State Schools Quality Assurance Monitoring System
Using Data Mining Technique

Ja’afar A. Nasir
SOCSIT, Linton University College
nasirjafer10@yahoo.com

Mia Torres-Dela Cruz
SOCSIT, Linton University College
mia@leg.edu.my

Abstract

Ensuring quality education is seen to be one of the most essential responsibilities that State Education Boards maintain to its schools in order to produce graduates of secondary school students who are higher-education ready and have received relevant knowledge to pursue their chosen careers. This is achieved through monitoring of the schools based on standard sets of formalized features like, methodology of teaching, classifying of students, types of support, educational resources available and fair attention to all students. State school boards will use resources in the form of administrative and academic information, particularly teachers’, staff’s and students’ information, in the various schools they monitor. The management board has an issue of keeping up-to-date as to how their schools are operating. They are also concerned with keeping their schools’ records (students’, teachers’ and staff’s records inclusive) and how these can be conveniently and effectively accessed to keep track of the progress and performance of the schools under their management. The amount of data stored in educational databases is huge and rapidly increasing which makes it more challenging to the board. It is important to maintain these databases that contain records of all concerns of the schools, thus, a system of management is developed to ensure systematic organization and control of these important information. The system utilized data mining, the science of filtering data for information and knowledge retrieval, which has developed a new set of applications and an emerging discipline, called educational data mining. It is concerned with developing methods for discovering useful and valuable knowledge from data found in educational database domains. It would enable organizations to facilitate better resource operation to achieve quality education and improve students’ performance. The system is applied as a case of Katsina State Science and Technical Education Board (KSSTEB). Classification and cluster techniques of data mining were used in this research to evaluate school performances in educating their students and it addresses the applications of information management in the educational sector to extract valuable information from the available data set that will serve as analytical tool to ensure quality education among schools as monitored by state education boards like KSSTEB.

1. Introduction

The paper will study and propose to develop and implement an intranet system that will allow or help state management boards to monitor and control the number of the schools that operate under them. This system is needed to enable and allow the management board to gather schools’ valuable data about their students, teachers and staff from different perspectives and make use of this information for proper decision making to enhance schools productivity and performance (Lawrence et al., 2008).

Data mining techniques is employed in this research as it is applied to the educational sector to identify and enhance educational process to ensure quality – how to enhance it and how to evaluate it (Ayesha et al., 2010). The proposed system is however going to focus on
understanding how users interact with computer system to perform tasks and enhance user experience. It also introduces Human Computer Interaction (HCI) concepts and goals as it draws upon interest and expertise in many important disciplines such as ergonomics, psychology (both cognitive psychology and experimental psychology), sociology, anthropology, computer science and more, to design the system that is safe and efficient, comfortable and even enjoyable for humans who use the system (Mackenzie, 2013).

Different researches have been made which are associated with the educational system of different levels using data mining techniques in order to improve the productivity of the sector. Besides, nowadays computer systems are globally used in most certainly all the universities across the globe for different purposes to increase operation efficiency (Ayesha et al., 2010). Moreover, most of the countries have specific systems or departments that control and monitor their schools in different level. Most of them have a series of theories that serve for the same purpose.

This system is applied as a case of Katsina State Science and Technology Education Board located in Katsina State, Nigeria. The system is based on information taken from this board but will, in the long run, be beneficial to all state education boards around the globe. It will help the education management board to monitor and control the schools that they operate. The system will provide a user-friendly login interface for the board administrator and users to login, it will help the administrator to add new school to the system, collect and save schools information, it will also help him to add or register new student from different schools and manage their records and performances. It will also help the administrator to add or assign new staff (teacher) to their various schools with his records and transfer history and it should also enable the administrator to generate promotion and demotion exercise report on the academic session basis. Besides, it will also help the school administrators to submit their schools report to the system which can be accessed by the management board administrator, and also it should help the board chairman to generate reports for the various schools based on the data collected from the respective schools.

In this real world, people often say that there is nothing that could be regarded as a new, instead that all apparently new knowledge is based at least partly upon previous knowledge. In every subject area, literature review provides that previous knowledge which helps researchers and anchor to which to attach their new ideas (Olson and Delen, 2008).

The biggest challenge that faces the educational institution sector today is the exponential growth of educational data and the use of it to improve the quality of managerial decisions (Ayesha et al., 2010). However, according to Maimon and Rokach (2005), data mining techniques are the analytical tools that could be used to extract and get meaningful and useful information or knowledge from these large data sets (in databases, data warehouses or other information repositories).

Higher educational institutions, schools and colleges are interested in predicting the paths of students and alumni. Quality in higher educational systems nowadays is given most consideration in any educational system. There are numerous problems facing educational systems, these challenges act as obstacles to attain their quality objectives. Most of these problems are associated with the processes and operations in the educational system like planning, evaluation, counseling and monitoring of the higher education systems.

Widely range of data associated to schools and its students in schools forms, logs and databases can be used in educational research to improve educational system. The field of data mining is concerned with finding new patterns in large amounts of data widely used in business sectors, sciences (like biomedical and DNA analysis), web mining and engineering management by using different techniques to support decision making (Han and Kamber, 2011).
2. Schools Monitoring System using Data Mining Techniques

Gnanasundaram and Shrivastava (2012), they described monitoring as an internal process activity which provides necessary feedback to any organization management on the process of a particular subject matter, the problems that face it and the efficiency with which it is being implemented. Hence, there is a need of continuous collection of information and review of the entire storage infrastructure.

However, Maimon and Rokach (2005) stated that, monitoring could be regarded as a system which comprises of the three critical components: it requires regular collection of information, evaluation of the collected information and most importantly evaluation of the results in an institutional (project). These are the pioneers to determine what information should be collected and hence make judgment about the status of a system. Cristobal Remero and Sebastian Ventura on their paper titled Educational Data Mining: A Review of the State of the Art stated that, for the successful future of every country and the society progress, more attention must be given to promote the development rate of their education, improving its quality and advancing the management of its supervision.

Monitoring of education is the management of education quality, which will help an organization in decision making in the determination of its further development, directions, improvement and reconstruction.

Furthermore, Cristobal added that educational institutions need to adopt modern management practices and state of the art technology to manage their internal and external operations. The awareness of such demands motivated software industries and started developing automated solutions for educational administration (Cristobal et al., 2012).

However, different researches have been carried out on tracking and mining data stored in educational institutions in order to enhance the educational sector (in terms of student performance, attendance, students' assessment and alumni). The usefulness of mining this data is promising but still need to be proven and stereotypical analysis to be streamlined (Beck, ed). Different researchers already tried and set up some guidelines for ensuring that educational data could be mined and used out of their experience of mining data in project (Saxena and Saxena et al., 2009).

3. System Concept

The system is an intranet-based system that uses computer technologies to facilitate communication and access of information within an organization particularly, state education boards. An intranet system is basically an integrating mechanism for people, processes and information within an enterprise. Such systems can often be accessed within a restricted area. It helps to connect across disparate platforms and enable users to control their data (Vacca, 2014). According to Terplan (2000), intranet technology has proven be one of the effective ways of accessing and disseminating data or knowledge available within an organization. However, the only challenging problem which organizations face before and then is none other than decision-making capability. Top management level that are responsible for decision-making are unable to make appropriate decisions due to the immature of the required data to help to do so not being available. Due to the large amount of data obtained within organizations, these mandated organizations to strive hard to find appropriate tools and techniques to manage their data or knowledge. In order to overcome these kinds of situations within organizations, better solutions are evaluated and the term intranet evolved (Terplan, 2000).

Furthermore, this system will employ data mining techniques concept as a tool to extract valuable data from the information in the management board database to help in
decision-making and increase schools productivity. Data mining is an interdisciplinary field of economics, computer science business and other related fields to discover new pattern to extract useful knowledge and information from large data set stored in the databases, data warehouses and other repositories to analyze large quantity of data in order to discover previously unknown patterns (Knobbe, 2006). It is however known as 'knowledge discovery in databases', the science of filtering data for information and knowledge retrieval. This branch of science has developed new album of applications and an emerging discipline, called educational data mining. Educational data mining concerns with developing methods for discovering useful and valuable knowledge from data found in educational databases domain. Data mining is accepted as decision support tool that would enable organization to facilitate better resource utilization to achieve quality education and improve students’ performance.

4. Data Mining Techniques

Data mining, according to Berry (2008), has been applied in a great number of fields, including retail sales, bioinformatics, and counter-terrorism. In recent years, there has been increasing interest in the use of data mining to investigate scientific questions within educational research, an area of inquiry termed educational data mining. Educational data mining (also referred to as “EDM”) is defined as the area of scientific inquiry centered around the development of methods for making discoveries within the unique kinds of data that come from educational settings, and using those methods to better understand students and the settings which they learn in. This is the exact rationale for the paper as educational boards would want to monitor schools and study how these can be improved with regards to quality of education.

The paper used clustering and classification technique of data mining. In clustering, the goal is to find data points that naturally group together, splitting the full data set into a set of clusters. Clustering is particularly useful in cases where the most common categories within the data set are not known in advance. If a set of clusters is optimal, within a category, each data point will in general be more similar to the other data points in that cluster than data points in other clusters. Clusters can be created at several different possible grain-sizes: for example, schools could be clustered together (to investigate similarities and differences between schools), students could be clustered together (to investigate similarities and differences between students), or student actions could be clustered together (to investigate patterns of behavior) (cf. Amershi & Conati, 2006; Beal, Qu, & Lee, 2006). Clustering algorithms can either start with no prior hypotheses about clusters in the data (such as the k-means algorithm with randomized restart), or start from a specific hypothesis, possibly generated in prior research with a different data set (using the Expectation Maximization algorithm to iterate towards a cluster hypothesis for the new data set). A clustering algorithm can postulate that each data point must belong to exactly one cluster (such as in the k-means algorithm), or can postulate that some points may belong to more than one cluster or to no clusters (such as in Gaussian Mixture Models). The goodness of a set of clusters is usually assessed with reference to how well the set of clusters fits the data, relative to how much fit might be expected solely by chance given the number of clusters, using statistical metrics such as the Bayesian Information Criterion (Berry, 2008).

Classification technique, on the other hand, is a kind of prediction technique of which predicted variable is binary or categorical variable. In prediction, the goal is to develop a model which can infer a single aspect of the data (predicted variable) from some combination of other aspects of the data (predictor variables). Prediction requires having labels for the output variable for a limited data set, where a label represents some trusted “ground truth” information about the output variable’s value in specific cases. In some cases, however, it is important to consider the degree to which these labels may in fact be approximate, or incompletely reliable (Berry, 2008).

5. System Developmental Methodology
The author in this context chose to use *Throwaway Prototype* as the development methodology of the proposed system. This is because throwaway prototype allows both the developers and all the stakeholders to first sit and analyze the entire system and develop or build a system prototype (operational sample of the system) of the proposed system based on the potential requirements gathered, the prototype will be given to the users to use and evaluate and supply constructive feedback to the developers. Through this, users will understand the system and will know whether all that they want are covered or included in the developed prototype. The proposed system will be built and developed through this simple development methodology, any requirement or function either not needed by the users will be identified and thrown away and any which is not included will be detected and included in the system. This will be adopted throughout the development of the proposed system until the complete system is completed.

To identify the user requirements, the researcher interviewed educational board personnel and school heads. The data gathered were processed and analyzed to be integrated into the design and implementation of the intranet system for the project.

6. **Conclusion**

Education is considered wealth that everyone wants to invest in. Everyone wants to make sure that their children would be educated and ready for future careers. But this education should be of good quality in order to be competitive in the working world. In this context, education boards all over the globe would want to make sure that people are investing in quality education. This proposal is then being done to look into a more innovative and beneficial system for the education boards so they can monitor their schools in more efficient system. Data mining technique is used, particularly cluster and classification techniques, to be able to analyze and sort out the bulk of data that the boards are facing in order to do the task they are responsible of. The resulting algorithm and model developed will be used to design and implement an efficient and secure system for the educational monitoring board. An intranet system is developed and implemented so that monitoring for quality would be easier, more convenient and technologically sound. In this regard, optimum benefits will be received by users as the system will help to increase staff productivity and ease school managing and monitoring work; instant records can be generated from the system regarding a particular school, its progress and achievement; and schools, and all its elements, can easily controlled and monitored.

**References**


