

# 差分パターンマイニングによる重みづけを利用した fastText 特徴による成績予測

舘 良太<sup>1,a)</sup> 平川 翼<sup>1,b)</sup> 山下 隆義<sup>1,c)</sup> 藤吉 弘亘<sup>1,d)</sup>

**概要：**教育教材のデジタル化に伴い、学習ログデータを用いた成績の予測が期待されている。従来研究では、デジタル教材に対する操作回数ヒストグラム特徴を機械学習モデルに入力することで成績予測を行っている。しかし、この手法では時系列を考慮していないため、学習行動パターンを反映させた成績予測が困難である。そこで、我々は、学習行動を埋め込み特徴とした fastText を用いることで時系列を維持する。そして、差分パターンマイニングにより有意な差を持つ学習行動を検出し、その fastText に重みづけを行う。評価実験より、提案手法は従来手法よりも成績予測精度が向上し、学習行動の重みづけが有効であることが分かった。

## Grade Prediction by fastText Features Weighted by Difference Pattern Mining

### 1. はじめに

近年、急速なペースで進展しているデジタル化により、学生の操作ログデータや成績に関するデータを大規模に蓄積することが可能となった。これらを分析することで、学習者の学習進捗状況、学習者の特性、学習傾向などを包括的に把握することが可能である。これにより教育機関は、一般的な学習サポートの方法に加えて、従来では困難であった各学習者の特性やニーズに合わせた効果的な学習のサポートを提供することが可能となる。そのため、教育分野における機械学習モデルの活用が期待されている。

特に、早期退学者の検出や学習行動の改善を目的として、学習行動から成績を予測するの研究は多く行われている。しかし、成績予測のための操作ログデータの入力方法については十分に議論されておらず、その扱いについては研究者によって様々である。従来手法では、特徴を生成する場合に、操作回数ヒストグラムに変換し、モデルへ入力して予測している。しかし、操作回数ヒストグラム特徴では時系列情報が損失してしまうため、学生の学習行動出現順序

などの行動パターンを考慮できないという問題が存在する。

そこで、本研究では学習行動パターンの特徴表現を実現することを目的とし、差分パターンマイニングによる重みづけを利用した fastText 特徴による成績予測手法を提案する。具体的には、学習行動を埋め込み特徴とした E2vec の前処理と埋め込みモジュールを用いることで時系列を維持する。そして、差分パターンマイニングにより有意な差を持つ学習行動を検出し、fastText に重み付けを行うことで、2 クラス間の差分を捉えた学習行動特徴生成を行う。評価実験では、操作ログのデータセットを使用し、成績予測の精度評価を行う。また、差分パターンマイニングを通じて、成績上位者と下位者にどのような差があるかの分析を行う。実験結果より、提案手法は従来手法よりも成績予測精度が向上し、学習行動の重み付けの有効性を示した。

### 2. 関連研究

過去の研究では、デジタル教材における操作ログデータから、学生の学習行動を表す特徴を様々な方法で抽出している。

Okubo ら [1] は、デジタル教材やその他の学習管理システムにおける学生の行動特徴として、アクティブラーナーポイント (ALP) を提案した。ALP では、マーカーやメモの数、出席情報などの学習活動は 0 から 5 の範囲で採点さ

<sup>1</sup> 中部大学

Chubu University

a) tatti@mprg.cs.chubu.ac.jp

b) hirakawa@mprg.cs.chubu.ac.jp

c) takayoshi@isc.chubu.ac.jp

d) fujiyoshi@isc.chubu.ac.jp

れた。このスコアは、成績予測タスクのための RNN モデルの学習に使用された。

Kohama ら [2] は、デジタル教材の操作ログデータにおける学生の行動特徴として、操作回数ヒストグラムを提案した。デジタル教材の操作ログデータから教材 ID と各操作種類における操作回数を抽出し、正規化をすることで操作回数ヒストグラムを生成している。これは学生の最終成績 (A,B,C,D,F) を予測するタスクのための Transformer Encoder モデルの学習に使用された。

Miyazaki ら [3] は、デジタル教材の操作ログデータ埋め込み手法 E2Vec を提案した。この研究では、操作ログデータ 1 つ 1 つを文字として扱い、次の操作までの時間間隔を間に挿入した文字列を単語として扱うことにより、fastText [4] による学習を行った。また、それらを Bag-of-Visual-Words アプローチにより集約し、学生の学習行動特徴を生成している。

Li ら [5] は、差分パターンマイニングを用いて、学生の学習から自己主導的な広範囲読書の行動シーケンスを、自己主導的な行動ログからマイニングした。さらに、従来の行動頻度特徴量と行動順序特徴量を用いて、学生の学業成績を予測するモデルを構築した。

これらの手法では、共通して操作の回数やパターンの回数を最終的な特徴としているため、時系列情報が損失してしまい、学習パターンを反映させた特徴生成ができていない。

### 3. 提案手法

前述の通り、従来研究では、時系列情報を考慮して特徴量の生成しておらず、学習行動ログの出現順序などの学習パターンを反映させた成績予測ができていない。そこで、本研究では、時系列順の操作ログデータを E2Vec の前処理、埋め込みモジュールを用いて、時系列を維持した操作ログ埋め込み特徴を生成し、差分パターンマイニングにより特定した 2 クラス間で有意な差を持つ学習パターンによる重みづけを行う。これにより、時系列操作ログ埋め込み特徴に学習パターンを反映させた成績予測手法を提案する。

提案手法の概要を図 1 に示す。提案手法は主に差分パターンマイニング、操作ログデータ埋め込み、分類予測の 3 つのモジュールで構成される。差分パターンマイニングにより、操作ログデータから成績上位者と成績下位者の 2 クラス間で有意な差を持つパターンを検出し、各差分パターンについて効果量の計算を行う。次に差分パターンによる重みづけを行った E2Vec の前処理と埋め込みモジュールを用いて操作ログ埋め込み特徴を生成する。生成した特徴を分類モデルへ入力し、学生の成績予測を行う。

以下では、差分パターンマイニングおよび操作ログデータ埋め込みの詳細について説明する。

### Algorithm 1 Depth-First Search with Pruning (DFS-Pruning)

```
1: procedure DFS-PRUNING( $n, S_n, I_n$ )
2:    $S_{temp} \leftarrow \emptyset$ 
3:    $I_{temp} \leftarrow \emptyset$ 
4:   for  $i \in S_n$  do
5:     if  $(s_1, \dots, s_k, \{i\})$  is frequent then
6:        $S_{temp} \leftarrow S_{temp} \cup \{i\}$ 
7:     end if
8:   end for
9:   for  $i \in S_{temp}$  do
10:    DFS-Pruning( $(s_1, \dots, s_k, \{i\}), S_{temp}$ , all elements in
     $S_{temp}$  greater than  $i$ )
11:  end for
12:  for  $i \in I_n$  do
13:    if  $(s_1, \dots, s_k \cup \{i\})$  is frequent then
14:       $I_{temp} \leftarrow I_{temp} \cup \{i\}$ 
15:    end if
16:  end for
17:  for  $i \in I_{temp}$  do
18:    DFS-Pruning( $(s_1, \dots, s_k \cup \{i\}), S_{temp}$ , all elements in
     $I_{temp}$  greater than  $i$ )
19:  end for
20: end procedure
```

#### 3.1 差分パターンマイニング

差分パターンマイニング [6] とは、SPAM アルゴリズム [7] を用いて系列データから頻出パターンを抽出し、統計仮説検定の一種である t 検定を用いて、2 クラス間で有意な差を持つ、差分パターンを特定するマイニング方法である。

##### 3.1.1 SPAM アルゴリズム

SPAM アルゴリズムとは、時系列データから時系列パターンの検出を行うシーケンシャルパターンマイニングである。時系列データをビットマップ表現し、最小サポート率を満たす頻出アイテムを特定する。頻出アイテムを基に S-step や I-step を行い、深さ優先探索方式によるパターン拡張を行う。以下に深さ優先探索の流れを示す。

本研究では、オープンソースである SPMF [8] ライブラリ内の SPAM アルゴリズムを用いて頻出パターンの検出を行い、ある頻出パターンの各学生における出現回数に基づき、t 検定を行うことで差分パターンを特定した。

##### 3.1.2 t 検定

t 検定とは、2 つの集団やサンプルの平均値が統計的に有意な差があるかどうかを判断するための統計手法である。各サンプルの平均値や標準偏差を用いて、平均値の差が偶然によるものか、それとも有意な差であるかを判断するための指標である t 値の計算を行う。t 値の計算式を式 1 に示す。次に、t 値と自由度を基に、帰無仮説が正しいと仮定した場合に、観測されたデータが得られる確率である p 値を求める。p 値が有意水準よりも小さい場合、帰無仮説を棄却し、対立仮説を支持する。つまり、これは平均値の差が統計的に有意であり、偶然の誤差ではないと結論づけ

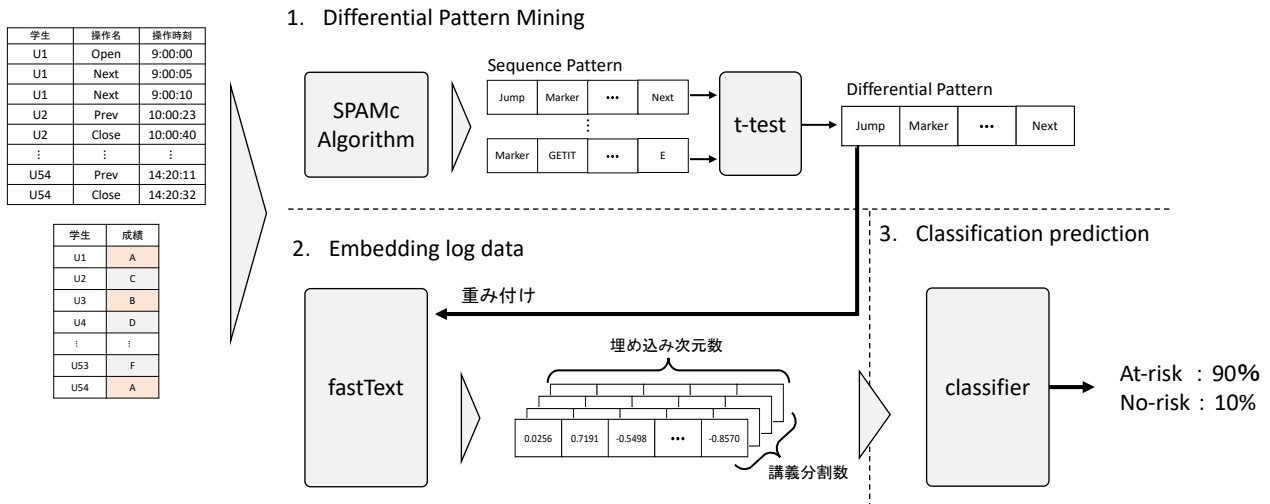


図 1 提案手法の概要

ることを意味する。

本研究では、使用するデータに 2 クラス間で学生数に差があり、かつ平均の差が正の場合と負の場合の両方に対して検定を行うため、t 検定の 1 種である対応なしウエルチ検定の両側検定を用いて検定を行う。ここで、 $n_1$  は成績上位の学生数、 $n_2$  は成績下位の学生数、 $s_1^2$  は成績上位者の不偏分散、 $s_2^2$  は成績下位者の不偏分散、 $\bar{x}_1$  は成績上位者のサンプル平均、 $\bar{x}_2$  は成績下位者のサンプル平均である。サンプルは各学生におけるパターンの出現回数である。

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (1)$$

### 3.2 操作ログデータ埋め込み

本研究では、時系列に並んだ操作ログデータを文章として扱い、E2Vec の前処理、埋め込みアプローチによって分散表現へと埋め込む。

#### 3.2.1 前処理

自然言語処理では、1 つの単語は複数の文字から構成され、1 つの文は複数の単語から構成され、複数の文は文章を形成する。E2vec は、fastText を操作ログデータに適用するために、自然言語処理の “character”, “word”, “sentence” に対応するプリミティブ、ユニット、アクションをそれぞれ定義した。操作ログとプリミティブシンボルとの対応関係を表 1 に示す。前処理のルールは以下の通りである。

- (1) 操作ログデータを学生 ID と教材 ID で分割する。したがって、教材に対する学生の操作からユニットが作成される。
- (2) 各操作ログデータをプリミティブに変換する。
- (3) 2 つのプリミティブの間に、操作間隔を表したプリミティブである “s”, “m”, “l” を挿入する。
- (4) 1 分間の操作かつ 15 プリミティブ以内でユニットを生

表 1 操作ログとプリミティブシンボルとの対応関係

操作名	説明	プリミティブ
NEXT	次のページへ移動	N
PREV	前のページへ移動	P
OPEN	教材を開く	O
ADD MARKER	マーカーを引く	A
CLOSE	教材を閉じる	C
PAGE JUMP	指定したページへ移動	J
GET IT	ページ内容について理解した	G
OTHERS	低頻出の操作	E
short interval	1 から 10 秒の時間間隔	s
medium interval	10 から 300 秒の時間間隔	m
long interval	300 秒以上の時間間隔	l

成する。

- (5) 次の操作までの間隔が 5 分以上の場合、アクションを分離する。

#### 3.2.2 埋め込み

文を埋め込む際、単語をそれぞれ埋め込み、分散表現の平均を行うことが一般的であるが、本研究では、特定した差分パターンが 2 クラス間でどの程度の差があるかを定量的に評価できる効果量をもとに加重平均を求める。このように重みを付与して文の分散表現を獲得する。効果量には “Cohen の d” を用いる。求めた効果量は Min-Max スケーリングにより正規化し、重みとして用いる。 “Cohen の d” の計算式を式 2 に示す。ここで、 $n_1$  は成績上位の学生数、 $n_2$  は成績下位の学生数、 $s_1^2$  は成績上位者の不偏分散、 $s_2^2$  は成績下位者の不偏分散、 $\bar{x}_1$  は成績上位者のサンプル平均、 $\bar{x}_2$  は成績下位者のサンプル平均である。サンプルは各学生におけるパターンの出現回数である。

$$d = \frac{\bar{x}_1 - \bar{x}_2}{S_p} \quad (2)$$

表 2 6 コースに関する情報

Course	学生数	操作回数	講義回数	講義時間
A-2020	60	142,754	7	200
A-2021	54	130,330	8	200
A-2022	52	197,389	8	200
D-2020	105	272,749	14	90
D-2021	106	288,216	15	90
D-2022	93	287,088	16	90

表 3 成績の分布

Course	A	B	C	D	F	No-risk	At-risk
A-2020	22	23	5	3	7	45	15
A-2021	9	11	10	18	6	20	34
A-2022	17	6	5	22	2	23	29
D-2020	61	7	1	2	34	68	37
D-2021	60	3	6	4	33	63	43
D-2022	50	10	8	8	17	60	33

$$S_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \quad (3)$$

## 4. 実験

本実験では、提案手法により生成した、差分パターンマイニングによる重みづけを利用した fastText 特徴を入力とし、分類器による成績の分類予測を行う。

比較実験として、関連研究において、操作ログデータのみを用いた特徴生成を行っている Kohama らと Miyazaki らの手法との比較を行う。

### 4.1 データセット

本研究では、BookRoll [9], [10] と呼ばれるデジタル教材システムを用いて収集された、6 コースの学生の操作ログデータと学生の最終成績情報を使用する。表 2 に各コースの詳細を示す。A, D は担当した教員を表し、2020 年, 2021 年, 2022 年はコースが開催された年を表す。これらのコースは九州大学のコンピュータサイエンスを受講した学生を対象に行われた。成績は A, B, C, D, F の 5 段階で評価されており、成績 F は単位不認定を意味する。表 3 に成績の分布を示す。本研究では、成績 A, B を No-risk, 成績 C, D, F を At-risk とした 2 クラスに統合したデータを使用する。また、A-2020 と D-2020 のデータは差分パターンの特定と効果量の計算にのみ使用し、その他のコースは分類予測タスクに使用する。

### 4.2 実験条件

本研究で利用する SPMF ライブラリの SPAM アルゴリズムは、パラメータとして、パターンの最小の長さ、最大の長さ、最小サポート率、最大ギャップを設定することが可能である。E2vec 前処理ルールとしてユニットの最大の長さが 15 までと定義されている。差分パターンとユニッ

トを対応させるため、パターンの長さの最小値は 1、最大値は 15 とする。また、操作ログデータを扱う上で、学生の学習活動の根底にはパターンがあるものの、学生が必ずしもパターンに従っているとは限らないという点がある。その点を考慮するため、Li らの手法に習い、最小サポート率は 40 %、最大ギャップ数は 2 とする。使用するコースは、A-2021, A-2022, D-2021, D-2022 であり、A-2020 と D-2020 は差分パターンマイニングと効果量の計算にのみ使用される。分類モデルとして、RandomForest, SVM, XGBoost を用いる。分類モデルのハイパーパラメータを調整するために、訓練データを用いた 3 重クロスバリデーションによる GridSearch を用いる。スコアには F1-Score を用いる。評価には最もスコアが高いハイパーパラメータを用いる。

Kohama らの手法では、成績 A, B, C, D, F の 5 クラス分類を行っているが、条件を揃えるため、評価上で At-risk, No-risk の 2 クラスへの統合を行う。分類モデルには Transformer Encoder を用いる。Miyazaki らの手法では、使用する E2Vec モデルとして、E2Vec<sub>k100</sub> を用いる。分類モデルには Random Forest を用いる。本実験、比較実験ともに、評価指標として F1-Score を用いる。

### 4.3 実験結果

以下では訓練データと評価データのセットを (Train, Test) と表現する。各手法とモデルにおける分類精度の比較を表 4 に示す。表 4 より、モデル RandomForest を用いたとき、データセット (A-2021, D-2021), (D-2021, A-2022), (D-2022, A-2022), (D-2022, D-2021) において従来手法より高い分類精度が得られた。また、データセット (D-2022, D-2021) を用いた場合に、本実験において、一番高い精度である 0.8927 が得られた。モデルに XGBoost を用いたとき、データセット (A-2021, A-2022), (A-2021, D-2021) において従来手法より高い分類精度が得られた。モデルに SVM を用いたとき、データセット (A-2021, D-2021), (A-2022, D-2021), (A-2022, D-2022), (D-2021, A-2022), (D-2022, A-2022) において従来手法より高い分類精度が得られた。

訓練データごとの平均精度比較を表 5 に示す。表 5 より、提案手法と従来手法を比較すると、訓練データ A-2021 において 0.0228pt, D-2021 において 0.0387pt, D-2022 において 0.0515pt の精度向上が確認できた。以上のことから、差分パターンを利用した学習行動の重み付けの有効性を確認した。

### 4.4 分析

本研究において、差分パターンマイニングを行った結果、差分パターンが 7,785 個特定された。本節では、その中からどのような差分パターンが 2 クラス間で差を持つかの分

表 4 分類精度の比較

Train	Test	F1-score				
		Histogram	E2Vec	Proposed <sub>RF</sub>	Proposed <sub>XGB</sub>	Proposed <sub>SVM</sub>
A-2021	A-2022	0.5486	0.6756	0.5825	<b>0.6862</b>	0.5778
	D-2021	0.5891	0.6029	0.7537	0.7144	<b>0.7965</b>
	D-2022	0.5350	<b>0.6019</b>	0.5232	0.5482	0.5739
A-2022	A-2021	0.5559	<b>0.7088</b>	0.4139	0.4035	0.3636
	D-2021	0.5604	0.7959	0.6364	0.5920	<b>0.7980</b>
	D-2022	0.5062	0.6666	0.5553	0.5662	<b>0.7062</b>
D-2021	A-2021	<b>0.5867</b>	0.5000	0.3704	0.4292	0.3755
	A-2022	0.5438	0.2941	<b>0.6181</b>	0.5143	<b>0.6181</b>
	D-2022	0.3841	<b>0.6666</b>	0.5934	0.6123	0.6371
D-2022	A-2021	0.2916	<b>0.5600</b>	0.4899	0.5046	0.5046
	A-2022	0.5744	0.4324	<b>0.6065</b>	0.5685	0.5876
	D-2021	0.3727	0.8421	<b>0.8927</b>	0.7898	0.7547

表 5 訓練データごとの平均精度比較

Train	Histogram	E2Vec	Proposed <sub>RF</sub>	Proposed <sub>XGB</sub>	Proposed <sub>SVM</sub>
A-2021	0.5575	0.6268	0.6198	<b>0.6496</b>	0.6494
A-2022	0.5408	<b>0.7237</b>	0.5352	0.5205	0.6226
D-2021	0.5048	0.4869	0.5273	0.5186	<b>0.5435</b>
D-2022	0.4129	0.6115	<b>0.6630</b>	0.6209	0.6156

析を行う。効果量の値が大きかった差分パターン 10 個を表 6 に示す。また、効果量の値が小さかった差分パターン 10 個を表 7 に示す。表 6, 表 7 より、効果量が大きい差分パターンはほとんど“NEXT”という操作で構成されており、効果量が小さい差分パターンはほとんど“NEXT”と“PREV”という操作で構成されていることが確認できた。

次に、表 8 に差分パターンを構成する操作の出現回数を示す。また、図 2 に差分パターン作成に用いたデータセットによる各操作回数と成績の関係性を示す。図 3 に拡大した各操作回数と成績の関係性を示す。表 8 より、本実験において特定した差分パターンには、“ADD MARKER”や“OTHER”といった特徴的な操作が含まれていないことが確認できた。これは、図 2 より、特徴的な操作の回数は、No-risk, At-risk とともに“NEXT”と比較すると極端に少ないため、2 クラス間で帰無仮説を棄却できるほどの十分な差がなかったからであると考えられる。図 3 より、特徴的な操作は At-risk の学生にも現れていることが確認できる。そのため、本研究で使用した差分パターンマイニングにおいて、操作回数が少ないが特徴的な操作パターンの特定ができていないことがわかった。また、表 6 のようになる理由として、差分パターンを特定する際の t 検定において、あるパターンが各学生に何回出現するかの回数を値として使用しているため、図 2 より、2 クラス間で操作回数に特徴

表 6 効果量の大きい差分パターン

pattern	effect size
N	0.7397
NN	0.7341
NNN	0.7255
NNNN	0.7109
NNNNN	0.6577
NNNNNN	0.6707
NNNNNNN	0.6453
NNNNNNNN	0.6160
NNNNNNNNN	0.5813
NNNNNNNNNN	0.5416

表 7 効果量の小さい差分パターン

pattern	effect size
NNNNNPPPPPPPPNP	0.0200
PNNPPNNNNPNNNN	0.0376
PNNNNPPNNPNPP	0.0403
NNNNNNPPNPNNPNP	0.0406
PPNPNNNNPNPPPN	0.0423
NPNNNNNNPPPPPPP	0.0443
NNNNNPPNPNNPNP	0.0446
NPPPNPNPNPNP	0.0459
PPNPNNPNNNPNP	0.0477
NNNNPPPPPPPPNP	0.0504

的な差がある場合、その差に強く影響されるからであると考えられる。

## 5. おわりに

本研究では、関連研究の問題点である時系列情報を考慮すべく、E2Vec の前処理、埋め込みモジュールを用いた時系列操作ログ埋め込み特徴を生成し、差分パターンマイニングを利用した重み付けを行った。評価実験により、本手法は従来手法と比べ、成績予測精度の向上が確認できた。また、分析により、本実験により特定された差分パターン

表 8 差分パターンを構成した操作の出現回数

NEXT	PREV	ADD MARKER	OPEN	CLOSE	PAGE JUMP	GET IT	OTHER
59,759	33,006	0	387	218	0	0	1

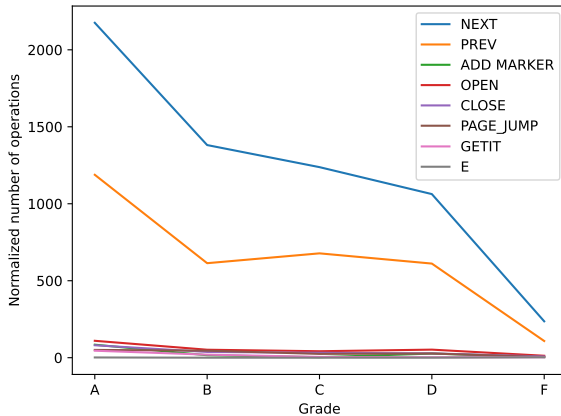


図 2 各操作回数と成績の関係

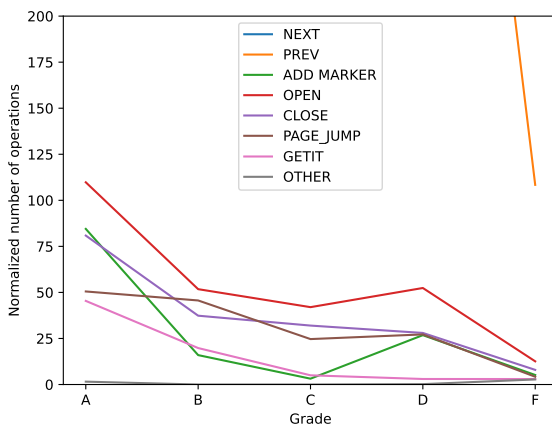


図 3 200 回以下の各操作回数と成績の関係

はほとんど“NEXT”と“PREV”によって構成されていることがわかった。これは、使用したデータの操作回数に2クラス間で特徴的な差を持っており、それに強く影響したためだと考えた。

今後の研究では、“ADD MARKER”や“OTHER”といった特徴的な操作に焦点をあて、学生固有の行動プロセスを捉えた特徴の抽出が必要であると考えられる。

## 6. 謝辞

本稿は、JST CREST グラント番号 JPMJCR22D1 により支援されました。

## 参考文献

[1] Okubo, F., Yamashita, T., Shimada, A. and Konomi, S.: Students' performance prediction using data of multiple

courses by recurrent neural network, *Proceedings of the 25th International Conference on Computers in Education, ICCE 2017 - Main Conference Proceedings*, pp. 439–444 (2017).

- [2] Kohama, H., Ban, Y., Hirakawa, T., Yamashita, T., Fujiyoshi, H. and Akitoshi Itai, H. U.: Recommending Learning Actions Using Neural Network, *International Conference on Computers in Education 2023* (2023).
- [3] Goren, O., Cohen, L. and Rubinstein, A.: Early Prediction of Student Dropout in Higher Education using Machine Learning Models, *Proceedings of the 17th International Conference on Educational Data Mining*, Atlanta, Georgia, USA, International Educational Data Mining Society, pp. 349–359 (online), DOI: 10.5281/zenodo.12729834 (2024).
- [4] Bojanowski, P., Grave, E., Joulin, A. and Mikolov, T.: Enriching Word Vectors with Subword Information (2017).
- [5] Li, J., Li, H., Majumdar, R., Yang, Y. and Ogata, H.: Self-directed Extensive Reading Supported with GOAL System: Mining Sequential Patterns of Learning Behavior and Predicting Academic Performance, LAK22, New York, NY, USA, Association for Computing Machinery, pp. 472–477 (online), DOI: 10.1145/3506860.3506889 (2022).
- [6] Kinnebrew, J. S., Loretz, K. M. and Biswas, G.: A Contextualized, Differential Sequence Mining Method to Derive Students' Learning Behavior Patterns, *Journal of Educational Data Mining*, Vol. 5, No. 1, pp. 190–219 (online), DOI: 10.5281/zenodo.3554617 (2013).
- [7] Ayres, J., Flannick, J., Gehrke, J. and Yiu, T.: Sequential Pattern mining using a bitmap representation, *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '02, New York, NY, USA, Association for Computing Machinery, p. 429–435 (online), DOI: 10.1145/775047.775109 (2002).
- [8] Fournier-Viger, P., Lin, J. C.-W., Gomariz, A., Gueniche, T., Soltani, A., Deng, Z. and Lam, H. T.: The SPMF Open-Source Data Mining Library Version 2, *Machine Learning and Knowledge Discovery in Databases* (Berendt, B., Bringmann, B., Fromont, É., Garriga, G., Miettinen, P., Tatti, N. and Tresp, V., eds.), Cham, Springer International Publishing, pp. 36–40 (2016).
- [9] Ogata, H., Yin, C., Oi, M., Okubo, F., Shimada, A., Kojima, K. and Yamada, M.: e-Book-based learning analytics in University education, *Doctoral Student Consortium (DSC) - Proceedings of the 23rd International Conference on Computers in Education, ICCE 2015*, Asia-Pacific Society for Computers in Education, pp. 401–406 (2015).
- [10] Flanagan, B. and Ogata, H.: Learning Analytics Platform in Higher Education in Japan, *Knowledge Management and E-Learning*, Vol. 10, pp. 469–484 (2018).

## 正誤表

下記の箇所に誤りがございました。お詫びして訂正いたします。

訂正箇所	誤	正
6 ページ 参考文献	[3] Goren, O., Cohen, L. and Rubinstein, A.: Early Prediction of Student Dropout in Higher Education using Machine Learning Models, <i>Proceedings of the 17th International Conference on Educational Data Mining</i> , Atlanta, Georgia, USA, International Educational Data Mining Society, pp. 349–359 (online), DOI: 10.5281/zenodo.12729834 (2024).	[3] Miyazaki, Y., Švábenský, V., Taniguchi, Y., Okubo, F., Minematsu, T. and Shimada, A.: E2Vec: Feature Embedding with Temporal Information for Analyzing Student Actions in E-Book Systems, <i>Proceedings of the 17th International Conference on Educational Data Mining</i> (Paaßen, B. and Epp, C. D., eds.), Atlanta, Georgia, USA, International Educational Data Mining Society, pp. 434–442 (online), DOI: 10.5281/zenodo.12729854 (2024).