Mathematics in Applied Sciences and Engineering https://ojs.lib.uwo.ca/mase
Volume **5**, Number **2**, June **2024**, pp.**149-164** https://doi.org/10.5206/mase/17138

MEDICAL DIAGNOSIS USING ARTIFICIAL NEURAL NETWORKS

AFSANA BEGUM, MD. MASHIUR RAHMAN, AND SOHANA JAHAN

Abstract. Medical diagnosis using Artificial Neural Networks (ANN) and computer-aided diagnosis with deep learning is currently a very active research area in medical science. In recent years, for medical diagnosis, neural network models are broadly considered since they are ideal for recognizing different kinds of diseases including autism, cancer, tumor lung infection, etc. It is evident that early diagnosis of any disease is vital for successful treatment and improved survival rates. In this research, five neural networks, Multilayer neural network (MLNN), Probabilistic neural network (PNN), Learning vector quantization neural network (LVQNN), Generalized regression neural network (GRNN), and Radial basis function neural network (RBFNN) have been explored. These networks are applied to several benchmarking data collected from the University of California Irvine (UCI) Machine Learning Repository. Results from numerical experiments indicate that each network excels at recognizing specific physical issues. In the majority of cases, both the Learning Vector Quantization Neural Network and the Probabilistic Neural Network demonstrate superior performance compared to the other networks.

1. Introduction

Artificial neural networks (ANNs) have been extensively used for medical diagnosis in recent years. For disease diagnosis systems, a number of neural network structures have successfully replaced conventional pattern recognition techniques [19, 1, 27, 13, 2]. Additionally, neural network classification systems have been effectively employed for the diagnosis of chest and cancer diseases, among other clinical diagnosis problems. It is evident that early diagnosis of any disease is vital for successful treatment and improved survival rates. Detecting symptoms at an early stage enhances the effectiveness of treatment options including close monitoring, surgery, radiation therapy, and chemotherapy.

In this research, the performances of five ANNs, such as MLNN,PNN, LVQNN, GRNN, and RBFNN in diagnosing different medical issues have been explored. Patients with various medical issues such as autism, diabetes, cancer, and sepsis are considered.

Autism is a complex, lifelong, neuro-developmental disorder that appears during early childhood and that affects social communication and interaction. It is identified by a range of challenges in social skills and repetitive behaviors. The healthcare needs of people with autism are complex and require a range of integrated services. The symptoms of autism often appear during early childhood and according to the Centers for Disease Control and Prevention(CDC), a reliable diagnosis can be made as early as 2 years old. However, many people do not receive a diagnosis until much later. Sometimes during their developmental years, early diagnosis may help a child to receive support and that will benefit them throughout their life.

Received by the editors 30 November 2023; accepted 26 May 2024; published online 4 June 2024.

2020 *Mathematics Subject Classification.* 9010, 68T07, 68T09, 68T10.

Key words and phrases. Multilayer neural network (MLNN), probabilistic neural network (PNN), learning vector quantization neural network (LVQNN), generalized regression neural network (GRNN), and radial basis function neural network (RBFNN).

This research is supported by National Science and Technology Fellowship 2021, Bangladesh.

Cancer is a very complex and diverse group of diseases that is characterized by the uncontrolled growth and spread of abnormal cells in the body. Cancer originates from the conversion of normal cells into tumor cells. It is a multi-stage process that normally progresses from pre-cancerous damage to a lethal tumor. It is the result of interaction between a person's genetic factors and three kinds of external agents. They are physical carcinogens: Such as ionizing radiation and ultraviolet; chemical carcinogens: Like asbestos, components of tobacco smoke, alcohol, aflatoxin (a food contaminant), and arsenic; and biological carcinogens: Such as infections from viruses, bacteria, or parasites. The common cancers are breast, lung, colon, and rectum cancers. Early diagnosis of cancer is vital for successful treatment and improved survival rates. Detecting cancer at an early stage enhances the effectiveness of treatment options including surgery, radiation therapy, and chemotherapy. Early-stage cancers are generally more localized and have not spread to surrounding tissues or distant organs, making treatment more feasible and potentially curative.

The probabilistic neural network structures, based on Bayesian Classifiers, provide a general solution for solving pattern problems. PNN uses a supervised training set to develop a distribution function within a pattern layer. The training process of PNN is comparatively simpler than others. However, in some cases, the pattern layer can be quite large depending on the structure of data of different classes [23]. The PNN is appropriate for disease diagnosis systems because it offers a general solution to classify pattern problems. [9] . Like PNN, the training process of GRNN also uses only one pass of the training set using sparse data sets. However, the applications of GRNN provide continuous-valued outputs. GRNN is useful for problems that require continuous function approximation. It can fit multidimensional surfaces through data [24]. RBF neural networks [5, 12] are suitable for the implementation of multi-class and high-dimensional classification problems [7]. RBFNN, LVQNN, and MLNN are also reported to be used for disease diagnosis problems.

Use of ANN for medical diagnosis is increasing gradually. Nowadays every sector of medical science involves deep learning techniques for better results in disease diagnosis. As a result, in the last few years scientists are developing new ideas to apply ANN in this area. Several diseases like Diabetes, brain tumors, kidney stones,s and even diseases like Alzheimer's can also be diagnosed using ANN. In a recent article [17], the author showed the successful implementation of MLNN to extract features accurately from input and achieve the best result to diagnose Alzheimer's. MLNN is seen to outperform some other algorithms in predicting Diabetes Mellitus Risk [22]. Several studies are reported focusing on brain tumor diagnosis using PNN. In [14], the author proposed a system for screening brain tumors which is developed using deep learning probabilistic neural networks. Here the author focused on hybrid clustering for the segmentation of brain images to classify normal and abnormal forms of the brain.. PNN is also seen to be successfully being used in dental demineralization diagnosis in [18]. Here the author used PNN together with three multiscale entropy analysis methods to extract features from the Optical coherence tomography(OCT) one-dimensional echo signal of normal and demineralized teeth. Artificial Neural Networks have also shown excellent performance in identifying Kidney stone diseases. LVQNN and RBFNN are seen to perform very well in identifying kidney stones in [15]. Recently in [21] LVQNN is also seen to be used in detecting factitious disorder, a mental problem related to severe emotional disorders. In [20] RBFNN is used for thyroid disease diagnosis. Here the author proposes a method to enhance thyroid disease diagnosis by combining Multi-Layer Perceptron (MLP) and RBF neural networks. The proposed method was implemented on a data-set from UCI Machine Learning Repository. GRNN is known for its ability to diagnose various diseases. In a recent article [16], GRNN is seen to be useful in distinguishing hepatocellular carcinoma from normal controls using electronic noses (eNose) which collect the exhaled gas signals that are contained in Hepatocellular carcinoma patients due to abnormal metabolism. Here the author designed a model based on multidimensional features and GRNN. The result shows that GRNN performs very well with a limited training set. Some studies report that GRNN structure is also successful on the diagnosis of chest diseases [8] .

Motivation of this Research:

ANN algorithms are successfully being implemented in identifying several physical and mental disorders like brain tumors, diabetes, kidney stones, thyroid problems, Alzheimer's and many more. Observing the accuracy of individual neural networks MLNN, PNN, GRNN, LVQNN, RBFNN and GRNN in different sectors, we realized it would be interesting to analyze a comparative study of these algorithms. So in this paper, we have explored their comparative analysis.

Our Contribution:

- Comparative Overview of these 5 Neural Networks are documented.
- The algorithms are applied on 3 benchmarking data sets collected from UCI machine learning repository, Autitsm data, Breast cancer data, and Haberman's Survival data. Autism data is examined to determine whether a patient has Autism or not. Similarly, Cancer data is analyzed to check whether a patient has Cancer. In the same manner, models identify the diabetic patient or person with Sepsis.
- To justify the results split of training testing data is considered including 5-fold cross-validation technique.
- The performance of each algorithm is analyzed and why some method works better than others is discussed in terms of computational time and complexity.

Paper Organization:

The rest of the article is structured as follows. In the next section, we discuss the basic structure of the artificial neural networks MLNN, PNN, LVQNN, GRNN and RBFNN that are used in this research. A comparative overview of these neural networks is also documented. In section three, first, a brief description of the several medical data sets are given which are used for numerical experiments. These data sets are acquired from the (UCI) repository[26]. Experimental results and comparisons between the algorithms are also demonstrated in this section. Finally, the paper is concluded in Section 4.

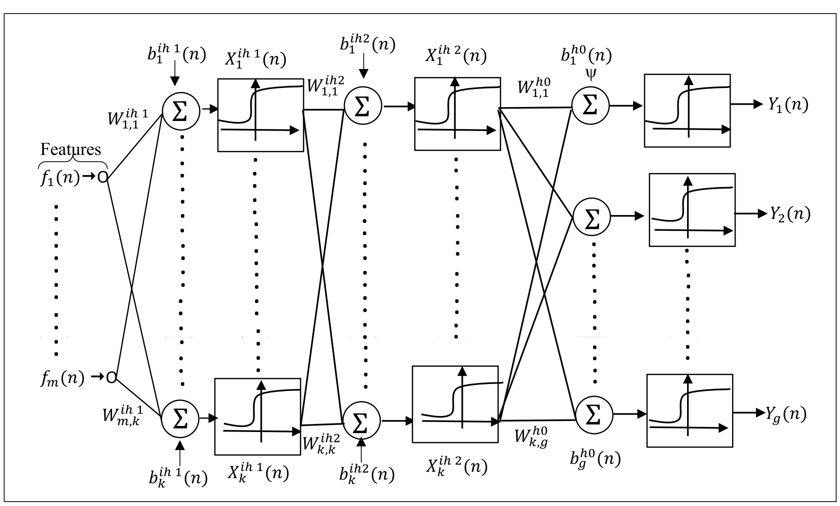

Figure 1. Multineural neural network with g class

2. Materials and Methods

2.0.1. *Multilayer Neural Network (MLNN).* An artificial neural network containing many layers of neurons is called a multilayer neural network, often known as a multilayer perceptron (MLP). In a one-hidden layer MLP, An input layer, a hidden layer, and an output layer are the three layers that make up the network. The input layer receives the input features, and each neuron represents a specific feature. The input layer is coupled to a hidden layer that handles intermediate calculations. Every neuron in the hidden layer takes inputs from the input layer, processes those inputs using an activation function to create outputs, and then sends those outputs to the layer above. The output layer, connected to the hidden layer, creates the network's final output. It typically uses a different activation function suitable for the task, such as a sigmoid or softmax function for classification or a linear function for regression [3]. In a two-hidden layer MLP, an additional hidden layer is added between the input and output layers. This layer allows for more complex representations and can capture higher-level features in the data. Similar to the first hidden layer, the extra hidden layer's neurons carry out calculations by receiving input from the input layer, applying an activation function, and sending the results to either the output layer or the next hidden layer [10]. Using methods like backpropagation, the links between the layers' weights and biases are changed throughout the training phase to enhance the performance of the network. Equations utilized in the building of an MLNN with two hidden layers are displayed in Fig. 1. The first hidden layer neurons' outputs [11] are:

$$\vec{X}^{ih1}(n) = \frac{1}{1 + e^{(}W^{ih1}(n) * \vec{f}(n) + \vec{b}^{ih1}(n))} \tag{2.1}$$

The second hidden layer neurons' outputs are:

$$\vec{X}^{ih2}(n) = \frac{1}{1 + e^{(}W^{ih2}(n) * \vec{X}^{ih1}(n) + \vec{b}^{ih2}(n))} \tag{2.2}$$

Outputs of the network are:

$$\vec{Y}(n) = \frac{1}{1 + e^{(}W^{h0}(n) * \vec{X}^{ih2}(n) + \vec{b}^{h0}(n))} \tag{2.3}$$

Here, symbols used in the formation of MLNN are represented in table 1

Table 1. Symbols MLNN

n	training pattern index
$\vec{f}(n)$	features
$W^{ih1}(n)$	weights from the input layer to the first hidden layer
$W^{ih2}(n)$	weights from the first hidden layer to the second hidden layer
$W^{h0}(n)$	weights from the second hidden layer to the output layer
$\vec{b}^{ih1}(n)$	biases of the first hidden layer
$\vec{b}^{ih2}(n)$	biases of the 2nd hidden layer
$\vec{b}^{h0}(n)$	the biases of the output layer
$\vec{Y}(n)$	outputs for the class index

2.0.2. *Probabilistic Neural Network (PNN).* Probabilistic Neural Network (PNN) takes a probabilistic approach to modeling and prediction. It functions by calculating the likelihood that an input belongs to a specific class and is especially well suited for classification jobs.

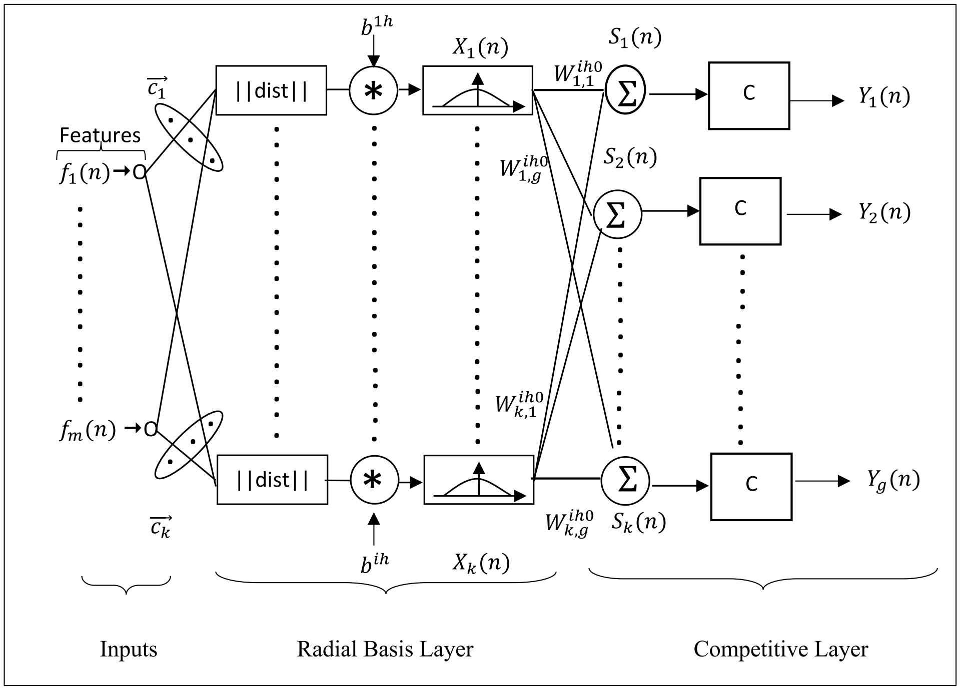

Figure 2. Probabilistic neural network with g class

In a PNN, the network consists of four layers: the input layer, pattern layer, summation layer, and output layer. During training, the PNN builds a density function for each class using a Parzen window estimation technique. This density function represents the probability distribution of the training samples for each class.

To make a prediction, the network is given an input pattern, and the PNN calculates the Euclidean distance between the input pattern and each training pattern in the pattern layer. The distances are then transformed into similarity values using a Gaussian function.

Next, the similarity values are passed to the summation layer, where they are summed for each class. The summed values represent the evidence for each class given the input pattern. The output layer then employs a winner-takes-all strategy, choosing the class with the strongest supporting data as the projected class for the input pattern. Fig.2 represents the basic structure of a PNN. The equations that make up the neural network model are [11, 25]

$$X_j = \phi(\|\vec{f} - \vec{c_j}\| * b^{ih}) \tag{2.4}$$

$$\phi(x) = e^{-x^2} \tag{2.5}$$

$$b^{ih} = \frac{0.833}{s} \tag{2.6}$$

$$S_i = \sum_{j=1}^{h} W_{ji}^{h0} * X_j \tag{2.7}$$

$$Y_j = \begin{cases} 1, & \text{if } S_i \text{ is max of } \{S_1, S_2, S_3\} \\ 0, & \text{otherwise} \end{cases} \tag{2.8}$$

where $i = 1, 2...g$ and $j = 1, 2, ..., h$; The symbols used in the model of PNN are represented in table 2

Table 2. Symbols used in PNN

n	training pattern index
$\vec{f}(n)$	features
Y_i	i^{th} output (classification in the index)
$\vec{W}^{h0}$	weight between the hidden layer and the output layer
$\vec{c_j}$	center vector of the j^{th} hidden node
S	spread factor
b^{ih}	bias term of radial basis layer
$\phi(.)$	radial basis function (rbf) (Gaussian)

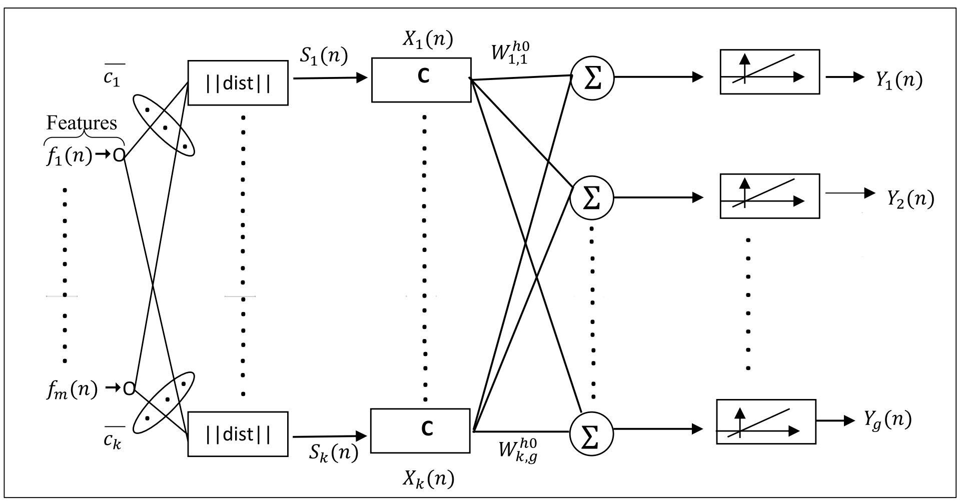

Figure 3. Learning Vector Quantization Neural Network with g class

The newpnn function from MATLAB was used with the PNN structures used in the study. [28] contains comprehensive details on the implementation of PNN structures.

2.0.3. *Learning Vector Quantization Neural Network (LVQNN).* A sort of artificial neural network that excels at classification tasks is the Learning Vector Quantization (LVQ) Neural Network [4, 6]. It is designed to learn and recognize patterns by organizing the input data set into predefined classes. The LVQ network consists of three main layers: the input layer, the codebook layer, and the output layer. A prototype or codebook vector associated with a certain class is represented by each neuron in the codebook layer, which receives the input features or variables. The codebook vectors are initialized randomly or based on some heuristics. During training, the network adjusts the codebook vectors based on a learning algorithm that aims to minimize the classification error. The output layer compares the input pattern with the codebook vectors and determines the closest match. The output neuron with the highest activation represents the predicted class for the input pattern. The basic structure of an LVQNN model is given in Fig.3. The equations that describe the model are (2.9), (2.10) and (2.11) given as follows:

$$S_j = \|\vec{f} - \vec{c_j}\| \tag{2.9}$$

$$X_j = \begin{cases} 1, & \text{if } S_j \text{ is max of } \{S_1, S_2, ...S_h\} \\ 0, & \text{otherwise} \end{cases} \tag{2.10}$$

$$Y_i = \sum_{j=1}^{h} W_{ji}^{h0} * X_j \tag{2.11}$$

where $i = 1, 2...g$ and $j = 1, 2, ..., h$; The symbols used in the model of LVQNN are described in table 3

Table 3. Symbols used in LVQNN

n	training pattern index
$\vec{f}(n)$	features
Y_i	i^{th} output (classification in the index)
$\vec{W}^{h0}$	weight between the hidden layer and the output layer
$\vec{c_j}$	center vector of the j^{th} hidden node

The newlvq function from MATLAB was used with the LVQ structures used in the investigation. The neural network toolbox section of the MATLAB Documentation contains comprehensive information regarding the realization of the LVQ structures.

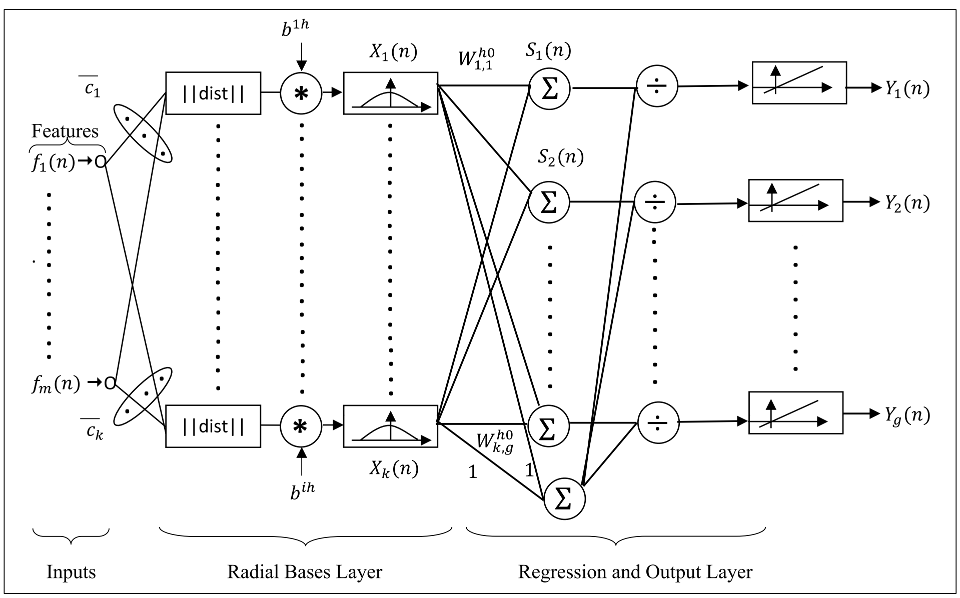

Figure 4. Generalized regression neural network with g class

2.0.4. *Generalized Regression Neural Network (GRNN).* The Generalized Regression Neural Network (GRNN) has been employed in diagnosing various diseases. The network architecture utilized for this task is depicted in Bayesian networks, commonly referred to as Generalized Regression Neural Networks (GRNNs), as introduced by [23, 24]. It comprises four main layers: the input layer, pattern layer, Gaussian layer, and summation layer. The input layer receives the input features or variables. The pattern layer stores the training patterns, where each node represents a specific training sample with its corresponding feature values. The Gaussian layer calculates the similarity between the input pattern and the training patterns using the Euclidean distance and a Gaussian function. Indicating the degree of similarity between the input and training patterns, this layer assigns similarity values. Finally, the summation layer aggregates the weighted outputs from the Gaussian layer, computing a weighted average based on the similarity values. The output from the summation layer represents the predicted regression result for the given input pattern. The GRNN structure is given in Fig. 4. It allows for effective modeling of complex relationships between input variables and target outputs, making it a valuable tool for regression analysis in various domains. The following equations are utilized in the neural network model (2.12), (2.13), (2.14) and (2.15)

$$X_j = \phi(\|\vec{f} - \vec{c_j}\| * b^{ih}) \tag{2.12}$$

$$\phi(x) = e^{-x^2} \tag{2.13}$$

$$b^{ih} = \frac{0.833}{s} \tag{2.14}$$

$$Y_i = \frac{\sum_{j=1}^{h} W_{ji}^{h0} * X_j}{\sum_{j=1}^{h} X_j} \tag{2.15}$$

where i=1,2,...,g and j=1,2,...,h; Table 4 represents the symbols used in the formation of GRNN.

Table 4. Symbols used in GRNN

n	training pattern index
$\vec{f}(n)$	features
Y_i	i^{th} output (classification in the index)
W_{ji}^{h0}	regression layer weights
$\vec{c_j}$	center vector of the j^{th} hidden node
S	spread factor
b^{ih}	bias term of radial basis layer
$\phi(.)$	rbf (Gaussian)

The newgrnn function from MATLAB was used with the GRNN structures used in the study. You can get more specific details on how the GRNN structures were implemented in [28]

2.0.5. *Radial Basis Function Neural Network (RBFNN).* The RBFNN consists of three main layers: the input layer, the hidden layer, and the output layer. The input layer receives the input features or variables. The hidden layer contains a set of rbf neurons, each with a center and a spread parameter. These neurons calculate the similarity between the input pattern and their respective centers using rbfs, such as Gaussian or inverse multiquadric functions. The output layer combines the outputs from the hidden layer neurons using weighted connections to produce the final output of the network. Using methods like least squares or gradient descent, the weights in the links between the hidden and output layers are established throughout the training phase. The RBFNN architecture shown in Fig.5is capable of capturing complex nonlinear relationships between the input and output, making it suitable for tasks such as function approximation, time series prediction, and classification. It offers flexibility in choosing appropriate rbfs and can effectively model complex patterns in the data. Equations that describe the RBFNN model are represented by (2.16), (2.17), (2.18) and (2.19) as follows:

$$X_j = \phi(\|\vec{f} - \vec{c_j}\| * b^{ih}) \tag{2.16}$$

$$\phi(x) = e^{-x^2} \tag{2.17}$$

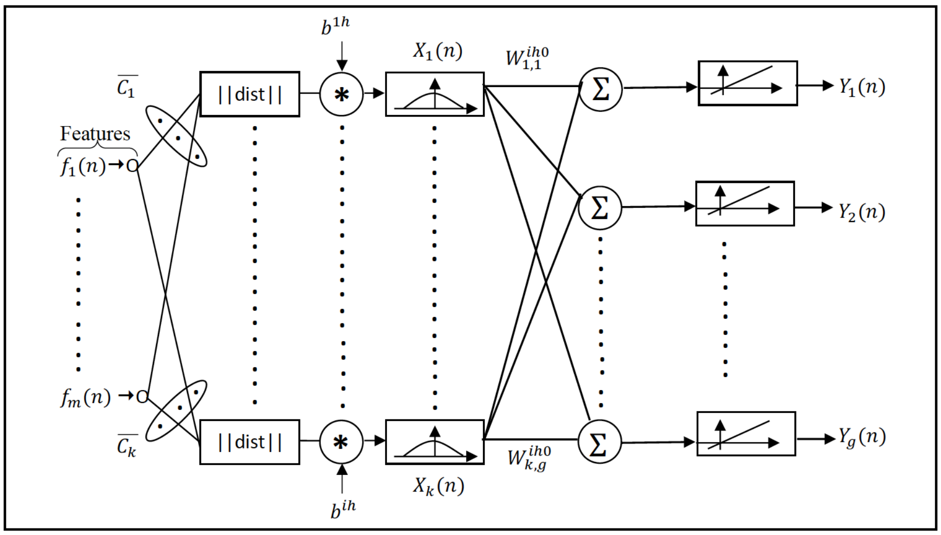

Figure 5. Radial basis neural network with g class

$$b^{ih} = \frac{0.833}{s} \tag{2.18}$$

$$Y_i = \sum_{j=1}^{h} W^{h0}_{ji} * X_j \tag{2.19}$$

where $i = 1, 2, ..., g$ and $j = 1, 2, ..., h$5 represents the symbols used in the model of RBFNN.

Table 5. Symbols used in RBFNN

n	training pattern index
$\vec{f}(n)$	features
Y_i	i^{th} output (classification in the index)
W^{h0}_{ji}	output layer weights
$\vec{c_j}$	center vector of the j^{th} hidden node
S	spread factor
b^{ih}	bias term of radial basis layer
$\phi(.)$	rbf (Gaussian)

The newrb function from MATLAB was used with the RBF structures used in the study. You may get more details about the RBF structures' implementation in [28].

To facilitate understanding of the selection and capabilities of each neural network type used in this study, Table 6 provides a comparative overview highlighting their key features, typical use cases in medical diagnosis, and reasons for their selection based on the specific needs of our datasets.

3. Experimental Results

3.1. **Data set Description.** This research focuses on the medical diagnosis of patients with different health issues. The data sets Autism data, Breast Cancer data, Diabetes data, and Haberman's Survival Data are collected from the UCI machine learning repository.

Table 6. Comparative Overview of Neural Network Types

Network Type	**Architecture**	**Learning Algorithm**	**Implementation in areas**	**Advantages**
MLNN	Multiple layers with one or more hidden layers	Backpropagation with gradient descent	General diagnosis for both binary and multi-class classification	High accuracy with capability for complex pattern recognition
PNN	Four-layer network with input, pattern, summation, and output layers	Direct computation based on statistical probability	Pattern classification with statistical basis, ideal for quick decision-making	Fast training, high precision with probabilistic approach
LVQNN	Three main layers including input, codebook, and output layers	Competitive learning with prototype vector matching	Categorical data classification, particularly where prototype exemplars are useful	Efficient for small datasets with clear categorical boundaries
GRNN	Four-layer network similar to PNN with a regression focus	One-pass learning that sets weights based on training data	Regression and continuous data predictions, useful for outcome forecasting	No need for iterative training, making it suitable for quick deployment
RBFNN	Three layers including input, hidden with radial basis neurons, and output	rbf for activation and typically trained using gradient descent or least squares	Function approximation, time series prediction, and classification where smooth interpolation of functions is beneficial	Fast convergence and can approximate any continuous function

Table 7. List of data sets used in this research

Data set	no.of instance	no. of attributes	no. of classes
Autism (A)	699	21	2
Breast Cancer (BC)	699	10	2
Haberman's Survival(HS)	306	3	2

Selection of Training Sets:

Each of the data sets is divided into training and testing sets. It is observed that the performance of a network is highly dependent on the choice of the training set. The training-testing split can be done in several ways. Here first, the selection of the training set is considered in three ways.

- Choice 1: Random selection without specifying a fixed portion of data from each class. (75% from the total data)
- Choice 2: Specifying the number of data from each class (first 75% of each class).
- Choice 3: Random selection of data with a specific portion (but not fixed data) from each class (any 75% of each class).

Later we considered a 5-fold cross-validation technique for the training-testing split of each data set. All the calculations were done using MATLAB.

3.2. **Autism Data set.** Autism data also known as Autistic Spectrum Disorder Screening Data for Adults is a two-class data having 704 instances and 20 attributes. According to the table (8), The LVQNN provides the greatest outcome in terms of recognition rate for Autism data whereas RBFNN yields the lowest accuracy compared to other neural networks.

Table 8. Final result for Autism Data

Method Name	Recognition rate
MLNN	97.8873%
PNN	99.2857%
LVQNN	**99.5261**%
GRNN	99.4083%
RBFNN	93.2927%

3.3. **Breast Cancer Data set.** Wisconsin Breast Cancer data is a two-class data having 699 instances and 9 attributes. The summary of our findings regarding the Breast Cancer Data is given in table 9:

Table 9. Final result for Breast Cancer Data

Method Name	Recognition rate
MLNN	83.3333%
PNN	**99.6700**%
LVQNN	99.0868%
GRNN	96.2963%
RBFNN	99.5146%

The results from the table (9) clearly indicate that the Probabilistic Neural Network is the most effective method among the five neural network methods presented in the table. It achieved an outstanding recognition rate of 99.6700%, surpassing all other methods in accurately classifying breast cancer data. Although the MLNN had the lowest accuracy among the five methods, both the LVQNN and the RBFNN exhibited exceptional performance with recognition rates above 99%.

3.4. **Haberman's Survival Data set.** Haberman's Survival Data set is a 2-class data with 306 instances and 3 attributes.

Experimental results regarding Haberman's Survival Data are provided in table 10:

Table 10. Final result for Haberman's Survival Data

Method Name	Recognition rate
MLNN	80.6452%
PNN	98.6667%
LVQNN	**99.4318**%
GRNN	97.1429 %
RBFNN	97.9592%

The results of the table (10) clearly show that the LVQNN is the best choice for Haberman's Survival Data. In terms of successfully predicting survival outcomes for this Data, it outperforms other techniques with a recognition rate of 99.4318%. In contrast to the MLNN, which has the lowest accuracy of the five techniques, the PNN performs better.
Table (11) and the two bar graphs Fig.(6), Fig.(7) represents the comparison of the five models in prediction class information of above mentioned three data sets.

Table 11. Comparison with all methods for three Data set

DS	Method				
	MLNN	*PNN*	*LVQNN*	*GRNN*	*RBFNN*
A	97.9%	99.3%	**99.5**%	99.4 %	93.3%
BC	83.3%	**99.7**%	99.1%	96.3%	99.5%
HS	80.6%	98.67%	**99.4**%	97.1%	97.9%

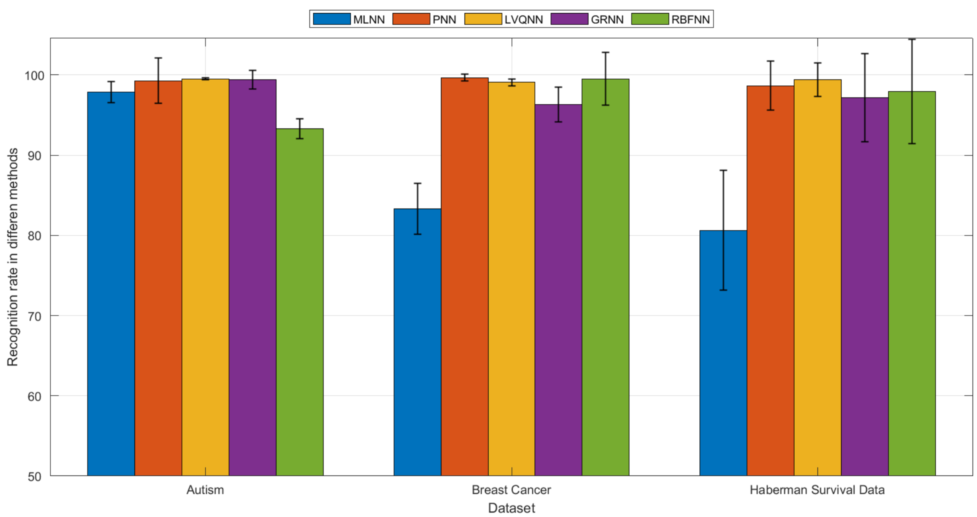

Figure 6. Bar graph for all Methods of three data sets

From the table (11) it can be observed that for autism data, all techniques exhibit high recognition rates, with PNN, LVQNN, and GRNN having recognition rates above 99%. However, with an identification percentage of 99.5261%, the LVQNN approach is clearly the most accurate. Despite having somewhat lower recognition rates of 97.8873% and 93.2927%, the MLNN and RBFNN nevertheless perform well.

The PNN approach is the undisputed victor in the dataset for breast cancer, with a remarkable identification rate of 99.6700%. Following closely behind, with recognition rates of 99.0868% and 99.5146%, respectively, are the LVQNN and RBFNN techniques. The MLNN approach has the lowest accuracy of the five methods, with a recognition rate of 83.3333%, while the GRNN method does quite well with a recognition rate of 96.2963%.

With a recognition rate of 99.4318% for Haberman's Survival dataset, the LVQNN algorithm once more leads the pack. The PNN also showed a good performance attaining a recognition rate of 98.6667%.

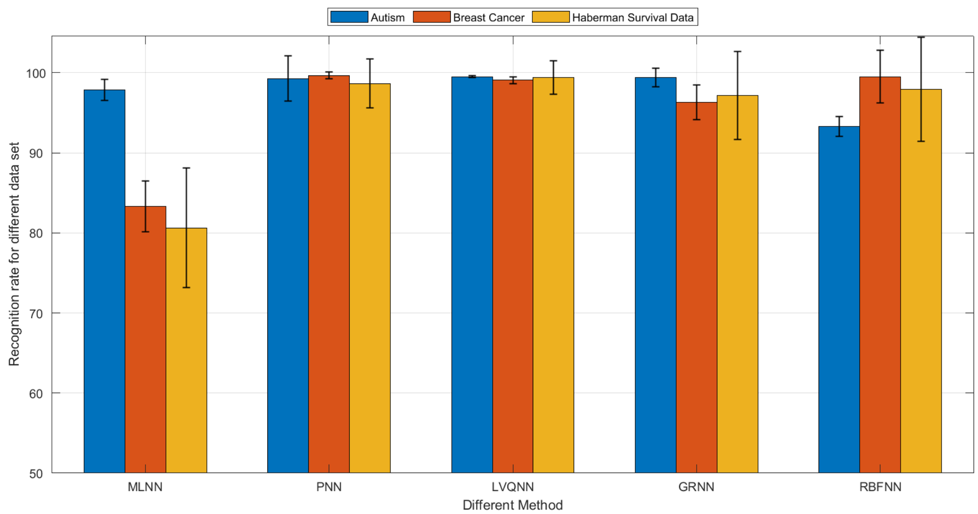

Figure 7. Bar Graph for each Method

With recognition rates of 97.1429% and 97.9592%, respectively, the GRNN and RBFNN techniques function rather well. The MLNN approach, however, falls short with the lowest recognition rate (80.6452%).

The error bar in the bar diagram (6) and Fig.(7) indicates the stability of the LVQNN model. It implies that, unlike the other 4 techniques, the performance of LVQNN does not vary that much for different training samples. Moreover, it remains stable for each of these three datasets.

We have also used 5-fold cross-validation for training testing split. The accuracy score for each of these data sets is given in table 12. It can be observed that the average accuracy of each methods obtained using 5-fold cross-validation techniques coincides with the result documented in table 11

From the experiments we have observed that the performance of neural network models on the various datasets can be attributed to their unique architectural and computational characteristics which align differently with the intrinsic properties of each dataset.

For instance, PNN is based on a statistical algorithm that excels in classifying data with a clear probabilistic distribution. The inherent ability of PNN to model the probability density function of the input data makes it exceptionally suitable for datasets where the categorical distinction is prominent, as evidenced by its superior performance on the Breast Cancer (BC) dataset. Its architecture allows for capturing subtle statistical nuances that can be critical in medical diagnosis.

Conversely, LVQNN utilizes a competitive learning algorithm that effectively identifies and adapts to the representative prototypes of the input space. This characteristic is particularly beneficial for disorders like Autism, where diagnosis relies on a spectrum of behavioral patterns. The LVQNN's robustness in pattern classification likely contributes to its impressive performance on the Autism (A) dataset, where recognizing these patterns is essential for accurate diagnosis.

Similarly, GRNN and RBFNN are designed to approximate functions that can map input features to output predictions smoothly. Their ability to handle continuous data with complex underlying relationships can be observed in their performance across all datasets, with RBFNN performing particularly well on the Breast Cancer (BC) dataset, where the underlying biological processes are continuous by nature and require sophisticated modeling capabilities.

Dataset	test data	MLNN	PNN	LVQNN	GRNN	RBFNN
Autism Data	fold-1	98.2	99.3	99.7	99.2	94.1
	fold-2	98.5	99.2	99.5	99.7	94.7
	fold-3	98.3	99.7	99.8	99.3	94.2
	fold-4	97.8	99.2	99.4	99.6	94.5
	fold-5	98.5	99.5	99.5	99.9	94.3
	Average Accuracy	**98.26**	**99.38**	**99.58**	**99.54**	**94.36**
Cancer data	fold-1	85.6	99.8	99.0	98.1	99.2
	fold-2	85.5	98.9	98.8	97.5	99.8
	fold-3	84.9	99.8	98.9	97.8	99.5
	fold-4	84.8	99.7	99.5	98.1	99.7
	fold-5	84.9	99.8	99.8	97.8	99.3
	Average Accuracy	**85.14**	**99.6**	**99.2**	**97.86**	**99.5**
H. Survival data	fold-1	82.1	99.5	99.7	97.5	98.4
	fold-2	82.6	99.3	99.6	99.0	98.7
	fold-3	82.8	99.1	99.8	98.9	98.2
	fold-4	81.9	98.9	99.2	98.2	98.6
	fold-5	82.5	99.5	99.3	98.5	98.3
	Average Accuracy	**82.38**	**99.26**	**99.52**	**98.42**	**98.44**

Table 12. Accuracy of 5-folds split using 5 methods implemented on the data sets

The Multilayer Neural Network (MLNN), with its deep architecture and back propagation learning, demonstrates versatility across a range of datasets. However, the potential for overfitting and the need for extensive training are limitations that might account for its lower performance in comparison to more specialized networks like PNN and LVQNN for certain medical diagnosis tasks.

Overall, these insights into the differential performance of neural network models highlight the importance of model selection in medical diagnosis tasks, underlining the necessity to match the network's strengths with the dataset's specific characteristics to achieve optimal diagnostic accuracy.

Computational Complexity and Training Times of different method:

To optimize each neural network, a systematic approach is employed to tune hyperparameters such as learning rates and layer sizes. Techniques such as grid search were utilized, leveraging MATLAB's optimization toolbox to iteratively explore parameter spaces. Computational complexity was a consideration, with networks like GRNN and RBFNN requiring significant resources due to their complex calculations, especially noticeable in larger datasets. Training times varied significantly across networks, with PNN demonstrating the quickest training due to its simpler calculations, whereas MLNN required the longest training period due to its iterative backpropagation and convergence processes.

4. Conclusion

This research provides a comprehensive exploration of applications of neural networks such as MLNN, PNN, LVQNN, GRNN, and RBFNN in medical diagnosis of patients with different diseases. A number of benchmarking datasets are collected to evaluate the performances of these networks. The research emphasizes the importance of selecting an appropriate training set as it has a great impact on network performance. Numerical experiments show the performance of different types of neural networks varies for different datasets. The LVQNN technique consistently displays better performance across a variety of datasets, obtaining the greatest identification rates in both the Autism and Haberman's Survival datasets. The PNN approach performs better than others for the Breast Cancer dataset. These techniques might be regarded as the best options by academics and professionals for precise categorization and prediction assignments in their respective fields.

References

[1] M. S. Abdullatef, F. N. Alzubaidi, A. Anees Al-Tamimi, Y. A. Mahmood, *Fatigue Life Estimation of High Strength 2090-T83 Aluminum Alloy under Pure Torsion Loading Using Various Machine Learning Techniques*, Fluid Dynamics & Materials Processing, **19**(8), (2023), 2083-2107. DOI: http://doi.org/10.32604/fdmp.2023.027266

[2] S. Bari, M.R. Arefin and S. Jahan, *On Gradient Descent and Co-ordinate Descent methods and its variant*, AJSE, **19**(3)(2020), 107-115. DOI: https://doi.org/10.53799/ajse.v19i3.103

[3] R. P. Brent, *Fast Training Algorithms for Multi-layer Neural Nets*, IEEE Trans. on N. Networks, **2**(3)(1991), 346–354. DOI: https://doi.org/10.1109/72.97911

[4] Z. Chen, Z. Yang, L. Zhu, P. Gao, T. Matsubara, S. Kanaya, M.A.U Amin, *Learning vector quantized representation for cancer subtypes identification*, Computer Methods and Programs in Biomedicine,**236**(2023):107543. DOI: https://doi.org/10.1016/j.cmpb.2023.107543

[5] F. C. Chen, M. H. Lin, *On the learning and convergence of the radial basis networks*, IEEE Int. Conf. on N. Networks, San Francisco, CA, USA, **2**(1993), 983–988. DOI: https://doi/org/10.1109/ICNN.1993.298691

[6] Yu-An Chung, H. Tang, J. Glass, *Vector-quantized autoregressive predictive coding*, Interspeech-2020, 3760-3764. DOI: http://doi.org/10.21437/Interspeech.2020-1228

[7] G. Daqi, W. Shuyan, J. Yan, *An electronic nose and modular radial basis function network classifiers for recognizing multiple fragrant materials*, Sensors and Actuators B – Chemical, **97**(2-3)(2004), 391–401. DOI: https://doi.org/10.1016/j.snb.2003.09.018

[8] A. A. El-Solh, C.-B. Hsiao, S. Goodnough, J. Serghani, B. J. B. Grant, *Predicting active pulmonary tuberculosis using an artificial neural network*, Chest, **116**(1999) 968–973. DOI: https://doi.org/10.1378/chest.116.4.968

[9] O. Er, C. Sertkaya, F. Temurtas, A. C. Tanrikulu, *A comparative study on chronic obstructive pulmonary and pneumonia diseases diagnosis using neural networks and artificial immune system*, J. of Medical Systems, **33**(6)(2009), 485–492. DOI: https://doi.org/10.1007/s10916-008-9209-x

[10] A. Gulbag, F. Temurtas, *A study on quantitative classification of binary gas mixture using neural networks and adaptive neuro fuzzy inference systems*, Sensors and Actuators B: Chemical, **115**(l)(2009), 252–262. DOI: https://doi.org/10.1016/j.snb.2005.09.009

[11] A. Gulbag, F. Temurtas, I. Yusubov, *Quantitative discrimination of the binary gas mixtures using a combinational structure of the probabilistic and multilayer neural networks*, Sensors and Actuators B: Chemical, **131**(1)(2008), 196–204. DOI: https://doi.org/10.1016/j.snb.2007.11.008

[12] N. H. H. M. Hanif, W. H. Lan, H. B. Daud, J. Ahmad, *Classification of control measures for asthma using artificial neural networks*, Scientific and Technical Publishing Company, Malaysia, ACTA Press, (2009), pp. 163-168.

[13] S. Jahan , *Discriminant analysis of regularized multidimensional scaling*, Num. Alg. Cont. & Opt, **11**(2)(2021), 255-267. DOI: https://doi.org/10.3934/naco.2020024

[14] M.D. Javeed, R. Nagaraju, R. Chandrasekaran, G. Rajulu, P. Tumuluru, M.F. Ramesh, S.K. Suman, R. Shrivastava, *Brain Tumor Segmentation and Classification with Hybrid Clustering, Probabilistic Neural Networks*, Journal of Intelligent & Fuzzy Systems, **45(4)** (2023), 6485-6500.

[15] D. Kumar, R. Rani, N. Vivek, N. Kumar, *Compare the Performance of Distinct Neural Networks Techniques to diagnose the kidney stone disease*, International Transactions on Electrical Engineering and Computer Science, **3(1)**(2024), 69–78. DOI: https://doi.org/10.62760/iteecs.3.1.2024.74

[16] H. Lijun, Z. Min, Z. Geng and W. Xuan, *Application of Multidimensional Features and GRNN in Electronic Nose for Detection of Hepatocellular Carcinoma*, Journal of Physics: Conference Series, **1971**(2021), 3rd International Conference on Electronic Engineering and Informatics (EEI 2021).

[17] C. R. Nagarathna, M. Kusuma, K. Seemanthini, *Classifying the stages of Alzheimer's disease by using multi-layer feed-forward neural network*, Procedia Computer Science, **218**(2023), 1845-1856. DOI: https://doi.org/10.1016/j.procs.2023.01.162

[18] Z. Peng, S. Okaneya, H. Bai, C. Wu, B. Liu, T. Shiina, *Proposal of dental demineralization diagnosis with OCT echo based on multiscale entropy analysis*. Math Biosci Eng. **21(3)**(2024), 4421-4439. DOI: https://doi.org/10.3934/mbe.2024195. PMID: 38549334.

[19] D. Popescu, M. El-Khatib, H. El-Khatib, and L. Ichim, *New Trends in Melanoma Detection Using Neural Networks: A Systematic Review*, Sensors, **22**(2)(2022), 1-41. DOI: https://doi.org/10.3390/s22020496.

[20] M. A. Samani, F. Khazaee, S. Ebadollahi, *Multi-Layer Perceptron and Radial Basis Function for Thyroid Disease Diagnosis*, 05 March 2024, PREPRINT (Version 1). DOI: https://doi.org/10.21203/rs.3.rs-3994161/v1

[21] S. Samsiana, S. Arifin, *Machine Learning Algorithms for Predicting Factitious Disorder Using the Learning Vector Quantization Method*, In E3S Web of Conferences, **500**(2024). 03010. EDP Sciences.

[22] J.J. Sonia, P. Jayachandran, Md, A.Q.; S. Mohan, A.K. Sivaraman, K.F. Tee, *Machine-Learning-Based Diabetes Mellitus Risk Prediction Using Multi-Layer Neural Network No-Prop Algorithm*, Diagnostics, **13**(2023): 723. DOI: https://doi.org/10.3390/diagnostics13040723

[23] D. F. Specht, *Probabilistic neural networks*, Neural Networks, **3**(1)(1990),109–118. DOI: https://doi.org/10.1016/0893-6080(90)90049-Q

[24] D. F. Specht, *A generalized regression neural network*, IEEE Trans. Neural Networks, **2**(6), (1991)568–576. DOI: https://doi.org/10.1109/72.97934

[25] F. Temurtas, *A comparative study on thyroid disease diagnosis using neural networks*, Expert Systems with Applications, **36**(1)(2009), 944-949. DOI: https://doi.org/10.1016/j.eswa.2007.10.010

[26] UCI Machine Learning Repository, http://www.ics.uci.edu/mlearn/MLRepository.html

[27] X.C. Zhang, J.G. Gong, F.Z. Xuan, *A deep learning-based life prediction method for components under creep, fatigue, and creep-fatigue conditions*, Int. J. of Fatigue., **143**(2021), 1-8. DOI: https://doi.org/10.1016/j.ijfatigue.2021.106236

[28] Neural network toolbox part of MATLAB Documentation (Matlab Documentation, 2004)

A. Begum, Department of Mathematics, University of Dhaka, Dhaka, Bangladesh
Email address: afsana-2015418469@math.du.ac.bd

M. M. Rahman, Department of Mathematics and Statistics, Bowling Green State University, Ohio, USA
Email address: mdmashr@bgsu.edu

S. Jahan, corresponding author, Department of Mathematics, University of Dhaka, Dhaka , Bangladesh
Email address: sjahan.mat@du.ac.bd