

自己増殖型ニューラルネットワークによる属性値学習

日大生産工(数理情報工学科) ○中島 涼太 日大生産工 山内 ゆかり

1 まえがき

近年, ITの分野だけでなく様々な分野で人工知能の利用が検討されており, 人工知能の種類や活用方法は多岐にわたるようになった. 人間の脳は「大脳」「小脳」「脳幹」などいくつかの部位から構成されており, 部位ごとに役割が異なってくる. また, 人間の脳自体完全に解明されているわけではない. そのような状況の中で, 人間の脳を単体の人工知能で表現することは難しいと言える. したがって, 目的に応じていくつかの人工知能を組み合わせることで人間の脳を表現しなくてはならない.

人工知能が使われている具体的な例としては, 「自動運転」「株取引」「チャットボット」「音声認識」など様々である. それに対して機械学習(Deep Learning), ニューラルネットワーク, 遺伝的アルゴリズム(GA)をはじめとした多彩な人工知能の手法が提案されている. その提案されている手法の一つとして自己増殖型ニューラルネットワーク(Self-Organizing Incremental Neural Network)があり SOINN と呼ばれている. SOINNはSelf-Organizing Map(SOM)とGrowing Neural Gas(GNG)に着想を得て構築した教師なし学習である. 教師なし学習は, クラスタリングとも呼ばれ, 代表的な手法としてk-means法やEMアルゴリズムなどが挙げられる. また, SOINNはノイズを自動的に除去するためノイズ耐性が高く, 逐次学習可能な人工知能である.

画像, 音声, テキストなど幅広い種類のデータを複合的に処理することができるため, 東京工科大学やインド工科大学, 米軍, アメリカ科学財団など海外からの注目を浴びている.

実世界のデータは分布の形状が非定常であり, さらにデータを継続的に入力される場合が非常に多い. したがって, それらを処理するためにはオンラインに対応していることと逐次的に処理する能力が求められる.

このような事態を想定し, 初期のSOINNには2層のネットワークから構成されていた. これにより, 非定常なデータからデータを抽出す

る能力や, 複雑なデータに対する適切なクラスの割り当てを実現することを可能にした.

2 研究背景および従来手法

山崎らは先に述べたようなSelf-Organizing Incremental Neural Networkを提案し, SOINNによる実世界のデータ処理の効率化を目指した.¹⁾

以下の図1にSOINNアルゴリズムの初期化から結果の出力までの手順を示す.

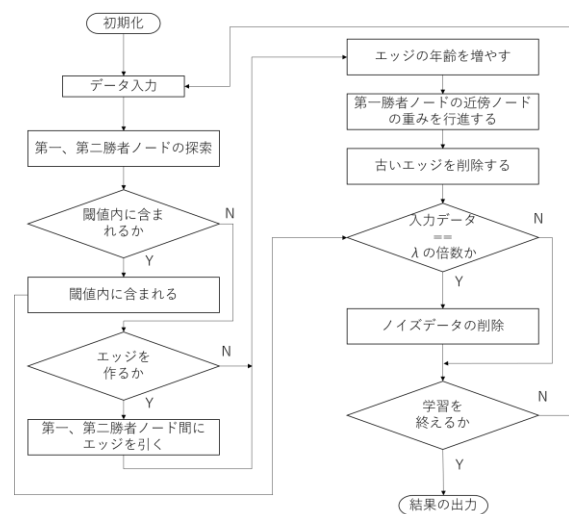


図1 SOINNアルゴリズムのフローチャート
木村らはKSOINNという手法を提案した.²⁾

KSOINNとはSOINNを改良したもので, 大きく2つの部分を改良している.

1つ目は, ノードの挿入や削除, およびノードの重みベクトルを調節する処理の改良である. これは従来のSOINNの入力に対する不安定性を軽減する目的で行われた. 具体的な手法としては, ノードに年齢を導入することで解決した. ノードに年齢を導入することで全ノードに対するノイズ削除処理のステップ数が同じになる. このようにすることで削除処理の中で早い段階で入力されたデータと遅い段階で入力されたデータとでノイズの処理をする際に起こる不安定性の軽減に成功した.

2つ目は, ノードにクラスの統計情報を追加した点である. 従来までのSOINNではクラス

Attribute value learning by Self-Organizing Incremental Neural Network

Ryota NAKASHIMA and Yukari YAMAUCHI

と同数のSOINNが必要となり,学習後にはデータの量や学習時間がクラス数と比例して増加してしまうという問題点があった.そこで,クラスに関する教師信号を「ラベル」としてノードに追加することと,ネットワークの変化に応じてラベル情報の変更を行うことを機能として追加した.こうすることで,1つのネットワークですべてのクラスの情報を管理することが可能になり,ネットワークの削減ができ,従来よりも効率的に学習することが可能となった.

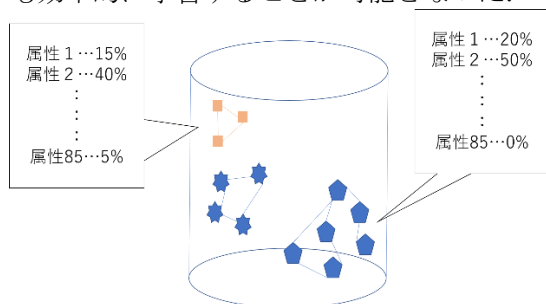


図2 ラベル情報の概念

3 提案手法

従来研究における問題点としては,認識率の低さが低いということが挙げられる.

様々な SOINN の手法が提案されている中,SOINN の認識率は依然として高くても 4 割程度である.しかし,従来研究ではそれを下回る 2 割程度である.この認識率の低さとして○が原因であることが挙げられる.導入する SOINN の数を減らしエッジを引く条件を変更することで,処理速度を向上させながらも認識率の維持をすることに成功していた.しかし,認識率は依然として 2 割前後であった.

本研究では,SOINNの学習する際に使用する,属性値パラメータに関する式の見直しを提案する.そして,見直しをしつつその都度数値を測定することで認識率の向上を目指す.

そこで,提案手法では処理速度はそのまま認識率の向上を目標とした.

具体的な手法としては,属性値のラベルに関する式の見直しを主として認識率の向上を目指す.

ラベルに関する式 (1)である.

$$g = \underset{z \in Z}{\operatorname{argmax}} \left[\sum_{m \in M} \left(\sum_{q \in Q} \left[P^m(V_q) - \frac{1}{2} (R_q^{m+} + R_q^{m-}) \right] \right) * (2a_z^m - 1) \right] \quad (1)$$

ここで,推定したいクラス, Z : クラスの集合, M : 属性の集合, Q : 特徴の集合, V_q : 入力画像の特徴 q , $P^m(V_q)$: 特徴パターンの近傍重みベクトルを求め,そのノードが持つラベルの平均, R_q^{m+} : 学習クラス中の属性 m と関連度が高いクラス, R_q^{m-} : 学習クラス中の属性 m と関連度が低いクラス, a_z^m : クラスと属性の関連度,を表す.

(1)を変更していきながら,提案手法と従来手法お互いの認識率を記録する.そして,記録した数値がなぜそうなったのかを考察しながら数値の向上を目指し改良していく.

4 実験結果および検討

本実験では,3823個のデータを用いて SOINN空間に挿入,削除,エッジを描画するなどの操作を行った.挿入したデータは64次元である.

用いたデータに対してラベルの付加を行う学習フェーズと,その学習フェーズで学習させたSOINNを使用してテストデータと近いノードのラベルを用いて行う推定フェーズを行う.

5 まとめ

本研究では,SOINNの学習する際に使用する,属性値パラメータに関する式の見直しを提案した.

しかし,従来手法を完全に実装できたとはいえないので,従来手法の実装が終わり次第再度考察を行う.

【参考文献】

- 1) 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. The Brain & neural networks 17-4(2010-12)
- 2) 木村 大毅, PICHAI Kankuekul, 長谷川修,属性知識の高速オンライン学習と転移による未知クラス推定, 知能と情報 26(5), pp.830-843,2014