

電子教材の閲覧データと コンテンツを用いた点数推定

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できれば
はやめのアプローチが学生にできるのでは？

問題設定

電子教材の閲覧データと コンテンツを用いた点数推定

入力：週ごとに取得した閲覧データ、コンテンツ内容

出力：小テストの点数（週ごと）

評価：RMSE

最終成績や小テストに頼りがちな理解度推定

⇒ 閲覧データとコンテンツでどこまで推定できるのか

使用データ

九州大学講義（2020年 サイバーセキュリティ基礎論）

閲覧データ（7週間、100名、200,818ログ）

コンテンツ画像・内容（npzファイル）

小テストのデータ（問題文・正解かどうか・提出時間）

全体像

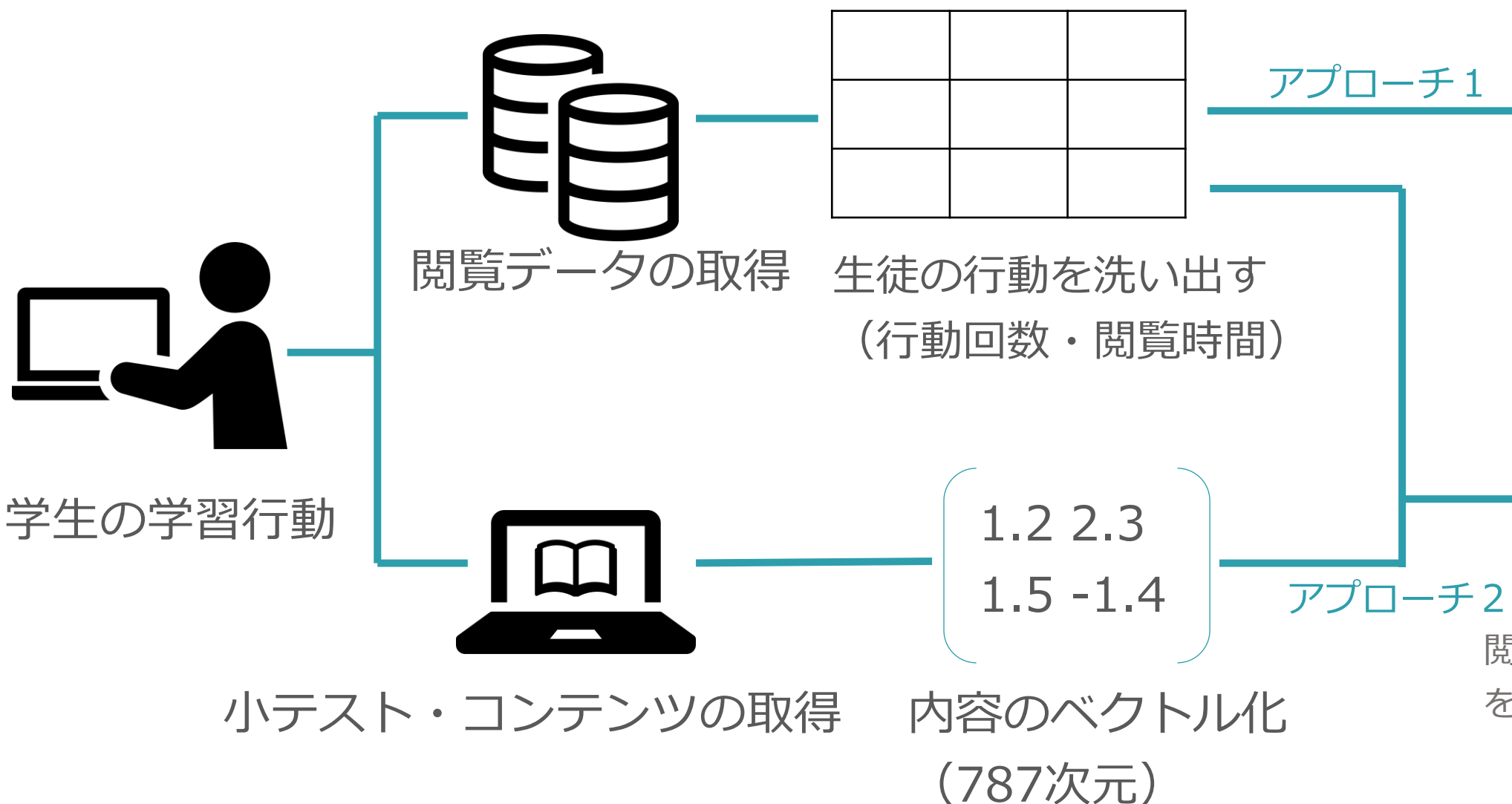
閲覧データのみを使用した場合

0:4	2:5
1:3	3:2

LightBGM等で
点数の推定
RSMEで評価

0:5	2:5
1:3	3:4

閲覧データとコンテンツ
を使用した場合



アプローチ

アプローチ 1

閲覧データのみを使用

週関係なく40点満点

週ごとにわけて5(10)点満点

行動を講義時間内に絞って

etc.

アプローチ 2

閲覧データとコンテンツ内容を使用

方針 1: 特徴量にベクトルを入れる

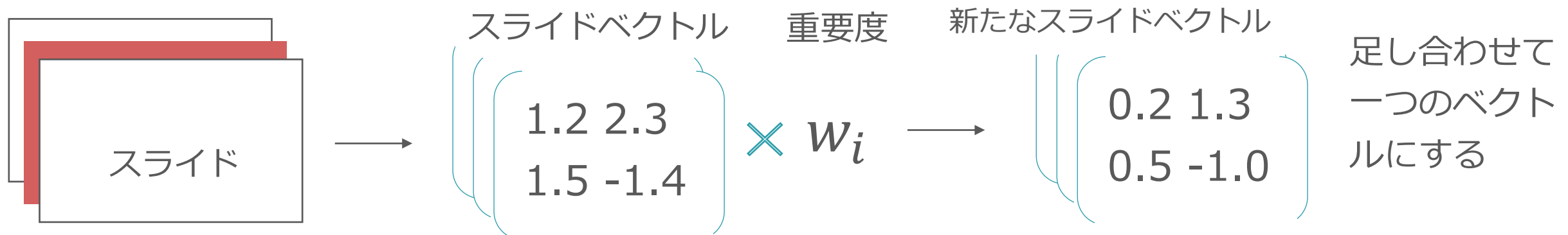
方針 2: 行動特徴量に重みづける
(重み計算のみに使う)

方針 3: 行動の重要度及びスライドの
重要度によって重みを変える

方針 1 : 特徴量にベクトルを入れる

- 1, 学生ごとに閲覧時間が長いスライドほど重要度が高くなる
閲覧時間が長すぎる部分は省く
- 2, 小テストによく関係しているスライドほど重要度を高くする

コサイン類似度を使用



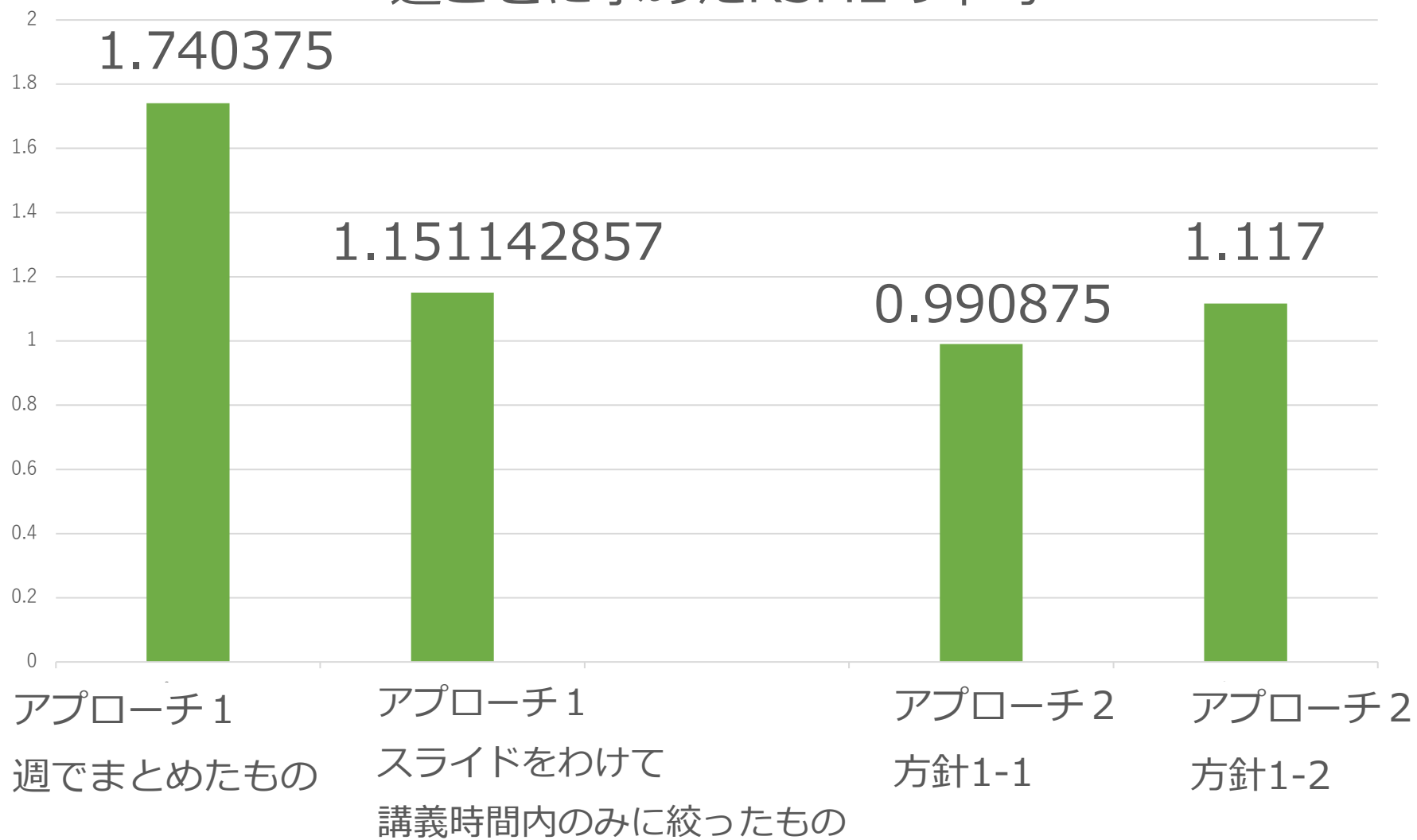
方針 2 : 行動特徴量に重みづける

方針 1 で求めたスライドごとの重要度を行動特徴量に重みづける

userid	Open_1	Close_1	Next_1	...	Open_4	Close_4	...	Open_15	Close_15
0	3	0	3		0	2		0	3
1	2	1	2		1	1		0	2
2	0	0	1		0	0		0	1
3	1	2	4		1	0		0	2
	×	×	×		×	×		×	×
重要度	w_1	w_1	w_1		w_4	w_4		w_{15}	w_{15}

結果

週ごとに求めたRSMEの平均



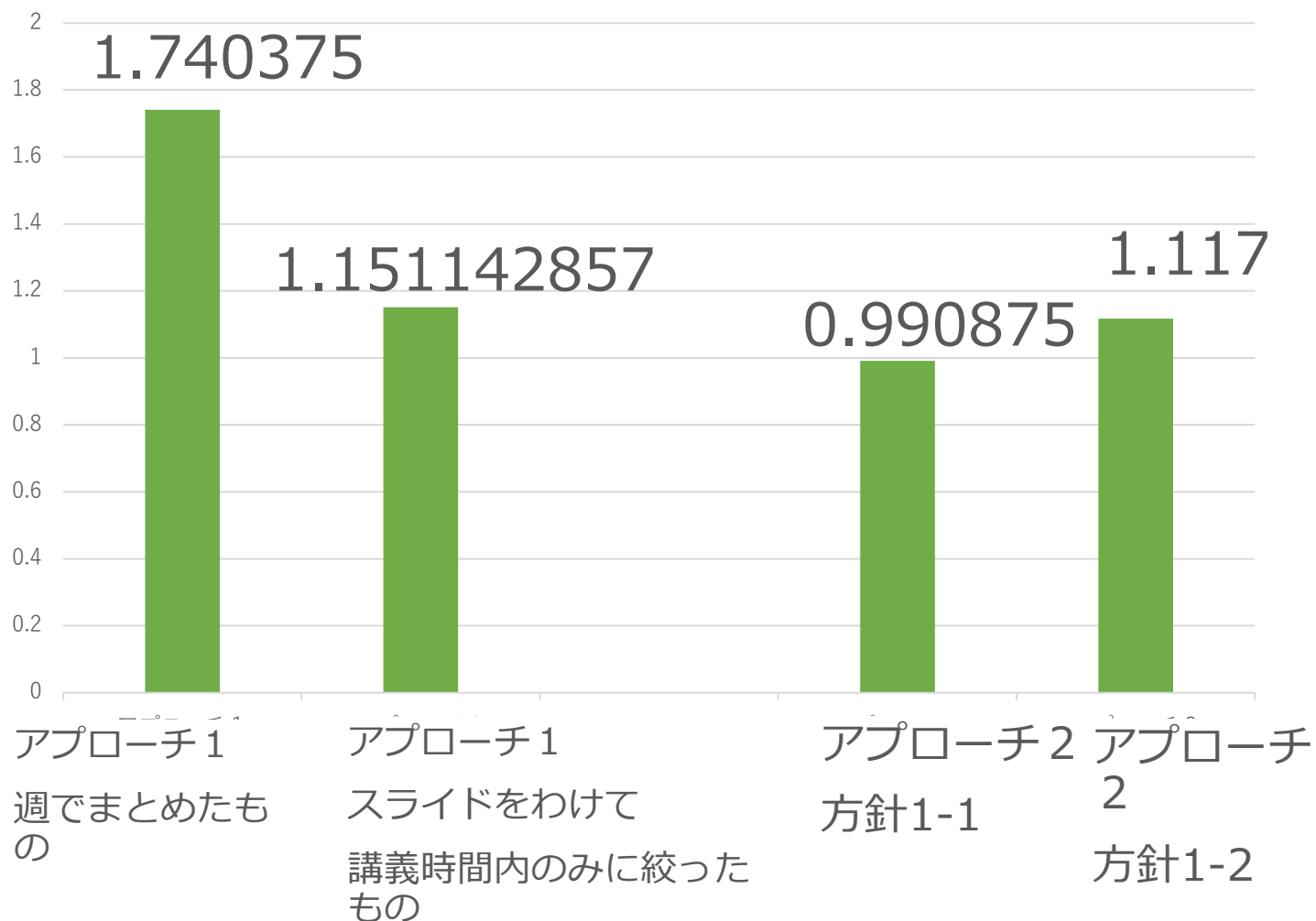
結果

閲覧時間がより重要？

スライドごとでの行動も
大切

方針 3 に期待

週ごとに求めたRSMEの平均



今後の課題・計画

課題

方針 3 について

問題点・明らかにしたい部分がふわふわしてきた

既存のものと比較ができていない

計画

10月 方針 3 をすすめる

11月 評価実験、論文執筆

12月 論文執筆、修正

結果(P9の補足)

小テスト	アプローチ 1	アプローチ 1		アプローチ 2	アプローチ 2
1	1.554	0.905		0.989	1.708
2-1	1.789	1.780		1.172	0.739
2-2	1.809			0.808	1.346
3	1.888	1.165		0.494	1.175
4	1.738	1.001		1.082	0.867
5	1.747	0.926		0.883	0.987
6	1.604	0.830		0.969	0.790
7	1.794	1.451		1.530	1.324