WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Modelling and Design of Artificial Intelligent based Patient Monitoring System for Measuring Vital Parameters for Diabetes Mellitus Prognosis

Achebe, P.N.[a], Akpado, K. A.[b], Obioma, P. C.[b]

[a] *Department of Electrical and Electronic Engineering, Chukwuemeka Odumegwu Ojukwu University, P.M.B. 02, Uli, Nigeria*
[b] *Department of Electrical and Electronic Engineering, Electronic and Computer Engineering, Nnamdi Azikiwe University, Awka, Nigeria*

A B S T R A C T

This paper has presented modelling and design of artificial intelligent based patient monitoring (PM) system for measuring vital parameters for diabetes mellitus prognosis. The study carried out concise study of diabetes and artificial intelligence (AI) model based on machine learning (ML) algorithm of Fully connected Neural Network (FNN). The paper augmented data from four different databases. Data pre-processing was carried as a critical aspect of the model training process, which involves the preparation and subsequent transformation of data set. Learning classification architecture was designed with four convolution blocks that generates the feature maps, which was for model building and involves reshaping of data into the format that can be input into the linear classifier layer that finally outputs the predictions for the two classes. The model training used was the stochastic gradient descent approach. The model was validated by testing the dataset from trained dataset by randomly picking 20% of the data in each of the database employed. The model was implemented as part of the microcontroller. The system uses six core parts for power supply purposes (using DC battery source), data gathering (using wearable blood pressure sensor, temperature sensor, and ECG sensor), data processing (using microcontroller unit), and outputting of information (or processed data using LCD display unit). The proposed system works in such a way that it reduces computational complexity in a rather simple but efficient approach to the diagnosis of diabetes mellitus.

Keywords: Artificial intelligent, Diabetes mellitus, Machine learning, Neural network, Patient monitoring system

1. Introduction

Artificial intelligence (AI) is the process of gathering data, process data, by making necessary predictions and produce a well-defined output to the end user using machine learning algorithms and software. Principles of machine learning have been used to build algorithms to support predictive models for the risk of developing diabetes or its consequent complications. Digital therapeutics have proven to be an established intervention for lifestyle therapy in the management of diabetes. Patients are increasingly being empowered for self-management of diabetes, and both patients and health care professionals are benefitting from clinical decision support. AI allows a continuous and burden-free remote monitoring of the patient's symptoms and biomarkers.

Despite the rapid development of science and technology in healthcare regarding AI applications, diabetes remains an incurable lifelong illness. Diabetes mellitus (DM) is a long-term chronic disease. According to recently published data from the International Diabetes Federation, the number of people worldwide suffering from diabetes has reached 451 million. Globally the prevalence of diabetes was estimated to be 463 million in 2019 and has been projected to reach 578 million in 2030 and 700 million by 2045. Most increases in the prevalence of diabetes occur in low- and middle-income countries of the world including Nigeria which has the largest number of people living with diabetes in the sub-Saharan region. In Nigeria diabetes is evidently becoming a national problem and the difficulties of adequately treating all groups within the population are well known. Due to the limited number of diabetes specialists required in the management of diabetes and its myriad complications, many of the individuals with diabetes who developed complications are not adequately managed especially those who reside in rural or remote places.

Advances in digital technology and especially mobile smartphone technology have led to a plethora of innovative strategies aiming to improve the self-management skills of patients with chronic diseases and especially diabetes. Diabetes is associated with various complications and a significant morbidity and mortality. It is important to intervene not only to treat but also to monitor, prevent and make a timely detection of diabetes. Management of diabetes is challenging because 1 of 2 adults with diabetes are undiagnosed, yet 10% of global health expenditure (US$760 billion) are spent on diabetes.

In 1996, Gruman and Von Korff defined it as “engaging in activities that protect and promote health, monitoring and managing the symptoms and signs of illness, managing the impact of illness on functioning emotions and interpersonal relationships and adhering to treatment regimens. A more basic definition, by Nease et al, was offered by Lorig and Holman in 2003: “patients’ efforts to engage in behaviors to manage their chronic illness. Indeed, self-management programs on the whole have been successful in improving outcomes regarding serious chronic illnesses such as diabetes and hypertension. This research will focus on designing patients’ self-management tool which comprises of instruments of self-care, mobile and home care

using a proto type device to measure vital parameters like Electrocardiogram (ECG), temperature etc. with Artificial intelligence based health monitor, they help patients deal with their own medical conditions, or those of their loved ones, outside the walls of formal institutions.

Nomenclature

2. Overview of diabetes and artificial intelligence in medicine

The term diabetes is the shortened version of the full name diabetes mellitus. Diabetes mellitus is derived from the Greek word diabetes meaning siphon - to pass through and the Latin word mellitus meaning honeyed or sweet. This is because in diabetes excess sugar is found in blood as well as the urine. It was known in the 17th century as the "pissing evil" (GizemKoca, 2020).

The term diabetes was probably coined by Apollonius of Memphis around 250 BC. Diabetes is first recorded in English, in the form diabetes, in a medical text written around 1425. It was in 1675 that Thomas Willis added the word "'mellitus'" to the word diabetes. This was because of the sweet taste of the urine. This sweet taste had been noticed in urine by the ancient Greeks, Chinese, Egyptians, Indians, and Persians as is evident from their literature (GizemKoca, 2020).

Approximately all healers or medical practitioners noted that the disease was more common in heavy, wealthy people of the population more than other parts of the community. The inference was based on the fact that wealthiest people ate more than healthy people and are less active. In 1916, Boston scientist Elliot Joslin established himself as one of the world's leading diabetes experts, and reported that fasting diet combined with regular exercise could significantly reduce the risk of death in diabetes patients (GizemKoca, 2020). After discovering the disease, the search for the treatment has started. The Greek physicians suggested that the treatment of the disease was an exercise on horseback, preferably, and they believed that this would decrease the excessive urination needs. The modern treatment methodologies are to use animal-based insulin injections and other oral medicines, such as met forming (GizemKoca, 2020).

Studies conducted in Nigeria indicated that the prevalence of diabetes ranged from low level of 0.8% among adults in rural highland dwellers to over 7% in urban Lagos with an average of 2.2% nationally. As already pointed out, the sixth edition of IDF diabetes Atlas, shows that Nigeria is the leading country in Africa in terms of the number of people with diabetes, 3.9 million had diabetes with 105,091 diabetes-related deaths in 2013 which is estimated to increase annually by 125,000 between 2010 and 2030 even though the prevalence of 4.99% is far less than that of Reunion (15.38%), Seychelles (12.11%), Gabon (10.71%), Zimbabwe (9.73%), and South Africa (9.27%); in addition, there are still about 1.8 million Nigerians with undiagnosed diabetes in 2013 (GizemKoca, 2020).

2.1 Types of diabetes Mellitus

- Prediabetes

Prediabetes is a condition in which blood glucose levels are too high to be considered normal but not high enough to be labeled diabetes. Pre-diabetes is a stage, which patients have a high risk of developing diabetes mellitus. The term refers to Impaired Fasting Glucose (IFG) and Impaired Glucose Tolerance (IGT) (Goldenberg &Punthakee, 2013), (GizemKoca, 2020). Pre-diabetes is a condition of elevated blood glucose, including impaired fasting glucose and impaired glucose tolerance, which often precedes the onset of type 2 diabetes mellitus (Watson, 2017). All patients, who suffer from pre-diabetes, do not have to develop diabetes mellitus and disease complications in the future. If the patients with pre-diabetes are not treated properly, there is a considerable chance to suffer from type 2 diabetes mellitus within five years of onset. Type 2 diabetes mellitus development and pre-diabetes progress can be prevented with changing the lifestyle and receiving a proper treatment (GizemKoca, 2020).

People have prediabetes if their fasting blood glucose level is between 100 mg/dL (5.6 mmol/L) and 125 mg/dL (6.9 mmol/L) or if their blood glucose level 2 hours after a glucose tolerance test is between 140 mg/dL (7.8 mmol/L) and 199 mg/dL (11.0 mmol/L). Prediabetes carries a higher risk of future diabetes as well as heart disease. Decreasing body weight by 5 to 10% through diet and exercise can significantly reduce the risk of developing diabetes (Goldenberg &Punthakee, 2013).

- Type 1 Diabetes

In type 1 diabete (formerly called insulin-dependent diabetes or juvenile-onset diabetes), the body's immune system attacks the insulin-producing cells of the pancreas, and more than 90% of them are permanently destroyed. The pancreas, therefore, produces little or no insulin. Only about 5 to 10% of all people with diabetes have type 1 disease. Most people who have type 1 diabetes develop the disease before age 30, although it can develop later in life (Goldenberg &Punthakee, 2013). Scientists believe that an environmental factor—possibly a viral infection or a nutritional factor during childhood or early adulthood—causes the immune system to destroy the insulin-producing cells of the pancreas. A genetic predisposition makes some people more susceptible to an environmental factor.

- Type 2 Diabetes

In type 2 diabetes (formerly called non–insulin-dependent diabetes or adult-onset diabetes), the pancreas often continues to produce insulin, sometimes even at higher-than-normal levels, especially early in the disease. However, the body develops resistance to the effects of insulin, so there is not enough insulin to meet the body's needs. As type 2 diabetes progresses, the insulin-producing ability of the pancreas decreases. Type 2 diabetes was once

rare in children and adolescents but has become more common. However, it usually begins in people older than 30 and becomes progressively more common with age. About 26% of people older than 65 have type 2 diabetes (Goldenberg &Punthakee, 2013).

Obesity is the chief risk factor for developing type 2 diabetes, and 80 to 90% of people with this disorder are overweight or obese. Because obesity causes insulin resistance, obese people may need large amounts of insulin to maintain normal blood glucose levels.

People of African, Asian American, American Indian, Alaskan native, and Spanish or Latin American ancestry are at increased risk of developing type 2 diabetes. Type 2 diabetes tends to run in families. Certain disorders and medications can affect the way the body uses insulin and can lead to type 2 diabetes. Examples of common states (conditions) that result in impaired insulin use are: High levels of corticosteroids (most commonly due to use of corticosteroid medications, such as prednisone, or Cushing syndrome).

- Gestational Diabetes

Diabetes also may occur in people with excess production of growth hormone (acromegaly) and in people with certain hormone-secreting tumors. Severe or recurring pancreatitis and other disorders that directly damage the pancreas can lead to diabetes.

Diabetes mellitus also may develop as a secondary condition linked to another disease, such as pancreatic disease; a genetic syndrome, such as myotonic dystrophy; or drugs, such as glucocorticoids. Gestational diabetes is a temporary condition associated with pregnancy. In this situation, blood glucose levels increase during pregnancy but usually return to normal after delivery. However, gestational diabetes is recognized as a risk for type 2 diabetes later in life. Gestational diabetes is diagnosed when blood glucose concentrations measure between 92 and 125 mg per 100 ml (5.1 and 6.9 millimoles [mmol] per litre) after fasting or when blood glucose concentrations equal or exceed 180 mg per 100 ml (10 mmol per litre) one hour after ingesting a glucose-rich solution.

2.2. Symptoms of diabetes mellitus

Symptoms of type 1 diabetes can start quickly, in a matter of weeks. Symptoms of type 2 diabetes often develop slowly—over the course of several years—and can be so mild that you might not even notice them. Many people with type 2 diabetes have no symptoms. Some people do not find out they have the disease until they have diabetes-related health problems, such as blurred vision or heart trouble. There are various symptoms of diabetes and they vary according to types of diabetes mellitus. The symptoms that are predominant among all of them are (Goldenberg & Punthakee, 2013): dehydrated skin or Itching skin, delayed healing of wounds, cuts, or sores, mood swing, reduced sex drive, frequent urination, erectile dysfunction (ED), extreme hunger, blurry vision, poor muscle strength, presence of ketones in the urine, numbness in the hands or feet, increased thirst, unexplained weight loss or gain, nausea and vomiting, extreme fatigue, lack of energy, or feeling very tired, and dry mouth.

2.3 Methods for diagnosing diabetes mellitus

Medical experts diagnose diabetes mellitus by measuring blood glucose levels. Generally, there are four (4) different kinds of methodologies that are used for measuring blood glucose levels in the body. They include: a Fasting Plasma Glucose (FPG), a 2-hour Plasma Glucose (2hPG), a 75-gram Oral Glucose Tolerance Test (OGTT) and a glycatedhemoglobin (AIC) (GizemKoca, 2020). Willis, a London physician, epitomized the true spirit of scientific enquiry by his bold action of tasting the urine of his patients possibly because the passage of copious urine seemed to be the hallmark of the disease! This was a supreme and extreme example of bedside testing leading to labelling a patient as diabetic if his urine was ‘honeyed’.

Urine strips in the 1960s and the automated ‘do-it-yourself’ measurement of blood glucose through glucometers, produced by Ames Diagnostics in 1969, brought glucose control from the emergency room to the patient’s living room. It imbued diabetic patients with a new sense of freedom, making the disease more comprehensible and manageable (GizemKoca, 2020).

Routine blood sugar tests at prescribed intervals continued for a long time until the introduction of the glycosylated haemoglobin (HbA1c) estimation. That test, which measured blood glucose control over the previous three months (linked to the life of red blood cells), defined an extremely important aspect of diabetes management—tight control of blood glucose levels. The latter directly determined the risk of the occurrence of devastating complications of target organs like the eyes, vessels, nerves and kidneys that ultimately influenced morbidity and mortality (GizemKoca, 2020).

Doctors use a variety of tests to diagnose diabetes and prediabetes. The doctor may recommend different tests depending on whether you have symptoms or not, or whether you are pregnant. The tests include:

- Fasting plasma glucose test

The fasting plasma glucose (FPG) test measures your blood glucose level at a single point in time. For the most reliable results, your doctor will give you the test in the morning after you have fasted for at least 8 hours. Fasting means having nothing to eat or drink except sips of water.

- AIC test

The fasting plasma glucose (FPG) test measures your blood glucose level at a single point in time. For the most reliable results, your doctor will give you The A1C test is a blood test that provides your average levels of blood glucose over the last 3 months. Other names for the A1C test are hemoglobin A1C, HbA1C, glycatedhemoglobin, and glycosylated hemoglobin test. You can eat and drink before this test. Before using the A1C test to diagnose diabetes, your doctor will consider factors, such as whether you are in your second or third trimester of pregnancy or whether you have certain types of anemia NIH external link or another problem with your blood. The A1C test might not be accurate in those cases.

Certain types of hemoglobin, called hemoglobin variants, can interfere with measuring A1C levels. Most A1C tests used in the United States are not affected by the most common variants. If your A1C test results and blood glucose levels do not match, your doctor should consider that the A1C test may not be a reliable test for you. The doctor will report your A1C test result as a percentage, such as an A1C of 7%. The higher the percentage is, the higher average blood glucose levels.

- Random plasma glucose test

Sometimes doctors use the random plasma glucose test to diagnose diabetes when you have symptoms of diabetes and they do not want to wait until you have fasted for 8 hours. You may have this blood test at any time.

- Glucose challenge test.

If one is pregnant, the doctor might test for gestational diabetes with the glucose challenge test. Another name for this test is the glucose screening test. In this test, a health care professional will take a sample of your blood 1 hour after you drink a sweet liquid containing glucose. One does not need to fast for this test. If your blood glucose level is too high—135 mg/dL to 140 mg/dL or higher—you may need to return for an oral glucose tolerance test while fasting.

- Oral glucose tolerance test

The oral glucose tolerance test (OGTT) helps doctors detect type 2 diabetes, prediabetes, and gestational diabetes. However, the OGTT is a more expensive test than the FPG test and the glucose challenge test, and it is not as easy to give.Before the test, you will need to fast for at least 8 hours. A health care professional will take a blood sample to measure your glucose level after fasting. Next, you will drink a liquid that is high in sugar. Another blood sample is taken 2 hours later to check your blood glucose level. If your blood glucose level is high, you may have diabetes.

If you are pregnant, your blood will be drawn every hour for 2 to 3 hours. If the blood glucose levels are high two or more times during the OGTT, you may have gestational diabetes.

2.4 Machine learning algorithm

The artificial intelligent system designed in this paper is based on machine learning. Depending on the learning task, the field offers various classes of ML algorithms, each of them coming in multiple specifications and variants, including regressions models, instance-based algorithms, decision trees, Bayesian methods, and ANNs.

The family of artificial neural networks is of particular interest since their flexible structure allows them to be modified for a wide variety of contexts across all three types of ML. Inspired by the principle of information processing in biological systems, ANNs consist of mathematical representations of connected processing units called artificial neurons. Like synapses in a brain, each connection between neurons transmits signals whose strength can be amplified or attenuated by a weight that is continuously adjusted during the learning process. Signals are only processed by subsequent neurons if a certain threshold is exceeded as determined by an activation function. Typically, neurons are organized into networks with different layers. An input layer usually receives the data input (e.g., product images of an online shop), and an output layer produces the ultimate result (e.g., categorization of products). In between, there are zero or more hidden layers that are responsible for learning a non-linear mapping between input and output (Bishop 2006; Goodfellow et al. 2016). The number of layers and neurons, among other property choices, such as learning rate or activation function, cannot be learned by the learning algorithm. They constitute a model's hyperparameters and must be set manually or determined by an optimization routine.

Deep neural networks typically consist of more than one hidden layer, organized in deeply nested network architectures. Furthermore, they usually contain advanced neurons in contrast to simple ANNs. That is, they may use advanced operations (e.g., convolutions) or multiple activations in one neuron rather than using a simple activation function. These characteristics allow deep neural networks to be fed with raw input data and automatically discover a representation that is needed for the corresponding learning task. This is the networks' core capability, which is commonly known as deep learning. Simple ANNs (e.g., shallow autoencoders) and other ML algorithms (e.g., decision trees) can be subsumed under the term shallow machine learning since they do not provide such functionalities. While some shallow ML algorithms are considered inherently interpretable by humans and, thus, white boxes, the decision making of most advanced ML algorithms is per se untraceable unless explained otherwise and, thus, constitutes a black box.

DL is particularly useful in domains with large and high-dimensional data, which is why deep neural networks outperform shallow ML algorithms for most applications in which text, image, video, speech, and audio data needs to be processed (LeCun et al. 2015). However, for low-dimensional data input, especially in cases of limited training data availability, shallow ML can still produce superior results (Zhang and Ling 2018), which even tend to be better interpretable than those generated by deep neural networks (Rudin 2019). Further, while DL performance can be superhuman, problems that require strong AI capabilities such as literal understanding and intentionality still cannot be solved as pointedly outlined in Searle (1980)'s Chinese room argument.

ML algorithms are generally classified into three, based on the way the models learn. These are:

Supervised learning - This algorithm includes a target or outcome variable (also known as dependent variable) which will be predicted from a set of predictors (independent variables). The algorithm attempts to model the relationship and dependencies between the dependent and independent variable with the help of the given set of variables, in a way that the output values for a new data set based on those relationship can be predicted This training process goes on till the model is able to achieve the desired accuracy level on the training data. Example: Regression and classification tasks

Unsupervised learning - In this algorithm, no target or outcome variable is available to make predictions. It is used for detecting patterns and descriptive modeling. Clustering of population in different groups generally uses unsupervised technique for segmenting customers in different groups for specific intervention.

Semi-supervised learning – This method falls in between the above two. The cost of getting labeled dataset is usually high and, in such situations, we obtain some labeled data along with lots of unlabeled data for training.

Reinforcement learning - Here, the machine is exposed to an environment where it trains itself repeatedly through trial and error and has to determine the ideal behavior within the given context to maximize performance. Q-learning, temporal difference and deep adversarial networks are some of them. ML can be further categorized into three based on the objective the model is trying to achieve namely, regression: for prediction of a continuous real number, classification: for prediction of a discrete number, and clustering: for segmentation.

Similarly, instead of codifying knowledge into computers, machine learning (ML) seeks to automatically learn meaningful relationships and patterns from examples and observations (Bishop 2006). Advances in ML have enabled the recent rise of intelligent systems with human-like cognitive capacity that penetrate our business and personal life and shape the networked interactions on electronic markets in every conceivable way, with companies augmenting decision-making for productivity, engagement, and employee retention (Shrestha et al. 2021), trainable assistant systems adapting to individual user preferences (Fischer et al. 2020), and trading agents shaking traditional finance trading markets (JayanthBalaji et al. 2018).

The capacity of such systems for advanced problem solving, generally termed artificial intelligence (AI), is based on analytical models that generate predictions, rules, answers, recommendations, or similar outcomes. First attempts to build analytical models relied on explicitly programming known relationships, procedures, and decision logic into intelligent systems through handcrafted rules (e.g., expert systems for medical diagnoses) (Russell and Norvig 2021). Fueled by the practicability of new programming frameworks, data availability, and the broad access to necessary computing power, analytical models are nowadays increasingly built using what is generally referred to as ML (Brynjolfsson and McAfee 2017; Goodfellow et al. 2016). ML relieves the human of the burden to explicate and formalize his or her knowledge into a machine-accessible form and allows to develop intelligent systems more efficiently.

During the last decades, the field of ML has brought forth a variety of remarkable advancements in sophisticated learning algorithms and efficient pre-processing techniques. One of these advancements was the evolution of artificial neural networks (ANNs) towards increasingly deep neural network architectures with improved learning capabilities summarized as deep learning (DL) (Goodfellow et al. 2016; LeCun et al. 2015). For specific applications in closed environments, DL already shows superhuman performance by excelling human capabilities (Madani et al. 2018; Silver et al. 2018). However, such benefits also come at a price as there are several challenges to overcome for successfully implementing analytical models in real business settings. These include the suitable choice from manifold implementation options, bias and drift in data, the mitigation of black-box properties, and the reuse of preconfigured models (as a service).

Beyond its hyped appearance, scholars, as well as professionals, require a solid understanding of the underlying concepts, processes as well as challenges for implementing such technology. Against this background, is to convey a fundamental understanding of ML and DL in the context of electronic markets. In this way, the community can benefit from these technological achievements – be it for the purpose of examining large and high-dimensional data assets collected in digital ecosystems or for the sake of designing novel intelligent systems for electronic markets. Following recent advances in the field, this article focuses on analytical model building and challenges of implementing intelligent systems based on ML and DL. As we examine the field from a technical perspective, we do not elaborate on the related issues of AI technology adoption, policy, and impact on organizational culture (for further implications as in Stone et al. 2016). The hierarchical relationship between those terms is summarized in Venn diagram of Fig. 1.

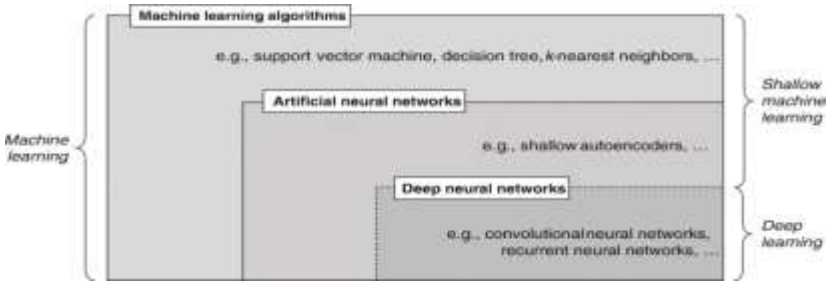

Fig. 1 – Venn diagram of machine learning concepts and classes (Goodfellow et al., 2016)

3. Method of system design

The method that will be adopted in this research would be based on design of a hardware prototype of a patient monitoring system, simulations and experimentations carried out using machine learning based library including Tensor flow, MATLAB, and Jupiter notebook. To build the inference system, the research flow for the software aspect will follow the step-wise approach illustrated in Fig. 2.

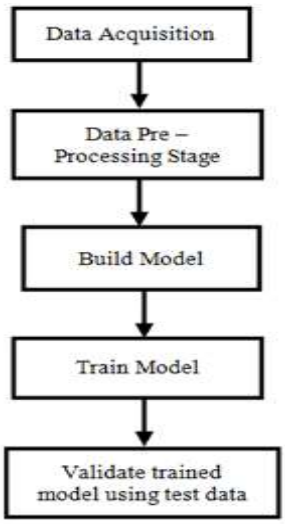

Fig. 2 – Research approach block diagram

3.1 Data acquisition and pre-processing

Data-acquisition –The goal of data acquisition is to find datasets that can be used to train machine learning models. There are largely three approaches to this, which are: data discovery, data augmentation, and data generation. For this paper, data augmentation approach was used. Two options were used to obtain data to train the machine learning model, these are: data generated by the developed Patient Monitoring (PM) prototype device, and data obtained from other medical databases. The work uses an augmentation of data from four (4) different databases. The databases include – The Pima Indian Diabetes Dataset, Bio-Statistics Diabetes Dataset, Nnamdi Azikiwe University Teaching Hospital (NAUTH), and Enugu State University Teaching Hospital (ESUTH).

Data pre-processing –This stage is one of the most critical aspects in the training process. It prepares and transforms the basic dataset. The data were also collected manually from federal teaching hospitals in eastern Nigeria. The dataset from Nnamdi Azikiwe University Teaching Hospital has 67 patients with four different attributes, which are – plasma glucose concentration, diastolic blood pressure, Gender, and age. In the dataset, 42 patients out of 67 patients indicated as a type 2 diabetes mellitus patient.

The dataset from ESUTH has records from 2163 patients with ten different attributes, which are –sex, cholesterol, pregnancy situation, alcohol consumption, Physical activity, plasma glucose concentration, diastolic blood pressure, Gender, and age.

Improving accuracy or reducing computational cost are the main approaches of machine learning techniques, but it depends heavily on the test data used. Even more so when it comes from real-world data that contain a high level of missing values. It is very important to select a method that is capable of replacing these missing values with plausible values.

In this paper, the dataset is checked to see if there are any correlated features. Correlation can be defined as a measure of how strongly one input feature depends on another. By removing any correlated features, one can increase the speed of learning of an algorithm. It can also reduce the bias in the neural network. Also, the Multiple Imputation method was used to handle the missing values in the original dataset. The multiple imputation technique by Rubin (2004) was selected based on the percentage and pattern of the missing values. The Multiple Imputation is an approach that replaces each deficient or missing value with more than one acceptable value representing a distribution of possibilities. It looks at the pattern of the available data, and based on probability judgment, attempts to find the best matches, replacing the missing values with imputed values. Replacement is performed repeatedly in order to find the perfect fit. Normalization is also carried out to ensure that the data is appropriate for the training process. In this process, the data is scaled in some specific range for every input feature to reduce the bias in the neural network. The system transforms all the features by scaling them in to a given range. The transformation is given by the following equation:

$$X_{std} = \frac{x - x_{\min}}{x_{\max - x_{\min}}} \quad (1)$$

3.2 Model building

The deep learning classification architecture used in this work has four convolutional blocks which generate the feature maps. That data is then reshaped into the format so that it can be input into the linear classifier layer, which finally outputs the predictions for the 2 classes. In a classification problem, the output layer can either have one node or can have nodes equivalent to the number of the classes or the categories. The network built here is called Fully Connected Network (FNN) or Dense Network since every neuron has a connection with the node of the previous layer output. It is also known as the Feedforward Neural Network or Sequential Neural Network.

The equation for the neural network is a linear combination of the independent variables and their respective weights and bias (or the intercept) term for each neuron. The neural network equation is expressed by:

$$Y = \sum (\vec{x} * [w^{[n]}]) + b \tag{2}$$

where Y is the symbol for denotation of the graphical representation of ANN, $\vec{x}$ is the input data, $w^{[n]}$ is the weight of the n^{th} input, and b is the bias.

There are three steps to perform in any neutral network:

- Take the input variables and the linear combination equation: $Y = w^{[0]} + w^{[1]}x_1 + w^{[2]}x_2 + \cdots + w^{[n]}x_n$ to compute the output or the predicted Y values, called the Ypred.
- Calculate the loss or the error term. The error term is the deviation of the actual values from the predicted values.
- Minimize the loss function or the error term.

The neural network designed for training consists of feedforward, fully connected (dense) layers and the recurrent layers. The input layer has 8independent variables to accept the input features and 2 neurons at the output layer to provide an estimate of the gain values in each band. It is a 5-layer network with 3 hidden layers (L_2, L_3 and L_4) and 1 output layer (L_5). Hidden layer L_2 has 12 neurons, while hidden layer L_3 and L_4 has 15 and 7 neutron respectively. The output layer nodes are dependent on their immediately preceding hidden layer, which is coming from the first hidden layer and those nodes are further derived from the input variables. These middle hidden layers create the features that the network automatically creates and we don't have to explicitly derive those features. In this manner, the features are generated in Deep Learning models and this is what makes them stand out from Machine Learning.

So, to compute the output, it will require the calculation of all the nodes in the previous layers. In this paper, the neurons in hidden layer L_2 are referred to $a_1, a_2, a_3, a_4, \cdots a_{12}$ and for the hidden layer L_3, the neutrons are referred to $a_{13}, a_{14}, \cdots a_{28}$ respectively in the linear order of their occurrence. In a classification problem, the output layer can either have one node or can have nodes equivalent to the number of the classes or the categories.

Note that each neuron cannot have the same general equation for the output as shown in (3.2). We will have one such equation per neuron both for the hidden and the output layer. The nodes in the hidden layer L_2 are dependent on the values of X present in the input layer therefore; the equation will be the following:

$$\left.\begin{aligned}
a_1 &= w^{[11]}x_1 + w^{[12]}x_2 + \cdots + w^{[18]}x_8 + w^{[10]} \\
a_2 &= w^{[21]}x_1 + w^{[22]}x_2 + \cdots + w^{[28]}x_8 + w^{[20]} \\
a_3 &= w^{[31]}x_1 + w^{[32]}x_2 + \cdots + w^{[38]}x_8 + w^{[30]} \\
a_4 &= w^{[41]}x_1 + w^{[42]}x_2 + \cdots + w^{[48]}x_8 + w^{[40]} \\
&\vdots \\
a_{12} &= w^{[121]}x_1 + w^{[122]}x_2 + \cdots + w^{[128]}x_8 + w^{[120]}
\end{aligned}\right\} \tag{3}$$

Similarly, the nodes in the hidden layer L_3 are derived from the neurons in the previous hidden layer L_2 hence their respective equations will be:

$$\left.\begin{aligned}
a_{13} &= w^{[131]}a_1 + w^{[132]}a_2 + \cdots + w^{[1312]}a_{12} + w^{[130]} \\
a_{14} &= w^{[141]}a_1 + w^{[142]}a_2 + \cdots + w^{[1412]}a_{12} + w^{[140]} \\
a_{15} &= w^{[151]}a_1 + w^{[152]}a_2 + \cdots + w^{[1512]}a_{12} + w^{[150]} \\
a_{164} &= w^{[161]}a_1 + w^{[162]}a_2 + \cdots + w^{[1612]}a_{12} + w^{[160]} \\
&\vdots \\
a_{28} &= w^{[281]}a_1 + w^{[282]}a_2 + \cdots + w^{[2812]}a_{12} + w^{[280]}
\end{aligned}\right\} \tag{4}$$

This also continues till the output layer. The output layer nodes are coming from the hidden layer L_4 which makes the equations as:

$$\left.\begin{aligned}
o_1 &= wo^{[129]}a_{29} + wo^{[130]}a_{30} + \cdots + wo^{[136]}a_{36} + wo^{[10]} \\
o_2 &= wo^{[229]}a_{29} + wo^{[230]}a_{30} + \cdots + wo^{[236]}a_{36} + wo^{[20]}
\end{aligned}\right\} \tag{5}$$

The value of Y can be -∞ to +∞. So, the neuron can't decide whether it will fire or not. Here the activation function is used to decide the neuron will fire or not. This paper used ReLU as an activation function. The rectified linear activation function or ReLU is a non-linear function or piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero.It is the most commonly used activation function in neural networks, especially in Convolutional Neural Networks (CNNs) & Multilayer perceptrons.It is simple yet it is more effective than its predecessors like sigmoid or tanh.

The Activation functions play a vital role while designing a neural network. To understand the activation function we have to understand what artificial neurons do. It evaluates the weighted sum of all the inputs it receives and adds bias and then decides whether it should fire or not. So, this value can range between + ∞ to - ∞. Thus, activation functions bound this value produced by neurons and decide if the external connection is fired or not.

Mathematically, it is expressed as:

$$f(x) = \max(0, x) \quad (6)$$

3.3 Model training

Training on a whole dataset using batch gradient descent takes longer, and weight updates are less frequent. For training the DNN, the stochastic gradient descent approach was used, and the functions for the optimizer, loss, and scheduler were designed to dynamically alter the learning rate as training progressed, allowing training to converge in fewer epochs. The training had a 0.2 dropout rate. The momentum rate was set at 0.5 for the first five epochs, and then raised to 0.9. In the last layer, the cost function was the mean square error (MSE). The batch size was set to 1024, the learning rate was set to 0.001, and the network trained for 20 epochs. During training, the model is trained over numerous epochs, with each iteration processing a batch of data. The system monitors a basic accuracy metric that calculates the proportion of right predictions. During training, the function uses gradient descent to find the least value of the error function in weight space. A solution to the learning issue is therefore regarded to be the weights that minimize the error function. Fig. 3 depicts the flow chart for this process:

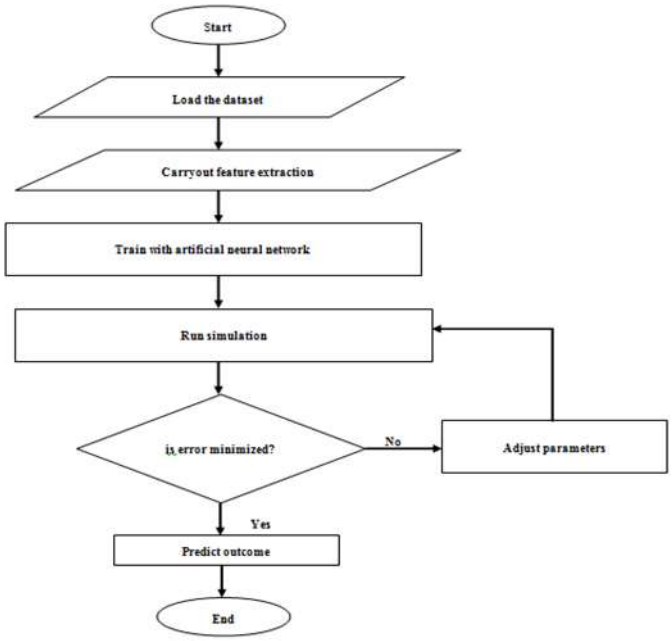

Fig. 3 - Flow chart of training process

3.4 Model validation

To validate the model, the testing dataset is obtained from the training dataset by randomly picking 20% of the data in each of the database used. During the test, the result of the test can either be positive (classifying the patient as being diabetic when it he/she is diabetic) or negative (classifying the patient as being not diabetic when he/she is not).The result of the test for each input signal may or may not match the actual status. To accommodate these scenarios, the following instances after the simulation are postulated:

- True positive (TP): Diabetic patient correctly identified as diabetic
- False positive (FP): Non-diabetic incorrectly identified as diabetic
- True negative (TN): Non-diabetic correctly identified as non-diabetic
- False negative (FN): Diabetic patient incorrectly identified as non-diabetic

The performance metrics used to evaluate the performance of the fuzzy expert system for the incidence of diabetes are – accuracy metric, specificity metric, sensitivity metric, and precision metric.

The accuracy metric shows the fraction of true results (both true positives and true negatives) among the total number of cases examined. The specificity metric (also called true negative rate) refers to the test's ability to correctly detect patients who do not have diabetes, whereas the sensitivity metric (also called recall, or true positive rate) relates to the test's ability to correctly detect patients who do have diabetes. In other words, sensitivity is the proportion

of correct positive classifications (TP) from cases that are actually positive. On the other hand, precision is the proportion of correct positive classifications (TP) from cases that are predicted to be positive. The equations of the performance metrics are as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \quad (7)$$

$$Specificity = \frac{TN}{FP + TN} \times 100 \quad (8)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \quad (9)$$

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (10)$$

4. Hardware design outline

The design process adopted in this work is as illustrated in Fig. 4.

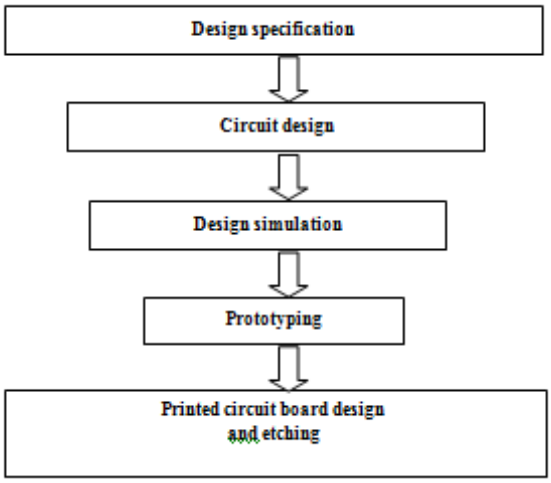

Fig. 4 - Block diagram showing the design outline

4.1. Design Specification

Design specification as it concerns this work will be discoursed based on the following: functional, power and physical structure specification.

- Functional specification: Due to the fact that the main components of the system are to be interfaced to a microcontroller, the choice of a programmable and reprogrammable controller with low power consumption is made. It must have enough internal flash memory to accommodate the written program for monitoring the measurements of the blood pressure sensor, Electrocardiogram (ECG) sensor, and temperature sensor. It has its own circuit board with serial output data for external processing or display. It also has its own LCD that displays its output first before transmitting to the microcontroller; this enables the programmer to be sure that the written program or the entire hardware wiring is in other when the same information is displayed on the entire circuit LCD. It has its own memory that holds the last reading from the patient's body even after transmitting to the controller. 16x2 LCD type was chosen for the entire system display because of its small size with respect to the portability feature of this system and also its low power consumption (same with the controller). It has backlight adjustment for more visibility. It can display 16 characters per line and there are 2 such lines. In this LCD each character is displayed in 5x7 pixel matrix. This LCD has two registers, namely, Command and Data.
- Power specification: Having considered the voltage and current consumption of the individual main components, a conclusion was reached that the power source of this system will be a 5VDC supply (since all the major parts of this system require a maximum of 5V DC supply).
- Physical structure design: With respect to the significance of this project, starting with the internal structure, the circuit board is a Printed Circuited Board (PCB) format, because it help to miniature the circuit size as will be explain later in detail. The display would be mounted on the top outer surface of the entire system.

- Circuit design: This part of the design outline gives the display system schematic with respect to the design specification. It starts by producing the block diagram of different branches of the entire system with arrows that show how they function together to form the PM system as shown in Fig.5.

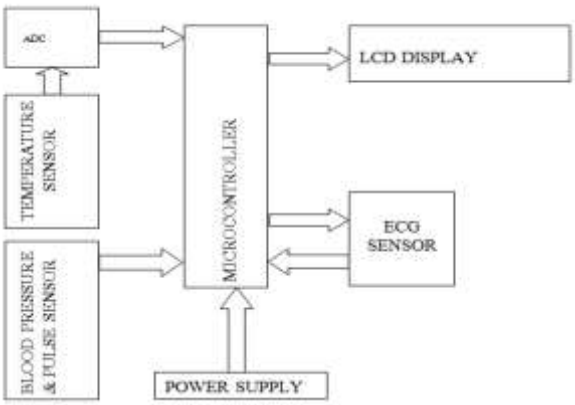

Fig. 5 - Block diagram of the hardware prototype

The block diagram (Fig.5) helps to analyze each block or main part of the circuit so that one can effectively select other minor circuit components suitable to drive each of the main components according to their individual data sheet. The electrical interconnection of the entire system components and their physical layout were done using Proteus 8 professional (ISIS schematic capture). The entire circuit diagram is shown in Fig. 6.

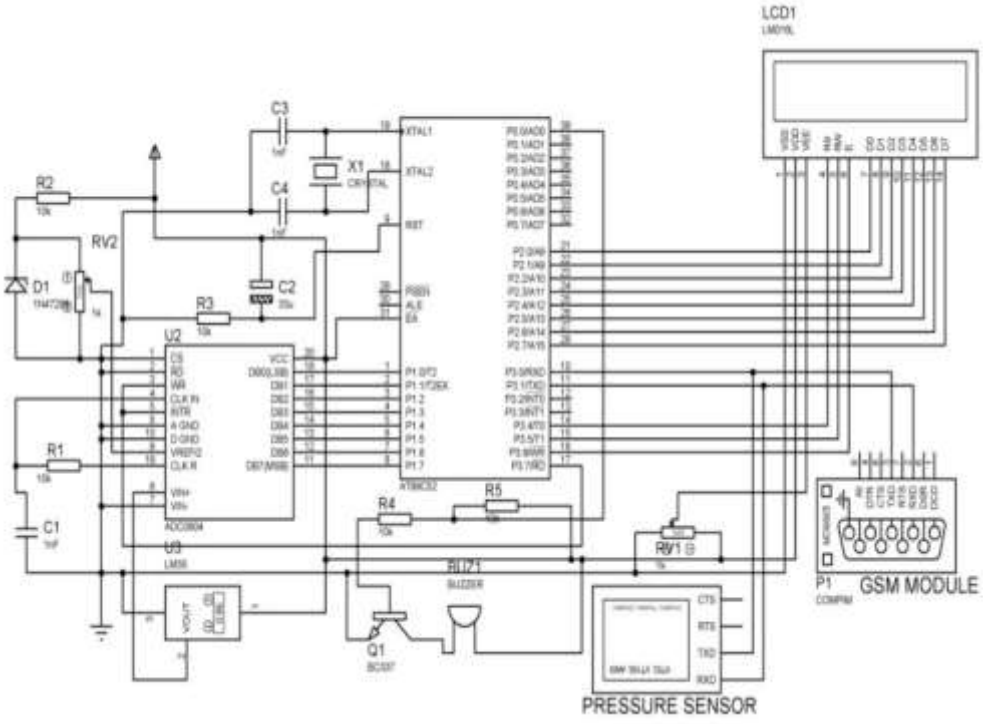

Fig. 6 - The circuit diagram of PM system

The core components are wearable Blood pressure sensor unit, temperature sensor and ADC unit, ECG sensor, microcontroller unit LCD display unit, and power supply unit.

- Wearable blood pressure sensor: This wearable blood pressure sensor is a kind of monitoring device that allows a patient to measure his or her blood pressure and heart rate at a time and store them in its memory for reference purpose. Blood Pressure & Pulse reading are shown on display with serial out for external projects of embedded circuit processing and display. It shows Systolic, Diastolic and Pulse Readings. Compact design fits over your wrist like a watch. Easy to use wrist style eliminates pumping. The sensor features are intelligent automatic compression and decompression, easy to operate (switching button to start measuring), 6o store groups memory measurements, can read single or all measures, 3 minutes automatic power saving device, intelligent device debugging, automatic power to detect, local tests for: wrist circumference as 135-195 mm, large-scale digital liquid crystal display screen (easy to read display), fully automatic including clinical accuracy and high accuracy, powered by external +5 V DC, serial output data for external circuit processing or display, working voltage: +5

V, 200 mA regulated, output format: serial data at 9600 band rate (8 bits data, no parity, 1 stop bits) and output three parameters in ASCII (from 000 to 255), and sensing unit wire length is 2 metres.

- Temperature sensor and ADC: The temperature sensor which is the second sensor is only used to monitor the body temperature of a patient. The choice type of temperature sensor used for this system is LM35 because of its advantages over other temperature measuring sensors. LM35 Features include: calibrated directly in 0 Celsius (Centigrade), linear + 10.0 mV/oC scale factor, 0.5ºC accuracy guarantee able (at +25oC), rated for full −55o to +150oC range, suitable for remote applications, low cost due to wafer-level trimming, operates from 4 to 30 V, less than 60 µA current drain, low self-heating, 0.08oC in still air, nonlinearity only ±1⁄4oC typical, and low impedance output, 0.1 Ω for 1 mA load. Since the output of the temperature sensor is an analog signal and what we want is a digital output, we made use of an analog to digital converter (ADC0804, which is a device of CMOS 8-bit successive approximation converter (ADC) that use a differential potentiometric-ladder — similar to the 256R products) that will convert the analog signal to a digital type for easy understanding of the microcontroller.

- AD8232 ECG sensor: This sensor is a cost-effective board used to measure the electrical activity of the heart. This electrical activity can be charted as an ECG or Electrocardiogram and output as an analog reading. ECGs can be extremely noisy, the AD8232 Single Lead Heart Rate Monitor acts as an op-amp to help obtain a clear signal from the PR and QT Intervals easily. The AD8232 module breaks out nine connections from the IC that you can solder pins, wires, or other connectors to. SDN, LO+, LO-, OUTPUT, 3.3V, GND provide essential pins for operating this monitor with an Arduino or other development board. Also provided on this board are RA (Right Arm), LA (Left Arm), and RL (Right Leg) pins to attach and use your own custom sensors. Additionally, there is an LED indicator light that will pulsate to the rhythm of a heartbeat. The AD8232 Heart Rate Monitor breaks out nine connections from the IC. We traditionally call these connections “pins” because they come from the pins on the IC, but they are actually holes that you can solder wires or header pins to. The ECG is the graphical recording of the electrical signals of the heart. It is useful in determining a person’s heart rate and rhythm. Some of the things an ECG reading can detect are: cholesterol clogging up your heart’s blood supply, a heart attack in the past, enlargement of one side of the heart, and abnormal heart rhythms.

- Microcontroller: The choice of microcontroller used in this work is AT89S52. The AT89S52 is a low-power, high-performance CMOS 8-bit microcontroller with 8K bytes of in-system programmable Flash memory. The device is manufactured using Atmel’s high-density nonvolatile memory technology and is compatible with the industry-standard 80C51 instruction set and pin out. The on-chip Flash allows the program memory to be reprogrammed in-system or by a conventional nonvolatile memory programmer. By combining a versatile 8-bit CPU with in-system programmable Flash on a monolithic chip, the Atmel AT89S52 is a powerful microcontroller which provides a highly-flexible and cost-effective solution to many embedded control applications. The Features of AT89S52 are: Compatibility with MCS-51 products, 8K Bytes of In-system programmable (ISP) flash memory – endurance: 10,000 Write/Erase Cycles, 4.0V to 5.5V operating range, fully static operation: 0 Hz to 33 MHz, three-level program memory lock, 256 x 8-bit internal RAM, 32 programmable I/O Lines; three 16-bit timer/counters; eight interrupt sources; full duplex UART serial channel; low-power idle and power-down modes; interrupt recovery from power-down mode; watchdog timer; dual data pointer; power-off flag; fast programming time; flexible ISP programming (Byte and Page Mode); green (Pb/Halide-free) packaging option.

- LCD display: This is the display unit of this work. With the purpose of making this system portable, we decided to use a 16X2 liquid crystal display. LCD (Liquid Crystal Display) screen is an electronic display module and find a wide range of applications. A 16x2 LCD display is very basic module and is very commonly used in various devices and circuits. These modules are preferred over seven segments and other multi segment LEDs. The reasons being: LCDs are economical; easily programmable; have no limitation of displaying special & even custom characters (unlike in seven segments), animations and so on.

- Power supply: Having considered the individual components of the system and their current and voltage usage, the power supply of this circuit was made to be a 5V rechargeable DC battery.

- Design simulation: This is where the developed circuit diagram is being simulated using a schematic capture tool known as Proteus Virtual System Modeling. At this point, the programming codes that will drive and coordinate the activities of the various components were developed and burn into the controller. The essence of this simulation is to ascertain with fact that the real-life workability of the developed circuit diagram. The simulated design outcome is shown in Fig. 7.

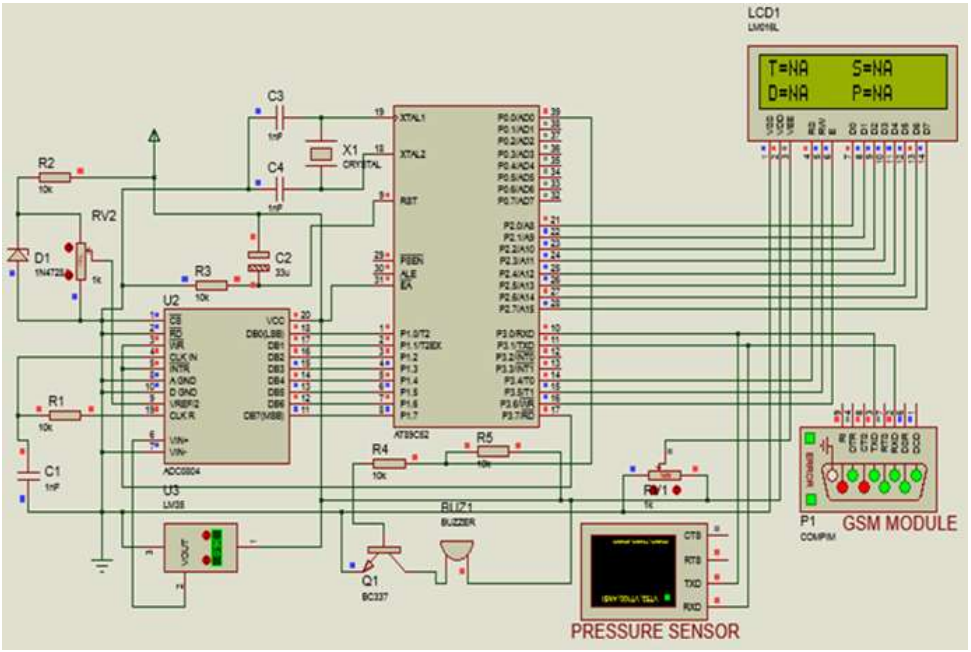

Fig. 7 - The simulated design of "The PM System using proteus

- Prototyping: The prototyping model is a systems development method in which a prototype (an early approximation of a final system or product) is built, tested, and then reworked as necessary until an acceptable prototype is finally achieved from which the complete system or product can now be developed. It is an iterative, trial-and-error process that takes place between the developers and the users. The essence of this prototype stage is to make sure that the simulated design can actually become a real-life working system. The working prototype of this project is shown in Fig. 8.

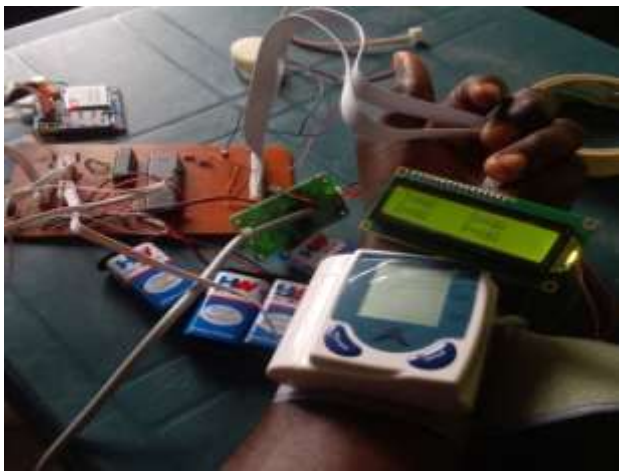

Fig. 7 - The working prototype of RPM System

In this paper attempt has been made to model and design an artificial intelligent aided patient monitoring system for measuring vital parameters for diabetes mellitus diagnosis. The proposed system is developed in such a way that reduces the computational complexity in a rather simple but efficient approach to the diagnosis of Diabetes Miletus. The simplicity of the algorithm would make for faster and more accurate results than previous models. Also, there is hardly any available dataset for usage in this part of the world (developing nations like Nigeria) for this research. This research stands out as it develops an embedded system for generating its own data which would also make them available for further research, thus making the solution differently and easily applicable to people from this region. The algorithm is also.

Acknowledgements

An example appendix

References

Gizem Koca (2020) .An Intelligent System For Type II Diabetes Mellitus Diagnostic. A Thesis Submitted to the Faculty of Graduate Studies and Research. In Partial Fulfillment of the Requirements For the Degree of Master of Applied Science in Industrial Systems Engineering, University of Regina.

Goldenberg, R., & Punthakee, Z. (2013). Definition, Classification and Diagnosis of Diabetes, Prediabetes and Metabolic Syndrome. *Canadian Journal of Diabetes, 37*, 8-11. doi:10.1016/j.jcjd.2013.01.011

Rubin, D. (2004). Multiple imputation for nonresponse in surveys. *John Wiley & Sons*, vol. 81, p. 253.

Watson, C. S. (2017). Prediabetes: Screening, Diagnosis, and Intervention. *The Journal for Nurse Practitioners - JNP, 13*(3), 216 - 221. doi:10.1016/j.nurpra.2016.08.005