

数学リメディアル教育における期末試験の得点予測に活用できる学習者データの分析

Analysis of Useful Learner Data for Final Exam Score Prediction in Remedial Mathematics Education

手塚 祐樹^{*1}, 高木 正則^{*1}, 山田 敬三^{*1}, 佐々木 淳^{*1}, 森本 康彦^{*2}
Yuki TETSUKA^{*1}, Masanori TAKAGI^{*1}, Keizo YAMADA^{*1}, Jun SASAKI^{*1}, Yasuhiko MORIMOTO

^{*1}岩手県立大学大学院 ソフトウェア情報学研究科
^{*1}Graduate School of Software and Information Science, Iwate Prefectural University

^{*2}東京学芸大学
^{*2}Tokyo Gakugei University

Email: g231o021@s.iwate-pu.ac.jp, takagi-m@iwate-pu.ac.jp, k-yamada@iwate-pu.ac.jp,
jsasaki@iwate-pu.ac.jp, morimoto@u-gakugei.ac.jp

あらまし: ラーニングアナリティクスでは学習者の有用な学びの記録を収集・蓄積することが重要とされている。本研究では、数学リメディアル科目を対象とし、期末試験の得点予測に有用な学習データを明らかにすることを目的とし、平成 26 年度と平成 27 年度に収集した多様な学習者データの組み合わせで予測精度の比較実験を実施した。実験の結果、各回の授業終了直前に実施する確認テストの得点と e ラーニングの学習時間が期末試験の得点予測に有用であることが示唆された。

キーワード: 得点予測, 学習履歴分析, ラーニングアナリティクス

1. はじめに

近年、MOOC や e ポートフォリオなどの利用が急速に拡大し⁽¹⁾、膨大な学習行動ログや学習活動の記録（以下、学習者データ）が収集・蓄積されるようになってきた。これに伴い、学習状況を把握・最適化するために、学習者とそれを取りまく文脈に関わるデータを測定、収集、分析、報告する方法であるラーニングアナリティクス（Learning Analytics）が注目を集めている⁽²⁾。ラーニングアナリティクスでは、学習履歴データを分析、可視化することにより、学習者の学習到達度の把握、将来的な活動結果予測、学力を予測するための要因調査などに活かそうとする研究があり、学習者の有意味で有用な学びの記録を収集・蓄積することが重要とされている⁽³⁾。

一方、大学全入時代の到来や大学入試の多様化により、多様な学力を持った学生が大学に入学するようになり、各大学ではリメディアル教育の需要が高まっている。リメディアル教育の対象学生は、高校時の数学の履修状況が多様なため、教員が全学生に対して講義をする従来型の画一的な授業が困難である。そのため、岩手県立大学ソフトウェア情報学部では、e ラーニング教材を活用した個別学習やグループ学習、個別面談など多様な教育方法を駆使して授業を実施している⁽⁴⁾が、毎年、期末試験で不合格になる学生が少なからず存在した。これらの学生には、セメスターの早い時期から自身の学習を内省させ、適切な学習へと導く必要がある。本研究では、これらの学生に対し、期末試験の得点予測結果を活用することを検討しているが、既存研究では、期末試験の得点予測や予測結果に基づいた学習支援のために収集・蓄積すべき学習者データが明らかになっていない。そこで、本研究では数学リメディアル教

育でこれまで収集・蓄積してきたデータを活用し、期末試験の得点予測に活用できる学習者データを明らかにすることを目的とする。

2. 期末試験の得点予測実験

2.1 活用する学習者データとアルゴリズム

本研究では平成 26 年度と 27 年度に岩手県立大学ソフトウェア情報学部で実施された数学リメディアル科目で収集された 133 人分の学習者データを活用する。本科目で収集した学習者データを表 1 に示す。

表 1 活用可能な学習者データ

No	データ名	データの概要
(1)	プレースメントテストの得点	数学リメディアル科目の履修可否を判定するために入学前に実施されるテストの得点 (0 点~100 点)。
(2)	各授業の事前テストの得点	各授業の最初に実施される確認テストの得点 (0 点~10 点)。
(3)	各授業の事後テストの得点	各授業の最後に実施される確認テストの得点 (0 点~10 点)。
(4)	e ラーニングの学習時間	各回の授業で公開された e ラーニング教材を利用して学習(教科書の閲覧と演習問題の解答)した時間(秒)。
(5)	e ラーニングのヒント閲覧回数	各回の授業で公開された e ラーニング教材のうち、演習問題を解答する際にヒントを閲覧した回数。
(6)	e ラーニングの進捗率	各回の授業で公開された e ラーニング教材のうち、演習問題を解答して正解した問題の割合 (0~100%)。
(7)	作問演習の成果物	作問演習で作成された問題や相互評価時に投稿されたコメント。
(8)	主観的な理解度調査アンケート結果	各授業で学習する学習単元の理解状況を質問したアンケートへの回答結果(できる, できない, 不明など)。
(9)	高校時の数学の履修状況	各授業で学習する学習単元の高校時の履修状況を質問したアンケートへの回答結果(習った, 習わなかった, 不明)。
(10)	TA による個別面談結果	各学習単元の最終回に実施した TA との個別面談後に TA が入力した内容。

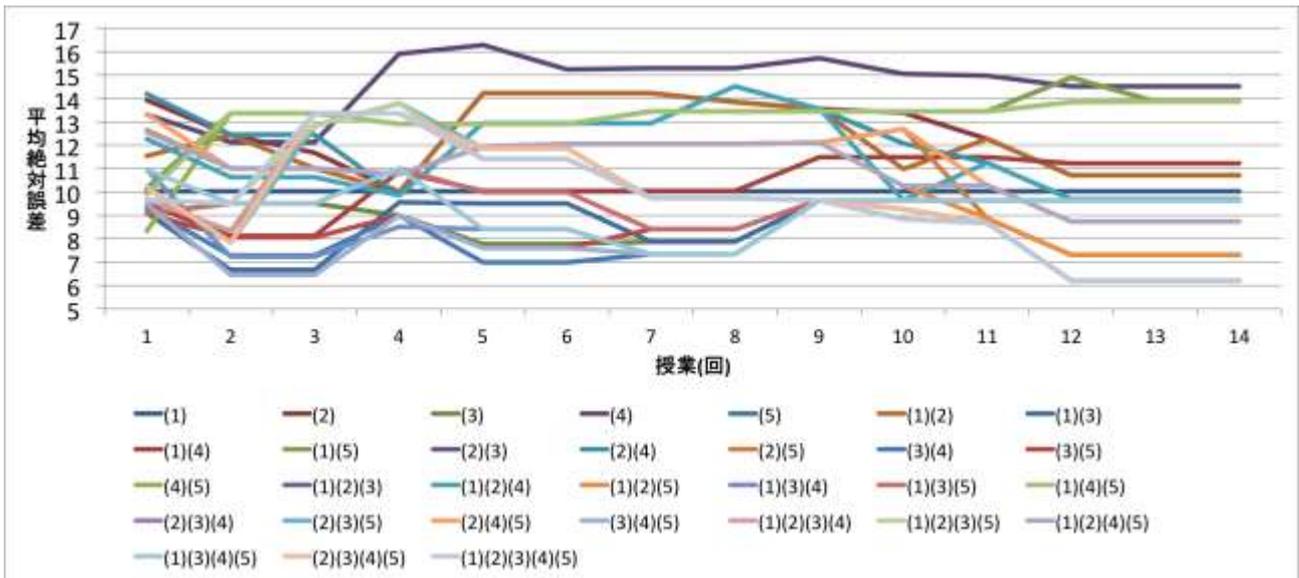


図1 期末試験の予測得点の平均絶対誤差の比較

表2 各回の授業終了時点で平均絶対誤差が最も低い学習者データの組み合わせ

回	1	2	3	4	5	6	7
学習者データ	(4)(5)	(3)(4)(5)	(3)(4)(5)	(1)(3)(4)	(3)(4)	(3)(4)	(3)(4)
回	8	9	10	11	12	13	14
学習者データ	(3)(4)	(3)(4)	(1)(2)(3)(4)(5)	(1)(2)(3)(4)(5)	(1)(2)(3)(4)(5)	(1)(2)(3)(4)(5)	(1)(2)(3)(4)(5)

本研究では、表1の学習者データのうち(1)~(5)を活用し、CARTアルゴリズムを利用して決定木を構築することで得点を予測する。

2.2 実験手順

対象科目では毎回の授業で収集するデータ(事前・事後テストの得点など)があるため、欠席などで学習者データの一部が欠損してしまうことがある。この欠損データが存在すると決定木を構築できなくなってしまうため、期末試験の得点を予測する前に、張らが提案した欠損値補完法⁶⁾を用いて欠損値を補完した。具体的には約30個の学習者データを補完した。

次に、交差検証により予測した得点と実際の期末試験の得点の平均絶対誤差を算出して予測精度を評価した。この評価は授業の第1回から第14回まで各授業終了時点で収集されたデータを活用し、表1(1)~(5)の学習者データすべての組み合わせ(31通り)で行った。なお、交差検証におけるテストデータ数はステージスの公式を活用して決定した。

2.3 実験結果

決定木により予測された期末試験の得点と実際の期末試験得点の平均絶対誤差の推移を図1に示す。また、各回において最も平均絶対誤差が低くなった学習データの組み合わせを表2に示す。表2より、最も平均絶対誤差が低くなる学習データの組み合わせは各回で異なることがわかる。また、ほとんどの回で、(3)事後テストの得点と(4)eラーニングでの学習時間が含まれている。これらは(1)プレースメント

テストの得点のみで予測した得点と比べても誤差は低く、期末試験得点の予測に有用な学習者データであると考えられる。

3. おわりに

本研究では数学リメディアル科目における期末試験の得点予測に活用できる学習者データを明らかにするため、平成26年度、27年度に収集・蓄積された学習者データを活用し、期末試験の得点予測の比較実験を実施した。実験の結果、事後テストの得点とeラーニングの学習時間が期末試験得点の予測において有用な学習者データであることが示唆された。今後はニューラルネットなど別のアルゴリズムでも同様の実験を行い、予測精度を高める学習者データとアルゴリズムの組み合わせを明らかにする。

参考文献

- (1) 森本康彦, 喜久川功, 宮寺庸造: “eポートフォリオ活用のための蓄積文法と支援システムの開発”, 日本教育工学会論文誌, 35巻, 3号, pp.227-236 (2011)
- (2) 1st International Conference on Learning Analytics and Knowledge. 2011. <https://tekri.athabascau.ca/analytics/> (2016.06.07 閲覧)
- (3) 森本康彦, eポートフォリオとしての教育ビッグデータとラーニングアナリティクス, コンピュータ&エデュケーション, Vol.38, pp.18-27, 2015
- (4) 高木正則: “数学リメディアル教育における反転授業の実践と評価”, 情報処理学会研究報告コンピュータと教育(CE), Vol.2015-CE-131, No.14, pp.1-6 (2015)
- (5) 張諾, 木村寛明, 高木正則: “データの特徴選択に基づいた教育データの欠損値補完法”, 電子情報通信学会技術研究報告, 114巻, 513号, pp.89-93 (2015)