

教学 IR 支援に向けた SOM による退学者の傾向分析

Trend Analysis of Student Dropout using Self-Organizing Map for Institutional Research of Education and Learning

高橋 駿嗣^{*1}, 小松川 浩^{*1}

Toshitsugu TAKAHASHI^{*1}, Hiroshi KOMATSUGAWA^{*1}

^{*1}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: cist.takahashi@gmail.com

あらまし：本研究の目的は、ICT 教育支援システム上に蓄積された学生の教学データを活用し、学生一人ひとりの教学に関する傾向を把握できる教学 IR 支援システムの確立にある。今回は、目的に向けて、SOM (Self-Organizing Map) による教学データの分析システムを試作した。そして、そのシステムを利用し、学生の教学に関する傾向、特に退学状況を把握できるか調査した。

キーワード：教学 IR, ICT 教育支援システム, SOM

1. はじめに

近年、大学の質保証への取り組みに伴い、IR (Institutional Research) 活動が注目され始めている。特に、学生の教学に関する側面に視点を当てた教学 IR 活動は、教育支援に役立つと考えられている[1]。

一方、多くの高等教育機関では、e ラーニングやポータルサイトなどの ICT 教育支援システムが普及してきている。ICT 教育支援システムには、学生の教学に関するデータ (以下、教学データと称す) が各学生を識別できる形で大量に蓄積されている。そのため、蓄積された教学データを分析することで学生一人ひとりの教学に関する傾向を把握できる可能性がある。さらに、分析結果を可視化し、活用することで教職員による学生へのきめ細かな指導・支援に役立つ可能性もある。しかし、現状そのような事例はあまり見られない。

そこで、本研究では、ICT 教育支援システム上にある教学データを活用し、学生一人ひとりの教学に関する傾向の把握に向けて、SOM による教学データの分析方法を提案し、それを適用したシステムを試作した。そして、作成したシステムを利用して、学生の教学に関する傾向、特に退学状況を把握できるか調査した。

2. 提案した分析方法

本研究では、目的に向けて SOM による教学データの分析方法を提案した。分析のプロセスは、まず、SOM を用いて多変量の教学データを分析・可視化した。そして、退学者が分布するグループの特徴を教学データと照らし合わせて確認し、退学者の傾向を把握した。

2.1 SOM

本研究では、学生の多変量な教学データ全体を考慮して、学生一人ひとりの教学に関する傾向を把握していく。また、退学者の傾向把握のために、ある

程度似た属性を持つ学生同士をグルーピングし、二次元上に可視化したい。これにより、グループごとに退学者の分布や特徴が確認でき、退学者の傾向が把握しやすくなると考える。そのため、本研究では、多変量のデータ分析やその結果を二次元上に可視化できる SOM アルゴリズムを採用した。

3. 教学データの分析システム

本研究では、退学状況の傾向把握のために 2 節の分析方法を適用した分析システムを試作した。作成した分析システムは、データベースにある学生の教学データを抽出し、それを標準化して csv ファイルで出力する機能とその csv ファイルを用いて SOM による分析を行い、その結果を png 形式で出力する機能を有している。可視化した二次元マップには、学生の学籍番号がプロットされる。そして、多変量データを解析した結果、類似度が高い学生同士は、グループ化され、色分けされるようになっている。SOM の計算処理には、データ分析ツールのデファクトスタンダードである R 言語を利用した。学生の教学データの標準化や標準化した数値の出力は、Java 言語を利用した。学生の教学データの管理・抽出は、PostgreSQL を利用している。教学データの分析システムでは、解析から出力までを一体的に行える仕様としている。

なお、作成した分析システムは、試作段階であるため、一部手動で行わなければならない作業がある。1 つ目は、二次元マップにプロットされる学籍番号から退学者と在学生在を分類分けする作業である。2 つ目は、二次元マップ上に色分けされているグループの特徴を解釈する作業である。現状、どちらも技術的な要因から実現できていない。

4. 退学者の傾向調査

3 節で作成した分析システムを利用し、学生の教学データの分析をした。そして、その結果から退学

者の傾向を把握できるか調査した。

まず、分析を行う前段階としてデータの整理を行った。ICT教育支援上には、多種多様な教学データが蓄積されているため、「学習関連」、「キャリア関連」、「その他」に分類した。今回は、「学習関連」の教学データを軸に分析し、退学者の傾向を把握できるか調査した。

次に、今回の傾向調査の対象と分析する学生の教学データを表1に示す。

表1. 退学者の傾向調査で利用したデータの一覧

フィールド	A 大学
対象者	2013 年度の入学前教育受講者
期間	入学前教育から1年次の春学期まで
教学データ	GPA 授業への出席率 演習への取り組み時間 プレースメントテストの点数

表1に示すA大学では、AO・推薦入学者向けにA大学のeラーニングシステムを用いた入学前教育での学習によって、学習履歴が豊富に蓄積されていることから入学前教育受講者とした。また、特に退学者が多かった2013年度とした。分析したデータの期間は、退学者の早期発見を視野に入れ、1年次の春学期までとした。分析する教学データは、客観的に評価できる数値とし、GPA、授業への出席率、プレースメントテストの点数、演習への取り組み時間となっている。GPAは、春学期に評価されたデータを用いた。授業への出席率は、選択科目数による誤差を考慮し、必修科目の授業への平均出席率とした。演習への取り組み時間は、入学前教育での学習で蓄積された演習への取り組み時間とした。プレースメントテストの点数は、入学前教育の最後に実施されるテストの点数としている。A大学は理系の大学であるため、演習の内容とプレースメントテストは、理数系の内容となっている。これらの教学データを作成したシステムを用いて分析した。その結果を図1に示す。なお、分析する学生のデータは、個人を特定できないよう学籍番号を隠蔽して行っている。

図1は、6×6の円が参照ベクトルを表している。その中にある数値は、学生を表している。学生は、類似度の高い参照ベクトルごとに6つのグループに分けられている。このグループと退学者の分布を分析に利用した学生の教学データと照らし合わせて退学者の傾向を把握できるか調査した。

図1に示すように、退学者はグループA、Bのそれぞれに分けられた。そして、それぞれのグループで退学者の傾向がわかった。グループAでは、授業への出席率のみ全体平均より高い、または、演習への取り組み時間のみ全体平均より高く、他の教学データの数値が全て平均以下の学生が退学傾向にあることがわかった。グループBでは、プレースメントテストの点数のみ全体の上位レベルで、他の教学デー

タが全体平均より低いといったある学生が退学傾向にあることがわかった。

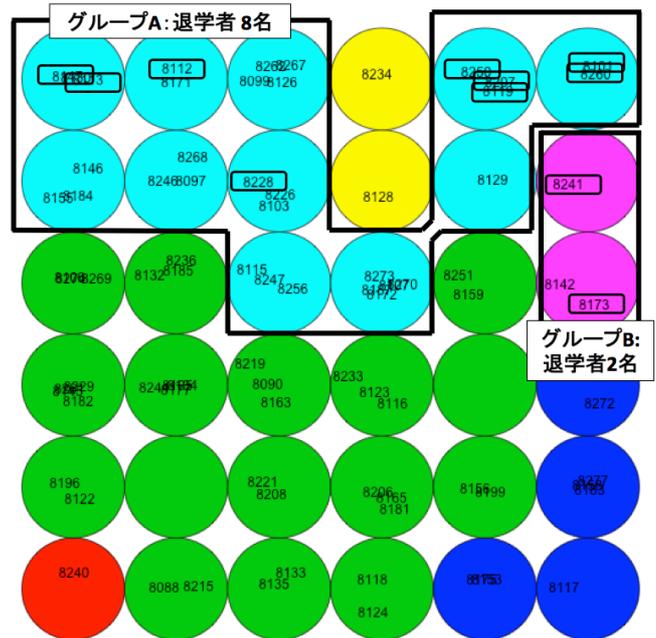


図1. 入学前教育受講者のデータ分析結果

5. まとめ

本研究の目的は、ICT教育支援システムに蓄積された学生の教学データを活用して学生一人ひとりの教学に関する傾向を把握できる教学IR支援システムの確立にある。今回は、目的に向けてSOMによる分析方法を提案し、それを適用したシステムを試作した。作成した分析システムで学生の教学データを分析した結果、グループA、Bの2つに退学者が分けられた。そして、それぞれのグループで退学者の傾向が見られた。

6. 今後の取り組み

今後は、退学者と同じグループに存在する在学学生、または、退学者の教学データと類似度の高い在学学生の傾向を調査し、在学学生と退学者の傾向の違いを把握していく。これにより、退学を未然に防ぐ教職員の指導・支援に役立つ情報が抽出できると考えている。また、授業の欠席間隔や演習への取り組み間隔などの数値の分析や「キャリア関連」、「その他」の教学データを組み合わせた分析、年度ごとの退学者傾向の比較などを行っていき、よりきめ細かな退学者の傾向も調査していく。

参考文献

- [1] 山田礼子：“学生の特性を把握する間接評価：教学IRの有用性” <Journal of JSEE 61(3), 3_27-3_32, 2013年>