

エッジ接続機会を増やす Gm-SOINN

日大生産工 ○井上 由唯 日大生産工 山内 ゆかり

1. まえがき

膨大な量のデータを生成し、その中から有用な情報を抽出するため多くのデータマイニングアルゴリズムが採用されている。例として、自己増殖型ニューラルネットワーク (Self-Organizing Incremental Neural Network: SOINN) [1] は自己組織化マップ (Self-Organizing Map: SOM) [2]とは異なり事前にネットワークの構造を決定する必要がなく、追加学習が可能という特徴がある。しかし、SOINNの学習によって得られた位相構造がデータ分布を忠実に表すとは限らない。安定性と可塑性のジレンマを解決するため、Hang Yuらが Gaussian関数に基づく自己増殖型ニューラルネットワーク (Gaussian membership-based Self-Organizing Incremental Neural Network: Gm-SOINN) [3]を提案した。Gm-SOINNではGaussian関数を使用することでノードが過度に削除されることがなくなり、ノイズノードが挿入されることがなくなる。よって、Gm-SOINNでは安定性と可塑性のジレンマがなく、他のSOINNベースの方法と比較しても優れたパフォーマンスを発揮したと報告されている。Gm-SOINNではエッジを接続するのに条件を設けている。エッジ接続の処理が行われるには第1勝者、第2勝者のラベルやクラス、ガウスモデルの重なり度合いを計算し、条件を満たしたときのみエッジ接続の処理が行われる。

本研究では、エッジ接続に判定条件を持っている Gm-SOINN にエッジ接続機会を増やす条件を与える手法を提案する。この手法により Gm-SOINN に比べ、分布の特徴を多く得ることにつながると考えられる。それにより、従来の手法より学習精度を向上させることを目指す。

2. 従来研究

Gm-SOINNはSOINNを継承、発展させたもので、特徴空間に隣接するパターンである類似パターンは同じノードまたはネットワーク内で隣接するノードにマッピングされる。Gm-SOINNでは、Gaussian関数を使用し、入力 $X(t)$ がノードに属する度合い、勝者ノードと勝者ノードの近傍ノードを更新する。1つのノードの重みベクトルだけでなく、そのノードが属するサブクラスの密度分布も記録する。

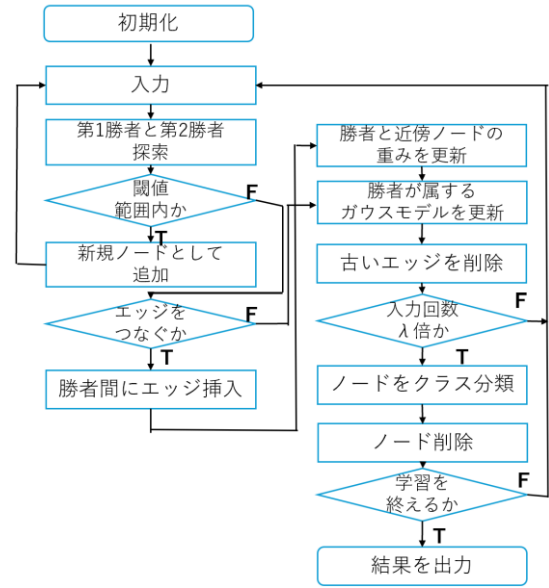


Figure 1 Gm-SOINN のフローチャート

Figure1 に Gm-SOINN のフローチャートを示す。新しい入力を与えられ、Gaussian 関数で μ_i を求める。 d_i^2 は入力と勝者のユークリッド距離を表している。第1勝者 s_1 、第2勝者 s_2 をとって見つける。

$$\mu_i = \left[d_i^2(t) \sum_{j=1}^{AN} \left(\frac{1}{d_j^2(t)} \right) \right]^{-1} \quad (1)$$

勝者の類似度閾値 T を求める。類似度閾値の範囲外である場合、現在のノードと大きく異なるため、新しいノードとして挿入される。連結ノードが存在しない場合、近傍ノードの最小距離を使用する。連結ノードが存在する場合は連結ノードの中で最大距離を使用する。類似度閾値内である場合、エッジの挿入判定や削除、ガウスモデルの更新の処理を行う。エッジは s_1 または s_2 とともにラベルがない場合と s_1 と s_2 が同じクラスに所属している場合は2つのノードをエッジで接続する。エッジの接続後に重みを更新する。重みの更新は式(2)、(3)を踏まえて式(4)を使用する。

$$W_i(t) = \frac{\sum_i^t \mu_i^2(t) X(t)}{\sum_i^t \mu_i^2(t)} \quad (2)$$

$$\begin{aligned} \Delta W_i(t+1) &= \frac{\mu_i^2(t+1)(X(t+1) - W_i(t))}{\sum_{k=1}^t \mu_i^2(k) + \mu_i^2(t+1)} \quad (3) \end{aligned}$$

$$W_i(t+1) = W_i(t) + \Delta W_i(t+1) \quad (4)$$

ガウスモデルにはこのモデルが勝利した回数 num_G 、平均ベクトル mv_G 、共分散行列 Cov_G 、入力されたデータがガウスモデルにどの程度、属するか判断するパラメータ r_G が含まれている。エッジ接続、重み更新後、式(5)、(6)、(7)、(8)を用いて s_1 、 s_2 それぞれのガウスモデルを更新する。

$$num_G(t+1) = num_G(t) + 1 \quad (5)$$

$$mv_G(t+1) = mv_G(t) + \frac{1}{num_G(t+1)} (Xt(t) - mv_G(t)) \quad (6)$$

$$Cov_G(t+1) = (1 - \frac{1}{num_G(t+1)}) Cov_G(t) + \frac{1}{num_G(t+1)} \times (X(t) - mv_G(t))(X(t) - mv_G(t))^T \quad (7)$$

$$r_G = x_{num_G(t+1)}^2(\alpha) \quad (8)$$

異なるラベルがついている場合はガウスモデルの重なり度合いを用いてエッジを接続するか判定する。重なり度合いを計算するために Bhattacharyya 距離を使用する。式(9)を用いて式(10)で求められる。これを計算し、得られた結果が0または1に近いほど重なりが大きくなる。式(11)を用いて s_1 、 s_2 のそれぞれのガウスモデルの体積を求める。モデルの重なりが大きく、式(12)の条件を満たした場合、エッジを接続する。

$$Cov^{-1} = (|Cov_{s1}^{-1}| \times |Cov_{s2}^{-1}|) / 2 \quad (9)$$

$$olap(G_{s1}, G_{s2}) = \frac{1}{8} (mv_{s1} - mv_{s2})^T \times Cov^{-1} (mv_{s1} - mv_{s2}) + \frac{1}{2} \ln \left(\frac{|Cov^{-1}|}{\sqrt{|Cov_{s1}^{-1}| \times |Cov_{s2}^{-1}|}} \right) \quad (10)$$

$$V_G = \sqrt{2}^{(P+1)} \left(\prod_{i=0}^{\left(\frac{P}{2}\right)-1} \frac{1}{P-2i} \right) \sqrt{|Cov_G|} r_G \quad (11)$$

$$V_{md} \leq V_{s1} + V_{s2} \quad (12)$$

重みを式(4)で更新する。エッジの接続判定の条件を満たしエッジを接続した場合、ガウスモデルそれぞれの情報を加味した方法で更新する。条件を満たさなかった場合はそれぞれを更新する。また、 s_1 とつながっているすべてのエッジの年齢を1ずつ増加する。エッジの年齢が設定していたパラメータ $agemax$ を超えたエッジを削除する。入力回数がパラメータ λ の整数倍の学習が終了後、Gm-SOINNはノードを異なるクラ

スに分類し、ノイズの多いノードを削除する。学習がすべて終了すると、最終的な位相構造を出力する。

3. 提案手法

従来のGm-SOINNではエッジ接続をするために条件があるため、分布の特徴を得られにくいと考えた。従来研究ではラベルやクラス、ガウスモデルの重なり度合いによってエッジの接続機会が与えられていた。そこで、本研究では与えられていた条件に加えて、エッジ接続の機会を増やす条件を与える。エッジ接続の機会を加えることにより本研究では、従来の手法より分布の特徴を得ることにつながる。多くの特徴を得ることで精度の高い学習が行われると考える。

4. 実験および検討

Optdigitの手書き文字データセットを用いて従来手法のGm-SOINNとエッジ接続機会を増やす提案手法を取り入れたものとの正解度の比較を行う。従来の手法であるGm-SOINNに比べ、提案手法では分布の特徴を得ることにつながると考えられるため精度が向上すると考えられる。

5. まとめ

従来手法のGm-SOINNにエッジ接続機会を増やす条件を与える手法を提案した。これにより多くの分布の特徴を得ることにつながり、従来手法より精度の高い結果を得られると考えた。

参考文献

- [1] S. Furao and O. Hasegawa, "An incremental network for on-line unsupervised classification and topology learning," *Neural Network*, vol. 19, no. 1, pp. 90–106, 2006
- [2] T. Kohonen, "The self-organizing map," in *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, Sept. 1990
- [3] H. Yu, J. Lu and G. Zhang, "Online Topology Learning by a Gaussian Membership-Based Self-Organizing Incremental Neural Network," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 10, pp. 3947–3961, Oct. 2020