

データを説明する論理規則の発見システム

吉田 浩幸 中野 智文 犬塚 信博

キーワード: データの分類 ルール発見 ファジィ データマイニング

目的

大量のデータを複雑な知識を使って分類する論理規則を発見する

データの分類

データ形式	分類手法	利用法	特徴
関係データベース	決定木	分類	
複合的なデータ	LP	分類, 順序付け, 選択	背景知識が利用可能
	FLP	数値割付け	LPを拡張, ファジィ

LPとは?

入力: 正 / 負事例, 背景知識
出力: 事例を説明可能な
論理プログラム



論理的な解釈が可能

分類問題

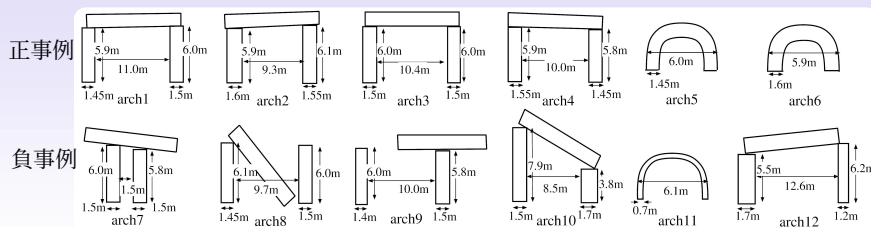
事例进行分类するための規則を発見する

バランス良く構成された
アーチを学習

- LPに正 / 負事例を与えて帰納
- 帰納したルールで未知事例进行分类

$\text{good_arch}(A) \leftarrow \text{has_archblock}(A,B) \wedge \text{well_proportioned}(B).$

$\text{good_arch}(A) \leftarrow \text{has_roofblock}(A,B) \wedge \text{has_supportblock}(A,C,D) \wedge \text{well_proportioned}(D) \wedge \text{very_balanced_distance}(C,D) \wedge \text{support}(D,B) \wedge \text{support}(C,B).$



選択問題

複数候補から1つを選択する規則を発見する

英文品詞付け



品詞付き英文

品詞付き英文データから

品詞をつける規則を発見

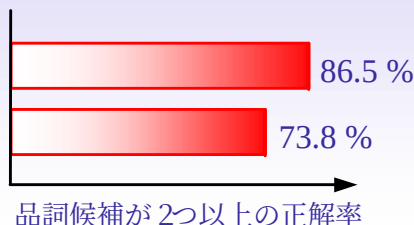
本システム

頻度辞書のみ

品詞ルール(プログラム)

ルール発見

ベイズ学習をもとにノイズを考慮



品詞候補が2つ以上の正解率

数値割付け問題

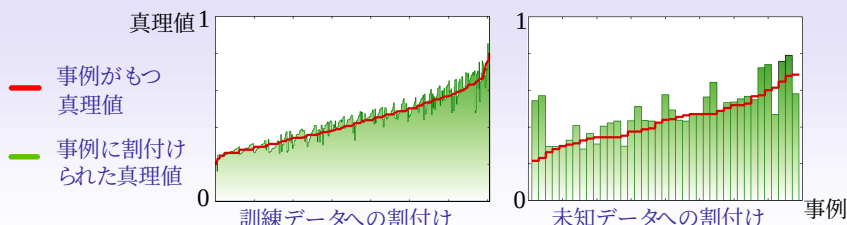
事例がもつ真理値と同程度満たす規則を発見する

車の排気量, 馬力, 重量等の
データを与えて, 燃費を予測

- 燃費の値を0から1へマッピング
- ファジィ**を導入したLPで帰納
- 帰納したルールで**数値割付け**

$\text{class}(X0) \leftarrow \text{dt}(X0,X1,X2,X3,X4,X5,X6,X7) \wedge \text{very_high_hor}(X0,X3) \wedge \text{low_dis}(X0,X2).$

$\text{class}(X0) \leftarrow \text{dt}(X0,X1,X2,X3,X4,X5,X6,X7) \wedge \text{very_low_wei}(X0,X4) \wedge \text{high_mod}(X0,X6) \wedge \text{very_low_hor}(X0,X3).$



応用

自然言語処理, WEB 文章の分類, 情報フィルタリング(メーラー等の知的カスタマイズ)
タンパク質の構造予測, 化学物質の発癌性予測, 将棋の戦略知識の発見
データマイニング(顧客情報抽出等), 自動プログラミング, ロボット行動最適化

お問い合わせ先:

名古屋工業大学共同研究センター助教授 犬塚信博

phone: 052-735-5050

e-mail: inuzuka@ics.nitech.ac.jp

http://www-inzklab.ics.nitech.ac.jp/