

教学データの学習エビデンスに基づくeポートフォーリオのための  
情報推薦システム開発

1815043 滝沢光介

## 要約

本研究では、大学の単位取得をより効率的に行うために過去の教学データを参照、分析し最適な単位取得のためのシステムを実装する。大学での単位の選択は膨大な数にもかかわらず生徒個人が自信で決定する必要がある。まだ入学した手で間もない1年生や将来の進路が決定しておらず、どの単位を取得すれば良いかわからないような生徒に対する意思決定支援を行う。

## キーワード：

教学ビックデータ,eポートフォーリオ,学習エビデンス,情報推薦

## 1 はじめに

近年、労働者が不足しており就活生にとって有利な状況である売り手市場であった就活情勢がコロナウイルスの影響から就活生に不利な買い手市場にシフトしつつある。そこで、学生は自分の希望する企業に就職するために効率的に大学での勉学を納める必要がある。

現状、様々な大学や機関で就職支援と称したシステムが実装されているが、その実態は求人情報や就職情報、学生に対するアンケートの結果をまとめただけのものが多い。これらの機能を不要だとは言えないがその情報を生徒個、人の価値として利用するには生徒が再び情報の取扱選択を行わなければならず、生徒が自発的にこの機能を使うとは考えづらい。また、学生の成績の向上を目的としたシステム「M2B」などが存在するがあくまでも授業中の理解度などから学習支援、教育改善を図るためにシステムであり、就職に対する支援としては十分な役割を果たせていません[1]。

大学における単位の選択は大学での成績、ひいては就職への影響をおよぼす重要な決定である。しかしながら大学での単位の種類は無数にあり、その組み合わせは多なものとなる。

そこで本研究では、過去の卒業生の就職先や教学データなどのデータ分析を行い、システムを使用する学生がどのように学校での単位を取得すれば効率の良い学習を行えるかをフィードバックすることを目的としたシステムの実装を行う。

## 2 教学データ分析

## 2.1 教学におけるビックデータ・アナリティクス

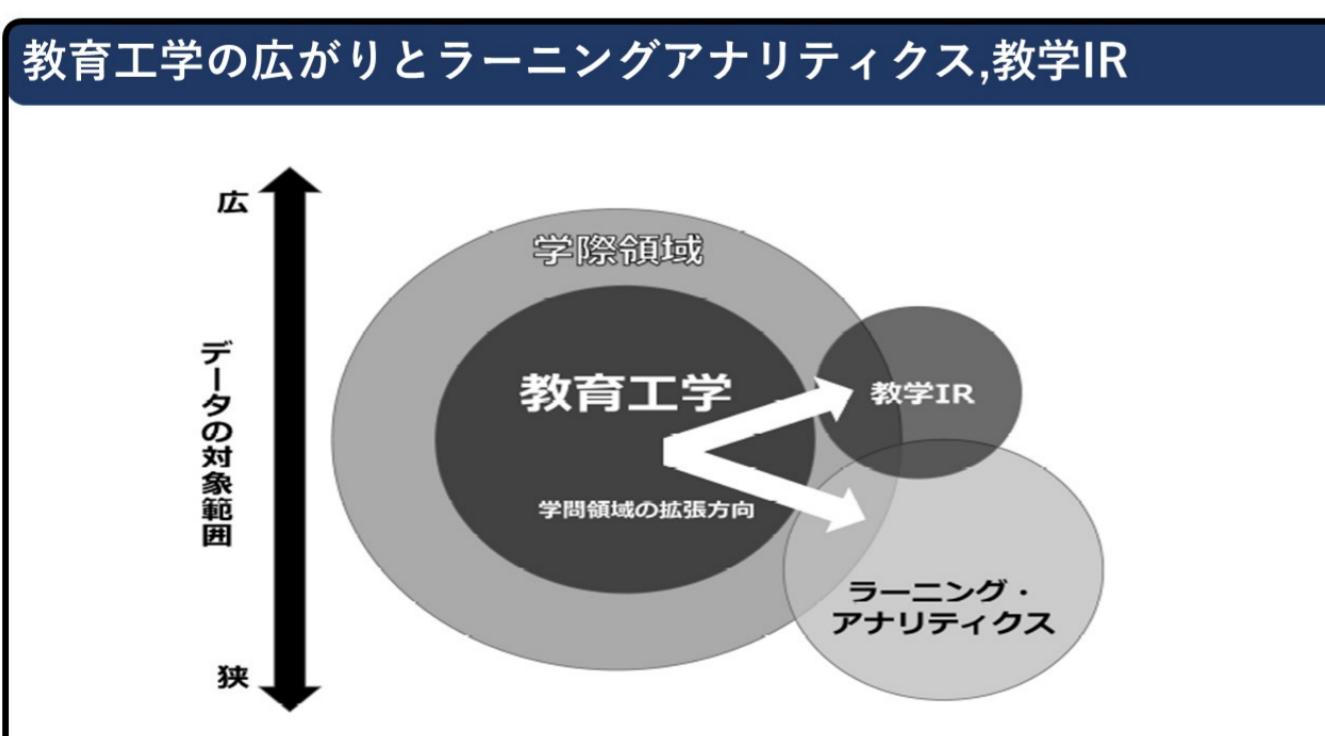
近年、大学などにおいて教学データに対する分析が盛んにされており、それらを総じて教育工学という。教育工学の広がりとして教学IRとラーニング・アナリティクス(LA)が存在する。

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[2]。

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としている[3]。

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、単位取得に悩んでいる学生の意思決定の支援を行えるようなシステムの提案を行う。



## 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[4]。

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ータを作成して分析を行う。

## 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[5]。

本研究では、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

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

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

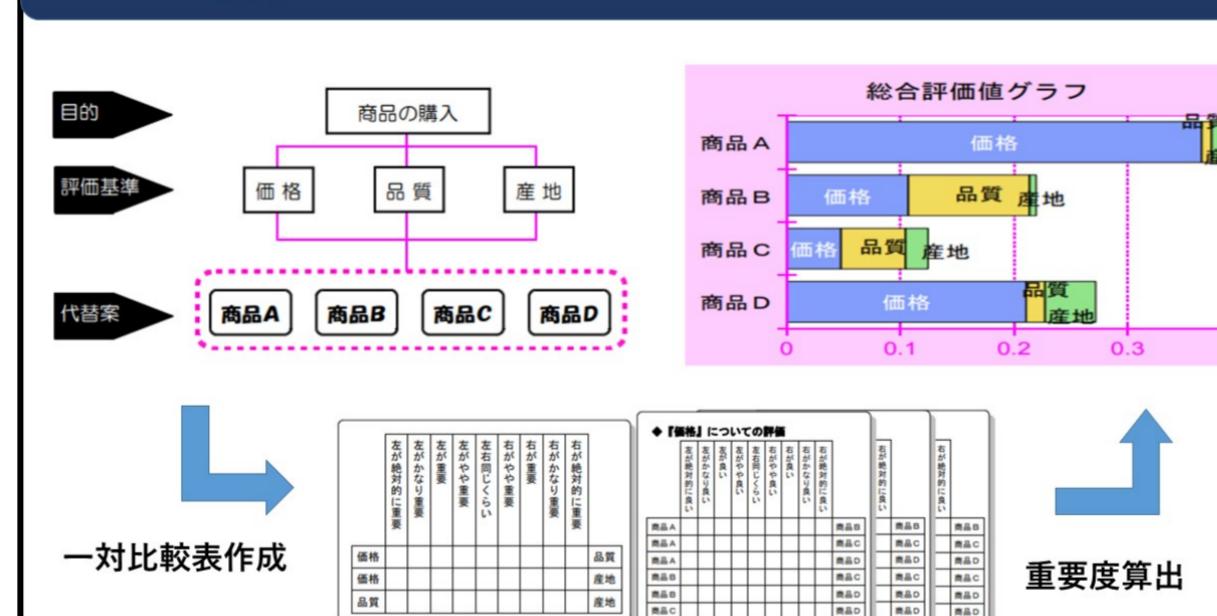
図2 ユーザーベース協調フィルタリング

## 3 情報推薦に基づく取得科目の最適化

## 3.1 教学データに対する大規模不完全AHPの適用

AHPとは階層分析手法と呼ばれ、世の中の事象を目的、評価基準、代替案の3つの階層に整理し、一对比較表を作成し、各代替案における比較対象者の主観的な重要度を算出する手法である。本来ならば、代替案における主観的な重要度を求めるものであるが本実験では、学生の成績データをもとにAHPを実装することで生徒同士の能力の優劣をはっきりとさせることができる。

## AHPの流れ



大規模とは代替案に対する評価を行う評価対象者が複数人であることを示しており、不完全とはデータの欠損を示している。本来のAHPならば全ての代替案同士についての一対比較を行うが、現実では全ての代替案について一対比較を行うことは難しい。そこで、全ての一対比較が行われていなくとも、AHPのようにそれぞれの代替案について重要度を算出する手法を不完全AHPと呼び、二つの特性を合わせたAHPを大規模不完全AHPと呼ぶ[6]。

不完全なデータから各代替案の重要度を算出するため、「単独の評価者( $L = 1$ )による全代替案に対する一对比較から得た一对比較値  $x_{ij}^l (i, j = 1, \dots, n)$ 」が完全整合した場合、それに限り、 $w_i / w_j = x_{ij}^l (i, j = 1, \dots, n)^T$  となる代替案の重要度ベクトル  $w = (w_1, \dots, w_n)^T$  が存在する。」という事実から、ある重要度ベクトル  $w_i / w_j = x_{ij}^l (i, j = 1, \dots, n)^T$  に対して比較者による一对比較値  $x_{ij}^l (i, j = 1, \dots, n)$  は  $w_i / w_j = x_{ij}^l$  であるときが理想と仮定し、偶然誤差的に生じていると仮定する。

情報基盤工学講座 指導教員 奥原浩之

なる代替案の重要度ベクトル  $w = (w_1, \dots, w_n)^T$  が存在する。」という事実から、ある重要度ベクトル  $w_i / w_j = x_{ij}^l (i, j = 1, \dots, n)^T$  に対して比較者による一对比較値  $x_{ij}^l (i, j = 1, \dots, n)$  は  $w_i / w_j = x_{ij}^l$  であるときが理想と仮定し、偶然誤差的に生じていると仮定する。

## 大規模AHP

$$\begin{aligned} & \text{各評価者 } l = 1, \dots, L \text{ が代替案 } (i, j) \text{ に対して与えた評価 } x_{ij}^l \text{ から一对比較グラフ } G \text{ を作成し、それを一对比較ネットワーク } N \text{ と呼び、一对比較ネットワークの接続行列 } A \text{ とカットベクトル } \tilde{x} \text{ を重要度ベクトル } w = (w_1, \dots, w_n)^T \text{ とする。} \\ & \min ||A^T \tilde{x} - \tilde{x}|| \text{ 誤差最小化問題} \\ & \downarrow \text{一般解} \\ & \tilde{x} = M \tilde{x} + (I - M A^T) y \\ & \downarrow \\ & \tilde{x} = Q \tilde{x} + \frac{\tilde{a}}{n} \mathbf{1} \end{aligned}$$

図4 大規模AHPの定式化

$w^T \mathbf{1} = 1$  となるように  $a$  を決定すれば、幾何平均法となる。重複評価がなく、一对比較されていない代替案対が存在する場合は、不完全AHPである。不完全データを含むAHPの重要度の算出法としてハーカー法が知られている。ハーカー法は導出できる重要度  $w_i (i = 1, \dots, n)$  が存在する物と仮定して、一对比較行列の  $(i, j)$  要素が欠けているならばその要素に  $\frac{w_i}{w_j}$  を補間して特殊法である。そこで本研究においても欠落した要素が存在するならば補間モデルを構築し、欠損値の補間を行った上で重要度ベクトルを算出する。

## 不完全大規模AHP

$$\begin{aligned} & \text{全一对比較を行う場合} \\ & \tilde{x} = \frac{1}{n} \sum_{j=1}^n \left( \frac{1}{L} \sum_{l=1}^L \tilde{x}_{ij}^l \right) + \frac{\tilde{a}}{n} \cdot (i = 1, \dots, n) \\ & \downarrow \text{逆対数変換} \\ & \tilde{x} = \left( a \prod_{j=1}^n z_{ij} \right)^{\frac{1}{n}} \quad z_{ij} = \left( \prod_{l=1}^L x_{ij}^l \right)^{\frac{1}{L}} \cdot (i, j = 1, \dots, n) \end{aligned}$$

図5 大規模不完全AHPの定式化

本研究では、大規模不完全AHPにおける評価基準=学生の成績、代替案=学生と置き換える。大学の授業はその年ごとに担当する教員が異なり、その年によって良い成績が取りやすいといったことが起きてしまうが、大規模AHPで学生の重要度を算出することで評価者による各学年の違いをなくすことができる。また、大学の授業はすべての単位が取得必須ではないので、学生によって取得している単位が異なる。といった穴抜きのデータフレームに關してもこの不完全AHPは有効である。

## 3.2 多目的最適化による情報推薦

多目的最適化とは単一の計画が多用途の目的で立案実施される場合、これらの複数の目的をどのようにして最適に達成されるかと言った問題である。多目的最適化の解説はいくつかあるが、ベクトル評価関数を用いた制約条件のもので最適化するという問題となる。以下に多目的最適化問題の定式化を示す。ただし  $x = (x_1, x_2, \dots, x_N)$  は決定変数ベクトル、 $f_i(x), (i = 1, 2, \dots, n)$  は目的関数、 $g_k(x) (k = 1, 2, \dots, m)$  は制約条件式である。

## 組み合わせ最適化

$$\begin{aligned} & \text{組み合わせ最適化} \\ & \text{目的関数} \quad \min_x \{f_1(x), f_2(x), \dots, f_n(x)\} \\ & \text{制約条件} \quad g_k(x) \leq 0 (k = 1, 2, \dots, m) \end{aligned}$$

図6 組み合わせ最適化

## 4 提案手法

本研究では、学生のデータから学生同士の類似度、またAHPによるランク付けを算出し、それらの結果を制約条件として最適化問題を行う。まず、データからクラスタリングを行いその後使用ユーザーと同じクラスタに属するユーザー群を抜き出す。そしてそのユーザー群について協調フィルタリングを行い使用ユーザーとの類似度を算出し、成績に対する類似度によるランク付けを行う。次に全ての過去の学生データに対して大規模不完全AHPを実行する。本研究の場合、AHPの評価者における評価を学生の成績と置き換えることで学生の成績から各学生の重要度を求め、重要度におけるランク付けを

行うそして、協調フィルタリングによる類似度と大規模不完全AHPにおける重要度を目的関数として設定する。さらに、大学での単位取得は1つの学期に取得できる単位に上限があるうえに各学年で取得可能な単位の種類も異なってくる。これらの条件を制約条件として加え最適化を行う。最終的にはその次の学期に学生が取ることのできる最適な科目選択を表示することで、単位選択の意思決定を行う。

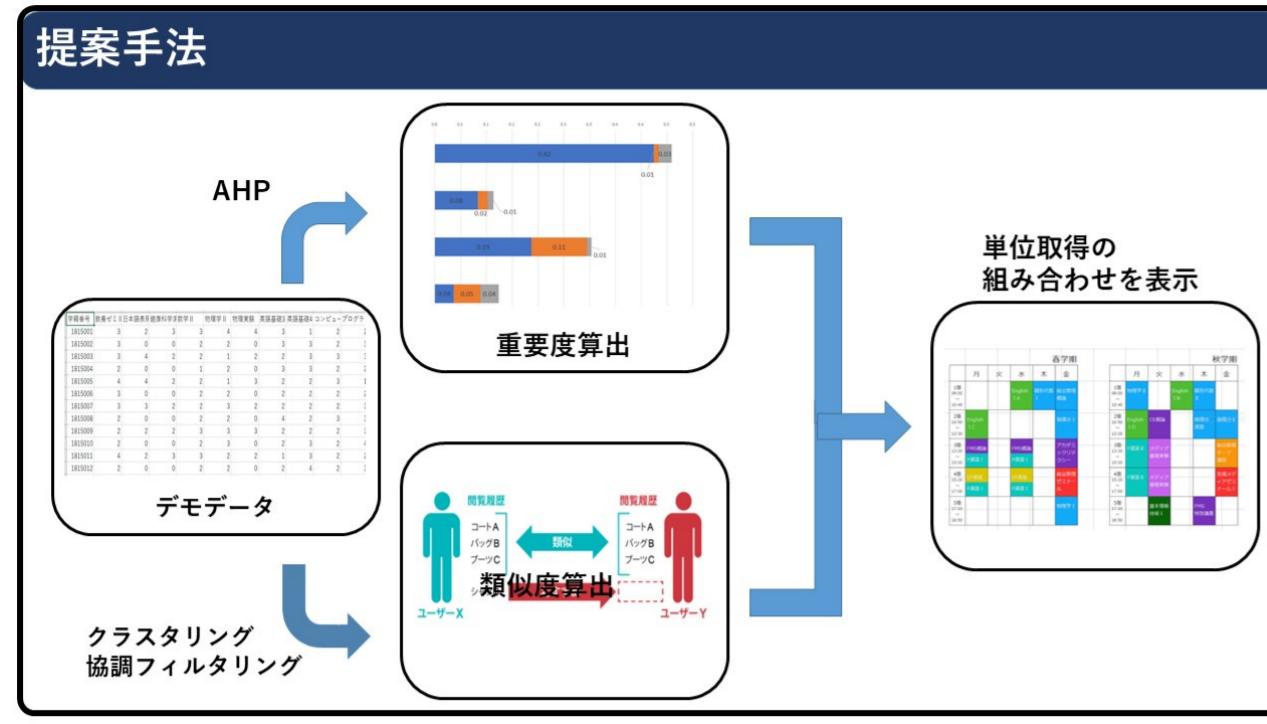


図7 提案手法の流れ

## 5 数値実験並びに考察

## 参考文献

- [1] ”M2B(みつば)学習支援システム”, Learning Analytics Center, Kyusyu Unibersity, Japan
- [2] 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)
- [3] 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
- [4] 森本康彦, ”eポートフォー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
- [5] 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
- [6] 八巻直一, 関谷和之, “複数の評価者を想定した大規模AHPの提案と人事評価への適用”, Journal of the Operations Research Society of Japan Vol. 42, No.4, December 1999

## 6 おわりに