Jewel Journal of Scientific Research (JJSR) 6(1&2): 150–163, 2021
©Federal University of Kashere-Nigeria (Print ISSN: 2384 – 6267, Online ISSN 2756-651X)
jjsr.fukashere.edu.ng

A Gaussian Radial Basis Function for Classification of Diabetic Patients: A Case of FMC, Yola

Ahmed, H.*[1], Muhammed, M.B.[2] and Baba, I. A.[2]

[1]Department of Statistics and Operations Research, Modibbo Adama University, Yola, Adamawa State, Nigeria.

[2]Department of Mathematics and Computer Science, Federal University of Kashere, Gombe, Gombe State, Nigeria

[3]Department of Mathematical Sciences, Taraba State University, Jalingo, Taraba State, Nigeria.

***Corresponding Author:** hassanahmed.official@gmail.com; +2348032417537

Abstract

Diabetes is a chronic illness that has become a worldwide epidemic. Modern clinical practice shifting towards a patient-centered approach, using computer-based knowledge in finding and understanding health and illness concepts is critical. Thus, several artificial intelligence techniques such as machine learning algorithm, regression tree, and logistic regression are used to build a prediction or classification model to improve effective understanding of this illness. Therefore, this study utilizes the Gaussian Radial Basis Function (RBF) to classify diabetic patients using data obtained from the Federal Medical Center, Yola, Adamawa State, Nigeria. The data consists of 533 patients and their demographic variables and laboratory variables, which include age, mass of a patient, glucose level, pressure, insulin, and class variable. The SPSS version 21 for windows 10.1 was used to analyse the data. The network has 10 units hidden layer with softmax activation function and the output layer has 2 units with identity activation function. Also, the sum of squares is used as the error function. The analysis shows that the suggested Radial Basis Function (RBF) techniques generally achieve 94.3% correct classifications with 92.2 % correct classification for diabetes patients and 96.6% correct classifications for non-diabetes patients. Remarkably, the results indicate the effectiveness of applying RBF as significant method to predict and classify diabetic and non-diabetic patients to help in understanding variables that may influence diabetic treatment.

Keywords: Radial Basis Function, Neural networks, classification, diabetes mellitus, medical data.

Received: 16th Aug, 2021 *Accepted:* 11th Oct., 2021 *Published Online:* 25th Dec., 2021

Introduction[1]

The healthcare industry generates and stores a massive amount of data that can be used to forecast and analyse the overall healthcare ratio. It is possible to extract hidden and usable information from datasets known as Knowledge Discovery in Database (KDD) and Computer-based information system (CBIS) utilizing data mining (DM) approaches (Sohail, 2017). The healthcare field is notorious for its ontological challenges and a mix of medical data standards and varied data excellence (Witten, 2011; Bhattacharyya, 2006; Thomas, 2005;

Karegar, 2008). The modern clinical practice is undergoing a significance shift not just in terms of diagnosis and treatment, but also in terms of understanding health and illness concepts, moving away from disease-oriented problem resolution, toward a patient-centered approach, with computer-aided knowledge finding techniques playing a significant role (Chandamona, 2016).

Hypertension, diabetes, and coronary artery disease are currently the most common chronic health disorders in Sub-Saharan Africa (Haibinf, 2004). Infectious diseases such as HIV, tuberculosis (TB), and malaria are the leading causes of death in Sub-Saharan Africa (SSA); however, with increased international attention to these issues, treatment options are increasing and mortality rates are down (Haibinf, 2004); Joint United Nations program on (HIV/AIDS and WHO). Although, curing infectious diseases has resulted in greater life expectancy and an increase in the prevalence of noncommunicable diseases (Goodwin, 1997). Diabetes is a chronic illness that has become a worldwide epidemic. Traditional tribal groups' in developing countries are embracing a modern lifestyle while suffering chronic health problems associated with developed countries (Illayaraja, 2013). In SSA, the direct and indirect disease load surpasses the healthcare system's financial and human resources (Prather, 1997).

Hence, several artificial intelligence solutions such as machine learning algorithm, regression tree, and logistic regression are used to build a prediction or classification model to improve effective understanding of this illness. Therefore, the main aim of this research is to predict diabetes mellitus using the Gaussian RBF model, with network information, algorithm, confusion matrix, and ROC for validation. Thus, the study is presented as follows; section 2 discussed the literature review, section 3 covers the research methodology, section 4 discussed the results, and section 5 covers the discussion of the study and future work.

Literature Review

A cross-domain knowledge is critical for achieving practical results. The rapid advancement of healthcare automation has resulted in a massive amount of heterogeneous, both structured and unstructured data that may be studied and used for secondary purposes. While, the use of data mining in healthcare presents new issues that necessitate the use of specific methods, tools, and methodologies. To classify, cluster, and find hidden patterns in data, a variety of algorithms have been implemented. However, domain-specific challenges in healthcare are still being worked out. As Thomas (2005) and Syed (2002) discussed it, some particular problems shall be resolved to successfully apply data mining methods. Accordingly, the data mining applications are sub-optimal or impossible to achieve without resolving depersonalization, multi-relational, media data pre-processing, clinical data heterogeneity, and quality concerns.

Interestingly, a neural-network (NN) is a classification algorithm in the field of artificial intelligence (AI). Artificial Neural Networks (ANN) is a very powerful tool with the capability of pattern recognition (Russell and Norvig, 2003). ANN were designed to model the functioning of human brain. The linear classifiers separate objects by the value of a linear combination of their features. The feature of an object is represented by a vector. A weight vector is another vector to be trained with known observations. Although, there are several algorithms that are used to address such problems, such as support vector machines (SVM), multi-layer perceptron (MLP), and the radial basis function (RBF).

For example, Sisodia and Sisodia (2018) explore how to predict diabetes using three different classifiers: Naive Bayes (NB), Support vector machine (SVM), and Decision tree (DT), which performed on Pima Indian Diabetes Database (PIDD). The results indicate that Nave Bayes outperformed the other three algorithms with an accuracy of 76.30%. Also, Ambilwade and Manza, (2016) applied the function of the Fuzzy Inference System (FIS) and Multilayer Perceptron (MLP) in forecasting the risk of prediabetes and Type-2 diabetes using blood glucose levels. The experiment is conducted on 385 patient datasets on Matlab. In addition, Wang *et al.* (2019) explore diabetic mellitus prediction analysis in the context of imbalanced data with missing values. The study employed Naive Bayes to normalise the data by accounting for missing values, while the prediction is performed using the Random Forest (RF), by using PIDD database (Wang *et al.*, 2019).

Moreover, Sarwar et al., 2018) conducted a comparative research of diabetes mellitus prediction using several machine learning algorithms and also analyse numerous statistical indicators on PIDD dataset. Accordingly, five classifiers were utilised for predictive analysis: logistic regression (LR), k-Nearest Neighbor (KNN), support vector machine (SVM), random forest (RF), and decision tree. The SVM and KNN produce superior outcomes. To determine the model's effectiveness, the study divide the dataset in half, using 70% of the data as a training set and 30% as a test set. Similarly, Perveen et al., (2018) explore the importance of metabolic syndrome (MetS) as a risk factor for diabetes development. A logistic regression was used to identify the MetS that are associated with type 2 diabetes. Likewise, the study utilised Naive Bayes, Decision Tree, and J48 classifiers to do predictive analysis which shows that the Nave Bayes has 79% mean accuracy.

The Radial Basis Function (RBF), proposed by Broomhead and Lowe in 1988, is arose from the need to precisely interpolate a series of data points in a multidimensional space (Powell, 1987). The RBF falls under the category of functional link nets (Klaseen and Pao, 1990), the Green's functions of the Gram's operator associated with the stabilizer. Hence, the RBF network architecture is identical to that of the classical regularization network (Poggio and Girosi, 1990). An RBF network is created if the stabilizer has radial symmetry. The regularization network has three desirable qualities from the standpoint of approximation theory (Poggio et. al (1990); Klassen *et. al*, 1990).

Initially, RBFs were developed as a supervised technique for classification and regression applications (Amirian and Schwenker, 2020). Broomhead and Lowe advocated selecting the cluster centres either randomly or uniformly from the training samples and then maximising the output weights via a pseudo-inverse analytic solution (Broomhead and Lowe, 1988). RBFs can be trained in a single phase by randomly initialising the cluster centres and only training the output weights. Two-phase RBF training employs a variety of techniques for initialising the cluster centres and maximising the output weights. Since 1988, research has utilised both supervised and unsupervised methods to establish the centres. To establish these cluster centres, Moody and Darken presented an unsupervised approach (Moody and Darken, 1989), while Schwenker et al. proposed supervised vector quantization (Schwenker *et al*., 1994). Before training the output weights, Kubat (1998) and Schwenker and Dietrich (2000) used decision trees to independently locate centres. Finally, Schwenker et al. developed a third step for optimising the complete RBF network end-to-end via gradient descent, including output weights, cluster centre, and activation function parameters (Schwenker *et al*., 2001).

Numerous applications and implementations have prompted the application of several activation functions for RBFs in the literature (Amirian and Schwenker, 2020; Stoffel *et al*., 2020). The Gaussian function is the kernel that is favoured when data are modelled using a multivariate Gaussian distribution (Broomhead and Lowe, 1988). Apart from mature basic research, RBFs have been deployed to a wide variety of pattern classification and regression applications in recent years (Amirian and Schwenker, 2020). Nicodemou *et al*. (2020) used RBF networks to estimate three-dimensional hand poses (Nicodemou *et al*., 2020), Dehghan and Mohammadi (2014) used RBFs to approximate the numerical solution of Fokker-Planck differential equations, Li et al. (2019) used sparse multiscale RBFs to detect seizures in EEG signals, and Zhao et al.

(2019) predicted interfacial interactions using RBFs. RBFs are used to train models for classification and regression in the quantification of discrete and continuous pain application (Amirian *et al*., 2016).

RBFs have also been applied to computer vision and image classification problems. Friedhelm et al. classified hand-written digits using raw photos as feature vectors (Schwenker et al., 2001). Er et al. used principle component analysis (PCA) to extract features from facial photos and then processed these features using Fisher's linear discriminant (FLD) technique before classifying the patterns using RBFs (Er et al., 2002). Interestingly, the success of recent technological solutions has resulted in a paradigm shift toward automated AI application and representation learning using neutral networks. Given a sufficient number of RBF unit which can estimate any multivariate continuous function on a compact domain to arbitrary precision; it possesses the best approximation property since the unknown coefficients are linear. To our knowledge, this is among the very few studies that RBFs have been integrated into current prediction model in health domain.

Model and Analysis

Between 1st August, 2016 to 31st October, 2017, a total of five hundred and fifty-three (553) women were tested for diabetes at FMC, Yola. Three hundred and six were diabetic while two hundred and forty-seven were non-diabetic. The data collected was from records of patients at FMC, YOLA.

Observation with missing data were dropped from the analysis. The final dataset consists of 553 subjects, described by several clinical characteristics.

The classification task consists of predicting whether a patient would test positive for diabetes. The class labels of the data are 1 for diabetes and 0 otherwise. There are 6 predictor variables for 553 patients.

The data set have the following numeric attributes and they are:

1. “glucose”: Plasma glucose concentration 2 hours in an oral glucose tolerance test.
2. “pressure”: Diastolic blood pressure (mm Hg).
3. “insulin”: 2-Hour serum insulin (mu U/ml).
4. “mass”: Body mass index (weight in kg/(height in meters).
5. “age”: Age in years.
6. Class variable (0 or 1). The Class variable (6) is treated as 0 (false), 1 (true – tested positive for diabetes).

Radial Basis Function

The Radial Basis Function (RBF) network is a three-layer ($J1$-$J2$-$J3$) feedforward neural network, as shown in Figure 1. Each node in the hidden layer uses a radial basis function (RBF), denoted (r), as its nonlinear activation function. The hidden layer performs a nonlinear transform of the input, and the output layer is a linear combiner mapping the nonlinearity into a new space. Usually, the same RBF is applied on all nodes; that is, the RBF nodes have the nonlinearity $\Phi_i(\vec{x}) = \Phi(\vec{x} - \vec{c_i})$ $i = 1, \ldots., J_2$ where $\vec{c_i}$ is the prototype or center of the ith node and $\phi(\vec{x})$ is an RBF. The biases of the output layer neurons can be modeled by an additional neuron in the hidden layer, which has a constant activation function $\Phi_0(r)$=1. The RBF network achieves a global optimal solution to the adjustable weights in the minimum mean square error (MSE) sense by using the linear optimization method.

The RBF network has its origin in performing exact interpolation of a set of data points in a multidimensional space (Powell, 1987). It can be considered one type of functional link nets (Klaseen and Pao, 1990). It has a network architecture similar to the classical regularization network (Klaseen and Pao, 1990), where the basic functions are the Green’s functions of the Gram’s operator associated with the stabilizer. If the stabilizer exhibits radial symmetry, an RBF network is obtained. From the viewpoint of approximation theory, the regularization network has three desirable properties (Powell, 1990) It can approximate any multivariate continuous function on a compact domain to an arbitrary accuracy, given a sufficient number of units; it has the best approximation property since the unknown coefficients are linear. The solution is optimal by minimizing a functional containing a regularization term.

In pattern classification applications the Gaussian function is preferred (Brown *et. al*., 1993). Mixtures of Gaussians have been considered in various scientific fields. The Gaussian activation function for RBF networks is given by:

$$\varphi_j(\mathrm{X}) = \exp\left[-\left(\mathrm{X}-\mu_j\right)^T \Sigma_j^{-1}\left(\mathrm{X}-\mu_j\right)\right] \quad (1)$$

for $j = 1, \ldots, L$ where $\mathbf{X}$ is the input feature vector, L is the number of hidden units, μ_j and Σ_j are the mean and the covariance matrix of the jth Gaussian function. In certain approaches a polynomial term is added to the expression (1) as in (Girosi *et. al*, 1990) while in others the Gaussian function is normalized to the sum of all the Gaussian components as in the Gaussian-mixtures estimation (Micchelli, 1986).

Geometrically, a radial basis function represents a bump in the multidimensional space, whose dimension is given by the number of entries. The mean vector μ_j represent the location, while Σ_j models the shape of the activation function. Statistically, an activation function models a probability density function where μ_j and Σ_j represent the first and second order statistics.

The output layer implements a weighted sum of hidden-unit outputs:

$$\psi_k(\mathbf{X}) = \sum_{j=1}^{L} \lambda_{jk}\phi_j(\mathbf{X}) \quad (2)$$

for $k = 1, \ldots, M$ where λ_{ijk} are the output weights, each corresponding to the connection between a hidden unit and an output unit and M represent the number of output units. The weights λ_{ijk} show the contribution of a hidden unit to the respective output unit. In a classification problem if $\lambda_{ijk} > 0$ the activation field of the hidden unit j is contained in the activation field of the output unit k. In pattern classification applications, the output of the radial basis function is limited to the interval (0,1) by a sigmoidal function:

$$Y_k(\mathbf{X}) = \frac{1}{1+\exp\left[-\psi_k(\mathbf{X})\right]} \quad (3)$$

For k = 1, …, M.

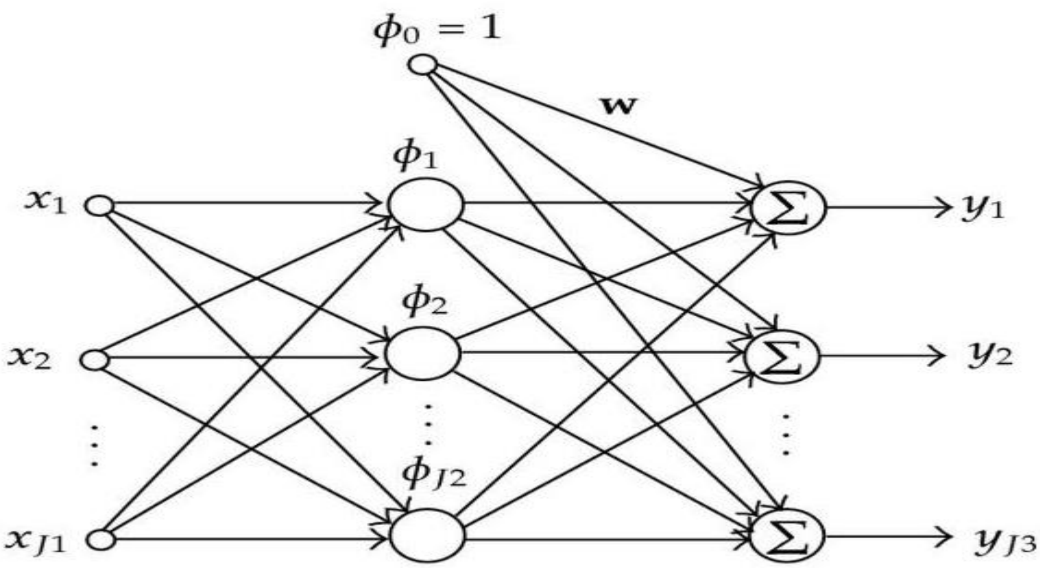

Figure 1: Architecture of the RBF network. The input, hidden, and output layers have $J1$, $J2$, and $J3$ neurons, respectively. $\phi_0(\bar{x}) = 1$ corresponds to the bias in the output layer, while $\phi_0(\bar{x})$'s denotes the nonlinearity at the hidden nodes.

Training algorithm.

By means of training, the neural network models the underlying function of a certain mapping. In order to model such a mapping we have to find the network weights and topology. There are two categories of training algorithms: supervised and unsupervised. RBF networks are used mainly in supervised applications. In a supervised application, we are provided with a set of data samples called training set for which the corresponding network.

Outputs are known. In this case the network parameters are found such that they minimize a cost function:

$$\min \sum_{i=1}^{Q} \left(\mathbf{Y}_k(\mathbf{X}_i) - \mathbf{F}_k(\mathbf{X}_i)\right)^T \left(\mathbf{Y}_k(\mathbf{X}_i) - \mathbf{F}_k(\mathbf{X}_i)\right) \qquad (4)$$

where Q is the total number of vectors from the training set, $Y_k(X_i)$denotes the RBF output vector and $F_k(X_i)$ represents the output vector associated with a data sample X_i from the training set.

Orthogonal least squares using Gram-Schmidt algorithm is proposed in (Lowe, 1995). An adaptive training algorithm for minimizing a given cost function is a gradient descend algorithm. Backpropagation adapts iteratively the network weights considering the derivatives of the cost function (2) with respect to those weights (Broomhead and Lowe, 1988). Expectation-maximization algorithm using a gradient descent algorithm for modeling the input-output distributions is employed. Backpropagation algorithm may require several iterations and can get stuck into a local minimum of the cost function (2). Clustering algorithms as k-means, or learning vector quantization have been employed for finding the hidden unit parameters in (Girona, 1990). The centers of the radial basis functions are initialized randomly. This algorithm is on-line and its first stage is unsupervised. For a given data sample Xi, the algorithm adapts its closest center:

$$\left\|\mathbf{X}_i - \hat{\boldsymbol{\mu}}_j\right\| = \min_{k=1}^{L}\left\|\mathbf{X}_i - \hat{\boldsymbol{\mu}}_k\right\| \qquad (5)$$

The Euclidean distance denoted by $||.||$ can be replaced with the Mahalanobis distance (Powell, 1990). In this situation we use the covariance matrix in the computation of the distance. In order to avoid the singularity of the covariance matrix a sufficient large number of data samples must be considered for the covariance matrix calculation used in the Mahalanobis distance. The center is updated as follows:

$$\hat{\boldsymbol{\mu}}_j = \hat{\boldsymbol{\mu}}_j + \eta(\mathbf{X}_i - \hat{\boldsymbol{\mu}}_j) \qquad (6)$$

where η is the training rate. For a minimal output variance, the training rate is equal to the inverse of the total number of data samples associated to that hidden unit. In this case the center corresponds to the classical first order statistical estimation. Similarly, second order statistical estimation can be employed for the covariance matrix. The output weights are evaluated in a second stage by means of Least Mean Square estimation. Outliers and data overlapping may cause bias in the parameter estimation.

The objective of this work is to evaluate the implementation and performance of Gaussian RBF in order to predict the presence of diabetes in a collected data from Federal Medical Center, Yola, Adamawa State, Nigeria. This paper describes how these techniques have been applied to the data and presents an analysis with classification, modelling, confusion matrix, network architecture, input, hidden and output layer of the RBF on SPSS version 21 for windows 10.2. The results are reported and discussed according to this technique.

Analysis and Discussion

Gaussian RBF was applied. In RBF, the Input layer has 5 factors with 315 units excluding the bias unit as seen from Table 1. The hidden layer has 10 units excluding the bias unit. The SoftMax was the activation function in the hidden layer. The output layer has 1 dependent variable which is 'class', with 2 units. Identity was the activation function. Sum of squares was the error function used.

368 cases were used in the training sample. the network weights that corresponded to the lowest mean squared error on the validation set were used for evaluation on the test data.

The training sample had 368 data points that

is about 67.6 % of the data while the testing sample had 176 data points which is 32.4% of the data. There were a total of 544 data point and 9 excluded cases.
It can be seen from Table 2 that the training data set 5.7% of in accurate prediction with 0.36 seconds taken to train the network. Also, the Testing data had 20.5% percentage of in correct classification with a Sum of Squares error of 24.351.

Table 1: Network Information

Input Layer	Factors	1	Glucose
		2	Pressure
		3	Insulin
		4	Age
		5	Weight
	Number of Units		315
Hidden Layer	Number of Units		10[a]
	Activation Function		Softmax
Output Layer	Dependent Variables	1	Class
	Number of Units		2
	Activation Function		Identity
	Error Function		Sum of Squares

Table 2: Model Summary

Training	Sum of Squares Error	25.740
	Percent Incorrect Predictions	5.7%
	Training Time	0:00:36.94
Testing	Sum of Squares Error	24.351
	Percent Incorrect Predictions	20.5%

Table 3: Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Glucose	553	4	9	6.52	1.476
Pressure	553	80	120	100.37	11.991
Insulin	553	16	166	92.01	43.860
Age	553	21	50	35.68	8.729
Weight	553	60	110	85.14	14.312
Valid N (listwise)	553				

Table 4: Confusion Matrix for classification

		Predicted		
Sample	Observed	0	1	Percentage of Correct
Training	0	170	6	96.6%
	1	15	177	92.2%
	Overall Percent	50.3%	49.7%	94.3%
Testing	0	51	14	78.5%
	1	22	89	80.2%
	Overall Percent	41.5%	58.5%	79.5%

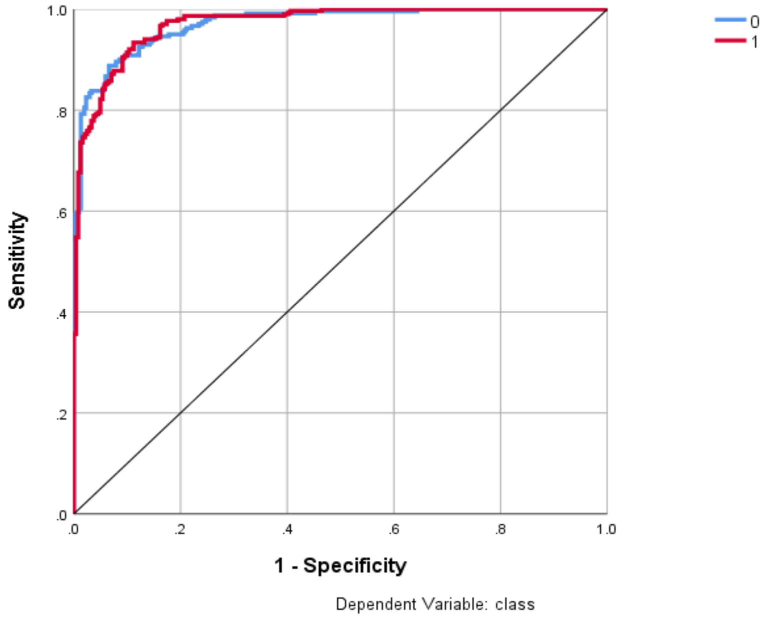

Figure 2. Receiver Operating Characteristics (ROC) plot

Table 5. Area Under the ROC Curve.

		Area
Class	0	0.971
	1	0.971

We observe from the ROC curve above that the curve is approaching the top left-hand corner which AUC = 1. From the table 5 above, for both class 0 and 1 the AUC = 0.971 which is strongly close to 1 and that means the diagnostic performance of the model (RBF) is strong.

For class 0 in the training sample, the network has 96.6 % correct classification and 92.2% correct classification for class 1. The testing sample has 78.5% correct classification and 80.2 % correct classification for classes 0 and 1 respectively. Overall, the training data was classified 94.3% correctly while the testing data was classified 79.5% correctly.

The study was carried out to see the classification power of the Gaussian RBF model. 553 records of data were collected on diabetic patients who were tested. The studied variables comprise of glucose level of each patient, diastolic pressure, insulin level, weight of each patient and their ages. The aim was to examine which of the two techniques classifies better. First, at the implementation stage, we chose to evaluate the method at its best performance, i.e. after optimization of the modeling specifications. This required to understand the meaning of each learning parameter and to test its influence on final results. In the analysis, (70%) of the data was used to train the network and 30% were used for testing the trained network. In training, the percentage of incorrect prediction is 5.7% while the percentage of incorrect prediction in testing is 20.5%. Remarkably, Gaussian RBF predicts Diabetic patients with 94.3% in this study.

The study proposed a machine learning model based on Gaussian RBF model which classified patients into two groups: diabetes and non-diabetic. Numerous articles in the literature are devoted to identifying high-risk factors for diabetes as well as sophisticated classification of the condition. Zou *et al.* (2018) used dataset derived from physical examination at a hospital in Luzhou, China. The dataset consisted of 220,680 patients and had 14 attributes. 69% of the patients were diabetic, whereas 31% were not. The study used three classifiers to classify diabetic patients, reporting the RF-based classifier with the highest classification accuracy of 80.84%. In Maniruzzaman *et al.* (2017) study, classified diabetic patients using difference classification method, Gaussian process classification (GPC) was one of them. The study demonstrated that using a GP-based classifier with a radial basis kernel (RBF) resulted in the highest classification accuracy of roughly 82%. Additional, Maniruzzaman *et al.* (2018) used eleven classifiers to classify diabetic patients. The result revealed that the RF-based feature selection technique with the RF-based classifier resulted in the best classification accuracy of 92.26%.

Table 5. Comparative analysis of previous works.

Authors	Best Performing Classifier	Database	Records	Performance
Zou *et al*. (2018)	RF	Hospital, Luzhou, China	220,680	80.84%
Maniruzzaman *et al.* (2017)	Gaussian RBF	PIDD		82%
Maniruzzaman *et al.* (2017)	RF	PIDD		92.26%
Ahuja *et al.* (2019)	MLP	PIDD	768	78.7%
Sisodia and Sisodia (2018)	NB			76.30%
Yu *et al.* (2010)	SVM	1999–2004 US NHANES	6214	83.50%
Semerdjian and Frank (2017)	GB GB	1999–2004 US NHANES	5515	
Mohapatra *et al.* (2019)	MLP			77.50%
Pei *et al.* (2019)	DT			94.20%
Maniruzzaman *et al.* (2020)	RF	2009-2012 US NHANES	6561	94.25%
Proposed Study	RBF	Hospital, Yola, Nigeria	553	94.30%

Ahuja *et al.* (2019) used PIDD dataser with 768 observations and ten attributes. The result demonstrated that the MLP had the maximum classification accuracy of 78.70%. Moreover, Sisodia and Sisodia (2018) applied SVM, NB, and DT classifiers and found that the NB classifier had the best accuracy of 76.30%. Yu *et al.* (2010) developed an SVM model to categorise diabetes patients using data from the 1999–2004 US NHANES. The results revealed that the SVMs with RBF kernels performed the best, with an accuracy of 83.50%. Similarly, Semerdjian and Frank (2017) analysed 5515 total samples from the 1999–2004 NHANES dataset. They determined the most significant risk factors using RF with the Gradient boosting (GB) classifier performing the best. Furthermore, Mohapatra *et al.* (2019) employed MLP and discovered that it provided 77.50% classification accuracy and Pei *et al.* (2019) applied DT and reported a classification accuracy of 94.20 percent. Recently, Maniruzzaman *et al.* (2020) applied RF-based classifiers and achieves the highest of 94.25% (see Table 1). However, in this work, we used the RBF classifier and achieved the highest classification accuracy of 94.3%. Interestingly, this study supported the findings by Maniruzzaman *et al.* (2017) who demonstrated that using RBF yield better performance.

Conclusions

Diabetes is one of the most commonly problems affecting people recently. It is a set of metabolic disorders characterised by elevated blood sugar levels. This study was tested in a machine learning-based system utilising an RBF classifier and demonstrated the highest classification accuracy. Our results indicated that our model achieved 94.25% classification accuracy. Additionally, a comparison analysis was performed with existing studies and our model slightly outperformed other classifiers. It would be intriguing to see other types of medical data classified similarly in the future, thereby establishing a cost-effective and time-saving solution for both diabetic patients and physicians.

Acknowledgment

We express our appreciation to Dr. S. S Abdulkadir, Department of Statistics and Operations research, Modibbo Adama University, Yola, Adamawa State, Nigeria, for guidance throughout the study. We also appreciate Federal Medical Center, Yola, for providing us with data to carry out the study.

References

Ahuja, R., Vivek, V., Chandna, M., Virmani, S. and Banga, A. (2019). Comparative study of various machine learning algorithms for prediction of insomnia. In *Advanced classification techniques for healthcare analysis.* IGI Global. pp. 234-257.

Ambilwade, R. P., & Manza, R. R. (2016, December). Prognosis of diabetes using fuzzy inference system and multilayer perceptron. In *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)* (pp. 248-252). IEEE.

Amirian, M., & Schwenker, F. (2020). Radial basis function networks for convolutional neural networks to learn similarity distance metric and improve interpretability. *IEEE Access*, *8*, 123087-123097.

Amirian, M., Kächele, M. and Schwenker, F. (2016). Using radial basis function neural networks for continuous and discrete pain estimation from bio-physiological signals. In *IAPR Workshop on Artificial Neural Networks in Pattern Recognition* Springer, Cham. pp. 269-284.

Bhattacharyya, D. and Hazarika (2006). Data mining & arti_cial intelligence: Trends & future directions (1st ed.), Narosa Pub House.

Bose, I. and Pal, R. (2006). Predicting the survival or failure of click-and-mortar corpora-tions: A knowledge discovery approach. Eur. J. Operat. Res., 174: 959-982.

Broomhead D. S. and Lowe D. (1988), “Multivariable functional interpolation and adaptive networks,” *Complex Systems*, 2(3): 321–355

Brown M., Harris C. J., and Parks P. C. (1993). The interpolation capabilities of the binary CMAC. *Neural Networks*, 6(3): 429–440.

Chandamona, M. and Ponperisasmy, D. (2016). Improved analysis of data mining techniques on medical data. *Int. J. Nano Corr Sci and Eng.* 3: 85-90.

Chao, S. and Wong, F. (2009). An incremental decision tree learning methodology regarding attributes in medical data mining. Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding.

Chinenye, S., Uloko, A.E., Ofoegbu, E., Fasanmade, F., Fasanmade, S., Ogbera, A. O. and Ogugua (2011). Profile of Nigerians with diabetes mellitus. Nigerian Medical Practitioner (in press).

Cotter N. E. and Guillerm T. J. (1992), The CMAC and a theorem of Kolmogorov. *Neural Networks*, 5(2): 221–228, 1992.

Dehghan, M., and Mohammadi, V. (2014). The numerical solution of Fokker–Planck equation with radial basis functions (RBFs) based on the meshless technique of Kansaʹ s approach and

Galerkin method. *Engineering Analysis with Boundary Elements*, 47: 38-63.

Er, M. J., Wu, S., Lu, J., and Toh, H. L. (2002). Face recognition with radial basis function (RBF) neural networks. *IEEE transactions on neural networks*, 13(3): 697-710.

Goodwin, L., Prather, J., and Schlitz, K. (1997). Data mining issues for improved birth out-comes. *Biomed. Science Instrum,* 34: 291-296.

Haibinf, M., Zhang, J. and Fan, Y. J. (2004). Mining frequent patterns based in IS+□ Tree. IEEE: 1208-1213.

Ilayaraja, M. and Alagappa (2013). Mining medical data to identify frequent diseases using Apriori algorithm. IEEE: 1-6.

Karegar, M. and Isazadeh, F. (2008). Data mining by probability-based patterns, 353-360.

Karegowda, A.G. and Jayaram, M.A. (2009). Cascading GA & CFS for feature subset selection in medical data mining. IEEE: 1-4.

Klaseen M, and Pao Y. H. (1990). The functional link net in structural pattern recognition," In *Proceedings of the IEEE Region Conference on Computer and Communication Systems (IEEE TENCON '90)*, Hong Kong, September 1990. vol. 2, pp. 567–571,

Kubat, M. (1998). Decision trees can initialize radial-basis function networks. *IEEE Transactions on Neural Networks*, 9(5), 813-821.

Li, Y., Cui, W. G., Huang, H., Guo, Y. Z., Li, K., & Tan, T. (2019). Epileptic seizure detection in EEG signals using sparse multiscale radial basis function networks and the Fisher vector approach. *Knowledge-Based Systems*, 164: 96-106.

Lowe D. (1995). On the use of nonlocal and non-positive definite basis functions in radial basis function networks," in *Proceedings of the 4th International Conference on Artificial Neural Networks*, Cambridge, UK, pp. 206–211.

Maniruzzaman, M., Kumar, N., Abedin, M. M., Islam, M. S., Suri, H. S., El-Baz, A. S., & Suri, J. S. (2017). Comparative approaches for classification of diabetes mellitus data: Machine learning paradigm. *Computer methods and programs in Biomedicine*, 152 23-34.

Maniruzzaman, M., Rahman, M. J., Ahammed, B., and Abedin, M. M. (2020). Classification and prediction of diabetes disease using machine learning paradigm. *Health Information Science and Systems*, 8(1): 1-14.

Maniruzzaman, M., Rahman, M. J., Al-MehediHasan, M., Suri, H. S., Abedin, M. M., El-Baz, A. and Suri, J. S. (2018). Accurate diabetes risk stratification using machine learning: role of missing value and outliers. *Journal of Medical Systems*, 42(5): 1-17.

Micchelli C. A. (1986), "Interpolation of scattered data: distance matrices and conditionally positive definite functions," *Constructive Approximation*, 2(1): 11–22.

Mohapatra, S. K., Swain, J. K., & Mohanty, M. N. (2019). Detection of diabetes using multilayer perceptron. In *International conference on intelligent computing and applications* (pp. 109-116). Springer, Singapore.

Moody, J., and Darken, C. J. (1989). Fast learning in networks of locally-tuned processing units. *Neural computation*, 1(2): 281-294.

Nicodemou, V. C., Oikonomidis, I. and Argyros, A. (2020). Single-shot 3D hand pose estimation using radial basis function networks trained on synthetic data. *Pattern Analysis and Applications*, *23*(1), 415-428.

Pampel, F. C. (2000). Logistic regression: A Prier. Sage publications, Thousand Oaks, CA.

Pei, D., Zhang, C., Quan, Y., and Guo, Q. (2019). Identification of potential type II diabetes in a Chinese population with a sensitive decision tree approach. *Journal of Diabetes Research*, 2019:1-7

Perveen, S., Shahbaz, M., Keshavjee, K. and

Guergachi, A. (2018). Metabolic syndrome and development of diabetes mellitus: Predictive modeling based on machine learning techniques. *IEEE Access*, 7: 1365-1375.

Podgorelec, V. and Maribor, H.M. (2005). Improving mining of medical data by outliers' predictions. IEEE: 1-6.

Poggio T., and Girosi F (1990), "Networks for approximation and learning," *Proceedings of the IEEE*, 78(9): 1481–1497.

Powell M. J. D (1987), "Radial basis functions for multivariable interpolation: a review," in *Algorithms for Approximation*, J. C. Mason and M. G. Cox, Eds., vol. 10 of *Institute of Mathematics & Its Applications Conference Series*, Oxford University Press, Oxford, UK, pp. 143–167.

Prather, J.C., Lobach, D.F., Goodwin, L.K. and Edward, W.H. (1997). Medical data mining: Knowledge discovery in a clinical data warehouse. American Medical Informatics Association.

Sarwar, M. A., Kamal, N., Hamid, W. and Shah, M. A. (2018, September). Prediction of diabetes using machine learning algorithms in healthcare. In *2018 24th International Conference on Automation and Computing (ICAC)*. IEEE. (pp. 1-6).

Schwenker, F., and Dietrich, C. (2000). Initialization of radial basis function networks using classification trees. *Neural Network World*, 10(3): 473-482.

Schwenker, F., Kestler, H. A., and Palm, G. (2001). Three learning phases for radial-basis-function networks. *Neural networks*, 14(4-5): 439-458.

Schwenker, F., Kestler, H. A., Palm, G., and Hoher, M. (1994). Similarities of LVQ and RBF learning-a survey of learning rules and the application to the classification of signals from high-resolution electrocardiography. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*. 1: 646-651.

Semerdjian, J., and Frank, S. (2017). An ensemble classifier for predicting the onset of type II diabetes. *arXiv preprint arXiv:1708.07480*.

Singh, A.S. and Masuku, M.B. (2014). Sampling techniques and determination of sample size in applied statistics research: An overview. *International Journal of Economics, Commerce and Management*, 2(11):1-22.

Sisodia, D., and Sisodia, D. S. (2018). Prediction of diabetes using classification algorithms. *Procedia Computer Science*, 132: 1578-1585.

Sohail, M.N., Ren, J.D., Uba, M.M. and Irshad, M. (2017). A comprehensive look at data mining techniques. Contributing to medical data growth: A survey of research reviewers. In: Patnaik S., Jain V. (eds), Recent Developments in Intelligent Computing, Communication and on the prevalence of diabetes mellitus in the North-Western Part of Nigeria 10 Devices. Advances in Intelligent Systems and Computing, Springer, Singapore. 752 pp

Stephen, V. (1997). Selecting and Interpreting measures of thematic classification accuracy. Remote Sensing of Environment, 77-89.

Stoffel, M., Gulakala, R., Bamer, F. and Markert, B. (2020). Artificial neural networks in structural dynamics: A new modular radial basis function approach vs. convolutional and feedforward topologies. *Computer Methods in Applied Mechanics and Engineering*, 364, 112-989.

Syed, S., Raza, A. and Meng, K. (2002). Symbolic exposition of medical datasets: A data-mining workbench to inductively derive data-de_ning symbolic rules. pp. 1-6.

Tang, P.H. and Tseng, M.H. (2009). Medical data mining using BGA & RGA for weighting of features in Fuzzy KNN classification. IEEE, 5: 1-6.

Thomas, H. and Paul, L. (2005). Statistics: Methods and Applications (1st ed). Stata Soft, Inc.

Uba, M.M., Asiribo, O.E. and Sohail, M.N.

(2017). Application of logistic regression modelling using fractional polynomials of grouped continuous covariates.

Uloma, A., Maurice, C. A., Godswill, A. N. and Dennis, O. U. (2014). Socio-demographic determinants of the knowledge of diabetes mellitus in Onitsha-North Local Government Area, Anambra State, Nigeria. *Determinants of Knowledge of Diabetes Mellitus Orient Journal of Medicine,* 26: 1-32.

Wang, Q., Cao, W., Guo, J., Ren, J., Cheng, Y., and Davis, D. N. (2019). DMP_MI: an effective diabetes mellitus classification algorithm on imbalanced data with missing values. *IEEE Access*, *7*, 102232-102238.

Wang, S. and Zhou, G.G. (2005). Application of fuzzy clusters analysis for medical image data mining. *IEEE*, 2: 1-6.

Witten I., Frank and Hall, M. (2011). Data Mining: Practical Machine Learning Tools & Techniques, Google e-book.

Xue, W. and Yanan S. Y. (2006). Research and application of data mining in traditional Chinese medical clinic diagnosis. *IEEE*, 4: 1-4.

Yu, W., Liu, T., Valdez, R., Gwinn, M. and Khoury, M. (2010). Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes *BMC Medical Informatics and Decision Making.* 10:16-23

Zhao, Z., Lou, Y., Chen, Y., Lin, H., Li, R. and Yu, G. (2019). Prediction of interfacial interactions related with membrane fouling in a membrane bioreactor based on radial basis function artificial neural network (ANN). *Bioresource Technology*, 282: 262-268.

Zou, Q., Qu, K., Luo, Y., Yin, D., Ju, Y. and Tang, H. (2018). Predicting diabetes mellitus with machine learning techniques. *Front Genet* 9: 5-15

Appendix I

SPSS Command:

```
RBF dependent variable [(MLEVEL = {S})] [dependent variable...]
                                  {O}
                                  {N}

  [BY factor list] [WITH covariate list]

[/EXCEPT VARIABLES = varlist]

[/RESCALE [COVARIATE = {STANDARDIZED**}] [DEPENDENT = {STANDARDIZED**}]]
                       {NORMALIZED    }              {NORMALIZED    }
                       {ADJNORMALIZED }              {ADJNORMALIZED }
                       {NONE          }              {NONE          }

[/PARTITION {TRAINING = {70**  } TESTING = {30**  } HOLDOUT = {0**   }}]
                        {number}           {number}           {number}
            {VARIABLE = varname                                      }

[/ARCHITECTURE [{[MINUNITS = {AUTO** } MAXUNITS = {AUTO** }]}]
                             {integer}            {integer}
                {NUMUNITS = integer                        }

               [HIDDENFUNCTION = {NRBF**}]]
                                 {ORBF  }
```

```
[/CRITERIA OVERLAP = {AUTO**}]
                     {number}

[/MISSING USERMISSING = {EXCLUDE**}]
                        {INCLUDE  }

[/PRINT [CPS**]  [NETWORKINFO**]  [SUMMARY**]  [CLASSIFICATION**]
        [SOLUTION]  [IMPORTANCE] [NONE]]

[/PLOT [NETWORK**]  [PREDICTED]  [RESIDUAL]  [ROC]
       [GAIN]  [LIFT]  [NONE]]

[/SAVE [PREDVAL[(varname [varname...])]]

       [PSEUDOPROB[(rootname[:{25     }] [rootname...])]]]
```