

# 受験向け動画サービスにおける合否結果を加味した教材の推薦手法の提案

The Proposal of Recommend Method of Learning Contents Considering The Result of Entrance Examination In the Service of Movie For Entrance Examinations

後藤 拓矢<sup>\*1</sup>   那須野 薫<sup>\*1</sup>   萩原 静厳<sup>\*2</sup>   井上 綾香<sup>\*2</sup>   伊藤 岳人<sup>\*2</sup>   浜田 貴之<sup>\*3</sup>  
 Takuya Goto   Kaoru Nasuno   Seigen Hagiwara   Ayaka Inoue   Taketo Itoh   Takayuki Hamada  
 川上 登福<sup>\*3</sup>   松尾 豊<sup>\*1</sup>  
 Takayoshi Kawakami   Yutaka Mtsuo

<sup>\*1</sup>東京大学  
 The University of Tokyo

<sup>\*2</sup>(株)リクルートマーケティングパートナーズ  
 Recruit Marketing Partners Co., Ltd.

<sup>\*3</sup>株式会社 経営共創基盤  
 IGPI, Inc.

In the movie streaming service for study for entrance examinations, we propose the method of recommending learning contents assumed that learners pass the university of the their first choice, evaluating transitions of learning contents by using Q-learning algorithm. The purpose of this study is not only evaluate transitions that more learners who pass the university of their first choice do, but evaluate that fewer learner who drop the university do by using Q-learning algorithm.

## 1. はじめに

### 1.1 背景

近年、教育系サービスにも情報通信技術が活用されるようになった。情報通信技術を活用した教育系サービスでは、従来の、教室における教師との対面形式による学習や、教科書を紙の教材を用いた学習とは異なり、多くの物理的な制約を受けない [Jones 04]。文字や映像といった学習の媒体の多様化だけでなく、これまで時間と場所によって制約を受けていた学習内容がその制約を受けなくなったことで、教材自体の数が増加した。これまで、教科書通りの学習を行う他無かった学習者は、同じ目的でも様々な学習が可能になり、目的が違う場合にも必要に応じて教材を選択でき、自由な学習ができるようになった。しかし、同時に多すぎる教材の選択肢の中から、自分の目的を達成するにはどの教材をどの順序で用いれば効果的であるかわからず、困惑する学習者も出てきた [Manouselis 11]。そこで、情報通信技術を活用した教育系サービスにおいて、学習者の目的に応じて、学習を効果を考慮した学習教材の推薦によって学習者の学習を補助することが求められていると考えられる。

これまで日本の大学受験を対象とした、情報通信技術を用いた受験向け動画サービスがこれまで展開されてこなかった理由としては、学習塾や予備校の存在が挙げられる。学習塾や予備校には大学受験に熟知した講師がいるため、合格の為に最適な学習はその講師に聞くことで目的は達成された。しかし、2013年の調査では、大学の一般入試の受験生の内 65% が「経済的な事情などで塾・予備校に通っていない」と回答しており<sup>\*1</sup>、学習と受験に関する知識と技術を有する講師のいる為に高額の費用を要する予備校や学習塾に代わる、安価な受験向け教育系サービスが求められているといえる。実際に、近年、情報通信技術による動画配信を用いた安価な受験向け動画サービスが増

加してきていることから、受験の合否結果を加味した教材の推薦が求められることになると考えられる。

### 1.2 目的

第 1.1 節の背景を踏まえ、本研究では、次の教材を推薦する際に、学習した結果の是非を加味することで、単に利用した学習者が多い教材を推薦するだけではなく、大きい学習の効果を期待できる教材を推薦することを目的とする。これは、利用した学習者が多い教材でも学習の効果が小さい教材は推薦するべきではないと考えられるからである。利用した学習者が多い教材でも、その学習者の多くが良い結果を得られなかったという教材があることを想定できる。単に利用した学習者が多い教材を推薦した場合には、このような学習効果が小さい教材も推薦されることが予想される。

本稿では大学受験の合否結果を用いた手法を提案する。受験向けの教育系サービスにおいて、学習者は第 1 志望に合格することが学習の目的としていられると考えられる。したがって、第 1 志望に合格することは学習の効果が大きいことと見なすことができる。本研究の提案手法では、ある教材の次に学習するべき教材を評価する際に強化学習の 1 つである Q-learning を用いる。Q-learning によって、より第 1 志望に合格する可能性が高くなると考えられる教材を高く評価し、直前に学習した教材と直後に学習する教材の依存関係を推定する。

本論文の新規性と有用性は以下の通りである。

- 学習の履歴を用いた教材の依存関係の推定において、未だどの研究にも用いられていない日本の大学受験における合否結果を加味した。
- Q-Learning による教材間の依存関係の推定は受験に合格するという目的だけでなく、様々な目的の達成を評価できるため、教育系サービスにおいて汎用性が高く、有用性が高い。

連絡先: 後藤 拓矢, 東京大学工学部システム創成学科,  
 〒 1138654 東京都文京区本郷 7-3-1 工学部 2 号館,  
 goto@weblab.t.u-tokyo.ac.jp

<sup>\*1</sup> 日本経済新聞 (2013 年 10 月 19 日)

<http://www.nikkei.com/article/DGXDZO61258500Y3A011C1W04001/>

## 2. 従来の手法の問題点

従来の教材の推薦方法として通常以下の2つの手法が挙げられる。

- ルールベースによる一意的な推定
- ユーザーベース協調フィルタリングを用いた推定

ルールベースとは教科書順や予備校の指導方針など事前に教育分野に精通している者によって合格するのに最適であるとした推定をそのまま使用する方法である。この方法は熟練の講師の経験や過去の研究もふまえた専門的な知識をもとに推定されており、結果の信頼度が高い。しかし、学習者は日常的に前に学習した内容を復習したり、学習分野によって講義の難易度を変更しているが、このルールベースによる方法では、教科書の順をみればわかるようにこのような遷移は考慮されていない[松澤 08]。さらに、人が直接判断しなくてはならないため、時間や人件費のコストが大きい。

ユーザーベース協調フィルタリングを用いた推定は過去の遷移をもとに過去の学習者との類似度を計算し、過去の学習者の遷移先にその類似度と合否結果によって重みをつけて計算を行う方法である。ユーザーベース協調フィルタリングを用いた方法は学習において目的を達成することに最適化した推薦システムの研究に多く用いられてきた[Capuano 12]。ユーザーベースの協調フィルタリングは実際の履歴に基づいて推定が行われるため復習や難易度の変更なども考慮でき、また、学習者それぞれに個別化できるため、多様な学習スタイルや学力の背景に対応できるという利点がある。しかし、ユーザーベース協調フィルタリングは計算を推薦の度に毎回行われ、結果が出るまでに時間がかかる。また、類似度の計算にある程度の履歴が必要であるため、サービスの利用して初期の頃は推定が適切に行えないという欠点があると考えられる。

## 3. 提案手法

本研究では、学習者が講義を受講している状態を1つの状態にあるとして、次の講義を受講することを状態の遷移が起きていると見なせると考えた。また、講義から講義へ遷移していった後で、受験に合格することは、状態が遷移していく中で報酬が与えられることと見なせると考えられる。したがって、受験勉強はマルコフ決定過程によるモデル化が可能であると考えた。そこで、マルコフ決定過程における、報酬獲得のために最適な遷移を学習する Q-learning によって、受験勉強における合格するのに最適な遷移を抽出できるのではないかと考えた。したがって、本研究では、過去の履歴データを用いて Q-learning によって合格するのに最適な遷移モデルを推定し、そのモデルに基づいて推薦を行う手法を提案する。本手法は推薦の前に事前にモデルを推定する為に計算が必要であるが、推薦の際には計算をほとんど行わなくていいため、計算コストが低いと言える。また、モデルに基づけばサービスの利用を始めただけの学習者に対しても推薦が行える。さらに、普通の学習で行っている学習をしていき定期テストや受験の成績によってそれまでの学習方法を改めるという過程は、Q-learning などのモデルとなるマルコフ決定過程のモデルによく合致していると考えられるため、Q-learning による学習方法の改善は可能であると考えた。

提案手法の具体的な計算方法を述べる。手法に用いる講義の集合を  $L$  とおく。本手法では「講義〇〇を受講している」ことを1つの状態とみなし、行動を「講義〇〇から講義××に

遷移する」としているため、Q 値は受講している講義  $a$  と次に受講する講義  $b$  によって定義され、 $Q(a, b)$  と設定する。したがって、講義の集合  $L$  の要素数を  $p$  とすると Q 値は  $p^2$  個設定されることになる。また、推定に用いる合否結果がわかっている学習者の集合を  $U = \{u_1, u_2, \dots, u_n\}$  とする。Q 値の推定は以下のアルゴリズムに従って行う。

1. 全ての Q 値を 0 に初期化を行う。
2. 学習者の集合  $U$  に含まれる学習者をランダムに並べる。このときに並べた学習者の列を学習者列  $u$  とする。
3. 学習者列  $u$  の先頭から1人の学習者  $u_i$  を取得し、 $u_i$  の合否結果  $r_i$  と  $u_i$  が取得した講義を順に並べた取得講義列  $l_i = (l_i^1, l_i^2, \dots, l_i^{m_i})$  を取得する。( $l_i^1, l_i^2, \dots, l_i^{m_i} \in L$ )
4. 取得した合否結果  $r_i$  と  $l_i$  を用いて Q-learning のアルゴリズムに基づいて Q 値を更新する。
  - (a)  $t = 0$  とする。
  - (b)  $l_i$  の  $t$  番目の講義  $l_i^t$  を受講している状態とする。
  - (c) 講義  $l_i^t$  から  $l_i^{t+1}$  に遷移したと見なして Q 値を更新する。
  - (d)  $t = t + 1$  とする。
  - (e) (b) から (d) を  $t = m_i$  となるまで繰り返す。
5. 学習者列  $u$  の先頭  $u_i$  を  $u$  から取り除く。
6. 学習者列  $u$  の要素がなくなるまで3から5を繰り返す。

$$Q(s_t, a_t) \leftarrow (1 - \alpha)Q(s_t, a_t) + \alpha r_t + \alpha \gamma \max_{a \in A} Q(s_{t+1}, a) \quad (1)$$

- $s_t$  : 時刻  $t$  における状態  
 $s_{t+1}$  : 時刻  $t + 1$  における状態  
 $a_t$  : 時刻  $t$  における行動  
 $r_t$  : 行動  $a_t$  によって得られる報酬  
 $A$  : 全行動の集合

なお、3. における合否結果  $r_i$  は合格していた場合 1、不合格の場合に 0 となる。また、4. の (c) の Q 値の更新は Q-learning の更新式 1 を本手法の場合に合わせた以下の式 2 によって更新を行う。Q 値は状態と行動の組み合わせごとに定義されるため、今回は遷移前の講義  $a$  と遷移先の講義  $b$  ごとに定義され、 $Q(a, b)$  と表すとする。合否結果が  $r_i$  の学習者  $u_i$  が  $t$  番目に講義  $l_i^t$  を受講し、 $t + 1$  番目に講義  $l_i^{t+1}$  に遷移したとすると

$$Q(l_i^t, l_i^{t+1}) \leftarrow (1 - \alpha)Q(l_i^t, l_i^{t+1}) + \alpha r_i R + \alpha \gamma \max_{l \in L} Q(l_i^{t+1}, l) \quad (2)$$

上記の式の右側の  $Q(l_i^t, l_i^{t+1})$  や  $Q(l_i^{t+1}, l)$  は更新前の推定されている Q 値を表している。

なお、本手法では学習率  $\alpha$  を 0.6、割引率  $\gamma$  を 0.8 とし、合格者の遷移の場合に与える固定の報酬  $R$  を 5 とした。

以上の方法によって推定された Q 値を元に、ある講義  $l'$  を受講した際により合格する可能性の高い遷移先の講義  $l''$  を推

定する。そのとき講義  $l'$  から講義  $l''$  への遷移についての  $Q$  値  $Q(l', l'')$  は以下の条件を満たす。

$$Q(l', l'') = \max_{l \in L} Q(l', l) \quad (3)$$

## 4. 実験

### 4.1 実験手法

提案手法によって、実際に合格するために最適な遷移を推薦できるかを検証するために、学習履歴と合否結果のデータに対して提案手法を適用して評価実験を行った。提案手法によって得られた  $Q$  値が最大の遷移が、合格するために最適な遷移であるならば、実際の合格者は不合格者より、この最適な遷移をより多く取っているはずである。したがって、評価実験では実際のサービス利用者のデータの内ランダムに 8 割を学習用データ、残りの 2 割をテスト用データにわけ、テストデータにおける合格者と不合格者それぞれで  $Q$  値が最大の行動をとっている割合を算出し、比較した。また、目安として第 1 講から第 2 講への遷移のように講義の整列順序の通りの遷移の割合とも比較する。講義の整列順序は教科書の通りと同じであるため、これはルールベースの手法との比較になる。

実験に使用するデータには、(株)リクルートマーケティングパートナーズが運営する受験サプリの 2013 年度データを使用する。受験サプリは 1000 を超える講義数を有しており、全ての遷移を考慮すると計算量が膨大な量となる上、教科をまたいだ遷移に強い相関関係が想定できないことから、実験に用いる講義は英語の講義に限定する。また、学習者には、2013 年度の受験生の会員で、サービス利用頻度が高い学習者のうち、アンケート調査によって進学大学等が得られた約 450 人のデータを用いた。同データでは、合格者、不合格者共にほぼ半数程度存在している。

### 4.2 実験結果

表 1: ルールベースと提案手法の結果の比較

講義の整列順に	合格者	0.426
遷移した割合	不合格者	0.403
Q 値が最大の	合格者	0.174
遷移をした割合	不合格者	0.167

表 1 に実験結果の集計を示す。結果の数値は同じ手順で行った 4 回の実験結果の数値の平均を表している。実験の結果、合格者で  $Q$  値が最大の遷移をした割合が、講義の整列順に遷移した割合より小さかったため、提案手法はルールベースより、合格するために適した遷移を抽出することができなかったと言える。

次に、提案手法の実験結果である  $Q$  値が最大の遷移をした割合を合格者と不合格者と比較すると、合格者の方が 1 % 程、不合格者より  $Q$  値が最大の行動をしていることがわかる。この結果から、提案する手法によって合格するのに適した遷移を評価しているといえるが、その差が 1 % 程度であり、この結果を、合格するのに最適な遷移であるということは難しいと考えられる。

ただし、講義の整列順に遷移した割合を見てみると、合格者も不合格者の同様に 4 割程度となっており、残りの 6 割は独自に判断した遷移を取っていると考えられる。したがって、

ルールベース以外の手法によって、教材の推薦する余地があると言える。

## 5. 考察

今回、提案手法の結果が悪かった原因の 1 つとして状態の設定に改善の余地があることが挙げられる。今回の手法では 1 つの状態を「ある講義を受講している状態」としたが、学習において同じ講義を受講している状態でも、サービス利用初期にその講義を受講している状態と、サービスを利用してしばらくたってからその講義を受講している状態は同じではないと考えられるため、これらの状態を区別するような状態の設定が必要である。

また、使用したデータが少なかつたために、学習が収束しきれていなかったことも考えられる。 $Q$ -learning の収束には全ての状態において、取り得る行動が十分回なされることが必要であるが、今回使用したデータでは全ての行動が十分回取られているというには不十分な量であった。

## 6. まとめ

本研究では  $Q$ -learning のアルゴリズムを取り入れることで、合格に最適な教材遷移の抽出を目的とした。しかし、データ数が不十分であることと手法のモデル設定に改善の余地があったため、有意な結果が得られなかった。したがって、今後はよりデータ数を増やすことと、過去の講義の受講履歴も状態に含めることで、手法の改善を試みたい。

## 謝辞

株式会社リクルートマーケティングパートナーズには、使用した受験サプリにおけるデータを提供いただきました。この場を借りてお礼申し上げます。

## 参考文献

- [Capuano 12] Capuano, N., Mangione, G. R., Pierri, A., and Salerno, S.: Learning Goals recommendation for self regulated learning, *future*, Vol. 1373, (2012)
- [Jones 04] Jones, V. and Jo, J. H.: Ubiquitous learning environment: An adaptive teaching system using ubiquitous technology, in *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference*, 468, p. 474 (2004)
- [Manouselis 11] Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., and Koper, R.: Recommender systems in technology enhanced learning, in *Recommender systems handbook*, pp. 387–415, Springer (2011)
- [松澤 08] 松澤俊典, 山口未来, 和田雄次, 土肥紳一: 教材コンテンツ双方向推薦システムの実装, 情報処理学会研究報告. コンピュータと教育研究会報告, Vol. 2008, No. 13, pp. 127–132 (2008)