

# ANNMDD: Strength of Artificial Neural Network Types for Medical Diagnosis Domain

Ahmed Hamza Osman

Department of Information System, Faculty of Computing and Information Technology  
King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

**Abstract**—The abundance of medical evidence in health institutions necessitates the creation of effective data collection methods for extracting valuable information. For several years, scholars focused on the use of computational techniques and data processing techniques in order to enhance the study of broad historical datasets. There is a deficiency to investigate the collected data of health disease in the data sources such as COVID-19, Chronic Kidney, Epileptic Seizure, Parkinson, Hard diseases, Hepatitis, Breast Cancer and Diabetes, where millions of people are killed in the world by these diseases. This research aims to investigate the neural network algorithms for different types of medical diseases in order to select the best type of neural network suitable for each disease. The data mining process has been applied to investigate the mentioned medical disease datasets. The related works and literature review of machine learning in the medical domain were studied in the initial stage of this research. Then, the experiments behind the initial stage have been designed with six neural network algorithm styles which are Multiple, Radial Based Function Network (RBFN), Dynamic, Quick and Prune algorithms. The extracted results for each algorithm have been analyzed and compared with each other to select the perfect neural network algorithm for each disease. T-test statistical significance test has been applied as one of the investigation strategies for the NN optimal selection. Our findings highlighted the strong side of the Multiple NN algorithm in terms of training and testing phases in the medical domain.

**Keywords**—Medical data; neural network algorithm; multiple; radial based function network; dynamic; quick; prune; accuracy

## I. INTRODUCTION

Currently the number of records in health databases is so huge thanks to technology which has made it possible to securely and effectively store and retrieve this large volume of data. The process of creating meaningful patterns or evidence from medical data sites is medical diagnosis [1]. The extract from these medical datasets helps the physician to diagnose disease in the early stages. The problem is largely solved by having adequate tools for dealing with such large data. A huge amount of research in this area has been carried out and it is still a very interesting area. There are numerous classification algorithms available, and it is worthwhile to do a more in-depth examination of these algorithms and their success on medical datasets such as [1-7]. We perform studies in this paper on a number of medical datasets using a variety of familiar prediction algorithms. The objective is to evaluate if preprocessing techniques prior to classification can improve classification performance. Materials containing the missing values, contours and noise are well known, and few papers

analyze the effect of pre-processing to the best knowledge of the author [3].

For future documentation and for preparing future procedures, most healthcare organizations and research institutes digitally store patient records. Because of the difficulty and size of information, the noise and lacking of values, mining is a tedious task and it is very difficult to analyze in this heterogeneous medical dataset [8]. Most healthcare organizations and medical research institutions store their patients' data digitally for future references and future treatment planning [8]. Healthcare is a field which is closely linked to the daily life of all owing to the high uncertainty [9]. Machine learning (ML) is a powerful and flexible tool for analyzing and predicting biological outcomes and clinical data[10]. Early diagnoses benefit from the detection of useful trends in the medical dataset[11].

Machine Learning (ML) and Artificial intelligence (AI) and master learning are key terms for a range of algorithms, allowing computers to detect and determine data patterns. Despite a quiet time, AI's abilities are an omnipresent part of mainstream culture in a number of activities, from automated digital assistants to self-driving vehicles. Given some positive advances in cardiovascular oncology, however, a major AI revolution has not yet occurred. Cardiovascular routine patient care accumulates large quantities of electronic health record (EHR) data. Integrating a large amount of diverse data in a busy clinical environment is a challenge, leading to marked underuse of data that could influence clinical decisions. Artificial intelligence (AI) applies in general to computational algorithms, such that machines can gain secret information without being programmed directly[12]. Machine learning techniques are attracting substantial interest from medical researchers and clinicians [13]. Data mining is the discovery process of patterns and trends from the enormous amount of data and examines the various existing data mining techniques [14]. Several studies helped to find the most appropriate neural network algorithms for the classification of medical data and also showed the value of pre-processing in improving classification performance. Therefore, this study has investigated the effectiveness and efficiency of the neural network algorithms (NN) in healthcare in order to select the significant NN method that helps the medical doctor to make a diagnosis decision for a specific disease. Most healthcare organizations and medical research institutions store their patients' data digitally for future references and future treatment planning [8]. In addition to being used by clinicians, decision support systems are also helpful in making medical

decisions; they're built on two main types, and the better ones assist with precise, consistent and prompt responses [15].

As machine training (ML) is commonly accepted as a technique to choose from in the common diseases pattern classification and predictive modelling, due to its specific advantages in critical features detecting in some diseases, the problems of this study can be summarized as:

- 1) Need for a clear method to identify the disease into disease stages, pattern or status. The number of patients who are infected is growing in specific disease.
- 2) The urgent need for an intelligent system that helps in the process of diagnosing diseases for themselves.

The goal of this research is to see the most useful disease features in predicting status and general patterns which can help us select models and select hyper parameters. The goal of this study is to identify the efficiency and effectiveness of some machine learning algorithms such as Neural Network algorithms in the healthcare domain. Our main objective is to investigate the neural network algorithms to adapt to a function that can predict the discreet new input class [16]. In particular, this study investigates and reviews the current healthcare systems in order to detect the shortcomings that need to be improved. The significant role in this study is to find and predict an early disease stage that can save patients' lives using neural network classification methods in order to help medical doctors to take suitable diagnosis decisions. Diagnosis at an early stage, when it isn't too large and hasn't spread, makes successful treatment more likely.

This study comprises six sections. The introduction to this study is explained in Section 1. The Section 2 discusses related works. The upgraded methodology and its fundamental method are discussed in Section 3. Section 4 will discuss data classification using neural networks. Section 5 will illustrate the experiments and results of the investigation. Finally, Section 6 will summarize the study's findings.

## II. RELATED WORK

For future documentation and for preparing future procedures, most healthcare organisations and research institutes digitally store patient records. Due to the complexity and volume of data and the noise and lack of values and noise, this mining is a tedious task, making it very difficult to analyse this heterogeneous medical dataset [8]. Most healthcare organisations and medical research institutions store their patients' data digitally for future reference and future treatment planning [8]. Healthcare is a field which is closely linked to the daily life of all owing to the high uncertainty [9]. Machine learning (ML) is a powerful and flexible tool for analysing and predicting biological outcomes and clinical data [10]. Early diagnoses benefit from the detection of useful trends in the medical dataset [11]. It's almost old news by now: big data is going to transform medicine. However, it is essential to remember that the data is useless by itself. Data must be analysed, interpreted and acted upon to be useful. Thus, it is algorithms that will prove transformative, not datasets. Obermeyer. and Emanuel (2016) therefore believed that attention needs to be shifted to new statistical tools in the field of machine learning, which are

critical for everyone in the 21st century who practises medicine [17]. Shameer, et al. discussed existing approaches that use machine learning and bioinformatics for behavioural analysis as well as well as clinical, genetic and climatic methodologies in their talks [18]. Artificial intelligence (AI) applies in general to computational algorithms, such that machines can gain secret information without being programmed directly [12]. The computational study of machine learning is drawing significant attention from academics and clinicians alike [13]. The best of these schemes assist doctors in making decisions while others play a crucial part in making them [15]. Machine learning is the discovery procedure of patterns and trends from the enormous amount of data This study examined the various existing data mining techniques [14]. This study helps to find the most appropriate neural network algorithms for the classification of medicinal data and also will show the value of pre-processing in enhancing prediction results.

Much of the computer-based rule sets that deal with healthcare situations are "expert systems". Computer simulations operate in the same manner as ideal medical students: they apply theories to new cases. In the other words, machine learning methods go through training programs. The algorithms start with the patient stage, working their way through large quantities of variables in the process of compiling information before they come to a conclusion that can be applied to a variety of different patients.

Standard regression models, such as the result, the covariates, and statistical functions, may be seen to be in one way an extension of this phase Nonlinear, and immersive forms will work with a large number of predictors. When statistics are not feasible, the integration and interpretation of nuanced biomedical and healthcare data using artificial intelligence (CITL), supervised machine learning, deep learning, and cognitive approaches may be employed. H. Shameer, et al. (2018) was one of the research groups that explored the fundamentals and the uses of master learning algorithms; and they investigated the possible shortcomings and obstacles that machine learning could present in cardiovascular care[12]. Computer-based artificial intelligence and artificial-learning algorithms, respectively, are not all that novel in the medical field. In medical practice, risks are typically found in databases and are used to stratify patients or to provide anticoagulation guidance on which drugs to use [19]. This process entails analysing medical datasets in order to uncover intriguing trends in decision-making [14]. Machine learning algorithms have the potential to significantly increase the quality of healthcare in a variety of ways. Prognosis modelling algorithms assist health authorities in allocating money efficiently and physicians in selecting the right care choices for patients [19]. The multilayer perceptron of neural networks (MLPNN), logistic regression (LR) and validation is applied to test the predictive models [20].

Evaluation of machine learning algorithms depends on their accuracy, specificity and error rate [8]. Some used well-known regression equations, such as the simple linear and logistic, widely seen in clinical and modern statistical models such as Bayesian analysis, such as the data [18]. Analytic forecasting techniques include decision trees (J48) and

Bayesian analyses. The three mass classification methods used machine learning approaches to build three statistical models for the diagnosis of breast cancer. A standardised General Linear Model (GLM), a standard Support Vector Machine (SVM) with radial basis function, and single-layer neural networks were included. Predictive models were trained on the sample prior to the validation dataset. They used the validation datasets to compute the precision, sensitivity, and specificity of the three models – a decision aid for detecting breast cancer based on three kinds of decision tree factors. Simple decision tree (SDT) and Boosted simple decision tree (BDT). Many analyses were performed on several major classifiers including C4.5 (J48), Naïve Bayes, SMO, and Random Forests and compared with the results of these [11]. The Decision Tree C4.5 is unpruned. C4.5 is an increase in the previous ID3 algorithm of Quinlan. J48 creates decision trees using the idea of information entropy from the collection of labelled training data. J48 explores the uniform data benefit that results in choosing a data splitting attribute. The highest information gain attribute is used to make the decision. Then the little subsets use the algorithm. If all occurrences in a subset have the same class, the dividing strategy stops. A leaf node for this class is then made in the decision tree. The algorithm of Naive Bayes depends on conditions. It is using Bayes Theorem, an equation which calculates a probability by inspecting the recurrence of historical data values and values mixes. Given the probability of another opportunity that already occurred, Bayes' theorem finds the possibility of an occasion [21]. The document 'Prediction of diabetes use of the Bayesian Networks,' suggested by Mukesh Kumari et al. [22]. This article proposed to predict people, whether diabetic, not diabetic or pre-diabetic, by classifying the Bayesian Network. The dataset used is collected from a hospital that collects information from people with and without diabetes. Weka is the tool used for the study and exam. On the dataset of persons collected from the hospital, classification algorithm is applied and the results have been obtained. The author analysed the attributes' values to determine whether a certain individual is diabetic, non-diabetic or pre-diabetic in a dataset. The fact that a man is diabetic, non-diabetic or pre-dentinal had led to the determination of attributes such as qfast Gtt, casual Gott and diastolic blood pressure values above a given amount. The author concluded that 99.51 was best accurate in the classification with the Bayesian system. The paper 'Improved J48 Classification Algorithm for Diabetes Prediction' was proposed by Gaganjot Kaur et al. [23]. This work manages successful data mining to predict diabetes in patients' medical records. Today, diabetes in all populations and all ages is an extremely regular infection increasing the risk of developing renal disease, nervous damage, damage to the venous tract and visual impairment is a result of coronary disease. This paper uses the Pima Indians Diabetes Data Set which collects data from and without diabetes patients. Using the modified J48 classifier, the data mining method accuracy rate is determined. The WEKA data mining tool was used for manufacturing the J48 graders as a MATLAB API. The results of the test showed that the calculation J-48 is considerably different. Precision up to 99.87 percent have been demonstrated in the proposed calculation. V. Karthikeyan et al. [24]. Suggested the paper titled 'Data Mining Algorithm in Diabetes Disease

Prediction Comparative Mining Classification (CDMCA)'. The data mining is an iterative development that is defined by discovery, by normal or manual techniques. Two types of supervised and uncontrolled classifications are classed in this paper, which uses the CDMCA data mining concept. This is the classification of the supervised diabetes-based data mining algorithms. It contains at least the plasma glucose diseases. This research describes C4.5, SVM, KNN, PNN, BLR, MLR, CRT, CS-CRT, PLS-DA as well as PLSLDA algorithmic discussion. The paper compares computer time performance, accuracy and data evaluated by means of a 10-fold Cross Validation error rate, and focuses on True Positive, True Negative, False Positive and False Negative and Accuracy. This shows that CS-CRT is the best algorithm. For this experiment, different data mining techniques are applied with the Pima Indian Diabetes Dataset. Pre-processing techniques convert raw data into useful and understandable formats to enable more precise results during algorithm execution. The features are extracted using permutation techniques from the pre-processed data and the classification techniques are performed in different combinations of features. The results achieved are evaluated for every combination [11].

Automatic risk prediction algorithms to guide clinical treatment; using unsupervised training techniques to more accurately phenotype complex conditions; and implementing algorithms to enhance the education of providers of healthcare intelligently [12]. Efficient classification of the medical dataset is then and now a big problem in data mining. Diagnosis, disease prediction and outcome accuracy can be enhanced if the relationships and trends are effectively extracted from these complex medical datasets [8]. The most serious scourge affecting the industrialised nations is CVD disease. Not only does CVD affect a large proportion of the population without warning but it also causes chronic suffering and disability in an even greater number [25]. The objective of this research was to collect the breast cancer data set for medical decision-making using clustering and data mining techniques [14]. Francis, F. and J. Saleema (2017) sought optimal features by combining the permutation input data attributes to improve the accuracy of the classifier [11]. These algorithms' performance assessments are based on precision, sensitivity, specificity and error rate. Heart statistics are the medical information used in this study[8]. BNs have been effective in developing powerful algorithms which can manage very large datasets and create predictive models of high quality from medical data and genomic [18]. A directed acyclic graph is a graph in which each node represents a variable and each arc represents a connection. Each arc is interpreted by BNs as a direct impact on a child node (variable) by a parent node (variable) [18]. The SVM, MLPANN, and LR models had accuracy, area under the curve (AUC), sensitivity, and specificity of 90.4 percent, 86.5 percent, 98.2 percent, and 49.6 percent, 85.9 percent, 76.9 percent, 97.3 percent, and 26.1 percent, and 84.7 percent, 77.4 percent, 97.5 percent, and 17.4 percent, respectively. Meanwhile, the independent predictors were discharge time creatinine, recipient age, donor age, donor blood type, etiology of ESRD, and post-transplant recipient hypertension [20]. The trained algorithms were capable of classifying cell nuclei with a high degree of accuracy (0.94 – 0.96), sensitivity (0.97 –

0.99), and specificity (0.97 – 0.99). (0.85 – 0.94). The SVM method produced the highest accuracy (0.96) and area under the curve (0.97) values. When algorithms were organised in a voting ensemble, prediction performance improved somewhat (accuracy=0.97, sensitivity=0.99, specificity=0.95) [13]. The results indicated that SDT and BDT obtained overall accuracies of 97.07 percent with 429 accurate classifications and 98.83 percent with 437 right classifications, respectively, during the training phase. BDT outperformed SDT on all performance indicators. The receiver operating characteristic (ROC) and Matthews correlation coefficient (MCC) values for BDT in the training phase were 0.99971 and 0.9746, respectively, which were superior to those of the SDT classifier. During the validation phase, DTF attained a classification accuracy of 97.51 percent, outperforming SDT (95.75 percent) and BDT (97.07 percent) classifiers. For DTF, the ROC and MCC values were 0.99382 and 0.9462, respectively [15]. That is the promise of medical machine learning: the wisdom contained in decisions made by almost all clinicians and the results of billions of patients should inform each patient's care. That is, every diagnosis, management decision and therapy should be customised based on all known patient information [26]. With machine learning located at the peak of inflated expectations, they considered softened as subsequent crash into a "trough of disillusionment" by encouraging a greater appreciation of the capabilities and limitations of the technology before they counter an idealised and unrealisable standard of perfection with computerised systems (or humans) [19]. Firstly, machine learning can improve the prognosis dramatically. Second, much of the work of radiologists and anatomical pathologists would be replaced by machine learning. These doctors mainly focus on the interpretation of digitised images. Firstly, machine learning can improve the prognosis dramatically. Second, much of the work of radiologists and anatomical pathologists would be replaced by machine learning. These doctors mainly focus on the interpretation of digitised images [17]. This would reduce the burden on doctors, increase and speed up access to care, reduce resources and cut costs for patients [10].

In the 18th century [27] Sir Galton introduced the first linear regression. Linear regression is a statistical method for modelling the connection between a variable dependent and one or more explicative variables. It assumes that weighted amounts of input variables can be predicted. Normally this is the very first model you would analyse when the outcome variable remains constant before moving into more complex models. Reed et al. [28] said that the association between highly accessible electronic health records (EHR) and ED visits, hospitalisations and office visits for diabetes mellitus patients has been researched. They used a linear regression model with patient-level effects and found that the use of EHR was associated with a small decline in ED and hospitalisation, but not with office visit rate among patients with diabetes. Yaffe et al. [29] have taken an account of the non-independence of the proportions as time series in the annual controls from the Kaiser Permanente Northern California hypertension registry, configuring a log linear proportion regression on time to make auto-correlated errors possible. They found that the application of a large-scale hypertension program was associated with a major increase in hypertension

control in comparison with national and state control rates among adults with hypertension. In patients treated with target agents, Yuasa et al. [30] investigated the correlations between original tumour size and the tamper reduction rate. They used linear regression analyses both univariable and multivariate to determine that only the initial tumour size was associated with the rate of tumour reduction. This could be beneficial for doctors who treat patients' metastatic renal cell carcinoma.

In certain ways the logistic regression is similar to the linear regression. Regression of logistics assumes that the result can be explained by a weighted amount undergoing a specific mathematical transformation named log it. This transformation allows for the mapping of all weighted sums into a value between 0 and 1, which can be interpreted as a chance of a binary result. Therefore, logistic regression is widely used in the result variable with two outcomes, for example, whether or not you have a disease. From Vries et al. [31] the relation between mortality and iatrogenic diseases occurring outside the operating room has been investigated. The investigators carried out a multidisciplinary safety examination list with the medication, surgical side and medication checked by six hospitals. The relationship between checklist and mortality was evaluated by logistic regression. The research demonstrated a link between the full checklist and a decrease in surgical complication and mortality and high-grade hospitals. The relationship between maternal risk factors and congenital urinary tract anomalies was examined by Shmorhavorian et al. [32]. The study was carried out in case-control. In cases in which children were diagnosed with urinary anomalies while controls did not display urinary tract anomalies, they received Washington state birth-hospital discharge records from 1987-2007. The analysis identified increased risk of renal anomalies for gestational diabetes, pre-existing diabetes, and maternal renal disease. The results of these incidents have been studied by Peterson et al. The most common medical conditions were described and the form of assistance given on board. Through means of logistic regression, they established syncope, respiratory symptoms and gastrointestinal symptoms in most medical emergencies in flight. In 1996, Somogyi and Sniegowski [33] first launched Boolean networks. Boolean networks were easily deployed as genetic networks with their convenient representation. But because Boolean networks do not specifically demonstrate the uncertainty the data may have, the vague characteristic of a bio-system cannot be modelled. Also note that no arrows are used when a Boolean network is created; thus, no path or cause of the model is clear.

In mathematics, differential equations have a long history of modelling a biological system [34, 35]. Chen et al. have modelled a simplified dynamic gene control system (with transcription feedback). Differential equations are more likely to model biological processes than boolean networks, but the computing costs of using differential equations are high and many of the parameters are often not available to use differential equation models. Since most of the genetic trajectory dynamics seem to be non-linear, a linear model appears to be working only on the limited genetic trajectories.

The BN model has been used extensively to learn data predictive models. BNs can model causality on the basis of the

knowledge, data or both of the researcher. It is also utilised in a large number of medical fields as it can easily infer [36, 37]. One practical limitation of BNs is the fact that inferences in BNs are virtually impossible with a large number ( $> 50$ ) of modelled variables [37], a frequently used limitation of many reasoning methodologies. A causal BN (or short causal network) is a BN, in which the parent variable and each arrow are interpreted as direct causal influence and the variable explicitly connected to which the variable is called the child variable [38]. Fig. 1 displays the structures of the five variables describing genes of a hypothetical causal BN system. The structure of the causal network in figure. For example, 1 indicates that Gene1 may regulate the level of expression (a cause of influence) of Gene3 which in turn may regulate the level of expression of Gene5. A variable is independent from its non-descendant since its parents occur (i.e., direct causes). It gives rise to conditional independence relationships defined in a causal BN.

### III. SUGGESTED MODEL

This section discusses the methodology that was used for the medical data classification and diagnosis. It was necessary to develop a sound methodology prior to the implementation of this research in order to improve classification of medical data. Research design of this study is a combination of several stages, each stage contains a number of steps. First access to the datasets of people infected with disease form data repertory such as UCI. Then, we define the second phase for six sub-stages of Knowledge data mining discovery (KDD) process, including data selection, data preprocessing and cleaning, data transformation, data mining, evaluation, and interpretation. In the third phase we design the experiments of this study using neural network algorithms. The fourth stage discusses the results and discussion phase. The fifth stage research is results analysis.

Operational guidance framework provides a structured manner and is used to help the researcher to achieve their goals. It is important that the operational framework is organization of a systematic process in this study. The framework has been divided into two phases: operational framework of the main action named planning phase, and the implementation phase. In this study, each of these stages is made up of different stages, starting with a review of the literature in the planning stage and the end of the written report. Fig. 1 below shows the operational framework.

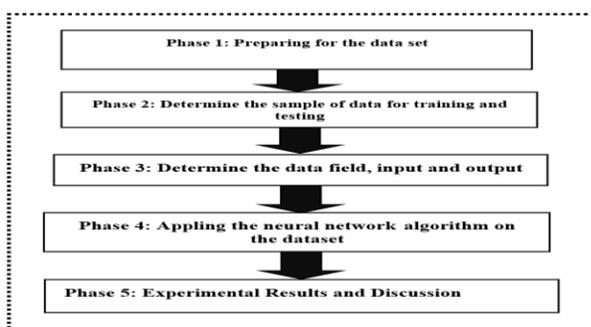


Fig. 1. Suggested Framework.

### IV. MEDICINE DOMAIN CLASSIFICATION USING NEURAL NETWORK

Artificial neural network (ANN) is a paradigm in the treatment of information based on biological nervous system data, such as the brain. The modern framework of the information management system is the core aspect of this model. It consists of many highly interconnected processing elements (neurons) which work together to solve specific issues. ANNs learn from examples, like people. For a specific application an ANN is configured by means of a learning process, like model recognition or data classification, Biology Systems Learning [39]. The goal is to create models of biological neural systems so that biological systems can understand how they work. Neuroscientists strive to link biological processing observed (data), neural theory (theory of statistical learning and theory of information) and biologically plausible neural methods for processing and learning (biological neuron network models) [40].

Neural networks are distinct from those of traditional computers in problem solving. Conventional computers use an algorithmic approach. In order to solve a problem, the machine follows a series of instructions. The computer can't solve the problem unless the specific steps the computer must take are known. This restricts conventional computers' problem solving capability to problems which we understand and are already able to resolve. But if computers could do things we do not know exactly how to do they would be much more useful [41].

ANNs are currently a hot area of medical research, and in the next few years they are expected to be widely used in biomedical systems. The research is currently primarily focused on modelling the human body parts and the recognition of diseases from various scans (for example, cardiograms, CAT scans, ultrasound scans).

Neural networks are suitable for identifying diseases using scans since a complex algorithm on how to recognise the disease does not need to be given. Neural networks learn by example to escape the need to classify the disease. A number of examples are needed which are representative of all the disease variations. The number of examples does not matter as much as the amount. The examples must be carefully selected for reliable and efficient performance of the system [42].

We provide some descriptions of the neural network types:

1) *Quick*: This approach chooses a topology using rules of thumb and features of the data the default number of hidden layers has changed in previous versions of clementine. The new process normally yields thinner layers that are quicker and more adaptable. if you get bad results at the default size, consider increasing the size of the "hidden" layer on the expert page. Fig. 2 shows the quick neural network structure.

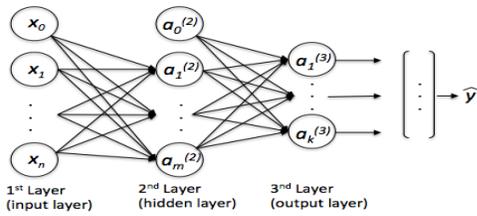


Fig. 2. Quick Neural Network.

2) *Dynamic*: It provides an initial topology, but adjusts it when training is underway. Fig. 3 shows the dynamic neural network structure.

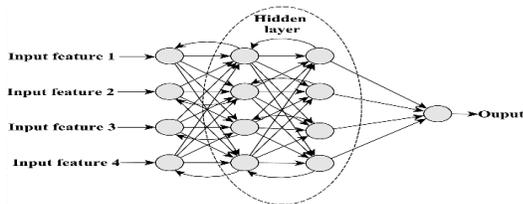


Fig. 3. Dynamic Neural Network.

3) *Multiple*: Several topologies are made (the exact number depends on the training data): pseudo-parallelization. The model with the lowest root mean square-squared error is declared the winner. Fig. 4 shows the multiple neural network structure.

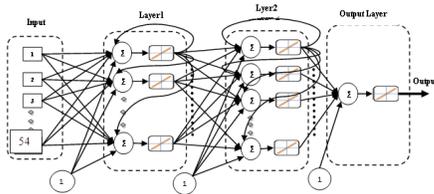


Fig. 4. Multiple Neural Network.

4) *Radial based Function Network (RBFN)*: Similar to k-means clustering, the radial basis algorithm partitions the data on values for the target area. Fig. 5 shows the rbf neural network structure.

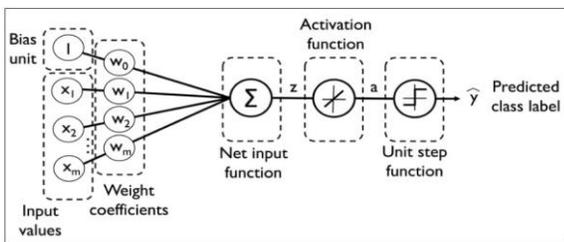


Fig. 5. RBF Neural Network.

5) *Prune*: Pruning is a compression technique that entails extracting weights from a learned model. Pruning is the process of removing unneeded branches or stems from a plant in agriculture. Pruning is the process of deleting superfluous neurons or weights in machine learning. Fig. 6 shows the prune neural network structure.

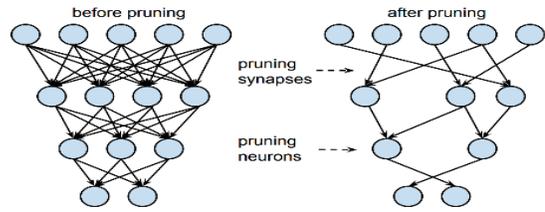


Fig. 6. Prune Neural Network.

The NN layout is motivated by the need to execute precise addition on a multidimensional array of information [43]. It can be thought of as a type of practical link network [37]. It employs a system design similar to that of classical regularization [44], in which the premise capacities correspond to the Green's components of the Gram's administrator associated with the stabilizer. The NN organization is obtained on the off-chance that the stabilizer exhibits outspread symmetry. From the estimation hypothesis's perspective, the regularization organization has three appealing properties [44, 45]: it can approximate any multivariate consistent capacity on a smaller space to a subjective degree of precision, given an adequate number of units; it has the best estimation property because the obscure coefficients are straight; and the arrangement is ideal by limiting a practical containment.

As illustrated in Fig. 7, the general function is a three-layer (J1-J2-J3) feed forward neural network. Each node in the hidden layer is activated by an activation function such as (RBF), denoted by ( $\phi$ ). The hidden layer performs a nonlinear transformation on the input, whereas the output layer is a linear integrator that maps the nonlinearity to a new space. In general, the activation function is applied to all nodes; that is, NN nodes have the nonlinearity ( $\phi(x) = \phi(x - ci)$ ,  $i = 1, \dots, J2$ , where  $ci$  is the prototype or centre of the  $i$ th node and  $\phi(x)$  is an NN. The output layer neurons' biases can be approximated with an additional neuron in the hidden layer with an activation function of  $\phi(0) = 1$ . Fig. 7 demonstrates the general architecture of the NN.

The  $J1$ ,  $J2$ , and  $J3$  neurons are used in the input, hidden, and output layers, respectively.  $\phi(0(x)) = 1$  denotes the output layer's function, whereas  $\phi_i(x)$  denotes the hidden nodes' nonlinearity.

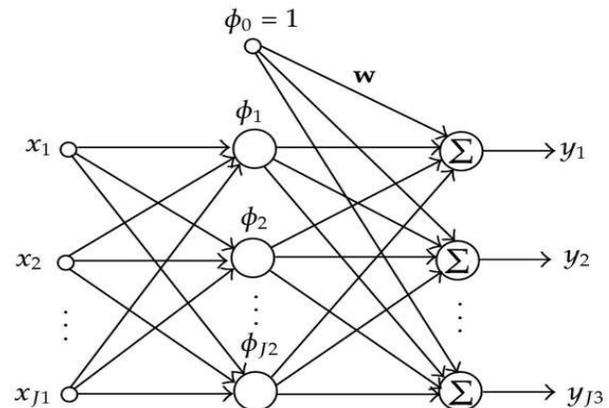


Fig. 7. Architecture of the Neural Network.

## V. EXPERIMENTAL RESULTS AND DISCUSSION

The purpose of this experiment was to filter and detect patient status as a positive or negative diagnostic of disease. Each sample from the dataset was classified into five groups. Each category had a predetermined sample size (disease cases). Four sets were deemed learning stages, while the remaining one was used as a cross-validation testing dataset. The purpose of this research was to demonstrate the robustness of neural network type prediction and classification when applied to medical information.

As previously stated, the dataset was prepared and tested using the 5-fold cross-validation approach. The examination connected the datasets using the outputs of a neural network classifier in order to assess the strength of NN Types for medical domain diagnosis. The cross-validation approach resulted in the following judgment accuracy outcomes:

$$Accuracy = \frac{(TN + TP)}{(TN + FP) + (TP + FN)} \times 100 \quad (1)$$

True Negative (TN): The number of wrongly identified positive and negative executable; True Positive (TP): The proportion of benign and positive executable that have been appropriately diagnosed; False Negative (FN): The number of positive executable that have been incorrectly classified as negative; False Positive (FP): The number of benign executable that have been incorrectly identified as negative.

In this study, the IBM SPSS modeler software has been used as the machine learning tool to design our experimental models. The IBM Modeler is the high-performance data mining tool for enterprise. Via an in-depth understanding of data, Modeler enables organizations to strengthen consumer and citizen partnerships. Businesses use Modeler insight to retain successful clients, find cross-selling opportunities, recruit new customers, track fraud, mitigate risk, and enhance government service delivery. The visual interface of Modeler encourages users to apply their domain knowledge, which results in more efficient predictive models and a shorter time to solution. Modeler includes a variety of simulation methods, including algorithms for prediction, sorting, segmentation, and association detection. Modeler Solution Publisher facilitates the enterprise-wide distribution of models to decision makers or to a database after they are developed. The experiments of the investigation study are divided into several steps:

- Step 1: Data collection and preparation from the (UCI dataset).
- Step 2: Divide the dataset into training and testing parts.
- Step 3: Apply different Neural Networks Algorithms (Quick, Dynamic, Multiple, and Radial Basis Function Network (RBFN), Prune, and Executive Prune).
- Step 4: Results analysis.

### A. Data Collection and Preparation

- Data collection: In this step, the medical datasets has been collected from University of California Irvine (UCI) Machine Learning data Repository as a source of

the datasets. The UCI repository currently maintains 559 datasets as a service to the machine learning community. You may view all datasets through the searchable interface and the following website:

- Data Preparation: We performed pre-processing steps during this procedure, which includes data cleansing, outlier values elimination, and missing values handling (?). The pre-processing has been handled with the following choice based on the dataset description.

1) Delete the corresponding row from the dataset (if the number of missing values in less).

2) Determine the average value for each function and then substitute it for the missing value (numeric values).

3) Counting the number of zeros and ones in each feature and then substitute the highest count values (0 or 1) in the missing value in each feature individually.

- Dataset Descriptions: this section discussed the medical datasets that have been used which are Breast cancer, Chronic Kidney, Diabetes, Parkinson, COVID-19, Epileptic Seizure, HCV, and Heart Disease.

#### • Breast Cancer Dataset

The breast cancer databases were obtained from the University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg. The description of the dataset is illustrated in Table I.

TABLE I. BREAST CANCER DATASET DESCRIPTION

Features Description	No of Samples	No of Features	Machine Learning Task
Integer	699	10	Classification

#### • HCV (Hepatitis C Virus) Dataset

The data collection includes laboratory results for blood donors and Hepatitis C patients, as well as demographic information such as the age of the patient. The description of the dataset is illustrated in Table II.

TABLE II. HCV DATASET DESCRIPTION

Features Description	No of Samples	No of Features	Machine Learning Task
Real & Integer	615	14	Classification & Clustering

#### • Parkinson Dataset

The dataset contains a variety of biomedical speech measurements taken from 42 people with early-stage Parkinson's disease who were enrolled in a six-month trial with a digital symptom progression tracking system. Automatic recordings were made in the patients' homes. The description of the dataset is illustrated in Table III.

TABLE III. PARKINSON DATASET DESCRIPTION

Features Description	No of Samples	No of Features	Machine Learning Task
Real & Integer	5875	26	Classification, Clustering, Regression

- Heart Disease

Heart disease database includes 76 attributes; all reported studies make use of a subset of 14. To date, only the Cleveland database has been used by machine learning researchers. The "goal" area indicates if the patient has heart disease. It is an integer number from 0 (no presence) to 4. The description of the dataset is illustrated in Table IV.

TABLE IV. HEART DISEASE DATASET DESCRIPTION

Features Description	No of Samples	No of Features	Machine Learning Task
Real, Integer, Categorical	303	75	Classification

- Diabetes

Diabetes medical records were gathered from two sources: an automated electronic recorder and paper records. The automated system had an internal clock that was used to timestamp events, while paper documents contained only "logical time" slots (breakfast, lunch, dinner, bedtime). Breakfast (08:00), lunch (12:00), dinner (18:00), and bedtime were both set times in paper documents (22:00). Thus, while paper documents contain fictitious standard tracking dates, electronic records contain more accurate time stamps. The description of the dataset is illustrated in Table V.

TABLE V. DIABETES DATASET DESCRIPTION

Features Description	No of Samples	No of Features	Machine Learning Task
Integer, Categorical	100000	55	Classification & Clustering

- Epileptic Seizure

The original dataset from the guide is divided into five directories, each of which contains 100 files, each of which represents a single subject/person. Each file contains a 23.6-second recording of brain activity. 4,097 data points are sampled from the corresponding time sequence. Each data point represents the value of the EEG recording at a certain time point. Thus, we have a total of 500 people, each with 4,097 data points over a period of 23.5 seconds. The description of the dataset is illustrated in Table VI.

TABLE VI. EPILEPTIC SEIZURE DATASET DESCRIPTION

Features Description	No of Samples	No of Features	Machine Learning Task
Real, Integer	11500	179	Classification & Clustering

- Chronic Kidney

The dataset can be used to classify chronic kidney disease and was obtained from a hospital over a period of almost two months. The description of the dataset is illustrated in Table VII.

TABLE VII. CHRONIC KIDNEY DATASET DESCRIPTION

Features Description	No of Samples	No of Features	Machine Learning Task
Integer, Categorical	100000	55	Classification & Clustering

- COVID-19

The X-ray data is saved in the png, jpg, and jpeg formats. Johns Hopkins University uses the two databases from the *Kaggle X-ray* competition as part of their medical database. The comparisons were made between cases with a virus causing bacterial pneumonia, healthy people, and cases caused by COVID19 [46]. One of the photographs in the dataset is of people who have had pneumonia. The COVID X-ray image database was created by Cohen JP using open access images from many different sources. This archive contains a lot of photographs that have been shared with the creators. This archive contains 125 photographs of an X-ray diagnosis of COVID. There are 43 women and 82 men who have demonstrated that they have a certain propensity for creative thinking within the study. Datasets without complete metadata do not include all users. The average age of COVID patients is 55, but the ages of the members range from 26 to 89 years. a database of chestxray8 has been developed by Wang et al. [47] in order to compensate for unbalanced images from this series of unbalanced data, we've produced an even distribution of 500 no-finding and 500-frontal chest X-rays that looks random.

### B. Divide the Dataset into Training and Testing Parts

After the data mining pre-processing steps, we then divide the dataset into two groups. The first group is training data and the data distribution. The investigation models were built based on training and testing dataset. The training dataset was divided into 90, 80, 70, 60, and 50 whereas the testing dataset was divided into 10, 20, 30, 40, and 50 respectively. The training results of the neural network algorithms were examined based on learning groups for each neural network algorithm (Quick, Dynamic, Multiple, and Radial Basis Function Network (RBFN), Prune, and Executive Prune). The accuracy measurement was calculated to examine all the collected medical datasets, including Breast cancer, Chronic Kidney, Diabetes, Parkinson, COVID-19, Epileptic Seizure, HCV, and Heart diseases. Then the average results were computed to select the optimal neural network algorithm for each dataset.

### C. Applying Neural Network Algorithms

Our investigated study used the neural network algorithms for data prediction to diagnosis different types of diseases. The model was built by selecting the original data after dividing the datasets into training and testing parts. The IBM SPSS Modeler contains six types of neural network algorithms

which are: Quick, Dynamic, Multiple, and Radial Basis Function Network (RBFN), Prune, and Executive Prune. The advantage of using the SPSS modeler is that the system applies integration and data visualization in order to show the predicted results explicitly. A sample of investigation model is demonstrated in Fig. 8.

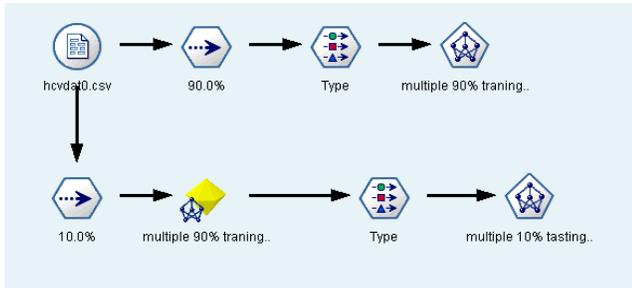


Fig. 8. Sample of Investigation Model using Neural Network Algorithms.

The experimental design process and running procedure has been examined based on the training and testing models.

#### D. Results Analysis

In the experiments, the medical datasets were used in order to determine the patient (injured or not injured) or predict the diagnosis percentage for some diseases. The dataset had each instance reported as either an injured or not injured case or labelled with target field (Class feature). The neural network algorithm was applied by training and testing the dataset using our investigation models. The main objectives of the learning model in this study is to investigate the diagnosis level by collecting the patients' samples with similar patterns together, thus the variation will be reduced and the diagnosis interpretation will be accurate.

The obtained results of the investigation models have been extracted and analyzed individually for training and testing results. The average results for each algorithm based on the specific disease have been calculated and figured out in different shapes. We noted that the multiple neural network algorithm is better than the other five neural network algorithms in many medical datasets for prediction diagnosis measure. Each fold in the training and testing model is examined with the individual dataset. Fig. 9 to 13 demonstrate the training results of each dataset using different types of neural network algorithms, while Fig. 14 to 18 shown the testing results for accumulated 5 folds cross validation.

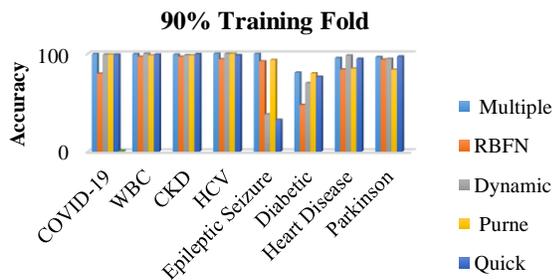


Fig. 9. Training Model with 90%.

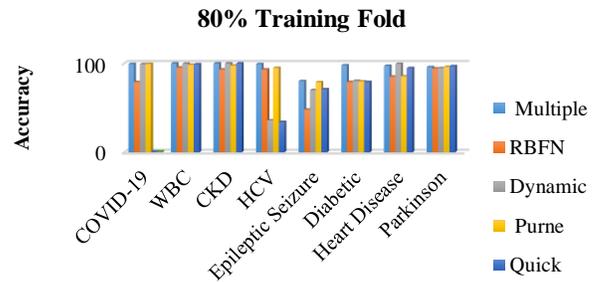


Fig. 10. Training Model with 80%.

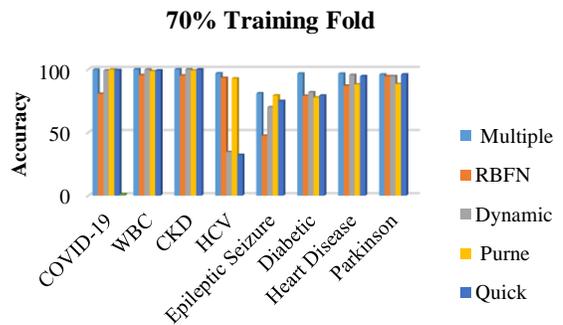


Fig. 11. Training Model with 70%.

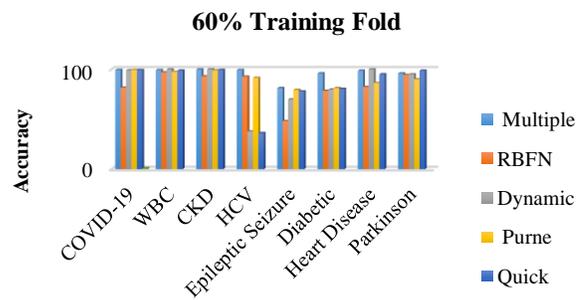


Fig. 12. Training Model with 60%.

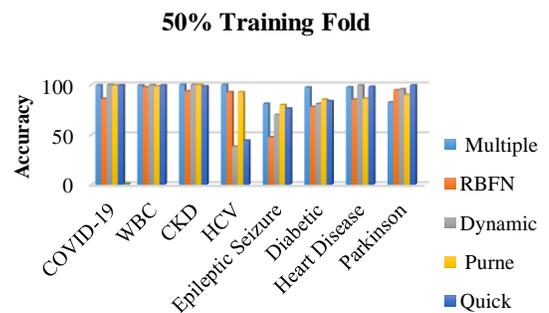


Fig. 13. Training Model with 50%.

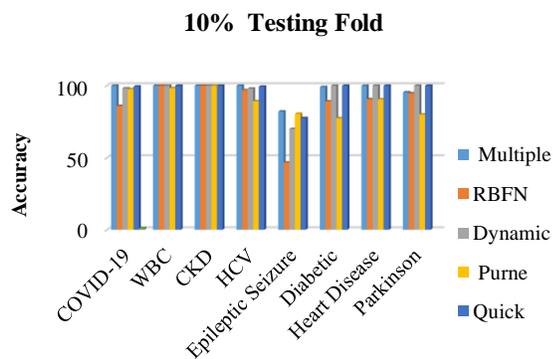


Fig. 14. Testing Model with 10%.

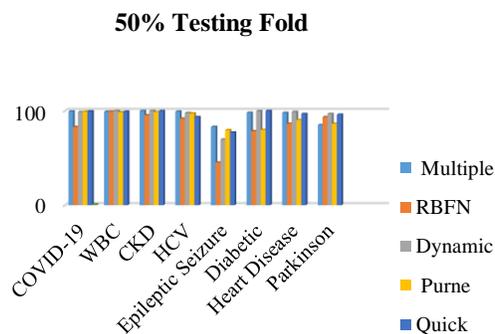


Fig. 18. Testing Model with 50%.

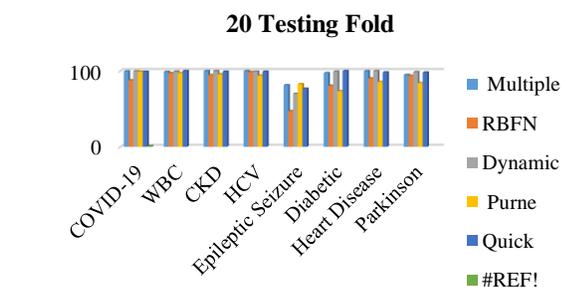


Fig. 15. Testing Model with 20%.

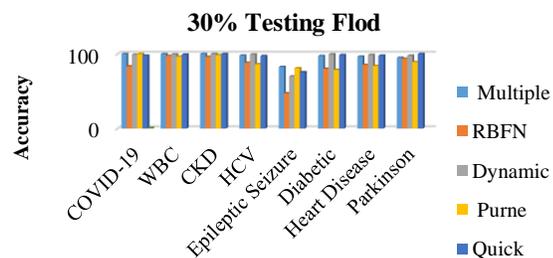


Fig. 16. Testing Model with 30%.

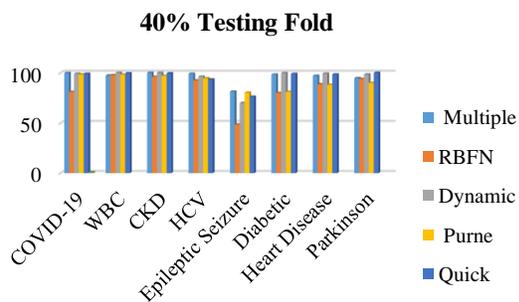


Fig. 17. Testing Model with 40%.

The accuracy of the results of the training and testing model using the five parts folds has been calculated for each dataset using the five main types of the neural network algorithms. Table VIII and Table IX demonstrate the training and testing results for each fold.

The findings of our prediction model trials indicated that neural network algorithms improved performance, and the t-test technique was utilized to quantify the improvement. Low t-test scores (usually less than 0.05) indicate that the two variables have been adjusted significantly. This criterion was emphasized in the assessment measures in light of the diagnosis accuracy values obtained in Table X between the multiple neural network method and the other five neural network algorithms. This demonstrates that Multiple NN outperformed RBFN, and Prune in terms of diagnostic performance. The results of the t-test statistical significance test are shown in Table X.

As shown in Table X, the significant T-test strategy has been applied between the investigated neural network algorithms in this study. We noted that the multiple neural network algorithm achieved significant results less than 0.05 when compared with RBF and Prune algorithms with P-value equal to 0.21 and 0.20 respectively. On the other hand, the Multiple neural network obtained better classification accuracy results as shown in Tables V and IV when compared with Quick and Dynamic algorithms without significant different results under the T-test investigation.

TABLE VIII. FOLD TRAINING RESULTS USING NEURAL NETWORK ALGORITHMS

90 Training Fold					
Dataset	Multiple	RBFN	Dynamic	Prune	Quick
COVID-19	99.7	79.6	99.11	99.1	99.18
WBC	99.6	96.9	100	98.4	99.21
CKD	99.1	97	98.57	98.4	99.64
HCV	100	94.5	100	100	98.54
Epileptic Seizure	99.8	92.2	38.05	93.8	32.47
Diabetic	80.77	47.7	70.086	80	76.49
Heart Disease	95.7	83.9	98.19	84.8	94.87
Parkinson	96.7	94.1	94.88	83.7	97.15
80 Training Fold					
Dataset	Multiple	RBFN	Dynamic	Prune	Quick
COVID-19	99.4	79	99.18	99.5	97.72
WBC	100	95.4	99.8	98.6	99.1
CKD	100	93.2	100	97.6	100
HCV	99.3	93.2	35.9	95.1	34.1
Epileptic Seizure	80.1	48	70	78.9	71
Diabetic	97.9	79.1	80.3	79.8	79.1
Heart Disease	97.3	85.1	99.5	85.5	94.8
Parkinson	95.9	94.4	94.8	96.1	96.9
Dataset	Multiple	RBFN	Dynamic	Prune	Quick
COVID-19	99.7	80.6	99.16	99.8	99.46
WBC	100	95.4	99.802	98.6	99.16
CKD	100	95.07	100	99.1	99.9
HCV	96.8	93.2	34.5	92.7	32.18
Epileptic Seizure	80.95	47.3	69.95	79.3	74.88
Diabetic	96.6	79	81.87	77.7	79.15
Heart Disease	96.5	87	95.53	88	94.57
Parkinson	95.8	94.5	94.57	88.3	95.9
Dataset	Multiple	RBFN	Dynamic	Prune	Quick
COVID-19	99.5	81.7	99.12	99.5	99.34
WBC	99.4	97.1	100	97.6	98.83
CKD	100	93.03	100	99.1	99.5
HCV	99.4	92.7	37.78	91.5	36.26
Epileptic Seizure	81.3	48.2	69.95	79.5	77.95
Diabetic	96.1	78.5	79.86	81.3	80.46
Heart Disease	98.7	82.4	100	86.4	95.08
Parkinson	96	94.4	95.07	90.1	98.71
Dataset	Multiple	RBFN	Dynamic	Prune	Quick
COVID-19	99.5	86.1	99.85	99.4	99.6
WBC	99.4	97.6	99.71	98.7	99.43
CKD	100	93.7	100	100	98.44
HCV	100	92.7	38.2	92.6	44.11
Epileptic Seizure	81.1	47.7	69.95	79.8	76.31
Diabetic	97.5	78	81.02	85.2	83.6
Heart Disease	97.6	85.3	99.37	86.2	98.07
Parkinson	82.5	94.6	95.7	90.1	99.4

TABLE IX. FOLD TESTING RESULTS USING NEURAL NETWORK ALGORITHMS

10 Testing Fold					
Dataset	Multiple	RBFN	Dynamic	Prune	Quick
COVID-19	100	85.9	98.14	97.6	99.37
WBC	100	100	100	98.3	100
CKD	100	100	100	100	99.98
HCV	100	96.9	97.97	89.2	99.33
Epileptic Seizure	82.1	46.7	69.93	80.4	77.55
Diabetic	99.1	89.1	100	77.4	100
Heart Disease	99.9	90.6	100	90.5	100
Parkinson	95.4	94.7	100	80	100
20 Testing					
Dataset	Multiple	RBFN	Dynamic	Prune	Quick
COVID-19	99.6	87.7	99.76	99.3	99.02
WBC	99	97.3	99.33	97.5	100
CKD	100	94.8	100	95.7	99.29
HCV	100	98.7	99.29	94.1	99.28
Epileptic Seizure	81.3	47.2	70.02	82.7	76.64
Diabetic	97.3	80.6	99.35	73.6	100
Heart Disease	99.9	90.4	100	85.6	98.27
Parkinson	94.9	93.7	98.72	84.2	98.07
30 Testing Fold					
Dataset	Multiple	RBFN	Dynamic	Prune	Quick
COVID-19	100	83.4	99.15	100	99
WBC	99.9	97.5	99.53	96.4	99.35
CKD	100	96.2	100	98.6	100
HCV	98	88.2	99.41	86.3	97.56
Epileptic Seizure	82.6	47.3	70.05	80.9	75.61
Diabetic	97.3	80.2	99.55	78.8	98.6
Heart Disease	96.5	85.4	98.76	84	98.8
Parkinson	95	93.6	97.8	89.2	100
40 Testing Fold					
Dataset	Multiple	RBFN	Dynamic	Prune	Quick
COVID-19	99.8	81	99.19	98.5	99.77
WBC	97.4	97.8	100	98.3	99.66
CKD	100	96.05	100	97	99.61
HCV	99.1	92.4	96.96	94.6	97.39
Epileptic Seizure	82.1	48.5	69.95	80.2	76.22
Diabetic	98.3	80	99.86	81.1	99.02
Heart Disease	97.1	88.66	99.2	88.1	98.4
Parkinson	94.7	93.8	98.4	90	100
50 Testing Fold					
Dataset	Multiple	RBFN	Dynamic	Prune	Quick
COVID-19	99.5	83	98.89	99.3	99.59
WBC	99.2	97.4	100	98.7	99.42
CKD	100	95.6	100	98.9	100
HCV	99.3	91.8	97.84	97.3	93.52
Epileptic Seizure	83.1	45.4	69.73	79.7	77.16
Diabetic	98.1	78.5	99.74	79.8	100
Heart Disease	98	86.4	98.96	90.4	96.86
Parkinson	85	93.5	96.86	86.4	96.09

TABLE X. T-TEST STATISTICAL SIGNIFICANCE RESULTS

	Differences between the correlation coefficient factor, mean absolute error, root mean squared error, relative absolute error, and root relative squared error before and after the improvement				t	df	P Value
	Mean	Std. Deviation	95% Confidence Interval of the Difference				
			Lower	Upper			
Multiple-RBF	5.6800	3.4339	1.41619	9.94381	3.699	4	.021
Multiple-Dynamic	1.17200	1.66354	-.89356	3.23756	2.785	4	.190
Multiple-Prune	6.98000	4.15295	1.82343	12.3657	3.758	4	.020
Multiple-Quick	1.86400	2.24282	-.92083	4.64883	1.858	4	.137

## VI. CONCLUSION AND FUTURE WORK

## REFERENCES

Machine-learning modelling is analysed to support the provision of the best possible care to all patients with medically relevant data used by millions of healthcare clinicians in decision making for trillions of patients. The speeding up of vast volumes of healthcare data would radically change the structure of the healthcare system. We firmly agree that the relationship between patient and doctor will be the fundamental cornerstone of treatment for many patients and that new developments into machine learning will contribute to this relationship. We expect a few early models, along with the development of regulatory frameworks and economic incentives for value-based services, to be published in the next several years to make us meticulously optimistic about machine education in the field.

This study attempted to analyse medical disease diagnostics in order to modify the prediction method for different types of neural networks. We highlighted the quality of illness prediction models by utilising Quick, Multiple, Dynamic, and RBFN neural network algorithms, as well as prune neural network algorithm. In this study, the experiments conducted were based on different types of medical datasets such as Breast cancer, Chronic Kidney, Diabetes, Parkinson, COVID-19, Epileptic Seizure, HCV, and Heart disease. Our investigation found that the diagnosis results can be predicted and achieved by the neural network algorithms with different types of medical datasets. Additionally, our extensive investigations revealed that the multiple neural network algorithm had the greatest results in terms of diagnostic accuracy for a variety of ailments. In addition, P-value scores have been computed and indicate that the multiple neural network algorithm has significantly better performance compared with other neural Network algorithms. In future work, our research has try to improve all the objectives that were addressed in this study. However, the quality of prediction methods in medical disease has been investigated using neural network algorithms. At some point, this study will plan to apply another types of the neural network algorithms on different medical datasets to achieve high prediction and diagnosis models.

## ACKNOWLEDGMENT

This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant No. (D-172-830-1440). The authors, therefore, gratefully acknowledge DSR technical and financial support.

- [1] Osman, A.H., et al., SOM-LWL method for identification of COVID-19 on chest X-rays. *PloS one*, 2021. 16(2): p. e0247176.
- [2] Osman, A.H. and H.M.A. Aljahdali, An effective of ensemble boosting learning method for breast cancer virtual screening using neural network model. *IEEE Access*, 2020. 8: p. 39165-39174.
- [3] Hassan, M.I., A.H. Osman, and E.M. Hussein. Heart Disease Diagnosis Using Diverse Neural Network Categories. in *International Conference of Reliable Information and Communication Technology*. 2019. Springer.
- [4] Osman, A.H. and A.A. Alzahrani, New approach for automated epileptic disease diagnosis using an integrated self-organization map and radial basis function neural network algorithm. *IEEE Access*, 2018. 7: p. 4741-4747.
- [5] Osman, A.H., An enhanced breast cancer diagnosis scheme based on two-step-SVM technique. *Int. J. Adv. Comput. Sci. Appl*, 2017. 8(4): p. 158-165.
- [6] Alshamrani, B.S. and A.H. Osman, Investigation of hepatitis disease diagnosis using different types of neural network algorithms. *International Journal of Computer Science and Network Security (IJCSNS)*, 2017. 17(2): p. 242.
- [7] Osman, A.H. and H.M. Aljahdali, Diabetes disease diagnosis method based on feature extraction using K-SVM. *Int J Adv Comput Sci Appl*, 2017. 8(1).
- [8] Nass, L., S. Swift, and A. Al Dallal, Indepth analysis of medical dataset mining: a comparative analysis on a diabetes dataset before and after preprocessing. *KnE Social Sciences*, 2019: p. 45-63-45-63.
- [9] Wang, F., et al., Guest editorial: Special issue on data mining for medicine and healthcare. 2015, Springer.
- [10] Gui, C. and V. Chan, Machine learning in medicine. *University of Western Ontario Medical Journal*, 2017. 86(2): p. 76-78.
- [11] Francis, F. and J. Saleema, Feature Selection in Data Mining using Permutation Combination. *International Journal of Advanced Research in Computer Science*, 2017. 8(3).
- [12] Shameer, K., et al., Machine learning in cardiovascular medicine: are we there yet? *Heart*, 2018. 104(14): p. 1156-1164.
- [13] Sidey-Gibbons, J.A. and C.J. Sidey-Gibbons, Machine learning in medicine: a practical introduction. *BMC medical research methodology*, 2019. 19(1): p. 64.
- [14] Sawhney, H. and H. Kaur, IMPLEMENTATION AND APPLICATIONS OF DATA MINING IN MEDICAL DECISION MAKING PREDICTIONS. *International Journal of Advanced Research in Computer Science*, 2017. 8(7).
- [15] Azar, A.T. and S.M. El-Metwally, Decision tree classifiers for automated medical diagnosis. *Neural Computing and Applications*, 2013. 23(7-8): p. 2387-2403.
- [16] Kourou, K., et al., Machine learning applications in cancer prognosis and prediction. *Computational and structural biotechnology journal*, 2015. 13: p. 8-17.
- [17] Obermeyer, Z. and E.J. Emanuel, Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine*, 2016. 375(13): p. 1216.
- [18] Yoo, C., L. Ramirez, and J. Liuzzi, Big data analysis using modern statistical and machine learning methods in medicine. *International neurology journal*, 2014. 18(2): p. 50.

- [19] Chen, J.H. and S.M. Asch, Machine learning and prediction in medicine—beyond the peak of inflated expectations. *The New England journal of medicine*, 2017. 376(26): p. 2507.
- [20] Nematollahi, M., et al., Classification models to predict survival of kidney transplant recipients using two intelligent techniques of data mining and logistic regression. *International journal of organ transplantation medicine*, 2017. 8(2): p. 119.
- [21] National Academies of Sciences, E. and Medicine, *Improving diagnosis in health care*. 2015: National Academies Press.
- [22] Berwick, D.M. and M.E. Gaines, How HIPAA harms care, and how to stop it. *Jama*, 2018. 320(3): p. 229-230.
- [23] Obermeyer, Z. and T.H. Lee, Lost in thought: the limits of the human mind and the future of medicine. *The New England journal of medicine*, 2017. 377(13): p. 1209.
- [24] Schwartz, W.B., *Medicine and the computer: the promise and problems of change*, in *Use and Impact of Computers in Clinical Medicine*. 1970, Springer. p. 321-335.
- [25] Kirmani, M., Heart disease prediction using multilayer perceptron algorithm. *International Journal of Advanced Research in Computer Science*, 2017. 8(5).
- [26] Rajkomar, A., J. Dean, and I. Kohane, Machine learning in medicine. *New England Journal of Medicine*, 2019. 380(14): p. 1347-1358.
- [27] Stanton, J.M., Galton, Pearson, and the peas: A brief history of linear regression for statistics instructors. *Journal of Statistics Education*, 2001. 9(3).
- [28] Reed, M., et al., Implementation of an outpatient electronic health record and emergency department visits, hospitalizations, and office visits among patients with diabetes. *Jama*, 2013. 310(10): p. 1060-1065.
- [29] Jaffe, M.G., et al., Improved blood pressure control associated with a large-scale hypertension program. *Jama*, 2013. 310(7): p. 699-705.
- [30] Yuasa, T., et al., Tumor size is a potential predictor of response to tyrosine kinase inhibitors in renal cell cancer. *Urology*, 2011. 77(4): p. 831-835.
- [31] de Vries, E.N., et al., Effect of a comprehensive surgical safety system on patient outcomes. *New England Journal of Medicine*, 2010. 363(20): p. 1928-1937.
- [32] Shnorhavorian, M., et al., Maternal risk factors for congenital urinary anomalies: results of a population-based case-control study. *Urology*, 2011. 78(5): p. 1156-1161.
- [33] Somogyi, R. and C.A. Sniegoski, Modeling the complexity of genetic networks: understanding multigenic and pleiotropic regulation. *complexity*, 1996. 1(6): p. 45-63.
- [34] Goodwin, B.C., Oscillatory behavior in enzymatic control processes. *Advances in enzyme regulation*, 1965. 3: p. 425-437.
- [35] Griffith, J., Mathematics of cellular control processes II. Positive feedback to one gene. *Journal of Theoretical Biology*, 1968. 20(2): p. 209-216.
- [36] Citro, G., G. Banks, and G. Cooper, INKBLOT: a neurological diagnostic decision support system integrating causal and anatomical knowledge. *Artificial Intelligence in Medicine*, 1997. 10(3): p. 257-267.
- [37] Lucas, P.J., et al., A probabilistic and decision-theoretic approach to the management of infectious disease at the ICU. *Artificial Intelligence in Medicine*, 2000. 19(3): p. 251-279.
- [38] Pearl, J., *Morgan Kaufmann series in representation and reasoning. Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo, CA, US: Morgan Kaufmann, 1988.
- [39] Azadeh, A., et al., Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption. *Applied Mathematics and Computation*, 2007. 186(2): p. 1731-1741.
- [40] McClelland, J.L., B.L. McNaughton, and R.C. O'Reilly, Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, 1995. 102(3): p. 419.
- [41] Calvo, R.A., et al., Autonomy in technology design, in *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. 2014. p. 37-40.
- [42] Kresse, G. and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 1996. 54(16): p. 11169.
- [43] Powell, M.J. Radial basis function for multivariable interpolation: a review. in *IMA Conference on Algorithms for the Approximation of Functions and Data*, 1985. 1985. RMCS.
- [44] Poggio, T. and F. Girosi, Networks for approximation and learning. *Proceedings of the IEEE*, 1990. 78(9): p. 1481-1497.
- [45] Girosi, F. and T. Poggio, Networks and the best approximation property. *Biological cybernetics*, 1990. 63(3): p. 169-176.
- [46] Cohen, J.P., P. Morrison, and L. Dao, COVID-19 image data collection. *arXiv preprint arXiv:2003.11597*, 2020.
- [47] Wang, X., et al. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.