ANALYSIS OF UTILIZATION OF LIBRARY RESOURCES USING BIBLIOMINING TECHNIQUE EMPLOYING DESCRIPTIVE ANALYTICS THROUGH KDD MODEL

[1]MARICEL LANDICHO MALABANAN, [2]JASMIN DE CASTRO NIGUIDULA

[1,2]Technological Institute of the Philippines, Manila, Philippines
E-mail: [1]maricellmalabanan@gmail.com, [2]jasniguidula@yahoo.com

Abstract - Library marks its place to be beneficial to every individual than any other sectors in the community. Library is like offering faith, hope and optimistic view of life that do not fail grasp by unseen possibilities [1]. However, support for life-long learning with its offered services needs substantial assistance from its people. Substantiation showing proof for the expenditures hike has been quite high and its burden are still underlying on the institutions' librarians. This apprehension is multipart by the growing numbers of unutilized services and accustomed users of library. The paper illustrates the utilization of library resources and its significance into perspective challenging librarians and educators to infuse the utilization of library resources. Produced patterns captured during evaluation were evaluated using bibliomining technique in the form of descriptive data analytics through the process of KDD Model. As the leading contributor to the education system, it was analyzed that still, teachers are the most users of library resources in terms of book borrowing. Produced results of the tests will then be presented to the school librarian for future library related decision making. Some recommendations were made and emphasize the use of data mining tool to help librarians of a faster way of analyzing reports which gives importance to every bit of raw data that was gathered and save into their library database systems.

Keywords - Data Mining, Bibliomining, Descriptive Analytics, Library Resources, Kdd Model

I. INTRODUCTION

According to Merriam Webster, a place where people can borrow or use the book and other related materials like digital resource for learning is no other than library which is free and accessible to anyone.

In support to Andrew Carnegie's statement relative to the library academic setting, it is an asset of every institution which can contribute to a higher education's transformation. As a physical entity, campus library is typically perceived as the exclusive ground of the on-campus learner. While library resources are professed as a free amenity because the valve of information is always there, the truth is, it costs gigantic for subscriptions to the electronic resources and reading collections that support learning and research. More so, venture of library resources cost too much because investing in libraries is an investment in education and lifelong learning [1].

In addition, having this expensive venture, support for discussion and documents for library resources expenditures must be well presented in terms of overall aspects of library systems that can significantly contribute positive outcomes leading to a large influence on the academic institutions such as student's perseverance and determination [2]. Above all, it is said that no matter what type of academic setup of library an institution has, there is no way to erase its undying purpose. That undying purpose which became the core of an academic setting continues to be an emblematic state of the institution. Unlike on the departmental setting, having all the support and services to community, undertakings occur in this passionate place of the school, the heart of the institution, can be more attainable and be successful because of this emblematic state than any other state-of-the-art facilities located inside the university campus [3]. Hence, with proper utilization of its available resources, libraries payback to institution in plentiful ways: (1) help in realizing user's perspectives of definition of success; (2) encourage all users' community to take part in the evaluation procedure and provide better services; (3) provide modified services and means to users' community; (4) help escalate student enrollment, accomplishment, success, graduation rates, learning and student commitment; (5) help improve faculty teaching and research, thereby increasing publications; (6) help increase institution ranking, accreditation status and help to receive sponsors and grant for projects [4].

However, based on research, college students do not possess positive behavior on information seeking habit. This information gave negative feedback to librarians given the fact that the school uses huge cost to make the library be conducive to anyone. Since scholars exhibited seldom visit to library, utilization of resources was not maximized as well [5]. Some factors of declining number of academic library usage in different criteria of library management like services offered, quality assessment, collection analysis, information seeking habit are the following: (1) internet access, (2) lack of awareness, (3) image of the librarian and (4) mobile reference services [1].

Having these researches of declining number of academic library usage and availment of its services, the senior librarian and/or staffs of the institution conduct orientation every opening of the semester to promote and encourage all employees and students of the library utilization.

The study aims to determine the utilization of library resources specifically on determining the most number of borrowed books, frequent book borrowers, and authors of those borrowed books. This study is not limited to accession of multimedia, kindle and online library resources like e-books as part of its online subscription in the utilization of library resources. The report then will be presented through data analytics thus providing recommendations for the improvement of library collection that can attract more users and visits of the library. The researchers also want to emphasize the use of data mining tools in dealing with large volume of data for better decision making.

II. LITERATURE REVIEW

Looking at the importance of library, through research, it is evident that there is an increasing number of students who do not use academic library.
This decreasing inclination of non-users has been observed globally. Factors that are looking into some consideration for this phenomenon are the sources of alternatives like electronic resources which is more easily accessible and convenient to students rather than go its physical setup [5].

Relative to the decreasing inclination of library non-users, library leaders can have the opportunity to seek for such data producing reasons of this instances and conduct study for this occurrence that could help them in providing solutions in library-related matters for better decision-making processes. Unluckily least number of librarians are using this opportunity to further deepen the study as tool for linkage to new solution to the problem for their universities. One way to start with the study is the advancement to applying tools for statistics and by producing valuable data in relation to library services usage that is not limited to library systems though the application of data mining techniques [6].
In addition, there is also an option to extend the study in application to library services, collection and resources most especially to the student who do not went to library regularly given the fact they are being taught on line or off campus learning set up with state-of-the-art instruction systems. Analysis in this off-campus setup emphasized the embedment of the requirement in the teaching plan of study that can result to a more concentrative learning and relationship to academic setting because of this information seeking requirement [7].
Alarming number of non-users of library continually increases therefore, it is needed to have an interminable review and assessment of library services offered as to maintain its emblem as the heart of every institution. The review and assessment should always be in accordance to the light of user's perceptions and satisfaction [5].

III. METHODOLOGY

To look for the pattern from the data gathered, this will be mined and tested using Tableau Desktop – a software for data set testing tool. Results will be produced with the help of data mining through the aid of bibliomining technique.
Collecting data from a large database then gone through the processes of data mining of library processes in accordance to its standards is called bibliomining. This data mining processes consists of several steps that are chunked into smaller parts to get the final output that is known to be called as Knowledge Discovery in Databases or KDD [8].
Those smaller chunks of processes cut from a large database of bibliomining technique were identified. This process covers the detailed stages of KDD Model [9].
The fusing goal of the KDD process is to excerpt knowledge from data in the context of large databases. Application of KDD in library atmosphere consist of the following: (1) identify topic; (2) create data warehouse; (3) refine data; (4) explore data; (5) evaluate results; and (6) report and implement [9].

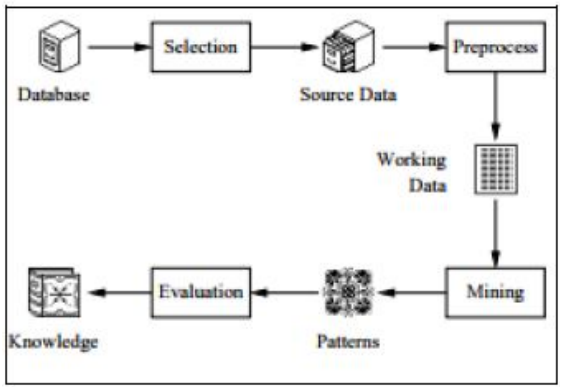


Figure 1. The KDD process model adapted by Williams and Huang, (1996) [10]

Showing the model in figure 1 which is adapted by Williams and Huang are the detailed processes of KDD Model in library systems as follows:

1) **Identify topic.** This is where data gathering takes place. Identified topic is the utilization of library resources.
2) **Create data warehouse.** To analyze the identified topic, data warehouse is needed. This is where raw data are being collected from a large database.
3) **Refine data.** This process increases awareness and understanding of data set as data cleansing take place.
4) **Explore data.** Refined data are now set to be analyzed. Series of test are made to produce needed output for the identified topic.
5) **Evaluate results.** To test the significance of explored data to the intended output, produced results should be evaluated.

6) **Report and implement.** Last process is the presentation and execution of final results.

To support the detailed processes of KDD Model in the light of bibliomining technique, researchers use descriptive bibliomining and descriptive analysis or statistics in the form graph representations based on the tested results pertaining to the current situation. Descriptive bibliomining pursues to pronounce current situation. While descriptive bibliomining focuses on the current situation, descriptive analysis uses the previous data to be used in describing the present situation of tested results also in the form graph representation as well. These two techniques are being combined producing one result only. Descriptive analytics are beneficial to be used because it permits us to learn from past behaviors and understand how they might impact future outcomes [10].

Department/Course	SY 2016	SY 2017
Grade 11	650	648
Grade 12	15	15
BSP	70	69
BSE/BEED	151	154
BPE-SPE	34	33
BSN	18	19
BSM	30	30
BAC	64	62
BSA	98	76
BSBA	103	104
BSHRM	40	40
BST	73	63
BSECE	106	101
BSIE	168	170
BSCS	22	20
BSIT	156	151
BSCPE	31	28
BSCRIM	48	45
ABMMA	7	7
BSACT	83	99
FACULTY	182	182
ADMIN	139	139
Total Population	**2287**	**2255**

Table 1: Population breakdown of the used dataset.

Based on the listed above population, table 1 shows the basis of data set used of this study which came from one academic institution.

IV. RESULTS AND DISCUSSION

The following figures below are the results that have been produced throughout the series of testing of the data set. Researchers used various types of graphs for a more graphical representation of different tested and evaluated results. To be able to have a well-presented graph, produced results of the following displayed graphs gone through data cleansing first since thousands of data was collected based on the intended output of each analysis.

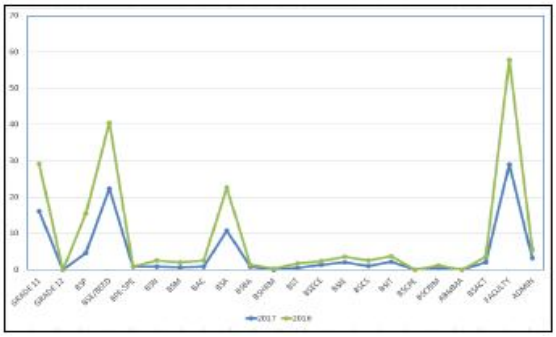


Figure 2. Statistics of library book borrowers for A.Y. 2016-2017

Figure 2 in the form of line graph shows the statistics of book borrowers from different units of the institution. It is observed that based on the borrowing percentage, it was from faculty department with 29%, education department with 22%, grade 11 with 16%, accountancy department with 11% and psychology department with 5% are the consistent and most frequent book borrowers of the library while Grade 12, admin, CS and IT, engineering, criminology, HRM and Tourism, Nursing and ABMMA are the departments ranging from 0-4% are the departments who come the least in book borrowing.

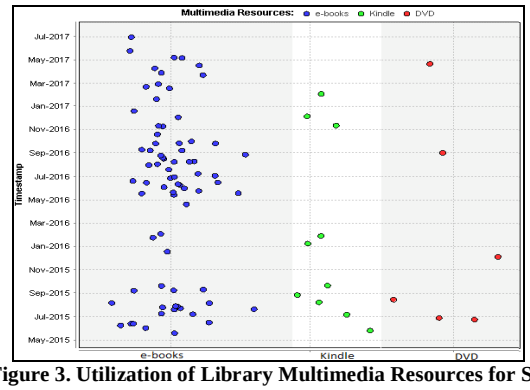


Figure 3. Utilization of Library Multimedia Resources for SY 2015-2017

In Figure3, it shows the utilization of multimedia resources of the institution. It shows that the most utilized multimedia resources are e-books by 3.28%, followed by kindle which utilized by 0.44%, and DVD as the most little utilized by 0.26%. Scatter graph also shows that the months of May to November of the year 2016 is the peak season of e-books utilization as the most utilized multimedia resource.

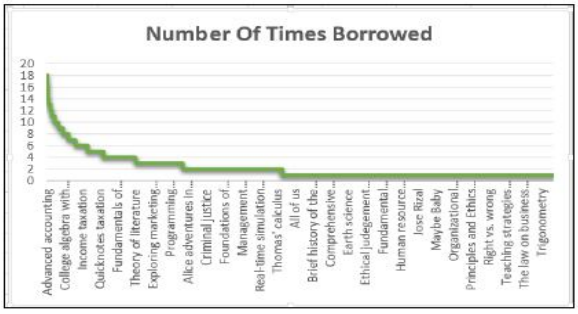


Figure 4. Book utilization for SY 2016-2017

Figure 4 shows in the form of line graph that the most borrowed books are advanced accounting which was borrowed 0.80%, followed by college algebra that was borrowed 0.71%, and income taxation by 0.27%. List belongs to bottom three are teaching strategies, the law on business, and trigonometry books placed at the bottom as the least frequency which were borrowed once or twice only. Researches chose to present data from top one down to top twenty of most utilized books.

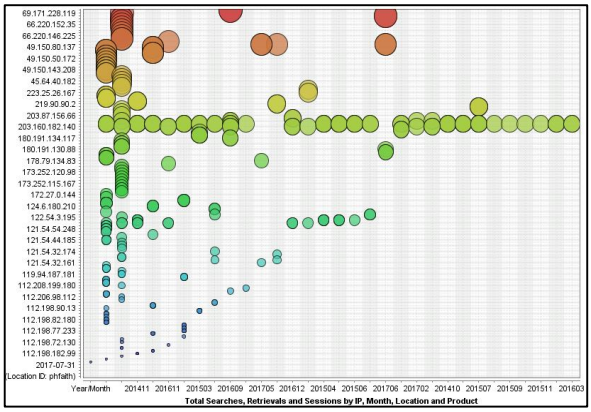


Figure 5. Utilization of library online resources on its subscription basis from A.Y. 2014-2017

In figure 5, bubble graph shows the utilization of online resources based from a user location tracked down through IP addresses, total searches, retrieval and sessions of different products based from the library online subscription. Color coding in the graph which consists of four different colors namely red, yellow, green and blue means different names or different products of the library online resources subscription. Graph also shows the frequency of users of different types of resources retrieved from these subscriptions based on its color coding. As we can see, this shows that the most access period of online resources is during the year of 2014. It is the time that is related to the accreditation endeavor of the school.

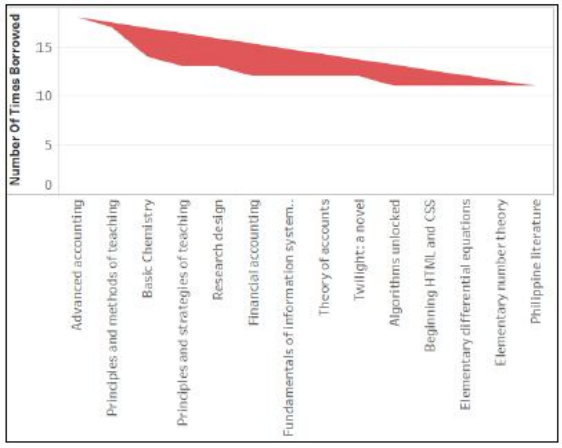


Figure 6. Title-based utilization of library resources for SY 2015-2017

Figure 6 in the form of a polygon graph shows title of the books that the users had borrowed. From a large database, researchers opt to display most frequent borrowed titles pertaining to the record gathered from database based on its accession. Same with the process, this graph gone through data cleansing first. Coming first on the list is the advanced accounting utilized by 0.80%, followed by principles and methods of teaching by 0.75% and basic chemistry by 0.62%. Placed on the middle of the graph is the famous twilight novel which was utilized by 0.53%. Came last is the Philippine Literature that was utilized 0.49%. We can see that based from this analysis, library users have different perspectives on the resources that they want to borrow whether the intention is for academic or entertainment.

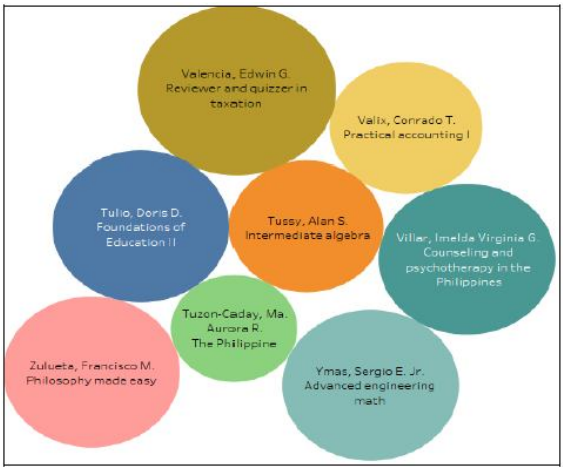


Figure 7. Utilization of books according to author for SY 2014-2017

Figure 7 shows in the form of packed bubbles graph that the most common searched books from an author-based analysis. Edwin Valencia ranks first. This report is based on the number of different accession that the researchers have been used. This is not based on how many times a title of the book has been borrowed but depends on how many times the author appears in different titles. This means that the researchers also depend of whom the author of the book they are borrowing. We may recall that figure 4 shows the utilization of borrowed books. The author-based analysis in this figure shows that it is linearly related to each other since the author of the books related to accounting and taxation is Mr. Edwin Valencia which was top utilized title and author of that book.

CONCLUSIONS

Different test results proved that library collection can affect the user's borro-wing behavior as utilization is concern. It was analyzed that most of the book borrowers of the library are faculty members of the institution. The analysis based on most borrowed books do not match to the results of the analysis for author-based results that shows visitors preference can also be based on the author of the book. While other users do not visit

the school library for book borrowing, it is evident that they are using and accessing the online subscription of the school specifically e-books. Through bibliomining, it is essential for every library management system to have data mining tools that can help them produce well described, graphically and statically – analyzed results.

RECOMMENDATIONS

From the results of this study, the researchers come up to recommend the following:

1. For gender equality study, it is better to have a portion in the database that determines male and female users of library.
2. Aside from utilization of resources, future researchers may conduct additional study in terms of overall services given from the library.
3. Produced results in data mining are coming from raw data. It is better to have all the needed information in the database (e.g. copyright, e-books title, books classification, etc.) for more detailed and well-managed database thus providing a complete report.
4. Librarians may also consider the use of data mining tools for a much easier report which can be one of the bases for library-related decision making.

REFERENCES

[1] S. Krashen, "Why Invest in Libraries," Los Angeles California, 2014.
[2] E. M. Mezick, "Return on Investment: Libraries and Student Retention," 2007.
[3] S. Weiner, "The Contribution of the Library to the Reputation of a University," The Journal of Academic Librarianship , vol. 35, 2008.
[4] S. R. Shimray and C. K. Ramaiah, Return on Investment in Academic Libraries, Puducherry India, 2015.
[5] P. W. Kiilu and J. Otike, "Non Use of Academic Library Services: A Literature Review.," International Journal of Library Science, vol. 5(1), pp. 7-13, 2016.
[6] S. Nicholson and S. J, Gaining Strategic Advantage through Bibliomining: Data Mining for Management Decisions in Corporate, Special, Digital and Traditional Libraries, Hershey Pennsylvania: Idea Group Publishing, 2003.
[7] L. J. Toner, "None use of library services by students in a U.K. academic library," Evidence-based Library and Information Practice , vol. 3, no. 2, pp. 18-29, 2008.
[8] The Bibliomining Processes, 2002.
[9] S.-Y. Hwang, P. Keezer and E. T. O'Neill, "The Bibliomining Process: Data Warehousing and Data Mining for Libraries," Proceedings of the American Society for Information Science and Technology , vol. 40, no. 1, pp. 478-479, 2005.
[10] "Descriptive, Predictive, and Prescriptive Analytics Explained," Halo Business Intelligence, [Online]. Available: https://halobi.com/blog/descriptive-predictive-and-prescriptive-analytics-explained/. [Accessed 30 August 2017].
[11] G. J. Williams and Z. Huang, "Modelling the KDD Process," in CSIRO Division of Information Technology, Canberra Australia, 1996.
[12] A. Carnegie, "Why should we invest in a new or expanded public library facility," 2015.
[13] G. Ong, "A Smarter Library: Using Data Analytics to Improve Resource Management and Services at NLB," Institute of Public Administration & Management (IPAM), Singapore, 2014.

★★★