Hybrid Model Combining BPNN and SVM-RBF Classifiers for Diagnosis

K.G.S. Venkatesan¹*, Allanki Sanyasi Rao², Sushma Jaiswal ³

¹Department of Computer Science & Engineering, MEGHA Institute of Engineering and Technology for Women, Hyderabad, Telangana, India.
²Department of Electronics & Communication Engineering, Christu Jyothi Institute of Technology & Science, Jangaon, Telangana, India.
³Department of Computer Science & Information Technology (CSIT), Guru Ghasidas Vishwavidyalaya (A Central University), Koni, Bilaspur, Chhattisgarh, India.

Received: 13.08.2023  ✡  Accepted: 23.08.2023  ✡  Published Online: 31.08.2023

Abstract: The global prevalence of cancer as a leading cause of death in women is undeniable. If we can get better at finding and treating cancer, it will be a huge boon to people’s health and longevity. As a result, having accurate cancer prediction capabilities of a high grade is very necessary in order to maintain up-to-date treatment and survival requirements for patients. The detection and treatment of breast cancer at an early stage has been a major focus of study in recent years, and techniques based on machine learning have been shown to be a successful methodology. In this investigation, we used a total of five distinct machine learning approaches. It is not a recent development for empirical investigations of breast cancer therapy to include the use of machine learning and other kinds of soft computing. The wide range of breast cancers as well as their prevalence all over the world are addressed. It cannot be overstated how important it is to be able to recognise anomalies and distinguish between benign and malignant breast cancer. Diagnostic models for breast cancer have been created that use both the Levenberg-Marquardt (LM) and the Sparse Representation (SR) algorithms. A great performance in terms of accuracy and mean square error was achieved as a result of the

*Correspondence: Professor, Department of Computer Science & Engineering, MEGHA Institute of Engineering and Technology for Women, Hyderabad, Telangana, India. Email:drkgsvenkatcse@meghaengg.ac.in

https://doi.org/10.58599/IJSMEM.2023.1802
Volume-1, Issue-8, PP:9-17 (2023)

This work is licensed under a Creative Commons Attribution 4.0 International License CC BY-NC-ND 4.0.
combination of LM’s resilience in distinguishing tiny deformations and SR’s correctness during training. This was made possible by the combination of the two factors. Information has been extracted from the Wisconsin Dataset, sometimes known as the WD. Using a combination of different classification algorithms for pre-clinical diagnosis is something that can be done, as this model indicates.

Key words: Prediction, Data Preprocessing, Hybrid Classifier, Multimodal Classification.

1. INTRODUCTION

In spite of the fact that cancer is one of the most fatal diseases that a person may have, there is currently no therapy for the condition. There is a high incidence of breast cancer diagnoses. The National Breast Cancer Foundation has developed some forecasts regarding the number of new instances of breast cancer in the United States in the year 2020. These projections indicate that there will be 276,000 new cases of invasive breast cancer and 48,000 new cases of non-invasive breast cancer. Cancer of the breast is the most common cause of mortality in women all over the world. According to the findings of the American Cancer Society, breast cancer has been responsible for the deaths of more than 41,760 women and more than 500 men in the United States in recent years. In typical healthcare programmes in LMCs, early detection and diagnostic techniques, as well as surgical operations, radiation therapy, and pharmaceutical therapy, are all easily combinable and coordinated with one another [1].

It is necessary to do further study on distribution and implementation in order to determine the strategy that is most successful for guaranteeing compliance with standards, particularly as it applies to LMCs [2]. The potential loss of fertility as a result of therapy is something that younger women in particular should be concerned about. At the time of diagnosis, younger patients should be counselled about their reproductive choices, and attempts should be made to identify methods for young breast cancer survivors to preserve their eggs viable. In addition, younger patients should be counselled about their reproductive possibilities. There is a dearth of evidence about how the diagnosis and treatment alter women’s attitudes towards fertility after they have gone through them. The rapid diagnosis and treatment of a broad variety of life-threatening disorders has been significantly facilitated by recent advancements in machine learning and artificial intelligence, which has resulted in an increase in the probability that the patient will have a favourable outcome [3]. In [4], the author analyses and breaks down a DNNS model into its component parts. In contrast to previous methods, this one centred on the Support value of the deep neural network. A normalisation approach was used to enhance the usefulness, efficacy, and quality of the data.

In [5], the pre-processing and classification phases, together with the procedure of selecting the appropriate approach and parameters, are automated [6]. The ultimate objective of this study is to provide the groundwork for the rapid development of improved machine learning techniques. A bespoke Wienmed filter was developed to preprocess the raw, noisy data from the MRI scans, since the MRI breast images are the primary focus of the system [7] sought to identify breast cancer using a neural network. Using the Wisconsin Breast Cancer Data (WBCD), specialists in neural networks have investigated the accuracy of clinical breast cancer diagnoses.
2. RELATED WORKS
The efficacy of multi-layered perceptron networks in the diagnosis of breast cancer is assessed using a variety of back propagation techniques. Several strategies for acquainting physicians with neural network models of vital patient measurements and outcomes were studied and reported in [8]. It is possible to adapt the proposed approaches to practically any conceivable clinical outcome model. Researchers at [9] used information from the Wisconsin Breast Cancer Dataset. In [10], an elaborate neural network design is described, the foundation of which is multi-layered feed-forward neural networks. How many entries are needed depends on the details of the offered examples. Using ANN and LM models [10], the accuracy was found to be 92%, while the sensitivity was found to be 98%. In the referenced publication. The study’s overarching objective was to provide a method for the autonomous detection of breast cancer. There is an urgent need to classify and gather data on breast cancer inequalities. This was made possible by a team of oncology and computer science experts [11].

In order to measure the accuracy of a classification used the K nearest neighbour (KNN) technique, which is included into the R programme. The original dataset used for the research may be obtained at the UCI machine learning repository [12].

The suggested using a Levenberg-Marquardt parameter in numerical experiments with non-linear equations. Simulations show that LM is effective in solving non-linear equations, and the method converges super-linearly. Here we evaluate two training methods, LM and Scaled Conjugate Gradient (SCG), for use in a Multilayer Perceptron (MLP) to detect breast cancerous tissue. They used features from the Wisconsin Database, a standard dataset for comparing the efficacy of various classification algorithms, to assess the best approaches to training use supervised and unsupervised filter banks to evaluate the impact of sparse representation on identification accuracy. The evaluation employs a lightweight modular structure. For the purpose of analyzing linearly connected pictures, SR has also found usage in medical imaging applications. The dataset used in was created for the express purpose of spotting phishing campaigns. Gamma correction is a nonlinear mathematical method used to modify the values of individual pixels in a picture. The gamma parameter is adjusted to get this effect. In order to solve the difficulty of using big data to detect breast cancer in young women, this study adapts a dataset initially established for phishing detection. To do this, it combines LM and SR techniques, which together provide a robust classification strategy. Wisconsin is the location of the original data utilised for this study [13, 14].

3. METHODOLOGY
The given hybrid classification strategy incorporates the unique features of both the LM and SR classifiers. To maximise the usefulness of the hidden layer’s retrieved features while limiting the layer’s size to a minimum, the SR is used not only as a classifier but also as a layer manager. Breast cancer diagnosis may be broken down into two distinct phases. Three separate methods of categorization are introduced first. In order to improve the features, size, and weights of the LM training units, each of
these categorization approaches use an evolutionary technique. The first group of classifiers is taught using a parallel training approach on the training dataset. The overall proposed approach is shown in Figure 1.

**Pre-processing:**
The need of pre-processing data stems from the understanding that bad input will lead to bad results. Therefore, the initial step in developing a machine learning method is often data pre-processing. Pre-processing is the stage of research that involves preparation and standardization.

**Figure 1.** Overall Proposed approach.

**Data Preprocessing:**
- **Data Collection:** Gather the dataset relevant to your problem.
- **Data Cleaning:** Remove duplicates, handle missing values, and outliers.
- **Data Transformation:** If possible, features should be normalized or standardized.
- **Data Splitting:** Separate the data into a test set and a training set.

**Feature Engineering:**
Select or engineer relevant features to improve model performance.

**Backpropagation Neural Network (BPNN):**
- **Design the architecture of your BPNN.**
- **Define the number of input nodes (features), hidden layers, and output nodes (classes or regression**
targets).
- Initialize the network’s weights and biases.
- Implement the forward propagation algorithm
- Calculate the weighted sum of inputs for each neuron in each hidden layer.
- Apply activation functions (e.g., sigmoid, ReLU) to the weighted sums.
- Pass the output through the network to compute predictions.
- Implement the backpropagation algorithm to update weights and biases:
  - Calculate the error between predictions and actual values.
  - Propagate the error backward through the network.
  - Update weights and biases using gradient descent or a variant (e.g., Adam, RMSProp).
- Train the BPNN using the training data.
- Validate the BPNN’s performance on the testing data.

**Support Vector Machine with Radial Basis Function Kernel (SR Classifier):**
- Choose or optimize hyperparameters for the SVM-RBF model (e.g., C, gamma).
- Train the SR classifier on the training data.
- Evaluate its performance on the testing data.

**Integration of BPNN and SR Classifier:**
Decide how you want to integrate the predictions from the BPNN and SR classifiers. Common approaches include:
- Stacking: Use the predictions as additional features for a meta-classifier.
- Voting: Combine predictions using majority voting or weighted voting.
- Averaging: Average the predictions from both models.
- Meta-Learning: Train another model to learn the optimal combination of BPNN and SR predictions.

**Performance Evaluation:**
Assess the overall performance of the integrated system using appropriate metrics (e.g., accuracy, F1-score, RMSE). Fine-tune the integrated system if necessary.

4. RESULTS AND DISCUSSION
The gathered information was used to create two sets of data: a training set and an evaluation set. The neural networks were not given the data that would be utilized in the testing phase until after the training phase had concluded. The primary focus is on testing the generalizability of the learned neural network structures. Eighty percent of the time was spent on classroom instruction, while twenty percent was devoted to assessment. Three layers make up a typical back propagation network, with nine units in the input layer, six units in the hidden layer, and one unit in the output layer. An ordinary cell in the input will result in a layer with a value of 1 being returned in the output. Before a network is taught for the first time, its connection weights are completely reshuffled. The upper and lower bounds of each input vector are fixed at 0 and 1, respectively. Every calculation is done in MATLAB 2013a. Incidence rates of several breast cancer subtypes diagnosed in Wisconsin are shown.
in Figure 2.

![Figure 2](image1.png)

**Figure 2.** Cancerous and noncancerous cases of breast cancer in Wisconsin.

In order to evaluate the device’s parameters, this article tries out a few various formulations, the outcomes of which are summarised below. We have calculated both the Mean Square Error and the Accuracy. The average accuracy rates achieved during training with the LM method are shown in Figure 3. Calibration of SR’s accuracy across a wide range of startup factors may provide an estimate of its typical precision. As can be seen, both algorithms achieved an overall accuracy of 80% or better.

![Figure 3](image2.png)

**Figure 3.** The standard deviation of the Levenberg-Marquardt (LM) training accuracy metric

The mean MSE is shown against the sample size in Figure 4. Less MSE values suggest that SR produced a good fit to the data and that the remainders are minimal; this is shown by the fact that the MSE values are less. After undergoing both types of training, it has been shown that certain SR structures eventually converged on an MSE value that is comparatively low. Because of the high MSE for Sample
3, it is likely that the sample size is insufficient to effectively reflect the input-output relationship in the data, which will restrict the applicability of any conceivable exclusions.

![Figure 4](image)

**Figure 4.** The LM and SR classifiers were used to get the average MSE.

Figure 5 presents the findings obtained by applying the suggested method to about 8 different samples in order to run them through the SR classification algorithm. On average, 8 samples will be analysed and evaluated during this procedure. Higher spatial allocations and segmentation precision lead to improved accuracy and other key parameter variables including sensitivity, precision, recall, and f-measure compared to baseline models in more realistic settings. Recall is another essential element.

![Figure 5](image)

**Figure 5.** Utilizing a composite LM-SR technique, an assessment of the usefulness of the proposed model was carried out on eight distinct data sets.
Despite the best efforts of several research teams from across the world, anomaly identification has proved challenging since some of the data sets used are complex and take into account a wide range of attributes. The former is abbreviated as LM, and the latter as SR. In this research, we provide a new approach to classification by fusing the LM and SR algorithms into a single procedure. The job is done with the support of the Wisconsin Database (WD), which enables precise labelling of data occurrences as benign or malignant. After carefully examining each model, we found that the LM-SR method yielded the most precise results. The success rate and the mean square error for this approach were 99.5% and 98.2%, respectively. Because the results will only apply to the WD database, the investigation may have to be narrowed in scope. Expanding the scope of these technologies to include new forms of data is, thus, vital. Developing harmonised breast cancer data sets is the top priority for the next several years. This will be useful for evaluating how different algorithms stack up against one another.

5. CONCLUSION

Despite the best efforts of several research teams from across the world, anomaly identification has proved challenging since some of the data sets used are complex and take into account a wide range of attributes. Using the Levenberg-Marquardt (LM) training algorithm and the Sparse Representation (SR) classification strategy, this study creates a technique for detecting breast cancer. The former is abbreviated as LM, and the latter as SR. In this research, we provide a new approach to classification by fusing the LM and SR algorithms into a single procedure. The job is done with the support of the Wisconsin Database (WD), which enables precise labelling of data occurrences as benign or malignant. After carefully examining each model, we found that the LM-SR method yielded the most precise results. The success rate and the mean square error for this approach were 99.5% and 98.2%, respectively. Because the results will only apply to the WD database, the investigation may have to be narrowed in scope. Expanding the scope of these technologies to include new forms of data is, thus, vital. However, when compared to the results of other modern training and classification algorithms like KNN, MLP, RBP, SVM, and NB, among others, the results reveal a higher degree of success. Developing harmonised breast cancer data sets is the top priority for the next several years. This will be useful for evaluating how different algorithms stack up against one another.

References


