



データサイエンス教育における 和歌山大学の取り組みの紹介

吉野 孝

和歌山大学

データ・インテリジェンス教育研究部門



1. データサイエンス教育の概要と特徴
2. 講義内容の紹介
3. 履修者を増やすための工夫

データサイエンス教育の 概要と特徴

和歌山大学のDS教育の特徴

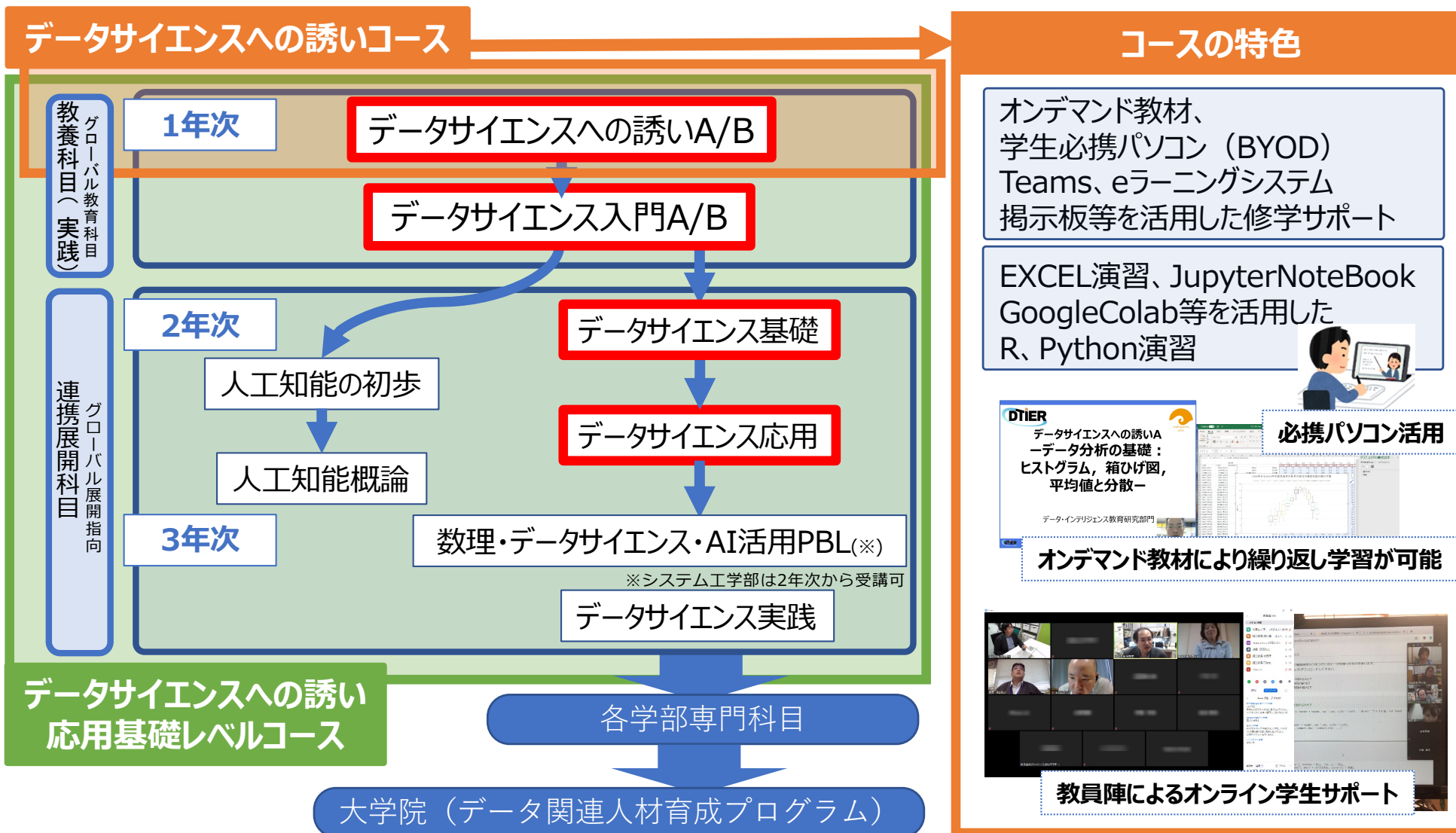
ゼロベースで全ての構成科目を立ち上げ
一貫した教育プログラムを構築

連携展開科目（教養教育と専門教育を接続する**全学対象の科目区分**）という新しい科目区分を利用
リテラシーレベルより**ハイレベルな内容を全学科目として展開**

部門教員が**連続して構成科目を担当**することで
レベルの異なる科目で**内容や実施形態に一貫性**を確保

一貫した教育プログラムを**オープンバッジ**というシステムで**可視化**
学生自身の**学修の管理及びプログラムの運用**に使用

- ✓ 「データサイエンスへの誘いコース」は、和歌山大学全学部必修科目のため受講率100%!
- ✓ 「データサイエンスへの誘い応用基礎レベルコース」は、教養科目・連携展開科目として全学部生が受講可



動画によるオンデマンド型授業

6

オンデマンドの講義

45分～60分

オンデマンドの演習

30分～45分

合計で
90分程度の動画

単回帰モデルのあてはめの考え方 (2/2) 11

直線Aと各データ点との差

↓

もっと良い直線はないだろうか?

数学的には

$$S_e = \sum (y_i - a - bx_i)^2$$

の最小値を求めることと同じ (最小二乗法)

実測データ: x_i, y_i
統計モデル: $\hat{y}_i = a + bx_i$
未知数: a, b

青い点線は残差

DTIER Data Intelligence Education and Research Division

オンデマンドは全て顔出し

DS_A5_1_v1_std.ipynb

単回帰分析と重回帰分析

フォントのインストール

GoogleColabでは、現在のところ、国に関して、日本語の表示がそのままでは出来ないため、まず、日本語を表示するためにフォントをインストールしてまいります。

下記の手順で日本語フォントの利用の準備をします。

```
install.packages("tidyverse")
install.packages("systemfonts")
remotes::install_github("Godevan-Aleksizide/fontregisterer", repos = NULL, type = "source")
require(tidyverse)
require(systemfonts)
system("apt-get install -y fonts-noto-cjk") # 日本語フォントのインストール
```

ここに代入して検索

GoogleColaboratoryや
Jupyter notebook上で実施

オンデマンド型講義のサポート体制

- Moodleの「質問コーナー」
(フォーラム)
- メールで個別の質問に対応
- オンラインサポート室
 - Teamsビデオ会議で実施
 - 毎週18:10～ (90分間)
 - 教員とTAによるサポート
- LINEチャットボットの導入
(1年生科目のみ)

オンラインサポート室
を一つに共通化し、
効率化

他の科目の受講生同士の交流も期待

多様なメディアを効果的に用いた
充実したサポート体制を構築

講義内容の紹介

データサイエンス 入門A/B

※この講義は、和歌山県データ利活用推進センターの教材の協力のもとに作成されています。

講義の概要（シラバス）

■ データサイエンス入門A

1. データサイエンスとは何か
2. **R言語**の基本操作と1変量の統計
3. 2変量の統計とクロス集計, 相関係数
4. データの演算, 基本統計量
5. データの視覚化 (1) : 棒グラフ, 円グラフ, 折れ線グラフなど
6. データの視覚化 (2) : 箱ひげ図, 散布図, その他のグラフ
7. 単回帰分析
8. 課題解説や総評等 (フォローアップ講義)

■ データサイエンス入門B

1. 重回帰分析
2. 決定木
3. k近傍法
4. アソシエーション分析
5. 主成分分析
6. k平均法
7. 階層クラスタリング
8. 課題解説や総評等 (フォローアップ講義)



受講生の内訳（2022年度後期）

■学部別受講生の内訳

□入門A 計315名

- 経済学部：41名
- 教育学部：7名
- システム工学部：256名
- 観光学部：11名

□入門B 計308名

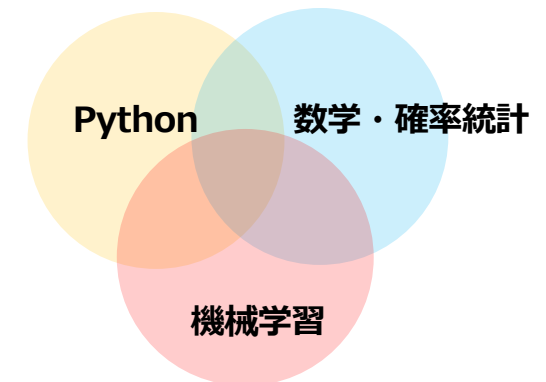
- 経済学部：42名
- 教育学部：9名
- システム工学部：251名
- 観光学部：6名

データサイエンス 基礎

講義の概要

Pythonを使ったデータサイエンスの基礎を学ぶ

- 基本的なPythonプログラミングとデータ分析方法を修得する**講義**と**演習**
- Pythonの基礎
 - 各種ライブラリ (NumPy, SciPy, matplotlibなど)
 - 数値計算, データの加工と可視化
- 数学の基礎および確率と統計の基礎
- 機械学習の基礎とPythonでの利用
 - 回帰分析
 - 決定木
 - K-最近傍法
 - サポートベクタマシン
 - クラスタリング
 - 主成分分析
 - アソシエーション分析



受講生の内訳（2023年度前期）

■学部別受講生の内訳（計285名）

- 経済学部：16名
- 教育学部：2名
- システム工学部：263名
- 観光学部：4名

DS誘いA/Bが必修となった効果で**受講生の増加**

※データサイエンスへの誘いA/Bは令和4年度は必修，令和5年度から必修

データサイエンス 応用

講義の概要

Pythonを用いたテキストマイニング（講義と演習）

- テキストからの特徴抽出やテキスト分類等の技術を理解し，実際に利用
- ソーシャル・ネットワーキング・サービス（SNS）や新聞記事，書籍，論文等の**テキストデータ**から様々な**発見**を得るための手法を修得
- 音声言語やビッグデータ，Web API等の取り扱いについて体験的に学修

テキスト（文章）から
お宝を発掘しよう！



授業中に利用する主なデータ等

- 青空文庫（著作権が切れた小説）
- Amazon Customer Review Dataset（日本語）
- Wikipedia日本語版のダンプファイル
- CiNii Articles - 国立情報学研究所
- 声優統計コーパス
- JVS（Japanese versatile speech）コーパス

※その他（受講生自身がデータを探して利用することを推奨）

受講生の内訳（2022年度後期）

- 学部別受講生の内訳（計129名）
 - 経済学部：3名
 - 教育学部：3名
 - システム工学部：118名
 - 観光学部：5名

データサイエンス 実践

講義の概要

- PythonとExcel を使い、**実データを用いたデータマイニング**を中心とした講義と演習を実施
- 実際の**POSデータを題材**として、データマイニングに取り組み、実践的なデータマイニングと可視化の手法の修得
- 業務把握、仮説立案、データ分析・検証、提案、プレゼンなど、データサイエンティストの一連の仕事の流れを体験

分析テーマとデータについて

- オークワ様から提示されたテーマ
 1. 環境変化による**購買変化**
 - **温暖化等災害**の増加
 - **コロナ禍から現在まで**の購買変化
 - **値上げ**
 2. **次世代食生活**への対応
 - 次世代ミート
 3. **会員施策**への提案
 - 会員カードを発行する企業は、常に会員数を増やしたいと考えている
- 提供データ (**直近のデータ, 複数年度, 複数店舗のデータ**)
 - 2017年～2022年 9月度～12月度×全国4店舗
 - **コロナ禍の分析**のために
 - 2019年～2022年の3月度～5月度の1店舗

受講生の内訳（2023年度前期）

■学部別受講生の内訳（計23名）

- 経済学部：3名
- 教育学部：0名
- システム工学部：20名
- 観光学部：0名

■教員

- 3名

■TA

- 1名

※2022年度は経済学部0名だったので、+3名

※経済学部の4年生は、「データサイエンスへの誘いA/B」の受講開始の学年

講義の様子

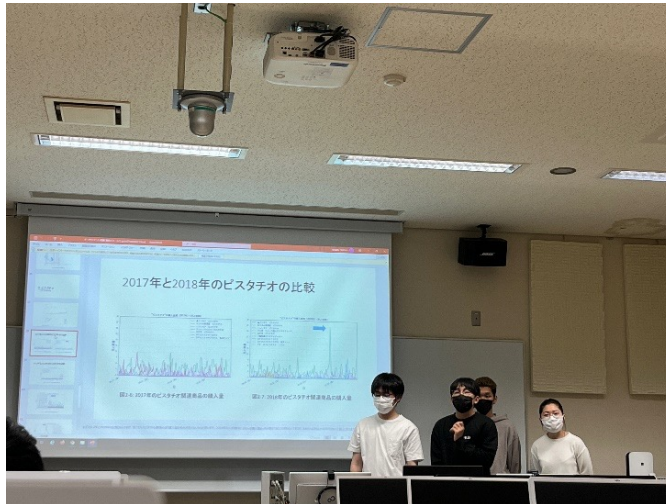


写真1 最終発表会の様子

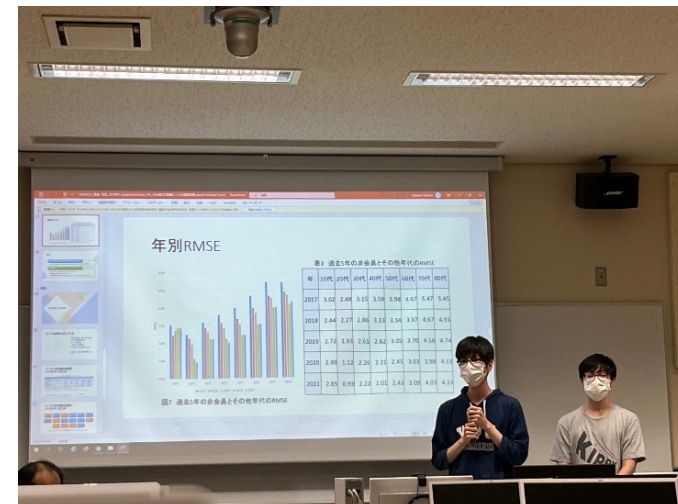


写真3 最終発表会の様子

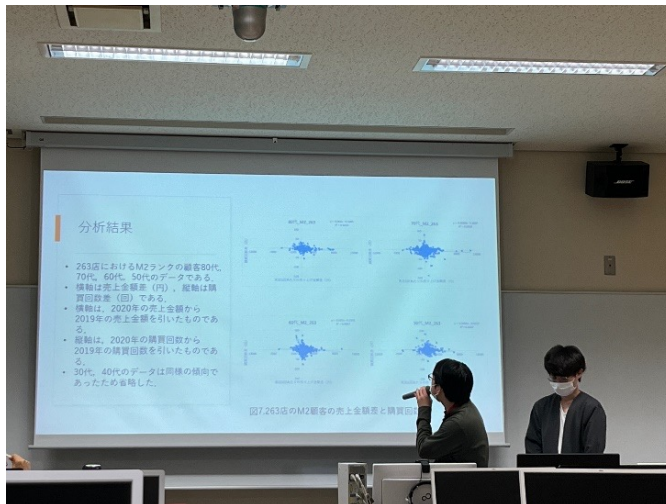


写真2 最終発表会の様子



写真4 最終発表会の様子

実践的データマイ ニング1/2 (大学院)

企業と連携した講義

- 実践的データマイニング1（2018年度～）
 - **オークワ**の匿名加工ID-POSデータを利用
 - 4店舗・直近数年分のデータ
- 実践的データマイニング2（2019年度～）
 - **紀陽銀行**の匿名加工ATM取引データを利用
 - 全店舗・直近数年分のデータ
- 企業が抱える問題等を本物のビジネスデータを使って分析，課題解決を提案（ロールプレイ型PBL）
- **企業担当者が授業に常駐**し，学生と意見交換・追加情報等を提供



鮮魚の廃棄ロスが多いと思うのだけど、調べてくれる？



授業実施中の様子 (1/2)



企業の情報課の方の
説明中の様子



議論中の様子



最終の発表中の様子

授業実施中の様子 (2/2)



2023年8月実施
実践的データマイニング1: 受講者26名

2022年9月実施
実践的データマイニング2: 受講者28名



**履修者を
増やすための工夫**

次のステップは？

**全学的に受講可能な
データサイエンス科目の整備**



**多くの学生がデータサイエンス科目
を受講する状況の実現
(特に文系学生の受講を促進したい)**

授業紹介動画の作成 と 視聴促進

授業紹介動画（1/2）

- 学生のモチベーション向上
- 学生インタビュー中心
 - 約1時間の長編！（意外とすぐ見終わります）



授業紹介動画 (2/2)

32

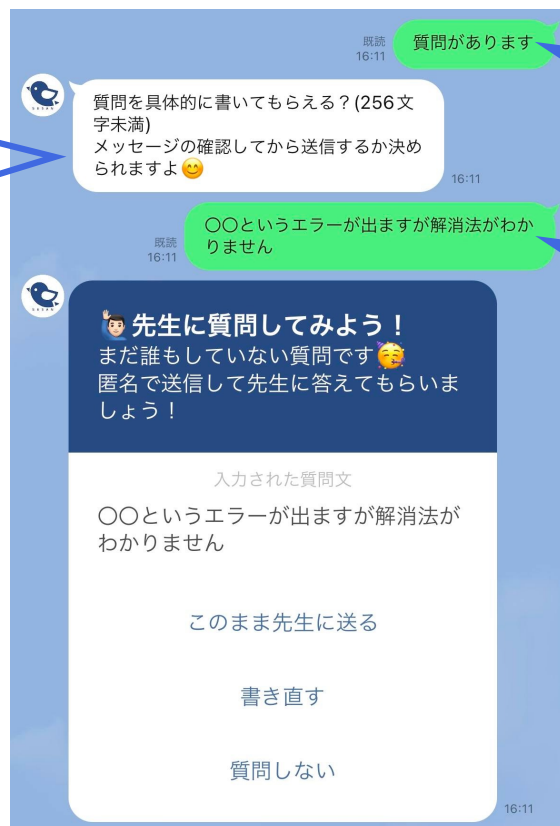


データサイエンス 和歌山大学
YouTube



LINEチャット ボットの導入

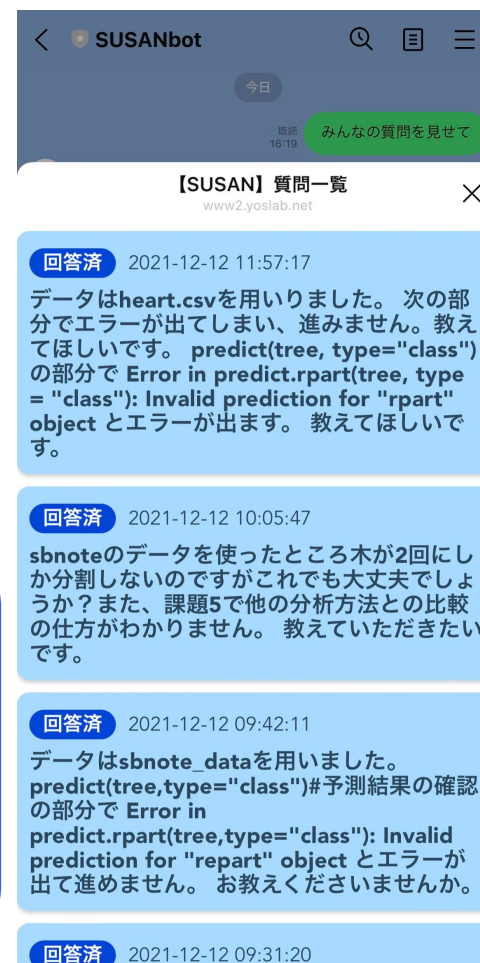
システムの
自動応答



学生の質問
(あいまい)

学生の質問
(具体的)

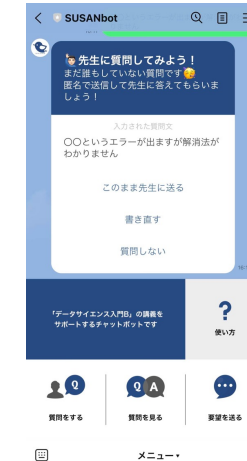
- 過去質問の閲覧
- ボットの自動応答との併用により、**同じ質問が減ることを期待**



LINEによるサポート（1/2）

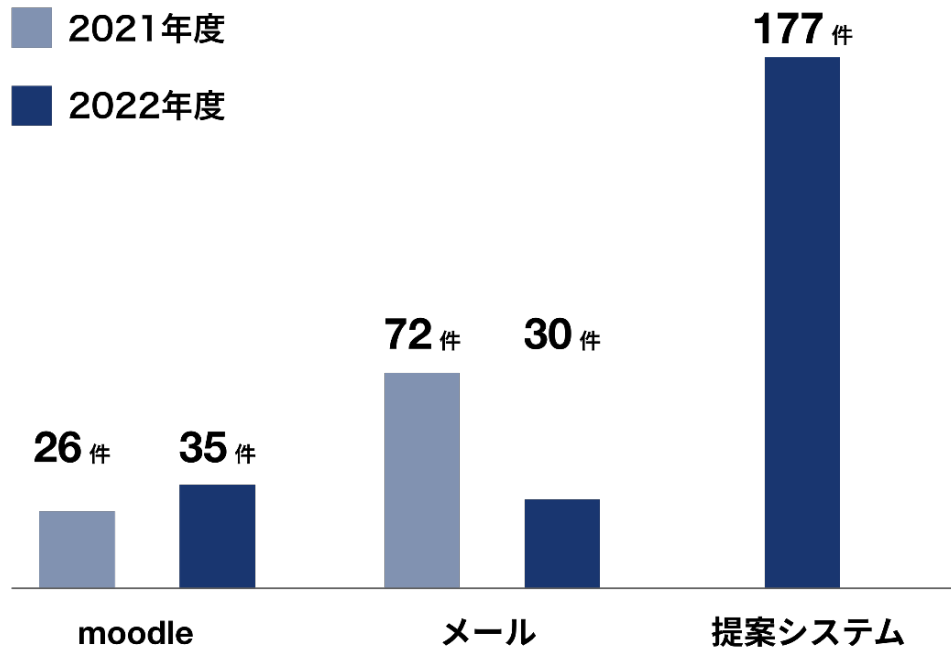
■ LINE利用の効果

- 質問投稿数，質問者数ともに増加
 - 匿名による質問
 - **LINEによる気軽に質問しやすい環境**の提供
- 他の人の質問を参考にできる効果
 - 同じ質問が減る **（効果大）**
 - ◎ フォーラム（電子掲示板）には，同じことが書かれることがあった
- **「他の人が質問している」という状況の共有**が「質問」を促す？

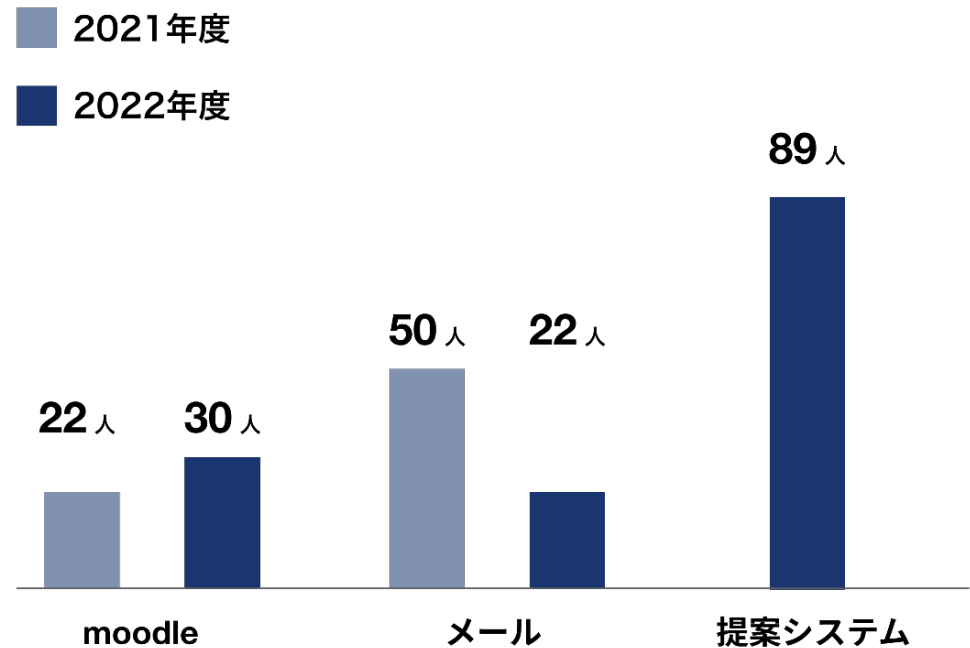


LINEによるサポート (2/2)

質問投稿数の比較



質問者数の比較



質問投稿数, 質問者数の増加

オープンバッジ の導入

オープンバッジ導入の経緯

- 当初から、数理・データサイエンス・AI教育プログラムの一定数（10単位程度）の単位取得者には**証明書を発行する予定**
- 「数理・データサイエンス・AI教育プログラム」を立ち上げ、**全学必修修化して、3年終了したため、発行のタイミング**



オープンバッジ！

和歌山大学の工夫点

■工夫点

- バッジとカリキュラム系統図を関連付け
 - 科目バッジと単位積み上げ式バッジの併用
 - ◎ バッジを取るために講義を受講
 - ||
 - ◎ 体系的な受講につながる

オープンバッジの紹介 (1/3)

■科目バッジ

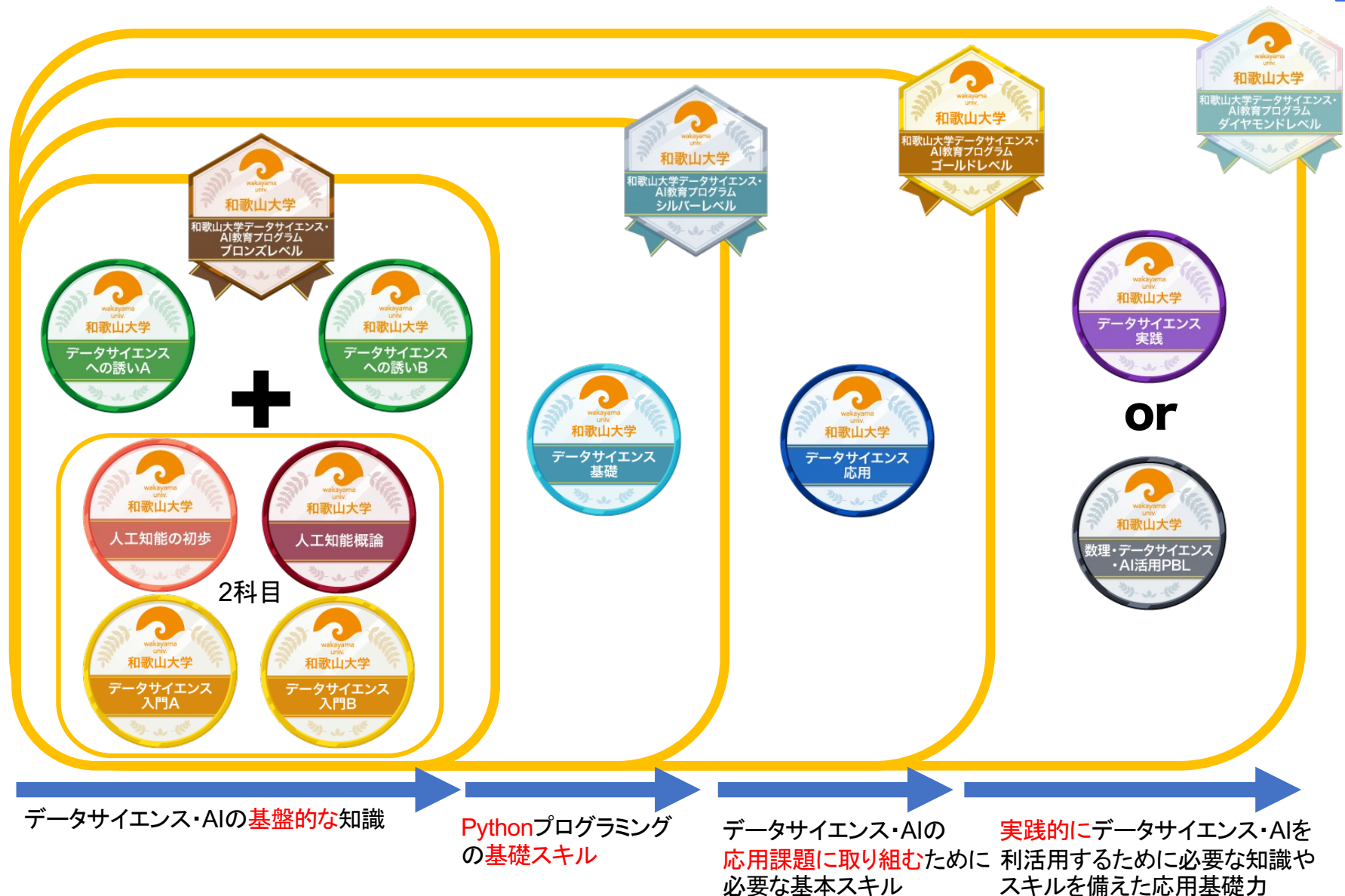


■単位積み上げ型バッジ



単位積み上げ型バッジを利用し、サーティフィケーションを実現

オープンバッジの紹介 (2/3)



オープンバッジの紹介 (3/3)



授与学生のツイート



物理バッジの 導入

本物のバッジの作成と授与



DTIER



和歌山大学 DTIER

