

Analyzing user Satisfaction with Academic Library Services using CRISP-DM Process

I. M. Nawarathne^{1*} and R. M. Kapila Tharanga Rathnayaka²

¹Main library, Sabaragamuwa University of Sri Lanka, Belihuloya - 70140, Sri Lanka; imihami9@gmail.com

²Department of Physical Sciences and Technology, Sabaragamuwa University of Sri Lanka, Belihuloya - 70140, Sri Lanka; kapilar@appsc.sab.ac.lk

Abstract

This research is a quantitative study and its aims were to evaluate users' satisfaction of the available services in Sabaragamuwa university library. We used Cross-Industry Standard Process for Data Mining (CRISP-DM) process which is an industry-proven way to build predictive analytics models across the enterprise. The process contains six phases as follows: business understanding, data understanding, and preparation of data, modeling, evaluation, and deployment. A structured questionnaire was used to collect data in the data understanding phase. These were distributed among stratified randomly selected 360 undergraduates from the university. The descriptive statistics, Pearson correlation and multinomial logistic regression were applied to analyze the obtained data in modeling phase. The results indicated that the selected variables have a satisfaction level and the mean of their ranking values is around 3 and 4. Furthermore, results suggested that selected variable is distributed normally. Pearson Correlation results revealed that library collection (0.662) and library services (0.658) are highly correlated with physical facilities u. Furthermore, library services highly correlated with library collection and library website services.

Keywords: Academic Library, Cross-Industry Standard Process for Data Mining (CRISP-DM), Pearson Correlation, User Satisfaction

1. Introduction

The common library of Sabaragamuwa University of Sri Lanka consists of thousands of books and provides study materials and resources for almost three thousand undergraduates who are studying in the university. Thousands of books are circulated daily. The current automated system only consists of book circulation data, student's data and penalty calculation mechanisms only. However, so far the common library of Sabaragamuwa University of Sri Lanka has not done a proper and systematic review to monitor and evaluate user feedback. There are many deficiencies in the current system and these hinder further expansion of the system.

In this paper, we evaluate the level of user satisfaction of the library and discuss the results.

2. Related Work

Many studies have been carried out to investigate user satisfaction in library services. In order to understand user satisfaction and expectation of services at the Louisiana State University, a survey was conducted by Perrault and Arseneau (Perrault & Arseneau, 1995) and they had found that there was a high level of satisfaction of the library services. User satisfaction with library resources and services in Nigeria Agricultural Research Institutes was conducted by Ezeala and Yusuff (2011). The finding shows that users were dissatisfied with the library orientation, photocopying services, bindery services, inter library loan services, weekend library services, electronic library information resources and internet services. It was also found that users were satisfied with duration of loan

*Author for correspondence

services, opening hours of library, the lighting system of the library and the number of computers available. Motiang, Wallis and Karodia (2014) conducted an evaluation of user satisfaction with library services at the University of Limpopo, based on his research the level of user satisfaction of library opening hours, library registration, duration of loan services and available resources such as books and journals, has been generally high. The study suggested that providing high quality services to library users encourages them to use the library for their information needs.

Lee (1999) conducted a research to investigate library usage and user satisfaction with library services. Circulation, collection utilization, environment and physical facilities and attitude of the staff were considered as major dimensions for evaluating user satisfaction. Kassim (2017) evaluated user’s satisfaction with academic library performance. Johari and Zainab (2007) in a study, proposed improvements in library website, user education programmes.

There have hardly been any studies of user satisfaction

with academic library services in Sri Lanka,. One such research has been conducted by Library users of the University of Colombo to measure the service quality. The other research has been conducted by Jayasundara Ngulube, and Minishi-Majanja (2009) using students from four universities in art faculties.

Nawarathne and Singh (2013) studied Users’ Satisfaction with Academic Library Services in Sri Lanka. The analysis indicated that the highest dissatisfaction was with regard to library opening hours.

3. Methodology

In our research, we used Cross-Industry Standard Process for Data Mining (CRISP-DM) (Kumara, Paik et al., 2015). The process runs in six phases as follows: business understanding, data understanding, and preparation of data, modeling, evaluation, and deployment. Figure 1 shows the phases.

Business Understanding: The ultimate goal of the study

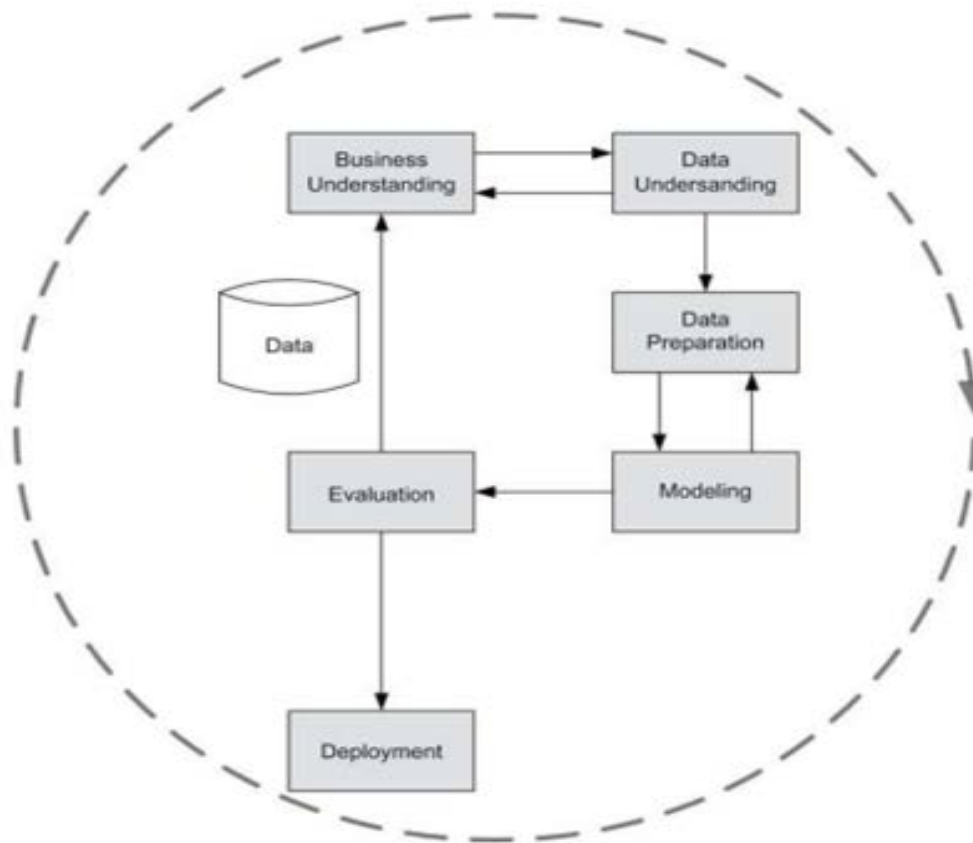


Figure 1. CRISP-DM process.

is to measure the users' satisfaction. Toward this goal, the main objective is to determine the causing factors.

Data Understanding: In this stage, we collect data using survey. We select the undergraduate students

of Sabaragamuwa University and we consider several methods to select the samples. We will discuss about this step in the next section.

Data Preparation: In this phase, we identify the core

Table 1. Descriptive statistics

Variable	Mean	Std. Deviation	Variance	Kurtosis Statistic	Cronbach's Alpha
	Statistic	Statistic	Statistic		
Physical facilities	3.2939	.6416	.412	-.227	.595
Library collection	3.2248	.6804	.463	-.418	.751
Library services	3.3500	.5596	.313	-.394	.531
Library website services	3.1066	.8082	.653	-.005	.671

Table 2. Pearson correlation results

	Physical facilities	Library collection	Library services	Library website services	Overall satisfaction
Physical facilities	1	.662**	.658**	.396**	.416**
Library collection	.662**	1	.651**	.474**	.403**
Library services	.658**	.651**	1	.581**	.389**
Library website services	.396**	.474**	.581**	1	.319**
Overall satisfaction	.416**	.403**	.389**	.319**	1

Note: ** Correlation is significant at the 0.01 level (2-tailed).

Table 3. Logistic regression results

Predictor	Co-efficient	SE Co-efficient	Z	P	Odds Ratio	95% CI	
						Lower	Upper
Const(1)	1.30838	0.814992	1.61	0.108			
Const(2)	4.87514	0.691333	7.05	0.000			
Const(3)	7.10113	0.748472	9.49	0.000			
Const(4)	9.91380	0.845479	11.73	0.000			
Physical facilities	-0.735650	0.234275	-3.14	0.002	0.48	0.30	0.76
Library collection	-0.548773	0.221383	-2.48	0.013	0.58	0.37	0.89
Library services	-0.312731	0.284755	-1.10	0.272	0.73	0.42	1.28
Library website services	-0.332389	0.157703	-2.11	0.035	0.72	0.53	0.98

influencing factors. The identified factors are then treated as variables upon which model analysis is performed. Further, during this phase, some data preprocessing tasks are performed including data format conversion, missing values handling, outliers handling, and converting data to a tabular format.

Modeling: Apply appropriate modeling techniques and calibrate the model to optimize the results. The results show patterns (i.e. the model) discovered in the data analyzed. If necessary, loop back to the preparation phase to bring the form of the data in line with the specific requirements for particular techniques. We used statistical analysis here.

Evaluation: Evaluate the model (the patterns discovered) from the quality and effectiveness perspectives before deploying and discover whether the model achieves the objectives set for it in the business-understanding phase.

4. Data Collection

The major data collection method used in the study was a questionnaire (Krejcie and Morgan, 1970) Participants answer the questions administered through questionnaires. After participants answer the questions, responses were analyzed.

Pilot study was conducted before finalizing the questionnaire. The selected population belonged to five major disciplines, viz., applied sciences, Social Sciences and Languages, Management studies, Geomatics and Agricultural sciences. The academic staff and the nonacademic staff of the university were not involved because our major objective was focused on Undergraduate students.

The questionnaire was distributed and data obtained were filtered, cleaned, duplicate data and outliers are removed, computed with the use of computer and software programs such as Statistical Package for the Social Science (SPSS) software and Minitab, edited, coded and then the prepared data was tabulated and analyzed.

5. Results and Discussion

Using above five major parameters mentioned, we ran basic descriptive statistics with Mean, Standard deviation, Variance and Kurtosis. As seen in Table 1, selected variables have a satisfaction level of around 3 and 4. Furthermore, variables have Cronbach's alpha value of above 0.5. So, the variables have satisfactory reliability for further analysis.

Furthermore results suggested that, selected variable are distributed slightly normally.

The correlation coefficient (r) measures the strength and direction of a linear relationship between two variables.

The value of r is always between +1 and -1. The correlation results are summarized in Table 2.

Clearly, library collection (0.662) and library services (0.658) are highly correlated with physical facilities. Furthermore, library services are highly correlated with library collection (0.651) and library website services.

6. Multinomial Logistic Regression

Once the behavior of the data set has been identified via basic descriptive statistics and the strength between separate variables is recognized, we decided the best method to exploit further hidden patterns was a logistic regression. But since the nonparametric variables we used had multiple levels or choices, we had to expand the logistic regression into multinomial regression instead of simple binomial logistic regression. As a next step logistic regression was run and results are summarized below (Table 3).

Based on our results, we can suggest that Library collection (OR=0.88), Library services (OR=0.73) and Library website services (OR=0.72) are highly affecting the Overall satisfaction rate than other selected variables.

7. Conclusion

While services of the university library were broadly satisfactory, a few services such as "delivery e-resources" "internet facilities for library users" "condition of the library materials" "attraction of the library building" are still not satisfactory and this fact needs to be taken into consideration. The library administrators should take necessary steps to provide appropriate services and adopt strategies to retain the users.

8. References

1. Perrault, A.H. and Arseneau, M. (1995). User satisfaction and interlibrary loan service: A study at Louisiana State. *University*, 35(1):90-100.
2. Ezeala, L. O. and Yusuff, E. O. (2011). User satisfaction with library resources and services in Nigerian Agricultural Research Institutes. *Library Philosophy and Practice (e-journal)*. <https://digitalcommons.unl.edu/libphilprac/564/>.
3. Motiang, I. P., Wallis, M. and Karodia, A. M. (2014). An evaluation of user satisfaction with library services at the University of Limpopo, Medunsa campus (Medical University of Southern Africa). *Arabian Journal of Business and Management Review (OMAN Chapter)*, 3(11):41-57. <https://doi.org/10.12816/0016519>.
4. Lee, S. (1999). The study of the library use satisfaction: An

- example of council for economic planning and development of the executive. *Yua. Suyan Periodicals*, 41: 64-85.
5. Kassim, N. A. (2017). Evaluating users' satisfaction on academic library performance. *Malaysian Journal of Library and Information Science*, 14(2): 101-115.
 6. Johari, R. and Zainab, A. (2007) Identifying what services need to be improved by measuring the library's performance. *Malaysian Journal of Library and Information Science*, 12(1):35-53.
 7. Jayasundara, C., Ngulube, P. and Minishi-Majanja, M. K. (2009). A theoretical model to predict customer satisfaction in relation to service quality in selected university libraries in Sri Lanka. *South African Journal of Libraries and Information Science*, 179(2):179-194. <https://doi.org/10.7553/75-2-98>.
 8. Nawarathne, I. M. and Singh, A. P. (2013) Users' Satisfaction of the Academic Library Services in Sri Lanka. *Pearl: A Journal of Library and Information Science*, 7(2):103-112. <https://doi.org/10.5958/j.0975-6922.7.2.016>.
 9. Kumara, B. T. G. S., Paik, I., Zhang, J., Siriweera, T. H. A. S. and Koswatte, K. R. C. (2015). Ontology-Based Workflow Generation for Intelligent Big Data Analytics, 2015 IEEE International Conference on Web Services; 495-502. <https://doi.org/10.1109/ICWS.2015.72>.
 10. Krejcie, R. V. and Morgan, D. W. (1970). Determining sample size for research Activities, Educational and Psychological Measurement. *SAGE Journal*, 30:607-610. <https://doi.org/10.1177/001316447003000308>.