

# **eBusiness braucht Data Mining**

**Dr. Helge Petersohn**

NHConsult GmbH

Arno-Nitzsche-Str. 43-45

04722 Leipzig

# eBusiness braucht Data Mining

*Mit eBusiness werden Geschäftsprozesse innerhalb der Wertschöpfungskette integriert, verzahnt und weitestgehend dv-gestützt abgewickelt. eProcurement, eCommerce und Customer Relationship Management beschreiben hierfür wesentliche Teilgeschäftsprozesse, die sehr stark datengesteuert ablaufen. Für die Optimierung dieser Prozesse ist es notwendig, in die Entscheidungen zur Prozeßsteuerung aktuelle Information aus den historischen und gegenwärtigen Daten einzubeziehen. Das erfordert, große Datenbestände gezielt auszuwerten, nach Zusammenhängen zu suchen, die Entscheidern bisher verborgen geblieben sind. Data Mining gewinnt eine neue Dimension. Zur qualifizierten Anwendung von Data Mining sind umfangreiche methodische Kenntnisse erforderlich. Dies betrifft Wissen zur Methodenauswahl und Ergebnisevaluation. Ist dies bei Analytikern vorhanden, können - neben einer Unterstützung zur Hypothesengenerierung für bspw. OLAP - ganz gezielt Webdaten, bspw. für Webshops ausgewertet und kundenorientiert verarbeitet werden. Es bieten sich aber auch neue Möglichkeiten der Prozeßanalyse über die Auswertung von Daten aus Prozeßdurchläufen. Einen Schwerpunkt dieses Beitrages bildet die Vorstellung von Methodenklassen und die Anwendung von Data Mining zur Portalanalyse. Data Mining wird immanenter Bestandteil für Decision Support im eBusiness.*

## 1 eBusiness und Daten

### 1.1 eBusiness produziert Daten

eBusiness beschreibt eine auf modernster Informations- und Kommunikationstechnologie (IuK) basierende Organisation und Abwicklung von Geschäftsprozessen. Der Einsatz vernetzter Hard- und Software orientiert gleichermaßen auf die wirtschaftlichen Ziele von Unternehmen und Kunden. Die in der Wertschöpfungskette integrierten Prozesse werden durchgängig dv-gestützt realisiert.

Nutzenpotentiale erschließen sich durch eine kostensparende, effektive und zielgerichtete Kommunikation und Kundenbearbeitung. Möglich wird dies vor allem durch die Verbesserungen der Datensituation.<sup>1</sup>

An Porters<sup>2</sup> strategischen Erfolgsfaktoren - Kostenführerschaft, Differenzierung und Konzentration - richten sich relevante, unternehmerische Entscheidungen aus.

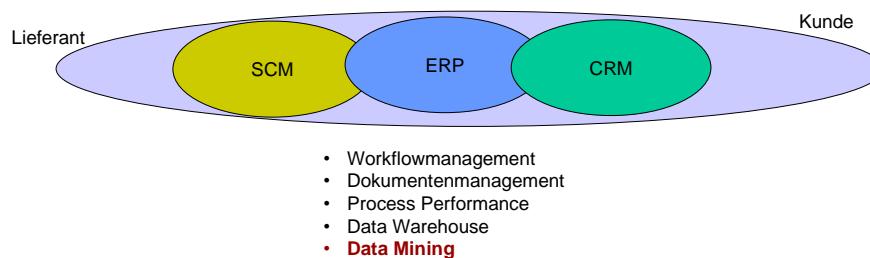
---

<sup>1</sup> vgl. Rebstock, M. (1999), S. 15

Information und Wissen sind Resultate einer zielgerichteten Bereitstellung und Analyse von verfügbaren Daten und führen zu Entscheidungen. eBusiness generiert diese Daten entlang der Wertschöpfungskette in bisher ungekanntem Ausmaß. Dies sind u.a. Bestands-, Verkaufs- und Logdaten, Prozeßzeiten und -kosten. Umfang, Aktualität, Zugang zu den Daten, Senkung der Kosten zur Datenaufbereitung und Möglichkeiten zur Just-in-Time-Datenanalyse bieten optimale Chancen für eine gezielte Informationsversorgung von Entscheidern.

## 1.2 Analytisches CRM verwendet Daten

eBusiness umfaßt die Prozesse vom Lieferanten des Lieferanten bis zum Kunden des Kunden. Auf dem Softwaremarkt haben sich zur Unterstützung der Aufgaben innerhalb dieser Prozesse Produkte etabliert, die unter den Bezeichnungen Supply Chain Management-Software, Enterprise Resource Planning-Software und Customer Relationship-Software geführt werden. Eine für Data Mining besonders relevante Produktgruppe stellt die zur Unterstützung und Verbesserung von Customer Relationship dar. Hier wiederum ist es insbesondere das analytische Customer Relationship Management, welches auf traditionellen und webbasierten Datenbeständen aufsetzt.



**Abbildung 1: Komponenten des eBusiness**

Die hierfür notwendigen und anfallenden Daten sind reif für die Anwendung von anspruchsvollen mathematisch-statistischen Methoden zur Datenanalyse in Größenordnungen. Eine Vielzahl an Methoden einerseits und die Problemvielfalt andererseits stellen hohe Anforderungen an die Analysten. Domänenwissen der einzelnen Wirtschaftsbereiche und das Fachwissen der Datenanalysten müssen koordiniert werden.

---

<sup>2</sup> vgl. Porter, M. (1999), S.71

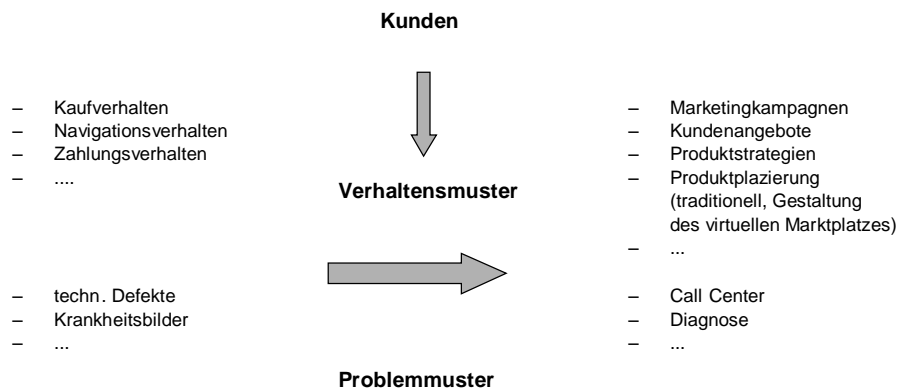
### 1.3 Anwendungsbeispiele

Resultate des Data Mining beeinflussen unternehmensstrategische Entscheidungen und tragen zur Verbesserung der Chancen im Wettbewerb bei.<sup>3</sup> Analysten verknüpfen ihr Wissen mit Domänenwissen. Das ist Voraussetzung für die technisch korrekte Anwendung der Methoden einerseits und die sinnvolle Interpretation der Ergebnisse andererseits. Eine wesentliche Aufgabe beim Data Mining ist die Suche nach Mustern.

Dies verspricht überall dort die Erschließung nützlicher Potentiale, wo viele Daten anfallen, so z.B. in Handel, Banken, Versicherungen, Telekommunikations- und Energieunternehmen oder in der Forschung.

Muster sind einerseits die mit ausreichender Sicherheit existierenden Beziehungen, insbesondere Regelmäßigkeiten zwischen Datensätzen, andererseits Objektbeschreibungen selbst. Als Muster können bspw. Verhaltensmuster oder Problemmuster aufgefaßt werden.

Ist es möglich, diese Muster zu identifizieren und durch Klassifikation aufzubereiten, dann kann der Kunde durch Klassifizierung zugeordnet und zielgerichtet behandelt werden (vgl. Abbildung 2).



**Abbildung 2: Muster für Data Mining mit Fokus auf Kunden**

<sup>3</sup> vgl. Chamoni, P., Budde, C. (1997), S. 1

Mit diesem Wissen lassen sich folgende typische Anwendungsfelder für das Data Mining im eBusiness aufzeigen:

#### *Webshopanalyse/ Portalgestaltung*

Mit Hilfe von Data Mining wird untersucht, wie sich Kunden im Webshop informieren und was diese kaufen.

Praktische Anwendung findet Data Mining in der Marktforschung bei der Auswertung von Meinungsumfragen. Diese Umfragen können sowohl als traditionelles Interview als auch webbasiert erfolgen. Die Vorstellungen potentieller Kunden über Produktgestaltungen werden analysiert. Für Kundengruppen mit gemeinsamen Wünschen und ähnlichem Informationsverhalten können gezielt Marketinginstrumente eingesetzt werden. So kann die Kundenansprache über Bannerwerbung, eMail, Brief, Anzeigen, auf Messen etc. gezielt erfolgen. Diese Erfahrungen, welche in den letzten Jahren in Handel und Marketing gesammelt wurden, lassen sich größtenteils auf den Bereich eCommerce übertragen. Hier stellen neben den traditionellen Aufgaben der Segmentierung auch Fragen der Portalgestaltung einen ganz aktuellen Aspekt dar. Marktplätze mit mehreren hunderttausend Kunden sollten so individuell und rentabel wie möglich angesprochen werden. Dazu ist es notwendig, typische Portale zu entwickeln, um möglichst vielen Kunden das Gefühl der Individualität zu vermitteln. Kundenverhalten läßt sich aus Umfragen, aber auch über die Auswertung von Navigationspfaden im Internet erkennen.

Im Unterschied zu Website-Tracking-Software werden mit Data Mining-Methoden n-dimensionale Clickstreams analysiert. Besuchergruppen und deren Aktionen im Web unterstützen die Profilerstellung von Online-Kunden.

Die Transaktionen werden ausgewertet, indem über alle Kunden der Zusammenhang von Dauer des Aufenthalts eines Kunden auf einer Webseite und dem Kaufbetrag analysiert wird.

#### *Kundenberatung*

Innerhalb des Schlagworts CRM finden sich beinahe branchenneutral hervorragende Anwendungsgebiete für Data Mining. Bspw. greifen Finanz- und Versicherungsdienstleister für die Unterstützung von Außendienstmitarbeitern oder Webportalgestaltern im besonderen Maße auf Analyseresultate aus Data Mining zurück. Während oder zur Vor- und Nachbereitung von Kundengesprächen können Vertriebsmitarbeiter durch Analogieschlüsse kundengruppenindividuelle Angebote offerieren, Portfolios zusammenstellen usw. Möglich wird dies aufgrund einer umfangreichen Kundendatenbasis zur Kundenklassifikation. Dies erleichtert die Klassifizierung von neuen Kunden auf Basis verschiedener Kriterien. Sind darüber hinaus durch webbasierte Kundenberatung über Navigationspfade auch die Informationsbedürfnisse der Kunden bekannt, stellt dies eine ideale Datenbasis für Data Mining zur Kundenbetreuung dar.

### *Prozeßoptimierung*

Über die Analyse von Prozeßinstanzen (Prozeßdurchläufen) kann Information gewonnen werden, die sowohl aus der Klassenbildung als auch der Klassifizierung mit Prognose resultiert. So lassen sich bspw. ähnliche Prozeßinstanzen bspw. zu optimalen, kritischen und durchschnittlichen Prozeßinstanzgruppen zusammenfassen. Über Analogieschlüsse werden dann aus den optimalen Prozessen Handlungsempfehlungen zur Gestaltung der kritischen Prozeßdurchläufe abgeleitet.

Weiterhin eröffnen sich Möglichkeiten, auf Basis vergangener Prozeßinformation frühzeitig die Entwicklung bspw. von Ressourcenverbräuchen zu prognostizieren. Dies erfolgt über die Bestimmung von Ähnlichkeiten zu Klassen von Prozeßinstanzen.

## **2 Vom Data Warehouse zum Data Mining**

### **2.1 Data Warehouse und Data Mining**

In den letzten Jahren hat sich diesbezüglich und auch im Zusammenhang mit dem Data Warehouse-Thema der Data Mining-Begriff etabliert (vgl. Abbildung 3). Daten aus operativen Quellsystemen, Webdaten wie bspw. stützstellenbezogene Prozeßdurchlaufdaten, Logdaten oder andere unternehmensrelevante aber externe Daten wie z.B. regionale Statistiken und Marktstudien werden mit Hilfe von Metadaten zusammengeführt, teilweise aggregiert und verwaltet. Für die Auswertung und Nutzung dieser Daten spielen aus datenanalytischer Sicht Online Analytical Processing (OLAP) und Data Mining eine besondere Rolle.

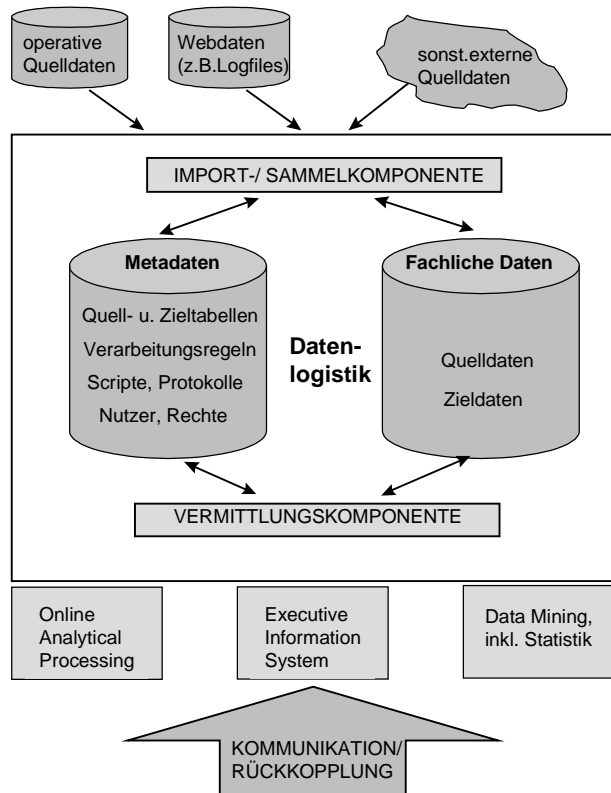


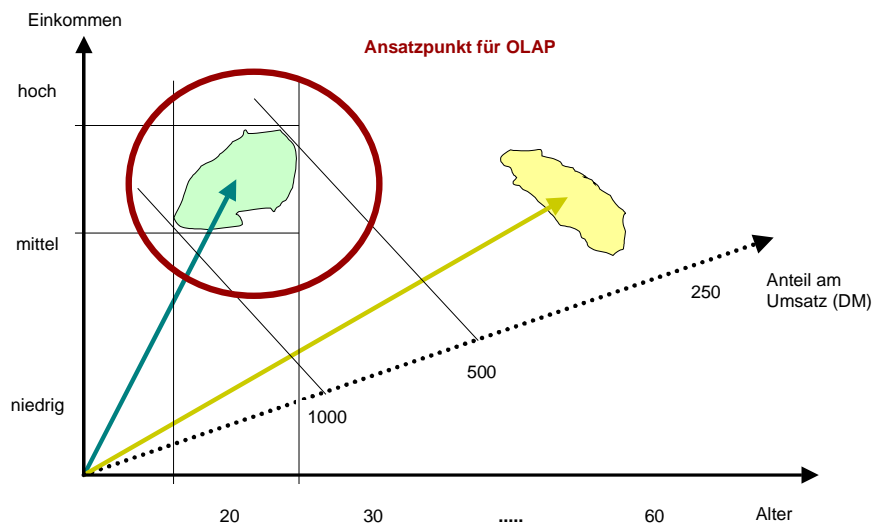
Abbildung 3: Data Warehouseadäquate Datenlogistik und -nutzung

## 2.2 Data Mining und OLAP

OLAP realisiert mehrdimensionale Datenanalysen auf Basis von formulierten Anfragen an die Datenbank. Der Analyst stellt auf Erfahrung beruhende Hypothesen auf, die er versucht mit den Ergebnissen von meist SQL-Statements zu beweisen. Data Mining steht in engem Zusammenhang mit Induktion in der Logik. Induktion in diesem Sinne ist die Ableitung von Information aus Daten. Wesentlich bei der induktiven Analyse ist, daß die entdeckten Zusammenhänge vorher nicht bekannt waren.

Data Mining ist ein Prozeß, mit dem Muster in Datenbeständen erkannt und interpretiert werden können. Innerhalb dieses Prozesses kommen einzeln oder kombiniert Methoden zur Klassenbildung und Klassifizierung (im Spezialfall zur Funktionsapproximation) zum Einsatz. Für die Interpretation werden deskriptive Statistiken erstellt. Die Ergebnisse des Data Mining können Hypothesen derart gene-

rieren, daß sich Zusammenhänge aus Klassifikationen herleiten lassen. Diese Zusammenhänge und die Interpretationsergebnisse können sehr sinnvoll für OLAP bzw. zur Generierung von SQL-Statements genutzt werden. Ein Beispiel hierfür liefert die folgende Abbildung, die zeigt, wie Vektoren typische Klasseneigenschaften repräsentieren (vgl. Abbildung 4). Die Vektorkomponenten werden beim OLAP zur Formulierung von Statements herangezogen.

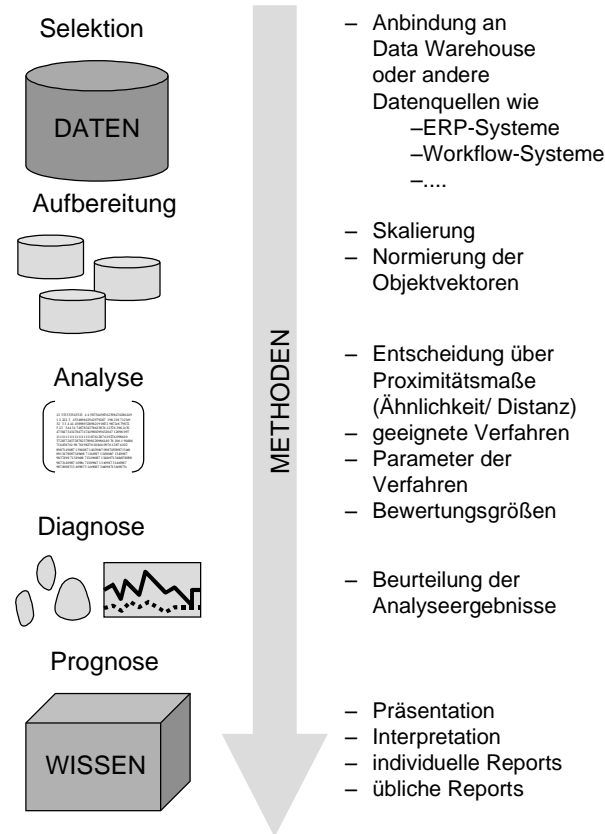


**Abbildung 4: Data Mining findet Klasseneigenschaften für OLAP**

Data Mining ist ein Prozeß, der aus Arbeitsschritten besteht, die auch bei klassischen statistischen Verfahren notwendig sind. Überhaupt müssen Data Mining und Statistik immer im Zusammenhang betrachtet werden. Im Unterschied zur Datenanalyse mit statistischen Verfahren sollten die Arbeitsprozesse weitgehend automatisiert sein. Die Prozeßschritte werden nicht sequentiell durchlaufen, sondern sie weisen Rücksprünge auf. Somit ist Data Mining ein iterativer Prozeß.

Das prinzipielle Vorgehen beim Data Mining ist in Abbildung 5 dargestellt.





**Abbildung 5: Ablauf des Data Mining**

Bei der Selektion werden nur diejenigen Daten ausgewählt, die für die Analyse als sinnvoll und geeignet eingeschätzt werden. Dabei sind Qualität, Struktur und Inhalt der Daten zu betrachten. Im Anschluß daran erfolgt eine Aufbereitung, um Objekte mit falschen oder fehlenden Eintragungen zu erkennen und zu beseitigen. Diese vorverarbeiteten Daten werden einer Transformation (Skalentranslation, Normierung) unterworfen. Dies erfolgt unter Beachtung des angestrebten Analysezziels. Die vorliegenden Daten bilden das Ausgangsmaterial für das eigentliche Data Mining, bei dem die Suche nach Klassifikationsregeln, Entscheidungsbäumen, Clusterungen etc. erfolgt. Eine Vielfalt an Methoden macht die Anwendung von Gütebewertungskriterien erforderlich. Die gewonnenen Erkenntnisse bedürfen anschließend einer Interpretation und Bewertung durch den Benutzer, um festzu-

stellen, ob die Zielsetzung erfüllt ist oder der Vorgang, beginnend bei einem der vorausgehenden Schritte, erneut durchlaufen werden kann und muß.<sup>4</sup>

## 2.3 Data Mining-Methoden

Ausgangspunkt des Data Mining sind Objektbeschreibungen. Objekte werden durch Attribute beschrieben. Attribute sind Merkmale oder Eigenschaften, die unterschiedliches Skalenniveau aufweisen können. Als Objekte können u.a. bezeichnet werden:

- ~ Webshopkunden,
- ~ Clickstreams,
- ~ Warenkörbe,
- ~ Leser von Online-Zeitschriften,
- ~ Nutzer von Finanzdienstleistungsportalen,
- ~ Beschaffungsvorgänge oder
- ~ Prozeßinstanzen.

Werden die Attribute zur Objektbeschreibung als unabhängige Variablen bezeichnet, so sind die Klassenzugehörigkeiten oder Funktionswerte von Objekten die abhängigen Variablen. Das Ziel beim Data Mining besteht in der Ermittlung von Klassenzugehörigkeiten oder in der Modellierung eines funktionalen Zusammenhangs zwischen unabhängigen und abhängigen Variablen.

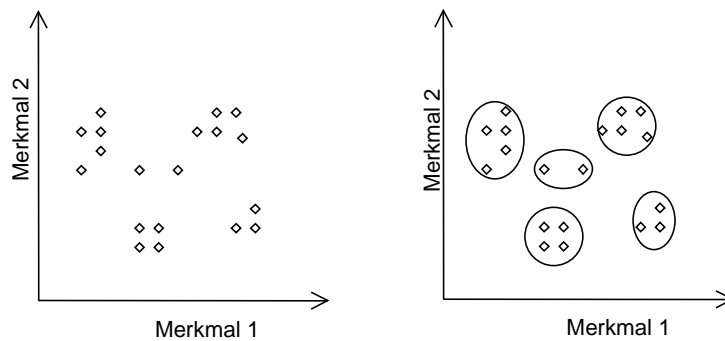
Dazu werden Methoden zur Klassenbildung und Klassifizierung angewandt.

### *Klassenbildung*

Methoden zur Klassenbildung (Clusteranalyse, dazu auch nicht überwacht lernende KNN und Assoziationen) identifizieren Gruppierungen oder Cluster innerhalb der Daten. Es wird angenommen, daß jedes Objekt durch einen Vektor von Meßwerten fester und bekannter Dimension charakterisiert werden kann (vgl. Abbildung 4).

---

<sup>4</sup> vgl. Nimis, J. (1996), S. 8 f



**Abbildung 6: Häufungen von Objekten mit proximativer Merkmalsausprägung**

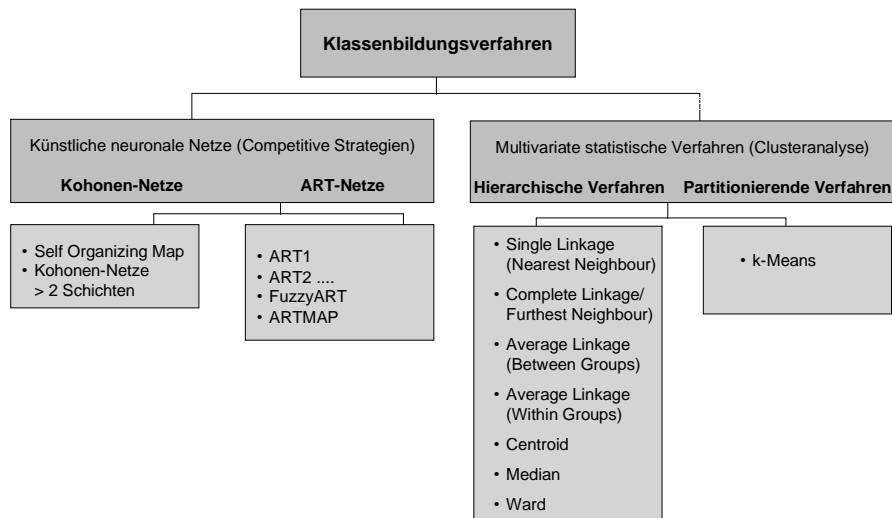
Ausgangspunkt der Clusteranalyse bildet eine vorgegebene konkrete Problemstellung. Diese kann aus dem Informationsbedarf innerhalb eines Entscheidungsproblems resultieren. Es wird beispielsweise Information über Zielgruppen benötigt, um die Werbung für ein neues Produkt zu verbessern. Daraufhin können dann die zu klassifizierenden Objekte ausgewählt werden. Dabei ist insbesondere zwischen Aufwand (Aufwand für Befragung, Verarbeitung) und Nutzen (zusätzliche Information, Sicherheit der Ergebnisse) abzuwägen. Weiterhin ist zu untersuchen, welche Merkmale das Objekt möglichst gut beschreiben, welches Skalenniveau sie aufweisen und welches Proximitätsmaß (Ähnlichkeits- oder Distanzmaß) auf sie anwendbar ist. Anschließend kann das Clusterverfahren ausgewählt werden.<sup>5</sup>

Ist eine Klassifikation gefunden, muß diese inhaltlich interpretierbar sein. Aus diesem Grund wird in der Praxis oft die Anzahl der zu bildenden Klassen vorgegeben.

Zur Ermittlung der Klassen existiert eine ganze Reihe unterschiedlicher Methoden. Abbildung 7 gibt einen Überblick über ausgewählte Klassenbildungsverfahren.

Die hier genannten Verfahren unterscheiden sich vor allem hinsichtlich der Art und Weise, wie Objekte zu einer Klasse zusammengefaßt werden.

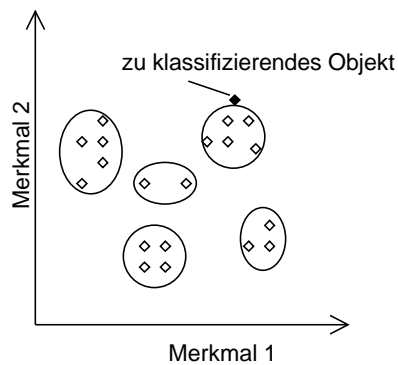
<sup>5</sup> vgl. Petersohn, H. (1997)



**Abbildung 7: Ausgewählte Klassenbildungsverfahren**

*Klassifizierung*

Methoden zur Klassifizierung (Diskriminanzanalyse, Entscheidungsbaumalgorithmen, überwacht lernende KNN) verwenden das Wissen über die Zuordnung von Objekten zu Gruppierungen, Clustern oder Klassen. Dieses Wissen liegt in den Daten vor. Klassifizierungsmethoden entwickeln anhand dieser Daten ein verfahrensspezifisches Modell, welches neue Objekte mit unbekannter Klassenzugehörigkeit bekannten Klassen zuordnet (vgl. Abbildung 7).



**Abbildung 8: Zuordnung eines Objektes zu einer bekannten Klasse**

## 3 Webportalanalyse mit Data Mining

### 3.1 Problemstellung

Für das Data Mining von webbasierten Daten wird auch der Begriff Web Mining verwendet. Web Mining beinhaltet genauer Web Content Mining, Web Structure Mining und Web Usage Mining. Web Content Mining analysiert Internetdaten nach bestimmten inhaltlichen Fragestellungen. Web Structure Mining umfaßt die Untersuchung der Struktur von Webdaten. Unter Web Usage Mining wird die Suche nach Verhaltensmustern der Benutzer von Websites zusammengefaßt. Folgende Fragestellungen stehen im Mittelpunkt:

- ~ Wieviel ernsthafte Besucher hat die Website?
- ~ Wie stark werden die einzelnen Seiten von den Benutzern frequentiert?
- ~ Wie lange informieren sich Besucher auf den von ihnen angeklickten Seiten?
- ~ Gibt es typische Navigationspfade?
- ~ Welche Größenordnungen von Besuchergruppen mit ähnlichem Interesse gibt es? Können diese Besucher, die auf die Website zugreifen, signifikant klassifiziert werden?
- ~ Welche Einstiegs- und Ausstiegspunkte wählten diese?
- ~ Bestehen signifikante Zusammenhänge zwischen dem Besucherverhalten einer Gruppe und deren Bedeutung für das portal anbietende Unternehmen?

Untersucht werden diese Aspekte bspw. bzgl. der Besucher von Websites einer Gesellschaft deren Internetauftritt hauptsächlich als Informations- und Kommunikationsportal gegenüber kapitalsuchenden Unternehmern, Geschäftspartnern, der Presse, potentiellen Investoren und Jobsuchenden dient.

### 3.2 Methodisches Vorgehen

Beim Web Usage Mining werden Logfile-Daten von Zugriffen auf Internetseiten ausgewertet. Der Web Mining-Prozeß umfaßt die typischen Etappen beim Data Mining.

Die Erwartungen an Web Mining umfaßt vor allem unkomplizierte Analysen der Zugriffe auf Websites, Unterstützung bei der Datenaufbereitung (Aufbereitung der Logfiles), eine Auswertungsmöglichkeit des Besucherverhaltens auf Websites und die Klassifikation nach Besucherprofilen. Im Zentrum steht die Suche nach einer Antwort auf die Frage nach Nutzergruppen für die Website der Firma. Diese können sein:

- ~ Unternehmer/ Unternehmensgründer,
- ~ Investoren (konkrete oder potentielle Geldgeber),
- ~ potentielle Mitarbeiter,
- ~ Wettbewerber und
- ~ Sonstiges Publikum.

Aus dem Logfile stehen alle Zugriffe auf die Elemente einer Website zur Verfügung. Je nach Format der Logfiles wird dabei folgende Information aufgezeichnet:

- ~ die URL bzw. die Domain des Providers der Besucher,
- ~ Datum und Zeit des Zugriffs,
- ~ der Name der aufgerufenen Datei,
- ~ einige technische Information (z.B. Anzahl der übermittelten Kilobyte),
- ~ die URL der Website, von welcher der Besucher auf die Seite gelangt ist,
- ~ der von dem Besucher benutzte Browser.

So kann für jeden Besucher ein Vektor mit n Dimensionen gebildet werden. Die Anzahl der Dimensionen entspricht der Anzahl der für die Analyse relevanten Websites. Voraussetzung für eine ordnungsgemäße Analyse ist die eindeutige Identifizierung eines Clickstreams.

Anhand von Gütekriterien wird eine geeignete Anzahl Klassen von Website-Besuchern gefunden, deren Profile interpretierbar sind.

Abbildung 9 zeigt beispielbezogen einen Screenshot des NUGGET HUNTER®. Über Segmentierung mit Competitive Strategien wurden Profile von Besuchergruppen ermittelt. Die Häufungen visualisieren zweidimensional ähnliche Verhaltensweisen von Besuchern in Nachbarschaft. Die schachbrettartige Matrix faßt je nach Anzahl von Elementen auf dieser Besucher zu Klassen zusammen.

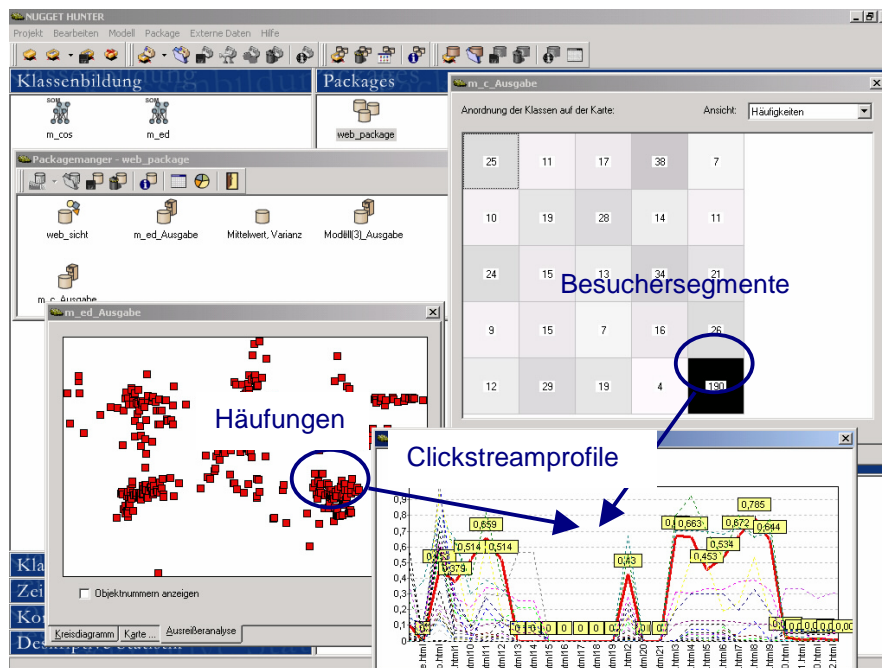


Abbildung 9: Typisierte Clickstreamprofile von Webbesuchern

Für die ausführliche Interpretation von Besuchergruppen und deren typische Eigenschaften mit entsprechendem Clickstream ist weitere Information notwendig. Solche Information gibt bspw. Auskunft über Referrer, Einstiegspunkte, Ausstiegspunkte, Browsertypen und Betriebssysteme (vgl. Abbildung 10).

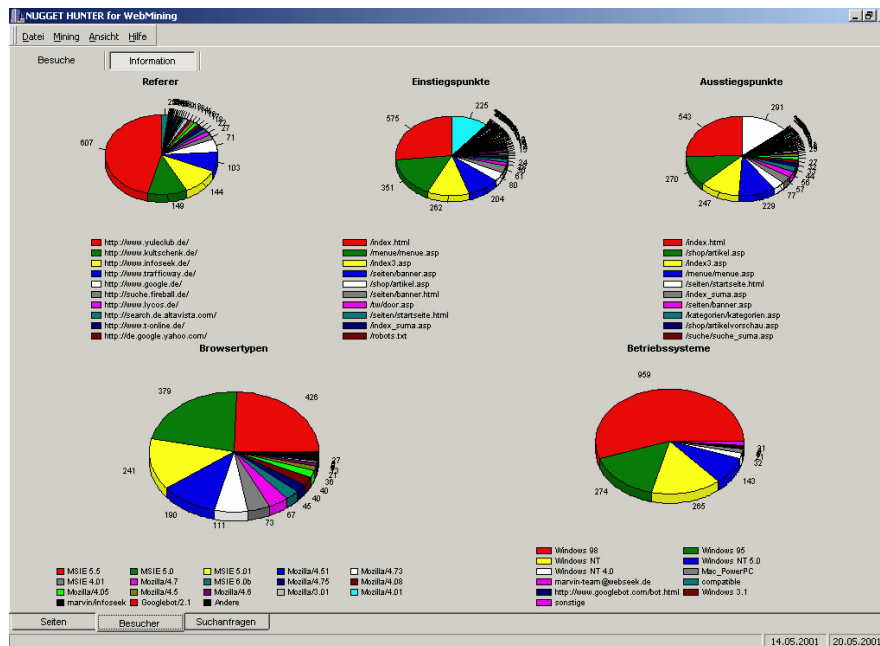


Abbildung 10: Deskriptive Analysen des NUGGET HUNTER® für Web Mining

All diese Information liefert dem Analysten bspw. für ein Relaunch des Webauftritts wertvolle Hinweise. Sind neben dieser Information bspw. bei einem Webshop noch die generierten Umsätze je Clickstream bekannt, so ist der Betreiber des Webshops in der Lage, brauchbare Schlüsse für Produktangebote, Werbungen und Webauftritt zu ziehen.

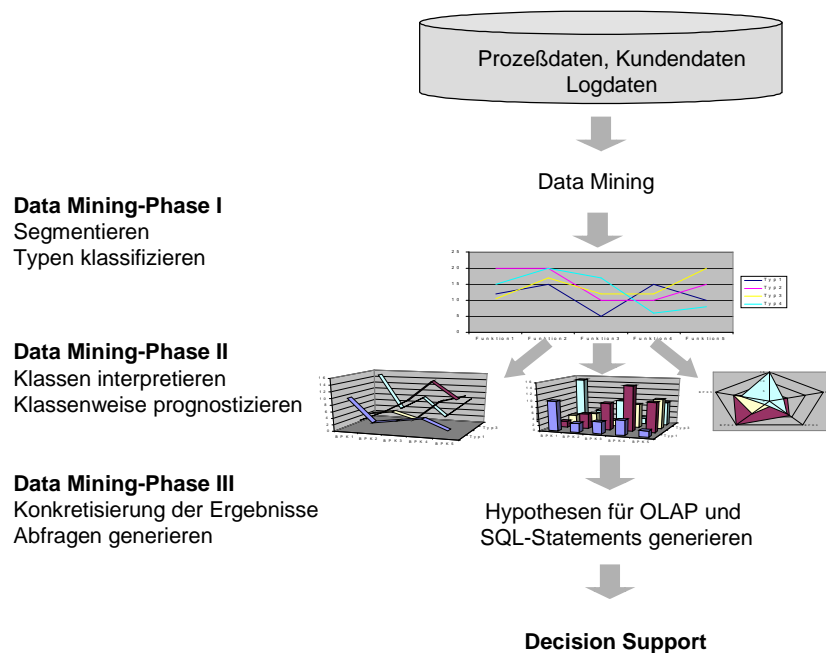
#### 4 Zusammenfassung

Im eBusiness entstehen entlang der Wertschöpfungskette umfangreiche Daten, deren Auswertung mit Data Mining und OLAP notwendig wird. Besondere Relevanz besitzt dies derzeit für die durchgängige Prozeßunterstützung, eProcurement, eCommerce und Webportale. Bspw. werden Prozeßdaten, Kundendaten und Logdaten analysiert.

Data Mining stellt Methoden zur Klassenbildung und Klassifizierung bereit, die für die Auswertung sehr umfangreicher Datenbestände von großem Interesse sind.

Die Daten müssen vor der Analyse bereinigt, transformiert und ggf. normalisiert werden. Es wird nach auffälligen Mustern in den Daten gesucht. Muster mit ähnlichen Eigenschaften werden zu Klassen zusammengefaßt. Die Vielfalt an Methoden erfordert die Anwendung von Gütekriterien zur Evaluierung der Data Mining-Resultate.

Data Mining leistet auch Unterstützung bei der Formulierung von Anfragen an die Datenbank. Über die Klassifikation mittels Data Mining werden Vektorkomponenten als Anhaltspunkte zur exakten mehrdimensionalen Datenanalyse geliefert. Data Mining wird zu einem immanenten Bestandteil entscheidungsunterstützender Systeme (vgl. Abbildung 11).



**Abbildung 11: Entscheidungsunterstützung durch Data Mining**



## Literatur

Chamoni, P., Budde, C. (1997): Methoden und Verfahren des Data Mining, Fachbereich Wirtschaftswissenschaft, Gerhard-Mercator-Universität Gesamthochschule Duisburg, Duisburg 1997.

Nimis, J. (1997): Einführung in die Methoden der Wissensgewinnung, In: Bol, G. et al. (Hrsg.): Wissensgewinnung aus großen Datenbasen, Seminar im Wintersemester 95/96, Universität Karlsruhe, Februar 1996, S.1-18, (<http://theseus.ubka.uni-karlsruhe.de/ira-techreport/1996/1996-08.html>, Okt. '97)

Petersohn, H. (1997): Vergleich von multivariaten statistischen Analyseverfahren und Künstlichen Neuronalen Netzen zur Klassifikation bei Entscheidungsproblemen in der Wirtschaft, Frankfurt am Main, 1997.

Petersohn, H. (1998): Beurteilung von Clusteranalysen und selbstorganisierenden Karten, In: Hippner, H.; Meyer, M.; Wilde, K.D.: Computer-Based Marketing, Göttingen 1998.

Petersohn, H.; Heine, P. (1998): Interdependencies between Data Warehouse Concept, Management Information System and Business Process Model. In: Proceedings of the 6th European Conference on Information Systems, 4.6. - 6.6. 1998, Aix-en-Provence, 1998, S. 1305-1319.

Porter, M. (1999): Wettbewerbsstrategie, Frankfurt, 10. Aufl., 1999.

Rebstock, M. (1999): Das Geschäftsmodell E-Business – Begriff, Anwendungsbereiche, Nutzen, in Rebstock, M./ Hildebrand, K. (Hrsg.): E-Business für Manager, 1. Aufl., Bonn 1999.