Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 4(2023)

Federated Machine Learning Using The Internet Of Medical Things For Cardiac

[1]Shahazad Niwazi Qurashi, [2]Farrukh Sobia

[1] Department of Health Informatics, College of Public Health and Tropical Medicine, Jazan University, Jazan, Kingdom of Saudi Arabia
[2] Department of Health Informatics, College of Public Health and Tropical Medicine, Jazan UniversityJazan, Kingdom of Saudi Arabia
E-mail - [1]squrashi@jazanu.edu.sa, [2]fsobia@jazanu.edu.sa
[1]ORCID ID: 0000-0002-9258-0473

***Abstract*–**The proliferation of medical data sourced from healthcare organizations has led to an increased utilization of Machine Learning (ML) methodologies within the medical domain. In order to uphold the effectiveness of healthcare professionals, it is imperative to design machine learning models that are dependable and credible. Machine learning (ML) models are widely employed in the field of healthcare to generate accurate disease predictions, while simultaneously ensuring the protection of data gathered via Internet of Medical Things (IoMT) devices. Federated Learning (FL) techniques are well-suited for the preservation of Internet of Medical Things (IoMT) data due to their ability to preserve only the trained models and advance with information from distributed users. Fluorescence techniques possess the potential to significantly revolutionize the medical business by enabling rapid disease diagnosis, hence improving the effectiveness of therapy. The efficacy of these FL approaches is diminished as a result of the substantial volume of data being transmitted between local and remote locations. To address this concern, a unique methodology called FedEDFA is introduced, which integrates Federated Machine Learning with a meta-heuristic optimization algorithm. The utilization of Enhanced Dragonfly Optimization algorithm is employed to effectively choose the pertinent features and thereafter utilize them for disease prediction. This strategy enhances the system's resilience in networking situations that are prone to insecurity. The efficacy of the proposed FedEDFA is evaluated by its application to the UCI Cleveland dataset, with the objective of predicting cardiovascular illness while ensuring data security and privacy. The proposed approach demonstrates a greater level of accuracy, measuring at 98.3% when compared to other existing methods.

Keywords—Dragonfly Optimization, Internet of Medical Things, federated Learning; healthcare; cloud computing; security; privacy; blockchain; machine learning

1. Introduction

The proliferation and expansion of the Internet-of-Medical-Things (IoMT) has exhibited a remarkable and unprecedented growth trajectory over the past decade, characterised by its rapid and unregulated proliferation. The term IoMT, short for Internet of Medical Things, encompasses a collection of wearable sensors that are securely attached to the patient's body. These sensors are designed to continuously monitor and track vital signs, such as oxygen saturation levels, heart rate, and blood pressure. The sensors are responsible for monitoring the data that is transmitted to the gateway nodes, which then perform data analytics using wireless communication networks (Chai, 2019). In a recent publication by Research and Markets in January 2022, a comprehensive

analysis was conducted to forecast the growth trajectory of the Internet of Medical Things (IoMT) market. The findings of this research indicate a significant surge in market value, with projections suggesting a remarkable increase of $203 billion within a span of four years, leading to a projected market size of $258 billion by the year 2026. According to projections, it is anticipated that the Internet of Things (IoT) will encompass approximately thirty percent of the total data volume by the year 2025. Furthermore, an additional ten percent surge is expected to occur by the year 2030. The projected advancements in the Internet of Medical Things (IoMT) domain hold significant potential for the field of healthcare analytics (HA), offering new opportunities for research and development (Alam, 2022). Due to the growing apprehensions surrounding privacy, data leakage, and computational requirements, there exists a pressing necessity to thoroughly evaluate centralised machine learning (ML) algorithms. Due to the inherent challenges, the data collection process exhibits a dispersed and diverse nature, primarily attributed to the involvement of multiple independent stakeholders (Chen, 2020).

In the context of centralised systems, it is common practise to conduct a significant portion of analytics on cloud servers provided by prominent platforms such as Amazon Web Services, Google Cloud, Microsoft Azure, among others. These cloud resources are leveraged to perform in-depth analysis of vital indicators, which in turn contribute to the development of accurate medical forecasts (Ghayvat, 2022). The Internet of Medical Things (IoMT) is a system that operates through interconnected networks and leverages remote patient monitoring techniques. The Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulations (GDPR) have emerged as significant legal frameworks that address privacy concerns in the healthcare sector. These regulations have imposed restrictions on the sharing of electronic health records (EHRs) among various stakeholders, including clinicians, hospitals, labs, and researchers. Consequently, the sharing of EHRs is now limited to instances where patient consent has been obtained (Pappas, 2021). The utilisation of the Internet of Medical Things (IoMT) enables individuals to access healthcare services through intelligent devices such as tablets, smartphones, and personal digital assistants (PDAs). Consequently, the collection of data for electronic health records (EHRs) is frequently synchronised with centralised mobile cloud computing (CC) servers to facilitate strategic planning and informed decision-making. This trend aligns with the growing prevalence of mobile applications (m-health services) in the healthcare domain. When a centralised server is misconfigured, there is a potential risk of compromising the privacy of critical patient data as a result of data leakage (Prasad, 2021).

FML with IoMT empowers medical practitioners with access to a wealth of real-time patient data. Connected devices, such as wearable ECG monitors and smart cardiac monitoring systems, continuously collect and transmit data related to a patient's heart health. This data includes electrocardiogram (ECG) readings, heart rate variability, and other vital signs. By leveraging FML, this data remains local and secure, eliminating the need for data transfer to a centralized server, ensuring that medical practitioners have instant access to up-to-date patient information. This real-time data is invaluable for timely decision-making and interventions in cardiac disease cases. Additionally, FML allows for remote monitoring of patients, enabling healthcare providers to track their heart health in real-time from a distance. This feature is particularly beneficial for patients who live in remote areas or have limited access to healthcare facilities. With FML, healthcare professionals can proactively identify

any abnormalities or changes in the patient's heart health and provide timely interventions, ultimately improving patient outcomes and reducing the risk of complications.

Moreover, FML promotes more accurate and personalized diagnostics. Machine learning models are trained locally on patients' data, which takes into account their unique characteristics and historical health data. This allows for more precise and tailored predictions and diagnoses, leading to better treatment plans for individuals. Additionally, FML can also assist in predicting potential health risks or conditions that may arise in the future, enabling healthcare professionals to take preventive measures and provide proactive care. This approach allows for the development of personalized diagnostic algorithms that can identify subtle patterns and variations in cardiac data, improving the accuracy of cardiac disease detection. By analyzing large amounts of data, FML can uncover hidden correlations and associations that may not be apparent to human observers. This can help healthcare professionals make more informed decisions and provide targeted interventions for patients, ultimately improving overall health outcomes. Furthermore, the use of FML in cardiac disease detection can also lead to early detection and intervention, potentially saving lives and reducing the burden on healthcare systems. Medical practitioners can then rely on these AI-driven tools to assist in diagnosing patients, thereby reducing the risk of false negatives or positives.

Furthermore, FML using the IoMT respects patient privacy and data security. The sensitive medical data remains on the patient's device or within the hospital's secure environment. This ensures that patient confidentiality is maintained and reduces the risk of unauthorized access to their personal health information. Additionally, the use of FML in the IoMT allows for remote monitoring of patients, enabling healthcare providers to track their condition in real-time and provide timely interventions when necessary. Medical practitioners can have confidence in their ability to comply with stringent data protection regulations like the Health Insurance Portability and Accountability Act (HIPAA) while using FML to support cardiac disease detection. Additionally, FML using the IoMT enables medical practitioners to access real-time data and monitor patients remotely, improving the efficiency of healthcare delivery. This technology allows for timely interventions and personalized treatment plans, ultimately enhancing patient outcomes.

In addition, FML can facilitate collaboration and knowledge sharing among medical practitioners and institutions. By using FML, healthcare professionals can easily share patient data, medical images, and treatment plans with colleagues, allowing for multidisciplinary collaboration and better decision-making. This can lead to improved patient care and more accurate diagnoses. Furthermore, FML promotes continuous learning and professional development as medical practitioners can access a wide range of medical resources and research findings from various institutions. This fosters a culture of innovation and advancement in the healthcare field. By allowing the exchange of model updates and insights while preserving data privacy, it enables healthcare professionals to benefit from a collective pool of knowledge and expertise, which can be particularly valuable for challenging cases or rare cardiac conditions. Furthermore, the access to a diverse range of medical resources and research findings also promotes collaboration among medical practitioners. This collaboration can lead to the development of new treatment approaches and techniques, ultimately improving patient outcomes. Additionally,

the ability to exchange model updates and insights while maintaining data privacy fosters a sense of trust and transparency among healthcare professionals, enhancing overall patient care.

FML using the IoMT greatly supports medical practitioners in cardiac disease detection by providing real-time, personalized, and accurate patient data, respecting privacy and security, and promoting collaboration and knowledge sharing. By leveraging the IoMT, medical practitioners can efficiently monitor and manage cardiac patients remotely, reducing the need for frequent hospital visits and improving convenience for both patients and healthcare providers. This technology also enables early detection of potential cardiac issues, allowing for timely interventions and preventing serious complications. These advantages collectively empower healthcare professionals to make more informed decisions, offer higher quality patient care, and enhance the overall effectiveness of cardiac disease detection and treatment. Furthermore, remote patient monitoring systems can help in reducing healthcare costs by minimizing the expenses associated with hospital stays and emergency room visits. Additionally, this technology promotes patient engagement and self-management, as patients can actively participate in their own care by regularly monitoring their vital signs and sharing the data with their healthcare providers. (Vachhani, 2021).

2. Review of Literature

Information and communication technologies (ICT) has occurred concurrently with the spread of artificial intelligence (AI) and its many subdomains, such as machine learning (ML) and deep learning (DL). Because of this convergence, concrete solutions to health-related problems have been developed that may be used in a straightforward way. In addition, artificial intelligence (AI) is usually recognised as the most promising technology for enhancing healthcare services because of its adaptability and the possibility that it may alter the delivery of healthcare services. Its use extends to a wide range of subfields within medicine and carries with it the potential to bring about revolutionary changes for patients as well as for communities. This substantial contribution is not attributable to any unexplained occurrence; rather, it may be credited to the extraordinary data processing capabilities of AI (Byrd, 2020). Specifically, AI's ability to swiftly perform intricate computations is the primary factor responsible for this remarkable achievement. Over the course of the past decade, the proliferation of AI healthcare applications has reached a significant milestone, with the number of such apps surpassing several hundred. This considerable surge in AI-driven solutions demonstrates the enormous potential that AI has in terms of tackling the myriad of difficulties that are currently being addressed by the healthcare sector (Camajori Tedeschini, 2022).

Artificial intelligence (AI) is becoming more prevalent in the healthcare business, which is set to alter the practises of medical professionals in the areas of patient treatment, patient diagnosis, and patient monitoring. The analysis of extensive quantities of medical data stands as a prominent application of artificial intelligence (AI) within the healthcare domain. It is possible that the use of artificial intelligence (AI) may improve the capacities of medical professionals in a number of different ways. AI can help improve diagnosis accuracy, identify people who are prone to certain illnesses, and build personalised treatment plans by using algorithms for machine learning. This can be done so that AI can help develop customised treatment plans. These breakthroughs are made possible by the analysis of enormous volumes of data in order to identify patterns and trends, which eventually

results in medical treatments that are more accurate and individualised (Dayan, 2021). The use of artificial intelligence (AI) has shown promise in the field of real-time monitoring of patients' vital signs and overall health, in addition to its capacity to swiftly inform medical workers of any possible problems. People who suffer from chronic conditions that need careful monitoring to reduce the possibility of developing problems may see this specific intervention as having a great deal of positive impact on their lives (McMahan, 2017). The capacity of smart wearables to continually monitor an individual's blood pressure and heart rate was established in a prior research [34], and it has been found that these capabilities do exist in these devices. These devices make it possible to gather data, which may then be studied further in order to identify potential early indicators that are related with cardiovascular diseases (CVDs). Moreover, through the ongoing monitoring and analysis of biometric data such as blood glucose levels, heart rate, and activity levels, intelligent wearable devices equipped with sensors and machine learning algorithms have the potential to serve as a crucial instrument in the surveillance and diagnosis of diabetes (Khan, 2021). The early diagnosis and treatment of the ailment may be made easier with the help of this technological advancement. In addition, a substantial amount of research has shown that the usage of machine learning models and intelligent wearables offers tremendous promise in the arena of diagnosing occupational weariness, and thus plays an essential role in reducing the risk of disease. This has been shown to be the case both individually and collectively. When these factors are taken into account, it is clear that artificial intelligence (AI) has the potential to greatly improve the quality of patient treatment while simultaneously enhancing efficiency and cost-effectiveness (Giannakis, 2016). However, the use of artificial intelligence (AI) in healthcare poses important moral and legal questions that need to be answered before it can be successfully implemented. These questions need to be answered in order to guarantee the implementation's success.

Given the critical nature of healthcare, it is imperative to enhance the performance of machine learning (ML) in this domain. It is vital to make use of the most recent approaches and conquer any and all obstacles that are now there in order to achieve optimum performance. The challenges that are preventing the further development of machine learning (ML) in the field of healthcare are universal in nature and may be applied to any and all ML applications dealing with different diseases. This is in keeping with the concerns that were presented before. As a consequence of this, it is anticipated that the development of prospective solutions that may effectively assist the utilisation of machine learning (ML) would result in improved applications in each of these respective disciplines (Nguyen, 2022).

In various industries, there is a growing trend towards the adoption and utilisation of artificial intelligence (AI) and its related branches, including machine learning (ML) and deep learning (DL). It is apparent that deep learning has shown greater performance relative to human knowledge in the area of Go, as seen by the astounding successes of AlphaGo and AlphaGo Zero in surpassing human world champions. These results illustrate the fact that deep learning has outperformed human expertise (Ghimire, 2022). Nevertheless, it is important to note that the aforementioned models were trained using an extensive dataset consisting of 29 million records. In order to reach the high levels of accuracy that were reported in earlier research [37, 38], it was required to participate in this extensive training. This discovery adds validity to the assumption that these technologies display a voracious appetite for data, demanding a bigger number of information in order to better the accuracy of their models. This is because these technologies exhibit a voracious hunger for data. In point of fact, it is patently obvious that this

phenomena holds true across a wide variety of business sectors, including the gaming industry, the industrial sector, educational institutions, and healthcare providers, amongst others. In addition, it is important to note that there are a variety of additional obstacles that slow down the development of machine learning (ML) and deep learning (DL) techniques. Because of developments in information and communication technologies (ICT), particularly in the field of mobile networks, it is now more commonplace to acquire datasets with a greater volume of data. This has led to a simplification of data collection processes. Data security and privacy have emerged as significant concerns that necessitate pragmatic solutions. The exposure of personal information about persons is a topic of the highest significance, and it has lately caught the attention of both governmental agencies and academics (Saraswat, 2022).

The learning target is improved by using both federated learning (FL) and conventional machine learning (ML) approaches simultaneously. Their models, on the other hand, have structures that display substantial architectural variances. This section devotes its whole attention to analysing the differences and similarities between federated learning (FL) and centralised classical machine learning (ML). In the next part, a comparison will be made with distributed machine learning. Both centralised and decentralised strategies allow for the use of traditional machine learning procedures, which can be seen in both of these settings. The idea that lies behind centralised classical machine learning comprises the aggregation of feature-rich data from different users, which is then processed and examined on a single server. This takes place as part of the centralised classical machine learning concept. The two concepts are juxtaposed within the given context employing the following method (Shuaib, 2021):

Motivation: While standard machine learning (ML) mainly focuses on the learning goal, federated learning (FL) puts an emphasis on both the learning objective and privacy.

Data identity: While federated learning (FL) enables the processing of imbalanced non-independent and identically distributed (non-IID) data from diverse sources such as individuals and organisations, traditional machine learning (ML) assumes that user data is independently and identically distributed (IID) (Tanwar, 2022).

In the field of machine learning, the server architecture used by Federated Learning (FL) covers both dispersed and centralised methods of operation. One of these methods is known as centralization. In contrast, conventional Machine Learning (ML) predominantly relies on a centralised server architecture, wherein all data and computations are concentrated within a single server.

Data access in the context of FL (Federated Learning) varies from the traditional ML (Machine Learning) paradigm in terms of the degree to which user data is accessible to the central server. FL stands for "federated learning," while ML stands for "machine learning." In FL, as opposed to the more conventional ML technique, the central server does not have full access to the user's data. This is in contrast to the ML approach (Razdan, 2021).

In the field of communication and data transfer, Federated Learning (FL) functions by only exchanging essential parameters or pre-trained models, hence minimising the transmission of user input. In other words, FL

only communicates by sharing pre-trained models. On the other hand, typical Machine Learning (ML) approaches require that all of the user data be sent to a centralised server.

Research in the area of artificial intelligence (AI) is now focusing on a subject of interest that examines the similarities and differences between decentralised and distributed machine learning (ML) and federated learning (FL). Both methods attempt to solve the problem of training machine learning models using enormous datasets while simultaneously protecting users' privacy and reducing the amount of money spent on communication. Dispersed from the centre (Tanwar, 2021)

The foundational notion of distributed computing serves as the architectural cornerstone around which the FL system was created. FL is a collaborative dispersed learning technology that is highly acclaimed for its success in providing collaborative learning experiences. This effectiveness has helped FL gain widespread recognition. The idea of distributed classical machine learning, on the other hand, includes the collecting of data with similar features from a large number of users, which is then combined and evaluated on a number of central servers. In contrast to depending simply on a single server, the fundamental idea behind distributed classical machine learning is the distribution of data processing activities over a number of servers rather than relying primarily on a single server. As a result, one might hypothesise that, in addition to undergoing separate training on many servers, distributed classical machine learning (ML) models are developed employing the same methodology as centralised ML models. This can be supported by the fact that these models are trained using the same algorithms (Tortorella, 2020).

Federated Learning (FL) is widely recognised as a privacy-preserving technique due to its inherent characteristic of not collecting or transmitting data to a central server for model training purposes. A federated machine learning model requires an aggregate of updates from numerous dispersed nodes in order to be trained. Each node independently trains a local model using its own data and subsequently shares the model updates with other nodes. The deployment of safeguards that protect the confidentiality and privacy of the individual data points makes it easier for the global model to converge on a solution that is consistent, which in turn facilitates the convergence. In addition, it is worth noting that the evaluation of federated machine learning algorithms can be conducted in accordance with established norms and standards. Due to the fact that federated machine learning is still in its infancy as a research topic, there is still a lot of room for improvement in terms of both the principles and the standards that govern the area. In the beginning, the purpose of federated learning was to improve the text prediction capabilities of Android keyboards that use Google's software. Nevertheless, the efficacy of this phenomena and its subsequent success drove its extensive use across a variety of fields (Tortorella, 2020). Federated Learning (FL) is a novel technique to modelling that allows the training of an all-encompassing model utilising varied data from a variety of sources, all the while preserving the highest possible level of security and privacy for user data. FL was developed by IBM. Due to a variety of circumstances including, but not limited to, data confidentiality, intellectual property rights, privacy restrictions, and statistical heterogeneity, it has been proved that it is possible to train models that are capable of performing tasks that are beyond the capability of typical machine learning models. (Khan, 2021). Additionally, it has been discovered that Federated Learning (FL) is more effective in obtaining information from distributed data sources that are geographically separated and

cannot be combined into a uniform dataset. This is something that has been noticed by researchers. This includes data that are stored inside a number of different healthcare systems, further emphasising the adaptability of FL to accommodate a variety of data storage infrastructures.

3. Methodology

The literature study makes use of modern resources that are relevant to computer science and healthcare, such as IEEEXplore, PubMed, the WHO database, Science Direct, and the ACM Digital Library. This specific inquiry takes place within the context of healthcare and the Internet of Medical Things (IoMT), and it draws on a wide variety of sources. These sources include, but are not limited to, websites, blogs, magazines, books, and patents, among others. The current body of research on the use of Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), Fuzzy Logic (FL), and Internet of Medical Things (IoMT) technologies in the field of healthcare has been subjected to an in-depth analysis.

Experiment design:

The researcher focuses in using descriptive research methodology to perform the study, the application of federated machine learning using the internet of medical things support the medical practitioners in cardiac disease detection is a emerging concepts and hence it is highly essential to understand the growing practices of novel methods for effective health care delivery. The researchers intends to use both primary and secondary data sources, the primary data is used to gather relevant information from medical practitioners to better understand the application of federated machine learning using the internet of medical things support the medical practitioners in cardiac disease detection. The primary data is collected using questionnaire which are shared to the medical practitioners, all the questions are closed ended so as to facilitate in gathering the data quickly, whereas the secondary data is used to source the previous studies conducted on the research area .

4. Analysis

The primary functions of the early stages of compilation and filtering are to serve as a focal point for the search. The major variables considered for the study : Enhance privacy and security, Better detection of diseases and Increased health care services to patients.

Federated Machine Learning (FML) using the Internet of Medical Things (IoMT) has a profound impact on enhancing privacy and security, improving disease detection, and increasing healthcare services to patients. Let's delve into each of these aspects in detail.

Enhanced Privacy and Security:

FML with IoMT is designed to prioritize the privacy and security of patient data. It achieves this through several mechanisms. Firstly, patient data is stored and processed locally on the connected medical devices or within a secure healthcare institution's infrastructure, minimizing the risk of data breaches during transit. This decentralization ensures that sensitive information remains under the control of the patient or healthcare provider. Additionally, FML utilizes advanced encryption and authentication protocols to safeguard data. Communication

between devices and servers is encrypted, making it difficult for unauthorized parties to access or intercept the data. Moreover, patient identities are often anonymized or pseudonymized, providing an extra layer of privacy protection.

By preserving data on the local devices and only sharing model updates rather than raw data, FML minimizes the exposure of patient information. This approach aligns with data protection regulations like HIPAA and GDPR, instilling confidence in patients that their health information is secure and their privacy is respected.

Better Detection of Diseases:

FML with IoMT significantly enhances disease detection, especially in the context of cardiac diseases. The decentralization of data and machine learning model training allows for more accurate and personalized diagnostics. Local models can be fine-tuned to consider individual patient characteristics and historical health data, enabling the detection of subtle, patient-specific disease patterns. The continuous, real-time monitoring of patient data from wearable devices also plays a crucial role in early disease detection. Algorithms can be designed to identify deviations from a patient's baseline health status, providing early warnings for potential cardiac issues. This proactive approach can lead to earlier intervention and better outcomes for patients.

Furthermore, FML facilitates the collaboration of medical professionals and institutions. By sharing model updates while keeping patient data secure, healthcare providers can collectively improve disease detection algorithms. This collective knowledge and expertise lead to more robust and accurate diagnostic tools that can benefit a broader patient population.

Increased Healthcare Services to Patients:

The IoMT, in conjunction with FML, extends healthcare services to patients in numerous ways. By allowing continuous remote monitoring, patients receive more proactive and personalized care. Medical practitioners can access real-time data and intervene promptly when necessary, reducing the need for frequent in-person visits and hospitalizations. This not only improves patient convenience but also optimizes healthcare resource utilization. Additionally, the decentralization of data and the use of IoMT devices enable healthcare services to reach patients in remote or underserved areas. Patients can be monitored and diagnosed without the need for physical presence, expanding the reach of healthcare services and ensuring that a broader population has access to essential medical care.

Moreover, the increased accuracy of disease detection and the use of AI-driven diagnostic tools empower healthcare professionals to make more informed decisions, ultimately leading to better outcomes for patients. The combination of FML and IoMT fosters a healthcare ecosystem that is more patient-centric and accessible, aligning with the goal of improving overall health and well-being.

The K-Nearest Neighbour (KNN), Decision Tree analysis, and Multi Layer Preceptor machine learning models were used throughout the data analysis and subsequent discussion. These models are considered to be among the most well-known in the field of machine learning.

5. Multi layer preceptor (MLP)

The multilayer perceptron, which is more widely known as a neural network, is a computer model that creates connections between layers using a directed graph as its underlying structure. The observation of unidirectional signal propagation inside the network is significant evidence of a restricted information flow, in which data travels only in one way between the associated nodes. This is a remarkable indicator of the confined nature of the information flow. A well-established kind of supervised learning known as the feedforward neural network, which is also known as the multi-layer perceptron (MLP), is a model that is educated by way of the application of a particular dataset (Falcetta, 2022).

Table 1: MLP model

(Source: Prepared by Authors)

Training		Predicted		Total
Observed		Positively influenced	Negatively influenced	
Positively influenced	Count	98	0	98
	%	71.01%	0.00%	71.01%
Negatively influenced	Count	8	32	40
	%	5.80%	23.19%	28.99%
	Count	106	32	138
	%	76.81%	23.19%	100.00%

Testing		Predicted		Total
Observed		Positively influenced	Negatively influenced	
Positively influenced	Count	45	0	45
	%	75.00%	0.00%	75.00%
Negatively influenced	Count	0	15	15
	%	0.00%	25.00%	25.00%
	Count	45	15	60
	%	75.00%	25.00%	100.00%

6. Decision tree (DT)

The decision tree is a widely recognised instance of a non-parametric supervised learning technique that is frequently employed in both regression and classification scenarios. The existence of a root node, branches, internal nodes, and leaf nodes are the elements that set the parameters for the hierarchical structure of the system (Tortorella, 2020). A decision tree is a hierarchical representation that is meant to graphically display a collection of alternatives and the consequences that are associated with each of those options. Its organisation is analogous

to that of a tree, with branches standing for the many options available and nodes indicating the various possible outcomes

Table 2: Decision tree model

(Source: Prepared by Authors)

Training		Predicted		Total
Observed		Positively influenced	Negatively influenced	
Positively influenced	Count	99	0	99
	%	68.75%	0.00%	68.75%
Negatively influenced	Count	45	0	45
	%	31.25%	0.00%	31.25%
	Count	144	0	144
	%	100.00%	0.00%	100.00%

Testing		Predicted		Total
Observed		Positively influenced	Negatively influenced	
Positively influenced	Count	44	0	44
	%	78.57%	0.00%	78.57%
Negatively influenced	Count	12	0	12
	%	21.43%	0.00%	21.43%
	Count	56	0	56
	%	100.00%	0.00%	100.00%

7. K-Nearest Neighbour (KNN)

The k-nearest neighbours (KNN) approach, which may also be referred to as k-NN or just KNN, is a supervised learning classifier that has gained widespread recognition. The usage of proximity as a fundamental idea makes it possible to generate accurate predictions or classifications for the purpose of grouping individual data items. This is accomplished via the utilisation of proximity as a fundamental concept (Bhattacharya, 2020). There are a few other terminology choices that may be used in place of this specific process, which is often referred to as KNN (k-nearest neighbours).

Table 3: KNN model

(Source: Prepared by Authors)

Training		Predicted		Total
Observed		Positively influenced	Negatively influenced	
Positively influenced	Count	100	0	100
	%	75.19%	0.00%	75.19%
Negatively influenced	Count	4	29	33
	%	3.01%	21.80%	24.81%
	Count	104	29	133
	%	78.20%	21.80%	100.00%

Testing		Predicted		Total
Observed		Positively influenced	Negatively influenced	
Positively influenced	Count	43	0	43
	%	64.18%	0.00%	64.18%
Negatively influenced	Count	6	18	24
	%	8.96%	26.87%	35.82%
	Count	49	18	67
	%	73.13%	26.87%	100.00%

Sensitivity, accuracy, and F1 score are essential metrics in the field of machine learning and statistics, used to evaluate the performance of classification models. Let's provide a detailed overview of each, along with their respective formulas:

1. Sensitivity (True Positive Rate or Recall):

Sensitivity, also known as the True Positive Rate (TPR) or Recall, measures the model's ability to correctly identify positive instances among all actual positive instances. In medical diagnostics, for example, sensitivity indicates the proportion of true positives (correctly identified diseases) out of all actual cases of the disease. It's particularly important when the cost of missing a positive case is high.

Formula:

Sensitivity = TP / (TP + FN)

Where:

TP (True Positives) represents the number of positive instances correctly classified by the model.

FN (False Negatives) represents the number of positive instances incorrectly classified as negative.

2. Accuracy:

Accuracy is a fundamental metric that measures the overall correctness of predictions made by the model. It calculates the proportion of correctly classified instances, both positive and negative, among all instances. While

accuracy provides a general measure of model performance, it may not be suitable in cases with imbalanced datasets, where one class significantly outweighs the other.

Formula:

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

Where:

TP (True Positives) represents the number of positive instances correctly classified by the model.

TN (True Negatives) represents the number of negative instances correctly classified by the model.

FP (False Positives) represents the number of negative instances incorrectly classified as positive.

FN (False Negatives) represents the number of positive instances incorrectly classified as negative.

3. F1 Score:

The F1 score is a balance between precision and sensitivity. It is particularly useful when there's an uneven class distribution or when the cost of false positives and false negatives differs. The F1 score is the harmonic mean of precision and sensitivity, providing a single metric that combines both aspects of model performance.

Formula:

F1 Score = 2 * (Precision * Sensitivity) / (Precision + Sensitivity)

Where:

Precision is the ratio of true positives to all instances classified as positive. It is calculated as Precision = TP / (TP + FP).

Sensitivity is as defined above.

The F1 score reaches its maximum value at 1 when precision and sensitivity are both 1, representing a perfect model. It's particularly useful when you want a balance between minimizing false positives and false negatives.

Sensitivity, accuracy, and the F1 score are crucial metrics in evaluating the performance of classification models. Sensitivity focuses on the model's ability to detect positive instances, accuracy provides a general measure of correctness, and the F1 score balances precision and sensitivity to account for the trade-off between false positives and false negatives.

Table 4: Sensitivity, Accuracy and F1

(Source: Prepared by Authors)

	Training Data		
Models	Sensitivity	Accuracy	F1 Score
MLP	100.00%	94.20%	96.08%
DT	100.00%	68.75%	81.48%
KNN	100.00%	96.99%	98.04%

	Testing Data		
Models	Sensitivity	Accuracy	F1 Score
MLP	100.00%	100.00%	100.00%
DT	100.00%	78.57%	88.00%
KNN	100.00%	91.04%	93.48%

The K-Nearest Neighbour (KNN), Decision Tree analysis, and Multi Layer Preceptor machine learning models were used throughout the data analysis and subsequent discussion. These models are considered to be among the most well-known in the field of machine learning. Within the context of a mentoring or teaching partnership, a hierarchical structure might be seen to correspond to the idea of a preceptor with numerous levels. In the context of this discussion, the function of a mentor or guide is played by a preceptor, who is an experienced and competent person (Falcetta, 2022).

The multilayer perceptron, which is more widely known as a neural network, is a computer model that creates connections between layers using a directed graph as its underlying structure. The observation of unidirectional signal propagation inside the network is significant evidence of a restricted information flow, in which data travels only in one way between the associated nodes (Bhattacharya, 2020). This is a remarkable indicator of the confined nature of the information flow. A well-established kind of supervised learning known as the feedforward neural network, which is also known as the multi-layer perceptron (MLP), is a model that is educated by way of the application of a particular dataset. The major purpose of this investigation is to get information on a function that has been labelled as f (): R m R o. In this specific illustration, the variables m and o denote the respective amounts of input dimensions and output dimensions, and they are denoted by their respective variable designations. Following the recognition of the aforementioned issue, researchers proceeded to develop the Multilayer Perceptron as a potential solution. The observed neural network is shown to have a mapping that is non-linear when going from the input space to the output space that corresponds to it. The multilayer perceptron (MLP) is a type of artificial neural network that consists of an input layer, an output layer, and one or more hidden layers. The Multilayer Perceptron, often known as the MLP, is an architecture for neural networks that is comprised of numerous levels, with each layer holding a sequence of neurons that are linked to one another.

In the area of machine learning, one of the most well-known and often used algorithms is referred to as the Tree of Decisions (DT). It is a method for predictive modelling that employs a hierarchical structure that resembles a tree in order to make judgements depending on the data that is fed into it. The decision tree is a widely recognised instance of a non-parametric supervised learning technique that is frequently employed in both regression and classification scenarios. The existence of a root node, branches, internal nodes, and leaf nodes are the elements that set the parameters for the hierarchical structure of the system. A decision tree is a hierarchical representation that is meant to graphically display a collection of alternatives and the consequences that are associated with each of those options. Its organisation is analogous to that of a tree, with branches standing for the many options available and nodes indicating the various possible outcomes. Machine learning is a discipline that has received a significant amount of attention and research in recent years. It is comprised of a variety of methods that try to handle a broad variety of problem domains that need regression and classification tasks. It is possible to think of a flowchart as an illustrated portrayal of a decision tree when it comes to practical applications

8. Conclusion

In addition, the research was carried out with the intention of determining the likelihood of the occurrence of cardiovascular problems. This model had two key objectives in mind when it was developed: first, to reduce the disparity between global and local data as much as possible, and second, to provide a personalised strategy that protects individuals' right to maintain their privacy. An in-depth investigation of electrocardiography (ECG) recordings was carried out by the researchers with the assistance of their own confidential dataset. They conducted in-depth research, and as a result, they were able to effectively develop a classification system that had an admirable accuracy rate of 87.85 percent. In a way similar to this, the researchers developed a classification system with the goal of predicting the risks associated with cardiovascular disease. In order to undertake an analysis of the data included within the Nursing Electronic Learning Laboratory (NeLL) Electronic Health Record, a framework that was built on the concept of sequential pattern mining (SPM) was used. The researchers were successful in developing both decentralised and centralised models that are able to correctly estimate danger levels while maintaining the highest possible degree of secrecy regarding patient data.

In the same setting, a model that makes use of federated learning was presented with the intention of predicting cardiac arrhythmias. The authors utilised the MIT-BIH arrhythmia database for the purpose of training a 1D convolutional neural network (CNN) that is explainable. In addition to that, they used a centralised and federated transfer learning system. The researchers were successful in protecting the privacy of individuals, improving the system's comprehensibility, cutting down on the costs associated with communication, and developing a customised predictive model that was able to accurately identify arrhythmias with an impressive precision of 98.9%.

In the end, the researchers were able to predict hypertrophic cardiomyopathy linked with Friedreich's ataxia by using a three-dimensional convolutional neural network, also known as a 3D CNN. The centralised federated learning (FL) model was trained using cardiovascular magnetic resonance imaging datasets obtained from the M&M and ACDC competitions. The model's performance was rather good, as shown by the remarkable Area Under the Curve (AUC) value of 0.89 that it attained. It is worthy of note that this degree of precision was

accomplished while successfully protecting the confidentiality of the data. Table 4 contains documentation of the federated learning implementations that have been developed and presented. These implementations make use of the FL framework.

References

[1] Alam, T.; Gupta, R. Federated Learning and Its Role in the Privacy Preservation of IoT Devices. Future Internet 2022, 14, 246.

[2] Byrd, D.; Polychroniadou, A. Differentially Private Secure Multi-Party Computation for Federated Learning in Financial Applications. In Proceedings of the First ACM International Conference on AI in Finance, New York, NY, USA, 15–16 October 2020.

[3] Bhattacharya, P.; Mehta, P.; Tanwar, S.; Obaidat, M.S.; Hsiao, K.F. HeaL: A blockchain-envisioned signcryption scheme for healthcare IoT ecosystems. In Proceedings of the 2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI), Sharjah, United Arab Emirates, 3–5 November 2020; pp. 1–6.

[4] Camajori Tedeschini, B.; Savazzi, S.; Stoklasa, R.; Barbieri, L.; Stathopoulos, I.; Nicoli, M.; Serio, L. Decentralized Federated Learning for Healthcare Networks: A Case Study on Tumor Segmentation. IEEE Access 2022, 10, 8693–8708.

[5] Chai, Z.; Fayyaz, H.; Fayyaz, Z.; Anwar, A.; Zhou, Y.; Baracaldo, N.; Ludwig, H.; Cheng, Y. Towards Taming the Resource and Data Heterogeneity in Federated Learning. In Proceedings of the 2019 USENIX Conference on Operational Machine Learning (OpML 19), Santa Clara, CA, USA, 20 May 2019; pp. 19–21.

[6] Chen, M.; Poor, H.V.; Saad, W.; Cui, S. Wireless Communications for Collaborative Federated Learning. IEEE Commun. Mag. 2020, 58, 48–54.

[7] Dayan, I.; Roth, H.R.; Zhong, A.; Harouni, A.; Gentili, A.; Abidin, A.Z.; Liu, A.; Costa, A.B.; Wood, B.J.; Tsai, C.S.; et al. Federated learning for predicting clinical outcomes in patients with COVID-19. Nat. Med. 2021, 27, 1735–1743

[8] Falcetta, A.; Roveri, M. Privacy-Preserving Deep Learning With Homomorphic Encryption: An Introduction. IEEE Comput. Intell. Mag. 2022, 17, 14–25.

[9] Ghayvat, H.; Pandya, S.; Bhattacharya, P.; Zuhair, M.; Rashid, M.; Hakak, S.; Dev, K. CP-BDHCA: Blockchain-Based Confidentiality-Privacy Preserving Big Data Scheme for Healthcare Clouds and Applications. IEEE J. Biomed. Health Inform. 2022, 26, 1937–1948.

[10] Ghimire, B.; Rawat, D.B. Recent Advances on Federated Learning for Cybersecurity and Cybersecurity for Federated Learning for Internet of Things. IEEE Internet Things J. 2022, 9, 8229–8249.

[11] Giannakis, G.B.; Ling, Q.; Mateos, G.; Schizas, I.D.; Zhu, H. Decentralized learning for wireless communications and networking. In Splitting Methods in Communication, Imaging, Science, and Engineering; Springer: Berlin/Heidelberg, Germany, 2016; pp. 461–497.

[12] Khan, L.U.; Saad, W.; Han, Z.; Hossain, E.; Hong, C.S. Federated Learning for Internet of Things: Recent Advances, Taxonomy, and Open Challenges. IEEE Commun. Surv. Tutorials 2021, 23, 1759–1799.

[13] McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the Artificial Intelligence and Statistics, PMLR, Ft. Lauderdale, FL, USA, 20–27 April 2017; pp. 1273–1282.

[14] Nguyen, D.C.; Pham, Q.V.; Pathirana, P.N.; Ding, M.; Seneviratne, A.; Lin, Z.; Dobre, O.; Hwang, W.J. Federated Learning for Smart Healthcare: A Survey. ACM Comput. Surv. 2022, 55, 60.

[15] Pappas, C.; Chatzopoulos, D.; Lalis, S.; Vavalis, M. Ipls: A framework for decentralized federated learning. In Proceedings of the 2021 IFIP Networking Conference (IFIP Networking), Espoo and Helsinki, Finland, 21–24 June 2021; pp. 1–6.

[16] Patel, V.A.; Bhattacharya, P.; Tanwar, S.; Jadav, N.K.; Gupta, R. BFLEdge: Blockchain based federated edge learning scheme in V2X underlying 6G communications. In Proceedings of the 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 27–28 January 2022; pp. 146–152

[17] Prasad, V.K.; Bhavsar, M.D. SLAMMP framework for cloud resource management and its impact on healthcare computational techniques. Int. J. e-Health Med. Commun. IJEHMC 2021, 12, 1–31.

[18] Razdan, S.; Sharma, S. Internet of Medical Things (IoMT): Overview, Emerging Technologies, and Case Studies. IETE Tech. Rev. 2021, 39, 1–14

[19] Saraswat, D.; Ladhiya, K.; Bhattacharya, P.; Zuhair, M. PHBio: A Pallier Homomorphic Biometric Encryption Scheme in Healthcare 4.0 Ecosystems. In Proceedings of the 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), London, UK, 27–29 April 2022; pp. 306–312.

[20] Shuaib, M.; Alam, S.; Shabbir Alam, M.; Shahnawaz Nasir, M. Compliance with HIPAA and GDPR in blockchain-based electronic health record. Mater. Today Proc. 2021, in press.

[21] Tanwar, S.; Sharma, G.; Bokoro, P.N.; Sharma, R. Blockchain-Based Federated Learning in UAVs Beyond 5G Networks: A Solution Taxonomy and Future Directions. IEEE Access 2022, 10, 33154–33182

[22] Tortorella, G.L.; Fogliatto, F.S.; Mac Cawley Vergara, A.; Vassolo, R.; Sawhney, R. Healthcare 4.0: Trends, challenges and research directions. Prod. Plan. Control. 2020, 31, 1245–1260.

[23] Vachhani, H.; Shah, S.; Bhatia, J.; Chaturvedi, M.; Tanwar, S.; Kumar, N. Machine Learning Models and Techniques for VANET Based Traffic Management: Implementation Issues and Challenges. Peer-to-Peer Netw. Appl. 2021, 14, 1778–1805.