SYNOPSIS OF THE PROPOSED RESEARCH WORK

ON

A Novel Approach for Recommender Systems in Education using Machine Learning

Submitted By
Name:S Chandra Sekhar

TABLE OF CONTENTS

ABSTRACT

Recommender systems are widely used in many areas, especially in e-commerce. However, there is limited information of the impact of recommender systems in other domains like education. Recently, they are also applied in e-learning tasks such as recommending resources (e.g. papers, books,..) to the learners (students). In this work, I propose a novel approach which uses recommender system techniques for educational data mining, to best utilize the time and predict the most important questions .Generally, students have to study the entire syllabus to perform very good in academic exams. Not many perform well in doing so, and end up getting less marks. With this kind of recommender system student can practise only those questions which are mostly certainly to be coming in the exam. Given the time limit, the system will detail all the important questions to be learned so that he will gain maximum marks under the given time constraint. This is achieved by ranking all the questions of a topic according to the probability of their occurrence in the exam. By this, within less amount of time the student can learn all the important questions and their by the overall student performance will be hugely increased.

Keywords: recommender systems, matrix factorization, ,educational datamining, probability of important questions.

1. INTRODUCTION

Systems that retrieve and filter the data through content and similar profiles are known as recommendation systems (RS). These systems are usually used within the e-commerce domain. For example, some websites, such as Amazon, through the application of RS allow offering the user recommendations for products that users do not know and could be of their interest. Suggested recommendations help to overcome the distressing search problem for the user. But this technology is not only used to sell products, but it is also used to suggest videos (YouTube), movies (Netflix), friends (Facebook), among others.This demand spans across several domains, among which is the educational domain. RS, which are applied in education, have the role of supporting teaching and learning activities through enhanced information retrieval. Nevertheless, there is limited information of the application of recommender systems in educational environments.

Educational data mining has also been taken into account recently. Since the universities desire to improve their educational quality, the usage of data mining in higher education to help the universities, instructors, and students to improve their performance has become more and more attractive to both university managers and researchers. For example, to help improving the student performance, some research questions should be explored such as how the students learn (e.g. generally or narrowly)? How quickly or slowly the students adapt with the new problems? This work proposes a novel approach which applies recommender system techniques for improving student performance.

2.LITERATURE SURVEY

As surveyed in Manouselis et al. [3], many recommender systems have been deployed in technology enhanced learning. Concretely, Garc´ıa et al. [4] uses association rule mining to discover interesting information through student performance data in the form of IF-THEN rules, then generating the recommendations based on those rules; Bobadilla et al. [9] proposed an equation for collaborative filtering which incorporated the test score from the learners into the item prediction function; Ge et al. [10] combined the content-based filtering and collaborative filtering to personalize the recommendations for a courseware selection module; Soonthornphisaj et al. [11] applied collaborative filtering to predict the most suitable documents for the learners; while Khribi et al. [12] employed web mining techniques with content-based and collaborative filtering to compute the relevant links for recommending to the learners.

In predicting student performance, Romero et al. [8] compared different data mining methods and techniques to classify students based on their Moodle usage data and the final marks obtained in their respective courses; Bekele and Menzel [6] used Bayesian Networks to predict student results; Cen et al. [15] proposed a method for improving a cognitive model, which is a set of rules/skills encoded in intelligent tutors to model how students solve problems, using logistic regression; Thai-Nghe et al. [7] analyzed and compared some classification methods (e.g. decision trees and Bayesian networks) for predicting academic performance; while Thai-Nghe et al. [6] proposed to improve the student performance prediction by dealing with the class imbalance problem. (i.e., the ratio between passing and failing students is usually skewed).

Different from the literature, instead of using traditional classification or regression methods, we propose using state of the art techniques in recommender systems (e.g. matrix factorization) for predicting student performance.

3. OBJECTIVE OF THE PROPOSED RESEARCH WORK

1. The main objective of the work is to guide the student with all the important questions he/she has to learn in a specified amount of time thus improve the student performance.
2. To view how many peer members are currently reading which topic and which question.

4. STUDY AREA AND METHODOLOGY

let S be a set of student IDs, T be a set of task IDs and, $f \subseteq R$ be a performance measure, then $D \subseteq (S \times T \times f)$ is the triple data collected from the study application which the student uses. f indicates the probability of each question under a topic . Given $s \in S$ and $t \in T$, our problem is to predict f. Obviously, in a recommender system context, s,t, and f would be user, item, and rating/probabilty, respectively. The recommender system task at hand is thus rating/probability prediction.

Proposed Technique:

Traditionally, recommender systems focus on reducing the information overload and act as information filters. The most famous recommender system and indeed one of the first commercial recommender system at all is the Amazon's "Customers Who Bought This Item Also Bought". The aim of recommender system is making vast catalogs of products consumable by learning user preferences and applying them to items formerly unknown to the user, thus being able to recommend what has a high likelihood of being interesting to the target user.

The two most common tasks in recommender systems are Top-N item recommendation where the recommender suggests a ranked list of (at most) N items $i \in I$ to a user $u \in U$ and rating prediction where the aim is predicting the preference score (rating) $r \in R$ for a given user-item combination. For item recommendation the training data is currently usually unary information on items being viewed, clicked, purchased etc. by the respective users. Rating prediction mainly uses rating information itself as training data.

Collaborative Filtering :

In the early days of recommender systems, content was deemed very valuable training data and research data sets contained lots of attribute information for algorithm training. But since the late nineties the so called collaborative filtering approach prevails. Collaborative filtering is based on the assumption that similar users like similar things and, being content-agnostic, focuses only on the past ratings assigned. In this work, we make use of matrix factorization (Rendle and Schmidt-Thieme [13], Koren et al. [2]), which is known to be one of the most successful methods for rating prediction, outperforming other state-of-the-art methods (Bell and Koren [14]).

Matrix Factorization :

Matrix factorization is the task of approximating a matrix X by the product of two smaller matrices W and H, i.e. $X \approx WHT$ (Koren et al. [2]). In the context of recommender systems the matrix X is the partially observed ratings matrix, $W \in RU\times K$ is a matrix where each row u is a vector containing the K latent features describing the user u and $H \in RI\times K$ is a matrix where each row i is a vector containing the K features describing the item i. Let wuk and hik be the elements of W and H, respectively, then the rating given by a user u to an item i is predicted by:

RMSE = $\sqrt{\sum ui(rui - r\hat{}ui)^2 / n}$. where n is the number of test cases.

Data Set: There is no existing data set that we can use to analyse the results.we have to build our own data set once the students use the application which we develop.

Mapping Educational Data to Recommender Systems

In traditional recommender systems settings, it is unambiguous how the available information is mapped to users, items, and ratings, respectively. At least for all major recommender system data sets used (Jester, Movielens 100k, and Netflix) there is a unique assignment3.

In our Model there is There is an obvious mapping of users and probability of question being asked in the exam.
student ⇒ prepares the questions which have higher probability=>Probability of question
The student becomes the user, and the probability of a question would become the probability of the question , bounded between 0 and 1.
For mapping the item, several options seemed to be available . One possible option is an item was supposed to be the combination (concatenation) of problem hierarchy (PH), problem name (PN), step name (SN), and problem view (PV),Available time(AT),Frequency of the problem(FP).

5. EXPECTED OUTCOME

The proposed Matrix Factorization model will classify the questions and predict all the important questions to be studied by the student in order to secure more marks in the exam.

REFERENCES

[1]Nguyen Thai-Nghe∗, Lucas Drumond∗, Artus Krohn-Grimberghe∗, Lars Schmidt-Thieme,Recommender System for Predicting Student Performance ,Procedia Computer Science.

[2] Y. Koren, R. Bell, C. Volinsky, Matrix Factorization Techniques for Recommender Systems, IEEE Computer Society Press 42 (8) (2009) 30–37, ISSN 0018-9162.

[3] N. Manouselis, H. Drachsler, R. Vuorikari, H. Hummel, R. Koper, Recommender Systems in Technology Enhanced Learning, in: Kantor, P.B., Ricci, F., Rokach, L., Shapira, B. (eds.) 1st Recommender Systems Handbook, Springer-Berlin, 1–29, 2010.

[4] E. Garc ́ıa, C. Romero, S. Ventura, C. D. Castro, An architecture for making recommendations to courseware authors using association rule mining and collaborative filtering, User Modeling and User-Adapted Interaction 19 (1-2).

[5] H. Cen, K. Koedinger, B. Junker, Learning Factors Analysis A General Method for Cognitive Model Evaluation and Improvement, in: Intelligent Tutoring Systems, vol. 4053, Springer Berlin Heidelberg, ISBN 978-3-540-35159-7, 164–175, 2006.

[6] R. Bekele, W. Menzel, A Bayesian Approach to Predict Performance of a Student (BAPPS): A Case with Ethiopian Students, in: Artificial Intelligence and Applications, Vienna, Austria, 189–194, 2005.

[6]N.Thai-Nghe,A.Busche,L.SchmidThieme ImprovingAcademicPerformancePredictionbyDealingwithClassImbalance ,in:Proceed- ing of 9th IEEE International Conference on Intelligent Systems Design

and Applications (ISDA'09), Pisa, Italy, IEEE Computer Society, 878–883, 2009.
[7] N.ThaiNghe,P.Janecek,P.Haddawy,AComparativeAnalysisofTechniquesforPredictingAcademicPerformance,in:Proceedingof37th IEEE Frontiers in Education Conference (FIE'07), Milwaukee, USA, IEEE Xplore, T2G7–T2G12, 2007.
[8] C.Romero,S.Ventura,P.G.Espejo,C.Hervs,DataMiningAlgorithmstoClassifyStudents,in:1stInternationalConferenceonEducational Data Mining (EDM'08), Montral, Canada, 8–17, 2008.
[9] J. Bobadilla, F. Serradilla, A. Hernando, Collaborative filtering adapted to recommender systems of e-learning, Knowledge-Based Systems 22 (4) (2009) 261–265, ISSN 0950-7051. *T.-N. Nguyen et al. / Procedia Computer Science 1 (2010) 2811–2819* 2819 *N. Thai-Nghe, L. Drumond, A. Krohn-Grimberghe, and L. Schmidt-Thieme / Procedia Computer Science 01 (2010) 1–9* 9
[10] L. Ge, W. Kong, J. Luo, Courseware Recommendation in E-Learning System, in: International Conference on Web-based Learning (ICWL'06), 10–24, 2006.
[11]N.Soonthornphisaj,E.Rojsattarat,S.Yim-ngam,SmartE-LearningUsing RecommenderSystem, in:InternationalConferenceonIntelligent Computing, 518– 523, 2006.
[12]M.K.Khribi,M.Jemni,O.Nasraoui,AutomaticRecommendationsforE-LearningPersonalizationBasedonWebUsageMiningTechniques and Information Retrieval, in: Proceedings of the 8th IEEE International Conference on Advanced Learning Technologies, IEEE Computer Society, 241–245, 2008.
[13] S. Rendle, L. Schmidt-Thieme, Online-updating regularized kernel matrix factorization models for large-scale recommender systems, in: Proceedings of the ACM conference on Recommender Systems (RecSys'08), ACM, New York, USA, 251–258, 2008.
[14] R. M. Bell, Y. Koren, Scalable Collaborative Filtering with Jointly Derived Neighborhood Interpolation Weights, in: Proceedings of the 7th IEEE International Conference on Data Mining (ICDM'07), IEEE CS, Washington, USA, ISBN 0-7695-3018-4, 43–52, 2007.
[15] T. Tang, G. McCalla, Smart Recommendation for an Evolving E-Learning System: Architecture and Experiment, International Journal on E-Learning 4 (1) (2005) 105–129, ISSN 1537-2456.