

PAPER • OPEN ACCESS

Method Extreme Learning Machine for Forecasting Number of Patients' Visits in Dental Poli (A Case Study: Community Health Centers Kamal Madura Indonesia)

To cite this article: E M Sari Rochman *et al* 2018 *J. Phys.: Conf. Ser.* **953** 012133

View the [article online](#) for updates and enhancements.

Method Extreme Learning Machine for Forecasting Number of Patients' Visits in Dental Poli (A Case Study: Community Health Centers Kamal Madura Indonesia)

E M Sari Rochman, A Rachmad, M A Syakur, I O Suzanti

Department of Multimedia and Network Engineering, Faculty of Engineering,
University of Trunojoyo Madura, Jl. Raya Telang, Kamal,
Bangkalan, Madura 69162 Indonesia

Email: ekamalasari3@gmail.com

Abstract. Community Health Centers (Puskesmas) are health service institutions that provide individual health services for outpatient, inpatient and emergency care services. In the outpatient service, there are several polyclinics, including the polyclinic of Ear, Nose, and Throat (ENT), Eyes, Dentistry, Children, and internal disease. Dental Poli is a form of dental and oral health services which is directed to the community. At this moment, the management team in dental poli often has difficulties when they do the preparation and planning to serve a number of patients. It is because the dental poli does not have the appropriate workers with the right qualification. The purpose of this study is to make the system of forecasting the patient's visit to predict how many patients will come; so that the resources that have been provided will be in accordance with the needs of the Puskesmas. In the ELM method, input and bias weights are initially determined randomly to obtain final weights using Generalized Invers. The matrix used in the final weights is a matrix whose outputs are from each input to a hidden layer. So ELM has a fast learning speed. The result of the experiment of ELM method in this research is able to generate a prediction of a number of patient visit with the RMSE value which is equal to 0.0426.

1. Introduction

The Center for Public Health is a functional organization with health efforts that are comprehensive, integrated, equitable, acceptable and affordable by the community. Health services provided by puskesmas is a comprehensive service that includes curative services (treatment), preventive (prevention), promotive (health improvement) and rehabilitative (health restoration). The service is addressed to all residents with no distinction between sex and age group. Puskesmas have working areas covering one sub-district or part of the kecamatan. Population density factors, area size, geography condition and other infrastructure conditions are considered in determining the working area of puskesmas. Health development efforts are selected from the list of primary health care efforts of existing puskesmas that are school health efforts, health efforts by the body, public health care efforts, work health efforts, dental and oral health efforts, mental health efforts, eye health efforts, and efforts to foster traditional medicine

Puskesmas Kamal is a government-owned health service that provides health services for the Kamal sub-district. In the outpatient service there are several poly clinics, including Nose and Throat Ear Pole (ENT), Eye Poly, Dental Poly, Child Poly, and Poly Disease. Dentistry is a form of dental and oral health services directed to the community. Actions in dentistry are removal, patching, non-specialist neurological treatment and cleansing of tartar. In addition, it also provides counseling to



patients about the importance of maintaining oral health as part of Maintain personal health, and improve knowledge and public awareness in the field of oral health.

Important factor to be considered is the availability of resources, because with the lack of resources will reduce the awareness and quality of service to patients. The existing resources, especially the material must be prepared in order to maintain the quality of services Puskesmas. Currently, the management of dental clinic of Kamal Puskesmas often experience difficulties in preparing and planning. The number of fluctuating patient visits makes it difficult for the planning party to predict how many patients will be coming so that the resources that have been provided are not in accordance with the requirements.

Forecasting technique is a technique to estimate a value in the future by taking into account past and current information. This forecasting technique has been studied in the last period which has many applications used to support in modeling forecasts such as stock forecasting, weather forecasting and so forth[1]. Standardized quantitative frameworks or techniques and mathematically explained rules are the definitions of a forecast. Forecasting is a vital part of any business organization and for any significant management decision-making because forecasting can be the basis for long-term planning. Each forecasting method has its limits and shortcomings. For example, traditional statistical methods are heavily dependent on time series data features that greatly affect the accuracy of forecasting [2].

Artificial Neural Network (ANN) is able to perform the introduction of data-based activities of the past. The past data which is studied by ANN has the ability to make decisions on data that has not been studied. ANN does not require a mathematical model but the problem is solved based on existing data.

ANN has a learning method which is called as Extreme Learning Machine (ELM), Single Hidden Layer Feedforward Neural Networks (SLFNs)[2]. ELM has attracted much attention from many researchers because it has advantages in terms of learning speed and good accuracy [3]. In addition to having universal capabilities ELM also has the ability to classify. So with this method, the output produced can approach the optimal settlement and computation time is relatively short. ELM performance can produce good generalizations in many cases and can do training faster than the current popular conventional learning algorithm [4]. ELM can be implemented with ease and the smallest weights so as to produce the smallest error value in the training process. While on the conventional network, all the parameters for the learning process then local minimum is set iteratively.

In the ELM method, input and bias weights are initially determined randomly. After that, Generalized Invers can be used to look for final weights. The matrix used in the final weights is a matrix whose outputs are from each input to a hidden layer [3]. Based on this, the study will predict how much the number of patient visits dentist in Kamal Community Health Center using ELM learning method. So by knowing how much the number of patient visits in the next period, it will help the puskesmas in providing resources needed by patients so that the optimal service.

2. Methodology

Data collection

The source of data in this study was obtained from Puskesmas Kamal. The data used is daily dental patient visit data for the last year that is year 2016 from January to December with total counted 248 data. Several studies divided the training and testing process with the composition of 80% of training data and 20% of the total data testing [5]. Therefore in this study, the data is divided into two data training and testing. January-October data is data training as much as 198 daily data, while November-December is used as data testing with the amount of 50 data daily.

Extreme Learning Machine (ELM)

ANN is one of the many applied methods in forecasting. Like the human brain, ANN consists of several neurons. Neurons will transform information received through the outgoing connection to other neurons. In neural networks, this relationship is known by the name of weight. The information is stored at a certain value on the weight [6].

Table 1. Visitors of Dental Poly in January 2016

No.	Date	Number of Patient Visits Daily
1	04/01/2016	23
2	05/01/2016	17
3	06/01/2016	12
4	07/01/2016	14
5	08/01/2016	12
6	11/01/2016	11
7	12/01/2016	8
8	13/01/2016	10
9	14/01/2016	14
10	15/01/2016	6
11	18/01/2016	2
12	19/01/2016	9
13	20/01/2016	18
14	21/01/2016	2
15	22/01/2016	15
16	25/01/2016	19
17	26/01/2016	5
18	27/01/2016	21
19	28/01/2016	16
20	29/01/2016	8

ELM is one of the models of ANN which uses Single Hidden Layer Feedforward Neural Networks (SLFNs) applications. ELM is an evolving learning technique of ANN by providing an efficient and unlimited solution to the feedforward network [7]. ELM shows how important the number of neurons in the hidden layer, but the number can be done randomly. ELM also has the ability to approach and universal classification. So ELM can produce high learning speed process. ELM has a simple level of complexity that unites popular learning algorithms and can be an appropriate solution for classification, regression, binary and multikas [2]. In the classification feature, features are randomized using a matrix so that the ELM can produce a final decision [8].

The process of normalization is the process of changing the value of a data into a value with a certain range. This process must be done first before the data is input into the input neuron in the ELM. Normalization is required because the activation function used will produce output with the data range [0,1] or [-1,1]. In this study the training data is normalized so that it has a range of values [-1.1]. The formulation of normalization is shown by equation (1),

$$X = 2 \times (xp - \min\{xp\}) / (\max\{xp\} - \min\{xp\}) - 1 \quad (1)$$

X Is the result of normalization which ranges between [-1.1], xp the actual data, $\min(xp)$ Is the minimum value, whereas $\max(xp)$ Is the maximum value in the data set. The results of normalization can be seen in table 2.

Table 2. Normalization of data

Pattern	Actual data						Normalization					
	X1	X2	X3	X4	X5	Target	X1	X2	X3	X4	X5	Target
1	23	17	12	14	12	11	0,9167	0,4167	0,0000	0,1667	0,0000	-0,0833
2	17	12	14	12	11	8	0,4167	0,0000	0,1667	0,0000	-0,0833	-0,3333
3	12	14	12	11	8	10	0,0000	0,1667	0,0000	-0,0833	-0,3333	-0,1667

4	14	12	11	8	10	14	0,1667	0,0000	-0,0833	-0,3333	-0,1667	0,1667
5	12	11	8	10	14	6	0,0000	-0,0833	-0,3333	-0,1667	0,1667	-0,5000

The most important thing in the forecasting process with the ELM is the process of training and testing. The training process aims to gain input weight, bias and output weight with a small error rate. While the testing process is to forecast based on the weight of input and output is, obtained from the training process. In the ELM training, it is used to develop the model, while in the testing process is used to evaluate the ability of ELM as a method of forecasting.

At the training stage, input and output are connected by the network. The input is supplied to the input neuron, the input layer processing by giving the activation function. The output generated from this layer is the input to the next neuron down to the neuron in the output layer. The connected relationship between neurons has a weight, which will always be updated to get a minimal error value. The error rate depends on the learning algorithm, the quality of the data and the type of network used [9].

Input on the network will be processed by a function that will add up the values of all weights. The result of the weighted sum will be compared with a threshold value through the activation function of each neuron [6]. The number of hidden neurons and activation functions must be determined to perform the training process. The trials in this study used the sigmoid log activation function because the function is most often used for forecasting problems. For the number of hidden neurons, ELM produces stable forecasting output with the number of neurons in the hidden layer 5-30 by evaluating every 5 neurons[10]. But if the output obtained from the ELM is less than optimal, it will be used to alternate transfer function or change the number of hidden neurons.

The input and output weights and biases of the hidden neurons with low error rates measured by MSE are the outputs of the ELM method. Input weight is determined randomly, while the output weight is the inverse of the hidden layer matrix and output. The new weight value is shown by table 3, while calculating the mathematical weight value can be seen in equation (2):

$$\beta = H^T T \tag{2}$$

returning the value of the output to its true value. while denormalization is shown by equation (3).

$$X = 0.5 \times (xp + 1) \times (\max\{xp\} - \min\{xp\}) + \min\{x p\} \tag{3}$$

Table 3. Weight and Bias

Weight					Bias
w1	w2	w3	w4	w5	
0,1996	0,2140	0,4003	0,4705	0,4305	0,1046
0,5208	0,8745	0,6328	0,5144	0,2818	0,7350
0,0931	0,1667	0,0552	0,6172	0,0540	0,6508
0,9123	0,3128	0,7046	0,7248	0,2397	0,7966
0,2380	0,7907	0,3283	0,2306	0,0452	0,2643

Table 4. Activation and new weight

Activation	Error	New Weight				
		w1	w2	w3	w4	w5
0,098367	0,1817	0,199858	0,218018	0,400088	0,470163	0,430179
		0,521994	0,87576	0,634027	0,515599	0,28301
		0,094538	0,166669	0,055201	0,617198	0,053805
		0,914432	0,312758	0,704574	0,724752	0,237104
		0,238698	0,790741	0,328328	0,230555	0,044491

After doing the forecasting process, the next step is to do the process of denormalization.

Denormalization is the process of

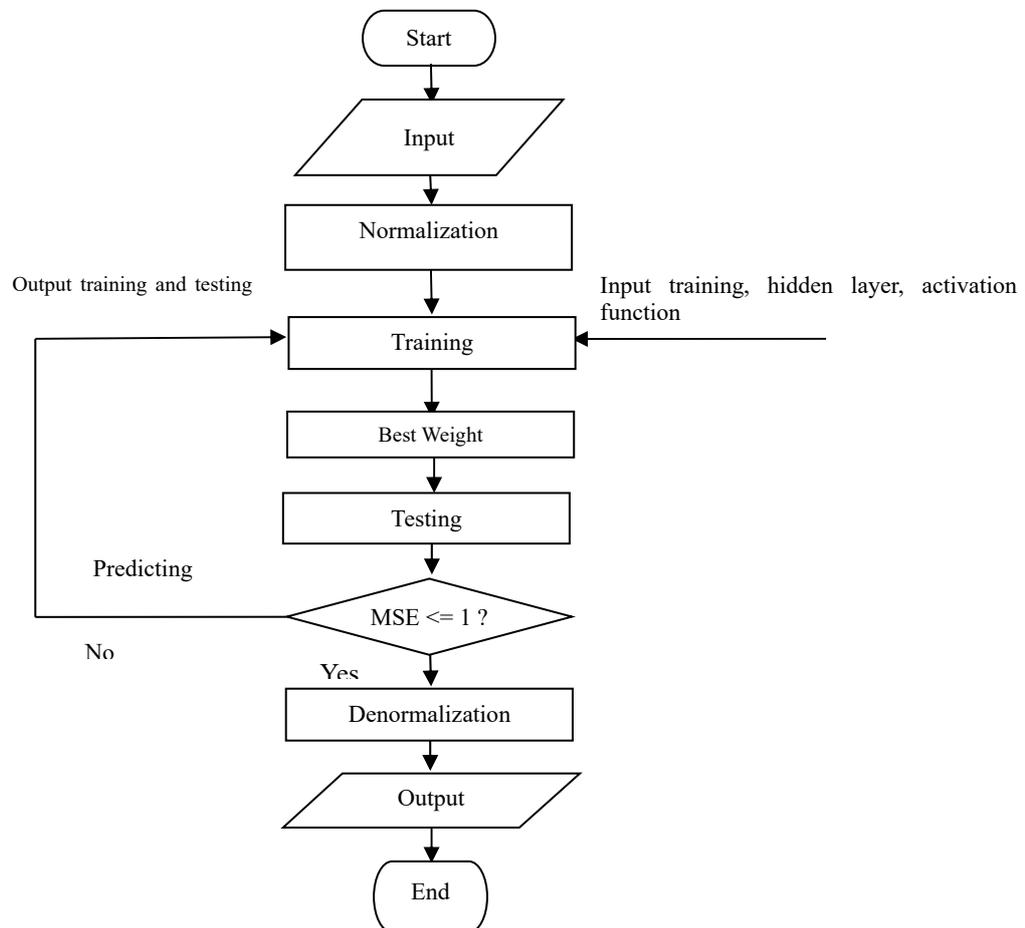


Figure 1. Flowchart *ELM*

Size Errors Forecasting

To evaluate the performance of the ELM method, this study used Root Mean Square Error (RMSE) [5]. To see if the method has been used is sufficient to predict a data then made an error measurement, because there is no forecasting method that can predict future data appropriately. The more precise a method of generating predictions, then the resulting error rate is smaller.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (X_t - F_t)^2}{n}} \quad (4)$$

The RMSE is obtained by subtracting the value of X_t which is the expected forecast value of period t with the F_t value which is the forecasting value of the system in the squared period t . Last is leveling as much n the amount of data.

3. Results and Discussion

The methods performed and the data used are described in the methodology section. This section describes accurate and effective forecasts that can assist decision makers in planning the amount of resources which needed in puskesmas especially dental poly. So that the service to patient

can be maximum. In the implementation stage, it is tested by using the sigmoid activation function with the hidden number of combining 5,10,15, and 20. In table 3, it shows the comparison of learning speed and error size using RMSE.

Table 5 RMSE comparison uses a combination of hidden layer neuron changes

Hidden Layer	Time	RMSE
5	0.0468	0.0851
10	0.0313	0.0638
15	0.0312	0.0426
20	0.2184	0.1064

4. Conclusions

The conclusions that can be drawn from this research are as follows:

1. The test results using different number of hidden neurons, then the ELM produces the optimal output with the number of hidden fifteen neurons.
2. ELM produces forecasting output with low error rate of 0.0426
3. Learning speed required by ELM is very short, that is average 0.0312 seconds.
4. The output of the ELM is determined by parameter determination such as the activation function and the number of hidden neurons.

5. Acknowledgment

The authors state that there are no interest issues with the institution listed in this paper. We would like to thank our colleagues at the university of Trunojoyo Madura - Indonesia, Network and Multimedia laboratory team who helped author in completing this research. So, the authors has a chance to use the equipment in the laboratory.

References

- [1]. N. Liu, S. Ren, T.-M. Choi, C.-L. Hui and S.-F. Ng, "Sales Forecasting for FashionRetailing Service Industry: A Review," Hindawi Publishing Corporation Mathematical Problems in Engineering, p. 9, 2013.
- [2]. H. Guang-Bin, Z. Hongming, D. Xiaojian and Z. Rui, "Extreme Learning Machine for Regression and Multiclass Classification," IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, pp. 513-528, 2012.
- [3]. B. Yadav, Sudheer, S. MATHUR and J. ADAMOWSKI, "Assessing the suitability of extreme learning machines (ELM)for groundwater level prediction," JOURNAL OF WATER AND LAND DEVELOPMENT, Vols. I-III, no. 32, pp. 103-112, 2017.
- [4]. S. Sajjadi, S. Shamsirband, M. Alizamir, P. L. Yee, Z. Mansor, A. A. Manaf, T. A. Altameem and A. Mostafaeipour, "Extreme learning machine for prediction of heat load in district heating systems," Energy and Buildings 122 , p. 222–227, 2016.
- [5]. J. Zhang and W. Ding, "Prediction of Air Pollutants Concentration Based on an Extreme Learning Machine: The Case of Hong Kong," International Journal of Environmental Research and Public Health, vol. 14, p. 114, 2017.
- [6]. S. Kusumadewi, Artificial Intelligence, Jogjakarta: Graha Ilmu, 2003.
- [7]. E. Cambria and G.-B. Huang, "Extreme Learning Machines," in TRENDS & CONTROVERSIES, M. M. L. erik cambria, Ed., IEEE Computer Society, 2013, pp. 1541-1672.
- [8]. J. Tang, C. Deng and G.-B. Huang, "Extreme Learning Machine for Multilayer Perceptron," IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, vol. 27, no. 4, pp. 809-821, 2016.
- [9]. V. Gosalia, M. P. Andhare, P. Patil, A. Chorage and R. Gound, "Electricity Price Forecasting Using

ELM-Tree Approach,” International Research Journal of Engineering and Technology (IRJET), vol. 3, no. 1, pp. 1209-1215, 2016.

- [10]. F. Mateo, J. J. Carrasco, M. Millán-Giraldo, A. Sellami, P. Escandell-Montero, J. M. Martínez-Martínez and E. Soria-Olivas, “Machine Learning Techniques for Short-Term Electric Power Demand Prediction,” Bruges (Belgium), 2103.

Export all Add all to list Set document alert Set document feed

Document title	Authors	Year	Source	Cited by
Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster	Syakur, M.A., Khotimah, B.K., Rochman, E.M.S., Satoto, B.D.	2018	IOP Conference Series: Materials Science and Engineering	0
View abstract ▾ Related documents				
Method Extreme Learning Machine for Forecasting Number of Patients' Visits in Dental Poli (A Case Study: Community Health Centers Kamal Madura Indonesia)	Sari Rochman, E.M., Rachmad, A., Syakur, M.A., Suzanti, I.O.	2018	Journal of Physics: Conference Series	0
Open Access				
View abstract ▾ Related documents				
Determination System of Food Vouchers for the Poor Based on Fuzzy C-Means Method	Anamisa, D.R., Yusuf, M., Syakur, M.A.	2018	Journal of Physics: Conference Series	0
Open Access				
View abstract ▾ Related documents				

Display: 20 results per page 1 ^ Top of page

The data displayed above is compiled exclusively from documents indexed in the Scopus database. To request corrections to any inaccuracies or provide any further feedback, please use the [Author Feedback Wizard](#).

Activate Windows
Go to PC settings to activate Windows