

## 出欠データと学習データを用いた学生の修学傾向分析

伊藤 宏隆 堀江 匠 舟橋 健司 内匠 逸 松尾 啓志

名古屋工業大学 情報基盤センター

## 1. はじめに

名古屋工業大学では、2007年4月にICカード出欠管理システムとコースマネジメントシステムを導入した[1, 2]. ICカード出欠管理システムでは、ICカード化された学生証により学生の出欠状況を把握する。コースマネジメントシステム(CMS)は、e-Learningを支援するシステムであり、教材の作成支援、課題の提出管理、小テストの実施、学生の受講管理を行う機能を持っている。本学では、ICカード出欠管理システムで収集された出欠データやCMSに蓄積された学習データを分析し、学生の修学指導支援に役立てることを考えている。例えば、過去のデータを分析することにより、学生の成長可能性の予測や落第候補者を発見し早期に指導することが可能となる。著者らは、これまでにある授業の出欠データとCMSでの課題提出の有無のデータと成績データの3つを用いて分析を行い、結果の報告を行っている[3]。それ以降、さらに課題提出時間のデータを加え分析を行っている。さらに、3つの授業についても同様の分析を行った。すなわち、計4つの授業について、修学傾向分析を行い、より信頼性の高い結果を得られている。本論文では、それらの分析結果について報告する。

## 2. 出欠、学習データを用いた修学傾向分析

著者らはこれまでに授業Aの出欠データ、課題の提出の有無のデータ、成績データを用いて分析を行っている[3]。授業Aは講義科目であるが、授業時間外にCMSでの課題提出を義務付けている。出欠データは15回分、課題のデータは14回分存在する。成績評価はS(秀), A(優), B(良), C(可), D(不可), X(失格)の6段階である。分析ツールとして、Microsoftのデータマイニングアドインを用い、主要な影響元の分析、ルールの抽出やカテゴリの抽出を行った。分析

の結果は以下のとおりであった。

- 上位成績者(S, A, B)は全体的に出席率が高い。
- 何名かの成績Bの学生は遅刻、欠席が多い。
- 下位成績者(C, D, X)は全体的に出席率が低いわけではなく、前半から徐々に出席率が下がっている。ただし、C, Dの学生とXの学生との違いは下位の学生の出席率が中盤から後半にかけてまた高くなるのに対し、最下位の学生は後半には完全に出席しなくなっている。
- 成績Aの学生は終盤の出席の影響が大きく、成績Bは中盤の出席の影響が強い。
- 課題を提出していなければ下位の成績(D, X)になりやすい。逆に課題を提出していても、上位の成績になるとは限らない。
- 下位成績者(C, D, 一部のX)は、徐々に課題の提出率が悪くなる。また、それ以外の成績者Xは常に課題提出率が悪い。

著者らは今回、課題提出時間のデータも用いて、成績との関連性を主要な影響元の分析を行い調査した。図1に結果を示す。

列	値	優先	影響値
提出値	$\geq 0.8135$	S	86
提出値	$0.4989 - 0.8135$	S	34
提出値	$0.3239 - 0.4989$	S	23
提出値	$0.0880 - 0.3239$	A	2
提出値	$0.0880 - 0.3239$	B	24
提出値	$< 0.0880$	C	100
提出値	$< 0.0880$	D	100
提出値	$< 0.0880$	X	100

図1. 課題提出時間と成績との関連

図1で提出値とは、課題をどれだけ早く提出しているかを表し、提出期間開始時を1とし、提出期間終了時を0としている。1に近いほど早く、0に近いほど提出期限ギリギリに提出していることになる。図1より、課題提出時間が早いほど、成績が良いことがわかる。

著者らは、さらに、授業B, C, Dの3つについても授業Aと同様の分析を行った。それぞれの授業の分析結果について述べる。

まず、授業Bであるが、授業Bも講義科目であり、出欠データは15回分、課題提出データは14回分存在している。分析結果は以下のとおり

Analysis of student's schooling trend using attendance data and learning data.

Hirota Itoh, Takumi Horie, Kenji Funahashi, Ichi Takumi and Hiroshi Matsuo

Information Technology Center, Nagoya Institute of Technology

である。

- ・上位成績者(S, A, B)は全体的に出席率が高い。
- ・下位成績者(C, D, 一部のB)は、徐々に欠席や遅刻が多くなる。
- ・最下位成績者(X)は、徐々に欠席率が高くなり終盤もあまり出席しない。
- ・成績が良いと出席の影響が大きく、悪いと欠席及び遅刻の影響が大きい。
- ・出席が続くと欠席しにくく、逆に欠席が続くと出席しにくい。後者の方が影響が大きく、終盤になるにつれ、その影響が大きくなる。
- ・課題を提出しなければ成績D, Xになりやすい。
- ・成績Aは、中盤～終盤(08～14回目)の提出の影響が大きく、Bは終盤(11～14回目)の影響が大きい。Cは中盤(07, 08, 10回目)の影響が大きい。
- ・下位成績者(C, D, 一部のX)は、徐々に課題の提出率が悪くなる。また、それ以外の成績Xの学生は常に課題提出率が悪い。
- ・課題を早く提出するほど成績が良い。

次に、授業Cの分析結果である。授業Cも講義科目であり、出欠データは15回分、課題提出データは3回分存在している。分析結果は以下のとおりである。

- ・出欠状況には、連続性が見られる。すなわち、出席が続くと欠席しにくく、欠席が続くと出席しにくくなる。
- ・成績が良いと出席の影響が大きく、悪いと欠席及び遅刻の影響が大きい。
- ・上位の成績者(S, A, B)は出席率が高く、特にSは高い。
- ・下位成績者(D, X)は徐々に出席率が低くなる。成績Xの大多数はDに比べ出席率が低く、中盤以降はほとんど出席していない。
- ・終盤の授業を欠席するとXになりやすい。
- ・一部の成績C, Dの者は中盤までほとんど出席せず、終盤のみ出席している。
- ・成績は中盤～終盤にかけての出欠の影響が大きい。特に中盤の影響が強い。
- ・課題を提出していなければ下位成績(D, X)になりやすい。特にXは3回目の提出率が低い。
- ・3回すべて課題を提出していれば、上位成績(S, A, B)になりやすい。
- ・課題を早く提出するほど成績が良い。

次に授業Dの分析結果である。授業Dは演習中心の科目である。講義と演習の繰り返しで、奇数回目は講義、偶数回目が演習となっている。15回目は試験である。出欠データは講義と試験時の8回分、課題提出データは演習時に提出する7回分が存在している。分析結果は以下のと

おりである。

- ・他の授業と比べて出席率が高く、成績も良い。
- ・成績Xは全体的に出席率が低い。
- ・授業はバラバラに出席するよりも連続して出席する方が成績は良くなる。
- ・成績Dは課題提出の影響が大きいが成績Xは未提出の影響が強い。すなわち、DとXの違いは課題を提出しているか否かである。
- ・どの成績でも課題提出は中盤～終盤にかけての影響が強い。
- ・成績X以外はほぼ課題を提出している。X以外の未提出時期は中盤～終盤にかけてである。

### 3. まとめ

ICカード出欠データ、CMSの課題提出データ、成績データを用いて学生の修学傾向分析を行った結果について述べた。講義科目では、ほぼ同様の傾向が見られた。特に、出欠は連続性があり、成績C, Dは中盤に欠席が連続し、成績Xは中盤以降出席しなくなる。また、課題を提出しないと下位の成績になりやすく、課題の提出が早いほど、成績が良い。演習科目では、出席率が高く、課題の提出率が良かった。

今後は、過去のデータをもとに、学期途中での成績予測を行うことを考えている。早期の成績予測により、下位成績候補者や失格候補者に警告を行うことで、落第者の減少が見込める。

### 謝辞

本研究は文部科学省特別研究経費『充実した「学びの場」の構築—教員の教育力向上および双方向型教育支援システムの整備』の支援により行われた。ここに感謝の意を表します。

### 参考文献

- [1] 伊藤宏隆, 舟橋健司, 中野智文, 内匠逸, 松尾啓志, 大貫徹, “名古屋工業大学におけるMoodleの構築と運用”, メディア教育研究, Vol. 4, No. 2, pp. 15-21 (2008)
- [2] 伊藤宏隆, 舟橋健司, 中野智文, 松尾啓志, 内匠逸, 大貫徹, “コースマネジメントシステム Moodle と IC カード出欠システムとの連携”, 平成 19 年度電気系学会東海支部連合大会講演論文集, 0-376 (2007)
- [3] 伊藤宏隆, 舟橋健司, 中野智文, 内匠逸, 松尾啓志, “IC カード出欠データと CMS 学習データを用いたデータマイニング”, 日本 e-learning 学会 2008 年秋季学術講演会論文集 pp. 41-47 (2008)