

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

堀田遥斗

富山県立大学 電子情報工学科

January 14, 2022

はじめに

分析方法の概要

協調フィルタリングによる科目の推薦

学習支援のための教材作成

提案手法

引継ぎ

今後の課題

背景

大学には成績評価方法の一つとして Grade Point Average(GPA) 制度がある.GPA は大学院入試や, 研究室配属の際に重要な評価基準となることが多い. しかし, 科目を履修するまでは, その科目で良い成績を修めることができるかはわからない. また, 卒業要件単位を満たすためには, 自身が不得意とする科目も履修する必要がある.

目的

過去の卒業生の成績データに対して分析を行い, まだ履修していない科目において成績評価の予測を行うことで,GPA が高くなるように単位選択を推薦する. また, 全ての科目において良い成績が取れるように,Web 上から関連情報を推薦するシステムの開発を行う.

教学 IR

IR(Institutional Research) とは, 「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」と定義されている. こ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.

はじめに

分析方法の概要

協調フィルタリングによる科目の推薦

学習支援のための教材作成

提案手法

引継ぎ

今後の課題

eポートフォリオは、大きく二つの意味で使い分けられている。

広義の定義 … 電子的な形式で扱われたすべてのポートフォリオ。

狭義の定義 … ポートフォリオを作成するためのソフトウェア、または、ポートフォリオをマネジメントするためのシステム。

この研究におけるeポートフォリオは、**生徒自身の学習の過程のデータ**のことを指す。

しかし、eポートフォリオは個人情報として扱われ、一個人が扱えるものではないので、デモデータを作成し、そのデータを用いて協調フィルタリングを行う。

ユーザベース協調フィルタリング (UBCF) の流れ

- 1, ユーザのアイテムに対する評価値をもとに, ユーザ同士の類似度を求める.
- 2, 類似度の高いユーザを何名か抽出.
- 3, 対象ユーザの未評価アイテムについて類似度の高いユーザの評価値から予測評価値を算出して, 予測評価値の高いアイテムを対象ユーザに推薦.

・「ユーザ A に似た人ならユーザ A が好むものを好むであろう。」
というのが, UBCF の基本的な考え方となっている.

教学データにおける UBCF の実装

- ・ UBCF における利用者を学生, アイテムを科目, 評価値を成績と置き換え, UBCF を実装する.
- ・ 協調フィルタリングにより, 予測評価値を算出し, 得られた予測評価値が良い科目を選択する.
- ・ しかし, 卒業要件を満たすためには, 予測評価値の低い科目も履修しなければならない. これらの総合的な GPA を下げるような科目については学習支援を行う.

教材作成の仕方

予測評価値の低い科目の授業計画をサーバ空間上から抽出して、授業の回数ごとに自然言語処理を行う。抽出した単語について、Web上で検索を行い、上位数件を教材として提供する。また,youtube などを利用して動画でもその分野について教材の提供を行う。



図 1: シラバスからの教材作成

提案手法

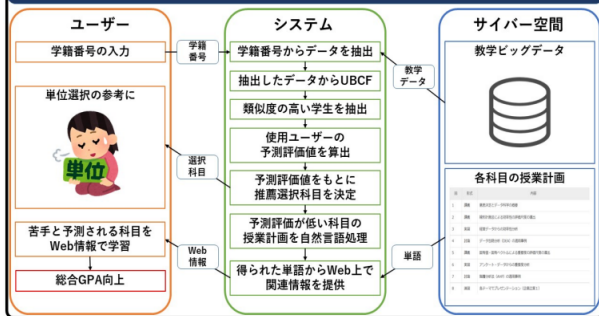


図 2: 提案手法の流れ

これまで取り組んできたこと

- ・まずはじめに取り組んだことは、滝沢さんのプログラムの有用性を示すための必要なデータの収集

```

C:\Users\HARUTO\Downloads\watasu>^Ss

B2前期
[ '論理回路', '数学Ⅰ', '電子物性', 'トピックゼミⅠ', '総合英語Ⅰ', '電子回路Ⅰ', '環境論Ⅱ',
  '計測工学', '環境論Ⅰ', 'アルゴリズムとデータ構造', '社会学Ⅰ', '総合英語Ⅱ', '工業数学Ⅲ',
  '経済学Ⅱ', '中国語Ⅱ', '工業数学Ⅳ', '芸術学', '線形代数', '数学物理学演習Ⅱ', '文学Ⅰ', '電
  気回路Ⅰ', '電磁気学Ⅰ' ]

B2後期
[ 'デジタル信号処理', '総合英語Ⅲ', 'トピックゼミⅡ', '電子情報工学実験Ⅰ', 'コンパイラ',
  '生物情報学概論', 'プログラミングⅡ', '制御工学Ⅰ', '総合英語Ⅳ', '芸術学Ⅱ', '電磁気学Ⅱ', '技
  術英語', '電子回路Ⅱ', '電気回路Ⅱ' ]

B3前期
[ 'プレゼンテーション演習', '半導体素子工学', 'インターンシップA', '電子情報工学実験Ⅱ', 'コ
  ミュニケーションの社会学', '国際関係論', '企業経営概論', '健康科学Ⅱ', '制御工学Ⅱ', '電送工
  学Ⅰ', '情報電子デバイス工学', '英語特別演習Ⅰ' ]

B3後期
[ '情報システムと地球環境', '技術者倫理', '電子情報工学実験Ⅲ', '専門ゼミ', 'データ処理工学',
  'ネットワーク設計論', 'パワーエレクトロニクス', 'ユビキタス通信工学', '伝送工学Ⅱ', '文学Ⅱ',
  '電子情報工学特別講義', '法学' ]

B4前期
[ '卒業研究' ]
  
```

図 3: 推薦科目の出力結果

出力科目の比較

・同じ学科でこれまでに取得した科目が同じで成績だけが異なる生徒同士の出力結果にどれほど違いが出るかを調べた。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	1回目	1915001	1915002 全体			2回目	1915001	1915002 全体			3回目	1915001	1915002 全体	
2	B2前期	10/18	10/15	10/23		B2前期	9/15	9/13	9/19		B2前期	8/13	8/12	8/17
3	B2後期	9/14	9/10	9/15		B2後期	9/13	9/10	9/14		B2後期	10/13	10/15	10/18
4	B3前期	11/14	11/15	11/18		B3前期	7/9	7/14	7/16		B3前期	6/14	6/14	6/22
5	B3後期	10/11	10/15	10/16		B3後期	9/13	9/14	9/18		B3後期	7/12	7/10	7/15
6														
7														
8	4回目	1915001	1915002 全体			5回目	1915001	1915002 全体			1-5回目の全体の平均			
9	B2前期	15/21	15/15	15/21		B2前期	13/20	13/16	13/23		B2前期	53%		
10	B2後期	13/13	13/15	13/15		B2後期	11/13	11/13	11/15		B2後期	68%		
11	B3前期	13/17	13/16	13/20		B3前期	9/13	9/13	9/17		B3前期	50%		
12	B3後期	10/16	10/13	10/19		B3後期	8/14	8/10	8/16		B3後期	52%		
13														
14														
15	10/23	3/5	11/18	5/8										
16	9/19	9/14	7/16	1/2										
17	8/17	5/9	3/11	7/15										
18	5/7	13/15	13/20	10/19										
19	13/23	11/15	9/17	1/2										

図 4: 出力結果の比較

出力結果比較のまとめ

- ・取得している単位の数と種類は同じで、成績だけが違う生徒がシステムを使った場合の出力結果を条件を変えて計 5 回, 比較した.
- ・結果,B2 前期が 53 パーセント,B2 後期が 68 パーセント,B3 前期が 50 パーセント,B3 後期が 52 パーセントであった.
- ・つまり,B2 前期を例にすると, 出力結果のうち 53 パーセントは同じ結果が出力されるが, 残り 47 パーセントは違う科目が推薦されるということである. よってシステムは, 有用であると確認できた.

データ収集

- ・シラバスに掲載されている科目の授業計画から、単語をスクレイピングして、講義データを収集した。

授業計画

- ①化学という学問
- ②原子・分子の世界の理論（量子論）
- ③原子軌道、量子数
- ④周期表、原子の電子配置
- ⑤化学結合、分子軌道
- ⑥混成軌道、化合物の構造
- ⑦電気陰性度、結合の極性
- ⑧分子間力、水素結合
- ⑨イオン結合、イオンの結晶構造
- ⑩金属結合、金属材料
- ⑪様々な結晶、半導体
- ⑫配位結合、金属錯体
- ⑬酸化物、水素化物
- ⑭原子核の構造と反応、放射能
- ⑮まとめ

図 5: シラバスの授業計画

今, やっていること

- ・システムのレビューについて, 信頼性の向上のために, レビューに対して, 類似性, 集中性, 情報性の3つの評価基準を導入する.
- ・今はその一つである情報性についてのプログラムを組んでいる.

情報性

- ・レビュー本文中に特徴的な名詞の割合が少ない程スパムらしいといえるという考えの元, レビューがどの程度 informative なレビューであるかを測る指標として, 情報性スコアを算出する.

```
[10 rows x 14 columns]
[[[残念, 趣味, ケルト, 音楽, 範疇, 幅, の], [残念, 趣味, 正直, 退屈, 眠気], [パ
フォーマンス, ミュージカル, 映画, 版, 歌唱], [専門, 的, 事件, レ, コメ, 象徴, の, 文才, コン
無題, ヒップ, ホップ, カルチャ, 度, ヒップ, ホップ, アルバム, エン, コード], [今, 環境, 時間
Core, 2, Duo, 4600, 十分, 快速, アルバム, エン, コード], [mp3, Divx, 時間
以前, br, 半分, 以下, 時間, br, 1, ミサイル, の, 現状, 不満, 十分], [figma, クオリティ,
目, br, 最高潮, 2, 巡, 武蔵, 一つ, どう, どう, br, 他, レビューア, 方, 数, 年, 発言, 交
一], [アナログ, デジスタラジ, 保存, の, 音, レ, ステレオ, 30', ブレーヤー, 曲, オーバ
音感, 音, 時, ランスタラジ, 音, 方, レ, 非常, 残念, もの, 別, 事, 楽曲, 曲, 交
像, br, 交換, すべて, 満足, 方, D, V, D, 手, 入, 久しぶり, フラモデル, 子供
CP, D, V, D, ], [小学, 2, 年, 子ども, 2, br, 入, 久しぶり, フラモデル, 子供
要望, 塗装, 星, 入れ, 2, 日, br, 最近, フラモデル, 造形, 本棚]]
[3, 2, 1, 1, 1, 1, 4]
[3, 2, 1, 1, 1, 1]
[2, 1, 1, 1, 1]
[1, 0, 2, 1, 1, 4, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1]
[1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 2, 1, 1, 4, 1, 1, 0]
[1, 1, 1, 2, 1, 1, 1, 1, 1, 1]
[1, 0, 0, 1, 2, 0, 1, 1, 1, 1, 2, 1, 1, 2, 1, 1, 1, 1]
[1, 1, 1, 4, 0, 1, 1, 2, 1, 1, 2, 1, 1, 0, 1, 2, 0, 1, 1, 3, 1, 1, 1, 0, 2, 1, 0, 1, 2]
[1, 1, 0, 2, 1, 1, 2, 1, 1, 2, 0, 1, 1, 1, 1, 1, 1, 0]
[1, 2, 2, 1, 1, 2, 1, 1, 2, 1, 1, 2, 1, 1, 0, 1, 1]
[2, 6347654053716636, 2, 3722199179873487, 2, 46977425413792, 3, 85597834454859, 3, 9218489965923133, 3, 14985
59708478178, 3, 489397120046654, 3, 962981954601569, 3, 510330816497697, 3, 7035274880041573]
[1, 675779204202107, 2, 007028596181098, 1, 883944612376835, 0, 13500390776285676, 0, 051896474019386685, 1, 0
259016986055158, 0, 5975107128673873, 0, 0, 0, 5710991663465453, 0, 3273475246287072]
```

図 6: 組んだプログラムの実行結果

情報性の算出

$$I_{score}(r_i) = \ln(1 + \sum_{j=1}^{|K_i|} \ln(\frac{n}{df(terms_j)}))$$

$$I_{score_{norm}}(r_i) = 5(1 - \frac{I_{score}(r_i)}{\max(I_{score}(r_i) | j=1,2,\dots,n)})$$

はじめに

分析方法の概要

協調フィルタリン
グによる科目の
推薦

学習支援のための
教材作成

提案手法

引継ぎ

今後の課題

今後行っていくこと

- ・三つ目の評価基準である集中性のプログラムを, どのように組むか考える.
- ・評価実験の手伝いをする事.