

出欠データと学習データを用いた成績予測と その教育効果の検証

伊藤宏隆[†] 舟橋健司[†] 山本大介[†]
内匠逸[†] 松尾啓志[†]

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。著者らはこれらのシステムに蓄積されたデータを学生指導に有効活用することを計画している。今回、ある授業において出欠データとCMS学習データを用いて成績予測を行い、結果を基に学生のレベル別に学習指導を行った。本文では、成績予測とその教育効果の検証について論述する。

Grade Prediction using Attendance and Study Data and Verification of Educational Effect

Hiroataka Itoh[†] Kenji Funahashi[†] Daisuke Yamamoto[†]
Ichi Takumi[†] Hiroshi Matsuo[†]

At Nagoya Institute of Technology, smart card roll call system and course management system were introduced in April, 2007. Authors are planning to use the data accumulated by these systems for the student guidance effectively. This time, we executed the grade prediction using the attendance data and the CMS study data in a certain class. We performed study guidance based on the prediction result according to student's level. In this paper, the prediction result and the verification of the education effect are stated.

1. はじめに

名古屋工業大学（以下、名工大と略す）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（以下、CMSと略す）[1]を導入した。ICカード出欠管理システムは、学生の学生証をICカード化し、教室に設置されたICカードリーダーに学生証を近づけることによって、学生の出席を記録し、出欠を管理するシステムである。学生には、ICカードリーダーの設置されたすべての教室において出席打刻を義務付けている。CMSとして、名工大では、オープンソースのMoodle [2]を導入している。導入以来、Moodleは授業のサポートシステムとして、主に電子的資料の配布、課題の提出管理、小テストや授業アンケートの実施、授業のスケジュールの連絡等に用いられている。

システム導入後、3年が経過し大量の出欠データやCMSでの学習データが蓄積されている。著者らはそれらのデータを成績評価のデータとしてだけではなく、有効活用し、教育指導にも役立てることを考えた。例えば、データマイニングにより過去の学生のデータの傾向を分析することで、学生の修学傾向を把握ができる。また、過去のデータをもとに学生の将来の成長可能性を予測することで、落第候補者を早期に発見することが可能となる。データマイニングや予測により、早期にきめ細かな指導を行うことが可能となる。

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、課題提出時間のデータと成績データを用いてデータマイニングを行い、各データ間の関係から学生の修学傾向について有用な結果を得ている[3][4]。また、早期に成績を予測し、落第候補者に対して学習指導を行うことにより落第者の減少を図ることを将来的な目的として、2つの成績予測実験を行っている。1つ目の実験として、ある授業のデータにおいて、学期途中までの出欠データや学習データから成績を予測する実験を行っている[5]。2つ目の実験として、ある学科の2年生の半期における単位修得数を3年生の過去のデータを用いて予測した実験を行っている[6]。

今回、著者らは1年生の情報基礎科目の授業において15回の授業のうち7回目が終了した時点で出欠状況や課題の提出状況から学生の成績を予測し、予測された成績のレベル別に課題を課した。レベル別課題の遂行により学生の理解度や学習進度に合わせた学習指導ができることを想定している。予測成績の低い学生はそれ以上の成績にアップすることを、予測結果の良い学生は最終的に最高成績Sとなることを期待した。レベル別課題締め切り後に、もう一度成績予測を行い、各学生の予測成績の推移を検証した。また、最終的な成績から本実験の有効性を検証した。

[†] 名古屋工業大学
Nagoya Institute of Technology

本論文では、学期途中での成績予測と、予測結果を学習指導に適用した場合の教育効果について論述する。

2. IC カード出欠管理システムとコースマネージメントシステム Moodle

2.1 IC カード出欠管理システム

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を使用して学生の授業への出欠を管理するシステムである。教職員証および学生証は接触型と非接触型の2種類のICチップを搭載している。ICカードとしてFelicaを採用した。ICカード出欠管理システムでは、非接触型ICチップを利用し、打刻を行う。すべての講義室(実験室を除く)の出入口には非接触型ICカードリーダーが設置されており、ICカードを近づけると、リーダーがカード情報を読み取り、打刻時間がサーバへ送られ蓄積される。学生証の代わりに事前に登録した携帯電話により打刻することも可能である。学生にはすべての授業において、受講時に授業の開始前と終了後の2回打刻することを義務付けている。出欠管理サーバにはすべての授業の授業情報(授業コード、授業時間、講義室、授業名、担当教員、受講生など)、教職員および学生情報(カード番号、職員番号または学籍番号、氏名など)やICカードリーダー情報(設置場所など)の情報が登録され、管理されている。講義室に設置されているICカードリーダーのインターフェース画面には、その時間に行われる授業の授業名と担当教員名が表示される。出欠管理サーバに登録されている情報と打刻情報をもとに学生の出欠情報が自動的に生成される。

出欠管理システムにはWebインターフェースが提供されており、Web上で、学生は自身の受講している授業での出欠状況を、教員は担当している授業の出欠状況を確認することができる。さらに、教員はWeb上で講義室変更や出席データの修正、学生へのメッセージを登録できる。登録されたメッセージは学生が打刻した時にカードリーダーの画面上に表示される。また、急な講義室変更の際には、カードリーダーの画面上(タッチパネル)で情報を変更することも可能である。

2.2 コースマネージメントシステム Moodle

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、CMSとしてMoodleを採用している。Moodleはオープンソースのソフトウェアで全世界で広く利用されている。Moodleのバージョンは1.9.5+である。名工大では、学部(一部、二部)、大学院のすべての授業はMoodleにコース登録されている。2007年度~2010年度で約8500の授業がコース登録されている。教員は申請なしで自身の担当してい

る授業のコースを利用することができる。Moodleに登録されているユーザは現在約11000ユーザである。MoodleはICカード出欠管理システムと連携し、出欠管理システムの出欠データを取り込むことができる。

3. 出欠データと学習データを用いた成績予測とその教育効果の検証

3.1 対象授業の概要

今回、成績予測とその教育効果の検証を行った授業は平成21年度後期の1年生の情報基礎科目の授業である。内容は、簡単なWebページの作成とJavascript, Excelであり、受講生は38名であった。最終的な成績評価は、出席状況、課題の提出状況と授業の終盤で課される最終課題2題の採点結果で判定している。出席回数、課題の提出数が一定以上ない場合は失格となる。成績予測を行うために前年度である平成20年度と同じ授業の受講生のデータを予測のためのサンプルデータとして用いた。前年度の受講生の数は64名であり、成績分布を表1に示す。成績のSは秀、Aは優、Bは良、Cは可、Dは不可、Xは失格である。S~Cが合格、D、Xが不合格である。

表1 平成20年度の受講生の成績分布

成績	S	A	B	C	D	X	計
人数(人)	18	16	13	9	0	8	64
比率(%)	28.1	25.0	20.3	14.1	0	12.5	100.0

3.2 成績予測手法

出欠データとCMSの学習データを用いて行った成績予測について説明する。成績予測に用いたCMSの学習データは課題の提出数と課題の提出値である。課題の提出値とは課題提出の早さを表し、課題提出受付開始時を1とし、課題提出締め切り時を0としている。課題提出が早いほど提出値は高くなる。ただし、未提出は-1とした。以前のデータマイニング[3][4]の結果によって、提出値が高いほど、成績が良くなる傾向が得られているので、予測のための一つの指標として有効であると考えられる。成績予測は成績が既知である前年度と同じ授業の受講生のデータをサンプルデータとして用いている。今回予測に用いた計算機環境は下記のとおりである。

- (1) ハードウェア
 - CPU Xeon クワッドコア 2CPU (8コア)
 - メモリ 8GB
 - ディスク容量 136GB
- (2) ソフトウェア

- OS Microsoft Windows 2003 Server
- データベース Microsoft SQL Server 2005
- マイニングツール Microsoft Office Excel 2007 + Microsoft SQL Server 2005 Data Mining Add-ins for Office system[7]

上記のようにツールとして、Microsoft のデータマイニングアドイン[7]を使用した。図1のように、予測のサンプルデータとして用いる成績の入った平成20年度の受講生のデータと成績欄を予測のために空白にした平成21年度の受講生のデータをつなげたテーブルを用意する。データマイニングツールのメニューの「例の全体適用」を成績に対して行くと、成績以外の列の各項目が成績にどのように影響を与えているかを分析し、それを基に推論し空欄の成績を埋めてくれる[8]。

	番号	出席点	課題提出数	課題提出値計	成績
平成二十年 度	23987	10 8		7.326	S
	⋮	⋮	⋮	⋮	⋮
	67159	10 8		5.741	A
平成二十一年 度	72135	3 2		-3.657	
	⋮	⋮	⋮	⋮	
	34784	9 7		1.342	

図1 例の全体適用による成績予測

3.3 成績予測を基にした学習指導の効果の検証

授業の中間、第7回終了時点までの出席回数、課題提出数と課題提出値の合計値を用いて成績予測を行った。予測成績の分布を表2に示す。

表2 平成21年度第7回終了時点での予測成績の分布

成績	S	A	B	C	D	X	計
人数(人)	15	6	3	7	0	7	38
比率(%)	39.5	15.8	7.9	18.4	0.0	18.4	100.0

表2より、この時点では、約4割が最高成績Sと予測され、全体の約8割が合格圏内、2割弱の受講生が不合格圏内と予測された。この時点では授業時に出題したいくつかのレポート課題は締め切り前であり、これらのレポートが未提出であった受講生のうち何名かがXと予測されたようである。

予測された成績をもとに受講生を3つのグループにグループ分けした。予測成績がSとAの受講生をグループ1、予測成績がBとCの受講生をグループ2、予測成績がXの受講生をグループ3とし、グループ1には難易度高の課題、グループ2には難易度中の課題、グループ3には難易度低の課題を出題した。学生には、“現在の進捗、理解度から応用力の調査として難易度の違う課題を課します”と伝えた。

次に、レベル別課題の提出締め切り後の第11回授業終了後にもう一度、成績予測を行った。予測成績の分布を表3に示す。表3より、レベル別課題実施後、Sの人数が増え、逆にXの人数が減っていることがわかる。第7回から第11回にかけての予測成績の推移を表4に示す。表4より、第7回終了時の予測がSだった受講生はそのままSをキープし、さらに、Sではなかった7名がSにランクアップしている。また、Cだった7名とXだった4名の計11名がBにアップしている。不合格のXだった7名のうち5名も合格圏内のBにアップした。38名中Aだった2名のみがBに1段階ダウンした。レベル別課題実施後、36名が予測成績をアップさせたかキープしている。

表3 平成21年度第11回終了時点での予測成績の分布

成績	S	A	B	C	D	X	計
人数(人)	22	0	14	0	0	2	38
比率(%)	57.9	0.0	36.8	0.0	0.0	5.3	100.0

表4 平成21年度予測成績推移

単位：人

予測成績		第11回終了時					
		S	A	B	C	D	X
第七回 終了 時	S	15	0	0	0	0	0
	A	4	0	2	0	0	0
	B	2	0	1	0	0	0
	C	0	0	7	0	0	0
	D	0	0	0	0	0	0
	X	1	0	4	0	0	2

次に、全授業終了時の最終成績の分布を表5に示す。半数がSとなった。Xは3名であった。Xの3名のうち2名は履修登録はしたが、授業の序盤からずっと出席していなかった。また、残りの1名は7回まではほぼ出席していたものの、課題の提出が3回しかなく、8回以降はずっと欠席となった受講生であった。

レベル別課題実施前の第7回終了時の予測成績と最終成績との関係を表6に示す。

表5 平成21年度最終成績分布

成績	S	A	B	C	D	X	計
人数(人)	19	9	2	5	0	3	38
比率(%)	50.0	23.7	5.3	13.2	0.0	7.9	100.0

表6 平成21年度第7回終了時の予測成績と最終成績
 単位：人

成績	最終成績						
	S	A	B	C	D	X	
第七回終了時予測成績	S	12	0	2	1	0	0
	A	3	1	0	0	0	1
	B	1	1	0	1	0	0
	C	0	3	0	1	3	0
	D	0	0	0	0	0	0
	X	3	2	0	0	0	2

表6より、Sだった15名のうち、12名はそのままSとなった。Sでなかった7名がSにランクアップした。うち3名はXからの5段階アップである。Xからは2名もAに4段階アップしている。他1段階アップしたのは5名、2段階アップしたのは4名であった。計14名が予測成績より高い成績となった。一方、ランクダウンした受講生は全体で6名であった。途中でドロップアウトした1名を除けば、1段階ダウンが2名、2段階ダウンが1名、3段階ダウンが1名となった。

次に、レベル別課題実施後の第11回終了時の予測成績と最終成績との関係を表7に示す。

第11回終了時点では、まだ4回の授業が残っており、成績評価の大きな比重を占める最終課題も実施されていない。しかしながら、成績と予測が一致したのは半数の19名、1ランクの違いは13名、2ランクの違いは4名であり、ほぼ予測通りであった。

表1と表5より、平成20年度と平成21年度の成績を比較すると、平成21年度では、Sが大幅に増加し、Xが減少した。

上述の結果より全体を通して、最終的に大半の受講生が予測通りか予測よりも良い成績となり、成績予測結果は直接伝えていないが、予測を基にしたレベル別課題の実施は受講生の学習姿勢の改善を図る指導の効果があったと考えられる。

表7 平成21年度第11回終了時の予測成績と最終成績
 単位：人

成績	最終成績						
	S	A	B	C	D	X	
第十一回終了時予測成績	S	16	4	1	1	0	0
	A	0	0	0	0	0	0
	B	3	5	1	4	0	1
	C	0	0	0	0	0	0
	D	0	0	0	0	0	0
	X	0	0	0	0	0	2

4. まとめ

出欠管理システムとCMSに蓄積されたデータを有効に活用し、学生の学習指導に役立てることを目的として、今回、1年生の情報基礎科目の授業において授業の中盤で出欠データやCMS学習データを用いて受講生の成績を予測し、予測された成績のレベル別に課題を課した。レベル別課題実施前と実施後で、予測成績を比較したところ、Sだった受講生のほぼ多数はそのままSを維持し、Sではなかった受講生の多数もランクアップしていた。また、レベル別課題実施前の予測成績と最終的な成績を見ると、多数の受講生で予測成績からランクアップしており、レベル別課題の効果がみられた。レベル別課題実施後の予測成績と最終的な成績を比較すると、多数の受講生の成績が予測と完全一致かほぼ一致しており、予測の有効性も確認できた。

前年度の成績との比較でも、Sの受講生は大幅に増加し、Xの受講生の割合も減少しており、成績予測を基にしたレベル別指導は効果があったといえる。

今後は、さらなるデータマイニングや予測結果を学生に直接的に伝え、現在の自身の学習理解度や他の学生との相対的な学力を示すことで学生の学習意欲の向上を図ることを計画している。

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) Moodle.org : open course community-based tools for learning
<http://moodle.org/>
- 3) 伊藤宏隆, 堀江匠, 舟橋健司, 内匠逸, 松尾啓志: 出欠データと学習データを用いた学生の修学傾向分析, 情報処理学会第 71 回全国大会講演論文集第 4 分冊, pp.357-358 (2009)
- 4) 伊藤宏隆, 舟橋健司, 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 学会誌, Vol.9, pp.95-108 (2009)
- 5) 伊藤宏隆, 舟橋健司, 山本大介, 内匠逸, 松尾啓志: 出欠データと学習データを用いた学生の成績予測, 日本 e-Learning 学会 2009 年度秋季学術講演会論文集, pp.52-57 (2009)
- 6) 伊藤宏隆, 舟橋健司, 山本大介, 内匠逸, 松尾啓志: 学生の修学データを用いた修学指導支援システムの検討, 情報処理学会第 72 回全国大会講演論文集第 4 分冊, pp.473-474 (2010)
- 7) Microsoft SQL Server 2005 Data Mining Add-ins for Office system
<http://www.microsoft.com/japan/sql/dmaddin/default.aspx>
- 8) 平井明夫, 松井浩輔: ひと目でわかる Microsoft Office Excel 2007 データマイニングアドインを使用したビジネスデータ簡単分析術”, 日系 BP ソフトプレス (2007)