ISSN (O): 2455-5738
GADL Journal of Inventions in Computer Science and Communication Technology (JICSCT)

Available Online at: www.gadlonline.com
Volume –4, Issue – 2, March– April - 2018, Page No. : 63 - 67

An Evolution of Activity Sequence Modeling with Active Clustering

Keshav Dev Gupta
Ph.D Scholar
Department of Computer Science & Electronics Engineering
Jayoti Vidyapeeth Women's University, Jaipur
Rajasthan, India

Dr. Kavita
Associate Professor
Department of Information Technology Electronics Engineering
Jayoti Vidyapeeth Women's University, Jaipur
Rajasthan, India

Abstract

Acquisition and preprocessing of data is a fundamental step in the process of searching for educational data and also represents the initial stage of the adaptation cycle. The nature of the monitoring data is a decisive factor for later stages of the cycle and for further analysis. Most adaptive education systems share a strong dependency of this early stage of the process. However, the accuracy of the data itself and the way they are being effectively monitored data may differ.

Keywords: sequence modeling, clustering, dependency, monitoring, problem solving.

1. Introduction

The original pool of recorded data, for example, each user use for navigation purposes, and contains very fine work. However, the data is analyzed at the original granularity, summarized and converted to more aggregate form (for example, students, the number of failed quizzes, the number of allocated tasks done in the number of past investigations, etc.) The ultimate goal is to evaluate the usefulness of the student's final grade and the performance of a different classification algorithm.

2. Student Modeling Based on Clustering:

In this research, we will introduce cluster in the context of student and explore more specific parts of the related work. We can find a detailed explanation of the classification of authors and clustering methods for modeling users. Their concepts of fields are more traditional instructive systems that students have to a more structured understanding, their database coming from more exploratory learning environments. The data they use is transformed into feature vectors supplied to the integration phase. The feature vector represents an aggregated version of the student's activities. Therefore, a low overall number of vectors are obtained, and there is only one feature vector for each student. The authors explain a similar approach and grouping is used to automatically recognize a group of learners in an explorative learning environment [Amershi and 2006 Conati],

Collaborative behavior based clustering methodology allows authors to find out how statistical measures in learning activity data are used to determine cluster membership [Anaya and Boticario, 2009]. Data was monitored for students at UNED [dotLRN, 2010] via the dotLRN platform. The described monitoring process began with the initial and required individual task questions completed by each learner. Each result, the manual group was used for learners, 3 members of each team. At a later stage, the team, for example, each member needs to solve some of the specific problem, given an additional task, a team; we had to integrate the individual solutions. One expert observed the processes and then used the learner's collaborative results to label the statistical data supplied to the algorithm (ie the aggregated version of the documented learning activity) In order to clarify the relationship with the joint behavior with the statistical indicator, clustering.

3. Sequence-Based Clustering Approaches:

This section reflects the sequential information in the activity log of the data and therefore summarizes the clustering method in the e-learning context comparable to our work. The use of Markov models in machine learning is important if domains require that sequences be expressed and analyzed, providing a convenient way to model interdependent data. In the field of EDM, this advantage has been used recently in various searches. For example, we can find a coordinated analysis based on Hidden Markov Model (HMM) in [Solar and Lesgold 2007] and [Solar, 2007]. In [Solar and Lesgold 2007], the modeling process is described in the case of such knowledge sharing, which defines "a series of knowledge sharing episodes of member initiation contributions and conversation actions. The group is in a group conversation Introducing new knowledge, sometimes the discussion of new knowledge will end on a continuous basis. "

Subsequent analysis is knowledge sharing if the recipient has assimilated new knowledge New knowledge is to explain and evaluate method analysis, role distribution (knowledge sharing for the receiver). By recording the activity sequence, the communication interface includes a tagging function which helps classify the individual activities. The marking process, that is, it is a manual that requires human effort. In the experiments described by the author, trained HMMs provide very good precision to identify the role of knowledge sharing. However, it is not completely clear as the number of states of reasons using the hidden model is known in advance. In our study, as we will explain below, we used a predefined number of discrete Markov models (DMM) of the state (as indicated by possible learner's behavior on the platform) I will.

Another pattern detection technique can be found used to model the performance of students to solve the HMM problem [Beer et al., 2007]. The model is adapted to the activity array of students with three virtual hidden states corresponding to "commitment level" students. The resulting HMM is used for group students in groups with specific patterns of subsequent actions. In addition, the model is a prediction based on the later stage of the process. As they are explicitly used to model unobservable effects, in this case the HMM is obviously suitable as indicated by the highest prediction accuracy compared to the simple Markov chain It is an analytic approach.

A different approach to sequential pattern extraction [Perera et al., 2009] The authors use surveillance activity data by the system to support mirroring, that is to say they are described in the model that characterizes the actions of successful groups Then, now. They are not restricted to the suggested tools, but specific rules on their use are available, but are designed to monitor genuine as a collaboration process of possibilities including the frequency of tool selection and use. The main purpose of this work is to allow them to interpret it, "use my knowledge to extract profiles and other information from the group of articles and put them in a hope model for stakeholders It is to present Group tasks and activities. "The underlying concept "Big Five" group based on the theory of labor and defines five major factors [Saras et al., 2005.]: Leadership, mutual performance, monitoring of backup behavior (eg reassigned work between members), adaptability and team direction. After the monitoring process, the final data pool contained user actions and traces of both the progressive brand and the final group. Groups were classified according to their performance (ie D. Notes). Classification was then used to determine the strongest distinction behavior type of the weakest group. This was achieved by a simple statistical analysis in the first stage and by applying the following data mining techniques (grouped into groups and students). A Loop-level clustering Basic data includes aggregation group activities as the average number of events like a specific tool, grouped database-level students include similar information for individual students. Detected at the student's level with the phase of aggregation Clusters define different types of users such as "administrator" or "loafer." Analysis of the activities by the model extraction process revealed the most important activities that made the group "strong", "weak," the authors said their approach is not perfectly mature, There are some restrictions in terms of data (wild type) and warn you that the release was interpreted.

In [Chon Biswas, 2008], another approach to behavioral modeling is presented. The author explains the research between the teacher's operating officer and high school student. They are using HMM again to represent the order of activities, to reveal patterns that lead to successful learning. The specific goal in this case was whether to provide better learning opportunities compared to other parameters to learn by teaching (self-control system and coaching). Thus, the array of models was mainly used as an aid to evaluate the concept of learning with this approach.

In [Li and Yu, 2006], the authors describe the modeling of learning behavior of students with Bayes Markov chains used in the adaptive teaching system Atre [Yura, 2005]. Their approach only takes two basic observations, including the specific task format, including the level of specific difficulty, which are saying correct and wrong answers to consideration. Also, the authors regularly postulate that they are just the three types of basic students based on three types of learning, which is the type of building, saying it (type of challenge and kind of defining type) . Their goal is to model clustering of students' behavior and use clustering using a model of the result to predict the learning style of new users. They show that the base sequence information can be used successfully in the integration process, for example, to improve the static model derived initial search so this result is interesting in the context of this article. What is proposed in this work, but goes one step further and does not count the number of models, but rather, as explained in more detail below, they are to be dynamically determined.

Another interesting use of the Markov model for business analysis can be found in [Mart'ın et al., 2011], in this issue. In this case, instead directly, the authors will search the behavior of learner's classification students based on the initial user characteristics and context. And they will build Markov A model of learning activity transition within each cluster to derive an activity sequence of recommendations for future users. As such, this approach encourages a sequence of activities that are proven to be useful to learners (and contexts), as in the past, rather than not trying to analyze behavior based on a direct model.

4. Data Analysis and Activity Sequence Modeling:

In order to facilitate discussion on the proposed approach, we [2005,. VanLehn et al.] Here we describe its practical use in modeling and ITS of Andes and its practical use in analysis of problem solving and can be used via PSLC DataShop [Koedinger Toru, 2008], [Koedinger et al., 2010]. Specifically, the experiment was conducted on data from physics courses at the US Naval Academy (USNA) and repeated data from different semesters (Spring 2007, 2008 and 2009).

Extract active sequence to extract raw data, sequence Markov model secret (transformation modeling department: the software we used in this study embodied the proposed approach consists of 3 main components and explained in more detail in thesis and learning library machine grouping based on clustering algorithm embedded in machine Weka [Hall et al. , 2009].

5. Multi-targeted Clustering Approach:

As defined by converting the genre array representation and the model corresponding to the database, this section describes the different direction from the integrated phase. Figure 2 provides a brief overview of the overall approach.

The process begins at the preprocessing stage, including the definition of the model used, depending on the database and possible user activities (DMM here), the structure of the data transformation (ie, ie, Sequence transformation) and expression power parameters later evaluate the quality of the cluster. Then, at the experimental stage of consolidation, the number of consideration features in the next step is limited by identifying and evaluating its characteristics (eg their discriminative ability) and the optimal number of clusters for each scenario I will decide. Solving styles If a predefined specific problem must be identified in the user's activities, in the integration phase, firstly, clustering concrete goals need to be defined, which means that he We have to decide, to say that if a predefined problem is to recognize or if the integration process needs to autonomously detect size and the potential style of user behavior must be resolved .the first In the case, each style MUST be defined, the most relevant functions and expectations selected for them MUST be identified. Based on this information, the appropriate data set is created, which is done in the clustering process. In the second case, you need to define each dimension, and the most important functions to be selected before the data set can be created to be supplied to the clustering process. In the third case, specific characteristics are not preset, but constraints such as the maximum number of features that can be used with such data set are identified. Different sets of data conforming to the above are automatically created and running the integration process.

6. The General Process:

We found a way to model the sequence of learning activities; we draw people's attention to the analysis of such sequences to discover the characteristic behavior of the example (People with behavioral characteristics.) Learning Style or problem solving), or do process or activity, for example (context being done, project progress of collaborative learning class). The first step in this direction is a group of earlier business model derivatives, generally combined with other monitoring data related to the activities themselves, if necessary.

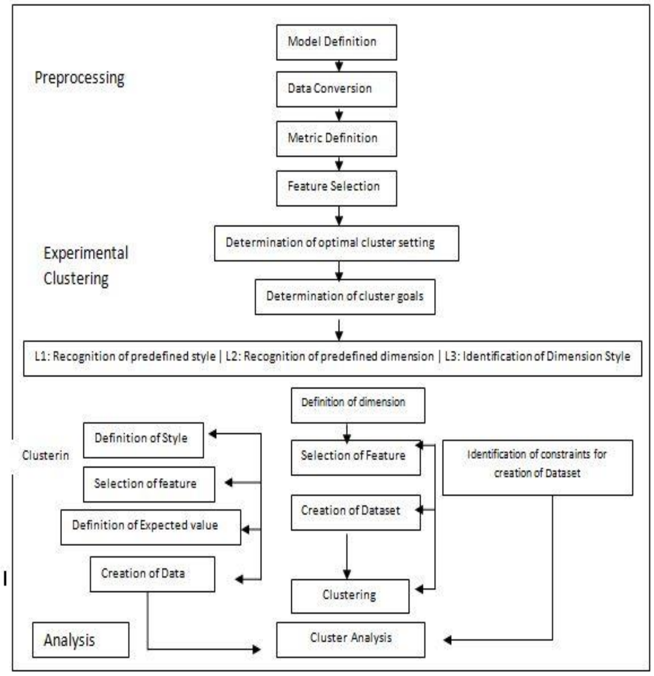

Figure: Overall process of the proposed approach to sequence modeling and subsequent dynamic clustering**.**

References

[1] Yoshua Bengio and Paolo Frasconi. Input-Output HMMs for Sequence Pro- cessing. IEEE Transactions on Neural Networks, 7(5):1231–1249, 1996.

[2] A.K. Jain, M.N. Murty, and P.J. Flynn. Data Clustering: A Review. ACM Computing Surveys, 31(3):264–323, 1999.

[3] Matthew Jarvis. The Psychology of Effective Learning and Teaching. Nelson Thornes, 2005.

[4] H. Jeong and G. Biswas. Mining Student Behaviour Models in Learning-by- Teaching Environments. In Proceedings of the 1st International Conference on Educational Data Mining, pages 127–136, 2008.

[5] Essi Kanninen. Learning Styles and E-Learning. Master's thesis, Tampere University of Technology, 2008.

[6] Mirjam K¨ock. Towards Intelligent Adaptive E-Learning Systems – Machine Learning for Learner Activity Classification. In Proceedings of the 17th Workshop on Adaptivity and User Modeling in Interactive Systems (ABIS 09), pages 26–31, 2009.

[7] K. Koedinger, K. Cunningham, A. Skogsholm, and B. Leber. An Open Repos- itory and Analysis Tool for Fine-Grained, Longitudinal Learner Data. In Proceedings of Educational Data Mining 2008: 1st International Conference on Educational Data Mining, pages 157–166, 2008.

[8] Kenneth R. Koedinger and Vincent Aleven. Exploring the Assistance Dilemma in Experiments with Cognitive Tutors. Educational Psychology Review, 19 (3):239–264, 2007.