

題目 機械学習を用いた議論の質評価指標の自動判定

氏名 村井 柊

指導教員 大沼 進

本研究では、議論の質評価指標(Discourse Quality Index: DQI)の機械学習を用いた自動判断の精度を高めることを目的に、様々なモデルの作成方法で機械学習を行い、その精度を比較した。まず予備実験1として、標準的な機械学習と先行研究で行われた機械学習を比較し、本実験で用いる手法を確認した。具体的には標準的な機械学習の手法と先行研究で行われた手法を比較しつつ、実際に標準的な手法での機械学習を行い、出力された精度がどのようなものであるかを確認した。次に予備実験2として、先行研究において、データのノイズ処理がされていたが、そのことについて十分な説明がなされていなかったため、そのノイズ処理の妥当性を検討した。具体的には、ノイズ処理前とノイズ処理後の精度を比較した。その後、ノイズと思われるデータを抜き出し、個別に吟味した。その結果、先行研究において排除されていたデータは実際に機械学習に悪影響を与えており、データのノイズ処理には妥当性があると結論付けられた。その後、本実験として、人間が判定しているDQIに関して、機械学習を用いてモデルを作成し、妥当な判定が可能であるかを検討した先行研究において作成されたモデルより、高い精度での判定を行うモデルの作成方法を探索的に検討した。モデルの精度を向上させる方法の仮説として、データ数の増加、ラベル数の減少、同系統ラベルの階層化、学習させる順番のランダム化の4つを設定した。具体的には、まず二人の評定者の一致の全体像を確認し、ラベル数の減少において減らすラベル、同系統ラベルの階層化において同系統とみなすラベルを決定した。その後、4つの仮説それぞれに対応した機械学習を行い、その精度を比較した。本実験の結果、データ数の増加により大きく精度が上昇した。また、ラベル数の減少、学習する順番のランダム化で精度は微増した。同系統ラベルの階層化では、精度は全体的には減少したが、特定の条件下では精度が上昇した。本研究では機械学習を用いた議論の質評価指標の自動判定の精度を高めるモデルの作成方法を模索し、一定の成果を得ることができた。これによって合意形成支援ツールとしてAIエージェントが議論の質評価を行うという社会実装への貢献について論じた。