International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V73I1P110

Volume 73 Issue 1, 80-90, January 2025
© 2025 Seventh Sense Research Group®

Original Article

Federated Edge Computing for Privacy-Preserving Analytics in Healthcare and IoT Systems

Ravi Kumar Vankayalapati[1], P.R. Sudha Rani[2], Shashikala Valiki[3], Venkata Krishna Azith Teja Ganti[4]

[1]*Cloud AI ML Engineer, Equinix Dallas USA.*
[2]*Department of CSE, SVECW, Bhimavaram, AP, India.*
[3]*Independent Research, USA.*
[4]*Sr Data Support Engineer, Microsoft Corporation, Charlotte NC, USA.*

[1]*Corresponding Author : ravikumar.vankayalapti.research@gmail.com*

Received: 24 November 2024 Revised: 30 December 2024 Accepted: 15 January 2025 Published: 31 January 2025

Abstract - *Federated Learning and Edge Computing, also known as Federated Edge Computing, have emerged as breakthrough technologies that enhance the privacy of patients and overall data privacy compliance. Federated Learning has been developed as a distributed learning approach to train machine learning models that exchange and learn only the model parameters from different databases. At the same time, the data remains local to the clients, hence avoiding the possibility of privacy erosion. The combination of Federated Learning and Edge Computing essentially extends the aggregation of learning parameters happening in Federated Learning from traditional edge nodes to increasingly underutilized resources at mobile phones, IoT, or other devices acting as smaller-scale edge nodes. A lot of work has been published to improve machine learning in a federated learning environment. Privacy, especially in the healthcare domain and in IoT, is a vital issue. More and more data is collected by IoT sensors, wearables, tablets, networks, compute resources, and even facilities. The collected data can range from energy management, visitor monitoring, equipment performance analytics, home/building security, surveillance, and even patient vitals, among others. There are many challenges in storing and processing this data. Given its sensitive nature, this data must be securely stored. They should possibly also be computed upon, creating insights or making decisions based on the data before importing it to less secure environments. In addition, compliance with privacy rules and regulations transmitted in healthcare communication acts must be met when storing and processing medical data/protected health information. Federated Learning provides a model set to instantiate and execute. In contrast, federated edge computing provides an instantiation on edge hardware to run the workloads and functionalities in a federated learning environment. With federated learning, all the processing of data, specifically medical data, will be kept within pockets of edge nodes themselves. Thus, this approach addresses safe and secure processing and the privacy of the user, patient, or community being observed.*

Keywords *- Federated Learning, Edge Computing, Federated Edge Computing, Data Privacy Compliance, Distributed Learning, Model Parameters Exchange, Local Data Processing, Privacy Erosion Prevention, Healthcare Privacy, IoT Sensors, Wearables, Mobile Edge Nodes, Data Security, Medical Data Processing, Protected Health Information, Privacy Regulations, Edge Node Aggregation, Secure Data Storage, Healthcare Communication Compliance, Patient Vitals Monitoring.*

1. Introduction

Merging technology, data analytics, and privacy have led to federated learning and edge computing to preserve patient privacy in healthcare. This focus has increased over recent years with new technologies that connect wearables, sensors, medical technology, and health informatics. To improve management, processing, and diagnostics in healthcare, these environments require analyzing sensitive data and patient information. Considering privacy requirements, these domains present significant data analytics challenges. This paper addresses knowledge gaps by utilizing edge computing and federated learning paradigms to provide a complete system overview, including system components, architecture, data flows, knowledge extraction, and results. We argue that a decentralized and privacy-preserving approach using edge computing and federated learning in the face of connected medical technologies will provide scalable and efficient solutions to the data analytics challenges for these IoT-enabled healthcare environments.

Health and biomedical informatics are critical components of data mining, knowledge management, and artificial intelligence. These health informatics industries, which seek to enhance and automate data processing and diagnostics and increase efficiency, face myriad issues, such as huge data volumes from connected medical and sensor devices that require real-time processing and analytics. Patient privacy and confidentiality by design are crucial requirements in the connected healthcare ecosystem, where medical data are analyzed with IoT and other intelligent systems. In general, IoT, particularly in the realm of health and medicine, is anticipated to generate over 30,000 petabytes of new healthcare data per year. Healthcare data are diverse and possess vagaries due to sensing devices, data entry errors, and complex patient-specific diagnosing constraints. Data privacy and security must be addressed in this setting. Data can be manipulated and used for identity theft and fraud.

1.1. Background and Motivation

The concept of edge computing initially stems from the industrial field and is currently becoming quite popular in the contexts of IoT systems and healthcare. In IoT, wearable devices and smart thermostats can collect personal data and communicate with powerful devices near users (at the edge) to reduce latency and enhance their capabilities. Smart hospitals and health management systems are also examples of systems that use edge computing nodes to carry out routine tasks in various environments. This architecture can provide several advantages, such as limited traffic, less centralized processing, and increased confidentiality. Recent progress in chip design and communication protocols, along with the growing attention of different application providers to the edge computing phenomenon and an increasing number of edge computing nodes installed in society, indicates that IoT and healthcare environments require edge computing to fulfill their requirements.

The sensitive nature of data, especially in healthcare systems, indicates that sharing information from devices with the cloud platform without applying data security mechanisms is not practical. Individual concerns about the use of personal data indicate an emerging trend toward developing software that can process IoT data without the need for access to raw private data collection, achieving significant privacy and security benefits. In addition to unauthorized access to private healthcare and IoT physical measurements, digital attacks or fraud are concerns in various activities. Authentication, encryption, and firewall mechanisms have been utilized in several studies to secure data connections between devices or parts of the system. However, these studies focused on the necessary hardware resources integrated into IoT devices required for security and preventing unauthorized physical access to private data collections. The use of edge computing devices as a secure intermediary to ensure the privacy of IoT and health data during data processing is the main consideration of the present study. We aim to conduct private data analysis in edge nodes by processing anonymous sensors, and to achieve this, two edge systems and a cloud system must be included in our architecture. This research is carried out as part of a larger international research project where a remote monitoring and advisory system for chronic obstructive pulmonary disease (COPD) patients utilizing smartphone and smartwatch technologies was developed. With technological trends to support preventive and personalized healthcare, the ambition of the system is to enhance the activity level of the patients and improve knowledge about the evolution of the health status of these elderly COPD patients. Such an approach aligns with organizations focusing on trends toward promoting health and trend data indicating continuous growth in health informatics research. With increasing computational power and emerging security threats in the digital transformation period, novel privacy-preserving data analysis methods for processing raw data will be of high interest to society. From an engineering point of view, the general considerations for privacy-preserving algorithms are the setup of both timing and computational requirements.

Local Model Update

$$\Delta \mathbf{w}_k = \nabla \mathcal{L}_k(\mathbf{w}_k; \mathbf{D}_k)$$

$\Delta \mathbf{w}_k$: Local model update for edge device k.

$\nabla \mathcal{L}_k$: Gradient of the loss function $\mathcal{L}_k$ on local data $\mathbf{D}_k$.

$\mathbf{w}_k$: Current model parameters.

1.2. Research Objectives

We seek to investigate the intersection between the concept of federated edge computing and privacy-preserving analytics and how it can be applied in healthcare and IoT systems. The objectives are as follows: 1. To evaluate the most recent state-of-the-art framework in federated systems as well as edge computing and IoT networking. 2. To investigate existing robust and privacy-preserving analytics frameworks proposed for various applications. 3. To propose a new framework that not only evaluates the system performance through comprehensive experiments but also takes into account data security as one of the main concerns. 4. To conduct performance analysis on edge devices and networks by employing machine learning models and finally present future research directions. The importance of edge computing lies in making decisions closer to where data is generated while having distributed networks share data, thereby potentially enhancing the final analytics. To protect user privacy, federated learning techniques are employed, and a robust learning framework is designed to ensure that authentic knowledge of data is used from end devices. The goal of this research is to develop scalable systems that can securely perform advanced analytics for IoT and healthcare applications, the two main focuses of this research, while ensuring privacy.

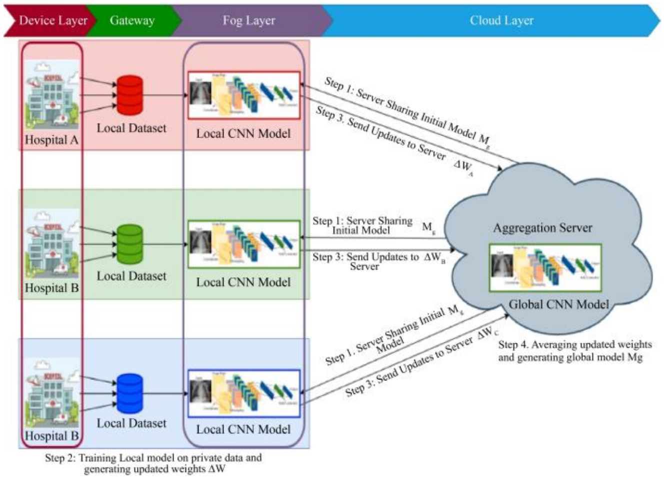

Device Layer
Gateway
Fog Layer
Cloud Layer
Hospital A
Local Dataset
Local CNN Model
Step 1: Server Sharing Initial Model M_g
Step 3. Send Updates to Server ΔW_A
Hospital B
Local Dataset
Local CNN Model
Step 1: Server Sharing Initial Model M_g
Step 3: Send Updates to Server ΔW_B
Aggregation Server
Global CNN Model
Hospital B
Local Dataset
Local CNN Model
Step 1. Server Sharing Initial Model M_g
Step 3: Send Updates to Server ΔW_C
Step 4. Averaging updated weights and generating global model Mg
Step 2: Training Local model on private data and generating updated weights ΔW

Fig. 1 A Fog-Based Privacy-Preserving federated learning system for smart healthcare applications

2. Edge Computing in Healthcare and IoT Systems

Edge Computing (EC), as compared to cloud computing, operates in a centralized data center located at far distances from the nodes that generate data, transform, and move data processing closer to the data sources. This holds high significance in transforming healthcare and Internet of Things (IoT) systems, usually data-rich domains. Low-latency solutions, such as fog and mist computing, share the computing responsibility with the cloud and do not perform the computations at the last hop for IoT devices. The decentralization of processing is one of the features that distinguishes EC from a centralized cloud computing architecture. An EC architecture enables data processing at lower aggregation levels and is more closely related to IoT network deployment. The concept of edge computing is characterized by the deployment of small data center-like facilities in the periphery of a network that is situated closer to digital devices and leverages embedded data centers. The data processing at the edge plays a critical role in the real-time data processing. Patients suffering from chronic diseases can benefit significantly from timely data processing at the edge. Importantly, slow data processing affects the timely distribution of insights through real-time or non-real-time health telemetry to primary care providers. Data analytics based on edge computing can exploit instantaneous large and small-scale data directly from health wearables and other sensory collectives, improving health telemetry. A data center at the edge can allow distributed processing of data, making EC capable of secure data storage and feasibly processing the data to generate knowledge in real-time. When it comes to IoT systems, EC supports advances related to reliability, longer battery life, scalability, mobile access, cloud merging, and ubiquity. Data-heavy domains can overcome the challenges of collaboration and optimizing IoT operations through edge computing models. Home health platforms are examples of IoT applications that can greatly benefit from edge computing.

2.1. Overview of Edge Computing

Edge computing brings computation and data storage closer to the sources of data generation. It is capable of providing reduced latency, enhanced speed, and improved reliability in data communication. Edge nodes or devices situated at the edge are principal components in the design of edge computing architectures. These edge computing principles have been in the industrial control systems domain, typically in the process control methodology, which aims to minimize latency in process control. The cloud computing infrastructure has been traditionally accessed by human-operated systems and applications through the internet. By allowing machines, devices, and applications to participate in computation directly at the edge, edge computing is beginning to enable a system that can automate the calling of microservices and resources as part of its application service set. These microservices can be distributed throughout the infrastructure, not necessarily running within any data center edge. A goal of edge computing is to minimize the delay of calling these microservices to sub-1ms using compute resources located anywhere in the infrastructure.

Collectively, technologies that address the requirements of edge computing are of vital importance. Their value is heightened because of a clear transition toward edge and hierarchical forms of cloud computing to address market demands. Tens of thousands of edge data centers are being deployed to meet infrastructure-as-a-service demands. These are often referred to as ‘middle price data centers’ to imply they serve markets that are somewhere between traditional enterprise and hyper-scale data centers. Healthcare is an industry poised for major transformation due to the adoption of edge computing on top of its existing cloud computing infrastructure. With broad and widespread use possible in sectors such as healthcare, further opportunities for improvements in operational efficiency can be realized through investments in edge computing ecosystems.

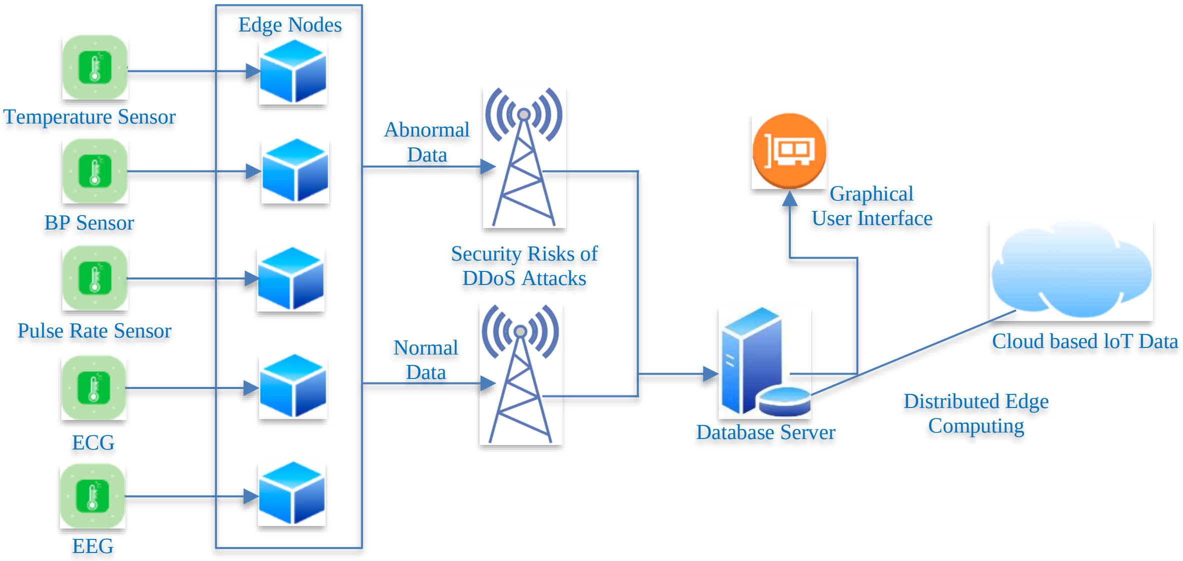


Fig. 2 Edge computing healthcare solutions

2.2. Applications in Healthcare and IoT

Advances in sensor technologies, data science, and telecommunications have led to a wide array of IoT-based applications to support patient care. IoT in healthcare is also commonly referred to as the Internet of Medical Things, whereby numerous types of medical devices and connected Point of Care sensors rely on Internet connectivity to establish communication necessary for data collection and sharing. Diagnostic data, patient-generated health data, and even daily biological signals can be continuously sensed and monitored with the diffusion of technology derivation, including lab-on-a-chip and wearable sensors. Typically, the acquired metrics are necessary for real-time analytics and for accurate subject-dependent modeling to assess systemic physiological responses or individual well-being states. The processed data often serve varying aims, such as data storage for safekeeping, providing information to clinical care team members, guiding a patient in self-management, identifying chronic disease sufferers, or informing the research community.

Given the heterogeneity of medical data generation with myriad application domains, modern medical infrastructure necessitates scalable and adaptable technologies to support a suitable volume of data transport and timely analytics. The ever-increasing number of connected IoT devices is anticipated to outpace population growth by 2030, with not only humans directly relying on connected devices but also autonomous devices, such as wearable exoskeletons, falling within this scope. As opposed to the traditional purview of IoT—connectivity between sensors and mainframes—the edge-based analytics paradigm aims to perform relevant data processing and decision generation using the computing, storage, and networking resources of edge assets, such as smart gateways and micro data centers, at the proximity of the data source located at the peripheries. By providing intelligence on or close to the data sources, the vast quantity of data collected from IoT devices can be judiciously analyzed at its source, saving valuable networking resources and time formerly expended to transport the raw information to the cloud for warehousing and, finally, analytics. Thus, healthcare IoT with the edge-based paradigm can offer cost-effective and timely access to medical treatments, such as telemedicine and connected medical devices. Additionally, three promising benefits of IoMT device-embedded edge-based computing are: (1) providing confidence to patients and practitioners with the refusal of diagnostic or treatment data to be warehoused elsewhere, (2) confidentially analyzing patient-individualized data to provide early medical pilot findings indicative of a chronic condition or overall physiological design, and (3) the effusion of excitatory pathway changes signifying comprehensive neuronal signaling. A systematic investment in IoMT devices operates at the cusp of smart-edge analytics.

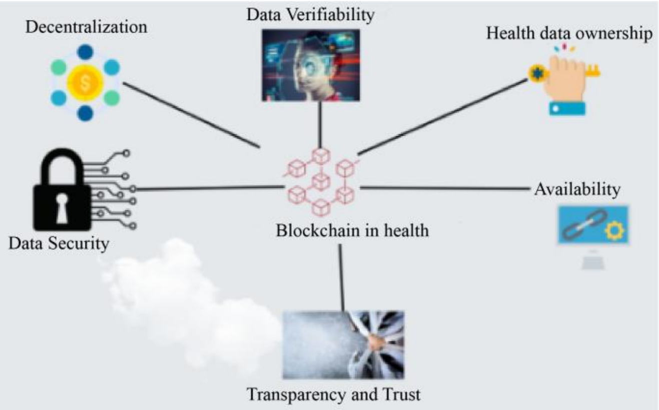


Fig. 3 Blockchain to secure internet of medical things

3. Privacy-Preserving Analytics

The idea behind privacy-preserving analytics is to provide analytics about the data without compromising the privacy of patients in healthcare or IoT systems. Doctors rely heavily on analyzing electronic health records data to provide appropriate treatments to the patient. Sharing and making available rich analytical findings based on EHR data helps not only doctors but also healthcare policymakers in many ways, such as patients' outcome prediction, presentation of evidence-based medicine, tailored treatment plans, disease outbreak monitoring and prediction, and reducing disease spread. In most of these cases, EHR data is sensitive and needs to be shared without revealing the patient's identity. However, handling sensitive data when mining it for useful information has several risks, including data aggregation, mining results, and re-identification based on them.

EHR data, in particular, is sensitive because any re-identification through the open data can lead to publicly known breaches. These breaches have different direct and indirect impacts on doctors, service providers, potential partners, and especially patients. One way to regulate the usage of sensitive data is to de-identify the attributes of the data. Traditional data analytics models, such as privacy-preserving data mining and differential privacy, nevertheless, mainly focus on privacy preservation in storing and sharing the data instead of privacy preservation during the analytics process. Close to our work is federated learning, which shares model updates between distributed data servers. It does not address lightweight IoT devices or patients in our context and pushes raw data to the cloud to be processed, which is impractical for privacy-sensitive applications.

3.1. Importance in Healthcare and IoT Systems

Privacy-preserving analytics constitute a critical aspect of healthcare and IoT systems. Health data is particularly sensitive, and unauthorized access to it can have serious implications for a myriad of individuals' lives. Privacy infringements can result in unwillingness to share personal or health data, limiting the potential for analytic-enabled therapies. Furthermore, data privacy issues arising in IoT systems can incur high costs and potentially pose direct safety risks, potentially disrupting the healthcare system and overall public health. In the specific healthcare case, it is not only the healthcare sector that may face such problems. Patients who either accidentally or purposefully disclose to unauthorized parties their health records stored locally on a gadget on their smartphone may exacerbate stigma against them, consequently being averse to possibly having been affected by a disease or being potentially affected by a disease due to suspected access to their data, and risk being subject to discrimination. In the health sector, as in any other sector, there is a consistent need for innovation where digital solutions can improve both services and products, but new digital health solutions must place privacy issues and increased public demands for data protection at the forefront of their construction, or they risk neither engaging the needs of the public nor the government demands sanctioning availability via the nature of the data proceeds systems.

In the IoT healthcare context and other scenarios, IoT device data can yield a wide variety of insights, many of which fall within privacy-sensitive areas. IoT device data can also be combined with location or similar data provided by the user, which extends the level of analytics that are observable from the uploaded data. The patient or the investigated individuals of the device-generated data are generally not cognizant that other potential data uploads could lead to competitors discovering their location or health analytic information. Sensitive information about the period of observation could lead to the identification of the likely periods of data uploads, which could be used to estimate more precise time windows of arrivals and departures from points of sale in a retail location, potentially leading to profile creation for commercial targeting.

Hence, to support the deployment of novel services and applications through facilitated data exfiltration, it is also common to require that data continues to be processed and analyzed beyond the lifetime of services, leading to questions about what value a service can inherently receive from the need to design for data whose utility expires with the service. It is becoming increasingly necessary to limit data usage even as innovations continue to be made in data analytics and data storage. In healthcare, the problem for all new conceptual ideas related to analytic development in health data sharing and processing is cemented in the potential infringement of patient data misuse, leading to fast adoption. Legal and policy restrictions are surfacing in dealing with this issue as the only way to ensure compliance with the law and publicized contracts with the patients receiving a service using their data. In this sense, privacy-preserving analytics are employed as systems are constructed rather than post-processing analytics, which assess the nature of the data use once data has been shared or exfiltrated.

3.2. Techniques for Privacy Preservation

There are several techniques developed for preserving privacy while carrying out data analytics. We discuss the most popular ones below:

Data anonymization techniques remove or generalize personal information from the records; however, they keep clinically related information such as illnesses, medications, etc. Encryption can be used to store data that is only decrypted by the service owner or the recipient. Secure Multi-Party Computation always performs computation among multiple data holders without revealing the owned data to each other. Differential privacy is specifically developed to quantify privacy assurance to an individual once their record is queried from the shared dataset. Although all techniques seem to provide good privacy, they still have some concerns about data utility, performance, or privacy assurance. Sometimes, multiple approaches need to be combined to achieve suitable privacy assurance with data utility or performance.

Realizing the significance of protecting patient information, many privacy-preserving schemes for clinical data have been proposed in the recent past. However, most of the existing privacy-preserving schemes for clinical data have considered only some aspects of patient information and treatment, whereas they remain limited in protecting patients' treatment from malicious attacks or providing security against labeling stochastic techniques. Providing guarantees for user confidentiality is a continuous evolutionary process, and there are still challenges. Moreover, the utilization of new technological solutions, such as IoT devices is posing new threats to the identity of users.

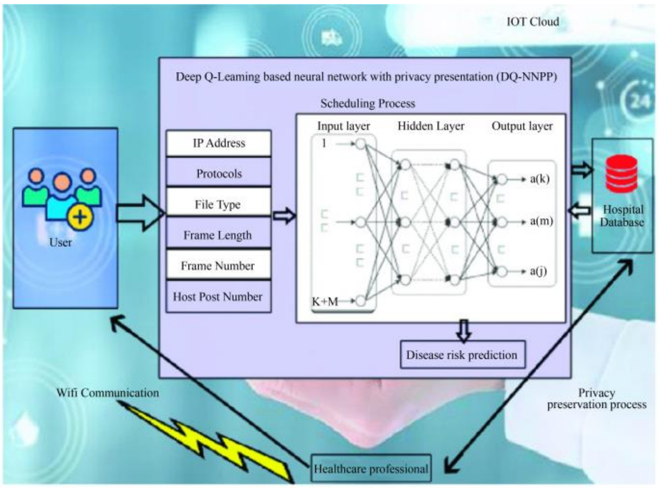


Fig. 4 privacy preservation method in IoT healthcare

4. Federated Learning

Federated learning is an ML-based innovation that enables the collective learning of a shared model from edge and server endpoints' local data without exchanging data samples. The central idea of federated learning is to update a common model based on the private local data of each edge device and then share the updated model. Over multiple model updates, the server averages out the updates against noise caused by initialization and training on locally varying noise. It arrives at a federated mean value after several work rounds. The server uses the trained model for batch or ensemble testing against global validation and testing held-out datasets or scoring against remote in-situ edge sensory data streams. Federated learning can improve client data privacy and the transmission thereof over an untrusted network.

Federated learning can guard against security threats like poisoning attacks, backdoor attacks, model inversion attacks, and differential privacy attacks. Data protection regulations in various countries or regions may provide potential benefits for data-sharing organizations in terms of data protection. Customers stricken with data privacy fears may be more inclined to collaborate with companies since such safeguards are in place. Federated learning takes different forms and has different applications. It can be used in healthcare to reduce data leakage, fraud detection, and security threat management. Different organizations are attempting to deploy federated learning for IoT, healthcare monitoring, and manufacturing applications. However, federated learning still has limitations, such as issues of heterogeneous and time-varying data. Furthermore, it may introduce a high communication cost, particularly when the data stream is dense, and may be ineffective.

Aggregated Global Model

$$\mathbf{w}_{\text{global}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{w}_k$$

$\mathbf{w}_{\text{global}}$: Aggregated global model parameters.

$\mathbf{w}_k$: Local model parameters from device k.

K: Number of participating edge devices.

4.1. Definition and Principles

In this work, we are particularly interested in federated learning, which has emerged as an important domain within data mining and machine learning. The term "federated learning" refers to techniques for training distributed machine learning models that are trained across a decentralized network of multiple devices or servers holding local data samples. The broad definition of federated learning can be seen from two different perspectives: it can be considered as a concept itself or as an approach that can be implemented. Conceptually, federated learning has principles that differentiate its nature from another approach in traditional machine learning. In this context, federated learning introduces the following principles: (i) Distribution: the data is distributed among different data owners through a communication mechanism; (ii) Decentralization: the models are trained on each piece of data in a distributed way, eliminating the need for central storage of data; (iii) Privacy preservation: data holders are the only owners who can share model updates; and (iv) Model adaptation: machine learning models can be adapted to and stored in each piece of data.

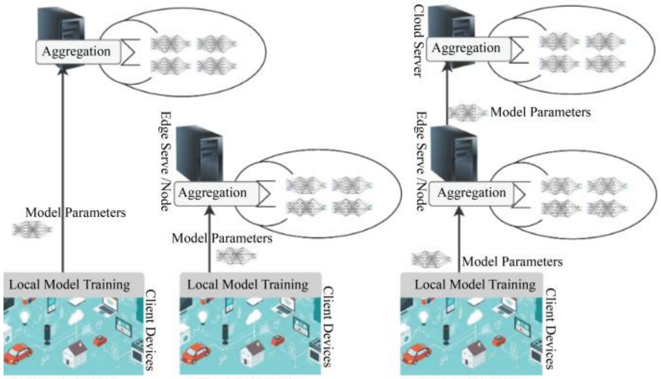
Federated learning allows instance-based learning. The learning process uses models from individuals' data for decision-making as a collaborative task. Each participant in the role of data owner is involved in training the model within their environment through an iterative process. The best performance is produced through optimization conducted using the respective local datasets. The progress update is then continuously evaluated and transferred to the other members of the federated group. The predictions or decisions can then be aggregated to make decisions based on the real context. While the concept of federated learning relies on distributed learning, it is consistent with other decentralized learning approaches, such as model-agnostic learning analysis based on user preferences or data that is locally distributed by some mechanism other than communication. As federated learning is privacy-aware, it is also in sync with privacy-by-design principles that refer to formalism.

4.2. Applications in Healthcare and IoT

Healthcare research and technologies store a large amount of valuable digital data, including clinical measurements, patient records, and biomedical images from genome projects and smartphone-assisted health monitoring. However, the ever-growing consumer fear of privacy misuse linked to data sharing slows down technological development and data utilization. Consequently, healthcare institutions and IoT stakeholders regret sharing their data either directly or via a third party. Several partnerships may require sensitive data to be coupled to leverage the potential of such industrial topics. A predictive model for hypertension was developed using federated learning, showing how the model exceeded the generalizability of data used for validation due to federated results from regions with different spatial distributions of factors. Advances in federated analytics for the discovery as upstream of models, the joint learning of centers, standardization of protocols, and kernel-based multi-center clustering collectively design optimized federated methods for the analysis conducted in real multi-center functional magnetic resonance imaging studies. Similarly, federated outcomes were used for researching fast malignant tumor drug discovery methods, particularly the prediction of tumor growth by a multistage multi-scale tumor development model for validation. The results expand the possibility of utilizing federated learning in various healthcare and IoT data analytics. Edge-based computing needs and promises enhanced algorithms and intelligence to gather and analyze the data rather than in remote data centers. Model personalization could support medical outcome personalization when outcomes are affected by the condition of patients.

5. Integration of Federated Learning and Edge Computing

Federated learning has been suggested as an enabling technology to perform analytics using decentralized learning without sharing any sensitive IoT data, which is a desirable feature in use cases such as IoT in healthcare. Regarding privacy, federated edge learning complements cloud-based federated learning and federated learning at the edge as follows. Cloud-based federated learning mainly focuses on providing privacy guarantees, disregarding strict latency requirements. In contrast, federated edge learning allows for preserving more privacy than federated learning at the edge, which generally needs to collect statistical information related to the input data at the edge device and/or exchange intermediate or final computation states with the edge device. Such statistical information can reverse-engineer or infer part of the data while exchanging intermediate or final computation states, revealing detailed information on the input data. The integration between edge computing and federated learning leads to a beneficial synergy. In the domain of machine learning, many algorithms can be adapted to execute in a distributed or decentralized way, which reduces the data centrality offered by edge computing. Consequently, the potentials of the two techniques overlap, leading to an increased overall performance and privacy guarantees. From a privacy perspective, whenever modeling and inference can be performed at the edge, processed data and predictions leave the edge to reach system management, decision support, or leak to the cloud. Landmarking and Internet of Things processing directly performed by oblivious care receivers bring the data one step closer to the best possible privacy. In case of failure in the system, e.g., computing devices eventually unresponsive, backups can be provided by a skilled physician or IT professional with any necessary credentials. Finally, let's stress that bringing privacy to the edge fosters cooperation among physicians and between physicians and ICT. Fixed-edge servers offer the necessary computing power to run analytics offline. The set of possible devices is much broader than those for wearable and portable devices. Such diversity introduces an issue that does not impact performance but hinders the solution's scalability. While it is possible to train on data acquired by a worn device, improving the model starting from the early stages of data processing, the analytical jump from one to the other may require the intervention of experts. As sweet spots, edge servers can, therefore, be used to improve the scalability of the federated project, addressing only the first type of the three aggregated cases or exploiting late fusion in healthcare applications. In both cases, stakeholders are envisaged to collaborate to obtain useful knowledge efficiently.



(a) Cloud-Based FL (b) Edge-Based FL (c) Hierarchical FL

Fig. 5 Federated learning in edge computing

5.1. Benefits and Challenges

Edge analytics with deep learning at the edge makes feedback predictions quickly using few or single samples and sends the remaining data that could be used to reduce errors in the local model to the cloud. Besides edge analytics, federated learning can also be used at the edge to streamline data aggregation and analysis, reduce communication overhead, and, most importantly, continuously improve local models without revealing personal information from model updates. By pairing federated learning with edge computing via an approach called federated edge learning, the use of edge computing and federated learning can be tightly integrated. Such a union overall has the following benefits and challenges.

a) Local processing efficiency: In edge analytics with deep learning at the edge, a pre-trained neural network generates coarse-grain insights for local processing, with only a little update of the neural network model in the cloud to improve the confidence of the generated coarse-grain predictions. b) Federated learning to ensure privacy: Our use of federated learning at the edge network can encapsulate personal information in the coarse data predictions, thus allowing a cloud server to improve the original cloud service based on the locally encapsulated information received from the edge networks in a privacy-preserving way. A fine-grained analytics engine running on the cloud processes the data but uses any received coarse-grain insights (generated by the pre-trained neural network used for edge analytics) to determine if adaptive local processing of the deep analytics model is needed. Federated learning at the edge would be ideal here, as it continuously improves the coarse analytics running at the edge with minimal changes in the entire distributed system. The edge analytic model can self-regulate the local confidence estimate. The main challenges we see in the applicability of this approach to a real-world scenario are the balance between privacy and adaptivity: maintaining high adaptivity without leaking too much privacy while accounting for error variance in the local predictions could be computationally expensive and may not be supported by the infrastructure in practical use cases. Heterogeneity across large-scale distributed systems could reduce from edge to fog/cloud. For example, at the edge, there are IoT devices/sensors that might vary differently from the IoT at the edge or fog or cloud, leading to differing processing capabilities and throughput. There is a trade-off between the accuracy of the insights from the pre-trained neural network model deployed at the edge and the time needed to obtain them, which, in turn, would differ based on the distribution of the IoT device's data across the edge-cloud infrastructure at their operating condition. In practice, a good strategy would be to allow the local device to have the final say on adjusting simple and energy-aware sensor parameters such as sampling rate. We believe the difference in the above use cases and platforms warrants further theoretical and practical exploration.

5.2. Use Cases and Implementation Strategies

Sense-making: Federated Edge Computing for Healthcare Monitoring Sensors

We presented the concept in a feasibility study in co-creation with a social startup founded by healthcare professionals. Participants were graduate students pursuing master's programs in project management, business information systems, applied computer science, and business administration. We introduced terminology for a basic understanding of federated learning and edge computing. We wanted to know whether they perceived the convergence as feasible and provided us with insights on possible challenges. In general, the students valued the proof of concept and agreed on the issues raised. They also proposed suitable use cases and noted potential research opportunities. For example, the combination of federated learning and edge computing can be used to analyze supply chains or sensory data in sports. The use cases investigated in the interviews are from the fields of healthcare and the Internet of Things in various sectors, including agriculture and manufacturing. Two use cases are also investigated: health monitoring of machine operators and electricity metering.

Edge computing and federated learning are prone to similar implementation obstacles, regardless of the implementation strategy. The case analysis identified the strategies used to tackle the convergence challenges. To implement the convergence of federated learning and edge computing on a large scale in healthcare and IoT, we recommend that practitioners and scholars primarily look at the degree of scalability and adaptability of the selected implementation strategy. An overview of the technology's application spectrum could encourage organizations to invest in the transition or optimization of their operations. In the case of "Storing Data on the Edge," we differentiate between the following implementations of federated learning: a pure machine learning platform, a distributed machine learning platform extended with wallets, and distributed machine learning without blockchain. The implementation can be extended by a combination of the installation implementation strategy and the software bundle sector-related extension.

An important field for further research would be additional unspecified use cases or convergence of federated learning with other emerging technologies, for example, federated learning in combination with tokenization or privacy-preserving analytics democratization. We have mapped out further suggestions to complement future research agendas. Our results enable practitioners to understand which use cases from more than one sector can be empowered by the convergence of federated learning and edge computing. The presented use cases from healthcare and IoT have been identified in the interviews as practical implementations and can be used as a framework for use cases relevant for practice to be added for organizations considering the transition to the convergence of federated learning and edge computing in their (health) IoT operations.

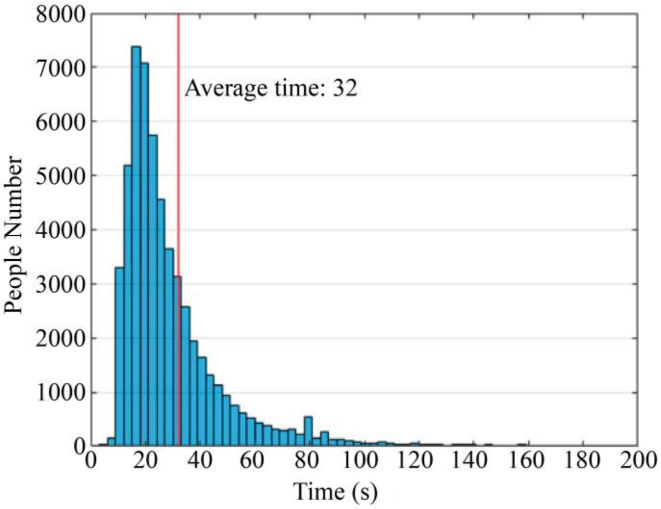


Fig. 6 Healthcare Internet of Things system implementations

6. Case Studies

This section presents several case studies based on real-world applications in healthcare and IoT systems. These sections provide detailed information about each case study, including techniques employed, specific challenges encountered, and the results obtained from the analysis. The case studies provide a comprehensive survey of the capabilities of federated edge computing. While applied to different use cases, the overall challenge in each of them is to deploy machine learning and data analytics techniques in a way that preserves data privacy, minimizes latency, and maximizes operational efficiency. In all the presented studies, the experience of deploying the machine learning models operationalized through the proposed privacy-preserving federation platforms is evaluated through various simulations or real scenarios. The results of the centralized and federated approaches are compared to show the benefits obtained from privacy-preserving federated learning and privacy-preserving computation. A strong focus is given to privacy-preserving computing techniques, including differential privacy and homomorphic encryption. Overall, the case studies provide detailed applications targeting different industrial sectors, healthcare, and IoT systems. They include further developments and optimizations concerning initial descriptions and provide indications to other potential users on the possible practical operation of the privacy-by-design techniques presented herein. These case studies are analyzed in five sections: Government and Mental Health Services, Clinical Research and Medical Sites Cooperation, and Privacy-Preserving Analytics on Diabetes Data.

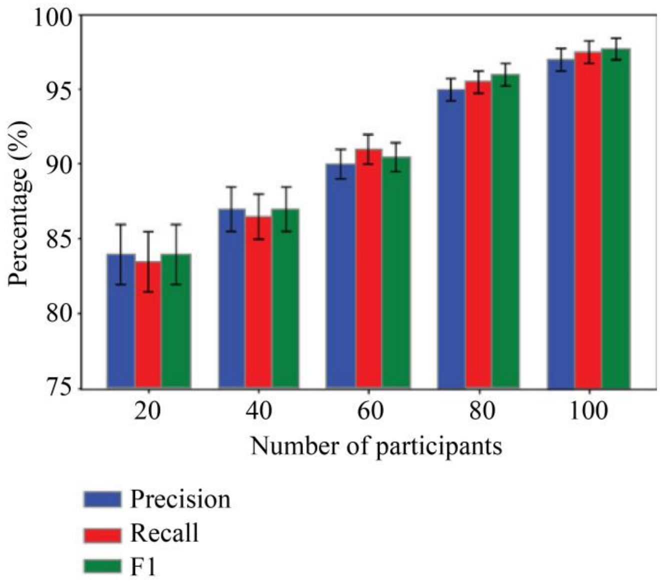


Fig. 7 Secure IoT healthcare architecture with deep learning-based access control system

6.1. Real-World Examples in Healthcare and IoT

While research has so far extensively dealt with the issues and methodologies around federated learning, System A is a full implementation of a real-world, privacy-preserving, and secure analytics system for hospital data.

System B is another example of a distributed, privacy-aware learning system to analyze data obtained from multiple contributing systems, such as a dose planning system, an MR imaging system, and a positron emission tomography scanner that are operated by different organizational entities that have agreed to share the data to optimize clinical workflow in oncology.

System B assimilates results across the participating clinics and allows the person to improve repository-wide accuracy. Five departments of System B came together for five months to form a learning health system, where stakeholders from different disciplines analyze data from three systems: chemotherapy, radiotherapy, and surgery, including functional MRI and histological reports. Currently, 270 datasets from 107 patients were collected. Since 10 May 2021, six meetings took place with a median of 18 participants.

Wherever possible, for System A and System B, we provide a structured analysis of how an organization can embark on this reform, which processes are involved, which legal, ethical, data protection, organizational, and technological frameworks are available, and which challenges any new project incurs – and how they may be approached themselves or in interdisciplinary collaboration. Both System A and System B have instituted a secure distributed data analytics platform that integrates data from an array of sources.

Privacy-Preserving Mechanism

$$\mathbf{w}_k' = \mathbf{w}_k + \mathcal{N}(0, \sigma^2)$$

$\mathbf{w}_k'$: Noised model parameters for privacy.

$\mathbf{w}_k$: Original model parameters.

$\mathcal{N}(0, \sigma^2)$: Gaussian noise with variance σ^2.

7. Conclusion

In this article, we provided a comprehensive overview of federated edge computing and privacy-preserving analytics for IoT and healthcare systems, two interconnected fields where preserving private data represents a top challenge. We analyzed and classified state-of-the-art works in the two areas by considering multiple factors that affect the design of a federated edge computing system. Further, we provided examples of systems and empirical studies conducted by the research community. Based on the analysis, we discussed the main potential advantages and challenges of integrating federated learning and edge computing, as well as provided insights into the role of edge devices in preserving privacy and the multi-party computation algorithms to decentralize the data. We believe our review can inspire new research and development activities, potentially paving the way to novel and innovative strategies for processing and analyzing data in a privacy-preserving fashion. In summary, federated edge computing merges advances in both areas and offers a broader environment where users, researchers, and the entire emerging edge-empowered ecosystem, such as healthcare services, can further enrich and speed up the solution of problems, new developments and applications that will benefit from a more sophisticated analytic approach and new touchpoints with the consumer or the patient. We expect notable innovation as this intertwined value unfolds as an opportunity to deeply improve healthcare and biomedicine sciences through artificial intelligence and machine learning, leveraging resources and throughput from distributed device agglomerates. Most of these devices can be the thing or object of interest and can be potentially decentralized with privacy by design. This phenomenon will allow the industry to offer safe, marketable services. There are also important challenges that are still open to research. For example, some issues include temporality with datasets, feasible server-based aggregation through cryptographic functions, poor real-world measurements, deployments, proofs for research, vendor lock-in because of collectively trained infrastructure and data format, and fusion costs.

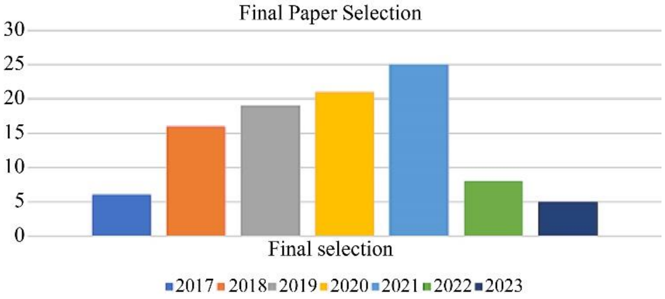


Fig. 8 Enhancing patient healthcare with mobile edge computing

7.1. Future Directions and Research Opportunities

In this paper, we described federated edge computing as a distributed architecture for privacy-preserving analytics and collaborative learning that aims to aggregate data and intelligence from edge devices and utilize such components for localized and personalized citizen healthcare in pervasive and responsive healthcare systems and large-scale operations of IoT analytics systems. We presented three case studies of federated edge computing use cases in healthcare and IoT. The paper also offers a detailed insight into the supportive architecture for federated edge computing scenarios and the challenges researchers and practitioners face. It underscores the need for further collaborative studies, especially in continuous collaboration with other researchers and practitioners. The ensuing content identifies new emerging opportunities for our edge-federated privacy-preserving work as well as promising research directions.

Important emerging directions in federated edge computing and privacy-preserving analytics include Improving the adoption and analysis of large-scale implementation strategies through protocols, algorithms, performance measurement with spatial or temporal analytics, efficiency, and scalability of the learning and analytics. Ethical and regulatory technology, policy engineering, law, and professional engineering are the future products of these systems. More collaborative efforts across the board to overcome challenges with operational testing and translation leveraging mechanisms such as formative studies in the healthcare operation deployment and GDPR-compliant data analytics. Future investigations into microservice-based approach innovations and data infrastructure challenges.

8. References

[1] Ramanakar Reddy Danda, “Financial Services in the Capital Goods Sector: Analyzing Financing Solutions for Equipment Acquisition,” *Library Progress International,* vol. 44, no. 3, pp. 25066-25075, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Rama Chandra Rao Nampalli, and Shakir Syed, “AI-Enabled Rail Electrification and Sustainability: Optimizing Energy Usage with Deep Learning Models,” *Letters in High Energy Physics,* vol. 2024, pp. 822-831, 2024. [Google Scholar] [Publisher Link]

[3] Shakir Syed, “Enhancing School Bus Engine Performance: Predictive Maintenance and Analytics for Sustainable Fleet Operations,” *Library Progress International,* vol. 44, no. 3, pp. 17765-17775, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Manikanth Sarisa et al., “Stock Market Prediction Through AI: Analyzing Market Trends With Big Data Integration,” *Migration Letters,* vol. 21, no. 4, pp. 1846-1859, 2024. [Google Scholar] [Publisher Link]

[5] Rajesh Kumar Malviya et al., “Evolving Neural Network Designs with Genetic Algorithms: Applications in Image Classification, NLP, and Reinforcement Learning,” *Global Research and Development Journals,* vol. 9, no. 12, pp. 9-19, 2024. [Publisher Link]

[6] Ramanakar Reddy Danda et al., “AI and Deep Learning Techniques for Health Plan Satisfaction Analysis and Utilization Patterns in Group Policies,” *International Journal of Medical Toxicology & Legal Medicine,* vol. 27, no. 2, pp. 422-431, 2024. [Google Scholar] [Publisher Link]

[7] Rama Chandra Rao Nampalli, “Leveraging AI and Deep Learning for Predictive Rail Infrastructure Maintenance: Enhancing Safety and Reducing Downtime,” *International Journal of Engineering and Computer Science,* vol. 12, no. 12, pp. 26014-26027, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Shakir Syed, “Planet 2050 and the Future of Manufacturing: Data-Driven Approaches to Sustainable Production in Large Vehicle Manufacturing Plants,” *Journal of Computational Analysis and Applications*, vol. 33, no. 8, pp. 799-808, 2024. [Google Scholar] [Publisher Link]

[9] Hemanth Kumar Gollangi et al., “Data Engineering Solutions: The Impact of AI and ML on ERP Systems and Supply Chain Management,” *Nanotechnology Perceptions,* vol. 20, no. S9, pp. 1-13, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Shaik Abdul Kareem et al., “Neural Transformers for Zero-Day Threat Detection in Real-Time Cybersecurity Network Traffic Analysis,” *International Journal of Global Innovations and Solutions,* 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Ramanakar Reddy Danda et al., “Smart Medicine: The Role of Artificial Intelligence and Machine Learning in Next-Generation Healthcare Innovation,” *South Eastern European Journal of Public Health,* vol. 25, no. 1, pp. 1693-1703, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Rama Chandra Rao Nampalli, “Moderlizing AI Applications in Ticketing and Reservation Systems: Revolutionizing Passenger Transport Services,” *Journal for ReAttach Therapy and Developmental Diversities,* vol. 6, no. 10, pp. 2547-2554, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Shakir Syed, “Big Data Analytics in Heavy Vehicle Manufacturing: Advancing Planet 2050 Goals for a Sustainable Automotive Industry,” *Journal for ReAttach Therapy and Developmental Diversities,* vol. 6, no. 10, pp. 2555-2563, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Chandrakanth Rao Madhavaram et al., “The Future of Automotive Manufacturing: Integrating AI, ML, and Generative AI for Next-Gen Automatic Cars,” *International Multidisciplinary Research Journal Reviews,* vol. 1, no. 1, pp. 20-28, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] Pradeep Chintale et al., “Levy Flight Osprey Optimization Algorithm for Task Scheduling in Cloud Computing,” *International Conference on Intelligent Algorithms for Computational Intelligence Systems,* Hassan, India, pp. 1-5, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Ramanakar Reddy Danda, “Decision-Making in Medicare Prescription Drug Plans: A Generative AI Approach to Consumer Behavior Analysis,” *Journal for ReAttach Therapy and Developmental Diversities,* vol. 6, no. 10, pp. 2587-2598, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Rama Chandra Rao Nampalli, “Neural Networks for Enhancing Rail Safety and Security: Real-Time Monitoring and Incident Prediction,” *Journal of Artificial Intelligence and Big Data,* vol. 2, no. 1, pp. 49-63, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Shakir Syed, "Zero Carbon Manufacturing in the Automotive Industry: Integrating Predictive Analytics to Achieve Sustainable Production," *Journal of Artificial Intelligence and Big Data*, vol. 3, no. 1, pp. 17-28, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Siddharth Konkimalla et al., "A Comparative Analysis of Network Intrusion Detection Using Different Machine Learning Techniques," *Journal of Contemporary Education Theory & Artificial Intelligence*, pp. 1-7, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] Shaik Abdul Kareem, Ram Chandra Sachan, and Rajesh Kumar Malviya, "AI-Driven Adaptive Honeypots for Dynamic Cyber Threats," *SSRN*, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] Ramanakar Reddy Danda, "Generative AI in Designing Family Health Plans: Balancing Personalized Coverage and Affordability," *Utilitas Mathematica*, vol. 121, pp. 316-332, 2024. [Google Scholar] [Publisher Link]

[22] Rama Chandra Rao Nampalli, and Balaji Adusupalli, "Using Machine Learning for Predictive Freight Demand and Route Optimization in Road and Rail Logistics," *Library Progress International*, vol. 44, no. 3, pp. 17754-17764, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Shakir Syed, "Sustainable Manufacturing Practices for Zero-Emission Vehicles: Analyzing the Role of Predictive Analytics in Achieving Carbon Neutrality," *Utilitas Mathematica*, vol. 121, pp. 333-351, 2024. [Google Scholar] [Publisher Link]

[24] Janardhana Rao Sunkara et al., "Optimizing Cloud Computing Performance with Advanced DBMS Techniques: A Comparative Study," *Journal for ReAttach Therapy and Developmental Diversities*, vol. 6, no. 10, pp. 2493-2502, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[25] Pradeep Chintale et al., "Leveraging Aiml Ops for Fraud Detection and Prevention in Fintech," *Journal of Harbin Engineering University*, vol. 45, no. 9, pp. 70-75, 2024. [Google Scholar] [Publisher Link]

[26] Chandrashekar Pandugula et al., "Omni-Channel Retail: Leveraging Machine Learning for Personalized Customer Experiences and Transaction Optimization," *Utilitas Mathematica*, vol. 121, pp. 389-401, 2024. [Google Scholar] [Publisher Link]

[27] Seshagirirao Lekkala, Raghavaiah Avula, and Priyanka Gurijala, "Next-Gen Firewalls: Enhancing Cloud Security with Generative AI," *Journal of Artificial Intelligence & Cloud Computing*, vol. 3, no. 4, pp. 1-9, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[28] Ravi Kumar Vankayalapati et al., "Unifying Edge and Cloud Computing: A Framework for Distributed AI and Real-Time Processing," *Journal for ReAttach Therapy and Developmental Diversities*, vol. 6, no. 9, pp. 1913-1926, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[29] Tulasi Naga Subhash Polineni et al., "AI-Driven Insights Into End-of-Life Decision-Making: Ethical, Legal, and Clinical Perspectives on Leveraging Machine Learning to Improve Patient Autonomy and Palliative Care Outcomes," *Migration Letters*, vol. 19, no. 6, pp. 1159-1172, 2022. [Google Scholar] [Publisher Link]

[30] Kiran Kumar Maguluri et al., "Advancing Pain Medicine with AI and Neural Networks: Predictive Analytics and Personalized Treatment Plans for Chronic and Acute Pain Managements," *Journal of Artificial Intelligence and Big Data*, vol. 2, no. 1, pp. 112-126, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[31] Srinivas Kalisetty, Chandrashekar Pandugula, and Goli Mallesham, "Leveraging Artificial Intelligence to Enhance Supply Chain Resilience: A Study of Predictive Analytics and Risk Mitigation Strategies," *Journal of Artificial Intelligence and Big Data*, vol. 3, no. 1, pp. 29-45, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[32] Lakshminarayana Reddy Kothapalli Sondinti et al., "Towards Quantum-Enhanced Cloud Platforms: Bridging Classical and Quantum Computing for Future Workloads," *Journal for ReAttach Therapy and Developmental Diversities*, vol. 6, no. 10, pp. 492-504, 2023. [CrossRef] [Publisher Link]

[33] Seshagirirao Lekkala, Raghavaiah Avula, and Priyanka Gurijala, "Big Data and AI/ML in Threat Detection: A New Era of Cybersecurity," *Journal of Artificial Intelligence and Big Data*, vol. 2, no. 1, pp. 32-48, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[34] Dilip Kumar Vaka, "Procurement 4.0: Leveraging Technology for Transformative Processes," *Journal of Scientific and Engineering Research*, vol. 11, no. 3, pp. 278-282, 2024. [Google Scholar] [Publisher Link]

[35] Seshagirirao Lekkala, and Priyanka Gurijala, "Establishing Robust Perimeter Defenses," *Security and Privacy for Modern Networks: Strategies and Insights for Safeguarding Digital Infrastructures*, pp. 133-142, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[36] Dilip Kumar Vaka, and Rajesh Azmeera, "Transitioning to S/4HANA: Future Proofing of cross industry Business for Supply Chain Digital Excellence," *International Journal of Science and Research*, vol. 13, no. 4, pp. 488-494, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[37] Dilip Kumar Vaka, "Enhancing Supplier Relationships: Critical Factors in Procurement Supplier Selection," *Journal of Artificial Intelligence, Machine Learning and Data Science*, vol. 2, no. 1, pp. 229-233, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[38] Priyanka Gurijala, and Seshagirirao Lekkala, "Securing Networks with SDN and SD-WAN," *Security and Privacy for Modern Networks: Strategies and Insights for Safeguarding Digital Infrastructures*, pp. 121-131, 2024. [Publisher Link]

[39] Dilip Kumar Vaka, "From Complexity to Simplicity: AI's Route Optimization in Supply Chain Management," *Journal of Artificial Intelligence, Machine Learning and Data Science*, vol. 2, no. 1, pp. 386-389, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[40] Tulasi Naga Subhash Polineni et al., "AI-Driven Automation in Monitoring Post-Operative Complications Across Health Systems," *Global Journal of Medical Case Reports*, vol. 2, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[41] Seshagirirao Lekkala, and Priyanka Gurijala, “Cloud and Virtualization Security Considerations,” *Security and Privacy for Modern Networks: Strategies and Insights for Safeguarding Digital Infrastructures,* pp. 143-154, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[42] Dilip Kumar Vaka, “Integrating Inventory Management and Distribution: A Holistic Supply Chain Strategy,” *International Journal of Managing Value and Supply Chains,* vol. 15, no. 2, pp. 13-23, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[43] Seshagirirao Lekkala, and Priyanka Gurijala, Leveraging AI and Machine Learning for Cyber Defense,” *Security and Privacy for Modern Networks: Strategies and Insights for Safeguarding Digital Infrastructures,* pp. 167-179, 2024. [CrossRef] [Google Scholar] [Publisher Link]