

会話を通して成長する人工知能「pirka」の開発

東京都立戸山高等学校 SS情報 1年F組 糸井主歩

1 動機・目的

人は言語という複雑なコミュニケーションツールを持っており、生まれたばかりの赤ちゃんは保護者や周りの人を真似しながら、徐々に言語を獲得していく。それは「日本語が話される環境で育ったから日本語を話すようになった」というマクロな視点から、「親の小さい言葉遣いの癖を子も真似するようになった」というミクロな現象にまで言えることである。

この現象に興味を持ったため、私は中学の頃から趣味で「会話を通してユーザの言葉遣いを真似し、ユーザに似る」pirkaという人工知能を開発している。SS情報の授業では主に深層学習による会話の応答生成と、その応答を「ユーザらしく」するためのプログラムの開発を行うことにした。

2 深層学習について

ここ数年で「人工知能」と呼ばれているものの多くは主に深層学習という技術を使っている。深層学習とは、人の脳の神経細胞から着想を得たニューラルネットワークという構造を用いて与えられた問題を解くというものである。

多くの場合、問題を解くためには大量の教師データが必要となる。例えば、画像に写っているのが犬なのか猫なのかを判別したいとすれば、教師データは「犬または猫の画像」と、答えである「画像に写っているもの」のペアになる。

画像を入力し、それが犬であれば犬と、猫であれば猫と出力できるようなニューラルネットワークをコンピュータに作らせることを学習という。

自然言語処理（コンピュータで言葉を扱う分野。自動翻訳やニュース記事の要約、今回のような対話など）においても深層学習が広く用いられている。

以前はそれぞれの課題ごとにモデルを作る必要があったが、近年、1つの汎用的なモデルを基に、それぞれの課題に最適化するための2回目の学習（ファインチューニング）をするだけで複数の課題を解くことができるBERTやT5といったモデルがGoogleなどから提案されていた。

3 今回使用したモデル

今回は、2020年5月にOpenAI社が提案したGPT-3というモデルを用いて対話の応答生成を行った。

GPT-3は、文章を途中まで入力するとその続きを予想して出力するという形をとっており、例えば「バレンタインデーは、」と入力すると、

バレンタインデーは、今となってはチョコだけではなく、スイーツなどのプレゼントを贈る人も多い時代です。しかし、せっかくするなら美味しいスイーツで感謝の気持ちを伝えたいですね

などと自然な続きを出力できる。これを応用すると、例えば

AIはおしゃべりが好きで、とても賢いです。以下は人間とAIの会話です。人間:「こんにちは！」AI:「

と文章を入力して続きを予想させれば「文章の続き」として「こんにちは！」に対する会話の応答を得ることができると考えられる。一見無茶に見えるが、これがGPT-3で用いられるfew-shotという方法である。ファインチューニングすら要らずに複数のタスクに対応できるため応用が非常に簡単だ。

しかしながら事前に個人レベルでは難しいほどの大量の学習を行わなければならないため、rinna株式会社がC4、CC100（どちらも大量の文章データ）、それにWikipediaを学習データに用いてAIコミュニティサイトHugging Faceに公開したモデルjapanese-gpt-1bを使用した。

モデルを用いて前述のような「どのような文章の続きを生成すればいいのか」という会話の例を示した上でユーザがコマンドラインに文章を入力すると、応答が生成されるように実装した。

```
>>> 今日はいい天気だね
そうね
>>> どこに行かない？
近場のお店を見て回ると良いよ
>>> どんなお店がいいかな
少し遠くまで行くと気持ちいいと思うわ
>>> え、イタリアンでも行く？
少し遠いけど、いいお店があるの
>>> じゃあそこに行こうか
ありがとう
>>> ピザとスバゲッティどっちにしようか
スイートポテトとハッシュドポテトも好き？
>>> まあ好きやけど
じゃあスペシャルを5ドルで頼んでいいよ！
>>> やったあ（）
すごいね！
>>> それはなにより
さ、ピザも来て！
>>> やったぜ
OK! それでは着いてきてね！
>>> 歩いていくんかい
```

4 結果

生成された会話をいくつか示す。「>>>」に続いているのが私の入力した文章で、それに対する応答が下に1行出力されている。

```
>>> どんな本が好き？
小説
>>> 僕も好きだな
え、どうして？
>>> いろんな世界に行けるから
そっか。いろんな世界。いいな。
>>> 音楽は好き？
僕、ピアノ弾けるよ。ドレミで弾ける。
>>> 僕もピアノ弾ける
ああ。羨ましいな。僕も弾ける。
>>> ギターはどう？
僕もギターも弾ける
>>> 天才じゃん
```

5 考察

かなり自然な文章が表示された。文脈も読んだ応答が生成されているように思う。

6 展望

以前の研究でユーザの文章の特徴を基に文を「ユーザらしく」するためのプログラムを途中まで開発しているので、今後はそれと合わせて実装したい。また最終的にはWebサービスとしてデプロイできるよう開発を進めていく。

7 参考文献

rinna/japanese-gpt-1b - Hugging Face

<https://huggingface.co/rinna/japanese-gpt-1b>