

大規模オンライン講座における自己適応学習者に着目した 学習項目の理解度予測

Predicting Understanding of Learning Items
Focusing on self-Adaptive Learners in Massive Online Course

那須野薫 *1
Kaoru Nasuno

萩原静蔵 *2
Seigen Hagiwara

井上綾香 *2
Ayaka Inoue

伊藤岳人 *2
Gakuto Ito

浜田貴之 *3
Takayuki Hamada

川上登福 *3
Takayoshi Kawakami

松尾豊 *1
Yutaka Matsuo

*1 東京大学
The University of Tokyo

*2 株式会社リクルートマーケティングパートナーズ
Recruit Marketing Partners Co., Ltd.

*3 株式会社 経営共創基盤
IGPI, Inc.

Considering not only one's learning activities but also one's understanding of learning items is important for recommendation of learning materials. But with respect to massive online course, it is difficult to evaluate the degree of understanding of all learners for all learning items from the result of assessment tests because their learning materials and learning pace is different from each other. In this paper, we predict whether or not learners can understand its content if they tackle on learning material, focusing on self-adaptive learners who choose their learning materials adaptively by themselves in massive online course.

1. はじめに

近年、大規模オンライン講座が注目を集めている。従来の教室で時間割通りに一斉授業形式で提供される学習機会とは異なり、大規模オンライン講座は誰もがいつでもどこでも自分のペースで自分の学びたいことを学習できるという、これまでにない学習機会を提供しており、次世代の学習環境を提供する可能性が期待され、その注目が高まっている。

しかし、大規模オンライン講座はまだ発展段階であり、いくつか課題を抱えている。例えば、自発的な勉強が中心であるため、逆に、つまずきの影響が大きい学習環境となっていることは大規模オンライン講座の抱える大きな課題の一つである。学習者は大規模オンライン講座で自発的に勉強するなかですぐには理解できない学習項目に遭遇しつまずいた場合、必ずしも任意のタイミングで先生や他の学習者から助言を受けられるというわけではないこともあり、その内容を理解するために繰り返し勉強するか、もしくは、その内容を理解に必要な別の教材を探すことになりがちで、理解するまで少なくない時間がかかってしまい効率が悪く、このようなつまずきの影響を減らすことは大規模オンライン講座での効率的な学習環境の提供に寄与し、ひいては学習者の信頼の獲得にも繋がるため重要である。

このようなつまずきの影響を減らすためには学習教材の推薦が有効であると考えられる。学習教材の推薦方法はいくつか考えられるが、特につまずきの影響を減らすものとして、1) 学習者がある教材を勉強した後にその内容を理解できているかを評価し、理解できていれば、次の教材を推薦し、そうでなければ、その教材を理解するために必要な別の教材を推薦するという方法や、2) 学習者がある教材を勉強する前にその教材を勉強したらその内容を理解できそうかを予測し、理解できそうであれば、そのまま勉強させ、そうでなければ、その教材を理解するために必要な別の教材を推薦するという方法が考えられる。1) の方法はつまずいた後にフォローしてつまずきの影響を減らすというものであり、2) の方法はつまずかないよう

にフォローしてつまずきの影響を減らすというものである。

これらの推薦方法のいずれも理解度の評価や予測が必要であるが、理解度の評価や予測に用いられてきた従来手法の適用可能性は大規模オンライン講座では限定的である。従来の主に紙を利用する学習環境やeラーニング環境では、学習者の理解度を評価するために主にテストが利用されてきた。主に紙を利用する学習環境とeラーニング環境ではコンピュータでの自動採点の有無の違いから利用できる出題形式の違いはあるものの、定期的に生徒全員にテストを解かせ、前者であれば論述問題や選択問題等の解答の正誤から理解度が評価され、後者であればコンピュータによる採点が比較的簡単な選択問題や穴埋め問題などの正誤から理解度が評価されてきた [菅沼 05, 難波 11]。しかし、大規模オンライン講座では利用教材や学習ペースの選択は学習者に委ねられるため、テストの結果から統一的、網羅的に理解度を評価することは難しく、従って、テストの結果のみから評価された理解度に基づいて個々の学習者に個別に教材推薦を行うのは難しい。

それでは、大規模オンライン講座では全ての学習者がつまずきの影響を受け効率の悪い学習をしているのか、ということではないと考えられる。教育心理学の研究 [野上 05] では、学習者には自分の認知状態に関する認知であるメタ認知の制御力が強い学習者があり、このような学習者は自分に適した学習計画を立て実行することで効率的に学習できることが指摘されており、また、大規模オンライン講座は数多くの多様な学習者が利用していることを踏まえると、大規模オンライン講座上の学習者の中にもメタ認知の制御力が強く自分自身で適応的に利用教材を選択する学習者（以下、自己適応学習者）はいると考えられる。すなわち、自己適応学習者とは自分で学習項目の理解度を認知し、十分理解していると判断すれば次の学習項目に着手し、そうでなければ、その項目の理解に必要な別の学習項目に着手し自ら学習を進めることができる学習者のことである。

このように効率的な学習を試みる自己適応学習者は自分の理解度に応じて学習行動を選択するため、逆に学習履歴データを分析することで、理解度の評価や予測ができる可能性がある。

連絡先: 那須野薫, 東京大学, nasuno@weblab.t.u-tokyo.ac.jp

大規模オンライン講座の特徴として、学習者の学習履歴データを分析することで個々の学習者の学習方法や学習ペースを抽出できることが挙げられる。学習方法は同じ人でも取り組む教科や取り組む項目の難易度などで必ずしも同じではないため厳密には分類できないが、例えば、教科ごとに学習履歴データを分析することで自己適応学習者を抽出できる可能性はある。また、自己適応学習者は自分の理解度に応じて学習行動を選択するため、逆に抽出した自己適応学習者の学習履歴データを分析することで、ある教材を勉強したらその内容を理解できるかといった将来の理解度を予測できる可能性がある。

本稿で報告する研究は、大学入学試験対策向けの大規模オンライン講座「受験サプリー」*1 を運営する(株)リクルートマーケティングパートナーズ*2 との共同研究プロジェクトの一貫で行われている。受験サプリーは2013年度には数学、国語、英語などに関する130以上の講座とそれを構成する1600以上の講義動画を提供し、日本の全47都道府県の30万人以上の学習者が利用した大規模オンライン講座である。本研究プロジェクトでは受験サプリーのユーザに、より効率的な学習環境を提供することを目指し、適応学習システムを実装することを主要な目的の一つとしている。

以上の全体の背景を踏まえ、本研究では、受験サプリー上の学習者の理解度を予測する手法の開発を目指し、手法の評価や精度向上に貢献する知見の獲得を通して、今後の研究方針を整理することを目的とする。2013年度に受験生であったユーザを対象とし数学の自己適応学習者を抽出し、ある講義の視聴後に連続してその講義の内容の理解に必要な講義(以下、依存先講義)を視聴するか否かをその講義を視聴する前に予測するというを試みる。予測は教師あり学習のアルゴリズムを利用し、素性には視聴する前の学習状況に関する指標、学習者の性別や現役/浪人等の属性に関する指標、学習者の高校や志望大学のレベルに関する指標を利用する。本研究の問題設定は、視聴後に連続してその講義の依存先講義を視聴するとまえて予測される場合には、視聴しても十分な理解は得られないと判断し、視聴前に依存先講義を推薦するという適応学習システムを意識したものである。

評価実験では、視聴後にその講義の依存先講義を視聴するか否かの2値分類を教師あり学習のアルゴリズムを利用して行う。予測に寄与する指標を特定するために、各指標を素性とした場合の予測性能を比較し、また、最も予測性能の高い予測モデルの素性の重みを分析することで、学習者が依存先講義に戻りやすいパターンの抽出を試みる。

本研究の貢献は以下の通りである。

- 大規模オンライン講座のユーザを対象に学習項目の理解度を予測した事例を報告した。
- 本研究で抽出した自己適応学習者は自分の学力に適した教材を選択して利用している可能性が高いことを示唆した。
- 講義間の依存関係を詳細に定義することで理解度予測の精度が向上する可能性を示唆した。

2. 背景

2.1 関連研究

理解度評価や理解度予測を行う先行研究として、菅沼ら[菅沼 05]は練習問題の正誤をもとに理解度を評価するeラー

ニングシステムを報告し、Shiら[Shi 11]はオンライン講座を対象とする典型的なファジィ理解評価モデルの事例を報告した。しかし、大規模オンライン講座のユーザを対象として理解度予測を試みる研究は著者らが知る限りまだ報告されていない。

2.2 受験サプリーの概要

受験サプリーは大学入学試験対策向けの大規模オンライン講座であり(株)リクルートマーケティングパートナーズが運営している。受験サプリーは2013年度には数学、国語、英語などに関する130以上の講座とそれを構成する1600以上の講義動画を提供し、日本の全47都道府県の30万人以上の学習者が利用した大規模オンライン講座で、無料で提供も行われている。受験サプリーのオンライン講座としての設計を明示するという目的で、Bates[Bates 14]が抽出したxMOOCsに共通する特徴的な設計と比較すると、受験サプリーには、専用に設計されたプラットフォーム、講義動画、自動採点課題、補助教材、学習分析の要素は実装されているが、ピアアセスメント、議論や質問の共有掲示板、バッジや修了証などは設計されていない。また、特に数学の講座は、講座ごとの内容(数学IA、数学IIBなど)や難易度レベル(トップレベルやハイレベルなど)で概ね分類されており、各講座は順番のある複数の講義から構成される。

3. 問題設定

本研究では、数学の自己適応学習者を対象に、学習項目の理解度をそれまでの学習状況や学習者の属性などを素性として教師あり学習のアルゴリズムを利用して予測する。まず、自己適応学習者について、次に、本研究で予測する理解度について、最後に、予測に用いる素性について説明する。

3.1 自己適応学習者

ある講義Xを視聴した後の視聴パターンは、A)講義Xの理解に必要な別の講義(以下、依存先講義)を視聴する、B)講義Xの理解を必要とする別の講義(以下、被依存先講義)を視聴する、C)講義Xの理解とは無関係の講義(以下、依存関係のない講義)を視聴する、の3つがある。自己適応学習者は自分の理解度を考慮して学習するため、理解度を考慮していないと表れにくい視聴行動を多く行っている可能性が高い。特に、上記のA)、B)、C)の中では、A)は理解度を把握していないと起こりにくいが、他のB)やC)は理解度を把握していないなくてもしばしば起こりうると考えられる。なぜなら、B)は単に次の講義に進むだけの行動であり、また、C)は学習者が講義で扱われる学習項目の理解度が不十分であるとオフラインで判断した場合にも生じうり、後者の場合も珍しいというわけではないからである。従って、本研究では、ある講義Xを視聴した後に特にA)講義Xの依存先講義を視聴するという視聴行動の割合が高い学習者を自己適応学習者とする。

3.2 理解度

自己適応学習者は自分で教材の理解度を認知し、十分理解していると判断すれば次の教材に着手し、そうでなければ、依存先講義に着手し自ら学習を進める学習者であるため、ある講義Xの視聴後に連続して講義Xの依存先講義を視聴した場合は講義Xを十分理解できなかったと推察され、その他の場合は講義Xを十分理解できたと推察される。

講義視聴の遷移の間に多くの時間が空いている場合、オフライン上で学習した可能性があるため、連続して視聴したか否かも考慮する。視聴の遷移は遷移先が講義Xの依存先講義か否

*1 <https://jyukensapuri.jp>

*2 <http://www.recruit-mp.co.jp>

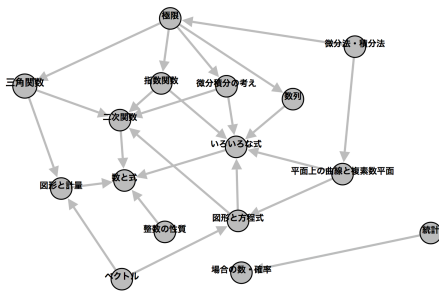


図 1: 系統間の依存関係ネットワーク図．シンクノードが依存先の系統を表す．

か、とすぐに視聴を行ったか否かで分類すると、1-a) 遷移先が講義 X の依存先講義で視聴後すぐに遷移した、1-b) 遷移先が講義 X の依存先講義で視聴後時間を空けて遷移した、2-a) 遷移先が講義 X の依存先講義でなく視聴後すぐに遷移した、2-b) 遷移先が講義 X の依存先講義でなく視聴後時間を空けて遷移した、の 4 つに分類される。1-a) と 1-b) の場合は講義 X は十分理解できなかったと推察され、2-a) の場合は講義 X は十分理解できたと推察されるが、2-b) の場合は理解度が十分であった可能性もあるがオフラインで学習し理解してオンライン上の非依存先講義や依存関係のない講義に遷移した可能性も考えられるため、講義 X の理解度は不明である。そこで、本研究では、1-a) と 1-b) に該当する視聴後に十分な理解が得られずに依存先講義に戻る視聴を戻り視聴とし、2-a) に該当する視聴後に十分な理解が得られ依存先講義でない講義に進む視聴を進み視聴とし、これらの 2 つの視聴を理解度として扱いそれまでの学習状況や学習者の属性等から予測するというを試みる。

3.3 素性

まず、学習状況に関する指標について説明する。ある講義 X の理解の有無には依存関係のある講義の理解が大きく関わっていると考えられ、依存先講義や非依存先講義などの理解度の指標は講義 X の理解度をよく説明する可能性があるため、これらを考慮して学習状況に関する指標を作成する。まず、講義間の依存関係の定義について説明する。本研究の対象教科である数学には一次関数やベクトル等の系統が存在し、系統間には図 1 のような依存関係が定義されているため、講義間の依存関係の定義には図 1 から再帰的に定義した系統間依存関係を利用する^{*3}。再帰的に定義した系統間依存関係と、受験サプリの講義に定義されている難易度レベルを合わせて、タイプ 1) 同系統 AND 同レベル AND 同講座 AND 講義番号が小^{*4}、タイプ 2) 依存系統 AND 同レベル、タイプ 3) 同系統 AND 下位レベル、タイプ 4) 依存系統 AND 下位レベル、の 4 つの依存タイプを定義した。依存先講義群で 4 つの依存タイプと非依存先講義群で 4 つの依存タイプがあり、対象講義の過去の利用を含めると講義は合計で 9 種類に分類される。計算のため、戻り視聴における講義 X の理解度を 0、進み視聴にお

*3 系統間の依存関係ネットワークは [東京書籍 (株) 12] を参考に作成し、講義の所属系統は講義名や補助教材の内容から著者が判断した「再帰的に」というのは、例えば、三角関数は二次関数や図形と計量だけでなく二次関数や図形と計量が依存する数と式にも依存することを指す。

*4 同系統 AND 同レベルの講義間には同じ講座内のものについては講義番号の大小に基づいて定義できるが、そうでない場合は定義が難しいため、このように定義した。

表 1: ROC-AUC による予測モデルの評価。

素性の種類	アルゴリズム		
	ロジスティック回帰	SVM	AdaBoost 分類器
全指標	0.705	0.699	0.722
学習指標	0.704	0.695	0.721
属性指標	0.503	0.500	0.508
ランク指標	0.507	0.500	0.514

る講義 X の理解度を 1、それ以外の視聴における講義 X の理解度を 0.5 として、これら 9 種類の講義群に対してそれぞれ、着手講義数、着手した講義の理解度の平均、着手した講義の理解度の標準偏差、着手した講義の理解度の最大値、着手した講義の理解度の最小値の 5 つの指標を定義し、以降では、これらの合計 45 個の指標を学習状況指標と呼ぶこととする。

次に、学習者の属性やランクの指標について説明する。学習者の属性に関する指標として、性別 (男=1, 女=0)、現役/浪人 (現役=1, 浪人=0)、文理 (文系=0, 理系=1, 不明=0.5) の 3 つの指標を定義し、以降ではこれらを属性指標と呼ぶこととする。また、学習者のランクに関する指標として、学習者の所属高校の学力ランク、第一志望大学の学力ランク、ランク差、の 3 つの指標を定義し、以降ではこれらをランク指標と呼ぶこととする。

以上、合計 51 の指標を素性として理解度の予測を試みる。なお、指標の詳細な定式化については紙面の都合で割愛する。

4. データ

本章では、評価実験で用いる受験サプリのデータについて説明する。まず、受験サプリのデータは、特にサービス運用のために保存されているデータ (以下、サービスデータ) と、特に分析のために保存されているデータ (以下、ログデータ) の 2 つに大別される。サービスデータは、ユーザに関するデータ、教材に関するデータ、学校に関するデータ等から構成されている。ログデータは、アクセスログ解析ツールを用いて取得された各ユーザがいつどのページに訪れたかというアクセスログデータである。ユーザ数やログデータはレコード数という点では、比較的大きな規模のデータが得られた。2013 年 4 月から 2014 年 3 月に受験サプリを利用していたユーザは 30 万人以上で、特に、本研究では、比較的活発に利用していたユーザの中から、約 3 万弱のユーザのデータを分析に利用する。

先の章で説明した手法で、受験生のユーザから数学の自己適応学習者を抽出したところ 380 人抽出され、戻り視聴は 4 千件弱、進み視聴は 8 千件強抽出された。

5. 評価実験

本章では、評価実験について説明する。先の章で説明した問題設定に従い、予測手法の評価を行う。進み視聴と戻り視聴を合わせた約 1 万 2 千件の視聴データを 80% が訓練データに 20% がテストデータになるようにランダムに分割して訓練データで予測モデルを構築しテストデータに適用して予測精度の評価を行う。進み視聴を正例とし、予測精度の評価指標には ROC-AUC を使い、教師あり学習のアルゴリズムはロジスティック回帰、SVM、AdaBoost 分類器の 3 つを利用する。各アルゴリズムの実装には、Python の機械学習ライブラリである Scikit-learn [Pedregosa 11] を利用した。

表 2: AdaBoost 分類器より構築した予測モデルにおける素性の重みと各クラスの平均値。重みの降順に 10 個を記載。

指標名		重み	戻り視聴 の平均値	進み視聴 の平均値
依存/タイプ 1/着手講義数	-	0.120	0.865	0.592
依存/タイプ 3/着手講義数	-	0.0650	1.00	0.787
被依存/タイプ 2/着手講義数	+	0.0650	1.00	1.95
同講義/着手回数	+	0.0500	0.512	0.565
依存/タイプ 1/理解度平均	-	0.0500	0.320	0.257
依存/タイプ 2/理解度平均	-	0.0500	0.472	0.367
被依存/タイプ 3/理解度標準偏差	+	0.0500	0.0258	0.0344
依存/タイプ 2/着手講義数	-	0.0450	3.92	2.16
被依存/タイプ 3/着手講義数	+	0.0450	0.426	0.603
依存/タイプ 3/理解度標準偏差	-	0.0400	0.0488	0.0372

まず、各指標と各アルゴリズムを用いた時の実験結果を表 1 に示す。全ての場合において AdaBoost 分類器を用いた手法の予測性能が最も高かった。全指標を用い AdaBoost 分類器を利用する手法の ROC-AUC は 0.722 と中程度の予測性能が得られ、過去の学習行動から将来の理解度を反映する学習行動を予測できることが示唆された。属性指標のみを用いた手法は精度が低く、性別や現役生か浪人生であるかは自己適応学習の傾向にはほとんど関係ないことが示唆された。また、ランク指標のみを用いた手法も精度が低く、自己適応学習者はそのアイデア通り自分の学力に適した教材を選択して利用していることが示唆された。

次に、最も予測精度が高かった AdaBoost 分類器より構築した予測モデルにおける素性の重みと各視聴における指標の平均値を表 2 に示す。表の符号は進み視聴の平均値 - 戻り視聴の平均値 ≥ 0 の場合に + とし、それ以外は - としている。全体として依存先講義に関する着手講義数や理解度の平均の指標は符号が - であることから、後ろの方の講義を着手した場合理解度が得にくいことが推察され、逆に、被依存先講義に関する着手講義数や理解度の平均の指標の符号が + であることから、後ろの方の講義を着手した後に前の講義に視聴した場合復習での利用であることが推察される。

6. 考察

本研究手法の限界と今後の拡張可能性について考察する。まず、本研究で扱う理解度が真に理解度であるかについては本稿では主張しない。真に理解度足り得るかは古くから理解度の評価に利用されてきたテストの点数との比較や成績向上度合いや学習の継続度合いとの比較を通して議論されるべきであると考えている。次に、本研究での学習行動の予測性能は自己適応学習者についてのものであり、自己適応学習者でない学習者への適用可能性については本稿では主張しない。なぜなら、自己適応学習者でない学習者は学習項目の理解度に応じて学習行動を選択していない場合が多いと推察されるからである。しかしながら、仮に視聴ごとの理解度向上が学習者の属性に依存せず学習者の学習状況にのみ依存するのであれば、従来のオフライン学習で考慮されなかった些末な学習行動もデータの分析を通して考慮できるということもあり、自己適応学習者でない学習者についても大規模オンライン講座上の学習行動から理解度を予測できる可能性はある。

次に、本研究の今後の拡張可能性について、特に、1) 講義間依存関係の詳細化、2) 過去の確認テストの点数の考慮、3) 他教科への拡張、の 3 つの拡張を検討している。1) 講義間依

存関係の詳細化について、本研究では講義間の依存関係は 4 つの依存タイプで定義されていたが、実際には図 1 の系統間の依存関係ネットワーク図のノードを講義にしたような細かい依存関係を定義できるはずである。2) 過去の確認テストの点数の考慮について、本研究では講義視聴のみから理解度の予測を試みたが、受験サプリーには各講義ごとに確認テストが存在するため、その点数を考慮することで、さらに一般性や精度の高い予測モデルを構築できる可能性がある。3) 他教科への拡張について、本研究では、数学のみを対象としたが、受験サプリー上で提供される講義は数学だけではないため、他教科への適用可能性の高い手法を考案することも今後の課題である。

7. まとめ

本研究は、受験サプリー上で、自分自身で適応的に利用教材を選択する学習者である自己適応学習者に着目し、ある講義の視聴後にその内容の理解に必要な講義（以下、依存先講義）を視聴するか否かをその講義の視聴前に予測するという試みを行った。予測は教師あり学習のアルゴリズムを利用し、素性には視聴する前の学習状況に関する 45 指標、学習者の属性に関する 3 指標、学習者のレベルに関する 3 指標を利用した。評価実験の結果、予測モデルの ROC-AUC は 0.722 と中程度の予測性能が得られ、また、依存先講義を視聴するか否かには学習者の属性や学力レベルや志望大学のレベルは無関係であり、本研究で抽出した自己適応学習者はそのアイデア通り自分の学力に適した教材を選択して利用していることが示唆された。

今後は受験サプリーへの適応学習システムの実装に向けてさらに研究を進める予定である。本研究が適応学習システム実装の足がかりとなり、ひいては人間の学習メカニズム解明の一助となれば幸いである。

参考文献

- [Bates 14] A. W. T.: *Teaching in a Digital Age*, BC Open Textbooks (2014)
- [Pedregosa 11] F., G., A., V., B., O., M., P., R., V., J., A., D., M., M., and E.: Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, Vol. 12, pp. 2825–2830 (2011)
- [Shi 11] F., L., and Y.: The fuzzy comprehensive evaluation for online courses, in *Multimedia Technology (ICMT), 2011 International Conference on*, pp. 842–844 (2011)
- [菅沼 05] 菅沼明, 峯恒憲, 正代隆義: 学生の理解度と問題の難易度を動的に評価する練習問題自動生成システム (学習支援), *情報処理学会論文誌*, Vol. 46, No. 7, pp. 1810–1818 (2005)
- [東京書籍 (株) 12] 数学編集部: 数学科 新学習指導要領 指導内容 中高関連表 (2012)
- [難波 11] 難波道弘: 3 値出力 CNN を用いた理解度診断システムの評価, *日本教育工学会論文誌*, Vol. 35, pp. 133–136 (2011)
- [野上 05] 野上俊一, 生田淳一, 丸野俊一: テスト勉強の学習計画と実際の学習活動とのズレに対する認識, *日本教育工学会論文誌*, Vol. 28, pp. 173–176 (2005)