A Latent Factor Model for Instructor Content Preference Analysis

Jack Wang¹, Andrew Lan², Phillip Grimaldi¹, Richard Baraniuk¹

1. Rice University; 2. Princeton University

1. Rice University; 2. Princeton University

Project Overview

- Focus on data that instructors exclude certain questions from students’ homework assignment in Biology classes they teach.
- Develop model to predict instructors’ preference (exclude or no) for questions.
- Incorporate Bloom’s Taxonomy to enhance model interpretability.

Background & Prior Work

- Bloom’s Taxonomy describes the cognitive processes by which learners encounter and work with knowledge.
- We hypothesize instructors have certain Bloom’s Taxonomy preference, and thus incorporate this information into our model.
- Prior work focuses primarily on interactions between students and learning resources to provide learners’ analytics, such as SPARFA, but overlooks interactions between instructors and learning resources.

Latent Factor Model

- We formulate the problem as a matrix factorization problem and propose the following probabilistic model:

  \[
  Y_{ij} \sim \text{Ber} \left( \phi \left( p_i^T a_j + g_i^T h_j \right) \right),
  \]

  where \( \phi(x) = \frac{1}{1 + e^{-x}} \).

- Y is binary (takes value 1 and 0) matrix of size N (number of instructors) x Q (number of total excluded questions).
- \( Y_{ij} = 1 \) means that the instructor i explicitly expresses exclusion preference for question j.
- \( a_j \) is the Bloom’s Taxonomy vector for question j. Each \( a_j \) is of dimension K (number of distinct Bloom’s tags), and satisfies:

  \[ a_{j k} \in (0, 1) \text{ and } \sum_k a_{j k} = 1 \forall j. \]

- \( p_i, g_i, h_i \) are model parameters to be estimated. \( p_i \) is the instructor – bloom’s tag association vector, \( g_i, h_i \) are of dimension M, chosen later via cross validation.
- Parameters estimation via block coordinate descent method.

Motivation

Mine instructor– learning resources interaction data to understand instructors’ preferences and how they utilize learning resources. Benefits include:
- Providing personalized educational content recommendation for instructors.
- Improving personalized learning plan recommendation for students.

Evaluation

- Show superior prediction performance of the proposed model by comparing it with three popular methods in recommender systems: user-based / item-based collaborative filtering (UBCF / IBCF), and Funk singular value decomposition (FSVD).

Table 1: Comparison between proposed model and collaborative filtering methods.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Full Model</th>
<th>UBCF</th>
<th>IBCF</th>
<th>FSVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>0.9013±0.0045</td>
<td>0.8961±0.0048</td>
<td>0.8905±0.0048</td>
<td>0.8896±0.0045</td>
</tr>
<tr>
<td>F-1</td>
<td>0.6483±0.0128</td>
<td>0.6007±0.0158</td>
<td>0.5960±0.0137</td>
<td>0.6155±0.0158</td>
</tr>
<tr>
<td>Precision</td>
<td>0.7136±0.0222</td>
<td>0.7037±0.0214</td>
<td>0.6928±0.0254</td>
<td>0.6964±0.0236</td>
</tr>
<tr>
<td>Recall</td>
<td>0.6153±0.0227</td>
<td>0.5226±0.0190</td>
<td>0.4854±0.0159</td>
<td>0.5601±0.0248</td>
</tr>
</tbody>
</table>

Table 2: Comparison between \( p_{ik} \) and the percentage of questions instructor i actually excluded under each bloom’s tag k.

<table>
<thead>
<tr>
<th>Bloom’s Taxonomy Tag</th>
<th>Instructor</th>
<th>( k = 1 )</th>
<th>( k = 2 )</th>
<th>( k = 3 )</th>
<th>( k = 4 )</th>
<th>( k = 5 )</th>
<th>( k = 6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>i = 3</td>
<td>0.95%</td>
<td>1.6%</td>
<td>0.5%</td>
<td>1.8%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>i = 5</td>
<td>16.9%</td>
<td>16.3%</td>
<td>16.9%</td>
<td>5.5%</td>
<td>21.1%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>i = 9</td>
<td>63.1%</td>
<td>63.8%</td>
<td>72.4%</td>
<td>67.3%</td>
<td>42.1%</td>
<td>33.3%</td>
<td>33.3%</td>
</tr>
</tbody>
</table>

- Demonstrate model interpretability by visualizing latent factor learnt with table and multidimensional scaling (MDS).

Figure 2: 2D projection of instructor Bloom’s tag preference vector using MDS.

References and Acknowledgements

- Special thanks to Phillip Grimaldi for collecting data, Mr. Lan and Prof. Baraniuk for providing feedback and guidance.

Future Work

- Incorporate factors (depth of knowledge, questions’ context in textbook) other than Bloom’s Taxonomy to explain instructors’ question exclusion preferences.
- Understand each instructor’s objectives from interaction data.

For questions, please contact: Zichao.Wang@rice.edu