كلية الدراسات العليا

Sudan University of Science and Technology

College of Graduate Studies

A thesis submitted in partial fulfillment of the master degree in biomedical engineering

Diabetes prediction by using artificial neural network

التنبؤ بمرض السكري عن طريق إستخدام الشبكات العصبية الإصطناعية

By:

Eman Ibrahim Abdelbagi Ibrahim

Supervisor:

Dr. Eltahir Mohamed Hussein

Dec.2022

Appendix A

Code

Dedication

To my parents, my husband, my children, my brothers, my sisters and my friends for always loving and supporting me.

And to my sister in low Fatima Khalid, Thank you very much for your support. This work would not have been possible without them.

Acknowledgement

I would like to express my deep gratitude to Dr. Altahir Mohamed Hussein, my research supervisor, for their patient guidance, enthusiastic encouragement and useful critiques of this research work.

I would also like to thank Eng. Saddam Saed Ali Ahmed, for his advice, assistance and help me in doing the programing section and data analysis.

I would also like to extend my thanks to all staff in Sudan University of science and technology, biomedical engineering department.

Finally, I wish to thank my mother for their support and encouragement throughout my study.

List of Figures

Figure 1: Proposed methodology …………………………………….. 2
Figure 2: The branches and sections of the computational methods...... 13
Figure 3: One hidden layer MLP……………………………………… 20
Figures 4: Flow diagram of proposed model………………………….. 24
Figures 5: The first five observation units of the data set……………... 26
Figures 6: Outcome classes……………………………………………. 26
Figures 7: histogram of the variables …………………………………. 27
Figures 8: Correlation matrix …………………………………………. 28
Figures 9: confusion matrix ……………………………………………... 35

List of Tables

Table 1.1: Summary of Literature review……………………………… 5
Table 5.1: The results……………………………………………….... 32
Table 5.2: Classification report……………………………..………. 34

Table of contents

1 Chapter one: Introduction

1.1 General views…………………………………………….............. 1

1.2 Problem statement……………………………………………………. 2

1.3 Objective………………………………………………………………. 2

1.4 Methodology…………………………………………………………. 2

1.5 Outline……………………………………………………………... 2

2 Chapter two: literature reviews

2.1 literature reviews……………………………………………………… 3

3 Chapter three theoretical backgrounds

3.1 Introduction to diabetes…………………………………………… 8

3.2 Symptoms………………………………………………………….. 9

3.3 Diabetes categories………………………………………………… 9

3.4 Artificial Neural Network………………………………………….. 11

3.5 Artificial neural network structure…………………………………. 13

3.6 Artificial Neural Network topologies……………………………… 15

3.7 Machine Learning………………………………………………... 16

3.8. Deep learning ANN systems………………………………………. 17

3.9 Preprocessing data…………………………………………………. 18

3.10 Neural network model (supervised)……………………………… 20

3.11 Multi-layer Perceptron (MLP) ……………………………………... 20

3.12 Classification…………………………………………………….. 21

3.13 Mathematical formulation……………………………………….. 22

4 Chapter four: Methodology

4.1 Methodology …………………………………………………. 24

4.2 Data set……………………………………………………….. 25

4.3 Data analysis………………………………………………….. 26

4.4 Correlation matrix…………………………………………….. 28

4.5 Split data……………………………………………………… 28

4.6 Preprocessing data……………………………………………. 29

4.7 Train data……………………………………………………… 29

4.8 Test data……………………………………………………….. 30

4.9 Applying algorithm…………………………………………… 30

5 Chapter five: Result and discussion

5.1 The results……………………………………………………….. 32

5.2 Cross validation result…………………………………………… 33

5.3 Evaluation……………………………………………………….. 33

5.4 Confusion matrix……………………………………………… 33

5.5 Discussion……………………………………………………... 35

6 chapter six: Conclusion and recommendations

6.1 Conclusion……………………………………………………... 36

6.2 Recommendations……………………………………………… 36

References………………………………………………………………….. 38
Appendix A: Code

Abstract

Diabetes is an illness caused because of high glucose level in a human body. Diabetes should not be ignored if it is untreated then Diabetes may cause some major issues in a person like: heart related problems, kidney problem, blood pressure, eye damage and it can also affects other organs of human body. Diabetes can be controlled if it is predicted earlier. To achieve this goal this proposed model built to early prediction of Diabetes in a human body through applying artificial neural network by used Bython software. ANN Provide better result for prediction by constructing models from datasets collected from patients. In this work MLP Classifier used on a dataset to predict diabetes. The best hyper parameter combination obtained by used CV technique is :{'batch_size': 100, 'hidden_layer_sizes': 5, 'learning_rate_init': 0.001, 'max_iter': 500} with the best accuracy: 0.7662337662337663. The accuracy is varied by hyper parameters. The Proposed model gives a higher accuracy with the best parameter combination _manually entered_ :{'batch_size': 1
00, 'hidden_layer_sizes': 5, 'learning_rate_init': 0.005, 'max_iter': 600}. The best accuracy: 0.81168831168831.

المستخلص

مرض السكري هو مرض يحدث بسبب ارتفاع مستوى الجلوكوز في جسم الإنسان. لا ينبغي تجاهل مرض السكري فإذا لم يتم علاجه ، فقد يسبب بعض المشكلات الرئيسية لدى الشخص مثل: مشاكل القلب ومشاكل الكلى وضغط الدم وتلف العين ويمكن أن يؤثر أيضًا على أعضاء أخرى من جسم الإنسان. يمكن السيطرة على مرض السكري إذا تم التنبؤ به في وقت مبكر. لتحقيق هذا الهدف ، تم بناء هذا النموذج المقترح للتنبؤ المبكر لمرض السكري في جسم الإنسان من خلال تطبيق الشبكات العصبية الإصطناعية بواسطة برنامج بايثون. الشبكات العصبية الإصطناعية تقدم نتيجة أفضل للتنبؤ بالمرض من خلال بناء نماذج من البيانات التي تم جمعها من المرضى. في هذا البحث تم استخدام MLP classifierللتنبؤ بمرض السكري.

وجد ان افضل تركيبة للمتغيرات تم الحصول عليها باستخدام تقنية CVهي : :{' :'batch_size 100, 'hidden_layer_sizes': 5, 'learning_rate_init': 0.001, 'max_iter': 500. }

و التي أعطت افضل دقة كانت ٠,٧٦٦٢٣٣٧٦٦٢٣٣٧٦٦

تتغير دقة النموذج تبعا لتغير المتغرات المدخلة له. يعطي النموذج المقترح دقة أعلى ٠,٨١١٦٨٨٣١١٦٨٨٣١ مع مجموعة المتغيرات :{' batch_size': 100, 'hidden_layer_sizes': 5, 'learning_rate_init': 0.005, 'max_iter': 600}. المدخلة يدويا.

Chapter One

Introduction

1.1 General views:

Diabetes is a chronic disease that occurs either when the pancreas does not produce enough insulin or when the body cannot effectively use the insulin it produces. Insulin is a hormone that regulates blood glucose. Hyperglycemia, also called raised blood glucose or raised blood sugar, is a common effect of uncontrolled diabetes and over time leads to serious damage to many of the body's systems, especially the nerves and blood vessels.

In 2014, 8.5% of adults aged 18 years and older had diabetes. In 2019, diabetes was the direct cause of 1.5 million deaths and 48% of all deaths due to diabetes occurred before the age of 70 years. Another 460 000 kidney disease deaths were caused by diabetes, and raised blood glucose causes around 20% of cardiovascular deaths [1]. Between 2000 and 2019, there was a 3% increase in age-standardized mortality rates from diabetes. In lower-middle-income countries, the mortality rate due to diabetes increased 13%.

By contrast, the probability of dying from any one of the four main noncommunicable diseases (cardiovascular diseases, cancer, chronic respiratory diseases or diabetes) between the ages of 30 and 70 decreased by 22% globally between 2000 and 2019 [2].

According to the current classification there are two major types: type 1 diabetes (T1DM) and type 2 diabetes (T2DM), Type 1 and type 2 diabetes are the most common forms of the disease, but there are also other kinds, such as gestational diabetes, which occurs during pregnancy, as well as other forms [3].

1.2 **Problem statement**:

Diabetes is a major cause of blindness, kidney failure, heart attacks, stroke and lower limb amputation. The costs are also correlated with chronic conditions.

1.3 **General Objective**:

General objective is Diabetes prediction by using artificial neural network

1.4 **Methodology:**

The proposed methodology is presented as in the diagram below in figure 1.

Figure 1: Proposed methodology

1.5 Theses layout:

This research consists of six chapters. Chapter one is an introduction. Related studies is described in chapter two. Chapter three is about theoretical background. Methodology is presented in chapter four. Chapter five is about result and discussion. Finally conclusion and recommendations is presented in chapter six.

Chapter Two

Literature reviews

In the literature there are various approaches for diabetes prediction the analysis of related work gives results on various healthcare datasets, where analysis and predictions were carried out using various methods and techniques. Various prediction models have been developed and implemented by various researchers using variants of data mining techniques, machine learning algorithms or also combination of these techniques:

KM Jyoti Rani developed a system which can perform early prediction of diabetes for a patient with a higher accuracy by combining the results of different machine learning techniques. The algorithms like K nearest neighbor, Logistic Regression, Random forest, Support vector machine and Decision tree are used. The accuracy of the model using each of the algorithms is calculated. Then the one with a good accuracy is taken as the model for predicting the diabetes. [3]

Aditya Saxena,   Megha Jain, Prashant Shrivastav at study used data mining tools to predict diabetes. Different automated information systems have been developed for the prevention and diagnosis of diabetes using various classifiers. [10]

Mitushi Soni and Dr. Sunita Varma used Machine Learning Classification and ensemble techniques on a dataset to predict diabetes. Which are K-Nearest Neighbor (KNN), Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), Gradient Boosting (GB) and Random Forest (RF). The accuracy is different for every model when compared to other models. The Project work gives the accurate or higher accuracy model shows that the model is capable of predicting diabetes effectively. thier Result shows that Random Forest achieved higher accuracy compared to other machine learning techniques[11].

Olta Llahaa and Amarildo Ristab the purpose of their paper is to evaluate data mining methods and their performances that can be used for analyzing the collected data about the diabetes. They identified the most appropriate data mining methods to analyze the data by comparing them theoretically and practically. Some attributes of this dataset are: Age, Body Mass Index, Insulin, Glucose, etc.

Methods are applied on these data to determine their effectiveness in analyzing and preventing diabetes. Evaluations on the data showed that the method with a higher performance is "Decision Tree". This was achieved by some performance measures, such as the number of instances correctly classified, accuracy, precision, recall and F-measure, that has brought better results compared to other methods. [12]

Nesreen Samer El-Jerjawi, Samy S. Abu-Naser used artificial neural networks to predict whether a person is diabetic or not. The criterion was to minimize the error function in neural network training using a neural network model. After training the ANN model, the average error function of the neural network was equal to 0.01 and the accuracy of the prediction of whether a person is diabetics or not was 87.3%. [13]

Aishwarya Mujumdara, Dr. Vaidehi Vb proposed a diabetes prediction model for better classification of diabetes, in this study, various machine learning algorithms are applied on the dataset and the classification has been done using various algorithms of which Logistic Regression gives highest accuracy of 96%. These algorithms include Support Vector Classifier, Random Forest Classifier, Decision Tree Classifier, Extra Tree Classifier, Ada Boost algorithm, Perceptron, Linear Discriminant Analysis algorithm, Logistic Regression, K-Nearest Neighbour, Gaussian Naïve Bayes, Bagging algorithm, Gradient Boost Classifier. [14]

GeetaRani[a] Vijaypal, Singh Dhaka[a] Ramesh, Chandra Poonia[b] identify challenges in existing techniques, namely Naïve Bayes, decision tree, and support vector machine, and they propose effective solutions for these. The model applies artificial neural networks for detecting diabetes and identifying its type. For evaluation of the proposed model, the authors use the “Pima Indian Diabetes” dataset. The highest accuracy of 85.09% proves the efficacy of the proposed work. [15]

Parastoo RAHIMLOO, Ahmad JAFARIAN tried to combine statistical models and neural networks, create a new compound that has at least error and maximum reliability and is analyzed. The accuracy and efficiency of the method has been investigated and acceptable results compared to the neural network and logistic regression methods were obtained. In this research, the criteria are the performance to minimize the error function in neural network training using a neural network in a hybrid model which eventually came to the conclusion that the error function of the neural network is equal to 0.1 and combined neural network model is equal to 0.0002.[16]

Table 2.1: summary of Literature reviews

No	Author	Approach	Nb. Data	Result
1	KM Jyoti Rani, July-August-2020	five machine learning classification algorithms are studied	2000 cases	99%
2	Aditya Saxena, Megha Jain,Prashant Shrivastava, April 2021	data mining tools	-	-

3	- Mitushi Soni - Dr. Sunita Varma, September-2020	SVM, Knn, Random Forest, Decision Tree, Logistic Regression and Gradient Boosting classifiers	768	77%
4	Olta Llahaa Amarildo Ristab, May 2021	classification data mining methods: Naive Bayes, Decision Tree, Support Vector Machine (SVM), Logistic Regression	270	79%
5	Nesreen Samer El-Jerjawi, Samy S. Abu-Naser, 2018	Artificial Neural Networks	1004	87.3%
6	Aishwarya Mujumdara, Dr. Vaidehi Vb, 2019	Support Vector Classifier, Random Forest Classifier, Decision Tree Classifier, Extra Tree Classifier, Ada Boost algorithm, Perceptron, Linear Discriminant Analysis algorithm, Logistic Regression, K-Nearest Neighbour, Gaussian Naïve Bayes, Bagging algorithm, Gradient Boost Classifier.	800	96%.
7	NiteshPradhan[a]Geeta Rani[a]Vijaypal SinghDhaka[a]Ramesh ChandraPoonia[b], 2020	They identify challenges in existing techniques, namely Naïve Bayes, decision tree, and support vector machine	768	85.09 %

8	Parastoo RAHIMLOO , Ahmad JAFARIAN2, 2016	hybrid neural network mode: combining the statistical logistic regression model and neural networks	180	E.f is 0.1 , Hnnm :0.0002.

E.f: Error function.

Hnnm: hybrid neural network model.

Chapter Three

Theoretical background

3.1 Introduction to diabetes

The term diabetes mellitus describes a metabolic disorder of multiple aetiology characterized by chronic hyperglycaemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action, or both. The effects of diabetes mellitus include long– term damage, dysfunction and failure of various organs. Diabetes mellitus may present with characteristic symptoms such as thirst, polyuria, blurring of vision, and weight loss. In its most severe forms, ketoacidosis or a non–ketotic hyperosmolar state may develop and lead to stupor, coma and, in absence of effective treatment, death. Often symptoms are not severe, or may be absent, and consequently hyperglycaemia sufficient to cause pathological and functional changes may be present for a long time before the diagnosis is made. The long–term effects of diabetes mellitus include progressive development of the specific complications of retinopathy with potential blindness, nephropathy that may lead to renal failure, and/or neuropathy with risk of foot ulcers, amputation, Charcot joints, and features of autonomic dysfunction, including sexual dysfunction. People with diabetes are at increased risk of cardiovascular, peripheral vascular and cerebrovascular disease. Several pathogenetic processes are involved in the development of diabetes. These include processes which destroy the beta cells of the pancreas with consequent insulin deficiency, and others that result in resistance to insulin action. The abnormalities of carbohydrate, fat and protein metabolism are due to deficient action of insulin on target tissues resulting from insensitivity or lack of insulin [20].

In this disease is destroy the ability to produce insulin in the patient's body or the body becomes resistant to insulin the and therefore the produced insulin cannot perform its normal function. The primary role of insulin to lower blood sugar by different mechanisms. There are two main types of diabetes. In Type I diabetes, destruction of beta cells in pancreatic leads to impaired insulin production and in type II, there is a progressive insulin resistance in the body and eventually may lead to the destruction of pancreatic beta cells and defects in insulin production. In type II diabetes it is known that genetic factors, obesity and lack of physical activity have an important role in a person [17].

3.2 Symptoms:

- Excessive excretion of urine (polyuria).
- Thirst (polydipsia).
- Constant hunger.
- Weight loss.
- Vision changes.
- Fatigue.

These symptoms may occur suddenly [9].

In general, a higher risk of diabetes infection correlates with various indicators such as female gender, age over 35, and unhealthy-weight individuals. But with the rise of approaches to machine learning we have the potential to find a solution to this problem.

3.3 Diabetes categories

Diabetes is an epidemic disease that occurs due to the decrease or absence of insulin in the body. There are different types of diabetes that usually are distinguished at diagnosis; so determine the type of diabetes is dependent on conditions that their disease manifest itself. Since the old division two types of diabetes is insulin-dependent and non-insulin dependent, the new classification of diabetes was developed by the America Diabetes Association: Type I diabetes, type II, gestational diabetes and other types.

Type 1 diabetes

Type 1 diabetes (previously known as insulin-dependent, juvenile or childhood-onset) is characterized by deficient insulin production and requires daily administration of insulin. In 2017 there were 9 million people with type 1 diabetes; the majority of them live in high-income countries. Neither its cause nor the means to prevent it are known.

Type 2 diabetes:

Type 2 diabetes (adult diabetes or Non-insulin-dependent diabetes), is one of the most common types of diabetes and constitutes about 90 percent of the patient. Unlike type 1 diabetes, the body produces insulin in type 2 diabetes, but produced insulin by the pancreas is not enough or body does not use insulin properly. When there is not enough insulin or the body does not use insulin, glucose (sugar) in the body, cannot move to the body's cells and causes to an accumulation of glucose in the body and the body would be in trouble and deficiencies. Unfortunately, there is no cure for this disease, but with the healthy diet, exercise and keep fit, can enhance it. If diet and exercise are not enough, you need medication or insulin treatment.

According to the Diabetes Center Research that has shown that with early diagnostics of patients at risk can be prevented 80 percent of chronic complications of type II diabetes or delayed them [23]. Delay in diagnosis and prediction of diabetes due to inadequate control of blood glucose increases Capillaries and macrovascular complications risk, ocular diseases and Kidney failure [22, 21]. So proposed a model to predict diabetes that doctors can be useful as a model to help predict diabetes. In this research, examined the relationship between complications in diabetic patients and their properties such as blood glucose, blood pressure, height, weight, and hemoglobin and body mass index of the patients. Based on the

Diabetes Research Center reports, the incidence of diabetes has doubled in the last ten years in the worldwide and About 200 million people are infected and about six percent increase in the annual prevalence of diabetes in the world. Since diabetes is a chronic disease and import irreparable damage to the limbs and vital organs in the body, using intelligent tools can improve detection methods and disease control and is a great help to the doctors.

A technique called, Predictive Analysis, incorporates a variety of machine learning algorithms, data mining techniques and statistical methods that uses current and past data to find knowledge and predict future events. By applying predictive analysis on healthcare data, significant decisions can be taken and predictions can be made. Predictive analytics can be done using machine learning and regression technique. Predictive analytics aims at diagnosing the disease with best possible accuracy, enhancing patient care, optimizing resources along with improving clinical outcomes [37]. Machine learning is considered to be one of the most important artificial intelligence features supports development of computer systems having the ability to acquire knowledge from past experiences with no need of programming for every case. Machine learning is considered to be a dire need of today's situation in order to eliminate human efforts by supporting automation with minimum flaws. Existing method for diabetes detection is uses lab tests such as fasting blood glucose and oral glucose tolerance. However, this method is time consuming [14].

3.4 Artificial Neural Network

Adaptive Artificial Neural Network is a non-parametric method to classify that in the medical field based on input variables to classify subjects into sick or healthy. Classification and prediction of the patient's condition based on risk factors are an application of artificial neural networks [24]. In artificial neural networks is inspired by the mixed structure of the human brain. Billion nerve cells (neurons) through the communication that with each other (synapses) creates a biological neural network in the human brain that is dedicated to human activities such as reading, comprehension, speaking, breathing, movement, voice recognition, face detection, also resolve issues and data storage. Artificial neural networks, in fact, simulate a part of brain functions [25.24].

(ANNs), or more simply neural networks, are new systems and computational methods for machine learning, knowledge demonstration, and finally the application of knowledge gained to maximize the output responses of complex systems [26].

An Artificial Neural Network (ANN) is a data processing model based on the way biological nervous systems, such as the brain, process data. They're focused on the neuronal structure of the mamalian cerebral cortex, but at a much smaller scale.

Many artificial intelligence experts believe that artificial neural networks are the best and perhaps the only hope for designing an intelligent machine. The branches and sections of the computational methods as well as Artificial Neural Networks in the chart is shown in the figure (2).

Artificial neural networks are designed in the same way as the human brain, with neuron nodes interconnected in a web-like fashion. Neurons are billions of cells that make up the human brain. Each neuron is made up of a cell body that processes information by bringing it to and from the brain (inputs and outputs) [27]. The main idea of such networks is (to some extent) inspired by the way the biological neural system works, to process data, and information in order to learn and create knowledge. The key element of this idea is to create new structures for the information processing system. The Artificial neural network architecture is shown in the figure 4 [28].

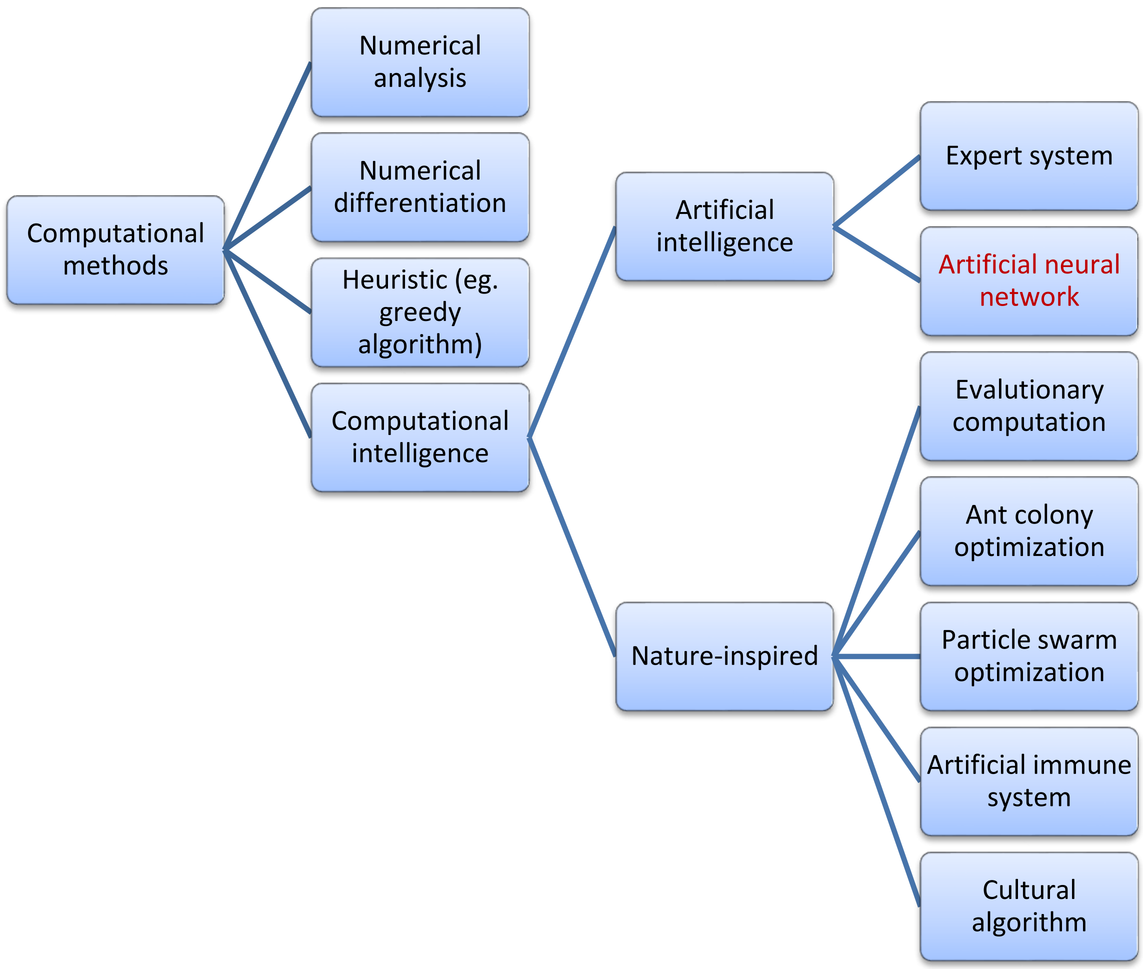

Figure 2: The branches and sections of the computational methods [28].

3.5 Artificial neural network structure

Neural networks are nonlinear modeling of intelligent computational techniques which in recent years as advances in computing and information processing tools obtained an important and advances position in science, and the results have been favorable. Feedforward neural networks, are useful type of artificial neural networks, because feedforward neural network with a hidden layer, suitable activation function in the hidden layer and the enough hidden layer neurons are able to approximate any function with arbitrary precision. In general, the artificial neural networks are three types of neuronal layers as each layer is as follows:

1. Input layer: Get the raw data that has been fed to the network.
2. Hidden layers: the function of this layers is determined by inputs and weight and the relationship between them and the hidden layers. Weights between input and hidden units determine when a hidden unit to be activated.
3. Output layer: output unit function depending on activity and weight of the hidden unit and the connection between hidden units and output.

The system is made up of a large number of highly interconnected processing elements called neurons that work together to solve a problem and transmit information through synapses (electromagnetic connections). The neurons are interconnected closely and organized into layers. The input layer receives the data, while the output layer generates the final result. Between the two, one or more secret layers are typically sandwiched. This arrangement makes predicting or knowing the exact flow of data difficult. Each connection has a connection weight, and each neuron has a threshold value and an activation function [29]. It is calculated if each input has a positive or negative weight based on the sign of the input's weight. The weight affects the signal intensity at a connection [30]. Neurons which have a threshold above which a signal is only transmitted if the aggregate signal exceeds it. The Activation Value is the weighted sum of the summing unit, and the output is generated based on the signal from this activation value.

In these networks, if one cell is damaged, other cells can make up for its absence and contribute to its regeneration. These networks are capable of learning. Basically, the ability to learn is the most important feature of an intelligent system. A learning system is more flexible and easier to program, so it can better respond to new problems and equations. Artificial neural networks, like humans, learn by using the different examples, and a neural network is set up to perform specific tasks, such as identifying patterns and categorizing information, during a learning process. For example, by injecting tactile nerve cells, the cells learn not to go to the hot body, and with this algorithm, the system learns to correct its error.

Artificial neural networks are increasingly used in the control or modeling of systems that have unknown or very complex internal structures. For example, a neural network can be used to control the input of an engine, in which case the neural network itself will learn the control function. Learning in these systems is adaptive, that is, using parables, the weight of the synapses changes in such a way that the system produces the correct response if new inputs are given [31]. A neural network is given a set of inputs and their corresponding outputs when it is being trained (using one of the training methods).

3.6 Artificial Neural Network topologies:

There are two Artificial Neural Network topologies **as FeedForward** and **Feedback**. In the feed forward systems, there are no feedback loops, because a unit sends information to another unit from which it receives none. Inputs and outputs are fixed. Each unit receives input information from its units on the left, and the inputs are multiplied by the weight of each connections. So, the output results related to the weight of each connections can be obtained. Pattern generation, identification, and classification are some application of the method.

In **the Feed Back** ANN systems, content addressable memories are used. Learning neural networks using a feedback process is by comparing the output of a network with the output that is desired and expected. The difference between these two outputs is used to change and modify the weights of the connections between the network units. [32].

3.7 Machine Learning:

Algorithms Types Machine learning algorithms are organized into taxonomy, based on the desired outcome of the algorithm. Common algorithm types include:

Supervised learning: where the algorithm generates a function that maps inputs to desired outputs. One standard formulation of the supervised learning task is the classification problem: the learner is required to learn (to approximate the behavior of) a function which maps a vector into one of several classes by looking at several input-output examples of the function.

Unsupervised learning: which models a set of inputs: labeled examples are not available.

Semi-supervised learning: which combines both labeled and unlabeled examples to generate an appropriate function or classifier.

Reinforcement learning: where the algorithm learns a policy of how to act given an observation of the world. Every action has some impact in the environment, and the environment provides feedback that guides the learning algorithm.

Transduction: similar to supervised learning, but does not explicitly construct a function: instead, tries to predict new outputs based on training inputs, training outputs, and new inputs.

Learning to learn: where the algorithm learns its own inductive bias based on previous experience.

Supervised Learning Approach is fairly common in classification problems because the goal is often to get the computer to learn a classification system that we have created. Digit recognition, once again, is a common example of classification learning. More generally, classification learning is appropriate for any problem where deducing a classification is useful and the classification is easy to determine. In some cases, it might not even be necessary to give predetermined classifications to every instance of a problem if the agent can work out the classifications for itself.

Unsupervised learning, all the observations are assumed to be caused by latent variables, that is, the observations is assumed to be at the end of the causal chain.

Supervised learning the most is the most common technique for training neural networks and decision trees. Both of these techniques are highly dependent on the information given by the pre-determined classifications in the case of neural networks, the classification is used to determine the error of the network and then adjust the network to minimize it, and in decision trees, the classifications are used to determine what attributes provide the most information that can be used to solve the classification puzzle.

Inductive machine learning is the process of learning a set of rules from instances (examples in a training set), or more generally speaking, creating a classifier that can be used to generalize from new instances.

The first step is collecting the dataset. If a requisite expert is available, then s/he could suggest which fields (attributes, features).

For instance, many elements used in the objective function of a learning algorithm may assume that all features are centered on zero or have variance in the same order. If a feature has a variance that is orders of magnitude larger than others, it might dominate the objective function and make the estimator unable to learn from other features correctly as expected [36].

3.8 Deep learning ANN systems:

It should be more complex in order to represent more complex features and "read" increasingly complex models for prediction and classification of data based on thousands or even millions of features. Deep learning is a machine learning subfield that focuses on learning successive "layers" of increasingly meaningful representations while learning representations from data [33]. It is concerned with artificial neural networks (ANNs), which are algorithms based on the structure and function of the brain. Deep learning allows computational models with multiple processing layers to learn multiple levels of abstraction for data representations. They are neural networks that have more than three layers of neurons (including the input and output layers). These layered representations are learned using models known as "neural networks," which are organized into literal layers that are placed one on top of the other [34]. Simply increasing the number of hidden layers and/or the number of neurons per hidden layer accomplishes this. More layers and neurons can represent increasingly complex models, but they also require more time and power to compute. The architecture of Deep learning technologies is shown in the figure 7 [35]. Deep learning technologies and their possible advantages/benefits are various. Using self-driving cars as an example, the “need to train a machine to take over key parts (or all) of driving is addressed.

3.9 Preprocessing data:

For various reasons, many real world datasets contain missing values, often encoded as blanks, NaNs or other placeholders. Such datasets however are incompatible with scikit-learn estimators which assume that all values in an array are numerical, and that all have and hold meaning. A basic strategy to use incomplete datasets is:

Data transformations (standardization): Standardization of datasets is a common requirement for many machine learning estimators. In practice we often ignore the shape of the distribution and just transform the data to center it by removing the mean value of each feature, then scale it by dividing non-constant features by their standard deviation.

Making a data set with mean=0, and standard deviation =1. This scaling method is useful when the data follows a normal distribution (Gaussian distribution), if the data does not follow normal distribution then this will make problems.

Example: -20, -6, 0, 40, 70,120

$$\text{Mean} = \frac{-20 + \ -6 + 0 + 40 + 70 + 120}{6} = 34 \qquad (1)$$

$$sd = \sqrt{\frac{(-20-34)^2 + (-6-34)^2 + (0-34)^2 + (40-34)^2 + (70-34)^2 + (120-34)^2}{6}}$$

$$sd = \ 48.98979$$

z-score standardization:

$$\text{x}'' = \frac{\text{x mean}}{\text{sd}} \qquad (2)$$

$$= \frac{-20-34}{48.98979} = \ -1.1022$$

Other values are changed too,

Accordingly, values are changed to:

-1.10227

-0.8165

-0.69402

0.122474

0.734847

1.755468

Now, if you calculate the average and sd of these new values you will see that the mean is zero and sd=1. [37]

3.10 Neural network model (supervised)

The Supervised Learning/Predictive Models Supervised learning algorithms are used to construct predictive models. A predictive model predicts missing value using other values present in the dataset. Supervised learning algorithm has a set of input data and also a set of output, and builds a model to make realistic predictions for the response to new dataset. Supervised learning includes Decision Tree, Bayesian Method, Artificial Neural Network, Instance based learning, Ensemble Method. These are booming techniques in Machine learning [38].

3.11 Multi-layer Perceptron (MLP):

Is a supervised learning algorithm that learns a function :

$$f(\cdot):Rm \rightarrow Ro \qquad (3)$$

by training on a dataset, where (m) is the number of dimensions for input and (o) is the number of dimensions for output. Given a set of features $X = x_1, x_2, \ldots . x_m$ and a target y, it can learn a non-linear function approximator for either classification or regression. It is different from logistic regression, in that between the input and the output layer, there can be one or more non-linear layers, called hidden layers. Figure 3 shows a one hidden layer MLP with scalar output.

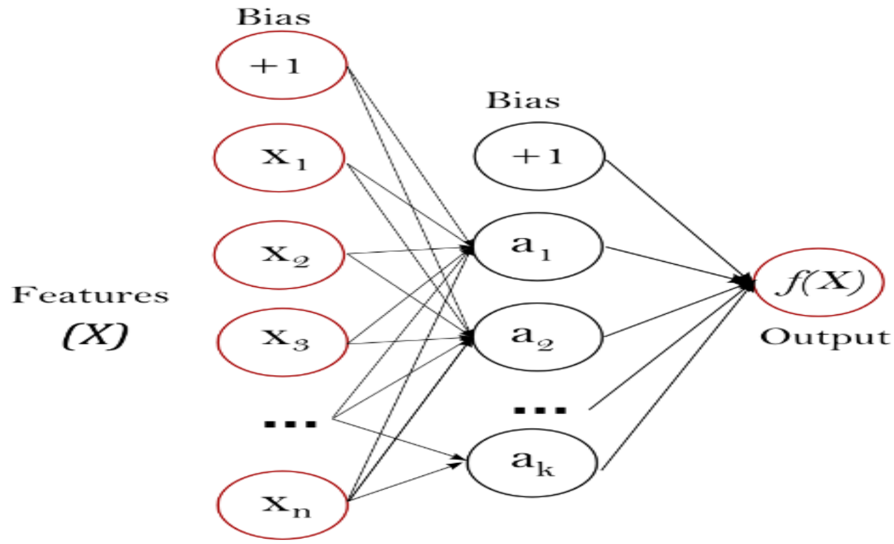

Figure 3: One hidden layer MLP. [38]

The leftmost layer, known as the input layer, consists of a set of neurons $\{xi|x_1, x_2, \ldots, x_m\}$ representing the input features. Each neuron in the hidden layer transforms the values from the previous layer with a weighted linear summation $w_1 x_1 + w_2 x_2 + \cdots + w_m x_m$, followed by a non-linear activation function $g(\cdot): R \rightarrow R$ like the hyperbolic tan function. The output layer receives the values from the last hidden layer and transforms them into output values.

The module contains the public attributes coefs_ and intercepts_. coefs_ is a list of weight matrices, where weight matrix at index i represents the weights between layer i and layer i+1. Intercepts_ is a list of bias vectors, where the vector at index i represents the bias values added to layer i+1.

The advantages of Multi-layer Perceptron are:

- Capability to learn non-linear models.
- Capability to learn models in real-time (on-line learning) using partial_fit.

The disadvantages of Multi-layer Perceptron (MLP) include:

- MLP with hidden layers have a non-convex loss function where there exists more than one local minimum. Therefore different random weight initializations can lead to different validation accuracy.
- MLP requires tuning a number of hyperparameters such as the number of hidden neurons, layers, and iterations.
- MLP is sensitive to feature scaling.

3.12 Classification:

Class MLPClassifier implements a multi-layer perceptron (MLP) algorithm that trains using Backpropagation.

MLP trains on two arrays: array X of size (n_samples, n_features), which holds the training samples represented as floating point feature vectors; and array y of size (n_samples,), which holds the target values (class labels) for the training sample.

3.13 Mathematical formulation

Given a set of training examples $(x_1, y_1)(x_2, y_2), \dots \dots (x_n, y_n)$ where

$$xi \in R^n$$

and $y_i \in \{0,1\}$ a one hidden layer one hidden neuron MLP learns the function:

$$f(x) = W_{2g}(W_1^T + b_1\ ) \qquad (4)$$

Where $W_1 \in R_m$ and $W_2, b_1, b_2 \in R$ are model parameters.

$w1, w2$ Represent the weights of the input layer and hidden layer, respectively; and $b1, b2$ represent the bias added to the hidden layer and the output layer, respectively.

$$g(0): R \rightarrow R \qquad (5)$$

is the activation function, set by default as the hyperbolic tan. It is given as,

$$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \qquad (6)$$

For binary classification, f(x) passes through the logistic function

$$g(z) = {}^{1}/_{(1 + e^{-z})} \qquad (7)$$

to obtain output values between zero and one. A threshold, set to 0.5, would assign samples of outputs larger or equal 0.5 to the positive class, and the rest to the negative class.

If there are more than two classes, $f(x)$ it self would be a vector of size (n_classes,). Instead of passing through logistic function, it passes through the softmax function, which is written as,

$$\mathrm{softmax}(z)_i = \frac{exp(z_i)}{\sum_{i=1}^{k} exp(z_i)} \qquad (8)$$

where z_i represents the i th element of the input to $softmax$, which corresponds to class i, and K is the number of classes. The result is a vector containing the probabilities that sample x belong to each class. The output is the class with the highest probability.

In regression, the output remains as f(x); therefore, output activation function is just the identity function.

MLP uses different loss functions depending on the problem type. The loss function for classification is Average Cross-Entropy, which in binary case is given as,

$$loss(\hat{y}, y, W) = -\frac{1}{n} \sum_{i=0}^{n} (y_1 \ln \hat{y}_i + (1 - y_i) \, ln \, \hat{y}_i \, (1 - y_i) + \frac{\alpha}{2n} ||W||_2^2 \qquad (9)$$

Where:

$\alpha||W_2^2$ is an L2-regularization term (aka penalty) that penalizes complex models; and α>0 is a non-negative hyperparameter that controls the magnitude of the penalty.

For regression, MLP uses the Mean Square Error loss function; written as,

$$Loss(\hat{y}, y, W) = \frac{1}{2n} \sum_{i=0}^{n} | |\hat{y}_i - y_i||_2^2 + \frac{\alpha}{2n} ||W||_2^2 \qquad (10)$$

Starting from initial random weights, multi-layer perceptron (MLP) minimizes the loss function by repeatedly updating these weights. After computing the loss, a backward pass propagates it from the output layer to the previous layers, providing each weight parameter with an update value meant to decrease the loss.

In gradient descent, the gradient $\nabla Loss_W$ of the loss with respect to the weights is computed and deducted from W. More formally, this is expressed as,

$$W^{i+1} = W^i - \epsilon \nabla Loss_W^i \qquad (11)$$

where i is the iteration step,

and ϵ is the learning rate with a value larger than 0.

The algorithm stops when it reaches a preset maximum number of iterations; or when the improvement in loss is below a certain, small number.

Chapter Four

Methodology

4.1 Methodology

This proposed model for Diabetes Prediction by using artificial neural network. MLP (Multi Layer Perceptron) Classifier implemented by sklearn.neural_network libraries in python. MLPClassifier is used to predict whether a patient has diabetes based on a set of diagnostics. The figure 9 below show the Flow diagram of proposed model.

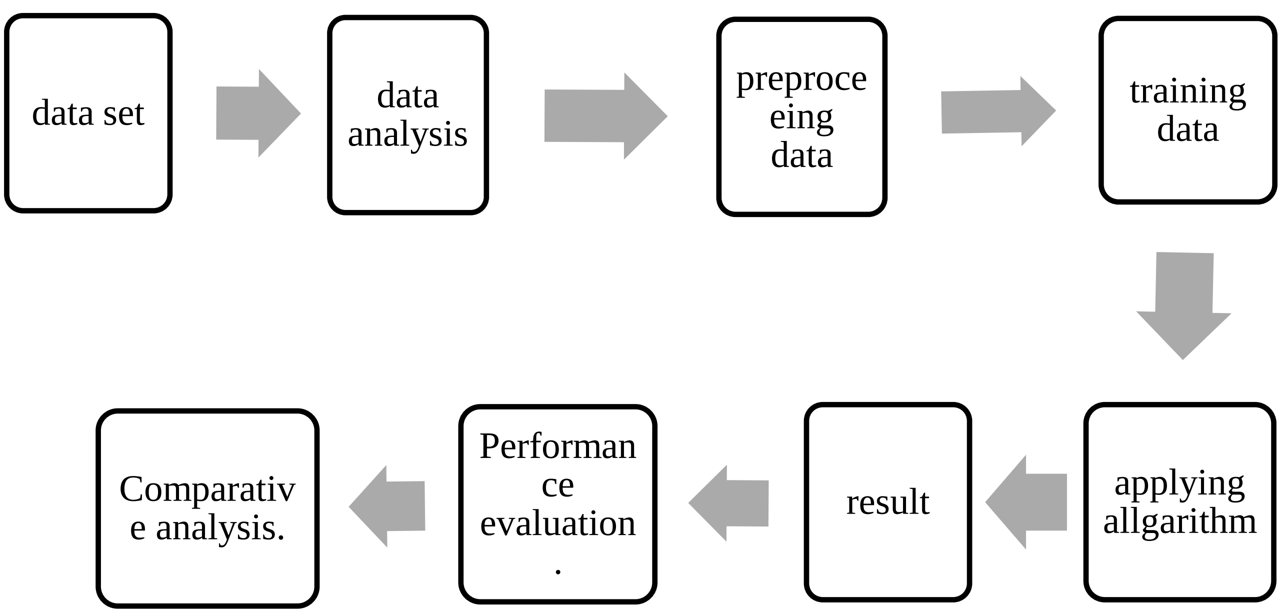

Figure 4: Flow diagram of proposed model

4.2 Data set:

This dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective of the dataset is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. Several constraints were placed on the selection of these instances from a larger database. In particular, all patients here are females at least 21 years old of Pima Indian heritage.

[https://www.kaggle.com/code/ahmetcankaraolan/diabetes-prediction-using-machine-learning/data]

Details about the dataset:

The datasets consists of several medical predictor variables and one target variable, Outcome. Predictor variables includes the number of pregnancies the patient has had, their BMI, insulin level, age, and so on.

Pregnancies: Number of times pregnant

Glucose: Plasma glucose concentration a 2 hours in an oral glucose tolerance test.

Blood Pressure: Diastolic blood pressure (mm Hg)>

BMI: Body mass index (weight in kg/ (height in m) ^2

Diabetes Pedigree Function: Diabetes pedigree function

Age: Age (years)

Outcome: Class variable (0 or 1)

Number of Observation Units: 768

Variable Number: 7

The first step is to loaded data set to program (python) as CVS copy, and defined by name :('dataset.csv')

4.3 Data analysis:

The first 5 observation units of the data set were accessed as follow:

	Pregnancies	Glucose	BloodPressure	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	33.6	0.627	50	1
1	1	85	66	26.6	0.351	31	0
2	8	183	64	23.3	0.672	32	1
3	1	89	66	28.1	0.167	21	0
4	0	137	40	43.1	2.288	33	1
5	5	116	74	25.6	0.201	30	0

Figure 5: The first five observation units of the data set

- The classes of the outcome variable were examined.

0 500

1 268

- Dataset shape

(768, 7)

The distribution of the outcome variable in the data was examined and visualized:

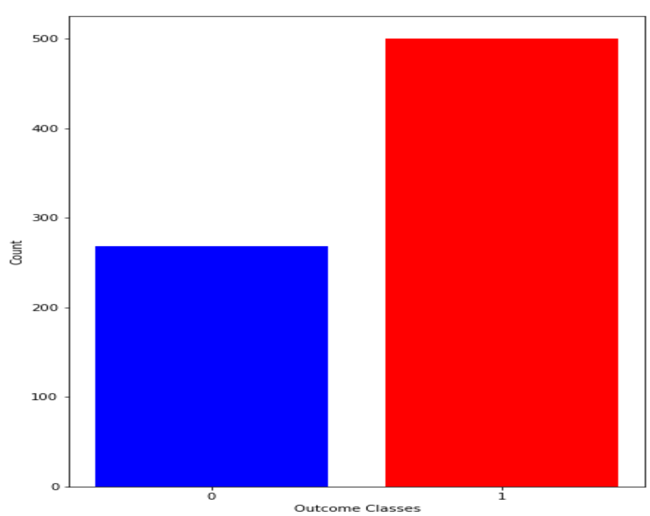

Figure 6: outcome classes

The histogram of the variables was reached:

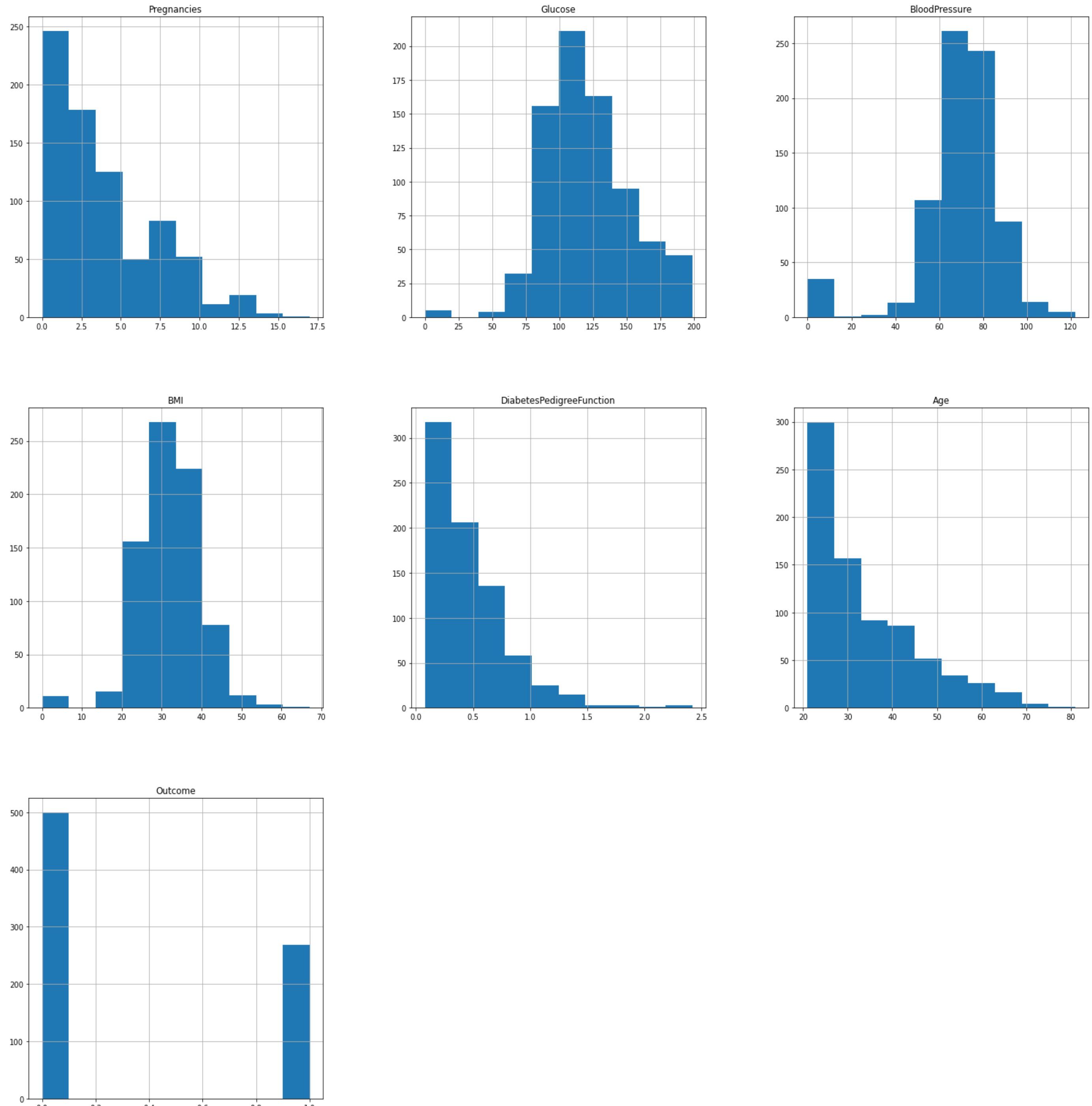

Figure 7: Histogram of the variables

4.4 Correlation matrix

Shows that there are mild correlations between **SkinThickness-BMI** and **Age-Pregnancies**. **Outcome** has the highest linear correlation with **Glucose.**

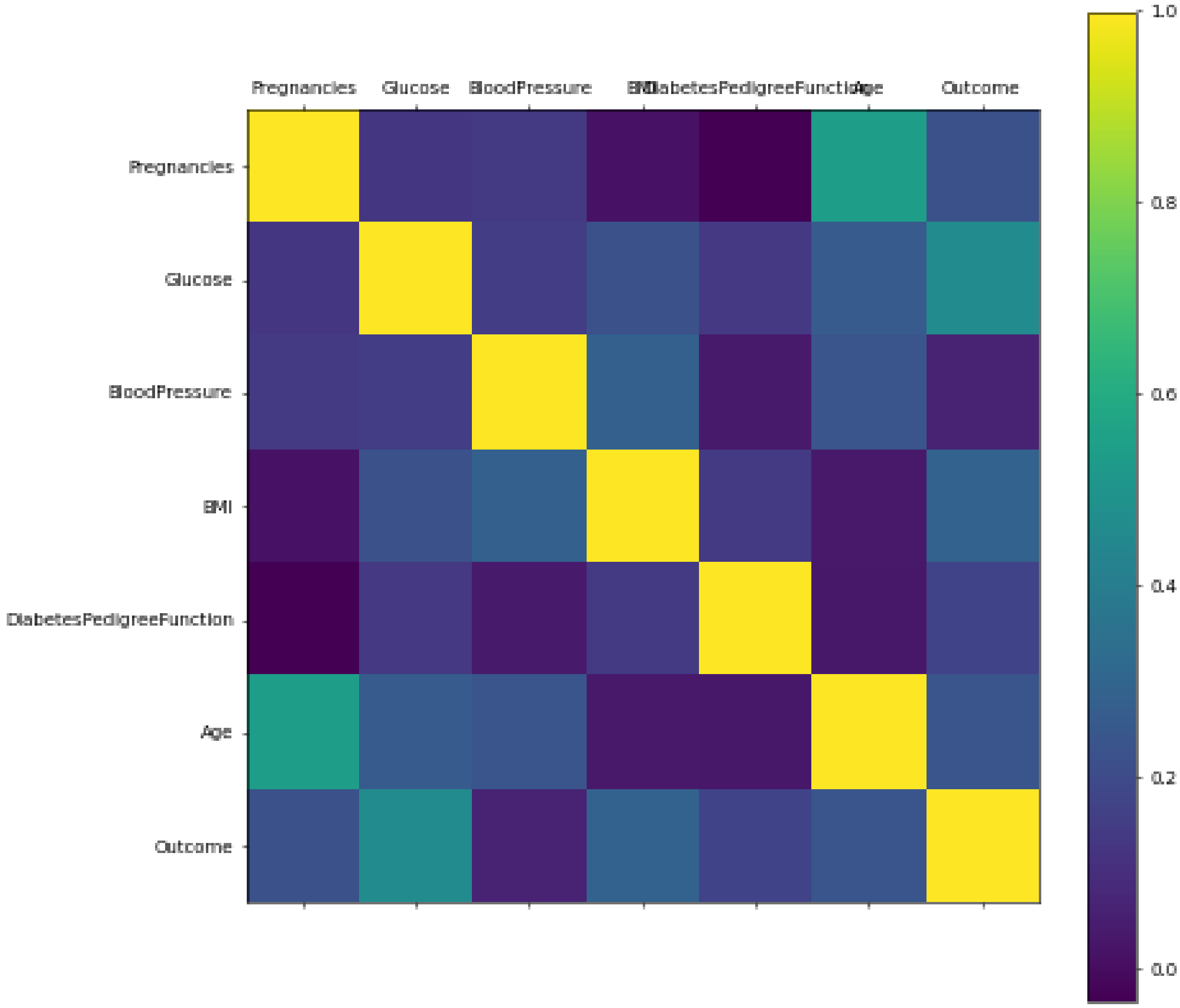

Figure 14: Correlation matrix

4.5 Split data:

In the dataset we defined the input variable as (x) and output variable as (y), and dataset is divided into train and test sets. We use stratify parameter of train_test_split function to get the same class distribution across train and test sets. Test data (20%) and train data (80%).

Test data 20% to:

- X_test
- Y_test

Train data 80% to:

- X_train.
- Y_train.

4.6 Preprocessing data:

1. A basic strategy to use incomplete datasets is Data reduction by visual inspection manually dismissed one units that's contain three missing data. Numbers of units that discarded is (428).
2. Data transformations (standardization): In practice we often ignore the shape of the distribution and just transform the data to center it by removing the mean value of each feature, then scale it by dividing non-constant features by their standard deviation.

Correlation Analysis:

Linear correlations between features and also between features and output are computed. **Pandas corr** function is used to compute correlation matrix and **Seaborn heatmap** is used for plotting.

4.7 Train data:

Multi-layer Perceptron (MLP) is a supervised learning algorithm used. From sklearn.neural_network import MLPClassifier, Gridsearch as follow:

Hyper_paramters:

batch_size: ["auto",100].

hidden_layer_sizes':[(5,5,5), (5), (6), (4)].

learning_rate_init:[0.001,0.005, 0.01].

max_iter:[500,600].

NN_Gridsearch_paramters:

Hyper_paramters,scoring='roc_auc'

n_jobs=-1

cv=5

return_train_score=False

verbose=0

NN_crossvalidation=NN_Gridsearch_paramters.fit(X_train,y_train).

Final model is obtained by NN cross validation best estimator and best AUC score was NN cross validation best score.

The variable summary contains all results of cross validation process in excel form. It appended at the end.

4.8 Test data:

y_pred = Final_Model.predict(X_test)

from sklearn.metrics import confusion_matrix to calculate it and classification report as target names=["Diabetes", "Healthy"]. Then the accuracy, sensitivity and specificity is calculated.

4.9 Applying algorithm

from sklearn.neural_network import MLPClassifier

Hyper_paramters={'batch_size':[auto,100],'hidden_layer_sizes':[(5)],

'learning_rate_init':[0.05],'max_iter':[600]}

NN_Gridsearch_paramters=GridSearchCV(MLPClassifier(),Hyper_paramters,scoring='roc_auc',n_jobs=-1, cv=5,return_train_score=False,verbose=0)

NN_crossvalidation=NN_Gridsearch_paramters.fit(X_train,y_train)

print ("The best paramter combination is ")

print(NN_crossvalidation.best_params_)

```
Final_Model=NN_crossvalidation.best_estimator_
print("The best AUC score was ")
print(NN_crossvalidation.best_score_)
```

Chapter Five

Results and discussion

5.1 The results

Table (5.1) contain six collections of hyper parameters selected randomly _ma nually_ to test the model and obtained their accuracy, sensitivity and specificity as follow:

Table 5.1: The result

NO/Parameters	Batch size	Hidden layer size	Learning _rat_init	Max _iter	Sensitivity	Accuracy	Specificity
1	100	5	0.005	500	0.9191919191919192	0.7792207792207793	0.5272727272727272
2	100	5	0.005	600	0.8888888888888888	0.8116883116883117	0.6727272727272727
3	100	5	0.05	500	0.8787878787878788	0.8051948051948052	0.6727272727272727
5	100	(5, 5, 5)	0.005	500	0.8585858585858586	0.753246753246753	0.5636363636363636

The best parameter combination is :{'batch_size': 100, 'hidden_layer_sizes': 5, 'learning_rate_init': 0.005, 'max_iter': 600}.

The best accuracy: 0.811688311688831.

5.2 Cross validation results:

The best hyper parameter combination obtained :{'batch_size': 100, 'hidden_layer_sizes': 5, 'learning_rate_init': 0.001, 'max_iter': 500}

The best AUC score : 0.8379627066158074

The best accuracy: 0.7662337662337663

Final model is obtained by NN cross validation best estimator, and best AUC score was NN cross validation best score.

The variable summary contains all results of cross validation process in excel form. It appended below.

5.3 Evaluation

This is the final step of prediction model. Here, we evaluate the prediction results using various evaluation metrics like classification accuracy, confusion matrix and f1-score.

5.4 Confusion matrix:

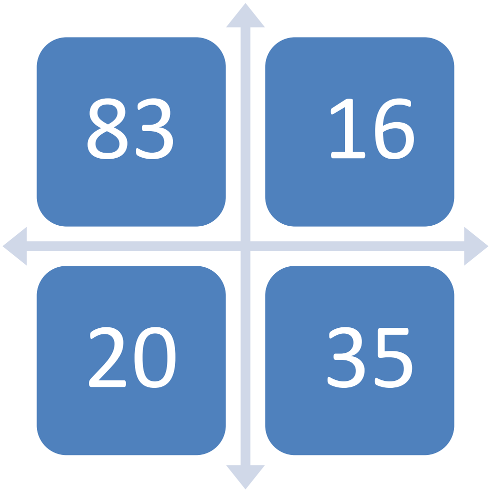

Table 5.2: Classification report

	precision	recall	f1-score	support
Diabetes	0.81	0.84	0.82	99
Healthy	0.69	0.64	0.66	55
accuracy	-	-	0.77	154
macro avg	0.75	0.74	0.74	154
weighted avg	0.76	0.77	0.76	154

Accuracy	0.7662337662337663
Sensitivity	0.8383838383838383
Specificity	0.6363636363636364

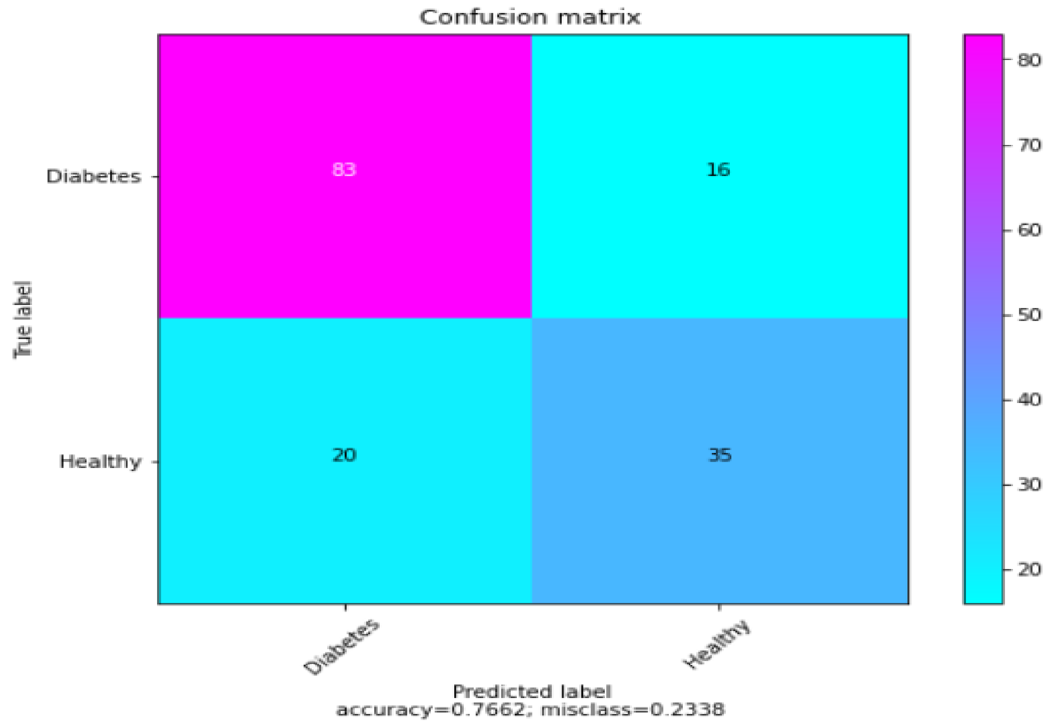

Figure 5: confusion matrix

5.5 Discussion:

The best parameter combination is :{'batch_size': 100, 'hidden_layer_sizes': 5, 'learning_rate_init': 0.005, 'max_iter': 600}.

The best accuracy: 0.81168831168831.

If we compared between accuracies in table 2 we found that's accuracy increased b y increasing of learning_rate_init and max_iter.

If we compare between table (2) and table (1) from chapter two we found that KM Jyoti Rani 99% accuracy by using Decision Tree algorithm. It is biter than our proposed model, because it used huge number of data and used decision tree.

Aishwarya Mujumdara, Dr. Vaidehi Vb used Logistic Regression gives highest accuracy of 96%

Chapter Six

Conclusions and recommendations

6.1 Conclusion:

The aim of this study was to create models for the diabetes prediction, by establishing models and to obtain maximum validation scores in the established models. The work done is as follows: Diabetes Data Set read. With Exploratory Data Analysis; the data set's structural data were checked. The types of variables in the dataset were examined. Size information of the dataset was accessed. Descriptive statistics of the data set were examined. Data Preprocessing section; The NaN values missing observations were filled with the median values of whether each variable was sick or not. The outliers were determined by LOF and dropped. The X variables were standardized. During Model Building; MLP classifier function used and Cross Validation Score were calculated, hyper parameters optimizations optimized to increase Cross Validation value. The model created as a result of NN cross validation best estimator and best AUC score. The best AUC score was 0.8261867335220048, Accuracy: 0.7662337662337663, Sensitivity: 0.8383838383838383, Specificity: 0.6363636363636364.

6.2 Recommendations:

The recommendation is to use huge number of data for more accuracy and build model by combined between neural network and fuzzy logic.

References

[1] Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019. Results. Institute for Health Metrics and Evaluation. 2020 (https://vizhub.healthdata.org/gbd-results/)

[2] WHO, fact sheet- details- Diabetes, www.who.com, 16 September 2022, (https://www.who.int/news-room/fact-sheets/detail/diabetes#:~:text=Overview,hormone%20that%20regulates%20blood%20glucose.

[3] KM Jyoti Rani **,** Diabetes Prediction Using Machine Learning, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, India,auga2020

[4] Rabina, Er. Anshu Chopra 'DIABETES PREDICTION BY SUPERVISED AND UNSUPERVISED LEARNING WITH FEATURE SELECTION'

[5] Rishab Bothra Department of Information Technology Motilal Nehru National Institute of Technology Allahabad, Prayagra.

[6] Falvo D, Holland BE. Medical and psychosocial aspects of chronic illness and disability. Jones & Bartlett Learning; 2017.

[7] Skyler JS, Bakris GL, Bonifacio E, Darsow T, Eckel RH, Groop L, et al. Differentiation of diabetes by pathophysiology, natural history, and prognosis. Diabetes 2017; 66:241–55. [CrossRef]

[8] Tao Z, Shi A, Zhao J. Epidemiological perspectives of diabetes. Cell Biochem Biophys 2015; 73:181–5. [CrossRef]

[9] https://www.who.int/news-room/fact-sheets/detail/diabetes

[10] Aditya Saxena, Megha Jain, Prashant Shrivastava, Data Mining Techniques Based Diabetes Prediction, Indian Journal of Artificial Intelligence and Neural Networking, April 2021.

[11] [Mitushi Soni and Dr. Sunita Varma, Diabetes Prediction using Machine Learning Techniques , International Journal of Engineering Research & Technology , September-2020.

[12] Olta Llahaa, Amarildo Ristab , Prediction and Detection of Diabetes using Machine Learning, South East European University, North Macedonia, Skopje, North Macedonia, may 2021.

[13] Nesreen Samer El-Jerjawi, Samy S. Abu-Naser, Diabetes Prediction Using Artificial Neural Network, International Journal of Advanced Science and Technology , Gaza, Palestine ,2018.

[14] Aishwarya Mujumdara, Dr. Vaidehi Vb, Diabetes Prediction using Machine Learning Algorithms, INTERNATIONAL CONFERENCE ON RECENT TRENDS IN ADVANCED COMPUTING, India, 2019.

[15] GeetaRani[a]Vijaypal, SinghDhaka[a]Ramesh, ChandraPoonia, Diabetes prediction using artificial neural network,[b]

[16] Parastoo RAHIMLOO, Prediction of Diabetes by Using Artificial Neural Network, Logistic Regression Statistical Model and Combination of Them, Department of Mathematics, Urmia Branch, Islamic Azad University, Urmia, Iran, 2016.

[17] Temurtas, H., Yumusak, N., Temurtas, F., "A comparative study on diabetes disease diagnosis using neural networks", Expert Syst, Vol. 36, pp. 8610–15, 2009.

[18] Nielsen, D., Krych, L., Buschard, K., "Beyond Genetics Influence of dietary factors and gut microbiota on type 1 diabetes", FEBS Lett, Vol.588, pp.4234–43, 2014.

[19] Pei, E., Li, J., Lu, C., Xu, J., Tang, T., Ye, M., et al, " Effects of lipids and lipoproteins on diabetic foot in people with type 2 diabetes mellitus: a meta-analysis", J Diabetes Complications, Vol. 28, pp.559–64, 2014.

[20] WHO Expert Committee on Diabetes Mellitus. Second Report. Geneva: WHO, 1980. Technical Report Series 646

[21] Chavey, A., Kioon, M., Bailbé, D., "programming of beta-cell disorders and intergenerational risk of type 2 diabetes Diabetes", Maternal Diabetes, Vol.40, No.5, pp. 323-30, 2014.

[22] Manzella, D., Grella, R., Abbatecola, AM., Paolisso, G., "Repaglinide Administration Improves Brachial Reactivity in Type 2 Diabetic Patients", Diabetes Care, Vol.28, pp. 366– 71, 2005.

[23] Mohamed, E. I., Linde,r R., Perriello, G., Di Daniele, N., Pöppl, S. J., De Lorenzo, A.,"Predicting type 2 diabetes using an electronic nose-based artificial neural network analysis", Diabetes nutrition & metabolism Vol.15, No.4, pp.222-215, 2002. [5] Pickup, J.C., Williams G., (Eds.). , Textbook of diabetes, Blackwell Science, Oxford, 2003

[24] Livingstone, D., Totowa, NJ, Artificial Neural Networks Methods and Application. 1th ed. Totowa, NJ: Hummana Press; 2008.

[25] Dunne, RA., Wiley, J., Inc, S.," A Statistical Approach to Neural Networks for Pattern Recognition", New Jersey: John Wiley & Sons Inc; 2007

[26] Chen, Mingzhe, Ursula Challita, Walid Saad, Changchuan Yin, and Mérouane Debbah. 2019. "Artificial neural networks-based machine learning for wireless networks: A tutorial." Review of. IEEE Communications Surveys & Tutorials 21 (4):3039-71.

[27] Van Gerven, Marcel, and Sander Bohte. 2017. "Artificial neural networks as models of neural information processing." Review of. Frontiers in Computational Neuroscience 11:114.

[28] Bre, Facundo, Juan M Gimenez, and Víctor D Fachinotti. 2018. "Prediction of wind pressure coefficients on building surfaces using artificial neural networks." Review of. Energy and Buildings 158:1429-41.

[29] Balakrishnan, Harikrishnan Nellippallil, Aditi Kathpalia, Snehanshu Saha, and Nithin Nagaraj. 2019. "ChaosNet: A chaos based artificial neural networkarchitecture for classification." Review of. Chaos: An Interdisciplinary Journal of Nonlinear Science 29 (11):113125.

[30] Liu, Jinjin, Yongchun Chen, Li Lan, Boli Lin, Weijian Chen, Meihao Wang, Rui Li, Yunjun Yang, Bing Zhao, and Zilong Hu. 2018. "Prediction of rupture risk in anterior communicating artery aneurysms with a feed-forward artificial neural network." Review of. European radiology 28 (8):3268-75.

[31] Wu, Yu-chen, and Jun-wen Feng. 2018. "Development and application of artificial neural network." Review of. Wireless Personal Communications 102 (2):1645-56.

[32] Mitra, Dipankar, and Ranjit Kumar Paul. 2017. "Hybrid time-series models for forecasting agricultural commodity prices." Review of. Model Assisted Statistics and Applications 12 (3):255-64.

[33] Grekousis, George. 2019. "Artificial neural networks and deep learning in urban geography: A systematic review and meta-analysis." Review of. Computers, Environment and Urban Systems 74:244-56.

[34] Schmidhuber, Jürgen. 2015. "Deep learning in neural networks: An overview." Review of. Neural Networks 61:85-117.

[35] Santos, Iria, Luz Castro, Nereida Rodriguez-Fernandez, Alvaro Torrente-Patino, and Adrian Carballal. 2021. "Artificial Neural Networks and Deep Learning in the Visual Arts: A review." Review of. Neural Computing and Applications:1-37.

[36] [https://scikit-learn.org/stable/modules/preprocessing.html].

[37] Gauri D. Kalyankar, Shivananda R. Poojara and Nagaraj V. Dharwadkar," Predictive Analysis of Diabetic Patient Data Using Machine Learning and Hadoop", International Conference On I-SMAC,978-1-5090-3243-3,2017.

[38] B. Nithya and Dr. V. Ilango,” Predictive Analytics in Health Care Using Machine Learning Tools and Techniques”, International Conference on Intelligent Computing and Control Systems, 978-1-5386-2745-7, 2017