

ICT 上の学習者特性を用いた中途退学者推論

Method of dropout students using student ICT-based Learner characteristics

高橋 大樹^{*1}, 山川 広人^{*1}, 小松川 浩^{*1}Hiroyuki TAKAHASHI^{*1}, Hiroto YAMAKAWA^{*1}, Hiroshi KOMATSUGAWA^{*1}^{*1} 公立千歳科学技術大学大学院光科学研究科^{*1} Graduate School of Photonics Science Chitose Institute of Science and Technology

Email: takahashi214@kklab.spub.chitose.ac.jp

あらまし：我々は先行研究にて Deep Learning を用いて、大学内のデータを分析し、在学生の中から将来的に中途退学する学生を推論するプログラムの開発を行ない、7 割から 8 割の推論精度を得た。本研究では、先行研究のデータに加え学習者特性を活用することで学生の学習履歴だけではなく学生の自己評価に基づく数値化された学習者特性からより良い推論精度の推論を行うことができると考えている。

キーワード：ICT 中途退学者推論 データ分析 プログラム

1. はじめに

学生のデータを活用して退学者動向の解析は、Institutional Research(IR)の観点でも重要なテーマになっている。2014 年に文部科学省が公表した中途退学者の実態調査の結果では、1163 校の大学・短期大学・高等専門学校に対し、2012 年度中途退学者の状況を調査したところ、同年代における退学者は、全学生数の 2.65%にあたる 79,311 人となっている⁽¹⁾。また近年、Deep Learning(以下 DL)が機械学習の新たなアプローチとして注目を浴びている。DLに関する話題は画像認識、音声認識、自然言語処理、ゲームなど様々な領域に広がり、教育への適用も期待されている。そこで本研究では先行研究⁽²⁾で行った DL を用いた中途退学者推論の精度の精度向上を目指し、学習者特性の活用の検証を行った。

2. 本研究の目的

これまでの分析では、学生に関する ICT 上のデータを活用して中途退学傾向を分析していた。しかし、ICT 上の学生の学習履歴だけを用いて分析を行っていたため、学習履歴以外の個人個人の学習に対する傾向などを分析することができていなかった。そのため、本研究では学習者特性のデータを活用し、中途退学傾向をさらに分析できると考えた。学習者特性とは学習者の学習に向けた考え方や志向などの特性である。今回の研究では学生の自己評価に基づく数値化された学習者特性を用いることとした。

3. 先行研究

3.1 先行研究 1

本研究グループによる先行研究では中途退学者推論プログラムを用いて、中途退学者の推論精度のさらなる向上を目的として推論を行った⁽²⁾。A 大学のデータを学習データとして入力し、分析を行った。取り扱う学生のデータに関しては、研究倫理委員会による確認のもと手続きを行い、学生番号や氏名などの個人情報をも特定できる情報を匿名化して取り扱

った。本研究で用いたデータは A 大学の複数のデータから取得し、データセットにした。データセットとは本研究で用いるために整えたデータ群である。本調査の対象は、1998 年度から 2017 年度に A 大学の学生から一部を抽出した計 3514 人となった。3514 人の内、中途退学をした学生の数は一定数存在している。本研究では、E ラーニングシステムと大学の講義を管理するシステムと学生の過去の成績データを管理するシステムの 3 つのデータベースから得た 37 列のデータを使用した。E ラーニングシステムは学習時間、学習の進捗率、E ラーニングを利用したテストの結果などの学習状況のデータを利用し、大学の講義を管理するシステムでは入学年、GPA、学科、出席率などのデータを利用し、学生の過去の成績データを管理するシステムでは高校のランク、入学方法、基礎学力テストの結果などのデータを利用した。分析の結果、学習比率を 80:20 とし、中間層が 5 層でノード数が 10 個>15 個>20 個>25 個>30 個と増加していく構造で、batchsize が 548、学習回数が 1000 回という条件において、推論精度が 7 割から 8 割となった。

3.2 先行研究 2

本研究グループでは、学習者特性を診断・可視化するシステムの開発と評価も行なっている⁽³⁾。先行研究では学習者特性を 5 件法による計 81 問の質問を学生に行い、その結果から学生がより容易に自身の特性を確認できる尺度となる学びのタイプという独自の分類に分けた。学びのタイプの 3 特性は「考えようとする意識」「行動しようとする意識」「周りの評価を気にする意識」の 3 つに分けられ、81 問の質問全ての合計数値と国内の 11 大学の平均値を比較し、平均値以下を L、平均値以上を H とし、LLL から HHH の 8 タイプについて分析を行なった。タイプごとにアドバイスを作成し、学生にフィードバックを行えるようになっている。

表1 学習者特性について

特性	質問数
考えようとする意識	33 問
行動しようとする意識	30 問
周りの評価を気にする意識	18 問

4. 本研究で利用した学習者特性について

本研究で用いる学習者特性のデータは A 大学の 2013 年から 2017 年のデータを先行研究と同様に 81 問の質問を行い、その結果を数値化した。その後先行研究と同様に 3 特性に分けた。考えようとする意識を特性 1、行動しようとする意識を特性 2、周りの評価を気にする意識を特性 3 とした。そして数値から 3 特性を H と L に閾値で分け、その後 LLL から HHH までの 8 タイプとした。表 2 に記す。その数値と 8 タイプから分析、推論を行った。

表2 学びのタイプと特性

タイプ	特性 1	特性 2	特性 3
LLL	L	L	L
LLH	L	L	H
LHL	L	H	L
LHH	L	H	H
HLL	H	L	L
HLH	H	L	H
HHL	H	H	L
HHH	H	H	H

5. 本研究

本研究ではまず学習者特性と中途退学傾向について分析を行った。具体的には、3 特性の数値を説明変数として中途退学傾向を目的変数として主成分分析を行った。分析の結果、表 2、より第一主成分は学習者特性が第二主成分は中途退学が影響している。図 1 より、第一主成分である横軸に対して 3 特性が、第二主成分である縦軸に対して中途退学傾向があることが表された。そのため中途退学傾向に対しては特徴量をあまり持たないと予想される。

表3 学習者特性と中途退学傾向の固有ベクトル

	第一主成分	第二主成分	第三主成分
特性 1	0.5848	-0.1546	0.5362
特性 2	0.6805	0.0275	0.0662
特性 3	0.4396	0.2486	-0.7898
中途退学傾向	-0.0393	0.9558	0.2903

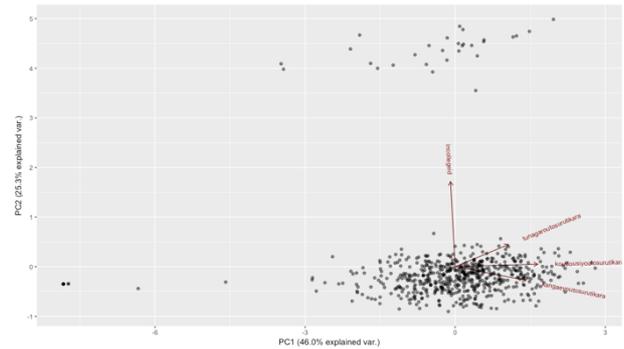


図1 主成分分析のプロット

次に粒度を下げて、これまでの先行研究のデータに学習者特性を LLL から HHH の 8 タイプの状態、さらにデータがない 2012 年以前のデータを加えた 9 タイプのデータ列として中途退学者推論を行なった。入力データは 38 列となった。先行研究と同様に中間層が 5 層でノード数が 10 個>15 個>20 個>25 個>30 個と増加していく構造の DL で推論を行なった。その結果これまでの結果と変わらない 7 割程度の結果となった。

6. 今後

本研究では学習者特性を中途退学者推論に焦点を当てて活用していたが、中途退学傾向に対しては特徴量をあまり持たないと予想された。そのため今後は、中途退学傾向の重要な特徴量である出席率や e ラーニングの学習状況など学習履歴のデータを補足し、中途退学者推論精度向上のために活用できる可能性を考えている。

参考文献

- (1) 文部科学省:”学生の中途退学や休学等の状況について” (http://www.mext.go.jp/b_menu/houdou/26/10/_icsFiles/afieldfile/2014/10/08/1352425_01.pdf)(2018 年 6 月 10 日アクセス)
- (2) 高橋 大樹, 小松川 浩, ” ICT 上の学生データを用いた中途退学者の分析手法の検討”, 2018 年度教育システム情報学会 全国大会, (2018)
- (3) 山川 広人, たなか よしこ, 小松川 浩, ” 学習者特性を診断・可視化するシステムの開発と評価”, 教育システム情報学会誌 Vol. 33, No. 2 pp. 104-109, (2016)