

柔軟な対話システムのための再帰的ニューラルネットワークによる文カテゴリ分類手法の検討

小室光広† 佐藤裕二†

法政大学情報科学部†

1. まえがき

人工知能(AI)研究において、初期の研究者達は「人間のように考え、会話し、見ることができるコンピュータ」を目指していた[1]。研究が進むにつれ、Strong AIとWeak AIの2つの概念が登場した[2]。Strong AIにおける技術課題を解決するためにはまずWeak AIにおける技術課題を解決する必要があると考える。また、Strong AIにおける主要なインターフェースはAIとの対話である。環境とやり取りをする機構は人間の対話と同じであると考えからである。システムの文表現理解における初期段階として文書分類の研究が挙げられる。従来研究では単語におけるPositiveな文とNegativeな文での出現頻度に基づく分類手法[3]などが報告されているが、語順を重視する研究例の報告はほとんどない。また、従来の文書分類手法ではコーパスを用いて分類する手法が提案されているが、対象とする言語に合わせてコーパスを用意する必要があった[4]。そこで本研究では、語順を重視し、コーパスを用いることなく、文がPositiveであるかNegativeであるかのニュアンスを判別させるための文カテゴリ分類器の構築を目指す。システムが大まかなニュアンスを掴む事で文解釈補助が可能であり、語順を重視することにより、対話システムにおけるユーザーとのやり取りの中で語順により文のニュアンスが変化しシステムから誤った応答が出る可能性を削減できると考えるためである。加えてコーパスを用いないことにより、多くの言語に容易に対応可能であると考えたからである。

2. 対話システムのための文カテゴリ分類器

本研究では人間とシステムが対話を行う前段階として文の大まかなニュアンスをシステムに理解させるために、会話文や評論等から取り出した英文を入力文とし、文がPositiveであるかNegativeであるかというカテゴリを判断する分類器のための学習機構を提案する。図1に本研究において提案するシステムの全体像を示す。提案するシステムでは始めにSourceとなる学習データセット等の文書集合からSentenceを抜き出しそれをWordsに分けvocabularyを生成する。その後入力として与えられたSentenceをWordごとに分け、学習データセットから生成されたVocabularyのインデックスで示されるベクトルに変換し、入力SentenceがPositiveであるかNegativeであるかを学習する。次にClassification Moduleにおいて学習したモデルと学習時に生成されたVocabularyにより未知の文についてPositiveであるかNegativeであるかを分類す

Sentence classification method with recurrent neural networks for flexible conversation systems

†Mitsuhiro Komuro, †Yuji Sato

†Faculty of Computer and Information Sciences, Hosei University

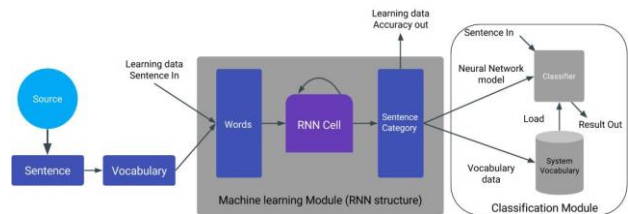


図1 提案する分類器システムの全体像

る。学習は図1のMachine Learning Moduleにおいて再帰的ニューラルネットワーク(RNN)の手法を用いる。これは第1節で述べた通り対話システム及び文の語順を念頭に置いているため時系列データを扱う事のできるRNNが適切であると考えたためである。RNNのパラメータ学習はBackpropagation through time [5]により行う。学習データセットはPositiveとNegativeに予めラベル付けされた10662件 [6]を用い、相互検証により400回学習を行う。前述のvocabularyは本データセットより生成され学習に用いられる。なお、本報告は図1のMachine Learning Moduleまでを扱い、プログラムの実装は単純・簡略化するため、多次元行列計算による機械学習に特化したライブラリであるTensorFlow[7]を用いて行う。

3. 評価実験

3.1 実験方法

本研究では、NNにおける勾配計算や各ハイパーパラメータ再帰的化に必要な学習率[8]の変更を実験1、中間層におけるCell数の変更を実験2、中間層数の変更を実験3として評価実験を行う。実験1では学習率を $1e-3$ から $1e-7$ に変更し、実験2ではCell数を16, 32, 64個に変更、実験3では層数を1, 2, 3層に変更して実験する。各実験では中間層におけるCellを通常のRNN cellだけではなく、LSTM[9] cellやその改良型である覗き穴付LSTM[10] cell (PLSTM), LSTM cellやPLSTM cellより内部構造を単純化したGRU[11] cellに変更し精度向上するかを比較検討する。本評価実験における実行環境は表1の通りである。

表1 実験環境

OS	Ubuntu 16.04 LTS
Linux Kernel	Linux 4.10.0-42-generic x86_64 GNU/Linux
CPU	Intel Xeon CPU X3450 @ 2.661GHz
Python	3.6.2 on Anaconda
Tensorflow	1.3.0

3.2. 実験結果

図2から図4にかけて各評価実験の検証データにおいて特徴的な結果となったグラフを示す。

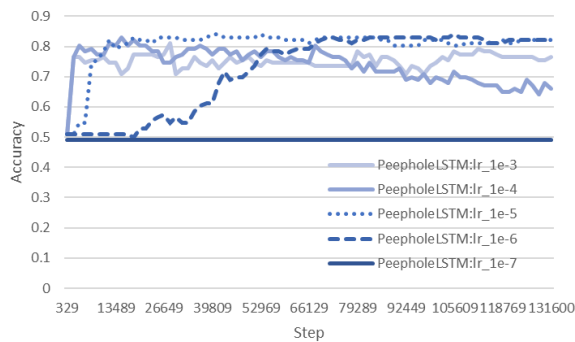


図2 学習率変更実験の精度値推移

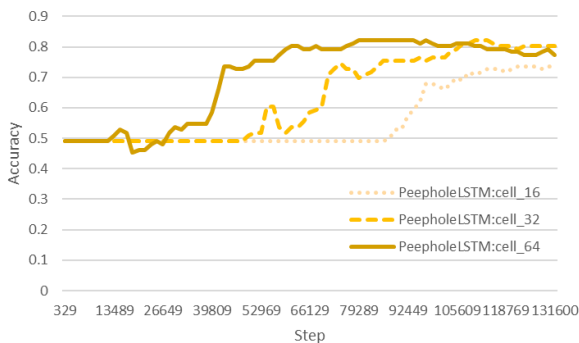


図3 Cell数変更実験の精度値推移

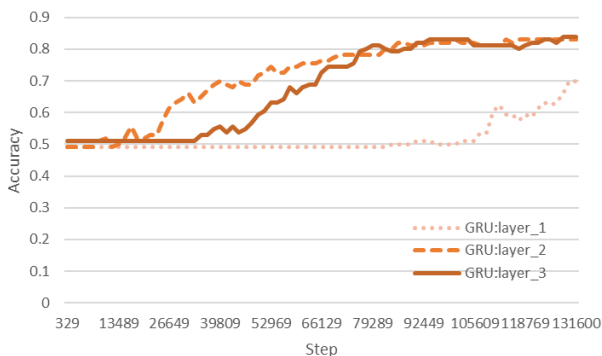


図4 中間層数変更実験の精度値推移

図2, 図3はPLSTMの結果, 図4はGRUの結果である. 図2のPLSTMは学習率 $1e-5$ 及び $1e-6$ で最大精度値0.8208である. 図3ではCell数32のとき最大精度値0.8019である. 図4のGRUでは3層のとき最大精度値0.8396である.

4. 考察

図2, 3より学習率が低く, Cell数が32の場合高精度で分類されていることから, 特定のハイパーパラメータを持つ場合に本システムが有用に作用したと考えられる. また, 中間層の層数では, 他のCellでは2層の場合が良精度である一方, 図4の通りGRU Cellでは3層の場合で精度がよく, 他のCellよりも高精度であることから, 提案システムにおける使用Cellに応じて中間層数も変更する必要があると考えられる. GRU Cellのみで層数が変化した理由の一つとしてCellの内部構造が単純化しネットワークモデルがLSTMやPLSTMよりシンプルで, 勾配計算が多層化に耐えることが原因と考えられる. 一方で図3,

4では低い値を取っている場合でも学習Step数が増える而变化が現れることが分かる. したがって学習回数を増やすことにより同等の精度まで向上する可能性があると考えられる. また, 本研究ではコーパスを用いることなく学習データセットに含まれる単語から vocabulary を生成し, 学習に用いているが, 3つの評価実験において総じて高い精度値が出ており, 学習データセットに使用される言語を変更することにより, 本研究で扱った英語以外の言語にも対応可能であると考えられる. 今後は学習データセット, 学習回数の増加が精度向上に影響を与えるかを調査, 検討する必要および学習データセットの言語を変更した場合に同程度の精度が得られるかの調査, 検討が必要である.

5. むすび

本研究では, 文の大きなニュアンスを理解するシステムの実現を目的として, 入力英文が Positive か Negative かを分類する文カテゴリ分類問題を NN の学習を用いて実現する手法を検討した. ハイパーパラメータや中間層の種類を変更するなどして精度を比較実験した結果, 提案システムは学習率が低く, 中間層が多層である場合に有用である可能性およびコーパスを用いないことによる容易な多言語対応の可能性を示した.

参考文献

- [1] 小林雅一, "クラウドから AI へ アップル、グーグル、フェイスブックの次なる主戦場", 朝日新聞出版, 東京, 2013.
- [2] John R. Searle, "Minds, brains, and programs", Behavioral and Brain Science, 3, pp.417-457, 1980
- [3] 姫野晋之介, 青野雅樹, "単語のポジティブ・ネガティブでの出現比率に着目したツイートの極性分類", 言語処理学会第23回年次大会 発表論文集, 2017.
- [4] 宅和晃志, 吉川大弘, 古橋武, "複数のコーパスから学習した分散表現の利用による文書分類制度向上に関する検討", The 30th Annual Conference of the Japanese Society for Artificial Intelligence, 2016.
- [5] P.J. Werbos, "Backpropagation through time: what it does and how to do it", Proceedings of the IEEE, vol.78, no.10, 1990.
- [6] Bo Pang, Lillian Lee, "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales", Proceedings of the ACL, 2005.
- [7] TensorFlow, <https://www.tensorflow.org/>, (accessed: 2018年1月12日)
- [8] Diederik P. Kingma, Jimmy Ba, "Adam: A Method for Stochastic Optimization", ICLR, 2015
- [9] Felix Gers, "Long Short-Term Memory in Recurrent Neural Networks", École Polytechnique Fédérale De Lausanne, 2001.
- [10] F.A. Gers, J. Schmidhuber, "Recurrent nets that time and count", IJCNN, 2000.
- [11] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, Yoshua Bengio, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation", EMNLP, 2014.