

指導教員: 川嶋 宏彰教授

電子教材の閲覧データとコンテンツ情報を用いた学習者のスコア予測

兵庫県立大学社会情報科学部

社会情報科学科

2019 年度入学

JB19S029

小岸 沙也加

2022 年 2 月提出

電子教材の閲覧データとコンテンツ情報を用いた学習者のスコア予測

JB19S029 小岸 沙也加

指導教員: 川嶋 宏彰教授

概要

講義で使用する資料の閲覧方法として、近年、デジタル教材の配信システム上に教師がアップロードした講義資料をオンラインで利用する形態が用いられるようになっていく。このようなシステムが使用されることにより、詳細な操作ログを取得できる。この閲覧データを解析することで各学生の理解度や、つまづき箇所を推定できれば、個々の学生に早い段階でアプローチすることができ、学生の学力向上に繋がると期待できる。学生の閲覧データを使用した先行研究はあるが、学生が講義で使用する講義資料そのものの情報を成績予測に用いることの有効性については十分検証されていない。

そこで本研究では、オンライン上で講義資料が閲覧できるシステムの一つである BookRoll から得られた閲覧データから、学生の行動からスコア予測を行うベースラインに加えて、実際に使用された講義資料のテキスト情報を含めて学習者のスコア予測を行うことで、スコア予測の精度向上を目指す。講義資料の情報を含めてスコア予測を行うアプローチとして、2つの手法を提案する。手法1では、それぞれの学生がよく閲覧したページに含まれるテキスト情報を用いる。また、手法2では、小テストに関係するページを使用する。講義資料の情報は学習済みの Sentence-BERT を使用して得られる、文の埋め込みベクトルを使用する。

スコア予測には勾配ブースティングを使用し、RMSE (Root Mean Squared Error) で評価する。小テストごとに予測を行い、ベースラインの RMSE 平均と各提案手法の RMSE 平均を比較すると、ベースライン、提案手法2に比べて、提案手法1が最も予測精度が良くなった。これより講義資料の情報を含めて予測することはスコア予測の向上に繋がるといえる。

目次

概要	i
第 1 章 序論	1
1.1 研究背景	1
1.2 先行研究	1
1.2.1 学生の行動から成績予測を行う研究	2
1.2.2 学生の学力向上を図る研究	2
1.3 問題設定	3
1.4 本論文の構成	4
第 2 章 電子教材閲覧データとその分析	5
2.1 BookRoll システム	5
2.2 BookRoll から得られる閲覧データ	5
2.3 本研究の使用データ	7
2.4 閲覧データの記述統計量	8
2.5 小テストデータの記述統計量	11
2.6 閲覧データとコンテンツ情報の分析	12
第 3 章 学習者のスコア予測	18
3.1 行動特徴ベクトルとコンテンツ情報を用いたスコア予測（提案手法 1）	18
3.2 小テストに関係するページを使用したスコア予測（提案手法 2）	20

第 4 章	評価	21
4.1	予測および評価方法	21
4.2	提案手法 1	22
4.2.1	スコア予測結果	22
4.2.2	考察	24
4.3	提案手法 2	24
4.3.1	スコア予測結果	24
4.3.2	考察	27
第 5 章	結論	29
	参考文献	32
付録 A	付録	34
A.1	各小テストのトピック	34

第 1 章

序論

1.1 研究背景

初等中等教育から高等教育まで、タブレットやノート PC が幅広く教育現場で普及し、学生の講義資料の閲覧や課題の提出等で用いられるようになっている。講義資料の閲覧や操作では、オンラインで配布された PDF 等の資料をダウンロードしてオフラインで閲覧する形態だけではなく、デジタル教材の配信システム上に教師がアップロードした講義資料をオンラインで閲覧する形態もある。オンライン上で講義資料を閲覧することで、いつ、どの学生が、どのコンテンツのどのページでどんな操作をしたかという学生の詳細な操作を記録した閲覧データが取得できる。この閲覧データを解析することで各学生の教材の理解度や、つまずき箇所を推定できれば、小テストや定期テストの前の早い段階で個々の学生に適したアプローチをすることができるようになり、学生の学力向上、教材のより深い理解に繋がることを期待できる。

1.2 先行研究

本節では、これまでにデジタル教材配信システムや学習支援システムなどのデータを用いて行われた研究として、本研究と同じく成績予測、スコア予測を行う研究、および学生の学力向上を図る研究を紹介する。

1.2.1 学生の行動から成績予測を行う研究

デジタル教材の配信システムのひとつである BookRoll システムから得た閲覧データを用いて、毎週、生徒の最終成績を予測し、受講期間中に随時リスクのある学生とない学生に分類する試み [1] がある。これ以前の研究でも、毎週の閲覧ログを利用して成績予測は行われてきたが、毎週機械的に収集していると講義が休みの週のデータも含まれており、予測精度を低下させる可能性があることに注目し、予測のためにデータを変換させることで、高性能な分類モデルを導出している。

ニューラルネットワークの解釈手法を用いて、閲覧データからみられるどのような学習行動が成績に関連するかを調べた研究 [2] では、特定の行動を要素として成績を予測する。予測した成績を上位、中位、下位に分類し、それぞれのグループの傾向と差を観察することで成績下位の学生によりの確なアドバイスをできるようになっている。

1.2.2 学生の学力向上を図る研究

[1] と同様に BookRoll システムで得たデータを使用し、学生の必修科目のコース割り当ての最適化により学生の講義の満足度、成績の向上を図る研究 [3] がある。この研究では、いくつかのコースに学生を分けて行われる必修科目において、学籍番号順などで機械的にコースを振り分けられることで学生と教師間にミスマッチが起こっていることに注目している。この機械的な割り当てに BookRoll システムから得たデータを使用することで学生の学習行動と教師の講義の進め方を観察し、学生と教師のよりよいマッチングを図っている。

講義スライドの自動要約による学生のプレビューシステムを提案している研究 [4] では、教師は講義内容を予習してほしいと考えているにもかかわらず、学生は講義以前に予習をすることが少ないことに注目し、予習効率や内容理解を高めるために、講義スライドを自動で要約するシステムを提案している。これにより、要約されたスライドを予習した学生は、予習時間が短いにもかかわらず、講義前の小テストでよりよいスコアを獲得で

きることが示されている。

学生の読書行動と特性を分析した研究学力予測や学生のタイプ分けは以前にも行われているが、[5]では、予測などのために作成する学生のデータの潜在的な表現は単純化されすぎて解釈が困難になっているため、変分オートエンコーダを用いて、より解釈しやすく複雑な潜在的な表現を作成することを検討している。

学生が講義後を書くリアクションペーパーや講義アンケート、小テスト、試験の点数などの情報を基に、各行動と学生の成績との関りを調べる研究 [6] では、機械学習を用いて、講義後の行動から学習者の成績に問題があるかどうかの推定を行っている。様々なアルゴリズムを使用し、どのような記入が成績に影響をあたえるかを調べている。

他にも、知識構造を可視化するコンセプトマップを BookRoll のデータから作成する研究 [7] や、作成したコンセプトマップを使用して学習ログを収集分析し、今後の学生の SRL(Self-Regulated Learning, 自己調整学習) 促進への示唆を提示する研究 [8] がある。

1.3 問題設定

1.2 節で述べたように、学生の成績予測を行う研究や学力向上を図る研究は長く行われてきた。多くの研究が学生の行動に注目している一方で、学生が講義で使用する講義資料そのものの情報も成績予測には有効である可能性があるが十分検証されていない。そこで本研究では、学習者の行動を記録した閲覧データに加えて、講義で使用されたスライドなどの講義資料（本論文では「コンテンツ」と呼ぶ）のテキスト情報を利用することで、各学生の理解度の推定精度向上を目指す。行動や小テストは基本的にコンテンツに基づいて行われるため、コンテンツ情報を利用して理解度予測を行うことはより高い予測に繋がると期待できる。ここで理解度とは、毎週講義後に行われる小テストの点数（スコア）とし、閲覧データは BookRoll システムで得られた操作ログを用いる。このとき、コンテンツ情報の利用が小テストのスコア予測にどれだけ貢献するかを検証する。

1.4 本論文の構成

本論文は 5 つの章で構成されている。2 章では本研究で使用するデータを取得した電子教材閲覧システムの BookRoll について説明をし、本研究で使用する閲覧データとそのデータの分析結果について述べる。3 章では本研究で行うアプローチについて、どのようにコンテンツ情報を使用して予測を行うかの説明をする。4 章では提案した手法の評価の方法とスコア予測を行った結果について述べる。最後に 5 章で本研究に対する結論について述べる。

第 2 章

電子教材閲覧データとその分析

2.1 BookRoll システム

BookRoll システム [9] は、教員がアップロードした講義資料や教材を学生がオンライン上で閲覧できる電子教材システムである。図 2.1 は BookRoll のコンテンツ閲覧画面の例である。コンテンツ閲覧画面では、各ページで、前後ページへの移動、ブックマークをつける、メモの記入、マーカーを引く、レコメンドを開く、理解したか否かの選択などの操作ができる。マーカーは赤色、黄色、青色の 3 色用意されており、各色にそれぞれ重要、わからない、わかったという意味が設けられている。閲覧ページにはブックマークをつけることができ、ブックマークをつけたページでは講義資料の選択画面から直接ブックマークをつけたページから開く、閲覧しているページからブックマークをつけたページへ移動することができる。メモはページ内にメモを張ることができ、タイトルと本文をわけて記入できる。他にもコンテンツ内をキーワード検索を行うことができ、検索したキーワードを含むページへ移動もできる。

2.2 BookRoll から得られる閲覧データ

閲覧データは、BookRoll のコンテンツ閲覧画面上で、コンテンツを開く、閉じる、次のページへ進む、一つ前のページへ戻るといった操作を行ったタイミングで、その操作タ



図 2.1 BookRoll コンテンツ閲覧画面例 [10]

表 2.1 記録される主な内容

記録名	内容
contents_id/name	コンテンツ ID/名
marker_position/color	マーカーをつけた場所/色
memo_title/text	メモのタイトル/内容
operation_date/name	操作した日時/操作内容
page_no	操作したページ番号
userid	学生に割り振られた番号

イプが，操作を行った学生の ID，日時，コンテンツ番号，ページ番号，操作タイプなどの情報と共に記録される．閲覧データに記録される主な内容を表 2.1 に示す．表 2.1 の operation_name にあたる，閲覧データに記録される操作の一部を表 2.2 に示す．

表 2.2 操作内容の一部

操作名 (operation_name)	内容
OPEN	コンテンツを開く
CLOSE	コンテンツを閉じる
NEXT	次のページへ移動
PREV	ひとつ前のページへ移動
PAGE_JUMP	特定のページへ移動
ADD/DELETE BOOKMARK	ブックマークをつける/消す
ADD/DELETE MARKER	マーカールをつける/消す
ADD/DELETE/CHANGE MEMO	メモをつける/消す/変える
SEARCH	コンテンツ内検索を行う
SEARCH_JUMP	コンテンツ内検索結果のページに移動
GETIT	わかったボタンを押す
NOTGETIT	わからなかったボタンを押す
CLICK/OPEN/CLOSE_RECOMMENDATION	関連サイトをクリック/開く/閉じる

2.3 本研究の使用データ

本研究では 2020 年に九州大学の情報系科目の講義で取得された BookRoll システムの閲覧データ、講義で用いられたコンテンツ、および講義中に行われた小テストの結果（小テストデータ）を用いる。講義は 100 名の学生に対し 90 分の講義が 7 週間に渡って行われ、各講義の最後の約 10 分間で小テストが行われた。講義から取得された閲覧データは計 200,818 行のログからなる。

7 週の講義のうち、1 週目と 2 週目では 2 つのコンテンツを使っている。1 週目のコンテンツのうち 1 つはガイダンス資料であり、2 週目はコンテンツに対応した小テストが 2 回行われている。

2.4 閲覧データの記述統計量

本節では，閲覧データから読み取ることのできる，各種操作の記述統計量を示し，その特徴について述べる．

マーカーについて

表 2.3 は ADD MARKER と DELETE MARKER の学生ごとの総操作回数の統計量，図 2.2, 図 2.3 はそれぞれの総操作回数の箱ひげ図である．表 2.3 より，よく使用している学生と使用していない学生にはかなり差がみられる．2.1 節で述べた通り，色は 3 色用意されているが，色分けして使っている学生もいれば色分けはせずに記号的に使っている学生もいる．学生がマーカーのつけた場所を確認すると，多くのマーカーは学生が個人的につけているものであるが，ページによってはほとんどの学生がマーカーを付けている場所もあるため，教師の発言からマーカーを引いているページもあると考えられる．

表 2.3 学生ごとのマーカーに関わる総操作回数の統計量

操作名	ADD MARKER	DELETE MARKER
平均	19.37	1.74
標準偏差	39.41	5.28
最小値	0.00	0.00
第一四分位数	0.00	0.00
第二四分位数	5.50	0.00
第三四分位数	19.00	1.00
最大値	229.00	42.00

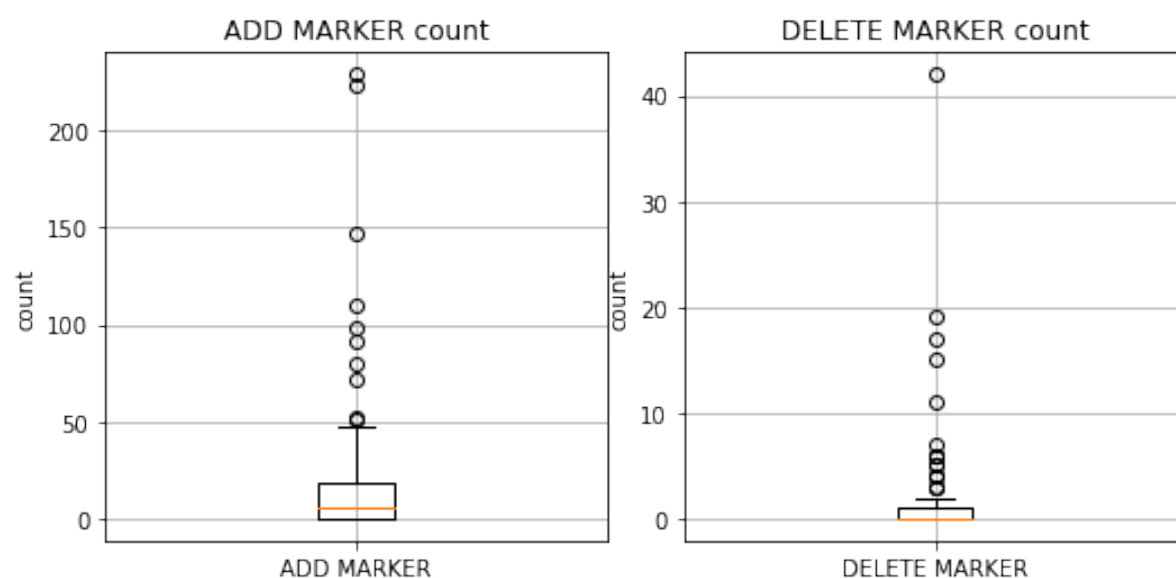


図 2.2 ADD MARKER の回数の箱ひげ図 図 2.3 DELETE MARKER の回数の箱ひげ図

GETIT, NOTGETIT について

表 2.4 は GETIT, NOTGETIT の学生ごとの総操作回数の統計量である。GETIT, NOTGETIT 共に、使用している人数に偏りがあることがわかる。GETIT と比べて、NOTGETIT の記録数の方が顕著に少なく、NOTGETIT の方が記録が多いページは 7 ページのみである。表紙のような情報のないページでも GETIT, NOTGETIT が記録されているコンテンツがあるため、どのテーマ、トピックがわからなかったかの目印として記録している学生がいるように見える。

メモについて

表 2.5 は ADD MEMO, DELETE MEMO, CHANGE MEMO の学生ごとの総操作回数の統計量である。表 2.5 より、メモはあまり使用されていおらず、使用している学生は偏りがあることがわかる。

表 2.4 学生ごとの理解に関わる操作回数の統計量

操作名	GETIT	NOTGETIT
平均	119.23	12.05
標準偏差	128.32	17.95
最小値	0.00	0.00
第一四分位数	2.00	0.00
第二四分位数	70.50	5.00
第三四分位数	206.00	18.25
最大値	460.00	82.00

表 2.5 学生ごとのメモに関わる操作回数の統計量

操作名	ADD MEMO	DELETE MEMO	CHANGE MEMO
平均	1.06	0.12	0.24
標準偏差	4.08	0.41	1.00
最小値	0.00	0.00	0.00
第一四分位数	0.00	0.00	0.00
第二四分位数	0.00	0.00	0.00
第三四分位数	0.00	0.00	0.00
最大値	33.00	2.00	7.00

ブックマークについて

表 2.6 は ADD BOOKMARK, DELETE BOOKMARK の学生ごとの総操作回数の統計量, 図 2.4 は ADD BOOKMARK の総操作回数の箱ひげ図である. 図 2.4 より, BOOKMARK を使用している学生にかなり偏りがあることがわかる.

表 2.6 学生ごとのブックマークに関わる操作回数の統計量

操作名	ADD BOOKMARK	DELETE BOOKMARK
平均	1.72	0.25
標準偏差	7.76	0.97
最小値	0.00	0.00
第一四分位数	0.00	0.00
第二四分位数	0.00	0.00
第三四分位数	0.00	0.00
最大値	65.00	6.00

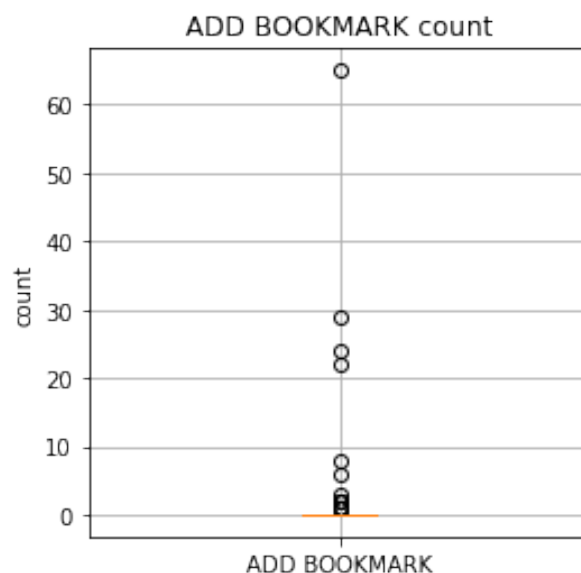


図 2.4 ADD BOOKMARK の回数の箱ひげ図

2.5 小テストデータの記述統計量

2.3 節で述べた通り，小テストは毎週講義後に 10 分間で行われている．小テストは 5 問の択一式の問題から成り，学生が小テストを提出したタイミングで，問題の文章，学生の選択した選択肢，正解か否か，および提出時間が記録される．表 2.7 は各講義回ごとの

表 2.7 各小テストのスコア分布 (人)

スコア (点)	1 週目	2 週目 (1)	2 週目 (2)	3 週目	4 週目	5 週目	6 週目	7 週目
0	0	0	0	0	0	0	0	1
1	1	1	0	0	1	1	2	3
2	3	1	0	0	1	2	5	6
3	12	11	3	1	14	12	15	28
4	42	23	40	9	30	33	23	28
5	34	54	47	80	43	42	45	24
合計人数	92	90	90	90	89	90	90	90

点数の分布である．表より多くの学生が満点をとっているが，小テストによっては点数にばらつきがあることが確認できる．講義に登録している学生は 100 名だが，全ての講義，もしくは途中の講義から講義に参加しなくなった学生がいるため，各小テストの分析は表 2.7 の合計人数で行っている．

出題される小テストの問題文は全ての学生に対して共通であるが，選択肢の一つである正解の文章や選択肢の順序は学生ごとに変化がつくようになっている．具体的には正解文は数通り用意されているが，文意は基本的には同じである．小テスト中もコンテンツの閲覧は許可されており，小テストの提出時間を確認すると，点数の低い人は提出時間が比較的早めであり，提出時間が遅くなるにつれて点数の高い人が増えているため，資料を確認しつつ答えている人がいることがわかる．

2.6 閲覧データとコンテンツ情報の分析

本節では閲覧データとコンテンツ情報について分析を行う．まず，コンテンツに含まれる各ページのテキスト情報をベクトル化し，ページ間の関係を観察する．続いて，操作された時間に注目し，いくつかの時間区分にわけて分析を行う．

コンテンツ情報の分析

各コンテンツにおいてページ間の関係があるか分析する。ページ間の類似度を確認するためにページに含まれるテキスト情報のベクトル化を行い、各ページベクトルのコサイン類似度を求める。

ベクトル化には事前学習済みの Sentence-BERT [11] *1を使用してコンテンツ c に含まれるページ $p^{(c)}$ ごとに 768 次元の「ページベクトル」 $\mathbf{v}_{(c,p)}$ を求める。ただし単位ベクトルに正規化を行った。コサイン類似度は対象となる 2 つのベクトルの距離を図る指標である。式 (2.1) で計算され、0-1 で表現される。

$$\cos_similarity(x, y) = \frac{x \cdot y}{\|x\| \|y\|} = \frac{\sum_n x_n y_n}{\sqrt{\sum_n x_n^2 + \epsilon} \sqrt{\sum_n y_n^2 + \epsilon}} \quad (2.1)$$

式 2.1 の分母に含まれる ϵ は分母が 0 にならないための、0 に近い整数である。

図 2.5 から図 2.13 は各コンテンツにおけるページ間のコサイン類似度をヒートマップに表したものである。1 に近いほど（ページ間の距離が近いほど）色が濃くなっている。これらの図では、どのコンテンツにおいてもいくつかのページがひと固まりになっている部分があることがわかる。実際にコンテンツを確認してみると、コサイン類似度においてひと固まりとされているページ間はある程度同じトピックが説明されているため、このページベクトルは各ページに含まれるテキスト情報をよく表現しているといえる。しかし、コンテンツ自体が 1 つのテーマをもっているため、離れたページでも強く関係があるとされているページも存在する。なお、コンテンツには図表を使用しているページもあるが、その図表に含まれる内容および位置関係についてはベクトル化を行っていない。

時間区分ごとの行動

学生の行動がされた時間を講義時間内 (inlec)、小テスト時間内 (intest)、講義時間外 (outlec) に分割し、分析を行った。講義時間内、小テスト時間内、講義時間外で学生の総

*1 日本語版 BERT モデル [12] を未公開のデータセットで MultipleNegativesRankingLoss を用いてファインチューニングを行った Sentence-BERT モデル

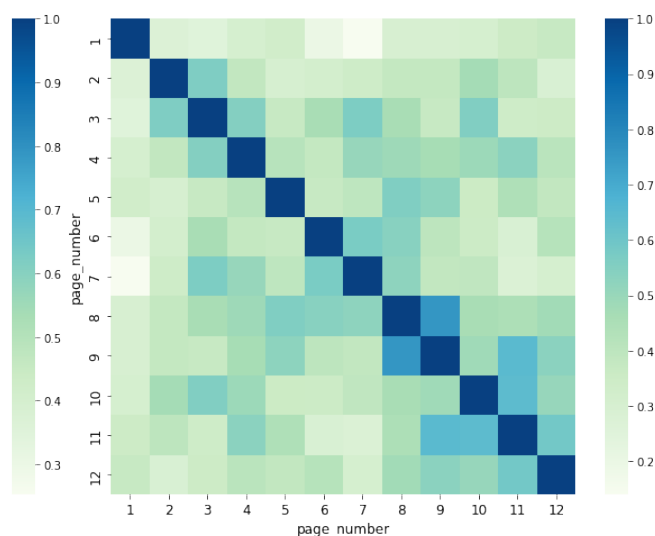
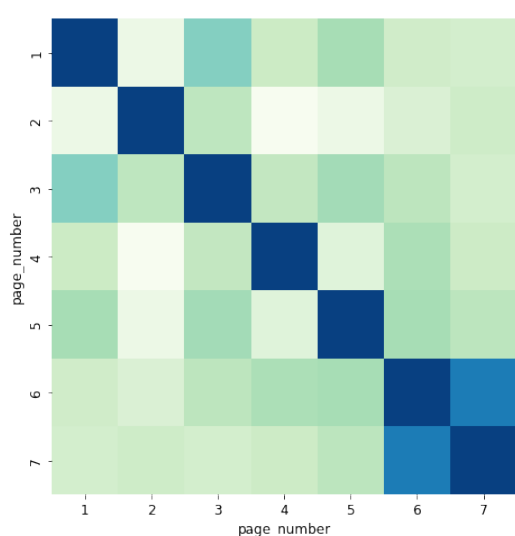


図 2.5 各ページ間の類似度 (1 週目ガイダンス) 図 2.6 各ページ間の類似度 (1 週目講義資料)

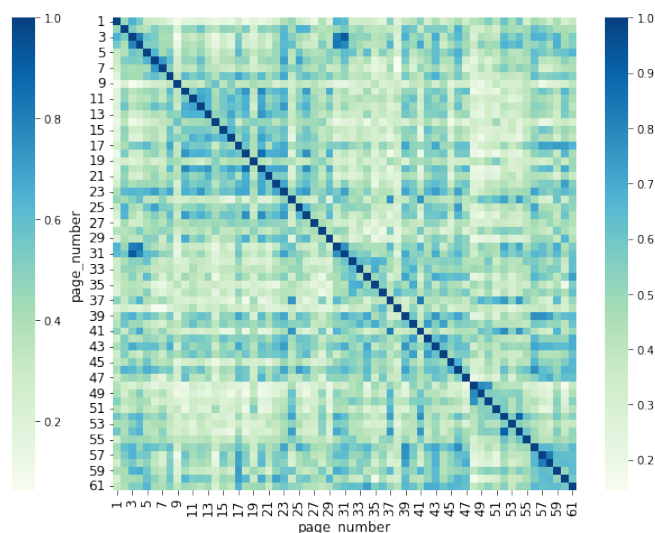
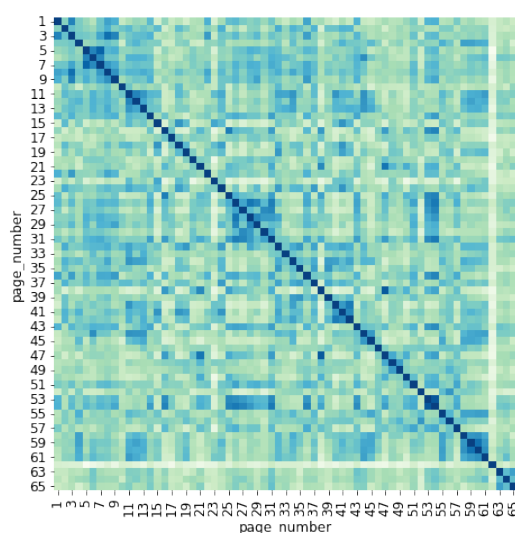


図 2.7 各ページ間の類似度 (2 週目 (1))

図 2.8 各ページ間の類似度 (2 週目 (2))

操作回数の 10 分ごとの時間平均をとった結果が表 2.8 であり，平均総操作回数が一番多い時間は小テスト時間内である．なお講義時間外の範囲は，全ての講義終了後 1 週間までとし，対象となる講義の予習復習のために活動可能な時間を各週 1 時間 30 分として時間平均を計算している．

ここで，時間区分ごとの行動を k-means 法 ($k = 4$) でクラスタリングを行った．その結果「全体的に活動にしている学生」「全体的に活動していない学生」「テスト時間中によ

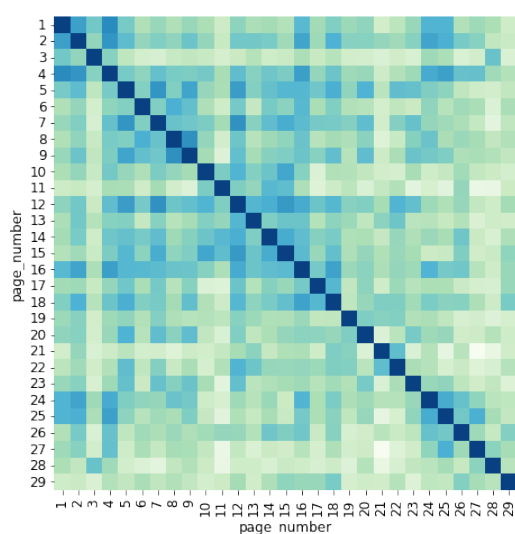


図 2.9 各ページ間の類似度 (3 週目)

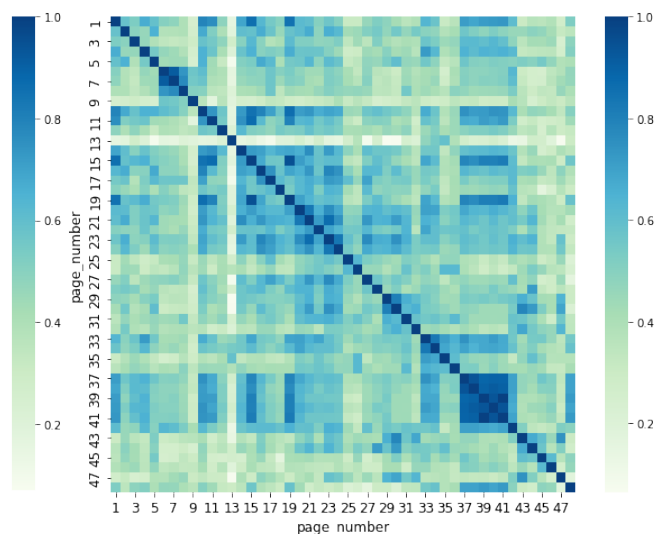


図 2.10 各ページ間の類似度 (4 週目)

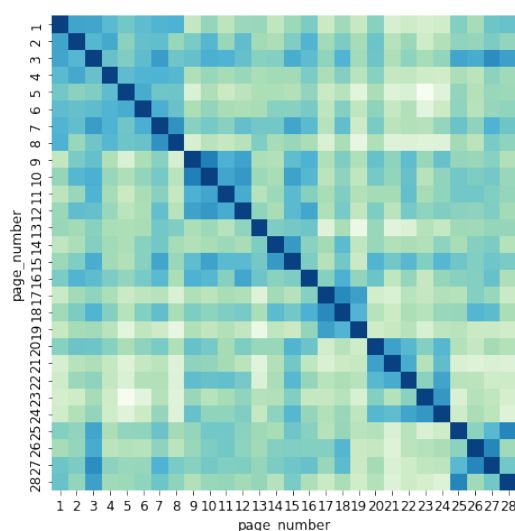


図 2.11 各ページ間の類似度 (5 週目)

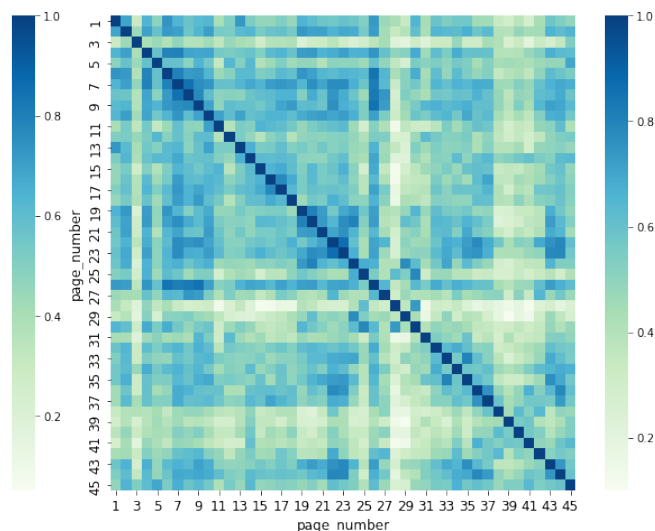


図 2.12 各ページ間の類似度 (6 週目)

く活動している学生」「全体的に活動量が中途半端な学生」というクラスタに分かれた。各クラスタの時間区分ごとの総行動数の箱ひげ図を図 ?? に示す。最も学生の多いクラスタは「全体的に活動量が中途半端な学生」のクラスタであり、44 人がこのクラスタに分類されている。さらに、講義時間内、小テスト時間内、講義時間外を各講義単位に分割すると「全体的に活動はしているが多くはない学生」は前半から中頃にかけて活動が減り、後半にかけて活動が増えていることがわかった。

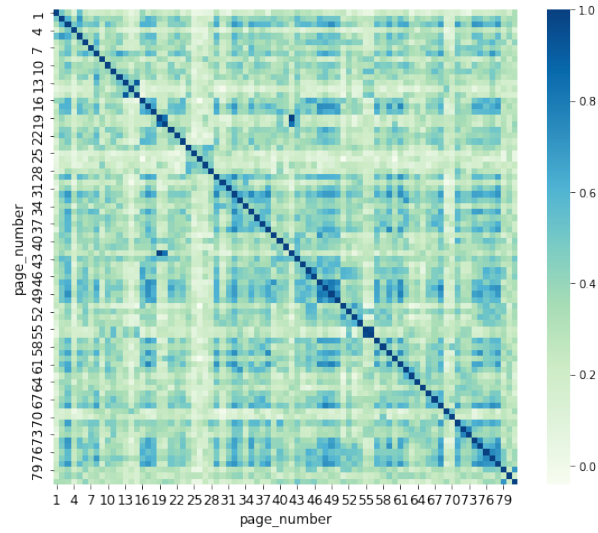


図 2.13 各ページ間の類似度（7 週目）

表 2.8 時間区分ごとの統計量（単位:10 分）

	講義時間内	小テスト時間内	講義時間外
平均	18.17	33.64	11.83
標準偏差	10.58	29.54	10.67
最小値	0.00	0.00	0.00
第一四分位数	12.37	9.79	3.42
第二四分位数	16.78	24.21	9.32
第三四分位数	23.26	48.93	17.29
最大値	57.43	128.00	63.54

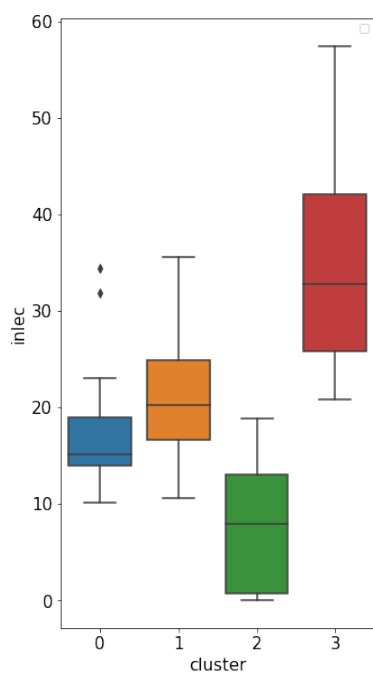


図 2.14 講義時間内

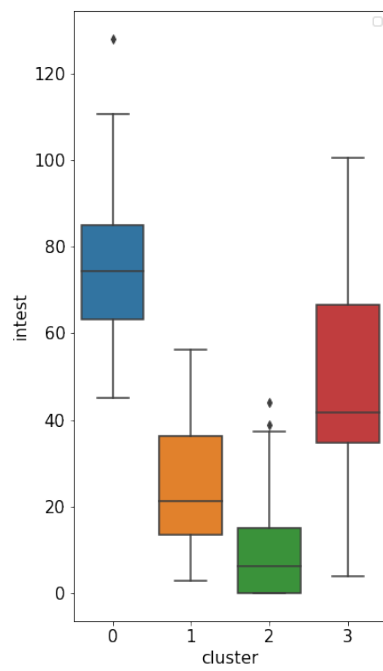


図 2.15 小テスト時間内

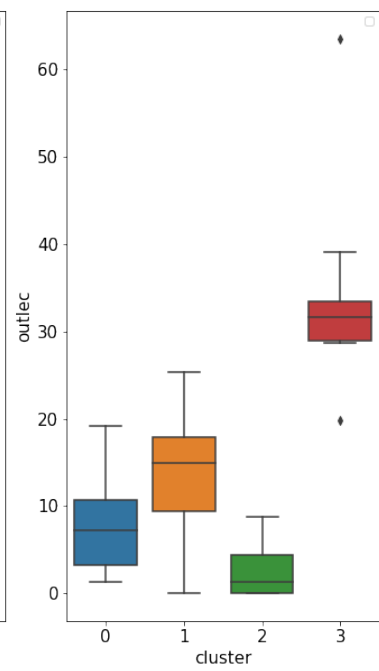


図 2.16 講義時間外

図 2.17 各クラスターの時間区分ごとの総操作回数の箱ひげ図

第 3 章

学習者のスコア予測

図 3.1 に本研究の全体像を示す．本研究では学生の行動からスコア予測を行うベースラインに加えて，講義で使用されたコンテンツの情報を含めて学習者のスコア予測を行う．コンテンツ情報を含めてスコア予測を行うアプローチとして，学生がよく閲覧したコンテンツに含まれるテキスト情報を使用する手法 1 と小テストに関係するページの情報を使用する手法 2 を提案する．

3.1 行動特徴ベクトルとコンテンツ情報を用いたスコア予測 (提案手法 1)

提案手法 1 では，学生の行動から得られる特徴ベクトルと，その学生がよく閲覧したコンテンツから得られる特徴ベクトルを連結して，スコア予測の特徴ベクトルとして用いる．

まず，BookRoll から得られた閲覧データより，各学生について，あるコンテンツの各ページにおける操作タイプごとの操作回数および閲覧時間を求め，これらの特徴量を要素とするようなベクトルを各学生の「行動特徴ベクトル」と呼ぶ．ここで，学生 i のコンテンツ c における行動特徴ベクトルを $\mathbf{u}_c^{(i)}$ と表す．行動特徴ベクトル $\mathbf{u}_c^{(i)}$ の次元数は「ページ数 \times (操作タイプ数 $+ 1$)」(最後の 1 はページ閲覧時間に対応)であり，ページ

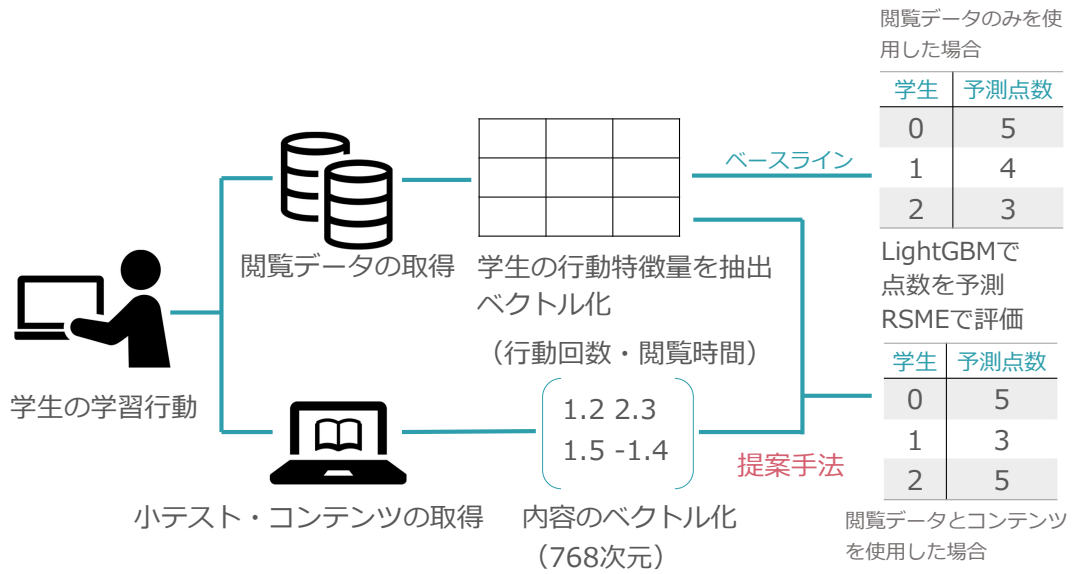


図 3.1 本研究の全体像

数はコンテンツ c によりそれぞれ異なる．ページ数は付録の表 A.1 に示す．

次に，コンテンツ c に含まれる文章のベクトル化を行い，そのベクトルを行動特徴ベクトルに連結してスコア予測を行う．コンテンツに含まれる文章のベクトルには 2.6 節のページベクトル $\mathbf{v}_{(c,p)}$ を使用する．ページベクトルをそのまま連結すると全ての学生が同じページベクトルをもつことになるため，重みづけをすることで学生ごとに違ったベクトルを作成する．本研究では各ページの閲覧時間に注目し，学生 i がコンテンツ c においてよく閲覧したページの内容情報を多く含む「閲覧コンテンツベクトル」 $\mathbf{v}_c^{(i)}$ を求め，学生 i のページ p に対する閲覧時間の総和 $t_{(c,p)}^{(i)}$ に基づいてページベクトルに重みづけを行って足し合わせる．具体的には，閲覧時間をそのまま各ページの重み $w_{(c,p)}^{(i)} = t_{(c,p)}^{(i)}$ としてページベクトルの線形和 $\mathbf{v}_c^{(i)} = \sum_p w_{(c,p)}^{(i)} \mathbf{v}_{(c,p)}$ を求め，単位ベクトルに正規化して閲覧コンテンツベクトルとする．ただし，同じページを継続して 5 分より長く開いていたログは，長時間放置されたものとして，各ページの閲覧時間の総和から除外した．

求めた行動特徴ベクトル $\mathbf{u}_c^{(i)}$ と閲覧コンテンツベクトル $\mathbf{v}_c^{(i)}$ を連結させ，以下の新た

なベクトルを得る.

$$\mathbf{f}_{1,c}^{(i)} = (\mathbf{u}_c^{(i)}, \mathbf{v}_c^{(i)}) \quad (3.1)$$

提案手法 1 ではこの特徴ベクトル $\mathbf{f}_{1,c}^{(i)}$ を用いてスコア予測を行う.

3.2 小テストに関係するページを使用したスコア予測（提案手法 2）

2 つ目の提案手法は、小テストに関係するページにおける行動をとりだしてスコア予測を行う方法である. 小テスト q に含まれる問題 j ごとに、問題文と正解文をベクトル化し、「小テストベクトル」 $\mathbf{Q}_{(q,j)}$ を作成する. ここで、テキストのベクトル化には 2.6 節で述べた Sentence-BERT に基づく手法を用いる. さらに、小テストベクトルとページベクトルの関係を見るために 2.6 節で用いた、コサイン類似度を使用する.

コンテンツ c に対応する小テストを $q(c)$ とする. このとき、コンテンツ c の各ページベクトル $\mathbf{v}_{(c,p)}$ について、 $q(c)$ に含まれる 5 つの小問から得られる小テストベクトル $\mathbf{Q}_{(q(c),j)}$ ($j = 1, \dots, 5$) それぞれとのコサイン類似度を求め、その類似度の和をページ p の重み $w_{(c,p)}$ とする. ただし、コサイン類似度が 0.4 未満である場合は和に含めないものとする. すなわち、重みは以下のように求める.

$$w_{(c,p)} = \sum_{j \in \mathcal{J}} \cos_similarity(\mathbf{Q}_{(q(c),j)}, \mathbf{v}_{(c,p)}) \quad (3.2)$$

ただし、 $\mathcal{J} = \{j \mid \cos_similarity(\mathbf{Q}_{(q(c),j)}, \mathbf{v}_{(c,p)}) \geq 0.4\}$ である.

$w_{(c,p)}$ は、各ページ p がどの程度小テスト $q(c)$ に関連するかを表す重みであり、ページごとの行動特徴ベクトル $\mathbf{u}_{(c,p)}^{(i)}$ をこの $w_{(c,p)}$ で重みづけることにより、以下の新たな行動特徴ベクトルを得る.

$$\mathbf{f}_{2,c}^{(i)} = \sum_p w_{(c,p)} \mathbf{u}_{(c,p)}^{(i)} \quad (3.3)$$

提案手法 2 ではこの特徴ベクトル $\mathbf{f}_{2,c}^{(i)}$ を用いてスコア予測を行う.

第 4 章

評価

4.1 予測および評価方法

スコア予測は小テストごとに LightGBM を使用して行う。LightGBM は勾配ブースティングに基づく機械学習手法である。高い予測精度で計算時間が速い点と、予測時結果への寄与度が要素ごとにわかる点から LightGBM を使用した。5-fold 交差検証を用いて各 fold の RSME(Root Mean Square Error) を計算し、各 RMSE の平均により評価する。評価には小テストごとに求めた学生のスコアを使用する。2.4 節で述べたように、2 週目は 2 つのコンテンツが使用され、小テストが 2 回分行われたため、2 週目 (1) と 2 週目 (2) に分けてスコア予測を行うものとする。

行動特徴ベクトルとして、全ての操作タイプを用いた場合、ベクトルが大きくなりすぎてしまうため、行動特徴ベクトルに含める操作タイプを限定した。事前に LightGBM で全ての小テストの合計値のスコア予測を行った際に、予測結果への寄与度 (Importance) が 0 であった操作タイプを除いた。その結果、ここで用いる行動特徴ベクトルに含める操作タイプは、表 4.1 に示す 12 タイプとなった。これに、ページの閲覧時間 (open_time) を加えた 13 次元の行動特徴ベクトルを実験では用いた。

行動特徴ベクトルは、講義時間外の操作を全て用いる場合と、講義時間内および前後 1 時間の操作に絞った場合の 2 通りで計算し比較する。2.6 節で述べた通り、講義時間外

表 4.1 講義時間外含む

操作名	予測結果への寄与度 (Importance)
OPEN	10.24
NEXT	42.92
PREV	78.11
CLOSE	34.41
PAGE_JUMP	76.88
GETIT	4.80
NOTGETIT	5.76
OPEN_RECOMMENDATION	32.00
CLOSE_RECOMMENDATION	20.67
ADD MARKER	4.27
DELETE MARKER	1.80
CLICK_RECOMMENDATION	0.86

の行動は特に講義時間内および前後 1 時間に偏りがあり、かつ本研究の目的の 1 つである、早めの学生へのレスポンスに繋がると考えたため、講義時間内および前後 1 時間の場合でもスコア予測を行う。

4.2 提案手法 1

4.2.1 スコア予測結果

提案手法 1 で重要度とした閲覧時間が予測にどれほど関わっているか、コンテンツ情報が予測向上に繋がっているかを確認する、

ここで、2.6 節のページベクトルの合計 $\mathbf{v}_c = \sum_p \mathbf{v}(c, p)$ をコンテンツごとに求め、行動特徴ベクトルに連結したベクトルを行動コンテンツベクトル $\mathbf{C}_c^{(i)} = (\mathbf{u}_c^{(i)}, \mathbf{v}_c)$ とする。

ただし、学生が一度も使用していないコンテンツのテキスト情報は省く。

予測結果の比較対象として、以下の特徴ベクトルを用いる。

行動特徴ベクトル（ベースライン）(b)	ページごとに各学生の各ページでの各操作の 回数を要素とするベクトル $\mathbf{u}_c^{(i)}$
閲覧時間のみ (b')	各ページの閲覧時間 $t_c^{(i)}$ を要素とするベクトル
閲覧コンテンツベクトル (c)	各ページベクトルに重みづけしたベクトル $\mathbf{v}_c^{(i)}$
行動コンテンツベクトル (b+c')	コンテンツごとに求めたページベクトルの合計を 行動特徴ベクトル $\mathbf{u}_c^{(i)}$ に連結したベクトル $\mathbf{C}_c^{(i)}$
提案手法 1(b+c)	行動特徴ベクトルと閲覧コンテンツベクトルを 連結したベクトル $\mathbf{f}_{1,c}^{(i)}$

閲覧時間のみ (b'), 行動特徴ベクトルのみ (c), 行動コンテンツベクトル (b+c') でそれぞれスコア予測を行い、比較する。それぞれの方法で小テストごとに求めた RMSE の平均を図 4.1 に示す。小テストごとの RMSE の値は表 4.2, 表 4.3 の通りであり、表 4.2 は講義時間外の操作を全て用いる場合、表 4.3 は講義時間内および前後 1 時間の操作に絞った場合である。

図 4.1 より、講義時間外の行動を含めた場合と講義時間内および前後 1 時間の行動に絞った場合のどちらの条件でも、閲覧時間のみ (b') が RMSE の平均が最も悪くなった。ベースラインの手法 (b) に比べて、行動コンテンツベクトルを用いた場合 (b+c'), 閲覧コンテンツベクトルのみを用いて予測する場合 (c) が RMSE の平均の値が小さく、予測精度がよかった、両方のベクトルを用いた提案手法 1(b+c) の RMSE の平均の値が小さく、予測精度がよかった。提案手法 1(b+c) で講義時間外の行動を含めた場合と講義時間内および前後 1 時間の行動に絞った場合の比較をすると、講義時間内および前後 1 時間の行動に絞った場合の方が RMSE の平均の値が小さく、予測精度がよかった。

小テストごとでは、講義時間外の行動を含めて予測した場合では、表 4.2 より、1 週目、6 週目がベースライン (b), 2 週目 (1) が閲覧時間のみ (b'), 2 週目 (2), 5 週目が閲覧コ

ンテンツベクトルのみ (c), 3 週目が行動コンテンツベクトル ($b+c'$) と提案手法 1($b+c$), 4 週目, 7 週目が提案手法 1($b+c$) が最も RMSE の平均の値が小さく, 予測精度がよく, 講義時間内および前後 1 時間の行動に絞って予測した場合では, 表 4.3 より, 1 週目がベースライン (b), 2 週目 (1), 3 週目が閲覧コンテンツベクトルのみ (c), 5 週目, 6 週目が行動コンテンツベクトル ($b+c'$), 2 週目 (1), 4 週目, 7 週目が提案手法 1($b+c$) が最も RMSE の平均の値が小さく, 予測精度がよかった.

4.2.2 考察

結果から, ベースラインと比べて閲覧コンテンツベクトルのみで予測した場合, コンテンツベクトルで予測した場合, 提案手法 1 で RMSE の値が下がっていることより, コンテンツ情報を含めることはスコア予測の精度向上に繋がるといえる, 加えて, コンテンツベクトルで予測した場合と比べて閲覧コンテンツベクトルのみで予測した場合, 同じ特徴ベクトルの次元であるにも関わらず, 提案手法 1 で RMSE の値が下がっていることより, 学生ごとによく閲覧したコンテンツ情報を用いることが, スコア予測により重要であるといえる. そして, コンテンツ情報を含めて予測する方法では, 講義時間外の行動を含めた場合に比べて講義時間内および前後 1 時間の行動に絞った場合の方が RMSE の値が小さくなっているため, 学生の行動を時間で絞ることはスコア予測の精度向上に繋がる.

4.3 提案手法 2

4.3.1 スコア予測結果

提案手法 2 の結果を図 4.2 に示す. ベースライン, 提案手法 1 と比較するため, それぞれの結果も示している. 小テストごとの RMSE の値は表 4.4, 表 4.5 の通りであり, 表 4.4 は講義時間外の操作を全て用いる場合, 表 4.5 は講義時間内および前後 1 時間の操作に絞った場合である.

図 4.2 より, 講義時間外の行動を含めた場合と講義時間内および前後 1 時間の行動に

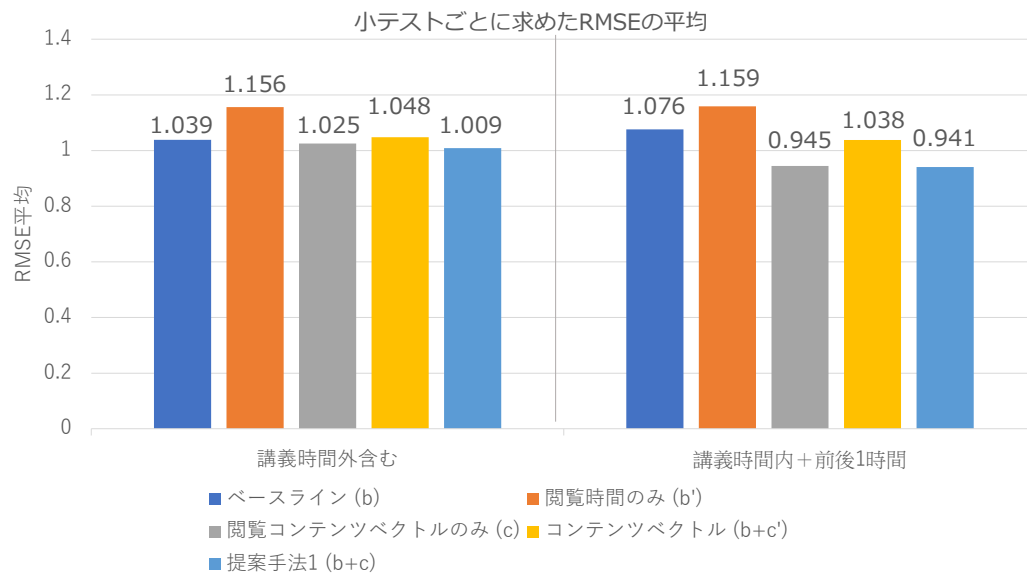


図 4.1 小テストごとに求めた RMSE の平均

表 4.2 小テストごとの RMSE (講義時間外含む)

小テスト	b	b'	c	b+c'	b+c
1 週目	1.144	1.167	1.219	1.158	1.217
2 週目 (1)	1.052	0.975	1.166	1.166	1.083
2 週目 (2)	1.088	1.110	0.777	0.887	0.796
3 週目	0.909	0.867	0.619	0.611	0.611
4 週目	1.191	1.243	1.109	1.181	1.107
5 週目	1.220	1.242	0.932	1.000	0.968
6 週目	1.075	1.229	1.199	1.094	1.132
7 週目	1.372	1.418	1.180	1.284	1.155
平均	1.039	1.156	1.025	1.048	1.009

表 4.3 小テストごとの RMSE（講義時間内および前後 1 時間）

小テスト	b	b'	c	b+c'	b+c
1 週目	1.051	1.084	1.072	1.165	1.103
2 週目 (1)	1.194	1.167	1.060	1.084	1.080
2 週目 (2)	1.089	1.047	0.707	0.896	0.690
3 週目	0.902	0.903	0.384	0.633	0.390
4 週目	1.283	1.201	1.068	1.272	1.007
5 週目	1.139	1.265	1.020	1.004	1.007
6 週目	1.085	1.199	1.004	0.942	0.949
7 週目	1.433	1.407	1.242	1.306	1.233
平均	1.076	1.159	0.945	1.038	0.941

絞った場合のどちらの条件でも，小テストと関係のあるページの行動を重要として予測する手法（提案手法 2）はベースライン，提案手法 1 と比べて，RMSE の平均が悪くなった．提案手法 2 で講義時間外の行動を含めた場合と講義時間内および前後 1 時間の行動に絞った場合の比較をすると，講義時間外の行動を含めた場合の方が RMSE の平均が小さく，予測精度がよかった．

小テストごとの結果を見ると，講義時間外の行動を含めて予測した場合では，表 4.2 より，1 週目がベースライン，2 週目 (1)，3 週目，4 週目，5 週目が提案手法 1，2 週目 (1)，6 週目，7 週目が提案手法 2 において最も RMSE の平均が小さく，予測精度がよかった．講義時間内および前後 1 時間の行動に絞って予測した場合では，表 4.3 より，1 週目がベースライン，2 週目 (1)，3 週目，4 週目，5 週目，6 週目，7 週目が提案手法 1 において最も RMSE の平均が小さく，予測精度がよかった．

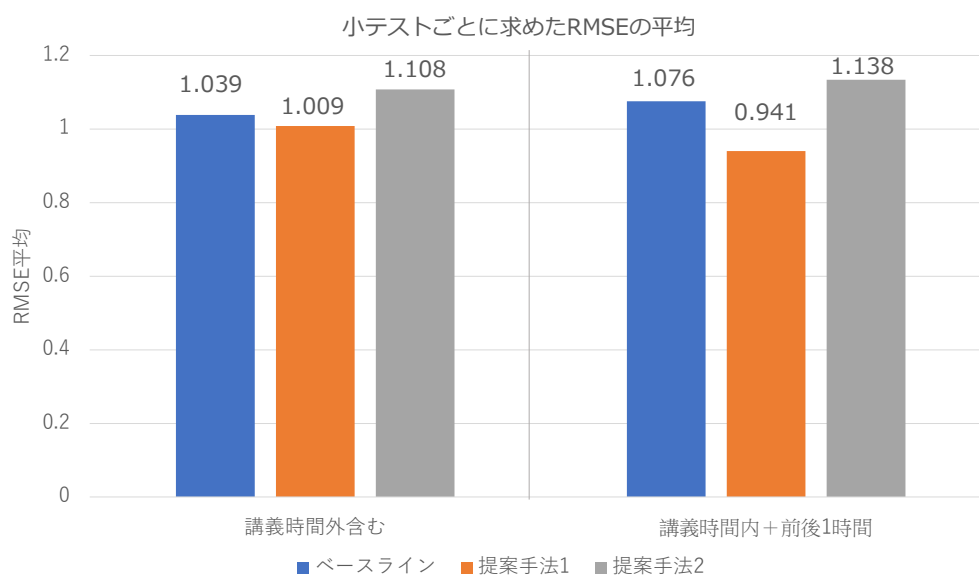


図 4.2 小テストごとに求めた RMSE の平均

4.3.2 考察

提案手法 2 ではベースラインと比較して、RMSE の平均が大きく、予測精度が悪くなった。加えて、提案手法 1 では学生の行動を時間で絞ることで予測精度が向上したが、提案手法 2 では、ベースラインよりも予測精度が悪くなった。これは、提案手法 2 ではコンテンツ情報が直接的に使用されていないからであると考えられる。

表 4.4 小テストごとの RMSE（講義時間外含む）

小テスト	ベースライン	提案手法 1	提案手法 2
1 週目	1.051	1.217	1.105
2 週目 (1)	1.194	1.083	1.002
2 週目 (2)	1.089	0.796	1.044
3 週目	0.902	0.611	0.879
4 週目	1.283	1.107	1.189
5 週目	1.139	0.968	1.212
6 週目	1.085	1.132	1.059
7 週目	1.433	1.155	1.372
平均	1.076	1.009	1.108

表 4.5 小テストごとの RMSE（講義時間内および前後 1 時間）

小テスト	ベースライン	提案手法 1	提案手法 2
1 週目	1.051	1.103	1.066
2 週目 (1)	1.194	1.080	1.197
2 週目 (2)	1.089	0.690	1.072
3 週目	0.902	0.390	0.853
4 週目	1.283	1.007	1.229
5 週目	1.139	1.007	1.137
6 週目	1.085	0.949	1.087
7 週目	1.433	1.233	1.428
平均	1.076	0.941	1.134

第 5 章

結論

本研究は、学習者の行動ログから得られる情報だけでなく、講義資料のコンテンツ情報も用いることで、小テストのスコア予測の精度向上を図ることを目的としている。コンテンツの利用がスコア予測の精度向上に繋がるのか、コンテンツに含まれるテキスト情報のベクトル化を行い、段階的に特徴ベクトルを使用することで検証した。この結果から、コンテンツ情報をスコア予測に含めることは有効であると考えられる。さらに、学生がよく閲覧したコンテンツ情報を用いることが、スコア予測の精度向上に繋がる可能性があることが示された。

今後の課題として以下のものが挙げられる。

ページベクトルの計算方法 ページベクトルに関して、本研究ではページに含まれるテキスト情報のみをベクトル化したが、ページ内の図表の位置関係、および図に含まれる文字も文章同様にベクトル化を行い、ページベクトルの要素とすることができれば、ページの情報を深く取得することになり、より高い精度が期待できる。

ページ重要度の計算方法 提案手法 2 では、ベースラインに比べ予測精度は低くなっているが、ページの重要度の求め方を変えることによって、予測精度が上がる可能性がある。具体的には、教師が長く説明したページほど重要度を高く設定する、学生の操作回数が多いページほど重要度を高く設定するなどである。

特徴量に対する重要度の利用 提案手法 1 では閲覧時間を重要度として特徴量に加え、提案手法 2 では小テストとの関係をページの重要度として特徴量に加えた。ここで二つの手法より、ページの重要度に加え、閲覧時間以外の行動の重要度を LightGBM で求めた Importance に沿って変えることで、予測に必要な情報が厳選され、より高い精度を期待できる。

ベクトルの次元数 本研究で使用した閲覧コンテンツベクトル、行動特徴ベクトルは使用する行動を削る以外の方法で次元数削減を行っていないため、高い次元数のまま予測を行っている。しかし、次元数を落とすことで予測精度が変化する可能性があるため、次元削減の手法を検討する必要がある。

予測結果についての考察 4.2 節の結果について、何が各週の予測結果の違いに影響を及ぼしているのかを十分に検証できていないため、今後の課題となる。

理解度予測 本論文では小テストごとに 5 点満点の点数予測を行ったが、1 問ごとに正解/不正解の予測を行うことで、より深い学生の理解度予測に繋がる可能性がある。

謝辞

本研究を行うにあたり，川嶋教授には研究のアイデアをいただき，また，研究，論文執筆に関して多大なる有益な助言をいただき，大変お世話になりました．深く感謝申し上げます．

九州大学島田教授，峰松准教授にはデータの提供，また，研究，論文への助言をいただきました．深く感謝申し上げます．

最後に，川嶋研究室の皆様には普段から研究に関してご助言，ご協力いただきました．本当にありがとうございました．

参考文献

- [1] Leelaluk Sukrit, Minematsu Tsubasa, Taniguchi Yuta, Okubo Fumiya, and Shimada Atsushi. Predicting student performance based on lecture materials data using neural network models. *CEUR Workshop Proceedings*, pp. 11–20, 2022.
- [2] 椎野徹也, 峰松翼, 島田敬士, 谷口倫一郎. デジタル教材の学習ログと成績の関連分析. 情報処理学会研究報告, Vol. 2020-CLE-30, No. 10, pp. 1–4, 2020.
- [3] Atsushi Shimada, Kousuke Mouri, Yuta Taniguchi, Hiroaki Ogata, Rin ichiro Taniguchi, Shin’ichi Konomi. Optimizing assignment of students to courses based on learning activity analytics. In *EDM*, 2019.
- [4] Atsushi Shimada, Fumiya Okubo, Chengjiu Yin, and Hiroaki Ogata. Automatic summarization of lecture slides for enhanced student preview technical report and user study. *IEEE Transactions on Learning Technologies*, Vol. 11, No. 2, pp. 165–178, 2018.
- [5] Erwin D. Lopez, Tsubasa Minematsu, Taniguchi Yuta, Fumiya Okubo, and Atsushi Shimada. Exploring the use of probabilistic latent representations to encode the students’ reading characteristics.
- [6] 東るみこ. 機械学習による学習者の理解度推定. 教育システム情報学会全国大会講演論文集 (CD-ROM) (教育システム情報学会全国大会 (CD-ROM)), Vol. 43rd, pp. 1–24, 2018.
- [7] Masanori Yamada, Atsushi Shimada, Misato Oi, Yuta Taniguchi, and Shinichi

Konomi. Br-map: Concept map system using e-book logs.

- [8] Hao Hao, Xuewang Geng, Li Chen, Atsushi Shimada, and Masanori Yamada. Learning analytics of the relationships among knowledge constructions, self-regulated learning, and learning performance.
- [9] Ogata Hiroaki, Oi Misato, Mohri Kousuke, Okubo Fumiya, Shimada Atsushi, Yamada Masanori, Wang Jingyun, and Hirokawa Sachio. Learning analytics for e-book-based educational big data in higher education. In *Smart Sensors at the IoT Frontier*, pp. 327–350. Springer, Cham, 2017.
- [10] <https://www.leds.ait.kyushu-u.ac.jp/ja/manual/bookrollh.8pz9mjly17c7>. last accessed on 18 January, 2023.
- [11] <https://huggingface.co/sonoisa/sentence-bert-base-ja-mean-tokens-v2>. last accessed on 11 January, 2023.
- [12] <https://huggingface.co/cl-tohoku/bert-base-japanese-whole-word-masking>. last accessed on 1 February, 2023.

付録 A

付録

A.1 各小テストのトピック

表 A.1 は各小テストのトピックと小テストに対応するコンテンツのページ数である。1 週目は 2 つのコンテンツが使用されているため、表のような記載となっている。

表 A.1 各小テストのトピックと対応するコンテンツのページ数

小テスト	トピック	ページ数
1 週目	セキュリティ対策について	8+13
2 週目 (1)	盗難・紛失・バックアップ	66
2 週目 (2)	アカウントの保護	62
3 週目	研究倫理	30
4 週目	脅威・暗号技術	49
5 週目	法律	29
6 週目	著作権	46
7 週目	ICT を使わない脅威・対策	82