Product-Impression Analysis Using Fuzzy C4.5 Decision Tree

Masataka Tokumaru and Noriaki Muranaka
Kansai University, Osaka, Japan

Outline

• Background of Our Research
• Product-Impression Analysis
• Impression Analysis of Golf Clubs
• Fuzzy C4.5 Decision Tree
• Experimental Result
• Conclusion

Background

• Designers are required to develop....
  • Functional and effective products
  • Valuable and reliable products
  • Recent needs to product designers
  • Easy to use
  • Attractive - consumer’s preference -

Product-Impression Analysis

• Product-Impression! *** emotion, feeling, sense, preference, subjectivity

Ease to Use? Attractive?

• Specification
  - Impressive Attributes (Various Aspect of feelings)
  - Driving Impression
  - Overall Impression
  - Feel Good?
  - Cool!
  - Attractive?
  - Engine Feel
  - Steering Feel
  - Comfortability
  - Rigidly of Chassis

Impression Analysis

• Questionnaire
  - Exterior Design Impression
  - Cute?
  - Modern?
  - Cool?
  - Elegant?

Design Factors

- factor a
- factor b
- factor c
- factor x

Quantity

Investigate relationship between Design Factors and Impression
Purpose of Our Research

**Specification of Product**
- factor a
- factor b
- factor c
- factor x

**Impressive Attributes for Evaluation**
- attribute a
- attribute b
- attribute c
- attribute n

**Overall Impression of Product**
- ease-of-use
- preference

*** is to investigate relationship between Impressive attributes and overall impression

---

**Ease-of-Use of Golf Clubs**

- Overall Impression: Ease-of-Hitting
  - How easily it can be used by people
- Impressive Attributes:
  - Easy to setup, Easy to control, timing to impact, drive stability, etc.
  - Hitting Sound (Feel, Level, Volume)
  - Length of Shaft (feel long or short), Weight (feel heavy or light)

---

**Impression Analysis of Golf Clubs**

**Impression Analysis Using Fuzzy C4.5 Decision Tree**

**SKIMA2006**

---

**Impression Analysis of Golf Clubs**

**Questionnaire**

- 24 attributes
- 7-point scale SD method

**Impression Analysis of Golf Clubs Using Fuzzy C4.5 Decision Tree**

- Fuzzy C4.5 Decision Tree
  - divides instances (questionnaire answers) into 3 subsets (easy to use, hard to use, neutral).
  - divides instances by value of a selected attribute
  - divides instances into fuzzy sets by membership functions

---

**Subjects:** 20 skilled golfers (42 years old and av. golf handicap of 15)

**11 drivers for the test**

- test for 5 min.
- rest and answer the questions for 5 min.
- repeat the task 15 times changing driver (11+4)

**Questionnaire**

- 24 attributes
- 7-point scale SD method

**273 instances (evaluations) of the various drivers in total**
Fuzzy C4.5 Decision Tree

Information entropy
\[ \text{info}(T) = - \sum \left( \frac{\text{freq}(C_i, T)}{|T|} \times \log_2 \frac{|T|}{\text{freq}(C_i, T)} \right) \]

Information entropy dividing the set \( T \) into subset \( T_j \) by an attribute \( X_p \)
\[ \text{info}_{T_j}(T) = - \sum \left( \frac{|T_j|}{|T|} \times \log_2 \frac{|T|}{|T_j|} \right) \]

Information gain
\[ \text{gain}(X_p) = \text{info}(T) - \sum_{j=1}^{n} \left( \frac{|T_j|}{|T|} \times \text{info}(T_j) \right) \]

The amount of split information
\[ \text{split info}(X_p)(T) = \text{info}(T) - \text{info}(T) \]

Information gain ratio
\[ \text{gain ratio}(X_p) = \frac{\text{gain}(X_p)}{\text{split info}(X_p)(T)} \]

Information entropy dividing the set \( T \) into subset \( T_j \) by an attribute \( X_p \)
\[ \text{freq}(C_i, T) = \sum \left( \text{possibility} \times \text{fitness value} + \cdots \right) \]

Membership Functions

answered value of questions about classes
\[ \text{freq}(C_i, T) = \sum \left( \text{possibility} \times \text{fitness value} + \cdots \right) \]

answered value of questions about an attribute \( X_p \)
\[ \text{freq}(C_i, T) = \sum \left( \text{possibility} \times \text{fitness value} + \cdots \right) \]

Branch and Leaf Node

Minimum Class Occupancy Ratio (MCOR): \[ \text{class}_{\text{min}} \]
If most of instances in a subset belong to one class, the node becomes a leaf node.

Minimum Data Content Ratio (MDCR): \[ \text{data}_{\text{min}} \]
If the number of data items contained in a subset is a little, the node becomes a leaf node.

Rules:
\[ R_1: X_5 \text{ is (+) and } X_2 \text{ is (+) and } X_3 \text{ is (+) then } C_1 \]
\[ R_2: X_5 \text{ is (-) and } X_7 \text{ is (-) and } X_6 \text{ is (+) and } X_2 \text{ is (-) then } C_3 \]

Experimental Result

Structural change of the constructed decision tree
Change in the number of rules

Classification Error Rate

The Condition for Reliable Decision Tree
\[ \text{MCOR} : \text{class}_{\text{min}} \text{ from } 0.6 \text{ to } 0.9 \]
\[ \text{MDCR} : \text{data}_{\text{min}} = 0.1 \]
An Example of Decision Tree for “Ease-of-Use” of Golf Clubs

Reliability Verification against Unknown Instances

- The number of instances: 273
- Example set
  - Training data: 200 (Selected from 273 instances)
  - Evaluation data: 70 (Selected from remaining 73 instances)
- Prepare 10 example sets
- Construct a decision tree using the training data
- Execute a classification test using the evaluation data
- Perform evaluations using 10 example sets

Experimental Results

- Proposed a method to investigate product impression using a fuzzy C4.5 decision tree.
- Conducted to impression analysis for the ease-of-use of golf clubs.
- Fuzzy decision tree could well consider user’s answered values with 7-level SD method.
- The proposed method showed distinct advantage in reliability of classifying unknown data.

Thank You for Your Attention