

# ベイズ境界らしさを最大化する分類器学習法

千田 将大<sup>†</sup> ア デイビッド<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

理想的な最小分類誤り確率(ベイズエラー)状態は、それに対応するクラス境界上の境界標本に対する事後確率値が等しくなる性質(ベイズ境界性<sup>1)</sup>と呼ぶ)と等価である。この性質に着目し、学習標本上の学習のみによって、分類器の出力(識別関数値)が直接的にこのベイズ境界性を満たすように分類器を学習する手法を提案する。

## 2. 分類器の出力

ベイズ境界性を分類器が表現できるようにするため、分類器の識別関数で構成される誤分類尺度を用いて事後確率の近似表現を定義する(式(1))。

$$\hat{P}(y | d_y(\mathbf{x}; \Lambda)) = \frac{1}{1 + \exp(\alpha_y d_y(\mathbf{x}; \Lambda))} \quad (1)$$

ここで、 $d_y(\mathbf{x}; \Lambda)$ はクラス $y$ における誤分類尺度であり、入力標本 $\mathbf{x}$ をモデルパラメータ $\Lambda$ を用いて分類する際、その負値が正分類を、正値が誤分類を示す。この定義は、境界付近では値がおよそ 0.5(最も分類判断が曖昧)になり、分類の正しさが高まるほど 1 に近づき、分類の誤り程度が高まるほど 0 に近づく。 $\alpha$ は、この変化の程度を制御するハイパーパラメータである。

## 3. 学習目標

学習は、式(1)の事後確率の近似値の近似精度を高めるように行う。そのため、目標値として、境界近傍標本のみを対象とする  $k$ -NN 法による事後確率推定値を用いる(式(2))。

$$\bar{P}(y | \mathbf{x}) = \frac{1}{N_{K_x}} \sum_{P_{kNN} \in K_x} P_{kNN}(y | \mathbf{x}) \quad (2)$$

ここで、 $K_x$ は、入力標本 $\mathbf{x}$ 付近の近傍標本からなる狭い領域であり、 $P_{kNN}$ はその領域内の標本のクラス比である。なおここで、事後確率推定値は本来、境界上にある境界標本を用いて行われるべきであるが、その取得がめったに実行可能ではない現実における近似として境界近傍標本を用いている。

境界近傍標本の抽出は、次の手順に従う。

1. 入力標本 $\mathbf{x}$ と異なるクラスの標本をランダムに選択。
2. 選択された 2 標本間を結び、その線上における誤分類尺度を計算。
3. 最も 0 に近い誤分類尺度値を示す(ほぼ境界上と思われる)点にアンカーを設置。
4. アンカーから最近傍の標本を対象として式(2)の学習目標を求める。

学習は、式(1)の値が式(2)の値に近づくように分類器パラメータ $\Lambda$ を更新することによって行う。この 2 式の差異を表す損失を次のように定義し、学習はその最小化を通して進める。式(3)の $L(\Lambda)$ の最小化は、式(2)の事後確率の近似精度を最も高める、すなわちベイズ境界性を最も高める $\Lambda$ をもたらす。

$$L(\Lambda) = \frac{1}{N_{NBS}} \sum_{\mathbf{x} \in NBS} \|\hat{P}(y | d_y(\mathbf{x}; \Lambda)) - \bar{P}(y | \mathbf{x})\|^2 \quad (3)$$

ここで、NBS は抽出した境界近傍標本(Near Boundary Sample)である。

## 4. 実験評価

人工データセット(2次元, 2クラス, クラス毎の正規分布に従う 2200 標本)にマルチプロトタイプ型分類器を適用して提案手法を評価した。式(3)の損失関数の最小化には最急降下法を用いた。標本は、学習用と試験用に等分割した。学習の目標は、ベイズ境界性を最も高めるプロトタイプ数  $k$  とその  $k$  個のプロトタイプの状態(値)との発見である。提案法の比較基準として、 $k$  平均法を用いる Cross Validation(CV)法によって最小分類誤り確率の推定値を求めた。

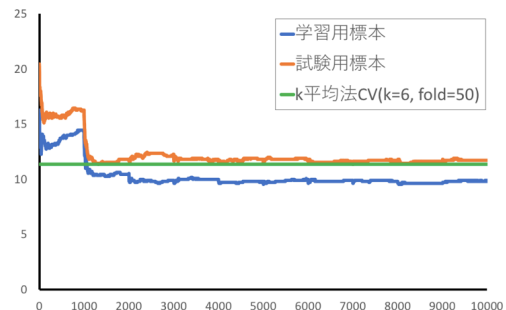


図 1: 分類誤り率(%)

図 1 に実験結果を示す。学習エポック数(横軸)が増えるにつれ、試験用標本に対する分類誤り率(縦軸)が CV 法による分類誤り率(縦軸)、即ち理想的な最小分類誤り確率に漸近していることがわかる。また、学習用標本に対する誤り率までもが比較的 CV 法の結果に近く、過学習が抑制されている可能性も読み取れる。以上より、提案学習法が、学習標本上における学習のみを通してベイズ境界を達成し得ることが推察される。

**謝辞:**本研究の一部は、科研費(18H03266)及び文科省平成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」の支援を受けた。

## 参考文献

- [1] David. Ha, et al.: IEEE MLSP2018 (Sep. 2018).