AI(人工知能)(2)

2023年1月21日作成

三木

◆音声や画像の認識

AIによる音声や画像の認識は、すでに一般的になった。例えば、スマホの音声入力や、デジカメの顔ピント自動調整など。さらに、個別の機器に組み込まれるだけでなく、単体のアプリケーションソフトとしても市販されている。

では、その原理とは？

◆認識の原理

AIによる認識機能は、人間による認識機能を模倣している。図1のように、AIの認識機能は、入力層、隠れ層、出力層で構成される。入力層はセンサーのようなもの、隠れ層はセンサーの信号を処理するもの、出力層は結果を表すものである。

入力層

隠れ層

出力層

図1

さらに、図2のように、入力層には1層×A列の要素があり、個々の要素は1つの出力をする。隠れ層にはN層×B列(B1・・BN)の要素があり、個々の要素は複数の入力を受けて1つの出力をする。出力層には1層×C列の要素があり、個々の要素は複数の入力を受けて1つの出力をする。

入力層

隠れ層

出力層

・

・

・

・

・

・

・

・

・

・

・

・

・

・

・・

・・

・

・

・

・

・

・

A

B1

C

BN

N

図2

隠れ層や出力層の要素は、パーセプトロンと呼ばれ、人間のニューロンに相当する。また、図2のように、パーセプトロンを多層にしたものは、多層パーセプトロンモデルと呼ばれる。

図3のように、要素は複数の入力を受けて、それぞれに重み付けをした結果を出力する。

7

5

×0.5

×0.3

5.0

図3

重み付けの初期値は、入力層と出力層に与えられるデータから誤差逆伝搬法と呼ばれる方法で計算される。図3のように、1つのパーセプトロンが2つ重み付けをもつ場合、2組の入力値と出力値があれば計算できる。しかし、要素が増えるほど、また要素が受ける入力値が増えるほど、データは急激に増える。例えば、イヌとネコの画像を識別する場合、少なくとも1万個程度のデータが必要とされる。

重み付けの初期値を計算することを、学習と呼ぶ。学習が終わると、入力値を与えて出力値を得られる。出力値に間違いがあると、重み付けが再計算され、精度が上がる。

現在では、コンピューターの性能が向上したため、多層パーセプトロンモデルをさらに多層にした、いわゆるディープラーニングモデルも使われるようになった。これにより、音声や画像を精度良く認識できるようになってきた。

◆AIの利用

現在のAIの要素数は、人間に比べて1/1000程度に過ぎないと言われる。しかし、人間の要素数は増えないが、AIの要素数は技術の進歩によって増やすことができる。そのため、将来はより複雑な認識もできるようになるだろう。

しかし、AIによる認識が進歩しても、AIは、それを人間の仕事や生活にどのように活用すればよいのかを教えてくれない。それは、あくまでも人間の範疇である。

◆参考図書

・人口知能育成ブック、染谷悠一郎、日経BPマーケティング、2019年3月

以上