

協調学習のための記述式課題評価支援システムの検討

情報学部情報科学科 綱川 隆司

本研究は静岡大学情報学部 福田 治輝(情報科学科)、大島 純(行動情報学科)、大島 律子(行動情報学科)、西田 昌史(情報科学科)、西村 雅史(情報科学科)との共同研究です。

はじめに

大学等の教育現場では、複数の学習者が一緒に課題に向き合うグループワークが取り入れられてきている。グループワークにおいて教師は各学習者に対する評価を行う必要があるが、レポートや最終発表といった成果物だけでなくグループワークの過程も評価しようとする相応の労力が要求される。本研究ではグループワークの一形態である協調学習において、グループ内の調整能力を測る手法における記述式課題のテキストの評価を支援するため、テキストを自動的に評価する方法について検討を行った。

研究の背景

協調学習と CSSER

協調学習とは複数の学習者がグループを作り一緒に学ぶ学習形態を指し、共に学習を助け合うことで理解の深化、モチベーションの向上、学びのあり方自体の習得等をねらいとする。その中で2つの調整活動が必要となる。

・社会認識的な調整活動

✓問題解決のための議論、専門的知識の共有等

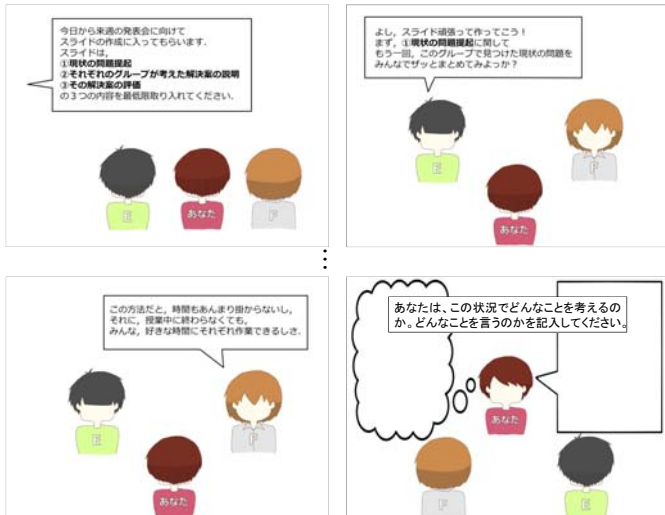
・社会感情的調整活動

✓グループメンバー内の衝突の調整等

Oshimaら(2015)は協調学習における個々の学習者の調整能力を測定するため、協調学習のある場面をマンガで描き、ある場面での“心の声”と“発言”を記述させて評価する指標CSSERを開発した。

CSSERによるテストの自動評価

CSSERは協調学習のある場面を想定した4つのシナリオを描いたマンガからなり、被験者は各シナリオを読んで最後の吹き出しにその時点で考えと発言を記述する。



評定者はルーブリック(評定基準)に従い社会認識的側面と社会感情的側面のそれぞれについて5段階評価のスコアを与える。本研究ではこのスコアを教師ありの深層学習により推定することを課題とする。

認識	感情	吹き出しの記述(発言)
5	5	確かにFさんの言う通り作業を分担したほうが効率がいいと思うよ。ただ、多少内容に齟齬が生じる場合もあると思うから、情報の共有だけはしっかりとしておくべきだと思う。
3	2	なるほど、じゃあ分担を決めてみんなでそれぞれスライドを作ろう。

BERTを用いたテキスト自動評価

汎用の言語モデルBERT (Devlin et al., 2018) を本課題にファインチューニングして入力テキストを5段階評価のいずれかに分類する分類器を構成した。

- ・事前学習: BERT日本語Pretrainedモデル(京都大学 黒橋・河原研)
- ・ファインチューニング: 入力吹き出しに入力されたテキスト、出力は5段階評価として学習

評価実験

実験設定

データ: 記入漏れ・解読不能等を除く吹き出し入力テキスト20760件(各シナリオの考え・発言の総数)とそれらに付与された5段階評価

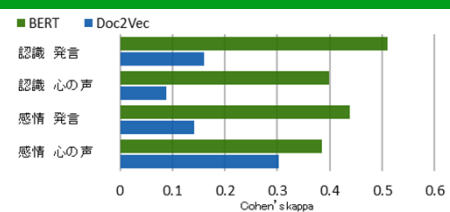
5分割交差検定: 訓練:開発:テスト=3:1:1

ベースライン: Doc2vec (Le and Mikolov, 2014)

1. 全ての入力テキストをdoc2vecでベクトル化
2. 訓練データ中で最もcosine類似度が高いテキストを抽出する
3. そのテキストの評価を入力テキストの評価として出力する

評価指標: Cohenの κ 値(評価者間の一致度を測る指標)を出力結果と正解値との間で求めたもの(人間の評価者間の κ は0.7~0.9程度)

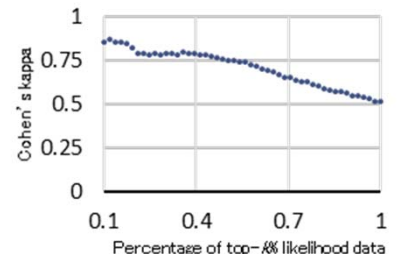
結果



BERTモデルはベースラインを大きく上回っているが、人間の評価者間一致度には至っていない

尤度によるフィルタリング

BERTにより出力される分類結果の尤度について、その高い順に $k\%$ (top- $k\%$)だけとったときの κ 値を調べた。尤度が高い出力結果については信頼性の高い結果であることが示された。



今後の課題

評定精度の改善のほか、入力テキストに対する類似例の提示やルーブリック改善の支援への応用方法を模索する。

謝辞 本研究はJSPS科研費 JP19H01714 の助成を受けたものです。

keyword: 協調学習、テキスト自動評価、BERT、自然言語処理