Evaluation on Services of Public Library using Supervised Classification System

Argyro Lekka, Eirini Kalogera and Marios Poulos
Department of Archives, Library and Museum Studies
Ionian University
Ioannou Theotoki 72 - Corfu
Greece
mpoulos@ionio.gr

Abstract: - This paper considers the case of a public library evaluation. To achieve this aim ISO standard indicators, which were adapted to MOPAB indicators, are used. A total of seven (7) indicators were used. The purpose of this work is to make a comprehensive assessment of the indicators, using the expert's opinion. The formulation of the above is obtained via a Supervised Classification System. The results of this embodiment have shown that the success rate outnumbered 99%. Therefore, the above percentage demonstrates that this method can reliably and safely be used for evaluation of a Public Library.

Key-Words: - Neural Network, learning Systems, Data mining

1 Introduction

The essential aim of this paper is to set an evaluation system in order to measure the efficiency on Public Libraries' services. For this purpose we relied on MOPAB standard and its already established indicators. We chose to use the most representative indicators to achieve our goal and these indicators were chosen because of the fact that they reliably represent all types of libraries, not strictly one kind. The indicators P, D are used to asses particular aspects of a library, for example the virtual accesses, the virtual visits and the percentage of people who use the library's services [1]. The reason we picked to utilize the MOPAB's indicators is that MOPAB is currently the only applicable assessment information in Greece. Furthermore, MOPAB is also known to be referred, as an extension to ARL (Association of Research Libraries) guide [2].
ARL [2] is a major task to develop a standardized measurement of library quality based on four dimensions: 1) influence service, 2) library as place, 3) personal control, 4) access to information-data [3] ISO (International Organization for Standardization), on the other side, is a non-governmental federation that prepares worlwide standards, with low cost, over any specialized foul field and branch of knowledge of science. The mission of the organization is the homogeneity, the consistency and the the agreeability of determinations on the measures in a worldwide level. ISO 11620 [3] has been created exclusively for public libraries and their evaluation. The purpose is to execute estimation on a library's services in any case of the kind and size of each library. It incorporates a set of indicators, that each one has an exceptional name, a comprehensive description and a calculating method. [3] It does not corporate execution indicators for the evaluation of impact on community library services.
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 [3, 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 recommend an alternative methodology, imparting the same proposition on measure execution alongside numbers.

2.1 Aims and Scopes

The scope of this study is to define a theoretical model for the combination of all individual assessment indicators into one single number. The tool proposed for the achievement of such an aim is a well-fitted Artificial Neural Network (ANN) [8], which will undertake to formulate the weighting relations between indicators, according to the expert's classification. The key challenge for this modelling is that every entry of individual assessment indicators will generate an overall value for the library's assessment, which will represent the expert's opinion. Additionally, it is known that the artificial ANN measures the performance of an evaluated system in numbers [8, 9], and this is in agreement with the previously mentioned intention, to measure performance in numbers. The objective of this paper is to collect the opinions of several experts in order to evaluate these opinions as a whole by an ANN [7-9]. In our case, the resolution of subjectivity will be based on the creation of an expert system, which will derive knowledge from the expert opinion. In this direction a supervised system is adopted. The proposed study is divided to the following sections: Formulating, describing and classifying the indicators, training a neural network via supervised model and performing an experimental part to collect data.

2 Methodology

2.1 Description and Classification indicators to a measurable process

In order to perform the practical part of this paper there had to be studied both the above two standards to be decided which indicators will be used in order to place a measurable process efficiency of public libraries, which is the subject of this work. Performance indicators are created by comparing quantitative data elements in different combinations. The purpose of the indicators is to analyze data in order to clarify the output and outcome of the library services and see how well the library is performing [10]. The methodology used is articulated into two stages. Stage 1 attempts to quantify and group any dependent individual variants into normalized single values. Every expert will determine the rationale for priority-setting in indicator weighting and the methodology followed to formulate the final scoring. A number of other experts follow the same procedure, leading to a indicator-weighting group. Finally, in stage 2, a supervised linear neural network will be constructed with four output neurons, reflecting the experts' opinion. Then, the ANN will be trained by a subtotal of samples; then, we will attempt to evaluate such training through an appropriately-formulated sigmoid function from the remaining untrained samples.

2.1.1 Expert Opinion on Formulation and Description

First, correlation of MOPAB indicators and the indicators of ISO 11620 was examined, so that the process would be as accurate and precise as possible, because the performance indicators included in this International Standard (ISO 11620) [3] are those seen to be most useful for libraries in general. Then, based on the model of MOPAB, seven indicators would be selected. These indicators certainly express the subjective opinion of researchers, regarding to the quality criteria of a public library, nevertheless they were selected carefully exclusively targeting to objective results [11, 12].

Table 1. List of performance indicators chosen

S/N	ISO 11620 Indicators	MOPAB Indicators	Calculation Method
b_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
b_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
b_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
b_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
b_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
b_{16}	B.2.4.2 - User Satisfaction	-	-
b_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

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 were 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.

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

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

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 "moderate" category and the "poor" category, which refer to high, medium and low efficiency.

B.1.1.1: If a library owns 100 items and helps 300 patrons, the ideal result for monthly loan of a patron is 2-3, the moderate result is 1-2 and the poor one is 0-1. So, the ideal range is 0.67-1.00, the moderate range 0.33-0.66 and the poor range 0.00-0.32.

B.2.1.3: After the division, the rate for the ideal category is 0.96-1.00, for the moderate category 0.92-0.95 and for the poor category 0.00-0.91.

B.1.3.5: If we suppose that the ideal is 14 hours every day, then the ideal range is about 10-14 hours, the moderate range 7-14 hours made the poor range 0-7. So, the result for the ideal category will be 0.71-1.00, for the moderate category 0.50-0.70 and for the poor category 0.00-0.49.

B.1.1.4: After the division, we found that the range for the ideal category is 0.10-0.00, for the moderate category 0.60-0.11 and for poor category 1.00-0.61.

B.2.2.2: 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 moderate is to answer 360-400 and the poor result is 0-360 answers. So the range for the ideal category is 0.90-1.00, for the moderate category 0.60-0.89 and for the poor category 0.00-0.59.

B.2.4.2, we suppose that the range for the ideal category is 0.80-1.00, for the moderate category is 0.50-0.79 and for the poor category 0.00-0.49.

B.4.2.2: 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 ideal category would be 20-50 hours and for the poor category 0-20. For the ideal category, the range would be 0.83-1.00, for the moderate category 0.33-0.82 and for the poor category 0.00-0.32. Finally, for every case, we constructed a vector **P** size 1x7 which contains the seven values of the selected indicators as described in Table 1, to wit $\vec{P} = \{b_1, ..., b_7\}$.

2.2 Description of Supervised System (Neural Networks)

In this study, a well-defined Neural Network (NN) was selected to provide an appropriate weight in the process of learning for input vectors (p) which present variability according to previous study [13]. Furthermore, this NN could be trained to distinguish and produce both spatial and temporal patterns that resolve the problem of the factors' threshold value variability in order to support a rule-based decision procedure. Thus, the recurrent Elman neural network was selected over BP and SOM, because according to previous work on similar problems, a comparison of their architectures (i.e., a multilayer perceptron (MLP) trained with the resilient backpropagation (RPROP)) showed that the best prediction accuracy was obtained with the extended Elman neural network [13]. The Elman network is based on a two-layer network with feedback in the first (hidden recurrent) layer and a second output-layer. This recurrent connection permits the Elman network to both detect and generate time-varying patterns [13] and in the training procedure which is based on propagation technique. In this stage the weights (w) between the elements of the vector are calculated using a continuous (epochs) calculation of weighted errors until the system convergences in a critical value. The hidden recurrent layer consists of a numerous of appropriate neurons. This architecture uses a sigmoid activation function (sigmf) in order to be calculated the critical distance which has each candidate for classification vector from the predetermined (supervised) class [14].

$$y = sigmf(w, p) = \frac{1}{1+e^{-f(p,w)}} \simeq 0 \qquad (1)$$

The output layer is characterized by a linear activation function. In this case, two different types of input vectors (ideal $0.8 < y < 1.2$, moderate $1.3 < y < 1.8$ and poor $1.9 < y < 2.5$ for developing type services of library) respond to the function. Also, in the testing procedure, the setting regarding to numbers of the neurons as well as the epochs are described in the experimental part.

3 Experimental Part and Results

To implement the above methodology we created a vector, comprising seven (7) indicators, those already described in the preceding paragraph of methodology. Therefor we created sixty (60) vectors for each category, 60 for the ideal, 60 for the moderate and 60 for the poor. So, according to paragraph 2.1.1 (Expert Opinion on Formulation and Description) and Table 1 and according to paragraph 2.1.2 (Indicators Range) and each single Indicator, 3 tables of random ranges for each category were made and put into Neural Network for training. Below is a sample vector of each category.

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

Indicator	Rate	R.R	R.R	R.R	R.R	R.R
B.1.1.1	0.67-1	0.80	0.90	0.68	0.69	0.71
B.1.1.4	0-0.10	0.05	0.04	0.01	0.09	0.08
B.1.2.3	0.96-1	0.98	0.97	0.99	0.96	0.99
B.1.3.5	0.71-1	0.90	0.88	0.72	0.78	0.80
B.2.2.2	0.90-1	0.95	0.91	0.92	0.95	0.91
B.2.4.2	0.80-1	0.83	0.84	0.85	0.98	0.81
B.4.2.2	0.83-1	0.89	0.99	0.86	0.90	0.83

Table 3. Random Ranges Moderate's Category for each Indicator.

Indicator	Rate	R.R	R.R	R.R	R.R	R.R
B.1.1.1	0.33-0.66	0.34	0.36	0.38	0.42	0.44
B.1.1.4	0.60-0.11	0.59	0.58	0.57	0.56	0.55
B.1.2.3	0.92-0.95	0.93	0.94	0.92	0.94	0.93
B.1.3.5	0.50-0.70	0.52	0.60	0.57	0.62	0.69
B.2.2.2	0.60-0.89	0.68	0.69	0.70	0.71	0.72
B.2.4.2	0.50-0.79	0.58	0.59	0.60	0.61	0.62
B.4.2.2	0.33-0.82	0.37	0.38	0.39	0.42	0.43

Table 4. Random Ranges Poor's Category for each Indicator.

Indicator	Rate	R.R	R.R	R.R	R.R	R.R
B.1.1.1	0-0.32	0.20	0.30	0.11	0.09	0.18
B.1.1.4	1-0.61	0.88	0.95	0.70	0.75	0.63
B.1.2.3	0-0.91	0.84	0.88	0.90	0.72	0.77
B.1.3.5	0-0.49	0.30	0.22	0.31	0.47	0.21
B.2.2.2	0-0.59	0.25	0.35	0.16	0.50	0.47
B.2.4.2	0-0.49	0.10	0.29	0.44	0.45	0.33
B.4.2.2	0-0.32	0.15	0.07	0.28	0.23	0.10

3.1. Neural Network

We designed an Elman neural network to classify the intracellular VEGF immunostaining into three classes. The architecture design as well as the learning and testing procedures occurred according to previous our knowledge. In order to achieve this, we trained 15 vectors of each class. In the testing procedure, we used as testing vectors, the rest vectors of Eq. (1).

Fig. 1 shows the architecture of the Elman network used to classify these vectors (model order $p = 7$, which are the indicators b1-b7 see Table1). We weighted and fed input vectors of dimensionality 7×1 to the first layer of neurons, known as the competitive layer. These neurons compete for inputs in a "greedy" way, hence their name. Four such neurons formed the competitive layer in our case. The output of the competitive layer, which is a grouping of the inputs into sub-classes, is fed to the second linear layer, which groups subclasses into target classes. The weights connecting the two layers take on binary values of zero or one, indicating mere class membership and not actual weighting (w). Three target classes exist here, the class of interest (A Ideal or B, Moderate or C, Poor).

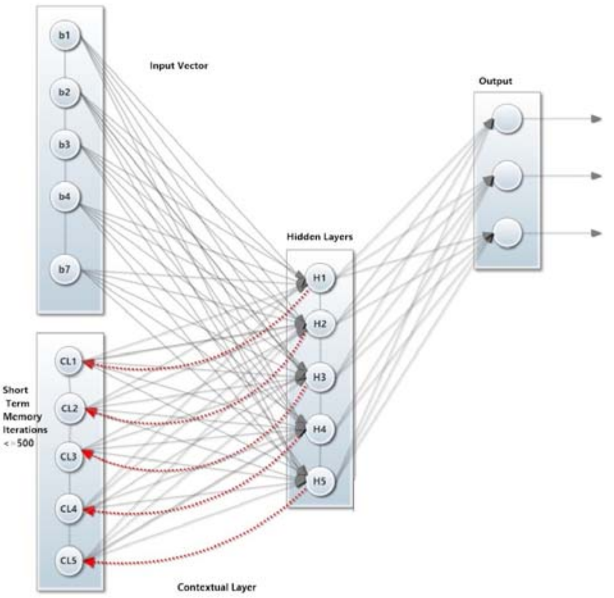

Figure 1. The architecture of Elman Neural Network

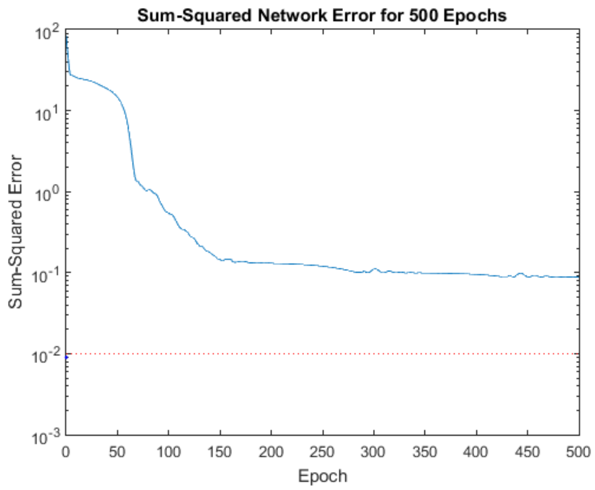

Figure 2. The error training procedure in 500 epochs

Fig. 1 shows the architecture of the Elman network used to classify these vectors (model order $p = 7$, which are the indicators b1-b7 see Table1). We weighted and fed input vectors of dimensionality 7×1 to the first layer of neurons, known as the competitive layer. These neurons compete for inputs in a "greedy" way, hence their name. Four such neurons formed the competitive layer in our case. The output of the competitive layer, which is a grouping of the inputs into sub-classes, is fed to the second linear layer, which groups subclasses into target classes. The weights connecting the two layers take on binary values of zero or one, indicating mere class membership and not actual weighting (w). Three target classes exist here, the class of interest (A Ideal or B, Moderate or C, Poor).

3.2 Results

In the testing procedure we used the calculated weighted for the training procedure and each candidate vector p_x is input in the sigmoid function (see equation 1 and figure 2) and then a value is estimated. For verification reasons, we tested 135 vectors which do not participate in the training procedure.

Table 5 Sample input vector data and create weighting in training procedure of Elman ANN

Classes	Vectors	Sgmf value
Class A 0.8-1.1	P1	0.9796
	P2	1.0091
	P3	0.9751
	P4	0.9886
	P5	0.9793
	P6	0.9970
	P7	0.9695
	P8	1.0039
	P9	1.0152
	P10	0.9479
Class B 1.2-1.6	C1	1.4090
	C2	1.3842
	C3	1.3602
	C4	1.3370
	C5	1.3354
	C6	1.3131
	C7	1.2916
	C8	1.2710
	C9	1.2694
	C10	1.2496
Class C 1.8-2.3	D1	2.2603
	D2	2.0728
	D3	1.8413
	D4	1.8848
	D5	1.8471
	D6	2.1263
	D7	2.0178
	D8	2.0308
	D9	2.0693
	D10	1.9729

For verification reasons, in Table 5 a part of the 135 tested vectors in which the ranging around the critical values is depicted. The results of this experimental test have shown that the success rate exceeds the 99%.

4 Conclusion and Discussion

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 could aid the training of an

ANN and how the virtual sets (vectors) of realistic random indicators will be produced for the neural network demands. These results demonstrated that the simulation model of the vectors can be adapted successfully to the proposed neural network. In the future, we would like to perform an extensive statistical evaluation of our model with real Library indicators obtained through experimental questioners [4]. In the future, we would like to perform an extensive statistical evaluation of our model with more Library indicators obtained through experimental questioners. The software environment proposed which provides a system aiding the recording of observed classroom events can support this process. Finally, we would like to extend our research by enhancing it with decision-making capabilities to help the expert identify problems and provide formative assessment.

References:

[1] 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

[2] 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.

[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] Ho, W., Xu, X., & Dey, P. K. (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. *European Journal of Operational Research*, *202*(1), 16–24.

[9] 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.

[10] Thimiakou, Vasileia, et al. "Αξιολόγηση Υπηρεσιών Βιβλιοθηκών μέσω Τεχνικού Νευρωνικού Δικτύου." *HealJournal* 1.2 (2015): 47-60.

[11] Glushkov, V. M. "Prediction based on expert opinions." *Cybernetics and Systems Analysis* 5.2 (1969): 127-129.

[12] 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.

[13] Kyrillidou, Martha. "An overview of performance measures in higher education and libraries." *Journal of Library Administration* 35.4 (2002): 7-18.

[14] Elman, J. L. (1993). Learning and development in neural networks: The

importance of starting small. *Cognition*, *48*(1), 71–99.

[15] Sopena, J. M., & Alquezar, R. (1994). Improvement of learning in recurrent networks by substituting the sigmoid activation function. In *ICANN'94* (pp. 417–420). Springer.