

# Artificial Neural Network (ANN) Model to Predict Depression among Geriatric Population at a Slum in Kolkata, India

ARKAPRABHA SAU<sup>1</sup>, ISHITA BHAKTA<sup>2</sup>

## ABSTRACT

**Introduction:** Depression is one of the most important causes of mortality and morbidity among the geriatric population. Although, the aging brain is more vulnerable to depression, it cannot be considered as physiological and an inevitable part of ageing. Various sociodemographic and morbidity factors are responsible for the depression among them. Using Artificial Neural Network (ANN) model depression can be predicted from various sociodemographic variables and co morbid conditions even at community level by the grass root level health care workers.

**Aim:** To predict depression among geriatric population from sociodemographic and morbidity attributes using ANN.

**Materials and Methods:** An observational descriptive study with cross-sectional design was carried out at a slum under the service area of Bagbazar Urban Health and Training Centre

(UHTC) in Kolkata. Among 126 elderlies under Bagbazar UHTC, 105 were interviewed using predesigned and pretested schedule. Depression status was assessed using 30 item Geriatric Depression Scale. WEKA 3.8.0 was used to develop the ANN model and test its performance.

**Results:** Prevalence of depression among the study population was 45.7%. Various sociodemographic variables like age, gender, literacy, living spouse, working status, personal income, family type, substance abuse and co morbid conditions like visual problem, mobility problem, hearing problem and sleeping problem were taken into consideration to develop the model. Prediction accuracy of this ANN model was 97.2%.

**Conclusion:** Depression among geriatric population can be predicted accurately using ANN model from sociodemographic and morbidity attributes.

**Keywords:** Cross-sectional design, Geriatric depression scale, Machine learning

## INTRODUCTION

The government of India, defines 'senior citizen' in "National Policy on Older Persons" (1999) as a person who is 60-year-old or above [1]. They are 8% of the total population as per 2011 Indian census. The numbers of elderly persons are increasing day by day not only in India but worldwide also, due to various factors like better health care delivery system, better diagnostic and treatment facility. At present, China has the largest geriatric population in the world followed by India [2]. By the year 2050 in India, elderly population will likely go up to 20% of total population [3].

Ageing is a physiological process, associated with bodily changes like wrinkling of skin, greying of hair, along with some internal organs related changes resulting decline in body function. Although, elderly people are more vulnerable to depression due to ageing brain, it cannot be treated as physiological and an inevitable part of ageing [4]. Depression is a common cause of morbidity, mortality among the geriatric population. Depression is merely presence of just a low mood; it is a serious mental health issue. It differs in severity from individual to individual. People with severe depression find it really hard to cope up with day-to-day problems. Mild depression will decrease quality of life, and without proper treatment and support, may become more serious problem in future. Depression is both under diagnosed and undertreated, especially among geriatric population, as symptoms of depression among them, like loneliness, sleep deprivation, are often overlooked because they coincide with other late life problems [5]. World Health Organization (WHO) estimates that depression occurs in 7% of the geriatric population globally, whereas, in India, 25% of elderly population is suffering from it [2]. Literature shows that prevalence of depression

among geriatric population were 46.9% at Chetla slum in Kolkata conducted by Dasgupta A et al., and 41.1% by Arumugam B et al., at a slum in Chennai [6,7]. As the number of elderly is increasing in the India, the burden of depression is going to be substantial on the society in near future.

Geriatric Depression Scale (30 item) [8] is a world-wide validated and accurate tool to diagnose depression among the elder population. This tool is mainly used by the doctors or psychologists. An easy, accurate alternative tool is necessary to predict depression at earliest at community by the grass root level health worker.

ANN model can be used to develop such an alternative tool. ANN is a data driven approach in machine learning technology for predictive analysis [9]. It is one of the most recent advancement in the field of nature inspired algorithms [10]. It is a sophisticated modelling and prediction tools capable of identifying complex relationships between predictor (input) and predicted (output) variables. It is inspired by biological system, specially the complex neuronal network of human brain. Like human brain, ANN analyse data in terms of pattern recognition [11]. For prediction purpose knowledge is stored and used in logical sequence by the brain due to the result of complex multiple inter-neuronal connective activities. Similarly, information is stored in ANN as weights of interconnections between neural processing units in distributed manner. Scientific literature shows that like other non-communicable diseases, depression is also multi factorial and various sociodemographic factors (like age, gender, living spouse, type of family, education, working status, personal income, and substance abuse) and morbidity condition (like hearing problem, visual problem, mobility problem and sleeping problem) significantly associated with it [5-7,12-14].

With this background knowledge, this research work has been carried out to develop an automated computer based tool to predict depression among the geriatric population with the help of various sociodemographic and morbidity factors. This can be used as a screening tool to identify those elderly who are probably depressed and to refer them to a psychiatrist for consultation at an early stage. To predict depression among geriatric population from socio demographic and morbidity attributes using ANN model.

## MATERIALS AND METHODS

It was an observational descriptive study with cross-sectional design. Then ANN was applied as a machine learning classifier for predictive modelling. The study was conducted at the service area of Bagbazar Urban Health and Training Centre (UHTC), Kolkata, West Bengal, India, between April 2016 to July 2016. It is also the urban field practice area of Department of Community Medicine, RG Kar Medical College and Hospital, Kolkata. All the elderly people living permanently at that slum during the study period were considered as the study population. As per records, available at Bagbazar UHTC, there were total 126 elderly people living there permanently during the study period. Individuals who were not able to comprehend the study schedule due to illness or dementia were excluded from the study. Those who expressed unwillingness to become a study participant were also excluded from the study.

Sample size was determined with finite population correction. The formula is  $n = Nz^2pq / \{d^2(N-1) + z^2pq\}$  [15], here  $n$ =sample size,  $N$ =population size/total no of person in geriatric age group under the service area of Bagbazar UHTC i.e., 126,  $z$ =1.96 at 95% confidence level,  $p$ =prevalence of depression among elderly person (46.9%) [6],  $q$ =(1- $p$ ), and considering  $d$ = Absolute precision/margin of error (5%). Using the formula and above values, calculated sample size was 95. Considering non response rate as 10%, the final sample size was 105. Each of them was selected by simple random sampling without replacement technique using random number table. The selected elderly persons were interviewed at their home using pre designed and pre tested schedule after taking informed consent.

Ethical clearance was obtained from the Institutional Ethics Committee of RG Kar Medical College and Hospital, Kolkata, West Bengal, India.

### Study Tool

A predesigned and pretested schedule was used for this study. It consisted of three parts. First part was sociodemographic information i.e., name, age, gender (Male/Female), living spouse (Yes/No), type of family (Nuclear/Joint), literacy (Literate /Illiterate), occupational status (presently working/ presently not working), personal income (Yes/No), and substance abuse (Yes/No); second part was information about presence or absence of co morbid conditions like hearing problem (Yes/No), visual problem (Yes/No), mobility problem (Yes/No) and sleeping problem (Yes/No); third part was 30 items Geriatric Depression Scale (GDS) [6] for assessment of depression (Normal/Depressed). GDS score 0-9 is normal and 10-30 is depressed [8].

Morbid conditions were assessed subjectively as per response made by the study participants. Hearing problem was considered to be present if there was any hearing difficulty with or without any hearing aids. Visual problem was considered to be present if there was any problem in visual acuity due to any reason after best possible correction. Mobility problem was considered to be present if there was difficulty in moving around for activity of daily living due to any reason. Sleeping problem was considered to be present if any problem related to falling asleep was reported by the study participants.

## Data Analysis and Disease Prediction System

ANN predictive modelling, its performance analysis and development of a depressing prediction system was done using WEKA version 3.8.0 software. Weka is open source data mining software in the field of machine learning technology [9]. Primary data, collected by interviewing the selected 105 elderly people was used as a training data set for data mining in Weka. Then, the predictive model was build using ANN (Multilayer Perceptron, A feed forward ANN Model) as a classifier in the Weka. Performance of the build model was automatically analysed using 10-fold cross validation method.

### ANN Model

Multilayer Perceptron (MLP), a feed forward ANN model was built in WEKA for classification with input variables i.e., age, gender (Male/Female), living spouse (Yes/No), type of family (Nuclear/Joint), Literacy (Literate/Illiterate), occupational status (presently working/ presently not working), personal income (Yes/No), and substance abuse (Yes/No), hearing problem (Yes/No), visual problem (Yes/No), mobility problem (Yes/No) and sleeping problem (Yes/No). The ANN model was designed with seven interconnected neurons in one hidden layer. The learning rate of the network was set at 0.3 for 500 iterations. Output was classified as normal i.e., non depressed and depressed.

### Prediction of Depression with a Test Data Set

A test data set were prepared selecting 30 geriatric individual by simple random sampling from the primary data set. Then this test data set were put into the ANN model in Weka with all input variables but output variable (Normal i.e., non depressed or depressed) were left blank. Then prediction was done by the ANN model which was previously trained with primary data set.

At the end of this research work a Microsoft Windows based application was designed using this trained ANN model to predict

| Variables                  | Depressed | Non-Depressed |
|----------------------------|-----------|---------------|
| <b>Gender</b>              |           |               |
| Male                       | 12        | 48            |
| Female                     | 36        | 9             |
| <b>Age</b>                 |           |               |
| Less than 70 years         | 38        | 50            |
| ≥ 70 years                 | 10        | 7             |
| <b>Living spouse</b>       |           |               |
| Yes                        | 18        | 51            |
| No                         | 30        | 6             |
| <b>Type of Family</b>      |           |               |
| Nuclear                    | 7         | 14            |
| Joint                      | 41        | 43            |
| <b>Literacy</b>            |           |               |
| Literate                   | 19        | 54            |
| Illiterate                 | 29        | 3             |
| <b>Occupational Status</b> |           |               |
| Presently working          | 41        | 21            |
| Presently not working      | 7         | 36            |
| <b>Personal Income</b>     |           |               |
| Yes                        | 15        | 47            |
| No                         | 33        | 10            |
| <b>Substance abuse</b>     |           |               |
| Yes                        | 23        | 32            |
| No                         | 25        | 25            |

**[Table/Fig-1]:** Comparison between depressed and non-depressed study participants with various sociodemographic factors. (n=105).

depression. This software was in its prototype version. This was easy to use; require very basic computer skill like data entry and computational time was less than one second.

## RESULTS

The prevalence of depression among the study population was 45.7% at Bagbazar Slum in Kolkata. Mean ( $\pm$ standard error) age of the study participants were 66.6( $\pm$ 5.6) years. Among them 57.1% were male, 66.7% had living spouse, 73.3% were living in a joint family, 69.5% were literate, 58.1% had personal income, 54.3% were engaged in some kind of occupational activity, and 52.4% had a habit of substance abuse. Morbidity profile of the study population showed that 51.4% were suffering from visual impairment, 54.2% suffering from mobility

| Variables               | Depressed | Non-Depressed |
|-------------------------|-----------|---------------|
| <b>Hearing Problem</b>  |           |               |
| Yes                     | 18        | 5             |
| No                      | 30        | 52            |
| <b>Visual Problem</b>   |           |               |
| Yes                     | 45        | 9             |
| No                      | 3         | 48            |
| <b>Mobility Problem</b> |           |               |
| Yes                     | 37        | 20            |
| No                      | 11        | 37            |
| <b>Sleeping Problem</b> |           |               |
| Yes                     | 11        | 9             |
| No                      | 37        | 48            |

**[Table/Fig-2]:** Comparison between depressed and non-depressed study participants with various morbidity factors. (n=105).

problem, 21.9% from hearing difficulty and 19.1% from sleeping disorder. A comparison between depressed and non-depressed or normal study participants with various sociodemographic [Table/Fig-1] and morbidity factors [Table/Fig-2] are tabulated.

### Training the ANN Model and its Performance Analysis

ANN model was trained and tested on the primary data set with 10-fold cross validation method. Result showed that among 105 instances it correctly classified 102 (97.14%) instances and incorrectly classified instances 3 (2.86%) with a Kappa statistic 0.94, Mean absolute error 0.03, and root mean squared error 0.14. Accuracy of the model showed weighted precision 97.2%, True positive rate 97.1%, false positive rate 3.1%, recall 97.1%, F measure 0.97, MCC (Matthews Correlation Coefficient) 0.943 and ROC were 0.99.

### Performance of ANN Model with the Test Data Set

Result showed that the trained ANN model correctly diagnosed all 30 instances (12 depressed and 18 non depressed) with a Kappa statistic 1.

## DISCUSSION

Automated disease prediction system using artificial intelligence is an upcoming concept in health care delivery system. Diagnosis of depression using machine learning technology is one such application of artificial intelligence. A study conducted in India by Mukherjee S et al., showed that ANN model can diagnose depression from various somatic symptoms, with reasonable accuracy with nominal root mean squared error (0.11) [16]. In our study, we are able to predict depression from socio economic and somatic symptoms with 97.2 % accuracy and root mean squared error was 0.14. Another study conducted by Khalili M et al., at Iran concluded that ANN modelling approach is a better way to diagnose depression and mania with acceptable accuracy with minimal error (0.13) [17]. Study conducted

by Tortajada S et al., in Spain predicted postpartum depression using ANN model with a reasonable accuracy (82%) [18]. A study, conducted by Bhakta I et al., to compare different machine learning classifiers for predicting depression among the senior citizens, found that multilayer perception, a feed forward ANN model, had the predicting accuracy of 87% [19]. In accordance with the published articles, our study also recognises the power of ANN model in prediction of depression, especially among the older individuals.

## LIMITATION

This study was carried out in only one slum area in Kolkata, so an ANN model based on primary data obtained from large cross-sectional studies involving various parts of the country would have been close to generalization. Moreover, this will help to built a robust prediction system.

## CONCLUSION

This ANN model based automated geriatric depression prediction system will help health care providers, to identify those geriatric persons who are more at risk of being depressed. Moreover, prediction can be made within seconds using sociodemographic and morbidity factors easily. At risk individuals can be referred to consultant psychiatrist for farther follow up and management. This will help to decrease the burden of morbidity and mortality due to depression among the geriatric population.

## ACKNOWLEDGEMENTS

We are thankful to the lady medical officer, members of the staff of Bagbazar UHTC and study participants for their kind and unconditional support. We are also thankful to our teachers for proper guidance.

## REFERENCES

- [1] National Policy on Older Persons. New Delhi: Ministry of Social Justice and Empowerment; Government of India; 1999.
- [2] Santosh A, Kumar A, Rao BV, Patil RS. Magnitude of depression among geriatric population and factors associated with it in the urban slum, Bashanagar, field practice area of SSIMS and RC Davangere"-a cross sectional study. *International Journal of Medical and Pharmaceutical Sciences*. 2014;4(7):20-26.
- [3] Ingle GK, Nath A. Geriatric health in India: Concerns and solutions. *Indian Journal of Community Medicine*. 2008;33(4):214.
- [4] Srivastava S, Kumar A, Khurana H, Tiwari SC, Akbar S. Short-term course and outcome of late-life depression. *Journal of Geriatric Mental Health*. 2015;2(2):96.
- [5] Chauhan P, Kokiwar PR, Shridevi K, Katkuri S. A study on prevalence and correlates of depression among elderly population of rural South India. *International Journal of Community Medicine and Public Health*. 2016;3(1):236-39.
- [6] Dasgupta A, Mukhopadhyay M, Das S, Majumdar S, Das A. Are our elderly really depressed?: a cross sectional study on depression among geriatric population in a slum of Kolkata. *International Journal of Health Sciences and Research (IJHSR)*. 2014;4(6):25-30.
- [7] Arumugam B, Nagalingam S, Nivetha R. Geriatric depression among rural and urban slum community in Chennai: A cross sectional study. *J Evol Med Dent Sci*. 2013;3:795-801.
- [8] Yesavage JA, Brink T, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatric Research*. 1983;17(1):37-49.
- [9] Alpaydin E. *Introduction to machine learning*: MIT press; 2014.
- [10] Yang X-S. *Nature-inspired computation in engineering*: Springer; 2016.
- [11] Roy S, Chakraborty U. *Introduction to soft computing*: Dorling Kindersley; 2013.
- [12] Gupta M, Lehl S, Boparoy NS, Katyaj R, Sachdev A. A study of prevalence of depression in elderly with medical disorders. *Journal of the Indian Academy of Geriatrics*. 2010;6:18-22.
- [13] Nair SS, Hiremath S, Nair SS. Depression among geriatrics: Prevalence and associated factors. *International Journal of Current Research and Review*. 2013;5(8):110.
- [14] Vishal J, Bansal R, Swati P, Bimal T. A study of depression among aged in Surat city. *National Journal of Community Medicine*. 2010;1(1):47-49.
- [15] Daniel WW, Cross CL. *Biostatistics: Basic concepts and methodology for the health sciences*. 10<sup>th</sup> ed. New York: John Wiley & Sons; 2015.
- [16] Mukherjee S, Ashish K, Baran Hui N, Chattopadhyay S. Modeling depression data: feed forward neural network vs. radial basis function neural network. *American Journal of Biomedical Sciences, American J Biomed Sci*. 2014;6(3):166-74.
- [17] Khalili M, Ghasemi MM. The effect of neural network models on diagnosis of bipolar disorder. *International Journal of Engineering Innovation & Research*. 2014;3(9):700-04.

[18] Tortajada S, García-Gomez JM, Vicente J, Sanjuán J, de Frutos R, Martín-Santos R, et. al. Prediction of postpartum depression using multilayer perceptrons and pruning. *Methods of Information in Medicine*. 2009;3:291-98.

[19] Bhakta I, Sau A. Prediction of depression among senior citizens using machine learning classifiers. *International Journal of Computer Applications*. 2016;144(7):11-16.

**PARTICULARS OF CONTRIBUTORS:**

1. Junior Resident, Department of Community Medicine, R.G. Kar Medical College and Hospital, Kolkata, West Bengal, India.
2. PhD Research Fellow, Department of Computer Science and Engineering, West Bengal University of Technology, Kolkata, West Bengal, India.

**NAME, ADDRESS, E-MAIL ID OF THE CORRESPONDING AUTHOR:**

Dr. Arkaprabha Sau,  
Professor Colony, Kanakpur, Panskura R.S., Purba Medinipur-721152, West Bengal, India.  
E-mail: arka.doctor@gmail.com

Date of Submission: **Aug 20, 2016**

Date of Peer Review: **Nov 28, 2016**

Date of Acceptance: **Dec 28, 2016**

Date of Publishing: **May 01, 2017**

**FINANCIAL OR OTHER COMPETING INTERESTS:** None.