Unsupervised Clustering Evaluation on Services of Public Library

Sofia Vlachou, Martha Georgaki, Charalampos Pagkratis and Marios Poulos
Department of Archives, Library and Museum Studies
Ionian University
Ioannou Theotoki 72 - Corfu
Greece
mpoulos@ionio.gr

Abstract: - This assignment examines the evaluation's case of public libraries services. For this purpose ISO-standards indexes were used, which we adjusted to the indexes of MOPAB (Total Quality Unit of Academic Libraries). Overall, we used seven indexes. This assignment's aim was for these indexes to be globally assessed using an expert's opinion. The formulation of the above is made by an unsupervised clustering classification, which is implemented by the K-Means algorithm. This implementation's results indicated that the success rate exceeded 99%. These results show that this method is valid and can be used for the assessment. The above purposes can definitely be achieved too

Key-Words: - K-mean Algorithm, clustering, library evaluation

1 Introduction

The need to create new services in order to answer typical more e-metrics issues in Greek Academic Libraries is considered as imperative. The role of this need has undertaken the "Total Quality Management Unit of Academic Libraries" which in Greece named as "MOPAB".

The purpose of the MOPAB is to focus on the most representative indicators for a library's evaluation, due to the fact that they can represent all types of libraries and they are not limited on a particular type. The indicators P and D are used to evaluate specific characteristics of a library, such as the virtual accesses, virtual visits and the percentage of people using the library's services in total. The reason why we chose to use MOPAB indicators has to do with the fact that MOPAB is the only readily available source for relevant evaluation data in Greece, plus MOPAB is known to be built on ARL (Assosiation of Research Libraries) [1,2].

ARL is a major project to develop a standardized measure of library quality based on four dimensions: 1) affect service, 2) library as place, 3) personal control, 4) access to information. On the other side, ISO (International Organization for Standardization) is a non-governmental federation that prepares international standards, with low cost, about any technical field and subject area of science. The mission of the organization is the homogeneity, the consistency and the compliance of specifications of the standards in a global level. This standard (ISO 11620: 2008) [3] has been created exclusively for public libraries and their evaluation. The purpose is the performance measurement of services regardless of the type and size of each library. It includes a set of indicators, that anyone has a unique name, a comprehensive description and a calculation method. The standard does not yet include performance indicators for the evaluation of impact of library services on community.

Any attempt to correlate and/or determine weighting would definitely lead to a subjective practice. A number of different methods, such as TOM, EFOM, SERVQUAL and LibQUAL [4] are used to evaluate traditional or digital libraries. The common prompt in the practice of quality management under all these methods is to measure the performance of these libraries, in numbers [5-7]. We propose a different approach sharing the same intention to measure performance in numbers.

1.1 Aims and Scope

The aim of this process is focused on the determination of a theoretical model that this could be helping us to cluster the selected indicators for evaluated purposes. To succeed in this process we had to use a tool, which is based on clustering algorithm. Until, now an attempt to classify the evaluation services according to expert opinion took place using supervised classification systems [8-10]. In contrast, this algorithm is responsible for the development of the indicators, according to the expert's classification but in unsupervised way. It is important to mention that every indicator that is defined has an entry, and each entry has to generate a value for the library's which will represent the expert's opinion. Furthermore, the aim of this algorithm is to perform an evaluated system in numbers, and this is the reason that this network is chosen for this study. The main purpose of this study is to collect data and after to evaluate them by this network. In this case, the problem of the subjectivity exists but using this system we will succeed in solving this problem. In this direction an unsupervised system based on k-mean algorithm is adopted.

2 Methods

2.1 Description and classification of indicators to a measurable process.

Both of the standards have been studied at length, because we had to find the most suitable indicators in order to place a measurable process efficiency of public libraries. We followed this process in order to perform the practical part of this study. The aim of the chosen indicators is to analyze data in order to help us understand how the library services are efficient. The methodology that is followed is articulated in two parts. During the first part, the effort to quantify and group the dependent individual variants into normalized single values is presented. Each expert will determine the rationale for priority-setting in indicator weighting and the way that is followed to reach the aim. Some other experts will follow the same process leading to an indicator-weighting group. Finally, in the second part, an unsupervised clustering procedure which is based on k-mean algorithm is described. After that, the k-mean will be trained by a subtotal of samples which came from the procedure which is described in part 1.

2.1.1 *Expert Opinion Formulation and Description*

First, correlation of MOPAB indicators and the indicators of ISO 11620 would be examined, so that the process would be as accurate and precise as possible, because the performance indicators included in this International Standard (ISO 11620) are those seen to be most useful for libraries in general. Then, based on the model of MOPAB, seven indicators would be selected (Table 1). These indicators certainly express the subjective opinion of researchers, regarding to the quality criteria of a public library, nevertheless they would be selected carefully exclusively targeting to objective results.

Once the ideal, as well as the moderate and poor efficiency values defined, the work would progress to the second stage, -through the choice of indicators- that of the calculation method, in order to make clear the value ranges that will be encountered in each level of efficiency. Finally, the indicators would be divided by category, and calculated by the number of statistics 'D'.

The data to be collected and the calculations to be performed shall be both described concisely [11, 12].

1) ISO B.1.1.1"Required Titles Availability" corresponds to MOPAB's P36= number of library documents collection per capita:

D5/D1, where D5= Size of library's collection and D1=Percentage of people using the library services in total

2)B.2.1.3"Percentage of Stock Not Used" corresponds to P33=collection use: 1(stock)-D3/D4, where D3= amount of lending during one year and D4= Library's lending collection size

Table 1 — List of performance indicators chosen

S/N	ISO 11620 Indicators	MOPAB Indicators	Calculation Method
1	B.1.1.1 - Required Titles Availability	P36 - Number of library documents collection per capita	D5/D1 Size of library's collection/Percentage of people using the library services in total
2	B.2.1.3 - Percentage of Stock Not Used	P33 – Collection use	1-D3/D4 1- Amount of lending during one year/Library's lending collection size
3	B.1.3.5 - Hours Open Compared to Demand	P47 - Hours of library operations daily	P47=D20 Hours of library operation daily= Total hours library's operation daily
4	B.1.1.4 - Percentage of Rejected Sessions	P35 - Percentage of material in disuse	(D6/D4)*100 (Number of documents into disuse/Library's lending collection size)*100

5	B.2.2.2 - Percentage of Information Requests Submitted Electronically	P55 - Number of information queries of users handled electronic monthly per capita:	D30/D1 Number of information inquiry/requests handled electronically/Percentage of people using the library services in total
6	B.2.4.2 - User Satisfaction	-	-
7	B.4.2.2 - Number of Attendance Hours at Formal Training Lessons per Staff Member	P53 - Intensive annual training library staff-training hours per staff per year	D28/D24 - Annual total hours of staff training /Library staff

3) B.1.3.5 “Hours Open Compared to Demand” corresponds to P47=Hours of library operation daily. P47=D20, where D20=Total hours library's operation daily.
4) B.1.1.4“Percentage of Rejected Sessions” corresponds to P35=Percentage of material in disuse:
(D6/D4)X100, where D6=Number of documents into disuse and D4=Library's lending collection size
5) B.2.2.2 “Percentage of Information Requests Submitted Electronically” corresponds to P55=Number of information queries of users handled electronic monthly per capita:
D30/D1, where D30=Number of information inquiry/requests handled electronically and D1=Percentage of people using the library services in total.
6) B.2.4.2 “ User Satisfaction”
7) B.4.2.2 “Number of Attendance Hours at Formal Training Lessons per Staff Member” corresponds to P53=Intensive annual training library staff-training hours per staff per year:
D28/D24, where D28=Annual total hours of staff training and D24=Library staff.

2.1.2 Indicators Range

First of all, it is important to mention that the range for all of the indicators fluctuates between 0.00-1.00. The “ideal” rate for all indicators is 1.00 and the “poor” rate is 0.00 except for the second indicator where the ideal rate is 0.00 and the bad rate 1.00. There are three categories for our results, the “ideal” category, the “good” category and the “poor” category, which refer to high, medium and low efficiency.
For the first indicator, B.1.1.1, we have to divide D5, which is the number of the items of a library and D1, which is the number of the patrons of a library. If a library owns 100 items and helps 300 patrons, the perfect result for monthly loan of a patron is 2-3, the good result is 1-2 and the bad one is 0-1. So, the ideal range is 0.67-1.00, the good range 0.33-0.66 and the poor range 0.00-0.32.
For the second indicator, B.2.1.3, we have to divide D3, which is the number of loans during a whole year and D4, which is the total number of the items of a library. After the division, the rate for the ideal category is 0.96-1.00, for the good category 0.92-0.95 and for the poor category 0.00-0.91.
For the third indicator, B.1.3.5, we get D20 which refers to the total hours that a library is open. If we suppose that the ideal is 14 hours every day, so the perfect range is about 10-14 hours, the good range 7-14 hours made the bad range 0-7. So, the result for the perfect category will be 0.71-1.00, for the good category 0.50-0.70 and for the bad category 0.00-0.49.
For the fourth indicator, B.1.1.4, we have to divide D6, which is the number of the items that are not used and D4, which is the total number of the items of a library. After the division, we found that the range for the ideal category is 0.10-0.00, for the good category 0.60-0.11 and for poor category 1.00-0.61.
For the fifth indicator, B.2.2.2, we have to divide D30 which is the number of questions that patrons did and the staff of the library answered online and D1, which is the number of the patrons of a library. If we suppose that a library has 100 patrons and they make about 4 questions per month, the ideal is that the staff will answer to all of them, the good is to answer 360-400 and the bad 0-360. So the range for the ideal category is 0.90-1.00, for the good category 0.60-0.89 and for the poor category 0.00-0.59.
For the sixth indicator, B.2.4.2, we suppose that the range for the ideal category is 0.80-1.00, for the good category is 0.50-0.79 and for the poor category 0.00-0.49.
For the seventh indicator, B.4.2.2, we have to divide D28 which is the total hours of staff training every year and D24, which is the staff of the library. If we suppose that a library has 10 people as staff the ideal number of training is 60 hours per person, so 600

hours for the whole staff. The number of training hours for the good category would be 20-50 hours and for the bad category 0-20. For the ideal category, the range would be 0.83-1.00, for the good category 0.33-0.82 and for the poor category 0.00-0.32.

2.2 Description of the K-Mean Algorithm

Before continuing with the processing of the data for the recommendation provision by learning the individual user profiles, we first check whether the data are homogeneous. The hypothesis that the groups of the points (data review) satisfies the condition of homogeneity is resolved by a well-fitted unsupervised clustering k-means method, where the partitions of the points in the (i-by-j) data matrix A (see the processing stage) are grouped into k clusters. This iterative partitioning minimizes the sum, over all the clusters, of the within-cluster sums of point-to-cluster-centroid distances. The rows of A correspond to points (see Eq. (1), while the columns correspond to variables. The k-means process returns an (i-by-1) vector containing the cluster indices of each point. By default, k-means uses squared Euclidean distances [13-14]. As a result, in our case we used as value k = 2 as
the initial number of clusters, and we intuitively classified a given data set using a certain number of clusters (assume k clusters) established beforehand. Furthermore, the K-means algorithm is simple and fast. The time complexity of K-means is $O(klN)$, where l is the number of iterations, k is the number of clusters and N the total sample. The main idea is to define one k centroids for each luster. These centroids should be placed carefully because different locations provide different results. A loop is generated, which shows that the k centroids change their location step by step until no more changes are made. In other words, the centroids do not move any more. In this case, we submitted the dataset from the indicators into k=3 clusters using the following objective function:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{N} \left\| A_i^{(j)} - c_j \right\| \quad (1)$$

3 Experimental Part

In Table 2 is presented a sample of indicators from all three categories. The part of the table 1 shows a sample of indicators from the category of “ideal”. The second part shows a sample of indicators from the category of “good”. And finally the third part depicts a sample of indicators from the category of “poor”.

Table 2. Random Ranges Ideal's Category for each Indicator.

Classes/Rate	Indicators	Rate	Sample
Class A Ideal	(b1) B.1.1.1	0.67-1	0.90
	(b2) B.1.1.4	0-0.10	0.04
	(b3) B.1.2.3	0.96-1	0.97
	(b4) B.1.3.5	0.71-1	0.88
	(b5) B.2.2.2	0.90-1	0.91
	(b6) B.2.4.2	0.80-1	0.84
	(b7) B.4.2.2	0.83-1	0.99
Class B Good	(b1) B.1.1.1	0.33-0.66	0.42
	(b2) B.1.1.4	0.60-0.11	0.56
	(b3) B.1.2.3	0.92-0.95	0.94
	(b4) B.1.3.5	0.50-0.70	0.62
	(b5) B.2.2.2	0.60-0.89	0.71
	(b6) B.2.4.2	0.50-0.79	0.61
	(b7) B.4.2.2	0.33-0.82	0.42
Class C Poor	(b1) B.1.1.1	0-0.32	0.30
	(b2) B.1.1.4	1-0.61	0.95
	(b3) B.1.2.3	0-0.91	0.88
	(b4) B.1.3.5	0-0.49	0.22
	(b5) B.2.2.2	0-0.59	0.35
	(b6) B.2.4.2	0-0.49	0.29
	(b7) B.4.2.2	0-0.32	0.07

3.1 Implementation of the K-Mean Algorithm

The experimental evaluation process involves three steps: data selection, normalization and clustering in order to find the closest type using a correlational approach. The main objective is to check whether the clustering process converges from the pre-processing stage, and to evaluate the coefficient of identification. The method k-means clustering is a partitioning method. The function k-means partitions data into k mutually exclusive clusters, and returns the index of the cluster to which it has assigned each observation. Unlike hierarchical clustering, k-means clustering operates on actual observations (rather than the larger set of dissimilarity measures), and creates a single level of clusters. The distinctions mean that k-means clustering is often more suitable than hierarchical clustering for large amounts of data [15-18]. Also the k-means treats each observation is adapted in our case as an object having a location in space. It finds a partition in which objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. In our case, we choose from three different distance measures which correspond in the three aforementioned values of vectors. Thus, testing procedure in total 180 set of all classes (60 per each class are participated.

Each cluster in the partition is defined by its member objects and by its centroid, or center. The centroid for each cluster is the point to which the sum of distances from all objects in that cluster is minimized. The k-mean procedure computes cluster centroids differently for each distance measure, to minimize the sum with respect to the measure that you specify. In the experimental part, the results showed that we have a very accurate successful score in the clustering issue. So this is verified, that is, each vector (after clustering procedure) has a minimize distance (95.38) for each clustered centroid in which belong (see Figure 1). Thus, we constructed a matrix with size (3 centroids x7 values of indicators) in which the distances from each indicator to every centroid are calculated and these are depicted in Table 3.

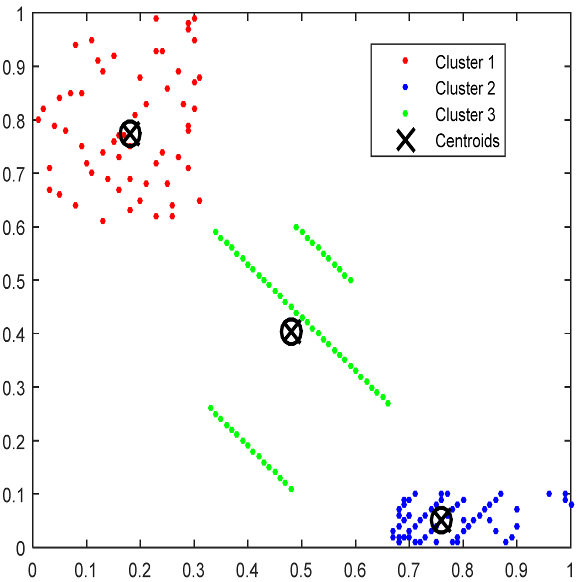

Figure 1. The clustering procedure regarding to vectors into k=3 clusters.

As, we can in the Table 3 the k-mean algorithm creates with clarity the boundary values of each indicators in the three (3) clusters.

Table 3. The k=3 clusters centroids locations according to 7 indicators

b1	b2	b3	b4	b5	b6	b7
0.1800	0.7750	0.6150	0.3050	0.3400	0.1850	0.1550
0.4800	0.4050	0.9300	0.6000	0.7500	0.6450	0.5250
0.7600	0.0500	0.9750	0.7700	0.9400	0.8500	0.9000

4 Conclusion

In this paper we focused on a method solving problems related to the normalization of measured data linked with significant relevant properties of the library services evaluation. In the first phase, we created a set of normalized weights of opinions of experts associated with the previous properties. Moreover, we described how these normalized weights which are expressed in Euclidean distance using a well fitted unsupervised k-mean clustering method could be used in a decision of evaluation library services system. These results demonstrated that the simulation model of the vectors can be adapted successfully to the proposed unsupervised method. In the future, we would like to perform an extensive statistical evaluation of our model with real Library indicators obtained through experimental questioners.

References
:

[1] Stubbs, Kendon, and Association of Research Libraries (USA). The ARL Library Index and quantitative relationships in the ARL. Washington, DC: Association of Research Libraries, 1980.

[2] Weiner, S.A.(2005). Library quality and impact: is there a relationship between new measures and traditional measures?. The Journal of Academic Librarianship, 31 (5), 432-437

[3] ISO, B. (2008). 11620: 2008. Information and documentation–Library Performance Indicators [on-Line] [dostęp 02.11. 2015]. Dostępny W: Http://portal. Psz. Utm.My/psp/wpcontent/uploads/2012/08/BS_ISO11620_2008. Pdf.

[4] Roszkowski, Michael J., John S. Baky, and David B. Jones. "So which score on the LibQual+™ tells me if library users are satisfied?." Library & information science research 27.4 (2005): 424-439.

[5] Garibay, Cecilia, Humberto Gutiérrez, and Arturo Figueroa. "Evaluation of a digital library by means of quality function deployment (QFD) and the Kano model." The Journal of Academic Librarianship 36.2 (2010): 125-132.

[6] Whitmire E.(2002), "Academic library performance measures and undergraduates' library use and educational outcomes", Library & Information Science Research, Vol 24, pp. 107-128.

[7] Babalhavaeji, Fahimeh, et al. "Quality assessment of academic libraries' performance with a special reference to information technology-based services: Suggesting an evaluation checklist." The Electronic Library 28.4 (2010): 592-621.

[8] Papavlasopoulos, Sozon, and Marios Poulos. "Neural network design and evaluation for classifying library indicators using personal opinion of expert." Library Management 33.4/5 (2012): 261-271.
[9] Thimiakou, Vasileia, et al. "Αξιολόγηση Υπηρεσιών Βιβλιοθηκών μέσω Τεχνικού Νευρωνικού Δικτύου." HealJournal 1.2 (2015): 47-60.
[10] Glushkov, V. M. "Prediction based on expert opinions." Cybernetics and Systems Analysis 5.2 (1969): 127-129.
[11] Ho, William, Xiaowei Xu, and Prasanta K. Dey. "Multi-criteria decision making approaches for supplier evaluation and selection: A literature review." European Journal of Operational Research 202.1 (2010): 16-24.
[12] Kyrillidou, Martha. "An overview of performance measures in higher education and libraries." Journal of Library Administration 35.4 (2002): 7-18.
[13] Singh, M., Nagar, H., & Sant, A. (2016). K-mean and EM Clustering algorithm using attendance performance improvement Primary school Student. Retrieved from http://ijsrset.com/paper/708.pdf
[14] Tyagi, L., & Trivedi, M. C. (2016). Hybrid K-Mean and Refinement Based on Ant for Color Image Clustering. In Proceedings of International Conference on ICT for Sustainable Development (pp. 779–789). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-981-10-0135-2_74
[15] Arthur, David, and Sergi Vassilvitskii. "K-means++: The Advantages of Careful Seeding." SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. 2007, pp. 1027–1035.
[16] Lloyd, Stuart P. "Least Squares Quantization in PCM." IEEE Transactions on Information Theory. Vol. 28, 1982, pp. 129–137.
[17] Seber, G. A. F. Multivariate Observations. Hoboken, NJ: John Wiley & Sons, Inc., 1984.
[18] Spath, H. Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples. Translated by J. Goldschmidt. New York: Halsted Press, 1985.