

大学内情報の二次利用に基づく職業／講義選択支援手法
Job/Lecture Selection Support Method based on Campus Information Reuse

齋藤 義人†
Yoshihito Saito

小澤 潤†
Jun Ozawa

藤本 貴之‡
Takayuki Fujimoto

松尾 徳朗†
Tokuro Matsuo

1. はじめに

大学をはじめとする教育機関は、膨大な量の情報を保持しており、その有効的な二次利用の重要性が認識されてきている[1]。大学が所有する情報は、大学の内部に関わる情報をはじめとして、附属図書館等の図書情報、所有する物品の管理情報、所属する大学スタッフの論文数、そして所属する学生の情報等多岐にわたる。本論文では、とりわけ学生の情報にターゲットを絞り議論する。学生の情報は、大学の保持する情報の中でも最も更新頻度が高く、学生が単位を取得するたびに更新する。また、学生は終身大学に所属するのではなく、毎年新しい学生が入学する一方、単位を取得した学生は卒業するため、大学は毎年累積した学生の数の情報を獲得することになる。しかし、これらの学生情報が再活用される場合は少なく、保管にとどまっている場合が多い。これらの情報を統合し、二次利用する事により、所属する学生の意思決定支援に効果を示す事が可能となると考えられる。そこで本論文では、学生情報である講義受講履歴情報、および就職先情報に焦点をあて、大学に在籍していた学生の履修パターンと就職先の関連性を分析し、そこから法則性としての知識を発見する推論手法を提案する。

2. キャリアサービス

キャリアという言葉は広く用いられ、キャリア教育といっても企業と大学が持っている観点は異なる[2]。一般的に企業等で言われているキャリア教育は、技術的部分に焦点を置いているが、大学等では職業観の構築に焦点が置かれている[3][4]。自分自身が希望する職種に対する職業観の構築は非常に重要なことである。なぜならば、職業観の食い違いがある場合、就職しても長続きせず退職してしまうため、高い離職率に繋がる。離職率の高さは近年753問題とされ、離職率を減少させることが重要とされている[5]。また、離職率の高さは、実際に働き始めた場合の理想とのギャップの大きさや知識不足、技術不足等が原因になっているため、これらを解消する必要がある。これらの原因を解消するために、キャリアサービスセンターや就職支援センターなどと呼ばれる部署が、大学において就職支援を行っている。

現在行われている就職支援活動は、精神的な部分に焦点をあてているため、技術や知識に関しての活動は行っていない。そのため、技術や知識に関しては、学生自身が企業や専門に詳しい教員に尋ねることになる。しかし、学生個人の調査では限界があり、適切な種類の知識や技術が身に付かない場合がある。就職支援活動の一環として技術や知識に関してもサポートし、適切な種類の知識や技術を身につける手助けをする必要がある。

著者らの調査により、次の2つの支援のニーズが高いことが明らかとなった。1つ目は、希望職種が決定している学生を対象にした知識獲得・技術向上支援である。支援の方法としては、過去に在学していた学生の受講履歴と就職情報を利用し、職種に必要な講義を推察し、その講義を学生に推薦することが挙げられる。このことにより、知識を深め、技術を身につける機会を作ることができる。また、希望している職種に関係がある講義を受講することで、その職種に関する理解を深めることもできる。2つ目は、希望職種が決定していない学生を対象にした意思決定支援である。学生の受講履歴と、各職種と関連性がある講義を比べることで、どのような職種が学生の関心領域に含まれるかを知ることができる。このことは、学生の職種決定支援に役立てることができる。

3. 関連性分析手法

3.1 定義

- 過去に在籍していた学生を $S = \{s_1, \dots, s_i, \dots, s_n\}$ と定義する。 s_i は i 番目の学生を示している。
- 講義を $A = \{a_1, \dots, a_j, \dots, a_m\}$ と定義する。 a_j は j 番目の講義を示している。
- 就職先の職種を $W = \{w_1, \dots, w_h, \dots, w_l\}$ と定義する。 w_h は h 番目の職種を示している。また、職種 w_h に就職した学生の人数を WN_h と定義する。
- 学生ごとに各講義を受講したかどうかの判定を行う。 $s_i^j = 1$ ならば学生 s_i が講義 a_j を履修したことを表し、 $s_i^j = 0$ ならば学生 s_i が講義 a_j を履修しなかったことを表す。
- 過去に講義 a_j を受講したことがある学生の人数を T_{all}^j と定義する。このとき、 $T_{all}^j = \sum_{i=1}^n s_i^j$ となる。
- 職種 w_h に就職し、過去に講義 a_j を受講したことがある学生の人数を T_h^j と定義する。このとき、 $T_{all}^j = \sum_{h=1}^l T_h^j$ となる。

これらの定義を利用し、職種毎に、各講義がその職種と関連性があるか分析する。関連性があるということは、その講義は職種にとって重要であることを表すため、関連性があるほど重要度が高いと言い、関連性がないほど重要度が低いと言うことにする。また、 G_h^j は、講義 a_j が職種 w_h に対し、どの程度重要であるかを表している。

3.2 分析手順

本節では、卒業生が就職した職種と受講した講義の関連性を分析する具体的な手順を述べる。

- 講義 a_j を受講した総人数 T_{all}^j を調査する。
- 職種 w_h に就職している、かつ講義 a_j を受講していた学生の人数 T_h^j を調査する。
- 職種ごとに講義の重要度を計算する。職種 w_h に就職するために講義 a_j がどの程度重要か、以下の式を用いて計算する。

$$G_h^j = (T_h^j / WN_h) / (T_{all}^j / n) \cdot \dots \cdot (1)$$

† 山形大学 大学院 理工学研究科 Graduate School of Science and Engineering, Yamagata University

‡ 東洋大学 工学部 Department of Computational Science and Engineering, Toyo University

(4) 重要度 G_h^j の高い講義ほど就職先との関連性がある講義とする。

式(1)は、実際に就職先と講義に関連性があるかを調査するために用いる。式(1)の右辺における分母 T_{all}^j/n は、過去に在籍していた学生 n 人のうち、どのくらいの割合の学生が講義 a_j を受講していたかを表す。また、右辺における分子 T_h^j/WN_h は、職種 w_h に就職した学生のうち、どのくらいの割合の学生が講義 a_j を受講していたかを表す。つまり、式(1)の右辺 $(T_h^j/WN_h)/(T_{all}^j/n)$ は、職種における受講割合と全体における受講割合の比を示している。 G_h^j を求める際に、 $T_h^j/WN_h < T_{all}^j/n$ となり、式(1)の右辺が $(T_h^j/WN_h)/(T_{all}^j/n) < 1$ となった場合、全体の受講割合の方が高いため、講義 a_j を先生の人気や単位の取りやすさなどの職種と関係ない理由で受講した可能性がある。逆に $T_h^j/WN_h > T_{all}^j/n$ となり、式(1)の右辺が $(T_h^j/WN_h)/(T_{all}^j/n) > 1$ となった場合、職種における受講割合の方が高いため、講義 a_j を職種に関係ある理由で受講した可能性がある。簡単な例を用いて説明すると、職種 w_h に過去 1000 人中 100 人が就職し、そのうち 80 人が講義 a_j を受講していたとする。職種を中心に見た場合、就職した学生の 8 割が講義 a_j を受講しているため、講義 a_j は職種 w_h と関連性があるように見える。しかし、過去 1000 人中 900 人が講義 a_j を受講しているのならば、講義 a_j は職種 w_h と関連性がある講義ではなく、先生の人気や単位取得のしやすさで多くの学生が受講している講義であると分析することができる。このように、式(1)を用いることで先生の人気や単位取得のしやすさで多くの学生が受講している講義を発見することができる。

講義が新設された場合に本手法を用いると $T_h^j \ll WN_h$, $T_h^j = T_{all}^j$ のように、全体の受講人数が少なく、職種 w_h に就職した人の中でも受講人数が少ない状況が発生し、 G_h^j が極めて高くなる可能性がある。 $T_h^j \ll WN_h$ かつ $T_h^j = T_{all}^j$ となった場合、受講した学生全員が職種 w_h に就職しているため、関係があると言えるが、母数が少ないため信頼性は低い。しかし、10年間開講していて少し職種と関連性がある講義と3年間開講していても関連性がある講義を比較した際に、10年間開講していた講義よりも3年間開講していた講義が関係あると計算された方が学生にとって有益である。そこで、実際の講義推薦/職業調査をシステムでは、講義が設立されてから何年目かを表記している。

式(1)を用いただけでは、分析するにあたって、学生情報が増え、分母 T_{all}^j/n が大きくなるにつれ、 G_h^j の値が小さくなり、比較が難しくなるという問題が発生する。対処法として、重要度 G_h^j がどのような値をとったとしても、最も関連性がある講義の重要度を 10、最も関連性がない講義の重要度を 0 と表記するような正規化を行う。このとき、最も関連性がある講義の重要度を 100, 1000 と変化させることでより精密な分析を行うことができる。正規化した重要度を G_h^j と表記する。この場合、正規化した重要度 G_h^j の高い講義ほど職種との関連性が高い講義とする。

3.3 適職推薦と科目推薦の例

図1は、職種 w_1 に就職したいという学生がシステムを利用した例である。この学生は、講義 $a_3, a_8, a_9, a_{11}, a_{13}, \dots$ を既に受講している。システムはこの学生に対し、講義 $a_7, a_{17}, a_{23}, a_{26}, \dots$ を受講することを示唆した。

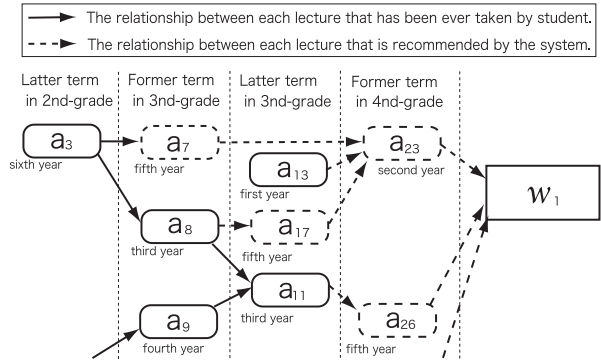


図1 講義を推薦する際に関連性を計算した結果

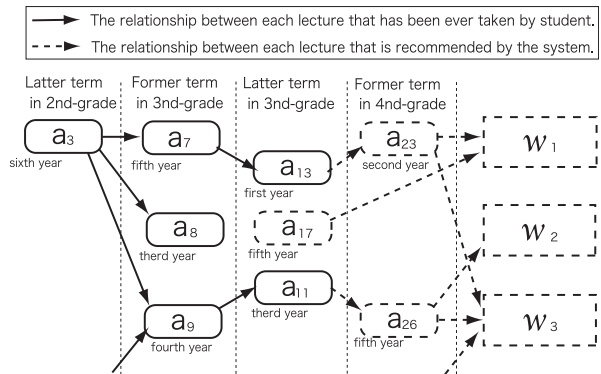


図2 職種を調査する際に関連性を計算した結果

図2は、講義 $a_3, a_7, a_8, a_9, a_{11}, a_{13}, \dots$ を受講している学生がシステムを利用した例である。システムは、この学生が受講した講義と職種 w_1, w_2, w_3, \dots が関連すると計算した。また、それぞれの職種に就職するために必要な知識を得る講義として、職種 w_1 に関しては講義 a_{17}, a_{23} を、職種 w_2 に関しては a_{26} を、職種 w_3 に関しては講義 a_{23}, a_{26}, \dots を推薦した。

4. おわりに

本論文で提案した関連性分析手法は、大学で開講している選択科目講義と学生が就職した職種の関連性を分析する手法である。本論文で示した計算においては、累積の卒業生数が増加すると利用するデータ数も増加する。このようにデータが多量にある場合、実際に講義推薦/職業調査のシステムを実装した際に、全データを用いた回帰分析を行うと計算終了まで時間がかかりすぎる。実際の利用を考えると計算時間がかかることは現実的ではない。本論文で提案した手法は、計算の手順上分析にかかる計算時間が短くてすむ。また、システムの管理者である大学が、詳細に関連性を分析したい場合や学生に詳細な情報を提供したい場合は、講義毎に設定されているシラバス記載のキーワードを用いることで、分析を細かく行うことができる。

参考文献

- [1] 竹田直樹, 正道寺勉, “就職先選定支援システムの試作”, 日本経営工学会誌, Vol.43, No.3, p. 209 (1992).
- [2] 宮武直樹, 長谷川洋介, “ファジィ理論を用いた就職指導支援システムについて”, 教育情報研究, Vol.4, No.3, pp. 101-109 (1988).
- [3] 厚生労働省, 生涯キャリア支援と企業のあり方に関する研究会報告書 (2007).
- [4] 佐藤龍子, “国立大学法人の中期目標・中期計画にみるキャリア教育と就職・学生支援”, 社会科学, Vol.75, pp. 53-73 (2006).