メンバーシップ関数の種類によるファジィ識別器への影響調査

第 8 グループ 瀧川 弘毅

# はじめに

識別性能が高い識別器が多く研究されているが，それらはもののブラックボックス化しており識別の判断理由を説明できない．しかし，ファジィ識別器は言語的に解釈可能なルール群から構成されており，どのように識別しているかが解釈可能であるという特徴を持つ．

実世界では識別されるデータセットも様々なものがあり，データセットに対して最適なメンバーシップ関数を用いた識別器を生成することで識別器の性能向上を目指す．本研究では前段階としてメンバーシップ関数を変更して識別器を生成し，識別器の性能への影響を調査した．

## ファジィ識別器

*n*次元*M*クラスのパターンが*m*個与えられたパターン識別問題に対して，ファジィ集合を条件部とする以下のIf-thenルールを用いてファジィ識別器を設計する．ある未知パターンは**x** = (*x*1, …, *xn*) のように表され，*xi*は第*i*次元 (*i* = 1, 2, …, *n*) における属性値を表す．

Rule *R*: If *x*1 is *A*1 and ... and *xn* is *An*

then Class *C* with *CF*  (1)

**A** = (*A*1, …, *An*)は条件部ファジィ集合，*C*は結論部クラス，*CF*はルールの重みを表す．また，学習用データを用いてルール重みと結論部クラスを決定する．未知パターンの推論は，適合度とルール重みの積が最大となるルールを勝者とする単一勝利ルールによって行う．

## メンバーシップ関数

適応させる条件部集合への帰属度を表現する関数．本研究では三角形型，ガウシアン型，台形型，均等区間型の4種類のメンバーシップ関数を用いる．各型については，2, 3, 4, 5分割のファジィ集合14種と，メンバーシップ値として必ず1を返す ”don’t care” の合計15種類のファジィ集合を同時に用いる．



三角形型

ガウシアン型

台形型

均等区間型

図 1:3分割の各メンバーシップ関数

## 多目的最適化

本研究では，ファジィ識別器の誤識別率の最小化とルール数最小化の2つの目的を用いる．本研究では代表的な進化的多目的最適化アルゴリズムであるNSGA- IIをファジィ識別器に適応し2目的最適化を図る．誤識別が低ければ識別性能が高く，ルール数が少なければ複雑性が少なくなり計算コストが少なく解釈性能が高い識別器となる．

# 数値実験

実験は以下の設定で行った．

|  |  |  |
| --- | --- | --- |
| 試行回数 | : | 30 (10-fold cross-validation 3) |
| 最大ルール数 |  | 60 |
| 個体群サイズ | : | 60 |
| EMOA | : | NSGA-II |

データセットには以下を用いた．

|  |  |  |
| --- | --- | --- |
| 試行回数 | : | 30 (10-fold cross-validation 3) |
| 最大ルール数 |  | 60 |
| 個体群サイズ | : | 60 |
| EMOA | : | NSGA-II |

|  |  |  |  |
| --- | --- | --- | --- |
|  | パターン数 | 属性数 | クラス数 |
| iris | 150 | 4 | 3 |
| wine | 178 | 13 | 3 |
| phoneme | 5404 | 5 | 2 |
| sonar | 208 | 60 | 2 |
| yeast | 1484 | 8 | 10 |

メンバーシップ関数には

三角形型，ガウシアン型，台形型，均等区間型

を用いた．

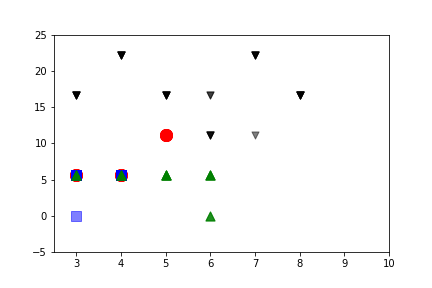
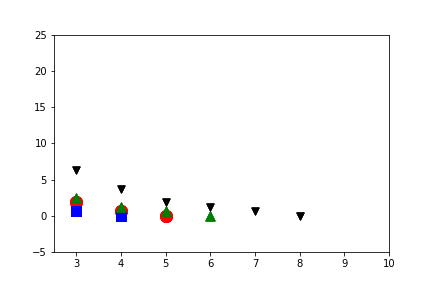
# 数値実験結果

各試行での学習用データ誤識別率最小個体の平均値を表にしたもの

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 学習用データ誤識別率 | | | | | |
|  | Iris | Wine | Phoneme | Sonar | yeast |
| 三角形型 | 0.72 | 0.00 | 16.12 | 6.126 | 38.66 |
| ガウシアン型 | 0.74 | 0.06 | 16.18 | 4.861 | 37.89 |
| 台形型 | 0.49 | 0.00 | 15.07 | 3.882 | 37.72 |
| 区間型 | 0.10 | 0.00 | 16.01 | 1.941 | 40.04 |

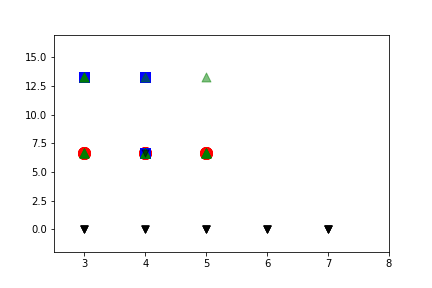
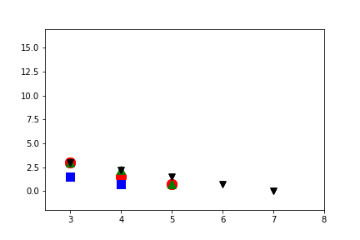
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 評価用データ誤識別率 | | | | | |
|  | Iris | Wine | Phoneme | Sonar | yeast |
| 三角形型 | 6.22 | 7.86 | 17.40 | 22.30 | 42.54 |
| ガウシアン型 | 4.89 | 5.46 | 17.42 | 23.82 | 42.92 |
| 台形型 | 5.33 | 5.80 | 16.05 | 22.16 | 42.25 |
| 区間型 | 5.33 | 9.89 | 17.46 | 25.23 | 43.87 |

次に各試行での学習用データ誤識別率最小個体のルール数を表にした．



Wine:学習用データ

Wine:評価用データ



Iris:学習用データ

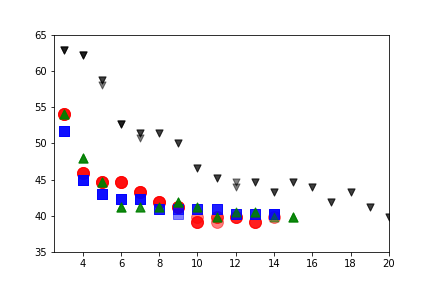
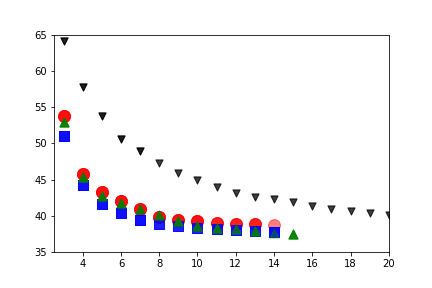
Iris:評価用データ

●三角形型 ■ガウシアン型 ▲台形型 ▼区間型

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ルール数 | | | | | |
|  | Iris | Wine | Phoneme | Sonar | yeast |
| 三角形型 | 5.10 | 5.00 | 13.50 | 8.767 | 12.70 |
| ガウシアン型 | 4.03 | 3.57 | 11.33 | 9.133 | 10.83 |
| 台形型 | 5.37 | 5.90 | 16.07 | 11.17 | 14.93 |
| 区間型 | 6.33 | 8.00 | 17.53 | 15.43 | 21.50 |

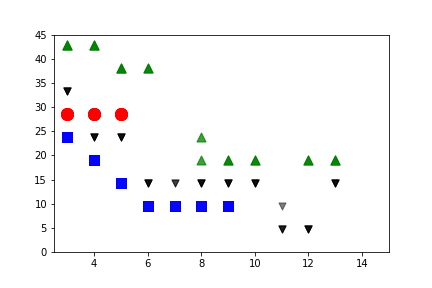
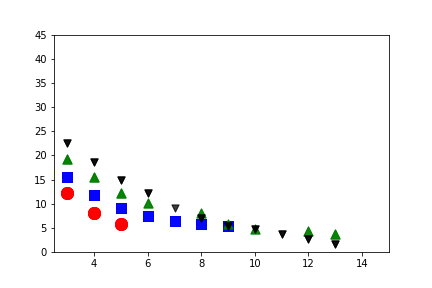
# おわりに

本研究では，移住操作を適用しない島型の並列分散実装をMoFGBMLに適用し，得られた各部分個体群から弱識別器を抽出するアンサンブル識別器を設計した．数値実験として，移住操作を適用した並列分散実装の識別性能と比較し，提案手法の汎化性能の高さを検証した．さらに，島数を変更した場合のアンサンブル識別器の識別性能への影響を調べ，島数が大きい場合の提案手法の有効性を検証した．



Yeast:学習用データ

Yeast:評価用データ



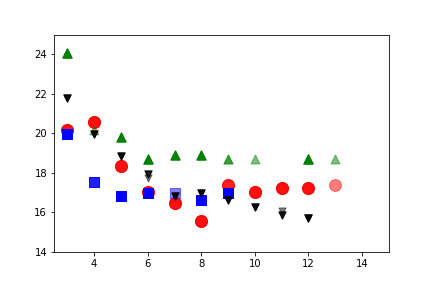
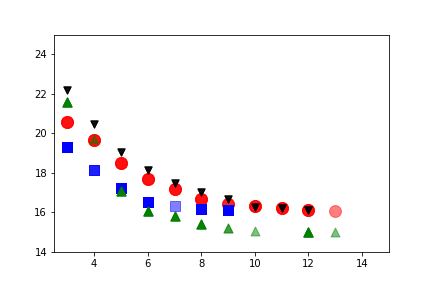
Sonar:学習用データ

Sonar:評価用データ

本手法で得られる非劣な弱識別器集合には，ルール数最小化に偏った分布が見られた．今後の課題として，トレードオフ曲線における分布の偏りを考慮した弱識別器の選択方法を考えることが挙げられる．

# 参考文献

1. Y. Nojima, Y. Takahashi, and H. Ishibuchi, “Application of parallel distributed implementation to multiobjective fuzzy genetics-based machine learning,” *Lecture Notes in Computer Science 9011: Intelligent Information and Database Systems – ACIIDS 2015*, Part I, pp. 462-471, Springer, Berlin, March 2015.
2. K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Trans. on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, 2002.
3. L. Rokach, “Ensemble-based classifiers,” *Artificial Intelligence Review*, vol. 33, pp. 1-39, 2010.
4. H. Ishibuchi, M. Yamane, and Y. Nojima, “Ensemble fuzzy rule-based classifier design by parallel distributed fuzzy GBML algorithms,” *Proc. Of 9th International Conference on Simulated Evolution and Learning – SEAL 2012*, pp. 93-103, Hanoi, Vietnam, December 16-19, 2012.



Phoneme:学習用データ

Phoneme:評価用データ

===MEMO===