

Recurrent Neural Networks

八十田 周作
Shusaku YASODA

1 はじめに

近年、ニューラルネットワークを利用した機械学習が普及している。ニューラルネットワークは、機械学習の性能を高め、ロボットの感情表現や自動車の運転支援技術などで成果をあげている。しかし、従来のニューラルネットワークは情報の前後関係を保持できず、音声や動画などの時系列データを利用するのは難しい。Recurrent Neural Networks(RNN) は情報の前後を考慮し、記憶しながら処理をするニューラルネットワークである。

2 Neural Networks

ニューラルネットワークは、人間の脳内にある神経回路網を模した数式的モデルである。従来のニューラルネットワークの構造を Fig. 1 に示す。

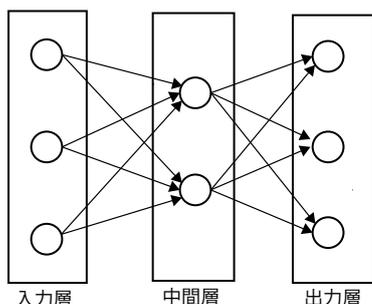


Fig.1 ニューラルネットワークの図

ニューラルネットワークは入力層、中間層、出力層から構成される。ニューラルネットワークは重みとバイアスを調整して学習する。重みとは層と層のつながりの強さを表す。バイアスはニューラルネットワークの好き嫌いや得意不得意といった偏りを表す。偏りを持たせることで複雑な意思決定の際に、答えを意図した方向への変化を加えることができる。ニューラルネットワークは入出力を繰り返すことで、重みとバイアスを修正し精度を上げていく。

ニューラルネットワークによってコンピューターが苦手とする画像認識や機械翻訳などの問題を解くことが可能である。しかし、従来のニューラルネットワークは入出力で扱えるデータの大きさが決まっている。そのため、音声や動画などの大きさが一定でないデータを扱うことが難しい。また、ニューラルネットワークは、複数の連続するデータの関係性を考慮することができない。

3 Recurrent Neural Networks

3.1 概要

RNN は、音声、動画、文章などの時系列データを扱うことのできるニューラルネットワークである。従来のニュー

ラルネットワークは時系列データを扱うことができない。一方、RNN は入力値、出力値に可変長のデータを扱うことができる。

RNN は中間層の情報を保持して、次の入力時に保存した中間層の情報を利用できる仕組みを持つ。文章を単語に分けた場合、時系列に順々に表れた複数の単語から次に出てくる単語の推測が可能となる。

3.2 RNN の仕組み

RNN は中間層を保存することで、次の入力時に保存した中間層の情報を統合して出力結果を出すことができる。一般的な RNN の構造を Fig. 2 に示す。

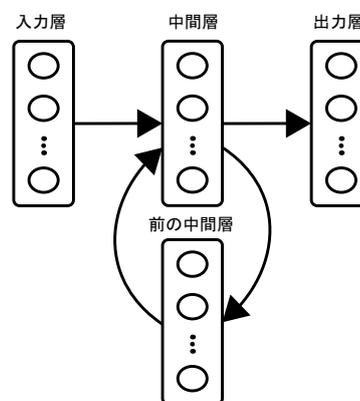


Fig.2 RNN の図

RNN は中間層の部分が従来のニューラルネットワークとは違い、前回の中間層の状態を保持する仕組みを持つ¹⁾。保持した中間層の情報は次回からの入力層からの状態に統合することができる。RNN の入力値を (1) 式に示す。

$$s_t = f(Ux_t + Ws_{t-1}) \tag{1}$$

s_t : 中間層の値
 f : 非線形関数
 U, W : 重み
 x_t : t ステップ時の入力値

(1) 式では時刻 t における入力値と時刻 t までの中間層の値を統合している。RNN では時刻 t に入力されたデータは時刻 $t-1$ のデータと統合し、中間層を経由した後出力される。また時刻 t の中間層の状態も保存され、次回からの入力と統合する。以上のように、前回の中間層の情報と入力層からの情報を合わせることで、情報に前後関係を持たせることができる。

例えば、「今日の天気」の次にくる単語は「が」、「を」、「は」が候補として挙げられる。従来のニューラルネットワークでは直前の単語の天気が続く単語しか予測できず、「今日

の」は考えることはできない。しかし、RNN の場合は「今日の」まで再帰して考えることができる。

3.3 問題点

ニューラルネットワークは重みとバイアスを修正する際に、誤差情報を出力層から入力層に向かって伝播する。誤差情報は最初の層まで遡るが、誤差が遡る層が多いと不必要な誤差情報が増える。

誤差伝播の際は誤差情報に層ごとの重みをかけ伝播する。かける重みの値が0に近いと誤差が急速に小さくなり消失、あるいは急速に大きくなって発散し、学習ができなくなる勾配消失問題が生じる²⁾。RNN は理論上は長期的なデータも利用できるが、実際は10ステップ以上遡ることが難しい。

3.4 解決策

勾配消失問題を解決策として Long Short Term Memory(LSTM) が提案されている。LSTM は情報を削除したり、セル状態に追加する能力を持ち、ゲートと呼ばれる機能を中間層に持つ。LSTM の中間層の構造を Fig. 3 に示す。

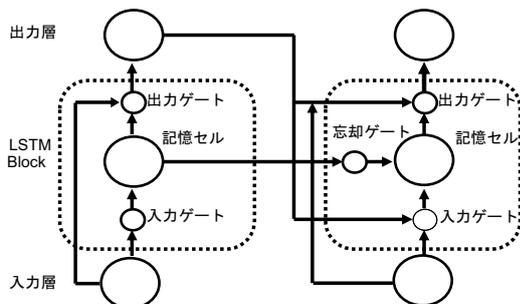


Fig.3 LSTMBlock の構造

LSTM は RNN の中間層が LSTMBlock と呼ばれるメモリに置き換えられたものである。LSTMBlock は入力、出力、忘却ゲート、記憶セルで構成される。

記憶セルはデータ全般を記憶する場所で、ゲートと組み合わせることで不要な誤差を除き、必要な誤差だけを通すことができる。ゲートは選択的に情報を取り出せる手段である。入力ゲートは前回の出力層の重みにより誤差を受け取るか判断し、記憶セルに通す情報を制御する。一方、出力ゲートは記憶セル内の誤差を次の層に必要な情報だけが通るように制御する。入力、出力ゲートでは必要な誤差だけを通すことで、誤差が不安定になるのを防止する。忘却ゲートでは中間層に記憶する必要がなくなった情報を除去することができる。忘却ゲートがあることで不要な情報が増えるのを防ぐことができる²⁾。LSTM では記憶セルとゲートにより誤差を制御することにより1000ステップ以上の長期的な学習が可能である。

4 RNN の応用事例

4.1 Google 翻訳

機械翻訳において英語から日本語に訳す際、入力サイズと出力サイズが可変になる。可変長のサイズを扱うこと

ができる RNN を用いることで Google 翻訳の精度が向上した。

ニューラルネットワークを用いた翻訳は、EncoderDecoder モデルを基本とする。EncoderDecoder モデルを用いて日本語から英語に翻訳する際の図を Fig. 4 に示す。

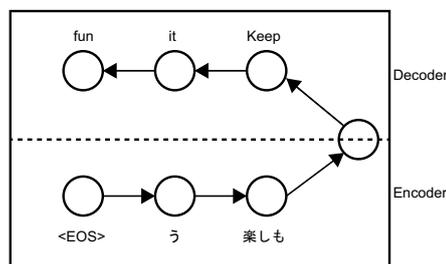


Fig.4 EncoderDecoder モデルの例

Fig. 4 の EOS は文章の終わりを意味する。EncoderDecoder には共に RNN が用いられている¹⁾。Encoder 部分の RNN で可変長の翻訳元文から Decoder 部分へ出力する。Decoder 部分で可変長の翻訳後の文を生成する。また LSTM を用いることで長文や複数文章の翻訳精度を向上させることに成功している。

4.2 画像の概要生成システム

Google は入力された画像に対して、画像の概要を自動で文章として出力するシステムを開発している。入力時は、従来のニューラルネットワークを用いた画像認識アルゴリズムで写真に映っている状況を解析させる。出力を RNN に変えることにより抽出された特徴から画像の説明文を生成する。また Google では大量のオープンデータベースの画像を処理させることで、説明文の質を向上させることに成功している。

5 今後の展望

現在も RNN や LSTM だけでなく新たな機能を持った RNN が開発されている。今後も RNN を中心として、機械学習の精度が向上すると考えられる。

また、医療分野ではニューラルネットワークを用いた画像認識技術の研究が行われている。RNN を導入することでレントゲン、CT、MRI から腫瘍や、ガンなどの検出や膨大な検査結果から異常値を見つけることが報告されている。アメリカではニューラルネットワークを用いた肺がん検出率が、医師を上回った研究結果も出ている。また RNN を利用して、時系列データである血液や脈拍などの測定データや投薬や検査データから最適な治療法を予測するモデルも開発されている。今後、医療分野はニューラルネットワークにより発展していくと考えられる

参考文献

- 1) 再帰型ニューラルネットワーク：入門
<https://qiita.com/kiminaka/items/87afd4a433dc655d8cfd>,
 参照 Apr.17, 2018
- 2) LSTM (LongShortTermMemory) 概要
<https://www.slideshare.net/KenjiUrai/kenji-urailstm>,
 参照 Apr.17, 2018