Student’s Admission and Retention Data Analysis System with Data Visualization (SARDAS-DV) for Marketing

Eve Pardilla-Colmo ¹, Alice M. Lacorte, MCS ²

¹ Technological Institute of the Philippines, San Antonio, Laguna, Philippines.
² First Asia Institute of Technology and Humanities, Philippines.

colmo.eve@uphsl.edu.ph; amlacorte@firstasia.edu.ph

Abstract - A study entitles Students Admission and Retention Data Analysis System with Data Visualization (SARDAS-DV) for Marketing were conducted for effective data interpretation to help the academic administrators/heads, and marketing director come up with a strong and innovative marketing strategy. Moreover, SARDAS-DV has able to provide visualized reports in a form of graphs, charts and google maps that greatly helps the Academic Heads, Marketing Department as well as the management to improve their marketing strategies and continuous quality improvement of the services provided to students. Evolutionary Prototyping was used for the software development of the system, PHP as its programming language, and an open source Google Map API for the visualized reports. SARDAS-DV was evaluated by 383 respondents which is composed of students, academic heads, marketing personnel, admission officers, CS/IT faculty and ITS personnel using the adapted software evaluation form from the International Organization Standardization 9126. The system obtained an overall weighted mean of 4.5 with a descriptive meaning of Strong Agree which denotes that SARDAS-DV was accepted for use by the institution’s marketing department to improve quality service and decision support for academic heads and administrators.

Keywords - Students Admission; Students Retention; Data Analytics; Data Visualization; Data Mining; Educational Data Mining.

I. INTRODUCTION

Students’ admission and retention to an academic institution plays a significant role in the entire enrollment process. It is likewise considered as a privilege granted to those students who are qualified under certain criteria set by an academic institution. It enables students to comply with the necessary requirements for admission to the field of study or degree they are going to pursue. Requirements and documents submitted by students undergo a thorough evaluation by the Admission Officer to validate its authenticity. Those data of admission or retention of students are useful if the management would undergo an in-depth study on it. It may be used as a reference and guide for them in making decisions in the future, address the current needs of the students as and most of all to develop a strong marketing strategy to ensure retention and capture admission of prospective entrants.

For instance, all universities located in Australia, in the year 2013, there were approximately 500,000 individuals pursue their education degree in college and is continuing to grow as years pass by [1,2,3] state that the need for a college degree is driven by the high labor requirement in relation with the knowledge economy, thus college education is considered difficult link into the knowledge society.

At present, education institutions pile up and store massive volumes of data like student admission and record of attendance as well as attendance results. According to mining these data sets will able to serve its handler by producing stimulating information.

However, handling and analyzing of student's admission and retention data and records for one big academic institution is difficult job for the part of the marketing department having limited number of personnel. Students' data of admission and retention were not given high value since academic institution tends to focus only on students personal information like, personal profile, address, educational background, parent/guardian profile and documents submitted. They forget to collect data on the reason why a new student tends to enroll in an institution and why old students prefer to re-enroll, over other academic institutions. In addition, for the past 40 years, the University of Perpetual Help System Laguna has a remarkable large number of students admitted for college, thus large amount of students' data also arises. In relation to this, [4] state that because of the voluminous number of data, it brings a difficulty to the marketing director and personnel in performing efficient data analysis using the current traditional method. Therefore, it is necessary to have a better data analysis of student admission and retention records as well as data gathered from prospected entrants to present a better data-driven which can be used by the upper management in the decision-making of what appropriate marketing strategy and approach they will have to develop.

This current situation encourages the proponent to develop a system that helps the marketing department to have a better presentation of students’ admission and retention and prospected student entrants’ data. The data will be later on forwarded to the different academic heads for analysis to develop better and strong marketing strategies for the institution and for quality improvement.

DOI 10.5013/IJSSST.a.20.S2.16 16.1 ISSN: 1473-804x online, 1473-8031 print
Hence, this work is necessary for academic planning, monitoring, decision-making and continuous improvement and development of effective marketing strategies that will attract possible enrollees to help increase enrollment population of the institution.

Additionally, this paper likewise discuss related literature and studies that presents related works to give the readers wide understanding and background on the what the paper is in to; methodology that presents the activities, methods which the author undertake in developing and conducting the study; results and discussion that clearly presents and discusses the result of the study and a conclusion which the author was able to drawn after conducting the study.

II. RELATED LITERATURE

A. Students Admission and Retention

Admission policies and guidelines to an academic institution differ from one institution to the other. It’s the institutions prerogative to come up with their own admission policy. Student admission is important to an educational institution. Like for instance, educational institutions in Thailand believes that students admission is very crucial because of education management, planning of budget for administration, institution and education management is directly affected because of the standard educational quality indicator [4]. Relatively[5] assert that the efficient student admission of students as well as nurturing them for the entire period they finish their study in high quality under a specified timeframe are therefore what the institution take into consideration and importance.

However, according to [6;7] engaging, attracting and retaining students has become a major challenge for academic institutions since the attrition of students specifically undergraduate engineering programs became a global issue. Financial constraints, unclear career directions or changing academic interests were some of the possible reason for attrition. Further, inconsistent or poor performance might also affect the motivation of students to pursue a particular degree or program [6]. Additionally, students' attrition can influence both an individual and the private or public institutions in terms of fund allocations, career opportunities, and capacity building.

On the other hand, there has been an increased interest in using nontraditional mechanism before being allowed for admission into an educational institution, for example, the use of personality and motivation questionnaires, test for trial-studying, and biodata [8;9;10]. Nowadays, the development of content-valid methods for selection and prediction in an academic institution has become a subject of interest. For instance, European countries, there is a trend towards students’ selection based on admission examinations that depict correspondence to specific program content.

HEI’s or Higher Educational Institutions nowadays set their interest in the paths of students as well as its alumni, thus finding if who among the students will enroll in a certain program or course and which of them calls for a series of debates and encouragements [11] Nowadays, [11] states that one of the leading challenges that HEI’s are facing is the volatile increase of educational data and how to utilize these data to improve future decisions of an upper management.

On the other hand, at present, we all know that the advancement of Information Technology has overlaid to produce a huge amount of databases and large data in a different area [12]. [13] assert that due to this scenario and current trends, a lot of different researches in relation to databases and the field of information technology to approach, to gather, store and use important data for future use. Additionally, [14] define data mining as the representation of processes which was being developed to scrutinize voluminous amount of regularly gathered data usually come from different areas and sector like marketing, communication, health etc.. Moreover, it also refers to a set of tools applied to execute the process. Relative to this, according to [15], data mining is the extraction of concealed information from big databases; it is an effective technology with a remarkable potential to assist organizations concentration on the most essential information in the data warehousing. The objective of data mining is to extract facts from a dataset in the form of human-understandable composition [16].

Thus, data mining which is an analytical tool can be used to dig up significant information from these large volumes of data sets. Additionally, [17] state that data mining is a technique of extracting implicit data and embedded set of information from the mass, raw, fuzzy, noisy and unsystematic data as well as understanding it well to be used in the future.

[18] identified the following advantages of data mining used in different applications like Banking, healthcare, marketing, manufacturing, and production, etc.

A1. Applications:

First, in banking application, data mining helps banking sector in the course of looking at a large databases to uncover the prior unfamiliar patterns; computerize the manner of finding analytical information. Additionally, it helps to predict levels of bad loans and deceitful credit cards utilization, forecasting credit card spending of new clients and forecasting the type of client best reply to the bank’s offer of new loans. Similarly, for banks and insurance companies, one of the success factors is the creditor’s worthiness detected and assessed in advance through data mining method which is being adopted to address credit scoring problem [19;20]. Moreover, according to [21], through data mining, banks, and other financial institutions
could detect fraudulent transactions for the purpose of the owner's protection.

Second, in Manufacturing and production applications, data mining improves prediction of machine errors/failures and seek key attributes that control manufacturing capacity optimization and ascertain control restrictions.

Third, in Marketing applications, data mining helps marketing sector through classification of the demographic of costumers that can be utilized to predict which among the costumers will buy the product or respond to mailing and is very much effective in the growth of a business. In relation to this, through data mining, consumers information are gathered by retailers, information of transactions made, and information of products that is important to improve the accuracy of forecasting product demands, the recommendation of a product, assortment optimization and ranking between manufacturers and retailers [22;23].

Fourth, Health-Care applications, data mining is vital in the healthcare sector by demographic correlation of patients having a critical disease, to develop effective insights in illness indicators/signs or symptoms and their causes and discovering proper treatments.

Fifth, in Insurance applications, data mining helps this sector in the prediction of fraudulent medical coverage and claims, classification of major factors that can influence

Sixth, in Law applications, data mining helps monitor criminal's behavioral patterns. It likewise helps track the pattern of crimes, behaviors, and locations of criminals, identifies different data mining attributes and assist in resolving various cases of crimes.

Seventh, in Government and Defense applications, data mining is used to calculate different military equipment’s cost and predict resource utilization and spending. Likewise, it helps potential military engagements strategy testing and homeland security improvement through data mining from various sources.

Eighth, in Brokerage and Securities Trading applications, data mining helps in the bond prices change prediction and predicting of stock range fluctuation in order to determine when to purchase to sell stocks.

Ninth, in Computer Hardware and Software applications, prediction of possible security defiance and disk –failures is possible in data mining.

Lastly, in Airlines applications, data mining supports the feasibility checking of additional routes in order to boost business profit and to reduce loss by capturing information of the possible ultimate passengers’ destination.

A2. Disadvantages of Data Mining:

In contrast, several authors has able to identified the disadvantages that data mining has brought.

First disadvantage is the Privacy Issues, in recent years, issue on personal privacy has increased incredibly due to the surge of internet. This issue is the man concern of most individual such as employees, customers and internet users because they are anxious that there is a possibility that unknown person may have see their personal data and later use those data in unethical means that they may be harmed. Though users are protected by several laws to trade personal data among two or more different organizations, yet selling it is still a practice at present [24;17].

Second, [24; 17] cited that another major disadvantage of data mining is the security issue especially in the field of information technology. Establishments, institutions and companies have a large set of personal information of the customers and employees like birthdate, social security number, payroll, etc. and it is as well obtainable online. The reality is, companies have no enough security systems available to safeguard this information. As observed and heard over the news, there are plenty of cases wherein hackers stole and access costumer’s personal data.

Third is the Misuse of information /inaccurate information. One example is the trend obtained from data mining which is intended for business or ethical purpose use. However, there is a tendency to misused it for other unethical intention. Unethical person or business will possibly use the information to benefit from weak people or discriminate against a group of individuals. Apart from that, techniques in data mining are not perfect in accuracy, thus errors may happen which lead to a grave consequence.

[25] pointed that there were a numerous challenges faced by data mining like scalability, heterogeneous and complex data, Network Settings, Ownership and Distribution of Data, Quality of Data, Dimensionality, Preservation of Privacy and Data Streaming.

B. Data Analytics, Data Visualization and Data Mining

Doggied up information from a large datasets can be better understood if analyzed thoroughly, hence, data analytics has began. It is a combination of computer programming, statistics and operational research. It is a tool commonly being utilized nowadays to alleviate process for decision-making by means of converting unprocessed precious data into a meaningful record of information [13]

At present, traditional reporting contain statistical data in numerical form. In this type of reporting, when a large number of data are involved, traditional reporting found to be a burden to be interpreted by human beings. Due to this situation, it is necessary that reports should be presented in way that human can easily recognize and understand it just by looking into them [26].

Thus, to address the problem of data reporting and analysis in traditional way, data analytics and visualization were introduced. [11] state that data analytics and data visualization are employed to emphasize useful information and decision making support. For instance, an educational environment, it could improve educators and academic administrators to analyze student’s course interest and activity as well as the usage of data to get student’s learning general view. In data analytics, information visualization
and statistics are the two widely used techniques. Visualization is a technique that uses graphics in order to help people easily analyze and understand data. Several studies were conducted which is oriented about visualizing various types of educational data like annual, seasonal, daily, and hourly patterns of user behaviors on online debate. More examples are statistical graphs to analyze assignment complements, admitted questions, score in examination, tracking data of students to analyze attendance, quizzes and assignments results, student and groups’ weekly activity information.

On the other hand, [11] defined statistics as a mathematical science that deals with the analysis, collection, interpretation or explanation and data presentation. Nowadays, it is easy to acquire basic descriptive type of statistics from statistical software like SPSS. In addition, statistical analysis is very helpful to get reports assessing how long does it takes for a student to finish his/her classroom task, number of resolved problems, and his correct percentage compare to our calculation about his performance level and score.

C. Educational Data Mining (EDM)

Education is one of the many sectors wherein data growth have been dramatically increasing as years passed by [27;28] . Presently, academic/educational institutions store and compile huge volume of data like, students’ attendance records and enrolment, enrolment results, etc. Mining those data sets yields information stimulation that will serves its handler efficiently and conveniently. To address this issue, Educational Data Mining emerge [29]. EDM focus on data analysis in an educational setting through desperate system that aims to improve institutional effectiveness and learning experience by developing models.

Educational Data Mining (EDM) is an interdisciplinary field of study that applies machine-learning, statistics, Data Mining (DM), retrieval of information, psycho-pedagogy, cognitive psychology and a method and techniques of recommender system using different educational data set to address and solve issues in the field of education [30].

[27] argues that Educational Data Mining as well as learning analytics had became the lingua franca for improvement seeking institutions in terms or strategy and decision-making operational abilities.

III. METHODOLOGY

In order to attain a successful software development, this project adopted the Evolutionary Prototyping Model by Bastanlar Y. [31] defines Evolutionary prototyping as the method the development team or the developer initially creates prototype. Subsequent prototypes are constructed after receiving response and comments from the end-user/customer, each with improvements, additional features and functionalities until the final system emerges. Additionally, evolutionary prototyping is alike to incremental development wherein parts of the software possibly be inspected or delivered to the end-user during the software life cycle model.

Moreover, evolutionary Prototyping is also known as breadboard prototyping which begins with the construction of actual functional but with minimal functionalities prototypes which later on become the heart of the future prototypes [31]. In addition, using this model, requirements that are well understood are included in the initial prototypes and the requirements that are not well understood are added as and when they are understood.

Data on student admission and retention in an academic institution were gathered through online surveys wherein they are asked to answer a survey. For data analysis on student retention, they are ask what are the factors and reasons of their continuous enrollment. Factors given are: employee’s dependent, scholarship grant, location’s accessibility, tuition fees and other fees, up-to-date facilities, board performance, quality instruction, autonomous status, community involvement activities, linkages and partners and employability rate. On the other hand, students admission data were analyze on how they get to know the academic institution with given different marketing strategies as follows: referred by friend or relatives, career talk/orientation, house to house marketing campaign, marketing ads/ tarpaulin posting/ fyers, website/ social media platforms, newspaper and publications, television ads/commercials, hosting on various invitational academic and non-academic competitions, academic and sports competition.

[14] likewise mentioned that software prototyping applications and methods are best use in the situations likes; systems that requires high level of interaction from users; online systems applications, systems that uses a lot of screens, and systems and applications wherein users need to fill out forms like online surveys.
Figure 1 shows the use-cases for Administrator, Students and Admission Officer/Staff. It was presented that the administrator, student and admission officer/staff logs-in and logs-out of the system. Both the administrator and the student can access the online survey for admission and retention. Also, both Administrator and Admission Officer/Staff can view the list of students who finished the online survey in order to issue clearance with the students. Moreover, the administrators will be the one to manage the account of the users, edit/modify survey instrument, can view survey results for individual programs and all programs. Lastly, the Administrator could generate and print survey results per program and all programs.

Figure 2 presents the use-cases for both Marketing Director and Academic Heads/Dean. It is clearly shown that both Marketing Director and Academic Heads/Deans can Logs-in and Logs-out of the system as well as can generate and print reports per program. Additionally, the Marketing Director can view survey results for all programs and can generate and print reports as well. On the other hand, Academic Heads and Deans will only have an access to generate and print reports of the program/s under their college.
IV. RESULTS AND DISCUSSION

Students Admission and Retention Data Analysis System with Data Visualization (SARDAS-DV) for Marketing was designed and developed to help the academic administrators/heads and marketing director to effectively interpret data in order to come up with a strong and innovative marketing strategy.

The developed system contains four (4) modules namely: users, survey, settings, and analytics. These modules are necessary in order to provide data and reports needed by academics heads and marketing director in decision-making in the future.

User module is the module wherein the system administrator manages accounts of different users. This module has four sub-modules namely: Employee, Student wherein the administrator can either add or delete a certain user while Prospected sub-module is a special module for prospective students who want to enroll in the institution. On the other hand, the sub-module Marketing Staff is a module where the administrator can be able to manage each user's access.

Survey Module allows the system administrator to create and publish surveys for students and create and add survey questions and categories. Additionally, in this module the student could be able to perform online surveys and submit once accomplished (see figure 3). Student’s answers to online surveys will be needed to generate visualized reports to help the academic administrators and marketing director easily interpret data.

Settings module allows the administrator to manage data in relation to admission and retention like: add course, add academic year, etc. While the analytics module provides comprehensive visualized reports on students admission and retention at UPHSL in the form of graphs and maps.

In the analytics module, SARDAS-DV has been able to provide map visualization using Google Map API to determine where the students came from represented by dots (see figure 4). Relatively, students enrolled were traced where they usually reside group per course, with corresponding color code assignment per course/program (see figure 5). The more colorful the area means a large number of individuals are coming from a certain area. This mechanism helps the academic heads as well as the
marketing department where and what area they will strengthen their marketing campaign to capture possible enrollees.

On the other hand, SARDAS-DV produces reports on the reason and factors why students enrolled in a form of bar graphs and pie charts (see figure 6). The categories are as follows: a) employees’ dependent/educational benefit; b) scholarship grant; c) location’s accessibility; d) tuition fee and miscellaneous fees; e) up-to-date facilities; f) board performance; g) quality instruction/competent faculty members; h) autonomous status recognition; i) community involvement and extension activities; j) linkages and partners; k) employability rate.

Given the different categories with corresponding assigned color code, it is now easy for the marketing department as well as the different academic heads/deans to determine what marketing strategy they need to improve and strengthen the system to make sure that the students would retain their admission at UPHSL.

Moreover, this visualized reports greatly help the concerned departments to easily analyze the data as compared to the typical reports provided by the marketing department which is presented in numerical form.
Relatively, SARDAS-DV produces reports on how students get to know the institution in a form of bar graphs and pie charts (see figure 7). The categories are as follows: a.) referred by friends or relatives/word-of-mouth; b) career orientation/talk; c) house-to-house marketing campaign; d) marketing ads/tarpaulin postings/flyers/banners; e) website/social media platforms; f) newspapers/publications; g) television ads/commercials; h) hosting on various invitational academic and non-academic competitions; i) academic and sports competitions.

![Figure 7. Sample Screen Shot of Graphical Report on How Students Get to Know UPHSL/Marketing Strategy Used](image)

Given the different categories with corresponding assigned color code, it is now easy for the marketing department as well as the different academic heads/deans to determine what method of marketing they will use to strengthen and popularize the institution. Moreover, this visualized reports greatly help the concerned departments to easily analyze the data as compared to the typical reports provided by the marketing department which are presented in numerical form.

A. Software Evaluation

The Evaluation of SARDAS-DV was conducted by the proponent using the ISO 9126 Software Evaluation Standard. Two groups of evaluators were asked to evaluate the conformity; End-User Evaluators evaluated the functionality aspect of SARDAS-DV while the Technical Evaluators evaluated both the functionality and technical aspects of the system. End-User Evaluators were composed of selected students, academic heads, marketing personnel, and admission officer with a total number of 376 respondents. On the other hand, Technical Evaluators were composed of Computer Science/Information Technology Faculty and ITS Employee making a total of seven (7) respondents.

<table>
<thead>
<tr>
<th>TABLE I. OVERALL RESULT OF SYSTEM EVALUATION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INDICATORS</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Functionality</td>
</tr>
<tr>
<td>Reliability</td>
</tr>
<tr>
<td>Usability</td>
</tr>
<tr>
<td>Efficiency</td>
</tr>
<tr>
<td>Maintainability</td>
</tr>
<tr>
<td>Portability</td>
</tr>
<tr>
<td>Overall Mean</td>
</tr>
</tbody>
</table>

Table I comprehensively presents the overall summary of the evaluation result for both technical and non-technical evaluation. As shown in the table, Usability characteristic obtained the highest weighted mean which was 4.62 with a descriptive meaning of STRONGLY AGREE. This denotes that both technical and non-technical evaluators strongly agreed that the system was able to provide attributes that bear on the effort needed for use and on the individual assessment of such use by technical users/evaluators. Relatively, respondents agreed that SARDAS-DV has interface and functions that are easy to understand, easy to learn and operate in a different environment.
On the other hand, functionality and portability characteristics both obtained a descriptive meaning of STRONGLY AGREE with an average weighted mean of 4.59 and 4.56 respectively. This data denotes that both technical and non-technical evaluators strongly agreed that SARDAS –DV found to be complete and able to provide functions needed by the user as well as maintain its level of performance understated period of time. In addition, both of them likewise agreed that the system has attributes that bear on the relationship of level of performance of the software and the number of resources used. Also, SARDAS-DV has the characteristics that bear on the efforts needed to make specific modifications and has the ability to easily transfer from one environment to the other.

Apparently, reliability, efficiency, and maintainability obtained a descriptive meaning of AGREE with an average weighted mean of 4.46, 4.41 and 4.39 respectively. This means that both technical and non-technical respondents agreed that SARDAS-DV has attributes that bear on the capability of software to maintain its level of performance under stated conditions for a specific period of time. Moreover, the system has attributes that bear on the relationship between the level of its performance and the number of resources used, under stated conditioned as agreed by both the technical and non-technical evaluators. In addition, both evaluators also agreed that SARDAS-DV obtains attributes that bear on the efforts needed to make specified modifications.

In general, SARDAS-DV obtained an overall weighted mean of 4.5 as evaluated by both technical and non-technical evaluators. It has a descriptive meaning of STRONG AGREE which denotes that the evaluators strongly agreed that the system was able to conform with the ISO 9126 International Software Quality Standard. Thus, the system can be used by the UPHSL marketing department to improve the quality of service, and would greatly help not only the marketing department but also the different academic heads for them to be able to come up with strong and effective marketing strategies to maintain the high number of enrollees in the institution.

Moreover, this project plays a significant role for:

a. Students – This project allows the students express their personal thought through online survey regarding the institutions services and marketing campaign strategy.

b. Marketing Director/Staff: This project plays an important role for marketing director and staff because it allows them to immediately generate reports and necessary data in terms of students’ admission and retention.

c. Admission Staff/Personnel: This project will be of great help to the Admission Officer / Staff because this software enables them to view comprehensive report of students list who have done the online survey instead of looking for the accomplished survey form before issuing clearance every end of semester. It will ease their burden of time-consuming effort in locating those records.

d. School Administrator/Academic Heads: This work enables the academic heads to have a visual presentation (graphical and map) of the students’ admission and retention in an academic institution. Moreover, those visualize data are vital for them that serve concrete guide to come up with good and correct decisions for continuous quality improvement in terms of service to students which ensures retention.

e. Proponent/Researcher: This work enhance the skills and knowledge of the proponent/researcher in system developing and designing which is important for an IT professional.

f. Future Proponents /Researchers: This work will be useful for future proponents/ researchers in which they can use this as reference for a related topic/ area.

V. CONCLUSIONS

Based on the conducted study entitled Students Admission and Retention Data Analysis System with Data Visualization (SARDAS-DV) for Marketing, it was concluded that it meets all the requirements as suggested and recommended by the adopted academic institution, was able to provide visualized reports in the form of graphs and charts as well as Google maps which greatly help the academic heads and marketing director easy interpretation of data. Moreover, SARDAS-DV has able to allow students to perform and answer online survey which greatly helps the marketing department easy generation of data. In addition in terms of conformity the Software Quality standards, SARDAS –DV has able to meet the minimum requirement as indicated in the ISO 9126 Software Quality Standards. Lastly, the developed system was acceptable to the end users and was able to perform all the stated function in the objectives of this study.

After the successful development and designing of Students Admission and Retention Data Analysis System with Data Visualization (SARDAS-DV) software, the proponent has encountered some issues and concerns in which she tends to resolve in other related future works as follows:

a. To include data on feeder schools, so that the institution will be able to determine from what secondary school students of UPHSL usually come from;

b. To design an additional module that data of student admission and retention for Basic Education, Secondary Education and Senior/Junior High School will be gathered.

REFERENCES


EVE PARDILLA-COLMO et al: STUDENT’S ADMISSION AND RETENTION DATA ANALYSIS SYSTEM WITH . . .


