

異常検知を用いたレポート評価システムの試作

山添 久稔*, 大園 忠親, 新谷 虎松(名古屋工業大学大学院)

Implementing a Semi-Automatic Report Evaluator Using Anomaly Detection

Hisatoshi Yamazoe, Tadachika Ozono, Toramatsu Shintani (Graduate School of Engineering, Nagoya Institute of Technology)

1. はじめに

本稿では、評価基準を与えずにレポートを評価する手法を提案する。本稿が対象とするレポートは、講義内容のポイントの復習と理解を確認するための宿題である。レポートの問題とその解答例は、オンライン上で公開されている講義資料内に含まれていることを想定する。受講者は、講義終了後に講義資料を見直しながらかレポートを回答する。受講者は、レポート作成後、指定された様式のレポートを、メールで提出する。本講では、レポート評価手法およびシステムの試作について述べる。

2. 評価基準の作成が不要な自動レポート評価

評価基準をシステムに与えることなくレポートを評価する手法の実現可能性について議論する。

対象とするレポートは、(1) レポート間の類似度が高い、(2) 正当が大多数を占める、および(3) 正当と誤答では含まれる語彙が異なる、の3点を満たすことを想定する。(1)~(3)は、受講者らが同じ講義資料を参考していることを考慮すれば、自然な仮定であるといえる。

(1)~(3)を考慮すると、レポートは、類似した正当が大多数となることが期待される。

上記の性質を持つ評価システムの目標と、既存の自動評価における研究とを比較する。石岡らは、大学入試センター試験やその模擬試験などに対してアプローチしている(1)。石岡らの手法は評価の事前準備として、模範解答に加え、模範解答内の語彙についての言い換え表現、部分点、一部の回答を用いて作成した評価モデルの回答の評価を行う。提案するシステムは、石岡らの目的と異なり、評価者が、受講者らの講義内容のポイントの復習と理解の確認を主目的とする。そのため、評価基準を事前に設定することなく、評価が可能な判断手法とその手法を用いたシステムについて述べる。

3. 異常検知を用いた誤答の判断と評価画面

エラー検知を利用した誤答の判断手法と、評価者が操作を行う画面について述べる。上記の性質も持つレポートは、レポートに用いられた語彙を考慮すると、密な正当の集合と、正当の周辺に点在する誤答が存在すると想定できる。そこで、エラー検出にOne-Class-SVMを用い、正当と誤答の判別を行う。また、One-Class-SVMのパラメータである誤答率を、評価者は評価画面を用い操作する。

具体的な評価の流れを説明する。評価は、レポートの提出期限が過ぎたことをきっかけに行われる。まず、各レポートに対し、形態素解析を用い、名詞のみを抽出する。抽出された名詞を元に、各レポートに対しBag-of-Wordsのベクトルを生成する。次に生成されたベクトルの重心を計算する。計算された重心からユークリッド距離が近い順にレポートをレポート全体の半数選び、選ばれたレポートに含まれる名詞のみを用いて、新たにレポート全体からBag-of-Wordsのベクトルを生成する。この生成されたベクトルに対し、One-

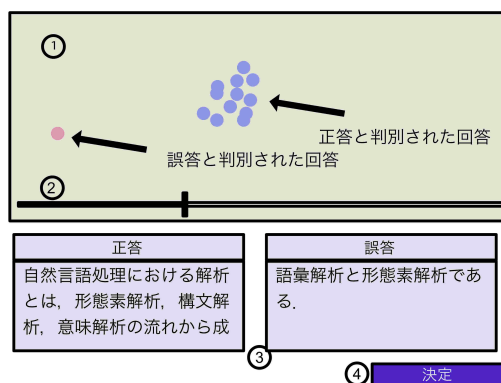


Fig. 1 評価画面

Class-SVMを用いた異常検知を行い、誤答を発見する。重心に近い半数のレポートに含まれる名詞のみを用いる理由は、受講者の独自調査を含むレポートが誤答と判別されるからである。受講者の独自調査を含むレポートとは、講義資料に含まれる正当に加え、その周辺知識を独自に調査し、レポートに含めたものである。これらのレポートは、ユニークな名詞が含まれている。ユニークな名詞は判別の際のノイズとなるため、除去を行う。

Fig. 1は評価者が操作を行う評価画面である。図中①は、各レポートのベクトルに対し主成分分析を行い、二次元に次元圧縮したベクトルを表している。誤答は異なる色で表示される。また、評価者は、それぞれのレポートをクリックすることで、レポートを閲覧できる。図中②はスライダーである。このスライダーを用いて、One-Class-SVMのパラメータである誤答率を調節する。スライダーを左側へスライドすることで、誤答率が下降し、正当の範囲がより広くなる。図中③はOne-Class-SVMの閾値に最も近い正当と誤答である。評価者は、表示された二つのレポートを見比べることで、レポートの評価を行う。図中④をクリックすることで、評価を終了し、結果の保存と受講者に対して評価結果をメールで送信する。また、一度保存した保存したレポートは、データベース上に保存し、講義全体を通した各受講者の平均等のデータを出力可能である。

4. おわりに

本稿では、評価基準を与えずにレポートを評価する手法とそのシステムを試作した。評価の対象となるレポートは3つの性質を想定される。想定される性質を利用した手法と、試作したシステムにより、評価者は評価画面上のスライダーを左右に調整することで、レポートの評価を可能にした。

文 献

(1)石岡恒憲 et al. "人工知能を利用した短答式記述採点支援システムの開発." 研究報告自然言語処理(NL) 16, pp. 1-6, 2016.