

高等教育におけるデータ活用

島田敬士

九州大学

大学院システム情報科学研究所 教授

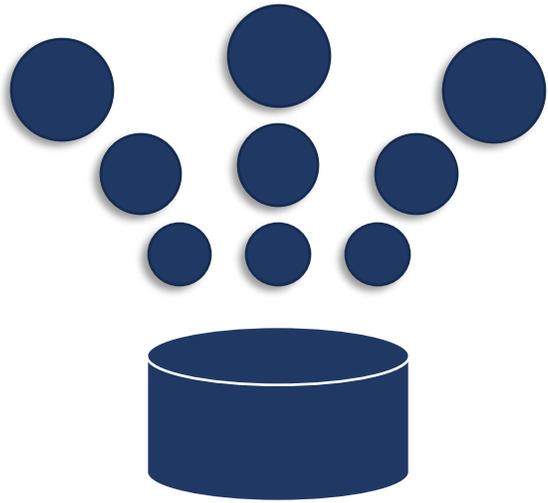
ラーニングアナリティクスセンター センター長



KYUSHU UNIVERSITY

教育データ活用の歩み

基盤・仕組みづくり



データ活用の黎明期



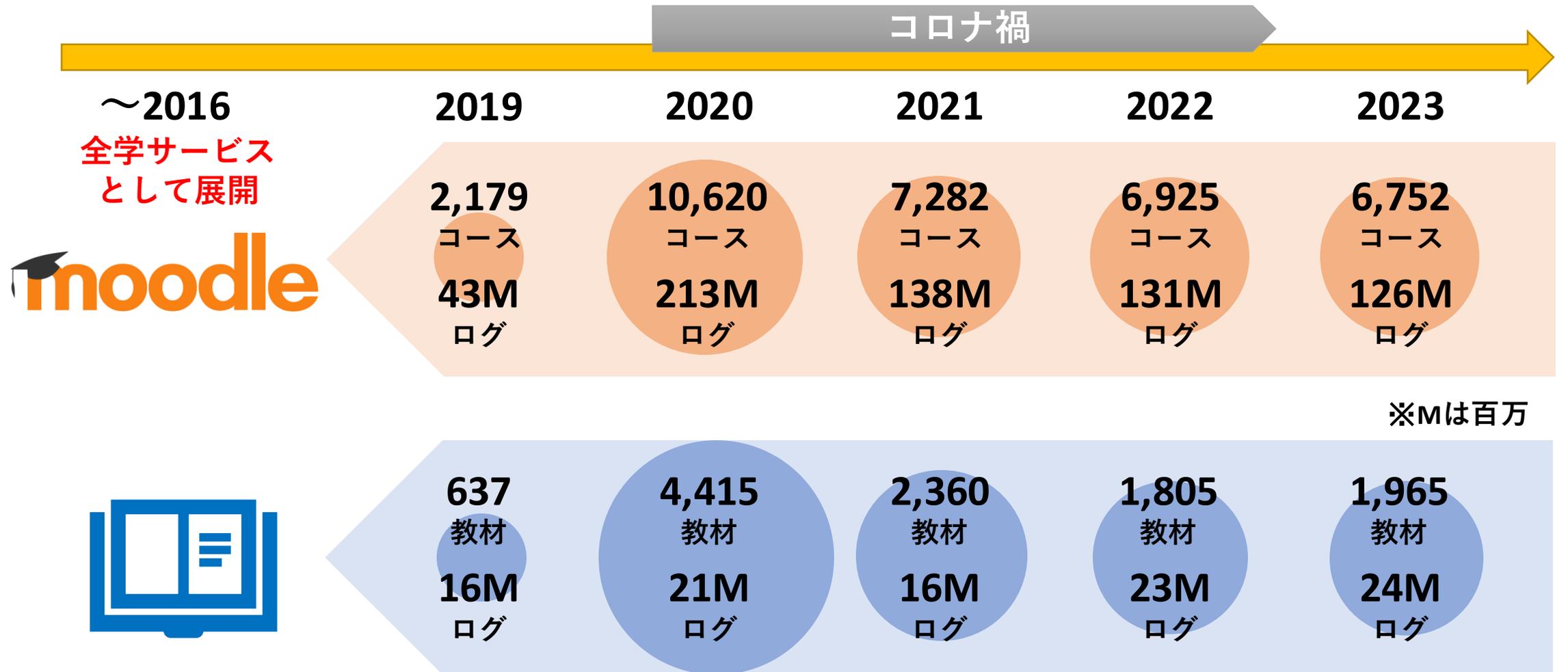
データ活用の拡大



~2022?

現在?

Moodle/eBookの利用状況



学習ダッシュボード : Metaboard



ツール群

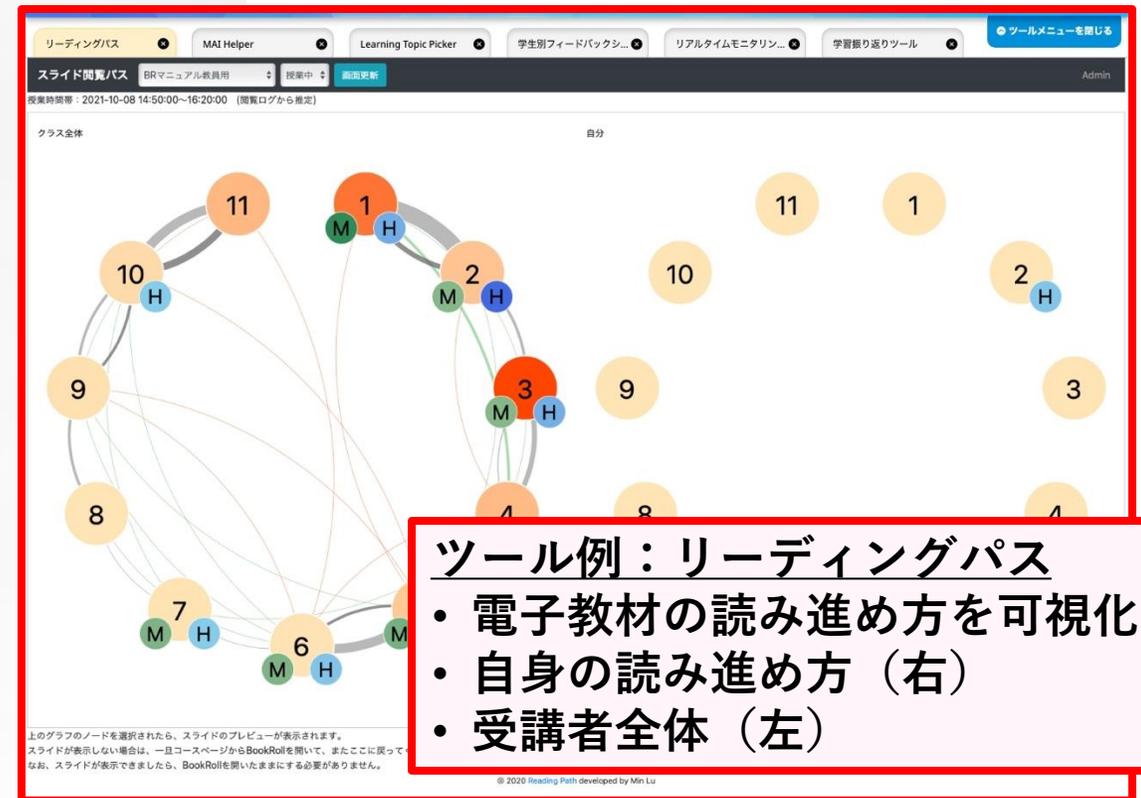


Metaboard

ツールを選択してください

- プラットフォーム
 - 分析機能を持たない
 - ツールを登録・管理
- ツール
 - データ分析
 - 分析結果の可視化

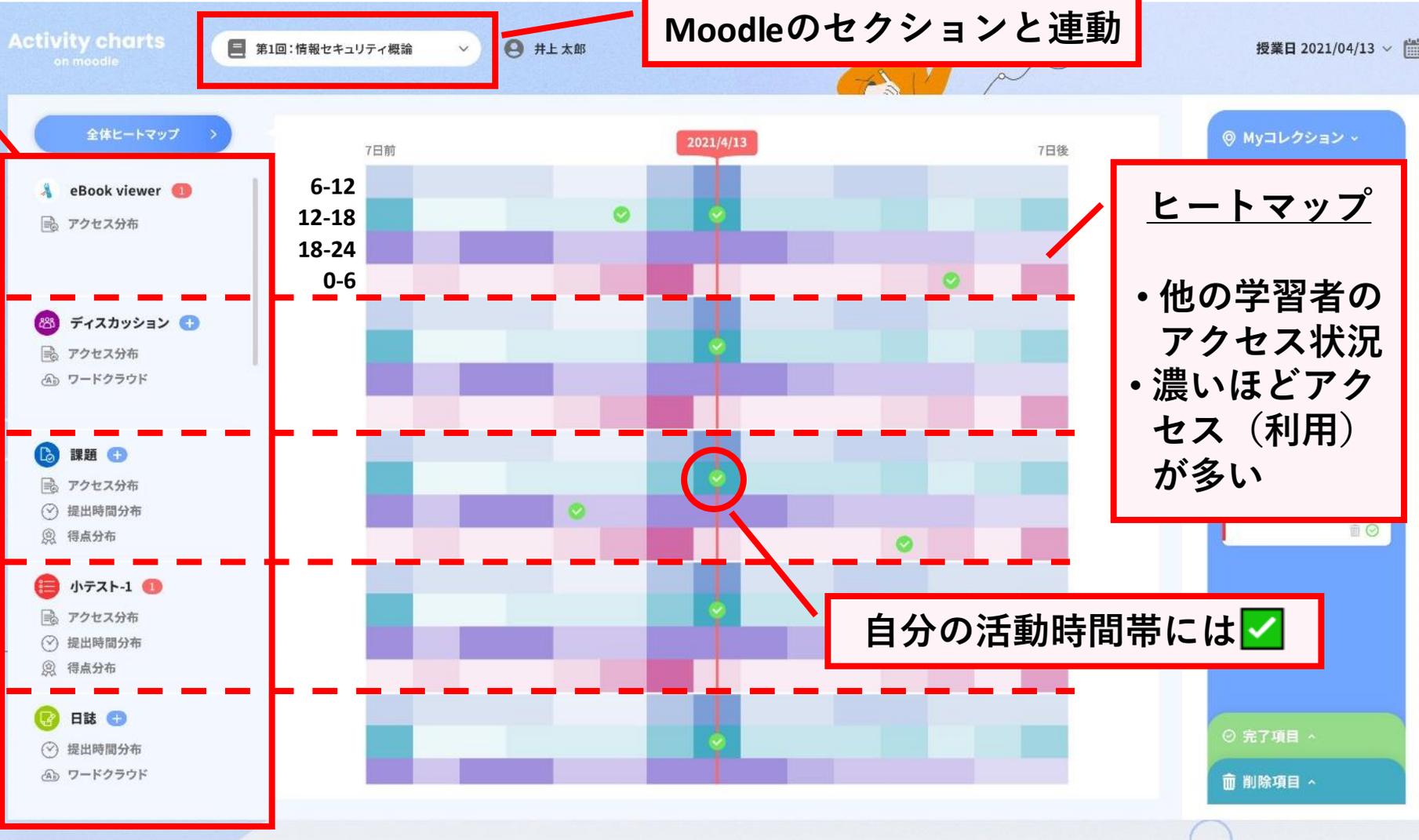
- データ連携
 - 学習管理システム
 - 電子教材配信システム



Activity Charts : 学習活動の可視化 (詳細)

Moodleのセクションと連動

各セクション内に配置された活動



ヒートマップ

- 他の学習者のアクセス状況
- 濃いほどアクセス (利用) が多い

自分の活動時間帯には

学習計画・リフレクション管理：MAI Helper

MAI Helper

山田 政寛

2018年度・他・福岡市教育委員会テストコース (山田 政寛)

Schedule List

登録されている現在以降のスケジュールはありません。

Home ホーム

Schedule スケジュール入力

Reflection MAI Helper入力

Home ホーム

Schedule スケジュール入力

Reflection MAI Helper入力

2024年09月02日(月)

BQ MB MH

BQube MetaBoard

9時

10時

11時

12時

13時

14時

15時

16時

17時

内省を入力

記録

● 今回の授業概要

● 今回の授業内容に対する自分の意見・考え

● 今回の授業受講の予習復習は計画通りにできたか?

● 今回の授業受講に向けて、計画通りにできなかった場合、なぜできなかったか記載してください

● 次の授業に向けて、意識したいことは何か?(予習時、授業中、復習時など)

● 質問があれば質問を書いて下さい

保存

2024年06月27日(木) 11:38

2024年06月18日(火) 11:16

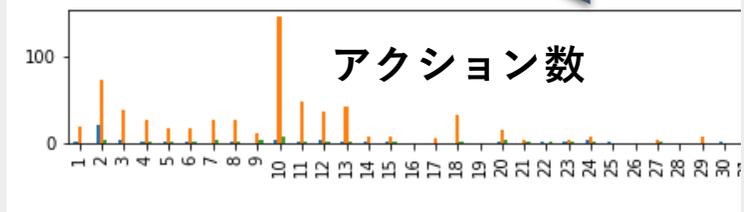
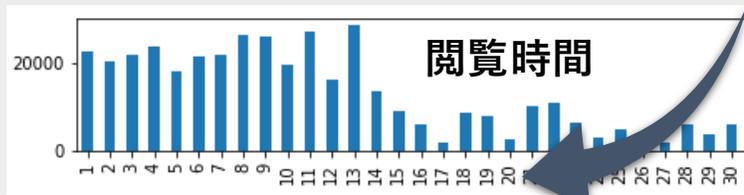
- 各システムへのアクセス状況を可視化
 - 電子教材配信システム
 - 学習ダッシュボード
 - MAI Helper

- 教員
 - リフレクションさせたいことを登録
 - 学生の記入状況を一覧
- 学生
 - 自身の学習活動の確認
 - 他者の状況を見ながら学習計画作成

データに基づく成績予測：従来課題

成績予測に用いるデータ種

Student ID	Contents ID	page	operation	Time	...
X-2022_Cn	X-2022_Ua	1	NEXT	2022-04-17 hh:mm:ss	...
X-2022_Cn	X-2022_Ua	2	NEXT	2022-04-17 hh:mm:ss	...



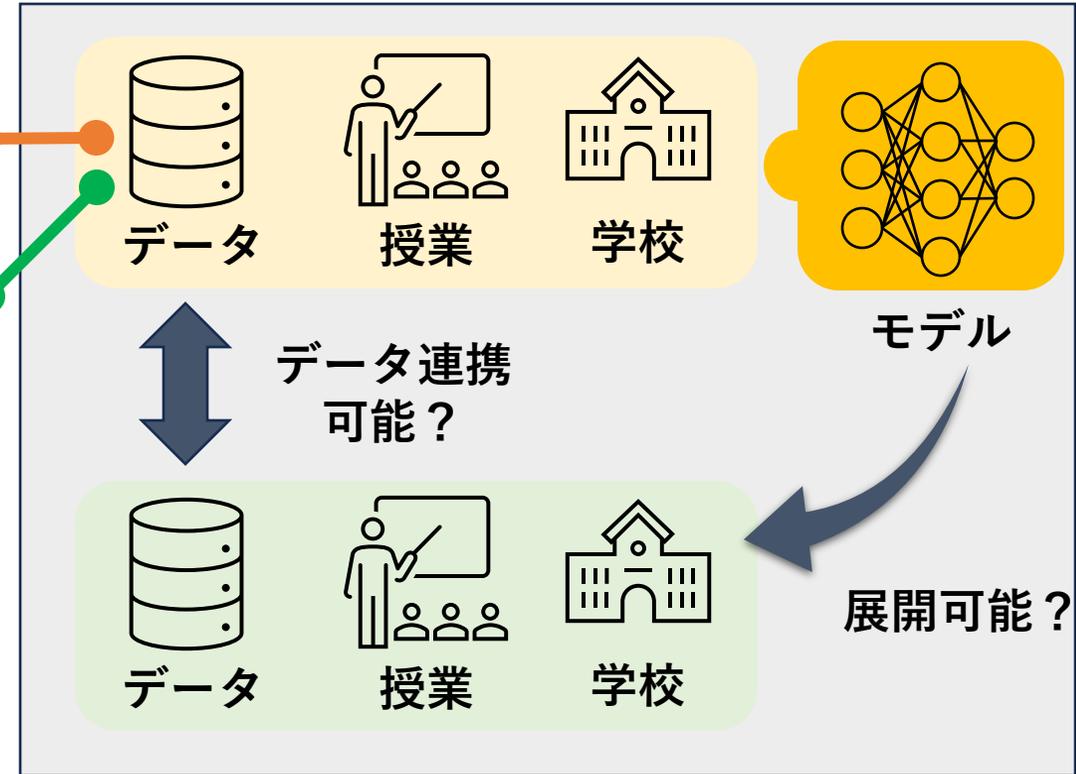
集計

...

...



技術展開／データ連携



△システムアクセスログが主流

○データ収集が容易

×記録形式がシステム依存



×組織、授業毎にモデル構築

×他に展開すると精度低下

×データ連携はプライバシー保護の観点から困難

自由記述データから成績予測

学習者 

- LLMを用いた自動リフレクション評価
- リフレクションスコアを用いた成績予測

学習日誌 

✗ RawテキストからBERTで成績予測

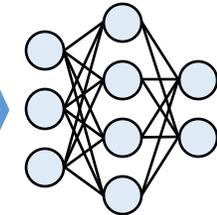
F1-score
Zero-shot : 66.7%

Single-Agent

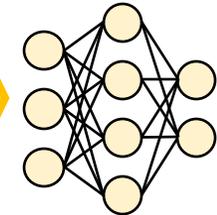
スコア	評価基準
3 (★★★★)	具体的で詳細な内容
2 (★★★☆☆)	一般的な内容
1 (★☆☆☆☆)	曖昧な内容
0 (☆☆☆☆)	記述無し, 無関係な内容

GT一致率
Zero-shot : 74.9%
Few-shot : 82.8%

F1-score
Zero-shot : 74.8%
Few-shot : 78.5%

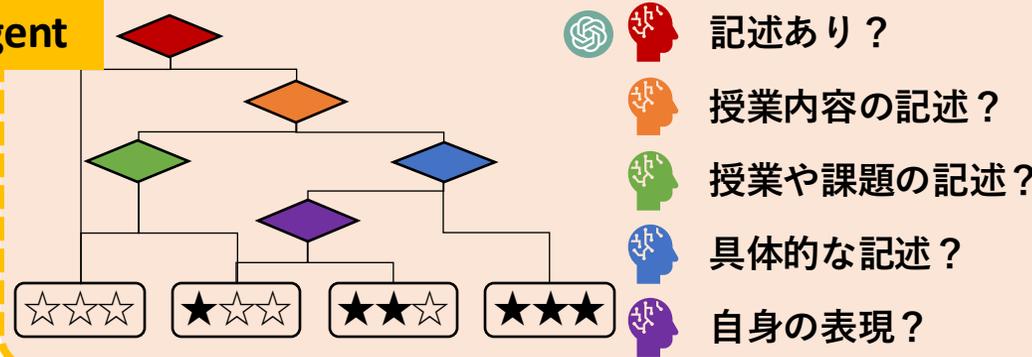


LSTMで
成績予測



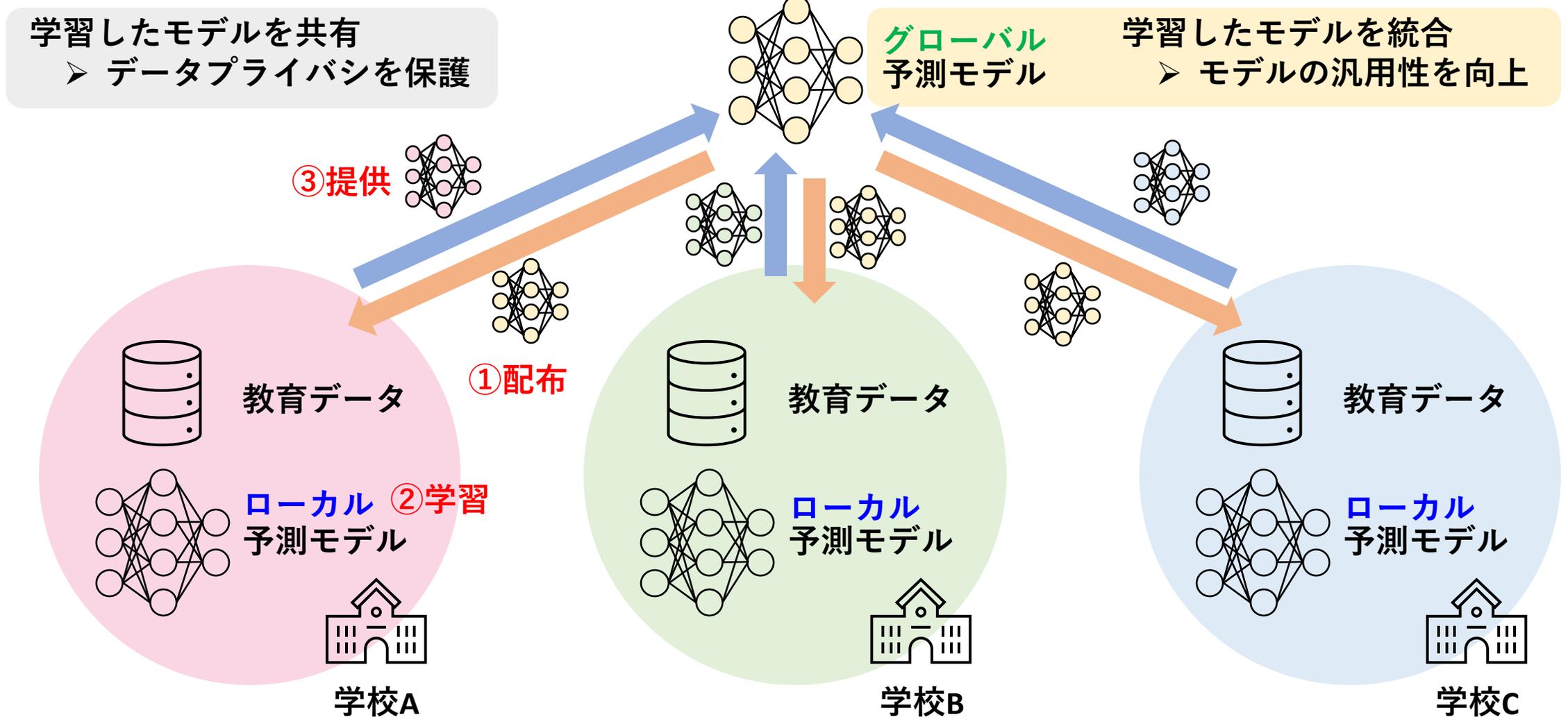
F1-score
Zero-shot : 75.6%
Few-shot : 74.1%

Multi-Agent



GT一致率
Zero-shot : 75.7%
Few-shot : 79.5%

プライバシーを考慮した組織間データ連携予測モデル



プライバシーを考慮した組織間データ連携予測モデル

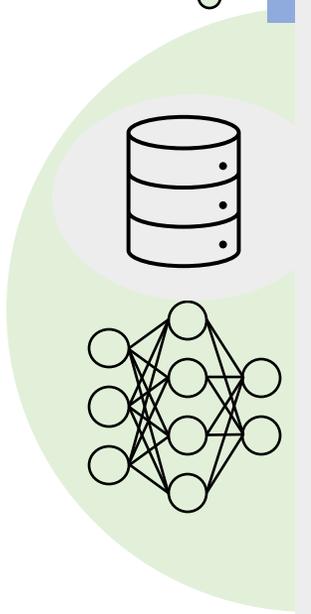
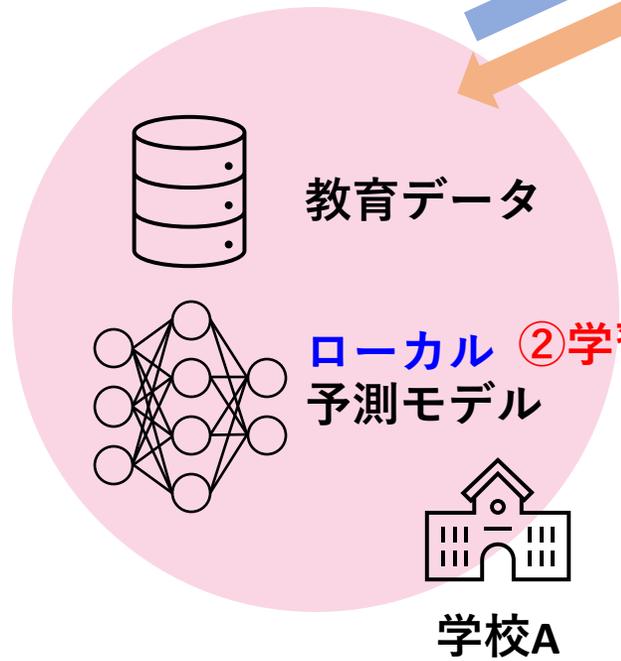
学習したモデルを共有
➤ データプライバシーを保護

グローバル
予測モデル
学習したモデルを統合
➤ モデルの汎用性を向上

③提供

①配布

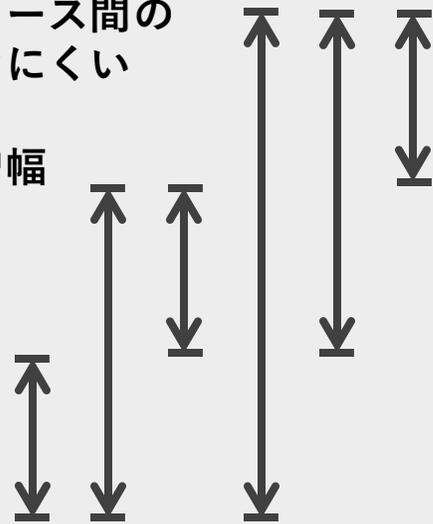
②学習



相対特徴を導入

学校間、コース間の
影響を受けにくい

データを増幅



	活動	成績
		90点
		60点
		85点
		65点

プライバシーを考慮した組織間データ連携予測モデル

学習したモデルを共有

- データプライバシーを保護

グローバル
予測モデル

学習したモデルを統合

- モデルの汎用性を向上

		相対	絶対	早期
Top n	n = 5	1.0	0.60	0.80
precision	n = 10	0.70	0.70	0.80
AUC		0.65	0.52	0.65

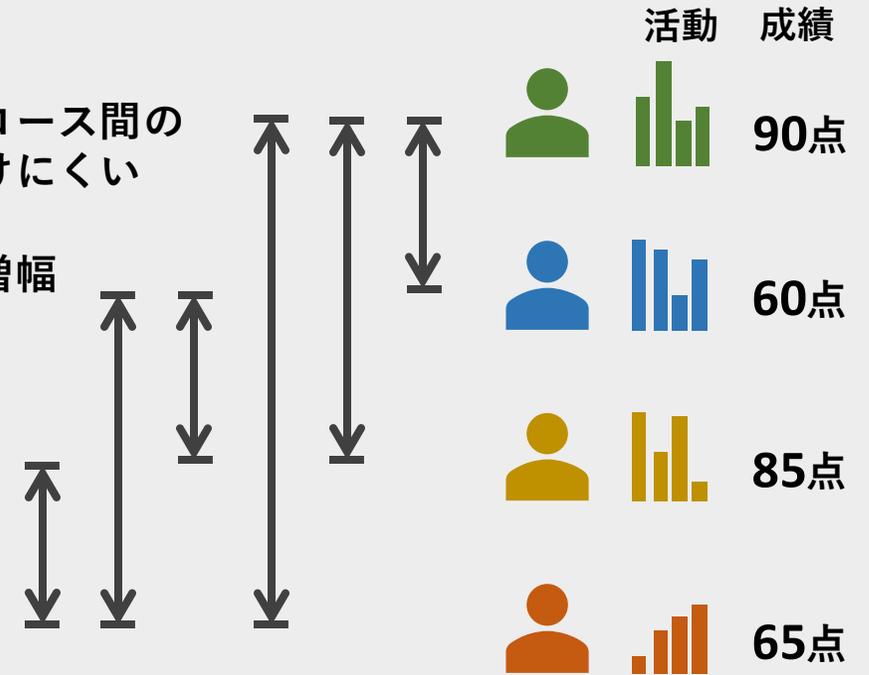
- 連合学習を導入しても一括型学習と同等の性能を実現可能
- 相対特徴の導入で性能向上
- 授業期間の途中（序盤）までのデータのみで予測しても高精度

汎用展開できる予測モデルで
早期のAt-risk学生検知が可能

相対特徴を導入

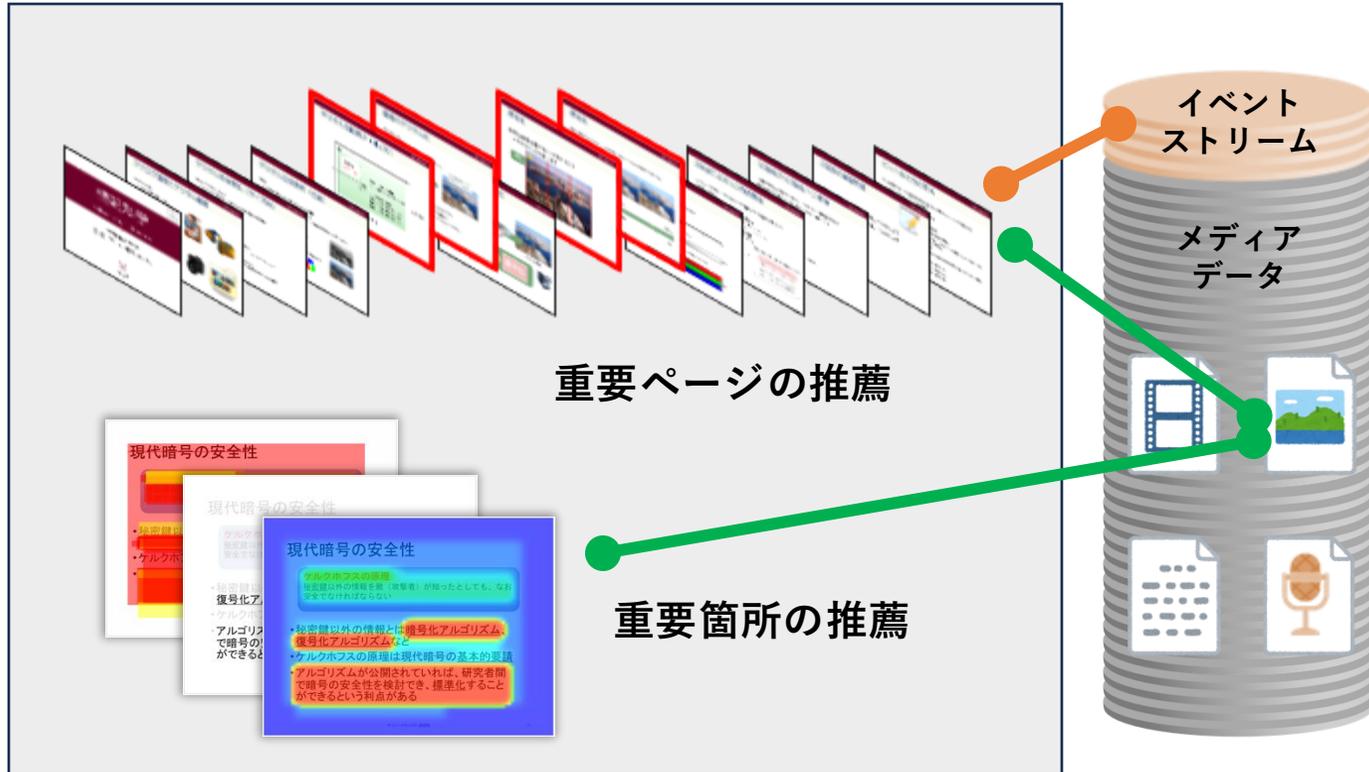
学校間、コース間の
影響を受けにくい

データを増幅



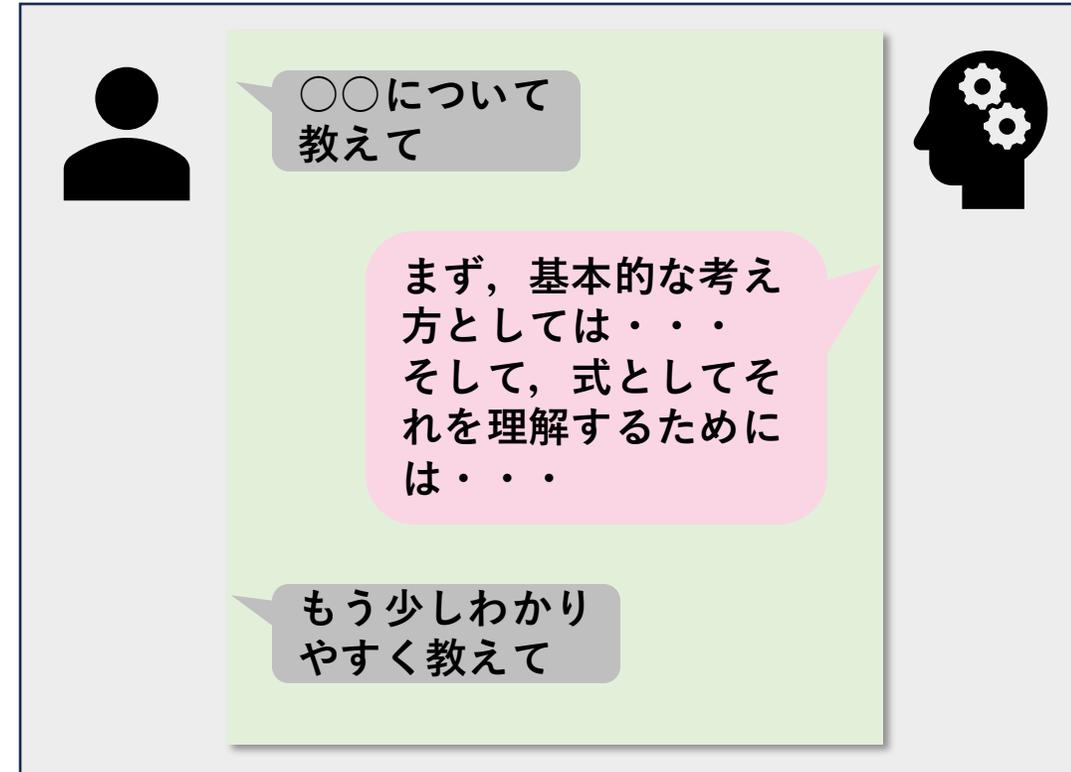
AIによる学習支援：従来課題

情報推薦



- ◎サポートは助かる
- ×内容を自力で理解する必要あり
 - もう少しかみ砕いてほしい

質問応答AI



- ◎なんでも答えてくれる
- ×回答が学習内容に合致しない
- ×依存性が高くなる可能性あり

AIと学習者が協働して学習支援

学習者



記事を投稿

他者が読むことを意識

ここでいう論理とは、思考の形式及び法則である。これに加えて、思考のつながり、推理の仕方や論証のつながりを指す。よく言われる「論理的に話す、書く」という言葉は、つながりを明確にし、論証を過不足なく行うということである。

論理学は、伝統的には哲学の一分野である。数学的演算の導入により、数理論理学（記号論理学）という分野ができた。現在では、数理論理学は数学と論理学のどちらであるとも（時にどちらでもないとも）される。現在の論理学は、（それを論理学であるとするなら）数理論理学と、数理論理学をふまえた論理学、数理論理学でない論理学に分化している。

弁証法なども、「論理」なのであるが、論理学における論理とは異なる。これらは、論理というよりむしろ理論 (Theory) である。

なお、日本語の「論理学」という語は西周によるものとされている。

デジタル画像表現 (色紙)

• 数式や図表を正確に再現する
• 印刷品質を向上させる
• 環境に優しい印刷方法を提供
• 印刷物のコストを削減する
• 印刷物の生産性を向上させる
• 印刷物の品質を向上させる
• 印刷物の生産性を向上させる

どのように学んだ？
どのように理解した？
を記事化して投稿

TAや教員によるレビュー

TA

記事の質を保証

投稿数: 3,600+
公開数: 2,700+
※2023年6月時点

大学図書館



学習記事

ノウハウの伝承

記事を参照

他者の観点から学習



学習者



- 同じ授業を受けている学習者同士で
- 同じ授業を受けた学習者から世代を超えて



- 大量のノウハウ記事をどのように活用する？

AIと学習者が協働して学習支援



学習者

記事を投稿

他者が読むことを意識

ここでいう論理とは、思考の形式及び法則である。これに加えて、思考のつながり、推理の仕方や論証のつながりを指す。よく言われる「論理的に話す、書く」という言葉は、つながりを明確にし、論証を過不足なく行うということである。

論理学は、伝統的には哲学の一分野である。数学的演算の導入により、数理論理学（記号論理学）という分野ができた。現在では、数理論理学は数学と論理学のどちらでもある（時にどちらでもないとも）される。現在の論理学は、（それを論理学であるとするなら）数理論理学と、数理論理学をふまえた論理学、数理論理学でない論理学に分化している。

弁証法なども、「論理」なのであるが、論理学における論理とは異なる。これらは、論理というよりむしろ理論 (Theory) である。

なお、日本語の「論理学」という語は西語によるものとされている。

どのように学んだ？
どのように理解した？
を記事化して投稿

*TAや教員によるレビューで記事の質保証

投稿数: 3,600+
公開数: 2,700+
※2023年6月時点

大学図書館



記事を参照

他者の観点から学習

学習記事

ノウハウの伝承



学習者

回答

質問

DSP_Q&A

Start Chat

POWERED BY Dify.

生成AIが回答を生成

- 教材知識に基づく回答で正確性を保証
- 記事知識に基づく回答で新たな観点, 洞察を提供
- 知識利用でNSSスコアが約15ポイント上昇

* NSS : Net Sentiment Score
- 100~100の範囲: 高いほど好反応



AI

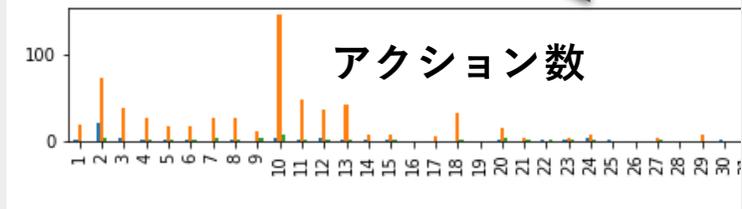
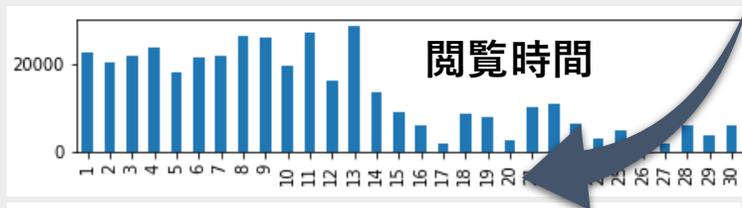
教科書

分析結果の教育現場への展開：従来課題

単純集計型

学習者の活動記録データ

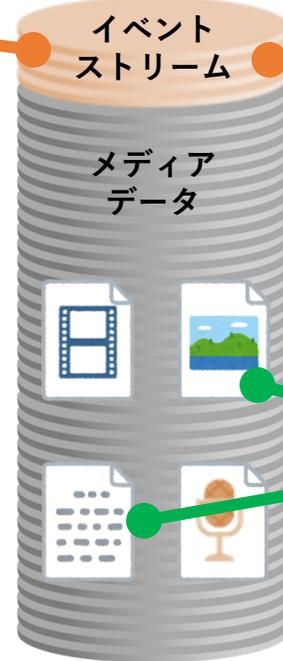
Student ID	Contents ID	page	operation	Time	...
X-2022_Cn	X-2022_Ua	1	NEXT	2022-04-17 hh:mm:ss	...
X-2022_Cn	X-2022_Ua	2	NEXT	2022-04-17 hh:mm:ss	...



集計

● ● ●

● ● ●



詳細分析型

学習者 ■ 閲覧 ■ 未閲覧



学習テーマ	第1因子	第2因子	第3因子	第4因子
周期信号	0.246	0.211	0.214	0.374
非周期信号	0.167	0.019	0.019	0.019
フーリエ級数	0.294	0.221	0.003	0.003
フーリエ変換	0.167	0.583	0.121	0.212
離散時間信号	0.544	0.337	0.039	0.348
Z変換	0.26	0.722	0.229	0.387
逆Z変換	0.791	0.122	0.091	0.41
離散時間システム	0.254	0.013	0.156	0.067
	0.308	0.046	0.134	0.211

教材閲覧ログ、授業内外の活動、成績などを対応付けて細かく分析

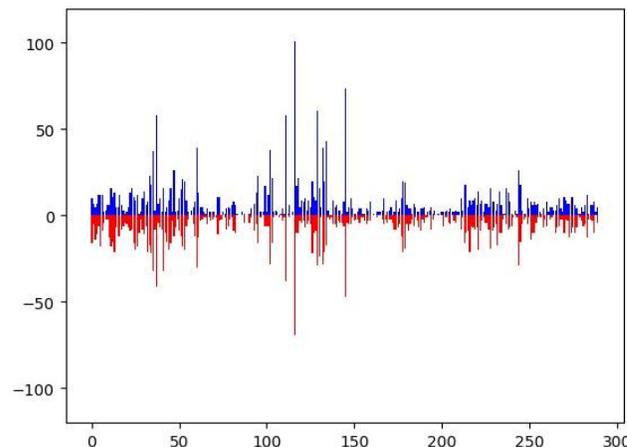
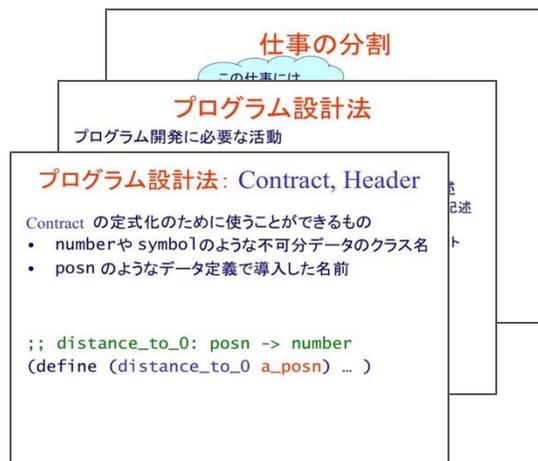


- ◎結果は分かりやすい
- ×細かいことは分からない
 - 学習内容, 学習の進め方など



- ◎細かいところまで分析
- ×情報が多すぎて解釈性が欠如
 - 要するに??

深層分析を通じた解釈性の高い分析結果のフィードバック



Contents cluster 9

プログラミング論Iの概要と基礎概念を学ぶ際に、成績の良い学習者は教材を多様かつ活発に探索し、成績の悪い学習者よりもじっくり学習する傾向があります。

Contents cluster 16

構造体データの定義と操作に関して、成績の良い学習者はページを素早く確認し、次のページに進む操作が多い傾向がある一方、成績の悪い学習者は特定のページに効率的にアクセスするジャンプ操作を多用しています。

Contents cluster 21

プログラミングにおける関数作成と抽象化に関連する教材の学習において、成績の良い学習者は多様な操作を行いながらじっくりと学び、成績の悪い学習者は短時間で急ぎ足でページを遷移して学習する傾向が見られました。

類似トピックに対して

成績群ごとに学習行動の違いを

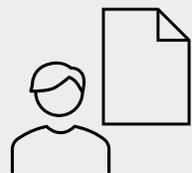
ユーザが解釈しやすい表現で出力



コンテンツ



閲覧行動



AI/生成AIを活用した分析技術を開発

深層分析を通じた解釈性の高い分析結果のフィードバック

仕事の分割

この仕事には

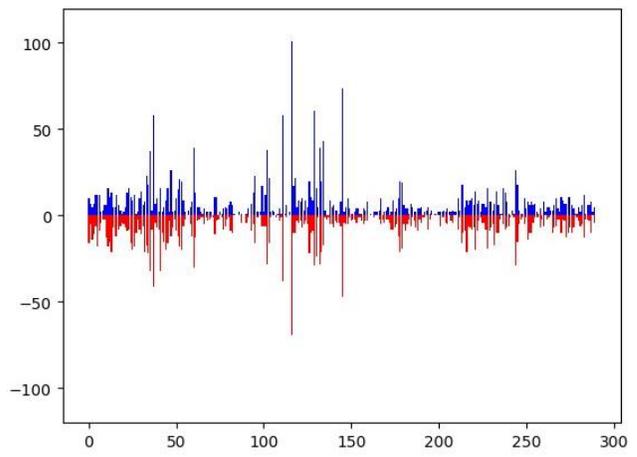
プログラム設計法
プログラム開発に必要な活動

プログラム設計法: Contract, Header

Contract の定式化のために使うことができるもの

- numberや symbolのような不可分データのクラス名
- posnのようなデータ定義で導入した名前

```
;; distance_to_0: posn -> number
(define (distance_to_0 a_posn) ...)
```



Contents cluster 9
プログラミング論Iの概要と基礎概念を学ぶ際に、成績の良い学習者は教材を多様かつ活発に探索し、成績の悪い学習者よりもじっくり学習する傾向があります。

Contents cluster 16
構造体データの定義と操作に関して、成績の良い学習者はページを素早く確認し、次のページに進む操作が多い傾向がある一方、成績の悪い学習者は特定のページに効率的にアクセスするジャンプ操作を多用しています。

Contents cluster 21
プログラミングにおける関数作成と抽象化に関連する教材の学習において、成績の良い学習者は多様な操作を行いながらじっくりと学び、成績の悪い学習者は短時間で急ぎ足でページを遷移して学習する傾向が見られました。

類似トピックに対して

成績群ごとに学習行動の違いを

ユーザが解釈しやすい表現で出力



コンテンツ

Contents ID	Page num	Content vector
X-2022_C1	1	[0.01, 0.43, ...]
X-2022_C1	2	[0.22, 0.52, ...]
....

教材の中身を細かく分析

説明文を生成

成績高群と低群で、学習方法が異なるトピックを特定

コンテンツや閲覧行動をベクトル表現化

閲覧行動

Student ID	Contents ID	pages	unit	Unit vector
X-2022_Ua	X-2022_C1	[2]	Nm	[0.21, 0.03, ...]
X-2022_Ua	X-2022_C1	[3,4]	NsNm	[0.25, 0.12, ...]
...

教材の読み進め方を細かく分析

成績

Contents ID	Grade
X-2022_Ua	A
X-2022_Ub	D
....	...

The screenshot shows the OpenLA documentation website. At the top, there is a search bar labeled "Search docs" and a home icon with the text "OpenLA". Below this, the site is organized into sections: "OPENLA:" with links for Introduction, OpenLA modules, Tutorial of modules, Code example, Re-implementation, Use OpenLA in Google Colaboratory, Contributors, and Release Notes; "OPENLA FEATURE REPRESENTATION:" with links for OpenLA Feature Representation Tutorial, Contributors, and Release Notes; and "LIMU:" with a link for Developed at LIMU. A "downloads 32k" badge is visible at the bottom of the sidebar.

モジュール・クラス・関数の詳細

モジュールの概要

クラス・関数の具体的な利用例

Google ColabでのOpenLA利用

特徴表現作成機能

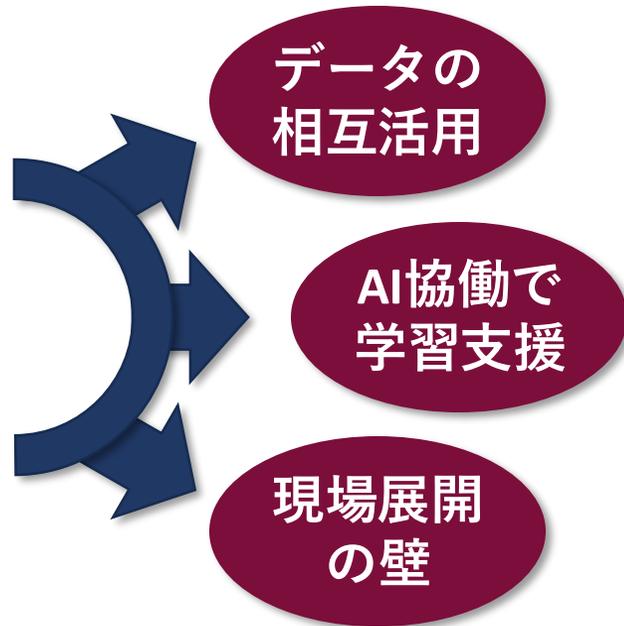


実装時の
コード量を

約6割
削減

まとめ：高等教育におけるデータ活用に向けた研究

データ活用の拡大



データの拡大

- 記述データを用いることでモデルの性能が向上
- 組織間データ連携でモデルの性能と安定性が向上

学習者参加型

- 学習者に合わせてAIが推薦資料を要約説明
- 資料の作成には学習者自身もコミット

緻密分析と要約

- 人手では時間を要する緻密な分析をAIが実施
- 分析結果を専門性に合わせて要約して提示

そのほかの研究成果についてはHPをご覧ください

<https://limu.ait.kyushu-u.ac.jp/>