

非同期オンライン授業における学習履歴をもとにした 先延ばし行動を予測するための特徴ベクトルの検討

Consideration for Characteristic Vectors to Predict the Procrastinating Learners based on Learning Logs in Online Class

久保田 真一郎^{*1}

Shin-Ichiro KUBOTA^{*1}

^{*1}熊本大学

^{*1}Kumamoto University

Email: kubota@cc.kumamoto-u.ac.jp

あらまし：課題の提出締切までに LMS へ課題を提出する必要がある場合で、先延ばし行動などで締切までに提出できない学習者に対して、効果的なリマインドができないか検討を進めている。本稿では、課題提出に限らず、授業テキストが LMS に設置されており、その授業テキストに課題の説明が記述されている場合を想定し、毎週の課題提出が設定されている授業において、テキストへのアクセスの仕方を前の週と比較して先延ばし行動を予測するための特徴ベクトルの検討を行った。検討した特徴ベクトルについてナイーブベイズによる予測結果を参照しつつ、特徴ベクトルの要素について議論した内容について報告する。

キーワード：ラーニングアナリティクス、先延ばし行動、特徴ベクトル、学習履歴

1. はじめに

課題の提出締切までに LMS(Learning Management System)へ課題を提出する必要がある場合で、先延ばし行動などで締切までに提出できない学習者に対して、効果的なリマインドができないか検討を進めている。藤田は先延ばし行動と失敗行動との関連性について調査しており、失敗行動との関連性を示唆している^①。本研究では学習履歴から学習者の行動を表す値を算出し、それらを要素とする特徴ベクトルを使い、先延ばし行動によるリスクのある学習者を見つける方法を検討する。

2. 研究方法

研究の対象としたのは著者を含む 7 名の教員で運営される受講者約 1800 名の情報リテラシーを学ぶ授業である。この授業では、教室に集合することなく、学習者がパソコンまたはタブレットなどを用いて授業テキストを閲覧して学習を進める。授業テキストは LMS で公開されており、学習者は Web ブラウザを使って閲覧する。さらに、授業テキストには学習内容の説明および毎週の提出課題に関する説明がある。学習者はその説明を読みながら課題を作成し、決められた締切までに毎週の課題を提出する。LMS には授業テキストへのアクセスが記録されており、本研究で扱う LMS である Moodle では、標準ログとして学習者が授業テキストを閲覧した履歴に限らず、授業テキストのページに相当する chapter 属性の id 番号(以降、chapterid と記す)が履歴として記録される。

2020 年度前期に行われた授業において、締切後に未提出者を確認し、未提出者宛に担当教員へ相談するよう通知を行い、相談があった学生に対して担

当教員が学習計画などの指導を行った。指導を行った学生の記録をもとに、先延ばし行動などで締切までに提出できなかった学習者にラベルをつけることができる。

2020 年度前期に行った授業を対象として、先延ばし行動などで締切までに提出できなかった学習者の授業テキストへのアクセス情報をもとに、締切までに提出できない学習者を予測することを目指して、予測に必要な特徴ベクトルを検討する。本研究では、締切日の 9 時までの学習履歴をもとに予測して該当者に通知することを想定しているため、学習している週(w 週)とその前の週(w-1 週)の授業テキストへのアクセス情報を利用して予測する。前の週のアクセス情報を利用する理由は、先延ばし行動により現在学習している週とその前の週とで行動が変化すると予想したためである。行動の変化という意味では、前の週に限らず、第 1 週からの変化を入力として予測した方が、その変化をより明確に検知できる可能性がある。この点については今後検討を進めたい。

3. 特徴ベクトル

前節で述べたように特徴ベクトルは行動の変化を表すベクトルとなるように、w 週と w-1 週との学習者の行動の違いを含むような要素を考える必要がある。表 1 に、検討した 9 個の要素について説明する。先延ばし行動をする可能性のある個人を特定するために、学習者につき 1 つの特徴ベクトルが対応する。なお、第 5 要素から第 8 要素は、第 1 要素から第 4 要素の値を w-1 週について集計した値となっている。

第 1 要素は開始時間の値で先延ばし行動などで取り掛かりが遅いような場合に値が大きくなる。第 2 要素は締切日時に近い学習活動になると大きくなる。

第3要素は、学習する速さに相当する値を扱う。授業テキストのページを表す `chapterid` はページが進むにつれて大きくなるように整理しているので、最大値($\max(\text{chapterid})$)と最小値($\min(\text{chapterid})$)およびそれにアクセスした時間の最大値($\max(t)$)と最小値($\min(t)$)から速さに相当する値とした。第4要素は、単位時間あたりのアクセス回数とした。第9要素は w 週と $w-1$ 週とのページ遷移の仕方の変化を扱うために、 w 週と $w-1$ 週の `chapterid` の値を時系列に並べた数列に対して DTW(Data Time Warping) ⁽²⁾ により算出される類似度を用いた。

表1 特徴ベクトルの各要素

要素	説明
1	学習している週(w 週)のうち初めて授業テキストにアクセスした時間から授業テキスト公開の日時を引いた値を w 週の授業テキスト公開から予測に利用する金曜日 9 時までの時間(以降、予測対象時間)で正規化した値
2	w 週のうち予測対象時間の間で最後に授業テキストにアクセスした時間から授業テキスト公開の日時を引いた値を予測対象時間で正規化した値
3	w 週に学習者が <code>chapterid</code> の最大値から最小値を引き、テキストにアクセスしていた時間で除した値($\max(\text{chapterid}) - \min(\text{chapterid}) / (\max(t) - \min(t))$)を学習者全体のうちの最大値を 1、最小値を 0 として正規化した値
4	w 週に授業ページへアクセスした回数をテキストにアクセスしていた時間($\max(t) - \min(t)$)で除した値を算出し、対象としている学習者全体のうちの最大値を 1、最小値を 0 として正規化した値
5 から 8	$w-1$ 週の第1要素から第5要素の値
9	DTW によって $w-1$ 週のページ遷移(<code>chapterid</code> を時系列に並べた数列)と w 週のページ遷移の類似度を算出し、対象としている学習者全体のうちの最大値を 1、最小値を 0 として正規化した値

4. 評価と考察

今回の取り組みでは、先延ばし行動を行う可能性のある学習者を 2 週間の学習履歴(例えば、第 6 週に予測するなら第 5 週と第 6 週の学習履歴)を使って予測する。その際、トレーニングデータはさらに 1 週間前に先延ばし行動などで締切までに提出できなかった学習者の 2 週間の学習履歴(例えば、第 6 週に予測するなら、第 5 週の締切までに提出できなかった学習者の第 4 週と第 5 週の学習履歴)になる。特徴ベクトルを評価するために、2020 年度第 5 週の締切までに提出できなかった学習者 7 名とそれと同数の締切に間に合った学習者の第 4 週と第 5 週の学習履歴から算出される特徴ベクトルをトレーニン

グデータとして、2020 年度第 6 週において締切に間に合わなかった学習者 26 名を予測できるか評価を行った。なお、テストデータは第 6 週の締切までに提出できなかった学習者(26 名)と第 6 週の締切までに提出できた学習者(68 名)の学習履歴から算出した。

すべての要素を使い、ナイーブベイズ分類器を用いて分類した結果を表 2 に示す。`at-risk` が締切までに提出できなかった学習者を表し、`not-at-risk` が締切に間に合った学習者を表している。縦に実際の学習者、横に分類器によって分類された結果を示している。正解率は 0.66 で、精度は 0.45、適合率は 0.96、F 値は 0.61 であった。

表2 特徴ベクトルのすべての要素を用いた結果

	at-risk(positive)	not-at-risk(negative)
at-risk	25	1
not-at-risk	31	37

すべての要素を用いず、要素を第2要素、第4要素、第5要素、第6要素にして同様に分類すると、正解率は 0.74 で、精度は 0.51、適合率は 1.00、F 値は 0.68 と改善が見られた。

5. まとめ

課題の提出締切までに LMS へ課題を提出する必要がある場合で、先延ばし行動などで締切までに提出できない学習者を支援するために、先延ばし行動を予測するための特徴ベクトルを検討した。すべての要素を用いた場合で 6 割程度の精度で、適合率が 0.96 とリスクのある学習者はほぼもれなく予測できそうであるが、リスクのない学生の多くを誤って推定しているため、方法の改善が必要である。すべての要素を用いず特定の要素を選択して同様に予測した結果で少し改善が見られたことから、特徴ベクトルの構成の再検討が必要と考えられる。また、今回は第 6 週を対象としたが、他の週の結果も参照し、過去週の特徴ベクトルを使いトレーニングすることで改善する可能性も考えられる。DTW は時系列データの類似性を表すがピークの位置のずれなどは数値として反映されない。先延ばし行動によりピークの時間は変化すると予想されるため、アクセスのピークを表すような要素の採用を検討したい。

謝辞

この研究の一部は JSPS 科研費 JP19K12273 の助成を受けたものです

参考文献

- (1) 藤田 正:“先延ばし行動と失敗行動の関連について”, 教育実践総合センター研究紀要,14, pp.43-46 (2005)
- (2) Berndt, D.J. and Clifford, J.: Finding Patterns in Time Series: A Dynamic Programming Approach, Advances in Knowledge Discovery and Data Mining, pp.229-248, AAAI/MIT(1996).