

RNN 型日本語対話モデルの推測精度の検証

情報科学科 森 のどか

指導教員：小林 邦和，鈴木 拓央

1 はじめに

近年、機械学習を用いた対話システムの開発が進んでいる。その中に、Recurrent Neural Network (RNN) の一種である Sequence to Sequence (Seq2Seq) [1] モデルを用いて、英語のチャットデータを学習し、入力文に対し違和感のない返答をする研究がある。

そこで本研究では、同モデルを用いて日本語対話モデルの構築を目指す。学習データにテキストや発話をデータベース化したもの(コーパス)を用いる。そして、学習パラメータの変更、入力文の先頭の特徴が反映されにくくなることを防ぐ Attention[2] を導入することにより、モデルが応答文を生成する精度の向上を図った。

2 関連研究 -Sequence to Sequence-

Seq2Seq モデルは単語列の Encoder (青色部分) と Decoder (黄色部分) の 2 つの RNN で構成される。まず Encoder 側で入力文(図 1 中“ABC”)から特徴ベクトルを求める。次に文の終わり(図 1 中“<EOS>”(end of sequence の略))を Decoder 側が受け取ったら、適切な応答文(図 1 中“WXYZ”)を出力するように学習する。

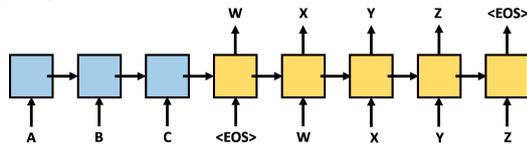


図 1 Seq2Seq モデルの構造

3 提案手法 -Attention-

Seq2Seq モデルは入力文が長い場合、最初に入力された単語の特徴が特徴ベクトルに反映されにくくなる。これを防ぐため、Encoder 側の隠れ層を全て考慮できるようにする。Attention の計算方法は式 (1) ~ 式 (3) の通りである。ここで \bar{h}_s , h_t はそれぞれ Encoder, Decoder の隠れベクトル, W_c は学習パラメータを指す。Encoder 側の隠れ層を荷重 $a_t(s)$ で加重平均した文脈ベクトル c_t を求め、 \tilde{h}_t より出力を予想する。score は Encoder と Decoder の隠れベクトルの関連度を求める関数である。

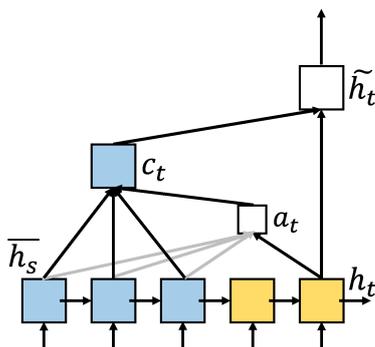


図 2 Attention を加えた図

$$a_t(s) = \frac{\exp(\text{score}(h_t, \bar{h}_s))}{\sum_{s'} \exp(\text{score}(h_t, \bar{h}_{s'}))} \quad (1)$$

$$c_t = \sum_s a_t(s) \bar{h}_s \quad (2)$$

$$\tilde{h}_t = \tanh(W_c [c_t; h_t]) \quad (3)$$

4 計算機シミュレーション

4.1 シミュレーション方法

Seq2Seq モデルに Attention を加えた場合と加えない場合と比較する。学習データは『名大会話コーパス』[3] と、『BTSJ による日本語話し言葉コーパス(トランスクリプト・音声) 2011 年版』[4] を用いた。2 つのコーパスを合わせた入力文と出力文は約 44000 ペアある。コーパスの形態素解析には京都大学の黒原・河原研究室で開発された JUMAN++ を用いた。

精度比較に用いる perplexity とは、言語モデルの予測性能を測る指標であり、次に来る単語の候補数に相当する。perplexity は式 (4) によって与えられる。学習には RNN より長期依存を学習できるよう、中間層に忘却・更新ゲートを導入した Gated Recurrent Unit (GRU) を用いた。またミニバッチ学習という、学習データ N 個の内、ランダムな n ($n \leq N$) 個のデータを用いて学習を行う方法を取り入れた。今回ミニバッチは 8 である。その他の条件は表 1 の通りである。

$$\text{perplexity} = e^{\text{loss}} \quad (4)$$

表 1 パラメータ設定

パラメータ	設定値	パラメータ	設定値
学習率	0.05	隠れ層のユニット数	512
ステップ数	60000	層数	4

4.2 結果

それぞれのモデルを学習させた結果、最終的に perplexity の値は Attention を加えたモデルが 15.88, 加えていないモデルが 38.58 となった。このことから Attention を加えた方が少ないステップ数で次に来る単語の候補をより絞れており、予測性能は高いと言える。図 3 は各ステップの perplexity の値をプロットしたグラフである。

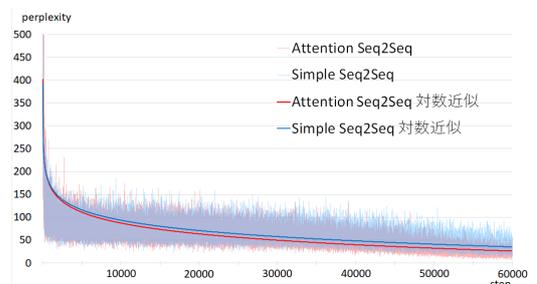


図 3 ステップ毎の perplexity の推移

5 おわりに

Seq2Seq を用いて日本語対話へ適用し、Attention の有無を比較した結果、Attention を加えたモデルの方が少ないステップ数で予測性能が向上したことが確認できた。他にも Beam Search などモデルを改良する方法はいくつか提案されているため、様々な手法を比較し、より良い精度が得られる手法を調査する。

参考文献

- [1] Ilya Sutskever, Oriol Vinyals, Quoc V. Le 『Sequence to Sequence Learning with Neural Networks』, 2014
- [2] Minh-Thang Luong, Hieu Pham, Christopher D. Manning 『Effective Approaches to Attention-based Neural Machine Translation』, 2015
- [3] 藤村逸子, 大曾美恵子, 大島ディヴィッド義和 『会話コーパスの構築によるコミュニケーション研究』藤村逸子, 滝沢直宏編 『言語研究の技法: データの収集と分析』 p. 43-72, ひつじ書房, 2011
- [4] 宇佐美まゆみ監修 『BTSJ による日本語話し言葉コーパス(2011 年版)』 『人間の相互作用研究のための多言語会話コーパスの構築とその語用論的分析方法の開発』平成 20-22 年度科学研究費補助金基盤研究 B (課題番号 20320072) 研究成果, 2011