



# 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・学習支援システム開発

1815043 滝沢光介

情報基盤工学講座 指導教員 奥原浩之

## 要約

大学において良い成績を修めるために過去の教学データを参照し、分析し最適な単位の組み合わせを提案するシステムを実装する。分析方法として、協調フィルタリングによる予測評価値を算出する。それらの結果をもとに使用ユーザーに対して、高いGPAを修めることができると予測された科目についてWeb上から教材を抽出しユーザーに提供を行う。

**キーワード：**  
教学ビッグデータ,GPA向上,UBCF,情報推薦

## 1はじめに

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。GPAは各科目における成績を基に決定されるもので、その人の大学における成績の総合評価といえる。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。さらには、高いGPAを取得しておくことで就職活動において有利に働くことがある。

GPAを上げる方法として、第一に履修した単位においてよい成績を修めるということである。しかし、大学生は一つの学期で多くの単位を取る必要がありすべての履修単位について上位の成績を修めることは困難である。また、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。さらには、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る必要がある。

そこで本研究では、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どれくらいの成績を修めるかを予測し、なるべく高い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その科目に関連した教材を推薦するようなシステムの開発を行う。

## 2 教学データ分析

### 2.1 教学におけるビッグデータ・アナリティクス

近年、大学などにおいて教学データに対する分析が盛んにされており、それを総じて教育工学という。教育工学の広がりとして教学IRとラーニング・アナリティクス(LA)が存在する。

IRとはInstitutional Researchの略であり、浅野はIRを「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」と定義しており、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]。

浜田らは教学IRの目的は「教育機関全体あるいは学部・学科の意思決定支援や質保証に必要な分析結果の提供」であり、LAの目的は「各授業やeラーニングコースの改善」としており両者は目的による違いである[2]。

このような教学データにおける分析の広がりから全国の多くの大学で十数年前から生徒の学習成績などの蓄積が行われている。

そこで本研究では、教学IRを広義に解釈し上記のような蓄積された学生のデータを元に、GPAを最もよくするような単位の組み合わせを提案するようなシステムを実装する。

### 2.2 eポートフォリオによる学習エビデンス

eポートフォリオは生徒の学習履歴や学習記録などの教学データである。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い、どのような成績を収めてきたかという証拠(エビデンス)となり得る。

教学IRではこのような教学データ、もとい、蓄積されたeポートフォリオを活用し分析を行う。

しかし、eポートフォリオは生徒個人の成績情報や学習情報を表しているといった性質から個人情報として取り扱われ、その扱いは極めて繊細なものであり、扱う上ではその目的をはっきりと明示し、本人の同意を得た上で細心の注意を払わなければならない[3]。

上記のように各機関におけるeポートフォリオは一個人が扱える情報ではない。そこで本研究では、過去の学生のデータに類似したデモデータを作成して分析を行う。作成したデモデータを以下に示す。

教学データのデモデータ									
学籍番号	教養ゼミ	教養ゼミⅡ経済学	社会学Ⅰ	環境論Ⅰ	日本語表現芸術	健康科学心理学Ⅰ	合否		
							数字	評価	点数
1915001	1	4			3	4	2	S	90点以上
1915002	2	2			2	2	2	A	80点以上~90点未満
1915003	2	3	3		2	2	1	B	70点以上~80点未満
1915004	3	0	2	4	4	2	2	C	60点以上~70点未満
1915005	2	2	2		4	3	4	不可	60点未満
1915006	3	1			1	3	4		履修なし
1915007	3	1	2	3	3	4	2		
1915008	2	2			4	4	2		
1915009	1	4		1	2	2	1		
1915010	1	1	0			3	1		
1915011	1	1	2	1	2	4	3		
1915012	4	1	2	1	4	4	3		
1915013	4	3			4	1	2		
1915014	1	1			4	2	3		
1915015	3	1	4	3	3	2	2		

図1 デモデータ

### 2.3 情報推薦と協調フィルタリング

情報推薦とは利用者にとって有用と思われる情報を選択し、それらを目的に合わせた形で提示することである。

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[4]。

本研究では、デモ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。UBCFは利用者のアイテムに対する評価値から類似した利用者を見つけ出し、利用者がまだ評価していないアイテムへの予測評価値(利用者がどれくらいそのアイテムを高く評価するか)を算出し、予測評価値の高いアイテムを推薦するといった手法である。本研究では上記の利用者=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だろうという考え方のもとUBCFを実装する。

UBCFにおいて利用者集合  $X = \{1, \dots, n\}$ , アイテム集合  $Y = \{1, \dots, m\}$  とすると評価値行列  $R$  は利用者  $x \in X$  のアイテム  $y \in Y$  への評価値  $r_{xy}$  を要素とする行列となり、利用者  $a$  と利用者  $x$  の類似度は、共通に評価しているアイテムについての Pearson 関数で、予測評価値は評価値の加重平均で予測を行う。以下に学生の成績に対するユーザーベース協調フィルタリングの概要を示す。

#### ユーザーベース協調フィルタリング

Pearson関数			
$\rho_{ax} = \frac{\sum_{y \in Y} (r_{ay} - \bar{r}_a)(r_{xy} - \bar{r}_x)}{\sqrt{\sum_{y \in Y} (r_{ay} - \bar{r}_a)^2} \sqrt{\sum_{y \in Y} (r_{xy} - \bar{r}_x)^2}}$			学生の成績から類似する学生を見つける
加重平均			
$\hat{r}_{ay} = \bar{r}_a + \frac{\sum_{x \in X} \rho_{ax} (r_{xy} - \bar{r}_x)}{\sum_{x \in X}  \rho_{ax} }$ 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 ユーザーベース協調フィルタリング

## 3 取得科目の情報推薦と学習支援

### 3.1 協調フィルタリングからの科目的推薦

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  $S$  または  $A$  になる科目を選択していく。しかし大学を卒業するには、大学が定めている卒業要件単位を満たす必要があるので必然的に予測評価値が低い科目も履修する必要がある。そういうたった総合GPAを下げるような要因については、次節で述べる学習支援を行う。

以下は協調フィルタリングによる推薦科目決定の模式図である。

#### 協調フィルタリングからの科目的推薦

図3 選択科目才決定

### 3.2 シラバスからの教材作成

協調フィルタリングによる予測評価値が低い科目を取る必要がある際には、その科目について学習できる教材も同時に提供する。

富山県立大学のシラバスにおける各科目的授業の目標、学生の到達目標、授業計画、キーワードなど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する。以下にシラバスからの提供教材作成の流れを示す。

#### シラバスからの教材提供

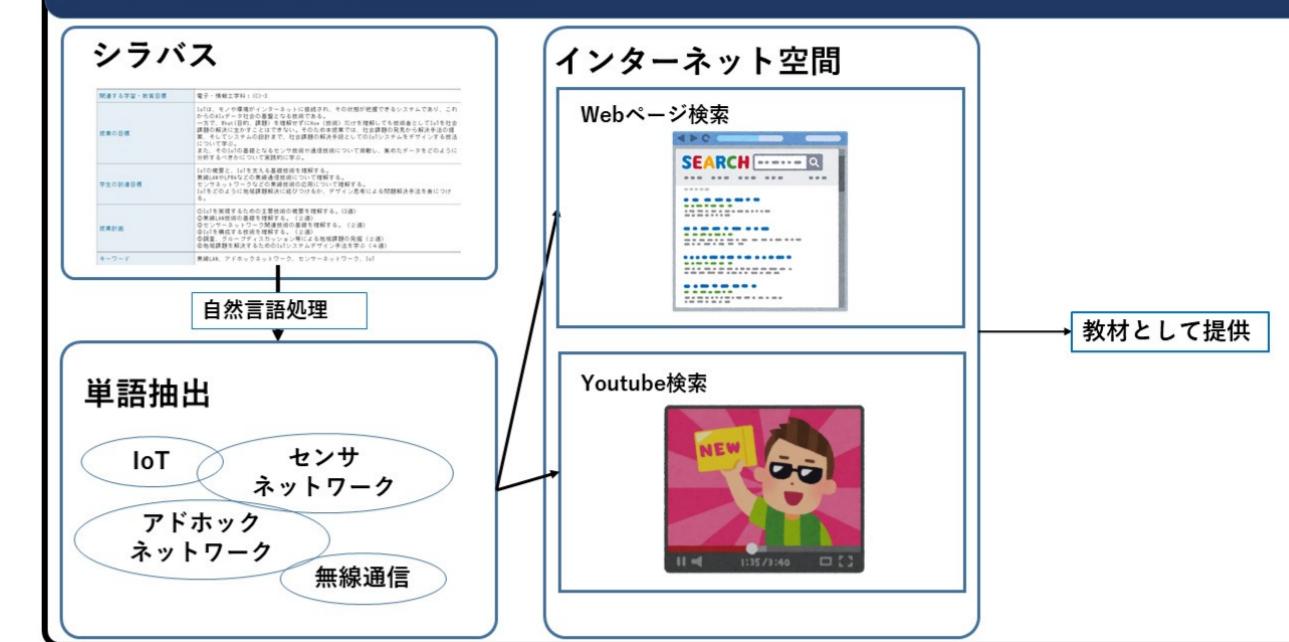


図4 シラバスからの教材作成

## 4 提案手法

本研究では、UBCFによる学生間の類似度算出、科目に対する予測評価値計算を行い、使用する学生が高いGPAをとれるような科目選択の情報推薦を行うシステムを実装する。

まず、学生には自信の学籍番号を入力してもらう。蓄積された教学データから入力された学籍番号の学生と過去の学生のデータを抽出し、UBCFを行う。UBCFの結果得られた予測評価値をもとに、使用するユーザーのGPAを最もよくするような単位の組み合わせを推薦する。この際に、大学における単位取得の影響から推薦する科目が必ずしも高い予測評価値であるとは限らない。そこで、低い予測評価値でありながら大学での卒業の関係で仕方なく選択する必要がある科目については、あらかじめその科目と関連した教材を提供するようなシステムの実装を行う。

低い予測評価値がでているにも関わらず、選択しなければいけない科目についてはその科目的授業計画を参考してはまる分野についての教材提供を行う。まず、予測評価値が低くなってしまった科目的授業計画をサイバー空間上から抽出し授業の回数ごとに自然言語処理を行う。そこで得られた単語について、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などのプラットフォーム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。最終的には、取得する単位と学習支援の二つの情報について推薦を行う。

本研究は、個人のGPAを向上させるといった目的での研究であるが、蓄積された過去の学生のデータをもとに分析を行っているという特徴からこのシステムを長年使用すればするほど学校全体の学生の質が上がることになり、教学IRという観点から大きな役割を果たすといえる。

#### 提案手法

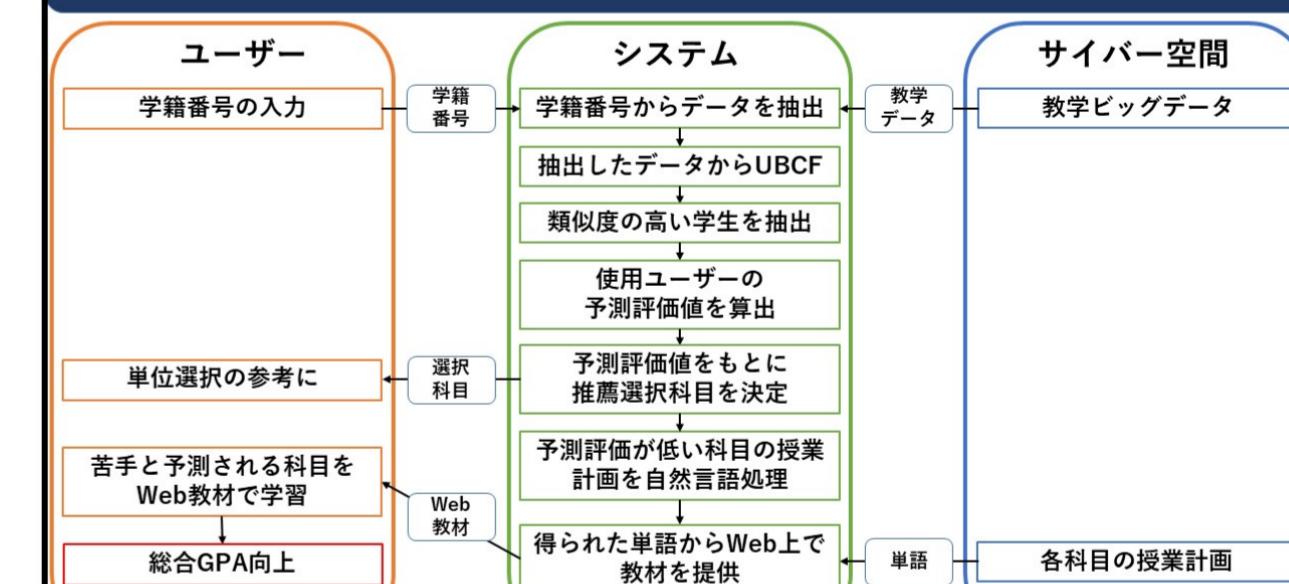


図5 提案手法の流れ

## 5 数値実験並びに考察

## 6 おわりに

今回は協調フィルタリングによる予測評価値の算出とそれに基づく単位の選択を行った。今後は選択された単位に応じて教材を提供する学習支援プラットフォームの構築を行う。

## 参考文献

- [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] 松田岳士, 渡辺雄貴, “教学IR, ラーニング・アナリティクス, 教育工学”, 日本教育工学会論文誌 41(3), 199-208, 2017
- [3] 森本康彦, “eポートフォリオとしての教育ビッグデータとラーニングアナリティクス”, コンピュータ&エデュケーションVOL.38 2015
- [4] 神島 敏弘, “推薦システムのアルゴリズム”, 2016-09-26 21:53:16 +0900; 9645c3b