

K-032

授業評価アンケートの回答結果を推定する
階層的アイテムバンクの推定精度
The hierarchical item bank for predicting student's answers
in class evaluation questionnaire and its accuracy

中村 修也[†] 石井 隆稔[‡] 藤森 進[・] 赤倉 貴子[‡]
Shuya NAKAMURA Takatoshi ISHII Susumu FUJIMORI Takako AKAKURA

1. はじめに

高等教育機関では 2008 年度から FD 活動 (Faculty Development) が義務化され、大学での組織的な授業改善が求められている[1]。FD 活動の一環として学生によるアンケートが行われている。授業評価アンケートは大規模かつ容易に学生の意見や感想を収集できることから、ほとんどの大学で実施されている。しかし、授業評価アンケートが形骸化している[2]など、FD 活動の一環として十分に活用できていないという指摘がある。その理由として、授業評価アンケート結果のフィードバックは、回答結果の平均値であることが多く[3]、具体的な改善に活かすことが困難であることが挙げられる。また、出題項目が多すぎると学生・教員双方の負担が大きくなるため、抽象的で少ない項目の出題が多いことも課題である。以上により、個々の授業の改善を、現状の授業評価アンケートにより直接的に実施することは非常に困難である。

理想的には、授業改善に役立つ情報を、少ない出題数でより多く教員に対してフィードバックすることが望まれる。そこで、本研究では授業評価アンケートの回答結果から、他の項目の回答結果を推定する階層的アイテムバンクを提案する。階層的アイテムバンクにより、少ない出題数で多くの回答結果情報を教員に提供することが可能になる。本稿では、授業評価アンケートの項目間の関係性に確率的な因果関係があることを仮定し、調査を行った。調査結果に対し、ベイジアンネットワークによる構造学習を行い、他の項目がどの程度推定可能かをシミュレーションした。その結果、出題項目数に合わせて推定精度は単純増加せず、頭打ちになる項目数があることを示した。

2. 授業評価アンケートの項目間の関係性

2.1 階層的アイテムバンク

ここで、評価項目間の因果関係とは、ある抽象的な評価項目を達成するためには、別の具体的な評価項目が必要となるような関係とする。具体的な項目とは、教員が項目の値を得た時に、直接的に改善に結び付けられる項目である。評価項目間の関係を考慮すれば、評価項目は階層的な関係を持っていると考えられる。そこで、本稿におけるアイテムバンクの構成を図 1 のような階層型モデルで表した。抽象的な項目ほど上位の項目に当たり、上位の項目から下位の項目への有向辺は因果関係を示している。同一のレイヤー (階層) にある項目同士は同じレベルの抽象度を持った項目であることを示している。項目間の有向辺は因果関係であり、上位の項目の値が把握できれば、下位の項目

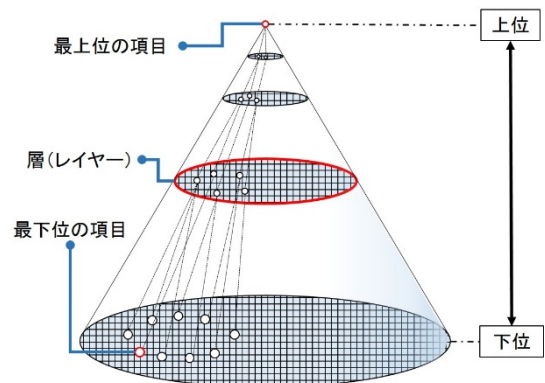


図 1 階層型アイテムバンクイメージ

の値を、またその逆も推定可能であることを示している。ここで最下位の項目は具体的な改善策と結びつく項目であり、具体的な項目ほど抽象度は低い。

本研究では、授業評価アンケートの項目が階層関係であると仮定し、項目の回答結果の推定に利用する。

2.2 先行研究

著者ら[4]は授業評価アンケートの項目間の関係性を、ベイジアンネットワークの構造学習によって示した。ベイジアンネットワークとは複数の確率変数の間の定性的な依存関係をグラフ構造によって表し、個々の変数の間の定量的な関係を条件付確率で表した確率モデルである。複雑な因果関係を事前確率と事後確率によって推定し、不確実性を伴う事象の推定に用いることが可能である。

先行研究においては、授業評価項目がおおよそどのような項目で構成されているのかを明らかにしており、本稿では抽出されたトピックを抽象的な項目として位置づけ、KJ法を用いて具体的な項目を作成した。質問項目は 75 題の 5 段階評価と、5 段階評価の理由について尋ねる 21 題に回答してもらった。

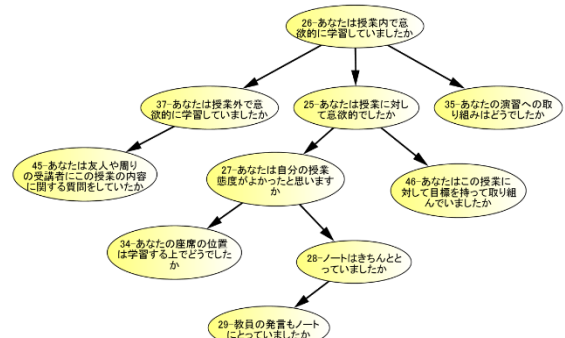


図 2 意欲に関する構造

[†] 東京理科大学大学院工学研究科経営工学専攻

[‡] 東京理科大学工学部情報工学科

[・] 東京理科大学工学部第二部経営工学科

表 1 問 26 と問 35 の事前事後確率

		問26. 授業内で意欲的に学習していたか				
		1.してなかつた	2.あまりしてなかつた	3.どちらともいえない	4.ややしていた	5.していた
問35 演習への取り組みはどうでしたか	1.してなかつた	0.304	0.125	0.056	0.031	0.037
	2.あまりしてなかつた	0.174	0.375	0.056	0.046	0.019
	3.どちらともいえない	0.217	0.219	0.286	0.092	0.056
	4.ややしていた	0.174	0.219	0.314	0.523	0.167
	5.していた	0.130	0.063	0.256	0.371	0.722

調査の結果、62 人から 66 種の科目について、欠損値を除いた有効回答数 184 件のデータを収集することができた。

図 2 は、構造学習をした結果の一部であり、質問項目が意欲に関する項目であることから、授業に対する意欲の因果関係と見て取れる。また、表 1 は問 26 と問 35 の事前事後確率である。表 1 より、問 26 の回答結果が「5.していた」のとき、演習への取り組みが「5.していた」の確率が 72.2%であることが推定できる。このように、項目の回答結果から、出題していない項目の回答結果を確率的に推定することが可能になった。

3. 実験概要と結果考察

構築した確率的因果関係を利用し、どの程度当てはまり良く推定できるかを検証する。先行研究で使用したデータの内、欠損値を含むデータを除いた 180 件のデータを対象にランダムに選択した 90% (162/180) のデータをトレーニングデータ、残りの 10% (18/180) をバリデーションデータとして扱い、トレーニングデータにより構造学習の結果とバリデーションデータとの当てはまりを評価する。

評価の方法として、バリデーションデータのうち 1~12 項目をエビデンス (出題した項目) として与え、残りの項目についてどの程度当てはまったかを確率の平均値で算出する。これにより、項目の出題数によって、他の項目の推定精度にどの程度影響するのかを明らかにする。

出題項目数を固定し、確率の平均値を基準に遺伝的アルゴリズム (以後 GA と称す) を用いて最もスコアの良い項目群を、「その項目数の中で最も推定精度の高い組み合わせ」と位置付ける。GA では、10 回のランダムサンプリングを 50 個生成し、出題項目をエビデンスとして与え、推定精度の高い上位 20% を次世代に利用した。計 10 世代生成した結果の中で、最も確率の高い項目群の確率の平均と項目数を図 3 に示す。

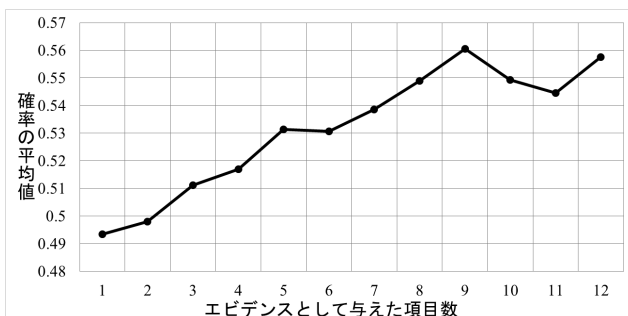


図 3 項目数と確率の平均値

確率の平均値は、出題していない他の項目の回答結果がどの程度当てはまっているかを示している。例えば、出題項目数 9 のとき、出題していない他の 66 項目のうち、37 項目 (約 56%) が正しく推定できることを示している。

図 3 より、エビデンスとして与えた項目数が 9 以下では、出題項目を増やすほど推定精度が向上傾向にあるが、以後項目数 10, 11 で下降する。さらに項目数 12 で向上している。このことから、項目数を単純に増やしても推定精度は向上しないと見える。これは、エビデンスの数が多すぎると、出題していない項目全体の推定確率を低下させる項目の組合せが発生しやすくなることが考えられる。よって、今回使用したアイテムバンクにおいては、出題数 9 項目が最も良い結果であることが示された。

4. まとめと今後の課題

本研究においては授業評価アンケートを具体的な改善に結びつけるために、項目の回答結果から、出題していない他の項目の回答結果を推定するアイテムバンクの作成を行った。授業評価アンケート項目が教員の具体的な改善に結びつくよう、具体的な項目の評価値を、少ない項目数で見出すことを目標に項目の関係性に着目し、階層型アイテムバンクを提案した。階層型アイテムバンクでは、抽象的な項目から具体的な項目の値を推定することにより、少ない項目数で具体的な改善に結びつけることができる。本アイテムバンク中の項目間の関係を調査するため、62 名の工学部大学生に対して授業評価アンケート調査を行い、推定精度の検証を行った。その結果、現状のアイテムバンクでは 9 項目出題により、少ない出題数でありながら、具体的な項目の評価値を推定できることを示した。

今後は、回答データを収集していく中で質問項目の追加及び削除をすることを適宜行う。さらに、項目の抽象度については、著者らが質問項目の作成にあたって定めた抽象度に過ぎない。今後は、項目の抽象度を定量的に定めることや、具体的な項目を得ることでどのような改善手法を提案できるかについても検討する。最後に、最終的にはこのアイテムバンクを利用した授業評価アンケートシステムの開発と評価を行なうことも今後の課題として挙げる。

謝辞

本研究は、科学技術振興機構 (JST) RISTEX 問題解決型サービス科学研究開発プログラム採択プロジェクト「高等教育を対象とした提供者のコンピテンシーと受給者のリテラシー向上による共創的価値の実現方法の開発」並びに、平成 28~31 年度科学研究費補助金基盤研究 (B) (課題番号 16H03086: 研究代表者赤倉貴子) の助成によるものである。

参考文献

- [1] 文部科学省, “学士課程教育の構築に向けて,” [http://www.mext.go.jp/component/b¥_menu/shingi/toushin/¥_icsFiles/afiedfile/2013/05/13/1212958¥_001.pdf](http://www.mext.go.jp/component/b_menu/shingi/toushin/icsFiles/afiedfile/2013/05/13/1212958¥_001.pdf), 参照 Jun. 15, 2016.
- [2] 文部科学省中央教育審議会, “大学分科会大学教育部会 配布資料 3-2: 「学士過程答申」の主な提言とその進捗の概要,” http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/015/gijiroku/icsFiles/afiedfile/2011/09/30/1311567_1.pdf, 参照 Jun. 15, 2016.
- [3] 阿久津洋巳, “授業評価アンケートは何を評価しているのか,” 岩手大学教育学部附属教育実践総合センター研究紀要, Vol. 13, pp. 245-252, 2014.
- [4] 中村修也, 石井隆稔, 藤森進, 赤倉貴子, “適応的授業評価アンケートシステム開発に向けた階層型アイテムバンク,” 電子情報通信学会技術研究報告, Vol. 115, No. 492, pp. 159-164, 2016.