

科目選択支援のための SDLRS に基づく成績予測に関する考察

A Study for Prediction of Academic Performance Based on SDLRS

近藤 伸彦^{*1}, 松田 岳士^{*1}, 渡辺 雄貴^{*2}, 重田 勝介^{*3}, 加藤 浩^{*4}

Nobuhiko KONDO^{*1}, Takeshi MATSUDA^{*1}, Yuki WATANABE^{*2}, Katsusuke SHIGETA^{*3}, Hiroshi KATO^{*4}

^{*1} 首都大学東京大学教育センター

^{*1} University Education Center, Tokyo Metropolitan University

^{*2} 東京理科大学教育支援機構

^{*2} Organization for Education Advancement, Tokyo University of Science

^{*3} 北海道大学情報基盤センター

^{*3} Information Initiative Center, Hokkaido University

^{*4} 放送大学教養学部

^{*4} Faculty of Liberal Arts, The Open University of Japan

Email: kondo@tmu.ac.jp

あらまし：学生が履修科目を選択する際に、従来のシラバス情報だけでなく、学生自身の特性と科目の特性とのマッチングをもとにした情報が示されれば、より充実した科目選択支援が可能になると考えられる。本発表では、そのような情報を示すための基礎研究として、自己管理学習のレディネス尺度 SDLRS に基づき機械学習の手法を用いて授業の成績を予測することを考える。実際のいくつかの授業における実験結果から、適当なアルゴリズムを用いれば一定の性能をもった予測ができる可能性が示唆された。

キーワード：科目選択支援、自己管理学習、SDLRS、成績予測、機械学習

1. はじめに

大学における履修科目選択は、多くの場合一定の自由度のもとで、学生の自己決定によりなされる。とくに初年次は一般に教養系の科目が多く、科目選択の幅が広いことが多い。学生個々に応じた適切な履修科目選択のためには適切な支援が必要であると考えられるが、履修指導の方法や質が大学によって大きく異なるなど、学生は不十分あるいは不適当な情報に基づいて、十分な検討を経ずに履修選択を行っていることも多いと考えられる。

これについて松田らは、学生の自己管理学習 (Self-Directed Learning, 以下 SDL) を支援するという観点から、SDL の Readiness Scale である SDLRS を用いて学生の科目選択支援を行うためのシステムについて検討している⁽¹⁾。ここでは、学生の有する SDL レディネスと授業が求める SDL レディネスとのマッチング情報、および成績分布や単位修得可能性などを提示することが学生の履修科目選択において有用であることが示唆されている⁽¹⁾。

本稿ではこれらをふまえ、科目選択支援のために SDLRS をもとにした成績予測を行うことについて考える。成績予測には機械学習の手法を使用し、実際のいくつかの授業におけるデータを用いた実験の結果をもとに、その適用可能性について考察する。

2. 科目選択支援と SDLRS

2.1 SDLRS

SDLRS は Guglielmino により開発された SDL のレディネス尺度として知られる⁽²⁾。SDLRS の日本語版は松浦らによって開発され、信頼性と妥当性も検討されている⁽³⁾。日本語版 SDLRS は 58 項目から構成

され、7 因子が抽出されている。各項目は、その内容が「全くあてはまらない」から「いつでもあてはまる」の 5 段階で回答されるものとなっている。本研究では、先行研究⁽¹⁾⁽⁴⁾と同様に、各因子に該当する質問項目のうち因子負荷量の高い項目をそれぞれ 1～3 項目抽出した合計 16 項目を用いる。表 1 に具体的な項目を示す。ここで、因子 I～VII は、I：学習への愛着、II：基礎学習技法の活用能力、III：学習に対する自己責任の受容、IV：探究心、V：学習における主体性、VI：エネルギーシユな自己イメージ、VII：将来に対する前向きな姿勢、と命名されている。

2.2 SDLRS に基づく科目選択支援と成績予測

1 で述べたように、学生の科目選択等の支援を行うための要件として、SDL レディネスについての学

表 1 本研究で用いた SDLRS の項目 (*は逆転項目)

因子	項目 (番号は質問紙における順番を示す)
I	①学ぶことが好きだ ⑫学べば学ぶほど、世界はおもしろくなる ⑬学習は楽しい
II	②自分一人ではうまく学習できない ^(*) ③他の多くの人達よりもうまく自己学習ができる ⑭クラスの中でも自分一人の時でも、効果的な学習ができる
III	④自分が学ぶことについて責任を負うのは、自分以外の何者でもない ⑥もし私が学習しなくても、自分のせいではない ^(*) ⑮自分の学習に責任を負うのは、自分であり他人ではない
IV	⑪うまくいくかどうかわからない学習状況は好きではない ^(*) ⑭いつも新しい方法を試すよりは、よく知られた学習方法どおりやる方がよい ^(*)
V	⑤学習しようと思ったことがあれば、たとえどんなに忙しくても、そのための時間を作ることができる ⑦知る必要があることは、ほぼ何でも勉強することができる
VI	⑧私は、何事にも好奇心がおう盛だ ⑨私は、何をやるにもユニークな方法を考え出すのが得意だ
VII	⑩将来について考えるのが好きだ

生と授業のマッチング情報や、単位修得可能性などを表示することが科目選択において有用であることが示唆されている⁽¹⁾。これに関連して、SDLRS と成績の関係についての調査も行われている⁽⁴⁾。

本稿では、SDLRS に基づく質問紙調査の回答をもとに、履修前の時点における科目の成績予測がどの程度可能であるかを検討する。具体的にはこれを成績の高低を判別するパターン分類問題とみなし、いくつかの機械学習のアルゴリズムを用いたときのふるまいを検証する。

3. 数値実験

3.1 予測モデルの設定と実験環境

本実験では、予測モデルの説明変数として表 1 で示した 16 の項目を、目的変数として授業の成績を用いた。授業の成績は、成績の素点 (100 点満点) をある閾値によって High/Low の 2 値に離散化するものとした。本稿ではこの閾値を 60 点、70 点、80 点と 3 種類設定し、それぞれについて調べた。

実験は Weka 3.6.10⁽⁵⁾を用いて行い、機械学習のアルゴリズムには代表的なものを複数用いた。本稿ではそのうちサポートベクターマシンとロジスティック回帰の結果について示す。

本実験は、ある大学の授業 A、別の大学の授業 B の 2 つの授業を対象とした。授業 A は 2015 年度後期および 2016 年度後期 (N=57)、授業 B は 2016 年度前期および 2017 年度前期 (N=41) のデータをそれぞれ用いた。授業 A、B はいずれも授業ごとに例年同一の内容で実施されている。

3.2 実験結果

分類性能は 5-fold cross validation により評価した。評価指標は、誤分類しない割合である適合率と、正解が正しく分類されている割合である再現率を用いた。表 2~5 に実験結果を示す。

紙面の都合上一部しか示せないが、機械学習のアルゴリズムによって結果は大きく異なる。また授業によっても予測結果が異なる。これは授業によって求められる SDL レディネスが異なり、また成績評価基準が異なることから当然に予想される結果である。

High/Low を分ける閾値の設定によっても分類性能は変化する。たとえば閾値が高くなるほど High クラスの適合率・再現率はともに低くなる一方、Low クラスの適合率・再現率は高くなる傾向にある。逆に閾値が低ければ High クラスの適合率・再現率は高く、Low クラスの適合率・再現率は低くなる傾向がある。ただしその傾向に外れる結果もあり、たとえば表 3 においては閾値を 60 点にしたとき、適合率は High クラスで 0.933、Low クラスで 0.500 となっている。つまり、High と予測すればほぼ正解、Low と予測すれば半数は正解することになる。

サンプルサイズが小さいため引き続き検証が必要であるが、適当なアルゴリズムと閾値を用いれば、成績の高低を一定の性能をもって予測可能であることが示唆される。

表 2 予測結果 (授業 A, サポートベクターマシン)

閾値	High クラス			Low クラス		
	N	適合率	再現率	N	適合率	再現率
60 点	48	0.870	0.979	9	0.667	0.222
70 点	44	0.772	1.000	13	0.000	0.000
80 点	35	0.673	0.943	22	0.750	0.273

表 3 予測結果 (授業 A, ロジスティック回帰)

閾値	High クラス			Low クラス		
	N	適合率	再現率	N	適合率	再現率
60 点	48	0.933	0.875	9	0.500	0.667
70 点	44	0.860	0.841	13	0.500	0.538
80 点	35	0.649	0.686	22	0.450	0.409

表 4 予測結果 (授業 B, サポートベクターマシン)

閾値	High クラス			Low クラス		
	N	適合率	再現率	N	適合率	再現率
60 点	37	0.902	1.000	4	0.000	0.000
70 点	32	0.789	0.938	9	0.333	0.111
80 点	24	0.590	0.958	17	0.500	0.059

表 5 予測結果 (授業 B, ロジスティック回帰)

閾値	High クラス			Low クラス		
	N	適合率	再現率	N	適合率	再現率
60 点	37	0.914	0.865	4	0.167	0.250
70 点	32	0.828	0.750	9	0.333	0.444
80 点	24	0.684	0.542	17	0.500	0.647

4. おわりに

本稿では、科目選択支援のための基礎調査として、SDLRS を用いた機械学習による成績予測について実験的に検討した。信頼性を高めるため今後さらに検証を進めるとともに、授業の求めるレディネスと予測性能との関連などについても検討を進めたい。

謝辞

本研究は JSPS 科研費 JP16K16331 および JP16H03082, ならびに JP17H01998 の助成を受けた。

参考文献

- (1) 松田岳士, 渡辺雄貴, 重田勝介, 加藤浩: “教学データと SDLRS を用いた科目選択支援システムの開発”, 日本教育工学会研究報告集, JSET15-5, pp.225-230 (2015)
- (2) L. M. Guglielmino, Development of the Self-Directed Learning Readiness Scale, Doctoral Dissertation, University of Georgia (1977)
- (3) 松浦和代, 阿部典子, 良村貞子, 神成陽子, 升田由美子, 阿部修子, 浜めぐみ: “日本語版 SDLRS の開発—信頼性と妥当性の検討—”, 日本看護研究学会雑誌, No.26, Vol.1, pp.45-53 (2003)
- (4) 松田岳士, 渡辺雄貴, 重田勝介, 近藤伸彦, 加藤浩: “科目選択支援に SDLRS を用いるシステム開発のための基礎研究—成績との関係の検討—”, 日本教育工学会研究報告集, JSET17-1, pp.63-68 (2017)
- (5) Weka 3: Data Mining Software in Java: <http://www.cs.waikato.ac.nz/ml/weka/> (参照 2018.6.12)