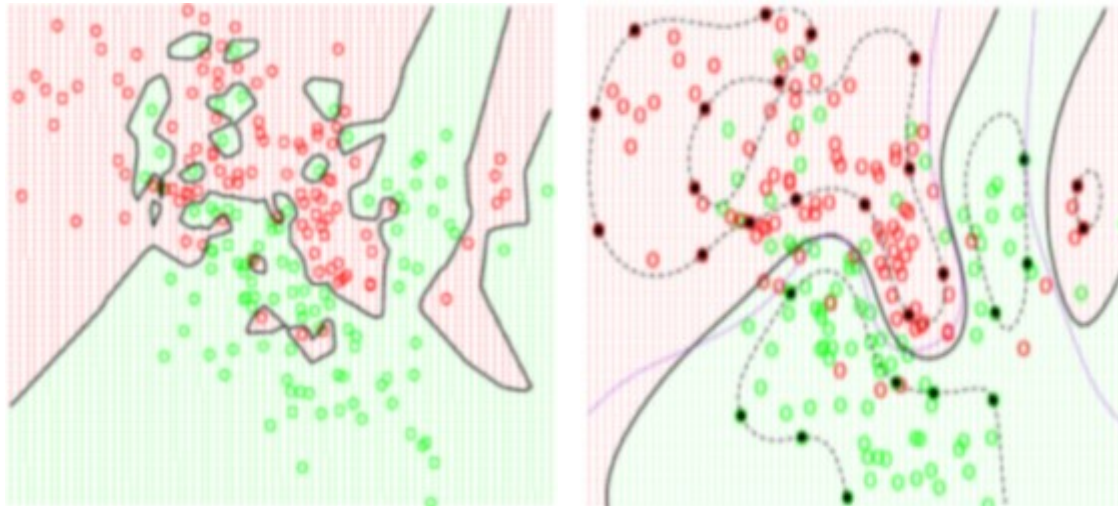


Behavioral Targeting und selbstlernende Kampagnen

Aktuelle Herausforderungen für
Data Mining



Dr. Alexander K. Seewald



Behavioral Targeting

Kognitive Neurowissenschaften

- Verhalten aussagekräftiger als Erklärung
- Viele Erklärungen retrospektiv konstruiert
- Systematische Fehlerquellen bei Umfragen

Interaktive Werbeformen (Online-Marketing)

- Banners & Webseiten-spezifische Werbung
- Handy/SMS/Email-Werbung
- Suchbegriff-Werbung

Behavioral Targeting (2)

Klassisches Online-Marketing

- Soziodemographische Vorselektion (10%)
- Laufende Optimierung der Selektion (90%)
- Abschließende Analyse der Performance

Behavioral Targeting

- Aus Kundenverhalten in Echtzeit lernen
- Automatisierung laufender Optimierung („selbstlernende Kampagnen“)
- Abschließende Analyse gelernter Modelle

MACHINE LEARNING

„The field of machine learning is concerned with the questions of how to construct computer programs that automatically improve with experience.“ (Tom M. Mitchell, 1997)

DATA MINING

„Data Mining is the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data.“
(Fayyad Piatetsky-Shapiro & Smyth, 1996)

Herausforderungen für BT...

.. in Data Mining & Maschinellem Lernen

- 1 Datenrepräsentation
- 2 „Sparse data“
- 3 Inkrementelles Lernen
- 4 Lernen in Echtzeit
- 5 „Responsiveness“
- 6 Rasche Konvergenz
- 7 Verständliche Modelle
- 8 „Predictive Power“ (Performance)

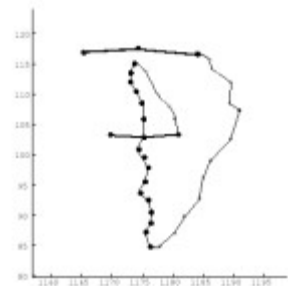
1. Datenrepräsentation

Problemlos verarbeitbar

- Kategorische, Ordinal-, Intervall- und Verhältnisswerte
- Text (einfache Sachverhalte)

Erfordern spezifisches Hintergrundwissen

- Zeitlich/räumlich veränderliche Daten
- (einfache Struktur & Inhalte)
- Sequenzen von einfachen Datentypen
- Text (komplexe Sachverhalte)



State-of-the-art

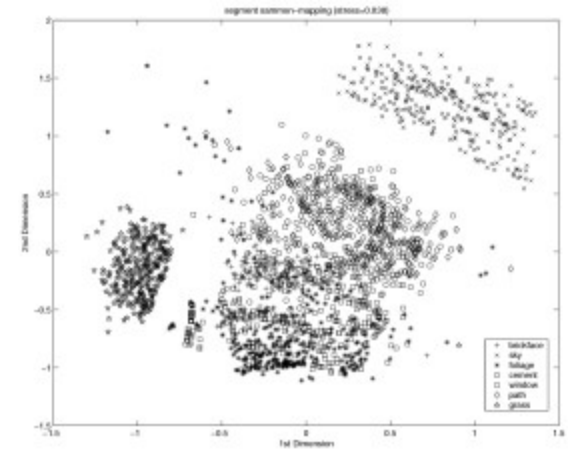
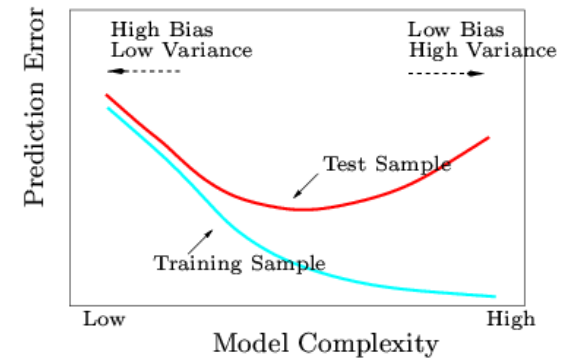
- Audio, Video- und Bilddaten
- Datenfusion
- Sequenzen von komplexen Datentypen

2. „Sparse data“

Dünn besetzte Datenmatrizen

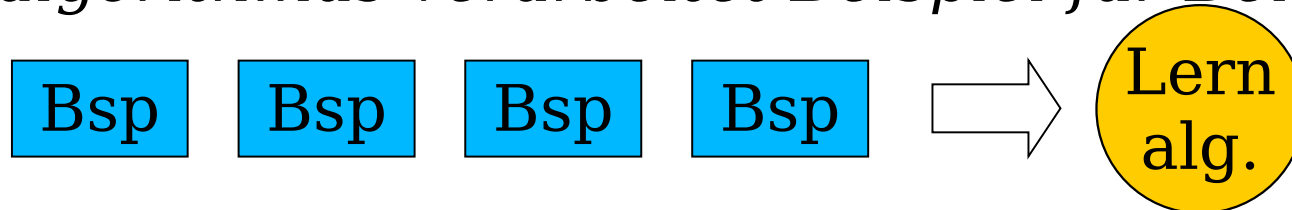
Gängige Lösungen

- Robuste Lernalgorithmen mit hohem Bias (SVM, Logistic, aktuelle Bayes-Algorithmen)
- Integration über längere Zeiträume (aber: langfristige Kundenzuordnung schwierig)
- Kompakte Datenrepräsentationen (zB PCA, MDS, Hintergrundwissen)



3. Inkrementelles Lernen

Lernalgorithmus verarbeitet Beispiel für Beispiel



Gegenteil: Alle Beispiele auf einmal (batch-learn.)

Aktives Forschungsgebiet

- Etliche geeignete Lernalgorithmen verfügbar
- Performance typischerweise schlechter als nicht-inkrementelle Lernalgorithmen
- Konvergenz meist viel langsamer als nicht-inkrementelle Lernalgorithmen
- Meiste Forschung im Bereich *batch-learning*

4. Lernen in Echtzeit

Lernen und Anwendung des Gelernten in Echtzeit

Aktives Forschungsgebiet

- Einige geeignete Lernalgorithmen verfügbar
- Typischerweise ist Anwendung des gelernten Modells viel schneller als Lernen
- Trade-Off Performance vs. Geschwindigkeit

5. „Responsiveness“

Wie schnell wird auf neuen Input reagiert?

Das „stability-plasticity“ Dilemma

- Wenn wir uns jedesmal anpassen, reagieren wir hauptsächlich auf Noise
- Wenn wir uns zu selten verändern, erkennen wir keine kurz- bis mittelfristigen Trends

Lösungsansätze

- Unter „Concept drift“ oder „Context sensitivity“ im inkrementellen Lernen
- Hintergrundwissen hilft meist, ist aber in vielen Domänen sehr aufwendig zu erhalten

6. Rasche Konvergenz

Mit wenigen Beispielen ein brauchbares Modell

Beginnendes Forschungsgebiet

- Inkrementelle Algorithmen konvergieren unterschiedlich schnell – aber noch nicht systematisch untersucht
- Rasche Konvergenz und „Responsiveness“ hängen sehr eng zusammen.
- Problem: Zu rasche Konvergenz kann zu suboptimaler Performance führen
- Ebenfalls lösbar durch Hintergrundwissen

7. Verständliche Modelle

Einsicht in das funktionierende BT System

Lösungsansätze

- Trade-Off Verständlichkeit vs. „Pred. Power“
- Kompetitive Regellerner existieren seit längerem, sind jedoch kaum bekannt
- Stiefkind kommerziell verfügbarer Systeme

```
(blank_ORV >= 42.13) and (PerturnLocal >= 9.441547) => appl=1 (1776.0/479.0)
(Loantermcov <= 0) and (no_accounts >= 3) => appl=1 (1054.0/413.0)
(Numpay6 >= 1) and (days_since_latest_reminder <= 1075) and
    (employed_months <= 197) => appl=1 (752.0/327.0)
(blank_ORV >= 0.45) and (no_accounts >= 2) and (Produktgruppe = CLF2F&Card)
    => appl=1 (103.0/38.0)
=> appl=0 (5652.0/1820.0)
```

8. „Predictive Power“

Wie gut funktioniert das Lernsystem?

Performance eines Lernsystem

- Trade-Off vs. Inkrementelles Lernen, Lernen in Echtzeit, „Responsiveness“, Rasche Konvergenz & Verständliche Modelle
- Noch unklar, welche Lernalgorithmen am besten für BT geeignet
- Kombination von Lernalgorithmen für verschiedene Anforderungen ist möglich

Kein „Silver Bullet“ Lernalgorithmus!

Herausforderung	Lin. & Log.R	SVM	Naïve Bayes	Inst. Based	Dec. Trees	Rule Learn.	Neural Netw.
Datenrepräsentation	-	-	0	-	+	+	-
„Sparse data“	0	0	+	+	0	0	-
Inkrementelles Lernen	+	0	+	+	0	-	+
Lernen in Echtzeit	+	0	+	-	+	0	-
„Responsiveness“	-	0	0	-	+	+	0
Rasche Konvergenz	+	+	+	-	0	0	-
Verständliche Modelle	+	0	0	-	0	+	-
„Predictive Power“	0	+	-	+	0	0	0

Übersicht bestehender BT-Systeme

Herausforderung	Nugg.ad	Peerius	OutThere Media	Double Click	Wunder Loop	Prudsys RE
Datenrepräsentation	-	-	0	0	-	+
„Sparse data“	0	+	0	0	0	0
Inkrementelles Lernen	+	0	0	+	0	0
Lernen in Echtzeit	-	+	+	+	+	+
„Responsiveness“	-	-	0	+	+	0
Rasche Konvergenz	-	-	+	0	0	0
Verständliche Modelle	-	0	0	-	0	0
„Predictive Power“	0	+	+	+	0	+

Vielen Dank für die Aufmerksamkeit!

**Für Fragen stehe ich jederzeit
gerne zu Ihrer Verfügung.**