Convolutional Knowledge Tracing: Modeling Individualization in Student Learning Process

Shuanghong Shen, Qi Liu, Enhong Chen, Han Wu, Zhenya Huang, Weihao Zhao, Yu Su, Haiping Ma, Shijin Wang
SIGIR 2020
Outline

1. Background
2. Problems and challenges
3. Model architecture
4. Experiments
5. Conclusions
Benifiting from the technological progress and large-scale application of online learning systems, such as Coursera and MOOCs, the vast majority of students have received or are receiving computer-assisted learning.
Knowledge tracing (KT): an emerging research area under this context that aims to assess students' changing knowledge state during learning process.

- teaching students in accordance with their aptitude
- maximizing the learning efficiency of students

Traditional methods

- BKT: a special case of Hidden Markov Model (HMM)
- DKT: introduces deep learning into KT, utilizing RNNs or LSTMs to model students’ knowledge state
- DKVMN: introduces external memory module to store the knowledge and update the corresponding knowledge state
Problems and challenges

- Individualization problem
  - Students have different prior knowledge
  - The learning rates differ from student to student

Without related information given in advance, it is very challenging to measure the individualization of students. We proposed CKT to model the individualization of students with only their learning interactions.
Model architecture

Problem Definition

Given the learning sequence $X_N = (x_1, x_2, ..., x_t, ..., x_N)$ with $N$ learning interactions of a student, KT aims to assess the student’s knowledge state after each learning interaction. In the learning sequence, $x_t$ is an ordering pair $\{e_t, a_t\}$, which stands for a learning interaction. Here $e_t$ represents the exercise being answered at learning interaction $t$ and $a_t \in (0, 1)$ indicates whether the exercise $e_t$ has been answered correctly (1 stands for correct and 0 represents wrong).
Model architecture

- **Embedding**
  
  \[
  x_t = \begin{cases} 
  [e_t \oplus a_t], & \text{if } a_t = 1, \\
  [a_t \oplus e_t], & \text{if } a_t = 0,
  \end{cases}
  \]

- **Individualized Prior Knowledge**
  
  ✓ **Historical Relevant Performance**: measuring student historical performance relevant to the exercise to be answered
  
  \[
  \begin{align*}
  r_t(i) &= \text{Masking}(e_i \cdot e_t), & i \in (t, N), \\
  w_t(i) &= \text{Softmax}(r_t(i)), & i \in (1, N),
  \end{align*}
  \]

  \[
  \text{HRP}_t(t) = \sum_{i=1}^{t-1} w_t(i)x_i.
  \]

  ✓ **Concept-wised Percent Correct**: accounting for the overall knowledge mastery of student on all knowledge concepts

  \[
  \text{CPC}_t(m) = \frac{\sum_{i=0}^{t-1} a_i^m}{\text{count}(e^m)},
  \]

- **Individualized Learning rate**
  
  design hierarchical convolutional layers to extract the learning rate features by processing several continuous learning interactions simultaneously within a sliding window

- **Objective Function**

  \[
  L = -\sum_{t=1}^{N} (a_t \log p_t + (1 - a_t) \log(1 - p_t)).
  \]
Model architecture

Input Module

Prior Knowledge

GLU

Z: knowledge state

Predictions

Output Module

Modeling Learning Rates

\( H \)

\( H \)
Experiments

• Datasets

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Students</td>
<td>Concepts</td>
</tr>
<tr>
<td>ASSIST2009</td>
<td>4,151</td>
<td>110</td>
</tr>
<tr>
<td>ASSIST2015</td>
<td>19,840</td>
<td>100</td>
</tr>
<tr>
<td>ASSISTchall</td>
<td>1,709</td>
<td>102</td>
</tr>
<tr>
<td>Statics2011</td>
<td>333</td>
<td>1,223</td>
</tr>
<tr>
<td>Synthetic-5</td>
<td>4,000</td>
<td>50</td>
</tr>
</tbody>
</table>

• Comparison methods

  • CKT-ONE with only one convolutional layer.
  • CKT-HRP measures prior knowledge only from HRP.
  • CKT-CPC measures prior knowledge only from CPC.
  • CKT-ILR only models individualized learning rate.
  • CKT-IPK only models individualized prior knowledge.
  • DKT leverages recurrent neural network to assess student knowledge state [12]. We utilized LSTM in our implementation.
  • DKVMN takes advantage of memory network to get interpretable student knowledge state.
Experiments

- Student performance prediction

![Graph showing AUC, ACC, and $r^2$ for different methods across various datasets.](image)

*Figure 3: Results of comparison methods on student performance prediction.*

- Visualization of knowledge tracing results

![Diagram showing a sequence of learning interactions and knowledge tracing results.](image)

*Figure 4: Visualization cases of individualized knowledge tracing result of student.*
Experiments

- Exercise embeddings learning

Figure 5: Visualization of exercise relevant coefficients
Conclusion

- We proposed a novel model called Convolutional Knowledge Tracing (CKT) to model individualization of students in KT task.
- We measured individualized prior knowledge from students’ historical learning interactions (i.e., HRP and CPC).
- We designed hierarchical convolutional layers to extract individualized learning rates based on continuous learning interactions.
- Extensive experiment results indicated that CKT could get better knowledge tracing results through modeling individualization in student learning process.

Code for CKT is available at: https://github.com/bigdata-ustc/Convolutional-Knowledge-Tracing
Thank you for listening

Q & A