



## Kundensegmentierung

*Dr. Patrick Bangert, algorithmica technologies GmbH*

Bei einem großen Europäischen Großhandelsanbieter finden Hoteliers, Gastronomen, Caterer, Kantinen, kleine und große Einzelhändler sowie Dienstleister und Unternehmen aller Art alles, was sie für ihren täglichen Betrieb benötigen. Jeder Kunde hat eine Mitgliedsnummer und -karte, dank derer es dem Großhändler möglich ist, jeden verkauften Gegenstand einem bestimmten Kunden zuzuordnen.

Im Allgemeinen verstehen wir unter Kundensegmentierung, dass Kunden sinnvollen Gruppen zugeordnet werden, etwa nach Branche/Beruf oder nach Kaufverhalten. In unserem speziellen Fall können wir die Kunden in die Gruppen einteilen, weil wir ihre Geschäftsidentität kennen. Diese Art von Erkennungsprogramm wird von vielen Einzelhändlern bei Treueprogrammen eingesetzt, dank derer die Kunden es dem Einzelhändler ermöglichen, ihre Identität dem Einkauf bestimmter Produkte zuzuordnen.

Es ist von besonderem Interesse, das Kaufverhalten einer bestimmten Kundengruppe zu bestimmen. Auf der Basis einer detaillierten Beschreibung der Gruppen und aufgrund einer Untersuchung von Ursache und Wirkung des Verhaltens dieser Gruppen, ist es dem Verkäufer möglich, seine Marketingstrategie den Kundeneigenschaften anzupassen, beispielsweise indem er mit Hilfe zielgerichteter Anzeigen bestimmte Produkte bestimmten Käufern anbietet, deren Kaufverhalten er kennt.



Eine derartige Untersuchung haben wir in zwei Läden über ein Jahr lang an einem Datensatz aller verkauften Produkte durchgeführt. Die Untersuchung erfasste über 31 Millionen Transaktionen. Bei der Untersuchung mussten keine bestimmten Fragen beantwortet und auch keine a priori-Hypothesen verifiziert oder falsifiziert werden. Das Ziel war lediglich herauszufinden, ob man allein aufgrund des Datensatzes irgendwelche Rückschlüsse ziehen könne, die aus Marketingsicht von Interesse sein könnten.

Dabei kamen mehrere Methoden zum Einsatz: (1) deskriptive Statistik, (2) nicht-lineare multidimensionale Regressionsanalysen in allen Dimensionen, (3) K-Means Clusteralgorithmen sowie (4) Markov-Chain-Modellierung. Die Ziele, die wir mit Hilfe dieser Methoden verfolgten, sind in dem folgenden Kasten beschrieben(fett):

1. Descriptive Statistik [1]: Mit dieser Methode will man zunächst ein allgemeines Gefühl für den Datensatz und seine diversen Sektionen erhalten, die von anderen Algorithmen identifiziert wurden: (Diese Methode beinhaltet Korrelationsanalysen. Dabei wird die Markovketten-Methode eingesetzt, was auch die Bayesische (a priori und a posteriori) Verteilungsanalyse beinhaltet, die uns beispielsweise sagen kann, in welcher zeitlichen Reihenfolge sich die Dinge ereignen – was dann zu Ursache-Wirkungs-Aussagen führt).
2. Nonlineare multidimensionale Regression [2]; Diese Methode hat den Zweck zu erkennen, welche Variablen von anderen Variablen abhängig sind. (Variablen durch andere Variablen zum Ausdruck zu bringen kann zu einem schnellen Verständnis führen und die Zahl der unabhängigen Dimensionen reduzieren.)
3. K-Means Clusteranalysen [3]: Hiermit soll herausgefunden werden, welche Käufe/Käufer zu derselben phänomenologischen Gruppe gehören. (um so die tatsächliche Segmentierung zu bestimmen, welche die anderen Methoden beschreiben).
4. Markovketten-Modellierung [3]: Diese hat den Sinn, die zeitabhängige Dynamik des Systems zu modellieren (und so herauszufinden, ob das Kundenverhalten sich über längere Zeiträume stabil verhält oder nicht).

Mehrere Schlussfolgerungen stehen uns nun zur Verfügung, die uns helfen, den Datensatz zu verstehen. Wir präsentieren diese Ergebnisse hier in anschaulicher (nicht-numerischer) Form, da man eigentlich nicht mehr braucht, um das Endergebnis zu verstehen. Bei der eigentlichen Fallstudie, wurden die Ergebnisse allerdings auch numerisch genau beschrieben:

1. Der (durchschnittliche) Geldbetrag, der pro Kundenbesuch ausgegeben wird, ist – statistisch gesehen – für jeden Kunden derselbe. Um also die Gesamteinkünfte zu steigern, kommt es entscheidend darauf an, den Kundenfluss zu verstärken – entweder indem man einen Kunden dazu bringt, öfters zu kommen, oder indem man neue Kunden anlockt.

2. Kunden gehen in der Regel zu dem Laden, der ihrem (Wohn)ort am nächsten ist (in dieser Studie ging es um den Ort, an dem die Einzelhändler(kunden) ihr Geschäfte betreiben, da es sich bei diesem Verkäufer um einen Großhandel dreht). Die Wahrscheinlichkeit, dass ein Kunde einen anderen Laden (als unseren Großhändler) besucht, wächst exponentiell mit der Distanz (zwischen Kundensitz und Großhändler).
3. Die geschäftliche Hochsaison liegt nach den Sommerferien und in der Vorweihnachtszeit. Die kaufschwachen Zeiten sind die Sommerferien und die frühe Nachweihnachtszeit.
4. Das Geld, das insgesamt pro Jahr und pro Besuch ausgegeben wird sowie die Zahl der gekauften Artikel hängen sehr stark vom Kundentyp und der geographischen Region ab. Das wirkt sich stark auf die Lager- und Logistikplanung aus.
5. Die Mehrheit der Kunden besucht den Laden nur selten. Es gibt aber eine Kerngruppe von Kunden, die mehr oder weniger regelmäßig kommt.
6. Die verkauften Artikel und Artikelgruppen hängen stark von regionalen Umständen und von der Besuchshäufigkeit eines Kunden ab.
7. Bestimmte Produkte werden im Allgemeinen in Kombination mit anderen Artikeln gekauft. Wir können deshalb von einer Art „Warenkorb“ sprechen, der als Ganzes erworben wird. Der Inhalt des Korbes hängt aber von der Kundengruppe und von der Geographie ab.
8. Mit Hilfe der Bayesischen Analyse und der Markovketten-Modellierung kann man herleiten, dass der Kauf eines bestimmten Produktes zum Kauf anderer Produkte führt. Beispielsweise führt der Kauf von Frischfleisch fast unweigerlich zum Kauf von Gemüse, Käse und anderen Milchprodukten.

Um die Ergebnisse zusammenzufassen, können wir sagen, dass das Kundenverhalten sehr von der Geographie, von der Produktverfügbarkeit, von der Jahreszeit und von bestimmten Schlüsselartikeln abhängt. Es zeigte sich, dass die folgenden Faktoren ein maßgebliches Potenzial haben, die Gewinnspanne des Handels zu verbessern (die wichtigsten zuerst):

1. Individuelles Marketing [4]: Kunden neigen dazu, sich nur für eine enge Spektrum von Waren zu interessieren. Es ist deshalb sinnvoll, Kunden in Interessengruppen aufzuteilen. Alle Kunden können in weniger als 10 Interessengruppen aufgeteilt werden, wobei davon auszugehen ist, dass diese Gruppen sich deutlich voneinander unterscheiden und dass die Mitglieder der jeweiligen Gruppen ein ähnliches Verhalten an den Tag legen. Diese unterschiedlichen Interessengruppen sollten unterschiedlich behandelt werden, beispielsweise indem man Werbematerial speziell für die jeweilige Gruppe verschickt.
2. Preisgestaltung. Für jede wichtige Produktgruppe gibt es ein bestimmtes Produkt, das als eine Art „Einstiegsprodukt“ fungiert. Das heißt, dass wenn der



Käufer dieses Produkt kauft, er dann auch verschiedene ähnliche Produkte in derselben Kategorie erwirbt. Diesen Ursache-Wirkungs-Mechanismus kann man sich zunutze machen, indem man das Einstiegsprodukt preislich attraktiver gestaltet, um den Verkauf der ganzen Produktgruppe zu steigern. Man könnte etwa den Preis des Einstiegsproduktes niedriger ansetzen (und die Preise der übrigen Produkte dieser Produktkategorie etwas höher gestalten). Es lässt sich zeigen, dass der ursächliche Zusammenhang ungeachtet der Preisgestaltung bestehen bleibt. (Allerdings ist die Identifizierung eines Artikels als Einstiegsprodukt nicht universal gleich gültig, weil es diesbezüglich regionale Unterschiede gibt.)

3. Geographie. Die meisten Anschaffungen werden von Kunden getätigt, deren Geschäftssitz zwischen 20 und 40 Minuten vom Großhandel entfernt ist. Der Großhändler kann seine Marketingbemühungen auf diesen Umkreis fokussieren. Die üblichen Werbemaßnahmen können in dieser Region etwa mit Plakaten entlang der Hauptverkehrsstraßen flankiert werden.
4. Saisonale Schwerpunkte: Die besseren jährlichen Kaufzeiten sind der Monat März, der August (gleich nach Ende den Sommerferien) sowie die Vorweihnachtszeit. Die schlechteren Kaufzeiten sind Januar, Februar und die Sommerferienzeit. Der Rest des Jahres zeigt ein durchschnittliches Kaufverhalten. Werbemaßnahmen sollten diese Trends widerspiegeln, indem man sich auf die Hochsaison(s) konzentriert und diese für sich ausnutzt.

Aus Gründen der Geheimhaltung haben wir diese Schlussfolgerungen nur in allgemeiner Sprache formuliert. Die Vorgehensweise des Data-Mining erlaubt es jedoch, die Ergebnisse auch in Zahlen quantitativ darzustellen (wenn auch mit gewissen Unsicherheitsspielräumen), was eine stabile Grundlage für Unternehmensentscheidungen bietet.

Auch wenn wir hier nicht näher in die Details einsteigen können, stellen die oben vorgetragenen Ergebnisse doch eine aussagekräftige Grundlage für Entscheidungen auf höchster Unternehmensebene dar. Die Ergebnisse zeigen sehr gut die Wirksamkeit des Data-Mining. Wir konnten eine riesige Datenansammlung in eine kleine Zahl von verständlichen Schlussfolgerungen umwandeln, die als Handlungsbasis dienen und dem Unternehmensmanagement zur Entscheidung vorgelegt werden können. Das alles nahm nicht viel Zeit in Anspruch. Überdies kann dieselbe Vorgehensweise automatisch monatlich wiederholt werden, um Veränderungen des Kundenverhaltens zu verfolgen. Ein Vorbehalt bleibt indes: Die Herausforderung besteht letztlich darin, die Ergebnisse des Data-Mining in eine erfolgreiche Handlungsstrategie umzusetzen.

## References

- [1] Mann, P.S. (1995): Introductory Statistics. Wiley.
- [2] Bates, D.M. and Watts, D.G. (1988): Nonlinear Regression Analysis and Its Applications. Wiley.



- [3] Bishop, C.M. (2006): Pattern Recognition and Machine Learning. Springer Verlag.
- [4] Beyering, L. (1987): Individual Marketing. Verlag Moderne Industrie.

