



A Novel Internet of Things Framework Integrated with Real Time Monitoring for Intelligent Healthcare Environment

A. Suresh¹ · R. Udendhran² · M. Balamurgan² · R. Varatharajan³

Received: 26 December 2018 / Accepted: 22 April 2019 / Published online: 3 May 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

During mammogram screening, there is a higher probability that detection of cancers is missed, and more than 16 percentage of breast cancer is not detected by radiologists. This problem can be solved by employing image processing algorithms which enhances the accuracy of the diagnostic through image segmentation which reduces the misclassified malignant cancers. By employing segmentation, the unnecessary regions in the breast close to the boundary between the breast tissue and segmented pectoral muscle can be removed, therefore enhancing the accuracy the calculation as well as feature estimation. In-order to enhance the accuracy of classification, the proposed classifier integrates the decision trees and neural network into a system to report the progress of the breast cancer patients in an appropriate manner with the help of technology used in healthcare system. The proposed classifier successfully demonstrated that it achieved more accurate prediction when compared with other widely used algorithms, namely, K-Nearest Neighbors, Support Vector Machine and Naive Bayes algorithm.

Keywords Internet of things · Body sensor network · Neural network · Decision trees · Breast cancer · Healthcare

Introduction

At present enough research work is going on developing scalable and dynamic techniques for classification and prediction which may be able to handle large dataset in distributed

environment. In distributed environment, a series of challenges have been emerged in the field of data mining, triggered in different real life applications. One of the subdivision of Artificial Intelligence is Machine learning that is to attain complete knowledge of the education and to establish learning capacity in computer system. This finds variety of application in the field of engineering and society. This approach can give indiscriminate solutions for many troubles economically and this could make the computers intellectual. Machine learning is also used in the field of medical science. It assists us in resolving the extrapolative and diagnostic problems [1]. Essentially this algorithm is used for prediction of certain diseases, for therapy planning, hold up largely for managing the patients [2]. As already mentioned breast cancer is very harmful, and therefore computerized systems are necessary that make the doctors to be an expert in detecting the disease and help in accurate diagnosis. Hence for this process machine learning plays a major role. There are different approaches for different problems and therefore there is a need for detecting the approach and then perform according to that approach.

From Fig. 1, it is seen that the case of women who are affected by the breast cancer is said to be very high among the developed and developing countries [3]. One of the data mining function is said to be known as the classification, which allocate items/instances from the data set in a gathering

This article is part of the Topical Collection on *Systems-Level Quality Improvement*

✉ A. Suresh
prisu6esh@yahoo.com

R. Udendhran
udendran@gmail.com

M. Balamurgan
mmbalmurgan@gmail.com

R. Varatharajan
varathu21@yahoo.com

¹ Department of Computer Science and Engineering, Nehru Institute of Engineering and Technology, T.M.Palayam, Coimbatore, TamilNadu 641105, India

² Department of Computer Science and Engineering, Bharathidasan University, Trichy, Tiruchirappalli, India

³ Department of Electronics and Communication Engineering, Sri Ramanujar Engineering College, Chennai, Tamil Nadu 600127, India

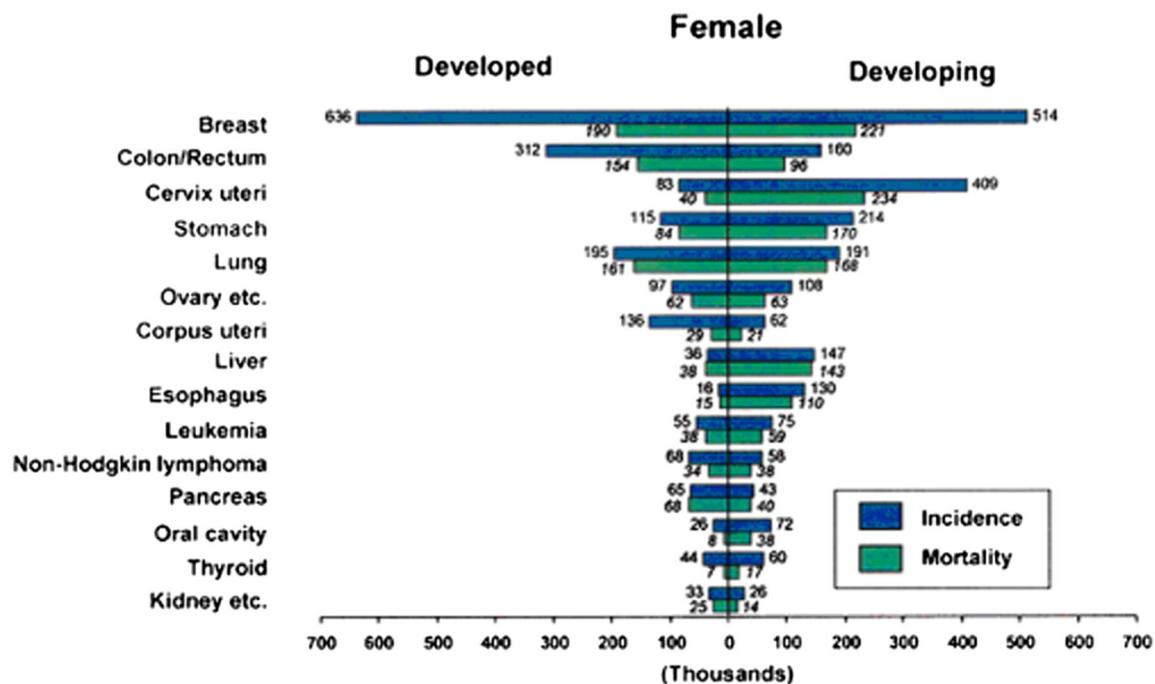


Fig. 1 Projected amount of people affected by cancer

of target categories or classes. For instance, a classification model can be used to detect loan applicants as low, medium, or high credit risks. Every Classifier process includes two steps:

1. Classifier Building: This method is to build a learning phase. The classifier, which evolved from training certain databases instances/tuples, is constructed by the classification algorithms [4]. Individual instance/tuple (is also known as objects or samples or data points) that is composed of the training set is referred to as a category or class.
2. Usage of Classifier - The training model/classifier generated using the training data set will classify the test data set objects/tuples. The chief concern is for the preparation of data for Classification and Prediction. There are certain activities in preparing the data:

- Cleaning the Data: This means removal of noise which is made possible by applying the smoothing techniques and treating the missed values which could be replaced by certain commonly occurred values from that attribute.
- Relevance Analysis: Database could consist of some unrelated attributes. Correlation [5] analysis will identify whether any two given attributes are related.
- Data Transformation and reduction: The following methods help in data transformation:

- Normalization: This transformation occupies scaling every value that will make them descend within a specific range. This method is mainly used in the learning step when the neural networks or the methods evolving measurements are used.
- Generalization: This is the major concept of transforming the data. The hierarchy concept is employed.

Related work

Statistics done at the year of 1973–1999, undertaken by ACR says that the women affected by breast cancer have significantly increased at a rate of 40% [6]. But this survey conducted at the year of 1989–1995 says that the death rate due to the breast cancer has significantly reduced. This is due to the early detection of cancer cells. ACR also produced a survey of 192,370 new cases of invasive breast cancer, and 62,280 cases of in situ breast cancer. The death rate at the year 2009 was also predicted that is shown in Fig. 2.

However, the possibility of survival could happen when the cancer is identified at an early stage. This progress was carried out in European countries at the year 1990. Finally, it is expected that people who die out of breast cancer has significantly reduced by

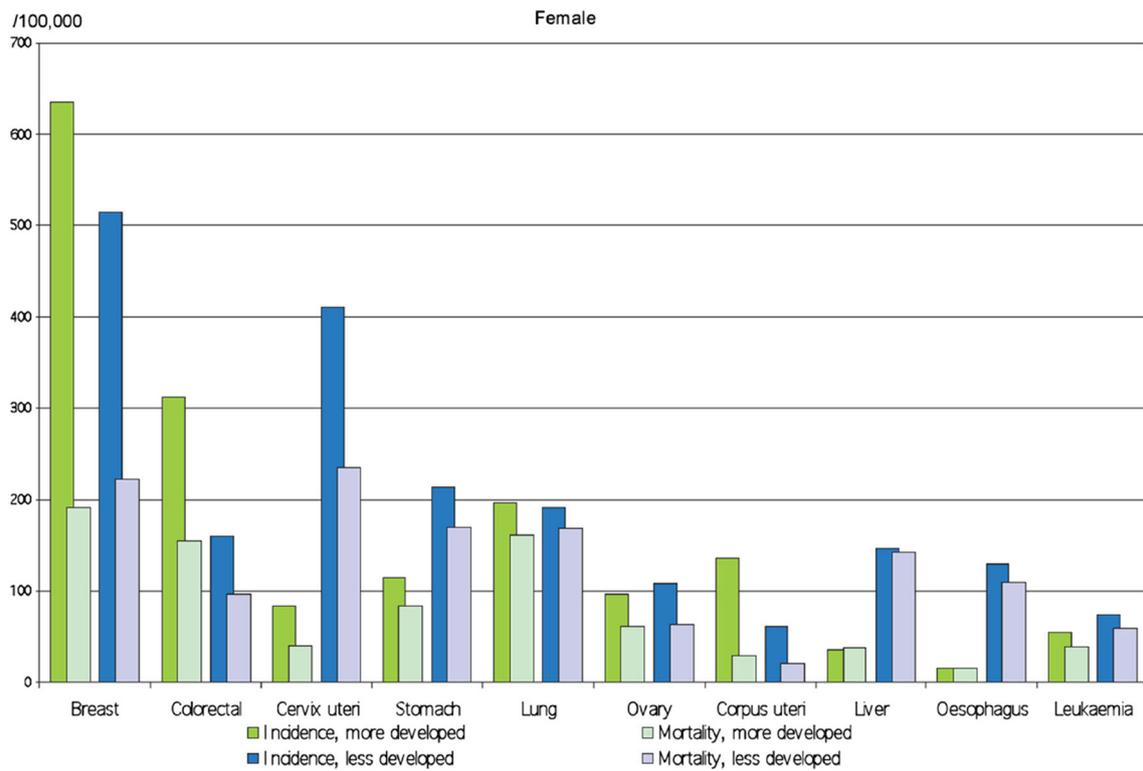


Fig. 2 Worldwide Mortality report of most affected cancers

detecting the cancer at the early stage with the improved treatment by introducing the digital mammography screening program throughout all countries of Europe. One of the best techniques for detecting cancer cells is through mammography [7]. The cancer cells detected by mammography are as follows: 1) mass

lesions 2) micro calcification clusters (MCCS) and other suspicious anomalies till 2 years before the cells start too profound. In fact, mammography could detect the cancerous cells that could not be identified by the physical examination [8].

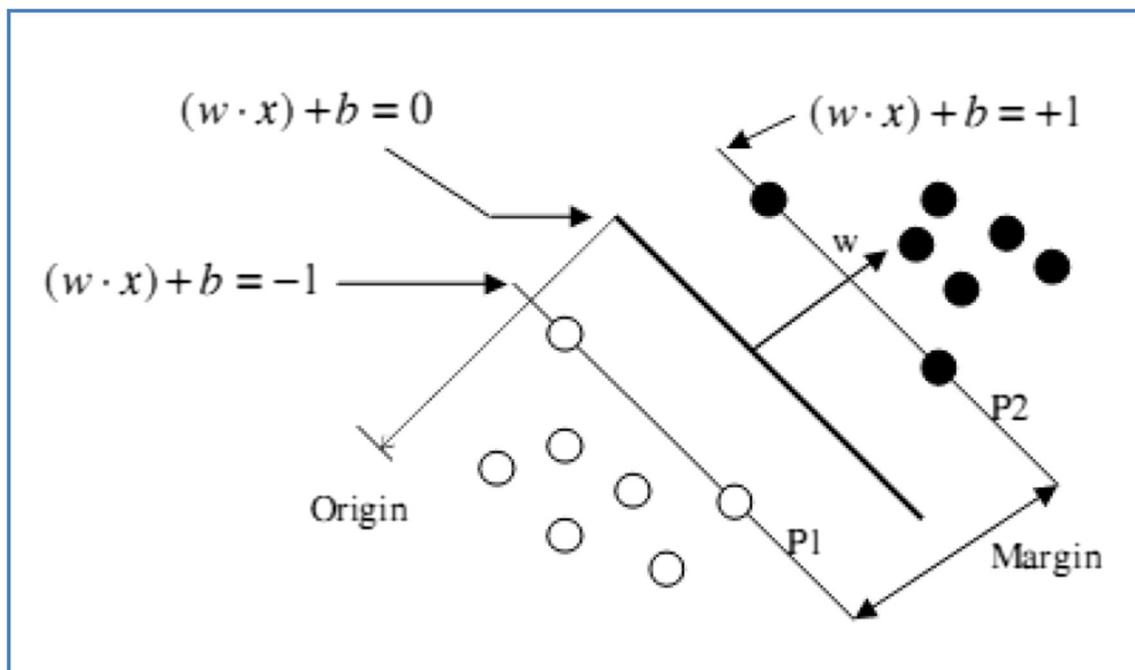
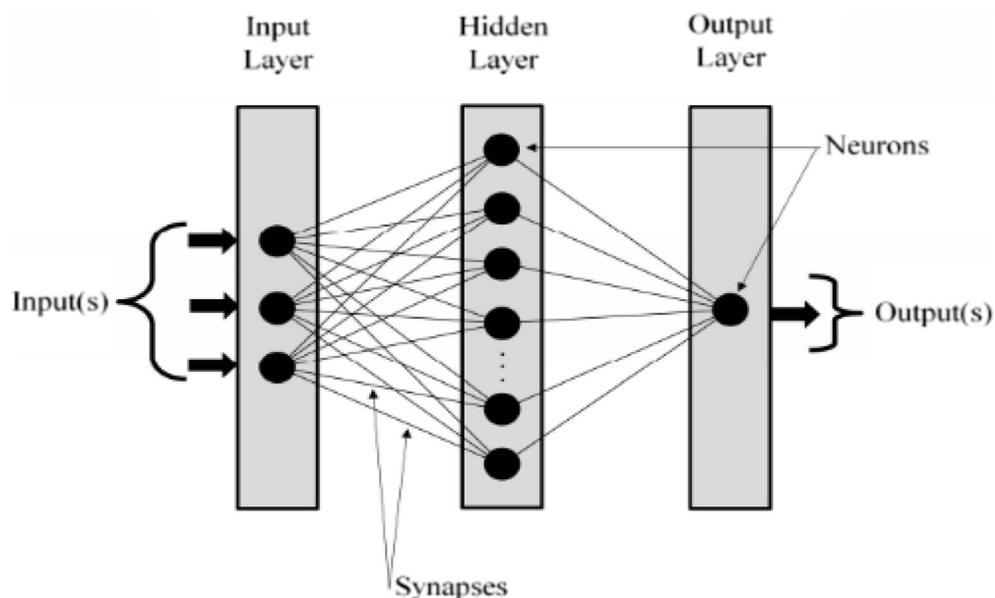


Fig. 3 Basics of SVM classification

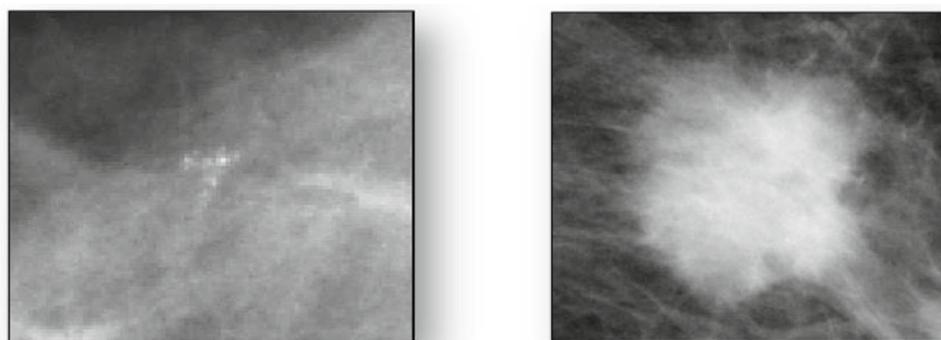
Fig. 4 Simple neural network



Digital Mammography can be operated effectively and the mammograms analysis can be done only by the radiologist experts [9]. Recent studies say that Positive Predictive Value (PPV) i.e. the ratio of total breast cancers found to the total number of biopsies is said to be only 15–30%. Receiver Operating Characteristic (ROC) curve analysis is done that measures the accomplishment of the medical diagnostic test, which measures the following performance: true positive (TP), false positive (FP), true negative (TN) and false negative (FN). The samples taken are tested with the malignant samples as well as the benign samples where the sensitivity and the specificity are determined through the following measures. Certain factors such as poor quality mammograms, defect in the physician eye, delicate radiographic nature and any other causes could lead to the misunderstanding of malignant tumour to the benign class that are generally said to be false positive (FP). These misunderstanding is said to be Type I/II error [10].

The indicators used in the radiography are Mass lesions (or masses) and micro calcification clusters (MCCs). Mass lesion could be more complicated since lesions are usually present at the depth glandular area of the breast tissue that is hard to find due to the variation in the dimension [11]. These are difficult to distinguish since they resemble like a normal inhomogeneous breast tissues [12]. Moreover, MCCs is also quite difficult, since these calcium deposits appear in a small area that forms a granular bright spot in the mammogram. The size could vary between 0.1 mm to 1 mm that is found to be circular with the diameter of 0.3 mm [13]. Smaller MCCs are tough to detect due to their superimposition on the breast parenchyma texture. They are visible with the structure of the breast tissue in an inhomogeneous background. Fatty tissues have certain MCCs that have a dense bright tissue. This is because the intense breast image could have high absorbent tissues. Hence, the contrast of the image should be reduced.

Fig. 5 **a** an example of non-invasive breast cancer, **b** an invasive breast cancer with weak borders



(a)

(b)

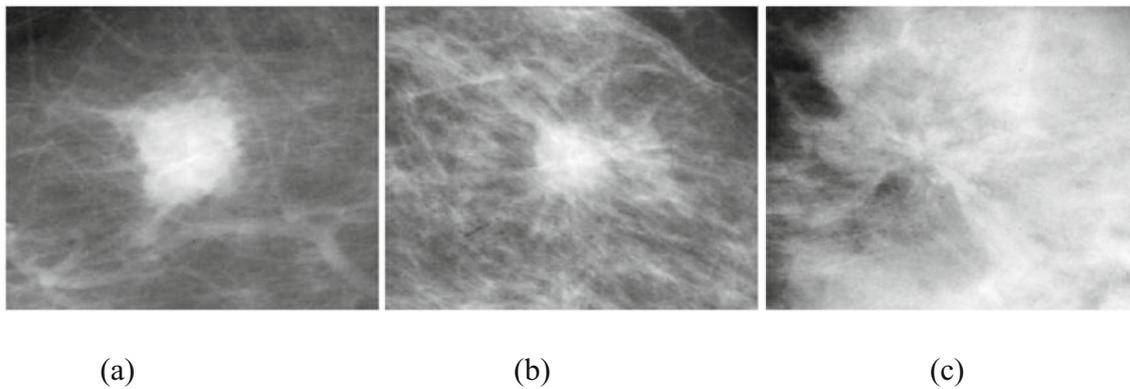


Fig. 6 Examples of the most common signs of malignant abnormalities a circumscribed lesion b stellate lesion c architectural distortion

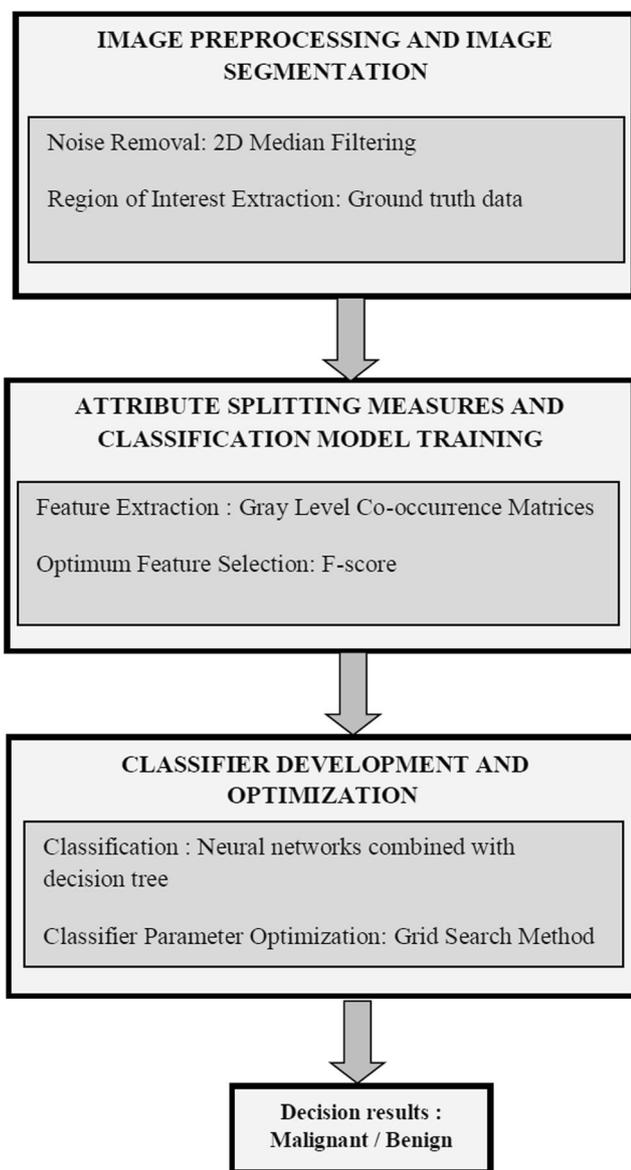


Fig. 7 Architecture and components of classifier

Reliable classification techniques

SVM

SVM utilizes a technique that identifies a hyper-plane with the most intense margin towards the sample object. Classification of feature vector is done by the probability as well as the margin value of the hyper-plane that is indirectly linked with each other as shown in Fig. 3.

The main objective is to make a separation between 2 classes. The open circle indicates the -1 class label and the solid circle indicates $+1$. The circle that is lying on the plane P1 and P2 indicates the support vector. The optimum plane should lie within the two planes that are parallel to one another. The formulae representing the hyper-plane are as follows:

$$(m.n) + c = +1 \text{ for classy} = +1 \tag{1}$$

$$(m.n) + c = -1 \text{ for classy} = -1 \tag{2}$$

The optimal hyper-plane is produced in the high dimensional feature space that utilizes the kernel function; $W(v_i, v_j)$ that is produced between two input vectors; v_i and v_j .

$$M(v_i, v_j) = \Phi(v_i) \cdot \Phi(v_j) \tag{3}$$

After applying the polynomial mapping,

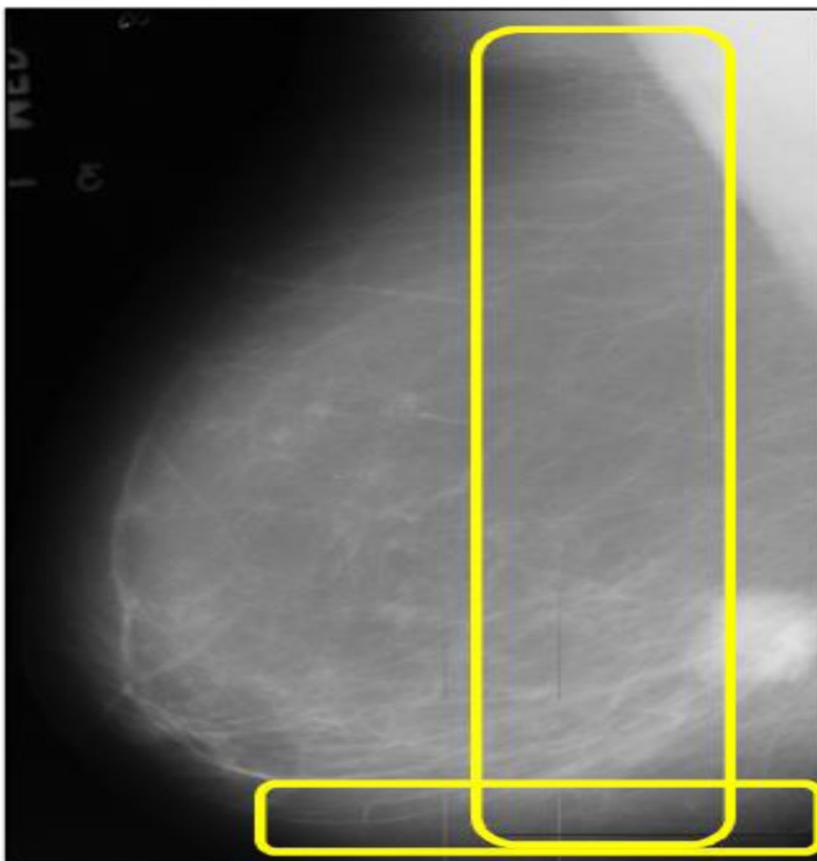
$$M(v_i, v_j) = (v_i \cdot v_j + 1)^c \tag{4}$$

Where, c is the polynomial degree. The above equation is applied with Radial Basis Functions as given below,

$$M(v_i - v_j) = e^{-\| \frac{v_i - v_j}{2\sigma} \|^2} \tag{5}$$

Where, σ is denoted as Gaussian sigma. With the help of SVM technique we could be able to solve the binary classification problem that has -1 and $+1$ respectively. Enhancing this technique further could be utilized for the multi-class problems [14, 15].

Fig. 8 After removing noise by applying two-dimensional median filtering approach



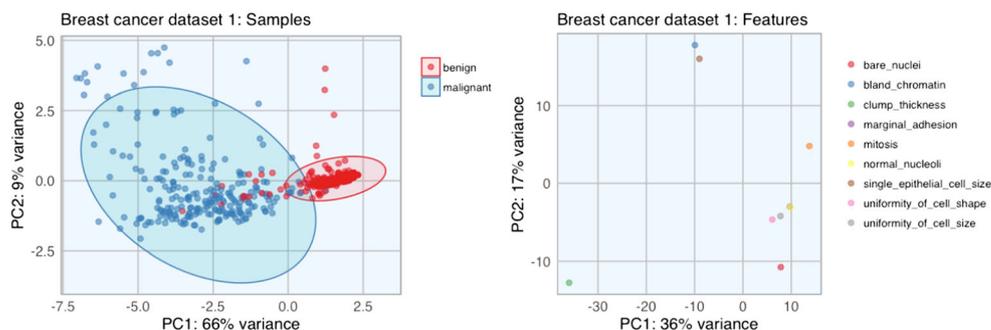
Neural networks

Neural Network is said to be the collection of neurons similar to the human brain system. This stands intermediate among the Artificial Intelligence (AI) and the approximation algorithm that is shown in Fig. 4. This is also known as nonlinear predictive model since it learns through training resemble structured biological neuron networks. Neural Networks are best suited for detecting the pattern, making predictions and it also learns from the past.

Cancerous breast tumours

- Non-Invasive breast cancer:** This is said to be the composition of malignant cells known as In situ cancer. The epithelial cells get restored by the malignant cells, which appears as a duct layer or lobules in the breast tissue. The breast stroma will not be attacked by these malignant cells since they form as a cramp at the basement of the membrane. Ductal Carcinoma In Situ (DCIS) and Lobular Carcinoma In Situ (LCIS) are said to be the two types of this cancer.

Fig. 9 Samples and features of the dataset



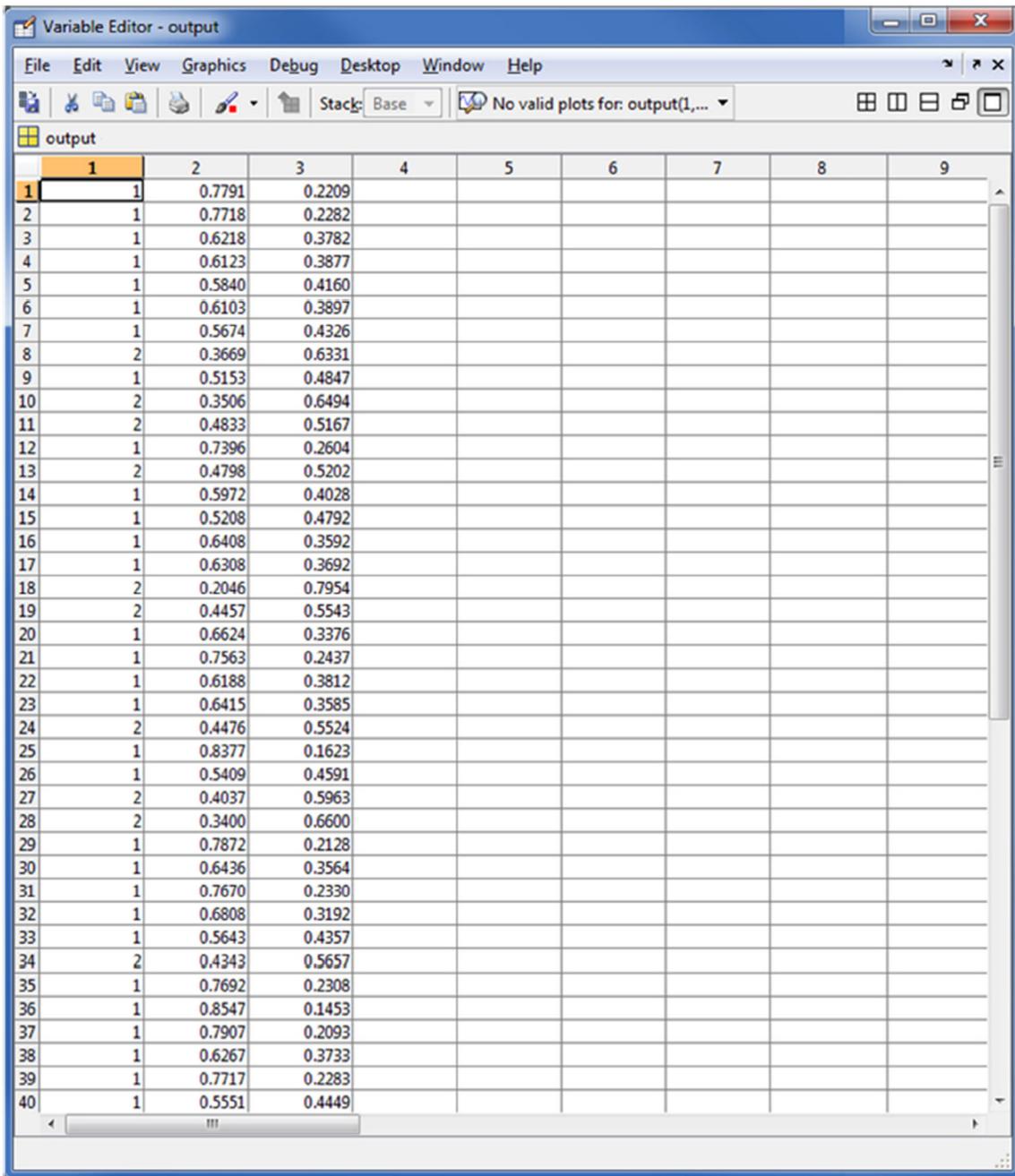


Fig. 10 Neural network testing results

- Invasive breast cancer:** When the malignant cells go beyond the duct then it is said to be infiltrating cancer. They spread to the entire part of the breast and in fact to the body. The size of the cell could vary around 80 mm.

Most breast tumors are in the form of focal mass lesion. As shown in the Fig. 5, malignant tumors appear in the form of speculation that is a stellate pattern of lines focused in the direction of the centre of the lesion.

MCCs is the another symptom of malignant tumor. Within the duct, they form a small or miniature cavity. Benign cancer

such as fibro-adenoma leads to MCCs [16]. This may be the early sign of the breast cancer that is shown on the Fig. 6.

Development of IoT based monitoring framework and optimization of the proposed classifier

Since the proposed architecture follows healthcare system, the system employs certain procedures like sensing of data, processing the data, analysing the data and finally delivering the

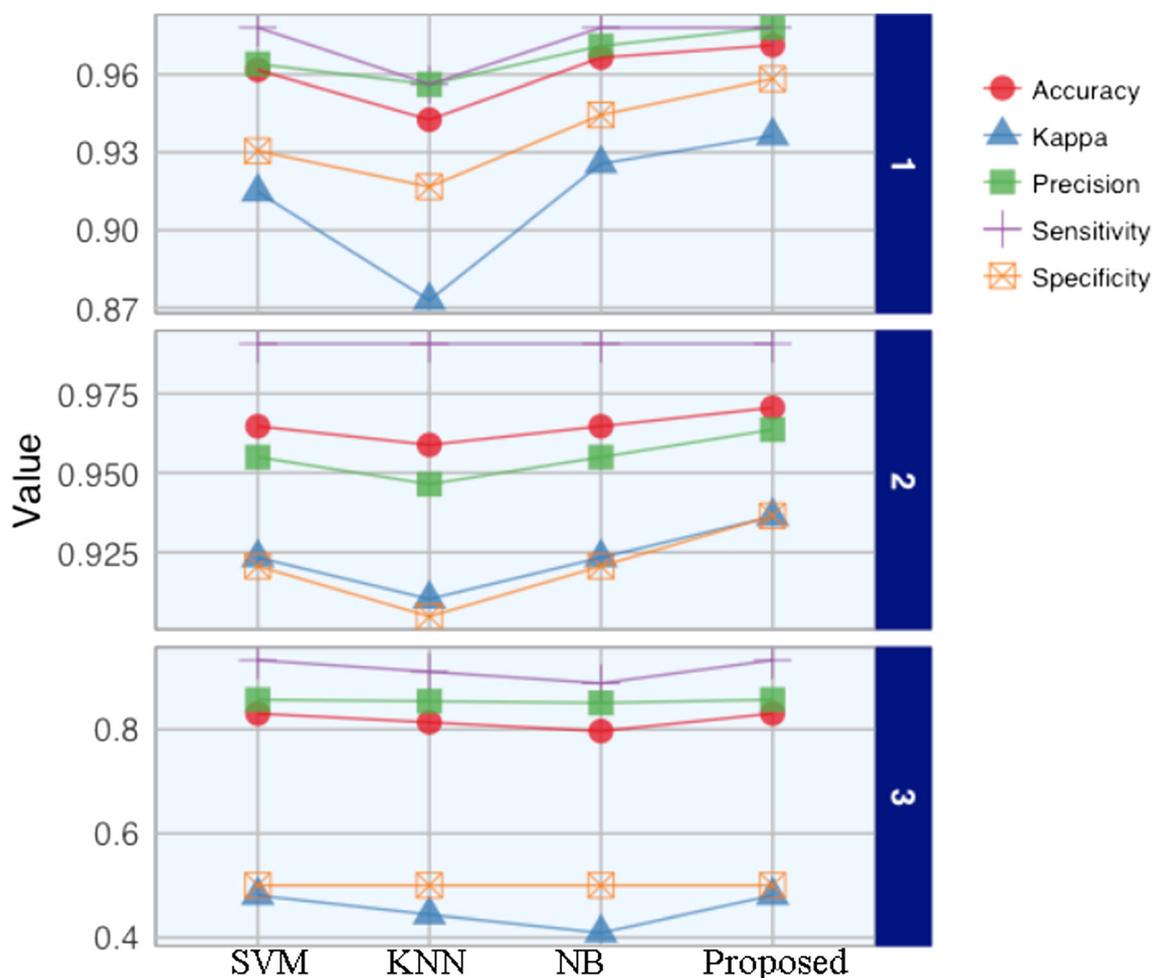


Fig. 11 Performance of our approach compared with support vector machines, K-NN, Naive Bayes

data. The sensors should possess the capability of awareness in any medical as well as environmental conditions. This could be achieved by employing tiny, cheap, intrusion smart home sensors with the BAN in a home environment. By incorporating these sensors, a comprehensive level of real-time home monitoring is possible, thereby achieving reliable and effective healthcare decisions and quality of life for self-reliant living. Every wireless sensor is considered as a node in a wireless sensor network, which is competent of collecting sensory information, working on the sensory information as well as interacting with other nodes in the network. While deploying, the layout of the home environment and the identification of ideal sensor locations must be considered. These ideal sensor locations can choose a wide range of the radios being employed such as, MiWi, as well as the sensitivity of sensors.

A wireless sensor consists of:

- Micro-controller: possesses computational abilities of the platform

- Radio transceiver: possesses low-power wireless communications
- Sensor board: possesses hardware interfaces to external sensors
- Power Layer: possesses power through batteries, capacitors.

Classification technique has a major trouble in minimizing the feature space dimension by identifying constricted amount of features, which yields to a better performance for classification. Feature selection has undertaken heavy work in these recent years. The time for the pattern classification could be completely reduced by the feature elimination [17]. This also enables the categorizing routine with the privilege of machine learning. At last a list could be created that shows the feature (feature contribution will be more) to be enabled and the features to be eliminated (feature contribution will be less). The architecture and the components of the proposed classifier are presented in Fig. 7. The texture feature extraction is carried out by employing Gray Level Co-occurrence

Matrices from the abnormal regions (Region of Interest). Certain features that are eliminated from the rank list due to their ineffectual discrimination process [18]. By employing MATLAB Image Processing Toolbox, the texture analysis capabilities were successfully proposed based on the correlation, energy, contrast as well as homogeneity texture descriptors which determine the statistics.

By employing ID3 algorithm, the decision tree is built along with samples. The target function value consists of the highest added value to the output value of the target function according to the samples. The attribute with all possible values are considered as edges and remaining attributes determines the target function and the optimal one will be chosen by the algorithm. The implementation of J48 needs a scan among the whole training set for every nodes of the tree, which shows a split on an attribute. $O(n)$ is denoted as the amount of nodes in the tree based on the data source, where n denotes the training instances that makes the time intricacy for the part $O(n^2)$. It is difficult and very much complex to merge the different local decision trees to form the global one. For the efficient merging process the decision tree rules have been converted into the simple decision rules. In-order to calculate the information gain which is required for the target-attribute, the ID3 sorts the attributes based on descriptive power as well as builds the tree according to the statistical value. The optimal solutions are employed for the creating the neural network structure based on the decision tree. In other words, the decision tree will be transformed into neural network in which the neurons of the hidden layer were obtained and built by the neural network. The process of scanning all the decision rules of different network is somewhat complex as well time consuming process. In-order to reduce the network overhead and complexity the decision rules are transformed into the decision table form.

Attribute splitting measures

The central choice of ID3 algorithm is the selection of attribute that is to be tested at every node in the tree. The attribute which is most helpful for classifying the instances will be selected based on information gain Eq. 6:

$$AH(Q, W) = J(W) - \sum p(t)H(t) \tag{6}$$

Where,

- $S(W)$ - Entropy of W
- t – Sub-sets obtained from W splitting set
- $p(t)$ - Total number of entities in set W
- $H(t)$ - Entropy of t

In-order to train the network by splitting the dataset is carried out by Eq. 7,

$$SplitInfo(D) = -\sum_{j=1}^v \frac{|D_j|}{|D|} \log \frac{|D_j|}{|D|} \tag{7}$$

Where D is the partition set,

Full split information: Same size partitions are processed.

Down split information: Different size partitions are processed.

The three input parameters are: (i) w represents input weight, (ii) x represents hidden neuron, and (iii) output weight μ .

Some arbitrary distinct samples is denoted by Z , (b_v, m_v) , where $b_v = [b_{v1}, b_{v2}, b_{v3}, \dots, b_{vn}]^T \in W^n$ and $m_i = [m_{v1}, m_{v2}, m_{v3}, \dots, m_{vm}]^T \in W^m$ with D hidden neurons and function. W represent the weight vector connecting input neurons and the v^{th} unknown neuron, f represents the threshold and μ is the weight vector connecting the hidden neuron and the output neurons [19]. The $b^* m$ denotes the inner product, if the classifier can approximate with zero errors. Usually the neurons are smaller when compared with the quantity of distinct training samples [20].

Evaluation

The data has been from the UC Irvine Machine Learning Repository. We observed two important classes, namely, benign and malignant classes. The characteristics of these datasets consist of cell nucleus properties and were produced from image analysis of fine needle aspirates (FNA) of breast masses. Mammogram images were extracted obtained from UMMC.

Horizontal and vertical lines noises were removed in the mammogram images by employing Two-dimensional Median Filtering approach as shown in Fig. 8.

Row depicts testing sample and predicted label is represented in the first column and samples and features of the dataset is presented in Fig. 9. Based on the probability, a sample is termed benign if its probability is higher as shown in Fig. 10.

In order to find the dimensionality and variance of the datasets, a function that performs the PCA employing *pcaGoPromoter* package which computes ellipses of the data points and yields the plot with *ggplot2* as shown in Fig. 11.

In test stage, 3-fold cross validation method and average values were calculated. From the Fig. 11, it is proven that our proposed classifier achieves accuracy rate of 96 when compared with KNN, SVM, Naive Bayes algorithms. The best classification performance of our proposed classifier was obtained with eight inputs and its correct classification rate is 97.4%. The correct classification rate of Naive Bayes (NB)

with 9 inputs is 95.2% and the correct classification rate of SVM is 95.6%. Therefore, the proposed classifier is best suited for the classification performance with the minimum number of input parameters.

Conclusion

For this research, the proposed classifier addresses successfully the problem of misclassified malignant cancers. Breast cancer is considered as most dangerous disease among women. Digital mammography is an effective method based on concentrated density of tissues employed for detecting shadows. Early detection and treatment of breast cancer contributes to higher chances of survival. In order to design of an effective classifier Neural Network and Decision trees were combined and trained employing digital mammogram for validation and testing. A decision tree is effective with the redundant attributes. The construction of decision tree could be affected when the redundant attributes are involved. Naive Bayesian offers low accuracy in classification since it assumes all the features that are independent. SVM automatically map the inputs with high dimensional feature/attribute spaces. In general SVM supports to construct hyper plane in any space and this can be employed for any tasks such as regression, prediction or classification. The only fact with neural network is that it takes time in training the data but has a large tolerance towards the noisy data.

Compliance with ethical standards

Ethical responsibilities of authors The authors follow the ethical information provided in the journal and hereby abide the same with the journal.

Conflict of interest In accordance with all authors, we are reporting that there is no conflict of interest in the research paper.

References

- Magoulas, G. D., and Prentza, A., Machine learning in medical applications. In: Paliouras, G., Karkaletsis, V., Spyropoulos, C. D. (Eds), *Machine Learning and its Applications, Lecture Notes in Computer Science*. Berlin: Springer-Verlag, 2010, 300–307.
- Hsieh, S. L., Hsieh, S. H., Cheng, P. H. et al., Design ensemble machine learning model for breast cancer diagnosis. *J. Med. Syst.* 36:2841, 2012. <https://doi.org/10.1007/s10916-011-9762-6>.
- Naghbi, S., Teshnehlab, M., and Shoorehdeli, M. A., Breast cancer classification based on advanced multi dimensional fuzzy neural network. *J. Med. Syst.* 36:2713, 2012. <https://doi.org/10.1007/s10916-011-9747-5>.
- Murakami, Y., and Mizuguchi, K., Applying the naive Bayes classifier with kernel density estimation to the prediction of protein-protein interaction sites. *Bioinformatics* 26(15):1841–1848, 2010.
- Peter, N., Enhancing random forest implementation in WEKA. In: *Machine Learning Conference*, 2005.
- Levi, F., Bosetti, C., Lucchini, F., Negri, E., and La Vecchia, C., Monitoring the decrease in breast cancer mortality in Europe. *Eur. J. Cancer Prev.* 14(6):497–502, 2005.
- Tyczynski, J. E., Plesko, I., Aareleid, T., Primic-Zakelj, M., Dalmas, M., Kurtinaitis, J., Stengrevics, A., and Parkin, D. M., EU member states: Mortality declining in young women, but still increasing in the elderly. *Int. J. Cancer* 112(6):1056–1064, 2004.
- Saritas, I., Prediction of breast cancer using artificial neural networks. *J. Med. Syst.* 36:2901, 2012. <https://doi.org/10.1007/s10916-011-9768-0>.
- Mehta, M., Agrawal, R., and Rissanen, J., SLIQ: A scalable parallel classifier for data mining. IBM Almaden Research Center, CA 95120.
- Li, J. B., Yu, Y., Yang, Z. M. et al., Breast tissue image classification based on semi-supervised locality discriminant projection with kernels. *J. Med. Syst.* 36:2779, 2012. <https://doi.org/10.1007/s10916-011-9754-6>.
- Nassif, H., Page, D., Ayvaci, M., Shavlik, J., and Burnside, E. S., Uncovering age-specific invasive and DCIS breast cancer rules using inductive logic programming. In: Veinot, T. (Ed.), *Proceedings of the 1st ACM International Health Informatics Symposium (IHI '10)*. New York: ACM, 2010, 76–82.
- Huang, M. L., Hung, Y. H., Lee, W. M. et al., Usage of case-based reasoning, neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. *J. Med. Syst.* 36:407, 2012. <https://doi.org/10.1007/s10916-010-9485-0>.
- Diz, J., Marreiros, G., and Freitas, A., Applying data mining techniques to improve breast cancer diagnosis. *J. Med. Syst.* 40(203), 2016. <https://doi.org/10.1007/s10916-016-0561-y>.
- Suresh, A., and Varatharajan, R., Recognition of pivotal instances from uneven set boundary during classification. *Multimed. Tools Appl.*, 2018. <https://doi.org/10.1007/s11042-018-5905-9>.
- Issac Niwas, S., Palanisamy, P., Chibbar, R. et al., An expert support system for breast cancer diagnosis using color wavelet features. *J. Med. Syst.* 36:3091, 2012. <https://doi.org/10.1007/s10916-011-9788-9>.
- Paulin, F., and Santhakumaran, A., Back propagation neural network by comparing hidden neurons: Case study on breast cancer diagnosis. *Int. J. Comput. Appl.* 2(4), 2010. (0975–8887).
- Ganatra, A., Panchal, G., Kosta, Y., and Gajjar, C., Initial classification through back propagation in a neural network following optimization through GA to evaluate the fitness of an algorithm. *International Journal of Computer Science and Information Technology* 3(1):98–116, 2011.
- Vapnik, V., and Vashist, A., A new learning paradigm: Learning using privileged information. *Neural Netw.* 22(5–6):544–557, 2009.
- Mahmoudabadi, H., Izadi, M., and Menhaj, M. B., A hybrid method for grade estimation using genetic algorithm and neural networks. *Comput. Geosci.* 13:91–101, 2009.
- Chattopadhyay, S., Kaur, P., Rabhi, F. et al., Neural network approaches to grade adult depression. *J. Med. Syst.* 36:2803, 2012. <https://doi.org/10.1007/s10916-011-9759-1>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.