Analysis of Data Mining Concepts in Higher Education with Needs to Najran University

Mohamed Hussain Tawarish[1], Farooqui Waseemuddin[2]

Department of Computer Science, Najran Community College.

Najran University. K.S.A

Abstract: *One of the biggest defy that higher education faces today is predicting the paths of current and alumni students, the purpose of this research is to focus the benefits of the current and alumni students of higher education, Najran University, through Data Mining concepts. In the present scenario textual information is being augmented by ever increasing amounts of non textual information such as images, audio, video. Thus mining of this non textual information plays a great role in its retrieval. Data mining is a process of extracting previously unknown knowledge and detecting interesting patterns from a massive set of data. In the following research approach a functional architecture of user customized audio information on demand system which is termed as Personal Audio Cast (PAC) is proposed, and some of the data mining techniques that might help the staff and the students to achieve the goal.*

Keywords: *Data Mining, Audio and Video Mining Techniques, Education, Najran University, Predictive Mining.*

Introduction

Educational data mining is emerging as a research area with a suite of computational and psychological methods and research approaches for understanding how students learn. New computer-supported interactive learning methods and tools—intelligent tutoring systems, simulations, games—have opened up opportunities to collect and analyze student data, to discover patterns and trends in those data, and to make new discoveries and test hypotheses about how students learn. Data collected from online learning systems can be aggregated over large numbers of students and can contain many variables that data mining algorithms can explore for model building. [8]

Many efforts have been established to help educators realize the benefits of technology and ways of implementing them in the classroom. Many different types of technology can be used to support and enhance learning. Some examples of technologies used in the education system are video content and digital movie making, laptops, computers, and handheld technologies. [1]

Many students are growing up in a digital age where they have constant exposure to a variety of media that is impacting on the way they interact and use information. [1], [2].

- Audio Mining:

The Web, Databases, and other digitized information storehouses contain a growing volume of audio content. Sources include newscasts, sporting events, telephone conversations, recording of meetings webcasts, documentary archives such as the visual history foundation's interviews with Holocaust survivors (http://www.vhf.org) and media file in libraries. Users want to make the most of this material by searching and indexing the digitized audio content.

Salient features of audio mining:

Audio mining software uses two basic methods, both consisting of indexing and searching phases. Large Vocabulary Continuous Speech Recognition (LVCSR) analyzes word or phrases to generate an index; phonetic audio mining analyzes sound patterns to generate an index and stores the phonetic content of the speech, rather than information about words. The audio mining software makes a rough transcript of the important words in the file (the index) and then searchers one or more index files for all matches to specified searched terms.

- Video Mining:

Due to MPEG-4 and MPEG-7 standards, there is a further overlap in research for computer vision, computer graphics, image processing and databases. In a typical model-based coding for MPEG-4, video is first analyzed to estimate local and global motion then the video synthesized using the estimated parameters. Based on the difference between the real video and synthesized video, the model parameters are updated and finally coded for transmission. This is essentially analysis followed by synthesis, followed by model update and followed by coding.

Salient features of video mining:

Modern computer technology, together with the proliferation of broadcast channels and of video-based surveillance systems, has enabled us to produce vast amounts of both raw and processed video data. The potential uses of this data are many and varied.

Monitoring and mining of the content of this already huge, rapidly growing mass of data calls for the development of major computational resources and the development sophisticated video understanding techniques.

Review of Literature

Data mining is a powerful tool for academic intervention. Through data mining, a university could, for example, predict with 85 percent accuracy which students will or will not graduate.

The university could use this information to concentrate academic assistance on those students most at risk.

Data mining enables organizations to use their current reporting capabilities to uncover and understand hidden patterns in vast databases. These patterns are then built into data mining models and used to predict individual behavior with high accuracy. As a result of this insight, institutions are able to allocate resources and staff more effectively. Data mining may, for example, give an institution the information necessary to take action before a student drops out, or to efficiently allocate resources with an accurate estimate of how many students will take a particular course. [4]

Data mining combines machine learning, statistics and visualization techniques to discover and extract
knowledge. for universities, the knowledge discovered by data mining techniques would provide a personalized education that satisfies the demands of students and employers. data mining techniques can be applied to provide further knowledge beyond the data explicitly stored. Compared to traditional analytical studies, data mining is forward looking and is oriented to individual students. [5]

Framework of Audio and Video Sequences

Methodology

The functional architecture of a user customized audio information on demand system which is termed as Personal Audio Cast (PAC). A user request is processed by the user interface and dispatched to the PAC selector. The selector then constructs a script for a user personalized audio package, employing the database and a profile for the user. Selections from the PAC database are assembled and packaged for the user as a continuous audio presentation. The selector then passes the script to the PAC scheduler which dispatches and delivers the package to the user. for example, several hours of broadcast lecture audio are stored in PAC database along with descriptions of audio (metadata). Let us assume a hypothetical student who is only interested in Windows and Linux Platform. These interests are included in the profile for that particular student. Now, suppose this student sends a request by specifying "Lecture on Operating system platforms". The user interface handles this request and sends it to the PAC selector. Consulting the student profile, the PAC selector generates database quires related to Windows and Linux platforms, and submits these to the database. Relevant audio lecture items are identified and the scheduler starts to play them. Thus, by employing a student profile unwanted lectures items are automatically filtered out.
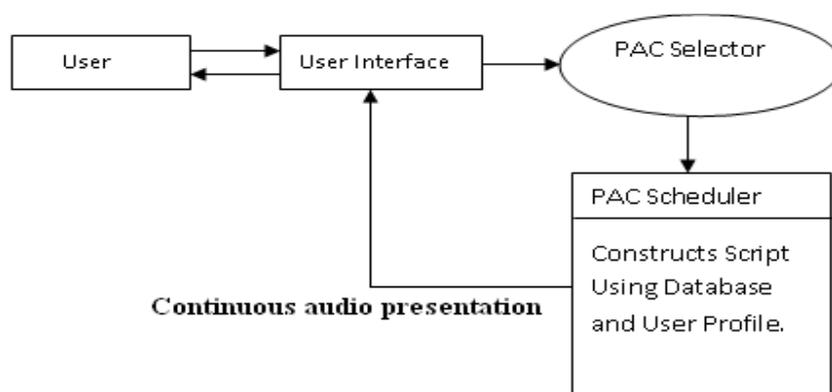
Figure 1: PAC Architecture

From the above framework for the PAC system, key research and engineering problems/tasks can be identified that must be addressed:

- Segmentation

- Metadata Acquisition

- Selection

- Scheduling

- User Profile Generation

Work Area

First, different customization techniques to achieve personalization are discussed. Second, different efforts for metadata acquisition are presented. Finally, key information modeling techniques for the selection of multimedia information are presented.

For metadata acquisition it is necessary to specify the content of media objects. Therefore, to ensure appropriate selection and presentation of audio information, it is to capture its semantic description. By using word spotting for selected content extraction speech recognition looks for a set of pre defined key words in audio where these keywords should covey semantic descriptions. These keywords are from ontology.

Although audio, is used the related work in the video domain which is closest to and which compliments our approach in the context of data modeling for the facilitation of information request.

Data Mining Techniques:

- *Student Personalization Data*

Data mining (DM) is a series of data analysis techniques applied to extract hidden knowledge from server log data Roiger & Geatz, 2003) by performing two major tasks: pattern discovery and predictive modeling (Panov, Soldatova, & Dzeroski, 2009). Pattern discovery involves extracting unknown interesting patterns. For example, online instructors can utilize pattern discovery techniques to classify students based on their shared learning preferences, to identify outlier students, and to depict the frequent navigational paths in the course. Predictive modeling involves analyzing current or historical facts to make predictions about future events. For example, online instructors can utilize predictive modeling techniques to identify key predictors of students' academic performance and then interventions can be developed for performance improvement.[7]

- Are the students ready to proceed to the next topic?

- When a student is sinking behind in session?

- When is a student at hazard for not completing a course?

- What score is a student likely to get without involvement?

- Which is the best following path for a given student?

- Must a student be referred to a predictor for help?

Students Personalization Technique
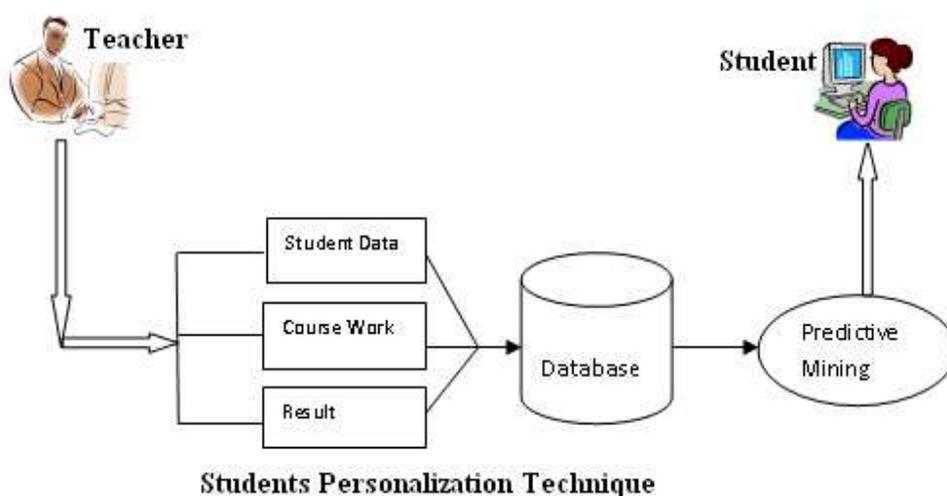
- *Student Metadata Acquisition*

Data that serves to provide context or additional information about other data. It may also describe the conditions under which the data stored in a database was acquired, its accuracy, date, time, method of compilation and processing, etc.

Effective management of metadata facilitates efficient storing and retrieval of audio, video and textual information.

Most of the educational metadata schemas have been projected over time in order to better distinguish learning objects.

Metadata can be organized into several levels ranging from a simple listing of basic information about available data to detailed documentation about an individual data set or even individual features in a dataset [9]

First, simply store the most general concept. But there are relevant objects for queries to specific concepts. Second, the most specific concept can be stored in the database. In this case, in order to support queries related to the most general concepts associated with this object are stored, thus facilitating efficient retrieval.

- *Alumni Services*

✓ Alumni Services: Help them further to progress their education with much ease and flexibility.

✓ Consultancy Service: To help students identify both their educational and career goals in life.

✓ Career Service: Helps students in landing their desired jobs. Match the talent of the students with the requirement of job market.

✓ Predictive Data Mining for the alumni students.

- *Key Information Modeling: Multimedia Information*

Data mining is a powerful new technology with great potential in information system. It can be best defined as the automated process of extracting useful knowledge and information including, patterns, associations, changes, trends, anomalies and significant structures from large or complex data sets that are unknown (Han and Kamber, 2001; Two Crows Corporation, 1999; Chen *et al.*, 1996) [10]

In this research the various new proposed techniques can be used to implement for key information modeling, such as, Database Query Techniques, Mathematical Model techniques, Web based Techniques. Textual Information Techniques, Program Based Techniques.

Conclusion

In this paper we discussed about data mining concepts and technologies in education which are useful for the current and the alumni students. Some of the main aspects discussed are audio, video sequences, PAC, predictive concept and questions, student metadata acquisitions and key modeling features.

In Higher Education, the Universities will gain more benefit, outline and links to reach a conclusion for current, alumni student performance and activities.

The skill to expose unknown pattern in databases, community colleges and universities can build replica that predict the performance by mining.

Our upcoming research include applying data mining techniques on an stretched data set with more unique feature to get better results.

This research will truly be helpful for all aspects in higher education, and especially in Najran University current and alumni students.

Acknowledgement

References

1] http://en.wikipedia.org/wiki/Impact_of_technology_on_the_educational_system

2] Geer, R., & Sweeney, T. (2012). Students' voices about learning with technology. Journal of Social Sciences, 8 (2). 294-303)

3] P. K. Singh and A. K. Majumdar. Semantic Content Based Retrieval in a Video Database. In Proc. Of International workshop on multimedia data mining (MDM/KDD'2001) pages 50-57 San Francisco, CA August 2001.

4] Jing Luan, PhD. Chief Planning and Research Officer, Cabrillo College Founder, Knowledge Discovery Laboratories.

5] Ying Zhang, Samia Oussena Thames Valley University, London,UK. Tony Clark, Hyeonsook Kim Middlesex University, London,UK

6] Structuring and Querying Personalized Audio Using Ontologies1 Ph.D. Thesis Proposal Latifur Khan Department of Computer Science and Integrated Media Systems Center University of Southern California Los Angeles, California 90089

7] Jui-Long Hung Kerry Rice Anthony Saba Boise State University

8] Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: U.S. Department of Education Office of Educational Technology.

9]   http://www.gitta.info/MetaDataQual/en/multimedia/MetaDataAcqua.pdf

10] Naeimeh DELAVARI, Somnuk PHON-AMNUAISUK Faculty of Information Technology, Multimedia University, Mohammad Reza BEIKZADEH, Faculty of Engineering, Multimedia University