Combining Data Mining and Decision Support

Example: Cars

DM + DS Integration?

DM + DS Integration!

Literature
Combining DM and DS

- "DS for DM":
  - ROC methodology
  - meta-learning
- "DM for DS":
  - MS OLE Db for DM
  - MS Analysis Services
  - model evolution (from data)
- "DM, then DS" (sequential application):
  - Decisions-At-Hand approach
- "DS, then DM" (sequential application):
  - using models in data pre-processing for DM
- "DM and DS" (parallel application):
  - combining through methods, e.g., DEK in HINT
  - considering different problem dimensions

ROC space

- True positive rate = TP / P
  - TP1 = 40, TP2 = 30, TP3 = 20
- False positive rate = FP / N
  - FP1 = 10, FP2 = 20, FP3 = 30
- ROC space has
  - FP on X axis
  - TP on Y axis

The ROC convex hull

Choosing a classifier

- FP cost = \( \frac{1}{2} \)
- Neg = 4
- Pos = 4
- slope = \( \frac{1}{2} = 2 \)
Choosing a classifier

FP cost = \( \frac{1}{8} \)
FN cost = 4
Neg = 4
Pos = 4
slope = \( \frac{3}{5} \)

“DM for DS”

Introducing DM methods into the DS process:
- MS SQL Server + Analysis Server
- Meta Object Profiler

“DM for DS”: Model Revision

“DM and DS”: Model Development

“DM and DS”: Model Development

 Modes of Operation

1. DEX only: from expert
2. HINT only: from data
3. Supervised: from data under expert supervision
4. Serial: HINT-developed model subsequently refined by the expert
5. Parallel: partial development of model(s) by DEX and HINT
6. Combined: combining sub-models developed in different ways
Sequential Application: “First DS, then DM”

- Decision Support
- Data Mining

“DS, then DM” in Data Pre-Processing

Sequential Application: “First DM, then DS”

- Data Mining
- Decision Support

Sequential Application: “First DM, then DS”

Parallel Applications: Multiple DM models, then DS
Problem: Prediction of Academic Achievement

Primary School | High School
---|---
1 | 1
... | 2
6 | 3
7 | 4

Prediction

DM + DS Integration: Academic

Data

DM: Weka

DS: DEXI

DM: HINT

Parallel Application: EC Harris

Models for Cost Value
Building Construction Project Workzone
Device Support
Value Zone
Feasible Buildings

Models for Building Feasibility

Conclusion

- DM & DS approaches are:
  - complementary
  - supplementary
- New and developing research area
- Typical combinations:
  - DS & DM
  - DM & DS
  - DM, then DS
  - DS, then DM
  - DM and DS
- Open questions:
  - formalization (template) of DM & DS integration
  - common methodologies and approaches