

# LCT の結果を使った期末試験の合否予測： 2017LCT から作成したモデルに2018LCT を適用

廣瀬 英雄\*

(令和元年8月6日受付)

## Success/Failure Prediction for Final Examination Using the LCT Results: A Model using 2017 LCT Data is Applied to 2018 LCT Data

Hideo HIROSE

(Received Aug. 6, 2019)

### Abstract

The author proposed a prediction method to find students with drop-out risks as early as possible by using the learning check testing (LCT) results, where the nearest-neighbor method was used in the previous paper. If we want to build a mathematical model using a part of 2018 LCT results and we want to predict 2018 end-term results using the model, it is impossible to do that in the middle of the term, because we will not have the final results until the end of the term. Thus, in this paper, we investigated whether the proposed method was valid or not by applying the model built using 2017 LCT data to newly obtained 2018 LCT results. Using 2017 model, the prediction accuracies and the corresponding confusion matrices applying to 2018 case were almost the same as that applying to 2017 case. Therefore, the previously proposed method is proved to be useful.

**Key Words:** online testing, learning check testing, item response theory, drop-out.

### 1. はじめに

先に報告した「大規模オンラインテストから得られるラーニングアナリティクス」では、フォローアッププログラム (FP follow-up program) の中に組み込まれている、授業時間でのオンラインテスト (LCT, learning check testing) を使ってドロップアウトリスクを抱えた学生を早期に発見しアラートを発するような一つの方法を提案した。このとき予測に用いた LCT のデータは2017年度前期のもので、検証に用いたデータも同じく2017年度前期のものであった。つまり、トレーニングデータとテストデータを同じにしていたようなものである。そのときにはまだ、2018年度のデータを使うことができなかったからである。

したがって、先の予測結果は甘い方向になされているかもしれないとも考えられる。そこで、今回は、2018年度のデータが使える状態になったため、2017年度 LCT データから作られたモデルを2018ケースに適用して予測を行なう検証を行なった。

2016年度から2018年度までの3年間に入学した学生を対象とすると、最近3年間でドロップアウトした学生のうち2/3は数学の期末試験に失敗していた。具体的には、このようなドロップアウト学生数は149人、そのうち期末試験に失敗した学生数は99人であった。そこで、先の報告と同様、ドロップアウトリスクを数学の期末試験に失敗することと同等に取り扱うこととし、ここでは、期末試験に不合格となる確率をそれまでに行った LCT のトレンドから

\* 広島工業大学環境学部建築デザイン学科

予測を行うこととした。

## 2. LCT からの期末試験不合格確率の予測

期末試験に失敗するリスクを抱えた学生を早期に特定するための基準は最近傍(Nearest Neighbor)である。図1に、トレンド間の距離を用いて類似度を求める考え方を図示した。

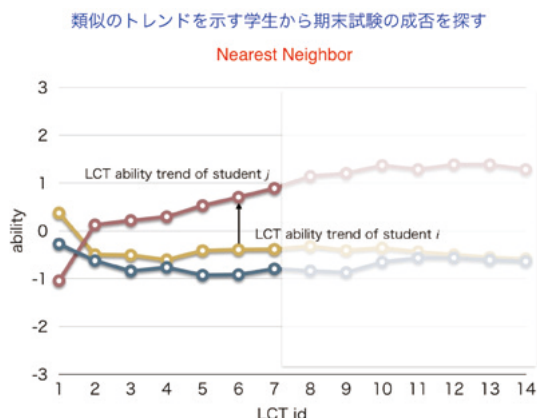


図1 トレンド間の距離を用いた類似度 (文献1の再掲)

図2には、距離を求める数式を示した。2017年度のある学生に最も似た学生を2017年度の学生の中から10人集めてそれらの学生の動向(期末試験の合否)から当該学生の期末試験の合否確率を求めたのが先に報告した内容であった。ここでは、2018年度のある学生に最も似た学生を2017年度の学生の中から10人集めてそれらの学生の動向(期末試験の合否)から当該学生の期末試験の合否確率を求めた。

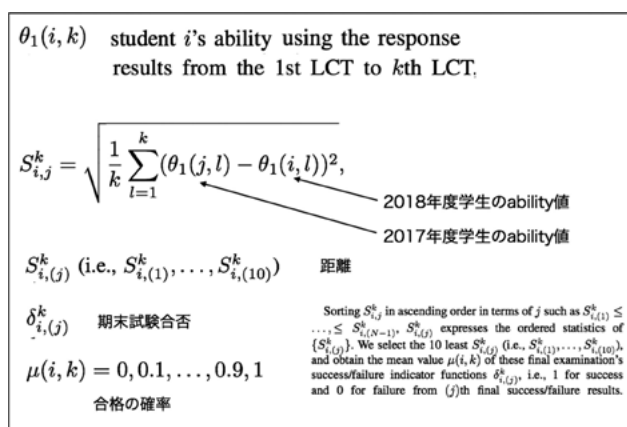


図2 距離を用いた類似度の求め方

この場合、確率は0、0.1、0.2、…、1.0の11パターンが得られる。そこで、リスク管理に重要になりそうな確率として不合格率0.3から0.5の間について予測してみた結果が図3である。図では、LCTの1回目から4回目まで、7回目まで、11回目までの3ケースを示している。例えば、

LCTの1回目から11回目までの結果を使った場合、期末試験の不合格率が0.4以上になる人数を約150人と予測しており、実際にその半数の80人が期末試験に失敗していることがわかる。

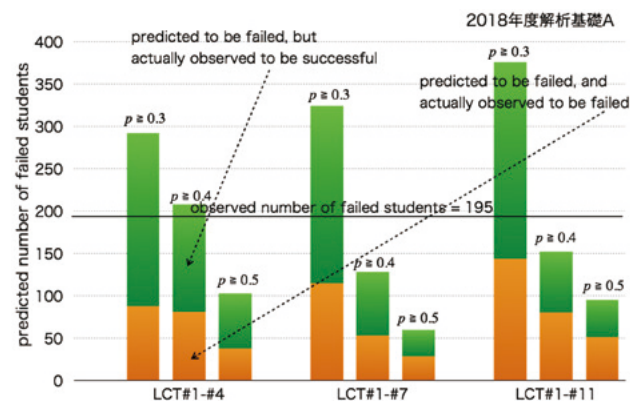


図3 期末試験の合格者予測数と実際の合格者と不合格者(2018年度の類似データを2017年度から求めた)

このときの  $p \geq 0.3$ 、 $p \geq 0.4$ 、 $p \geq 0.5$  の場合の confusion matrix を求めてみた結果が表1、表2、表3である。また、そのときの誤分類率と不合格の的中率も表の中に併記した。比較のために、トレーニングに2017年度データを使用した場合の2018年度 confusion matrix を表2に示す。

表1 2018年度 confusion matrix ( $p \geq 0.3$ ) の場合 (トレーニングに2017年度データ使用)

LCT#1-#4		成功	失敗	合計
	成功	831	204	1035
観測結果	失敗	107	88	195
	合計	938	292	1230

誤分類率：0.253、不合格の的中率：30.1%

LCT#1-#7		成功	失敗	合計
	成功	826	209	1035
観測結果	失敗	80	115	195
	合計	906	324	1230

誤分類率：0.235、不合格の的中率：35.5%

LCT#1-#11		成功	失敗	合計
	成功	803	232	1035
観測結果	失敗	51	144	195
	合計	854	376	1230

誤分類率：0.230、不合格の的中率：38.3%

表2 2018年度 confusion matrix ( $p \geq 0.4$ ) の場合  
(トレーニングに2017年度データ使用)

LCT#1-#4		予測結果		
		成功	失敗	合計
	成功	908	127	1035
観測結果	失敗	134	81	195
	合計	1042	208	1230

誤分類率：0.209、不合格の的中率：38.9%

LCT#1-#7		予測結果		
		成功	失敗	合計
	成功	960	75	1035
観測結果	失敗	142	53	195
	合計	1102	128	1230

誤分類率：0.176、不合格の的中率：41.4%

LCT#1-#11		予測結果		
		成功	失敗	合計
	成功	963	72	1035
観測結果	失敗	115	80	195
	合計	1078	152	1230

誤分類率：0.152、不合格の的中率：52.6%

表3 2018年度 confusion matrix ( $p \geq 0.5$ ) の場合  
(トレーニングに2017年度データ使用)

LCT#1-#4		予測結果		
		成功	失敗	合計
	成功	970	65	1035
観測結果	失敗	157	38	195
	合計	1127	103	1230

誤分類率：0.180、不合格の的中率：36.9%

LCT#1-#7		予測結果		
		成功	失敗	合計
	成功	1003	32	1035
観測結果	失敗	167	28	195
	合計	1170	60	1230

誤分類率：0.162、不合格の的中率：46.7%

LCT#1-#11		予測結果		
		成功	失敗	合計
	成功	991	44	1035
観測結果	失敗	144	51	195
	合計	1135	95	1230

誤分類率：0.153、不合格の的中率：53.7%

表4 2017年度 confusion matrix ( $p \geq 0.3$ ) の場合  
(トレーニングに2017年度データ使用)

LCT#1-#4		予測結果		
		成功	失敗	合計
	成功	728	193	921
観測結果	失敗	107	99	206
	合計	835	232	1127

誤分類率：0.266、不合格の的中率：33.9%

LCT#1-#7		予測結果		
		成功	失敗	合計
	成功	702	219	921
観測結果	失敗	96	110	206
	合計	798	329	1127

誤分類率：0.280、不合格の的中率：33.4%

LCT#1-#11		予測結果		
		成功	失敗	合計
	成功	710	211	921
観測結果	失敗	73	133	206
	合計	932	344	1127

誤分類率：0.252、不合格の的中率：38.7%

表5 2017年度 confusion matrix ( $p \geq 0.4$ ) の場合  
(トレーニングに2017年度データ使用)

LCT#1-#4		予測結果		
		成功	失敗	合計
	成功	826	95	921
観測結果	失敗	143	63	206
	合計	969	158	1127

誤分類率：0.211、不合格の的中率：39.9%

LCT#1-#7		予測結果		
		成功	失敗	合計
	成功	817	104	921
観測結果	失敗	137	69	206
	合計	954	173	1127

誤分類率：0.214、不合格の的中率：39.9%

LCT#1-#11		予測結果		
		成功	失敗	合計
	成功	821	100	921
観測結果	失敗	111	95	206
	合計	932	195	1127

誤分類率：0.187、不合格の的中率：48.7%

表6 2017年度 confusion matrix ( $p \geq 0.5$ ) の場合  
(トレーニングに2017年度データ使用)

LCT#1-#4			予測結果	
		成功	失敗	合計
	成功	872	49	921
観測結果	失敗	162	44	206
	合計	1034	93	1127

誤分類率：0.187、不合格の的中率：47.3%

LCT#1-#7			予測結果	
		成功	失敗	合計
	成功	878	43	921
観測結果	失敗	159	47	206
	合計	1037	90	1127

誤分類率：0.179、不合格の的中率：52.2%

LCT#1-#11			予測結果	
		成功	失敗	合計
	成功	865	56	921
観測結果	失敗	138	68	206
	合計	932	124	1127

誤分類率：0.172、不合格の的中率：54.8%

表1-3を表4-6と比較すると、誤分類率、不合格の的中率ともに表1-3の結果は表4-6の結果とそれほど変わらない。例えば、 $p \geq 0.4$ の場合、学期の中間付近で期末の結果を予測するのであればLCT#1-#7の結果を使って予測することになるが、その場合、トレーニングに2017年度データ使用して2018年度の予測を行ったとき、誤分類率：0.176、不合格の的中率：41.4%、トレーニングに2017年度データ使用して2017年度の予測を行ったとき、誤分類率：0.214、不合格の的中率：39.9%となっている。過学習のことを考えれば、先に提案した方法は、過去のLCTのデータを教師にして現在のLCTの途中のデータから期末の結果を予測する今回の試みから、十分に機能することが検証されたと言える。

#### 4. まとめ

先に、LCTを使ってドロップアウトリスクを抱えた学生を早期に発見しアラートを発する方法の提案を行ったが、そこでは、予測に用いたLCTのデータと検証に用いたデータは同じ2017年度前期のものであったため、予測精度への信頼性に疑問が残るものであった。つまり、トレーニングデータとテストデータを同じにしていたようなものである。そのときにはまだ、2018年度のデータを使うことができなかつたからである。したがって、先の予測結果は甘い方向になされているかもしれないとも考えられる。そこで、今回は、2018年度のデータが使える状態になったた

め、2017年度LCTデータから作られたモデルを2018ケースに適用して予測を行なう検証を行なった。

検証の結果、2017年度のLCTから得られたモデルを使って2018年度の期末試験の可否の確率を予測するときの、不合格予測の的中率と confusion matrix の値は、2017年度のLCTのモデルを使って2017年度の期末試験の可否の確率を予測するときの結果と変わらないことがわかった。したがって、前回提案した方法は有効に機能することが示された。

#### 文 献

- 1) 廣瀬、ラーニングアナリティクス：LCT成績と期末試験成績の関係、広島工業大学紀要教育編、pp. 59-63, Vol. 18, 2019.
- 2) 廣瀬、大規模オンラインテストから得られるラーニングアナリティクス、広島工業大学紀要研究編、pp. 159-166, Vol. 53, 2019.
- 3) 廣瀬、新入生全員を対象としたオンラインテストの実際、広島工業大学紀要教育編、pp. 27-35, Vol. 16, 2017.
- 4) 廣瀬、フォローアップクラスにおける授業設計について、広島工業大学紀要教育編、pp. 37-41, Vol. 16, 2017.
- 5) 廣瀬、大規模授業支援テストシステムとそのラーニングアナリティクス、統計数理、Vol. 66, No. 1, pp. 79-96, 2018.
- 6) 廣瀬、ラーニングアナリティクス指向学習支援、コンピュータ&エデュケーション (CIEC)、Vol. 45, pp. 23-30, 2018.
- 7) Hideo Hirose, Meticulous Learning Follow-up Systems for Undergraduate Students Using the Online Item Response Theory, 5th International Conference on Learning Technologies and Learning Environments (LTLE2016), pp. 427-432, 2016.
- 8) Hideo Hirose, Difference Between Successful and Failed Students Learned from Analytics of Weekly Learning Check Testing, Information Engineering Express, Vol 4, No 1, pp. 11-21, 2018.
- 9) Hideo Hirose, Prediction of Success or Failure for Examination using Nearest Neighbor Method to the Trend of Weekly Online Testing, International Journal of Learning Technologies and Learning Environments (IJLTLE), Vol 2, No 1, pp. 19-34, May 31, 2019.
- 10) Hideo Hirose, Key Factor Not to Drop Out is to Attend Lectures, Information Engineering Express, Vol 5, No 1, pp. 11-21, May 31, 2019.