Internet of Things Enabled Disease Diagnosis Model for Smart Healthcare Systems

Ms. A. Priyanka, M.Tech student,CSE:BVCE, priyankaainagadda@gmail.com

Dr. P. Devabalan, Professor,CSE:BVCE, devabalanme@gmail.com
Shubhashish Jena, CSE:OUTR sjena1998@gmail.com
Subhashree Jena, PhD Scholar, Annamalai University sjena96@gmail.com

Article Info
Page Number: 5398 - 5411
Publication Issue:
Vol 71 No. 4 (2022)

Abstract

Recent advances in the Internet of Things (IoT), cloud computing, and Artificial Intelligence (AI) have turned the traditional healthcare system into an intelligent healthcare system. Adding important technologies like IoT and AI could make medical care better. Putting IoT and AI together gives the healthcare business new ways to do things. In this way, the current work presents a new way for AI and IoT to work together to help intelligent healthcare systems find sickness. In this paper, AI and IoT convergence methods are used to try to build a model that can find heart disease and diabetes. The provided model has several steps, such as collecting data, preprocessing it, classifying it, and adjusting the parameters. IoT devices, like wearables and sensors, make it easy to collect data, which can then be used by AI to figure out what's wrong with someone. Cascaded Long Short Term Memory (CSO-CLSTM), which is based on the Crowd Search Optimization algorithm, is used in the proposed method for finding diseases. CSO is used to fine-tune the "weights" and "bias" parameters of the CLSTM model to improve the classification of medical data. The isolation Forest (iForest) method is also used to leave outliers from this study. The CLSTM model is much better at making diagnoses when CSO is used. Healthcare data were used to prove that the CSO-LSTM model works. A new version of LSTM, called Convolution 2D Neural Network (CNN2D), has been added to the CNN2D algorithm, along with a CSO features selection technique. Experiments using data from the Heart and Diabetes show that the extension strategy is the most accurate.

Article History
Article Received: *25 March 2022*
Revised: *30 April 2022*
Accepted: *15 June 2022*
Publication: *19 August 2022*

Words like "Internet of Things," "convergence," "cloud computing," "artificial intelligence," "intelligent healthcare," and "disease diagnostics" are all important.

AN INTRODUCTION:

In the past few years, the healthcare industry has used information technology to create new applications and improve diagnostic and treatment methods. Most of the time, complex procedures and scientific ideas are used to create large amounts of digital data. Because of recent improvements in information technology, better healthcare apps have been made.

Intelligent health care apps are also thought to be easy to use, look good, and be flexible. These changes are part of the extension of the clinical model (from care based on diseases to care based on patients), the development of information technology (from medical data to regional medical data), the extension of clinical management (from general management to personal management), and the change from a system for treating diseases to a system for preventing diseases [1]. The following changes are all about meeting people's basic needs. This will improve the health service's knowledge and show how intelligent medicine will be used in the future to make health care more effective.

Doctors, patients, and clinical and research institutions are all involved in modern medical care. There are many things that need to be taken care of, such as preventing and watching for illness, diagnosing and treating it, clinical management, making decisions about health, and medical research. Modern healthcare is thought to be different because of things like mobile internet, Cloud Computing (CC), big data, 5G systems, microelectronics, Artificial Intelligence (AI), and smart biotechnology. These methods are used at every level of modern medical care. Health monitoring technology that can be worn or carried can be used whenever the patient thinks it is important. They can get help with therapy through virtual assistance and control their homes through remote features. Intelligent clinical decision support systems can help doctors guide and improve their diagnostic procedures.

1.1 Health Information for Smart Medical Care:

Intelligent healthcare is more than just a feat of technology. Because of this fast growth, all levels of health care facilities are now based on information. Regional healthcare ecosystems, like homes and communities, need to be driven more and more by data. The health care industry has used data for many years.

By making them based on information, we can collect data that can be analysed and used to make healthcare systems more effective, personalised, and easy to use.

1.2 Controlling Diseases: Chronic diseases like diabetes, heart disease, and arthritis need ongoing care like diagnosis, medication, monitoring of symptoms, and changes in lifestyle.

With health software, smart devices, health bots/telemedicine, and standardised procedures, intelligent health care systems can automate most of these tasks.

This improves the quality of life for patients and makes care easier and cheaper to give.

Smart machines and the Internet of Things:

IoT stands for the use of artificial intelligence (AI) in the architecture of the internet of things (IoT). AIoT is meant to improve how IoT works, how people interact with machines, how data is managed, and how data is analysed.

Getting and analysing data:

There is no way to treat a certain health problem. When giving therapy, the medicines and medical history of each patient must be taken into account.

Again, not all medical situations are out of the ordinary. There are always cases of this happening to other people, and it's likely to happen again. Smart Healthcare uses networked devices that can gather a lot of useful information in real time.

1.3: People think that the most advanced healthcare apps are easy to use, have nice graphics, and can be used for more than one thing. These changes are part of the extension of the clinical model (from care based on diseases to care based on patients), the development of information technology (from medical data to regional medical data), the extension of clinical management (from general management to personal management), and the change from a system for treating diseases to a system for preventing diseases [1]. The following changes are all about meeting people's basic needs. This will improve the health service's knowledge and show how intelligent medicine will be used in the future to make health care more effective.

1.4 Statement of the Problem: In a world where technology and innovation affect every part of our lives, it seems crazy to let something as important as healthcare fall behind. The Internet of Things (IoT), improvements in Artificial Intelligence (AI), Big Data, and Cloud computing may all work together to create sophisticated global healthcare systems.

1.5 Function:\sThis paper tries to build a model for detecting heart disease and diabetes by combining AI and IoT convergence methodologies. The provided model has several steps, such as collecting data, preprocessing it, classifying it, and adjusting the parameters. IoT devices, like wearables and sensors, make it easy to collect data, which can then be used by AI to figure out what's wrong with someone.

1.6 Scope of Work:

AI and the Internet of Things are two types of cutting-edge technology that are used in the proposed research. The author of the proposed study says that IoT and wearable devices with cloud and internet connections could collect data from the patient's body. This data would then be analysed by artificial intelligence to figure out if the patient's health is normal or bad.

In the proposed research, Artificial Intelligence is a group of different algorithms, like Isolation Forest, that can be used to get rid of wearable data or datasets that don't fit with the rest or aren't needed.

The Crowd Search method will be used on the processed data to improve the properties of the dataset, picking out attributes that help improve the accuracy of predictions. The most accurate feature will be chosen by an algorithm that evaluates and records the results of each feature over and over again. This article talks about the Crow search method, which was based on how crows act.

2.2 Scientific Contributions: The Heart/Diabetes disease is treated with SVM, KNN, J48, Nave

Bayes, Crowd Search Algorithm (CSA), and Convolution 2D Neural Network (CNN2). Dataset got from github.com on train test ratio 303 records from dataset and then split dataset into train and test parts using 242 (80%) training records and 61 (20%) test records to Disease Diagnosis.

III. Plan for how to do it

In the proposed study, the author uses cutting-edge technologies like AI and the Internet of Things to improve the way healthcare is delivered. The author of the proposed study says that IoT and wearable devices with cloud and internet connections could collect data from the patient's body. This data would then be analysed by artificial intelligence to figure out if the patient's health is normal or bad.

3.1 Calculations:

In this study, we use the SVM, KNN, J48, Nave Bayes, CSO-CLSTM Algorithm, Convolution 2D Neural Network, and CSO-CLSTM Algorithm (CNN2D)

3.1.1 SVM Algorithm:

Depending on the dataset, machine learning might involve using different machine learning techniques to predict and sort data. A linear classification and regression model is the Support Vector Machine (SVM). It can be used in many different ways and can deal with both linear and nonlinear problems. The SVM method generates a line or hyperplane that divides data into classes. The RBF kernel, which is also known as the radial basis function kernel, is used in a number of kernelized machine learning methods. It is often used to classify things with a support vector machine. One way to think of a hyper plane is as a line that divides and sorts a set of data based on two attributes (like the figure above).

3.1.2 Bayes Standard:

Compared to other, more complex algorithms, the Naive Bayes classifier may work well. Because class distributions are different, each can be looked at on its own as a one-dimensional distribution. This, in turn, makes the dimension curse less of a problem.

Given the value of the class variable, the Nave Bayes classifier is one of the Bayes' theorem-based probabilistic classifiers that has strict requirements for the independence of features. This method is a set of algorithms for supervised learning.

Finds the K-neighbor whose distance is the shortest.

K-Nearest Neighbor is one of the most simple Machine Learning algorithms based on Supervised Learning.

The K-NN algorithm saves all of the available data and decides how to classify a new data point based on how similar it is to other data points that have already been saved. This means that the K-NN method can be used to quickly classify new data as soon as it becomes available.

It is an algorithm based on the population. Particle Swarm Optimization is like the Crowd Search Method (PSO). Crow Search Algorithm is like the way crows think. The CSA is based on what the crows know.

3.1.5 J48

We have been putting information into groups using the most common method, J48. The J48 technique is used to put a lot of applications into groups because it gives accurate results. The J48 method is one of the best algorithms for analysing both discrete and continuous data. When used as an example, it takes up more memory and makes it harder and less accurate to sort medical data into categories.

3.1.6 Convolution 2D Neural Network (CNN2D), which is an improved version of LSTM: We added the CSO features selection technique to the CNN2D algorithm, and testing with Heart and Diabetes data showed that the extension algorithm gave the most accurate results.

3.3 Framework/Architecture:

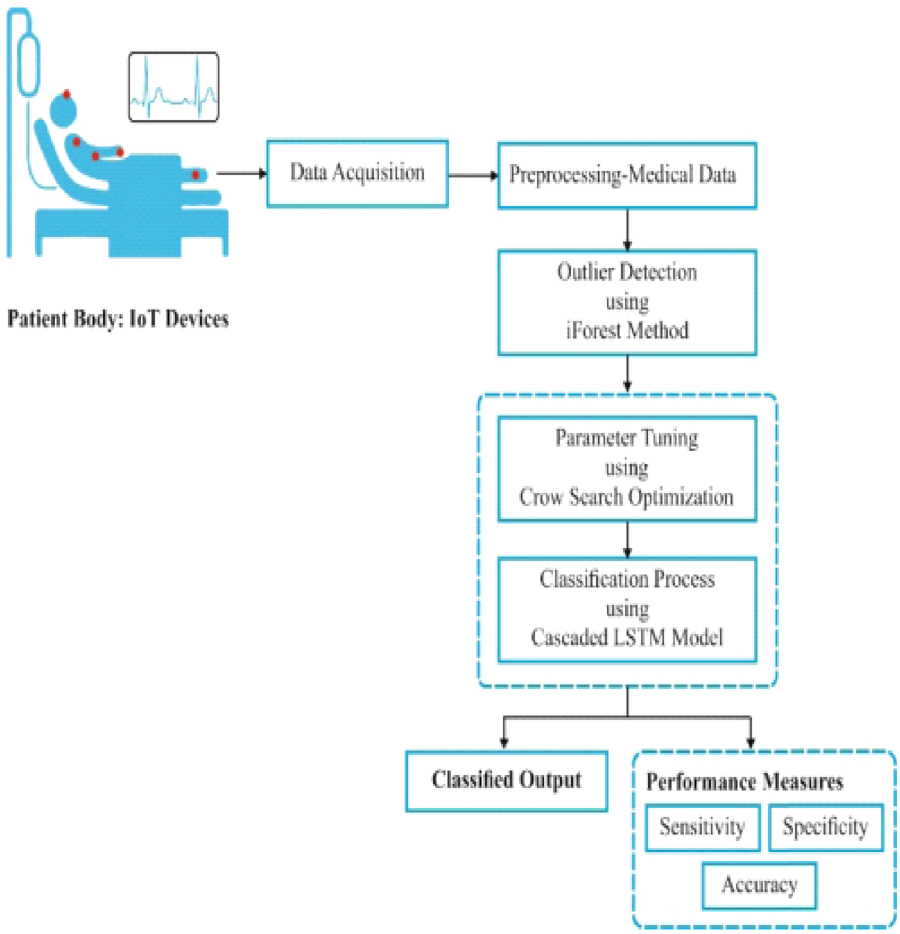


Figure1: Architecture/Framework.

3.4 Design of Algorithms and Processes:

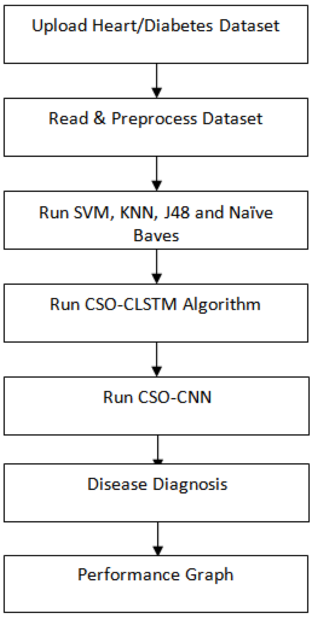


Figure2: Design of an Algorithm and a Process

Upload Heart/Diabetes Dataset: We will use this module to upload datasets about heart disease or diabetes.

Read and Preprocess Dataset: This module will be used to read the values in the dataset, get rid of any missing values, normalise the values so they are all between 0 and 1, and split the dataset into train and test parts.

We'll use this module to train SVM, KNN, J48, and Nave Bayes on training data, and then we'll use test data to see how well the models can predict.

CSO-CLSTM Method: In this module, we will use the Isolation Forest technique to get rid of OUTLIERS from the dataset, then the Crowd Search algorithm to improve the features, and finally the LSTM model will be trained. The trained model will be used to look at test data to figure out how accurate LSTM is.

Convolution 2D Neural Network (CNN2D) is an improved version of LSTM. We've added a CSO features selection strategy to the CNN2D algorithm and tested it with data from both the heart and diabetes, giving the extension algorithm the best accuracy.

Performance Comparison Graph: We will use this module to make a graph for each method that compares how well it works.

IV. How it was done and the results

4.1 Getting information

Researchers most often use https://github.com/rashida048/Datasets/blob/master/Heart.csv as their dataset.

It read the whole 303-record dataset and split it into train and test parts, using 242 (80%) of the records for training and 61 (the other 20%) for testing.

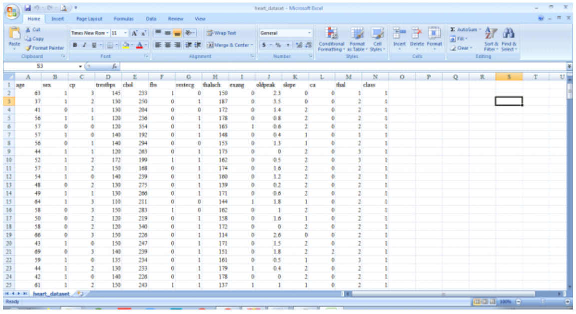


Figure3: Data Gathering

4.2 Evaluation Metrics:

F1-score, Accuracy, and Receiver Operating Characteristics-Area Under the Curve are used to measure how well our models work (ROC-AUC). To figure out F1-score and Accuracy, you have to look at Precision and Recall.

• False Positive Rate is what FPR stands for.

• The number of real positives (TPR)

• the accuracy • a high level of accuracy • remember

The F1-score is:

The following things affect how value is calculated:

• True positive (TP) is the number of correctly detected events.

• False negative (FN): The number of unnecessary events that were wrongly predicted.

False-positive (FP) refers to the number of wrongly predicted events.

• True negative (TN) is the number of events that could have happened but were not needed.

The False Positive Rate (FPR) is a measure of how well machine learning works. Here's how to understand this:

FPR=FP/(FP+TN)

Rate of True Positives (TPR): This is a word that means the same thing as recall.

TPR=FP/(FP+TN)

The most important performance metric is accuracy, which is measured as a simple ratio of the number of observations that were predicted correctly to the total number of observations.

Accuracy=(TN+TP)/(TP+FP+TN+FN)

Keep in mind that it is the number of correctly predicted positive observations out of the total number of original observations.

TP / (TP + FN) = remember

Precision is used to figure out how to calculate the right values. This means taking the number of expected positives away from the number of positives that were correctly predicted. It is said to be

Precision is equal to TP divided by (TP + FP).

The F1-score is a way to measure how accurate and easy to remember a model is. It is the average of how well the model works and how well it remembers things. This is also called an F-score. It is written as F1=2 (Precision Recall/Precision + Recall).

The area under the ROC curve can be used to get the ROC-AUC, which is a useful classification statistic.

4.3 Result:

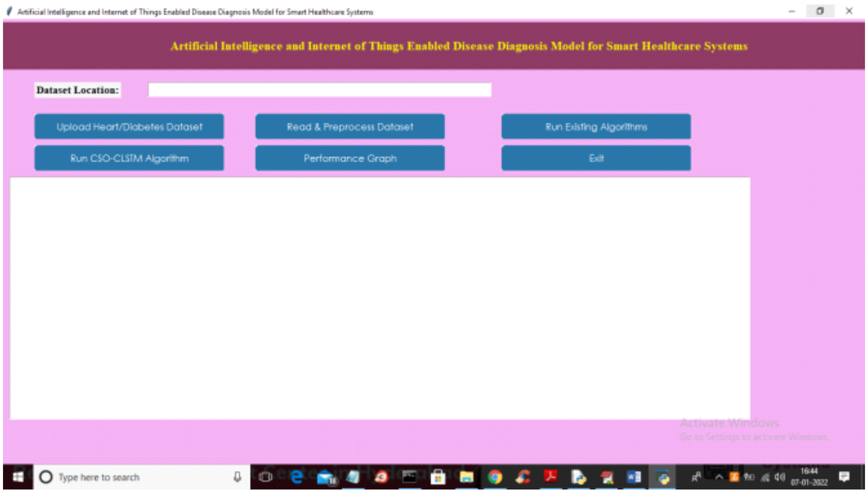


Figure4: Upload the Heart/Diabetes Dataset

To upload the dataset, click the "Upload Heart/Diabetes Dataset" button in the image above, and the figure below will appear.

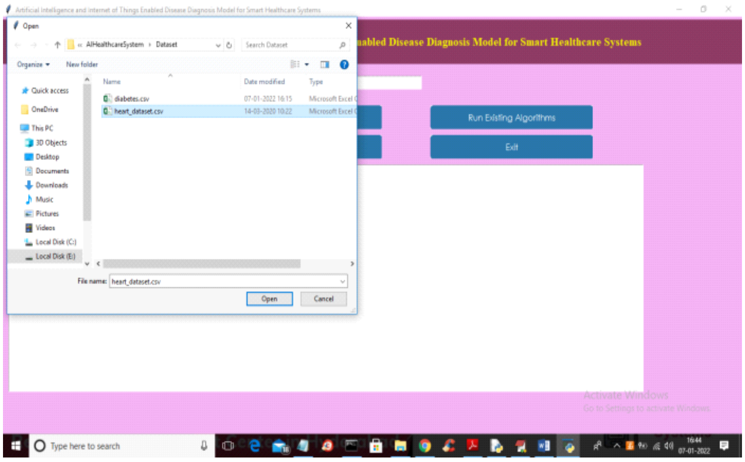


Figure5: Choosing and Uploading 'heart dataset.csv'.

In the image above, choose and upload the 'heart dataset.csv' file, then click the 'Open' button to load the dataset and get the figure below.

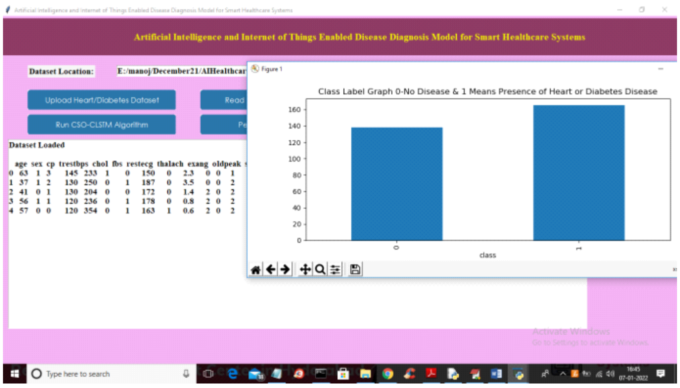
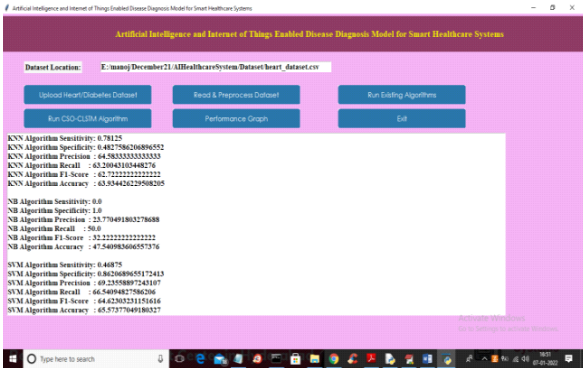


Fig.6:Dataset loaded and in graph' button in the above figure.

The dataset has been loaded in the image above, and the X-axis indicates 0 (no illness) and 1 (existence of heart or diabetes), while the y-axis reflects the number of entries in each class. Close the above graph, then click the 'Read & Preprocess Dataset' button to read and normalise the dataset, followed by the 'Heart/Diabetes Dataset' button to upload the dataset and receive the figure below.

Fig. 7 The application read the whole dataset of 303 entries.

The software read all 303 records from the dataset before splitting them into train and test halves, using 242(80%) of the records for training and 61 (the remaining20%) for testing. Now that the dataset is available, select 'Run Existing Algorithm' to train the existing algorithm and obtain the results shown below.



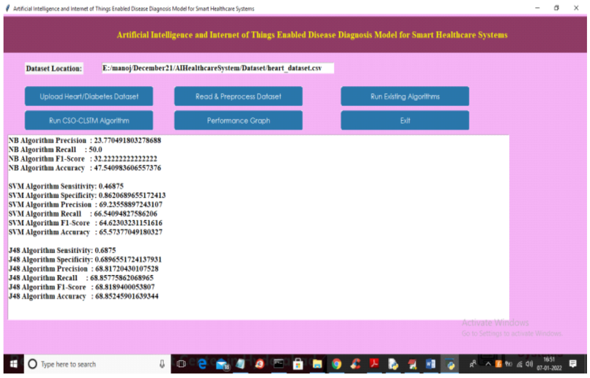


Figure8: 'Run the CSO-CLSTM algorithm.'

Each algorithm's accuracy, precision, recall, FSCORE, sensitivity, and specificity are shown in the diagram above. The results are shown below after clicking the "Run CSO-CLSTM Algorithm" button

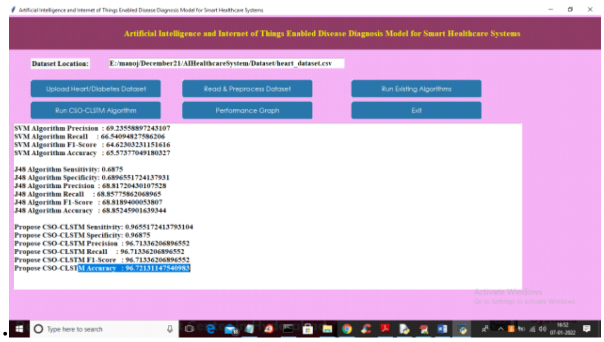


Fig. 9 CSO-CLSTM achieved 96% accuracy.

To see the graph below, click on the 'Performance Graph' button. As observed in the blue text emphasised in the image above, the suggested CSO-CLSTM offers a higher degree of accuracy than existing algorithms

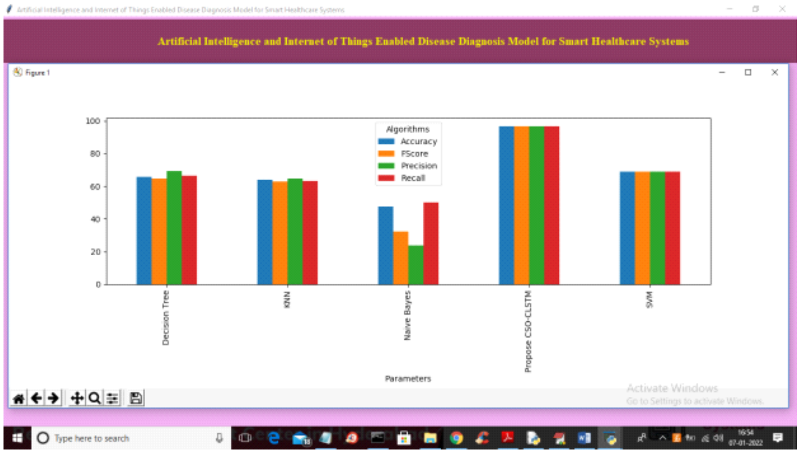


Fig.10 CSO-CLSTM demonstrated excellent performance.

The x-axis in the graph above represents the algorithm name, while the y-axis reflects the performance of each method, with each colour bar indicating a different statistic, such as accuracy or precision, and so on. The graph above shows that the suggested CSO-CLSTM algorithm outperforms the present technique in terms of performance.

Outcomes of extension:

As an expansion to this project, we have included Convolution 2D Neural Network (CNN2D), which is an enhanced version of LSTM. Furthermore, we introduced the CSO features selection approach to the CNN2D algorithm, and the trial with Heart and Diabetes data shown that the extension algorithm delivers the maximum degree of accuracy.

We've included a disease prediction model that uses the TEST data file from the 'Dataset' folder to forecast illness. Use the testHeart.csv file for disease prediction and the testDiabetes.csv file for diabetes prediction when training the Heart dataset.

The following are the results of each algorithm.

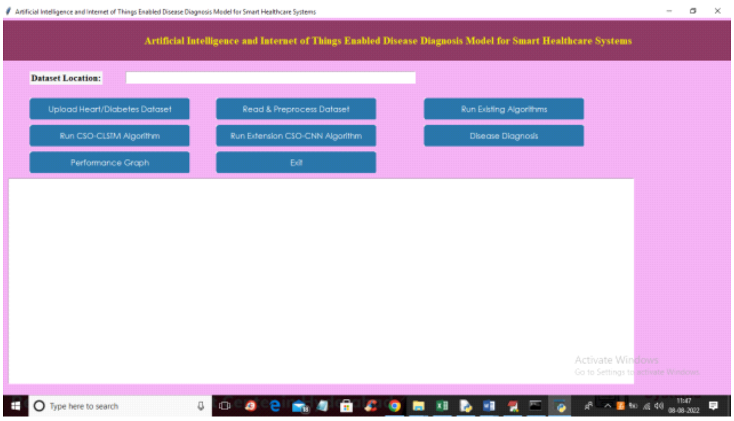


Figure11: Dataset Upload

In the graphic above, upload the dataset and then execute all buttons, including the extension.

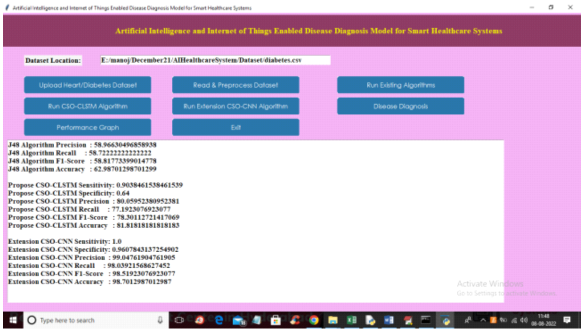


Fig. 12 CSO LSTM achieved 81% accuracy.

On the diabetic dataset, we achieved 81% accuracy with the initial CSO LSTM and 98% accuracy with the extension CSO-CNN in the picture above. To submit test data and get the result indicated below, click the 'Disease Diagnosis' button.

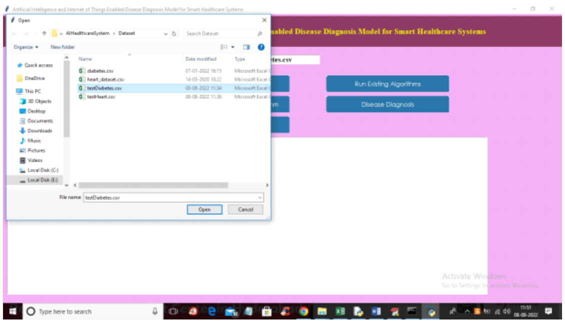


Figure13: Importing 'testDiabetes.csv'.

In the above example, selecting and uploading the testDiabetes.csv file, then pressing the 'Open' button, produces the results displayed below.

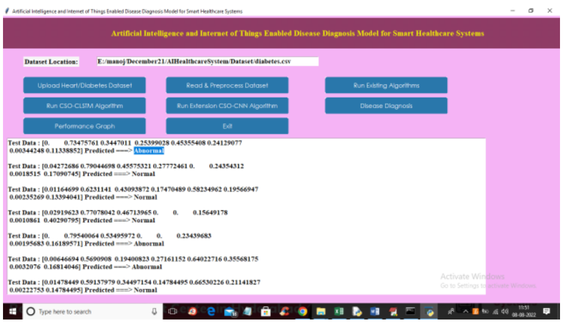


Figure14: TEST results

The TEST Data are marked by square brackets in the preceding picture, and the predicted output is labelled as Normal or Abnormal after the array symbol.

CONCLUSION:

In this work, an effective model for detecting sickness based on AI and IoT convergence was made for smart healthcare systems. The model shown has many steps, such as collecting data, preprocessing it, classifying it, and adjusting the parameters. IoT devices like wearables and sensors collect data that can be used by AI to find out if someone is sick. Then, the iForest method is used to remove outliers from the patient data. The CSO-CLSTM model is then used to sort the data based on whether or not the disease happens. CSO is also used to make the weights and bias parameters of the CLSTM model better. The CLSTM model is better at making diagnoses when CSO is used. Using data from the healthcare field, CSO-performance LSTM's was checked. During testing, the CSO-LSTM model got an accuracy of 81% at its best. Convolution 2D Neural Network (CNN2D), a more advanced version of LSTM, has been added to this project as an extension. Also, we added the CSO features selection technique to the CNN2D algorithm, and an experiment with Heart and Diabetes data showed that the extension algorithm has the highest accuracy of 98%.

Bibliography

[1].A. A. Abdellatif, A. Mohamed, C. F. Chiasserini, M. Tlili, and A. Abdellatif This study's authors are Abdellatif. Erbad, "Context-Aware Methodologies, Opportunities, and Challenges in Edge Computing for Smart Health," IEEE Network, vol. no.3, May2019, pp. 196-203

[2].A. M. K. Abd Ghani, N. Arunkumar, M. A. Mohammed, and O. Mutlag M. K. Abd Ghani, N. Arunkumar, and M. Mutlag Future Generation A. Mohammed Comput. Vol. of Syst. pp.62-78, Jan. 2019.

[3].A. M. A. Mohammed, M. S. Maashi, O. Mohd, S. A. Mostafa, K. H. Abdulkareem, G. Marques, and I. Mutlag Mutlag, M. K. A. Ghani, M. A. Mohammed, M. S. Maashi, O. Mohd, and M. A. Mohammed, MAFC: A multi-agent fog computing model for the administration of healthcare important activities. Sensors, 2nd ed. no.7, March2020, p. 1853.

[4].A. A. See Proc. Int. Conf. Computer, communication Autom. (ICCCA), Greater Noida, India, May2017, pages 1335-1340.

[5].F. F. Gong and X. Mod. Host Management, vol. no.2, pp.28-29, 2013.

[6].G. J. Bajo, J. De Paz, and J. Villarrubia "Monitoring and Detection Platform for Preventing Anomalies in Home Care," Corchado, Sensors, vol. no.6, June2014, pp. 9900-9921

[7].G. M. S. Hossain, Muhammad, and N. The study's authors are S. Iqbal. IEEE Journal of Special Topics Vol. no.2, Feb.2021, pp.603-610, doi:10.1109/ JSAC.2020.3020654.

[8].I. Azimi, Anzanpour, Rahmani, Pahikkala, Levorato, P. Liljeberg, and N. ACM Transactions on the Internet of Things Computing System, vol. October2017, Vol. No. 5s,p.

[9].J. T. Proc., P. Queralta 42nd Int. Conf. Telecommun. Pages 601-604 of Process Signals July 2019 (TSP).

[10].J. Ren, J. Guo, W. Qian, H. Yuan, X. Hao, and H. Zhu were the researchers. Jingjing, Secur,2018. "Building an Effective Intrusion Detection System Using Hybrid Data Optimization and Machine Learning Algorithms." Commun. Network, vol. pages. Jun. 2019.

[11].J. T. Proc., P. Queralta 42nd Int. Conf. Telecommun. Pages 601-604 of Process Signals July 2019 (TSP).

[12].K. H. Abdulkareem, M. A. Mohammed, A. Salim, M. Arif, O. Geman, D. Gupta, and M. Abdulkareem A. Mohammed and others all contributed to this research. Khanna, IEEE Internet Things J., early access, 11 January2021, doi:10.1109/JIOT.2021.3050775.

[13].M. G. and S. Hossain The report's authors are Zhu. "Emotion-aware connected healthcare large data for 5G," Muhammad, IEEE Internet Things Journal, vol. no.4, August2018, pp. 2399-2406.

[14].M. Y. Mengistu, H. Do, and W. Pham The Next Generation. Comput. Vol. of Syst. pp.129-140, Apr. 2018.

[15].M. G. and S. Hossain The report's authors are Zhu. Deep learning-based pathology diagnosis for smart-connected healthcare, Muhammad, IEEE Network, vol. no.6, November2020, pp. 120-125.

[16].M. Journal of the American Medical Informatics Association, vol. Z. Uddin, "A wearable sensor-based activity prediction system to enable edge computing in a smart healthcare system," vol. Distrib. in parallel Computer Science, vol. pp.46-53, Jan. 2019.

[17].N. "Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG inputs," Michielli, U. Comput., Molinari, Biol. Medical, vol. pp.71-81, Mar. 2019.

[18].O. K. and S. Alwan Lee is in charge of the design and execution of the "Dedicated real-time monitoring system for health care using ZigBee," Prahalad Rao, Healthcare Technol., 2012. Letters, vol. pp.142-144, August 2017.

[19].S. N., S. Ram, and B. Apduhan In the Proc. Int. Conf. Comput. Sci. Appl. Springer, Cham, Switzerland, pp.417-431, July 2019.