

Dynamic Questioning: 強化学習を用いた生徒の学習意欲維持と学習の効率化を両立する出題アルゴリズム

熊本県立宇土高等学校 3年 吉野泰生

研究論文URL: https://drive.google.com/file/d/1rplQpR0d_VS2jx62fVdVQRQtS-7jvR5u/view?usp=sharing

1. 導入

近年、教育分野ではeラーニングサービスや大規模公開オンライン講座(MOOCs)が普及し、大量の学習ログを収集・蓄積することが容易になった。それに伴い機械学習(AI)を用いて生徒一人一人に最適化された学習提供を行う技術が注目されている。

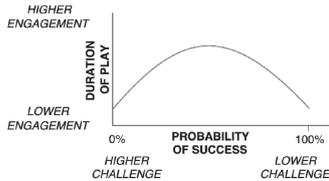


利点

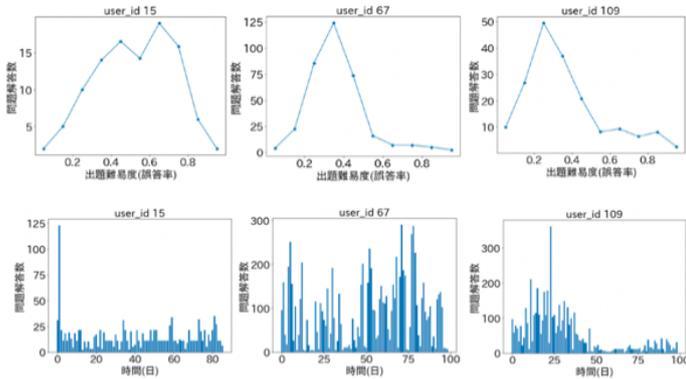
BKT [1] 等の Knowledge Tracing を基礎とする従来の出題手法では、各生徒の知識状態を推定し、その生徒が苦手とする問題を優先して出題することで学習効率を高める。

欠点

学習効率を上げるために苦手な問題(誤答予測の確率が高い問題)ばかりを出題すると、生徒の学習意欲が低下する。(左図: Inverted U-Hypothesis [2] より引用)

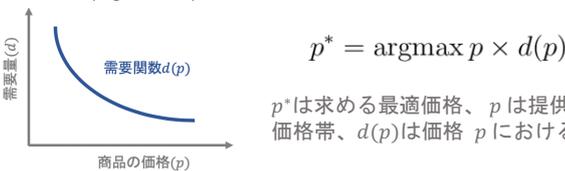


我々は、教育データセットEdNet [3]を用いて予備実験を行ったところ、問題解答数(1日に解答した問題の数)と出題難易度(1日の問題解答数のうち、誤答した問題数の割合)との間に逆U字の関係性があること、また学習意欲(問題解答数)は動的に変化することが確認された。学習効率を高めつつ、学習意欲を維持するためには、各生徒のその時のエンゲージメント状態に対応した最適出題難易度の出題を行うことが必要である。本研究では、Eコマース等で用いられる Dynamic Pricing という問題にアナロジーを見出し、この問題を解決する方法を提案する。



2. 関連研究 (Dynamic Pricing)

Dynamic Pricing [4] は、Eコマース環境における商品の購入数をもとにその時の需要量を推定し、最適な価格を動的に設定することで収益の最大化を図る手法である。典型的な Dynamic Pricing のアルゴリズムは、以下の最適化問題に帰着し、多腕バンディット問題 (Algorithm 1) として解くことで求める。



Algorithm 1: Active Algorithm for Dynamic Pricing

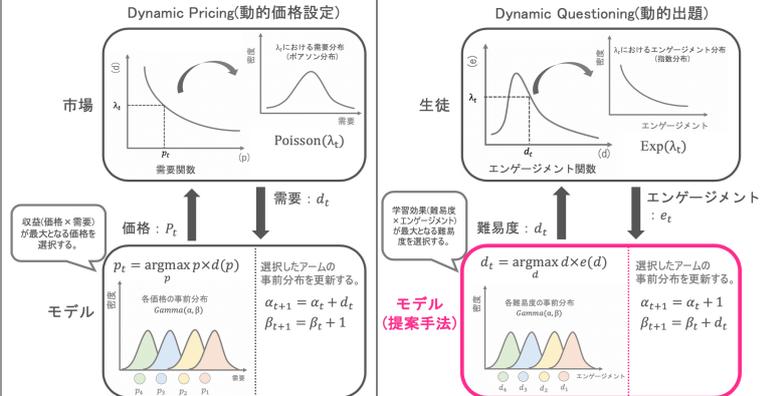
- α, β の初期値に対して事前分布を仮定する。 $p(\theta) = \text{Gamma}(\alpha, \beta)$
- For $t = 1$ to T :
 - $d \sim \text{Poisson}(\theta)$ について期待値 $E[d]$ を取得する。
 - 最適価格を見つける。:
 $p^* = \text{argmax}_p p \times d(p)$
 - 最適価格 p^* を提供し、実際の需要量 d_t を観測する。
 - Bayes 推定に基づきパラメータを更新する:
 $\alpha \leftarrow \alpha + d_t$
 $\beta \leftarrow \beta + 1$

3. 提案手法

本研究では、1日に解答した問題数を「問題解答数」、1日の問題解答数のうち誤答した問題数の割合を「出題難易度」と定義し、生徒のエンゲージメント分布を知らない状態で「学習効果(問題解答数×出題難易度)」が最大となる最適出題難易度を探す。この問題と Dynamic Pricing の間にあるアナロジーに着目した手法 Dynamic Questioning を提案する。生徒の学習意欲は「需要」、問題の出題難易度は「価格」と対応すると仮定すると、Dynamic Questioning における最適化関数を以下のように定義し、多腕バンディット問題として解くことが出来る。

$$d^* = \text{argmax}_d d \times e(d)$$

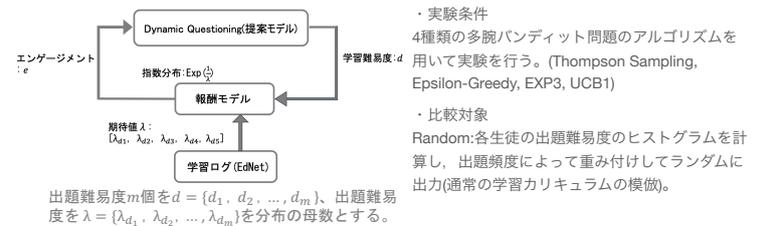
d^* は求める最適出題難易度、 d は提供された出題難易度、 $e(d)$ は難易度 d におけるエンゲージメント



Dynamic Pricing と Dynamic Questioning との対応関係 (Thompson Sampling)

4. 実験

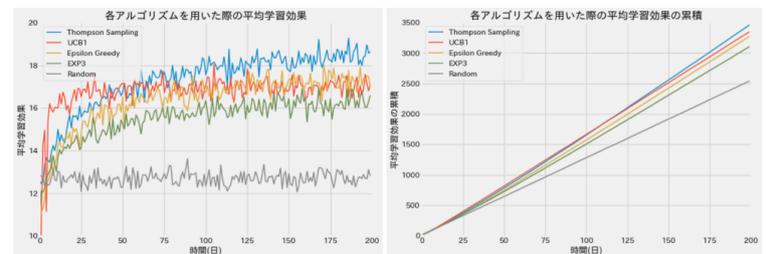
EdNet[3]から生徒の出題難易度と問題解答数の関係を抽出し、報酬の確率を表す報酬モデルを作成し、オフライン評価によって提案モデルを評価する。



- ・実験条件
4種類の多腕バンディット問題のアルゴリズムを用いて実験を行う。(Thompson Sampling, Epsilon-Greedy, EXP3, UCB1)

- ・比較対象
Random: 各生徒の出題難易度のヒストグラムを計算し、出題頻度によって重み付けしてランダムに出力(通常の学習カリキュラムの模倣)。

5. 結果と考察



Algorithm	平均学習効果		平均問題解答数		平均出題難易度	
	測定値	相対値	測定値	相対値	測定値	相対値
Random	12.81	1.00	39.15	1.00	0.34	1.00
Thompson Sampling	18.65	1.46	48.89	1.25	0.41	1.20
UCB1	17.10	1.34	44.82	1.14	0.41	1.19
Epsilon-Greedy	17.01	1.33	45.02	1.15	0.40	1.17
EXP3	16.60	1.30	44.08	1.13	0.39	1.15

多腕バンディット問題のいずれのアルゴリズムを用いても平均学習効果が大きく向上し、特にThompson Samplingでは、Randomと比較して1.25倍の問題解答数の増加と1.46倍の学習効果の向上が確認された。

6. 結論・展望

本研究では、オフライン評価において提案手法 Dynamic Questioning が生徒の学習意欲維持と学習の効率化を両立することが確認できた。展望として、実際の生徒へ出題することによる実証実験を行いたいと考えている。提案手法は、生徒のモチベーションを維持させることからeラーニングサービスを提供する企業の視点ではサブスクリプションの長期継続や広告収入増加などの便益をもたらすと予想される。

7. 謝辞・参考文献

本研究は、JSTグローバルサイエンス事業において東京大学グローバルサイエンスキャンパスの一環として行い、東京大学院工学系研究科の松尾研究室の岩澤有祐講師と三澤凜大様に指導いただきました。感謝申し上げます。

- [1] Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. User modeling and user-adapted interaction, 4 (4), 253-278.
- [2] Lomas, D., Patel, K., Forlizzi, J. L., & Koedinger, K. R. (2013, April). Optimizing challenge in an educational game using large-scale design experiments. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 89-90).
- [3] EdNet Dataset: <https://github.com/rlic/ednet>.
- [4] Ferreira, K. J., Simchi-Levi, D., & Wang, H. (2018). Online network revenue management using thompson sampling. Operations research, 66(6), 1586-1602.