandlung von Daten in betriebswirtschaftlichen Nutzen mithilfe von Predictive Analytics.

Es ist zunehmend offensichtlich, dass Daten eine wesentliche Rolle spielen, wenn es darum geht, den Entscheidungsfindungsprozess zu beschleunigen und schnell realistische Antworten und Erkenntnisse zu gewinnen. Predictive Analytics unterstützt Geschäftsentscheidungen, indem sie Chancen wie kommende Trends, Märkte oder Kunden vor der Konkurrenz erkennt.

Darüber hinaus können Daten in ihrer unbearbeiteten Form Herausforderungen stellen. Sie können über viele Quellen verteilt sein oder aus anderen Gründen durcheinander sein, um sie sofort bearbeiten zu können. Sie können auch Ihre eigenen Daten mit denen anderer verbinden. Daher sollten Sie damit rechnen, dass Ihre Datenwissenschaftler eine beträchtliche Zeit mit dem Untersuchen und Vorbereiten Ihrer Daten für die Analyse verbringen werden. Dieser Prozess der *Datenbereinigung* und *Datenaufbereitung* umfasst das Erkennen von fehlenden Daten, doppelten Aufzeichnungen und Ausreißern, das Erzeugen von abgeleiteten Werten und die Normierung. (Einzelheiten über diese Prozesse erfahren Sie in den Kapiteln 9 und 15.)

Big Data haben ihre eigenen herausfordernden Eigenschaften, zu denen Volumen, Geschwindigkeit und Vielfalt gehören: In der Tat kommen zu viele Daten zu schnell von zu vielen Orten in zu vielen verschiedenen Formaten. Daher besteht eine der wichtigsten Aufgaben darin, die relevanten Daten von den unwichtigen zu trennen.

In diesem Fall muss Ihr Team den Status und den Typ der Daten überprüfen und die für diese Art von Daten geeigneten Algorithmen auswählen. Solche Entscheidungen sind Teil einer Untersuchungsphase, in der die Datenwissenschaftler detailliertes Wissen über Ihre Daten erwerben, während sie auswählen, welche Attribute die meiste Vorhersagekraft besitzen.

## ***Den Markt begutachten***

Big Data und Predictive Analytics bringen große Veränderungen gleichermaßen für die akademische Welt, den Arbeitsmarkt und nahezu jedes konkurrenzfähige Unternehmen mit sich. Jeder wird die Auswirkungen spüren. Die Überlebenskünstler werden es als große Chance betrachten.

## ***Auf Big Data reagieren***

Zahlreiche Universitäten bieten Kurse und Abschlüsse in Predictive Analytics oder Big Data Analytics an. Einige dieser Studiengänge wurden erst innerhalb des letzten Jahres eingerichtet. Dies spiegelt den erstaunlichen Wachstum und die Popularität dieses Bereichs wider. Bekannte Jobbörsen und Webseiten in Amerika bezeichnen den Beruf »Datenwissenschaftler« als einen der spannendsten überhaupt.

Es ist zu erwarten, dass die Nachfrage nach Datenwissenschaftlern noch weiter steigt. Die Prognose besagt, dass die Zahl der Arbeitsplätze die Anzahl qualifizierter Bewerber übertreffen wird. Einige Universitäten verlagern ihre Angebote, um aus diesem Wachstum Nutzen zu ziehen und mehr Studenten anzulocken. Einige Wirtschaftsfachschulen bieten entsprechende Programme an, während naturwissenschaftliche und ingenieurwissenschaftliche Hochschulen ähnliche Angebote machen. Genau wie die Anwendungen von Predictive Analytics und Big Data in der realen Welt erstreckt sich die Disziplin, die von diesen Methoden Gebrauch macht, über viele Bereiche. Sie finden einschlägige Kursangebote in den Bereichen Wirtschaftswissenschaften, Mathematik, Statistik und Computerwissenschaften. Das Ergebnis ist das gleiche: Mehr einschlägige attraktive Studiengänge für die heutige Wirtschaft und mehr Studenten, die sich für ein im Wachstum befindliches Arbeitsfeld interessieren.

## ***Mit Big Data arbeiten***

Wir lesen jeden Tag Geschichten über den plötzlichen Aufstieg eines tollen neuen Unternehmens, das mithilfe von Predictive Analytics bestimmte Probleme löst, von der Vorhersage, was Sie Tag für Tag so treiben, bis zur Beurteilung, wie geeignet Sie als Freund sind. Ein ziemlicher Wildwuchs. Egal, wie haarsträubend die verschiedenen Geschäftskonzepte auch sein mögen, jemand scheint sie umzusetzen. Leute und Unternehmen machen dies aus einem einfachen Grund: Es gibt einen Markt dafür. Es gibt eine riesengroße Nachfrage nach sozialer Analyse, der Analyse von Menschen und der Analyse jeglicher Daten.

Statistiker und Mathematiker – deren Hauptaufgabe einmal darin bestand, an Schreibtischen zu sitzen und Zahlen für Finanzunternehmen zu verarbeiten – stehen jetzt an der Spitze der Datenrevolution, die verspricht, nahezu alles über nahezu jeden – einschließlich Ihnen – vorzusagen.

Warum erleben wir also diese plötzliche Verlagerung in der Analyse gerade jetzt? Im Grunde genommen gibt es Mathematik, Statistik und deren Teilgebiete, etwa Computerwissenschaften, maschinelles Lernen und Data-Mining seit Jahrzehnten. Die Algorithmen, die heutzutage zur Entwicklung von Vorhersagemodellen eingesetzt werden, wurden ebenfalls vor Jahrzehnten entwickelt. Die Antwort lautet schlicht »Daten« – Unmengen von Daten.

Wir erfassen und erzeugen jeden Tag riesige Datenmengen. Erst in jüngster Zeit sind wir in der Lage, Erkenntnisse aus diesen Daten effektiv zutage zu fördern. Die Rechenleistung und die Datenspeicherung sind exponentiell angewachsen, während gleichzeitig die Geräte schneller und billiger wurden. Wir haben herausgefunden, wie man mithilfe der Hardware große Mengen an Daten speichert und verarbeitet.

Dieses Gebiet, das Computer, Software-Entwicklung, Programmierung und eine gewinnbringende Nutzung des Internets umfasst, schafft ein Umfeld, in dem sich jeder kreativ austoben

kann. Die meisten Menschen in der Welt sind heutzutage durch das World Wide Web, soziale Netzwerke, Smartphones, Tablets, Apps oder was auch immer verbunden. Wir verbringen jeden Tag zahllose Stunden im Internet – und erzeugen dadurch jede Minute neue Daten. Somit ist es nur natürlich, dass Unternehmen diese Daten als eine Quelle von Informationen betrachten, die ausgewertet werden müssen, in denen man nach Mustern in unserem Online-Verhalten sucht und, was man findet, in der Hoffnung ausbeutet, es mithilfe der neuen Möglichkeiten zu kapitalisieren. Amazon (siehe den beigefügten Kasten) ist ein bekanntes Beispiel.

Um es kurz zu fassen, das ist erst der Anfang.

### ***Erstaunliches Wachstum***

Im Verlauf dieses Buchs stellen wir verschiedene Fallstudien vor, die die erfolgreiche Anwendung von Predictive Analytics veranschaulichen. In diesem Kasten betrachten wir die *Crème de la Crème* der Predictive Analytics: Amazon.

Als einer der größten Online-Händler ist Amazon wahrscheinlich eines der bekanntesten Unternehmen, das man mit Predictive Analytics verbindet. Amazon analysiert endlose Ströme von Kundentransaktionen in der Absicht, versteckte Kaufmuster sowie Verbindungen zwischen Produkten, Kunden und Einkäufen zu entdecken. Wenn Sie einem leistungsfähigen Empfehlungsdienst bei der Arbeit zuschauen wollen, dann werden Sie ihn bei Amazon finden. Predictive Analytics ermöglicht es Amazon, genau die Produkte zu empfehlen, die Sie schon immer haben wollten, auch den unglaublichen »Heiligen Gral« – das Produkt, von dem Sie noch gar nicht wussten, dass Sie es haben wollten. Das ist die Stärke der Analyse und der Vorhersagemodelle: Muster in riesigen Datenmengen erkennen.

Um seinen Empfehlungsdienst zu erstellen, verwendet Amazon *Kollaboratives Filtern* (engl. Collaborative Filtering) – einen Algorithmus, der nach Informationen über Nutzer und Produkte sucht. Indem er die augenblicklich im Einkaufswagen des Nutzers befindlichen Artikel sowie die Artikel betrachtet, die der Nutzer in der Vergangenheit gekauft, bewertet und »gelikt« hat – und ihn dann mit dem verlinkt, was andere Kunden gekauft haben –, liefert Amazon den Kunden diese Online-Empfehlungen, mit denen wir alle vertraut sind, wie etwa:

- ✓ Wird oft zusammen gekauft
- ✓ Kunden, die diesen Artikel gekauft haben, kauften auch

Amazon geht aber bei seiner Verwendung der Daten noch weiter: Über das zusätzlich durch Cross-Selling verdiente Geld und die Kaufempfehlung für Kunden hinaus verwendet Amazon die Daten, um Beziehungen zu seinen Kunden aufzubauen – kundenspezifische Ergebnisse, kundenspezifische Webseiten und einen individualisierten Kundendienst. Daten bestimmen jede Stufe der Wechselwirkung zwischen dem Unternehmen und seinen Kunden. Und die Kunden reagieren positiv darauf; Amazons Einnahmen erhöhen sich jedes Vierteljahr.

