Enhancing Education Quality Assurance Using Data Mining

Case Study: Arab International University Systems.

Faek Diko
Arab International University
Damascus, Syria
f-diko@aeu.ac.sy

Zaidoun Alzoabi
Arab International University
Damascus, Syria
z-zoabi@aeu.ac.sy

Mouhib Alnoukari
Arab Academy for Banking and Financial Sciences
Damascus, Syria
m-noukari@aeu.ac.sy

Abstract

In this paper we introduce a comprehensive educational quality assurance system for a university. The system takes into consideration the three main pillars of the educational process: content, delivery, and assessment. We will demonstrate a comprehensive system that enables quality control and quality assurance using data mining combining data from Quality Assurance Automated System QAAS, the Academic System, HR System, and Financial System focusing on various performance indicators in the aforementioned three pillars. We will explain the system through a real-life case, where this system produced valuable information in a way that helped Arab International University (AIU)-Syria to make use of the vast data produced by its main four systems to uncover hidden trends, knowledge, and quality deficiencies making it easier to the decision maker to plan, assure and control quality.

Keywords: *quality assurance, data mining, QAAS, higher education.*

1. Introduction

Quality assurance has had an increased attention for all higher education institutions. This has been forced by different factors such as the growing awareness for the need of quality assurance systems in different disciplines, pressure from government entities towards standardization, customer demand for better performance, and the need for more organizational efficiency and excellence (Gatfield et al. 1999, Fowler and Gilfillan 2003).

This has been emphasized through international recognition for the importance of assuring quality in higher education systems in order to facilitate student and instructor mobility, the need for joint and dual degree programs, and cross cultural programs

The use of information systems in developing quality assurance comprehensive systems has been the focus of software engineers in various industrial disciplines (Blackmur 2004, Spalter and Dam 2003). However, educational system was lacking such type of information systems which could be attributed for two main reasons:

First, the concept of quality assurance in educational systems is rather new and has not matured enough (Yang 2006).

Second, the concept of quality assurance in educational systems seemed to be difficult to computerize (Tait 1997). For example how can we computerize aspects related to assuring quality of the content of a program using information systems? How can we assure quality in the way instructors are teaching their classes with the help of information systems?

An exception to that could be the use of data mining and data warehouses in improving decision making process. One example is the work of Selmoune and Alimazighi (2008) where they developed a data mining technique to enhance the quality of the services provided by a university by analyzing its pedagogical results, discovering success and failure factors. The system traces indicators such as absence ratio and

difference in results in modules with their prerequisites.

Despite of the interesting approach used here, we still find it insufficient as it does not cover aspects such as delivery and assessment in order to have panoramic view of the entire pedagogical process.

In this paper we will demonstrate a Quality Assurance Automated System (QAAS system) developed at the Arab International University. The system is a full-fledged information system that stores almost every academic move for instructors and students, traces performance of both, and provides management at various levels with information that helps in discovering all factors affecting performance. This occurs in analytical view in a way that helps in making decisions to improve the educational process.

QAAS data and information are combined with information produced by the academic system that stores the academic history of the students in addition to the HR system which stores the history of instructors in a way that provides a panoramic view that helps Quality Assurance Council and the top management team in making decisions in an effective and efficient manner.

The following sections are organized as follows: first we will introduce the three pillars of the educational process quality assessment, and then we introduce the performance indicators that describe each pillar. After that we explain QAAS; the software system used to assure quality in AIU, and finally the data mining system with some of its reports showing how they are utilized in decision making.

2. The Three Pillars of Educational Process Quality Assessment

According to BS 7850 (1992) quality assurance is all those planned and systematic actions necessary to provide adequate confidence that a product or service will satisfy given requirements for quality. According to Juran & Gryna (1988), quality is fitness for use, which means the following two things: "(1) quality consists of those product features that meet the needs of the customers and thereby provide product satisfaction. (2) Quality consists of freedom from deficiencies". In higher education, the main customers are the students or their parents who pay for their education. These customers are interested in assuring that the content of the study programs fits for local, regional, and international markets in order to achieve the employability of the students after graduation, the lectures are delivered in a way that enhances students' learning capabilities, competences, skills, and knowledge, and that students are assessed in a way that helps them in fairly verifying their level of performance.

Hence, assuring quality in higher education system means: All planned and systematic actions required to provide adequate confidence in education quality, as a result of receiving the content that enhances employability, and receiving lectures in student-centered manner that enhances students' skills, knowledge, and competences, and finally assessed fairly (Alzoabi et al. 2008).

In order to achieve this we followed the following steps:

1. Identify goals of the study programs
2. Collect relevant data that helps in assessing the three pillars of higher education system.
3. Construct performance indicators that help in assessing every pillar.
4. Computerize data entry in order to achieve consistency, efficiency, and standardization.

In the following, we will see the performance indicators that help in assessing every pillar, and how QAAS and the data mining system will help in decision making using reports generated by these systems.

3. Quality Assurance Automated System: QAAS

The system was developed internally in the AIU in order to integrate the computerized academic system with quality assurance concepts to provide the management with a decision support system that helps in the effectiveness and efficiency of the decision making process.

The system allows instructors to do the following:

1. Entering the study plan of a specific course with every chapter to be covered in every class.
2. Entering the text book used throughout the course.
3. Entering all references- including electronic- that used for every chapter or topic to be covered in every chapter.
4. Entering the methodologies that will be used for every chapter.
5. Specifying the outcomes of the course.
6. Connecting chapters (topics) with every outcome.

In every class, the system allows the instructor to enter the following:

1. Attendance.
2. Chapter(s) to be covered in every class.
3. Methodology used: lecture, seminar, team work, field study etc.

The system then provides the academic management with many reports such as:

1. Plan completion (how many chapters covered divided by the chapters planned).
2. Punctuality of instructors (the time the attendance is taken is recorded).
3. Performance of the students in the subject as compared to their general performance in the previous semesters (to be explained in the following sections).
4. Absence ratios.
5. Performance indicator of the instructor in terms of student feedback.
6. Number of methodologies used throughout the course as compared to the planned ones.

4. Performance Indicators and QAAS

The performance indicators will be built upon several variables. We will describe these variables in the following:

4.1 Student Feedback

The student feedback is taken from an anonymous questionnaire that is published on the university's website at the end of every course. The questionnaire is built according to a statistical model that measures the overall satisfaction of the students with the three educational pillars. The results of student feedback are calculated by QAAS and all instructors get a mark between 0 and 100.

4.2 GPA Difference

The grade point average describes the student's performance ranking from 0 to 4 with 4 as the highest and reflects a mark between 95 and 100 out of 100. The accumulative GPA describes the students' performance in all subjects taken in all previous subjects. This variable is calculated as the difference between the mean AGPA (cumulative GPA) of all students before taking the subject and the mean of GPA in the measured subject, as in the following equation:

$$GPA_{diff} = avg(AGPA) - avg(GPA)$$

The higher the $GPA_{diff,}$ the better the performance is- i.e. delivery- of the instructor. However this indicator has to be taken in a cautious manner and balanced with external examiners reports, students' feedback, and other indicators, as instructors may be subjective and give high marks just for the sake of increasing their performance indicator. The GPA_{diff} has a maximum value of 4 that is when the students have the avg(AGP.A) as zero and the avg(GPA) as 4, and minimum value of -4 and that is when they have avg(AGPA) as 4 and the avg(GPA) as zero. This is of course hypothetical and almost impossible to happen.

4.3 Drop Ratio

The AIU academic system allows the students to withdraw from a subject one week after the beginning of the semester without losing money paid for the subject. It also allows the students to drop the course any time before the final exam without affecting the AGPA but with losing money paid for the subject. The percentage of students dropped the course is a good indicator of the performance of the instructor. That is, the lower the drop ratio the more attractive the classes are, and vice versa.

The drop ratio ranges from 0 (when none of students drop the course) to 100% (when all students taken the course dropped it).

4.4 Plan Completion

The plan completion refers to the number of chapters covered during the semester and divided by the number of chapters specified in the plan at the beginning of the semester and entered to the QAAS system. This indicator is important as it affects the student feedback. This is because students will like to have less workload. This variable helps in assessing the delivery pillar as it reflects the effort provided by the instructor and showing his/her commitment.

4.5 Industry Feedback

Industry feedback helps in assessing the suitability of content of the courses with the market needs. This is accomplished through a survey whose results are inserted to the QAAS system.

The outcomes of any course are inserted to the QAAS system at the beginning of the semester and hence the industry people will be able to see the content of the subject and its outcomes in terms of skills, knowledge, competences, and abilities of the students as well as the text books and references used in the course.

The university has a specified group of businesses who are granted access to specified courses in the QAAS according to the company's profile.

The rating of the industry feedback is ranked between 0 and 100 with 100 reflecting a high

correspondence between the content of the course and the market requirements.

As mentioned above, the results of the instructors feedback is entered to the QAAS system allowing a track of courses describing a set of competences related to a specific program to be evaluated and again to be provided as a mean result of all the subjects related to a specified track and ranked between o and 100.

This variable helps in assessing the content pillar as well as adding value to the process of improving the curriculum.

In addition to the GPA_{diff} which helps in a single course level, another indicator is helpful on faculty level that is the difference between the AGPA off all students as compared to their admission level. This helps in assessing the overall performance of the faculty bearing in mind that the target is improving or at least maintaining their level.

4.6 Second Examiners and External Examiners Reports

The second and external examiners role is to select samples of any exam papers and assess the following factors:

1. Comprehensiveness of the exam.
2. Clarity of questions and structure.
3. Fairness of the correction process.
4. Suitability of the level of complexity of the exam: simple, moderate, and difficult, with moderate as the target. This is an important variable that QAAS compares it with GPA difference so that the later variable is moderated accordingly.

5. Applying Data Mining to Enhance Education Quality

Using data mining for educational system is a new growing research discipline. Data mining can improve quality in higher education system. Most of higher education procedures such as assessment, evaluation, and counseling require knowledge. Knowledge can be extracted from huge educational data sets using data mining applications. Data mining applications can help both instructors and students to improve the quality of education.

Data mining techniques extract hidden patterns from huge educational data sets. The discovered hidden patterns enhance the procedures of decision making especially producing more advanced plans for directing students.

Data mining is the core component of educational business intelligence solution where pedagogic strategies can be experimented and evaluated. The main objective of using data mining in educational system is to improve learning (Romero and Ventura 2007).

Data mining techniques used for educational systems include (Romero and Ventura 2007, Shyamala and Rajagopalan 2006, Smith 2005): clustering, classification, association rules, decision trees, linear regression, and neural nets.

University lecturers and management can have deeply insights of the need of different groups of students by means of data mining methods especially clustering method (such as TwoStep and K-means) (Romero and Ventura 2007).

Data mining analysis can help also in better allocation of resources and staff, manage students' outcomes, and improve effectiveness of alumni development. Data mining results can be used for further step like adjustment of time table based on students' desires, this means that two courses with high association correlation don't overlap in the time table, also these two courses should be enrolled in the same semester, and not wait for one or two semesters.

These are some of the questions that can be answered and analyzed using data mining methods:

- How can data mining techniques be used to predict next semester GPA for each student?
- How can data mining techniques be used to identify the students likely to drop out?
- How can data mining techniques be used to help in providing counseling for students in timely manner?
- How can data mining techniques be used to identify students at risk of failures, in order to provide extra help?
- How can data mining techniques be used to classify students' results?
- Which students are taking the most "credit hours", or what are our profitable customers?
- What type of courses will attract more students?

Reponses to these questions can help improving educational quality by maximizing educational system efficiency, increasing student success and learning outcomes, and decreasing students drop outs (Shyamala and Rajagopalan 2006).

Web-based educational systems are new and emergent system. They replace the face-to-face communication, into virtual communications where the data collected by web servers is the only way to record the most of the learning behavior of the students (Romero and Ventura 2007.

6. Case Study: Arab International University

AIU is a new private university in Syria. Having 3 years old, the university began to find difficulties in managing the huge data deployed from its different information systems. The academic system, financial system, HR system, and QAAS system, are at the core of the university daily operations.

As most of the university information systems were provided from different sources, there was an urgent need to integrate data from all these sources into one data warehouse in a manner that could help the university in making use of all data to assure quality

The data gathered in order to produce the university data mining solution that can help in enhancing education quality are:

- Academic data (registration, examination, enrollment, etc).
- Financial data (student fees, staff salaries, orders, sales, etc.).
- Human Resources data (staff personal information).
- QAAS data (student feedback, GPA differences, drop ratio, plan completion, industry feedback, etc).

An enterprise data warehouse is built to hold all the previous data. The data sources were provided from different sources (Oracle databases, SQL Server data bases, excel). The data warehouse was built using Oracle 11g data base.

The ETL was built using ETL package in Oracle Warehouse Builder.

The solution developed (by AIU developers) followed ASD-DM paradigm for data mining solution. ASD-DM paradigm suggests three agile modeling steps (specifically Adaptive Software Development (ASD)) to developing successful data mining model (Figure 1) (Alnoukari et al. 2008).

The data mining solution could provide quality assurance officers as well as top management with reports to uncover trends in the students and instructors' performance in a manner that would be impossible or at least extremely tedious without data warehouses and data mining.

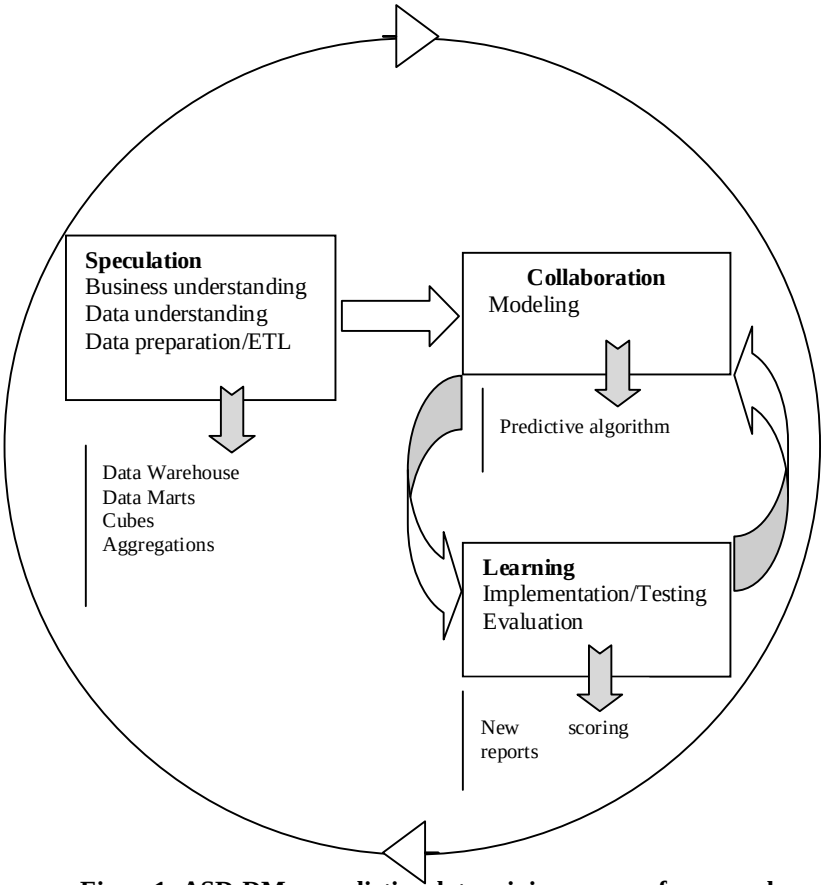

Figure1: ASD-DM a predictive data mining process framework based on ASD methodology (Alnoukari et al. 2008).

6.1. Examples of the AIU Data Mining System's Reports and how it was Helpful in Assuring Quality

In the following we list some of the reports that are generated by the system and how they are used by the AIU's Faculty of Business Administration:

1. Correlation of students' performance in different subjects. The system here tries to find some correlation between students' performance in different subjects and provides an indicator of how students should select the subjects for enrollment. For example, the system showed that some strong correlation (around $r^2 = 0.67$) exists between Business Ethics and Organizational Behavior despite the fact that none is a prerequisite for the other. So academic advisors were told to encourage students to enroll for Business Ethics after they have passed Organizational Behavior. This indicator promises a great aid when curriculum is to be redesigned later on.
2. Correlation of instructors' performance and countries of universities or specific universities they graduated from. This report is generated by combining data from HR system and QAAS. This report helped a lot in the recruitment process for new instructors.
3. Correlation of students' level of English and overall performance. This report helped a lot in evaluating the current system used in the English Language Center in the university and resulted in major changes in the system.

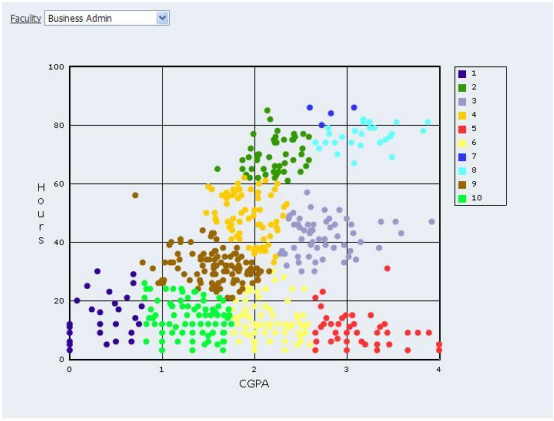

Figure 2: Business Administration Students clustered according to their accumulative GPA & their credit hours.

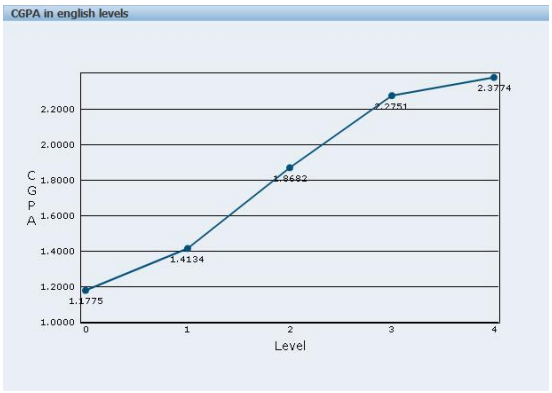

Figure 3: Business Administration Students clustered according to their accumulative GPA & their English level.

Figure 2 shows the correlation between student's accumulative GPA and their credit hours. It shows that students with higher accumulative hours are getting higher accumulative GPA. On the other hand it was very clear that there is strong correlation between students' level of English and accumulative GPA as shown in figure 3.

4. GPA_{diff} as related to various teaching aspects such as number of methodologies used in every course, number of chapters covered, absenteeism level, students' feedback, etc. These reports helped the deans in evaluating performance and finding out reasons for good or bad performance and to improve the overall performance of the teaching process.
5. The difference between admission level of students as explained in 4.5 helps in identifying the current position of any faculty. This when combined with the difference between students' accumulative GPA from semester to the other, and with the help of data mining prediction techniques can help in estimating the overall performance of a faculty for several following semesters.

7. Conclusion

In this paper we demonstrated a new approach of utilizing information systems in general, and data warehouses and data mining systems in specific, in improving the quality of the teaching process in a real-life case that is being used in the Arab International University. We also demonstrated some fairly new indicators that are used in the university, which could be used in any university after some customization according to the system adopted in the concerned

university.

Further investigation is going to be carried out in area of:

1. Building new performance indicators.
2. Building extra reports that can be generated by the system, such as alumni reports, and their reflection on the design of different courses, and student academic movement reports.
3. Putting the system into action at the faculty level with the appropriate security level to help the faculties taking the necessary correction actions.
4. Preparing the necessary QA procedures to control the work and documents.

References

Selmoune N., and Alimazighi Z., "A decisional tool for quality improvement in higher education", *International Conference on Information & Communication Technologies: From Theory to Applications, IEEE Conference*, Syria, 2008.

Fowler A., and Gilfillan M., "A Framework for Stackholder Integration in Higher Education Information Systems Projects", *Technology Analysis & Strategic Management*, Vol. 15, No, 4, December 2003.

Smith W., "Applying Data Mining to Scheduling Courses at a University", *Communication of the Association for Information Systems*, vol. 16, pp. 463-474, 2005.

Blackmur D., "Issues in higher education quality assurance", *Australian Journal of Public Administration*, pp. 105-116, June 2004.

Gatfield T., Barker M., and Graham P., "Measuring Student Quality Variables and the Implications for Management Practices in Higher Education Institutions: an Australian and International Student Perspective", *Journal of Higher Education Policy and Management*, Vol. 21, No. 2, 1999.

Spalter A. M., and Dam A. V., "Problems with using components in educational software", *Computers & Graphics* 27, pp. 329-337, 2003.

Tait A., "Quality Assurance in Higher Education: Selected Case Studies", *The Commonwealth of Learning*, Vancouver, 1997.

Juran, J. M., and Gryna F. M. (1988). "Juran's quality control handbook", *Mcgraw-Hill*, 1988.

Alnoukari M., Alhussan W., "Using Data Mining Techniques for Predicting Future Car market Demand", *International Conference on Information & Communication Technologies: From Theory to Applications, IEEE Conference*, Syria, 2008.

Alnoukari M., Alzoabi Z., and Saiid Hanna, "Using Applying Adaptive Software Development (ASD) Agile Modeling on Predictive Data Mining Applications: ASD-DM Methodology", *International Symposium on Information Technology, ITSIM 08*, Malaysia, 2008.

Alzoabi Z., Diko F., and Alnoukari M., "Enhancing Education Quality Assurance Using information Systems-QAAS System", *International Symposium on Information Technology, ITSIM 08*, Malaysia, 2008.

Shyamala K., and Rajagopalan S.P., "Data Mining for a Better Higher Educational System", *Information Technology Journal (5) 3*, pp. 560-564, 2006

Romero C., Ventura S., "Educational data mining: A survey from 1995 to 2005", *Expert Systems with Applications 33*, pp. 135–146, 2007.

European Higher Education available at: www.bologna-bergen2005. no/EN/Part_org/EU/050511_European_Commission.pdf.

Yang C., "Problems in Quality Assurance under Open Source Development Mode", 2006. available at: www.inf.fu-berlin.de/inst/ag-se/teaching/S-OpenSource-2006/YangC06_Quality_Assurance_Ausarbeitung.pdf.