

4Q-05

問題文の単元識別に基づく学習支援に関する基礎研究

A basic study on a learning support based on a unit classification of question contents

琉球大学工学部情報工学部：嘉陽桃子 指導教員：當間愛晃*

1 はじめに

ある教科を学習する際には多くの場合、以下の手順を踏む。学習するための最適な教材は何か、試験等で誤答があった場合、原因は何なのか、またはその教科のどの部分が苦手なのか、その部分を理解するためにはどのような知識が必要なのか、等である。このような工程を自動化し、最適な学習メニューを提供するための研究はすでに様々な試みが行われている。

例えば、e-learning の 1 つである STACK(数学オンラインテスト評価システム [1]) では、教師側が問題、解答、学習者への指導コメントを登録できる。教師側は正答と予測しうる誤答、誤答に対するコメントを登録できるので、学習者はインターネット上でもよりリアルに近い自主学習を行うことができる。

また、小学校低学年の算数における日本語文章題を対象に、文章解析と意味理解を試みた研究 [2] や、場合の数・確率の文章問題に対してコンピュータに解答させるなどの研究 [3] も行われている。

これらの研究はいずれも、充実した自主学習を実現するためのシステム構築の助けになると考えられる。

調査した先行研究は単元設定を手動で行っている。これは同じ教科でも単元や、単元内容を構成する公式に注目しても、問題文の文章構成は大きく変わることが理由として考えられる。

本研究での目的は、問題文の文章構成を解析して単元ごとの特徴を抽出し、入力された問題文から単元を識別することである。問題文から単元を自動的に識別できれば、先行研究の手法を自主学習後の自動採点に取り込むなど、より有効に活用できるのではないかと考えられる。

本研究は最終的には学習者の苦手単元を特定、可視化、再学習のための学習メニュー構築を目的としている。本稿で示した提案で実験が成功すれば、最終目標の基礎研究としても使用することができる。そのために、本稿では問題文の文章構成を解析し、自動的に単元を特定することを目的として基礎実験を行った。

2 研究計画

2.1 設問の用意

本研究では、本稿では単元識別の対象として高校数学の設問を採用した。問題文のサンプルには「高校数学の基本問題」[5] から、数 A「順列・組合せ」、「確率」を用いている。上記のサイトを利用したのは、本実験を行う上で単元ごとの問題数が重要になると考え、サンプル数がある程度確保する目的のためである。なお、図形問題は省き、文章題のみで解ける問題をのみを対象とした。

2.2 実験手順

問題文の文章構成から単元を判断するために、教師あり学習を行う。

まず、実験に使用する問題文の文章をそれぞれ形態素解析する。その際、元の問題文から、全ての英数字を半角に、句読点を「,」「。」から「,」「.」に統一した。MeCabを用いて問題文を分かち書きした後、Doc2Vec[4]を用いて問題文の特徴ベクトルを生成する。本実験では「順列・組合せ」と「確率」のそれぞれで1問ごとに特徴ベクトルを生成する。それらの特徴ベクトルを SVM(Support Vector Machine) で分類する。

Doc2Vec は文章の特徴を抽出するための技術であるが、他にも特徴抽出のための手法がある。本実験にとって Doc2Vec が有効であることを確かめるために、次項では性能比較を行う。

3 性能比較

従来の文章の特徴抽出のための手法の 1 つとして Bag-of-Words が挙げられる。ある文章の中に出現する単語の頻出度から特徴を得る手法である。とくに TF-IDF は、文書毎に出現する単語割合を用いて重み付けするための手法であり、ベースラインとして採用した。

これに対して Doc2Vec は文章の中の単語の位置と関係性を重視する手法である。Quoc Le、Tomas Mikolov らの論文 (Distributed Representations of Sentences and Documents[5]) では Bag-of-Words の 3 つの弱点 (1) 語の順序が失われること、(2) 語の意味が無視されること、(3) 特徴量のベクトル長が固定されてしまうことの克服を、論文中の実験で成功したことが報告されている。

*Momoko KAYO, Naruaki TOMA, Department of Information Engineering, University of the Ryukyus

本研究で Doc2Vec を用いる上で、問題文の特徴抽出に Doc2Vec を用いた場合に上記のような結果を得られるか事前に確認する必要がある。性能評価のための実験概要を次項に示す。

3.1 実験環境

- 特徴ベクトルを生成する問題文について
 - － 「順列・組合せ」22 問、「確率」20 問。合計 42 問。
- 比較手法について
 1. 特徴ベクトルの生成
 - － 従来の特徴ベクトル生成方法

特徴ベクトルの生成に Bag-of-Words、重み付けに TF-IDF を用いる。次元数はデフォルトの 200 に設定している。
 - － 本実験の特徴ベクトル生成方法

デフォルトの PV-DM(Distributed Memory Model of Paragraph Vecotors) を用いて語順を保存し、単語同士の繋がりから文章全体のベクトルを生成する。次元数は問題文の数を考慮し、20 に設定した。
 2. 交差検証

上記の手法で生成した特徴ベクトルを SVM を用いて分類する。その後、問題文テキストデータから順にテスト用に 1 サンプルずつ用意して交差検証を行う。

3.2 実験結果

- Bag-of-Words

averaged score = 0.975609756098
- Doc2Vec

averaged score = 0.52380952381

実験結果は Bag-of-Words が約 9 割、Doc2Vec が約 5 割の正答率になった。

Bag-of-Words で正しく分類されなかった問題文は 1 問である。「A,B,C の 3 人でじゃんけんを 1 回するとき、あいことなる確率を求めよ。」という問題だが、「じゃんけん」を含む問題文が「順列・組合せ」と「確率」に 1 問ずつあった。予測では「順列・組合せ」の問題であると出力されたが、正解は「確率」なのでこの 1 問は間違えた分類をした。

一方 doc2vec は約 5 割の正答率であったが、個々の推定結果を観察するとすべての設問を「順列・組合せ」と誤答するケースが多かった。教師データが 2 種類のためどちらか一方の値を出せばこの正答率まで達することができる。以上の結果から、本実験で使用した問題文の量と形なら Bag-of-Words のほうが良い結果が出ることが分かった。

3.3 考察

今回、このような結果になった理由として以下の要因が考えられる。

単元ごとに出現する単語に特徴がよく現れていた。例えば、「順列・組合せ」では「何通り」、「確率」では「確率」や「期待値」が挙げられる。単語同士の関係性を考慮せず、十分に分類可能だったと言えるだろう。また、本来 Doc2Vec は、最初に日本語の文章自体の特徴を学習する手順を踏むことが主流である。そのために Wikipedia などからコーパスを読み込むが、本実験では都合上その手順を省いた。さらに、用意した問題文の量が圧倒的に足りなかったことも原因の 1 つだと考えられる。

4 今後の課題

本実験の結果は、準備不足であったことが主要因であったと考えられる。本研究では今後、単体で特徴となるような単語が現れない問題文を扱うこともあるだろう。そのような場合には Doc2Vec が有効な手段となることもあり得る。今後は考察で述べた点を踏まえて実験を進めていきたい。

参考文献

- [1] 中村 泰之, 中原 敬広, 秋山 實 “STACK と Moodle で実践する数学 e ラーニング”, 数理解析研究所講義録, 2010 pp.41-44
- [2] 笹田 国博, 小高 知宏, 高賓 徹行, 小倉 久和 “算数問題文の解析とその知識表現”, 福井大学 工学部 研究報告, 1994 pp.54-58
- [3] 宮尾 祐介, 川添 愛, 松崎 拓也, 横野 光 “言語処理技術の統合的評価基盤としての大学入試問題” 言語処理学会 18, 2012 pp.995-998
- [4] “gensim models.doc2vec Deep learning with paragraph2vec”
<<https://radimrehurek.com/gensim/models/doc2vec.html>>
(アクセス日: 2017/1/13)
- [5] ”高校数学の基本問題”
<http://www.geisya.or.jp/mwm48961/koukou/index_m.htm>
(アクセス日: 2016/1/13)