



Knowledge Tracing の「単語学習」への応用

熊本県立宇土高校 2年

0.要約

MOOCをはじめとした e-Learning のコンテンツの充実に伴い、アダプティブラーニング (AL) と呼ばれる一人ひとりに最適化された学習が注目されるようになった。AL において、Knowledge tracing という機械学習の手法はほぼ必須である。Knowledge tracing とは、大量の学習ログをもとに、生徒が未知の問題に対して正解できるかを予測するアルゴリズムである。この手法は一般的に数学の「問題」などの学習ログで使用される。今回は、この knowledge tracing の手法を「単語学習」に応用したい。特に第二言語 (英語) の単語学習においては、ほとんどの学習者がかなりの時間を割いて行っており、この学習の効率化は非常にメリットが大きいと考えられる。また、「単語学習」に応用することで、今までは用いた表面的なアプローチだけでなく、自然言語処理的な手法や、言語学的手法などのアプローチが可能になり、論理的で高精度な予測ができる。

1.解決課題・現状

課題

学校教育では、1人の教員による均一な学習提供になっており、学習効果・効率が低いことが指摘されている。

アダプティブラーニングと呼ばれる一人ひとりに最適化された学習サービスの提供。

ALの実践例

研究例として、リクルートグループの「スタディサプリ」と東京大学松尾研究室との共同研究[8]がある。

大量の学習ログ(授業視聴履歴、テストの正誤データなど)を用いてディープラーニング(DL)を行う。

学習者に対して「解けない問題」を予測し、学習のつまずきの予防や個人に最適な学習パス(どの順番で学習を進めるか)を提案する。

参考[8]

現状

MOOCをはじめとしたオンライン教育コンテンツの充実に伴い、その学習ログをもとにディープラーニングを行い、ALを行う研究が進んでいる。

スラディサプリ
Udemy

コロナによる学校休校によってMOOCの利用が高まるとともに、教育ログのデジタル化が進んでいる。

2.機械学習について

Deep Learning とは

【図1 人工ニューロン】
x: 入力値, w: 重み, y: 出力値

【図2 ディープラーニング構造】
x: 入力値, w: 重み, y: 出力値

人工ニューロンは「入力→計算→出力」の順で処理を行う。X1, X2, X3を入力すると、それぞれに対して weight をかけ、それらを足し合わせた値を出力する。出力値にラベルの値と出力値の誤差を基に weight の値を調節する。

Recurrent Neural Network (RNN)

$x^{<t>}$ はある時刻 t における Input である。
1番目の式から、 $a^{<t>}$ にはそれ以前のデータの $a^{<t-1>}$ と、その時点での Input の値である $x^{<t>}$ の両方が考慮されている。(参考[6])

$x^{<t>}$ の Input だけでなく、それ以前のデータも考慮することによって、動的なデータの推移を推測することができる

$$a^{<t>} = g_1(w_{aa} a^{<t-1>} + w_{ax} x^{<t>} + a_0)$$

$$y^{<t>} = g_2(w_{ya} a^{<t>} + b_y)$$

※(w → 重み, g → tanh, p → sigmoid)

Knowledge Tracing(KT)

APにおいて「Knowledge Tracing (KT)」という知識状態を追跡する技術は必須である。

KTは、学習ログをもとにその生徒が未知の問題に対して正解できるかどうかを予測する。KTで伝統的な手法が、「Bayesian Knowledge Tracing(BKT)」[9]である。BKTはRNNを主としている。

生徒の学習機会におけるスキル状態(k_i)を計算し、その問題に正解できるか(y_t)を予測する。

【図1 knowledge tracing[10]】

$$P(k_i = true | t) = P(k_i = true) + \frac{P(k_i = true) - P(k_i = false)}{P(k_i = true) + P(k_i = false)}$$

$$P(k_i = true | t+1) = P(k_i = true | t) + \frac{P(k_i = true) - P(k_i = false)}{P(k_i = true) + P(k_i = false)}$$

$$P(k_i = true | t+1) = P(k_i = true | t) + \frac{P(k_i = true) - P(k_i = false)}{P(k_i = true) + P(k_i = false)}$$

$$P(k_i = true | t+1) = P(k_i = true | t) + \frac{P(k_i = true) - P(k_i = false)}{P(k_i = true) + P(k_i = false)}$$

3.本研究の目的・方法

本研究の目的

テーマ
knowledge tracing の単語学習への応用は可能か?

展望

- 単語学習において、学習者が理解していると予想される単語をスキップすることで効率的な学習につながる。
- 学習者に最適な単語を学ぶ順番の提案にもつながる。

本研究の目的

- knowledge tracingのためのword embeddingの手法考案。
- 機械学習モデルの開発。
- モデルのテスト精度から、実用性の検証。

比較

一般的なknowledge tracingと今回応用したい「単語学習」の比較。

	一般的なknowledge tracing	単語学習のknowledge tracing
適用目的(応用先)・アプローチ	主に数学分野で、基礎的な学習への応用。 数値や文字列の学習が多い。 数値や文字列の学習が多い。 数値や文字列の学習が多い。	単語学習の分野での応用。 単語の意味や文脈を学習する。 単語の意味や文脈を学習する。 単語の意味や文脈を学習する。
手法	学習ログをもとに、生徒の理解度を推測する。 学習ログをもとに、生徒の理解度を推測する。 学習ログをもとに、生徒の理解度を推測する。	学習ログをもとに、生徒の理解度を推測する。 学習ログをもとに、生徒の理解度を推測する。 学習ログをもとに、生徒の理解度を推測する。
期待する効果	学習者の理解度を推測し、適切な学習パスを提案する。 学習者の理解度を推測し、適切な学習パスを提案する。 学習者の理解度を推測し、適切な学習パスを提案する。	学習者の理解度を推測し、適切な学習パスを提案する。 学習者の理解度を推測し、適切な学習パスを提案する。 学習者の理解度を推測し、適切な学習パスを提案する。
学習データ	学習ログ(問題の正誤データ)。 学習ログ(問題の正誤データ)。 学習ログ(問題の正誤データ)。	学習ログ(問題の正誤データ)。 学習ログ(問題の正誤データ)。 学習ログ(問題の正誤データ)。

双方の最終的な活用方法や、そのための手法がほぼ同じであることから、「単語学習」におけるknowledge tracingも実現可能であることが示唆される。

表面的なアプローチだけでなく自然言語処理的なアプローチや語学学習的なアプローチが可能になる

単語に関しては様々なバリエーションがある。

仮説・手法

知っているor推測できる単語はスキップする。

$P(W)$ = 知らない単語だが、一度で理解できると予想される確率

$p(w)$ が最も高いものをサジェストして勉強を行う。

$P(W)$ を以下のそれぞれの手法を用いて単語の関係を定義し、どれが最も効果的かを調べる。

- Word embedding(多次元的な要素の取得)
- クラスタリング(ラベリングなど)
- テストの正誤データ(正答率など)
- Spelling(単語のスペル的な類似性)

Word Embeddingとは

	Man	Woman	King	Queen
Gender	-1.00	1.00	-0.95	0.97
Royal	0.01	0.02	0.93	0.95
Age	0.03	0.02	0.70	0.69
Food	0.09	0.01	0.02	0.01

【図4 Word Embedding[6]】

ManとWoman → Genderについて関係
KingとQueen → Gender & Royalという関係

Word Embeddingとは自然言語処理の分野で用いられる特徴抽出手法で、RNNを用いて単語やフレーズを多次元ベクトルで表現する。

多次元ベクトルで表現することによって、論理的に単語と単語の関係性を抽出できる。

Word Embeddingでは複数の視点でより論理的に2つの単語に関係性があるのかどうか分かる。

参考文献

- [1] Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean: Efficient Estimation of Word Representations in Vector Space.
- [2] Tomas Mikolov, Weng-tau Yih, Geoffrey Zweig: Linguistic Regularities in Continuous SpaceWord Representations.
- [3] Tolga Bulutcu, Kai-Chen, James Zou, Venkatesh Salganik, Aden Kalai: Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embedding.
- [4] Yang Yang, Jian Shen, Yanyu Gu, Yunfei Liu, Keneng Wang, Yaoming GKT: A Graph-based Interaction Model for Knowledge Tracing.
- [5] Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas Guibas, John Duchi: Knowledge Tracing.
- [6] Andrew Ng, Cozart: Deep Learning specialization: Machine learning.
- [7] Lijiang Xu, Mark A. Daventz: Dynamic Knowledge Embedding and Tracing.
- [8] 藤田 啓泰: こまごま! アプラーンによる個人学習: 1年度データとアダプティブラーニング-学習のつまずきをなくし、学習プロセスを最適化する-
- [9] Corbett, A.T., Anderson, J.R.: Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge.
- [10] 藤田 啓泰, 大村 真一: "Knowledge Tracing(KT)の組み合わせによる大量試験結果データからの学生モデルの構築"