

# 学習者のパフォーマンスを高精度に予測する Deep-IRT

## Deep Item Response Theory for high accuracy prediction of the student's performance.

堤 瑛美子<sup>\*1</sup>, 植野 真臣<sup>\*1</sup>  
 Emiko TSUTSUMI<sup>\*1</sup>, Maomi UENO<sup>\*1</sup>  
<sup>\*1</sup>電気通信大学情報理工学研究所  
<sup>\*1</sup>The University of Electro-Communications  
 Email: tsutsumi@ai.lab.uec.ac.jp

あらまし：近年、人工知能分野では教育ビッグデータに基づいて学習過程における学習者の能力値を推定し、未知の項目への反応予測を行う Knowledge Tracing が注目されている。本研究では二つの独立な学習者ネットワークと項目ネットワークによってパラメータの解釈性を高めた新しい Deep-IRT を提案する。提案手法は高い反応予測精度を示し、解釈可能な能力値・項目難易度パラメータをもつ。さらに、多次元のスキルに対する学習者の能力値を正確に推定することができる。

キーワード：適応的支援, 深層学習, 項目反応理論

### 1. はじめに

近年、オンライン教育の普及に伴い、大量の学習履歴データが容易に入手できるようになった。教育現場では学習者の発達を促すため、これらのデータに基づいて個々の学習者の理解度を把握し、適切な支援を与えることが課題となっている。人工知能分野では、機械学習手法を用いて過去の学習履歴から学習者の知識の習得状態を推定し、学習者の未知の項目への反応予測を行う Knowledge Tracing (KT) が注目を集めている。KT では、学習者の項目への反応予測を行うことにより、学習者が未習熟な知識やスキルを把握することができる。

### 2. 先行研究

KT で主に用いられる手法には確率的アプローチと深層学習アプローチがある。確率的アプローチには代表的な手法として Bayesian Knowledge Tracing (BKT) <sup>(1)</sup> と項目反応理論 (Item Response Theory; IRT) <sup>(2)</sup> がある。これらのモデルは課題の難易度や学習者の習熟度を表すパラメータをもち、学習者が未知の項目に正答する確率を求めることができる。しかし、BKT, IRT モデルは複数のスキル間の関係性を考慮した反応予測ができないという問題もある。実際の学習では、高難度の項目になるほど1つの項目に正答するために複数のスキルが必要であると考えられる。

近年では、学習者の多次元のスキルに対する習熟度の変化を考慮しながら反応予測を行うために、Deep Knowledge Tracing (DKT) <sup>(3)</sup>, Dynamic Key-Value Memory Network (DKVMN) <sup>(4)</sup>, Deep-IRT <sup>(5)</sup> などの深層学習アプローチが開発されている。DKT, DKVMN は学習者の多次元の習熟度を推定することで反応予測精度を向上させたものの、学習者の習熟度が隠れ変数行列に圧縮されており、項目の困難度パラメータや学習者の習熟度パラメータの解釈が難しいという問題がある。そこで、DKVMN のパラメータ解釈

性を向上させるために隠れ層を追加し、解釈可能なパラメータが得られるように設計された Deep-IRT が開発された。しかし、Deep-IRT で推定される習熟度は項目の特性に依存しており、同一スキル内の全ての項目が等質であると仮定しているために、異なる困難度をもつ項目からの習熟度推定値は同一尺度上で比較することが難しい。Deep-IRT においてもパラメータの解釈性には制約があり、IRT と同等の解釈性は得られなかった。

### 3. 提案手法

本研究ではパラメータの解釈性と高精度な反応予測を両立するために、Deep-IRT を改良し、習熟度の時系列変化を表現する学習者ネットワークと独立な項目ネットワークにより学習者の項目への反応を予測する新たな Deep-IRT を提案する。提案モデルは学習者の能力推定値が項目の特性に依存せず、複数のスキルに関する多次元の能力を表現することが可能である。さらに、既存の深層学習手法では同じスキ

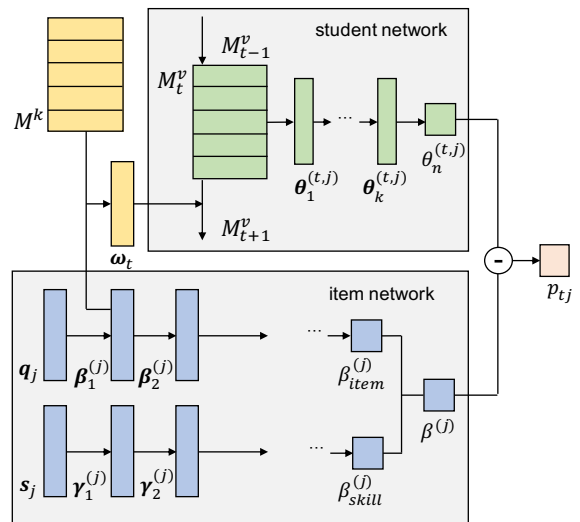


図1 提案手法の概要

ルを必要とする項目を全て同質とみなしているため、各項目の特性の違いが反映されておらず、反応予測を劣化させる原因となっている。そこで、提案モデルでは解答する項目とその項目が必要とするスキルの情報を入力とし、双方の特徴を考慮することで反応予測精度の向上を目指す。

提案モデルの概要図を図1に示す。提案モデルは学習者ネットワーク、項目ネットワークの2つの独立したニューラルネットワークから構成される。学習者ネットワークにはDKVMNと同様のメモリネットワーク構造を用いており、項目ネットワークでは解答する項目とその項目が必要とするスキルの双方を入力とし、項目の困難度を出力する。

## 4. 評価実験

### 4.1 実験方法

本章では、これまで Knowledge Tracing に用いられてきた代表的な手法(IRT, IDRT, DKT, DKVMN, Deep-IRT) と提案手法を用いて学習者の反応予測を行う。具体的には、10分割交差検証を用いて学習履歴データを訓練データ、検証データ、評価データに分割し、訓練データ、検証データから推定したパラメータを利用して評価データの反応予測を行う。予測精度の評価指標として Accuracy(一致割合), AUC スコア, F 値を算出する。本実験では、オンライン学習システムで収集された公開データセット<sup>(6)</sup> ASSIST2009, ASSIST2015, Statics2011, KDDcup を用いる。

### 4.2 実験結果

実験結果を表1に示す。結果より、提案手法は

AUC スコア, F 値では他の比較手法より高い予測精度を示すことがわかった。Acc においては提案手法が Statics2011 と KDDcup で最も高い精度を示し、ASSIST2009 では Deep-IRT, ASSIST2009 では DKT が最も高い精度を示した。これらの結果から、学習者と項目の独立なネットワークに加え、項目とスキルの特性を両方考慮した反応予測を行うことで高い反応予測が可能となることがわかった。

#### 参考文献

- (1) A.T. Corbett and J.R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge, User Modeling and User-Adapted Interaction, vol. 4, no. 4, pp. 253-278(1995)
- (2) M. Ueno and Y. Miyazawa, "IRT-based adaptive hints to scaffold learning in programming.", IEEE Transactions on Learning Technologies, Vol.14, No.8(2017)
- (3) C. Piech, et al, "Deep knowledge tracing." In C. Cortes, et al, editors, Advances in Neural Information Processing Systems 28, pages 505-513. Curran Associates, Inc., (2015)
- (4) J. Zhang, X. Shi, I. King, and D.-Y. Yeung, "Dynamic Key-Value Memory Networks for Knowledge Tracing.", Proceedings of the 26th International Conference on World Wide Web, pp.765-774, (2017)
- (5) C. Yeung, "Deep-irt: Make deep learning based knowledge tracing explainable using item response theory." Proceedings of the 12th International Conference on Educational Data Mining, EDM, (2019)
- (6) E. Tsustsumi, R. Kinoshita, M. Ueno, "Deep-IRT with independent student and item networks." Proceedings of the 14th International Conference on Educational Data Mining, EDM, (2021)

表1 実験結果

学習データ	指標	DKT	DKVMN	Deep-IRT	提案手法
	Acc	0.759	0.763	<b>0.768</b>	0.765
ASSIST2009	AUC	0.781	0.807	0.806	<b>0.810</b>
	F1	0.697	0.714	0.718	<b>0.722</b>
	Acc	<b>0.754</b>	0.732	0.727	0.752
ASSIST2015	AUC	0.730	0.749	0.747	<b>0.751</b>
	F1	0.433	0.541	0.54	<b>0.543</b>
	Acc	0.777	0.780	0.787	<b>0.822</b>
Statics2011	AUC	0.652	0.721	0.722	<b>0.821</b>
	F1	0.461	0.521	0.526	<b>0.690</b>
	Acc	0.784	0.773	0.792	<b>0.802</b>
KDDcup	AUC	0.538	0.594	0.588	<b>0.601</b>
	F1	0.439	0.439	0.455	<b>0.478</b>