Prediction by Normalized Radial Basis Function Artificial Neural Network and model Selection for the Risk Factors Most Affected on Kidney Transplantation

Shaymaa Mohammed Shakir Ahmed [1]

Assist.Prof .Dr Nazeera Sedeeq Barznji [2]

1Department of statistics and informatics, College of Administration and Economics, University of Salahaddin, , Kurdistan Reign, Iraq.

2 Department of statistics and informatics, College of Administration and Economics, University of Salahaddin, Kurdistan Reign, Iraq.

Abstract

Kidney transplants are performed to help patients with chronic kidney disease or end-stage kidney failure when the kidneys can no longer properly filter waste, requiring either dialysis or transplantation. This study used Artificial Neural Networks, specifically the Radial Basis Function (RBF) method, to predict outcomes for 150 kidney failure patients from a hospital in the Kurdistan region, including those who had received transplants or were preparing for them. Six key risk factors impacting access to transplantation results show by using analyzed normalized importance of each independent variables that congenital kidney diseases identified as the most significant, followed by polycystic kidney disease, systemic lupus erythematosus, renal artery stenosis, diabetes, and chronic uncontrolled high blood pressure. The classification analysis showed high accuracy in predicting patient outcomes, demonstrating that the model in training and testing is reliable and clinically valuable. Additionally, ROC curve analysis and model possesses a robust capacity to accurately classify and has fit statistics confirmed ,the strong performance and significant improvements in F-score, indicating that RBF is an effective approach to enhance the predictive power and clinical relevance of machine learning models in kidney disease classification ,The research participation is good RBF in predicting with risk factors for kidney transplantation and achieved higher accuracy compared to traditional methods.

IBM SPSS 27 with version 4.1of R programming Language were applied for analyze data.

Keywords:

Artificial Intelligence (AI) , Radial Basis Function, Gaussian Kernel, Classification, Normalized Importance and Kidney Transplantation

التنبؤ بواسطة شبكة الأعصاب الاصطناعية باستخدام دالة الأساس الشعاعي المنظم RBF واختيار نموذج لعوامل الخطر الأكثر تأثيراً على زراعة الكلى

شيماء محمد شاكر أحمد [1]
م.ب.د. نظيرة صديق برزنجي [2]

shaymaa.shakirahmed@su.edu.krd
nazeera.kareem@su.edu.krd

.1قسم الإحصاء والمعلوماتية، كلية الإدارة والاقتصاد، جامعة صلاح الدين، إقليم كردستان، العراق

.2قسم الإحصاء والمعلوماتية، كلية الإدارة والاقتصاد، جامعة صلاح الدين، إقليم كردستان، العراق

الملخص

تُجرى عمليات زراعة الكلى لمساعدة المرضى المصابين بأمراض الكلى المزمنة أو الفشل الكلوي في مرحلته النهائية، عندما تعجز الكلى عن تصفية الفضلات بشكل صحيح، مما يتطلب إما غسيل الكلى أو زراعة الكلى. استخدمت هذه الدراسة الشبكات العصبية الاصطناعية، وتحديدًا طريقة دالة الأساس الشعاعي ((RBF، للتنبؤ بنتائج 150 مريضًا يعانون من الفشل الكلوي من مستشفى في إقليم كردستان، بمن فيهم أولئك الذين خضعوا لعمليات زرع أو كانوا يستعدون لها. تم تحليل ستة عوامل خطر رئيسية تؤثر على عملية زرع الكلى، حيث اظهرت النتائج عن طريق تحليل المتغيرات الطبيعية المستقلة المهمة , كانت أمراض الكلى الخلقية على أنها الأكثر أهمية، تليها أمراض الكلى المتعددة التكيسات، والذئبة الحمامية الجهازية، وتضيق الشريان الكلوي، والسكري، وارتفاع ضغط الدم المزمن غير المنضبط. أظهر تحليل التصنيف دقة عالية في التنبؤ بنتائج المرضى، مما يدل على أن النموذج قيد التدريب والاختبار موثوق وقيم سريريًا. بالإضافة إلى ذلك، يمتلك تحليل منحنى ROC والنموذج قدرة قوية على التصنيف الدقيق وقد أكدت إحصائيات الملاءمة الأداء القوي والتحسينات الكبيرة في درجة F، مما يشير إلى أن RBF هو نهج فعال لتعزيز القوة التنبؤية والأهمية السريرية لنماذج التعلم الآلي في تصنيف أمراض الكلى. النتائج البحثية لطريقة RBF كانت جيدة في التنبؤ بعوامل الخطر لزراعة الكلى وحققت دقة أعلى مقارنة بالطرق التقليدية.

1. Introduction

The term Artificial Intelligence (AI) presents prospective methods to improve the identification and administration of several aspects within the different sectors [12]. Machine learning (ML), Machine learning models have been extensively utilized for predicting disease, with performance evaluated using diverse measures[22].

The Radial Basis Function (RBF) networks are a type of Artificial Neural Network(ANN) that uses Gaussian functions [5], RBF finds frequent application in the classification tasks performed by support vector machines SVM[7]. RBF procedure produces a predictive model for one or more target variables based on values of predictor variables. It structures consist of three layers: input layer, hidden layer and output layer. The steps of RBF contain the training process of a RBFNN involves determining the parameters required for the hidden layer and output layer, which has three key steps: selecting centers, determining spreads, and training the output weights.

The main statistical tasks of RBF are prediction and approximation, for this purpose the researcher used RBF to predict risk factors caused to kidney transplant, which kidney transplant is a well-established intervention for addressing end-stage chronic renal failure and is associated with enhanced patient survival rates and improved quality of life. In this paper we investigate the potential of applying (RBFNN) architecture for the prediction by Normalized independent variables Importance(regression algorithms),Classification and[ROC Curve] Both can be employed for prediction in machine analyses focusing on the six risk factors caused Kidney Transplant to emphasize significate these factors with using those three methods of RBF analysis The researcher show in result RBF is better for predicting with risk factors for kidney transplantation and achieved higher accuracy compared to traditional methods as (logistic regression) we describe that in practical aspect, hence facilitating clinical decision-making and targeted interventions. [Congenital kidney diseases has most effect follows polycystic kidney disease, Disease of the immunity then Renal artery stenosis, Diabetes and finally Chronic, uncontrolled high blood pressure, in a sequence, Classification Analysis emphasize that Kidney transplanted cases show overall a currency and finally ROC –Curve analysis the refers of(0. 835) of all area under the ROC –curve indicates that all risk factors have significantive effect on dependent variable.

2.Objective:

This study aims using RBF for Prediction the risk Factors caused Kidney Transplant, throw Classification, and Normalized independent variables Importance.

3. Literature Review

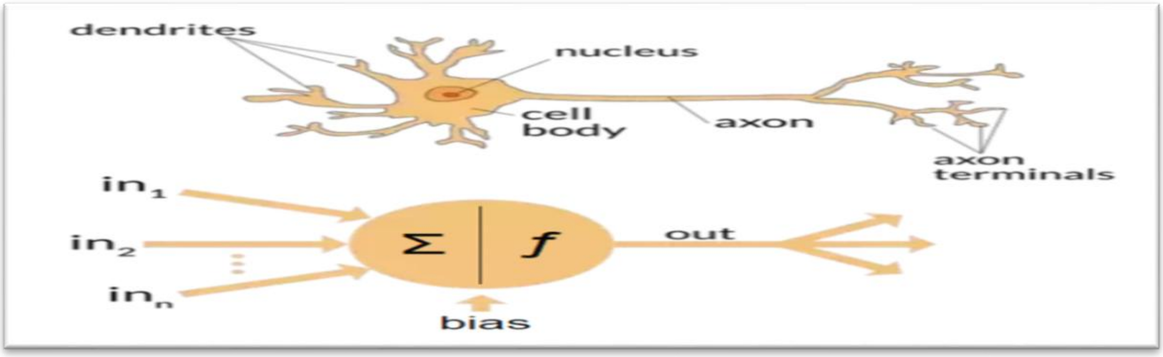
Author(s)	Year	Sector	Method	Result
Murat Kirişci[17]	2019	Neonatal health (birth weight)	ANN and Logistic Regression	Both models complement decision-making; assist clinicians in understanding and estimating birth weight risks.
Dong Chen[8]	2021	Physical education	Swarm Optimization Neural Network (RBFNN-PSO)	Faster convergence and lower training error compared to traditional RBF neural network.
Md. Imam Hossain & Mehadi Hasan Maruf[16]	2023	Cardiovascular disease	Data mining, AI, Random Forest (selected features)	Achieved 90% accuracy using selected features; outperformed other models; useful for early heart disease prediction.
Zhao, A. et al.[29]	2024	Control systems	RBF Neural Networks with Gaussian activation, HJB equation, Transfer learning	Effective optimal control, trajectory tracking, and system stabilization in dynamic systems.
Xin Zhang & Weihua Luo[27]	2024	Predictive modeling	PSO-based RBF network diffusion velocity function optimization; PSO-RBF intelligent coupling	Improved large-scale search efficiency and prediction accuracy at displacement increment mutation points.
Nguyen V. T. et al.[20]	2025	Robotics (ballbots control)	Adaptive nonlinear PID controller with RBFNN (NPID-RBFNN) and BCMO	Reduced chattering, strong robustness, smoother stable motion, verified stability, effective under disturbances.

4. Methodology (Theoretical Aspect)

4.1Artificial Intelligence (AI)

Artificial intelligence (AI) refers to computer systems capable of performing complex tasks that historically only a human could do, such as reasoning, making decisions, or solving problems. The criterion of Artificial Intelligence (AI) is fundamentally anchored in cognitive tasks, an idea of origins spanning several centuries and is revolved in various mythological narratives, The emergence of contemporary AI, as it is recognized today, while the term "Artificial Intelligence" was first articulated by John McCarthy in 1956 The cardinal aim of AI is this technology permits intelligent systems to realized, interpret, and assimilate information from data, accelerating the scope to make informed decisions established on cognizance derived from that data. [12]

Figure (1): A biological and an artificial neuron



4.2 Machine Learning

Within the domain of machine learning, referred to an artificial neural network ANN, and abbreviated as constitutes a model that draws inspiration from the architecture and operational dynamics of biological neural networks found in the brains of animals. An (ANN) artificial comprises interconnected units or nodes termed artificial neurons, which serve as the brain's neurons. Each artificial neuron receives input signals from affiliated neurons, subsequently processes these inputs, and transmits a signal to additional connected neurons. The "signal" is represented as a real number, and the fabrication generated by each neuron is determined through a non-linear function applied to the aggregate of its inputs, referred to as the activation function. The efficacy of the signal at each connection is governed by a weight that is subject to modification throughout the learning process The first use of artificial neural networks (ANNs) was made by psychologist Frank Rosenblatt, who also created the perceptron. However, some of the computational applications of ANNs are related to previous mathematical findings. Research on ANNs was performed in the 1970s and 1980s.The creation of Alex Net, deep neural network with several layers, occurred in the 2010s. as it shown in Figure (1)[22]. .

4.3: Model Types for Artificial Neural Nets:

Figure (2) shows Types for Artificial Neural as follows:

1. The Perceptron Single Layer.

2. Feed Forward Neural Network with Radial Basis Functional Neural Network.

3. A multilayer perceptron.

4- Neural Network with Recurrence.

5. Boltzmann apparatus

6. The Hopfield Network

7.Long Short-Term Memory, or LSTM

4.4 Definition of the Radial Basis Function (RBF)

In the field of mathematical modeling, a **radial basis function network** is an artificial neural network that uses radial basis functions as activation functions. The output of the network is a linear combination of radial basis functions of the inputs and neuron parameters. Radial basis function networks have many uses, including function approximation, time series prediction, classification, and system control. They were first formulated in a 1988 paper by Broom head and Lowe, both researchers at the Royal Signals and Radar is one of the simplest form of ANN consisting of exactly three layers, namely, input, hidden, and output layer. The restriction of only three layers makes it simplest and somehow efficient ANN architecture. The idea of RBFNN has been derived from function approximation. An RBF network positions one or more RBF neurons in the space described by the predictor variables. This space has as many dimensions as there are predictor variables. The Euclidean distance is computed from the point being evaluated to the center of each neuron. The RBF is so named because the radius distance is the argument to the function. Output of RBFNN depends on the distance of the input from a given stored vector.The RBF neural network is described by [4].

4.5. Basic functions in Radial Basis Function (RBF)

The RBF encompasses various basis interpolation functions as outlined as table (1) in Appendix [18].

4.5.1 Kernel: The kernel function serves the purpose of converting n-dimensional inputs into m-dimensional inputs, where m significantly exceeds n, thereby facilitating the efficient computation of dot products in a higher-dimensional space. The fundamental principle underlying the utilization of a kernel is that a linear classifier or regression line in elevated dimensions is transformed into a non-linear classifier or regression curve within lower dimensions[7].

4.5.2 Kernel Radial Basis Function

Within the domain of machine learning, the radial basis function kernel, commonly referred to as the RBF kernel, is a widely utilized kernel function employed across various kernelled learning algorithms. Notably, it finds frequent application in the classification tasks performed by support vector machines[7].

The RBF kernel on two samples$X \in R^k$ and X', represented as feature vectors in some input space, is defined as

$$K(x, x') = exp\left(-\frac{(\|x-x'\|)^2}{2\sigma^2}\right) \qquad \text{............. (1)}$$

where, $(\|X - X'\|)^2$ the squared Euclidean distance between the two data points is denoted by the expression above. The parameter represented by σ is commonly referred to as the bandwidth or width of the kernel, which regulates the smoothness of the decision boundary.

Upon the elaboration of the aforementioned exponential expression, it shall extend to an infinite degree of X and X' consistent with the expansion of the value of the Radial Basis Function (RBF) kernel diminishes with increasing distance and spans from zero (in the limit of infinite distance) to one (when x = x'), it readily lends itself to interpretation as a measure of similarity The feature space associated with the kernel possesses an infinite number of dimensions; for $\sigma = 1$, its expansion via the multinomial theorem is[10]:

$$exp\left(-\frac{1}{2}\|X - X'\|^2\right) = exp(X^T X')\, exp(-\frac{1}{2}\|X\|^2)\, exp(-\frac{1}{2}\|X'\|^2)$$

$$= \sum_{j=0}^{\infty} \sum_{n_1+n_2+\cdots.+n_k=j} exp\left(-\frac{1}{2}\|X\|^2\right) \frac{x_1^{n_1}\ldots x_k}{\sqrt{n_1!\ldots.n_k!}}\, exp\left(-\frac{1}{2}\|X'\|^2\right) \frac{x'_1{}^{n_1}\ldots x'_k{}^{n_k}}{\sqrt{n_1!\ldots.n_k!}}$$

$$=\langle \phi(X).\phi(X')\rangle \qquad \ldots (2)$$

$$\phi(X) = exp\left(-\frac{1}{2}\|X\|^2\right)\left(a_{\ell_0}{}^{(0)}.a_1{}^{(1)}.\ldots.a_{\ell_1}{}^{(1)}.\ldots.a_1{}^{(j)}.\ldots.a_{\ell_j}{}^{(j)}\right)$$

Where $\ell_{j=} \binom{k+j-1}{j}, a_{\ell_j}{}^{(j)} = \frac{x_1^{n_1}\ldots x_k^{n_k}}{\sqrt{n_1!\ldots.n_k!}} \; |n_1 + n_2 + \cdots.+n_k = j \wedge 1 \le \ell \le \ell_j$

.4.6Gaussian kernel in RBF

4.6.1 Gaussian Kernels

The Gaussian kernel is widely used in probabilistic models such as Gaussian Processes [21], where it plays a key role in defining the covariance structure of the distribution. also known as the squared exponential kernel – SE kernel – or radial basis function (RBF) is defined by

$$K(X.X') = \left(exp - \frac{1}{2}(x - x')^T\, \Sigma^{-1}(x - x')\right)$$

Σ the covariance of each feature across observations, is a p –dimensional matrix. When Σ is a diagonal matrix, this kernel can be written as

$$K(X.X') = exp\left(-\frac{1}{2}\sum_{j=1}^{p} \frac{1}{\sigma^2{}_j}\,(x_j - x_j')^2\right)$$

σ_j can be interpreted as defining the characteristic length scale of feature j. Furthermore, if Σ is spherical, i.e., $\sigma_j = \sigma, \forall j$,

$$K(X.X') = exp\left(-\frac{\|X-X'\|^2}{2\sigma^2}\right) \qquad \text{... (3)}$$

For this kernel, the dimension of the feature space defined by $\phi(\cdot)$ is D = ∞. Our methods will allow us to avoid explicit computation of ϕ (·). We can easily compute the $n \times n$ Gram matrix using this relative of the Mahalanobis Distance, even though we have implicitly projected our objects to an infinite dimensional feature space. A Gaussian handle with RBF kernel is a effective modeling instrument since it has expansive bolster and all-inclusive approximates.RBF kernels are the most generalized frame of kernelization and is one of the most broadly utilized parts due to its closeness to the Gaussian distribution. A kernel (covariance function) depicts the covariance of the Gaussian handle irregular factors.

The Gaussian kernel is a normal choice for part functions

$$(\bar{X}|\lambda) = \sum_{I=1}^{N} a_i\, b_i(\bar{x}), \quad where\ \sum_{I=1}^{N} a_i = 1$$

$$b_i(\bar{X}) = \frac{1}{(2\pi)^{D/2}(\sigma)^{1/2}} exp[-1/2\,(\bar{x}-C)^T \sum_{I=1}^{-1}(\bar{x}-C)] \qquad \text{...(4)}$$

A mean vector and a covariance matrix (c and σ) together provide the most comprehensive representation of a Gaussian kernel.. The inverse of the covariance matrix serves to elucidate the interrelations among various features, endowing each kernel with an n-dimensional ellipsoidal geometry. This approach is typically more adaptable than employing the straightforward distance to the kernel centroid, which presupposes rigid independence among the variables. [1].

$$\phi_i(\bar{x}) = \frac{1}{\left(1+\frac{\|\bar{x}-c_j\|}{\sigma_j^2}\right)^{1/2}} \qquad \text{... (5)}$$

Their actuation for designs distant from the centroid of the RBF which is more noteworthy than the actuation of the standard frame for these designs.. The q-Gaussian RBF for the i^{th} unit is characterized as:

$$\phi(\bar{x}) = \begin{cases} 1-(1-q)\left(\frac{\|\bar{x}-C_j\|}{\sigma_j^2}\right)^{1/1-q} \\ if\ 1-(1-q)\left(\frac{\|\bar{x}-C_j\|}{\sigma_j^2}\right) \geq 0\ \} \\ 0 \qquad\qquad otherwise \end{cases} \qquad \text{...(6)}$$

Where q is a genuine esteemed parameter[13].

4.6.2 Supervised Learning for Gaussian RBF

For the Gaussian RBF organize, the RBF at each nods can be relegated a different width σ_i The RBFs can be encourage generalized to permit for self-assertive covariance networks Σ_i

$$\phi(\vec{x}) = e^{-\left(\frac{1}{2}\right)(\vec{x}-\vec{c_i})^T \Sigma_i^{-1}(\vec{x}-\vec{c_i})} \qquad \text{... (7)}$$

where $\Sigma_i \in R^{J1\times J1}$ represents a positive definite, symmetric covariance matrix. When Σ_i^{-1} is expressed in its general form, the configuration and orientation of the axes of the hyper ellipsoid are arbitrary within the feature space. In the case that

$\Sigma_i^{-1} = \mathrm{diag}\left(\frac{1}{\sigma^2{}_{i,1}} / \ldots / \frac{1}{\sigma^2{}_{i,J_1}}\right)$, it is entirely defined by a vector $\vec{\sigma_1} \in R^{J1}$, and each $\phi_i$constitutes a hyper ellipsoid whose axes align with the axes of the feature space.

For the J_1-dimensional input space, each Radial Basis Function (RBF) employing a diagonal Σ_i^{-1} possesses a total of *J1(J1 + 3)/2* independent adjustable parameters, whereas each RBF utilizing a uniform σ across all directions and each RBF employing a diagonal Σ_i^{-1} contains merely *(J1 + 1) and 2J1* independent parameters, respectively. A trade-off exists between the deployment of a compact network with a multitude of adjustable parameters and the utilization of a larger network with a reduced number of adjustable parameters.

When employing the RBF with a uniform σ across all directions, the gradients are derived as

$$\frac{\partial E}{\partial \vec{c}_m} = \frac{2}{N}\sum_{n=1}^{N} \phi_m(\vec{x}_n)\frac{\vec{c}_m-\vec{x}_n}{\sigma_m^2}\sum_{i=1}^{J_3} e_{n,i}\,\omega_{i,m}$$

$$\frac{\partial E}{\partial \sigma_m} = -\frac{2}{N}\sum_{n=1}^{N} \phi_m(\vec{x}_n)\frac{\|\vec{x}_n-\vec{c}_m\|}{\sigma_m^2}\sum_{i=1}^{J_3} e_{n,i}\,\omega_{i,m} \qquad \text{... (8)}$$

Similarly ,for the RBF using diagonal Σ_i^{-1} ,the gradients are given by

$$\frac{\partial E}{\partial \sigma_{m,j}} = -\frac{2}{N}\sum_{n=1}^{N} \phi_m(\vec{x}_n)\frac{(x_{n,j}-c_{m,j})^2}{\sigma_{m,j}^3}\sum_{i=1}^{J_3} e_{n,i}\,\omega_{i,m} \qquad \text{... (9)}$$

Adaptations for $\vec{c}_I$ and Σ_i occur along the directions of negative gradients. The **W** parameters. To avert the imposition of unreasonable radii, the updating algorithms may also be derived by incorporating a constraint term into the Mean Squared Error (MSE) that penalizes diminutive radii.[27]

$$E_c = \sum_i 1/\sigma_i \quad or\ Ec = \sum_{i,j} 1/\sigma_{i,j} \qquad \text{... (10)}$$

4.7 Training RBF Neural Networks

The training of a neural network entails the identification of appropriate values for weights and biases. In the majority of instances, training is executed via a methodology commonly referred to as the train-test paradigm. As delineated by Dhubkarya (2010), Radial Basis Function (RBF) networks are predominantly employed in supervised contexts. The training dataset comprises specific conditions where the corresponding network outputs are predetermined. The works .The training procedure is utilized to as a certain the values of weights, centers, and widths. The optimization algorithm employed for minimizing a suitable error function (e) is applied to compute the weights, as referenced by Hamadneh (2012). RBF networks are generally trained utilizing pairs of input and target values

$x(t),y(t).t = 1.\dots.T$ through a two-stage algorithm.

RBF networks are typically trained from pairs of input and target values $X(t).y(t).1.2.\dots T$ by algorithm of simply fits a linear model with coefficients w_i to the hidden layer's outputs with respect to some objective function. A common objective function, at least for regression/function estimation, is the least squares function:

$$K(w) \stackrel{\text{def}}{=} \sum_{t=1}^{T} K_t(w) \qquad \dots (11)$$

Where

$$K_t(w) \stackrel{\text{def}}{=} [y(t) - \varphi(X(t).w)]^2 \qquad \dots (12)$$

There are occasions in which multiple objectives, such as smoothness as well as accuracy, must be optimized. In that case it is useful to optimize a regularized objective function such as

$$H(w) \stackrel{\text{def}}{=} K(w) + \lambda S(w) \stackrel{\text{def}}{=} \sum_{t=1}^{T} H_t(w) \qquad \dots (13)$$

Where $S(w) \stackrel{\text{def}}{=} \sum_{t=1}^{T} S_t(w)$ And $H(w) \stackrel{\text{def}}{=} K(w) + \lambda S(w)\ S(w) \stackrel{\text{def}}{=} \sum_{t=1}^{T} S_t(w)$

where optimization of S maximizes smoothness and $\lambda$is known as a regularization parameter [23].

4.8Testing

Upon the construction of a machine learning model (utilizing the training dataset), it is imperative to obtain unseen data to evaluate the efficacy of the model. This dataset is referred to as the testing data, that we find it in practical aspect which can be employed to assess the performance and advancement of the algorithms' training while allowing for adjustments or optimizations for enhanced outcomes. The objective of testing is to juxtapose the outputs generated by the neural network against the targets contained within an independent set (the testing instances). It is noteworthy that the testing methodologies are contingent upon the type of project (either approximation or classification).

4.9 Architecture:

The architecture of a Neural Network - RBF delineates the specific configuration of layers and nodes within the network as.

1- layers -2 weights 3- Bais 4-Activation Function

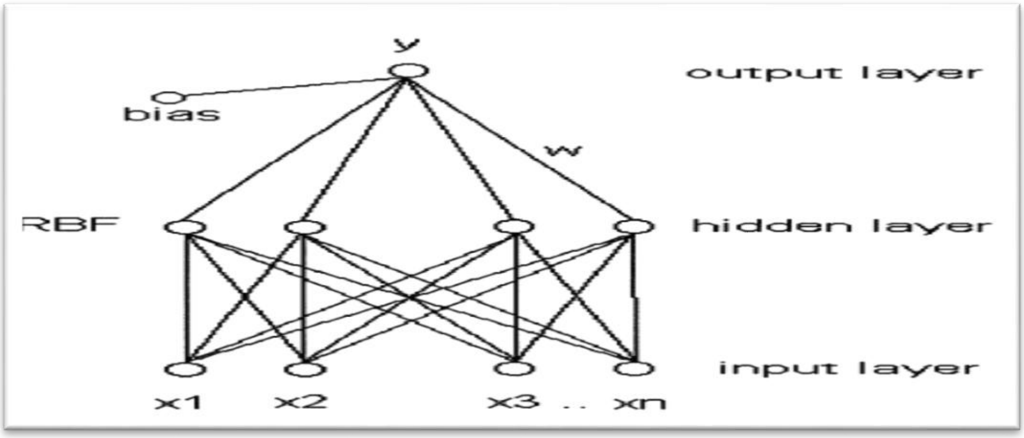
RBFN can be presented as a three-layer feed for-ward structure) see Fig. 2) The input layer serves only as input distributor to the hidden layer. Each node in the hidden layer is a radial function, its dimensionality being the same as the dimensionality of the input data. The output from the j^{th} function for the input object xi has the following form[23].

$$\phi_{ji} = \phi_j(x_i) = \phi\left(\|x_i - c_j\|\right) \qquad \text{... (14)}$$

$where$ $\|...\|$ is a distance measure and c_j describes the center of the j^{th} radial basis function and j=1, . . . , K.

At this point the notation of the centers of RBFs has been changed to stress that they are no longer constrained to be defined by the input data vectors

Figure (2): Architecture of radial basis function network



The output is calculated by a linear combination (i.e. a weighted sum. of the radial basis functions plus the bias, according to:

$$y(X_i) = \left[\sum_{i=1:k} w_j\, \phi\left(\|x_i - c_j\|\right) + w_0\right] \qquad \text{... (15)}$$

in matrix notation: $y = \phi w$ ...(16)

where

For multiple y the Eq. 15. has the following form:

$$y_b(\mathrm{X}_i) = \sum_{i=1:k} w_{bj}\, \phi(\|x_i - c_j\|) + w_{b0} \quad \text{... (17)}$$

In matrix notation:

$$Y = \Phi W \quad \text{... (18)}$$

In the RBFN, all basis functions have the same width parameter σ. In the more general case the width parameters can be tuned for each basis function individually. In the so-called hyper radial basis functions each argument is scaled by the matrix Q:

$$[\|x_i - c_j^i\|]^2 = [Q_{ij}^j(x_i - c_j)]^T \left[Q_{ij}^j(x_i - c_j)^j\right] \quad \text{... (19)}$$

4.9.1 Layers

a-Input Layer

The first layer is known as the input layer which accepts input for the network, The input layer are the input variables they referred to the visible layer, which includes multiple predictor variables, which accepts input for the network, with each variable corresponding to an independent neuron. At the input of hidden unit l, the input vector x is weighted by input weights W^h..

$$S_l = [X_1 W_1^h{}_l . X_2 W_2^h{}_l . \dots . X_n W_n^h{}_l \dots . X_N W_N^h{}_l] \quad \text{... (20)}$$

Where: n is the index of input; l is the index of hidden units; X_n is the n^{th} input; $W_{n.l}^h$ is the input weight between input and hidden unit; is the input weight. RBF network with N inputs, L hidden units, and M outputs [8]

b -Hidden Layer or a Processing Layer:

comprises a specific number of nodes with RBF non-linear activation function (conventionally it is implemented as Gaussian function), and the last layer is responsible to generate the final output of the network. The layers of nodes situated between the input and output layers are referred to as hidden layers. There can exist one or multiple instances of these layers. The hidden layer is responsible for processing the outputs derived from the input layer utilizing a distance metric of

$$d_l = \sqrt{\sum_{j=1}^{n} k_{1_l}^{i} [I_j - m_j^i]^2} \qquad ...(21)$$

and the transfer functions in. The resultant output of the network is represented as a weighted summation of the outputs of φ(di) emanating from the hidden layer and is articulated by the hidden layer, which comprises a RBF centered at a specific point, contingent upon the dimensionality of the input/output predictor variables. The activations of the hidden units are described by fundamental functions that typically encompass the activation function of the RBF. This activation function is contingent upon the Euclidean distance between the input pattern vector and the center of the hidden neuron, which possesses a multitude of neurons. Each neuron's output within the hidden unit l is computed by

$$\varphi_l(\overline{s_l}) = exp\left(\frac{-\|S_l - C_l\|^2}{\sigma_l}\right) \qquad ...\ (22)$$

where the activation function$\varphi_1(.)$ for the hidden unit is conventionally selected as a Gaussian function: C is the center of the hidden unit,1 and σ_1 is the width of the hidden unit$\|.\|$, with denoting the Euclidean distance. This indicates$\varphi_1(s_1)$ that possesses a feasible value solely when the distance $\|s_1 - c_1\|$ is less than the width [28].

c- Output Layer

The output layer is constituted of nodes that generate the output variables, which typically include a linear activation function that subsequently transmits the values to each neuron. The computation of the network output m is performed by:

$$o_m = \sum_{l=1}^{L} \varphi_l(s_l) W_{l,m}^{o} + W_{o,m}^{o} \qquad ...\ (23)$$

Where: m represented the index of output; $W_{1,m}^{o}$ is the output weight between the hidden unit1 and the output unit; $W_{o,m}^{o}$ is the bias weight of output unit$m$.Typically, the input weights are uniformly initialized to the value of '1'.

4.9.2-Weights

These are the real values related with the traits. They are significant as they tell the prominence of each trait which is thread as an input to the artificial neural network.

Neurons in neighboring layers are fully related to each other. Each connection has connected weight, which decide the strength of the connection. These weights are literate during the instruction operation [24] .

4.9.3-Bias Neurons

Bias in a neural network is desired to move the activation function towards the level either towards the left or the right. We will pall it in more posterior state .In the bonus to the input and hidden neurons, each layer (except the input layer) generally includes a bias neuron that allows a constant input to the neurons in the next layer. The bias neuron has its own weight related with each connection, which is also liberated during training. bias can mention to two connected but definite ideas: bias as a common expression in machine learning and the bias neuron , bias mentions to the error proposed by approaching to a real-world problem with a clarified model. Bias computes how well the model can arrest the underlying styles in the data. A high bias designates that the model is too naive and may under fit the data, while a low bias propose that the model is catching the underlying samples well.

4.10 Specialized algorithms normalized importance independent variables using neural networks

A higher normalized importance percentage he Gini's method. Having the higher importance percentage is meant the higher reliability and productivity, identification of significant independent variables through neural networks can be accomplished via a multitude of methodologies. One strategy involves evaluating the relative significance of each variable through an analysis of the connection weights and the sensitivity analysis inherent to the neural network model. An alternative approach entails employing an ensemble of neural networks rather than relying on a singular model, a strategy which enhances the consistency of the results derived from various methodologies Moreover, a technique that utilizes a combination of partial or wholly determined indices can be employed to ascertain the importance of specific nodes within the neural network. Independent variables serve as a crucial component of experimental design, allowing to investigate cause-and-effect relationships between variables.

Normalized architecture[6].

RBF networks can be normalized. In this case the mapping is

where $u(\|X - c_i\|) \stackrel{\text{def}}{=} \frac{\rho(\|X - c_i\|)}{\sum_{i=1}^{N} \rho(\|X - c_i\|)}$

is known as a normalized radial basis function.

There is theoretical justification for RBF architecture in the case of stochastic data flow. Assume a stochastic kernel approximation for the joint probability density

$$P(X \wedge y) = \frac{1}{N}\sum_{i=1}^{N} \rho(\|X - c_i\|)\, \sigma(|y - e_i|) \qquad \text{... (24)}$$

where the weights c_i and e_i are exemplars from the data and we require the kernels to be normalized $\int \rho(\|X - c_i\|)\, d^n X = 1$

And $\int \sigma(|y - e_i|)\, dy = 1$

The probability densities in the input and output spaces are

$$\int P(X \wedge y) dy = \frac{1}{N}\sum_{i=1}^{N} \rho(\|X - c_i\|)$$

And the expectation of y given an input X is

$$\varphi(x) \overset{\text{def}}{=} E(Y|X) = \int y P(Y|X) dy$$

where $P(y|X)$

is the conditional probability of y given X The conditional probability is related to the joint probability through Bayes' theorem

$$P(y|X) = \frac{P(X \wedge y)}{P(X)} \qquad \text{... (25)}$$

which yields $\varphi(X) = \int y \frac{P(X \wedge y)}{P(X)} dy$

This becomes

$$\varphi(X) = \frac{\sum_{i=1}^{N} e_i \rho(\|X - c_i\|)}{\sum_{i=1}^{N} \rho(\|X - c_i\|)} = \sum_{i=1}^{N} e_i u(\|X - c_i\|) \qquad \text{... (26)}$$

4.11 Performance Measurement [Classification, ROC and AUC]

In Machine Learning, performance measurement is an essential task. So when it comes to a classification problem, we can count on an AUC - ROC Curve. When we need to check or visualize the performance of the multi-class classification problem, we use the AUC (Area Under The Curve) ROC

(Receiver Operating Characteristics) curve. It is one of the most important evaluation metrics for checking any classification model's performance. It is also written as AUROC (Area Under the Receiver Operating Characteristics[2]. If you want to evaluate a model's quality across all possible thresholds, you need different tools

4.11.1Classification table and Confusion Matrix

During the data training phase, a confusion matrix was employed to assess the multilayer perceptron classification model presented in Table_2. that delineates the number of observation cases that exhibit a particular attribute alongside the number of observation cases that lack that characteristic, in contrast to the anticipated instances that display the attribute and those that do not. Consequently,

The fundamental premise of the analysis is to anticipate the accurate classification of cases based on a specified criterion, as this serves as evidence that the model aligns with the observational data. Table (2) illustrates the overarching structure of the classification table.

Confusion Matrix

Confusion matrix is a matrix which maps the predicted outputs across actual outputs.[1]

Table (2) Classification Table [Confusion Matrix]

Classification	Expected		
		Predictive Positive	Predictive Negative
Observation	Actual Positive	True Positive (TP)	False Negative (FN)
	Actual Negative	False Positive (FP)	True Negative Rate (TN)

Crucial metrics were derived from the classification matrix to facilitate a precise evaluation of the classification model, thereby enabling the prediction of one variable based on the known values of other variables. This process involves mapping the data item into a continuous real-valued variable. A plethora of soft-computing algorithms are available for data mining, including neuro-fuzzy computing, genetic algorithms, neural networks, rough sets, decision trees, and their respective hybridizations[3].

In statistical analysis pertaining to binary classification and information retrieval systems, the F-score, also referred to as the F-Measure as Eq.(27), serves as an indicator of predictive efficacy. This metric is derived from the precision and recall associated with the evaluation, wherein precision is defined as the ratio of true positive results to the total number of samples predicted to be positive, encompassing those incorrectly classified while recall is characterized as the ratio of true positive results to the total number of samples that should have been accurately classified as positive [19].

$$F_1\_score = \frac{2}{recall^{-1}+precision^{-1}} = 2\frac{precision.recall}{precision+recall} = \frac{2TP}{2TP+FP+FN} \qquad ...(27)$$

Where $Precision = \frac{TP}{TP} + FP$, $Recall = \frac{TP}{TP} + FN$

Since both precision and recall are rates (ratios) between 0 and 1 . This ensures that a low value in either precision or recall has a significant impact on the overall F1 score,

The highest possible value of an F-score is 1.0, indicating perfect precision and recall, and the lowest possible value is 0, if the precision or the recall is zero.

4.11-2 Receiver-operating characteristic curve (ROC)

The ROC curve is an analytical method, represented as a graph, that is used to evaluate the performance of a binary diagnostic classification method. The diagnostic test results need to be classified into one of the clearly defined dichotomous categories, such as the presence or absence of a disease[19].

4.11.2.1 Area under the curve (AUC)

The **area under the ROC curve (AUC)** represents the probability that the model, if given a randomly chosen positive and negative example, will rank the positive higher than the negative. The perfect model above means there is a 100% probability that the model will correctly rank a randomly chosen positive example higher than a randomly chosen negative example, In general, an AUC of 0.5 suggests no discrimination (i.e., ability to diagnose patients with and without the disease or condition based on the test),but (0.7 to 0.8) is considered acceptable,(0.8 to 0.9) is considered excellent, and more than (0.9)is considered outstanding. A ROC curve shows the relationship between clinical sensitivity and specificity for every possible cut-off. The ROC curve is a graph with[9].

$$\text{The x} - \text{axis showing (1 - specificity} = \text{false positive fraction} = \frac{FP}{(FP+TN)}$$

...(28)

$$\text{The y} - \text{axis showing (sensitivity} = \text{true positive fraction} = \frac{\text{TP}}{(\text{TP+FN})}$$

4.12. Model information criterion or **Model selection in machine learning**

4.12.1Model selection

is a machine learning operation utilized to select the finest model for a given task from a group of possible models? The possible models are estimated using different model selection methods The result of model selection depends on a strong validation plan and suitable estimation metrics. Model Selection task to Maximize achievement [11].

4.12.1.1-Akaike information criterion (AIC)

AIC count the type of a statistical model for a given dataset. It stabilities the trade-off between the goodness of the model's fit on the training data and the complexity of the model. It is an estimator of forecast error and just like comparative quality of statistical models for a given set of data. Given a groups of models for the data, evaluates the kind of each model, comparative to each of the other models. AIC penalizes models with more parameters, propitious the selection of easy that still illustrate the training data well(training dataset is the sample of data used to fit the model). A lower AIC value specifies a better-fitting model. AIC is focused to finding the finest approximating model to the unknown true data causing procedure and its applications[25].

$$\mathrm{AIC} = 2k - 2\log(L) \qquad \text{.........(29)}$$

Where L refers to the likelihood under the fitted model and

k is the number of parameters in the model. Specifically,

k = the number of parameters in the model. Specifically,

4.12.1.2. Bayesian Information Criterion (BIC)

Obtained from Bayesian probability and conclusion, BIC is a model selection statistic identical to AIC but contains a powerful penalty for model complication. BIC is especially fitting for models trained using maximum likelihood estimation.More compound models will have a larger BIC score, which means weak models Formula[25]:

$$\mathrm{BIC} = -2\log(\mathrm{L}) + \mathrm{k}\log(\mathrm{n}) \qquad \text{...(30)}$$

K = the number of parameters

L = the log-likelihood of the model for the training data

N = the number of data points/samples

4.12.2 Model Fitting Information -Likelihood Ratio Test

- The likelihood ratio test can be utilized to estimate the goodness of fit of a model of calculation if the sample is sufficiently large. In this case H1 identify to a 'filled' model in which the number of parameters equals the sample size n[14].

$$D = \text{likelihood ratio test} = -2 * \text{Ln}\frac{\text{Likelihood for null model}}{\text{Likelihood for alternative model}} \quad ...(31)$$

D is the test statistic.

4.13 Logistic Regression

is a supervised machine learning algorithm used for classification tasks, particularly binary classification, where the goal is to predict the probability of an outcome that can have only two possible values, such as yes/no, true/false, or 0/1. It models the relationship between one or more independent variables and the likelihood of the dependent variable using a logistic function (sigmoid function), which outputs a probability between 0 and 1.

Unlike linear regression, which predicts continuous outcomes, logistic regression is designed to handle categorical dependent variables by estimating the probability that a given input belongs to a specific class. It is widely used in various fields such as healthcare, finance, social sciences, and marketing to predict binary outcomes and identify important predictors.

In summary, logistic regression:

- Predicts binary or categorical outcomes.
- Uses independent variables (predictors) to estimate probabilities.
- logistic (sigmoid) function to produce probability values between 0 and 1.
- Helps in classification and understanding relationships between variables.

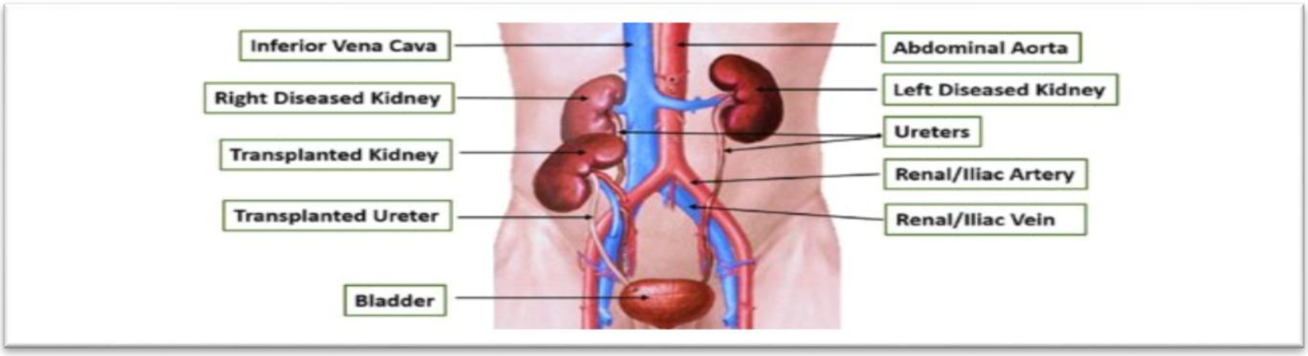
This model is foundational in machine learning and is often used to support decision-making and risk estimation in practical applications.

5-Practical Aspect

5.1 A brief explanation about kidney transplant

Kidney transplants are done to help people with chronic kidney disease or end-stage kidney failure. When the kidneys cannot filter waste properly, they need either dialysis (a device used to remove waste from the blood) or a kidney transplant. This medical procedure entails a surgical operation whereby the nonfunctional kidney is supplanted by a viable kidney sourced from either a living or deceased donor[15]. The inaugural successful live-related kidney transplant, performed in1953.

Figure (3): Kidney location after transplant[15]



The kidneys, which are two organs are situated bilaterally along the spine, just inferior to the ribcage. Each kidney is approximately the size of an adult fist. Their primary role is to filter and eliminate waste products, essential minerals, and excess fluid from the bloodstream via the production of urine. When the kidneys lose their capability to filter effectively, detrimental levels of fluid and waste begin to accumulate within the body, potentially leading to elevated blood pressure and eventual kidney failure, clinically referred to as end-stage renal disease (ESRD). End-stage renal disease is characterized by the loss of approximately 90% of the kidneys' functional capacity[26].

5.2 Practical Aspect Analysis

5.2.1 Data Description

The researchers contacted several institutions specializing in kidney disease, including the Shar hospital in Sulaimania governorate in Kurdistan, 267 patients were registered with kidney disease only the data from 150 patients in that hospital enrolled between February 2015 and May 2022 were eligible for this analysis, The (150 patients) is reasonable for this research , but it is important to point out that the results are preliminary and larger studies are needed to generalize them.

5.2.2 Variables Description

Dependent Variable (Y): defined as binary variable [patients with Kidney transplanted , Kidney prepare for transplant patients]

Independent Variables : In this paper we have deal with common risk factors caused Kidney transplant as independent variables for this purpose of the paper, many doctors specialized in kidney diseases were consulted, and confirmed that those variables are six risk factors caused Kidney transplantation (X_1 ,X_2,X_3,X_4,X_5 ,X_6) they explained in table(3)

5.2.3 Statistical Analysis

5. 2.3.1: first: Descriptive Statistics of Kidney dataset

Table (3): Shows descriptive statistics of Kidney transplantation dataset

Variables	Frequency	Percent

Dependent Y		Kidney transplanted	88	58.7%
		Kidney prepare for transplant	62	41.3%
Independents	X_1: Renal Artery Stenosis	Semi Control	54	36.0%
		Out Control	96	64.0%
	X_2: Polycystic Kidney Disease	Semi Control	57	38.0%
		Out Control	93	62.0%
	X_3:: Diabetes	Semi Control	62	41.3%
		Out Control	88	58.7%
	X_4: Congenital Kidney Diseases	Semi Control	48	32.0%
		Out Control	102	68.7%
	X_5 : Chronic, Uncontrolled High Blood Pressure	Semi Control	62	41.3%
		Out Control	88	58.7%
	X_6: Systemic Lupus Erythematosus	Semi Control	56	37.3%
		Out Control	94	62.7%

1-The dependent variable referred to the number of kidney transplanted are (88) with (58.7%) and it is higher than those number of Kidney prepare for transplant (62) with (41.3%)

2- independents variables out-control is higher than Semi Control in all variables

[X_1=64.0% , X_2= 62.0% , X_3=58.7% , X_4= 68.7% , X_5= 58.7% , X_6=62.7%]

it indicate that every risk factors has effect on the kidney transplant and the 62% of patients affected by risk factors

5.2. 3.2: Second: Advanced Statistics ANN -RBF of Kidney dataset

1 -Case Processing Summary

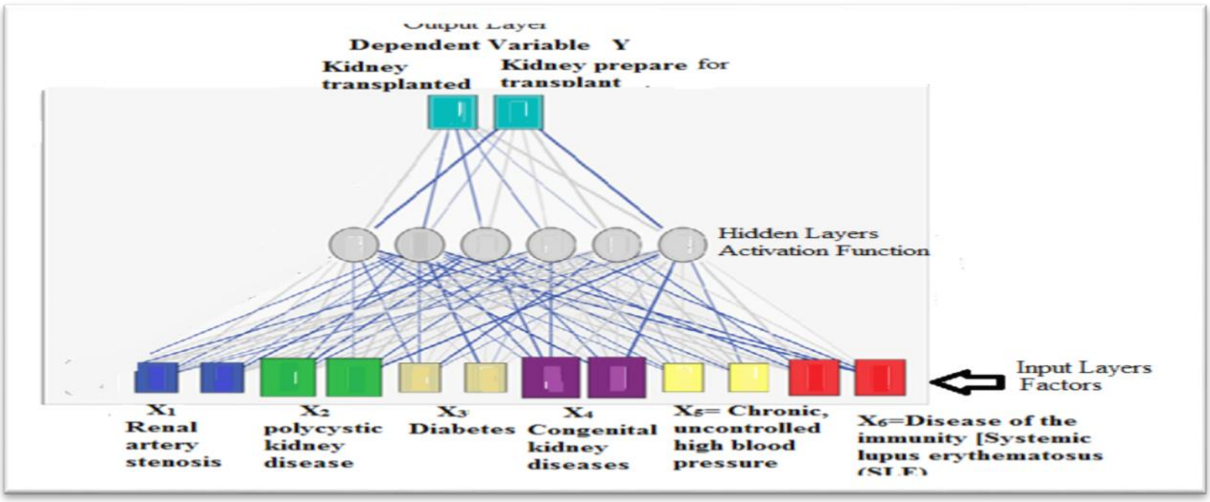
Table(4):Case Processing Summary

		Numbers	Percent
Sample	Training	115	76.7%
	Testing	35	23.3%
Excluded		0	
Total		150	100.0%

Table (4) gives information about the datasets used to build the RBF neural network model The Case Processing Summary shows that the training sample has 115 cases it equal to 76.7% was used it ensure to 76.7% is larger portion designed the model by the effect of risk factors on Kidney Transplantation The testing sample has 35 case it equal to 23.3% of sample used a smaller part of the dataset, used to estimate how well the model generalizes to new hidden cases. It means it's sufficient portion was used for testing to a generalization. But excluded has zero because no data records, this means the full dataset was used.

2-Network Information figure(4) The neural network is structured to classify kidney transplant status based on 6 clinical input variables risk factors considered by the model to predict the kidney transplant, It has one hidden layer with 12 neurons. The output layer uses Identity activation with 2 units, reflecting binary classification output (e.g., transplanted vs. prepare for transplant).

Figure (4): Network Information Diagram

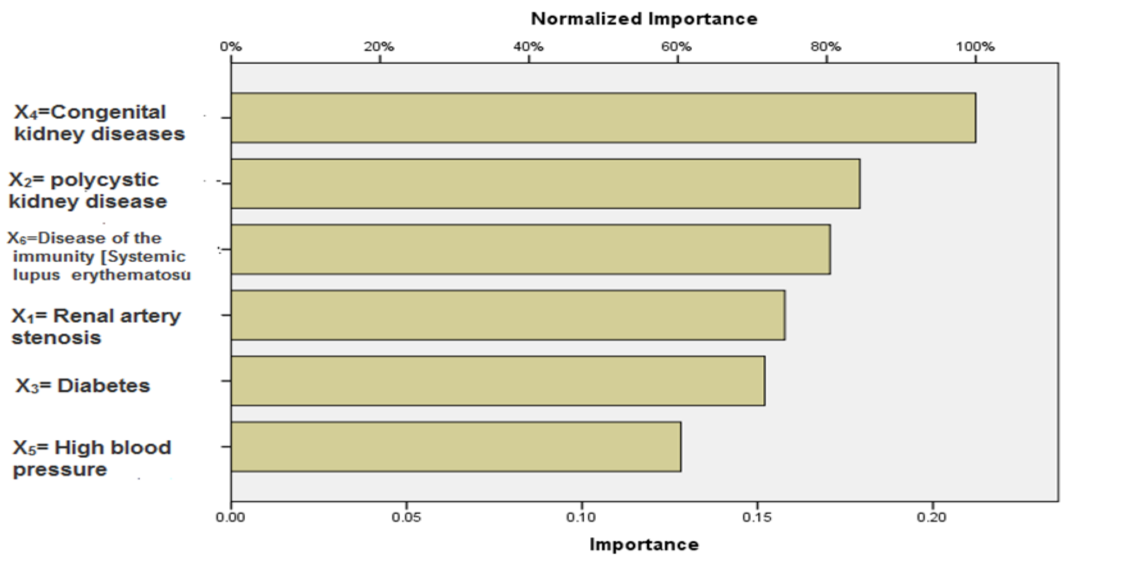


3–The Normalized Important Independent Variables

Table (5): The relative importance of each factor affecting Kidney transplant

Independent Variables	Importance	Normalized Importance
Renal Artery Stenosis	0.158	74.4%
Polycystic Kidney Disease	0.179	84.4%
Diabetes	0.152	71.7%
Congenital Kidney Diseases	0.212	100.0%
Chronic, Uncontrolled High Blood Pressure	0.128	60.4%
Systemic Lupus Erythematosus	0.171	80.4%

Figure (5): Feature Normalized importance of each factor affecting Kidney transplanted



From table (5) and the figure (5) and the relative normalized importance of each independent variables(risk factor affecting on the dependent variable Kidney transplant by using the RBF model, The model identifies "Congenital kidney diseases" (X_4) as the most significant factor, with a normalized importance of 100%, the highest relative effect on the dependent Kidney transplanted variable it's more than (polycystic kidney disease)which has 84.4%,then(Systemic lupus erythematosus) with 80.4%, then (Renal artery stenosis) 74.4%, Diabetes71.7% , (Chronic, uncontrolled high blood pressure) 60.4%. So, these values reflect how much each variable contributes to predicting the Kidney transplant in the RBF model

4- The Performance of Diagnostic Tests

a –Classification(Confusion matrix)

Table (6): Classification

	Observed	**Predicted for the dependent variable: Y**		
sample		**Kidney transplanted**	**Kidney prepared for transplant**	**Percent Correct**
Training	**Kidney transplanted**	67	2	97.1%
	Kidney prepare for transplant	9	37	80.4%
	Overall Percent	66.1%	33.9%	90.4%
Testing	**Kidney transplanted**	19	0	100.0%
	Kidney prepare for transplant	5	11	68.8%
	Overall Percent	68.6%	31.4%	85.7%

Table(6) displays a classification table or. confusion matrix for dependent variable Kidney transplant defined if the predicted probability is greater than 0.5 of the ratio of the correct prediction In the Training Sample the Kidney transplanted and Kidney prepared for transplant the results in the diagonal are (67 , 37 with overall accuracy: 90.4%, which is very strong for a medical prediction task and in the Testing sample Kidney transplanted and Kidney prepared for transplant the results in the diagonal are (19 , 11 with overall Percent 85.7%) still very acceptable it predict that the both kinds of patients affected by the risk factors. In general High overall accuracy (90.4% and 85.7%) suggests the model is reliable and clinically valuable.

$$F_1\_score = 2\left[\frac{precision * recall}{precision + recall}\right] = \frac{2TP}{2TP+FP+FN}$$

$$F_1\_training = \frac{2(67)}{2(67) + 9 \quad + 2} = \frac{134}{145} = 0.9241$$

$$F_1\_testing = \frac{2(19)}{2(19) + 5 \quad + 0} = \frac{38}{43} = 0.8837$$

So in both analyses the values are (0.9241 and 0.8837) it refer to perfect precision and to best predictive..

b -ROC curve (Receiver Operating Characteristics) :

Table (7): Area Under the Curve(AUC)

Dependent Variable		Area Under the Curve(AUC)
Y	Kidney transplanted	0.835
	Kidney prepare for transplant	0.835

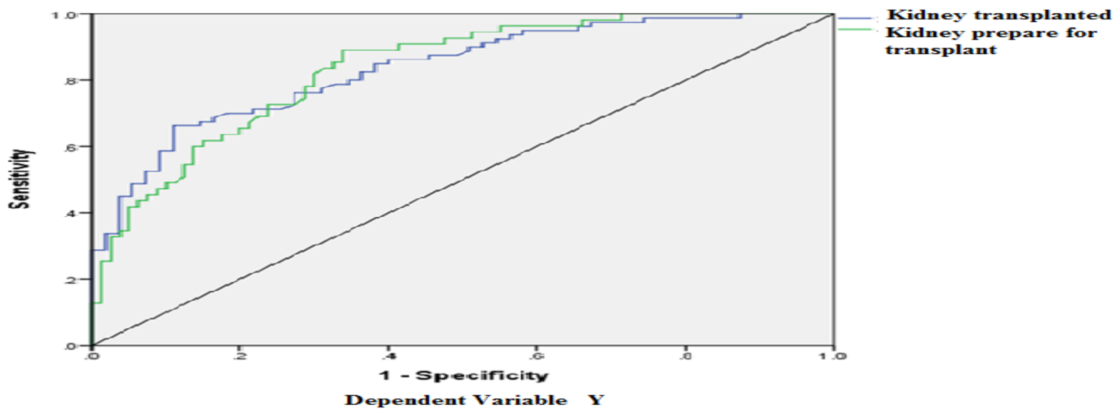


Figure (6) ROC curve

In both Table (7) and Figure (6), the area under the ROC curve (AUC) for both outcomes is 0.835, which is classified as "best" since an AUC in the range of 0.8 to 0.9 is deemed excellent, signifying that the model possesses a robust capacity to accurately classify whether a kidney is suitable for transplantation or is prepared for transplant.

5- Model Selection

a-Model Fitting Information

Table (8): Model Fitting Information

Model	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	Log Likelihood	Chi-Square	df	Sig.
Intercept	167.097	171.312	127.666			
Final Model	162.836	196.553	99.090	28.576	6	.000

This table presents the validation plan and appropriate Model Fitting Information, indicating that the chi-square test for AIC, BIC, and the likelihood ratio yields a P-value of 0.000, which is less than 0.05. This encapsulates model fit statistics, demonstrating the final model's improvement over the simpler model and assessing its statistical significance

.

b-Model Fitting Criteria

Table (9): Model Fitting Criteria Likelihood Ratio Tests of Reduced Model

Likelihood Ratio Tests						
Effect	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept	119.695	149.197	99.090	.000	0	.
Renal Artery Stenosis	125.803	151.091	108.579	9.489	1	0.032
Polycystic Kidney Disease	121.929	147.217	106.509	7.419	1	0.041
Diabetes	122.803	146.091	104.792	6.702	1	0. 046
Congenital Kidney Diseases	123.929	149.217	111.891	12.801	1	0.000
Chronic, Uncontrolled High Blood Pressure	124.803	150.091	104.168	5.078	1	0.049
Systemic Lupus Erythematosus	120.295	144.583	105.924	5.834	1	0.043

The chi-square statistic represents the disparity in log-likelihoods between the final model and a constrained model. The reduced model is created by excluding an effect from the final model. The null hypothesis posits that all effect parameters are equal to zero. The reduced model is equivalent to the final model, as the exclusion of the effect does not augment the degrees of freedom. As the p-values for all predictor variables (six risk factors) are >0.05, it demonstrates that the specified clinical variables significantly enhance the model fit compared to the intercept-only model. The statistical tests validate that these elements are crucial predictors in the model.

The bellows tables of comparison between RBF and logistic regression for showing which model is better

Table (10): comparison of Classification between RBF and logistic regression

Model	Observed	Predicted for the dependent variable: Y		
		Kidney transplanted	Kidney prepared for transplant	Percent Correct
Logistic regression	Kidney transplanted	73	15	83.0%
	Kidney prepare for transplant	15	47	75.8%
	Overall Percent	80.0 %		
RBF	Kidney transplanted	19	0	100.0%
	Kidney prepare for transplant	5	11	68.8%
	Overall Percent	85.7%		

the RBF model performs better overall in predicting 85.7% the correct status of kidney transplant patients, especially with perfect accuracy for the "Kidney transplanted" group, while logistic regression has slightly lower accuracy but more balanced performance across both groups. This suggests RBF may be more effective for this classification task.

Table (11): comparison of Model Fitting Criteria Likelihood Ratio Tests between RBF and logistic regression

Likelihood Ratio Tests				
Effect	Logistic regression		RBF	
	Chi-Square	Sig.	Chi-Square	Sig.

Renal Artery Stenosis	1.846	.174	9.489	0.032
Polycystic Kidney Disease	9.045	.003	7.419	0.041
Diabetes	2.253	.133	6.702	0. 046
Congenital Kidney Diseases	14.749	.000	12.801	0.000
Chronic, Uncontrolled High Blood Pressure	4.401	.036	5.078	0.049
Systemic Lupus Erythematosus	.946	.331	5.834	0.043

The Radial Basis Function (RBF) model better describes the influence of the risk factors on the outcome in this table RBF model identifies more statistically significant relationships (with lower significance values) for key risk factors compared to the logistic regression model, indicating greater sensitivity and capacity to capture complex, nonlinear patterns in the data. Additionally, the higher Chi-Square values in the RBF model for many factors suggest stronger effects being detected. Therefore, for this dataset, RBF provides a more comprehensive of the risk factors' impact.

6.Conclusion

this study confirms that the Radial Basis Function (RBF) classifier, coupled with normalized importance of risk factors, effectively identifies critical predictors influencing kidney transplant outcomes. The model demonstrates strong predictive performance with high accuracy and a robust AUC, emphasizing its capability to distinguish between transplanted and prepared-for-transplant kidneys. Key risk factors such as Congenital Kidney Diseases, Polycystic Kidney Disease, and Systemic Lupus Erythematosus were validated as statistically significant predictors, enhancing the reliability of transplant risk forecasting. The significant improvement in model fit over baseline confirms the value of integrating these normalized risk factors with RBF classification. Overall, this approach provides important insights to support clinical decision-making and targeted interventions, highlighting the advantages of advanced machine learning techniques in kidney transplant risk assessment ,RBF model better describes the influence of the risk factors on the outcome from a traditional methods.

Recommendations:

Recommendations:

1. Integrate the RBF-based prediction model incorporating normalized variable significance into clinical workflows to enhance pre-transplant risk assessment and optimize post-transplant patient monitoring.
2. Utilize key identified risk factors, such as congenital kidney disorders, polycystic kidney disease, and diabetes, to tailor individualized treatment plans and follow-up protocols for transplant patients.
3. Explore combining the predictive model with other clinical tools and biomarkers to develop a comprehensive kidney transplant risk management system.
4. Foster interdisciplinary collaboration among clinicians, data scientists, and researchers to continuously improve and update predictive models as new clinical data emerges.
5. Align this approach with advances in machine learning applications for kidney transplantation, which have demonstrated improved graft survival predictions and support targeted therapeutic interventions, leading to better long-term patient outcomes.
6. Conduct broader multi-center studies to validate and generalize findings beyond single-hospital data while incorporating survival analysis techniques for more dynamic and long-term risk predictions.
7. Emphasize early screening and proactive management by healthcare providers and nephrologists to enhance patient outcomes and investigate additional factors influencing transplant success.

References

[1] Al Barznji, N.S.K., Abdullah, D.F., Kareem, R. & Sedeeq, B.S., 2024. Assessing by model fitting criteria and comparing between Stirling and Aitken interpolation formulas to determine the preferred treatment of choice (surgery or surgery with chemotherapy) for lung cancer. Zanko Journal of Pure and Applied Sciences, 28(3), pp.266–281. https://doi.org/10.21271.

[2] Al Barznji, N.S.K., 2018. Using logistic regression analysis and linear discrimination analysis to identify the risk factors of diabetes. Zanko Journal, 22(6), pp.248–268.
[3] Amin, M.Z. & Ali, A., 2017. Application of Multilayer Perceptron (MLP) for Data Mining in Healthcare Operations. Proceedings of the 3rd International Conference on Biotechnology, University of Engineering and Technology, Lahore.
[4] Broom head, D.S. & Lowe, D., 1988. Multivariable functional interpolation and adaptive networks. Complex Systems, 2, pp.321–355.
[5] Broomhead, D.S. & Lowe, D., 1988. Radial basis functions, multi-variable functional interpolation and adaptive networks. RSRE Technical Report 4148. Archived from original 2013.
[6] Buhmann, M.D., 2003. Radial Basis Functions: Theory and Implementations. Cambridge University Press.
[7] Chang, Y.-W., Hsieh, C.-J., Chang, K.-W., Ringgaard, M. & Lin, C.-J., 2010. Training and Testing Low-degree Polynomial Data Mappings via Linear SVM. Journal of Machine Learning Research, 11, pp.1471–1490.
[8] Chen, D., 2021. Evaluation Model of Physical Education Effect: On the Application of Radial Basis Function Particle Swarm Optimization Neural Network (RBFNN-PSO). Computational Intelligence and Neuroscience, 2021, Article ID 5570685.
[9] Collins, N.A., 2005. Final Destination: Looks Could Kill. Black Flame. ISBN 1844163164.
[10] Dash, C.S.K., Behera, A.K., Dehuri, S. & Cho, S.B., 2016. Radial basis function neural networks: a topical state-of-the-art survey. Open Computer Science, 6(1), pp.33–63.
[11] Ding, J., Tarokh, V. & Yang, Y., 2018. Model Selection Techniques: An Overview. IEEE Signal Processing Magazine, 35(6), pp.16–34. https://doi.org/10.1109/MSP.2018.2867638.
[12] Dwivedi, Y.K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., et al., 2020. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, p.101994.
[13] Fernández-Navarro, F., Hervás-Martínez, C. & Martínez-Álvarez, F., 2012. Parameter estimation of q-Gaussian Radial Basis Functions Neural Networks using a hybrid evolutionary algorithm. Neurocomputing, 75(1), pp.123–134.

[14] Glover, S. & Dixon, P., 2004. Likelihood ratios: A simple and flexible statistic for empirical psychologists. Psychonomic Bulletin & Review, 11(5), pp.791–806. https://doi.org/10.3758/bf03196706.

[15] Grafals, M., Chandraker, A., et al., 2019. Noninfectious complications after kidney transplantation. In: Chronic Kidney Disease, Dialysis, and Transplantation, 4th ed., pp.1–12.

[16] Hossain, M.I., Maruf, M.H., Khan, M.A.R., Siddiqi Prity, F., Fatema, S., Ejaz, S. & Khan, M.A., 2023. Heart disease prediction using distinct artificial intelligence techniques: performance analysis and comparison. Iran Journal of Computer Science.

[17] Kirişci, M., 2019. Fibonacci statistical convergence on intuitionistic fuzzy normed spaces. Journal of Intelligent & Fuzzy Systems, 36(6), pp.5597–5604.

[18]. Montesinos López, O.A., Montesinos López, A. & Crossa, J., 2022. Multivariate Statistical Machine Learning Methods for Genomic Prediction. Springer International Publishing, Cham, pp.379–425. https://doi.org/10.1007/978-3-030-89010-0_10

[19]. Murad, S.H., Awlla, A.H. & Moahmmed, B.T., 2023. Prediction Lung Cancer based critical factors using machine learning. Scientific Journal of University of Zakho, 11(3), pp.447–452.

[20] Nguyen, V.T., Nguyen, Q.C., Van, M., Su, S.F., Garg, H., Duong, D.N. & Tan, P.X., 2025. Robust adaptive nonlinear PID controller using radial basis function neural networks for ballbots under external force. Engineering Science and Technology, an International Journal, 61, p.101914

[21] Rasmussen, C.E. & Williams, C.K.I., 2006. Gaussian processes for machine learning. Cambridge, MA: MIT Press.

[22] Samuel, A.L., 1959. Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development, 3(3), pp.210-229.

[23] Schwenker, F., Kestler, H.A. & Palm, G., 2001. Three learning phases for radial-basis-function networks. Neural Networks, 14(4–5), pp.439–458. https://doi.org/10.1016/s0893-6080(01)00027-2.

[24] Sejnowski, T.J., 2018. The Deep Learning Revolution. MIT Press.

[25] Stoica, P. & Selen, Y., 2004. Model-order selection: A review of information criterion rules. IEEE Signal Processing Magazine, 21(4), pp.36–47. https://doi.org/10.1109/MSP.2004.1311138.

[26] Syu, S.H. et al., 2019. Risk factors for complications and graft failure in kidney transplant patients with sepsis. Bosnian Journal of Basic Medical Sciences, 19(3), pp.304–309.

[27] Wu, Y., Wang, H., Zhang, B. & Du, K.L., 2012. Using radial basis function networks for function approximation and classification. International Scholarly Research Notices, Article ID 324194.

[28] Yadu, A.K. & Shrivastava, G., 2021. Study of Applications of Radial Basis Function Network in Forecasting. Advances in Materials and Lightweight Technologies. https://doi.org/10.5185/amlett.2021.051628

[29]. Zhao, A., Xing, S., Wang, X. & Sun, J.Q., 2024. Radial basis function neural networks for optimal control with model reduction and transfer learning. Engineering Applications of Artificial Intelligence, 136, p.108899.

Appendix

Table (1): Basic functions in Radial Basis Function (RBF)[5]

Basic function abbreviation	Formulas $\phi(r) = \phi(\|x - \mu\|)/\sigma$
Gaussian(GA)	e^{-cr^2}
Generalized Multiquadratic (GMQ)	$(C^2 + r^2)^\beta$, $c > 0$,$1 > \beta > 0$
Inverse Multiquadratic (IMQ)	$1/\sqrt{C^2 + r^2}$
Inverse Quadratics(IQ)	$(C^2 + r^2)^{-1}$
Multi Quadratics(MQ)	$\sqrt{C^2 + r^2}$
Hyperbolic secant (sech)	*Sech(cr)*
Cubic function	r^3
Linear function	r
Monomial(MN)	r^{2k-1}
Thin-plate spline	$r^2 \, log(r)$
Poly harmonic spline	$\begin{cases} \phi(r) = r^k , k = 1,3,5, \dots \\ \phi(r) = r^k \, ln(r) \, , k = 2,4,6, \dots \end{cases}$