

# 最小分類誤り学習法におけるベイズエラー推定能力の検証

鶴尾 明大<sup>†</sup> 西山 育宏<sup>†</sup> 渡辺 秀行<sup>††</sup> 片桐 滋<sup>†</sup> 大崎 美穂<sup>†</sup>  
<sup>†</sup>同志社大 <sup>††</sup>ATR

## 1. はじめに

分類器を最適化する手法に最小分類誤り(MCE: Minimum Classification Error)学習法<sup>[1]</sup>(以下, 関数マージン(FM: Function Margin)MCE 学習法)と大幾何マージン(LGM: Large Geometric Margin)MCE 学習法<sup>[2]</sup>がある. この二つの学習法のベイズ誤りの推定能力はまだ十分に調査されておらず, 本研究でその実験的な検証を行う.

## 2. FM-MCE 学習法と LGM-MCE 学習法

FM-MCE 学習法と LGM-MCE 学習法は, 共に分類誤り確率の最小化を目指す. 両者の主な違いは誤りの程度を表す誤分類尺度にあり, FM-MCE 学習法は関数マージン型尺度を, LGM-MCE 学習法は幾何マージン型尺度を採用している. 誤分類尺度空間における幾何マージンと標本空間における幾何マージンとは等しく, 理論的には LGM-MCE 法の方がベイズ誤りの推定能力が高いことが予想されている.

## 3. 評価実験

データセットには UCI Machine Learning Repository が提供する Letter Recognition データ(標本数:20000, クラス数:26, 次元数:16)を用い, 分類器はマルチプロトタイプ型とした. プロトタイプ数(/クラス)は3~50とした. 原理的に, 過少のプロトタイプは学習標本上でベイズ誤りの過大推定を, 過多のプロトタイプはその過小推定をもたらす. 従って, ベイズ誤り推定力が高い学習法は, この推定の過大さを抑制することが期待される. データセットのベイズ誤り(推定)値は,  $k$  平均法を用いる交差検証(CV: Cross Validation)法によって求めた( $k$ =プロトタイプ数).

### 3.1 実験 1: 全データを学習用標本とする評価

全データを学習に用いることで学習用標本を可能な限り増やして学習を行った(理想は, 無限個標本を用いることである). 学習用データに対して得られた分類誤り率を図1に示す. 図から, FM-MCE 法と LGM-MCE 法のいずれもが, プロトタイプ数が大きくなるとベイズ誤りの過小推定に陥ってしまったことがわかる. 差は小さいものの, FM-MCE 法の方がベイズ誤り推定値に近い推定を行っていた.

### 3.2 実験 2: CV 法による評価

5-fold-CV 法により評価を行った. 試験用データに対する分類誤り率を図2に示す. 図から両手法ともに, プロトタイプ数が大きともベイズ誤りの過小推定が抑制されていることがわかる. これは CV 法の効果と思われる. また, プロトタイプ数の広い範囲において, LGM-MCE 法の方がベイズ誤り推定値に近い分類誤り率を出して

いることがわかる.

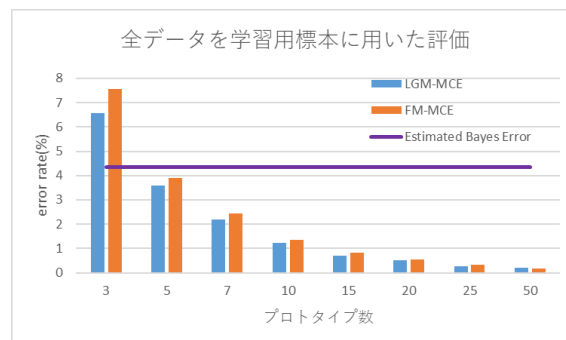


図 1. 全データを学習用標本に用いた評価

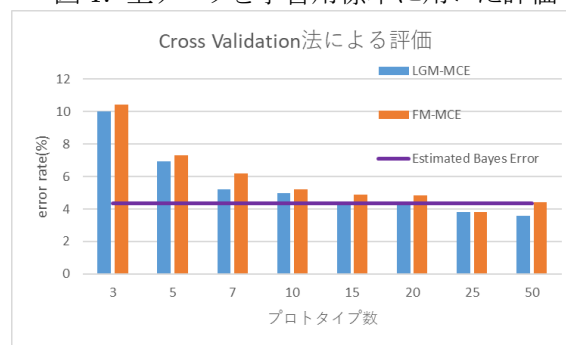


図 2. Cross Validation 法による評価

## 4. むすび

幾何マージンを増大させる機構を持つ LGM-MCE 学習法の方が, 特に多数のプロトタイプを持つ場合に優れたベイズ誤りの推定を行うことを予想していたが, 異なる結果が得られた. プロトタイプ数が小さいときに顕著にみられるように FM-MCE 学習法の分類誤り率が大きく出る傾向があり, これは学習不足に起因する可能性がある. 実際, 両手法が用いる平滑化分類誤り数損失の平滑度の値域はかなり異なり, これが学習の速度に影響したかもしれない. 今後, 両学習法のより詳細な効果の検証が必要である.

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] B.-H. Juang, et al., IEEE Trans. Signal Processing, vol. 40, no. 12, pp. 3043-3054, Dec. 1992.
- [2] 渡辺秀行, 他, 信学会論文誌 D, Vol. J94-D, No. 10, pp. 1664-1675 (2011 年 10 月).