Chapter 6

Factorization Techniques for Predicting Student Performance

Nguyen Thai-Nghe
University of Hildesheim, Germany

Lucas Drumond
University of Hildesheim, Germany

Tomáš Horváth
University of Hildesheim, Germany

Artus Krohn-Grimberghe
University of Hildesheim, Germany

Alexandros Nanopoulos
University of Hildesheim, Germany

Lars Schmidt-Thieme
University of Hildesheim, Germany

ABSTRACT

Recommender systems are widely used in many areas, especially in e-commerce. Recently, they are also applied in e-learning for recommending learning objects (e.g. papers) to students. This chapter introduces state-of-the-art recommender system techniques which can be used not only for recommending objects like tasks/exercises to the students, but also for predicting student performance. We formulate the problem of predicting student performance as a recommender system problem and present matrix factorization methods, which are currently known as the most effective recommendation approaches, to implicitly take into account the prevailing latent factors (e.g. “slip” and “guess”) for predicting student performance. As a learner’s knowledge improves over time, too, we propose tensor factorization methods to take the temporal effect into account. Finally, some experimental results and discussions are provided to validate the proposed approach.

DOI: 10.4018/978-1-61350-489-5.ch006
INTRODUCTION

Recommender systems are widely used in many areas, especially in e-commerce (Rendle, Freudenthaler, and Schmidt-Thieme, 2010). One of their main aims is to make vast catalogs of products consumable by learning user preferences and to apply them to items formerly unknown to the user. Thus they can learn which products have a high likelihood of being interesting to the target user. Recently, recommender systems have also been applied to e-learning, especially in technology enhanced learning (Manouselis, Drachsler, Vuorikari, Hummel, and Koper, 2010).

On the other hand, educational data mining has also been taken into account recently to assist the students in the learning process. One of the main educational data mining tasks, for instance, is to predict student performance. It is applicable when we would like to know how the students learn (e.g. generally or narrowly), how quickly or slowly they adapt to new problems or if it is possible to infer the knowledge requirements to solve the problems directly from student performance data (Feng, Heffernan, and Koedinger, 2009). Generally speaking, the prediction of student performance is the problem of predicting the student’s ability (e.g., estimated by a score metric) in solving tasks when interacting with a tutoring system. Cen, Koedinger, and Junker (2006) have shown that an improved model for predicting student performance could save millions of hours of students’ time and effort in learning algebra that they could otherwise have spent on other subjects or leisure. Moreover, many universities are extremely focused on assessment, thus, the pressure on teaching and learning for examinations leads to a significant amount of time spent for preparing and taking standardized tests. Any move away from standardized and non-personalized tests holds promise for increasing deep learning (Feng et al., 2009). From an educational data mining point of view, a good model which accurately predicts student performance could replace some current standardized tests and yield truly personalized, adaptive test.

To address the student performance prediction problem, many works have been published. Most of them rely on traditional methods such as neural networks (Romero, Ventura, Espejo, and Hervs, 2008), Bayesian networks (Bekele and Menzel, 2005), logistic regression (Cen et al., 2006), support vector machines (Thai-Nghe, Busche, and Schmidt-Thieme, 2009) and so on.

In the recommender system context, predicting student performance can be considered as a rating prediction problem since student, task, and performance information could be treated as user, item, and rating, respectively, which are the main objects recommender systems learn from, nowadays. Recently, Thai-Nghe, Drumond, Krohn-Grimberghe, and Schmidt-Thieme (2010) and Toscher and Jahrer (2010) have proposed the use of recommendation techniques, especially matrix factorization, for predicting student performance. The authors have shown that using recommendation techniques could improve prediction results compared to regression methods (Thai-Nghe et al., 2010) but they have not taken the temporal effect into account. Obviously, in the educational point of view, we always expect that the students (or generally, the learners) can improve their knowledge over time, so temporal information is an important factor for such prediction tasks.

Furthermore, in predicting student performance, two crucial user-dependent aspects need to be taken into account:

1. The probability of a student to guess correctly while not knowing how to solve the problem at hand or not having the required skills related to the problem (which we call “guess” for short); and the probability of a student to fail while knowing how to solve the problem or having all of the required skills related to the problem (which we call “slip” for short);
Related Content

An Investigation into the Parameters of Quantum Degeneration of an Ultra Cold Non–Neute Plasma of Identical Ions of Zero Spin in a Paul Trap

Wavelet and Unsupervised Learning Techniques for Experimental Biomedical Data Processing
www.igi-global.com/article/wavelet-unsupervised-learning-techniques-experimental/78330?camid=4v1a

Software Tool for Assessment of Complexity and Variability in Physiological Signals of Respiration
www.igi-global.com/article/software-tool-assessment-complexity-variability/58069?camid=4v1a

Novel Adaptive Exon Predictor for DNA Analysis Using Singular Value Decomposition
www.igi-global.com/article/novel-adaptive-exon-predictor-for-dna-analysis-using-singular-value-decomposition/201049?camid=4v1a