

令和元年6月24日現在

機関番号：32665

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K16326

研究課題名(和文)事例ベースに基づく学習者の潜在要因分析と個別教育支援システムの開発

研究課題名(英文) Analysis of learner's potential factors based on Case-based and development of individualized education support system

研究代表者

東 るみ子 (AZUMA, Rumiko)

日本大学・商学部・准教授

研究者番号：80550102

交付決定額(研究期間全体)：(直接経費) 2,100,000円

研究成果の概要(和文)：我が国の教育現場では、一人の教員が30名以上の生徒や学生を指導することが多く、一教員が生徒や学生の理解度や状況を把握し適切な対応をとることは難しい。

そこで本研究では、大学生のデータを基に、過去に学生が授業の中で作成した文章を事例ベースとみなし、学生の理解度を推定するシステムを検討した。本研究では、大学生のリアクションペーパーをもとにテキストマイニングにより定量化を行い、統計的手法や機械学習を用いて、理解度と関連の強い要素を分析し、成績の推定を試みた。その結果、学生の主観的理解度や毎回の小テストの点数よりもリアクションペーパーの内容の方が成績を推定する要素として重要であることが明らかとなった。

研究成果の学術的意義や社会的意義

多様化する生徒に対して、多角的な指標のもとで個々の学習問題を考慮し効率の良い教育指導を行うことは、生徒のみならず負担の多い教員にとっても必要である。またそのために、生徒の理解度を把握する教育補助システムの開発は重要である。

そのシステム開発の一助として、本研究では学生の真の理解度を推定するための要因を探り、推定精度の向上を目指した。これは先進的な教育改善策に繋がる可能性を秘めており、社会的意義は大きいと考える。

研究成果の概要(英文)： In Japan, it is hard for a teacher to take control of the understanding level of students and conduct personal instructing because a teacher is in charge of many students. Then, in order to respond to the wide variety of needs for education, it is necessary for teachers to get grasp student's understanding level in real time and provide individual instruction.

As a solving the problem, a reasoning system which can predict the understanding level of students with their reaction paper was proposed in this study. We analyzed some elements that are strongly related to understanding level using statistical method or machine learning algorithm from student's writing quantified by text-mining method and tried to extrapolate their understanding level. As a result, their final exam score had stronger correlation with the content of their reaction paper than their subjective understanding level or quiz score.

研究分野：教育工学

キーワード：学習分析 テキストマイニング データ分析 学習支援

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

(1) 現在日本では少子化が進む一方、学習障害や広汎性発達障害などの目に見えない障害を抱えた児童生徒が顕在化してきており、その数は年々増加傾向にある。そのため教育現場では、生徒一人一人の教育的ニーズを把握し、生活や学習上の困難を改善することが求められている。

2012年に行われた文部科学省の調査では、通常の学級に在籍する生徒のうち小学校では7.7%、中学校では4.0%の生徒が学習面または行動面で著しい困難を示すという報告がなされており、潜在的な生徒も含めるともはや担当教員だけでは対応しきれない状況にある。しかしながら現在の学習指導は、教員の「経験」や「努力」に依存するところが多く、必ずしも全ての教員が十分な知識や情報を持っているとは限らない上、教員が行う業務の増加に伴い、きめ細かな指導が行えない状況にある。

(2) 国内外の教育学の分野では、教育現場でのICT化に伴い、生徒の学習履歴などの学習データからつまづきを発見し、苦手な単元の学習を促す学習支援システムの開発が進められている。これらの研究で扱うデータはテストの点数などの量的データであり、教科学習における弱点をフィードバックすることを目的としている。しかしながら、学力だけでなく生徒の行動への気づきや学習以前の問題の発見、解決には更なる研究が必要不可欠である。

2. 研究の目的

本研究ではこの問題を工学的な立場で、過去の事例から最適な指導を選択する意思決定問題として捉える。見えない障害を抱えている生徒及びそうでない生徒も含め、様々な生徒の学習データや所見を分析し、生徒が抱えている可能性が高い問題や学習の改善ポイントを教員に提示することができる教育支援システムの開発(図1)を目指す。

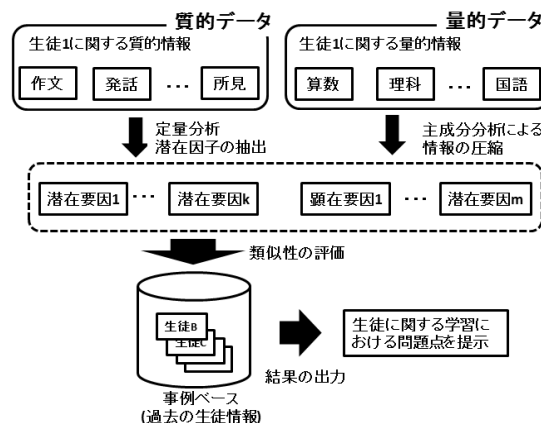


図1: 教育支援システムの流れ

また、従来使われてきた学習履歴などの量的データだけでなく、生徒の潜在的要素が反映されやすい、生徒の書いた文章(発話)などの質的データを定量化・分析し、生徒の学習に及ぼす潜在要因を解明することをも氣的とする。

3. 研究の方法

本研究では、大学で利用されているリアクションペーパー等の自由記述を質的データとしてその内容を定量的に分析し、記述内容のどのような要素が学生の理解度または成績と深い関連があるのかを分析した。

リアクションペーパーに関する先行研究では、その機能を検討した研究や、使用することによる効果を明らかにした研究が多く存在するが、一方、リアクションペーパーから得られる情報と学生の理解度の関連を分析している研究は少ない。本研究では、カリフォルニア大学バークレー校で実施された Minute Paper [1] (「今日学んだことで一番大切なことは何か。今日の講義における疑問点は何か。」を1分間で書かせるもの)をベースにし、学生の主観的評価項目を加えた学習の振り返りシートを用いて質的データ分析を行った。振り返りシートの内容は次の通りである。

設問1: 本日の講義の理解度(5段階評価)

設問2: 予備知識(予習)(5段階評価)

設問3: 講義に関する「振り返り」(その日に学んだ内容、理解した内容、分からなかったこと、質問など)(自由記述)

振り返りシートは、大学1年生向けの「統計学入門」の授業で使用した。毎授業の終わり10分前にLMSを使い、学生に入力をお願いした。設問3の質的データの分析には、テキストマイニングのソフトウェア KHCoder [2] を使用した。

また振り返りシートの情報に加え、毎授業で実施した小テストの点数、中間試験の点数、期末試験の点数も量的データとして扱い、分析対象とした。

以上のデータに対して、統計的手法、機械学習による手法を施し、学生のどのような要因が理解度、あるいは成績に影響を及ぼすのか、どのような学習データを用いることで理解度などに問題を抱えている学生を予測することができるのかを検討した。

4. 研究成果

(1) 基本統計量と質的データの特徴

はじめに設問1と2に関して単純集計を行った。その結果を表1に示す[3]。分布に関する内容を扱う第6回以降の講義から若干理解度の平均が低くなってきていることが分かる。また同時に第6回から標準偏差の値も大きくなってきていることから、理解度にばらつきが見え始めていることがうかがえた。

次に、それぞれのデータ間のピアソン積率相関係数を算出した(表2)。各データの内容は、13回分の講義における各学生の主観的な理解度の平均、予備知識の平均、小テストの平均、総合成績である。なお、総合成績は中間試験と期末試験の点数から算出される。主観的理解度は、予備知識、および小テストと正の相関がみられるものの、最終的な成績との有意な相関はみられなかった。毎回実施した小テストと成績における相関に関する有意水準5%での有意性はみられなかった。この結果から、学生の主観的な理解度は必ずしも成績に繋がっていないことが明らかになった。

表1 各回における理解度、予備知識の平均値と標準偏差

| 回 | 講義内容 | 理解度平均(SD) | 予備知識平均(SD) |
|----|-------------------------|-------------|-------------|
| 1 | 統計的方法の性質 | 4.17 (0.88) | 2.67 (1.33) |
| 2 | 度数分布表とヒストグラム | 4.34 (0.65) | 2.52 (1.39) |
| 3 | 代表値と散布度 | 3.90 (0.99) | 2.33 (1.45) |
| 4 | 標本空間、複合事象の確率 | 4.15 (0.78) | 4.29 (0.90) |
| 5 | ベイズの定理、順列、組合せ | 4.44 (0.74) | 3.84 (1.19) |
| 6 | 確率分布の平均と分散 | 3.48 (1.14) | 3.40 (1.14) |
| 7 | 2項分布、正規分布の標準化 | 3.54 (0.97) | 1.77 (1.23) |
| 10 | 無作為抽出、標本分布、不偏推定量、中心極限定理 | 4.00 (0.92) | 2.22 (1.43) |
| 11 | 点推定と区間推定、平均値の区間推定(大標本) | 3.71 (0.87) | 1.86 (1.31) |
| 12 | 割合の区間推定、平均値の区間推定(小標本) | 3.56 (1.09) | 1.69 (1.19) |
| 13 | 仮説検定、平均値の検定、割合の検定 | 3.34 (1.03) | 1.57 (1.13) |
| 14 | 平均値の差の検定、割合の差の検定、小標本の検定 | 3.04 (1.05) | 1.65 (1.19) |
| 15 | 相関係数、単回帰分析 | 3.45 (1.10) | 2.05 (1.43) |

第8回と9回の講義は、中間テストとその解説の回だった為、振り返りシートは実施していない。

表2 学生データの基礎統計量と相関係数

| | 平均値 | SD | 相関係数 | | |
|------|-------|-------|---------|------|------|
| | | | 理解度 | 予備知識 | 小テスト |
| 理解度 | 3.79 | .58 | 1 | - | - |
| 予備知識 | 2.48 | .72 | .49 *** | 1 | - |
| 小テスト | 78.75 | 12.55 | .49 *** | .13 | 1 |
| 成績 | 57.15 | 20.35 | .17 | -.14 | .21 |

*p<.05, **p<.01, ***p<.001

設問3の自由記述のテキストデータに対して定量分析を行った。テキストマイニングのソフトウェアであるKHCoder[2]を用いて、どのような記述が含まれているのか、自由記述の特徴を分析した。その際、複合語としてまとめられる語(例:標準偏差)は1つにまとめた。分析対象となる語の数14230、出現回数の平均は8.89であった。

分析の結果、全体的な形容詞・形容動詞の出現頻度をみると、「難しい」が最も多く、続いて「多い」「新しい」「必要」「複雑」「苦手」であった。これらの語の多くは、講義の感想として用いられていたことから、学生の講義に対する印象が「難しく、苦手意識の強い講義」になっていたことが伺える。「新しい」に関しては、「新しい公式が出てきた」など、講義の中で今までの予備知識にはない内容が出てきた時などに用いられていた。さらに名詞に関しては「確率」「高校」「標準偏差」「前回」「平均」の順で多く出現しており、主に専門用語が多く、その回で学んだことを振り返る内容の文章で用いられていた。

そこで本稿では、振り返りがしっかり行えている学生ほど真の理解度が高いのではないかという仮説をたて、設問3の文章(講義13回分)に含まれている専門用語の出現割合と成績の相関分析を行った。その結果、弱いながらも有意な正の相関がみられた($r=0.33$, $p<.001$)。また、専門用語の出現割合と小テストの平均に関しても、弱い相関がみられた($r=0.34$, $p<.001$)。

(2) 質的データの分類

自由記述に含まれる専門用語と成績に強い相関がみられなかったため、次に自由記述の文書

を分類し、その内容と学生の理解度や成績との関連がみられるのかどうかを検討した[4]。

分類は表3に示す8分類を設定した。各文書の分類方法として、文字数や文書に出現する単語などに基づき分類を行った。また、「軽い振り返り」「深い振り返り」「授業内容」「感想」「未分類」タグは、各文書にいずれか1つ付与されるようにし、複数のタグが付与された場合は、人間が直接目視で確認することで分類の修正を行った。そして、これらの分類に属する各文書にタグ付けを行い、どのタグが理解度、成績との相関が高いのかを分析した。

表3 分類の種類

| タグ名 | 分類内容 | 文書数 |
|--------|---|-----|
| 軽い振り返り | 自分の理解に関する単純な記述。単に「～を理解した」「～が分かった」「～を知った」「分からなかった」などの記述。 | 455 |
| 深い振り返り | 自分の理解度に関する記述で、理解した理由やどのようなところが理解できたのか、なぜ分からなかったのかなど、内容の記述が詳しく書かれているケース。 | 187 |
| 授業内容 | 単にその日学んだ内容のみが書かれている記述。 | 24 |
| 感想 | 「～頑張りたい」「もっと復習したい」「計算が面倒くさい」など感想のみの記述。 | 115 |
| 未分類 | 上記4つの分類に当てはまらない内容。 | 2 |
| 要望 | 授業や説明に対する要望の記述。 | 23 |
| 疑問 | 具体的にどの部分が、と指定した上で「"なぜ"そうなるのかがわからない」とか「～に疑問を感じた」などの記述。 | 88 |
| 質問 | 具体的に書いた内容の後で「～でしょうか?」「～知りたい」などの記述。 | 11 |

(3) 質的データの統計分析

相関分析の結果、表4に示すように「深い振り返り」は、期末試験と小テストにおいて弱い正の相関を示した。また「疑問」タグに関しても、同様に弱い正の相関がみられた。

表4 自由記述の分類と各テストの得点の相関係数

| | 期末試験 | 中間試験 | 小テスト |
|------|---------|----------|------------|
| 軽い理解 | 0.076 | 0.223 | -0.096 |
| 深い理解 | 0.314 * | 0.090 | 0.432 *** |
| 授業内容 | -0.062 | -0.191 | 0.004 |
| 感想 | -0.367 | -0.319 * | -0.466 *** |
| 要望 | 0.030 | 0.150 | 0.116 |
| 疑問 | 0.272 * | 0.091 | 0.431 *** |
| 質問 | 0.210 | 0.111 | 0.192 |

* $p < .05$, ** $p < .01$, *** $p < .001$

さらに、期末試験の得点を予測するのに最適な項目を分析するために、期末試験の得点を目的変数、理解度、予備知識および各分類タグの集計数を説明変数とし、ステップワイズ法による重回帰分析を行った。その結果、「軽い振り返り」($r = 0.422, p < .001$)と「深い振り返り」($r = 0.570, p < .001$)を説明変数としたときに、AICが最も小さく最適なモデルとなった。一方、決定係数は0.211 ($p < .001$)と低かった。

以上の結果から、自由記述の「深い振り返り」「疑問」の記述内容が、試験やテストの得点と関連があることが分かった。すなわち、「深い振り返り」を行う学生ほど、講義内容をより理解しようとしており、その結果得点も高くなる傾向にあった。

さらに興味深い結果として、「感想」のみ多く述べている学生ほど、中間試験や小テストの得点が低くなる傾向にあった。これは、「難しかった」や「簡単だった」などの単純な授業感想が、授業内容に深くコミットしていないことを表しており、その結果として試験や小テストの得点との間に負の相関が表れたのではないかと考える。

(4) 機械学習による分析

前節で示したように、学生の自由記述の文章が、学生の主観的理解度や予備知識、小テストよりも期末試験(成績)と関連があることが明らかとなったため、これらの情報を用いて成績不良者を推定することを試みた。推定には機械学習の手法を用いた[5]。

目的変数には定期試験の点数を使用した。説明変数には次の13変数を用いた。授業毎の(a1)小テストの点数、(a2)理解度、(a3)予備知識度の3つの変数。学習の振り返りの自由記述文章における(b1)専門用語の平均出現頻度、(b2)文字数の平均の2つの変数。さらに振り返り文章にタ

グ付した(c1)深い振り返り, (c2)軽い振り返り, (c3)質問, (c4)感想, (c5)一言, (c6)要望, (c7)疑問, (c8)授業内容のみ, の 8 つの平均出現回数を説明変数とした。これらの変数の中で, 互いに中程度の相関がみられるものもあったが, 本稿では成績不良の学生を推定することを目的とするため, 全ての変数をそのまま予測モデルの変数として扱うことにした。

教師あり学習でクラス分類ができる 7 つの機械学習のアルゴリズムを用い, 成績の推定を行った。ここでは, 主に成績不良の学生を推定することを目的としているため, 目的変数である定期試験の点数 (平均値: 58.7, SD: 18.7) を 2 つのカテゴリ (60 点以上: 問題なし, 60 未満: 問題あり) に分類し, バイナリデータとして扱った。

それぞれの機械学習アルゴリズムによる推定精度(F 値)を表 5 に示す。7 つの中で, Random Forest で成績を推定したときの精度が 53.3%と最も高かったため, 本研究では Random Forest による分析結果をもとに考察を行った。

表 5 成績推定における精度 (F 値)

| アルゴリズム名 | 精度 |
|--------------------------|--------------|
| Naive Bayes | 42.9% |
| Generalized Linear Model | 30.8% |
| Logistic Regression | 30.8% |
| Deep Learning | 40.0% |
| Decision Tree | 40.0% |
| Random Forest | 53.3% |
| Gradient Boosted Trees | 30.8% |

Random Forest のランダムサンプリングする説明変数の数 (木の深さ) を 4, 木の数を 100 とし, 成績が「問題あり」の学生を決定する際の影響度の高い説明変数を調べた。その結果を図 2 に示す。成績「問題あり」に正の影響を及ぼしている変数は, (c4)感想と(c2) 軽い振り返りであった。すなわち, 振り返り文の内容が「感想」または「軽い振り返り」に分類された回数が多かった学生ほど, 成績が悪かったことを示している。

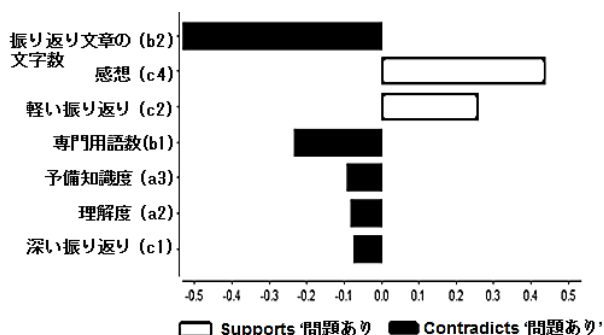


図 2 成績「問題あり」に対する説明変数の重要度

一方, (b2) 文字数の平均は負の影響度が最も高くなっていることから, 振り返り文章の文字数が多い学生ほど成績が悪くないといえる。また, (b1) 専門用語に関しても成績「問題あり」に対して負の影響があることがわかった。

以上の統計分析, 機械学習を用いた分析の結果から, 成績不良者を早い段階から見つけるには, 学生の自由記述の内容が有効であることが分かった。その中でも「感想」や「軽い振り返り」しか記述しない学生が成績不良につながる可能性が高いため, 早い段階でのケアが必要であると考えられる。

今後の本研究の課題は, 予測精度を向上させるためにデータ数を増やすことが考えられる。また, 学習者の自由記述の文章を, どの視点から分類をすれば, 推定精度が上がるのかを再検討する必要がある。

<引用文献>

- [1] B.G. Davis, R. Wilson, L. Wood: 授業をどうする! カリフォルニア大学バークレー校の授業改善のためのアイデア集”, 東海大学出版会 (1995)
- [2] 樋口耕一: “社会調査のための計量テキスト分析”, ナカニシヤ出版, (2014)
- [3] 東るみ子, 振り返りシートを用いた学習者理解度の分析 第 41 回教育システム情報学会全国大会, (2016)
- [4] 東るみ子, テキストマイニングによる学習者の特性と理解度の分析 第 42 回教育システム情報学会全国大会, (2017)

[5] 東るみ子, 機械学習による学習者の理解度推定, 第 43 回教育システム情報学会全国大会, (2018)

5 . 主な発表論文等

[学会発表](計 3 件)

東るみ子, 機械学習による学習者の理解度推定, 第 43 回教育システム情報学会全国大会, 2018.

東るみ子, テキストマイニングによる学習者の特性と理解度の分析, 第 42 回教育システム情報学会全国大会, 2017.

東るみ子, 振り返りシートを用いた学習者理解度の分析, 第 41 回教育システム情報学会全国大会, 2016.

6 . 研究組織

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。