

## SOINN+を用いた半教師学習の提案

日大生産工 ○福澤 隼人 日大生産工 山内 ゆかり

### 1. まえがき

近年機械学習において、ラベル付きデータの収集には高いコストが伴う一方で、ラベルなしデータは比較的容易に入手可能である。そのため、ラベルの有無にかかわらず効果的に学習可能な手法が必要とされている。そこで、少ないラベルを用いた学習法として自己教師学習やラベルを使用しない教師なし学習などが提案されてきた。その一つの自己教師学習であるSupervised Contrastive Learning(SCL) [1]は教師なし学習であるContrastive Learningから、自身で教師を生成し学習を行うモデルである。教師なし学習手法であるSelf-Organizing Incremental Neural Network (SOINN)[2]は自己増殖するノードとエッジを持つ逐次学習を可能にする手法として提案された。Chayut WiwatcharakosesらはそのSOINNを改良したSOINN+[3]を提案し、学習効率化やノード管理の改善が報告されている。しかし、SOINN+は教師なし学習を前提としており、ラベル付きデータを部分的に活用する仕組みを持たないため、限られたラベル情報を有効に利用できないという問題がある。

前述のSCLは教師あり学習からのアプローチであるが本研究では、ラベル付きデータを利用できない教師なしであるSOINN+に半教師あり学習の枠組みを導入したSemi-Supervised SOINN+(S-SSOINN+)を提案し、ラベル付きデータを活用し、クラス形成を誘導しつつラベルなしデータを取り込むことで汎化性能を向上させることを目的とする。数値データセットにおける実験により、従来のSOINN+と比較して分類精度及び学習効率の改善について報告する。

### 2. 従来研究

#### 2.1 Self-Organizing Incremental Neural Network

以下に従来研究につながる機械学習手法であるSelf-Organizing Incremental Neural Network (SOINN) [2]のアルゴリズムの更新部を示す式(1)、(2)に重み更新式を示す。

$$W_{s1}' = \frac{\varepsilon - W_{s1}}{T_{s1}} \quad (1)$$

$$W_i' = \frac{\varepsilon - W_i}{T_i \times 100} \quad (2)$$

ここで $\varepsilon$ は入力、 $T$ は勝利回数となっている。アルゴリズムをFigure 1に示す。

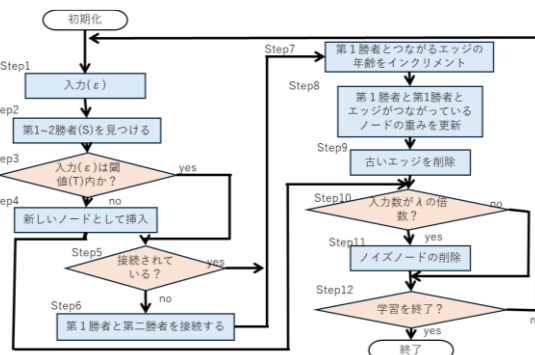


Figure 1 SOINNのアルゴリズム

#### 2.2 SOINN+

SOINN+[3]はSOINNのノード追加、重み更新の概念を継承するが、edgeの生成・削除方法およびノードの削除において根本的に異なる。より優雅な忘却を実現するために三つの概念を追加されている。

・アイドル時間・信頼性・ノードの有用性  
基本概念は反復回数 $t$ であり、ネットワーク更新サイクルを示す。

以下にSOINN+特有のアルゴリズムを示す。

#### アルゴリズム1:SOINN+の疑似コード

データ: $x$ :

- 1 ネットワーク初期化
- 2  $x =$  入力
- 3 類似度閾値の計算
- 4 if ノード追加条件が真 then
- 5     ノード追加
- 6 else
- 7     ノードの結合
- 8     ノードの連結
- 9     エッジ削除
- 10 ノード削除

#### ノードの連結

- 1 for すべてのノード $S$  do //信頼性を更新
- 2      $a_i$ に $i$ 番目のノードを代入
- 3      $\max(WT)$ にすべてのノードの最大勝利回数
- 4      $T(a) = \frac{WT(a)-1}{\max(WT)-1}$  (3)

```

5  $\tau_1$ に第一勝者(n1)の類似性閾値の平均を代入
6  $\tau_2$ に第二勝者(n2)の類似性閾値の平均を代入
7  $\sigma_1$ にn1の類似性閾値の標準偏差を代入
8  $\sigma_2$ にn2の類似性閾値の標準偏差を代入
9 条件1 ネットワーク内のエッジの数が3未満
10 条件2  $\tau(n1) \cdot (1 - T(n1)) < \tau_1 + 2 \cdot \sigma_1$ 
(4)
11 条件3  $\tau(n2) \cdot (1 - T(n2)) < \tau_2 + 2 \cdot \sigma_2$ 
(5)
12 if いずれかの条件が真 then //エッジを作成
13   if n1とn2の間にエッジがない then
14     エッジ作成
15      $\tau_1 \tau_2 \sigma_1 \sigma_2$ を更新
16   if 第1,2勝者間にエッジが存在 then
17     エッジ年齢LTに0を代入
18 Nにn1とエッジを持つノードの集合を代入
19 for 全てのN do
20    $LT(e_{n1ai})$ を1インクリメント

```

#### エッジの削除

```

1 Aに生きているエッジの寿命を代入
2  $A_{del}$ に削除済みのエッジの寿命を代入
3  $A_{0.75}$ にAの75パーセンタイルを代入
4  $\omega = A_{0.75} + 2 \cdot IQR(A)$  (6)
5  $\lambda_{edge} = A_{del} \cdot \frac{|A_{del}|}{|A_{del}| + |A|} + \omega \cdot (1 - \frac{|A_{del}|}{|A_{del}| + |A|})$  (7)
6 for N内のすべてのノード do
7    $a_i$ にNのi番目を代入
8   条件1  $LT(e_{n1ai}) > \lambda_{edge}$ 
9   if 条件1を満たす then
10      $A_{del}$ に $e_{n1ai}$ を加える
11      $e_{n1ai}$ を削除

```

#### ノードの削除

```

1 for すべてのノードS do
2    $a_i$ にiを代入//S内のノード
3    $U(a_i) = \frac{IT(a_i)}{WT(a_i)}$  (8)
4   条件1  $a_i$ が一つ以上エッジを持つ
5   if 条件1が真 then
6      $U$ に $U(a_i)$ を加える//つながり
        を持つノードの非有用度分布
        を作る準備
7   エッジの無いノードの集合  $I$  を作成
8    $\omega_{node} = median(U) + 2 \cdot sMAD(U)$  (9)
9    $R_{noise} = \left\lfloor \frac{I}{S} \right\rfloor$  (10)
10  $\lambda_{node} =$ 

```

$$\bar{U}_{del} \cdot \frac{|B_{del}|}{|B_{del}| + |S \setminus I|} +$$

$$\omega_{node} \cdot \left(1 - \frac{|B_{del}|}{|B_{del}| + |S \setminus I|}\right) \cdot (1 - R_{noise}) \quad (11)$$

```

11 for すべてのノードS do
12    $a_i$ にiを代入//S内のノード
13   条件1 一つ以上のエッジが存在
14   条件2  $U(a_i) > \lambda_{node}$ 
15   条件3  $a_i$ がエッジを持たない
16   if 条件すべてを満たす then
17      $a_i$ ノードを削除
18      $B_{del}$ に $a_i$ を加える
19      $U_{del}$ に $a_i$ を加える
20 for すべてのノードS do
21    $a_i$ にiを代入//S内のノード
22    $IT(a_i)$ を一つインクリメント

```

#### 3. 提案手法

本研究では、Semi-Supervised SOINN+ (SSOINN+)を提案する。まず、新たにラベル伝播率:LPを追加することで、ラベルありノードを中心に、自身のラベルを周りのノードに与える。SOINN+ネットワーク更新ではSOINN+の削除閾値 $\lambda$ を、ラベル情報をもとに変更を加える。そうすることで、よりラベルを有効活用したネットワーク構造にする。以下にラベル情報の伝播、 $\lambda$ の計算に使用する式を述べる。

$$Label_{next} = Label_i - LP \cdot n \quad (12)$$

ここで $Label_i$ はノードが保有しているラベルの濃度であり、値は1以下の実数であり、ネットワーク上のラベル保有ノードiから伝播先ノード $next$ に減衰率LPと距離nの乗算で得られた値をラベル濃度から引いた値を伝播する。

$$LabelRate = \frac{|Label|}{|S|} \quad (13)$$

ネットワーク全体のノード数 $|S|$ でラベル保有ノード数 $|Label|$ を割ることでLabelRateを求める。

$$LabelScore_i = \max(0, 2\max(Label_i) - \sum_c Label_{i,c}) \quad (14)$$

ラムダの計算に使用する $LabelScore_i$ はノードが保有しているラベルから最大濃度のラベルを抽出し、それ以外のラベルを保有しているときに相殺させることで、曖昧なノードの影響力を減少させている。

$$\lambda_{edge,node} = \lambda + \frac{1 + LabelScore_i}{1 + LabelRate} \quad (15)$$

SOINN+の削除閾値 $\lambda$ を計算した後、ラベル保有ノードの数、濃度で新たな閾値を計算する。これは、ネットワーク全体でラベルを一つも保有していない時、SOINN+の削除閾値そのものになる。Figure2にネットワーク上でのラベル伝播の様子を示す。

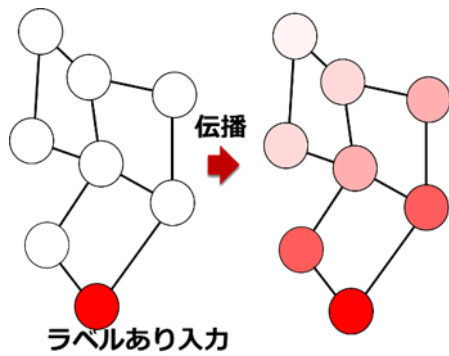


Figure 2 ラベルの伝播

#### 4. 実験および検討

データセットにyy5000とoptdigitsを用いたクラス分類タスクにおいて、従来手法のSOINN+との比較を行う。提案手法はS-SSOINN+(x)とし、xは教師データにラベルがついている割合を示している。yy5000とは、教師データ5000 ノイズデータ500 クラス数3の二次元データセットである。

Table 1 yy5000での実験結果

		node	Accuracy	Time(s)
SOINN+		502	1.000	5
SOINN+	S-S 0.50	294	1.000	2
	0.85	428	1.000	5

Table 2 optdigitsでの実験結果

		node	Accuracy	Time(s)
SOINN+		1026	0.968	7
SOINN+	S-S 0.50	1033	0.960	8
	0.85	965	0.969	6

従来手法のSOINN+と提案手法であるS-SSOINNの分類精度は大きく異なることはなかった。しかし、Table1ではラベル確率0.5、Table2ではラベル確率0.85のときにノード数が減少していることがわかる。これは、ラベルを持ったノードを効果的に活用し、より少ないノード数でネットワーク構築をしたからだと考えられる。その結果ネットワークの規模が縮小し、更新するための処理コストを抑えることができ、全体の処理時間が短くなった。だが、Table2のラベル付き確率が0.5のとき、従来よりもノード数と処理時間の増加、精度の低下している。これは各次元、データ数、クラス数によってラベル必要数やその他パラメータにウィークポイントが存在することが考えられる。

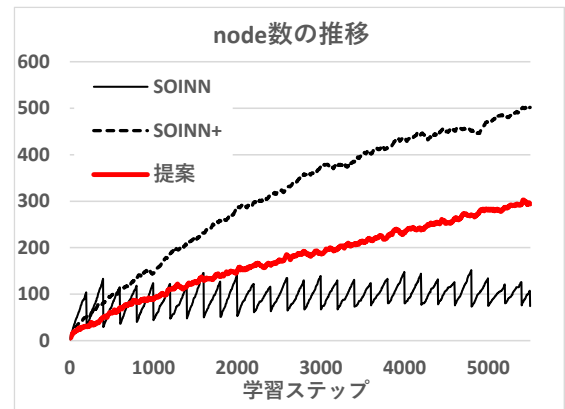


Figure 3 node 数の推移

S-SSOINNはパラメータの調整が必要ではあるが、ネットワークの縮小を実現しつつ、Table3からわかるようにSOINN+の安定した学習を継承した学習手法であるといえる。

#### 5. まとめ

本研究では、SOINN+の問題点であるラベル情報を用いた効果的なネットワーク構築ができないことに焦点を当て、半教師学習の枠組みをSOINN+に加えたS-SSOINN+を提案した。実験結果としてパラメータに左右されるものの、ノード数を削減することで処理時間を短縮しつつ分類精度を維持することができた。しかし、同じパラメータであってもデータセットによって従来手法よりも精度、処理時間が悪化することがわかった。今後はより高次元なデータにおけるS-SSOINN+での精度調査や、パラメータによってノード数の削減ができない問題点の改善を検討したい。

#### 参考文献

- 1) Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, "Supervised Contrastive Learning", arXiv:2004.11362v5 [cs.LG] (2021)
- 2) 山崎和博, 巻渕有哉, 申富饒, 長谷川修 「自己増殖型ニューラルネットワーク SOINN とその実践」 日本神経回路学会誌 17 巻, 4 号 (2010), pp. 187-196
- 3) Chayut Wiwacharakkoses and Daniel Berrar, "SOINN+, a Self-Organizing Incremental Neural Network for Unsupervised Learning from Noisy Data Streams", Expert Systems with Applications Vol. 143, (2020)
- 4) Chayut Wiwatcharakkoses, Daniel Berrar, "A self-organizing incremental neural network for continual supervised learning", Expert Systems with Applications Vol. 185, (2021)