

Automated Disease Diagnosis and Monitoring System Using Deep Learning in IoT Healthcare Applications

Dr. Mrunali Sonwalkar

Assistant Professor, College of Engineering, Ambajogai, Dist. Beed- 431517, Maharashtra, India

Email: [mrunali.sonwalkar\[at\]gmail.com](mailto:mrunali.sonwalkar[at]gmail.com)

Abstract: *Technology has become a fundamental aspect of modern life, enabling the automation of previously labor-intensive processes and making them easier and more efficient worldwide. Artificial Intelligence (AI), the Internet of Things (IoT), blockchain, and Deep Learning (DL) have all made significant contributions to the advancement of several industries. In the medical industry, where exact diagnoses are required daily, traditional approaches are often inefficient due to errors and excessive delays. A smart healthcare system, on the other hand, can be constructed using recent advances in DL and IoT, which are useful for providing accurate diagnoses promptly. This study focuses on an automated disease diagnosis and monitoring system that employs advanced technologies, including DL and IoT. The dataset, based on patients' health parameters, is generated via a Google search. This dataset is used to test disease prediction results using DL models such as Convolutional Neural Network (CNN), Multi-Layer Perceptron (MLP), and Bidirectional Long Short-Term Memory (Bi-LSTM). The Bi-LSTM model is identified as the best model, with 99% accuracy. The Bi-LSTM is then deployed in the cloud. The health parameters of individuals are acquired using wearable sensors. The data is transmitted to the cloud, where the model analyzes it and predicts the individual's health status. If an irregularity is discovered, the cloud sends an alert message to the authorized doctor.*

Keywords: Healthcare, Internet of Things, Deep Learning, Disease, Sensor, Performance Measure

1. Introduction

The healthcare system is a critical component of any country, responsible for developing healthy citizens who contribute to national wealth [1]. However, in emerging countries such as India, the healthcare system is fragmented and varied. Providing timely, specialized medical care is crucial, particularly during crises such as pandemics or other potential future calamities, as India has experienced. Clinically, the capacity to effectively collect patient information for therapy and diagnosis is critical to providing adequate care [2]. Diagnoses and therapies must be adaptable and tailored for each patient. A rigorous, inflexible strategy with few options not only raises the risk of misdiagnosing potentially fatal illnesses, but it also quickly becomes obsolete. A speedy, efficient, and objective strategy to disease detection is required to ensure that patients receive prompt treatment. Advanced technologies such as Big Data, cloud computing, IoT, and edge computing have grown in popularity due to their adaptable reaction capabilities for individual applications [3].

Furthermore, the convergence of IoT and AI is causing a shift in the healthcare sector [4]. IoT provides chances to reduce the load on healthcare systems by allowing bidirectional communication between patients and caregivers, resulting in more accountable and value-based patient care. Combining IoT and AI in healthcare allows for automated medical records, disease prediction, and real-time patient monitoring. AI improves Machine Learning (ML) and task performance by studying previous data patterns, resulting in better medical treatments and outcomes. Integrating artificial intelligence into the Internet of Medical Things has the potential to improve diagnosis, treatment, and

illness management for both patients and healthcare providers.

This study introduces a novel DL-based algorithm that seamlessly integrates multiple high-dimensional data sources, each offering unique insights into diseases, to automatically generate significant medical insights. Our objective is to develop a cost-effective, reliable health monitoring system that leverages AI and IoT to track vital signs continuously and accurately forecast health outcomes.

2. Literature Survey

The research [5] aims to propose a healthcare model based on ML that accurately and early detects various diseases. This study predicts nine lethal diseases, including diabetes, hepatitis, heart disease, cancer, liver condition, dermatology, and thyroid, using seven ML classification algorithms. Four performance measures are used to evaluate the model's effectiveness. The Random Forest classifier identifies diseases with the most positive metrics. Paper [6] describes an idea for remote health monitoring in an IoT environment using a lightweight block encryption approach to protect health and medical data. This model combines lightweight secure block encryption to safeguard sensitive patient data and data mining methods to analyze biological data collected by smart IoT sensors. The research findings are utilized to assess the health status of patients. Given the limited resources of IoT platforms, lightweight block encryption methods are crucial for their efficient operation. Experimental outcome indicates that the K-star classification strategy outperforms other classifiers. Thus, the proposed methodology effectively develops a remote health monitoring system using secure IoT platforms.

The article [7] presents a novel model for healthcare disease diagnosis based on biological ECG signals, facilitated by the IoT and DL. The proposed approach applies DL models to biological ECG signals to detect cardiovascular diseases (CVDs). Additionally, the model employs a Bi-LSTM feature extraction method to derive informative features from the ECG signals. Hyperparameter optimizers such as the artificial flora optimization (AFO) method enhance the performance of the Bi-LSTM approach. Furthermore, a fuzzy classifier accurately categorizes ECG signals. Biomedical ECG data are utilized to validate the model's performance, and the results are thoroughly analyzed. The experiments clearly demonstrate that this model is the most accurate and superior option. The purpose of study [8] is to review existing data on cardiovascular diseases to predict and prevent the onset of heart disease at an earlier stage. The cloud-based database provides information about cardiac disease patients in Jammu and Kashmir, India. To predict cardiac diseases, stored data must be preprocessed and analyzed using ML methods. Dataset analysis employs a wide range of ML methods. According to performance measures reported by ML approaches, Naive Bayes outperforms other methods when parameters are not adjusted. This study systematically develops a model for real-time prediction of cardiovascular diseases using clinical data collected via IoT and existing medical sensors.

Paper [9] proposes an Equilibrium optimized lightweight prediction model for diagnosing chronic health status (kidney or heart). Data acquired by a sensor fixed in the patient's body can be transmitted to the cloud via a gateway. The proposed model initiates the classification process to forecast chronic diseases based on the sensor data collected. The model is developed following a phase of training and testing. Disease prediction relies on datasets for cardiovascular and chronic diseases. During training, preprocessed data is employed for classification purposes. After training, sensor

data stored on the cloud server is analyzed and categorized as normal or abnormal. In case of an abnormality, an alert message is sent to the doctor for appropriate patient treatment. Paper [10] proposes a method for predicting heart disease using the IoT and a hybrid ML algorithm combining Artificial Neural Network (ANN) and Recurrent Neural Network (RNN). This study includes several constraints-increment record, blood vessel hardness, increment pressure, etc.- providing ample room for improvement. The proposed RNN outperforms previous ML approaches in terms of organization and infection prediction accuracy. This approach incorporates synthetic data typically used for predicting the likelihood of cardiac disease. The pilot study demonstrates that the proposed technique is more effective than competing approaches.

3. Methodology

To design an automated healthcare system, we employed advanced technologies such as AI and IoT. The dataset was manually created using normal ranges obtained from Google. The created data underwent preprocessing steps like normalization and checking for missing values. After preprocessing, DL models were utilized to detect disease and normal data samples. The DL models used include MLP, CNN, and Bi-LSTM. These models were evaluated based on accuracy, TPR, TNR, FPR, and FNR, and the best model was identified and finalized. The finalized DL model is deployed for real-time monitoring of the health system. Physiological data is collected using sensors, and the data are transmitted to the cloud via a WiFi module. The deployed model in the cloud analyzes the data: if the data is normal, no action is taken; if the data indicates a disease, an alert is sent to the doctor. An overview of the proposed healthcare system's operation is presented in Figure 1.

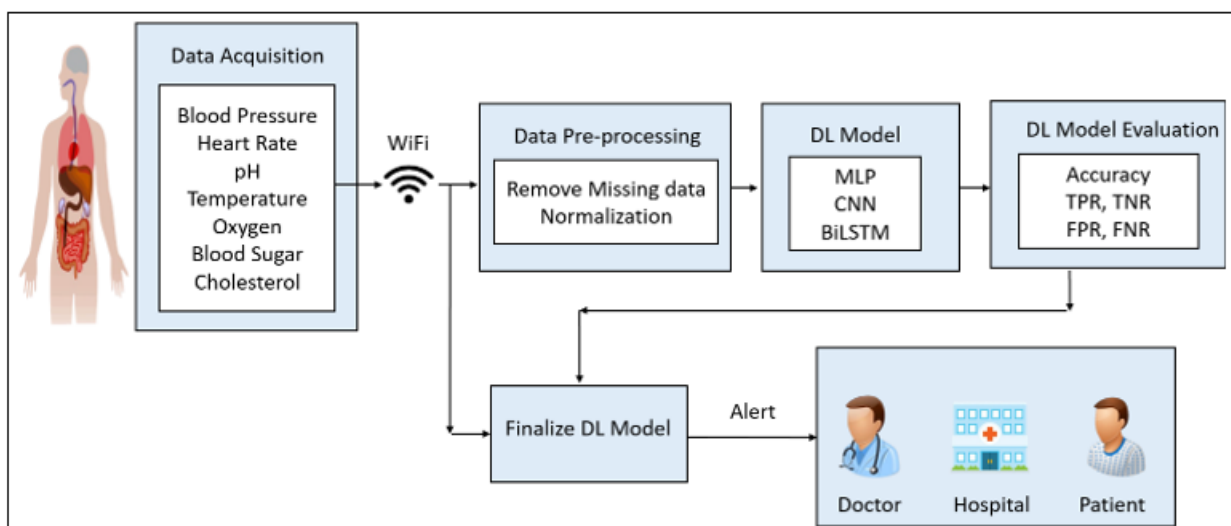


Figure 1: Block diagram of proposed AI and IoT healthcare system

a) Data Collection and Processing

To monitor the patient's health, we create a dataset containing values such as blood pressure, heart rate, pH, temperature, oxygen, blood sugar, and cholesterol. Normal values for each parameter are sourced from Google and research articles. Values within the normal range are labeled

as 0, while abnormal values are labeled as 1 in the dataset. The 0 and 1 in the dataset represent normal and diseased states, respectively. This process allows us to create a dataset comprising 1000 samples. Preprocessing steps, including checking for missing values and normalization, are applied to the dataset. The data is then split in the ratio

of 7:2:1 for train, validate, and test. To implement the DL model in real-time, the health parameters mentioned earlier are collected from corresponding sensors. A WiFi module is utilized to transmit the data to the cloud.

b) MLP

The MLP is a common ANN used for recognition, categorization, and prediction [11]. It is a Feed-Forward ANN due to its capacity to process input data into outputs. Figure 2 depicts a schematic of a MLP network that processes data unidirectionally. Data points are input into the network, processed by few hidden layers, and the result is produced through the output layer.

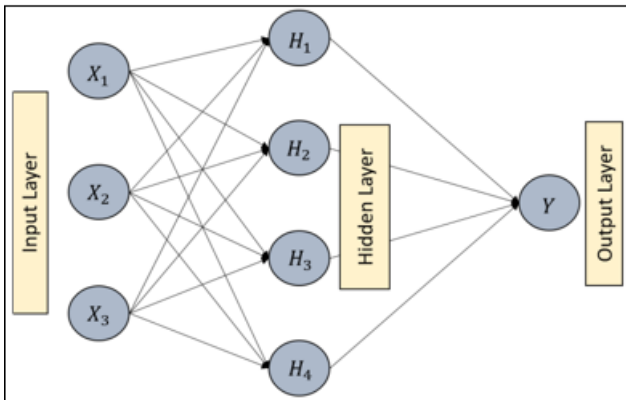


Figure 2: MLP architecture

The backpropagation learning technique [12] trains neurons as data flows from the input to the output layer. The output layer is responsible for disease prediction. Except for the input layer, all neurons in each layer receive data, multiply their outputs by respective weights, sum them, and then pass the sum through an Activation Function (AF) to produce the neuron's output. In each training iteration, the AF continues forward and does not revisit previously encountered neurons. Furthermore, feedforward neural networks assume that all input data instances are independent. The MLP consists of sensory lobe neurons. The perceptron accepts features ($X = X_1, X_2, X_3$) and applies weights to each. The weighted sum is calculated using these input features, all of which must be numerical.

$$S(x) = \sum_{i=1}^n w_{ij}X_i \quad [1]$$

Where, w_{ij} represents the weight that connects i to j neuron across two successive layers. The summation of the weight is sent into the AF, which then generates the perceptron's outcome. The hidden layer employs the sigmoid AF:

$$f(x) = \frac{1}{1+e^{-x}} \quad [2]$$

Since our data is binary, we used the *softmax* AF in the output layer.

c) CNN

An efficient neural network model, CNN can learn complex features using convolution, non-linear activation, dropout, and Pooling Layers (PL) [13]. Its initial applications focused on image segmentation and classification. Training CNN end-to-end enhances its efficiency. At its core, the paradigm aims to capture semantic information through Fully Connected Layers (FCL). By applying filters to the output of the preceding layer, this feed-forward network can map

features. A CNN model typically comprises Convolutional Layers (CL), PL or subsampling layers, a flattening layer, an AF, dropout, and FCL. After CL extracts features, the results feed into FCLs. PL reduces the risk of overfitting by condensing the mapped features from CL. Both maximum and average pooling in PL are viable, with the former selecting more precise features [14]. The data is flattened into an array by the flattening layer before entering the FCL. Rectified Linear Unit (ReLU) serves as the AF in this research.

$$y = \max(0, i) \quad [3]$$

The CLs extract both high and low level features by applying weights to kernels. CNN is increasingly used for diagnosing ailments. This study utilized the cross-entropy (CE) error loss function, which is common in binary classification. Equation (4) illustrates the computation.

$$CE = -(i \log(p) + (1 - i) \log(1 - p)) \quad [4]$$

The projected probability is denoted by \log , the natural logarithm, and i , indicating the class label. The sigmoid function serves as the output error function in CNN, as it is a variant of the backpropagation technique. In the CNN paradigm, each target class instance is represented by two neurons: the first neuron outputs 1 for the disease, while the second neuron outputs 0. In the case of normalcy, the neuron values are inverted. CNN has demonstrated reliability in medical classification tasks, with many scholars using CNN to tackle various categorization challenges.

d) Bi-LSTM

Bi-LSTM is a version of LSTM that utilizes bidirectional processing to learn sequential data by combining the influences of data both before and after the current sequence. It consists of two LSTM networks that operate in opposite directions. The LSTM proposed by Hochreiter and Schmidhuber [15]. LSTM enhances traditional RNNs by incorporating input, forget, and output gates. It addresses the issue of gradient vanishing or exploding, common in classic RNNs, by selectively retaining important information and recursively processing data based on temporal or spatial sequences. Equations (5-7) govern the update process for the input gate (i_t), forget gate (f_t), and output gate (O_t).

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad [5]$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad [6]$$

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad [7]$$

This study defines the weights as W , and the bias as b . When LSTM processes data in a unidirectional manner along the sequence direction, only the influence of preceding sequence data can be considered. To simultaneously incorporate both preceding and succeeding sequence data, a Bi-LSTM network is necessary. The Bi-LSTM structure, depicted in Figure 3, integrates both preceding and succeeding sequence data into its output. Bi-LSTM utilizes two hidden states- B for forward calculation and B'_i for backward calculation- for each sequence provided. The input vectors $x_i = (x_1, x_2, \dots, x_i)$ are linked to the hidden layer via the input layer using functions $B_i = f(AB_i + ZB_{i+1})$ and $B'_i = f(A'B'_{i+1} + Z'B_i)$. Finally, $y_i = g(CB_i + C'B'_i)$ connects to the output of the subsequent layer. For forward computation, the previous and current sequences determine the hidden

state B_i , and for backward computation, the next and current sequences determine B'_i .

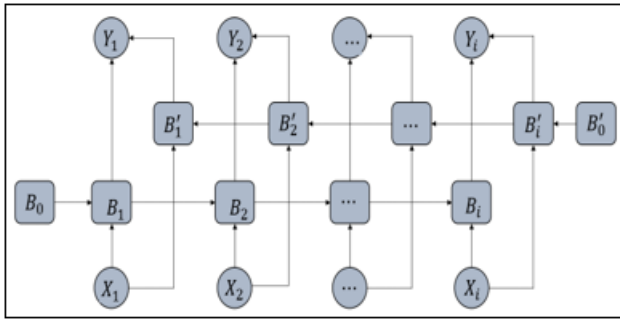


Figure 3: Bi-LSTM Architecture

In a specific layer, all neurons are fully connected to those in the preceding layer; an FCL integrates locally distinct class information and highly refined parameters. A typical deep neural network consists of three FCLs, with the Bi-LSTM layer following the first two. As the input layer for Bi-LSTM learning, the ReLU function preprocesses the data.

4. Results and Discussion

The experimental outcome of the designed healthcare system is discussed in this section. The DL model is designed, trained, and tested on google colab with python program language and hardware accelerator as T4 GPU. For storing, analyzing data, and deploying a DL model in the cloud, Firebase is used.

1) Training Outcome

The experimental outcome of the designed healthcare system is discussed in this section. The DL model is designed, trained, and tested on Google Colab using Python programming language with a T4 GPU hardware accelerator. Firebase is used for storing and analyzing data and deploying the DL model in the cloud. The training and validation of the three DL models are done on 70% and 20% of the dataset, respectively. The model epochs are fixed at 20, and the categorical CE function is used as the loss function. The performance of the models is evaluated using accuracy and loss values. Figure 4 illustrates the accuracy and loss plot of the MLP. The maximum accuracy and loss values reached by the MLP in the training phase are 0.92 and 0.31, respectively, and in the validation phase, the values are 0.96 and 0.14. The loss value is not satisfactory.

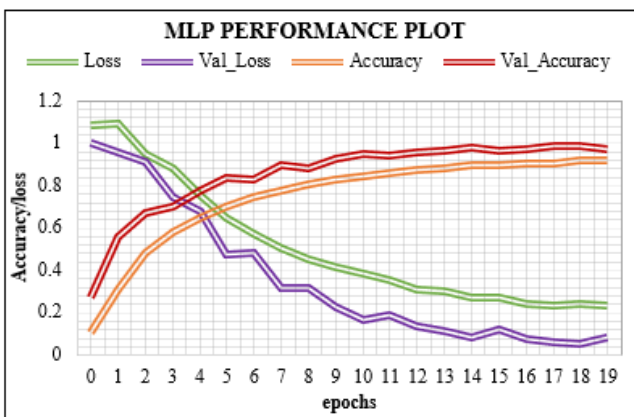


Figure 4: MLP performance plot

Figure 5 shows the accuracy and loss performance plot of the CNN. The maximum accuracy values reached in the training and validation phases are 0.98 and 0.97, with training and validation losses of 0.13 and 0.12. The validation loss values are oscillating significantly, as seen in the figure.

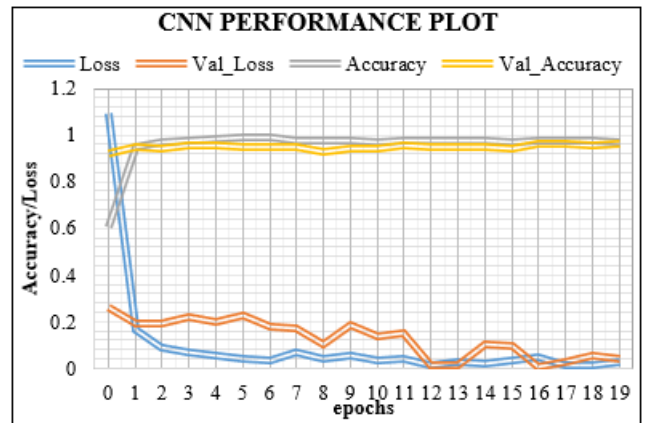


Figure 5: CNN performance plot

Figure 6 shows the performance plot of the Bi-LSTM model. The figure clearly shows that there is no oscillation in the accuracy and loss values. The accuracy reaches 0.99 in both phases, and the loss value is less than 0.05.

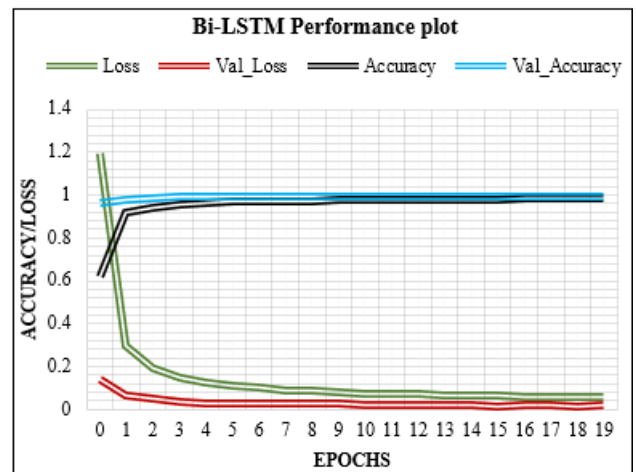


Figure 6: Bi-LSTM performance plot

2) Testing Outcome

After training the models, the three DL models are tested using the remaining data. The testing metrics include Accuracy, TNR, TPR, FRR, and FAR. The metrics values attained by MLP are 96.33, 96.95, 95.74, 4.26, and 3.05, respectively. The CNN gives values of 98.17, 97.37, 98.99, 1.01, and 2.63, respectively. Finally, the Bi-LSTM gives the highest accuracy, TNR, and TPR of 99.00, 98.68, and 99.33, and the lowest FRR and FAR of 0.67 and 1.32. The metric values for predicting diseases by the three DL models are given in Table 1.

Table 1: Performance metrics of the DL model on test data

Model	Accuracy	TNR	TPR	FRR	FAR
MLP	96.33	96.95	95.74	4.26	3.05
CNN	98.17	97.37	98.99	1.01	2.63
Bi-LSTM	99.00	98.68	99.33	0.67	1.32

Based on the performance measures, the Bi-LSTM model is finalized and deployed in the cloud. Sensors attached to the human body help collect physiological parameters and send data to the cloud. In the cloud, the data are processed using the Bi-LSTM model. If an abnormality is found, the Bi-

LSTM gives the output as 1. If the output is 1, an alert message of "immediate attention required" with the patient ID is sent to the doctor's mobile number. The working of the proposed system is given in Figure 7.

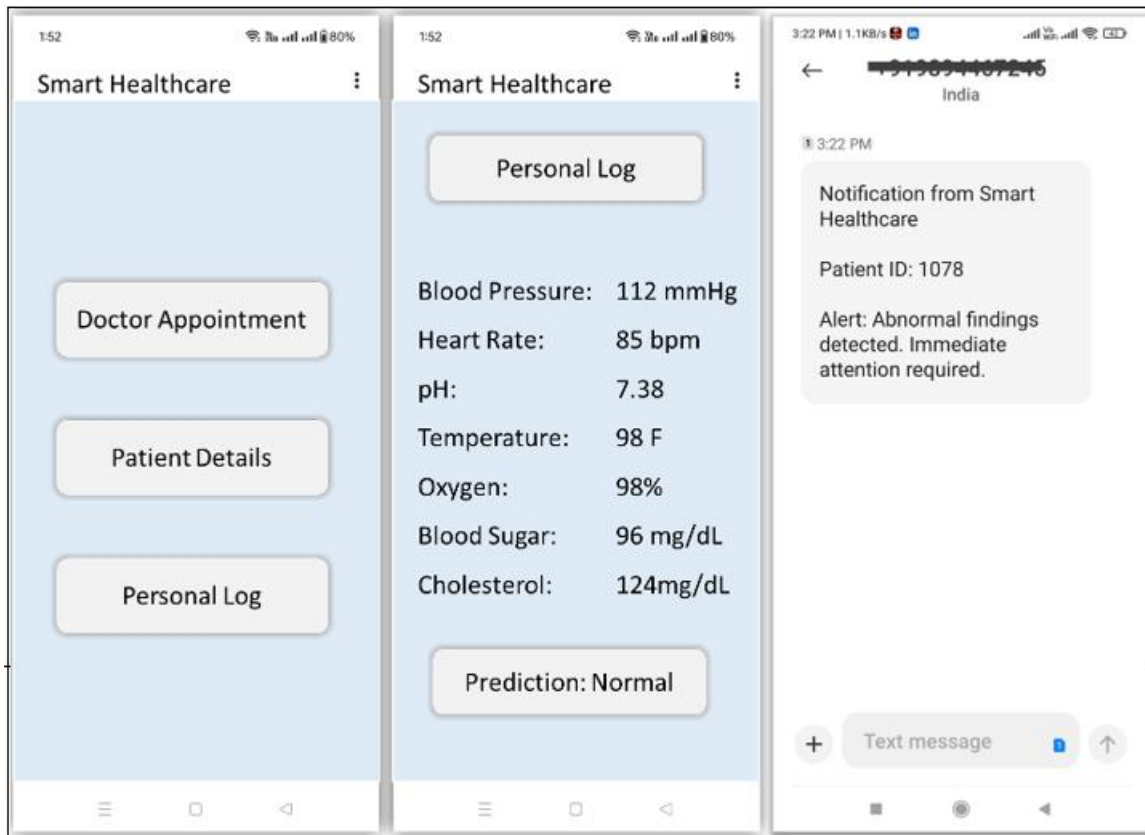


Figure 7: Working of the designed remote healthcare monitoring system

5. Conclusion

The research successfully designs an automated Disease Diagnosis and Monitoring System using DL and the IoT. The DL model is employed to predict the health condition of the patient, and IoT is utilized to transfer physiological data using sensors to the cloud. The system sends alert messages to the doctor in case of abnormalities. The Bi-LSTM model predicts the health condition with the highest accuracy, TNR, and TPR of 99%, 98.68%, and 99.33%, respectively, and the lowest FRR and FAR of 0.67% and 1.32%. Firebase Cloud is used to receive and analyze the data. Once the data is received from the sensors, the Bi-LSTM model predicts the health status of the person. The designed system correctly predicts the health status and sends the required message to the registered mobile number without delay. In this study, very basic health parameters were considered for analyzing health status. In the future, other parameters like ECG and EEG will be collected and processed to detect the accurate health prediction of the person. This will be helpful for people living in a busy world to avoid fatal situations.

References

- [1] World Health Organization. *Achieving well-being: a global framework for integrating well-being into public health utilizing a health promotion approach*. World Health Organization, 2024.
- [2] Daei, Azra, Mohammad Reza Soleymani, Hasan Ashrafi-Rizi, Ali Zargham-Boroujeni, and Roya Kelishadi. "Clinical information seeking behavior of physicians: A systematic review." *International journal of medical informatics* 139 (2020): 104144.
- [3] Firouzi, Farshad, Shiyi Jiang, Krishnendu Chakrabarty, Bahar Farahani, Mahmoud Daneshmand, Jaeseung Song, and Kunal Mankodiya. "Fusion of IoT, AI, edge-fog-cloud, and blockchain: Challenges, solutions, and a case study in healthcare and medicine." *IEEE Internet of Things Journal* 10, no. 5 (2022): 3686-3705.
- [4] Firouzi, Farshad, Bahar Farahani, and Alexander Marinšek. "The convergence and interplay of edge, fog, and cloud in the AI-driven Internet of Things (IoT)." *Information Systems* 107 (2022): 101840.
- [5] Kishor, Amit, and Chinmay Chakraborty. "Artificial intelligence and internet of things based healthcare 4.0 monitoring system." *Wireless personal communications* 127, no. 2 (2022): 1615-1631.
- [6] Akhbarifar, Samira, Hamid Haj Seyyed Javadi, Amir Masoud Rahmani, and Mehdi Hosseinzadeh. "A secure remote health monitoring model for early disease diagnosis in cloud-based IoT environment." *Personal and Ubiquitous Computing* 27, no. 3 (2023): 697-713.
- [7] Khanna, Ashish, Pandiaraj Selvaraj, Deepak Gupta, Tariq Hussain Sheikh, Piyush Kumar Pareek, and Vishnu Shankar. "Internet of things and deep learning

- enabled healthcare disease diagnosis using biomedical electrocardiogram signals." *Expert Systems* 40, no. 4 (2023): e12864.
- [8] Ahamed, Jameel, Abdul Manan Koli, Khaleel Ahmad, Alam Jamal, and B. B. Gupta. "CDPS-IoT: cardiovascular disease prediction system based on IoT using machine learning." (2022).
- [9] Malibari, Areej A. "An efficient IoT-Artificial intelligence-based disease prediction using lightweight CNN in healthcare system." *Measurement: Sensors* 26 (2023): 100695.
- [10] Velmurugan, A. K., K. Padmanaban, A. M. Kumar, H. Azath, and Murugan Subbiah. "Machine learning IoT based framework for analysing heart disease prediction." In *AIP Conference Proceedings*, vol. 2523, no. 1. AIP Publishing, 2023.
- [11] Basu, Jayanta Kumar, Debnath Bhattacharyya, and Tai-hoon Kim. "Use of artificial neural network in pattern recognition." *International journal of software engineering and its applications* 4, no. 2 (2010).
- [12] Lillicrap, Timothy P., Adam Santoro, Luke Marris, Colin J. Akerman, and Geoffrey Hinton. "Backpropagation and the brain." *Nature Reviews Neuroscience* 21, no. 6 (2020): 335-346.
- [13] O'shea, Keiron, and Ryan Nash. "An introduction to convolutional neural networks." *arXiv preprint arXiv:1511.08458* (2015).
- [14] Yu, Dingjun, Hanli Wang, Peiqiu Chen, and Zhihua Wei. "Mixed pooling for convolutional neural networks." In *Rough Sets and Knowledge Technology: 9th International Conference, RSKT 2014, Shanghai, China, October 24-26, 2014, Proceedings 9*, pp. 364-375. Springer International Publishing, 2014.
- [15] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9, no. 8 (1997): 1735-1780.