

特集：多様な端末と大規模学習データが拓く新たな学習支援環境

学士課程における大規模データに基づく 学修状態のモデル化

近藤 伸彦*, 畠中 利治**

Modelling of Students' Learning States Using Big Data of Students through the Baccalaureate Degree Program

Nobuhiko KONDO*, Toshiharu HATANAKA**

Institutional Research (IR) has been receiving much attention in Japanese higher education. In order to guarantee the educational quality of university, it has been discussed how to utilize the educational big data. In this paper, it is considered to construct models of students' learning states using large-scale students' learning data collected through the baccalaureate degree program based on some machine learning methods. In this research, data in 5 years are utilized in order to investigate the generalization ability of the models, and the performances of some machine learning methods are compared. From the experimental results, it is indicated that the models of students' learning states with high generalization ability can be constructed. Its capability of application to enrollment management is also discussed from experimental results.

キーワード：大規模学修データ，教育ビッグデータ，教学 IR，Learning Analytics，機械学習，学修支援

1. はじめに

さまざまな分野において、ビッグデータの活用が進められている。それは教育においても例外ではなく、教育ビッグデータという語が一般化しつつある。教育にまつわる横断的研究分野として、Learning Analytics や Educational Data Mining などが年々発展を見せている⁽¹⁾ ことも、こうした動きに呼応しているといえる。

一方、「大学全入時代」とも呼ばれるとおり、日本の高等教育はすでにユニバーサル化した状態にある。この現状において、学力や学習意欲などあらゆる面で学生は多様化しているが、その一方でこれまで以上の高等教育の質保証が求められ、学士としての質を担保するための教育改革が喫緊の課題となっている。

こうした背景のもと、データをもとに教学改善を支

援する機能として教学 IR (Institutional Research) の必要性が認識され、あらゆる高等教育機関において、組織、マネジメントから技術面に至るまで、さまざまな切り口で教学 IR の実質化が検討されている。教学 IR の重要な役割の一つとして、多様な学生の個に応じた適時的な学修支援を行うために、学士課程における学生の学修状況を把握し分析することが挙げられる。これを実現するためには、学内外に蓄積される教育ビッグデータを効果的に活用するという視点が欠かせないであろう。ここでの教育ビッグデータとは、LMS などに蓄積される学習履歴ログデータだけでなく、デモグラフィックデータや、入試、成績、出席状況、履修状況など、学修に関わるあらゆるデータを指す。

本稿では、学士課程を通して蓄積される大規模な学修データから、学生の学士課程にわたる学修状態を数

*大手前大学 CELL 教育研究所 (CELL Institute for Educational Development, Otemae University)

**大阪大学大学院情報科学研究科 (Graduate School of Information Science and Technology, Osaka University)

受付日：2015 年 6 月 19 日；再受付日：2015 年 9 月 20 日；採録日：2015 年 11 月 4 日

理的にモデル化し、これを学修支援へ活用することについて検討する。第2章では、研究背景として、教学IRと関連分野の動向について簡単に述べる。第3章では、学士課程において蓄積されるさまざまなデータを時系列的に整理した学修ライフログを用いた学修状態のモデル化について述べる。第4章では、機械学習による学習状態のモデル化について、ある大学（以後A大学と記す）に関する数値実験の結果を示したうえで、学修状態モデルの学修支援への活用について検討する。最後に本稿のまとめと今後の課題について述べる。

2. 研究背景

2.1 教学IRとエンrollment・マネジメント

ユニバーサル化段階にある近年の日本の高等教育においては、教学面・経営面での合理的な意思決定の必要性から、教育や経営の改善に資するデータの収集・分析などをもとに意思決定支援を行うIRが注目を集めている。IRには、経営改善、教育改善、認証評価・情報提供などさまざまな視点がある⁽²⁾が、日本においては、このうち教育改善に関する機能を指す「教学IR」が独自に発達している⁽³⁾。

教学IRにおける重要な視点の一つにエンrollment・マネジメントがある。これは、「戦略的な計画によって構造化され、IRによって支援された学生の大学選択、大学間移動、中途退学、在籍、アウトカムに係る活動」であるとされる⁽⁴⁾。多様な学生に対する教育の質保証や、健全な大学経営を行ううえで、こうした活動は必要不可欠である。入学前から学士課程を通してあらゆる学生データを紐づけ、高い学修成果につながる要因の分析や、逆に学習困難に陥っている学生の早期発見など、学修支援に結びつく知識発見を行うことがエンrollment・マネジメントにおいて重要である。

2.2 Learning AnalyticsやEducational Data Miningによる学修状態予測

この数年、大規模なデータを分析・活用して教育の改善を目指すLearning AnalyticsやEducational Data Miningと呼ばれる分野が急速な発展を遂げている⁽¹⁾。

その中には、ドロップアウトする、あるいは逆に高い学習成果を上げる学生のパターンを機械学習により発見する研究が散見される。

LMSやeラーニングのログデータを用いた例としては、次のようなものがある。たとえばJayaprakashらは、デモグラフィック属性とLMSのログデータを活用し、授業の離脱リスクのある学生（At-risk students）を早期発見して適切な介入を行うシステムを構築している⁽⁵⁾。Vihavainenらは、あるオンラインコースにおいて、学習ログから適切に特徴量を抽出し、機械学習によって数理モデルを構築することで、途中でドロップアウトする学生を全15回のうち最初の5回の学習ログから予測できることを示している⁽⁶⁾。

このほか、学士課程におけるデータから学修状態を予測する試みも行われている。雨森らは、入学前後から1年次の修学状況データにより3年次前期までの単位取得状況を推定する決定木分析を行っている⁽⁷⁾。大友らはサポートベクターマシンにより退学予備軍の予測を試みている⁽⁸⁾。Djulovicらは、決定木やニューラルネットワークなどの機械学習手法により、初年次の学生データから1年後の在学状況の予測を行っている⁽⁹⁾。

3. 学修状態のモデル化

3.1 学修ライフログ

本研究では、特定のeラーニングコースのような限定された学修状況ではなく、学士課程全体における総合的な学修状態を対象とする。

大学には、入学前から学士課程を通してさまざまなタイプのデータが存在する。本研究では、こうした学内に散在する大規模なデータを一元的に集約し、時系列に整理した学修ライフログを考える。学修ライフログのイメージを図1に示す。こうした時系列順のデータセットが各学生について存在し、学士課程が進行するに従ってライフログが蓄積されていくイメージである。このように整理しておくことで、学士課程全体にわたる学修状態やその変化を把握しやすくなる。こうした大規模な学修ライフログからいかに意味のある情報を抽出するかが問題となる。効果のある教育プログ

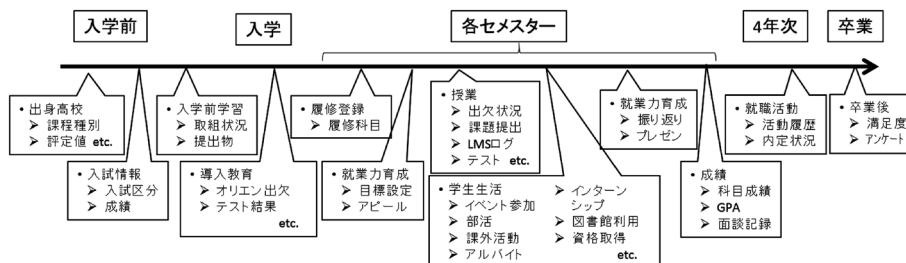


図1 学修ライフログのイメージ

ラムや、高い成果を上げる学生の成長フロー、支援すべき学生などを発見するにあたり、その非線形で複雑な構造のモデルをデータから推定しこれを活用していくためには、データマイニングや機械学習の手法が有効であると考えられ、本研究においてはそうした手法の活用を検討する。先行事例は第2章にて述べたとおりである。

3.2 対象とするデータと問題設定

本稿では、A大学における2009～2013年度入学生のデータを対象とする。編入学や再入学の学生は除き、1年次に入学した学生のみに着目する。

A大学は、リベラルアーツによる教養教育を使命とし、3学部からなる中小規模の文系私立大学である。2011年度にカリキュラム改編が行われたり、導入教育や初年次教育の内容が年々見直されていたりするなど、入学年度によって教育内容は少なからず異なり、学生にとっては学修の経験や機会が入学年度ごとに少なからず変化している。数値実験においては、こうした大学としての教育環境の変遷を考慮したうえで結果を検討する。

本稿では、図1に示したようなデータのうち、学士課程の初期段階と、経年後の学修状態について着目し、問題設定として、学士課程の初期段階における学修状態から、その後の学修状態を予測することを考える。

3.3 特徴量の性質

本稿で用いた特徴量は表1の10項目である。本稿では、 $X_1 \sim X_{10}$ を変数名として統一的に表現する。以下、いくつかの留意点について述べる。

X_2 (学部) は、3学部のいずれに所属するかを示す。

X_3 (入試種別) は、“AO入試”、“公募制推薦入試”、

“指定校推薦入試”、“一般入試”、“センター試験利用入試”、“外国人留学生入試”、“社会人入試”の7区分のいずれの入試により入学したかを示す。

X_4 (入学前学習課題提出度) は、入学前に課した学習課題の提出率から、“高提出率”、“低提出率”の2クラスのいずれかを割り当てたものである。2010年度までと2011年度以降で実施内容が大きく異なっているが、いずれの年度においても、提出率が70%以上の場合は“高提出率”、70%未満の場合は“低提出率”としている。この閾値の設定は、提出率の分布を見ながら経験的に行った。

X_5 (新入生オリエンテーション出席度) は、入学直後の1週間程度のオリエンテーション期間における出席状況である。ただし、オリエンテーションのプログラムは年度ごとに少しずつ変化している。これを踏まえたうえで、“すべてのプログラムに遅刻・早退なく出席”、“一つのプログラムに欠席”、“それ以上の欠席”の3クラスに振り分けた。なお、遅刻・早退は1/3回の欠席とカウントした。ただし、2013年度については、プログラムがそれ以前より長期化したこともあり、二つ目のクラスを“欠席したプログラムが2未満”とした。こうした「出席度」の設定は、欠席回数の分布を見ながら経験的に行った。

$X_6 \sim X_8$ は、1年次春学期必修科目の出席率である。A大学は1年2学期制をとり、春学期・秋学期それぞれ15週の授業期間をもつ。ここでは、春学期の15週を5週ごとに分け、第1～5週の出席率を X_6 、第6～10週の出席率を X_7 、第11～15週の出席率を X_8 とした。2010年度までは4科目、2011年度からは3科目と、科目数に変動があることに注意しておく。これらの出席率は、すべての必修科目の出席回数をすべての必修科目の授業回数で除したものである。

表 1 特徴量

特徴量	値
X_1 : 性別	名義尺度, 2 クラス
X_2 : 学部	名義尺度, 3 クラス
X_3 : 入試種別	名義尺度, 7 クラス
X_4 : 入学前学習課題提出度	順序尺度, 2 クラス
X_5 : 新入生オリエンテーション出席度	順序尺度, 3 クラス
X_6 : 1 年春学期 1~5 週必修出席率	連続値, [0, 1]
X_7 : 1 年春学期 6~10 週必修出席率	連続値, [0, 1]
X_8 : 1 年春学期 11~15 週必修出席率	連続値, [0, 1]
X_9 : 1 年次春学期 GPA	連続値, [0, 4]
X_{10} : 3 年次 4 月在籍状況	名義尺度, 2 クラス

なお、遅刻・早退は 1/3 回の欠席として計算している。

X_9 は 1 年次春学期の GPA (Grade Point Average) である。A 大学の成績評語は「A」「B」「C」「D」「F」の 5 段階であり、「A」~「D」の場合に単位が与えられる。これらの評語それぞれに対し、GP (Grade Point) を、「4」「3」「2」「1」「0」に対応させる。履修を学期途中で放棄した場合も「F」の成績が与えられ、GP は「0」となる。GPA は、対象とする学習期間におけるすべての履修科目について、単位数を重みとする GP の重み付き平均で定義される。 X_9 は、1 年次春学期の履修科目に対する GPA である。A 大学はリベラルアーツ教育のもと、とくに学士課程初期には幅広い分野の履修を奨励していることから、GPA の学部間格差は少ないものと考えられる。科目ごとの成績評価基準の違いが成績分布の科目間格差を生み出している可能性はあるが、GPA は複数の科目を平均しているためにある程度その違いが吸収され、本研究における学業成績の指標としては一定の妥当性があると考えられる。

X_{10} は、入学から 2 年が経過した時点における在籍状況を示し、在学している場合は“OK”，退学・除籍や休学をしている場合は“not OK”とした。

以上のように、カリキュラムや教育プログラムは年度により変化しているため、同じ意味をもたせた特徴量でも測定対象が若干異なることがある。

表 2 に各入学年度のサンプル数と、 X_{10} のクラスの内訳を示す。ここで、“OK”，“not OK”の値は、A 大

表 2 サンプル数と X_{10} の内訳

入学年度		2009	2010	2011	2012	2013
サンプル数		794	819	785	694	598
X ₁₀	OK	691	708	650	564	499
	not OK	103	111	135	130	99

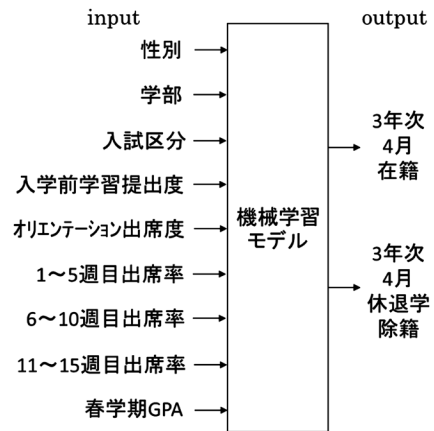


図 2 機械学習によるモデル化

学により情報公開された各年度の入学者数・在籍者数から推算されることを付記する。

3.4 機械学習手法による数理モデル化

本稿では、3.2 節にて述べた問題設定のように、学士課程初期のデータをもとに、経年後の学修状況を予測することを考える。これは、学士課程初期のデータと経年後の学修状況の間の写像関係をデータから学習することと同義であり、この写像関係は非線形性をもつと考えられる。ここでは、機械学習の手法を用いてこの非線形写像をモデル化することを考え、複数の手法を比較することによりその適用可能性を検討する。

$X_1 \sim X_9$ は入学前から 1 年次春学期までの学生の属性や学修状況を、 X_{10} は経年後の学修状況を示す。これらの間の非線形写像を機械学習でモデル化するイメージを図 2 に示す。 X_{10} が離散的な状態を表すことから、このモデルの学習はパターン分類問題と捉えられる。

4. 数値実験結果

4.1 実験手法

3.4 節で述べたモデル化について、複数の代表的な機械学習の手法を用いて数値実験を行い、そのパフォーマンスを比較した。モデルとしては、ロジスティック回帰（以下 Logistic）、ニューラルネットワークの一種である多層パーセプトロン（以下 MLP）および RBF ネットワーク（以下 RBFN）、決定木モデルの J48、サポートベクターマシン（以下 SVM）、アンサンブル学習手法の AdaBoost を用いた。数値実験には機械学習ツールの Weka⁽¹⁰⁾ を用い、すべてデフォルトの設定で実験を行った。用いた Weka のバージョンは 3.6.11 である。Weka は、2.2 節で紹介した Jayaprakash の例⁽⁵⁾ や Djulovic の例⁽⁹⁾ など、Learning Analytics にもしばしば用いられている。

4.2 1 年次春学期までのデータによる予測

データからのモデルの学習においては、未知のデータに対する予測性能の評価が重要である。本研究のケースにおいても、ある入学年度の学生のデータから構築したモデルを用いて、その先の入学年度の学生の学修状況予測が可能にならなければならない。つまり入学年度に対して汎化能力の高いモデル構築が必要である。しかし、3.2 節や 3.3 節で述べたように、入学年度が異なれば、大学が提供する教育内容も少しずつ異なり、また国の教育政策その他の変化による学生の全体的な傾向の違いも生じうる。そのような前提のもとでのモデルの学習について検討するため、ここでの実験においては、モデルの学習（training）と汎化能力のテスト（test）に用いるデータセットについて、次の 4 パターンの組み合わせを考えた。

[データセット 1a] 学 習：2011 年度生

テスト：2013 年度生

[データセット 1b] 学 習：2012 年度生

テスト：2013 年度生

[データセット 2] 学 習：2011～2012 年度生

テスト：2013 年度生

[データセット 3] 学 習：2009～2012 年度生

テスト：2013 年度生

[データセット 1a] [データセット 1b] は単年度のみを学習データとしているが、[データセット 2] および [データセット 3] は複数の年度を合わせて学習データとしている。これにより、年度を超えたデータ構造からの学習が汎化能力にどのように影響するかを検討する。また、3.2 節で述べたとおり、A 大学においては 2011 年度に大きなカリキュラム改編があったため、[データセット 2] と [データセット 3] の比較によって、そうした教育環境の違いを超えたデータ構造の学習の可能性を探る。

実験による分類結果を表 3～6 に示す。正確度（Accuracy）はデータを正しく分類した割合、適合率（Precision）はモデルの予測が正解である割合、再現率（Recall）は真のクラスが正しく分類されている割合を表す。適合率・再現率について図示したものを図 3～6 に示す。

ドロップアウトのリスクがある学生をなるべく多く早期発見するためには、“not OK” クラスにおける再現率を高めることが必要となる。多くのモデルにおいて、[データセット 2] では [データセット 1a] [データセット 1b] よりもテストデータにおける “not OK” クラスの再現率が高く、ばらつきも小さくなっている。単年度のデータによる学習よりも、複数年度にわたるデータによる学習のほうが、より本質的な “not OK” の構造を学習できると考えられる。

[データセット 1a] [データセット 1b] において、MLP は他のモデルよりも、学習データに対しては正確度、適合率、再現率ともに高い傾向にあるが、テストデータに対してはほかのモデルより低めである。これは過学習をしているものと捉えられる。一方で、[データセット 2] においては過学習の程度が抑えられており、[データセット 3] になるとほかのモデルと遜色がなく、むしろ “not OK” の再現率は高めである。

[データセット 2] と [データセット 3] の差はモデルによりまちまちである。[データセット 3] は、MLP のように過学習の傾向が強いモデルについては、過学習をより抑える効果があるように見えるが、ほかのモデルについては [データセット 2] よりもテストデータにおける分類性能がばらついているように見える。先述のように、A 大学では 2011 年度に大きな

表3 各モデルの分類結果 (データセット 1a)

モデル		正確度	クラス	適合率	再現率
Logistic	学習	0.873	OK	0.888	0.968
			not OK	0.727	0.415
	テスト	0.876	OK	0.906	0.950
			not OK	0.667	0.505
MLP	学習	0.948	OK	0.942	0.998
			not OK	0.990	0.704
	テスト	0.846	OK	0.889	0.932
			not OK	0.547	0.414
RBFN	学習	0.848	OK	0.883	0.942
			not OK	0.587	0.400
	テスト	0.873	OK	0.914	0.936
			not OK	0.632	0.556
J48	学習	0.871	OK	0.888	0.966
			not OK	0.718	0.415
	テスト	0.885	OK	0.901	0.968
			not OK	0.742	0.465
SVM	学習	0.866	OK	0.875	0.978
			not OK	0.759	0.326
	テスト	0.876	OK	0.896	0.964
			not OK	0.705	0.434
AdaBoost	学習	0.871	OK	0.888	0.966
			not OK	0.718	0.415
	テスト	0.883	OK	0.901	0.966
			not OK	0.730	0.465

表5 各モデルの分類結果 (データセット 2)

モデル		正確度	クラス	適合率	再現率
Logistic	学習	0.865	OK	0.884	0.962
			not OK	0.709	0.423
	テスト	0.871	OK	0.906	0.944
			not OK	0.641	0.505
MLP	学習	0.924	OK	0.923	0.990
			not OK	0.932	0.623
	テスト	0.851	OK	0.902	0.922
			not OK	0.557	0.495
RBFN	学習	0.853	OK	0.890	0.937
			not OK	0.617	0.468
	テスト	0.876	OK	0.921	0.932
			not OK	0.634	0.596
J48	学習	0.884	OK	0.893	0.975
			not OK	0.804	0.464
	テスト	0.878	OK	0.903	0.956
			not OK	0.686	0.485
SVM	学習	0.863	OK	0.879	0.966
			not OK	0.717	0.392
	テスト	0.883	OK	0.904	0.962
			not OK	0.716	0.485
AdaBoost	学習	0.863	OK	0.894	0.945
			not OK	0.658	0.487
	テスト	0.881	OK	0.915	0.946
			not OK	0.671	0.556

表4 各モデルの分類結果 (データセット 1b)

モデル		正確度	クラス	適合率	再現率
Logistic	学習	0.873	OK	0.887	0.963
			not OK	0.744	0.469
	テスト	0.876	OK	0.905	0.938
			not OK	0.617	0.505
MLP	学習	0.948	OK	0.960	0.982
			not OK	0.915	0.823
	テスト	0.846	OK	0.897	0.872
			not OK	0.434	0.495
RBFN	学習	0.848	OK	0.896	0.933
			not OK	0.645	0.531
	テスト	0.873	OK	0.918	0.924
			not OK	0.604	0.586
J48	学習	0.871	OK	0.904	0.996
			not OK	0.972	0.538
	テスト	0.885	OK	0.890	0.944
			not OK	0.594	0.414
SVM	学習	0.866	OK	0.883	0.973
			not OK	0.792	0.438
	テスト	0.876	OK	0.906	0.966
			not OK	0.742	0.495
AdaBoost	学習	0.871	OK	0.871	0.980
			not OK	0.814	0.369
	テスト	0.883	OK	0.890	0.968
			not OK	0.709	0.394

表6 各モデルの分類結果 (データセット 3)

モデル		正確度	クラス	適合率	再現率
Logistic	学習	0.878	OK	0.894	0.970
			not OK	0.695	0.376
	テスト	0.873	OK	0.906	0.946
			not OK	0.649	0.505
MLP	学習	0.915	OK	0.930	0.973
			not OK	0.804	0.599
	テスト	0.851	OK	0.907	0.916
			not OK	0.553	0.525
RBFN	学習	0.867	OK	0.898	0.950
			not OK	0.604	0.413
	テスト	0.868	OK	0.913	0.930
			not OK	0.611	0.556
J48	学習	0.902	OK	0.905	0.987
			not OK	0.860	0.436
	テスト	0.878	OK	0.896	0.966
			not OK	0.717	0.434
SVM	学習	0.879	OK	0.891	0.976
			not OK	0.724	0.351
	テスト	0.881	OK	0.899	0.966
			not OK	0.726	0.455
AdaBoost	学習	0.861	OK	0.914	0.923
			not OK	0.557	0.524
	テスト	0.836	OK	0.910	0.892
			not OK	0.505	0.556

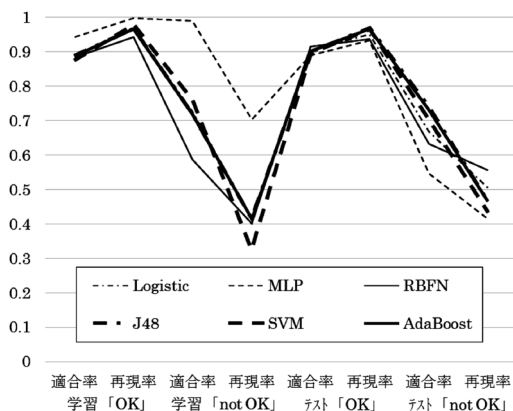


図3 各モデルの分類結果（データセット1a）

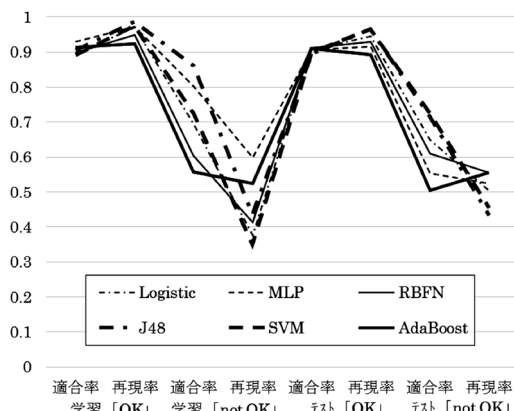


図6 各モデルの分類結果（データセット3）

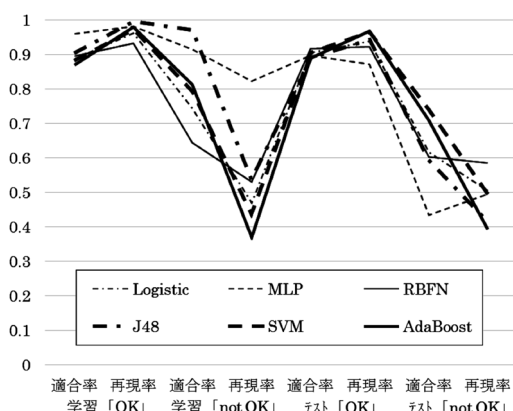


図4 各モデルの分類結果（データセット1b）

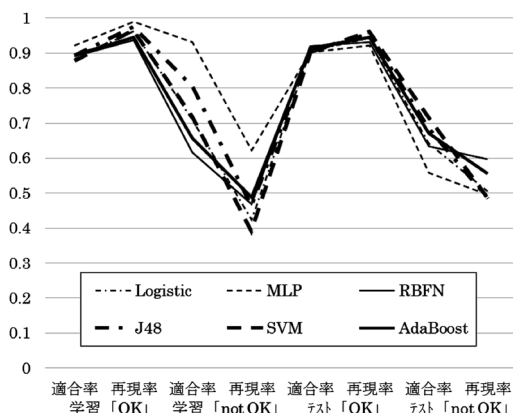


図5 各モデルの分類結果（データセット2）

カリキュラム改編があり、その前後では教育環境が大きく異なっている。[データセット3]の学習データはこのカリキュラム改編の前後を合わせたものである。[データセット2]の結果は、複数年度にわたるデータが汎化能力を高める可能性を示したが、[データセット3]の結果は、カリキュラム改編のような大きな環境変化を超えたデータ構造の学習が困難であることを示唆している。今後さらに検証を重ね、これらの可能性を探りたい。

[データセット2]においては、“not OK”の再現率はいずれのモデルも0.485～0.596の範囲にある。すなわち、A大学においては、適当なモデルを用いることにより、1年次の春学期が終了した時点で、3年次当初までにドロップアウトする学生のおよそ50～60%程度を予測可能であることがわかる。

また、いずれのデータセットに対しても、“not OK”の再現率はRBFNが最も高い値を示しており、数ある機械学習モデルの中でも、RBFNは比較的汎化能力の高いモデル化が可能であることが示唆される。

4.3 入学直後のデータによる予測

次に、より入学初期のデータから X_{10} をどの程度予測可能かを検証した。表7は、 $X_1 \sim X_5$ まで、すなわち1年次春学期の授業開始までにデータを入手可能な特徴量を使用したときの分類結果、表8は、それに加えて、最初の5週間の必修科目の出席率である X_6 を使用したときの分類結果である。いずれも[データセット2]を使用した。

X_5 までのデータでは、たとえば J48 や SVM はすべてを“OK”に分類しており、“not OK”の適合率・再現率はともに 0 である。その他のモデルも、学習データ・テストデータともに“not OK”の再現率は低い。ただし、テストデータの“not OK”の適合率を見ると、Logistic や AdaBoost では 0.5 を超えており、予測の半分以上が実際に“not OK”であるようなモデル構築は可能であることがわかる。

さらに X_6 （春学期第 5 週までの出席率）を用いると、いずれのモデルも分類性能は向上した。テストデータに対する“not OK”の再現率は、最も高い RBFN で 0.394 となり、実際に“not OK”である学生のおよそ 40% を春学期第 5 週が終了した時点で予測可能であることがわかる。

4.4 機械学習によるモデル化の効果

図 7～9 は、ある一つの特徴量と X_{10} の相関を見るために表した散布図である。例として X_6 （1 年次春学期 1～5 週必修出席率）、 X_8 （1 年次春学期 11～15 週必修出席率）、 X_9 （1 年次春学期 GPA）の 3 種類それぞれと X_{10} との散布図を示している。同一の点に複数のデータがある場合は、プロットする円の面積とデータの個数を比例させて表示している。これらの図からわかるように、とくに“not OK”クラスにおけるばらつきは大きく、単一の特徴量のみで精度よく分類することの困難さがうかがえる。

仮にこうした単一の特徴量から X_{10} を分類することを考える。具体的には、学習データにおいて正確度が最も高くなる点を閾値とするモデルを想定する。[データセット 2] を用いたときのテストデータに対する分類結果を表 9 に示す。同じように [データセット 2] を用いている表 5 と比較すると、とくに“not OK”の再現率について、 $X_1 \sim X_9$ を用いた多くの機械学習モデルより劣っている。表 8 との比較においても、 X_6 のみによる分類は、SVM を除き、 $X_1 \sim X_6$ を用いた各機械学習モデルと同等以下であることがわかる。これらの結果から、単一の特徴量からの分類に比べ、図 2 のように多変数からの非線形写像を機械学習によりモデル化することには一定の効果があると考えられる。

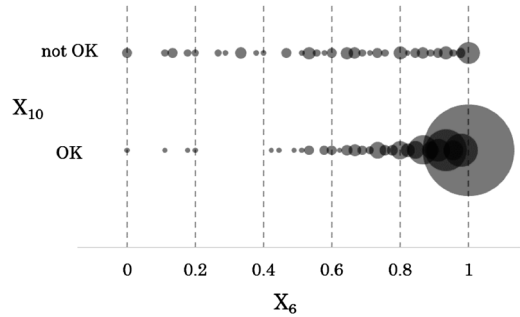


図 7 X_6 と X_{10} の散布図（2013 年度生）

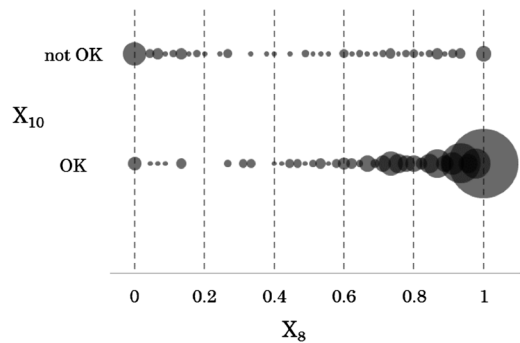


図 8 X_8 と X_{10} の散布図（2013 年度生）

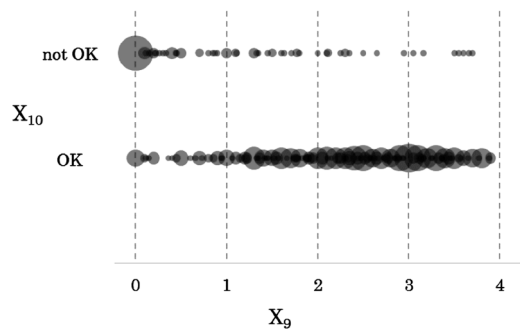


図 9 X_9 と X_{10} の散布図（2013 年度生）

4.5 学修支援への応用

予測された学修状況から、適当な閾値を設けることでドロップアウトリスクの高い学生を割り出し、実際に学生への支援や介入を行うことが可能であると考えられる。4.3 節で述べた結果からわかるように、A 大学のデータにおいては、学士課程のかなり初期の段階から、一定の予測性能をもつモデル構築が可能である。今後、少なくとも A 大学と同様の規模や形態、教育方針等をもつ大学において、このようなモデル構

表7 $X_1 \sim X_5$ のみによる分類結果 (データセット2)

モデル		正確度	クラス	適合率	再現率
Logistic	学習	0.821	OK	0.833	0.979
			not OK	0.500	0.098
	テスト	0.836	OK	0.861	0.958
			not OK	0.512	0.222
MLP	学習	0.848	OK	0.852	0.986
			not OK	0.770	0.215
	テスト	0.818	OK	0.852	0.946
			not OK	0.386	0.172
RBFN	学習	0.817	OK	0.828	0.982
			not OK	0.436	0.064
	テスト	0.828	OK	0.844	0.974
			not OK	0.409	0.091
J48	学習	0.821	OK	0.821	1.000
			not OK	0.000	0.000
	テスト	0.834	OK	0.834	1.000
			not OK	0.000	0.000
SVM	学習	0.821	OK	0.821	1.000
			not OK	0.000	0.000
	テスト	0.834	OK	0.834	1.000
			not OK	0.000	0.000
AdaBoost	学習	0.820	OK	0.825	0.992
			not OK	0.474	0.034
	テスト	0.836	OK	0.845	0.984
			not OK	0.529	0.091

表8 $X_1 \sim X_6$ のみによる分類結果 (データセット2)

モデル		正確度	クラス	適合率	再現率
Logistic	学習	0.837	OK	0.853	0.968
			not OK	0.618	0.238
	テスト	0.873	OK	0.882	0.978
			not OK	0.756	0.343
MLP	学習	0.887	OK	0.885	0.991
			not OK	0.908	0.411
	テスト	0.848	OK	0.871	0.960
			not OK	0.583	0.283
RBFN	学習	0.836	OK	0.862	0.954
			not OK	0.585	0.298
	テスト	0.860	OK	0.888	0.952
			not OK	0.619	0.394
J48	学習	0.867	OK	0.873	0.982
			not OK	0.805	0.343
	テスト	0.851	OK	0.880	0.952
			not OK	0.586	0.343
SVM	学習	0.821	OK	0.821	1.000
			not OK	0.000	0.000
	テスト	0.834	OK	0.834	1.000
			not OK	0.000	0.000
AdaBoost	学習	0.839	OK	0.841	0.991
			not OK	0.776	0.143
	テスト	0.863	OK	0.864	0.992
			not OK	0.840	0.212

表9 単一の特徴量による分類 ([データセット2] のテストデータに対する分類結果)

特徴量	正確度	クラス	適合率	再現率
X_6	0.863	OK	0.864	0.992
		not OK	0.840	0.212
X_8	0.876	OK	0.901	0.956
		not OK	0.681	0.475
X_9	0.883	OK	0.901	0.966
		not OK	0.730	0.465

築手法の一般化の可能性を検討したい。

また本稿では、予測対象として在学状況に着目したが、GPA や単位数、就職状況など、学修成果に着目することも同一のモデル上で可能である。さらには、特徴量を学士課程全体にわたって追加することで、1 年次終了時、2 年次終了時など、より細かく適時的な退学リスクや学修状況の推定も可能になると考えられる。

実用性を高めるためには優れた汎化能力が必要となるため、ネットワーク構造の決定やモデルパラメータの設定についてはさらに検討が必要である。

5. おわりに

本稿では、学士課程を通して蓄積される大規模な学修データから、学生の学士課程にわたる学修状態を数理的にモデル化し、これを学修支援へ活用することについて検討した。

数値実験からは、少なくとも本稿で用いたデータにおいては、機械学習手法によって一定の予測性能をもつ学修状態のモデル化が可能であることと、単年度のみデータではなく複数の入学年度を合わせたデータによる学習を行うことで、異なる入学年度に対する汎化能力に優れたモデル構築が可能であることが示された。ただし、どの程度まで過去のデータが有効にはたらくかについては、さらに検討が必要である。

また、本稿で用いたデータにおいては、RBFN を用いれば、1 年次春学期第 5 週の終了時点で、3 年次当初までにドロップアウトする学生のおよそ 40% を予測可能であり、春学期終了時点になると多くの機械学習手法において 50~60% 程度を予測可能であることも示された。

今後は、実際の学修支援への適用や、より信頼性の高いモデル構築について検討していきたい。特徴量の選択や、さらに学士課程全体へと引き伸ばしたモデルの構築なども今後の課題である。また、A大学のデータに基づいた本手法について、ほかの教育機関における適用可能性を探ることも課題である。

参 考 文 献

- (1) Siemens, G. and Baker, R. S. J. d.: "Learning Analytics and Educational Data Mining: Towards Communication and Collaboration", Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, pp. 252-254 (2012)
- (2) 沖 清豪, 岡田聡志 (編著): "データによる大学教育の自己改善—インスティテューショナル・リサーチの過去・現在・展望—", 学文社, 東京, pp. 139-157 (2011)
- (3) 松田岳士: "教学 IR の役割と実践事例", 教育システム情報学会誌, Vol. 31, No. 1, pp. 19-27 (2014)
- (4) リチャード・D・ハワード (編), 大学評価・学位授与機構 IR 研究会 (訳): "IR 実践ハンドブック 大学の意思決定支援", 玉川大学出版部, 東京 (2012)
- (5) Jayaprakash, S. M., Moody, E. W., Lauria, E. J. M. et al.: "Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative", Journal of Learning Analytics, Vol. 1, No. 1, pp. 6-47 (2014)
- (6) Vihavainen, A., Luukkainen, M. and Kurhila, J.: "Using Students' Programming Behavior to Predict Success in an Introductory Mathematics Course", Proceedings of the 6th International Conference on Educational Data Mining, pp. 300-303 (2013)
- (7) 雨森 聡, 松田岳士, 森 朋子: "教学 IR の一方略—島根大学の事例を用いて—", 京都大学高等教育研究, 第 18 号, pp. 1-10 (2012)
- (8) 大友愛子, 岩山 豊, 毛利隆夫: "学内データの活用—大学における IR (Institutional Research) への取り組み—", FUJITSU, Vol. 65, No. 3, pp. 41-47 (2014)
- (9) Djulovic, A. and Li, D.: "Towards Freshman Retention Prediction: A Comparative", International Journal of Information and Education Technology, Vol. 3, No. 5, pp. 494-500 (2013)
- (10) Weka 3: Data Mining Software in Java: <http://www.cs.waikato.ac.nz/ml/weka/> (参照 2015.6.19)

著 者 紹 介

近藤 伸彦

2003 年大阪大学工学部卒。2007 年同大学大学院博士後期課程修了。博士(情報科学)。2009 年より、大手前大学助教 (2011 年度まで), CELL 教育研究所研究員。2012 年より、大手前大学教学運営室室員, 現在に至る。教学 IR に従事。教育システム情報学会, 計測自動制御学会, 情報処理学会, 日本リメディアル教育学会, 大学教育学会, 大学行政管理学会各会員。



畠中 利治

1986 年大阪大学工学部卒。1988 年同大学大学院博士前期課程修了。2002 年より大阪大学助手, 現在に至る。博士(情報科学)。進化計算, システム同定の研究に従事。計測自動制御学会, 進化計算学会, IEEE 各会員。

