

# 学習行動ログデータを用いた成績予測における説明性向上に関する研究

TP24009 館良太

指導教授：山下 隆義

## 1.はじめに

教育現場における学習管理システム（LMS）やデジタル教材配信システムの普及に伴い、教育・学習活動に関するデータを収集・解析し、教育改善に活用する Learning Analytics の研究が活発に行われている。これにより、学習者の進捗状況、学習傾向などを把握することで、各学習者の学修状況に合わせた支援を提供することが可能となる。こうした背景から、操作ログのような大規模なデータを扱い、学習者ごとの特徴を捉える手法として、教育分野における機械学習モデルの活用が期待されており、早期退学者の検出や学習行動の改善を目的とした成績予測の研究が多く行われている。

宮崎らは、自然言語処理を用いて操作間の前後関係や時間間隔を保持した分散表現を生成する手法を提案している。しかし、学生特徴の生成には学習量を直感的に表現できる Bag-of-Words (BoW) によるヒストグラム特徴を用いており、操作頻度分布の偏りを十分に考慮できないため、希少だが重要な操作が特徴として反映されにくいという問題がある。そこで本研究では、BoW と同様に頻度情報に基づきつつ、文書長正規化と逆文書頻度を考慮可能な BM25 と、クラス間差の大きさを定量化できる効果量を組み合わせた重み付けにより、操作頻度分布の偏りとクラス間差の双方を考慮した成績予測手法を提案する。

## 2.先行研究

宮崎らは、学習操作間の前後関係や時間間隔といった情報を保持した分散表現の生成手法である E2Vec[1] を提案している。E2Vec の概要を図 1 に示す。E2Vec は、前処理、埋め込み、集約の 3 つのモジュールで構成される。

前処理では、操作ログを自然言語処理の“文字”、“単語”、“文章”に対応付け、文字列表現に変換する。文字は 1 文字 1 操作に対応し、操作名を 1 文字に変換して表現する。各操作名と文字の対応表を表 1 に示す。単語は最大 1 分間かつ 15 文字以内で構成され、隣接する文字間に操作の時間間隔に対応した文字を挿入して表現される。文章は複数の単語から構成され、操作間隔が 5 分以上となるまでを 1 つの文章とする。

埋め込みでは、前処理により文字列表現に変換した操作ログを用いて fastText を学習し、学習済み fastText により各単語を 100 次元ベクトルに埋め込む。さらに、単語ベクトルを平均化して文章埋め込みベクトルを生成する。

集約では、文章埋め込みベクトルに対して k-means によるクラスタリングを行い、codebook を生成する。生成した codebook を用いて BoW アプローチにより学生の特徴ベクトルを生成する。

E2Vec は BoW アプローチに基づく特徴量生成を行うため、操作頻度分布の不均衡を考慮できず、モデルが“希少だが重要な操作”を捉えることが困難となる。

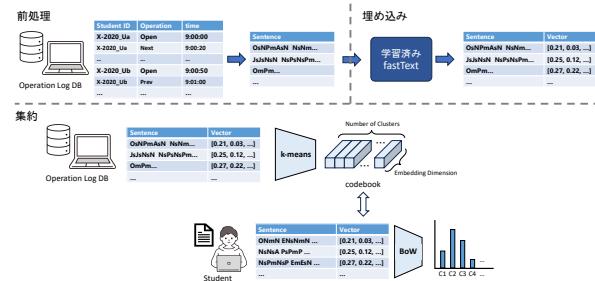


図 1 : E2Vec の概要

## 3.提案手法

本研究では、効果量に基づく重み付けを適用した、E2Vec ベースの BM25 特徴による成績予測手法を提案する。本手法の概要を図 2 に示す。本手法は、codebook 生成モジュール

表 1 : 操作名と文字の対応表

操作名	説明	文字
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

ル、特徴量生成モジュール、成績予測モジュールの 3 つで構成される。

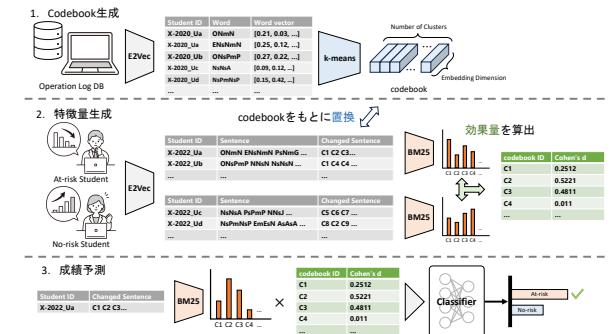


図 2 : 提案手法の概要

### 3.1. Codebook 生成モジュール

操作ログデータに E2Vec の前処理を適用し、操作単語の分散表現を獲得する。得られた操作単語ベクトルを k-means++ でクラスタリングし、codebook を作成する。各操作単語を最も近いクラスタに置換し、学生ごとのクラスタ列を生成する。

### 3.2. 特徴量生成モジュール

クラスタ列に BM25[2] を適用し、操作の頻度と希少度を考慮した特徴量を算出する。BM25 とは、文書集合における、ある単語の重要度を測るための尺度である。BM25 特徴  $BM25_{i,j}$  は式 (1) で定義される。

$$BM25_{i,j} = IDF_j \cdot \frac{f_{i,j} (k_1 + 1)}{f_{i,j} + k_1 (1 - b + b \frac{L_i}{L})} \quad (1)$$

$$IDF_j = \log \frac{N - n_j + 0.5}{n_j + 0.5} \quad (2)$$

ここで、 $N$  は学生数、 $n_j$  はクラスタ  $j$  を含む学生数、 $f_{i,j}$  は学生  $i$  のクラスタ  $j$  の出現回数、 $L_i$  は学生  $i$  のクラスタ総出現回数、 $k_1$  と  $b$  は BM25 のハイパーパラメータである。

BM25 特徴を用いてクラス間の行動差を定量化するために、式 (3) のように効果量 [3] を算出する。

$$d_j = \frac{\bar{x}_{at-risk, j} - \bar{x}_{no-risk, j}}{s_p} \quad (3)$$

ここで、 $\bar{x}_{at-risk, j}$  は低成績クラスにおけるクラスタ  $j$  の BM25 平均値、 $\bar{x}_{no-risk, j}$  は高成績クラスにおけるクラスタ  $j$  の BM25 平均値、 $s_p$  はプールされた標準偏差である。特徴量  $x_{i,j}$  は式 (4) で求める。

$$x_{i,j} = BM25_{i,j} \times |d_j| \quad (4)$$

### 3.3. 成績予測モジュール

式(4)により算出された特徴量  $x_{i,j}$  をクラスタごとに集約することで、各学生を  $K$  次元の特徴量ベクトルとして表現する。この特徴量ベクトルを機械学習モデルに入力し、At-risk と No-risk の二値分類を行う。本研究では、分類モデルとして Random Forest を用いる。

### 4. 評価実験

本実験では、定量的評価として、提案手法を用いて成績予測を行い、E2Vec と成績予測精度を比較する。定性的評価として、BM25 値と効果量の分析を行い、各成績において重要な操作を調査する。また、SHAP を用いて成績予測に寄与する特徴量の分析を行う。

#### 4.1. 実験条件

本実験では、九州大学で収集された LMS の操作ログデータを使用する。A, D はコースの種類を表し、2020 年、2021 年、2022 年はコースが開講された年を表す。A-2020 と D-2020 は、fastText の学習および codebook の生成にのみ使用し、A-2021, A-2022, D-2021, D-2022 の 4 コースを成績予測に用いる。1 コースを訓練、別の 1 コースを評価とした組合せを全通り実施し、計 12 通りの実験を行う。成績は A, B, C, D, F の 5 段階評価であり、F は単位不合格を意味する。本研究では、A, B を No-risk, C, D, F を At-risk として扱う。分類モデルには Random Forest Classifier を使用し、Grid Search によりハイパーパラメータ探索を行う。評価指標には F1-Score を用いる。

#### 4.2. 定量的評価

表 2 に、4 コースに対する E2Vec と提案手法の分類精度を示す。表 2 より、評価データにコース A のデータを用いた場合に最大 0.41 pt の精度向上が確認できた。これより、訓練データと評価データで異なるコースを用いた場合でも精度が向上し、汎化性能があることが確認できる。

表 2 : E2Vec と提案手法による分類精度比較

	train	test	E2Vec	Ours
A-2021	A-2022	0.72	<b>0.74</b>	
	D-2021	<b>0.60</b>	0.60	
	D-2022	0.53	<b>0.55</b>	
A-2022	A-2021	0.71	<b>0.77</b>	
	D-2021	<b>0.67</b>	0.58	
	D-2022	0.51	<b>0.52</b>	
D-2021	A-2021	0.53	<b>0.77</b>	
	A-2022	0.24	<b>0.65</b>	
	D-2022	<b>0.64</b>	0.48	
D-2022	A-2021	0.59	<b>0.79</b>	
	A-2022	0.38	<b>0.72</b>	
	D-2021	<b>0.85</b>	0.67	

#### 4.3. 定性的評価

定量的評価において最も高い予測精度であった D-2022 を対象として、BM25 値、効果量、および SHAP に基づき、各成績クラスの特徴的な操作パターンと、クラス間を識別するために重要な操作パターン、成績の向上・低下に寄与する操作パターンを分析する。表 3 に各クラスで BM25 平均値が高い操作パターン上位 5 件を示す。表 3 より、At-risk クラスは操作間隔が短い “s” を含む “Next” の連続操作パターンであること、No-risk クラスは操作間隔が短い “m” を含む “Next” の連続操作や “Page Jump” や “Add Marker” を含む操作パターンであることが確認できる。このことから、ページ移動効率や教材内容を理解しようとすると能動的な学習行動が成績に関係していると考える。

表 4 に各クラスで効果量が高い操作パターン上位 5 件を示す。表 4 より、At-risk を特徴づける操作パターンは、操作間隔が短い “s” を多く含む “Next” と “Prev” で構成された操作パターンであることが確認でき、No-risk を特徴づける操作パターンは、操作間隔が長い “m” と “l” を含む “Next” と “Prev” で構成された操作パターンであることや、“Add Marker” を含む操作パターンであること

が確認できる。このことから、ページごとの滞在時間や教材内容を整理する行動がクラス間の差であると考える。

表 3 : D-2022 における BM25 平均値が高い操作

At-risk Operation	No-risk Operation
CsCsCm	NNNNsNsCl
GNmGNl	JsNJPJJNPNJP
NNsNNsNNNsNsNs	NmNmNmNm
NsNsNsNsNsNsNs	NmNsNmNsNsNs
GsNmNsGmNsGm	AmAsAm

表 4 : D-2022 における効果量が高い操作

At-risk Operation	No-risk Operation
NsPmNmNm	PPPsPPPPPNNNNNN
NsNsNsNI	NsPPNsNPmNmNNN
NNNNNNNNPsPmPP	AsCm
PmPmPmNsNm	NNmCl
PPPsPsPPsPm	NNsNPsPl

さらに、予測結果の説明を目的として SHAP により特徴量の寄与度を分析した。図 3 に SHAP 値の分布を示す。その結果、操作パターン “OsJsNPPPPNNNNNN”, “AsCm”, “PPPPPPPPPsPmPP”, “NsNsNmOmNNNNNN” は特徴量値が高いほど負の寄与を示す傾向が見られ、一方で “NsNm”, “NNNNNsNNNNNNNN” は特徴量値が高いほど正の寄与を示す傾向が見られた。これより、成績に良い影響を与えるクラスは “Page Jump” や “Add Marker” を含むパターンであり、成績に悪い影響を与えるクラスは “Next” や “s” で構成されたパターンであることが分かる。以上より、希少な操作を含む学習行動が成績向上に寄与する一方で、短時間の連続ページ遷移が成績低下に寄与する可能性が示唆された。

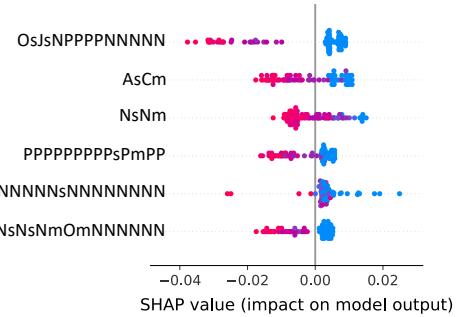


図 3 : SHAP による寄与度の可視化結果

#### 5. おわりに

本研究では、効果量に基づく重み付けを適用した E2Vec ベース BM25 特徴による成績予測を行った。結果から、従来手法と比べ、訓練と評価で異なるデータを使用した場合でも精度向上が確認できた。また、BM25 値と効果量、および SHAP による分析により、操作間隔の違いや “Page Jump”, “Add Marker” といった希少操作の有無がクラス間の差として現れており、教材の熟読やページ移動効率が成績に関係していると考えられる。今後は、重み付けを特徴量ではなく損失関数に適用した成績予測モデルの構築を行いう予定である。

#### 参考文献

- [1] Y. Miyazaki *et al.*, “E2Vec: Feature Embedding with Temporal Information for Analyzing Student Actions in E-Book Systems”, EDM, 2024.
- [2] S. E. Robertson *et al.*, “Okapi at TREC-3”, TREC-3, 1995.
- [3] J. Cohen, “Statistical Power Analysis for the Behavioral Sciences (2nd ed.)”, Lawrence Erlbaum Associates, 1988.

#### 研究業績

- [1] R. Tachi *et al.*, “Grade Prediction Using fastText Features Weighted Through Differential Pattern Mining”, LAK, 2025.

(他 2 件)