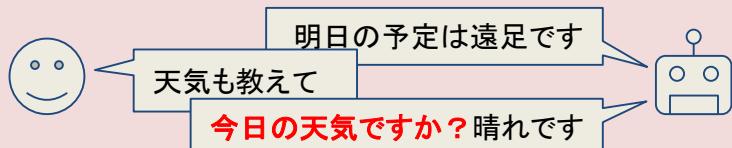


深層学習を使った AIアシスタントにおけるコンテキスト理解

株式会社クレスコ 森木研人 a-moriki@cresco.co.jp

AIアシスタントの問題点

理想的なAIアシスタントの実現にはコンテキストの理解が必要だが、現在はルールベースで実装されており、スケールしづらく限定的。



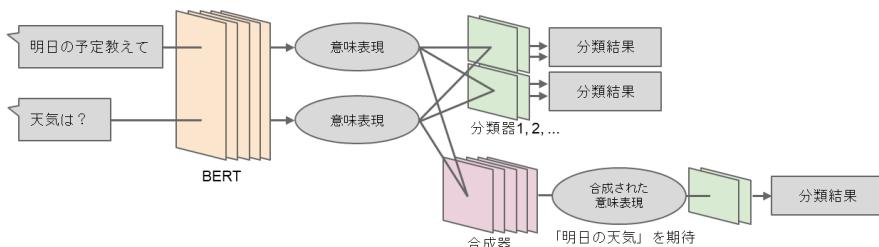
深層学習を使った解決を検討

スケールしづらい問題を回避するために、深層学習を使っての解決を検討。

エンコーダーモデルを通った情報はその意味を表現したものとなるため、これを重ね合わせるネットワークを構築し、コンテキストの理解が可能か検討する。

学習済みBERTモデルを使ったアプローチ

意味表現の抽出には BERT 学習済みモデル を利用。二つの意味表現合成するための合成器と、意味表現から情報を抽出するために分類器を作成。BERT出力の768次元配列を基にコンテキストを扱う。



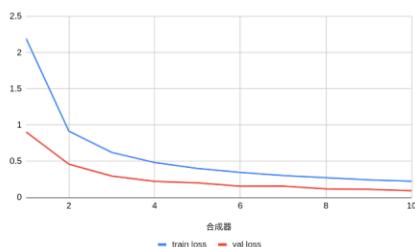
```
weather {date|place+の}天気教えて
weather {date|place+の}天気は?
weather {date|place+は}晴れる?
schedule {date+の}予定教えて
schedule {date+の}予定は?
schedule {date+は}暇?
continue {date|place+は}?
continue {date|place+は}どう?
...
```

```
weather here tomorrow 明日のこの天気は? -3.519543e-01,-1.139347e-01,....
weather next_week 来週の天気教えて -3.543214e-01,-1.43234e-01,....
schedule tomorrow 明日の予定は? -2.519543e-01,-3.139347e-01,....
schedule next_week 来週の予定は? -2.431234e-01,-4.111234e-01,....
continue tomorrow 明日は? -6.519543e-01,-7.139347e-01,....
continue next_week 来週は? -8.431234e-01,-7.111234e-01,....
```

大量の依頼文コーパスを用意することが難しかったため、テンプレート文から依頼文を作成、ネット上に公開されている雑談コーパスと合わせ 120,000 程度のデータを用意し学習。

結果

学習データとともに、検証データのLOSSも減少、正しく学習できた。



実際に利用した様子

正しく「明日」「天気」と判定

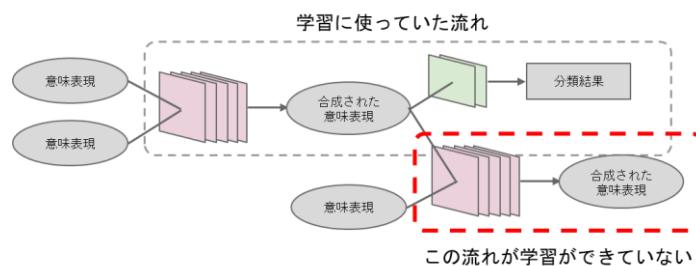
正しく「明日」「予定」と判定

```
> 天気は?
合成結果の分類
intent: weather
place: unknown
datetime: tomorrow

> 明日は?
合成結果の分類
intent: schedule
place: unknown
datetime: tomorrow
```

課題

合成器から出力した意味表現を再度合成器の入力として利用した場合、大きく精度が低くなるのが分かった。学習させ方に問題があったと考えられる。



その他課題としては、今回テンプレート文から生成したデータを使ったため、学習しやすい状況となっていたと考える。今後は実データに近いデータを用意し、精度が保てるかを検討したい。