

## Data Mining und Marketing am Beispiel der explorativen Warenkorbanalyse

Von Thomas Reutterer, Michael Hahsler & Kurt Hornik

Techniken des Data Mining stellen für die Marketingforschung und –praxis eine zunehmend bedeutsamere Bereicherung des herkömmlichen Methodenarsenals dar. Mit dem Einsatz solcher primär datengetriebener Analysewerkzeuge wird das Ziel verfolgt, marketingrelevante Informationen aus großen Datenbanken (sog. Data Warehouses) zu extrahieren und für die weitere Entscheidungsvorbereitung in geeigneter Form aufzubereiten. In vielen Situationen des unternehmerischen Routinebetriebs — vor allem dann, wenn möglichst rasch möglichst effektive Entscheidungen zu treffen sind — ist eine datengestützte, „intelligente“ (im Sinne von wissensbasierte) Komplexitätsreduktion für die Beurteilung von Handlungsalternativen von größter Bedeutung. Im vorliegenden Beitrag diskutieren wir weit verbreitete Vorbehalte, aber auch konkrete Berührungspunkte zwischen Data Mining und Marketing.

Der Einsatz ausgewählter Data-Mining-Methoden wird am Beispiel der explorativen Warenkorb- bzw. Sortimentsverbundanalyse für einen Transaktionsdatensatz aus dem Lebensmitteleinzelhandel demonstriert. Zur Anwendung gelangen dabei ausgewählte Techniken aus dem Bereich der klassischen Affinitätsanalyse, ein robustes *K*-Medoid-Clusterverfahren sowie Werkzeuge zur effizienten Generierung und anschließenden Beurteilung von Assoziationsregeln zwischen den im Sortiment enthaltenen Warengruppen. Die Vorgehensweise wird jeweils anhand des mit der Open-Source-Software R unter <http://CRAN.R-project.org/> frei verfügbaren Erweiterungspakets **arules** illustriert. Es zeigt sich, dass bestimmte Data-Mining-Techniken durchaus in der Lage sind, die von traditionelleren Zutritten der explorativen Warenkorbanalyse auferlegten Restriktionen und Einsatzbeschränkungen zu mildern oder komplett zu beseitigen. Damit eröffnen sich auch bis dato meist ungenutzte Verwertungsmöglichkeiten im Rahmen moderner CRM- bzw. Direktmarketing-Konzepte. Zusammenfassend dürften daher insbesondere in Kombination mit statistischen Modellierungszutritten genutzte Data-Mining-Methoden nützliche Erweiterungen des konventionellen Methodenvorrats darstellen.