

機械学習技術の活用

1270381 宮本武

2024 年 10 月 29 日

1 はじめに

機械学習とは、収集したデータの集合に対して機械などの計算資源を用いた分析を行うことでデータ間の関係性、意味を抽出し、物事の予測や価値の創出を自動的に行う手法、方法論である。収集したそのままのデータはただの数値の羅列である。このようなデータに数学的な処理を施すことでデータ間の線形的な相関や二次曲線的な相関といった意味を見出すことができる。この時、データ量は膨大であり、人間がこのデータを操作することはできない。そのため、代わりとして機械がデータを操作する必要がある。この操作により明らかになった関係から法則を導くことで、様々な形で応用することができる。このような価値を生み出すのが機械学習である。

2 実習内容

機械学習の実習では、手書き文字認識を行った。手書き文字認識は手書きした文字の画像から書かれた文字を認識させる技術である。実体は画像データであるため、テキストデータとは質的に異なる。この技術を用いることで、手書きの文字を識別させ、コンピュータが扱える文字のビットデータに変換することができる。これによって、手書きした文章を撮影した写真をコンピュータに取り込ませることによって、キーボードなどによる入力無しでコンピュータ上で文章を取り扱ったり、物に書かれた単語や識別子によって物を自動的に分類したりすることができる。また、手書き文字認識技術を文字の特徴を抽出する用途として使用すれば、筆跡を識別することもできる。これによって筆跡鑑定や平仮名をより丁寧に書くコツなどを知ることができる。このような手書き文字認識の技術を使用するためには、認識させるモデルを選択する必要がある。モデルに使用する技術として、ロジスティックス回帰、サポートベクトルマシン (SVM)、ニューラルネットワーク、深層学習などがある。使用する技術によって、精度や出力方法に違いが生まれる。そして、認識させる際にはモデルが求める入力に合致した形式に整えて、妥当な方法で出力データを取得する。文字を識別させる場合は入力文字が想定される各文字にどれほど確からしいかの値を取得できる。この情報を用いることで、文字の識別を行うことができる。

3 考察

手書き文字認識の実習では 0~9 の数字を識別するモデルを単層パーセプトロン、多層パーセプトロン、深層学習の複数の技術を用いて作成し、手書き文字を識別させた。いずれもニューラルネットワークを使用したモデルである。テストデータを使用して得たモデルの識別率によって性能を比較すると、単層パーセプトロンが最大約 0.73 で最も低く、深層学習が最大約 0.93 で最も高い結果となった。深層学習はより深い深層学習を合わせた 2 種類のモデルを使用した。より深い深層学習がより高い性能を発揮する傾向があった。0.93 はより深い深層学習の結果である。このような結果になった理由の一つとしては、層の深さによる違いが挙げられる。層が浅いと、一つのピクセルの画像全体に対する関係が直接的になり、思考プロセスを十分に行えない。それに対して、層が深いと段階を追った複雑な思考プロセスが可能となる。そして、二つ目の理由として、重みの個数の違いが考えられる。層が多いとニューロンの接続が増えて、重みの数もそれに応じて増える。重みが多いほど記憶容量が増えるため、より汎用性の高い識別が可能になると考えられる。実際に手書き文字を複数書いて識別させてみたところ単層パーセプトロンが識別できなかった文字を深層学習によるモデルが識別できることが見受けられた。

以上のような手書き文字を識別するモデルを作成する過程で発生した注意点として、データの構造と内容の妥当性が挙げられる。画像データは普通 2 次元リストの構造で保存されるが、処理段階では高次元のベクトルに変換したり、更に画像データと教師データを結合したりと、様々な場面でデータ構造の変換を行う必要があった。モデルの実体はプログラムであり、適切にデータ構造を変換する必要があるため、各処理段階のデータ構造が適切であるかに注意する必要がある。そして、入力データの内容はモデルに対して妥当である必要がある、つまり、入力させる文字画像が確実に目的とする形式に沿っている必要がある。識別させる文字画像に識別可能な文字が書かれていたとしても、文字のサイズが極端に小さかったり、位置がずれていたたりするとその画像をそのまま入力したとしても正しく識別されない可能性がある。それを防ぐために、モデルができるだけ識別しやすいように切り取ったり、移動したりといった画像の加工を行う必要がある。手書き文字の識別では総合的な性能はモデルの仕組みだけではなく入力データの形式も大きく関係することに注意する必要がある。

4 まとめ

識別率の観点からみると、基本的にモデルの性能はよりニューラルネットワークの層が深いものが高いということが分かった。しかし、層の数が多い分、より多くの計算を必要とするため、十分な性能の計算機と時間が必要となる点では常に最適であるとは言い切れない。

現在、文字認識技術は文字を識別するという目的で使用する場合がメジャーだが、将来、計算機の性能や機械学習の技術が向上し、より高精度の手書き文字認識が可能となれば、識別に加えて文字そのものがもつ特徴を抽出する目的で利用されるようになって考えられる。「実習内容」で述べ

たことも含めるが，例えば文字の特徴を抽出して個人の筆跡を特定したり，様々な特徴をミックスした文字を作り，デザインに生かしたといった用途が考えられ，今後文字認識が更に幅広い目的で利用されるようになって考えられる．

参考文献

- [1] 須山敦志, 「ベイズ推論による機械学習入門」, 講談社, 2017
- [2] 須藤秋良, 「スッキリわかる Python による機械学習入門」, インプレス, 2020