

Recommending Learning Materials to Students by Identifying their Knowledge Gaps

Konstantin Bauman
Stern School of Business
New York University
kbauman@stern.nyu.edu

Alexander Tuzhilin
Stern School of Business
New York University
atuzhili@stern.nyu.edu

ABSTRACT

We propose a new content-based method of providing recommendations of educational materials to the students by identifying gaps in their knowledge of the subject matter in the courses they take. We experimentally validate our method by conducting an A/B test on the students from an online university.

Keywords

content-based recommendations; technology-enhanced learning; knowledge gaps

1. INTRODUCTION

Due to the recently increased interest in online educational technologies and educational delivery methods, the topic of recommendations in the educational domain has become increasingly important lately. In particular, it has been studied in various communities, including RecSys, UMAP, Advanced Learning Technologies, and the Technology-Enhanced Learning communities, and many approaches have been proposed on how to recommend learning materials to the students to improve their learning performance [2].

One of such recommendation methods is based on the idea of identifying and filling the “gaps” in students’ knowledge in the subjects that they are studying. The idea of gap identification is not new, however. For example, Ciuciu and Demey referred to it in [1] and proposed an initial approach on how to deal with it. Unfortunately, they stopped short of describing the specific recommendation algorithm, leaving it as a topic of future research. Also, [3, 4, 5] proposed methods that are somewhat related to the “gap filling” idea, but the authors mainly focused on developing their frameworks and not on presenting specific recommendation algorithms.

In this paper, we present a novel method of identifying gaps in students’ knowledge and propose specific algorithms to fill-in these gaps by providing recommendations of remedial learning materials to the students. In contrast to many prior learning recommendation methods that are predominantly rating-based [2], our method, described in Section 2, is content-based. In addition to developing this method, we also performed A/B testing on the students of a leading online university to validate our approach. We present our experiments and the preliminary results in Sections 3 and 4.

Copyright is held by the author/owner(s).
RecSys 2014 Poster Proceedings, October 6-10, 2014, Foster City, Silicon Valley, USA.

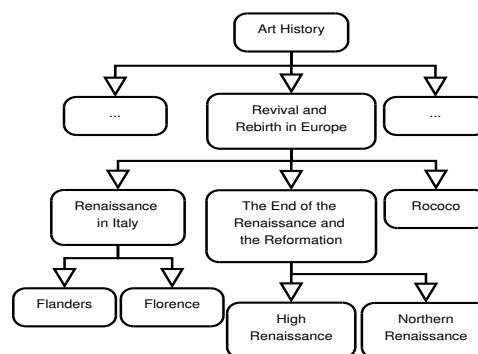


Figure 1: Part of Taxonomy for Art History Course

2. RECOMMENDATION METHOD

Our recommendation method is based on the “gap filling” idea discussed in Section 1. In particular, for each course in a curriculum, we build taxonomy of the topics covered in that course. For example, Fig.1 shows a part of the *Art History* course taxonomy where each node represents a topic. A node in the taxonomy has a set of obligatory reading materials chosen by the instructor and associated with this topic.

For each student and a course offering we determine how well the student understood all the topics specified in the course taxonomy by analyzing the student performance data in that course. At the end of this analysis, each student gets a certain performance score for each topic in the course taxonomy specifying how well the student understood a particular topic. For example, in course *Art History* for topic Rococo Joe got the score 0.94 while John got 0.67. This means that Joe understood Rococo well, while John did not. Although this score can be computed in many different ways, in our experiments described in Section 3 we have done it as follows. For each test performed by the student and each question on the test, we determine the list of topics in the course taxonomy to which this test question corresponds. Then for each topic we determine the list of questions corresponding to it and see how well the student answered these questions. For example, if there are 10 questions in the test corresponding to topic Rococo and Joe answered 9 of them correctly, then Joe’s score for this topic is 0.9.

After we determine students’ performance scores for each topic in the course taxonomy, we identify their *knowledge gaps*, i.e., identify those topics on which they performed poorly. In particular, a student has a knowledge gap for a topic if either (a) the performance score of a student for this topic is low (i.e., below a certain threshold level) or (b)

the student has knowledge gaps for a sufficient number of subtopics of that topic (and therefore needs remedial actions for these subtopics).

After we identify the knowledge gaps, we determine what types of *remedial materials* should be recommended to the students in order for them to close these gaps. We accomplish this task as follows. First, we build a library of related reading materials for each course consisting of (but not limited to) the most popular textbooks, online articles and various web pages related to the course. Each document in this library can have its own taxonomy that is based on the document's table of content. For example, a textbook is divided into chapters, sections and subsections. In contrast, some other documents, such as short articles, may not have any taxonomy and therefore are not "divisible" into smaller pieces. Also, we establish the relationship between the materials in this library and the course taxonomy as follows. For each node in the course taxonomy we identify the "unit of knowledge" in the library (e.g., book chapter) corresponding to it in the best way, thus establishing the link between the node and the reading material. In particular, we do this identification by using the TF-IDF-based measure of correspondence between the book unit and the textual description of the topic.

Given the structure of the course, the identified gaps in student knowledge in the class, and the links between the topics in the course taxonomy and the supplemental reading materials from the library that we described in the previous paragraph, we next provide recommendations of these supplementary reading materials to the students in order to close these knowledge gaps. In particular, for each knowledge gap topic node in the taxonomy, we recommend those supplementary reading materials linked to that node.

3. EXPERIMENTAL SETTINGS

To validate our approach, we tested it on students of an on-line university by conducting an A/B test. In particular, we worked with 527 students from all over the world taking one or more courses in that university over a period of one semester that lasted 9 weeks (8 weeks of studies and one week for the final exams). There were 25 different courses offered during that semester covering the areas of Computer Science (10 courses), Business (10 courses) and General Studies (5 courses). In total, we had 692 enrollments of all these students in the courses (i.e., 692 student/course pairs) during that semester. Studies during each week are carefully structured in that university and consist of (a) a set of obligatory reading materials, (b) various assignments, (c) questions to be discussed on the discussion forums and (d) a self-testing quiz (not contributing to the overall grade for the course). There are also two quizzes administered by the university during the semester that contribute to the final grade for the course. There is also the final exam given at the end of the semester during week 9.

In our experiments, we split the students into the following three groups. The first group received personalized recommendations as described in Section 2. The second group received the standard set of (non-personalized) recommendations where all the students got the same set of recommendations as the worst students in the personalized group who failed all their tests (and therefore needed help for all the topics in the course). The third group is the controlled group of students who did not receive any recommendations.

We provided recommendations to the first and the second groups up to three times. The first recommendation of the supplementary reading materials was provided shortly before they took graded Quiz 1. The second one was provided before students took graded Quiz 2, and the last one shortly before students took the final exam.

The goal of this experiment is to test two hypotheses: (1) recommendations (personalized and non-personalized) lead to better performance results, as measured by student's total score on the final exam; (2) personalized recommendations, as described in Section 2, lead to better performance results vis-à-vis providing non-personalized recommendations (as measured by the final exam score).

In addition, we also sent a survey to those students who have received at least one recommendation at the end of the semester in order to see how well they perceived our recommendations and also to detect possible biases and problems with the experimentation. In particular, we asked the students how much they liked our recommendations, i.e., what was their overall impression about the recommendations (vis-à-vis individual recommendations, as is normally done in recommender systems).

4. RESULTS

The results of the survey revealed that the vast majority of the students indeed liked our recommendations and found them to be very useful in their studies. However, when we measured the actual performance of the students on the final test (as opposed to how much they liked the recommendations), our *preliminary* results showed that our recommendations were not uniformly effective to all the students across all the courses. In particular, the recommendations worked the best for the mediocre students and were less effective for the excellent and good students. Also, they were most effective for the poorly performing students taking business courses where statistically significant performance differences on the final exam were detected in comparison to the control group. Further, we have also observed real performance differences on several other segments of students and types of courses. However, we could not demonstrate that these differences were statistically significant because of the sizes of our samples and the preliminary nature of our data and results. As a part of the future work, we plan to enhance our data and provide more extensive analysis on it to demonstrate that personalized recommendations indeed lead to better performance results.

5. REFERENCES

- [1] I. Ciuciu and Y. Demey. An evaluation methodology for c-foam applied to web-based learning. In *AWBL*. 2012.
- [2] N. Manouselis, H. Drachler, V. Katrien, and D. Erik. *Recommender Systems for Learning*. Springer, 2013.
- [3] A. Mavroudi and T. Hadzilacos. Broadening the use of e-learning standards for adaptive learning. In *Advances in Web-Based Learning*. Springer Berlin, 2012.
- [4] S. Saman, B. Seyed, Z. Nor, and N. Shahrul. Ontological approach in knowledge based recommender system to develop the quality of e-learning system. In *Australian J. of Basic and Applied Science*. 2012.
- [5] X. Zhou, J. Chen, and Q. Jin. Discovery of action patterns in task-oriented learning processes. In *Advances in Web-Based Learning*. Springer, 2013.