

教学データの学習エビデンスに基づく GPA 向上のための 情報推薦・学習支援システムの開発

滝沢 光介

富山県立大学 電子・情報工学科

November 19, 2021

- 1 はじめに
- 2 教学データ分析
- 3 取得科目の推薦
と教材の最適化
- 4 提案手法
- 5 実験結果並びに
考察
- 6 おわりに
進捗

1.1 研究の概要

2/14

背景

大学における成績評価方法の一つとして Grade Point Average(GPA) 制度が存在する. 学生は上位の GPA を収めることで, 自分の希望する研究室に入りやすくなったり大学院入試で大きく有利を得ることができる. しかし, 自分がよい成績を修めることができるかどうかはその科目を履修してみないとわからない. さらに, 大学における卒業要件単位を満たすために自分の不得意とする科目もある程度取得する必要がある.

目的

過去の卒業生の教学データに対して分析を行い, システムを使用する学生がまだ取得していない科目において成績評価の予測を行い, 高い GPA を獲得できるように単位選択を推薦するとともに, 全ての科目について良い成績がとれるように予測成績が低い科目については Web 上から関連情報を推薦するようなシステムの開発を行う.

- 1 はじめに
 - 2 教学データ分析
 - 3 取得科目の推薦と教材の最適化
 - 4 提案手法
 - 5 実験結果並びに考察
 - 6 おわりに
- 進捗

教学 IR

IR とは Institutional Research の略であり，IR は「ある特定の目的に沿って情報を収集し，それらを加工・統合して分析し，計画立案や意思決定を支援するために展開される活動の総称」と定義されており，それらを教学データに応用させたものを教学 IR と呼んでいる [1].

教学データの利用

このような教学データ分析の広がりから全国の多くの大学で生徒の学習成績などの蓄積が行われている．本研究では，教学 IR を広義の意味で捉え，GPA 向上を目的とした推薦システムの実装を行う．

¹浅野茂，“データベースの構築と IR の課題”，高等教育研究，第 19 集，2016

2.2 eポートフォリオによる学習エビデンス

4/14

eポートフォリオ

eポートフォリオは生徒自身の学習の過程であり、その生徒がいかに学習を行ったかという証拠(エビデンス)となり得る。しかし、eポートフォリオは個人情報として扱われる[2]。一個人が扱える物ではないのでデモデータの作成を行った。

学術データのデモデータ

卒業済み学生データ

学級番号	姓	名	性別	学年	入学年	入学月	入学日	入学地	入学経路	入学理由
1910001	1	1	0	0	1	1	0	1	0	0
1910002	1	1	1	1	1	1	1	1	1	1
1910003	1	1	2	2	2	2	2	2	2	2
1910004	1	1	3	3	3	3	3	3	3	3
1910005	1	1	4	4	4	4	4	4	4	4
1910006	1	1	5	5	5	5	5	5	5	5
1910007	1	1	6	6	6	6	6	6	6	6
1910008	1	1	7	7	7	7	7	7	7	7
1910009	1	1	8	8	8	8	8	8	8	8
1910010	1	1	9	9	9	9	9	9	9	9
1910011	1	1	10	10	10	10	10	10	10	10
1910012	1	1	11	11	11	11	11	11	11	11

在学中の学生データ

学級番号	姓	名	性別	学年	入学年	入学月	入学日	入学地	入学経路	入学理由
1910013	1	1	0	0	1	1	0	1	0	0
1910014	1	1	1	1	1	1	1	1	1	1
1910015	1	1	2	2	2	2	2	2	2	2
1910016	1	1	3	3	3	3	3	3	3	3
1910017	1	1	4	4	4	4	4	4	4	4
1910018	1	1	5	5	5	5	5	5	5	5
1910019	1	1	6	6	6	6	6	6	6	6
1910020	1	1	7	7	7	7	7	7	7	7
1910021	1	1	8	8	8	8	8	8	8	8
1910022	1	1	9	9	9	9	9	9	9	9
1910023	1	1	10	10	10	10	10	10	10	10
1910024	1	1	11	11	11	11	11	11	11	11
1910025	1	1	12	12	12	12	12	12	12	12
1910026	1	1	13	13	13	13	13	13	13	13
1910027	1	1	14	14	14	14	14	14	14	14
1910028	1	1	15	15	15	15	15	15	15	15
1910029	1	1	16	16	16	16	16	16	16	16
1910030	1	1	17	17	17	17	17	17	17	17
1910031	1	1	18	18	18	18	18	18	18	18
1910032	1	1	19	19	19	19	19	19	19	19
1910033	1	1	20	20	20	20	20	20	20	20
1910034	1	1	21	21	21	21	21	21	21	21
1910035	1	1	22	22	22	22	22	22	22	22
1910036	1	1	23	23	23	23	23	23	23	23
1910037	1	1	24	24	24	24	24	24	24	24
1910038	1	1	25	25	25	25	25	25	25	25
1910039	1	1	26	26	26	26	26	26	26	26
1910040	1	1	27	27	27	27	27	27	27	27
1910041	1	1	28	28	28	28	28	28	28	28
1910042	1	1	29	29	29	29	29	29	29	29
1910043	1	1	30	30	30	30	30	30	30	30
1910044	1	1	31	31	31	31	31	31	31	31
1910045	1	1	32	32	32	32	32	32	32	32
1910046	1	1	33	33	33	33	33	33	33	33
1910047	1	1	34	34	34	34	34	34	34	34
1910048	1	1	35	35	35	35	35	35	35	35
1910049	1	1	36	36	36	36	36	36	36	36
1910050	1	1	37	37	37	37	37	37	37	37
1910051	1	1	38	38	38	38	38	38	38	38
1910052	1	1	39	39	39	39	39	39	39	39
1910053	1	1	40	40	40	40	40	40	40	40
1910054	1	1	41	41	41	41	41	41	41	41
1910055	1	1	42	42	42	42	42	42	42	42
1910056	1	1	43	43	43	43	43	43	43	43
1910057	1	1	44	44	44	44	44	44	44	44
1910058	1	1	45	45	45	45	45	45	45	45
1910059	1	1	46	46	46	46	46	46	46	46
1910060	1	1	47	47	47	47	47	47	47	47
1910061	1	1	48	48	48	48	48	48	48	48
1910062	1	1	49	49	49	49	49	49	49	49
1910063	1	1	50	50	50	50	50	50	50	50
1910064	1	1	51	51	51	51	51	51	51	51
1910065	1	1	52	52	52	52	52	52	52	52
1910066	1	1	53	53	53	53	53	53	53	53
1910067	1	1	54	54	54	54	54	54	54	54
1910068	1	1	55	55	55	55	55	55	55	55
1910069	1	1	56	56	56	56	56	56	56	56
1910070	1	1	57	57	57	57	57	57	57	57
1910071	1	1	58	58	58	58	58	58	58	58
1910072	1	1	59	59	59	59	59	59	59	59
1910073	1	1	60	60	60	60	60	60	60	60
1910074	1	1	61	61	61	61	61	61	61	61
1910075	1	1	62	62	62	62	62	62	62	62
1910076	1	1	63	63	63	63	63	63	63	63
1910077	1	1	64	64	64	64	64	64	64	64
1910078	1	1	65	65	65	65	65	65	65	65
1910079	1	1	66	66	66	66	66	66	66	66
1910080	1	1	67	67	67	67	67	67	67	67
1910081	1	1	68	68	68	68	68	68	68	68
1910082	1	1	69	69	69	69	69	69	69	69
1910083	1	1	70	70	70	70	70	70	70	70
1910084	1	1	71	71	71	71	71	71	71	71
1910085	1	1	72	72	72	72	72	72	72	72
1910086	1	1	73	73	73	73	73	73	73	73
1910087	1	1	74	74	74	74	74	74	74	74
1910088	1	1	75	75	75	75	75	75	75	75
1910089	1	1	76	76	76	76	76	76	76	76
1910090	1	1	77	77	77	77	77	77	77	77
1910091	1	1	78	78	78	78	78	78	78	78
1910092	1	1	79	79	79	79	79	79	79	79
1910093	1	1	80	80	80	80	80	80	80	80
1910094	1	1	81	81	81	81	81	81	81	81
1910095	1	1	82	82	82	82	82	82	82	82
1910096	1	1	83	83	83	83	83	83	83	83
1910097	1	1	84	84	84	84	84	84	84	84
1910098	1	1	85	85	85	85	85	85	85	85
1910099	1	1	86	86	86	86	86	86	86	86
1910100	1	1	87	87	87	87	87	87	87	87
1910101	1	1	88	88	88	88	88	88	88	88
1910102	1	1	89	89	89	89	89	89	89	89
1910103	1	1	90	90	90	90	90	90	90	90
1910104	1	1	91	91	91	91	91	91	91	91
1910105	1	1	92	92	92	92	92	92	92	92
1910106	1	1	93	93	93	93	93	93	93	93
1910107	1	1	94	94	94	94	94	94	94	94
1910108	1	1	95	95	95	95	95	95	95	95
1910109	1	1	96	96	96	96	96	96	96	96
1910110	1	1	97	97	97	97	97	97	97	97
1910111	1	1	98	98	98	98	98	98	98	98
1910112	1	1	99	99	99	99	99	99	99	99
1910113	1	1	100	100	100	100	100	100	100	100

年度ごとの
科目名の違い
については
別途対応表を
作成

未履修のため空白

数字	評価	点数	可否
4	S	90点以上	合格
3	A	80点以上 90点未満	
2	B	70点以上 80点未満	不合格
1	C	60点以上 70点未満	
0	不可	60点未満	不合格
空白		履修なし	

2.3 情報推薦と協調フィルタリング

5/14

協調フィルタリング

情報推薦のアルゴリズムとして協調フィルタリングがある。協調フィルタリングは与えられたデータから規則性を見つけ出し、利用者がまだ知らない情報を予測するものである [3]。本研究では、教学データに対してユーザーベース協調フィルタリング (UBCF) を適用する。

ユーザーベース協調フィルタリング

	国語	数学	理科	社会
学生a	A	B	S	C
学生b	C	B	C	A
学生c	S	C	B	C
⋮	⋮	⋮	⋮	⋮
学生x	A	B		B

Pearson関数

$$\rho_{ax} = \frac{\sum_{y \in \mathcal{Y}_{ax}} (r_{ay} - \bar{r}_a)(r_{xy} - \bar{r}_x)}{\sqrt{\sum_{y \in \mathcal{Y}_{ax}} (r_{ay} - \bar{r}_a)^2} \sqrt{\sum_{y \in \mathcal{Y}_{ax}} (r_{xy} - \bar{r}_x)^2}}$$

学生の成績から類似する学生を見つける

加重平均

$$\hat{r}_{ay} = \bar{r}_a + \frac{\sum_{x \in \mathcal{X}_y} \rho_{ax} (r_{xy} - \bar{r}_x)}{\sum_{x \in \mathcal{X}_y} |\rho_{ax}|}$$

類似している学生aは理科で良い成績を修めているのでおすすめできそう

図 2: 協調フィルタリングで用いられる評価指標

³神嶋敏弘, “推薦システムのアルゴリズム”,

(<https://www.kamishima.net/archive/recsysdoc.pdf>, 2021.10.28)

3.1 協調フィルタリングからの科目の推薦

6/14

教学データにおける UBCF の利用

本研究では UBCF における利用者＝学生，アイテム＝科目，評価値＝成績と置き換えて，同じ科目で高い成績を修めている生徒同士は，互いにまだ取得していない科目でも片方が良い成績を修めていれば，もう片方も良い成績を取ることが出来るという考えのもと UBCF を実装する．

予測評価値からの科目の推薦

UBCF による予測評価値から学生に対して，科目の推薦を行う．この際，大学の卒業要件単位を満たすように科目の推薦を行う．

- 1 はじめに
- 2 教学データ分析
- 3 取得科目の推薦と教材の最適化
- 4 提案手法
- 5 実験結果並びに考察
- 6 おわりに
進捗

3.2 シラバスからの教材作成

7/14

教材の作成

富山県立大学のシラバスにおける各科目のキーワードの中から自然言語処理を行いその科目と関連付けられるような単語を抽出し、あらかじめデータを蓄積しておく。そして教材提供が必要になった場合には、抽出された単語について、Web ページ，YouTube の双方向に対して検索を行い，検索結果の上位数件を使用ユーザーに科目教材として提供し，学習を促す。

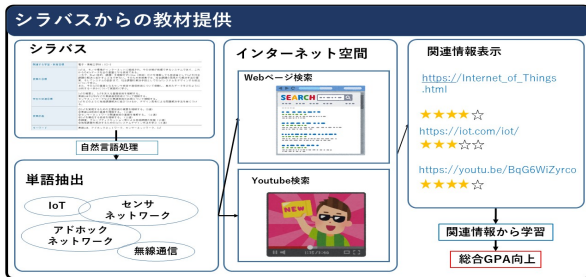


図 3: GPA 向上のためのシラバスからの教材作成

教材の洗練

その関連情報についてユーザーに評価してもらい、評価が低い関連情報は排除し価値のある情報だけを残す機能を実装する.

もう少し書く

今後の課題

- オリジナリティをどうやってだすか

4.1 提案手法の概要

9/14

提案手法

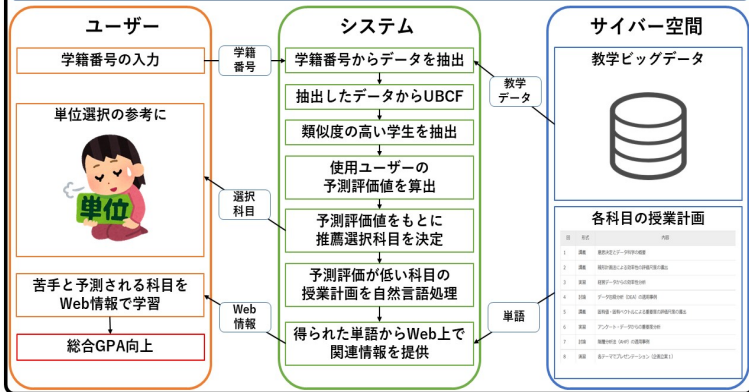


図 4: 本研究で構築するシステムの流れ

5.1 実験結果

10/14

できてからのお楽しみ

- 1 はじめに
 - 2 教学データ分析
 - 3 取得科目の推薦
と教材の最適化
 - 4 提案手法
 - 5 実験結果並びに
考察**
 - 6 おわりに
- 進捗

6.1 おわりに

11/14

まとめ

今後の課題

- 1 はじめに
- 2 教学データ分析
- 3 取得科目の推薦
と教材の最適化
- 4 提案手法
- 5 実験結果並びに
考察
- 6 おわりに
進捗

シラバスから授業のキーワードを取得した

スクレイピング

取得した情報は csv ファイルで別途保管. 今回はキーワードだけだが, 授業計画, 学生の到達目標などの取得も可能.

	0	1	2	3	4	5	6	7	8	9
コンピュー	階層モデル	TCP/IP	インターネット	WWW	LAN	プロトコル				
アルゴリス	計算量	ソートアル	探索アル	グラフ・ネ	Python					
オペレーテ	並行性	メモリ管理	ファイルシ	セキュリティ	リアルタイ	組み込みシ	並列・分散	関係デー	タデー	SQL
コンパイラ	字句解析	構文解析	構文木	意味動作	コード生成					
コンピュー	コンピュー	データ表現	論理回路	プロセッサ	記憶装置	入出力装置				

図 5: シラバスから単語抽出

単語による web ページと youtube のスクレイピングをやった

スクレイピング

web ページはそのページのタイトルとページ URL を取得, youtube はその動画のタイトルと動画 URL を取得.

```
----- webページ検索結果 -----
QLとは？データベース言語の基礎知識をわかりやすく解説！
https://hnavi.co.jp/knowledge/blog/sql/#:~:text=%E3%83%87%E3%83%BC%E3%82%BF%E3%83%99%E3%83%BC%E3%8
%B9%E8%A8%80%E8%AA%9E%E3%81%AFDBMS%E3%81%AB,%E4%BB%95%E6%A7%98%E3%81%8C%E3%81%BE%E3%81%A3%E3%81%9
%E3%81%8F%E7%95%B0%E3%81%AA%E3%82%8A%E3%81%BE%E3%81%99%E3%80%82
データベース言語 - Wikipedia
https://ja.wikipedia.org/wiki/%E3%83%87%E3%83%BC%E3%82%BF%E3%83%99%E3%83%BC%E3%82%B9%E8%A8%80%E8%A
%9E
データベース言語とは何？ Weblio辞書
https://www.weblio.jp/content/%E3%83%87%E3%83%BC%E3%82%BF%E3%83%99%E3%83%BC%E3%82%B9%E8%A8%80%E8%A
%9E
----- youtube検索結果 -----
QLとは？ | データベース言語のSQLについて、できることなど含めて3分でわかりやすく解説
https://www.youtube.com/watch?v=0Dngv89ra4I
【2021年版】データベース+SQL入門 | MySQL/PostgreSQL/Oracleなどデータベースの使い方や役割・SQL
について◆初心者向けに6分で解説
https://www.youtube.com/watch?v=liX6J0FfGng
【分析】2021年に学ぶべきプログラミング言語TOP5
https://www.youtube.com/watch?v=cp_MTBbQkys
```

図 6: スクレイピングの結果

いまやっていること

作った python プログラムを apache で動かす.

これからやること

スクレイピングした web ページと動画を評価によって洗練化するシステムを作る.

- 1 はじめに
 - 2 教学データ分析
 - 3 取得科目の推薦と教材の最適化
 - 4 提案手法
 - 5 実験結果並びに考察
 - 6 おわりに
- 進捗