Automatika
Journal for Control, Measurement, Electronics, Computing and Communications

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/taut20

A novel framework for multiple disease prediction in telemedicine systems using deep learning

Divya R. Unnithan & J. R. Jeba

To cite this article: Divya R. Unnithan & J. R. Jeba (2024) A novel framework for multiple disease prediction in telemedicine systems using deep learning, Automatika, 65:3, 763-777, DOI: 10.1080/00051144.2024.2301889

To link to this article: https://doi.org/10.1080/00051144.2024.2301889

Published online: 26 Feb 2024.

AUTOMATIKA
2024, VOL. 65, NO. 3, 763–777
https://doi.org/10.1080/00051144.2024.2301889

A novel framework for multiple disease prediction in telemedicine systems using deep learning

Divya R. Unnithan and J. R. Jeba

Department of Computer Applications, Noorul Islam Centre for Higher Education, Kumarakovil, India

ABSTRACT

Telemedicine systems are gaining popularity due to their ability to provide remote medical services. These systems produce a lot of data, which may be used for a variety of purposes, including quality improvement, decision-making, and predictive analytics. Deep learning is an effective data mining method that may be applied to this data to bring out significant findings. Telemedicine systems, which allow patients to receive medical consultation and treatment remotely, generate vast amounts of data. Analyzing this data can provide valuable insights for improving patient care and optimizing the telemedicine system. Data mining techniques can be incredibly valuable for telemedicine systems, as they can help to identify patterns and insights in large amounts of patient data. Data mining techniques can assist telemedicine systems in making better decisions and offer better care to patients. In this paper, a novel framework for multiple disease prediction in telemedicine system using an effective deep learning algorithm was developed. The proposed multiple disease prediction system is composed of Long Short Term Memory (LSTM) unit.The experimental results revealed that the suggested disease prediction model exceeded the present models with an accuracy of 98.51%.

ARTICLE HISTORY
Received 16 September 2023
Accepted 10 December 2023

KEYWORDS
Telehealth system; telemedicine system; data mining techniques; clustering; clustering techniques; classification techniques

1. Introduction

The literal meaning of Telemedicine is "healing from distance". Telemedicine has been defined by the World Health Organization (WHO) as "The delivery of healthcare services, where distance is a critical factor, by all healthcare professionals using information and communication technologies for the exchange of valid information for diagnosis, treatment and prevention of disease and injuries, research and evaluation, and for continuing education of healthcare providers, all in the interests of advancing the health of individuals and their communities".

The provision of healthcare services to patients through the use of information, communication, and technology has been referred to as telemedicine. Provision of telemedicine functions by catering to the patient's needs through calls and video conferencing. Globally, physicians use telemedicine for digital imaging transmission, video consultations, and remote medical signals [1]. The origin of telemedicine has been a consequence of the short of time required for consultations. The term "telemedicine" is largely used in conjunction with "telehealth". The term "telehealth" refers to the use of electronic information and communication technology to support and promote long-distance clinical healthcare, for the medical education of clinical professionals, and to help maintain and manage public health [2]. The contrast between the two terms can be drawn as telemedicine purely clinical application of the information technology in healthcare whereas telehealth is being used as an umbrella term for nonspecific healthcare-related activities. With the advancement in technology and application in the diverse fields, evolved concepts like telemedicine have been successful in their implementation globally. The application of telemedicine has been considered to be useful in both developed and developing nations. Core benefits of the application of integrated telemedicine in the healthcare system are inclusive of increment in the revenue, maximization of patient reach, provision of convenience, increased feasibility, high diversity and availability of several healthcare professionals on a single platform and improved healthcare quality [3].

Figure 1 depicts the general framework of a basic telemedicine system. The incorporation of numerous technologies into the telemedicine system enables the creation of a health network connecting medical facilities dispersed across a city, a country, or even a region. This network enables the interchange of health data, including clinical findings, radiographic pictures, laboratory data, and sound recognition, among other things. Because it is more likely that specialists will be available in places with dense populations, those who reside in remote and rural areas have very limited access

CONTACT Divya R Unnithan divyagireesh@ymail.com Department of Computer Applications, Noorul Islam Centre for Higher Education, Kumarakovil, India, 629180

to high-quality healthcare when they are in need [4]. Many aspects of medical practise can now be carried out even when the patient and healthcare professional are separated by a distance due to advancements in computing and telecommunications technologies.

Telemedicine has a variety of uses, including teleconsultation, tele education, telepathology, teleradiology, and telecardiology, all of which have demonstrated their efficacy in maximizing resource utilization [6]. Essentially, telemedicine is performed in two ways: first, by employing the store and forward method, and second, by implementing it in real-time. The store and forward method involve recording patient data on any storage media or actual paper records. Later, the information is transmitted via various media for the medical expert's assessment. In the second method, the expert examines the records concurrently while treating patients from a distance. The telemedicine data includes a variety of data categories, including personal data, historical data, clinical data, and patient reports. Telemedicine must be implemented with the growing population and insufficient medical resources.

Data mining is the process of identifying patterns in massive datasets using methods that incorporate statistics, machine learning, and database systems. The main objective of data mining is to apply cutting-edge data analysis techniques to enormous amounts of data stored in databases, data marts, or data warehouses in order to identify significant patterns, trends, rules, relationships, and correlations that are present in the data but that would be impossible or very difficult to identify using conventional data analysis techniques [7]. Due to their capacity to draw useful conclusions and patterns from vast quantities of data, data mining techniques have grown in significance in the telemedicine industry. Data mining techniques can be used in a telemedicine system to examine patient data, including medical records, diagnostic pictures, and sensor data, in order to spot potential health hazards, predict the evolution of diseases, and create individualized treatment plans. Figure 2 displays the data mining framework visualization for the telemedicine system.

Data mining techniques are particularly useful in telemedicine because they can reveal hidden links and patterns that human analysts would miss. By determining risk variables and creating tailored therapies, data mining techniques can also be utilized to enhance patient outcomes. A deep learning-based model was suggested in this paper for the efficient prediction of multiple diseases. The multiple disease prediction model is associated with real-time interactive telemedicine system to enable the communication between healthcare providers and the patients.

The following is the paper's contribution:

- In order to find patterns and insights in vast amounts of patient data, data mining techniques can be quite helpful for telemedicine systems. Data mining techniques can help telemedicine systems make better decisions and provide patients with better treatment.
- Using a powerful deep learning algorithm, a novel framework for multiple disease prediction in telemedicine systems was created.
- Long Short Term Memory (LSTM) units make up the suggested multiple disease prediction system.
- The testing findings showed that the proposed disease prediction model outperformed the current models.

The remainder of the document is structured in this manner. The pertinent initiatives for the identification and classification are covered in Section 2. Section 3 describes the approach that was used for this investigation. The results and analysis of the experiment are presented in Section 4. Section 5 marks the conclusion of this work.

2. Literature review

The analytical, experimental and concluding opinions made by authors in respect to various strategies for optimizing telemedicine using various data mining techniques are discussed in the following section.

Manoranjan Dash et al. [8] aimed at identifying the elements that will encourage patients in India to utilize telemedicine during the COVID-19 pandemic. In order to analyze the information gathered from 146 patients using a structured questionnaire, multiple regression and ANN techniques are applied. According to the experimental findings, the ANN model outperformed multiple regressions in terms of nonlinearity and linearity and predicted outcomes with a high degree of accuracy. Syed Thouheed Ahmed et al. [9] developed a dynamic user clustering method based on heterogeneous multi-input multi-output data. The suggested methodology employs networking nodes to add machine learning concepts for dynamic user grouping and classification, resulting in the construction of clusters reflecting similarity indexing ratios. The experimental findings revealed that the proposed method is effective for transmitting delicate medical datasets with pre-processed data. However, the proposed method cannot handle noisy data. Praveen Kumar Sadineni [10] presented on how big-data analytics and machine learning may be combined to improve the quality of healthcare services using techniques like decision trees, SVM, and KNN. The provision of individualized solutions to specific issues, such as the detection and treatment of epidemics, the enhancement of life value, the reduction of needless care, etc., is made possible by enhancing the quality of healthcare services. The outcomes of the experiment show that combining machine learning methods with Big-Data Analytics raises the

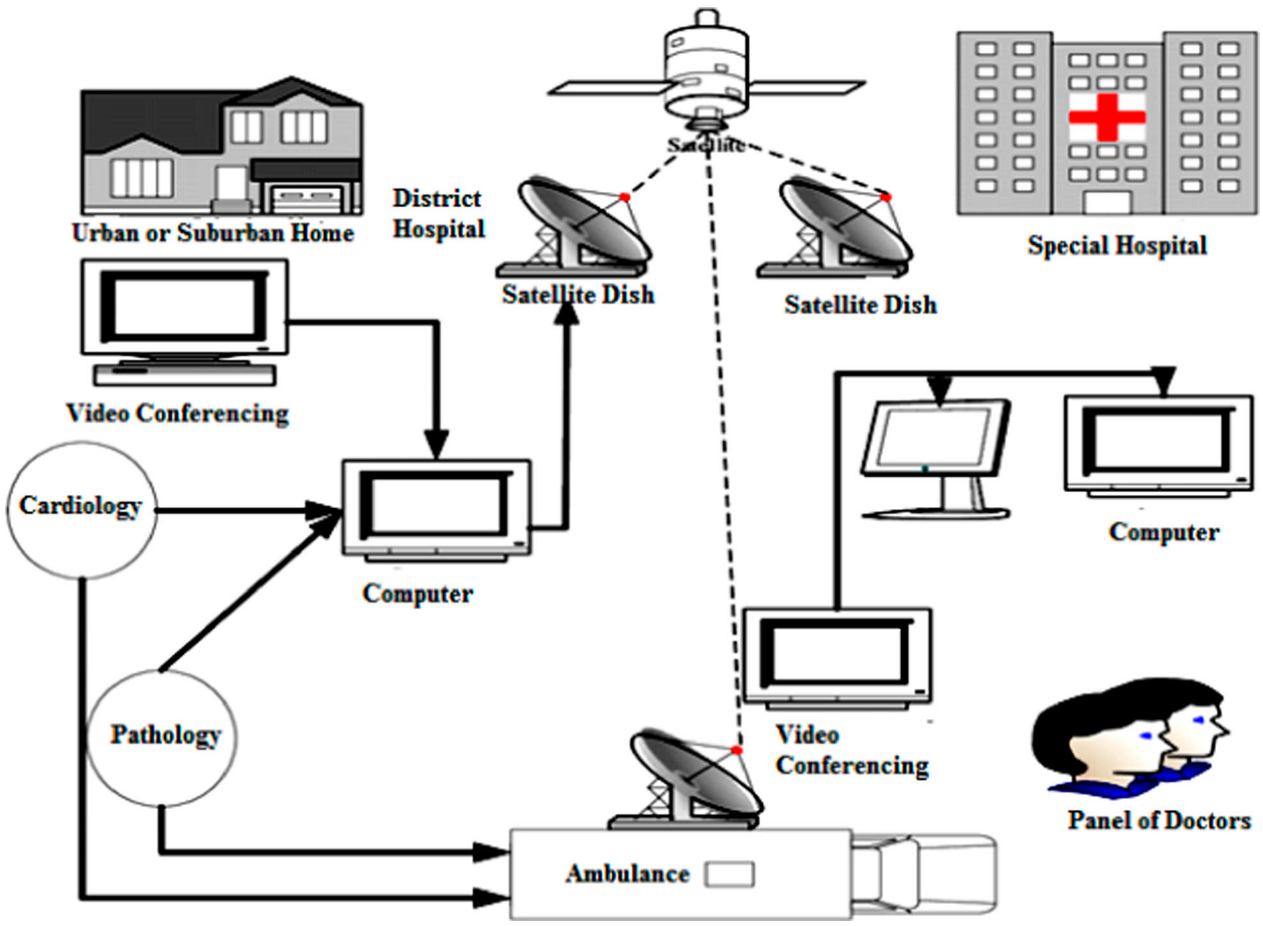

Figure 1. Telemedicine system [5].

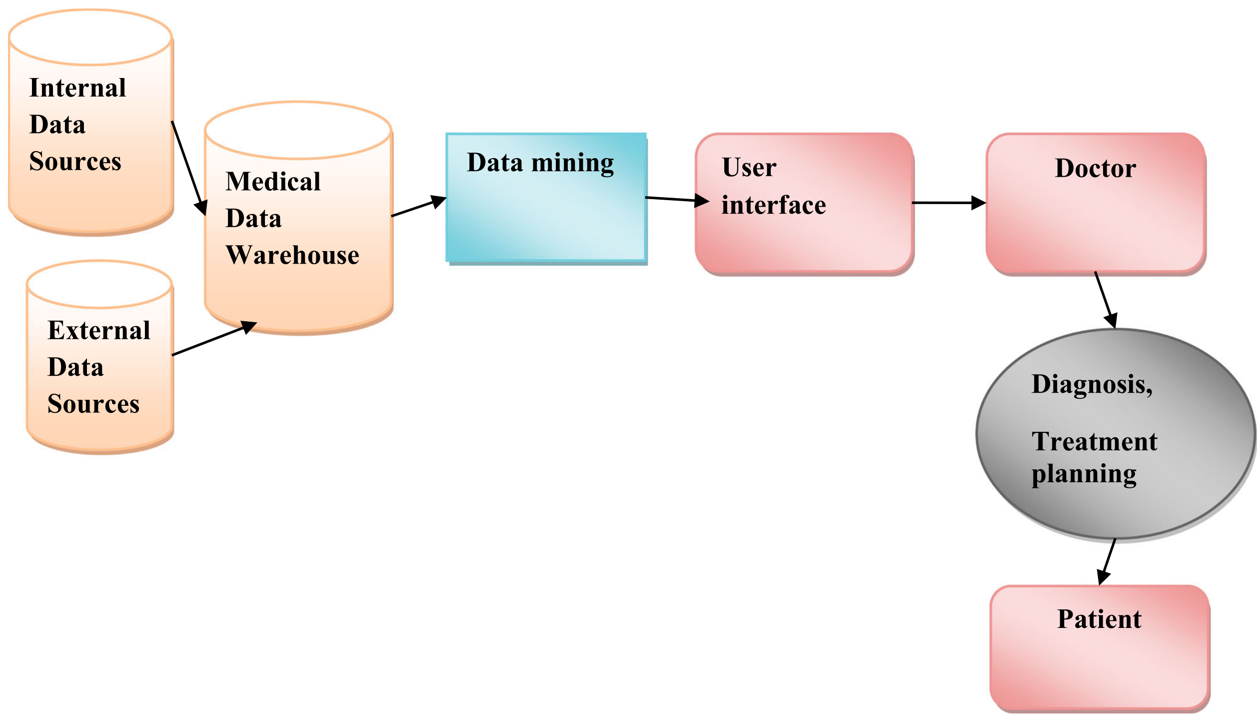

Figure 2. Datamining framework for telemedicine.

excellence of healthcare services. However, in order to deliver accurate results, the suggested method needed high-quality data. M. Sornalakshmi et al. [11] proposed an approach that coupled the context ontology and enhanced apriori algorithm for mining and modelling physiological data utilizing the concepts and connections established by the rules that were generated. A growing number of rules are obtained by combining the EAA with the context ontology. According to the performance analysis, the proposed method produces better support and confidence. The comparison analysis shows that the suggested EAA-SMO technique achieves maximum accuracy and requires the least amount of time to execute than the semantic ontology. The scalability of the suggested approach is constrained. So-Young Choi and Kyungyong Chung [12] presented a big-data knowledge procedure for the health sector using association mining and Hadoop's MapReduce technology. By combining WebBot and the common data model to gather and process heterogeneous health information, the suggested solution offers effective health management knowledge services. Documents that are periodically generated by dynamic linking and distributed file processing are assembled into a corpus for the purpose of finding relationships between data. The processing of large amounts of health-related data using MapReduce-based association mining can aid in disease prevention, the detection of hazards, and post-management using a common data model. As a result, healthcare services that are more advanced can be provided, which helps to enhance people's health and quality of life.

D.M. JeyaPriyadharsan et al. [13] presented machine learning techniques for keeping track of human health. The UCI dataset is used for the initial training and validation of ML algorithms. In the testing phase, anomalies in the health state are predicted using sensor data collected to use an IoT framework. IoT device data that has been stored in the cloud is statistically analyzed to determine the accuracy of the prediction percentage. Also, according to the results, the K-Nearest Neighbour beats other traditional classifiers. The major limitation of the study is that, when the training set is large, it takes a lot of space. R. Sandhiya and M. Sundarambal [14] created a clustering model with enhanced semantic smoothing that is based on ontology and domain knowledge. The model used TF-IGM and modified n-grams to enhance the clustering process. Hierarchical and partitional clustering techniques are used to assess the model's performance. The proposed method outperformed the semantic smoothing model in almost 80% of the quality criteria, proving its efficacy. A drawback of using n-gram overlap to assess document similarity is that it performs poorly when the original document has been updated. T. K. Anusuya and P. Maharajothi [15] designed a method to manage various multimedia medical databases in the telemedicine system. The primary objective of this work is to convey the medical services to the patient, instead of transporting the patient to the medical care services. This is accomplished through the use of web-based solutions, such as Modern Medical Informatics Services, which are simpler, quicker, and less expensive. The fragmentation of databases, clustering of network locations, and allocation of fragments were three enhanced services that were added into this method. In order to calculate the cost of communications, an estimating model was also put forth, which aids in the search for efficient data allocation strategies. The outcome demonstrated that the suggested technique considerably raises the level of satisfaction with services requirements in web systems. The main shortcomings of this proposed study are the lack of standardization and privacy issues. Syed Thouheed Ahmed and M. Sandhya [16] provided a cutting-edge method and presented about recursive image reduction in the cloud/server. The method depends on pixel value density matching with edge extraction for the suggested Real-Time Biomedical Imaging Recursion Detection. The suggested method reduced initial processing by 60% while achieving time optimization. According to time and space optimization, the suggested system has a 97.8% efficiency rate. It is difficult to schedule the suggested system's total synchronization. P. Sukumar et al. [17] proposed an ontology-oriented architecture that utilized a knowledge base (KB), enabling the integration of data from several heterogeneous sources. The proposed strategy has been used in the area of personalized medicine. In order to find knowledge concealed in diverse data sources, the AI approach is also utilized to mine data in the healthcare industry. The suggested system was subjected to three ontology phases. According to the findings, the textual documents were successfully grouped using the suggested system. The suggested system has many drawbacks, including limited language support and an inability to manage unstructured data.

BikashKanti Sarkar and ShibSankar Sana [18] created a disease decision support system in which the initial stage deals with determining the best training set in parallel with the best data-partition for each illness data set. The second stage investigates a general predictive model over the learned data for a precise disease diagnosis. The suggested method performs admirably on all of the selected medical data sets and can be a useful alternative for the well-known ML techniques. The findings demonstrated that, for the initial identification of the disorders, the suggested hybrid model consistently outperformed the basic learners. However, the quality of the data employed for training the model affects the accuracy of the model. Atta- Ur- Rahman and Mohammed Imran Basheer Ahmed [19] examined a telemedicine plan for a virtual clinic that would provide medical care in remote locations of developing nations. The suggested approach combines a fuzzy rule-based approach to rank the top doctors with a clinical decision support system that aids in selecting the best physician for a certain patient based on his prior prescriptions. The apriori algorithm and the inductive learning algorithm serve as the foundation for the clinical decision support system. The evaluation findings demonstrated that inductive learning performed better than the Apriori algorithm. Syed Thouheed Ahmd et al. [20] suggested a Real-Time Signal Re-Generator and Validator method based on neural networking and machine learning. The primary goal of the suggested design is to obtain a higher order of signal optimization for secure and reliable telemedicine consultation of biological samples via low line transmission channels. The RTSRV method has considered feature extraction and layered decomposition of a signal. The features are grouped using the KNN method to categorize each attribute based on the frequency with which it occurs within a certain time span. The suggested algorithm has a higher rate of accuracy. However, the effectiveness of the device is dependent on the signal quality. R. Sandhiya and M. Sundarambal [21] created an effective clustering approach for biomedical documents and health data based on chicken swarm optimization with dynamic dimension reduction to support telemedicine applications. The data are initially preprocessed using concept mapping and semantic annotation, which increase document representation, frequency, and inverse gravity moment factor, and the modified n-gram, which rectifies for substitution and deletion errors. The outcomes of the experiments demonstrated that the proposed methodology can be very effective in telemedicine applications and remote monitoring of medical treatment. The benefits of the suggested model include a reduction in time complexity, accurate clustering findings regardless of dataset rescaling or normalization, and independence from document order. The proposed system is sensitive to localization. Jesus Peral et al. [22] To enhance the telemedicine system's triage process for patients who live distant from hospitals and utilize it, taking into account the patients' diverse chronic conditions, such as diabetes, hypertension, hypotension, and chronic heart disease [26]. Utilizing AI and ML algorithms based on speech records, handwriting patterns, impairments in gait, and neuroimaging methods to identify Parkinson's disease (PD) early. The paper also discusses how the Internet of Things, the metaverse, and electronic health records may be used to manage Parkinson's disease (PD) more effectively and enhance quality of life [27]. The process involves producing highly informative features called Feature Engineering (FE). We used the Pima Indian Diabetes Dataset (PIDD) to experiment with and examine the effectiveness of ML models' ability to predict diabetes [28]. A smartphone-integrated, low-cost, and high-quality metadata production method for chronic wound picture collecting at the patient's side is also available. This approach facilitates seamless remote patient-doctor interaction and end-to-end routine diagnostics, as well as the maintenance of patient histories [29]. To improve the system's overall performance, create a TWM system model that can gather, process, and monitor chronic wound-related issues utilizing a cheap smartphone [30].

Sensor networks in telemedicine systems for various patient monitoring scenarios. The integration of current and upcoming advancements in communications with advancements in microelectronics and embedded system technologies will significantly influence the development of patient monitoring and health information delivery systems [29].

In order to combine data from several heterogeneous sources, an ontology-oriented architecture was developed, with a core ontology serving as the knowledge base. The strategy has been used in the area of personalized medicine. AI techniques have been employed with the goal of data mining in the healthcare industry to uncover information buried in diverse data sources. The viability of the suggested approach has been demonstrated using a case study using the diabetic situation. The method enables the doctor to enhance the DM guidelines by incorporating the knowledge acquired from specialized Web documents. The system is less flexible and more complicated.

Some of the limitations of the existing methods are discussed below. Both multiple regressions and the artificial neural network (ANN) approach require high-quality data to produce accurate and reliable results. Inaccurate or noisy data can hinder their effectiveness. Furthermore, scalability is limited, and handling large datasets can be time-consuming as distances need to be computed between test and training data for each test case. Additionally, privacy concerns arise, risking the compromise or exposure of patient data. Modifications to the first document can also negatively impact performance. Standardization and overall synchronization scheduling pose challenges, while the inability to handle unstructured data and bias in the data can lead to inaccurate predictions. The sensitivity to initial conditions further affects outcomes, and the rendering methods employed can influence the results of the ML algorithms. Considering these factors, it becomes evident that addressing data quality, scalability, privacy concerns, synchronization, unstructured data, and bias is crucial for effective utilization of multiple regression and ANN techniques.

3. Materials and methods

The collection of datasets was the initial stage in the proposed work. It is followed by pre-processing. The prediction method was carried out using the preprocessed data. The prediction approach made use of LSTM. Finally, the prediction performance was analyzed and compared with existing methods. The basic framework for the suggested system is displayed in Figure 3.

3.1. Data collection and data preparation

In a telemedicine system, patient data can be collected through various methods to facilitate remote

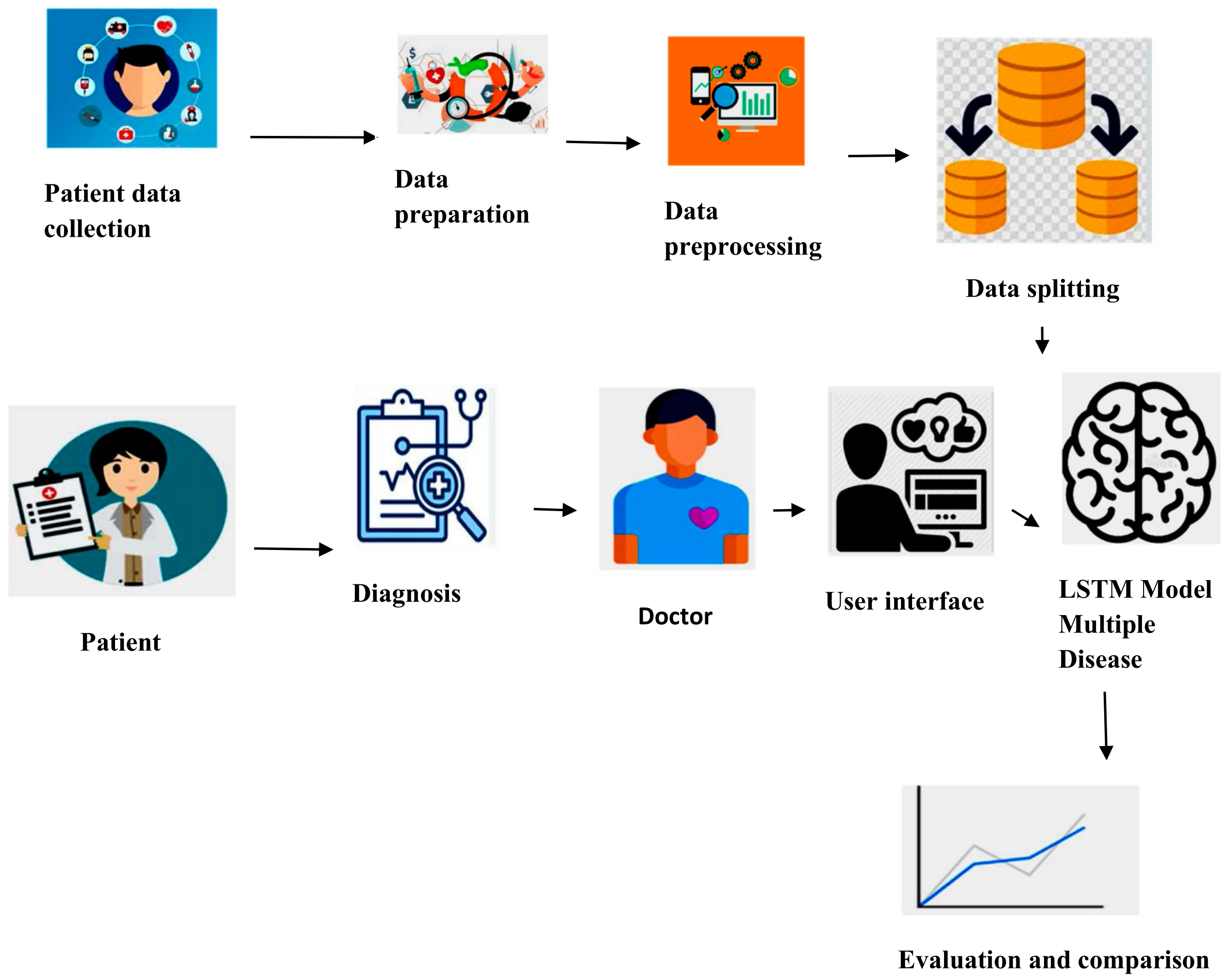

Figure 3. Block diagram of the proposed methodology.

	itching	skin_rash	nodal_skin_eruptions	continuous_sneezing	shivering	chills	joint_pain	stomach_pain	acidity	ulcers_on_tongue	...	blackheads	scurring	skin_peeling
0	1	1	1	0	0	0	0	0	0	0	...	0	0	0
1	0	1	1	0	0	0	0	0	0	0	...	0	0	0
2	1	0	1	0	0	0	0	0	0	0	...	0	0	0
3	1	1	0	0	0	0	0	0	0	0	...	0	0	0
4	1	1	1	0	0	0	0	0	0	0	...	0	0	0

Figure 4. Sample data for dataset.

healthcare delivery. Integration with electronic health record systems enables telemedicine platforms to access and retrieve existing patient data from the EHRs. This includes past medical records, test results, prescription history, and other relevant healthcare information. The multiple disease prediction dataset [23] was collected from the repository of the YBI Foundation. The collected dataset consists of 4920 rows and 133 columns. The first 132 columns in the dataset represents different symptoms of various diseases like itching, skin rashes, sneezing etc. The last column of dataset represents prognosis type. The sample data of dataset are depicted in Figure 4. The sample data of dataset contains values of 0 and 1.

	itching	skin_rash	nodal_skin_eruptions	continuous_sneezing	shivering	chills	joint_pain	stomach_pain	acidity	ulcers_on_tongue	...	pus_filled_pimpl
itching	1.000000	0.318158	0.326439	-0.086906	-0.059893	-0.175905	-0.160650	0.202850	-0.086906	-0.059893	...	-0.0598
skin_rash	0.318158	1.000000	0.298143	-0.094786	-0.065324	-0.029324	0.171134	0.161784	-0.094786	-0.065324	...	0.3208
nodal_skin_eruptions	0.326439	0.298143	1.000000	-0.032566	-0.022444	-0.065917	-0.060200	-0.032566	-0.032566	-0.022444	...	-0.0224
continuous_sneezing	-0.086906	-0.094786	-0.032566	1.000000	0.608981	0.446238	-0.087351	-0.047254	-0.047254	-0.032566	...	-0.0325
shivering	-0.059893	-0.065324	-0.022444	0.608981	1.000000	0.295332	-0.060200	-0.032566	-0.032566	-0.022444	...	-0.0224
...	...	...	...	...	...	...	...	...	...	...	...	
small_dents_in_nails	-0.061573	0.331087	-0.023073	-0.033480	-0.023073	-0.067765	0.359845	-0.033480	-0.033480	-0.023073	...	-0.0230
inflammatory_nails	-0.061573	0.331087	-0.023073	-0.033480	-0.023073	-0.067765	0.359845	-0.033480	-0.033480	-0.023073	...	-0.0230
blister	-0.061573	0.331087	-0.023073	-0.033480	-0.023073	-0.067765	-0.061889	-0.033480	-0.033480	-0.023073	...	-0.0230
red_sore_around_nose	-0.061573	0.331087	-0.023073	-0.033480	-0.023073	-0.067765	-0.061889	-0.033480	-0.033480	-0.023073	...	-0.0230
yellow_crust_ooze	-0.061573	0.331087	-0.023073	-0.033480	-0.023073	-0.067765	-0.061889	-0.033480	-0.033480	-0.023073	...	-0.0230

132 rows × 132 columns

Figure 5. Correlation between attributes in dataset.

	itching	skin_rash	nodal_skin_eruptions	continuous_sneezing	shivering	chills	joint_pain	stomach_pain	acidity	ulcers_on_tongue	...	pus_filled_pimples
count	4920.000000	4920.000000	4920.000000	4920.000000	4920.000000	4920.000000	4920.000000	4920.000000	4920.000000	4920.000000	...	4920.000000
mean	0.137805	0.159756	0.021951	0.045122	0.021951	0.162195	0.139024	0.045122	0.045122	0.021951	...	0.021951
std	0.344730	0.366417	0.146539	0.207593	0.146539	0.368667	0.346007	0.207593	0.207593	0.146539	...	0.146539
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000
25%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000
50%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000
75%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000
max	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	...	1.000000

8 rows × 132 columns

Figure 6. Descriptive statistics of dataset.

3.2. Data preprocessing and data splitting

The data is pre-processed after various records have been gathered. The data contain missing values, outliers, duplicate entries, or inconsistent formatting. Data preprocessing allows for identifying and handling these issues, ensuring data integrity and improving the quality of the dataset. It enables scaling or normalization of features, bringing them to a common scale and facilitating fair comparisons between variables. To maintain data quality and rectify any gaps or inconsistencies in the dataset, it is crucial to check for missing values during the data preprocessing process. Any data that is missing a value indicates that the event did not occur, the data were unavailable, or the data were not applicable. Select the most effective approach based on the quantity and pattern of missing values. There are several options, such as eliminating rows or columns with missing values, imputing missing values using methods like mean, median, or regression imputation, or utilizing sophisticated imputation algorithms like K-nearest neighbours (KNN) or multiple imputation. This dataset doesn't contain any missing values. Therefore, data cleaning is not required.

A statistical tool for quantifying the connection between two variables is correlation. It is employed to evaluate the significance and direction of the linear relationship between the variables. Correlation analysis can assist in identifying highly correlated variables. If two variables are strongly correlated, it indicates that they carry similar information, and including both in a predictive model may lead to redundancy. In such cases, one of the variables can be removed during feature selection to simplify the model and reduce dimensionality. The correlation between all columns in the dataset are visualized in Figure 5.

A direct or positive association between two variables is indicated by a positive correlation coefficient. The other variable tends to rise along with the first one as it rises. An inverse or negative association between two variables is indicated by a negative correlation coefficient. The other variable usually tends to decrease when one variable rises.

Exploratory Data Analysis (EDA) is a crucial stage in the data analysis process. In order to properly model the data, it is necessary to first examine and comprehend its features, patterns, and correlations. EDA helps in identifying insights, detecting anomalies, and formulating hypotheses. The commonly used EDA techniques are descriptive statistics, data visualization etc. The descriptive statistics of the dataset is given in Figure 6. Descriptive statistics provide concise summaries of the data. Measures such as mean, median, mode, and standard deviation give a sense of the central tendency, variability, and distribution of the data. Summarizing the data in this way helps in quickly grasping its overall behaviour and characteristics.

Data visualization is a powerful tool for exploring and communicating insights from a dataset. It allows to visually represent the data in a way that makes patterns, trends, and relationships more apparent. Heatmap is a commonly used data visualization technique. Heatmaps use colour gradients to represent the magnitude of values in a matrix or two-dimensional dataset. They are particularly useful for visualizing correlations or patterns in large datasets. The heatmap visualization of the dataset is represented in Figure 7. Histograms are also a powerful tool for data visualization. Histograms display the distribution of a single continuous variable. The height of each bar shows the number of data points lying within each bin, while the width of each bar represents the variable's range. Histograms help in effectively summarizing and communicating complex data distributions, making them an essential tool in data analysis and visualization. Figure 8 represents the histogram visualization of dataset.

The dataset undergoes data splitting after preprocessing the data. It is possible to separate the preprocessed data into training and testing sets. In this case, 70% of the data was used for training while the remaining 30% for testing. The model is trained by exposing it to labelled or known data in the training set. The testing set offers an objective assessment of the model's effectiveness on unobserved data.

3.3. *Long short term memory (LSTM) for multiple disease prediction*

A type of Artificial Neural Network which follows the concept of time series or sequential data is known as a Recurrent Neural Network. It uses training data to learn the same like CNNs and feedforward networks. RNNs are peculiar for their memory. The web can copy the information from previous inputs and influence current input and output. In CNN, input and output variables are independent of each other. Depending on the previous element, the output is decided in RNN. There are two types of sequence. One is unidirectional, and another one is bidirectional. It can predict the output while running the current task.

A standard Recurrent Neural Network is a feedforward neural network with no time independency. Short-term memory is provided for the hidden nodes connected to each other. From the last time point, some data will be stored for a short duration. As the network is training continuously, the predicted data may be erased, leading to a very short memory. Backpropagation learning through many hidden layers becomes the solution for the vanishing gradient problem. The activation function is continuously multiplied with an error signal. In backpropagation, limited time is allotted for memory to store. Thus, long-term memory is avoided. RNN can process a sequence of data that comes in the temporal form, whereas CNN cannot interpret temporal information. Thus, CNN and RNN are used for entirely different applications. Reusing activation functions in RNN to generate the sequence, whereas CNN has filters.

One of the artificial neural networks of RNN is a Long and short-term memory neural network that is mainly used in deep learning [24]. In sequence prediction problems, it can learn order dependencies. As it can solve complex problems, it is a peculiar type of RNN that handles long-term dependencies. LSTM has feedback connections. It finds applications where the order of input is an important factor. In Deep RNN, many variations were developed to solve the vanishing and exploding gradients problem. Among them, one special network is LSTM. The introduction of distinct activation functions for each Gate allows the long- and short-term memory to forget irrelevant information and attempt to remember the past knowledge via which the network comes into contact. Additionally, it includes a vector called internal cell state, which serves as a practical description of the data that the previous unit had kept. The general architecture of LSTM shown in Figure 9.

LSTM maintains four different gates. They are:

- Forget Gate (f)
- Input Gate (i)
- Input Modulation Gate(g)
- Output Gate(o)

The Forget Gate defines how much past data can be forgotten. The amount of data to be written onto the internal cell gate is determined by the input gate. The

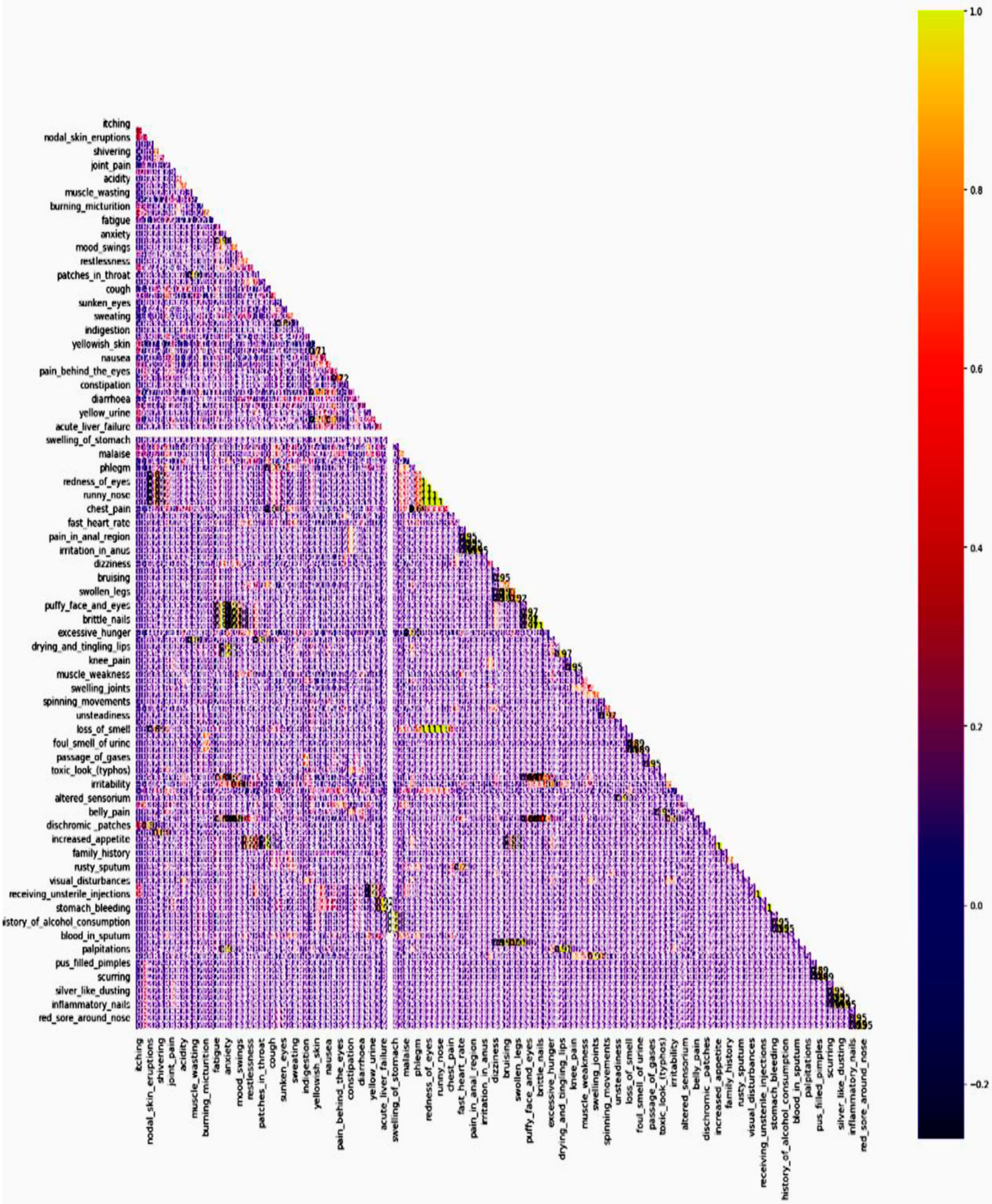

Figure 7. Heatmap visualization of dataset.

information is made zero mean by adding non-linear data with input data and is modulated by the input modulation gate and written onto the internal cell state. As zero-mean input converges, faster learning time is reduced. This Gate is treated as fineness providing concept even the involvement of these gates is very less. This modulation gate is hence frequently regarded as a component of the input gate. The output gate decides what output will be produced from the current internal state of the cell. In LSTM, the internal cell state is forwarded along with the hidden state. The mathematical algorithm that describes the forward pass of an LSTM network is explained below.

The proposed multiple disease prediction model consists of two LSTM layers. The first LSTM layer has 128 units. The first LSTM layer receives input sequences of length 132 and returns sequences as output. The second LSTM layer has 64 units. The second LSTM layer receives the output sequences from the first layer and returns a single output. A dropout layer is inserted

1. ***Initialization:***
    - *Initialize the hidden state h_o and cell state c_o to zeros or small random values.*
    - *Initialize the weight metrics, W and U, and the bias vectors, b, for each LSTM.*
2. ***Forward pass through each time step, t***

    For each time step, t
    - *Calculate the input gate, i_t, using the current input, x_t, the previous hidden state, h_{t-1}, and the current cell state, c_{t-1}:* $i_t = \boldsymbol{\sigma}\,((W_i \times x_t) + (U_i \times h_{t-1}) + b_i)$.
    - *Calculate the forget gate, f_t, using the same inputs:*

      $f_t = \boldsymbol{\sigma}\,((W_f \times x_t) + (U_f \times h_{t-1}) + b_f)$.
    - *Calculate the output gate, o_t, using the same inputs:*

      $o_t = \boldsymbol{\sigma}\,((W_o \times x_t) + (U_o \times h_{t-1}) + b_o)$.
    - *Calculate the candidate cell state, g_t, using the same inputs and the activation function, such as the hyperbolic tangent (tanh):*

      $g_t = \boldsymbol{tanh}\,((W_g \times x_t) + (U_g \times h_{t-1}) + b_g)$.
    - *Update the current cell state, c_t, by combining the previous cell state, c_{t-1}, with the candidate cell state, g_t, using the input and forget gates:*

      $c_t = f_t \odot c_{t-1} + i_t \odot g_t$.
    - *Update the current hidden state, h_t, by applying the output gate to the cell state:*

      $h_t = o_t \odot tanh(c_t)$
    - *Pass the hidden state, h_t, to the next time step.*
3. ***Output:***
    - *After processing all time steps, the final hidden state, h_T, can be used for prediction or passed to a dense layer for further processing.*

between the LSTM layers. Dropout serves as a regularization technique by preventing the model from relying too heavily on any specific feature or combination of features during training. It helps reduce overfitting by introducing randomness and forcing the network to learn more robust and generalized representations. Finally, a dense layer with sigmoid activation is added, producing the final output with 41 units. The dense layer is a fully connected layer where each input unit is connected to every output unit. The model summary of the proposed model is tabulated in Table 1.

The model architecture is shown in Figure 10.

3.4. Real-time interactive system

A real-time interactive telemedicine system is a technology-enabled platform that allows healthcare providers and patients to connect and communicate remotely in real-time for medical consultations, diagnosis, treatment, and monitoring. It leverages telecommunication tools and digital technologies to bridge the gap between patients and healthcare professionals, enabling access to healthcare services regardless of physical location. The system should support high-quality video and audio communication between healthcare providers and patients in real-time. It allows for face-to-face interaction, visual examination, and discussion of medical concerns, providing a more personalized and interactive experience.

Table 1. Model summary.

Total parameters	118,633
Trainable Parameters	118,633
Non- trainable Parameters	0

4. Results and discussion

4.1. Hardware and software setup

The proposed model was executed after the dataset has been prepared. The acquired dataset can be split into two sets: training set (70%) and test set (30%). The model was built, trained, and tested on Google Collaboratory using LSTM, and the entire process was carried out using Python and TensorFlow. The Adam optimization method was used for prediction. The model performed.

20 epochs of training with a batch size of 50. The hyperparameters utilized for this work is tabulated below (Table 2).

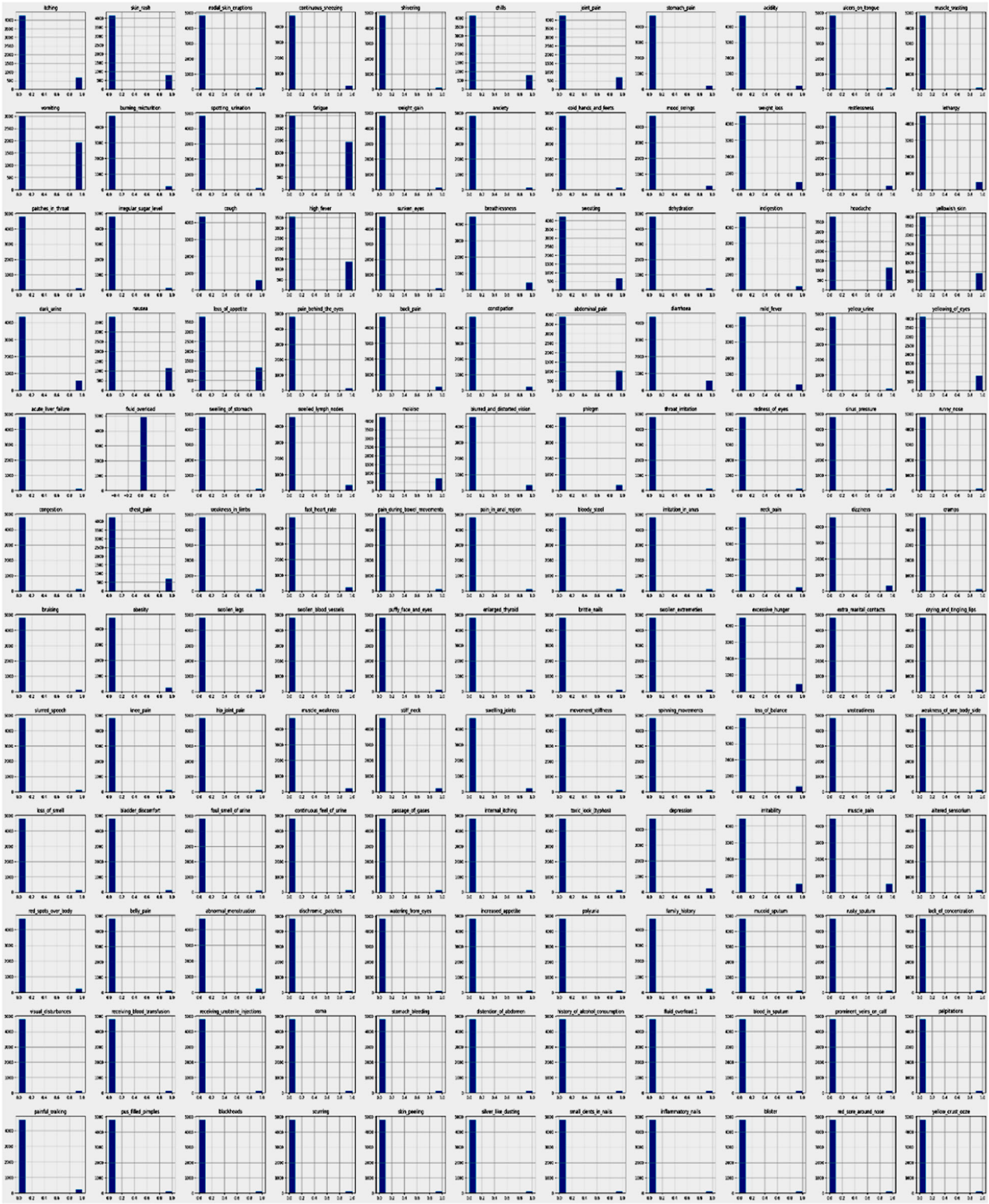

Figure 8. Histogram visualization of dataset.

Table 2. Hyperparameters.

Loss	Categorical crossentropy
Optimizer	Adam
Activation Function	Sigmoid
Batch Size	50
Number of Epochs	20

4.2. Experimental results

The performance of a model is often assessed using the metrics of accuracy and loss. A measure of accuracy is the percentage of instances that were correctly predicted out of all the instances in the dataset. The accuracy is determined by dividing the number of right predictions by the total number of predictions. An overall evaluation of the model's performance in terms of accurate prediction is provided by accuracy. However, it occasionally can be deceptive, particularly when the dataset is imbalanced, which means that some classes have a disproportionately large number of instances compared to other classes. In such situations, accuracy

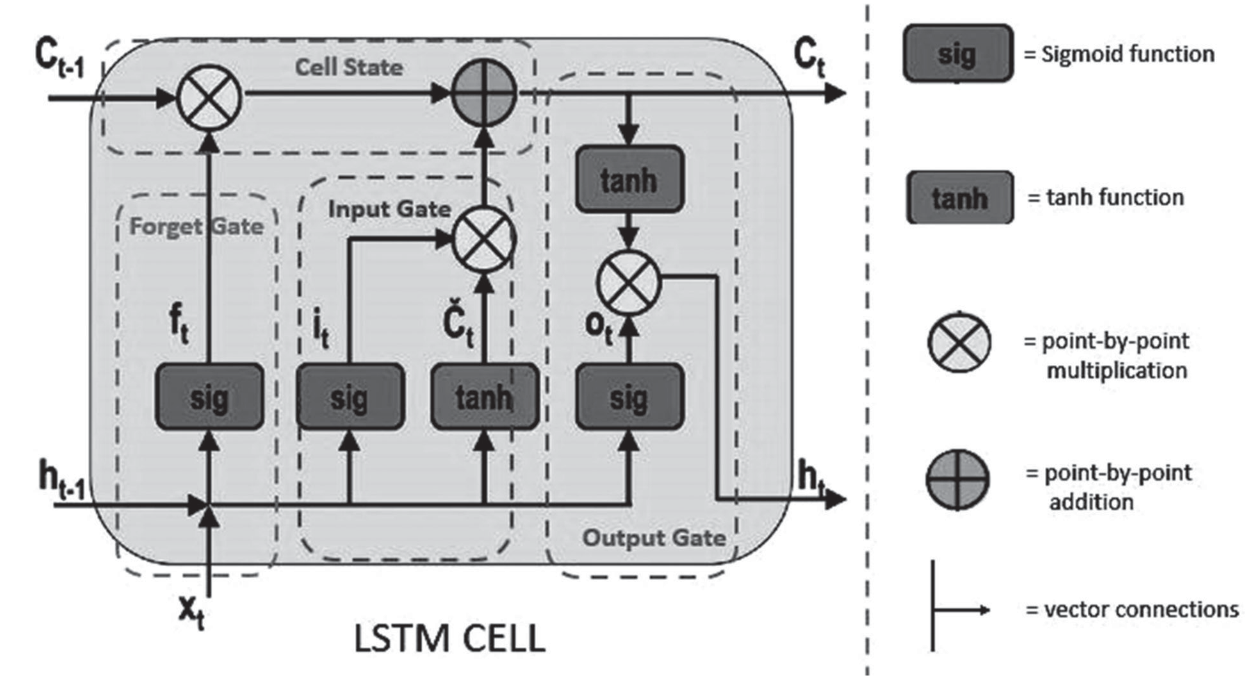

Figure 9. Architecture of LSTM.

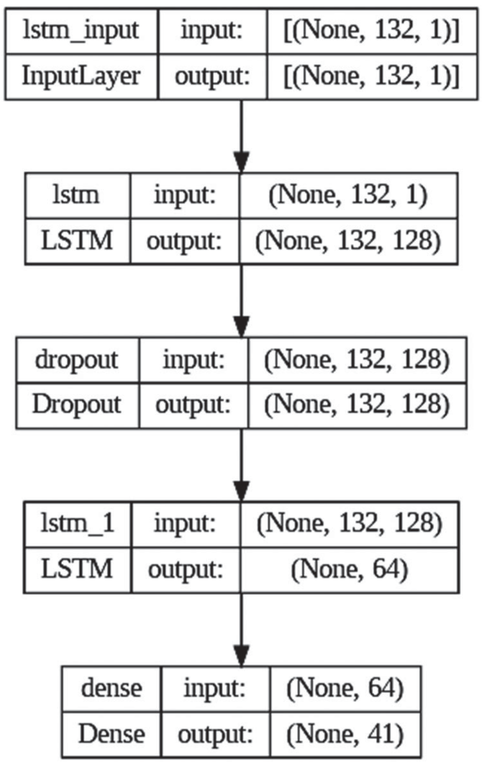

Figure 10. Architecture of proposed model.

may not give a complete picture of the model's performance. The difference between the model's predicted outputs and the actual targets is measured by loss, which is sometimes referred to as the cost or objective function. The objective of training is to reduce this loss function. Depending on the kind of issue being resolved, a particular loss function may be employed such as mean squared error or cross-entropy loss. In order to evaluate and enhance a model's performance, it is essential to consider both accuracy and loss. The proposed model got an accuracy of 98.51% and loss of 0.0842.

The accuracy plot of a model is significant as it provides valuable insights into the performance and learning progression of the model during training. The accuracy plot allows you to visually assess how well the model is learning from the training data. It shows the trend of accuracy improvement over epochs or iterations. A rising accuracy curve indicates that the model is effectively learning and making more accurate predictions over time. The accuracy plot helps determine whether the model has converged or not. The accuracy plot can indicate if the model is over fitting or underfitting the training data. When the model gets highly specialized to the training data and performs poorly on unobserved data, over fitting has taken place. On the other side, under fitting happens when the model fails to capture the underlying patterns in the data, leading to low accuracy overall. The model architecture or hyper parameters may need to be adjusted as a result of these problems. The performance of a model during training is significantly illustrated by the accuracy plot. It aids in measuring performance, gauging convergence, spotting over- or under fitting, contrasting models, directing optimization, and clearly articulating the model's development. The accuracy plot of the proposed model is illustrated in Figure 11.

The loss plot is a crucial tool in deep learning for monitoring model optimization, detecting overfitting or underfitting, guiding early stopping decisions, tuning hyperparameters, comparing models, and

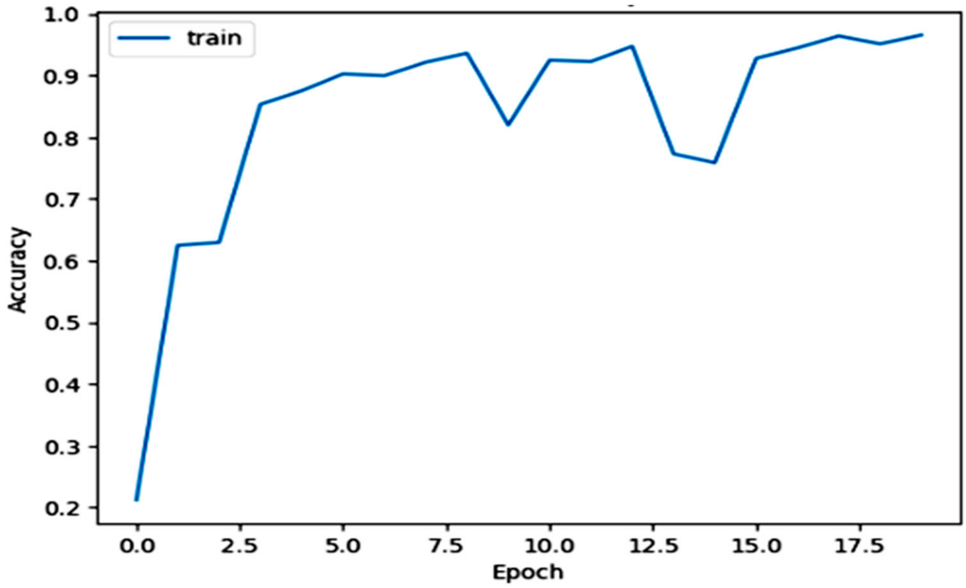

Figure 11. Accuracy plot of proposed model.

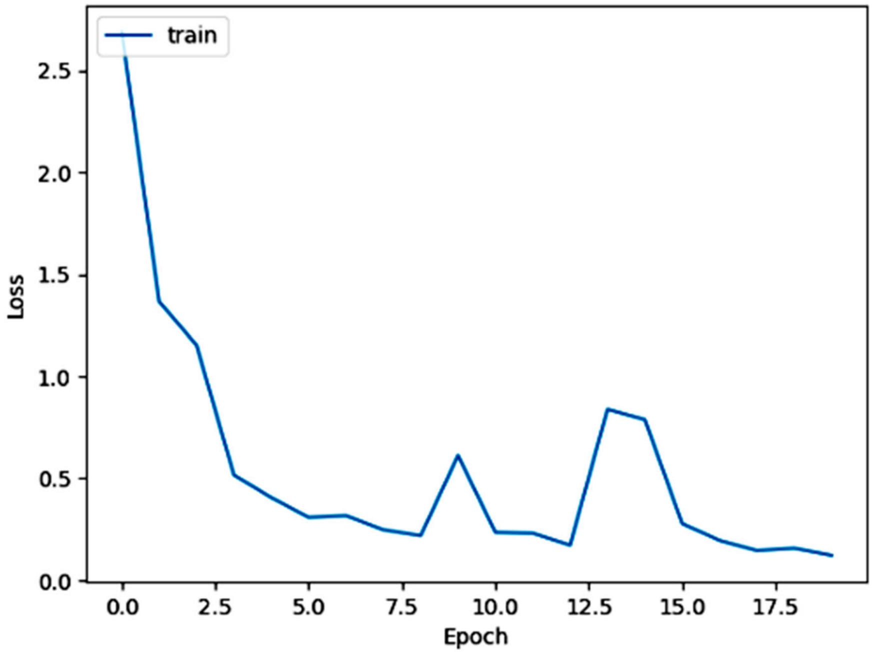

Figure 12. Loss plot of proposed model.

effectively communicating the training process. It aids in understanding the behaviour and performance of the model during training and facilitates the improvement of model accuracy and generalization. The loss plot is a significant visualization tool in deep learning that provides valuable insights into the training process and the performance of a model. The loss plot helps monitor the optimization process during training. The loss function measures the discrepancy between predicted and actual values. By observing the loss plot, you can determine whether the model is converging or if further training is necessary. A decreasing loss curve indicates that the model is optimizing and learning from the data. The loss plot provides a concise summary of the model's optimization trajectory and can support discussions, decision-making, and reporting. The loss plot of the proposed model is visualized in Figure 12. Finally, the proposed model predicts various diseases based on the given input. Table 3 represents some of the obtained prediction outputs.

The obtained prediction results are accessed by the healthcare providers. The real-time interactive system provides an effective communication between the patients and healthcare providers for diagnosis and treatment planning. It allows patients to receive quality healthcare services without the need to travel long-distances, saving time, costs, and inconvenience associated with transportation.

Table 3. Obtained prediction outputs.

Input data	Predicted disease
[1,1,10,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]	Fungal Infection
[0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]	AIDS
[1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,1 1,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0 0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0, 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]	Chronic Cholestasis
[0,0,0,0,0,0,0,1,1,1,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]	GRED
[1,1,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]	Drug Reaction
[0,0,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]	Allergy
[0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,1,0,0,1,0,0,1 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]	Malaria

5. Conclusion

Telemedicine systems hold significant importance in modern healthcare by revolutionizing the way healthcare services are delivered. It eliminates the need for long-distance travel by bridging the gap between patients and healthcare professionals, especially in rural or underserved areas. This improves healthcare accessibility, especially for individuals who face geographical or logistical barriers to traditional in-person care. The real-time interactive telemedicine systems facilitate timely diagnosis and treatment, allowing patients to connect with healthcare professionals promptly. This immediate access to healthcare services is crucial in emergency situations and helps expedite medical interventions, leading to better health outcomes. Data mining techniques play a significant role in telemedicine systems by extracting valuable insights and knowledge from large volumes of healthcare data. A multiple disease prediction model in telemedicine system using deep learning approach was proposed in this paper. The proposed model is evaluated on the multiple disease dataset. The proposed model achieved better prediction performance with an accuracy of 98.51%. A real-time interactive system was introduced for upcoming projects for the purpose of diagnostic and treatment planning, the system facilitates efficient communication between patients and healthcare professionals. Patients can obtain high-quality medical care without having to go far, which saves money, time, and trouble related to transportation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Wootton R, Craig J, Patterson V. Introduction to telemedicine. United Kingdom: CRC Press; 2017.

[2] Lilly CM, Motzkus C, Rincon T, et al. ICU telemedicine program financial outcomes. Chest. 2017;151(2):286–297. doi:10.1016/j.chest.2016.11.029

[3] Mehrotra A, Jena AB, Busch AB, et al. Utilization of telemedicine among rural Medicare beneficiaries. Jama. 2016;315(18):2015–2016. doi:10.1001/jama.2016.2186

[4] Maheu M, Whitten P, Allen A. eHealth, Telemedicine & Telehealth: A comprehensive guide, New York; 2004.

[5] Chowdhury SM, Kabir MH, Ashrafuzzaman K, et al. A telecommunication network architecture for telemedicine in Bangladesh and its applicability. Intern J Digit Content Technol Applic. 2009;3(3):4), doi:10.4156/jdcta.vol3.issue3.20

[6] Huang EY, Knight S, Guetter CR, et al. Telemedicine and telementoring in the surgical specialties: A narrative review. Amer J Surg. 2019;218(4):760–766. doi:10.1016/j.amjsurg.2019.07.018

[7] Larose DT. Discovering knowledge in data. An introduction to data mining. New Jersey: John Wiley & Sons Publisher; 2005; ISBN 0-471-66657-2.

[8] Dash M, Shadangi PY, Muduli K, et al. Predicting the motivators of telemedicine acceptance in COVID-19 pandemic using multiple regression and ANN approach. J Stat Manage Syst. 2021;24(2):319–339. doi:10.1080/09720510.2021.1875570

[9] Ahmed ST, Sandhya M, Sankar S. TelMED: dynamic user clustering resource allocation technique for MooM datasets under optimizing telemedicine networks. Wirel Person Commun. 2020;112(2):1061–1077. doi:10.1007/s11277-020-07091-x

[10] Sadineni PK. Developing a model to enhance the quality of health informatics using big data. In 2020 fourth international conference on I-SMAC (IoT in social, mobile, analytics and cloud)(I-SMAC) (pp. 1267–1272). IEEE; 2020.

[11] Sornalakshmi M, Balamurali S, Venkatesulu M,… Muthu BA. Hybrid method for mining rules based on enhanced Apriori algorithm with sequential minimal optimization in the healthcare industry. Neural Comput Applic. 2020: 1–14.

[12] Choi SY, Chung K. Knowledge process of health big data using MapReduce-based associative mining. Pers Ubiquitous Comput. 2020;24:571–581. doi:10.1007/s00779-019-01230-3

[13] Priyadarshan DJ, Sanjay KK, Kathiresan S, et al. Patient health monitoring using IoT with machine learning. Intern Res J Eng Technol (IRJET). 2019;6(03.

[14] Sandhiya R, Sundarambal M. Clustering of biomedical documents using ontology-based TF-IGM enriched semantic smoothing model for telemedicine applications. Cluster Comput. 2019;22:3213–3230. doi:10.1007/s10586-018-2023-4

[15] Anusuya TK, Maharajothi P. A survey of telemedicine services using data mining. International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN, 2456–3307; 2019.

[16] Thouheed Ahmed S, Sandhya M. Real-time biomedical recursive images detection algorithm for Indian telemedicine environment. In: Cognitive informatics

and soft computing: proceeding of CISC 2017. Springer Singapore; 2019. p. 723–731.
[17] Sukumar P, Monika G, Gokila D, et al. An NLP based ontology architecture for dealing with Heterogeneous data to telemedicine systems. South Asian J Eng Technol. 2019;8(1):89–92.
[18] Sarkar BK, Sana SS. An e-healthcare system for disease prediction using hybrid data mining technique. J Model Manage. 2019;14(3):628–661. doi:10.1108/JM2-05-2018-0069
[19] Ahmed MIB. Virtual clinic: A CDSS assisted telemedicine framework. In: Telemedicine technologies. Academic Press; 2019. p. 227–238.
[20] Ahmed ST, Sandhya M, Sankar S. An optimized RTSRV machine learning algorithm for biomedical signal transmission and regeneration for a telemedicine environment. Procedia Comput Sci. 2019;152:140–149. doi:10.1016/j.procs.2019.05.036
[21] Sandhiya R, Sundarambal M. Chicken swarm optimization-based clustering of biomedical documents and health records to improve telemedicine applications. Intern J Enterpr Netw Manage. 2019;10(3-4):305–328. doi:10.1504/IJENM.2019.103158
[22] Peral J, Ferrández A, Gil D, et al. An ontology-oriented architecture for dealing with heterogeneous data applied to telemedicine systems. IEEE Access. 2018;6:41118–41138. doi:10.1109/ACCESS.2018.2857499
[23] https://github.com/ybifoundation/Dataset/raw/main/MultipleDiseasePrediction.csv
[24] Sherstinsky A. Fundamentals of recurrent neural network (RNN) and long shortterm memory (LSTM) network. Physica D: Nonlin Phenomena. 2020;404:132306), doi:10.1016/j.physd.2019.132306
[25] https://www.pluralsight.com/guides/introduction-to-lstm-units-in-rnn.
[26] Kadum SY, Salman OH, Taha ZK, et al. Machine learning-based telemedicine framework to prioritize remote patients with multi-chronic diseases for emergency healthcare services. Netwo Model Anal Health Inform Bioinform. 2023;12(1):11), doi:10.1007/s13721-022-00407-w
[27] Gupta R, SmitaKumari A, Ambasta RK, et al. New era of artificial intelligence and machine learning-based detection, diagnosis, and therapeutics in Parkinson's disease. Ageing Res Rev. 2023: 102013), doi:10.1016/j.arr.2023.102013
[28] Bhat SS, Banu M, Ansari GA, et al. A risk assessment and prediction framework for diabetes mellitus using machine learning algorithms. Healthc Analyt. 2023: 100273), doi:10.1016/j.health.2023.100273
[29] Chakraborty C, Gupta B, Ghosh SK. A review on telemedicine-based WBAN framework for patient monitoring. Telemed e-Health. 2013;19(8):619–626. doi:10.1089/tmj.2012.0215
[30] Chakraborty C, Gupta B, Ghosh SK. Mobile metadata assisted community database of chronic wound images. Wound Med. 2014;6:34–42. doi:10.1016/j.wndm.2014.09.002
[31] Chakraborty C. Mobile health (m-Health) for tele-wound monitoring: Role of m-Health in wound management. In: Mobile health applications for quality healthcare delivery. IGI Global; 2019. p. 98–116.
[32] Verma S, RishabhaMalviya M, Tripathi BD. Tele-health monitoring using artificial intelligence deep learning framework. In: Deep learning for targeted treatments: transformation in healthcare; 2022. p. 199–228.
[33] Kadu A, Singh M, Ogudo K. A novel scheme for classification of epilepsy using machine learning and a fuzzy inference system based on wearable-sensor health parameters. Sustainability. 2022;14(22):15079. doi:10.3390/su142215079