Future Generation of Communication and Internet of Things (FGCIOT)

Journal homepage: http://fgciot.semnaniau.ac.ir/

Research article

IoT-Based Disease Prediction and Diagnosis Systems

Sayyed Mustafa Sarkabeiri, Javad Safaei Kuchaksaraei

Department of Electrical and Electronic Engineering, Shahmirzad Branch, Islamic Azad University, Semnan, Iran

Article Info

Article History:
Received: September 10, 2023
Revised: October 28, 2023
Accepted: December 30, 2023

Keywords:
Internet of Things,
Machine Learning,
Artificial Intelligence,
ML

*Corresponding Author's Email Address:
mostafasarkabiri7@gmail.com

Abstract

Today, most people use the Internet for web searching, accessing multimedia services, and engaging in social networks, providing smart communication between machines and electronic devices. In fact, the goal of the Internet of Things (IoT) is to connect anything, anytime, anywhere, using any path or network and serving any purpose ideally. The application of IoT in the field of medicine reduces waiting times, tracks patients, doctors, equipment, and more. The aim of this paper is to investigate IoT-based disease prediction and diagnosis systems, artificial intelligence, and machine learning methods.
Methods and techniques such as Machine Learning (ML) on IoT data, healthcare datasets, model evaluation, and machine learning description are mentioned for disease prediction and diagnosis systems. Real-world machine learning models for healthcare applications are then discussed in this paper. Some successful applications of machine learning in disease diagnosis through IoT data are presented. Finally, future trends in machine learning for disease diagnosis, collaboration between Artificial Intelligence and the Internet of Things in disease diagnosis, were introduced.

Introduction

Today, most people use the Internet for web searches, accessing multimedia services, and engaging in social networks. We are currently witnessing a revolutionary event in information technology and communication, enabling smart communication between machines and electronic devices [1]. The concept of the Internet of Things (IoT) was first introduced in 1999 by Kevin Ashton. In this era, objects equipped with unique addresses or sensors can communicate with each other. Kevin Ashton viewed IoT as a tool to overcome limitations of time and space. This perspective was initially popularized by the Auto-ID Center [2] .These objects collaborate to create new services and achieve common goals. In essence, the aim of the Internet of Things is to establish connections at any time and place, with anything and anyone, utilizing any pathway or network, and providing an ideal service [3]. Many countries and healthcare organizations are facing challenges in providing adequate and suitable healthcare services. These challenges include the physical distance between doctors and patients, the complexity of communication between isolated medical systems, and the costs associated with healthcare equipment and infrastructure [4]. The applications of the Internet of Things in the field of medicine are expanding day by day, including but not limited to: "Reducing Wait Times in the Emergency Room Queue" Tracking patients, healthcare staff, and equipment improving drug supply and demand management ensuring the availability of necessary and essential equipment [5].

New advancements in disease prediction and diagnosis make use of data from Electronic Health Records (EHR), Internet of Things (IoT) sensor devices, wearables, and social media. In modern systems, Artificial Intelligence (AI) is utilized for improving disease prediction and diagnosis, particularly through machine learning, which is used to develop analytical models. These models collect data from the Internet of Things to learn and identify patterns and disease conditions. Artificial intelligence is reshaping our modern life, and in the field of healthcare, it encompasses both virtual and physical aspects [10]. The physical branch includes surgery and robotics, which can assist in surgical procedures and rehabilitation. The virtual branch encompasses informatics, which is expected to aid physicians in clinical diagnosis and treatment decisions [11]. Before delving into IoT-based disease prediction and diagnosis, understanding the relationship between the Internet of Things, artificial intelligence, and machine learning is crucial [12]

A. Artificial Intelligence

The term "ificial Intelligence" emerged in 1956 by John McCarthy, who is also known as its father Artificial Intelligence means creating intelligent machines that can make decisions on their own. You might consider it as a science fiction fantasy, but with recent advancements in technology and computing power, this idea is becoming closer to reality day by day. Fuzzy logic, neural networks, machine learning, evolutionary computing, pattern recognition, and various artificial intelligence algorithms and methods are utilized as solutions to enable computers to make decisions and facilitate learning [12].

B. Machine Learning

Essentially, machine learning refers to a form of artificial intelligence that endows computers with the ability to learn without explicit programming by exposing them to a large volume of data. The two primary challenges in machine learning are as follows:

1- Traditional ML algorithms are not useful when working with high dimensions. This is where we have a large number of inputs and outputs. For instance, in disease detection, we have a significant amount of inputs involving various types of IoT devices with different diseases.

2 -The second main challenge is to tell the computer what features it should have, as this will play a crucial role in predicting the outcome [12].

C. Deep Learning

Deep learning is a multi-layered machine learning algorithm inspired by the neural networks of the human brain. Similar to the deep neural networks found in our brains, deep learning architecture plays a crucial role in processing and completing processed information in the field of artificial intelligence. Deep Learning is a subset of machine learning; various techniques employed in deep learning involve incorporating many scalable layers when producing results, and with the addition of more data, they exhibit improved performance [14].

Singh and colleagues (2020) state that in recent years, the development of the Internet of Things along with interconnected physical objects and their virtual representations has seen significant growth. This has led to the creation of a wide range of potentially new products and services in various domains. One of the industries that has extensively benefited from IoT technology applications is the healthcare sector [6]

Jiang and colleagues [8] have proposed the use of a device in the Internet of Things environment to remotely detect an individual's stress level by reading their heart rate. This is achieved by using a heart rate sensor to measure the person's heartbeats. By mapping stress and heart rate, many aspects can be identified, such as whether the individual is nervous, anxious, or scared. However, traditional measurements only had limited predictive capability.

Ebrahimzadeh and colleagues [7] conducted a study on an approach for predicting Sudden Cardiac Death (SCD) by analyzing ECG signals and examining HRV features. They developed appropriate approaches to assist physicians in accurately predicting Sudden Cardiac Death by developing time-frequency (TF) and nonlinear features of HRV signals from ECG signals. To distinguish normal ECG from a person without any heart issues and a person at risk of Sudden Cardiac Death (SCD), a Multi-Layer Perceptron (MLP) neural network and k-Nearest Neighbors (k-NN) classification have been utilized. These features are extracted from one-minute HRV intervals and compared to one-minute normal HRVs. The challenge with this approach is that there is no noticeable difference between normal ECG and those patients susceptible to SCD, and SCD symptoms can be observed even up to 4 minutes before SCD occurs. In other words, even experts in cardiology and electrocardiography cannot differentiate between normal ECG and patients prone to SCD. The proposed extracted features for predicting SCD can be used, highlighting that time intervals closer to SCD have a greater predictive capability. The goal of this paper is to investigate prediction and diagnosis systems for diseases based on the Internet of Things, artificial intelligence, and machine learning, among other approaches.

Methods and Approaches

A. ML Enhances IoT Performance

In the Internet of Things (IoT), machine learning delves into searching for specific relationships within the vast amounts of data collected by large companies, with the aim of making meaningful connections. By examining and analyzing the gathered data, patterns and similarities can be identified to assist us in making informed decisions. For example, wearable devices that contribute to health improvement have rapidly advanced in the industry. In the near future, this evolution will transform into an internet-connected healthcare service with the ability to update each individual's information and changes. Another advantage of the Internet of Things in healthcare is cost reduction. With IoT in healthcare, some patients will be able to stay at home while various IoT devices monitor and transmit all essential data.

B. Utilizing ML in IoT Data for Disease Detection Models

This section elaborates on the specifics of disease prediction and diagnosis models based on the machine learning process.

Figure 1-1 illustrates a conceptual model consisting of multiple modules, such as the dataset (both training and new data), feature extraction, machine learning algorithms, disease detection, and outcomes. It also describes the modeling process stages and subcomponents. Initially, defining the problem and assembling a required dataset is necessary. For instance, diabetes detection can be a problem, and the needed database should be defined as a process starting point.

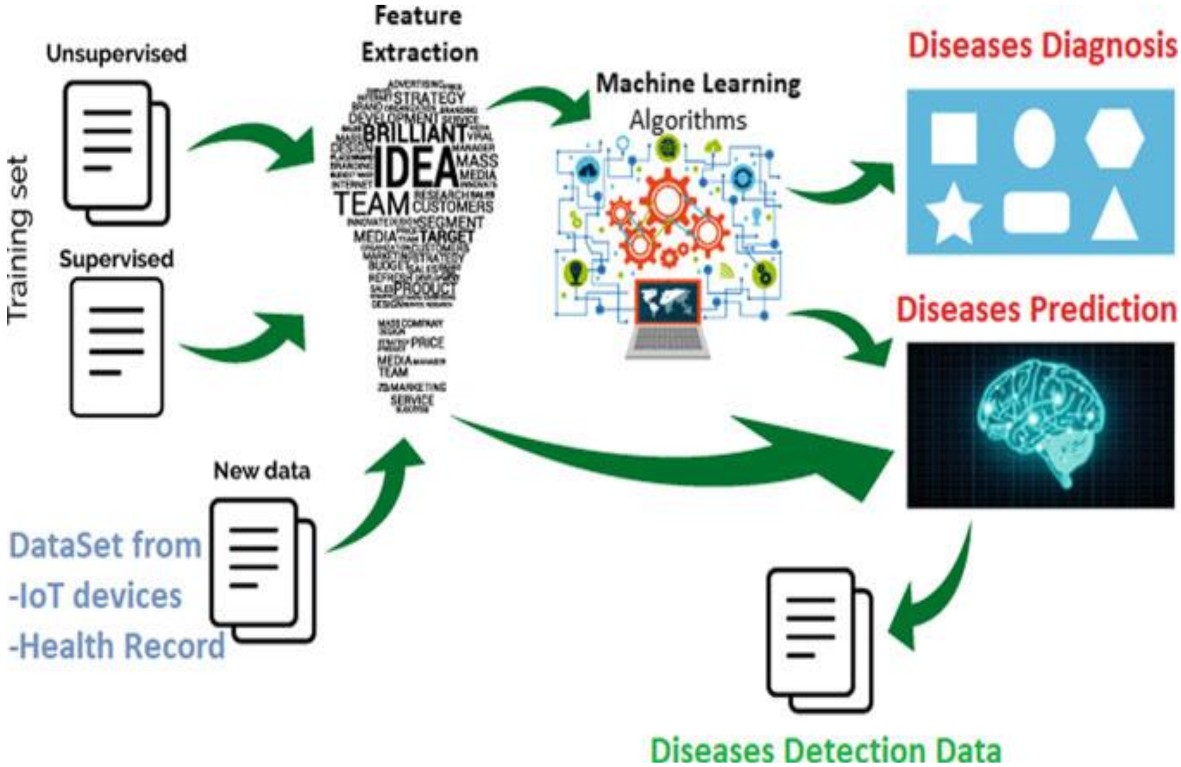

Fig. 1: The Disease Detection Process Model

C. Types of Datasets in Healthcare

The datasets used in this model are collected from IoT devices and health records. The dataset is currently divided into training and testing samples, which are suitable for machine learning models (both supervised and unsupervised). Quantity and quality of your data determine how accurate our model is. The result of this stage is usually obtained from displaying the data used for database analysis. Depending on the nature of the variables, measures such as mean and standard deviation are calculated, and extreme values for quantitative variables are explained. Absolute and relative frequencies for qualitative variables have identified the number of missing values. Data randomization, which removes specific effects, is carried out in the order that data is collected and/or prepared differently, facilitating data visualization through the identification of relevant relationships between variables or class imbalance, or conducting other exploratory analyses, and is finally divided into training and evaluation sets.

D. Feature Selection

The next step involves feature selection. Identifying key features from a dataset that can assist in constructing a model for disease detection is essential. Initial exploratory data analysis is performed, based on which a set of features is selected to assess the performance of various existing algorithms.

Furthermore, approaches are evaluated and compared to reach the optimal idea for disease detection from the heterogeneous dataset. [16]

E. Model Evaluation

In this stage, some evaluation data is used to measure the performance and test the model. The performance of models relies on certain metrics or a combination of metrics for "measurement." The concrete performance of the model is tested against previously unseen data. This step helps to assess the model's ability to generalize its predictions. However, real-world data still helps fine-tune the model (unlike the test data, which is not like that). Model hyperparameters might include the number of

training steps, learning rate, initial values, distribution, and more. Evaluation is often based on prediction accuracy.
There are several techniques used to calculate the model's accuracy. One of these techniques is called splitting, where the training dataset is divided with 66% for training and the remaining data for evaluation.In another technique known as cross-validation, the training dataset is divided into two exclusive and equal-sized subsets, and the model is trained on each subset. Thus, the average error rate of each subset is equal to an estimate of the model's error rate.

F. The selected model

The selected model is a representation of the relationships between variables used to describe the system. Different algorithms are used for various tasks. For prediction, additional data (test set), which has been concealed from the model up to this point, is utilized, and class labels are known. Classification is one of the most important aspects of supervised learning for disease detection, and various classification algorithms such as logistic regression, naive Bayes, decision trees, random forests, and many others can be employed as machine learning algorithms. Machine learning is a way to identify patterns in data and utilize them for automated prediction or decision-making.
In this section, some popular algorithms have been introduced as good selectable algorithms. In machine learning, decision tree algorithms are used for both prediction and classification. By using a decision tree with a set of inputs, one can map various outcomes that result from consequences or decisions. Random forests, unlike decision trees where each node is split based on the best feature, combine the outputs of multiple decision trees to improve predictive accuracy.
Random forests aim to minimize errors by randomly selecting features to make the best splits. However, the random selection of features for creating optimal splits reduces the correlation between the predictions of sub-trees, ultimately minimizing errors. Naive Bayes is a straightforward and fast method for predicting the class of a dataset. It is particularly useful for multi-class predictions. When the assumption of independence is valid, Naive Bayes outperforms similar algorithms. Using Naive Bayes, you can perform multi-class predictions. When the assumption of independence is valid, Naive Bayes outperforms similar algorithms. In addition to this, logistic regression requires fewer training data [16].

Applications and Results

In this section, some successful applications of machine learning in disease detection through Internet of Things data are presented.
The global population is aging, with one billion people aged 65 or older, many of whom are retired or no longer working. Therefore, the Internet of Things can significantly improve the quality of life for elderly individuals. For example, by using a small wearable device, vital signs of elderly individuals can be continuously monitored. If these vital signs go beyond the normal threshold or if the person falls and is unable to get up, the device can send an alert to a healthcare provider or medical center [9].

A. Diabetes and high blood pressure

Early disease prediction plays a significant role in improving healthcare quality. It can help individuals avoid risky health situations before it's too late. In research conducted, a Disease Prediction Model (DPM) was proposed to provide an initial prediction for Type 2 diabetes and high blood pressure based on the risk factors of an individual. The DPM proposal includes a method for outlier detection based on forest isolation to remove outlier data points. It also involves oversampling of the minority class to balance the data distribution and a group-based approach for disease prediction. Four datasets (monitoring device data, health records, personal information, and physician experiences) were used to build the model and extract the most significant risk factors. The machine learning algorithm is executed on the data, and the results show that the proposed DPM (Disease Prediction Model) achieves the highest accuracy compared to other models and previous studies. Furthermore, a mobile application has been developed to provide practical utility for the proposed DPM. The DPM mobile app collects data on risk factors and sends it to a remote server to assess an individual's current health status using the proposed DPM. Subsequently, the prediction results are sent to the mobile phone application. Therefore, for reducing and preventing individual health risks when unexpected health conditions occur, immediate and appropriate actions can be taken [17].

B. Kidney patients

A new predictive model for Diabetic Kidney Disease (DKD) has been developed using artificial intelligence, natural language processing, and large-scale machine learning on electronic medical records (EMR) from 64,059 diabetic patients. Artificial intelligence extracted features from the past 6 months as the reference period. "24 factors were selected to identify time-series patterns related to the aggravation of 6-month DKD (Diabetic Kidney Disease). Using an artificial intelligence convolutional autoencoder with 3,073 features, including time-series data, a predictive model was constructed through logistic regression analysis." "Artificial intelligence can predict the aggravation of DKD with an accuracy of 71 percent. The new predictive model by artificial intelligence can detect the progression of DKD, potentially aiding more effective and precise intervention to reduce hemodialysis" [18].

C. The diagnosis of cardiac arrhythmia

Cardiac arrhythmia is a threatening condition that can lead to severe health complications. Timely diagnosis of arrhythmia in patients is crucial for saving lives. The Internet of Things (IoT) ensures that healthcare can be revolutionized by providing continuous, remote, and non-invasive monitoring of cardiac arrhythmias.

An Internet of Things (IoT) platform for predicting cardiovascular diseases using ECG (Electrocardiogram) equipped with IoT capabilities receives ECG signal data, processes it, and sends emergency alerts to healthcare providers. Rapid analysis of heart disease is beneficial for physicians. This research has developed an ECG monitoring system with IoT capabilities for ECG signal analysis. Statistical features of raw ECG signals are calculated. These ECG signals are analyzed using the Pan-Tompkins QRS detection algorithm to extract dynamic features of the ECG signal. The system uses this information to identify RR intervals in the ECG signal for capturing heart rate variability features. Statistics and then dynamic features are applied in the classification process for heart classification. In cases of arrhythmia, individuals can monitor their heart status by acquiring ECG signals even from the comfort of their homes. The system is compact in size and requires minimal maintenance and operational cost [15].

The characteristics of chronic wound tissue

The output of chronic wound images processed by an automated model has been evaluated and approved by a specialist physician. They must visually inspect the results and confirm them based on the percentage of tissue characteristics through machine learning-based automated observation. Validation and proper verification are necessary before using them in clinical trials to provide useful services to doctors and patients. Considering the observations of the physician, the machine learning process provides a more precise percentage of different tissue types when comparing these two methods." Here's the translation of the remaining part of the text into English*:* "The results can be easily monitored using a smartphone, and the outcome can be predicted more easily, accurately, and quickly. This assessment can provide an accurate evaluation of chronic wounds remotely to the treating physician for monitoring wound improvement [13]."

Conclusion

One of the essential aspects of life is healthcare. Health and medical treatment can be improved through disease prediction and diagnosis. Collaboration between humans and machines plays a vital role in healthcare. Artificial intelligence, with various innovative technologies, has a significant impact on the medical field. This technology helps nurses, doctors, and surgeons perform their tasks in a simpler way. AI technologies can provide personalized treatment recommendations for patients .Artificial intelligence can be used in medical diagnostic support systems, particularly for the identification of congenital heart diseases. This technology plays a significant role in the electronic storage of patient health records. Artificial intelligence leads to the improvement of accuracy, speed, and reliability in diagnosis. These technologies accurately predict disease outcomes and make use of the data overlooked by physicians. In large healthcare organizations, AI technologies are employed for proper healthcare system management by monitoring treatment costs, healthcare expenses, and treatment response .Additionally, machine learning approaches are used in many applications, including glaucoma detection, Alzheimer's disease diagnosis, bacterial sepsis diagnosis, ICU readmission, and pearl water detection. ANN, SVM, and deep learning models, especially CNN, are the most common algorithms in machine learning, as they have proven to have high evaluation performance in most cases [19]. Algorithms, techniques, and machine learning tools are currently available in the market for reasonable execution of disease detection and prediction processes. Therefore, this technology is sometimes described as supervised or predictive machine learning. In general, most articles in the field of disease diagnosis and treatment indicate that technologies like the Internet of Things and artificial intelligence provide clear benefits to patients and physicians. The growth of the Internet of Things is undoubtedly beneficial, but the key question is how much of the data collected by IoT devices is actually useful. To answer this question, efficient technology and software for data analysis must be utilized. Machine learning and the Internet of Things should work together to strive towards creating better technology that ensures efficiency and productivity for all sectors of healthcare and treatment.

Discussion and Conclusion

Today, healthcare professionals can identify and treat a much wider range of diseases than in the past. However, even after years of practice, they can still strive for more accurate diagnoses. This is where technologies such as the Internet of Things and artificial intelligence can play a crucial role in providing reliable support for determining a diagnosis and the best treatment path. Artificial intelligence technologies like machine learning can quickly analyze a vast amount of available data, assist in the diagnostic process, and help prevent errors by integrating historical data and analyzing specific patient information.

The Internet of Things doesn't stand alone in the healthcare industry. All IoT devices and networks must be integrated with other technologies to truly enhance healthcare capabilities. As mentioned earlier, the Internet of Things is set to transform the healthcare industry, but it also requires data, high-speed and secure communications, and proper adaptation. It is expected that machine learning in the Internet of Things for healthcare will continue with more and better model designs. Researchers anticipate that artificial intelligence applications will see more from IoT data by assisting healthcare professionals as their virtual medical assistants. Researchers also expect further research into other robots for the detection of more diseases and patient treatment recommendations. In conclusion, the integration of artificial intelligence and the Internet of Things in healthcare is poised to revolutionize the industry. These technologies provide valuable tools for healthcare professionals, enabling them to make more accurate diagnoses, create personalized treatment plans, and enhance overall patient care. As these technologies continue to advance, we can expect to see even more innovation and improvements in healthcare, ultimately benefiting both patients and healthcare providers.

References

[1] Singh RP, Javaid M, Haleem A, Suman R. Internet of things (IoT) applications to fight against COVID-19 pandemic. Diabetes & metabolic syndrome. 2020; 14(4): 521-4.

[2] Smart Generation 2019 (pp. 27-51). Springer, Cham.Sharma N, Shamkuwar M, Singh I. The history, present and future with IoT. InInternet of Things and Big Data Analytics for

[3] Falehkerry Z. Research Title: Analysis of IoT Challenges in the Field of Health (Case Study: Using Technology to Optimize Devices Used by the Disabled). Journal of New Research Approaches in Management and Accounting. 2021; 5(72): 137-48

[4] Lima-Camara TN. Emerging arboviruses and public health challenges in Brazil. Revista de saude publica. 2016; 50: 36

[5] Kelly JT, Campbell KL, Gong E, Scuffham P. The Internet of Things: Impact and Implications for Health Care Delivery. Journal of medical Internet research. 2020; 22(11): 20135

[6] Jurcut, A., Niculcea, T., Ranaweera, P., & Le-Khac, N. A. (2020). Security considerations for Internet of Things: A survey. SN Computer Science, 1, 1-19. https://doi.org/10.1007/s42979-020-00201-3.

[7] E. Ebrahimzadeh, M. Pooyan, and A. Bijar, "A novel approach to predict sudden cardiac death (SCD) using nonlinear and time-frequency analyses from HRV signals," PloS one, Vol. 9, No. 2, p. e81896, 2014.

[8] L. Zhao, Y. Zhou, H. Lu, and H. Fujita, "Parallel computing method of deep belief networks and its application to traffic flow prediction," KnowledgeBased Systems, Vol. 163, pp. 972-987, 2019

[9] "The Internet of Things: How the Next Evolution of the Internet Is Changing Everything" (PDF). Dave Evans, (April 2011). Cisco. Retrieved 4 September 2015.

[10] Hamet, P.,&Tremblay, J. (2017).Artificial intelligence in medicine. Metabolism, 69, S36–S40.

[11] Urban, G. (2018). Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. Gastroenterology, 155, 1069–1078.e8.

[12] Vanani, I. R., & Amirhosseini, M. (2019). Deep learning for opinion mining. Extracting knowledge from opinion mining (pp. 40–65). Hershey, PA: IGI Global.

[13] Chinmay, C. (2019). Computational approach for chronic wound tissue characterization. Informatics in Medicine Unlocked, 17, 1–10.

[14] Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. Foundations and Trends in Signal Processing, 7(3–4), 197–387. https://doi.org/10.1561/2000000039.

[15] Devi, R. L., & Kalaivani, V. (2019). Machine learning and IoT-based cardiac arrhythmia diagnosis using statistical and dynamic features of ECG. Journal of Supercomputing. https://doi.org/10.1007/s11227-019-02873-y.

[16] Domingos, P. (2012). A few useful things to know about machine learning. Communications of the ACM. https://doi.org/10.1145/2347736.2347755.

[17] Fitriyani, N. L. (2019).Development ofDPMbased on ensemble learning approach for diabetes and hypertension. IEEE Access. Special section on data-enabled intelligence for digital health. https://doi.org/10.1109/ACCESS.2019.2945129.

[18] Mitchell, T. (1997). Machine learning (p. 2). New York, NY: McGraw Hill. ISBN: 978-0-07- 042807-2.

[19] Rayan, Z., Alfonse, M., & Salem, A.-B. M. (2019). Machine learning approaches in smart health. In 8th International Congress of Information and Communication Technology, ICICT 2019. Elsevier.