

ディープラーニングの話

江越 航

1. 人工知能

近年、AIや人工知能、ディープラーニングという言葉が毎日のように話題になっています。AIとはArtificial Intelligence、つまり人工知能のことで、人工知能は今やすっかり私たちの生活の中にとけ込み、様々な製品に使われています。

画像からその中に映っているものを自動的に判別し、工場での品質管理に利用されたり、顔認証で通る改札が作られたりしています。画像処理の技術は著しく向上し、いずれ車は自動運転で走るようになるでしょう。

音声認識の技術も進んでいます。スマホに話しかけると、声を正確に文字に変換するだけでなく、文章の意味を解析し、答えを返してくれます。さらには翻訳技術も進み、日本語を英語に直すことも、精度よくできるようになりました。

囲碁や将棋では、コンピューターがプロ棋士を負かすまでに強くなりました。近頃は、囲碁・将棋番組を見ると、どちらが優勢か、コンピューターが判断した局面が表示されています。

つい最近ではChatGPTというサービスが登場しました。これは、分からないことや調べて欲しいことを問かけると、まるで人間が答えてくれたかのように、詳細な答えを返してくれます。

人工知能は急速に進歩して、あらゆる分野に応用されています。AIを応用した製品は、あたかも魔法が使われているように感じるかもしれません。そこでは一体、どのような技術が使われているのでしょうか。

AIによる応用例とともによく聞くのが、機械学習やディープラーニングという言葉です。近年の人工知能の進歩は、ディープラーニングという技術に支えられています。今回は、ディープラーニングとはいったいどのような技術なのか、紹介したいと思います。



図1. 顔認証機能を使った自動改札機

2. ディープラーニング

ディープラーニングとは、人間の脳の働きをまねたニューラルネットワークという手

法を利用して、複雑なパターン認識や予測を行う技術です。

ディープラーニングでは、機械学習が行われます。機械学習とは、データからルールを獲得する、という仕組みです。多数のデータからコンピューター自身がルールや知識を学習し、学習の成果に基づき、予測や認識を行います。

この学習を行う時に、ニューラルネットワークという技術が用いられます。ディープラーニングは、ニューラルネットワークを多層に重ねて用い、複雑な学習を可能にしています。ニューラルネットワークの層を深く重ねることから、日本語では深層学習とも呼ばれています。

近年、ディープラーニングが広く使われるようになった背景として、1つは計算機の指数的な性能向上が挙げられます。半導体技術の進歩に関する有名な経験則に、「ムーアの法則」というものがあります。これは、半導体回路の集積密度は1.5年で倍になる、というもので、簡単に言えばコンピューターの性能は1.5年で倍の速度に進歩しています。20年前のスーパーコンピューターの性能は、今、使っているスマートフォンにも劣ってしまいます。かつては計算性能の制約で解けなかった問題も、現実的なコスト・時間で解けるようになってきました。

特に最近のパソコンには、GPU(Graphics Processing Unit)と呼ばれる演算装置が使われていますが、これがディープラーニングの計算にも使われます。GPUは、もともとはCGの描画専用装置で、ゲーム専用PCなどに搭載されています。これが並列計算処理を高速に実現することから、ディープラーニングの計算にも有用なのです。

ディープラーニングが広まったもう一つの背景として、データの爆発的な増加も挙げられます。ディープラーニングを行う際には、大量のデータが必要になります。例えば、犬と猫を見分けるなど、画像認識のためには、数百万枚の画像を学習させる必要があります。

現在はビッグデータという言葉にも表れているように、インターネットにより膨大なデータがやり取りされています。さらに磁気ディスク、フラッシュメモリなどの記憶装置の値段の低下、大量のデータを蓄積できるクラウドの登場も、ディープラーニングの発展につながっています。

3. ニューラルネットワーク

ニューラルネットワークは、生体の脳の情報処理の仕組みをコンピューターに応用したものです。

図2は脳の神経細胞(ニューロン)の構造を示した模式図です。人間の脳には1000億個ものニューロンがあり、相互に電気信号をやり取りしています。脳全体の巨大なニューロンのネットワークの中で電気信号がやり取りされることで、体全体を制御し、様々な人間活動を行うことが可能になります。

ニューロンでは細胞体から樹状突起と呼ばれる組織が伸び、他のニューロンとつな

がって、信号を受け取るようになっています。他のニューロンとの接合部をシナプスといいます。シナプスでは、いくつかのニューロンからの信号を受け取ります。受け取った信号は統合され、軸索を通して、さらに他のニューロンへと送られます。

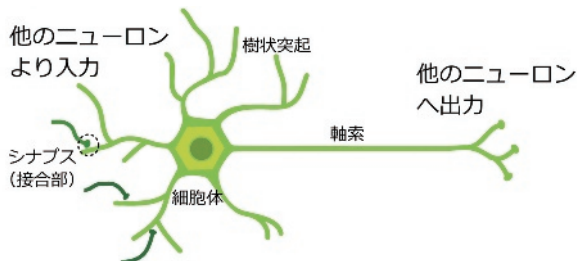


図2. 神経細胞(ニューロン)の模式図

シナプスを通して受け取る他のニューロンからの信号が大きければ、軸索を通して別のニューロンへ送る信号の大きさも大きくなります。

このニューロンの働きを、コンピューターの中で再現したのが、図3のニューロンのモデルです。他のニューロンからの信号が、図中のニューロンに与えられます。それぞれのニューロンからの信号が加わり、他のニューロンへと信号が出力されます。

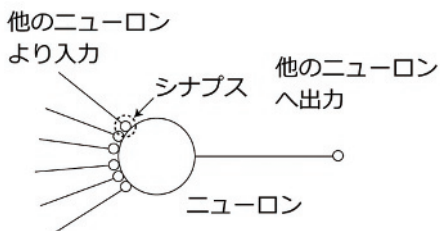


図3. ニューロンのモデル

このニューロンのモデルを多数組み合わせ合わせたのが、図4のニューラルネットワークです。図4は、入力層・中間層・出力層の3つの層からなるニューラルネットワークになっています。単純な働きをするニューロンを多数合成することで、全体としてより複雑な制御が可能になります。

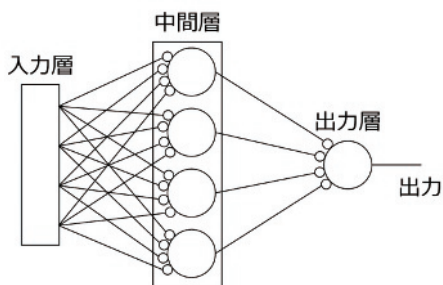


図4. ニューラルネットワークのモデル

図4では中間層は1つだけですが、さらに中間層を増やすこともできます。中間層が2つ以上あるニューラルネットワークのことを、ディープニューラルネットワークと呼んでいます。ディープラーニング(深層学習)とは、ニューラルネットワークのなかでも、層が深いネットワークを使って学習を行うものなのです。

4. アヤメの分類

ニューラルネットワークでは、コンピューター自身がネットワークを調整する、学習という操作が行われます。どのように学習を行うのか、アヤメの分類を通して紹介しま

す。機械学習の教科書を見ると、しばしばこのアヤメの分類の例が取り上げられています。これは、実は機械学習のプログラムのライブラリの中に、既にサンプルデータとして含まれているためです。

アヤメにはいくつかの種類があります。今回は、Iris setosa(ヒオウギアヤメ)とIris versicolor(ブルーフラッグ)の2つに分類することを考えます。分類のために、外花被片と呼ばれる部分の長さと幅を測ることにします。

図6のグラフは、様々なサンプルに対して、外花被片の長さと幅を測り、プロットした結果です。黄色の丸がIris setosa、青い丸がIris versicolorの結果です。グラフをみれば分かるように、中央の直線を境に、左上にIris setosa、右下にIris versicolorが集まっています。外花被片の長さから、2種類のアヤメを分類することが可能であることが分かります。

この分類を、ニューラルネットワークを使って行うことにします。使うニューラルネットワークは、図7のように、入力が2つ、出力が1つのごく単純なモデルです。

2つの入力には、各サンプルの、外花被片の長さの幅の値を入れます。入力された値は、そのまま足し合わせるのではなく、重みを掛け算して足し合わせます。式で書くと

$$[\text{出力}] = [\text{入力1}] \times [\text{重み1}] + [\text{入力2}] \times [\text{重み2}] + [\text{定数}]$$

となります。この結果の値が、例えば0以下なら、Iris setosa、0より大きければ、Iris versicolorと判断することになります。

サンプルのデータを順番に入れて行き、間違った判断がされた場合は、重みの値を少し調整します。これが学習に相当します。この操作を繰り返して少しずつ重みの値を調整していけば、最終的



図5. Iris setosa(ヒオウギアヤメ)
<https://ja.wikipedia.org/wiki/ヒオウギアヤメ>
の図に一部加筆

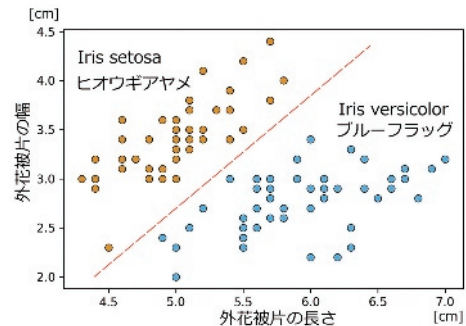


図6. Iris setosa(ヒオウギアヤメ)とIris versicolor(ブルーフラッグ)の分類

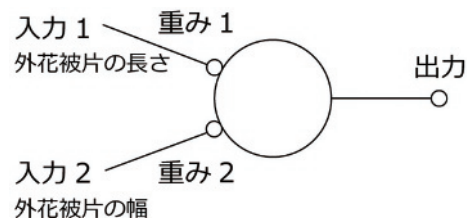


図7. アヤメを分類するニューロン

には、外花被片の長さから、Iris setosaなのか、Iris versicolorなのか、2種類のアヤメを正しく分類できるようになります。

これは図6のグラフでは、2つのアヤメを分ける直線を見つけることに相当します。重みの値を図6の直線の係数にすれば、図7のニューラルネットワークで、正しく分類できることになります。

5. 手書き数字の分類

次に、図8のような手書き数字を分類することを考えてみましょう。この分類も、機械学習の教科書でしばしば登場します。やはり、サンプルデータとして含まれているためです。

さて、アヤメの分類の際には、外花被片の幅と長さを入力しましたが、手書き数字の場合はどうすればいいでしょうか。図9の左側は、手書き数字の「0」のサンプルの数字の1つを拡大したものです。この手書き数字

は、8×8のマスキに書かれた、全部で64個の小さな四角形を組み合わせることで表現されています。64個の四角形は、白から灰色、黒と色の濃さが違ってきます。この濃さは16段階で表現されており、それを数値に直したのが、図9の右側です。

ニューラルネットワークでの分類は、入力としてこの64個の数字を一列に並べたものを使います。アヤメの分類の場合と同じように、64個の数字それぞれに重みの

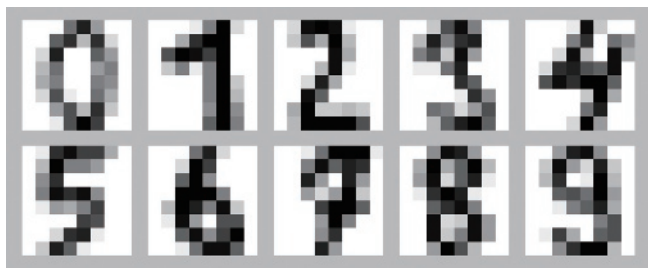
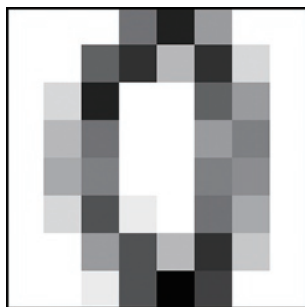


図8. 手書き数字



0	0	0	9	14	6	0	0
0	0	10	13	4	13	2	0
0	2	14	0	0	10	6	0
0	4	9	0	0	6	8	0
0	5	8	0	0	8	7	0
0	2	11	1	0	9	5	0
0	0	6	11	4	13	3	0
0	0	1	11	16	12	0	0

図9. 手書き数字の0とそれを表す数字

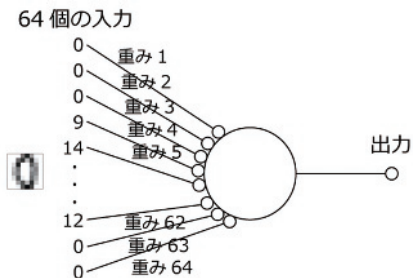


図10. 手書き数字の分類ニューロンのモデル

値を掛け算して足し合わせ、それがある値より大きければ、この手書き数字は「0」、そうでなければ、「0」ではない、と判断します。多くの手書き数字を用意し、それが「0」と書かれたものか、そうでないかを学習することで、コンピューターが手書きで書かれた数字を判断できるようになるのです。

手書き数字が「0」であるか、そうでないかを判定するだけなら、図10のように、ニューロンが1つだけのニューラルネットワークでも可能です。しかし、書かれた数字が「0」から「9」のどれに相当するのか、ということ判断するためには、さらに多数のニューロンを組み合わせたニューラルネットワークが必要になります。でも、こうした単純な仕組みを多数組み合わせることで、いろいろな分類をすることが可能になります。ニューラルネットワークの特徴は、簡単な関数を多数組み合わせることで、複雑な関数を実現することにあります。

このとき、重みの値を自動的に探すことが、ニューラルネットワークをうまく働かせるポイントになります。図7のアヤメを分類するような単純なニューラルネットワークなら、重みの値を見つけるのは比較的簡単です。しかし図4のような複数の層をもつネットワークになると、それぞれニューロンにつながる重みの値を見つけるのが、非常に難しくなります。ニューラルネットワークの学習には、出力された値と正解となる値を比べて、その差を重みに反映する、誤差伝搬法という方法が使われています。

6. 進化するディープラーニング

実はニューラルネットワークは、つい最近まで、学習が難しいと考えられていました。実際に実行させても、途中で学習が止まってしまう、しかもどこを修正すればいいかも分からなかったのです。しかし、層の数を増やしても適切に学習させる仕組みが見つかったことから、一気に注目を集めるようになりました。

ディープラーニングは、単に計算性能の向上だけでなく、学習が早く進む方法や、学習した結果をいろいろな場面に応用する方法が開発されてきました。ディープラーニングが広く使われている分野の1つに画像認識がありますが、ここではニューラルネットワークをさらに進化させた、畳み込みニューラルネットワーク(CNN)という手法が開発され、広く使われています。また最近新しい手法として、言語処理のためのトランスフォーマー(変換器)と呼ばれる手法が脚光を浴びています。このトランスフォーマーを用いたGenerative Pre-trained Transformer(事前学習済み文章生成トランスフォーマー:GPT)と呼ばれる高性能な言語モデルが開発されています。この言語モデルを用いたのが、冒頭で触れたChatGPTというサービスです。

ディープラーニングの分野は、日々進化をしています。すでに多くの分野で、人間を凌駕する性能を発揮しています。こうした機能を脅威と感じるのではなく、むしろ人間の生産性を向上させる道具として、うまく使っていくことが必要になると考えられます。

江越 航(科学館学芸員)