

# 佐久間まゆを機械学習で作ってみた

## ～ありえん層みの深い学習～

BV16055 中本 幸佑

平成 29 年 11 月 3 日

### 研究背景

夏休み中に数理研内でプログラミングコンテストが開催されたのだが、全然出来なかったため悔しくてプログラミングの勉強をしようと思った。「勉強するならそれを使ってなにかしよう。」と思い、最近話題の深層学習を用いて何かしようと思った。

### 1 研究内容

機械学習とニューラルネットワークについて学び、その後ディープラーニングについて学んでいき実際にマルコフ連鎖と LSTM を用いて佐久間まゆのセリフを自動で生成する。

### 2 佐久間まゆについて

アイドルマスターシンデレラガールズに登場するアイドルで、とても一途な女の子で、年齢は 16 歳、9 月 7 日生まれの乙女座で身長は 153cm、体重は 40kg、血液型は B 型で趣味はお料理と編み物。利き手は両利きの仙台出身のアイドルである。

### 3 機械学習

機械学習 (machine learning) とは、人間の脳が自然に行っている「学習」の能力を、コンピュータを用いて実現しようという手法のことである。ある程度の量のあるサンプルデータを入力してデータの中に、一定の規則やルールがないかを調べる。見つけ出した規則や特徴・ルールを基にして、別のデータを分類したり、将来を予測したりするものである。

### 4 深層学習

深層学習 (Deep Learning) とは、機械学習の一分野であり、ニューラルネットワークを用いて学習できるようにデザインされたものである。

### 5 ニューラルネットワークについて

ニューラルネットワークとは人の神経を模したネットワーク構造のことである。人間の脳には多数の神経細胞 (ニューロン) が存在していて、1 つのニューロンは、他の複数のニューロンから信号を受け取り、他のニューロンに対して信号を渡す。脳は、このような信号の流れによって、さまざまな情報を伝達している。この仕組みをコンピュータで再現したのが、ニューラルネットワークである。

### 6 佐久間まゆを作る

実際に佐久間まゆを作る。現在まで公開されているセリフをもとにまゆのセリフを自動生成する。

#### 6.1 マルコフ連鎖によるセリフ生成

実際に生成したところ以下のような文章が得られた。

- そっちの悪魔退治はいいですけど、お弁当がありますからね。に入り込んで自然ですし
- 車内って意外と狭いんです、うふふ…積極的。

#### 6.2 LSTM について

LSTM とは RNN を改良したもので、RNN に長期的に情報を記憶するためにブロックを採用したものである。下図は、LSTM の構造を図にあらわしたものである。

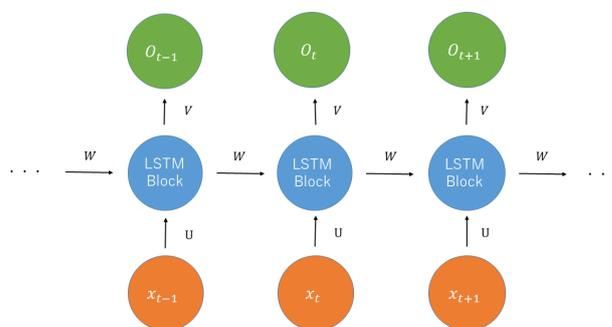


図 1 LSTM の構造

#### 6.3 LSTM によるセリフ生成

実際に生成したところ以下のような文章が得られた。

- プロデューサーさんの正ゆのお家をこのかて一緒に必♪
- 頑張らなきゃ！プロデューサーさんのためだものっ

### 今後の課題

実際にセリフは生成できたが、元のセリフが日本語になっていないの二択にしかならないので、どうしたらよりまゆっぽいせりふを生成できるかを調べようと思う。

### 参考文献

- [1] Ian Goodfellow and Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press, 2016, <http://www.deeplearningbook.org>, <http://www.deeplearningbook.me/>, 2017/09/21 アクセス。
- [2] クジラ飛行機, Python によるスクレイピング&機械学習開発テクニック, ソシム株式会社, 2016。
- [3] まゆの日記帳-佐久間まゆベディア@wiki, [http://www65.atwiki.jp/sakumamayu\\_wiki/pages/70.html](http://www65.atwiki.jp/sakumamayu_wiki/pages/70.html), 2017/10/27 アクセス。