

# 人工知能アルゴリズム探検隊

第8回 答えを学習しなくても特徴を予測できる  
「自己組織化マップ」

ご購入はこちら

牧野 浩二, 寺田 英嗣

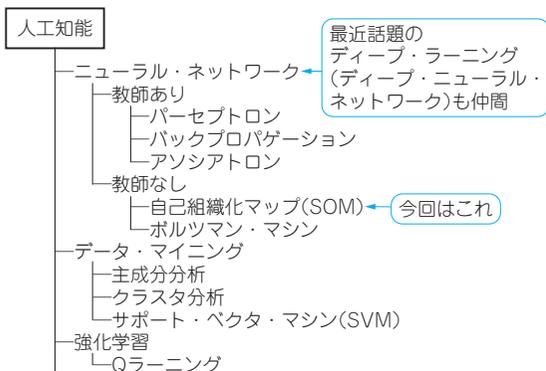


図1 人工知能のアルゴリズムあれこれ

## ● これまでに紹介したアルゴリズムとの違い

図1に人工知能のアルゴリズムを整理しました。自己組織化マップ(以降、SOM)は、ニューラル・ネットワークの教師なし学習に位置付けられます。ニューラル・ネットワークについては、今後の連載で解説していきます。ここでは、これまで紹介したアルゴリズムとの違いを説明します。

▶主成分分析…多くの評価項目を2次元の散布図で表す  
主成分分析は、「どのように分布させれば人間に分かりやすくグループを示せるか」ということが主題となっています。この分析手法は、たくさんの評価項目をうまくまとめて新しい指標を作るため、データを2次元の散布図で表したときの縦軸と横軸には意味を付けることができます。

▶クラスタ分析…設定した数のグループに分ける  
一方、クラスタ分析は、設定した数の「グループに分ける」ことが目的となっています。これまでのクラスタ分析の記事で、2次元で表現できていたのは主成分分析の機能を併用していたからです。

この2つは全く異なる観点でデータを整理・分類しているのですが、結果として2つには関係している点が面白いです。

主成分分析とクラスタ分析は、主にアンケート結果の分析やプロジェクトの評価に利用されています。

▶SOM…マップを作り新たなデータをマップに当てはめることが可能

今回のSOMは、似ているデータを集めて「2次元の表(マップ)」を作ることが目的となっています。2次元で表現する点は主成分分析に似ていますが、縦軸や横軸に意味はありません。似ているデータが集まるので、クラスタ分析の結果に関連性があります。

SOMによって作られたマップ上に新たなデータを

## ● 実際の需要予測等に使われていて進化中

自己組織化マップは英語でSelf-Organizing Mapといい、その頭文字をとってSOMと略されています。これは1980年代前半にT.コホネン(Teuvo Kohonen)が提案した手法で、自己組織的に(教師データがなく与えられたデータだけで自動的にという意味)いろいろなデータが分類されることで話題になりました。

しかし、当時のコンピュータが非力であったため、分類できるデータの数に限界がありました。そのため人間が想像できる程度の分類にしか役立ちませんでした。

2000年代に入りコンピュータの性能が飛躍的に向上したのをきっかけに、多数のデータが扱えるようになったため、第2次ブームが起きました。これにより分類するだけでなく、学習しておいた分類結果に新たなデータを入力すると、そのデータの性質や特徴を予測できるようになりました。例えば、脈拍解析やゲノム解析、電力需要の予測や気象予測、土砂生産量など、工学の分野を超えて利用されています。そしてさらに、今回紹介する2次元の平面マップだけでなく、マップを球面にしたりトーラス(ドーナツ型)にしたりなど、今でも改良が進んでいます<sup>注1</sup>。

注1: これを応用した研究として、本誌2016年7月号 特集 第8章「研究! 生体センシング×機械学習」で触り方を判別する方法を紹介しました。その紹介では、T.コホネンのグループが公開している自家本元のプログラムを使ったものでした。