Applied Computing and Informatics xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Applied Computing and Informatics

journal homepage: www.sciencedirect.com

Original Article

Genetic-neuro-fuzzy system for grading depression

Kumar Ashish [a], Anish Dasari [a], Subhagata Chattopadhyay [b], Nirmal Baran Hui [a,*]

[a] Department of Mechanical Engineering, National Institute of Technology, Durgapur, West Bengal 713209, India
[b] Sr. Faculty, Indus Training & Research Institute (ITARI), Billapura Cross, Sarjapur, Bangalore 562125, Karnataka, India

ARTICLE INFO

Article history:
Received 11 April 2017
Revised 15 May 2017
Accepted 18 May 2017
Available online xxxx

Keywords:
Depression modeling
Neuro-fuzzy system
Genetic algorithm
Automated tuning
Identification of redundant rules

ABSTRACT

Main aim of this study is to develop a software prototype tool for grading and diagnosing depression that will help general physicians for first hand applications. Identification of key symptoms responsible for depression is also another important issue considered in this study. It involves collection of data taken from patients through doctors. Due to several reasons, collection of data in Indian scenario is extremely difficult and thus this tool will be very handy and useful for general physicians working at remote locations. Also, it is possible to collect a data pool through this software model. An intelligent Neuro-Fuzzy model is developed for this purpose. Performance of the said model has been optimized through two approaches. In Approach 1, where a back-propagation algorithm has been considered and in Approach 2, Genetic Algorithm has been used. The model is trained with 78 data and validated with 10 data. Approach 2 superseded Approach 1 in terms of diagnostic accuracy. Therefore, it can be said that the soft computing-based diagnostic models could assist the doctors to make informed decisions. Data for training and validation for this purpose has been collected during 2004–2005 from a Government mental hospital in India.

1. Introduction

WHO says "depression is a common illness worldwide" and it is a growing risk [1–3]. There could be several factors contributing to depression and it is closely related to physical health [1–3]. There exist known and effective treatment procedures, which reach less than half of the sufferers [1,4]. There are several reasons such as lack of trained professionals and lack of resources [1], silent progression [5], social stigma [6] and perception difference among doctors leading to 'under' or 'over' diagnoses [7]. So, diagnosis has to be proper to reduce the disease load [8].

SC techniques are applied for screening and diagnosing the different ailments [9,10] in the last couple of decades. Clustering techniques alone are applied for grading of depression [11]. However, Chattopadhyay et al. [12] combined fuzzy logic with clustering for capturing the symptoms psychiatric diseases of human being. They later developed a NN-based toolbox for grading of adult depression [13] and classified it into three categories namely, 'mild', 'moderate', and 'severe' [13]. Tai and Chiu [14] used RBFN to understand the reasons for suicidal tendencies of Taiwanese soldiers. Chattopadhyay [15] developed a fuzzy-based automated model for grading depression. Later stage, Chattopadhyay et al. [16] tried to minimize the overlapping among the three different grades of depression. Regression analysis is also used for the similar purpose by Chattopadhyay and Acharya [17].

It is understood from the literature that grading of depression is a difficult task. It is mainly because of the non-availability of the data. It may of different reasons which are indicated below.

(i). social stigma,
(ii). majority of the patients do not approach to psychiatrists first, they approach to general physicians and when it becomes severe then comes to the psychiatrists,

Abbreviations: BP, Back Propagation; DB, Data Base; FS, Feeling Sad; FL, Fuzzy Logic; GP, General Physician; GA, Genetic Algorithm; H, Hypersomnia; IN, Insomnia; LA, Loss of appetite; LP, Loss of Pleasure; m, Mild; M, Moderate; MSE, Mean Squared Error; NN, Neural Network; PA, Psychomotor Agitation; RBFN, Radial Basis Function Network; RB, Rule Base; s, Severe; SC, Soft Computing; WL, Weight Loss; WHO, World Health Organization.

* Corresponding author.
E-mail addresses: ashish5srivastava@gmail.com (K. Ashish), anish.d47@gmail.com (A. Dasari), subhagata.chattopadhyay@itari.in (S. Chattopadhyay), nirmalhui@gmail.com (N.B. Hui).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.aci.2017.05.005
2210-8327/

(iii). non existence of systems of collecting data by the doctors. For the first time reporting to the doctors, majority of the patients are prescribed medicines. As a result true picture of the patients are not available to the doctors,
(iv). inferior health management system.

Therefore, basic purpose of our study is to develop a prototype tool which forecasts depression better than the general physicians. In this context, SC-based techniques are found to provide formidable solutions. However, performance of SC-based techniques depends on its formulation and optimization of the same. Out of different SC-based models, neuro-fuzzy models are finding wide applications in the recent times in different field with high prediction accuracies [18,19]. Development of a neuro-fuzzy-based software prototype model has been made in this study. It has been observed that, the performance of neuro-fuzzy systems depends on its DB and RB [20]. In majority of the case, user defined the same, which in no sense becomes optimal. Some researchers tried to combine different SC-based techniques. Therefore, the main contribution of this work is the development of a neuro-fuzzy-based prototype model for grading depression. Later, performance of the developed approach was made through two optimization methods, namely back-propagation algorithm and genetic algorithm. The model will be used by the general physicians for collecting more data. After proper validation with large set of data, it will be given for use to the psychiatrists/experts of mental health problems. Presently, the model is validated using data given by the clinicians. However, the model is evolutionary in nature and thus, it might handles uncertainty, impreciseness present (which are predominant) in the system in much broader way.

The developed neuro-Fuzzy model and its optimization procedures using two approaches are described in Section 2 and results have been analyzed in Section 3. Final outcomes of the work are discussed and future directions are mentioned in Section 4.

2. Methodology

Collecting depression data is very difficult. It is because of different reasons such as (i) prevailing social stigma, especially in the developing nations, (ii) difficulties in accessing the psychiatrists as they are very few in numbers, (iii) non-existence of system for collecting data by the doctors and so on. Together it leads to real a challenge in obtaining the clean data. The basic aim of this study is to build a software toolbox that will be of help for the GPs, who is the first point of contact to the patients and data will be possible to be collected.

2.1. Data collection

A study is made on 312 patients reported to one of the famous mental hospital in India in the year 2004–2005. Those patients reported to the hospital for the first time and they did not take any medicines before reporting to the hospital. They are then grouped in two. First group belongs to the patients who have suicidal ideations and require immediate treatment. However, patients belonging to the second group (total 88 patients) were further analyzed towards the severity of depression. Therefore, there exist 88 data; out of which 78 has been used for training/optimizing the neuro-fuzzy architecture and rest 10 have been used for testing the performance of optimized architecture. Seven common symptoms (FS, LP, WL, IN, H, LA, PA) are then observed by three senior psychiatrists. Doctors are then requested to quantify the symptoms and severity of depression with which those 88 patients are suffering.

2.2. Neuro-fuzzy model construction

A Mamdani-type [21] NN-FL model was developed by Hui et al. [20] is used in this study for grading depression (refer to Fig. 1). Interested readers may go through the paper of Hui et al. [20,22] for detailed discussion on the development of NN-FL system. In the present study, seven inputs representing the seven different symptoms and one output representing the depression is used. Three different grades of the inputs are considered (*m*, *M* and *s*). The data base of the NN-FL system is shown in Fig. 2. Since there are three possibilities (i.e., mild, moderate and severe) of the load of each of seven inputs, there will be $3^7 = 2187$ numbers of possible input combinations, for which grades of depression would vary as mild, moderate and severe, based on the symptom load. Manually constructed rules are presented in Table 1 and one particular rule (say the first rule of the above table) looks like:

If *FS* is *m* & *LP* is *m* & *WL* is *m* & *IN* is *m* & *H* is *m* & *LA* is *m* & *PA* is *m* THEN *Depression* is *M*.

For more details related to rule based selection, identification of relevant rules, optimization using GA and gravitational search algorithm, interested readers may go through [23–31].

The performance of the above NN-FL system largely depends on those rules and different weights such as V_1, V_2, V_3, V_4, V_5, V_6, V_7 and W_1. The NNFL system will have variable structure only when rule base of the FLC is optimized i.e., some rules are deleted. There exist a large number of literatures for developing an optimal NNFL system. In the present study two different approaches have been adopted. Both these two approaches have been explained below in brief.

2.3. Approach 1: Tuning using a BP algorithm

In this approach, RB is kept unaltered during its training. Only the weights V_1, V_2, V_3, V_4, V_5, V_6, V_7 and W_1 are optimized using BP algorithm [32]. It is usually implemented through the minimization of Mean Squared Error (E) in prediction, as given below.

$$E = \frac{1}{2C}\sum_{c=1}^{C}(T_{5c} - O_{5c})^2 \tag{1}$$

where T_{5c} and O_{5c} represent the target and model predicted outputs of c^{th} training scenario.

The change in W_1 that is ΔW_1 is determined as

$$\Delta W_1 = -\eta \frac{\partial E}{\partial W_1} \tag{2}$$

where η indicates the learning rate and $\frac{\partial E}{\partial W_1}$ is expressed as follows:

$$\frac{\partial E}{\partial W_1} = \frac{\partial E}{\partial O_{5c}} \frac{\partial O_{5c}}{\partial I_{5c}} \frac{\partial I_{5c}}{\partial W_1} \tag{3}$$

Similarly, the change in V_i that is ΔV_i can be calculated as follows

$$\Delta V_i = -\eta \frac{\partial E}{\partial V_i} \tag{4}$$

where $\frac{\partial E}{\partial V_i}$ can be calculated as:

$$\frac{\partial E}{\partial V_i} = \frac{\partial E}{\partial O_{5c}} \frac{\partial O_{5c}}{\partial I_{5c}} \frac{\partial I_{5c}}{\partial O_{4P}} \frac{\partial O_{4P}}{\partial I_{4P}} \frac{\partial I_{4P}}{\partial O_{3K}} \frac{\partial O_{3K}}{\partial I_{3K}} \frac{\partial I_{3K}}{\partial O_{2M}} \frac{\partial O_{2M}}{\partial I_{2M}} \frac{\partial I_{2M}}{\partial V_i} \tag{5}$$

2.4. Approach 2: Tuning using genetic algorithm

Approach 1 suffers from the local minima problem. Also, in Approach 1, only the weights are optimized. Therefore, in Approach 2, optimization of both DB and RB is attempted using a GA. GA with 4454-bits long string is used for this purpose. Data