

Seeds

キーワード: データマイニング、ラフ集合、ルール抽出、推論システム
ラフ集合を用いた定性的データマイニングとその応用

Yasuo Kudo



しくみ解明系領域・知能情報学ユニット

くどう やすお

工藤 康生 教授

Phone:0143-46-5469 Fax:0143-46-5499

E-mail:kudo@csse.muroran-it.ac.jp

URL <http://www.muroran-it.ac.jp/>



データから感性が見える

研究の目的

	形状	大きさ	厚さ	重さ	嗜好
s1	横長	小	薄い	軽い	好き
s2	縦長	中	薄い	軽い	好き
s3	横長	大	厚い	重い	嫌い
s4	正方形	小	薄い	軽い	嫌い
s5	縦長	中	薄い	軽い	好き
s6	正方形	大	厚い	重い	嫌い

「好き」「嫌い」の違いは?



相関分析、多変量解析など統計的手法では、データの線形性、項目間の独立性、サンプル数などの制約があるため、分析が難しいデータも多数存在する。そこで、そのような制約のないラフ集合理論を用い、統計的手法とは異なる側面からデータマイニングを行い、その結果を推論するシステムを構築する。

研究の概要

データからのルール抽出

一般的に、データ内にある項目(例えば車の「イメージ」)に関する規則性は、データ内の他の項目の組み合わせとして記述される。ラフ集合を用いることで、データを正しく分類するための必要最小限の属性集合(相対縮約)を抽出し、データに潜む規則性をIf-Then形式の決定ルールとして表現する。

関係データベースに対する検索要求に基づき、ラフ集合を用いて生成した決定ルールから、ユーザが好むと思われる商品を推測し、「おすすめ商品」として提示

	排気量	色	TM	WD	イメージ
C1	2000cc	黒	MT	4	スポーティ
C2	1800cc	白	AT	2	ファミリー
C3	1500cc	赤	AT	2	コンパクト
C4	1500cc	黄	AT	2	コンパクト
C5	1500cc	黒	MT	4	ファミリー

応用事例

① 検索条件を入力



② 決定ルールの生成

レコメンデーションシステム

- 1500cc∧AT⇒コンパクトカー∧2WD∧黄色.
- 1500cc∧AT⇒コンパクトカー∧2WD∧赤色.
- etc.



検索条件と一致



一致しないけどおすすめ

④ 「おすすめ」の提示

③ 「暗黙の条件」と「提案内容」を推論

検索条件	Q	暗黙の条件	B
決定ルール1	B→Q	決定ルール2	B→Q'
暗黙の条件	B	提案内容	Q'

Step 1. 発想推論 Step2. 演繹推論

ラフ集合を用いたレコメンデーションシステム

Seeds ラフ集合を用いた定性的データマイニングとその応用

研究(開発)のアピールポイント

◆研究の新規性、独自性

ラフ集合を用いてデータに潜む規則性を抽出するだけでなく、抽出された規則性自体の重要度およびデータ内の各項目の重要度なども評価可能。

◆従来研究(技術)と比べての優位性

統計的手法とは異なり、データの線形性、分析項目間の独立性、サンプル数等の制約がなく、多数の項目間の関連性を同時に分析可能。

表形式のデータ

	排気量	色	TM	WD	イメージ
C1	2000cc	黒	MT	4	スポーティ
C2	1800cc	白	AT	2	ファミリー
C3	1500cc	赤	AT	2	コンパクト
C4	1500cc	黄	AT	2	コンパクト
C5	1500cc	黒	MT	4	ファミリー

◆研究に関連した特許の出願、登録状況

なし

研究(開発)のビジョン、ステージ

◆適応分野

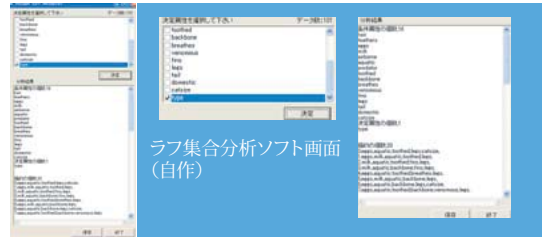
人の感性に関わるアンケート分析など。

◆製品化、事業化のイメージ

ラフ集合を利用したソフトウェアの開発など。

◆研究のステージ

基礎研究 応用段階



企業等へのご提案、メッセージ

◆研究(開発)に関連して、あるいはそれ以外に関われる業務

マーケティングによるデータ(感性データ含)からの助言、ラフ集合によるアンケート解析など。

表形式のデータ

被験者	Q.1	Q.2	Q.3	...	Q.97	Q.98	Q.99	結果
001	1	3	4	...	4	2	3	○
002	3	1	1	...	5	1	5	×
...				...				
029	5	5	2	...	4	2	2	×
030	2	2	4	...	2	5	3	○

データに潜む規則性を抽出

[Q.12=3] ∧ [Q.36=2] ∧ [Q.69=1] → [結果=○]

[Q.12=3] ∧ [Q.36=2] ∧ [Q.69=5] → [結果=×

[Q.12=5] ∧ [Q.36=5] ∧ [Q.69=3] → [結果=×

◆利用可能な設備、装置など

◆教員からのメッセージ

分析にお困りのデータがございましたら、お気軽にお問い合わせください。

工藤 康生

