

学習指導要領に基づいた設問自動分類タスクにおける語の集約による有効性の評価

名嘉 真之介[†] 當間 愛晃[‡] 赤嶺 有平[‡]
山田 孝治[‡] 遠藤 聡志[‡]

琉球大学大学院理工学研究科情報工学専攻[†]

琉球大学工学部情報工学科[‡]

1 はじめに

学校では生徒の学力向上を目的として定期試験が行われている。しかし、その結果から生徒の得手不得手の傾向分析が十分に行われていないことが、平成 23 年度の沖縄県教育委員会の調べ [1] で分かっている。

この問題に対して、年度を跨いだ試験結果のデータを蓄積することで得手不得手の学習カルテを作成し、問題解決を図りたい。しかし、そのためには、基となる試験結果からこれまで以上にデータの収集が必要不可欠である。

本研究では、試験結果から新たな学習カルテを作成するための準備段階として、学習指導要領 [2][3] に基づいて機械学習を用いた社会科の設問自動分類を行う。その結果、十分な分類精度を得ることが出来れば、データから生徒が正解した設問と間違えた設問について得手不得手の傾向を示す学習カルテを生成することができ、生徒の学習支援に生かすことができる。

本稿では、設問自動分類タスクにおける誤分類事例に関する分析 [4][5] を基に、設問中に出現する語を集約することによって新たなデータセットを用意し、F 値の向上を図った。また、語の集約前後で分類結果に差が出た設問を調べ、その要因を分析した。

2 提案手法

本研究では設問に出現する名詞を素性として TFIDF(1) を計算し、それを特徴量として特徴ベクトルを生成する。その後、特徴ベクトルからマルチラベル分類を行う分類器を構築し、10 分割交差検定により分類実験を行い、その F 値を算出して評価する。

$$TFIDF_{ij} = tf_{ij} \times \log\left(\frac{N}{df_i}\right) \quad (1)$$

tf_{ij} = 文書 j 中の単語 i の出現頻度

df_i = 単語 i を含む文書の数

N = 全ての文書の数

なお、分類のために用意したラベルは学習指導要領の見出しの項目を採用し、今回はその 8 ラベルのうち 4 ラベル (k2-1, k2-2, k3-1, k3-2) を使用する。

¹An evaluation of feature summarizing method for classification based on course of study

²†Shinnosuke NAKA · Graduate School of Eng. and Sci. Univ. of the Ryukyus

³‡Naruaki TOMA, Yuhei AKAMINE, Koji YAMADA, Satoshi ENDO · Faculty of Engineering Univ. of the Ryukyus

今回使用する各ラベル名の一覧

- k2-1:市場の働きと経済
- k2-2:国民の生活と政府の役割
- k3-1:人間の尊重と日本国憲法の基本原則
- k3-2:民主政治と政治参加

次に今回の提案手法を各段階に分けて説明する。

2.1 2種類のデータセットの用意

今回は 2 種類のデータセットを用意した。一つ目は、沖縄県立高校入試 [6] の社会科の問題と公民問題集 [7] から、設問中の問題説明文を省いた問題文と問題文中の指示代名詞から引用できる情報を取り入れたデータセットである。二つ目は、一つ目のデータセットに加えて語の集約を行ったデータセットである。語の集約ルールは、予め教科書などの参考資料を用いて定義した。

正解データに関しても、参考資料を使いどのラベルが適切なかの調べ、手作業で正事例を用意した。

2.2 特徴ベクトルの生成

設問に出現する名詞がその文章を特徴付けているとして、各データセットに対して MeCab[8] を用いて形態素解析を行い、名詞を抽出する。そして、抽出した名詞群から TFIDF を算出し、それを基に特徴ベクトルを生成する。

2.3 機械学習

WEKA[9] を利用して、機械学習を行った。分類器は NaiveBayes を使用し、ラベル毎に二値分類器を構築して分類 (T or F) を行う。なお、各ラベル毎に用意した正事例数は異なるために、それぞれ正事例数に合わせた数の負事例数を学習させている。

2.4 評価方法

10 分割交差検定を用いており、再現率 (recall) と適合率 (precision) の調和平均である F 値を基準として評価する。よって、F 値が大きければ大きいほど正しく分類されていると考えることができる。

3 実験

実験 1 では、語を集約する前後のデータセットの F 値を比較することで、今回の提案手法がどの程度の F 値を向上させる効果があったのかを検証した。実験 2 では、語の集約前後で分類に失敗した設問を比較し、その変化の傾向を分析した。

3.1 実験 1(語の集約前後での各ラベル F 値の比較)

各ラベル毎に 10 種類の負事例のパターンのデータセットを用意して機械学習を行った。そして、語の集約前後の各ラベルの F 値の平均を調べると次のようになった(図1)。

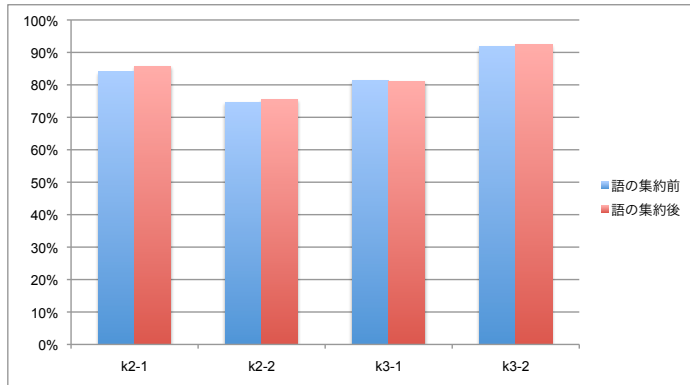


図 1: 語の集約前後の平均 F 値

この結果から、4 ラベル中 3 ラベル (k2-1, k2-2, k3-2) の F 値がわずかに向上しており、語の集約を行うことで従来よりも F 値が良くなる可以考虑することができる。

しかし、k3-1 ラベルのようにほぼ変化の無いラベルがあるために、置き換える語によってその効果が異なる可能性が考えられる。よって、語の集約することで設問にどのような影響を与えているか、その要因を調べる必要がある。

3.2 実験 2(語の集約前後での間違い分析)

語の集約前後で分類に失敗した設問がどのように変化したのかを、設問が語の集約前後で分類に失敗した回数之差を取ることで調査した(集約前-集約後)。次にその差分が大きい設問の例と小さい設問の例を挙げ、語の集約の前後でどのような影響があったのかを分析する。

語の集約をすることで良くなった設問の例

- ・ 公民問題集の 4 章 10 節の標準問題大問 4 の (3)
(3) 日本の賃金や労働時間などの最低基準を定めた法律を何というか。

置き換え前の出現単語
(日本、賃金、労働時間、最低、基準、法律)
置き換え後の出現単語
(日本、賃金、労働、最低、基準、法律)

この設問は k2-1 ラベルの正事例であり、語の集約前では 10 回中 8 回も分類に失敗していたが、語の集約を行うことで一度も失敗することなく正しく分類できた例である。

語の集約をすることで悪くなった設問の例

- ・ 公民問題集の 3 章 8 節のチェック問題第 9 問
地方財政の格差をなくすため、国から地方公共団体に支出されるお金を何というか。

置き換え前の出現単語
(地方財政、格差、国、地方公共団体、支出、お金)
置き換え後の出現単語
(財、格差、国、地方公共団体、支出、お金)

この設問は k3-2 ラベルの正事例であり、語の集約前では

10 回中 1 回だけ分類に失敗していたが、語の集約を行うことで 10 回中 6 回も分類に失敗するようになった悪い例である。

これらの要因としては、語の集約によって置き換えられた単語間でラベルに対する重要度が異なっているためだと考える。まず良い例として挙げられている設問で置き換えられた単語は「労働」である。これは予め定義した語の集約ルールで k2-1 ラベルのために置き換えた単語であり、置き換えた語のラベルに対する重要度が高かったために分類に成功しやすくなったと考えることができる。そして、悪い例として挙げられている設問で置き換えられた単語は「財」である。これは予め定義した語の集約ルールでは k2-1 ラベルのために置き換えた語であり、本来分類して欲しい k3-2 ラベルのために語の集約をした単語では無い。よって、置き換えた語が k3-2 ラベルに対する重要度が低かったために分類に失敗しやすくなったと考えることができる。

4 おわりに

本稿では、設問自動分類タスクにおける誤分類事例に関する分析を基に、設問中に出現する似た語の集約を行うことで F 値の向上を図った。その結果、語の集約を行うことで F 値の向上するラベルとそうでないラベルに分かれた。そして、その違いの要因を調べるために行った間違い分析では、語の集約を行うことで分類に成功しやすくなった例と失敗しやすくなった例について調査した。その結果、語の集約の際に置き換える単語に良し悪しがあることが発見できた。よって改善策としては、語を集約することによって分類結果が良くなった設問群と悪くなった設問群にそれぞれ共通する特徴を調べ、置き換える単語を調整することが挙げられる。

以上より、今後は更に語の集約前後における間違い分析を続け、誤分類する設問に共通する特徴を調べ、置き換える単語のリストを調整することで F 値の向上を目指していく。

参考文献

- [1] 沖縄県教育委員会: 学校教育における指導の努力点, 2012.
- [2] 文部科学省: 中学校学習指導要領, 2008.
- [3] 文部科学省: 中学校学習指導要領解説 社会編, 2008.
- [4] 名嘉ら: 学習指導要領に基づいた設問の自動分類における誤分類事例に関する分析, JSiSE, 2012, 沖縄地区, pp.208-209.
- [5] 名嘉ら: 学習指導要領に基づいた設問の自動分類タスクにおける設問外の情報付与における有効性の評価, FAN2012, 予稿集 CD-ROM, 2C3-1.pdf.
- [6] 富士教育出版社: 平成 23 年度受験 沖縄県 県立高校入試問題, 2011.
- [7] 文英堂: くわしい公民ステップアップ問題集, 2009
- [8] MeCab: <http://mecab.sourceforge.net/>
- [9] WEKA3(Data Mining Software in JAVA): <http://www.cs.waikato.ac.nz/ml/weka/>