

電子教材の閲覧データとコンテンツ 内容を用いた学習者のスコア予測

2023/2/24

兵庫県立大学 川嶋研 小岸沙也加

背景

講義ではオンライン上で講義資料が閲覧できる機能が使われる

詳細な閲覧データを取得することができる



学生の行動から理解度が推定ができれば
はやめのアプローチが学生にできるのでは？

閲覧データを使用した先行研究

毎週生徒の最終成績を予測し、学生のリスクの有無を分類する試み [1]

⇒ **情報のある週とない週の行動**を考慮し、データを変換させることで、高精度な分類モデルを導出している

どの学習行動が成績に関連するかを調べた研究 [2]

⇒ 特定の行動を要素として成績を予測し、上位・中位・下位に分類、**それぞれの傾向と差を観察する**ことでよりの確かなアドバイスをできる

[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.

アプローチ

コンテンツそのものが予測に繋がるか明らかでない
学生の行動はコンテンツが前提にあるのでは??

入力：閲覧データ、コンテンツ内容（講義資料）

出力：小テストごとに予測した点数

評価：小テストごとに求めたRMSE(Root Mean Square Error)の平均

コンテンツ内容を使用することで予測精度があがるのか

使用データ

九州大学講義（2020年 情報系科目）

閲覧データ（講義回数：7回、対象学生：100名、200,818ログ）

コンテンツ（スライド）

2週目に2回
行われている

小テストのデータ（小テスト回数：8回、5択択一式）

（学生番号・問題文・学生の選択・正解かどうか・提出時間）

スコア予測手法

ベースライン

閲覧データから各ページにおける各操作の操作回数、閲覧時間を求める

⇒ 行動特徴ベクトル

行動特徴ベクトルのみを使用

userID	Open 1	Close 1	...	Next 15	Prev 15
0	3	0		4	3
1	2	1		3	2

次元数はページ数 \times (12 (行動) + 1 (ページ閲覧時間))

↑ 行動特徴ベクトル

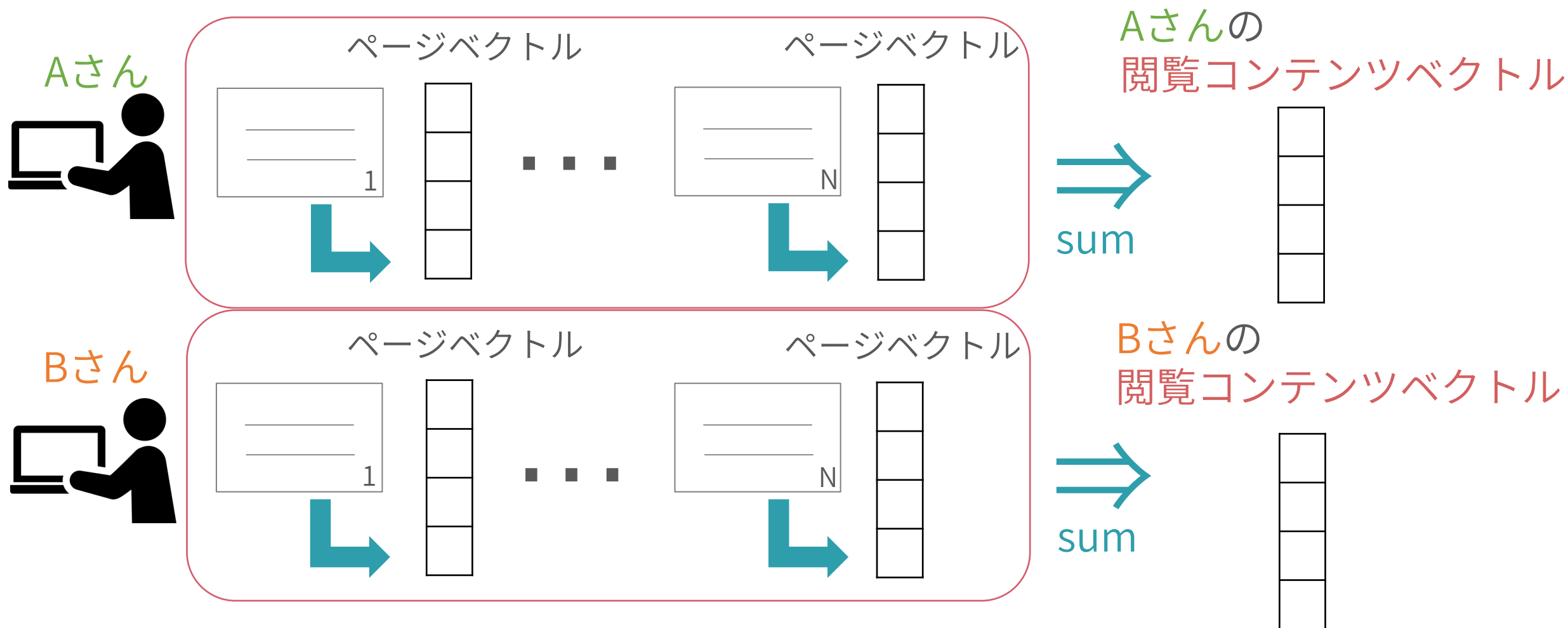
学生ごとに閲覧時間の長いページの情報を含むコンテンツベクトル

⇒ 閲覧コンテンツベクトル

行動特徴ベクトル と 閲覧コンテンツベクトルを使用

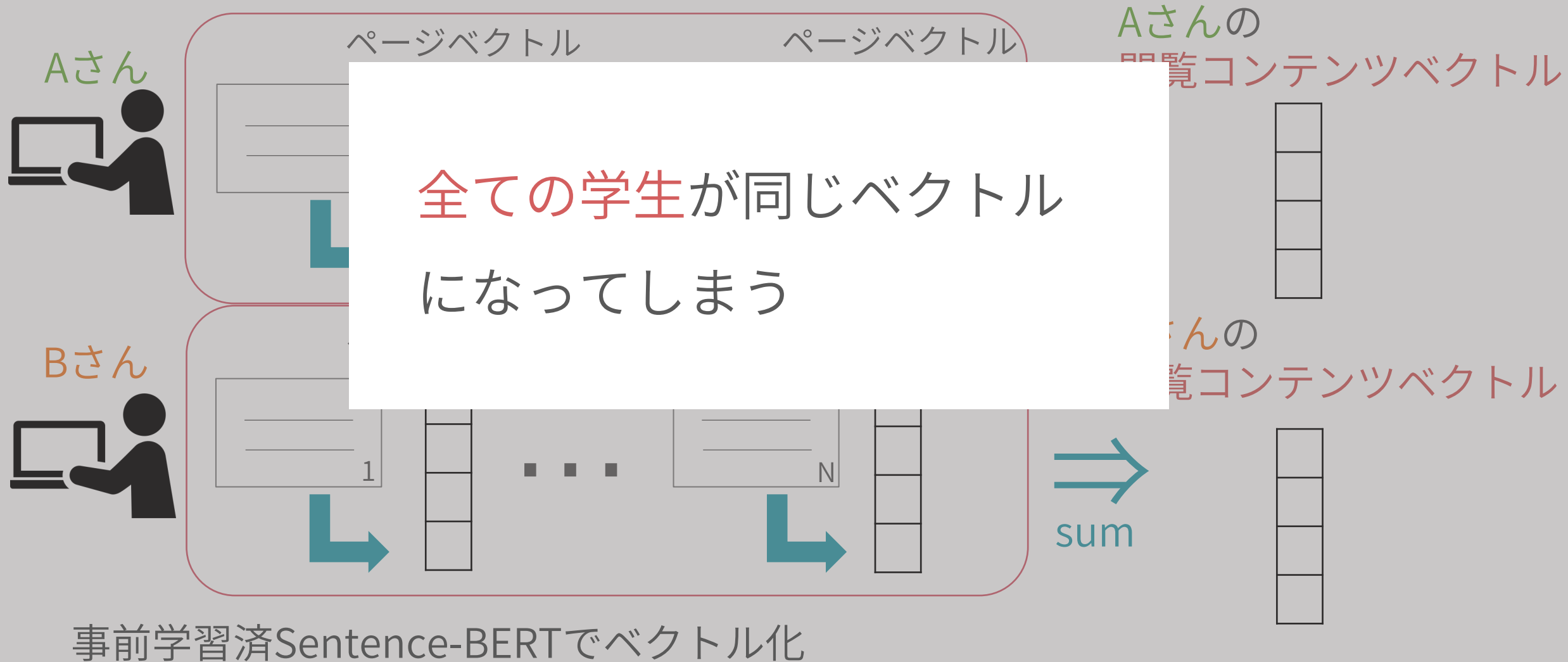
提案手法

閲覧コンテンツベクトル

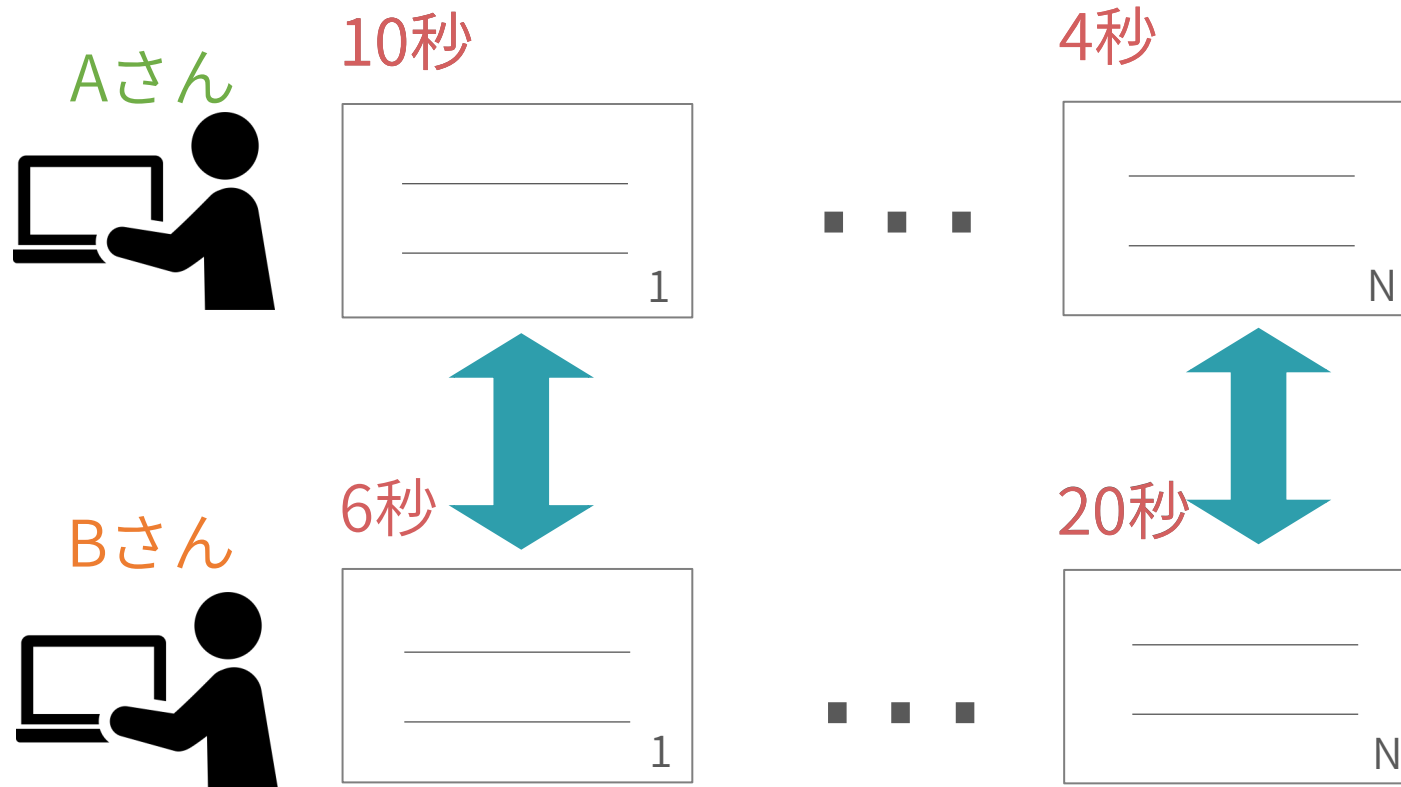


事前学習済Sentence-BERTでベクトル化

閲覧コンテンツベクトル



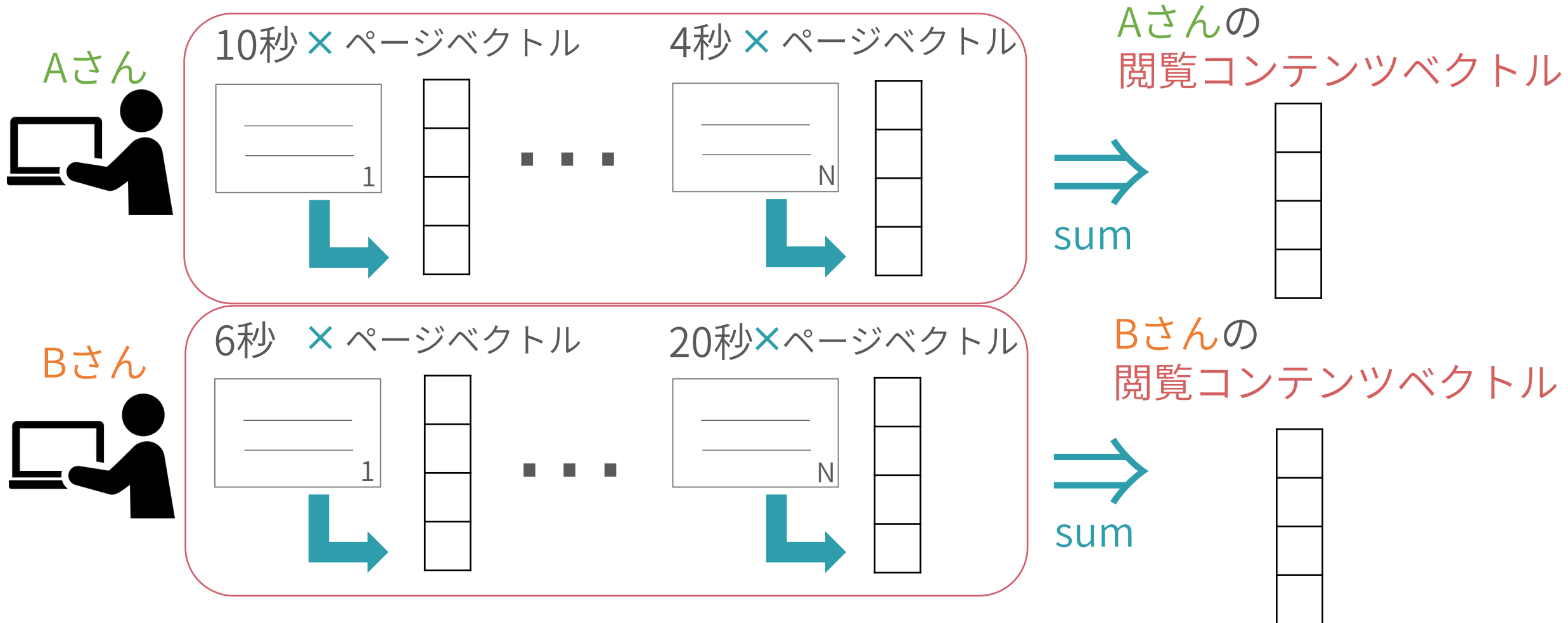
閲覧コンテンツベクトル



同じページでも閲覧時間は
学生ごとに異なる

閲覧時間に注目

閲覧コンテンツベクトル



1回の閲覧時間が5分より長いページは放置されたものとして省く

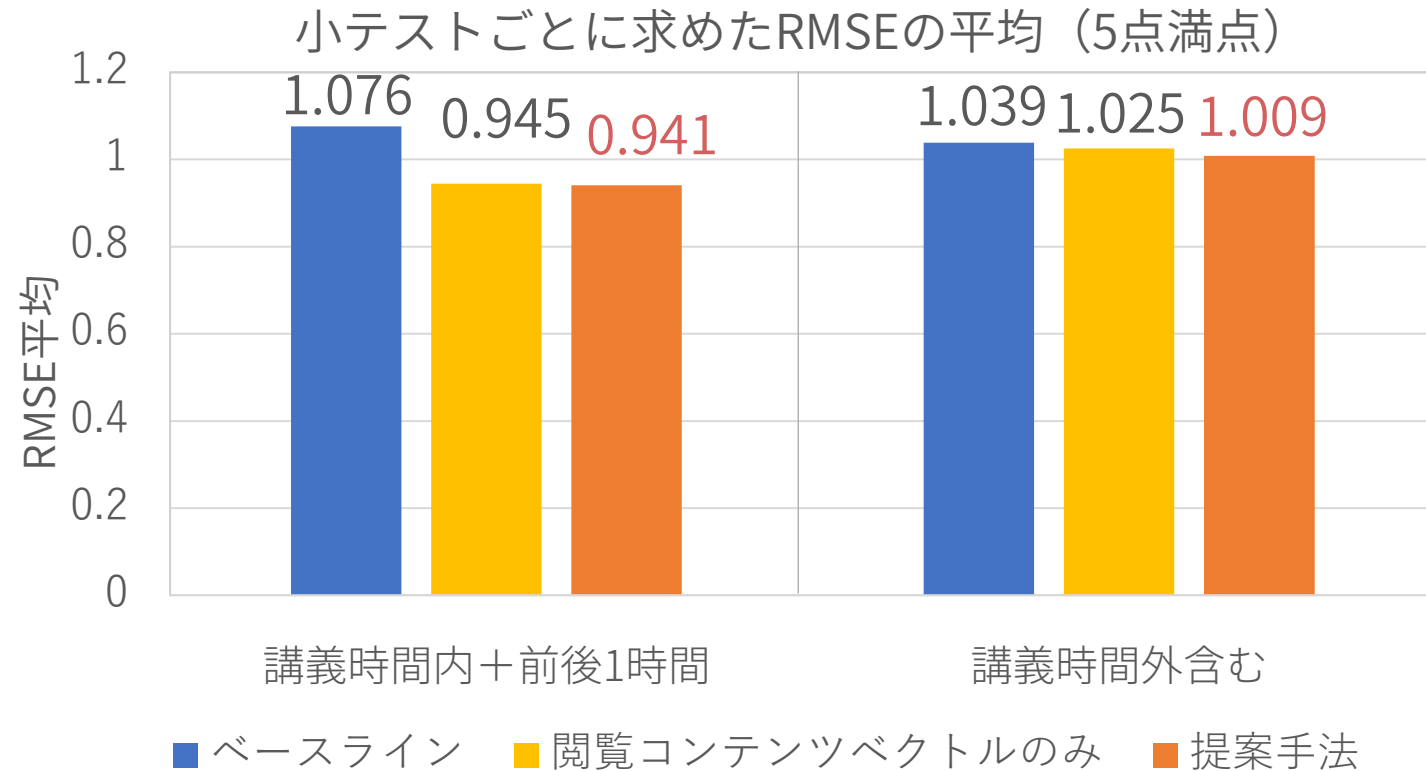
結果

5-foldの交差検証でRMSEを計算 手法別に比較

- 行動特徴ベクトルのみ
(ベースライン)
- 閲覧コンテンツベクトルのみ
- 行動特徴ベクトル
& 閲覧コンテンツベクトル

時間別に比較

- 講義時間内+前後1時間
- 講義時間内+講義時間外全て



コンテンツ情報を含めることが
スコア予測に貢献する

まとめ

コンテンツ情報を含めることでスコア予測精度向上を目指す

コンテンツ情報を含めることは予測精度向上に繋がる

長く閲覧したコンテンツ情報が予測精度向上に繋がる可能性がある

今後の課題

ベクトルの次元数を減らす

ページベクトルに画像の情報や配置の情報を含める

問題1問ごとに正解 / 不正解を予測する