

ASSESSMENT OF LIBRARY USERS' FEEDBACK USING MODIFIED MULTILAYER PERCEPTRON NEURAL NETWORKS

K.G. Nandha Kumar¹ and T. Christopher²

¹Department of Computer Science, Government Arts College, Udumalpet, India

²Department of Computer Science, Government Arts College, Coimbatore, India

Abstract

An attempt has been made to evaluate the feedbacks of library users of four different libraries by using neural network based data mining techniques. This paper presents the results of a survey of users' satisfactory level on four different libraries. The survey has been conducted among the users of four libraries of educational institutions of Kovai Medical Center Research and Educational Trust. Data were collected through questionnaires. Artificial neural network based data mining techniques are proposed and applied to assess the libraries in terms of level of satisfaction of users. In order to assess the users' satisfaction level, two neural network techniques: Modified Multilayer Perceptron Network-Supervised and Modified Multilayer Perceptron Network-Unsupervised are proposed. The proposed techniques are compared with the conventional classification algorithm Multilayer Perceptron Neural Network and found better in overall performance. It is found that the quality of service provided by the libraries is highly good and users are highly satisfied with various aspects of library service. The Arts and Science College Library secured the maximum percent in terms of user satisfaction. This shows that the users' satisfaction of ASCL is better than the other libraries. This study provides an insight into the actual quality and satisfactory level of users of libraries after proper assessment. It is strongly expected that the results will help library authorities to enhance services and quality in the near future.

Keywords:

Academic Libraries, Users' Satisfaction, Data Mining, Data Classification, Artificial Neural Networks

1. INTRODUCTION

In this digital era, infinite numbers of e-books are available on internet. But people prefer visiting libraries physically to acquire relevant resources. Libraries play an essential role in the field of higher education and they do not lose their importance even in this internet age. In such scenario, modernization and improvement of libraries are inevitable. So libraries have to be enhanced and extended beyond traditional ways. To measure the existing quality and to understand users' requirement, a study is needed. By conducting a survey type study, the quality of service and user satisfaction level can be evaluated. Results of such type of study will reveal some unknown facts. Because users' expectations are changing every day and libraries are in a position to fulfill users' requirement with high quality. So this study has a significant role in assessment of quality of service and satisfaction level.

Normally such studies are carried out by applying statistical methods and nowadays many statistical software tools are available. They are highly efficient and are used by traditional researchers. But in recent decades, many data analysis methods,

techniques and tools have been invented by researchers of different functional fields: among them data mining deals with knowledge extraction from raw data. Data classification is a data mining task which categorizes the large amount of data under known labels. This is suitable to analyze survey data of academic libraries. To achieve better data classification hundreds of techniques including statistics, traditional algorithms and machine learning algorithms are available. In the last few decades machine learning methods have gained more attention among the research communities and they are in vast range of fields. Artificial neural network (ANN) is one of the machine learning methods. In data analysis researches ANNs are applied equally with regular statistical methods. They are proven in terms of accuracy and reliability. ANNs are applied for various data mining problems and tasks irrespective of fields.

Assessment of academic libraries will help us to improve the quality of service and will ensure user satisfaction. A successful feedback and evaluation system will lead to high level utilization of any service. Hence a regular assessment is required to academic libraries, so that the core objective of a library will be achieved. Generally statistical methods and tools are used to assess the quality of service and user satisfaction level. Here two novel methods are proposed and applied for assessment. The proposed methods are artificial neural network based data mining techniques. Data mining is a set of data analysis methods which discovers unknown facts from raw data. They are used to extract knowledge from large data sets. Data classification, data clustering and mining associations are primary data mining tasks. A well framed questionnaire is used to collect the data. The responses are processed through data mining technique. Data classification task is accomplished by artificial neural network technique. ANN does the assessment on collected data based on classification principle.

The purpose of these proposed techniques is to analyze library users' feedback data sets of institutions of Kovai Medical Center Research and Educational Trust (KMCRET) effectively in order to obtain better assessment results. ANN techniques are inspired by biological neurons of human brain and one of the machine learning techniques. They are applied in various fields such as data analytics, engineering, finance and business and so on. Such type of computational intelligence methods are proven in terms of reliability, accuracy, and predictability.

This paper is organized as follows. In the section 2, literature survey is presented. In the section 3, methods of data collection and data preprocessing is discussed. In the section 4, novel methods for library users' feedback assessment are explained. In addition to that, detailed algorithms and flow diagrams of proposed techniques are presented. In the section 5, results and

performance of proposed algorithms are discussed with respective parameters. Concluding comments and thought for further research are included in the section 6.

2. RELATED WORK

Data mining techniques are applied to analyze different kinds of library data [1]. Library reader management models are used to identify characteristics of readers by using book lending history [2]. Decision tree algorithms are applied to analyze the needs of users of university library [3]. K-means clustering algorithm is applied on books circulation records to analyze readers' characteristics [4]. Artificial neural network techniques are applied for library management models [5]. A neural network based assessment method is developed and applied to assess college library website [6]. The performance of association rule mining techniques are proven in the analysis of circulation records of the library. It is also used to develop book recommendation model for the library [7], [8].

A comparative analysis is made among students of medical sciences, astrophysics, astronomy and mechanics to analyze the information seeking behaviour of readers by using traditional statistical methods [9], [10] and [11]. A novel approach is proposed in [12] to improve the classification performance of a polynomial neural network (PNN) which is also called as a higher order neural network. Real coded genetic algorithm (RCGA) is applied to improve the efficiency of PNN. Tenfold cross validation is performed by the researchers by using Irvine benchmark datasets.

Two deep learning neural network methods for metagenomic classification are introduced in [13]. A recursive neural network and a deep belief network are implemented and tested with metagenomic data. In the recursive neural network, a tree is used to represent the structure of data. The main tasks are learning hierarchical structure in a metagenomic sample and classification of phenotypes. It is concluded that traditional neural networks models are more powerful than baseline models on genomic data classification.

Classification of imbalanced protein data by using ensemble classifier technique EnFTM-SVM is proposed in [5]. This is an ensemble of fuzzy total margin support vector machine (FTN-SVM). It is a three stage framework. In the first stage, protein feature extraction and representation is made. In the second stage, large numbers of distinct data sets are created. Protein sequences have multiple classes to classify and receiver operating characteristic curve is used to evaluate the classification model.

A modified technique for the recognition of single stage and multiple power quality disturbances is proposed and implemented in [14]. An algorithm which combines S-Transform based artificial neural network classifier and rule based decision tree is proposed. Different types of disturbances to the power quality are classified based on IEEE-1159 standard. To classify power quality events, two-layered feed forward neural network with sigmoid function is used. Scaled conjugate gradient back propagation algorithm is used for network training. After thorough investigation, this proposed algorithm is implemented in real time events and the validity is confirmed.

A hybrid constructive algorithm for single layer feed forward networks learning (SLFN) which is widely used for classification

and regression problems is proposed in [15]. The SLFN learning has two tasks and they are determining the network size and training the parameters. The proposed hybrid constructive algorithm can train all the parameters and determine the size of the network simultaneously. In the beginning stage, they have applied the proposed hybrid algorithm which combined Levenberg-Marquardt algorithm and least square method to train the SLFN with fixed network size. Later, they have applied the proposed hybrid constructive algorithm which follows incremental constructive scheme. In this proposed method a new neuron is randomly initialized and added with the network when the training entrapped into local minima problem. Training is continued on previous results with the added new neurons. This hybrid constructive algorithm starts the training with no hidden neurons and increases the hidden neurons one by one every time. The performance and efficiency of this novel algorithm is proved through experiments.

Implementation of a novel method for vehicle type classification by using semi-supervised convolutional neural network is proposed in [16]. Sparse Laplacian filter learning is introduced for network training in the convolutional layer and it is an unsupervised method. Beijing Institute of Technology vehicle data set which includes 9850 frontal view images of vehicle is used for the experiments. The neural network classifies the images based on vehicle types such as bus, minivan, truck etc. A literature survey on applications of data mining techniques in libraries is carried out. Applications of three core tasks of data mining i.e. classification, clustering and association rule mining were discussed in the context of library data [6], [9].

3. DATA COLLECTION AND PREPROCESSING

Questionnaires were distributed and responses were collected from the users of the four libraries of KMCRET. They are: Dr. NGP Arts and Science College Library (ASCL), Dr. NGP Institute of Technology Library (ITL), KMCH College of Nursing Library (CNL) and KMCH College of Pharmacy Library (CPL). Collected data were processed and analyzed by using artificial neural network technique. The questionnaire contains twenty four questions and distributed over four major attributes of library services such as collection of study materials, services, electronic information sources and infrastructural facilities. The responses were collected in five point scale. Responses related to a particular attribute were merged in order to simplify the data sets. The final version of the datasets consist numeric values which represent four attributes of library service.

4. PROPOSED METHODOLOGY

For this research, two multilayer feed forward neural networks were constructed by using Python programming language. The conventional multilayer perceptron neural network is modified. The Modified Multilayer Perceptron Network-Supervised (MMPN-S) and Modified Multilayer Perceptron Network-Unsupervised (MMPN-U) are proposed and implemented here for effective data classification. Generally the conventional multilayer perceptron network will have single hidden layer, but in the proposed methods two hidden layer are built and the

learning techniques are modified. Classification is a technique to assign values under certain named classes. Here, the satisfaction of user on the library service is represented by binary values. The value zero indicates dissatisfaction and one indicates satisfaction. There are four layers in the proposed neural network architecture. The first layer contains a set of input neurons. Users' responses are fed in this layer and each input neuron receives a response related to particular kind of attribute (users' feedback value). The second and the third layers are referred as hidden layers and the learning processes are completed here. The fourth layer is the output layer in which the final decision that is the classification of user's satisfaction is made.

4.1 MODIFIED MULTILAYER PERCEPTRON NETWORK SUPERVISED (MMPN-S)

Since the proposed technique is a machine learning method, training is important for a newly built neural network. After certain amount of training is given the network learn and work accordingly. MMPN-S consists four layers and all the neurons of consecutive layers are interconnected. This is a supervised artificial neural network. Every layer is built by certain number of neurons and each layer has its own task. The first layer is an input layer and the last one is output layer. Intermediate layers were used to analyze the data. The Fig.1 represents the flow of the proposed technique.

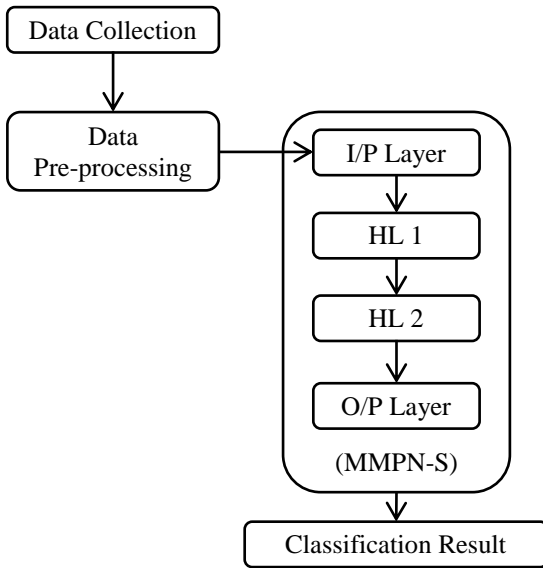


Fig.1. Functional flow of MMPN-S

4.1.1 Algorithm of MMPN-S:

Input: Library users' feedback

Output: Library users' satisfaction level

Procedure:

1. Feeding the input values to the first layer and initializing the weights.
2. Computing the square of Euclidean distance as in Eq.(1) for each input vector and finding the final unit index J, so that $D(J)$ is minimum.

$$D(j) = \sum_{i=1}^n \sum_{j=1}^m (x_i - w_{ij})^2 \tag{1}$$

3. Updating weights for all j as in (F2) and updating the learning rate α by using $\alpha(t+1)=0.5\alpha(t)$. Sending the weights to the hidden layer 1 (HL1).

$$w_{ij}(new) = (1-\alpha)w_{ij}(old) + \alpha x_i \tag{2}$$

4. Calculating the net input as in Eq.(3) and sending them to the hidden layer 2 (HL2).

$$y_{in} = b + \sum_{i=1}^n x_i w_i \tag{3}$$

5. Calculating output by applying activation function over the net input as in Eq. (4).

$$y = f(y_{in}) = \begin{cases} 1 & \text{if } y_{in} > \theta \\ 0 & \text{if } y_{in} < \theta \end{cases} \tag{4}$$

6. Updating weight and bias as in Eq.(5) and Eq.(6).

If y is not equal to t , then;

$$w_{ij}(new) = w_i(old) + atx_i$$

and

$$b(new) = b(old) + at \tag{5}$$

Else,

$$w_i(new) = w_i(old) \text{ and } b(new) = b(old) \tag{6}$$

7. Stopping the process.

4.2 MODIFIED MULTILAYER PERCEPTRON NETWORK UNSUPERVISED (MMPN-U)

MMPN-U is an unsupervised artificial neural network, hence training is not required. As in MMPN-S, MMPN-U also consists four layers; neurons of consecutive layers are interconnected, built by using certain number of neurons, and each layer has assigned separate task. The first layer is input layer and the last is output layer. Intermediate layers are used to analyze the data. Fig.2 represents the flow of the proposed MMPN-U.

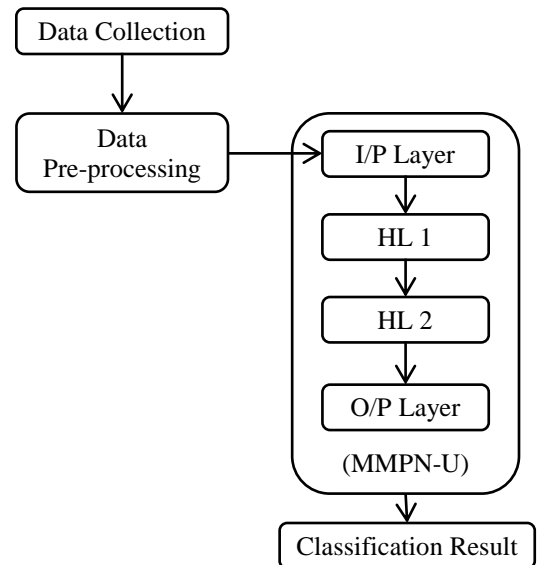


Fig.2. Functional flow of MMPN-U

4.2.1 Algorithm of MMPN-U:

Input: Library users' feedback

Output: Library users' satisfaction level

Procedure:

1. Feeding the input values to the first layer and initializing the weights.
2. Computing the square of Euclidean distance as in Eq.(1) for each input vector and finding the final unit index J, so that $D(J)$ is minimum.
3. Updating weights for all j as in Eq.(2) and updating the learning rate α . Sending the weights to the hidden layer1 (HL1).
4. Initializing the weights W and bias B as in Eq.(7).

$$w_{ij} = \frac{e_i(j)}{2} \text{ and } B_j = \frac{N}{2} \tag{7}$$

5. Calculating the net input as in Eq.(8) and Initializing the activation by using $Y_{inj}(0) = Y_{inj}$, $j = 1$ to M in the hidden layer 2 (HL2).

$$Y_{inj} = B_j + \sum_{i=1}^N x_i w_{ij} \tag{8}$$

6. Repeating step 4 and step 5 up to finding the exemplar value and stopping the process.

5. RESULTS AND DISCUSSION

The efficiency of a classifier is determined by many parameters such as accuracy, precision, recall and f-measure. In the field of artificial neural networks, the performance of a neural network is determined by its performance accuracy, error rate and processing time. Mean squared error (MSE), Root mean squared error (RMSE) and Mean absolute error (MAE) are different type of errors which are frequently used to measure the error rate of neural networks. Hence, the classification accuracy and MSE is chosen as parameters for determining the efficiency of the proposed methods.

where, TP = True Positive Classification, TN = True Negative Classification, FP = False Positive Classification, and FN = False Negative Classification.

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{X}_i - X_i)^2 \tag{10}$$

where, \hat{X}_i is a vector of classifications, and X_i is the vector of observed values corresponding to the inputs to the function which generated the classifications. The Fig.3 represents the performance accuracy of all the three neural network classifiers. MPNN-S and MPNN-U are proposed neural networks. They are modified versions of the traditional multilayer perceptron networks (MPN). The Fig.4 represents the error rate of all the three neural network classifiers. To find out the error rate, mean squared error method is followed.

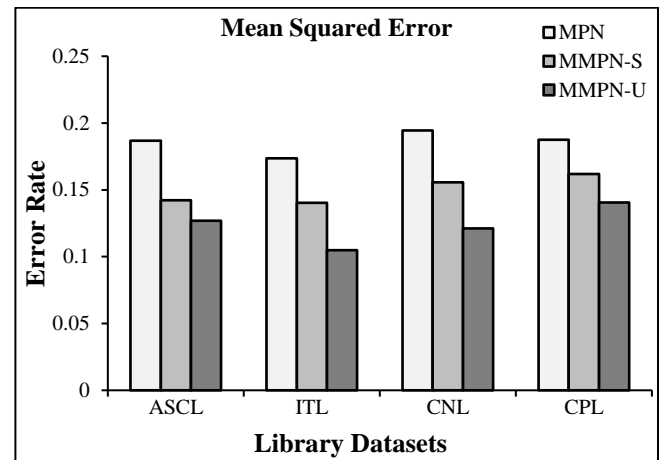


Fig.4. Comparison on Error rate of Classifiers

The average classification accuracy of MPN, MMPN-S, and MMPN-U which are derived from the experiments are 81.4%, 84.95%, and 87.63% respectively. Results indicate that the proposed techniques are more efficient in classification when compared with the traditional neural network. The average error rate of MPN, MMPN-S, and MMPN-U are 0.1856, 0.1501, and 0.1233 respectively.

6. CONCLUSION AND FUTURE SCOPE

The proposed neural network classifiers are performed efficiently with all the four data sets and produced better results than the traditional method. They have reduced the error rate also. The results of neural network approach were compared with the results of statistical approach. Apart from the classification techniques, the user satisfaction level of four libraries is found as follows. ASCL has gained 96%, ITL has gained 92%, CNL has gained 95% and CPL has gained 71%. This study indicates that, the service quality and the satisfaction level of users of KMCH college of Pharmacy is poor when compared with the other three libraries of the same group. The inference is clear, that the facilities of CPL have to be improved and upgraded in all the aspects. In addition to the regular statistical methods, neural network based data mining techniques are proposed and applied here as a novel attempt. The proposed methods have produced better results on par with statistical counterparts. Further, hybrid techniques, ensemble methods, and other soft computing

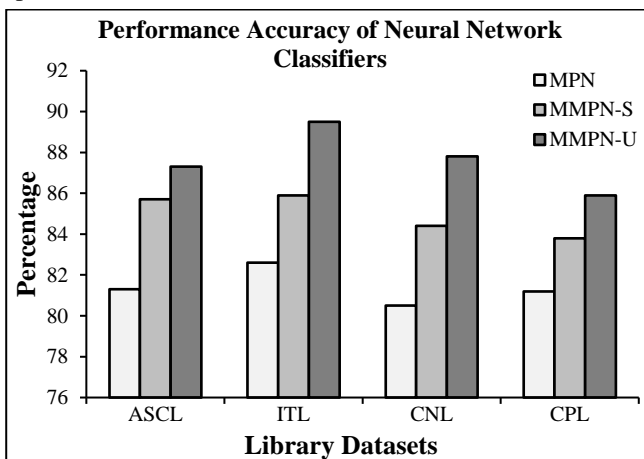


Fig.3. Comparison on Accuracy of Classifiers

The classification accuracy is calculated using Eq.(9) and the error rate is calculated using mean squared error as in Eq.(10).

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FN)} \tag{9}$$

techniques may be applied to in order to obtain more accurate classification performance.

ACKNOWLEDGMENT

The work has been supported by Mr. M. Muthu Krishnan, Librarian of Dr. NGP Arts and Science College, Coimbatore. He has supported in designing of questionnaire and data collection. The authors express their gratitude for his timely support.

REFERENCES

- [1] Pijitra Jomsri, "Book Recommendation System for Digital Library Based on User Profiles by using Association Rule", *Proceedings of 4th International Conference on Innovative Computing Technology*, pp. 130-134, 2014.
- [2] Runhua Wang, Guoquan Liu, Yi Tang, and Yan Li, "K-means Clustering Algorithm Application in University Libraries", *Proceedings of 10th IEEE International Conference on Cognitive Informatics and Cognitive Computing*, pp. 419-422, 2011.
- [3] Xing Wu, Pawel Rozycki and Bogdan M. Wilamowski, "A Hybrid Constructive Algorithm for Single-Layer Feed Forward Networks Learning", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 26, No. 8, pp. 1659-1668, 2015.
- [4] Yanhua Sun, "An Assessment Method for College Library Web Site Based on Neural Network", *Proceedings of 2nd International Conference on Intelligent Systems Design and Engineering Application*, pp. 773-775, 2012.
- [5] Hong-Liang Dai, "Imbalanced Protein Data Classification using Ensemble FTM-SVM", *IEEE Transactions on NanoBioscience*, Vol. 14, No. 4, pp. 350-359, 2015.
- [6] Veepu Uppal and Gunjan Chindwani, "An Empirical Study of Application of Data Mining Techniques in Library System", *International Journal of Computer Applications*, Vol. 74, No. 11, pp. 42-46, 2013.
- [7] Keita Tsuji, Erika Kuroo, Sho Sato, Ui Ikeuchi, Atsushi Ikeuchi, Fuyuki Yoshikane and Hiroshi Itsumura, "Use of Library Loan Records for Book Recommendation", *Proceedings of IIAI International Conference on Advanced Applied Informatics*, pp. 30-35, 2012.
- [8] Raj Kumar, Bhim Singh, D.T. Shahani, Ambrish Chandra and Kamal Al-Haddad, "Recognition of Power-Quality Disturbances using S-Transform-based ANN Classifier and Rule-based Decision Tree", *IEEE Transactions on Industry Applications*, Vol. 51, No. 2, pp. 1249-1258, 2015.
- [9] K.G. Nandha Kumar and T. Christopher, "Application of Data Mining Techniques in Academic Libraries", *International Journal of Applied Engineering Research*, Vol. 10, No. 55, pp. 1500-1502, 2015.
- [10] A.K. Pareek and Madan S. Rana, "Study of Information Seeking Behaviour and Library Use Pattern of Researchers in the Banasthali University", *Journal of Library Philosophy and Practice*, pp. 1-9, 2013.
- [11] Ping Yu, "Data Mining in Library Reader Management", *Proceedings of International Conference on Network Computing and Information Security*, pp. 54-57, 2011.
- [12] Chin-Teng Lin, Mukesh Prasath and Amit Saxena, "An Improved Polynomial Neural Network Classifier using Real-Coded Genetic Algorithm", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 45, No.11, pp. 1389-1401, 2015.
- [13] Gregory Ditzler, Robi Polikar, and Gail Rosen, "Multi-layer and Recursive Neural Networks for Metagenomic Classification", *IEEE Transactions on NanoBioscience*, Vol. 14, No. 6, pp. 608-616, 2015.
- [14] Runhua Wang, Yi Tang and Lei Li, "Application of BP Neural Network to Prediction of Library Circulation", *Proceedings of 11th IEEE International Conference on Cognitive Informatics and Cognitive Computing*, pp. 420-423, 2012.
- [15] Xing Wu, Pawel Rozycki and Bogdan M. Wilamowski, "A Hybrid Constructive Algorithm for Single-Layer Feed Forward Networks Learning", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 26, No. 8, pp. 1659-1668, 2015.
- [16] Zhen Dong, Yuwei Wu, Mingtao Pei and Yunde Jia, "Vehicle Type Classification using a Semisupervised Convolutional Neural Network", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, No. 4, pp. 2247-2256, 2015.