

LSTMを用いた質疑応答システムによる 人工知能知識習得の可能性の評価

An Evaluation of Q-A systems for Acquiring Knowledge of Artificial Intelligence via LSTM

根岸 儀和 *1
Yoshikazu Negishi

小田 徳光 *1
Norimitsu Oda

岩井 健二
Kenji Iwai

中村 彰宏
Akihiro Nakamura

亀田 雅之
Masayuki Kameda

浅川 伸一 *2
Shin Asakawa

*1株式会社ロボケン
RoboKen Co. Ltd.

*2東京女子大学
Tokyo Woman's Christian University

We investigated two question-and-answer (Q-A) processing systems via LSTM. A Sequence-2-Sequence model was compared to a vanilla LSTM to process the machine learning thread in a reddit.com. Based on our findings, we proposed several possibilities to deal with this problem. We discussed further requirements to process Q-A specific problems that remain unsolved. Suggestions might be obtained on sequential information processings, combined with our results and requirements.

1. はじめに

囲碁での人間への勝利 [Silver 16] 等, 人工知能技術が衆目を集め 関連する技術, 知識に関する教育の需要が高まっている。しかし, 人工知能の実装には統計や機械学習の知識が必要であり, 需要に対して十分な技術を持った技術者が不足している状態である。我々は「AIを教える AI」を作ることによってこの問題の解決ができないかを検討した。現在, 機械による学習を行う方法として, さまざまな e-learning システムが利用されている。しかし, それらのシステムは学習内容の提示と演習問題の組み合わせからなっており, 学習者に応じた適応的な教示ができるわけではない。現在の人間の講師の利点は, 学習者の理解の度合いや, 欠落している知識などの学習状態を推測し, 学習の順序や演習問題の形式を適応的に変化させられる点とであると言える。この適応性を実現する上では, 学習の過程で学習者「AIを教える AI」が頻繁にやり取りを行う質疑応答による学習, いわば人間の教師における個人指導のようなシステムであることが望ましいと考えられる。

人工知能による質疑応答には [Kalchbrenner 13, Wen 15, Lowe 15, Vinyals 15, Shang 15, Sordani 15] といった事例が挙げられる。学習という目的から, 事前に質疑応答を選択肢やパターンとして限定することは難しい。そのため, 学習者が任意の表現で質問可能な自然言語による質疑応答システムが望ましいといえる。そこで我々は, 事前に特徴量の選択を必要としない, ニューラルネットワークによる自然言語処理を考え, 一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 LSTM(Long short-term memory, 図 1) [Hochreiter 97] を用いたシステムの構築を想定した。我々は「AIを教える AI」実現の第一歩として, 1 問 1 答の質疑応答システムとして LSTM による対話モデルが適用可能か否かを検討した。

本稿では, 2 節で提案手法とための LSTM モデル, 実験方法について述べ, 3 節で実験の結果, 4 節で実験結果に対する考察と今後の課題について述べる。

2. 方法

質問文—回答文を一つの系列情報とみなして一つのリカレントニューラルネットワークに学習させる場合(一時系列モデル)と質問者の問いと回答者の答えを別個のシステムに与え, 質問文に対する回答文を生成させた場合(対話モデル)とを比較した。

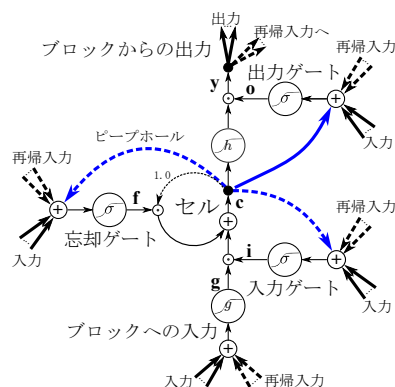


図 1: LSTM の概念図

入力層, 単語埋込み層, LSTM(128 ユニット, 図 1)2 層, ソフトマックス層からなる 4 層のニューラルネットワークを用いた。勾配クリップを 1 に設定しドロップアウト率 0.5 とした。BPTT の時間窓は 5 とした。

2.1 一列モデル

LSTM を用いた一列モデルの概要を図 2 に示した。

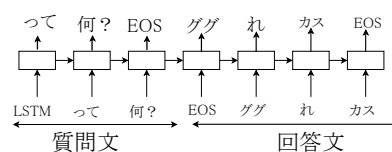


図 2: 一列モデル

連絡先: 浅川伸一, 東京女子大学, 167-8585 東京都杉並区善福寺 2-6-1, 03-5382-6746, asakawa@ieec.org.
亀田 雅之, m.kamedax@gmail.com

2.2 対話モデル

[Sutskever 14] による自動翻訳モデルを用いた。 Q と A とを個別の LSTM として実装した。LSTM を用いた対話モデルの概要を図 3 に示した。

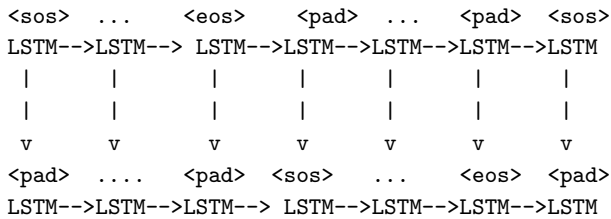


図 3: 対話モデル

Q を単語ごとに入力し、その間の出力としては埋草トークン $\langle \text{PAD} \rangle$ を出力とする。 Q 終了時に Q 終了と示すトークン $\langle \text{EOQ} \rangle$ を入力し、 A の先頭の単語を出力として回答文の予測学習を行った。

2.3 実験

2.3.1 データ

Web 掲示板 <https://www.reddit.com/r/MLQuesitons> スレッドをスクレイピングしてデータを用いた。スレッドのタイトルと本文とを質問文と見なし、回答のついているスレッドのみをデータとして用いた。html タグを除去した後、NLTK を用いてトークン化した。一つの質問分に対して、複数の回答があるスレッドについては、その都度質問文を繰り返して (Q, A) 対を作成した。Reddit という掲示板の性質上データセットには “:-p” や “WTF” などの特殊な文字列が存在した。加えて URL のみが記述された回答文も存在した。URL は arXiv などの論文を示しており、意味があると考えられる。しかし、処理上は低頻度語トークンとして扱った URL も存在する。データ採取の時期柄 AlphaGo の話題を示す URL などは低頻度語トークンに分類されずに残った。これらの傾向を期間を限定したデータ採取では、恣意的にならずに取り扱うことは困難である。参照先の URL が有益か否かについての判断は本研究の枠組みを超えるため取り扱わないこととした。

総スレッド数 289 を訓練データ (45064 語)、検証データ (10492 語)、テストデータ (9509 語) に分割した。出現頻度 5 以下の低頻度語を一括して UNK トークンとして扱った。

3. 結果と考察

学習済みのモデルに対してテストデータを用いた評価では、データ数の関係もあり正解データとの TD/IDF cos 類似度はそれ程高くなかった。一列情報モデルは (Q, A) 対についての完全な情報を学習することを意味するので通常は有利に働いた。一方対話モデルは Q の発話を受けて、 A の系列を学習することから学習基準に達するまでの繰り返し回数が多い傾向が見られた。ただし、TD/IDF cos 類似度による指標では学習中の結果では一列情報モデル 0.041 に対して対話モデル 0.079 となっており、有意な差異は認められなかった。

本研究の結果は、一列情報モデルと対話モデルとの差異が認められなかったことから、部分情報からの系列再生能力を用いることである程度の質疑応答システムの開発への可能性を示すものと考えられる。QA システム、ひいては教育システム構築においては、学習者の理解度に合わせた適応と低頻度語

であっても有益な情報をいかに処理するかとの問題が指摘できる。今後は、質問者の知識状態の推定から、回答方法の唆など柔軟な対応を可能とするシステムの構築が望まれる。

参考文献

- [Hochreiter 97] Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, *Neural Computation*, Vol. 9, pp. 1735–1780 (1997)
- [Kalchbrenner 13] Kalchbrenner, N. and Blunsom, P.: Recurrent Continuous Translation Models, in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, EMNLP*, Seattle, USA (2013), Association for Computational Linguistics
- [Lowe 15] Lowe, R., Pow, N., Serbany, I. V., and Pineau, J.: The Ubuntu Dialogue Corpus: A Large Dataset for Research in Unstructured Multi-Turn Dialogue Systems, *arXiv* (2015)
- [Shang 15] Shang, L., Lu, Z., and Li, H.: Neural Responding Machine for Short-Text Conversation, in *The 53rd Annual Meeting of the Association for Computational Linguistics (ACL)* (2015)
- [Silver 16] Silver, D., Huang, A., Arthur Guez, and C. J. M., Sifre, L., Driessche, van den G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., and Hassabis, D.: Mastering the game of Go with deep neural networks and tree search, *Nature*, Vol. 529, No. 28, pp. 484–492 (2016)
- [Sordoni 15] Sordoni, A., Galley, M., Auli, M., Brockett, C., Ji, Y., Mitchell, M., Nie, J.-Y., Gao, J., and Dolan, B.: A Neural Network Approach to Context-Sensitive Generation of Conversational Responses, in *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology (NAACL)* (2015)
- [Sutskever 14] Sutskever, I., Vinyals, O., and Le, Q. V.: Sequence to Sequence Learning with Neural Networks, in Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., and Weinberger, K. eds., *Advances in Neural Information Processing Systems (NIPS)*, pp. 3104–3112, Montreal, BC, Canada (2014)
- [Vinyals 15] Vinyals, O. and Le, Q. V.: A Neural Conversational Model, *arXiv* (2015)
- [Wen 15] Wen, T.-H., Gašić, M., Mrkšić, N., Su, P.-H., Vandyke, D., and Young, S.: Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems, in *Proceedings of Empirical Methods in Natural Language Processing (EMNLP)*, Libson, Portugal (2015)