International Journal of Multidisciplinary and Current Educational Research (IJMCER)
ISSN: 2581-7027 ||Volume|| 5 ||Issue|| 6 ||Pages 94-110 ||2023||

An Optimised Deep Learning-based Prediction Model for Smart Healthcare Systems

[1,] Dr Sunny Vinnakota, [2,] Mr Mohammed Shafiuddin, [3,] Dr Mohan Dass Mohan, [4,] Dr Walied Askarzai, [5,] Mr Johnson Boda, [6,] Ms Navya Brahmaiah.
[1,2,3,4,5,6,]Higher Education Department, Academies Australasia Polytechnic, Melbourne & Sydney, Australia

ABSTRACT: The healthcare industry is currently on the verge of a significant transformation, as it incorporates very sophisticated and comprehensible machines and deep learning models. Interpretable models play a crucial role in improving healthcare outcomes by enabling the timely discovery of diseases, hence promoting prompt and effective therapies. As a result, implementing this predictive methodology promotes individualized patient care, enhancing medical operations' overall effectiveness and productivity. The introduction of intelligent gadgets in the healthcare sector enhances this fundamental change. These gadgets, equipped with predictive functionalities, can warn about future health concerns, frequently obviating the necessity for supplementary diagnostic procedures. The primary focus of this study is creating an enhanced prediction model utilising deep learning techniques, with a specific emphasis on its applicability to advanced healthcare systems. The technique employed in this study was thorough, incorporating key stages such as data pre-processing, feature extraction, and classification. To enhance the model's resilience and precision, including an optimisation step was a crucial component of our study methodology. For benchmarking, our findings were compared with recognised methodologies, such as the random forest and Convolutional Neural Networks (CNN). The present study highlights the improved efficacy of our deep learning model in healthcare predictive analytics. The simulations were conducted using Python 3.9.7 and Jupyter Notebook 6.4, ensuring a uniform and replicable study methodology. This undertaking seeks to establish the significance of deep learning in influencing the future personalised healthcare trajectory.

KEYWORDS – Healthcare, Machine Learning, Optimisation, Random Forest, Convolutional neural networks

1. INTRODUCTION

This section briefly introduces the deep learning-based prediction model for smart healthcare systems. This section provides the background of the research paper, including a detailed introduction to classification and optimisation algorithms utilised in smart healthcare systems. The section will further elaborate on the study's objectives and research significance. There has been a recent surge in interest in Internet of Things (IoT) technologies and deep learning at both the industrial and academic levels. Data may be gathered, and effective judgments may be made in an IoT environment, ranging from adjusting the amount of medication for patients [7]. Background data is created by recording and checking the account of data. An IoT system consists of computers, sensors, and machines that can take action based on sensor data. As with people, machine learning uses data from their environments and natural laws to mimic the learning process. Industry's machinery could communicate with each other during the fourth industrial revolution since it was connected to the Internet. Sensors and actuators with an electronic unit are a few components utilised in this process.

A cyber-physical system incorporates physical and embedded systems [8] onto a unified platform. A network of "Internet of Things" (IoT) devices is formed when these cyber-physical systems are connected to the Internet. A highly centralised system design and computing and communication capabilities were used to build the current Internet of Things platforms, which allow access to real-world environments. Here, massive, embedded strategies were attached, which aid in providing services to end-users [9]. Connectedness is a potential source of innovation in the economy, society, and the personal and professional lives of those participating. IoT systems can monitor the health of the elderly and notify emergency services in real time if something goes wrong [10]. Security issues in IoT systems expand in complexity and breadth, in addition to data management capabilities. The limited processing, storage, and communication resources of IIoT field device systems must be considered while developing cryptographic algorithms for the Internet of Things. Elliptic curve cryptography (ECC) has been employed in certain studies, and this ECC is more secure than previous public-key techniques [11]. Among the several chronic diseases, diabetes is a risky disease that could lead to other severe issues like heart attack, kidney failure, etc.

It is broadly classified as Type I and Type II. Both these types depend on the insulin level in the blood [1]. Typically, type I is found in children and adolescents but could also affect senior citizens. Type II is usually milder than type I, and 90% of humans have type II. If diabetes can be known earlier and cured, several other issues can be prevented [2]. Artificial intelligence in diabetes care could be deployed in the management process to structure targeted data-driven precision care [3]. IoT could provide smart healthcare systems in disease prediction paradigms [4]. Machine learning and deep learning are AI-based techniques that have the potential to enhance efficiency and minimise the cost factor of treatment in healthcare systems. Based on data mining and machine learning means, several prediction measures were suggested for diabetes disease. Automatic diabetes detection is possible by employing prediction measures and management with the support of deep learning criteria [5] [6]. A person's health comprises complete physical, mental, and social goodness other than solely depending on a lack of illness [26]. Nowadays, global-level health issues concentrate more on early prediction of several diseases, which could assist in early treatment and cure.

Moreover, compared to rural and urban areas, doctors and nurses might not always be available for consultation, a drawback in healthcare systems [27]. To overcome this issue and assist patients with primary medication or knowing their health status, several IoT-based smart meters and devices are introduced for taking self-preventive measures such as first aid [28]. For example, there is no need to visit the hospital to measure the sugar level of a diabetes patient. They could get a smart meter to measure sugar levels and maintain their diet accordingly. In healthcare, smart IoT systems provide innumerable outcomes like early illness and treatment prediction [29]. IoT systems link the computers with the Internet, which could include sensors and networks. These connected elements could be used on several devices for health monitoring systems [30].

II. LITERATURE REVIEW

This section embraces the study of several present research approaches used in smart healthcare systems. The associated research study analyses current literature about diverse systems to diagnose diseases. Misuse detection, anomaly detection, and formal complete protocol analysis are the primary methods used to examine and organise network traffic data [22]. The misuse detection identifies the already identified explicit attack arrays by recognising the established static signatures and filters. The signatures and filters are constantly updated by humans using this method, which relies heavily on human inputs. Misuse detection looks for known attacks, whereas anomaly detection looks for new or unusual attack setups using heuristic techniques [23].] As a result, the rate of false positives may be considerable. Most institutes employ a mixture of static and heuristic approaches, referred to as a hybrid method, to minimise this. Like anomaly detection, state complete protocol evaluation uses vendor-defined criteria and requirements to identify protocol state deviations; generally, they are considered harmless network traffic. This has emerged as a viable method widely discussed because of its ability to operate at the network, application, and transport layers [24].

Several machine learning methods may be used to help examine normal and malicious network data to identify the complicated unidentified designs of assaults automatically. Furthermore, they have the capability of identifying limitless arrays of attacks that have been planned of the attack. The current favourable way of ID is ML methods, which can more precisely identify and arrange complex network traffic measurements with a low false positive rate [25]. Using this strategy, the most significant dangers can be examined and detected in real-time, and the necessary countermeasures may be taken intelligently. There has been an increase in the importance of machine learning (ML) solutions to ID. Investigators have developed various ML-based ID solutions throughout the last few years. Machine learning (ML) has the potential to turn gadgets and human-made knowledge into valuable data. It may also be described as the capacity of a smart device to change or automate a condition or behaviour based on knowledge considered fundamental to an IoT setup. Categorisation, relapse, and thickness assessment have been performed using Ml methods. Artificial intelligence (AI) methods and techniques may be used to solve various industry problems. Additionally, machine learning may be used in the Internet of Things, providing practical support [20].

There are four kinds of algorithms in machine learning: directed, single, semi-supervised, and learning. Supervised learning is carried out when goals must be reached using a specific arrangement of data sources. First, the information is labelled, and then the preparation of named knowledge is performed according to this learning technique (with inputs and expected outputs). Data from free datasets and various classes characterise each Unsupervised Learning. This enables the placement of components for each Unsupervised Learning: Nature only offers contributions unrelated to the desired aims in unsupervised learning situations. No specific data is required, and it may look for patterns in unlabeled data and assemble data for various groupings [21].

In contrast to the emphasis on information research on supervised learning and unsupervised techniques, fortification learning is more heavily weighted on correlated and dynamic challenges. Accessible information is the underlying concept that underpins the ML strategy's order and decision-making. Guided learning is employed when the kind of knowledge and the ideal output (names) are already known. For optimal yields, this system is ready to plan contributions. Some examples of administered learning procedures are relapse and arrangement, where relapse works with unceasing yields and order works with discrete ones, respectively. Support vector regression (SVR), direct relapse, and polynomial depression are a few of the often-utilised depression approaches. A broad view of several deep learning technologies generally applied to healthcare systems was presented [12]. Generally, an ANN achieves the opinions through iterative training known as back-propagation but lacks generalisation for supervised tasks [13]. By adding new hidden layers, deep learning covers ANN assembly to DNNs for enhanced generalisation, which excerpts data features and acquires demonstrations with thousands or millions of constraints [14]. The discoveries of computational hardware and software infrastructures principally sped up the expansion of deep learning by increasing the size and deepness of DNN prototypes in the last two decades. Standard software outlines to implement deep learning systems embrace Theano [15], Caffe [16], TensorFlow [17], CNTK [18], and PyTorch [19]. These structures support numerous programming languages and hardware acceleration, which assist people in professionally shaping DNN prototypes.

III. PROPOSED METHODOLOGY

The advancements in deep learning are significant due to their potential to extend and incorporate diverse services across multiple domains. The primary motivation for developing this application stems from the potential to effectively utilise large volumes of data for data analytics, facilitated by recent processing and storage capabilities developments. The healthcare industry is a critical area in which using deep learning algorithms for prediction and estimate is significant. Multiple healthcare institutions have implemented various phases of deep learning methodologies. The clinical healthcare unit is the vital region in which decision-making and early precaution aid a lot in addressing complications of diseases or avoiding the occurrence of risky diseases such as diabetes. This section presents an overview of the proposed research methodology employed to enhance the efficiency of smart healthcare systems.

This section further elucidates the specifics of the implementation plan. The primary factor driving the extensive implementation of deep learning is its inherent ability to process and manage large quantities of data effectively. With the exponential growth of data generation, traditional approaches to data processing and analytics are becoming increasingly ineffective. Nevertheless, deep learning, characterised by its multilayer neural network utilisation, excels in data-rich situations and extracting significant patterns and insights from such datasets. Combining these capabilities and recent advancements in computing power and storage options has driven deep learning to the forefront of scientific advancements. The utilisation of deep learning has demonstrated significant advantages across several domains, with the field of healthcare emerging as a notable beneficiary of its positive impacts. In an industry where accurate and timely forecasting can determine the difference between success and failure, applying deep learning algorithms has been shown to have a significant and positive effect.

The implications of these developments are far-reaching and hold considerable importance, encompassing the capacity to predict disease outbreaks and customise patient care at an individualised level. Numerous modern healthcare systems have successfully included diverse tiers of deep learning techniques in their operational frameworks. The scope of these integrations encompasses primary data analysis and advanced predictive modelling for intricate diseases. The clinical healthcare unit is a significant focus point within these applications. In this context, judgments must be made carefully to avoid potential errors. The importance of early identification, proactive interventions, and precision decision-making cannot be overstated. For diseases like diabetes, which can have severe complications if not controlled or recognised early, such enhanced analytical techniques can avoid onset, decrease risks, and improve patient outcomes. This study extensively investigates the convergence of deep learning and healthcare.

This study aims to provide a comprehensive analysis of the approaches that can potentially improve the performance of smart healthcare systems, hence increasing their responsiveness, accuracy, and efficiency. The strategy outlined in this part is not solely a theoretical concept but a real framework that may be feasibly implemented in healthcare units worldwide. Moreover, our analysis extends beyond the mere proposition of a methodology, delving into the complexities associated with its practical use. Acknowledging the fundamental principle that the quality of an idea is contingent upon its successful implementation, our research endeavours to furnish a thorough and all-encompassing plan for execution. The proposed plan encompasses the necessary conditions, a systematic procedure, potential obstacles, and corresponding resolutions, providing healthcare units that adopt our technique with a well-defined roadmap to adhere to. The ongoing development and transformation of deep learning have significant implications for various sectors. Notably, integrating deep learning with the healthcare field can revolutionise medical practices, resulting in a future characterised by enhanced personalisation, predictive capabilities, and precision in healthcare delivery. The objective of our study is to contribute to the realisation of a future in which technology and healthcare intersect to establish a comprehensive wellness ecosystem. The subsection includes:

Data Set Extraction : This section contains the framework of the proposed technique deployed for improving the performance of smart healthcare systems. This section also discusses the details of the implementation plan. The subsection includes:

In the dynamic realm of healthcare, incorporating state-of-the-art technology plays a pivotal role in augmenting the calibre of care and resulting achievements. This section presents a complete framework designed to enhance the effectiveness of smart healthcare systems. At the core of this undertaking is a carefully curated dataset supplied by the renowned "National Institute of Diabetes and Digestive and Kidney Diseases." The information under consideration consists of precise measurements that facilitate the identification of potential cases of diabetes. Significantly, the compilation of data in this study targets explicitly female patients of Pima Indian heritage who are 21 years of age or older, to conduct a thorough and accurate analysis. The subsections of the document provide a comprehensive overview of the dataset's origin, emphasising its credibility. Additionally, the varied properties of the dataset are discussed, highlighting their significance in diabetes diagnostics. The logic behind the dataset's unique demographic focus is also explained, providing a clear justification. Lastly, a methodical implementation strategy for the suggested technique is outlined. This section presents a comprehensive framework to enhance healthcare systems, focusing on increasing their predictive capabilities, personalisation, and accuracy.

The data set comprises 768 observations and nine (9) variables. It is available in the package mlbench.
Data Descriptions for the nine (9) variables are as follows.

pregnant - Number of times pregnant glucose - Plasma glucose concentration (glucose tolerance test) pressure - Diastolic blood pressure (mm Hg) triceps - Triceps skin fold thickness (mm) insulin - 2-hour serum insulin (mu U/ml) mass - Body mass index (weight in kg/(height in m)^2) pedigree - Diabetes pedigree function Age - Age (years) diabetes - Class variable (test for diabetes)

The dataset can be taken from the below link.
https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database
Software requirement:
Python 3.

9.7

Jupyter Notebook 6.4. Here, the necessary packages are imported using CSV data.

Figure 1: Imported Python packages

Importing Packages

Here, we import the necessary pandas packages for Machine Learning

```
import os
import pandas as pd
import numpy as np
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline

#ignore warning messages
import warnings
warnings.filterwarnings('ignore')
```

Adding CSV data in our project

Using pandas.read_csv(PATH) function to add csv dataset from our local system

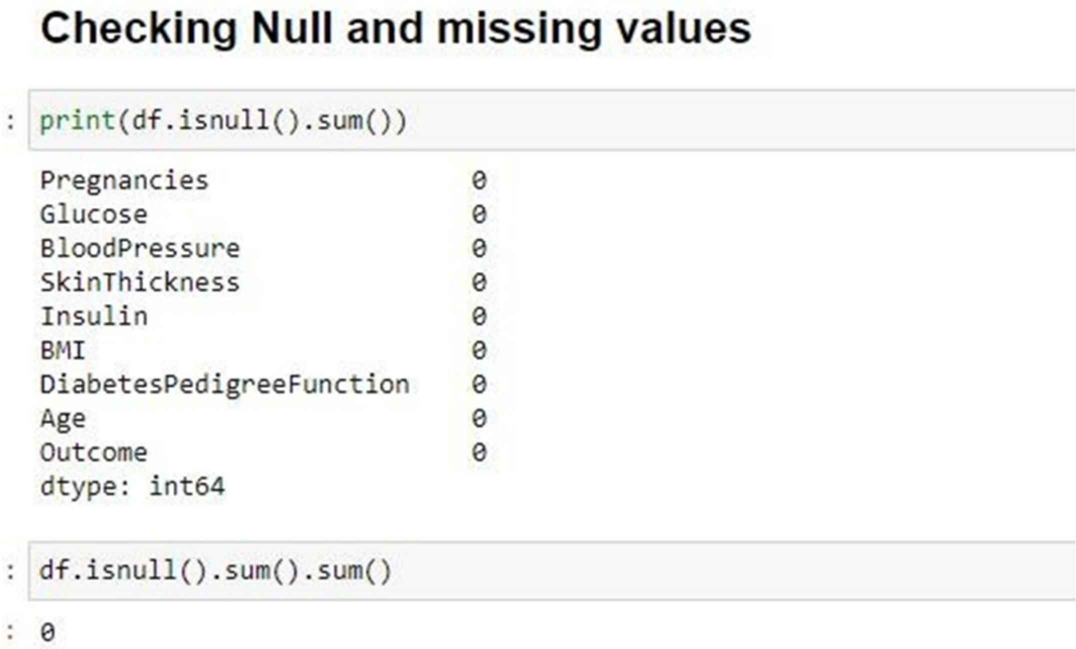
```
df=pd.read_csv("diabetes.csv")
```

1.1. Data Pre-processing

Most data scientists allocate much of their time to the data cleaning and pre-processing stage, as practical datasets consistently present complex challenges.

Within the complex domain of data science, significant attention and effort are dedicated to the initial phases of data management. Data cleansing and pre-processing are not solely preliminary procedures but the fundamental basis for the complete analytical framework. Because real-world datasets are rarely perfect, they inevitably present numerous difficulties that require careful examination and purification. Before delving into the extensive realm of data, it is imperative to adopt a systematic methodology. An essential initial phase is comprehending the underlying mechanisms contributing to data generation. The capacity to identify and understand the key factors that drive business outcomes can significantly impact the quality of data and its subsequent usefulness. The potential results can range from using graphical visualisations to enhance immediate comprehension to developing advanced predictive models that anticipate future patterns by analysing past data. Interacting with business teams may significantly enhance the context of numerical datasets, providing additional depth and insight.

A commonly observed issue in datasets is missing or misplaced values. Such holes reflect an absence of data and might influence interpretations and conclusions if not treated correctly. Many machine learning algorithms are typically ill-equipped to manage these gaps in data, resulting in biased outcomes. However, simply recognising the presence of missing data only represents a superficial understanding of the issue. It is of utmost necessity to conduct a thorough investigation to identify the underlying factors contributing to these instances of absence. Identifying the reasons behind the absence of particular data points and precisely identifying their positions within datasets can yield significant insights into the broader context. The comprehensive management of data, encompassing tasks such as data cleansing and data comprehension, is fundamentally essential to leverage the potential of data analytics fully.



Checking Null and missing values

```
: print(df.isnull().sum())

Pregnancies                 0
Glucose                     0
BloodPressure               0
SkinThickness               0
Insulin                     0
BMI                         0
DiabetesPedigreeFunction    0
Age                         0
Outcome                     0
dtype: int64

: df.isnull().sum().sum()

: 0
```

Figure 2: Pre-processing of data

Outlier Detection : The algorithms belonging to the nearest-neighbour family are based on a fundamental principle: data points comparable to each other are typically located nearby within the data space, but dissimilar points, often referred to as outliers, tend to be located far away from any cluster of similar observations.

The k-Nearest Neighbours (kNN) algorithm, although categorised as a supervised machine learning technique, takes an unsupervised approach when employed for anomaly detection. This distinctive attribute arises from the inherent nature of kNN, which does not acquire knowledge in the traditional sense. Instead of depending on pre-labelled data instances categorised as "outliers" or "non-outliers," this approach classifies data solely based on proximity and predefined threshold values.

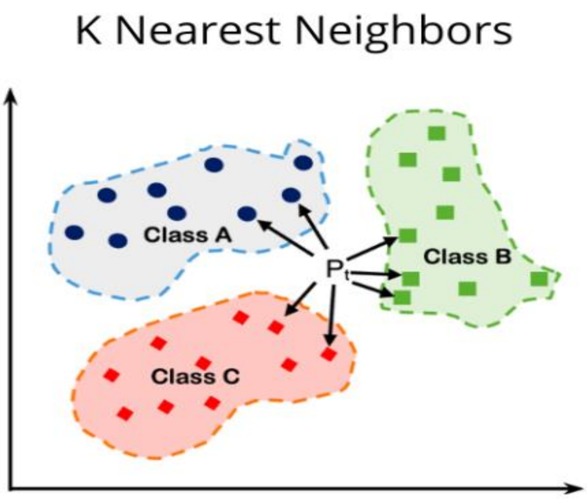


Figure 3: K Nearest Neighbors

However, the method of determining these threshold values is not always straightforward. Frequently, data scientists rely on empirical approaches, establishing these threshold values to some extent in a subjective manner. Once these thresholds are surpassed, any data point is considered aberrant. Instead of pre-existing labels, the tremendous flexibility and reliance on distance measures render kNN a potent yet demanding algorithm. Fundamentally, the k-nearest neighbours (kNN) algorithm relies on utilising similarities and distances between data points. However, when applying this algorithm for anomaly detection, it is crucial to comprehend the specific dataset comprehensively and make careful decisions regarding threshold values. This is essential to achieve precise identification of outliers.

```
X = df.values
```

```
# instantiate model
nbrs = NearestNeighbors(n_neighbors = 3)
# fit model
nbrs.fit(X)
```

```
NearestNeighbors(n_neighbors=3)
```

```
# distances and indexes of k-neaighbors from model outputs
distances, indexes = nbrs.kneighbors(X)
# plot mean of k-distances of each observation
plt.plot(distances.mean(axis =1))
```

Figure 4: kNN-based outlier detection.

Feature Extraction : The "Feature Selection by Pearson Correlation Coefficient" procedure provides a methodical strategy for identifying and utilising the most pertinent qualities in a given dataset. The key focus of this approach is the utilisation of filtering techniques to choose and retain the most influential elements in the data modelling process. The primary instrument in this methodology is the correlation matrix, utilised to ascertain the magnitude and orientation of the association between different variables. The Pearson correlation is the most generally employed among several methods to quantify correlation. This specific coefficient offers a metric that quantifies the extent of the linear association between two variables of a quantitative kind. To provide additional clarification, the term "correlation" within the domain of statistics and data analysis refers to the evaluation of the degree to which two variables exhibit a close relationship in terms of their movements relative to each other. This analysis offers valuable information regarding the strength and nature of the relationship between the variables in question. For instance, in the context of the above illustration involving the price of an automobile and its corresponding engine size, a positive correlation may be observed, suggesting that an increase in engine size is associated with a corresponding rise in the car's price. Acknowledging that the Pearson correlation coefficient values span from -1 to 1 is essential. A value near 1 indicates a robust positive correlation, indicating that both variables exhibit parallel movements. On the contrary, a score near -1 signifies a robust negative correlation, implying that when one variable experiences an increase, the other variable is likely to exhibit a drop. A number near 0 frequently signifies the absence of correlation, indicating that the variables do not necessarily exhibit a simultaneous movement. Using the Pearson Correlation Coefficient, data scientists can employ feature selection techniques to discover and maintain solely those variables that exhibit significant connections. This process enables the construction of models that are not only efficient but also less susceptible to problems such as multicollinearity. This methodology guarantees the development of streamlined models that exhibit enhanced interpretability and prediction accuracy.

Split the data into dependent and independent, then create a random forest model: The phrase "independent variable" denotes a distinct and autonomous characteristic. It denotes a variable that remains unaltered by variations in other variables within a specified context or research endeavour. Variables of this sort, by their predictive qualities, assume a fundamental function in many experiments and research endeavours, frequently exerting influence over the trajectory or path of the inquiry. Due to their inherent capacity for prediction, independent variables are frequently used interchangeably as predictors or features.

Correlation with Heatmap

```
]: #correlations of each features in Data
   corrmat = data.corr()
   top_corr_features = corrmat.index
   plt.figure(figsize=(12,9))
   #plot heat map
   g=sns.heatmap(data[top_corr_features].corr(),annot=True,cmap="RdYlGn")
```

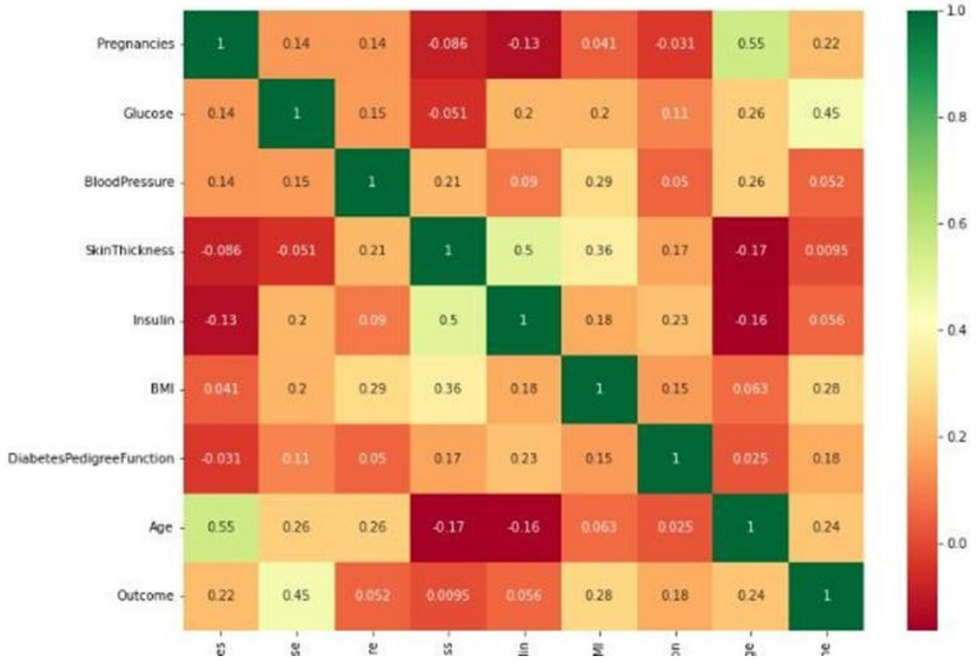


Figure 6: Parameter comparison table

In contrast, the dependent variable assumes the role of an outcome or result in the context of research and experiments. The dependent variable is a construct that signifies the topic of measurement or observation, with its value being established or changed by the actions or interactions of the independent variable(s). The dependent variable measures the potential outcomes or consequences resulting from changes in the independent variable. The dependent variable can be defined as the outcome or result of any actions, modifications, or manipulations performed on the independent variable. In an empirical investigation exploring the influence of exercise as an independent variable on weight loss as a dependent variable, the magnitude of weight loss would exhibit variability contingent upon the intensity or nature of the exercise regimen undertaken by the participants.

The distinction between independent and dependent variables holds significant importance within research contexts. The independent variable establishes the context by introducing a condition or component deliberately altered or classified to examine its impact on the dependent variable. In the research context, the dependent variable is crucial in providing significant and practical findings by shedding light on the impacts or implications of the manipulations being studied. Collectively, these variables provide the fundamental components of hypothesis testing, enabling researchers to make inferences, substantiate ideas, and reveal novel perspectives within their particular domains of inquiry.

Split of Dependent variable and independent variables

```
x=data.drop(columns=['Outcome'])
y=data['Outcome']
```

Figure 7: Dependent and Independent variables

Feature Scaling --> StandardScaler

StandardScaler: It converts the data in such a way that it has a mean of 0 and a standard deviation of 1. In short, it standardises the data. Standardisation is suitable for data that has negative values. It assembles the data in a standard normal distribution.
StandardScaler follows Standard Normal Distribution (SND). Therefore, it makes mean = 0 and scales the data to unit variance.
In the incidence of outliers, StandardScaler does not promise sensible feature scales due to the impact of the outliers while calculating the practical mean and standard deviation. This hints at shrinkage in the series of feature values.
By using RobustScaler(), it is possible to remove the outliers and then use either StandardScaler or MinMaxScaler to pre-process the dataset.

Feature Scaling --> StandardScaler

StandardScaler : It transforms the data in such a manner that it has mean as 0 and standard deviation as 1. In short, it standardizes the data. Standardization is useful for data which has negative values. It arranges the data in a standard normal distribution.

```
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x= sc.fit_transform(x)
```

```
y = y.to_numpy()
```

Figure 8: Feature Scaling

Split the data into train and test, then create the model:

Split the dataset into two pieces: training and testing sets. This consists of randomly selecting about 75% (you can vary this) of the rows, putting them into your training set, and putting the remaining 25% into your test set. Note that the colours in "Features" and "Target" indicate where their data will go ("X_train", "X_test", "y_train", "y_test") for a particular train test split.

https://towardsdatascience.com/understanding-train-test-split-scikit- learn-python-ea676d5e3d1

Train and Test data split

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(x,y,test_size=0.3,random_state=0)
```

Figure 9: Train and test data split

Implementation of Proposed System
Create the Existing model in a Random Forest: The ensemble learning technique known as random forest has emerged as a leading approach in classification methods. The strategy is characterised by its capacity to create a diversified and resilient collection of decision trees, known as a "forest," by combining several decision trees. This technique is based on the fundamental premise of aggregating these trees to enhance the overall decision-making process. The random forest operates based on two fundamental techniques: bagging and feature randomness. Bagging, also known as bootstrap aggregating, generates several datasets by randomly selecting samples with replacements from the original dataset. This practice guarantees that every decision tree inside the forest is trained on a marginally distinct sample of data, augmenting the model's variety and resilience. Simultaneously, incorporating feature randomness guarantees that every division in the decision tree is executed using a randomly selected subset of characteristics instead of evaluating all the available features. This additional procedure serves to diminish the association observed among trees further. The incredible potency of random forest stems from its unique approach to concluding. Instead of depending on the forecast of a single tree, the model aggregates the outputs from all the trees within the forest. In classification tasks, the final prediction is determined using a majority vote procedure, wherein the class that garners the most votes from individual trees is selected. In regression problems, the outcome is determined by the average prediction of all the trees.

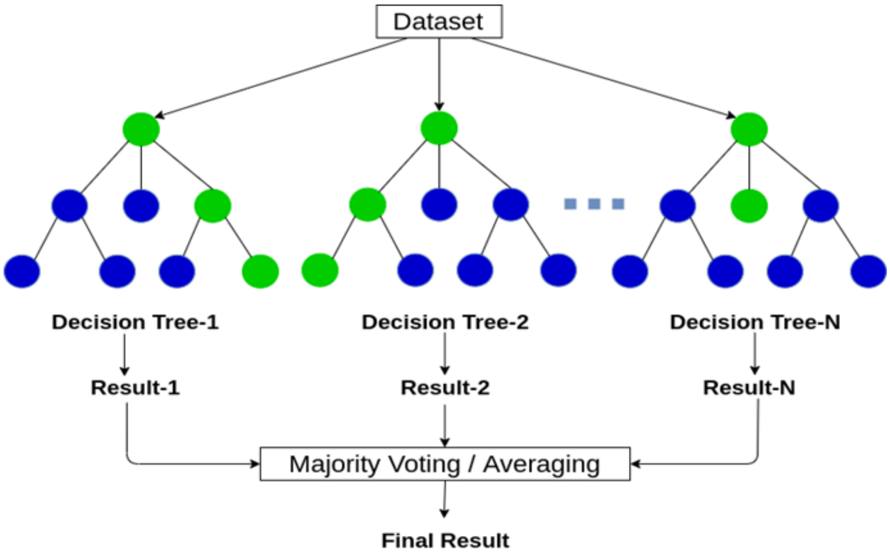


Figure 10: Random Forest Model

The random forest algorithm exhibits notable strengths in terms of scalability and adaptability. The system adeptly handles large datasets containing a wide range of variables, including those that number in the thousands. This characteristic renders it flexible and relevant in various fields and datasets. In brief, the random forest algorithm combines the capabilities of many decision trees, utilises bagging and feature randomness approaches, and applies aggregation methods to generate predictions that frequently exhibit higher accuracy and reliability than those produced by individual trees.

Model Building

Random Forest

```
from sklearn.ensemble import RandomForestClassifier

rf = RandomForestClassifier(n_estimators=100,random_state=64,criterion='gini',max_depth=None,  min_samples_split=6,
    min_samples_leaf=1,  min_weight_fraction_leaf=0.5, max_features='auto',  max_leaf_nodes=None,  min_impurity_decrease=0.4,
    bootstrap=True,  oob_score=False)

rf.fit( X_train, y_train)
predict_rf = rf.predict(X_test)
```

```
from sklearn.metrics import accuracy_score
existing_accuracy = accuracy_score(np.round(y_test),np.round(predict_rf))
print ('Existing Accuracy : ',existing_accuracy )
```

```
Existing Accuracy :  0.6857142857142857
```

Figure 11: Random forest method classifier

Create the Existing model in a Convolutional Neural Network. Artificial Intelligence has been perceived as an immense progression in linking the gap between the abilities of humans and machines. Researchers and supporters alike work on abundant phases of the field to create remarkable things that occur. One of many such extents is the field of Computer Vision.

The CNN approach performs sophisticated means of computations on vast amounts of data. This is a machine-learning tool that works on the structure and function of the human brain. It has three specific layers: an input layer, a hidden layer, and an output layer. Once the datasets are, the results are validated, and performance metrics are measured. The schema for this field is to empower machines to assess the world as humans do, notice it in an analogous approach, and even use the information for an assembly of tasks like Image and video recognition, Image Analysis and classification, Media Recreation, Recommendation Systems, Natural Language Processing, etc. The improvements in Computer Vision with Deep Learning have been created and achieved with time, principally over one specific process — a Convolutional Neural Network. CNN is an advanced form of multilayer neural network which mainly entails input, hidden, and output layers. A neuron is an elementary form of the information processing unit of a CNN, which comprises a system of synapses or links. Every connection is classified as weight W1, W2,..., Wm, an adder function (linear combiner Eq. (1)) which calculates the weighted sum of the inputs,

(1)Moreover, activation function f() is used for controlling the amplitude of the neuron's output. The typical model of neuron could be observed as Eq. (2)

(2) Where Xi(i=1,2,3,...,n) specifies the input vector. Wi signifies the weights among two (2) consecutive neurons. θ is the threshold. f() is the activations function, the frequently used function is the sigmoid function Eq.

(3) y is the desired output.

Construction of Novel Cascaded CNN Model : Create a proposed Cascaded Convolution Neural Network. The CNN model performs the one-step aliasing, which cannot be iteratively deployed to de-alias. While CNNs are powerful enough to learn single-step reconstruction, such networks could create some signs of overfitting concern unless there is a tremendous amount of training data. Moreover, these networks require more extended time and cautious fine-tuning steps.

.

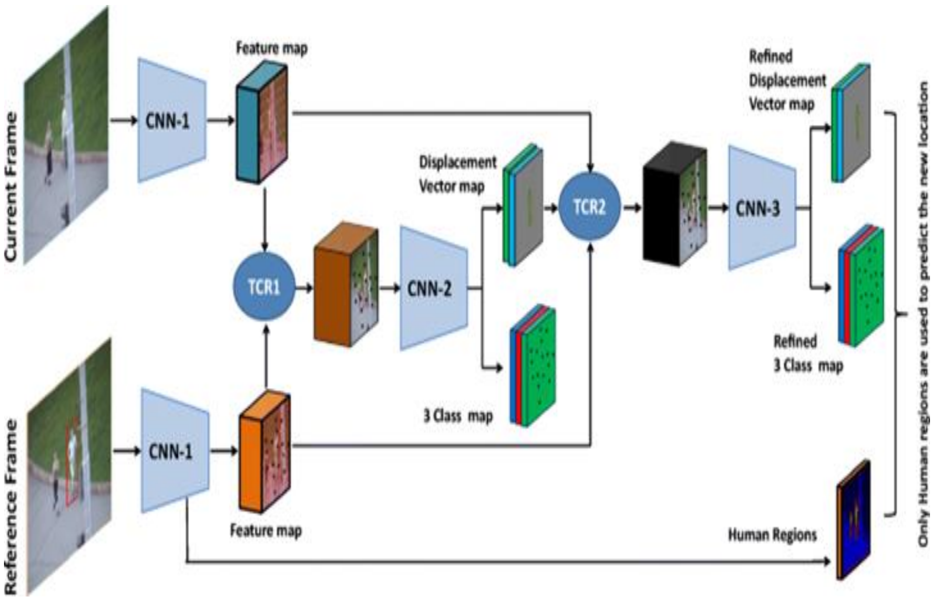


Figure 13: Cascaded CNN model

```
model_1D = Input((8, 1))
# 1 CNN
model_1D = Conv1D(512,kernel_size= 2 , strides=1, activation='relu')(model_1D)
model_1D = MaxPooling1D(pool_size= 1, strides=2)(model_1D)
model_1D=BatchNormalization()(model_1D)

model_1D = Conv1D(256,kernel_size= 2 , strides=1, activation='relu')(model_1D)
model_1D = MaxPooling1D(pool_size= 1, strides=2)(model_1D)
model_1D=BatchNormalization()(model_1D)

model_1D = Conv1D(124, kernel_size= 2 , strides=1, activation='relu')(model_1D)
model_1D = MaxPooling1D(pool_size= 1, strides=2)(model_1D)
model_1D=BatchNormalization()(model_1D)

model_1D = Dropout(0.2)(model_1D)


model_1D = Flatten()(model_1D)
model_1D = Dense(8 , activation='relu')(model_1D)


#2nd CNN
model_2D = Input((8, 1))

model_2D = Conv1D(512,kernel_size= 2 , strides=1, activation='relu')(model_2D)
model_2D = MaxPooling1D(pool_size=1, strides=1)(model_2D)
model_2D=BatchNormalization()(model_2D)

model_2D =  Conv1D(256,kernel_size= 2 , strides=1, activation='relu')(model_2D)
model_2D = MaxPooling1D(pool_size=1, strides=1)(model_2D)
model_2D=BatchNormalization()(model_2D)
```

Figure 14: Novel cascaded CNN model

Therefore, it is essential to use CNN as a reconstruction-based system. The most straightforward approach is to train a second form of CNN, which could learn to reconstruct from the output parameters of the first CNN. It is more like concatenating a new CNN model on the output base of the previous one to create an advanced model capable of reconstructing the deep networks. This is termed a cascaded convolution neural network.

Bayesian Optimisation of CNN Model

Using Bayesian Optimization for CNN model: Optimising the hyperparameter is an enhanced approach for this proposed model, and some of the traditional approaches include grid, random, and Bayesian optimisation. Despite its high time complexity level, the Bayesian approach is considered the best fit. The Bayesian optimisation method has become an effective way to optimise further the hyperparameters in any such machine learning algorithm.

For a Bayesian approach, two more choices are required, including predictions regarding the functions to be optimised, and then, the acquisition function is deployed to create a utility-based function to estimate subsequent points to evaluate. Here, a prior function should be considered over an objective function. Then, the probability of improvement and the expected improvement are performed, and the final output could find the optimal solution.

```
def black_box_function(C):
    # C: SVC hyper parameter to optimize for.

    model_final.fit([X_train,X_train],y_train, batch_size=5,epochs=50,validation_split=0.2)
    proposed =model_final.predict([X_test,X_test])
    proposed=np.argmax(proposed,axis=1)
    f=accuracy_score(np.round(y_test),np.round(proposed))
    return f
# Set range of C to optimize for.
# bayes_opt requires this to be a dictionary.
pbounds = {"C": [0.1, 10]}
# Create a BayesianOptimization optimizer,
# and optimize the given black_box_function.
optimizer = BayesianOptimization(f = black_box_function,
                                 pbounds = pbounds, verbose = 2,
                                 random_state = 4)
optimizer.maximize(init_points = 5, n_iter = 10)
print("Best result: {}; f(x) = {}.".format(optimizer.max["params"], optimizer.max["target"]))
```

Figure 15: Optimisation using the Bayesian model.

Convolutional Neural Networks (CNNs) Vs Random Forests (RFs) : CNNs are generally superior for image and video processing tasks because they capture spatial hierarchies in data. They are the go-to for computer vision applications, such as image classification, object detection, and segmentation. On the other hand, RFs are excellent for problems involving structured data, like tabular datasets found in various domains, including finance, biology, and marketing. They are preferred when the relationship between features is more important than the spatial or temporal structure in the data. CNNs are excellent in processing data with a topology resembling a grid, like pictures, by employing layers of convolutions to identify features automatically and pyramidally. As a result, CNNs perform better on tasks involving visual identification. RFs, on the other hand, combine judgments from several decision trees to offer resilience and avoid overfitting, making them excellent for tasks involving regression and classification on tabular data. Because of their decision-path structure, RFs are faster to train than CNNs and provide more interpretable results. CNNs, on the other hand, require a significant amount of computer power and data. The appropriateness depends on the context: RFs work well with various datasets, whereas CNNs perform better with high-dimensional data, particularly in cases where computational resources are few. In summary, there is no objective "best" model; it depends on the task at hand. CNNs are favoured for complex pattern recognition in visual data, while RFs are versatile, interpretable, and efficient for diverse predictive tasks with structured data.

Performance Evaluation : The classification report visualiser shows the proposed research model's precision, recall, F1, and support scores.

Classification Report

```
from sklearn.metrics import accuracy_score
proposed = model_final.predict([X_test,X_test])
proposed=np.argmax(proposed,axis=1)

from sklearn.metrics import classification_report

print(classification_report(y_test, np.round(proposed)))
```

Figure 16: Classification report

	Precision	Recall	F1Score	Support
0	0.84	0.92	0.87	144
1	0.77	0.61	0.68	66
Accuracy	-	-	0.82	210
Macro average	0.80	0.76	0.78	210
Weighted average	0.81	0.82	0.81	210

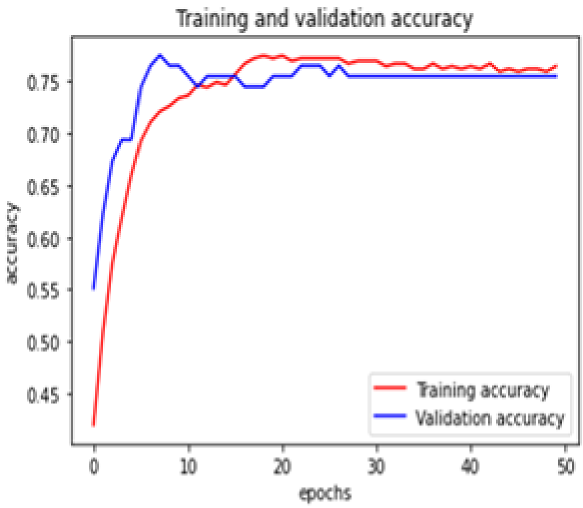


Figure 17: Training and validation of accuracy

Figure 17 describes the training and validation accuracy of the proposed method model and a new dataset model. Figure 18 denotes the loss parameter of the training and validated parameter, and the loss value is found to be minimum, which shows the accuracy of the proposed approach (Figure 18).

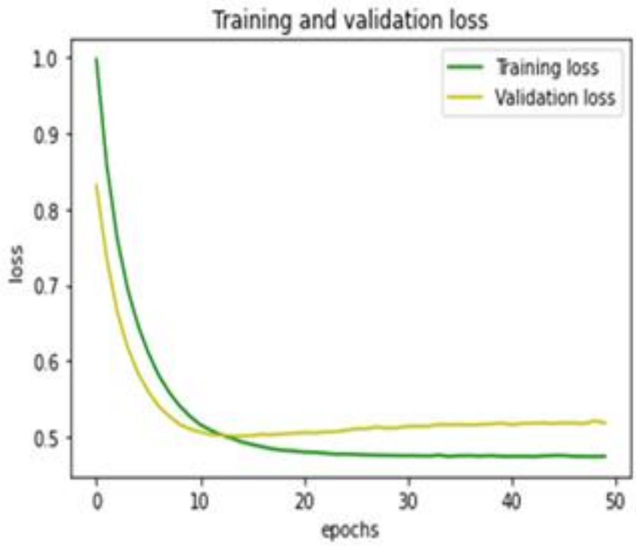


Figure 18: Training and validation of losses

Comparison of Proposed Work with Existing System

Comparing All Models

```
print('Proposed Accuracy ',Accuracy)
print('Existing Accuracy ',existing_accuracy)
print('Existing Accuracy2 ',Existing_Accuracy1)
```

```
Proposed Accuracy  0.82
Existing Accuracy  0.6857142857142857
Existing Accuracy2  0.6857142857142857
```

```
data = {'Existing Accuracy_RF':existing_accuracy,'Existing Accuracy_CNN ':Existing_Accuracy1, 'Proposed Accuracy_CCNN':Accuracy}

label = list(data.keys())
model = list(data.values())
color = ['red','blue','green']
fig = plt.figure(figsize = (12, 6))

plt.bar(label, model, color =color,width = 0.4)

plt.xlabel("Models")
plt.ylabel("Score level")
plt.title("Accuracy comparison")
plt.show()
```

Figure 19: Comparison of proposed and existing models

After comparing existing methods with the proposed technique, the graph (Figure 19) shows that the proposed method, by incorporating a cascaded convolution neural network-aided prediction system for diabetes patients, is found to be more appropriate in healthcare systems. The comparison is shown in Figure 20, which mentions the accuracy parameter. Using the proposed methodology, diabetes patients could be identified more precisely, and it would benefit healthcare systems by enhancing their service in a good way.

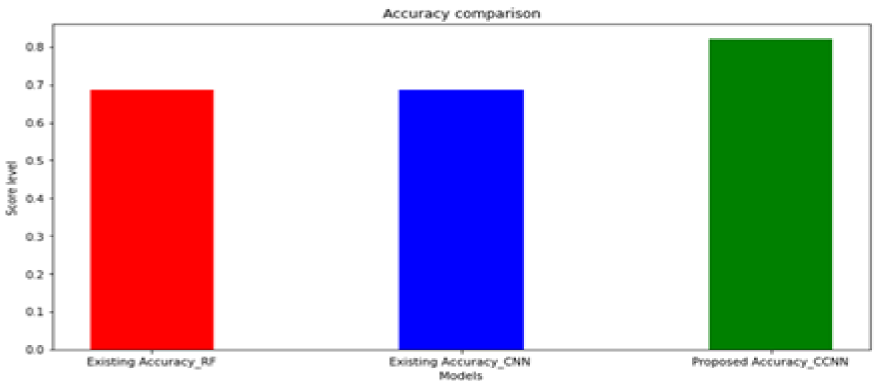


Figure 20: Accuracy comparison graph of proposed and existing approach

IV. CONCLUSIONS

The field of medical research is in a constant state of change, with technology playing a crucial role in transforming the methods of disease diagnosis and prediction. Our current study is evidence of our investigation into the complexities of utilising deep learning methodologies to construct a predictive model for diabetes, a prevalent ailment with a global impact on millions of individuals. A methodical and comprehensive approach characterised the methodology employed for this ambitious undertaking. The initial step in our research involved the acquisition of a comprehensive dataset consisting of pertinent information that would serve as the foundation for our predictive model. However, the raw version of this data frequently exhibits discrepancies, incomplete values, or anomalous observations. Concerning the matter at hand, the subsequent stage, known as pre-processing, involves the cleansing and structuring of the data, rendering it appropriate for analysis. Feature selection is crucial in ensuring the efficacy of machine learning or deep learning models. In this study, we conducted variable selection and isolation to identify the most relevant factors from the dataset that significantly influence the outcome variable, specifically diabetes prediction.

Following selecting the most pertinent characteristics, we proceeded to the step of feature scaling. Considering that deep learning models exhibit sensitivity to the scale of input data, it becomes imperative to standardise and normalise the data to enable the model to process it effectively and impartially. The core focus of our research was centred on the categorisation phase. Our model underwent benchmarking against existing predictive techniques but also against established classifiers such as the random model classifier and the convolution neural network (CNN). Utilising a comparative strategy yielded valuable insights regarding each respective method's relative merits and limitations. Nevertheless, our efforts did not cease at that point. The authors of this study have acknowledged the potential of convolutional neural networks (CNNs) and have proposed a cascaded convolutional neural network. To further improve its capabilities, they have employed a Bayesian optimisation approach. Various methodologies were integrated to capitalise on Convolutional Neural Networks (CNN) advantages and Bayesian optimisation. This approach was employed to ensure that the predictions generated by our model exhibited both high accuracy and reliability. Assessing the efficacy of our model was of utmost importance. Measurements included precision, recall, f1 score, and support. Each indicator provided distinct viewpoints on the model's performance, encompassing its capacity to forecast positive situations accurately and its effectiveness in differentiating between true positives and negatives. In conclusion, our study highlights the possibility of employing deep learning methodologies for medical prognostication. Although our primary emphasis was on diabetes, the approaches and insights obtained from this study have the potential to serve as a fundamental basis for predicting various other medical diseases. Consequently, this represents a notable advancement in personalised and predictive healthcare.

REFERENCES

[1] V. Agrawal, P. Singh, and S. Sneha, "Hyperglycemia Prediction Using ML: A Probabilistic Approach," In International Conference on Advances in Computing and Data Sciences, vol. 1046, pp. 304-312, April 2019, Springer, Singapore.
[2] P. Singh and N. Singh, "Blockchain with IoT and AI: A Review of Agriculture and Healthcare," International Journal of Applied Evolutionary Computation (IJAEC), vol. 11, pp. 13-27, 2020.
[3] S. Ellahham, "Artificial Intelligence in Diabetes Care", The American Journal of Medicine, 2020.

[4] P. Singh, and R. Agrawal, "A customer-centric best-connected channel model for heterogeneous and IoT networks. Journal of Organizational and End User Computing (JOEUC), vol. 30, pp. 32-50, 2018.
[5] J. Chaki, S. T. Ganesh, S. K. Cidham and S. A. Theertan, "Machine learning and artificial intelligence-based Diabetes Mellitus detection and self-management: A systematic review," Journal of King Saud University-Computer and Information Sciences, 2020.
[6] Chauhan, T., Rawat, S., Malik, S., & Singh, P. (2021, March). Supervised and unsupervised machine learning-based review on diabetes care. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 581-585). IEEE.
[7] T. M. Alam, M. A. Iqbal, Y. Ali, A. Wahab, S. Ijaz, T. I. Baig, and Z. Abbas, "A model for early prediction of diabetes," Informatics in Medicine Unlocked, 16, 100204, 2019.
[8] D. Sisodia and D. S. Sisodia, "Prediction of diabetes using classification algorithms," Procedia computer science, vol. 132, pp. 1578-1585, 2018.
[9] M. Alehegn, R. Joshi and P. Mulay, "Analysis and prediction of diabetes mellitus using machine learning algorithm," International Journal of Pure and Applied Mathematics, vol. 118, pp. 871-878, 2018.
[10] N. Sneha and T. Gangil, "Analysis of diabetes mellitus for early prediction using optimal features selection," Journal of Big Data, vol. 6, pp. 13,2019.
[11] N. P. Tigga and S. Garg, "Prediction of Type 2 Diabetes using Machine Learning Classification Methods," Procedia Computer Science, vol. 167, pp. 706-716, 2020.
[12] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning. MIT Press, 2016.
[13] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," nature, vol. 323, no. 6088, pp. 533–536, 1986.
[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.
[15] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley, and Y. Bengio, "Theano: a CPU and GPU math expression compiler," in Proceedings of the Python for scientific computing conference (SciPy), vol. 4, no. 3. Austin, TX, 2010.
[16] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in Proceedings of the 22nd ACM international conference on Multimedia, 2014, pp. 675–678.
[17] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., "Tensorflow: A system for large-scale machine learning," in 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16), 2016, pp. 265–283.
[18] F. Seide and A. Agarwal, "CNTK: Microsoft's open-source deep-learning toolkit," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 2135–2135.
[19] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al., "Pytorch: An imperative style, high-performance deep learning library," in Advances in Neural Information Processing Systems, 2019, pp. 8024–8035.
[20] Ullah, I., & Mahmoud, Q. H. (2020). A Two-Level Flow-Based Anomalous Activity Detection System for IoT Networks. Electronics, 9(3), 530.
[21] Alhakami, W., ALharbi, A., Bourouis, S., Alroobaea, R., & Bouguila, N. (2019). Network anomaly intrusion detection using a nonparametric Bayesian approach and feature selection. IEEE Access, 7, 52181-52190.
[22] Mendonça, R. V., Teodoro, A. A., Rosa, R. L., Saadi, M., Melgarejo, D. C., Nardelli, P. H., & Rodríguez, D. Z. (2021). Intrusion Detection System Based on Fast Hierarchical Deep Convolutional Neural Network. IEEE Access, 9, 61024-61034.
[23] Almiani, M., AbuGhazleh, A., Al-Rahayfeh, A., Atiewi, S., & Razaque, A. (2019). Deep recurrent neural network for IoT intrusion detection system. Simulation Modelling Practice and Theory, 102031.
[24] Saranya, T., Sridevi, S., Deisy, C., Chung, T. D., & Khan, M. A. (2020). Performance analysis of machine learning algorithms in intrusion detection system: A review. Procedia Computer Science, 171, 1251-1260.
[25] Aung, Y. Y., & Min, M. M. (2017, June). An analysis of random forest algorithm-based network intrusion detection system. In 2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD) (pp. 127-132). IEEE.
[26] Rahaman A, Islam M, Islam M, Sadi M, Nooruddin S. Developing IoT-based smart health monitoring systems: a review. Rev Intell Artif. 2019;33:435–40.
[27] Riazul Islam SM, Kwak Daehan, Humaun Kabir M, Hossain M, Kwak Kyung-Sup. The Internet of Things for health care: a comprehensive survey. IEEE Access. 2015;3:678–708.

[28] Lin T, Rivano H, Le Mouel F. A survey of smart parking solutions. IEEE Trans Intell Transp Syst. 2017;18:3229–53.
[29] Al-Ali AR, Zualkernan IA, Rashid M, Gupta R, Alikarar M. A smart home energy management system using IoT and big data analytics approach. IEEE Trans Consum Electron. 2017.
[30] Zanella A, Bui N, Castellani A, Vangelista L, Zorzi M. Internet of Things for smart cities. IEEE Internet Things J. 2014;1:22–32.

BIOGRAPHIES AND PHOTOGRAPHS

Short biographies (120-150 words) should be provided that detail the authors' education, work histories, and research interests. The authors' names are italicised. Small (3.5 X 4.8 cm), black-and-white pictures/digitised images of the authors can be included.

Note:

Since the Camera Ready copy of the paper is the final one, no further modification is entertained. So please make sure that the contents and format are fit for the Journal.