

# ベイズ境界性に基づく分類器パラメータの最適選択手法の実験的評価

蔭山昌幸<sup>†</sup> David Ha<sup>†</sup> 友利宥也<sup>†</sup> 千田将大<sup>†</sup> 渡辺秀行<sup>††</sup> 大崎美穂<sup>†</sup> 片桐滋<sup>†</sup>  
<sup>†</sup> 同志社大学 <sup>††</sup> ATR

## 1. はじめに

パターン認識の最終目標である、最小分類誤り確率(ベイズリスク)状態に対応するベイズ境界の性質を基に、分類境界の曖昧性尺度を定義し、学習された分類器(パラメータ)群から最適な分類器を選択する手法(BB法と略記)の研究を行なっている<sup>[1]</sup>。本稿では、このBB法において実行時間が長く手続きも複雑な境界近傍標本の選択法の単純化を目指した評価実験を行う。

## 2. 曖昧性尺度

理想的なベイズリスク状態においては、クラス境界上にある境界標本の事後確率の値は等しくなる<sup>[2]</sup>(ベイズ境界性と呼ぶ)。BB法は、このベイズ境界性を基に、隣接クラスの境界標本の近似値、境界近傍標本から推定クラス境界の曖昧性尺度を計算する。この尺度値は、その絶対値が大きい程、推定境界がベイズ境界に近いことを、言い換えれば、対応する学習された分類器パラメータがより望ましいものであることを示す。ここで、境界標本の代わりに境界近傍標本を用いるのは、有限個標本しか利用できない現実問題において、境界標本はめったに存在しないためである。

BB法は主に2つのステップから成る。ステップ1では境界近傍標本の選択を行う。ステップ2では選択された境界近傍標本を用いて、各標本位置におけるエントロピーを求め、曖昧性尺度を計算する<sup>[1]</sup>。

## 3. 幾何マージンを用いる境界近傍標本選択

従来、ステップ1は、境界近傍標本を求めるために、まずアンカーと呼ばれる擬似的な境界標本を生成していた。生成の計算量が多く、またアンカーあるいはそれに基づく境界近傍標本の生成数の設定はやや合理性を欠いていた。この解決のため、本稿では幾何マージン<sup>[3]</sup>と標本密度分布との関係に着目し、境界近傍標本選択の停止基準を明確にした手法を採用した。幾何マージンの値から昇順に候補となる標本を選択し、各標本位置における密度分布を計算する。そして境界近傍標本全体の密度が均一になるように、標本選択を行い、境界近傍標本が選択されなくなるまで処理を継続した。

## 4. 評価実験

次の条件において交差検証(CV)法(80 fold)との比較を行い、提案手法の性質を調査した。  
 データセット: 2クラス・2次元のガウス分布によって生成

した合成データ(GMM)。標本数は2200。

- ・分類器: マルチプロトタイプ分類器
- ・プロトタイプ数(/クラス): 1, 2, ..., 10, 20, ..., 170
- ・学習法:  $k$ 平均法

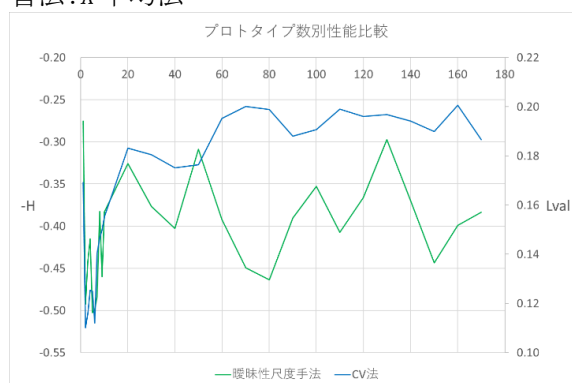


図1. プロトタイプ数別の性能評価

図1にプロトタイプ数別の結果を示す。右側の縦軸(Lval)はCV法による分類誤り確率の推定値を、左側の縦軸(-H)は提案手法による曖昧性尺度値(符号反転値)を示している。横軸は、プロトタイプ数である。Lvalの最小値は、最小分類誤り確率の推定値に対応する。図から、提案手法による-Hの最小値とLvalの最小値とはほぼ同じプロトタイプ数において対応していること、即ち提案手法がほぼ正確に最小分類誤り確率状態(あるいはベイズ境界)に対応するプロトタイプ数を示していることがわかる。しかしその一方で、-HとLvalとの2つの曲線全体は必ずしも類似の傾向を示しておらず、提案手法による-Hの計算には不安定さの問題があることもわかった。

**謝辞:** 本研究の一部は、科研費(18H03266)および文科省平成26年度私学戦略的研究基盤形成支援事業「進化適応型自動車運転支援システム「ドライバ・イン・ザ・ループ」」の支援をうけて行われた。

## 参考文献

- [1] David Ha, et al., “Optimal Classifier Model Status Selection Using Bayes Boundary Uncertainty”, IEEE MLSP2018 (Sep. 2018)
- [2] K. Fukunaga, Introduction to Statistical Pattern Recognition (2nd Ed.), Academic Press (1990).
- [3] 渡辺秀行, 片桐滋, 山田幸太, 中村篤, 渡部晋治, 大崎美穂ほか. 幾何マージンに基づく誤分類尺度を用いた最小分類誤り学習法. 電子情報通信学会論文誌 D, Vol.94, No.10, pp.1664-1675, 2011.