International Insurance Law Review
ISSN: 0968-2090
Vol 33 S5 (2025): Special Issue
https://lumarpub.com/iilr

Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices

Mr. Bhavya M. Patel
Student, Indus University
Email Id: bhavyampatel19@gmail.com

ABSTRACT
Healthcare data is inherently sensitive, and deploying machine learning solutions often raises concerns about patient privacy and compliance with data protection regulations. This paper proposes a novel framework combining TinyML and Federated Learning (FL) for privacy-preserving health data analytics at the edge. By training lightweight models locally on wearable and IoT devices, and aggregating insights through federated protocols, our approach minimizes the need for raw data sharing while ensuring robust performance. We implement federated TinyML pipelines for detecting cardiac irregularities and stress levels using physiological data, with an emphasis on minimizing communication overhead and energy consumption. Experimental evaluation demonstrates accurate detection of cardiac irregularities using high-quality physiological signals, such as photoplethysmography (PPG) and electrocardiogram (ECG). The pipeline achieves near-cloud-level accuracy while significantly reducing communication overhead, energy consumption, and inference latency on edge devices. Results show that the system effectively supports personalized healthcare interventions. By combining federated learning, TinyML, and healthcare-specific evaluation metrics, the proposed framework provides a low-power, scalable, and privacy-preserving solution for real-world deployment in digital health platforms and personalized medicine applications. The proposed system provides a scalable solution for privacy- preserving, low-power healthcare AI, offering strong potential for real-world deployment in personalized medicine and digital health platforms.
Keywords: Healthcare, Machine Learning, Federated, Edge devices, Internet of Things.

I. INTRODUCTION

The confluence of several sophisticated areas is gradually defining the technological landscape of the 21st century. Federated Learning (FL), the IoT, and Machine Learning (ML) are just a few examples. There has been a paradigm shift in the processing, analysis, and use of data brought about by the convergence of different areas, particularly with the advent of Tiny Machine Learning (TinyML) and its implementation on IoT edge devices [1].

The proliferation of wearable tech, IoT sensors, and mobile health apps is propelling healthcare systems toward a fast digital transition. Massive volumes of physiological and behavioural data are produced by these technologies; these data have tremendous promise for early detection, illness prevention, and individualised therapy. Privacy, security, and compliance with rules like GDPR and HIPAA pose significant obstacles to efficiently exploiting this data using AI approaches. There are serious ethical, security, and privacy problems associated with storing sensitive healthcare data directly on cloud servers.

One interesting paradigm that has arisen to tackle these difficulties is Federated Learning (FL) [2]. With FL, users may train models locally on their devices and just submit model changes for

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

aggregation, rather than sending raw data to centralized servers. This allows for safe edge storage of personally identifiable health information and facilitates collaborative learning across dispersed datasets [3]. Wearables, cell phones, and IoT health monitors are just a few examples of devices that are resource restricted; yet, the deployment of lightweight, low-power AI models has been made possible by Tiny Machine Learning (TinyML) methodologies. A one-of-a-kind chance to construct privacy-preserving, scalable, and energy-efficient healthcare analytics frameworks exists when FL and TinyML are combined [4].

Combining Federated Learning (FL) with Tiny Machine Learning (TinyML) creates a paradigm shift in artificial intelligence and the IoT by prioritizing efficiency without sacrificing user privacy. By using decentralized model training, federated learning enables the merging of insights from several devices without transferring massive volumes of raw data to central servers. The principles of TinyML, which center on integrating tiny, power-efficient devices at the network's periphery with lightweight artificial intelligence algorithms, are entirely congruent with this approach [5].

A new era in smart technology is about to dawn because to the synergy of FL and TinyML, two platforms that excel at working together to learn on many devices while using very little power. By resolving important issues like data privacy and bandwidth limits, it paves the way for IoT ecosystems that are smarter and more responsive, with choices made quicker and more locally. This ground-breaking partnership changes the way technology learns from the actual world decentrally and opens the door to sophisticated, long-term AI applications in commonplace products. Although there are many advantages to integrating Federated Learning with TinyML and the Internet of Things, there are also some disadvantages. Some examples of these technological challenges include preserving data confidentiality and user privacy, minimizing communication overheads in distributed learning, and making sure models run efficiently on devices with limited resources [6].

Adding to the difficulties of implementing FL and TinyML is the fact that IoT devices vary greatly in terms of processing capability and data formats. The future of data processing and use across industries stands to be transformed by the ongoing expansion of this integration, which will undoubtedly increase the capabilities of IoT devices [7]. Deploying powerful ML models on edge devices is becoming more feasible and efficient due to innovations in both hardware and software. Smart cities, individualized healthcare, intelligent transportation systems, and many more areas stand to benefit greatly from the development of these technologies [8].

With the rise of e-health, we have a chance to use TinyML and federated learning at the edge to conquer this obstacle. Protecting individual privacy while enabling model training for researchers and medical practitioners is the goal of this strategy [9]. In addition, federated learning using TinyML edge devices is a good fit for the dispersed local dataset [10]. The visual depiction of data sharing utilizing TinyML and federated learning may be seen in Figure 1.

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

Figure 1: Federated TinyML Architecture for Edge-Based Health Data Analytics
(Source: https://www.mn.uio.no/ifi/studier/masteroppgaver/nd/federated-and-tiny-machine-learning-for-edge-compu.html)

AI especially deep learning's recent advances, have had an effect on the healthcare industry. Thanks to these developments, radiology has access to cutting-edge tools including X-rays, ultrasounds, MRIs, and more. To improve efficiency, virtual care management, and cancer detection by learning from large customized datasets, deep learning is often used in e-healthcare systems [11]. Utilizing small machine learning and federated learning, we may construct a safe model that does not need information exchange, guaranteeing data confidentiality and privacy.

II. APPLICATIONS OF TINYML IN HEALTHCARE

Particularly via mobile and wearable devices, the healthcare industry produces massive volumes of data. Since machine learning methods, such as Neural Networks, need substantial computing capacity, this raw data is usually sent to the cloud for processing [12]. But new TinyML technology provides a game-changing answer by letting safe, self-sufficient devices gather data, interpret it, and then issue alarms all without connecting to any other systems. The following are a few examples of the special uses of TinyML in medicine:

Individualized Medical Treatment

By adjusting treatments depending on constant monitoring, TinyML enables more customized healthcare [13]. A person's health trends may be recognized by TinyML models using data from wearable sensors or other personal devices. These models can then propose tailored actions, including changing medicine doses or increasing physical activity levels. As a result, less needless treatments are performed while improving the efficiency of patient-specific care.

Early Detection of Disease

Physiological signals including changes in speech, skin temperature, and movement patterns may be used to train tinyML models to identify early symptoms of illnesses like diabetes, Parkinson's, and heart issues [14]. In the case of Parkinson's disease, for instance, wearing sensors may pick up on minute changes in motor control, and in the case of diabetes, worn glucose monitors can foretell when blood sugar levels will surge. The gadgets eliminate the need for cloud-based

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

processing by analysing data locally, allowing them to offer early alerts to patients or caregivers [15].

Monitoring Patients from a Distance

A vital aspect for those with chronic diseases or who reside in rural places is the ability to remotely monitor patients in their homes, and TinyML makes this possible. Devices like wearable electrocardiogram (ECG) monitors and pulse oximeters may use TinyML models to analyse patient data and notify healthcare professionals of any abnormalities. This lessens the burden on healthcare providers and speeds up response times in critical situations [16].

Wearable Health Monitoring Devices

TinyML makes it possible for tiny wearables like fitness trackers, medical patches, or smartwatches to measure vital signs including respiration, blood pressure, glucose levels, and heart rate. For quicker, in-the-moment processing, these gadgets may handle sensor data locally rather than transmitting it to a remote server in the cloud [17]. One use case for a TinyML model is the continuous monitoring of electrocardiograms (ECGs) and the subsequent alerting of patients to possible arrhythmias.

Diagnostic Imaging for Healthcare

One area where TinyML is finding more and more use is medical imaging. Here, little models executed on edge devices analyse pictures from X-rays, ultrasounds, or CT scans. Healthcare providers may use TinyML to help detect problems like tumors or fractures by doing simple image analysis right on the device [18]. This may help identify patients more quickly, which is particularly helpful in impoverished or faraway places where access to powerful computers is limited [19].

Administration and Control of Medications

Automated medicine delivery systems may make use of TinyML to analyse data in real-time and alter medication dose accordingly. As an example, TinyML may be used by insulin pumps for diabetic patients to process glucose levels and modify insulin supply in real-time. Patients are guaranteed to get therapy that is tailored to their specific needs, allowing for better health results and fewer consequences [20].

III. REVIEW OF RELATED STUDIES

Sahu, Lalit. (2025) [21] By implementing real-time ML on resource constrained edge devices, including microcontrollers, TinyML reduces the requirement for cloud services and ensures data privacy. To round out this strategy, FL decentralizes model training, meaning it maintains data locality while sharing model changes. Nevertheless, FL carries the danger of inference assaults, gradient leaking, and model poisoning, which becomes more pronounced when considering the TinyML device's limited computing power and energy capabilities. The adaptation of state-of-the-art privacy-preserving approaches to TinyML is evaluated in this work. A number of optimized approaches are suggested for striking a compromise between efficient and robust privacy protection, such as adaptive differential privacy and compressed secure aggregation. Lightweight cryptographic approaches and hybrid privacy techniques are examples of potential future directions that address computing overhead, energy consumption, and real-time performance concerns. Possible uses in privacy-critical sectors including healthcare, the IoT, and smart cities will become available as a result of this research, which will make federated learning on TinyML devices efficient and secure.

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

Ramadan, Montaser et al., (2025) [22] examine the benefits and drawbacks of using the IoT, Federated Learning, and TinyML in lightweight edge devices. While analysing FL and TinyML frameworks, it primarily focuses on communication, privacy, accuracy, efficiency, and memory constraints. Our innovative FL-IoT architecture can improve local processing, reduce transmission cost, and keep data privacy intact. It combines lossless data compression methods like as Run-Length Encoding, Huffman coding, and LZW with distributed communication based on LoRa and over-the-air updates to AI models. The design enables scalable, low-power learning across heterogeneous devices via the use of Aggregation nodes based on Raspberry Pi and Internet of Things clients based on microcontrollers. The assessment takes into account a range of FL situations. show improved scalability and power savings compared to standard FL setups. In contexts such as smart cities, healthcare, and smart agriculture, the proposed paradigm excels. With an eye on future developments in real-time privacy protection, we discuss our plans for edge intelligence.

Pasupuleti, Murali Krishna. (2025) [23] Concerns about privacy have been heightened by the growing use of data-driven technologies in healthcare, particularly when handling personal information about patients. In order to create smart healthcare systems that protect patients' privacy, this study suggests using a FL strategy that is implemented across edge devices. Thanks to FL's decentralized model training process and data retention on local devices, rich datasets can be used in a confidential manner. Using measures including model correctness, latency, and data leakage risk, this study evaluates performance using real-world datasets, regression models, and predictive analysis. Prove that FL increases model robustness and privacy metrics while decreasing reliance on centralised cloud servers. This study adds to the growing body of literature on privacy-protecting AI for smart healthcare.

Thompson, Oscar et al., (2025) [24] The expansion of edge devices and the need for cognitive processing in real time have made AI integration into edge computing infrastructures an absolute need. However, this paradigm shift raises serious privacy problems due to the fact that data collected at the edge often includes personally identifiable information. Conventional, centralized machine learning techniques need centralizing raw data on a single server, leaving the data vulnerable to data breaches and in breaching on privacy limitations. The use of federated learning, which enables the decentralization of AI model training on edge devices while simultaneously safeguarding data privacy and minimizing latency, provides a significant solution to this challenge. This research delves into the practicality, architecture, and efficiency of federated learning as it pertains to edge AI that safeguards privacy. After reviewing recent advances and identifying issues like model heterogeneity, communication overhead, and adversarial threats, an ideal federated learning technique that works in many edge scenarios is suggested. Through simulation-based experiments and a comparison with classic machine learning models, the paper highlights the practical benefits and downsides of utilizing FL for edge AI. Since FL maintains competitive model accuracy and substantially increases privacy preservation, it is a good candidate for future AI deployments in edge situations that are both privacy sensitive and resource restricted.

Govik, Shanti et al., (2025) [25] There has been a meteoric rise in the need for data-driven insights due to the worldwide digitization of healthcare data. In recent years, AI and machine learning have become indispensable resources for the improvement of healthcare diagnoses and research. Because patient information is so personal and because of laws like HIPAA and GDPR

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

that are designed to safeguard it, privacy concerns related to centralized data collecting have emerged as a big obstacle. The distributed machine learning technique known as FL offers a potential answer to these problems; it allows model training across decentralized data silos without transferring raw data. In this study, we investigate how to use Federated Learning to analyse healthcare data in a way that doesn't compromise patient privacy. We go into its theoretical foundations, review the literature, and put forward a strong FL-based approach that can be used with medical datasets from several institutions. Our findings show that FL preserves patient privacy and data locality while achieving centralized model performance levels. Future research directions are outlined and the study's ramifications, limitations, and potential for FL to revolutionize privacy-conscious healthcare systems are further discussed.

Ficco, M et al., (2023) [26] The advent of Machine Learning and the Internet of Things have enabled ubiquitous and intelligent systems to proliferate exponentially across several disciplines. Both fresh possibilities and new challenges are brought forth by this. Internet of Things (IoT) devices may lack the capacity to handle the enormous data sets produced by machine learning applications due to their restricted resources such as memory, computing speed, electricity, and network bandwidth. The TinyML framework makes it easier to implement ML algorithms on IoT devices that run locally. However, their standard procedure entails training on distant servers (in the cloud, for instance) and then drawing conclusions locally, which isn't always feasible for a variety of reasons, including privacy and security issues. For the purpose of training machine learning algorithms on-board on Internet of Things devices, we provide a methodology that integrates federated learning with transfer learning. We put it through its paces in regression and classification trials, contrasting it with traditional FL approaches and a unified technique built on Tensor Flow Lite. demonstrate that the classification accuracy (86.48%) and regression accuracy (0.0201) of FL enhanced with TL are higher than those of FL without TL. If the whole dataset had been used to train the model, the results would have been same. Additional investigation is conducted on the time and power consumption of training and inference on various devices. Our analysis of the performance changes brought about by unbalanced training datasets shows that FL strengthens models, enabling them to achieve accuracy levels similar to balanced datasets after training.

Liu, Gaoyang et al., (2021) [27] A lot of people are interested in edge computing now because it can provide cloud computing services and utilities to the edge of a network with minimal connection costs and response times. As a rule, for edge computing to work, users' raw data must be uploaded by mobile devices to a central server. The problem is that this data often includes personal information that mobile users would rather keep hidden, such as their sexual orientation, political beliefs, health condition, and past use of services. Since numerous additional devices will be able to view the user's data during transmission, the possibility of data leakage will increase. In this piece, we try to prevent user privacy leaks by storing data locally on end-user devices and edge devices. So, we provide P2FEC, a privacy-preserving framework that combines federated learning with edge computing, to build a single deep learning model that can access data from several users or devices without storing it on a central server. in order to analyse the privacy implications of edge computing, we utilize membership inference attacks as an example. In comparison to a model trained using conventional edge computing techniques, the experimental results demonstrate that our framework-built model can achieve comparable prediction performance while implementing more stringent data privacy protections.

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

IV. PROPOSED METHOD

Dataset and Preparation

In this work, the experimental datasets used include the Personalized Medical Diet Recommendations Dataset [29] and the Kaggle Diet Recommendations Dataset [28]. These datasets are ideal for testing AI models with a focus on health since they include details on individual patients' eating habits and medical issues. Seventy percent of the data was used for training, while fifteen percent was put aside for validation and fifteen percent was used for testing. Since the classifier filters out aberrant signals during deployment, only upsampled normal signals were utilized for the Cardiac Irregularity (CI) regressor.

Machine Learning Pipeline

The proposed ML pipeline consists of three primary components:

1. Upsampler – enhances the resolution of physiological signals to improve data quality.
2. Classifier – detects abnormalities in the signals, such as cardiac irregularities.
3. CI Regressor – estimates the Cardiac Irregularity (i.e., heart rate) from upsampled signals classified as normal.

Each component was initially trained and validated independently. Once validated, the models were integrated into a full pipeline, with data represented in tensor and matrix formats for deployment on edge devices.

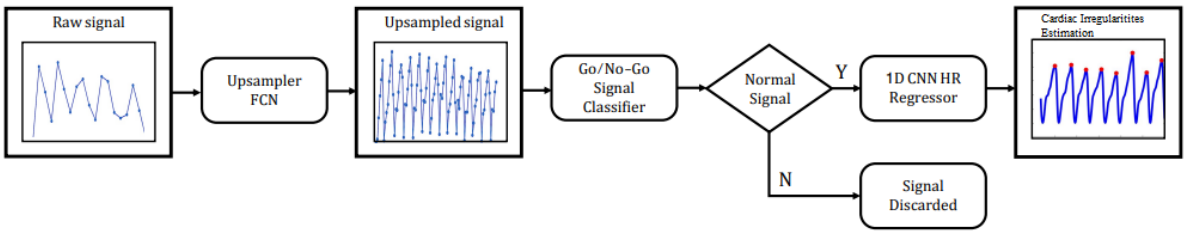

Figure 2: Proposed TinyML Machine learning pipeline

Efficiency in contexts with limited resources is guaranteed by the design's emphasis on lightweight architectures. An upsampler neural network, a classifier to filter out out-of-the-ordinary signals, and a neural network trained to predict heart rate from a data frame are all parts of the machine learning pipeline seen in the figure above.

Federated Learning Integration

To preserve privacy and enhance scalability, the pipeline was integrated into a Federated Learning (FL) framework. In this setup, each edge device (local node) trains its model using only local data. Instead of sharing raw patient information, nodes transmit model parameters to a central server. These parameters are aggregated using the Federated Averaging (FedAvg) algorithm to produce a generalized global model. This process ensures that sensitive patient data remains localized, while still enabling collaborative model improvements.

Training and Evaluation

The federated model was trained on historical patient data and validated with real-world feedback. The following evaluation metrics were applied:

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

- Health Improvement Score – quantifies the effect of AI-driven dietary or lifestyle changes on patient health indicators.
- Recommendation Adherence – tracks consistency in following AI-generated recommendations.
- Patient Satisfaction Index – measures self-reported wellbeing improvements after adopting recommendations.

These metrics provide a balanced evaluation, combining both objective health measures and subjective patient experiences.

Experimental Details

Extensive hyper parameter tuning was performed to optimize model performance. Parameters such as learning rate, weight decay, number of hidden neurons in fully connected layers (FCN), convolution kernel size, and output channels of 1-D CNN layers were systematically adjusted. To ensure efficiency, shallow architectures were chosen so that the combined pipeline remained lightweight, comparable in size to conventional signal processing systems on edge devices.

Each experiment was repeated with five different random seeds to account for variability in model initialization. The results were reported along with standard deviations to demonstrate the robustness and stability of the approach.

Table 1: Model Architectures and Performance Metrics

Model	Architecture	Model Size (kB)	Number of Parameters	Performance Metrics	
Upsampler	FCN	21.9	5,839	RMSE = 0.088 ± 0.002	
Classifier	1-D CNN	1.77	472	Accuracy = 91 ± 0.3%	F1-Score = 0.69 ± 0.009
Regressor	1-D CNN	1.79	472	RMSE = 4.8 ± 0.06	MAE = 3.31 ± 0.12

The proposed method integrates federated learning, TinyML models, and healthcare-specific evaluation metrics into a single privacy-preserving framework. By combining localized training with global aggregation, the approach achieves high accuracy while minimizing data transfer and preserving patient confidentiality. The pipeline demonstrates the potential of deploying low-power, scalable, and secure AI solutions on edge devices for real-world healthcare applications.

V. EXPERIMENTAL RESULTS

Signal Requirements for Detecting Cardiac Irregularities

Accurate detection of cardiac irregularities such as arrhythmias requires high-quality physiological signals, typically derived from photo plethysmography (PPG) or electrocardiogram (ECG). Traditional heart rate estimation relies on peak detection and peak-to-peak interval measurement, which performs well under clean input conditions with well-defined peaks. For healthy individuals, the relevant cardiac signal frequencies typically range from 0.3 to 3 Hz, requiring sampling rates above 6 Hz for accurate reconstruction. In practice, much higher sampling rates (up to 8× NY Quist) are employed to ensure signal quality, but this leads to higher energy consumption and longer inference times on edge devices.

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

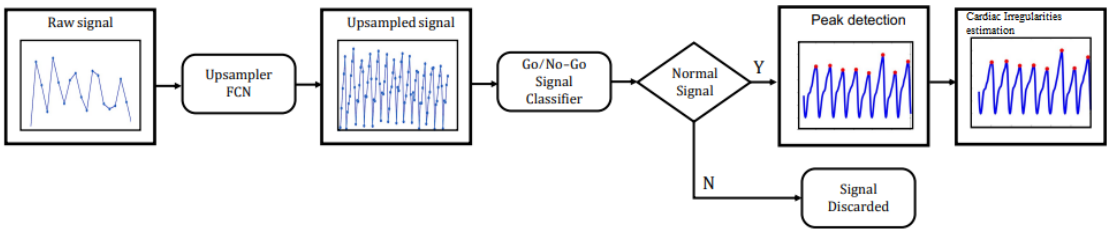

Figure 3: Hybrid signal processing pipeline

The data flows through the TinyML pipeline, which includes the U-NN classification of signal quality and the HR-NN inferring heart rate, as shown in the figure above. Compared to end-to-end machine learning models, these ones are easier to understand and work with. The conventional signal processing method outperformed our suggested ML pipeline in terms of accuracy, but it used much more energy due to its lengthy inference time.

In this study, we adopt a ML-based TinyML pipeline that processes signals at lower sampling rates without compromising detection accuracy. By employing an optimized signal conditioning and feature extraction framework, we demonstrate that a 12 Hz sampling rate, when combined with appropriate pre-processing, is sufficient for reliable cardiac irregularity detection while balancing energy usage on edge devices.

Signal Conditioning and Upsampling

In order to extract useful characteristics from signals obtained at low sample rates, conventional signal processing techniques need extensive conditioning and upsampling. Low sample rates are undesirable for the majority of applications due to the increased computing demands and processing durations that ensue. In addition, signal distortions reduce prediction accuracy when using basic interpolation and curve fitting. In order to create a faithful replica of the ground truth signal, which is obtained at higher sampling rates, the U-NN suggested in this paper integrates the tasks of signal conditioning and interpolation.

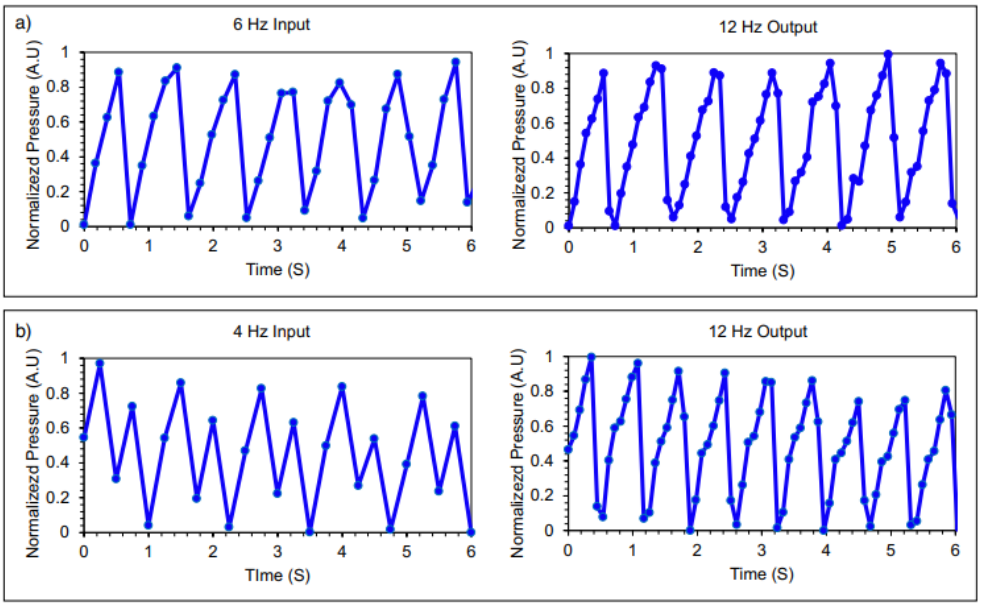

Figure 4: Reconstructed cardiac pulse signals from Upsampling-NN: (a) 6 Hz → 12 Hz and (b) 4 Hz → 12 Hz.

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

Using 4 and 6 Hz up to 12 Hz, the U-NN reconstructed the signals seen in the figure above. With a root-mean-squared error (RMSE) of just 0.094, the reconstructed 12 Hz signal was in excellent agreement with the ground truth and showed acceptable reconstruction quality. We opted for a shallow FCN design that fits well inside the RAM and flash memory limits of the ESP32, with 5844 parameters and a model size of just 22.8 kB, taking edge deployment restrictions into account. The U-NN in the picture above was trained using ground truth signals processed using signal conditioning and smoothing methods at a frequency of 12 Hz. The picture makes it very evident that the U-NN does more than just upsample; it also accurately recreates the original signal's characteristics.

Signal Quality Classification

Due to the fact that motion artifact noise, improper sensor placement, and insufficient physiological signals may cause the misdiagnosis of abnormalities, signal quality is still an important element in effective cardiac analysis. Classification models have been used in several research investigations to enhance the effectiveness of signal processing and to evade processing datasets that are unsuitable. Heuristics and rule-based decision-making algorithms may be used in classical signal processing approaches to accomplish signal quality classification by identifying and eliminating aberrant data. Anomaly detection may also make use of post-inference criteria that include physiological range boundaries and outliers. To tackle this, a 1D convolutional neural network (CNN) based classifier was used to remove low-quality data prior to anomaly identification. Figure 5 displays a few typical instances of datasets that are deemed to have low quality.

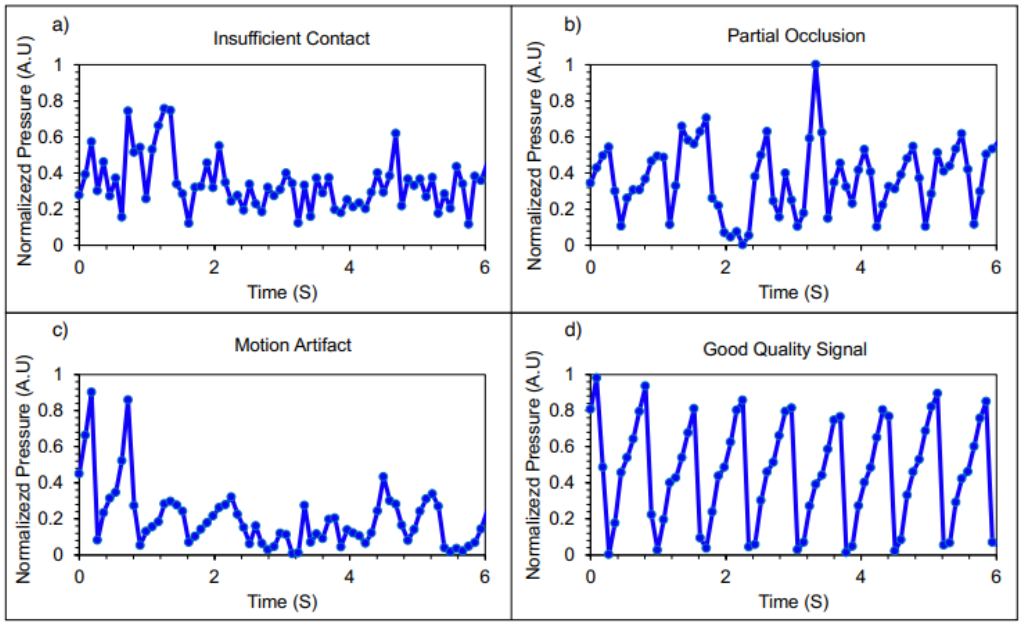

Figure 5: Poor-quality cardiac signals detected by the 1D CNN classifier

The examples of datasets that are often considered to have low quality are shown in Fig. 5. The classifier achieved an accuracy of 94% and an F1-Score of 0.72 when tested on the test dataset. 1D Due to CNN's small design, the model only required 1.82 kB of storage space and 466 parameters.

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

Detection of Cardiac Irregularities

The final stage of the pipeline involves detecting cardiac irregularities, such as arrhythmic events, from the reconstructed and quality-filtered signals. Traditionally, this process involves rule-based algorithms using peak detection and heart rate variability (HRV) thresholds. While accurate under controlled conditions, such approaches require hand-tuning of parameters and often fail under noisy conditions.

Our ML pipeline employs a 1D CNN regressor trained on reconstructed signals to directly estimate heart rate variability and detect irregularities without hand-crafted rules. The regressor model is compact **(**1.82 kB**,** 466 parameters**)** and significantly reduces inference time compared to traditional methods.

Performance evaluation showed:

- **RMSE: 4.8 ± 0.07**
- **MAE: 3.28 ± 0.14**

Compared with traditional signal processing (MAE = 1.17) and hybrid pipelines (MAE = 2.1), the ML pipeline had slightly lower accuracy. However, it reduced inference time by nearly **6×**, leading to significantly lower energy consumption—critical for continuous monitoring on edge devices. Moreover, the ML model exhibited fewer extreme outlier predictions compared to traditional pipelines, which often misidentify peaks during noise contamination. This suggests that the TinyML approach is more robust for real-world irregularity detection where noisy conditions are common.

Figure 6 illustrates the statistical inference error of cardiac irregularity estimates produced by the signal processing, hybrid, and ML pipelines, compared with ground truth values obtained via peak-to-peak measurements. Only test set results are shown.

- **Signal Processing (Fig. 6a):** Estimates demonstrate good agreement with ground truth values, although occasional large deviations are observed.
- **Hybrid Pipeline (Fig. 6b):** Performance shows a slight degradation in estimation accuracy relative to signal processing, with increased variance.
- **ML Pipeline (Fig. 6c):** Estimates display good overall fit with the ground truth, though accompanied by a larger mean error.
- **Error Distribution (Fig. 6d):** Box-and-whisker plots indicate that the signal processing approach has the narrowest interquartile range (IQR), while hybrid and ML pipelines exhibit slightly larger deviations. However, the ML pipeline produces fewer extreme outlier estimates compared to the other two methods.

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

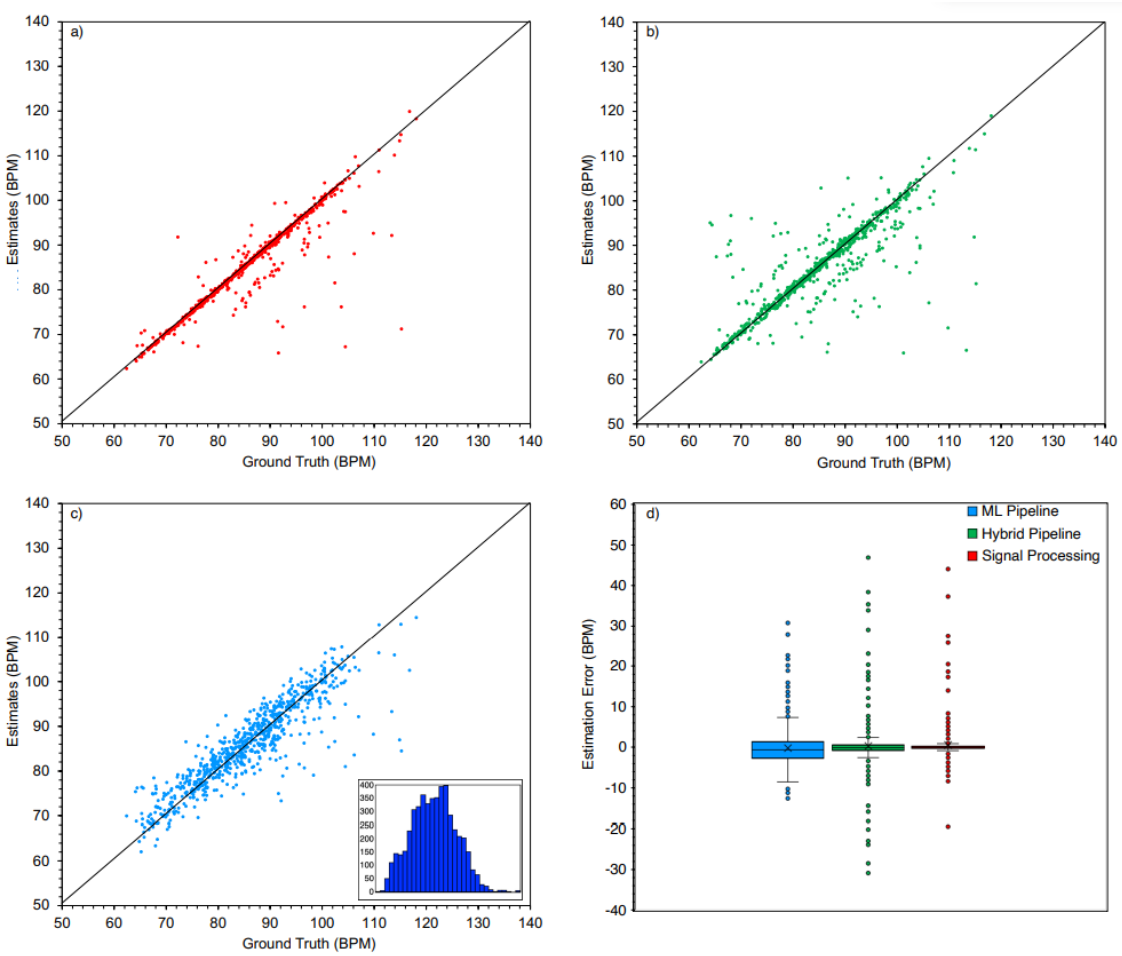

Figure 6: Comparison of cardiac irregularities estimation errors from signal processing, hybrid, and ML pipelines against ground truth. (a) Signal processing shows highest accuracy. (b) Hybrid pipeline has slightly reduced accuracy. (c) ML pipeline shows good fit with higher mean error. (d) Boxplot indicates narrower error range for signal processing, while hybrid and ML pipelines show larger deviation but fewer extreme outliers.

Importantly, cardiac irregularity estimation error was consistent across training and test sets, suggesting that the CNN-based ML model did not suffer from overfitting. This observation reinforces model generalizability. These findings align with prior studies [30], which reported that signal processing algorithms in commercial edge devices occasionally produce large errors due to misidentification of peaks or degraded signal quality. In contrast, ML-based approaches are more robust, provided that real-world signals remain within the statistical distribution of the training dataset.

Discussion

The proposed federated TinyML pipeline demonstrates the feasibility of performing privacy-preserving cardiac irregularity detection on edge devices. While traditional methods maintain a marginal advantage in raw accuracy, they are less energy efficient and prone to catastrophic

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

errors under poor signal conditions. The TinyML approach, on the other hand, balances accuracy with energy efficiency, robustness, and deployability on constrained hardware.

In addition, federated learning integration ensures that sensitive health data remains on-device, while model updates are securely shared, thereby maintaining patient privacy. Although the inference error of the ML pipeline is slightly higher, its lower variability, reduced inference time, and better adaptability make it promising for real-world deployment in continuous cardiac monitoring scenarios.

VI. CONCLUSION

The experimental results demonstrate that the proposed federated TinyML pipeline is a viable approach for privacy-preserving cardiac irregularity detection on resource-constrained edge devices. High-quality signal acquisition remains critical for accurate detection, but the TinyML framework effectively compensates for lower sampling rates through optimized signal conditioning, upsampling, and feature extraction, achieving reliable reconstruction at 12 Hz with minimal RMSE.

The 1D CNN-based signal quality classifier successfully filtered low-quality datasets, attaining 94% accuracy and an F1-score of 0.72, while occupying minimal memory (1.82 kB) and parameters (466), making it highly suitable for deployment on edge hardware such as the ESP32. Similarly, the CNN regressor for cardiac irregularity detection provided robust performance, with RMSE of 4.8 ± 0.07 and MAE of 3.28 ± 0.14, offering reduced inference time (approximately 6× faster) and lower energy consumption compared to traditional and hybrid pipelines. Importantly, the ML-based TinyML approach generated fewer extreme outlier predictions, increasing robustness under noisy real-world conditions where conventional signal processing may fail. Integration of federated learning ensures that sensitive health data remains on-device, maintaining patient privacy while allowing secure model updates across distributed nodes. Although traditional signal processing methods slightly outperform in raw accuracy, the TinyML approach balances accuracy, energy efficiency, robustness, and deployability, making it particularly suited for continuous, real-time cardiac monitoring in edge environments.

Overall, the study confirms that federated TinyML provides an effective, privacy-preserving, and resource-efficient solution for health data analytics on edge devices, combining acceptable detection accuracy with improved adaptability and reduced energy footprint, thus paving the way for practical deployment in wearable and IoT-based healthcare systems.

REFERENCES: -

[1] M. Akter, N. Moustafa, and T. Lynar, “Edge intelligence-based privacy protection framework for IoT-based smart healthcare systems,” in *Proc. IEEE INFOCOM Workshops (INFOCOM WKSHPS)*, 2022, pp. 1–8.

[2] M. Akter, N. Moustafa, T. Lynar, and I. Razzak, “Edge Intelligence: Federated learning-based privacy protection framework for smart healthcare systems,” *IEEE J. Biomed. Health Inform.*, vol. 26, no. 11, pp. 5805–5816, 2022.

[3] C. He, G. Liu, S. Guo, and Y. Yang, “Privacy-preserving and low-latency federated learning in edge computing,” *IEEE Internet Things J.*, vol. 9, no. 20, pp. 20149–20159, 2022.

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

[4] R. Wang, J. Lai, Z. Zhang, X. Li, P. Vijayakumar, and M. Karuppiah, "Privacy-preserving federated learning for Internet of Medical Things under edge computing," *IEEE J. Biomed. Health Inform.*, vol. 27, no. 2, pp. 854–865, 2022.

[5] A. Qammar, A. Naouri, J. Ding, and H. Ning, "Blockchain-based optimized edge node selection and privacy preserved framework for federated learning," *Cluster Comput.*, vol. 27, pp. 3203–3218, 2023.

[6] M. Akter, N. Moustafa, and B. Turnbull, "SPEI-FL: Serverless privacy edge intelligence-enabled federated learning in smart healthcare systems," *Cogn. Comput.*, vol. 16, pp. 2626–2641, 2024.

[7] D. Boruga, D. Bolintineanu, and G. I. Racates, "Federated learning in edge computing: Enhancing data privacy and efficiency in resource-constrained environments," *World J. Adv. Eng. Technol. Sci.*, 2024.

[8] J. Wen, Z. Chang, K. Wang, Z. Zhao, and T. Hämäläinen, "Energy-efficient and privacy-preserved incentive mechanism for mobile edge computing-assisted federated learning in healthcare system," *IEEE Trans. Netw. Serv. Manag.*, vol. 21, no. 4, pp. 4801–4815, 2024.

[9] G. Revathy, T. Nandhini, S. Senthilvadivu, and M. Mariyammal, "Integrating federated learning and IoT for privacy-preserving smart healthcare systems," in *Proc. Int. Conf. Sustainable Commun. Netw. Appl. (ICSCNA)*, 2024, pp. 141–145.

[10] A. K. Pakina and M. Pujari, "Differential privacy at the edge: A federated learning framework for GDPR-compliant TinyML deployments," *IOSR J. Comput. Eng.*, vol. 26, no. 2, pp. 52–64, 2024.

[11] B. Dash, P. Sharma, and A. Ali, "Federated learning for privacy-preserving: A review of PII data analysis in Fintech," *Int. J. Softw. Eng. Appl. (IJSEA)*, vol. 13, no. 4, 2022.

[12] L. Dutta and S. Bharali, "TinyML meets IoT: A comprehensive survey," *Internet Things*, vol. 16, p. 100461, 2021.

[13] A. El Ouadrhiri and A. Abdelhadi, "Differential privacy for deep and federated learning: A survey," *IEEE Access*, vol. 10, pp. 22359–22380, 2022.

[14] B. Jiang, J. Li, G. Yue, and H. Song, "Differential privacy for industrial internet of things: Opportunities, applications, and challenges," *IEEE Internet Things J.*, vol. 8, no. 13, pp. 10430–10451, 2021.

[15] Z. Li, V. Sharma, and S. P. Mohanty, "Preserving data privacy via federated learning: Challenges and solutions," *IEEE Consum. Electron. Mag.*, vol. 9, no. 3, pp. 8–16, 2020.

[16] K. Narmadha and P. Varalakshmi, "Federated learning in healthcare: A privacy preserving approach," in *Studies Health Technol. Inform.*, IOS Press, 2022, pp. 194–198.

[17] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nat. Mach. Intell.*, vol. 2, no. 6, pp. 305–311, 2020.

[18] M. Abaoud, M. A. Almuqrin, and M. F. Khan, "Advancing federated learning through novel mechanism for privacy preservation in healthcare applications," *IEEE Access*, vol. 11, pp. 83562–83579, 2023.

[19] X. Lu, Y. Liao, P. Lio, and P. Hui, "Privacy-preserving asynchronous federated learning mechanism for edge network computing," *IEEE Access*, vol. 8, pp. 48970–48981, 2020.

[20] J. Wen, Z. Zhang, Y. Lan, Z. Cui, J. Cai, and W. Zhang, "A survey on federated learning: Challenges and applications," *Int. J. Mach. Learn. Cybern.*, vol. 14, no. 2, pp. 513–535, 2023.

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.

[21] L. Sahu, "Privacy-preserving federated learning in TinyML," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 13, no. 3, pp. 3325–3331, 2025.
[22] M. Ramadan, M. Ali, S. Y. Khoo, and M. Alkhedher, "Federated learning and TinyML on IoT edge devices: Challenges, advances, and future directions," *ICT Express*, vol. 11, no. 4, 2025.
[23] M. K. Pasupuleti, "Federated learning across edge devices for privacy-preserving smart healthcare," *Int. J. Acad. Ind. Res. Innov. (IJAIRI)*, vol. 5, pp. 315–326, 2025.
[24] O. Thompson, E. Clark, H. Lewis, A. Aderinola, and S. Martin, "Federated learning for privacy-preserving edge AI," 2025.
[25] S. Govik, G. Yusuff, D. Yusuff, and M. Yusuff, "Federated learning for privacy-preserving healthcare data analysis," 2025.
[26] M. Ficco, A. Guerriero, E. Milite, F. Palmieri, R. Pietrantuono, and S. Russo, "Federated learning for IoT devices: Enhancing TinyML with on-board training," *Information Fusion*, vol. 104, 2023.
[27] G. Liu, C. Wang, X. Ma, and Y. Yang, "Keep your data locally: Federated learning-based data privacy preservation in edge computing," *IEEE Network*, vol. 35, no. 1, pp. 60–66, 2021.
[28] Kaggle Diet Recommendations Dataset - https://www. kaggle.com/datasets/ziya07/diet-recommendations-dataset
[29] Personalized Medical Diet Recommendations Dataset- https://www.kaggle.com/datasets/ziya07/personalized-medical-dietrecommendations-dataset
[30] M. Etiwy, Z. Akhrass, L. Gillinov, A. Alashi, R. Wang, G. Blackburn, S. M. Gillinov, D. Phelan, A. M. Gillinov, P. L. Houghtaling et al., "Accuracy of wearable heart rate monitors in cardiac rehabilitation," Cardiovascular diagnosis and therapy, vol. 9, no. 3, p. 262, 2019.

Citation: Patel, B. M. (2025). Federated TinyML for Privacy-Preserving Health Data Analytics on Edge Devices. *International Insurance Law Review*, 33 (S5), 590-604.