HINT: Learning DEXi Models From Data

Introduction

- Multi-Attribute Decision Making:
  - decompose the problem to less complex subproblems
- DEX:
  - An Expert System Shell for MADM
    - qualitative attributes
    - decision rules

Example

- What is the result of "Traditional" decision-tree learning, such as See5?
- How does this table look in DEX?
- How to create a hierarchical DEX model from this table?

Problem

Development of hierarchical decision models is difficult!

Given decision examples taken from:
- existing data sets of past decisions or
- provided explicitly by decision-maker, develop a corresponding model (hierarchy + functions)

Average weights

Decision Tree (See5)

DEXi Table & Rules

Scales

<table>
<thead>
<tr>
<th>Attribute Scale</th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>lo</td>
<td>lo</td>
<td>lo</td>
<td>lo</td>
</tr>
<tr>
<td>median</td>
<td>lo</td>
<td>med</td>
<td>lo</td>
<td>lo</td>
</tr>
<tr>
<td>high</td>
<td>lo</td>
<td>hi</td>
<td>lo</td>
<td>lo</td>
</tr>
<tr>
<td>low</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
</tr>
<tr>
<td>median</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
</tr>
<tr>
<td>high</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
</tr>
<tr>
<td>low</td>
<td>lo</td>
<td>lo</td>
<td>lo</td>
<td>lo</td>
</tr>
<tr>
<td>median</td>
<td>lo</td>
<td>med</td>
<td>lo</td>
<td>lo</td>
</tr>
<tr>
<td>high</td>
<td>lo</td>
<td>hi</td>
<td>lo</td>
<td>lo</td>
</tr>
<tr>
<td>low</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
</tr>
<tr>
<td>median</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
</tr>
<tr>
<td>high</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
</tr>
</tbody>
</table>
Example

<table>
<thead>
<tr>
<th>x_1</th>
<th>x_2</th>
<th>x_3</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>lo</td>
<td>lo</td>
<td>lo</td>
<td>lo</td>
</tr>
<tr>
<td>lo</td>
<td>med</td>
<td>lo</td>
<td>lo</td>
</tr>
<tr>
<td>lo</td>
<td>med</td>
<td>lo</td>
<td>lo</td>
</tr>
<tr>
<td>lo</td>
<td>med</td>
<td>hi</td>
<td>med</td>
</tr>
<tr>
<td>lo</td>
<td>hi</td>
<td>lo</td>
<td>hi</td>
</tr>
<tr>
<td>lo</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
</tr>
<tr>
<td>med</td>
<td>lo</td>
<td>med</td>
<td>lo</td>
</tr>
<tr>
<td>med</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
</tr>
<tr>
<td>hi</td>
<td>lo</td>
<td>lo</td>
<td>hi</td>
</tr>
<tr>
<td>hi</td>
<td>hi</td>
<td>lo</td>
<td>hi</td>
</tr>
<tr>
<td>hi</td>
<td>hi</td>
<td>hi</td>
<td>hi</td>
</tr>
</tbody>
</table>

- What is the result of “traditional” decision-tree learning, such as See5?
- How does this table look in DEK?
- How to create a Naive Bayes DEK model from this table?

Boolean Function Decomposition

Decomposition method

Extended Aitchison-Clarke decomposition of Boolean functions:
- adapted to model selection using interactive methods
- mixed-valued attributes
- unsupervised and supervised decomposition
- Partition selection measures
- generalization

Restriction:
- nominal attributes and utility

Implementation:
- miNT: Many-valued Naive Bayes Tool
  in C/Scala

Single-step decomposition

Partition matrix

Column compatibility

compatible columns

incompatible columns
Incompatibility graph

Finding function $H$

Finding function $G$

Single-step decomposition

Candidate decompositions

HINT Implementation: In ORANGE

\[ x_1, x_2, x_3 \]

\[ x_1, x_2, x_4 \]

\[ x_1, x_2, x_5 \]

\[ x_1, x_3, x_4 \]

\[ x_1, x_3, x_5 \]

\[ x_2, x_4, x_5 \]

\[ x_2, x_3, x_4 \]

\[ x_2, x_3, x_5 \]

\[ x_3, x_4, x_5 \]

\[ x_1, x_2, x_3, x_4, x_5 \]
Application: Housing Loan Allocation

- **User**: Housing Fund of the Republic of Slovenia
- **Task**: Allocating available funds to applicants for housing loans
- **Method**: Using a multi-attribute model for priority evaluation of applications
- **Supported by a DSS since 1991**:
  - Completed deals of loans 21
  - Applications: 44378 received, 27813 approved
  - Allocated loans: 254 million € (2/3 of housing loans in Slovenia)

Modes of Operation

1. **DEX-only from expert**
2. **HINT-only from data**
3. **Supervised**: from data under expert supervision
4. **Serial**: HINT-definition model subsequently refined by the expert
5. **Parallel**: parallel development of model(s) by DEX and HINT
6. **Combined**: combination sub-models developed in different ways

1. **DEX-Only Mode**

2. **HINT-Only Mode (1 of 2)**

   Reconstruction of the original model from unstructured data:
   - Real-life data from one file in 1994
   - 1932 applications
   - 12 attributes (2 to 5 values)
   - 722 unique examples
   - 3.7% coverage of the attribute space
   - Unsupervised decomposition

2. **HINT-Only Mode (2 of 2)**

   **Results**:
   - Reliably good overall structure
   - Inappropriate structure around c.3
   - Excellent classification accuracy:
     - **HINT**: 99.7 ± 2.5%
     - **C4.5**: 88.9 ± 3.9%

3. **Supervised Mode (1 of 4)**

   Unstructured dataset:

   - Predictors: cult_hist, fin_sources
   - Relevance: cult_hist, fin_sources
3. Supervised Mode (2 of 4)

All patterns with h=3 and minimal v (v=3) [11 of 120]

- suitab advantage employed marriages employed family
- suitab advantage employed employed health
- advantage employed employed married employed
- advantage employed employed health employed

New concept: status

3. Supervised Mode (3 of 4)

All patterns with h=3 and minimal v (v=4) [3 of 56]

- ownership suitab advantage
- advantage employed stage
- health family age

New concepts: social and then present

3. Supervised Mode (4 of 4)

Final structure

Results:
- Expert satisfied with the structure
- Improved classification accuracy:
  - supervised: 97.9 ± 1.8%
  - unsupervised: 94.7 ± 2.5%

4. Serial Mode

1. Develop an initial model by HINT from data
2. Extend/enhance the model manually using DEX

For example:
1. Take the model developed by HINT in supervised mode
2. Add the attributes cult-hist and fin-sources:
   - Extend the model structure
   - Define the corresponding decision rules

5. Parallel Mode

Develop two or more independent models by HINT and DEX for:
- comparison
- "second opinion"
- feasibility

For example, in this research we developed:
1. one DEX model
2. two HINT models in supervised and unsupervised mode

6. Combined Mode

Develop a single model using sub-models developed
- by different methods and
- from different sources

Hypothetical example:
1. Develop sub-tree for status by HINT
2. Develop soc-health by HINT from a different data set
3. A multi-disciplinary expert develops the house sub-tree using DEX
4. All three models are “glued” together in DEX by the gain-all definition expert
HINT: Conclusions

- Integration of DM and DS for model-based problem solving
- Requirements:
  - common model representation
  - expertise and data possession partial
  - methods for "at a model" (DA) and "implied" (DM) model development
- Offers a multitude of method combinations:
  - independent and parallel, combined, ...
- Specific scheme:
  - qualitative hierarchical multi-attribute models
  - DEX as a DA method
  - HINT as a DM method
- Weak points and limitations:
  - DEX requires substantial coverage of attribute space
  - qualitative attributes only
  - very sensitive to noise in data
  - high time complexity