

学習における評価の時系列変化に着目した評価変動要因の分析

Analysis of Grade Score Variation Factors Focusing on Time Series Change of Grade Score in Learning

戸田 綾佳^{*1}, 久野 弘暉^{*1}, 高橋 聡^{*2}, 吉川 厚^{*1}, 山村 雅幸^{*1}
Ayaka TODA^{*1}, Hiroki KUNO^{*1}, Satoshi TAKAHASHI^{*2}, Atsushi YOSHIKAWA^{*1}, Masayuki YAMAMURA^{*1}
^{*1}東京工業大学, ^{*2}関東学院大学
^{*1}Tokyo Institute of Technology, ^{*2}Kanto Gakuin University
Email: toda.a.aa@m.titech.ac.jp

あらまし: ある時点にて同程度の成績である児童生徒集団において、学年が上がるにつれて成績が伸びる人と伸びない人が存在する。この成績変動の要因となる項目を特定することは、児童生徒の成績の向上に繋がると考えられる。本研究では、ある児童生徒集団が経年で受検した算数・数学の学力調査の成績の変動に注目し、工学的手法を用いて分析を行った。そして、算数・数学の成績変動と相関の強い項目を抽出した結果、算数・数学の成績変動に国語の能力が関係していると示唆した。

キーワード: 小学校算数, 中学校数学, 小中学校国語, 成績評価, 経年変化

1. 問題の所在

児童生徒の能力を育むことが学校教育の目的である。その中で、特に科目ごとの能力については、成績が上がったり上位を維持できる児童生徒がいる一方、成績が下がったり下位を維持してしまう児童生徒がいる。しかし、このような経年変化による成績の変動を引き起こす要因の分析を試みる研究はあまりなされていない。

そこで、本研究では数学の成績に注目し、まず前提となる成績が上昇する児童生徒達と下降する児童生徒達がいることを実験により確認する。そして、数学の能力の変動に影響があると示唆される能力を実験により示す。

2. 研究の方法

2.1 手法

経年変化による成績の変動を引き起こす要因を分析するにあたり、以下の手順で分析を行う。

はじめに、どのような評価の変動が起こっているのか明確にする必要がある。時系列の数学の成績評価の変動パターンを明らかにする手順名をクラスタリングとする。クラスタリングには、類似度にデータ間のユークリッド距離を用いた、非階層的クラスタリングの k-means⁽¹⁾ を用いる。

次に、成績の変動の違いを引き起こす要因の分析を行う。この手順名を依存関係の抽出とする。本研究では、数学の成績変動に数学以外の能力が関わっている可能性を考慮し、国語と数学の項目において、複数の変数間の定性的な依存関係をグラフ構造によって表し、定量的な関係を条件付き確率で示したモデルである⁽²⁾、ベイジアンネットワークを用いる。

2.2 使用データ

本研究ではある市町村の公立小中学校に属する小学校第4学年から中学校第3学年を対象に実施され

た2014年から2018年の5年分の学力調査を用いる。この市町村では小学校第6学年、中学校第3学年では団体Bによる学力調査を行い、それ以外の学年では団体Aの学力調査を行う。そこで、今回は団体Aによる学力調査のデータのみを用いた。2014年に小学校第4学年であった学年の経年データを Group1、その1学年上の2014年に小学校第5学年であった学年の経年データを Group2 とする。本研究では、5年間全てこの市町村で受験した児童生徒のデータのみを使用する。回答者は Group1 が 168 人、Group2 が 201 人である。

2.3 団体Aによる学力調査

団体Aによる学力調査は、国語及び算数または数学とする。設問の情報として、その設問の解答形式や問題の内容、領域などが含まれる。また、児童生徒個人の情報として、全設問についての正誤情報が含まれる。選択問題については児童生徒が答えた選択肢番号とその正誤情報が示されており、記述問題については児童生徒の解答内容の解答型が示されている。また、各教科の合計点における全国偏差値がデータとして含まれる。

3. クラスタリング

3.1 解析

データを時系列データと見なし、評価の変動パターンを明らかにすることを目的として、k-means を用いて Group1 および Group2 のクラスタリングを行う。分析値には小学5年次の算数、中学1・2年次の数学の成績を用いる。しかし、k-means に3学年での成績をそのまま入力として与えても時系列的変動が加味されないため、時系列の差分も含めた。

今回は成績上位維持層、成績下位維持層、成績上昇層、成績下降層の4クラスタが現れると考え、クラスタ数4でクラスタリングを行った。使用したパ

パッケージは Python の scikit-learn である⁽³⁾。

3.2 結果

Group2 において、分かれた 4 クラスごとに時系列変化を示した図と、クラスごとの基本統計量を以下に示す。Group1 においても同様の結果を得た。

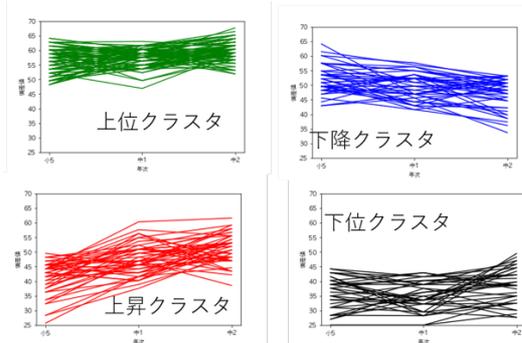


図 2 Group2 のクラスターリング結果

3.3 考察

Group1,2 共に、左上の図は成績上位維持層が得られたと解釈できる。次に、右上の図においては成績下降層が得られた。また、左下の図は成績上昇層が得られたと解釈できる。最後に右下の図では緩やかな上昇も見られるが、全体として成績が低いため、成績下位維持層と解釈する。

この 4 クラスの中で、上位クラスと下降クラス間、上昇クラスと下位クラスにおいて成績の差が広がっていると解釈できる。そこで次章では、数学の成績の差が広がる要因の分析を行う。

4. 依存関係の抽出

4.1 実験

前章の結果より、以下の 2 組のクラスに差が広がるような変動が見られた。

- ・ 成績上位クラスと成績下降クラス
- ・ 成績上昇クラスと成績下位クラス

上記 2 組において初期の段階でどういった項目に依存関係があり、変動が起こったのかをベイジアンネットワークを用いて分析する。目的変数にクラス番号、説明変数に小学校 5 年次の国語と数学の項目ごとの正誤 1/0 を用いる。小学校 5 年次の成績を使用する理由は、時系列データとして使用したデータの最初の時点が小学校 5 年次だからである。今回は BIC スコアを用いたスコアベースで、山登り法により構造学習を行った。また分析には R パッケージの bnlearn を用いる⁽⁴⁾。

4.2 結果

成績上位クラスと成績下降クラスの間では Group1,2 で共通して依存関係のある問題としてあげられた問題は以下の国語の 1 問であった。

- ・ 与えられた情報を読み取り、ポスターに補足する文章を書くことができる。

同様に成績上昇クラスと成績下位クラス間において依存関係のある問題としてあげられた問題

は以下の国語の 1 問であった。

- ・ 指定された長さで文章を書くことができる。

4.3 考察

成績上昇クラスと成績下降クラスの間では共通して「与えられた情報を読み取り、ポスターに補足する文章を書くことができる」という項目が得られた。この項目は問題を解く上で、文章とポスターの 2 つの対応関係を読み取れるかを問われる問題である。よって、数学の能力を上位に維持できるかは対応関係を見つけ出すことができているかがキーになる可能性があると考えられる。

一方で、成績上昇クラスと成績下位クラスの間では共通して「指定された長さで文章を書くことができる」という項目が得られた。この項目は作文問題において文字数制限を守って回答できているかが問われている。つまり、字数指定を守っていないものが成績下位クラスには多いことが伺える。

5. おわりに

本研究では、生徒個人の数学の評価が時系列でどのように変動しているかを分析し、成績変動の原因となる要素を推定した。

まずどのような評価の変動が起こっているのか明らかにするにあたり、個人の評価の変動を、時点間の変数を加味した k-means を用いてクラスターリングすることで解釈可能な変動パターンを抽出した。結果として、成績上位クラス、成績下位クラス、成績上昇クラス、成績下降クラスと解釈可能な 4 クラスが得られた。

また、成績変動の違いを引き起こす要因があるのかを明らかにするにあたり、変動パターンに依存関係のある項目を、ベイジアンネットワークを用いて抽出した。その結果、成績上位だった生徒が上位を維持できるか成績が下降してしまうかについては、小学校 5 年次の国語における対応関係を問う問題を解けるか否かが関係することが示唆された。また、成績下位だった生徒が下位を維持してしまうか成績が上昇するかについては、小学校 5 年次の国語における作文の文字数制限を満たして解けるか否かが関係することが示唆された。

参考文献

- (1) A.K.Jain: “Data clustering: 50 years beyond K-means”, Pattern recognition letters, Vol.31, No.8, pp.651-666 (2010)
- (2) 本村陽一, 佐藤泰介: “ベイジアンネットワーク: 不確定性モデリング技術”, 人工知能学会誌, Vol.15, No.4, pp.575-582 (2000)
- (3) F. Pedregosa, et al.: “Scikit-learn: Machine learning in Python”, Journal of machine learning research, Vol.12, Oct, pp.2825-2835 (2011)
- (4) M. Scutari, R.Ness: “Package ‘bnlearn’”, Bayesian Network Structure Learning, parameter learning and inference and inference, R package version 4.4, Vol.1 (2019)