

This is an Information series on Federated learning by Marktechpost. We have featured some of the papers that were trending in the last quarter of 2021 and we summarized them in a short article format.

## MARKTECHPOST

Marktechpost is an AI News Platform providing easy-to-consume, byte size updates in machine learning, deep learning, and data science research. Our vision is to showcase the hottest research trends in AI from around the world using our innovative method of search and discovery

Email: [Asif@marktechpost.com](mailto:Asif@marktechpost.com)

Website: [www.marktechpost.com](http://www.marktechpost.com)

## OUR CONTRIBUTORS

**Luca Arrotta**

*(Content writer)*

**Tanushree Shenwai**

*(Content writer)*

**Swapnil Dnyaneshwar More**

*(Content writer)*

**Nitish Kumar**

*(Content writer)*

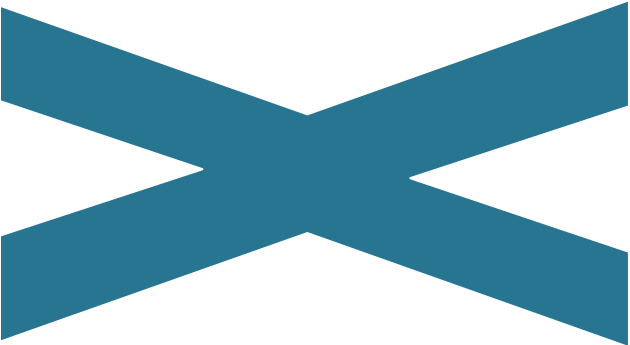
**G Chaithali**

*(Content writer)*

**Naveen Mahanwal**

*(Graphic designer)*






*Hierarchical Federated Learning-Based Anomaly Detection Using Digital Twins For Internet of Medical Things (IoMT)* \_\_\_\_\_ (03)

*Hierarchical Federated Learning-Based Anomaly Detection Using Digital Twins For Internet of Medical Things (IoMT)* \_\_\_\_\_ (06)



*Google AI Introduces ‘Federated Reconstruction’ Framework That Enables Scalable Partially Local Federated Learning* \_\_\_\_\_ (10)


*Researchers Propose ‘ProxyFL’: A Novel Decentralized Federated Learning Scheme For Multi-Institutional Collaborations Without Sacrificing Data Privacy* \_\_\_\_\_ (12)



*Google AI Improves The Performance Of Smart Text Selection Models By Using Federated Learning* \_\_\_\_\_ (15)



*NVIDIA Open-Source ‘FLARE’ (Federated Learning Application Runtime Environment), Providing A Common Computing Foundation For Federated Learning* \_\_\_\_\_ (17)

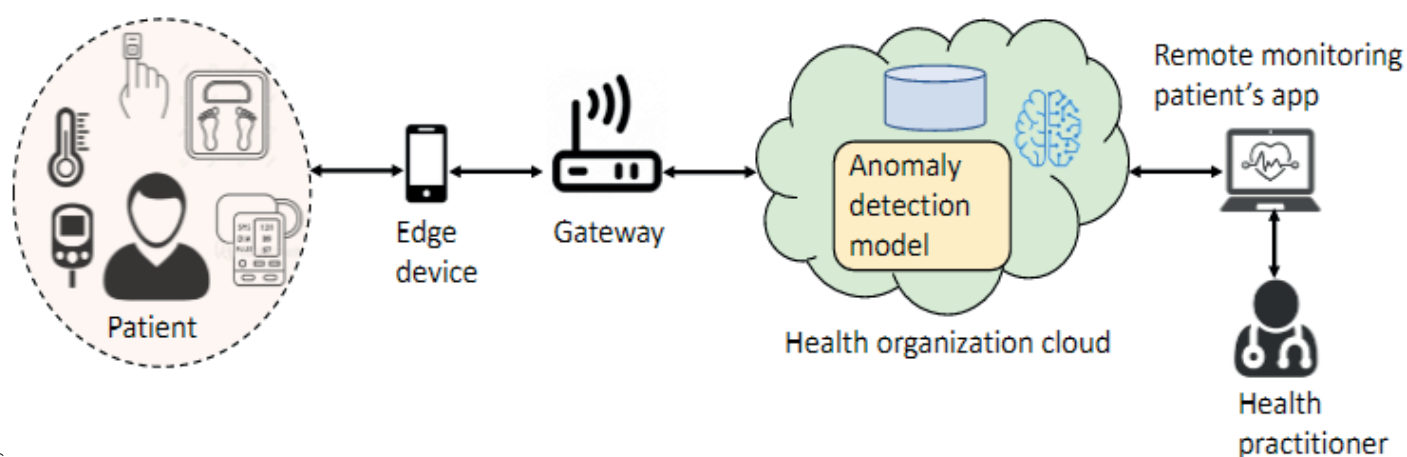


*Ericsson And Uppsala University Team Up To Research Air Quality Prediction Using Machine Learning And Federated learning* \_\_\_\_\_ (19)

# HIERARCHICAL FEDERATED LEARNING-BASED ANOMALY DETECTION USING DIGITAL TWINS FOR INTERNET OF MEDICAL THINGS (IOMT)

Smart healthcare services can be provided by using Internet of Things (IoT) technologies that monitor the health conditions of patients and their vital body parameters. The majority of IoT solutions used to enable such services are wearable devices, such as smartwatches, ECG monitors, and blood pressure monitors. The huge amount of data collected from smart medical devices leads to major security and privacy issues in the IoT domain. Considering Remote Patient Monitoring (RPM) applications, we will focus on Anomaly Detection (AD) models, whose purpose is to identify events that differ from the typical user behavior patterns. Generally, while designing centralized AD models, the researchers face security and privacy challenges (e.g., patient data privacy, training data poisoning).

To overcome these issues, the researchers of this paper propose an Anomaly Detection (AD) model based on Federated Learning (FL). Federated Learning (FL) allows different devices to collaborate and perform training locally in order to build Anomaly Detection (AD) models without sharing patients' data. Specifically, the researchers propose a hierarchical Federated Learning (FL) that enables collaboration among different organizations, by building various Anomaly Detection (AD) models for patients with similar health conditions.



Source:

*Hierarchical Federated Learning based Anomaly Detection using Digital Twins for Smart Healthcare*  
<https://arxiv.org/pdf/2111.12241.pdf>

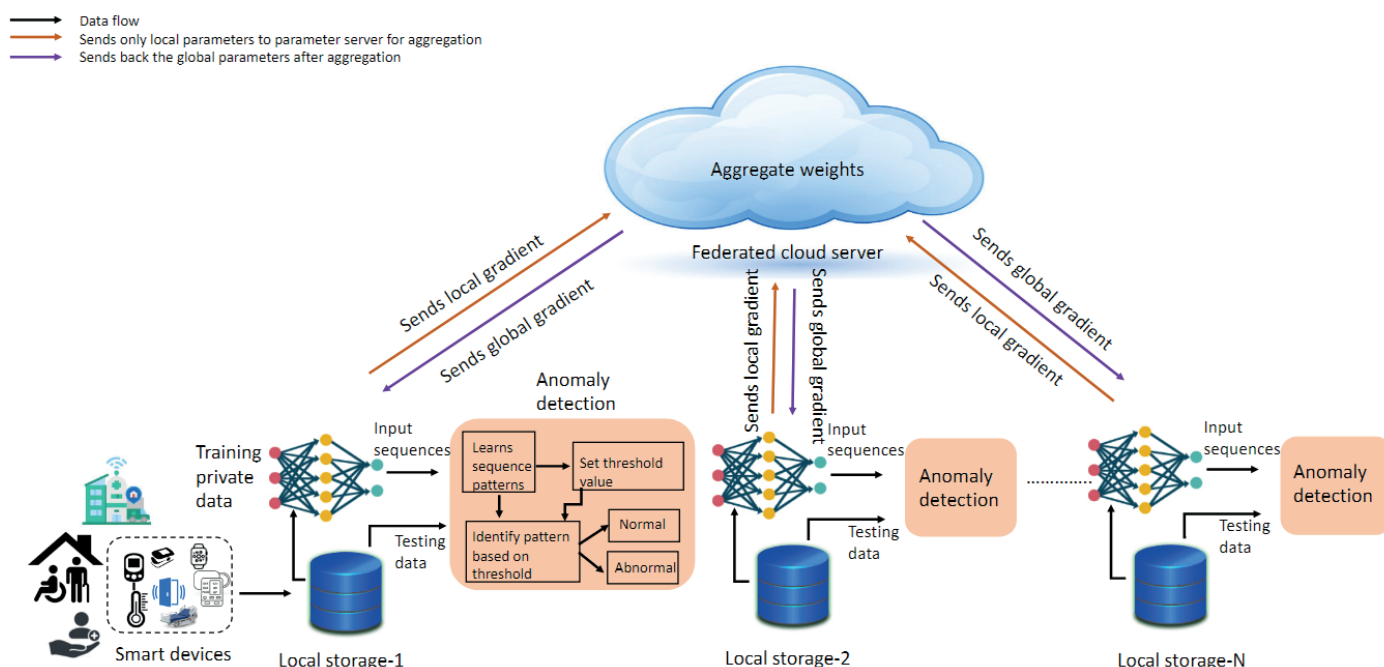
The figure above depicts a centralized Anomaly Detection (AD) model that can be used by clinicians to remotely monitor their patients. In this architecture, the Anomaly Detection (AD) model and the dataset are both in the same health organization cloud, a solution that facilitates model training and read/write operations. However, this approach leads to different weaknesses, such as having a single point of failure and data privacy issues since data from multiple sources are stored at the same location. Let's discuss some of the possible threats scenarios associated with centralized Anomaly Detection (AD) solutions.

1. **Privacy Leakage:** malicious users may gain access to the sensitive health information of any patient. At the same time, the patients should have control over their data. The proposed hierarchical Federated Learning (FL) solution ensures that only the local Anomaly Detection (AD) model assigned to a group of patients has access to their data, without sharing them with the Anomaly Detection (AD) models of the other patients.

2. Training Data Poisoning: an attacker could poison the training set by introducing data samples that impact the recognition rate of the model. Since with the Federated Learning (FL) paradigm the training data are stored locally on each client, it is harder to poison such localized data

3. Model Drift: generally, model drift occurs when the underlying statistical structure of the data changes over time. An attacker could alter this statistical structure by introducing specific data points within the training set. The proposed hierarchical FL solution relies on a disease-based grouping mechanism, ensuring that data generated by each patient will not dramatically affect the statistical structure of the data.

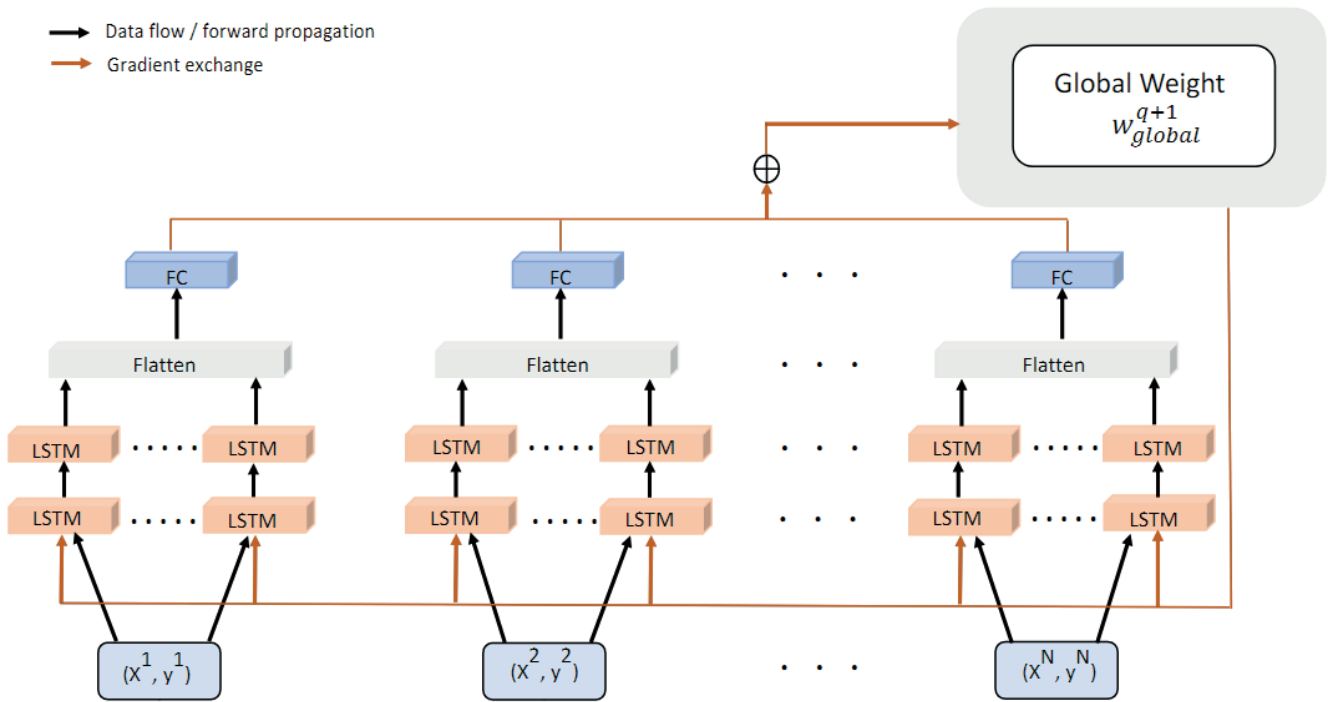
4. Performance Overhead: a single point of failure could bring to the loss of data and high response time that can affect the performance of an Anomaly Detection (AD) model. A Federated Learning (FL) solution, for example, can improve the response time by performing decentralized training on each client's device.



Source:

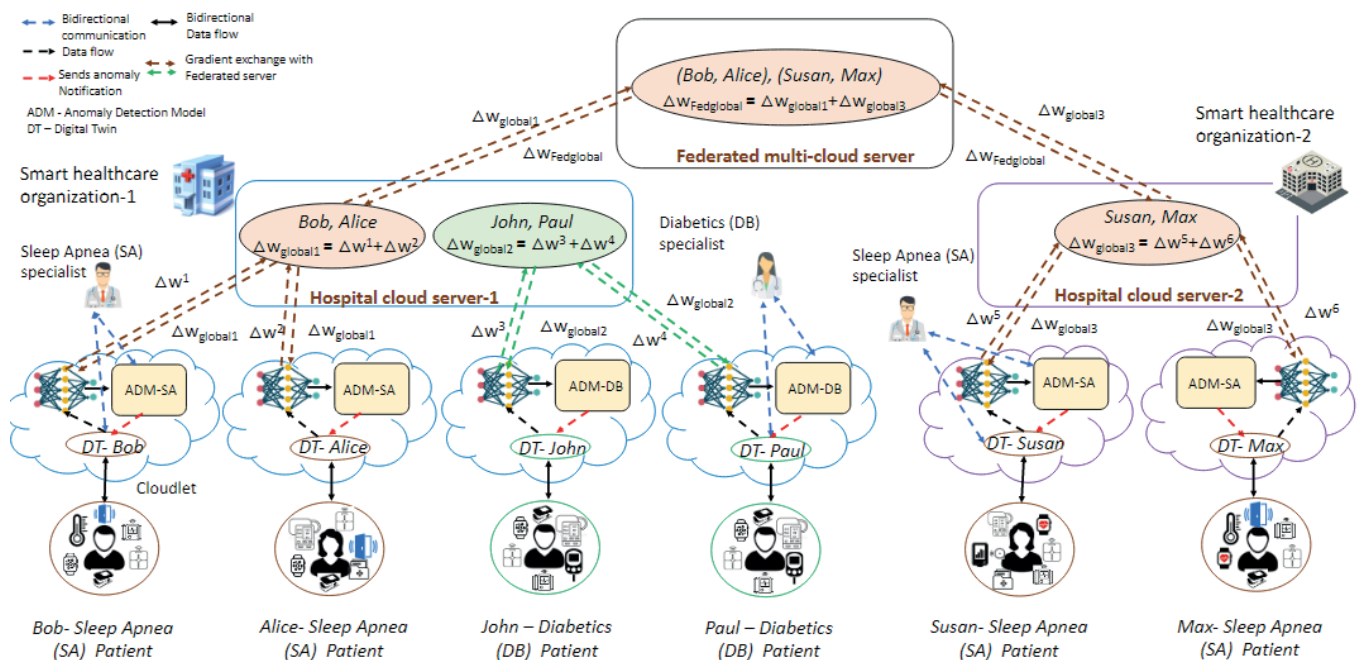
*Hierarchical Federated Learning based Anomaly Detection using Digital Twins for Smart Healthcare*  
(<https://arxiv.org/pdf/2111.12241.pdf>)

The figure above depicts the proposed Anomaly Detection (AD) model based on Federated Learning (FL). Each of the N participants has its own locally stored dataset that includes data collected by a set of smart IoT devices (e.g., wearable glucose meter, smartwatch). The samples of the dataset are labeled to detect “normal” and “abnormal” observations. Each participant trains a local model and then sends the model weights to the federated cloud server. Hence, this server receives the weights from all the participants’ local models. Then, it aggregates these weights based on specific attributes (e.g., user’s age, disease name) and sends them back to the participants as a global weight.



Source:  
 Hierarchical Federated Learning based Anomaly Detection using Digital Twins for Smart Healthcare  
 (<https://arxiv.org/pdf/2111.12241.pdf>)

The neural network used for Anomaly Detection (AD) is a Time Distributed LSTM, depicted in the figure above. The proposed model presents two stacked layers of four LSTM cells arranged sequentially. Hence, local training on each device is performed through sequences of four input samples given as input to the neural network. This training process is carried out until the recognition error is below a specific threshold, or after H epochs. As already described before, the weights of the local models are then uploaded to the server for aggregation. After receiving the global weights from the federated server, each device continues the training. The whole process shown in the figure below continues until the N local models are optimized.

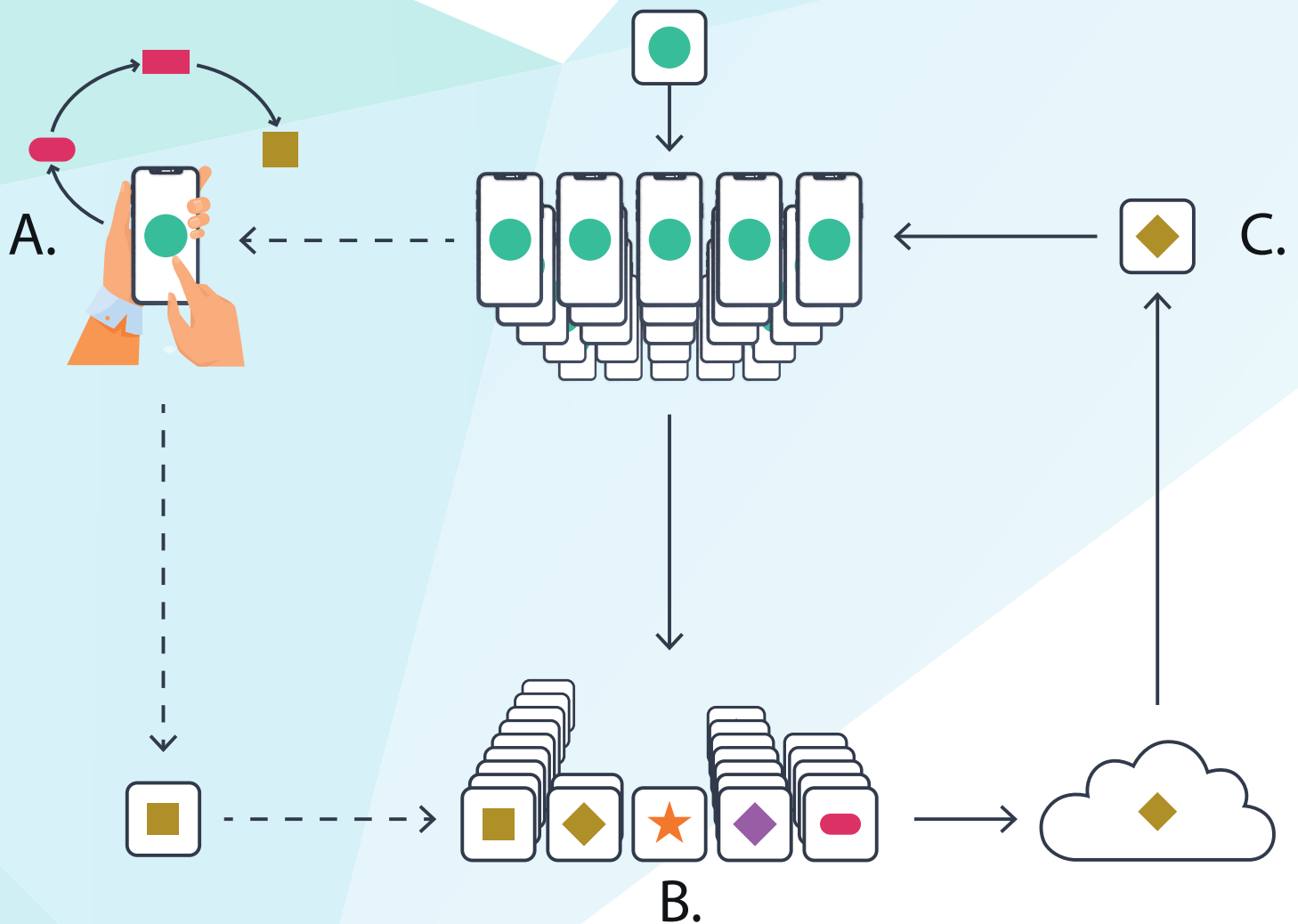


Source:  
 Hierarchical Federated Learning based Anomaly Detection using Digital Twins for Smart Healthcare  
 (<https://arxiv.org/pdf/2111.12241.pdf>)

Finally, let's focus on the RPM use case that uses the proposed Federated Learning (FL) framework, as illustrated in the figure above. This use case presents a scenario where patients are continuously monitored by clinicians. We consider two smart healthcare organizations. Bob, Alice, John, and Paul belong to smart healthcare organization-1, while Susan and Max belong to smart healthcare organization-2. Bob, Alice, Susan, and Max have been diagnosed with Obstructive Sleep Apnea (OSA) disease, while John and Paul are Diabetics (DB). For each patient, the data captured by the IoT devices are sent to a Digital Twins service that builds a Digital Twin (DT), a digital representation of the patient. The clinicians have access to their patients' data through the DTs. The proposed hierarchical framework allows collaboration among multiple health organizations. For instance, Bob and Alice, who are OSA patients of the smart healthcare organization-1, collaborate by exchanging their local weights to build an Anomaly Detection (AD) model. At the same time, they also collaborate with the OSA patients of smart healthcare organization-2 (Susan and Max). This approach allows participants to enhance the recognition rate of their local models by feeding the neural network with new samples, provided by other patients with similar characteristics.

### Paper:

Hierarchical Federated Learning based Anomaly Detection using Digital Twins for Smart Healthcare  
(<https://arxiv.org/pdf/2111.12241.pdf>)



# Federated Learning Research

Series-1

## GOOGLE AI INTRODUCES 'FEDERATED RECONSTRUCTION' FRAMEWORK THAT ENABLES SCALABLE PARTIALLY LOCAL FEDERATED LEARNING

Federated learning is a machine learning technique in which an algorithm is trained across numerous decentralized edge devices or servers, keeping local data samples without being exchanged. This prevents the collecting of personally identifiable information. It is frequently accomplished by learning a single global model for all users, although their data distributions may differ. Due to this variability, an algorithm that can personalize a global model for each user has been developed.

However, privacy concerns may prevent a truly global model from being learned in some cases. While sending user embedding updates to a central server may reveal the preferences encoded in the embeddings, it is required to train a completely global federated model. Even if models do not include user-specific embeddings, having some parameters local to user devices reduces server-client communication and allows for responsible personalization of those parameters for each user.

Google AI introduces an approach that enables scalable partially local federated learning in their work "Federated Reconstruction: Partially Local Federated Learning". Some model parameters are never aggregated on the server in this approach. This strategy trains a piece of the model to be personal for each user while eliminating transmission of these parameters for models other than Matrix Factorization. In the case of matrix factorization, a recommender model is trained. The model retains user embeddings local to each user service.

### **Google AI's scalable approach:**

In large-scale federated learning situations, approaches based on stateful algorithms tend to degrade. Most users do not attend training and those who do most likely do so only once. As a result, this state is rarely available and can turn stale over time. Furthermore, all non-participating users are left with untrained local parameters, preventing practical use.

Federated Reconstruction has been developed to address this problem. Federated Reconstruction is stateless, which means it does not require user devices to maintain local parameters because it reconstructs them as needed. The local parameters can be randomly initialized and trained using gradient descent on local data with the global parameters frozen. Furthermore, updates to global parameters can be calculated with local values frozen.

This primary strategy allows for large-scale training because it does not presuppose users have a state from earlier training rounds. To avoid staleness, local parameters are constantly recreated from scratch. Users not present during the training can obtain trained models and execute inference by simply recreating local parameters from local data. Compared to other approaches, Federated Reconstruction trains better performance models for unseen users.

Federated Reconstruction develops global parameters that enable unseen users to rebuild local parameters quickly and accurately. Federated Reconstruction, in other words, is learning to learn local parameters. This creates a link between Meta-Learning and the rest of the system.



Federated Reconstruction also allows developers to customize models for different users while decreasing model parameter transmission – even for models that do not have user-specific embeddings. When used at a fixed communication level, Federated Reconstruction outperforms other personalization methods.

### **Deployment of Federated Reconstruction in GBoard:**

GBoard is a popular mobile keyboard app that has millions of users. Gboard users use expressions such as stickers and GIFs to connect with others. Since these expressions have a wide range of preferences among users, it is ideal for utilizing matrix factorization to forecast new expressions that a user may want to share.

Federated Reconstruction was used to train a matrix factorization model over user-expression co-occurrences. This ensures that each Gboard user's embeddings are unique. The model was then implemented on GBoard, resulting in a 23.9 percent increase in click-through rate..

### **The Road Ahead:**

Federated Reconstruction is still a topic that is being explored on various levels. According to preliminary findings, Federated Reconstruction enables tailoring to diverse users while decreasing communication of privacy-sensitive factors. Following Google's AI Principles, the research team scaled this technique to Gboard. As a result, several people benefitted from improved recommendations.

Paper: <https://arxiv.org/pdf/2102.03448.pdf>

Github: <https://github.com/google-research/federated/tree/master/reconstruction>

Reference: <https://ai.googleblog.com/2021/12/a-scalable-approach-for-partially-local.html>

## RESEARCHERS PROPOSE 'PROXYFL': A NOVEL DECENTRALIZED FEDERATED LEARNING SCHEME FOR MULTI-INSTITUTIONAL COLLABORATIONS WITHOUT SACRIFICING DATA PRIVACY

Tight rules generally govern data sharing in highly regulated industries like finance and healthcare. Federated learning is a distributed learning system that allows multi-institutional collaborations on decentralized data while also protecting the data privacy of each collaborator. Institutions in these disciplines are unable to aggregate and communicate their data, limiting research and model development progress. More robust and accurate models would result from sharing information between institutions while maintaining individual data privacy.

For example, in the healthcare industry, histopathology has undergone increasing digitization, providing a unique opportunity to improve the objectivity and accuracy of diagnostic interpretations through machine learning. The preparation, fixation, and staining techniques utilized at the preparation site, among other things, cause significant variation in digital photographs of tissue specimens.

Because of this diversity, medical data must be integrated across numerous organizations. On the other hand, medical data centralization involves regulatory constraints as well as workflow and technical challenges, such as managing and distributing the data. Because each histopathology image is often a gigapixel file, often one or more gigabytes in size, the latter is very important in digital pathology.

Researchers from Layer 6 AI, the University of Waterloo, and Vector Institute recently investigated the context of multi-institutional collaboration in highly regulated areas. They developed a strategy for decentralized model training that respects data privacy in a recent study. Distributed machine learning on decentralized data could be a way to address these issues and increase machine learning adoption in healthcare and other highly regulated industries.

Federated learning (FL) is a distributed learning framework that was created to train a model using non-centralized data. It trains a model on client devices directly where data is created, with gradient updates transmitted back to the centralized server for aggregate. However, because it includes a centralized third party controlling a single model, the canonical FL arrangement is unsuitable for multi-institutional collaboration. When considering hospital collaboration, creating a single central model may be undesirable. Each hospital may desire autonomy over its own regulatory compliance and specialty-specific approach.

While it is sometimes claimed that FL improves privacy by ensuring that raw data never leaves the client's device, it does not provide the level of security required by regulated organizations.

Proxy-based federated learning, or ProxyFL, is a decentralized collaboration between institutions that allows for the training of high-performance and robust models without losing data privacy or communication efficiency, according to researchers from the University of Waterloo. They have contributed a method for decentralized FL in multi-institutional collaborations that are adapted to heterogeneous data sources and preserves model autonomy for each participant, as well as the incorporation of DP for rigorous privacy guarantees analysis and the reduction of communication overhead.

The team found many significant issues. Clients may not wish to share the structure and parameters of their private model with others. Disclosing model structure exposes confidential information, increases the danger of adversarial assaults, and exposes private data about local datasets. In addition to model heterogeneity, clients may not wish to entrust the management of a shared model to a third party, preventing centralized model averaging approaches. Information sharing must also be efficient, robust, and peer-to-peer.

To overcome the issues mentioned above, the team adds an additional proxy model with specific settings for each client. It acts as a link between the client and the rest of the world. For compatibility, all clients agree on a standard proxy model architecture as part of the communication protocol. Each client trains its private and proxy models together in each round of ProxyFL so that they can benefit from one another.

The proxy can extract meaningful information from private data through differentially personal training, which can then be shared with other customers without breaking privacy limitations. Then, according to a communication network defined by an adjacency matrix and de-biasing weights, each client transmits its proxy to its out-neighbors and receives new proxies from its in-neighbors. Finally, each client collects all of the proxies they've received and changes their existing proxy.

Because the proxy model is the only entity that a client discloses, each client must verify that this sharing does not jeopardize their data's privacy. Because arbitrary post-processing on a DP-mechanism does not compromise its guarantee, the proxy can be released as long as it was trained using one. On a client-by-client basis, privacy guarantees are tracked. Every customer in a multi-institutional partnership has a responsibility to ensure the privacy of the data it has gathered. As a result, each client keeps track of the parameters for its own proxy model training and can opt-out of the protocol once its privacy budget has been met.

When each approach was trained on a multi-origin real-world dataset, namely The Cancer Genome Atlas, the sub-type classification results for internal and external data under two different DP settings (strong and weak privacy) were reported (TCGA). ProxyFL, FML, and FedAvg were the three FL approaches that were compared.

On the internal test data, ProxyFL outperforms FML and FedAvg in terms of overall accuracy for both privacy settings. All three approaches perform similarly on external test data, with FedAvg marginally ahead when implementing stricter privacy guarantees. Because of the smaller variance in both privacy settings, ProxyFL has considerably better convergence than FML. The FedAvg central model demonstrates no increase in performance when high privacy is implemented across both test datasets. Because they exchange lightweight proxy models rather than larger private models, both ProxyFL and FML are more communication efficient than FedAvg. Still, ProxyFL has the lowest communication overhead due to fewer model exchanges.

## Conclusion

ProxyFL protects data and model privacy while facilitating distributed training through a decentralized and communication-efficient method. Experiments show that ProxyFL is competitive in terms of model correctness, communication efficiency, and privacy preservation when compared to alternative baselines. Furthermore, the method was tested in a real-world scenario with four medical institutions working together to train pan-cancer classification models. Some FL approaches for private knowledge

transfer, unlike ProxyFL, require a public dataset. This distinction is crucial because large-scale public datasets are not readily available in highly controlled domains. Access to a publicly available dataset broadens the range of approaches that may be applied and opens up new avenues for future research.

Paper: <https://arxiv.org/pdf/2111.11343v1.pdf>

Github: <https://github.com/layer6ai-labs/ProxyFL>

## GOOGLE AI IMPROVES THE PERFORMANCE OF SMART TEXT SELECTION MODELS BY USING FEDERATED LEARNING

Smart Text Selection is one of Android's most popular features, assisting users in selecting, copying, and using text by anticipating the desired word or combination of words around a user's tap and expanding the selection appropriately. Selections are automatically extended with this feature, and users are offered an app to open selections with defined classification categories, such as addresses and phone numbers, saving them even more time.

The Google team made efforts to improve the performance of Smart Text Selection by utilizing federated learning to train a neural network model responsible for user interactions while maintaining personal privacy. The research team was able to enhance the model's selection accuracy by up to 20% on some sorts of entities thanks to this effort, which is part of Android's new Private Compute Core safe environment.

The model is trained to only select a single word to reduce the incidence of making multi-word selections in error. The Smart Text Selection feature was first trained on proxy data derived from web pages that had schema.org annotations attached to them. While this method of training on schema.org annotations was effective, it had a number of drawbacks. The data was not at all like the text users viewed on their devices.

With this new release, the model no longer uses proxy data for span prediction and instead employs federated learning to train on-device on real interactions. This is a machine learning model training method in which a central server organizes model training across several devices while the raw data remains on the local device.

The following is how a typical federated learning training process works:

1. The model is initialized first by the server.

2. Then, in an iterative process,

- devices are sampled,
- selected devices improve the model using their local data, and
- only the improved model, not the data used for training, is sent back.

3. The server then takes the average of the modifications and creates the model that is sent out in the following iteration.

For Smart Text Selection, Android receives accurate feedback for what selection span the model should have predicted each time a user taps to choose the text and corrects the model's suggestion. To protect user privacy, the choices are held on the device for a short time without being seen on the server and then utilized to enhance the model using federated learning techniques. This strategy has the advantage of training the model on the same data that it would encounter during inference.

Because raw data is not available to a server, one of the advantages of the federated learning strategy is that it allows for user privacy. Instead, only updated model weights are sent to the server. To empirically

validate that the model was not memorizing sensitive information, the team used methods from Secret Sharer, an analysis approach that assesses the degree to which a model mistakenly memorizes its training data. Furthermore, data masking techniques were also used to prevent the model from ever seeing certain types of sensitive data.

Initial attempts to use federated learning to train the model were unsuccessful. The loss did not converge, and the predictions were all over the place. Because the training data was collected on-device rather than centrally, debugging the process was impossible because it could not be examined or confirmed. To get around this problem, the research team built a set of high-level indicators to see how the model fared throughout training. Among the metrics employed were training examples, selection accuracy, and recall and precision measures for each object type.

Smart Text Selection may now be scaled to many more languages thanks to this new federated technique. This should ideally work without the need for human system tuning, allowing even low-resource languages to be supported, making lives easier for billions of users around the world.

Reference: <https://ai.googleblog.com/2021/11/predicting-text-selections-with.html>

## **NVIDIA OPEN-SOURCE 'FLARE' (FEDERATED LEARNING APPLICATION RUNTIME ENVIRONMENT), PROVIDING A COMMON COMPUTING FOUNDATION FOR FEDERATED LEARNING**

Standard machine learning methods involve storing training data on a single machine or in a data center. Federated learning is a privacy-preserving technique that is especially useful when the training data is sparse, confidential, or less diverse.

NVIDIA open-source NVIDIA FLARE, which stands for Federated Learning Application Runtime Environment. It is a software development kit that enables remote parties to collaborate for developing more generalizable AI models. NVIDIA FLARE is the underlying engine in the NVIDIA Clara Train's federated learning software, which has been utilized for diverse AI applications such as medical imaging, genetic analysis, cancer, and COVID-19 research.

Researchers can use the SDK to customize their method for domain-specific applications by choosing from a variety of federated learning architectures. NVIDIA FLARE can also be used by platform developers to give consumers the distributed infrastructure needed to create a multi-party collaborative application.

Participants in federated learning collaborate to train or evaluate AI models without needing to share or pool their private datasets. NVIDIA FLARE supports a variety of distributed architectures, including peer-to-peer, cyclic, and server-client techniques, among others.

NVIDIA FLARE has been used in two federated learning collaborations:

1. NVIDIA collaborated with Roche Digital Pathology researchers on a successful internal simulation using whole slide images for classification
2. It also worked with Erasmus Medical Center in the Netherlands on an AI application identifying genetic variants associated with schizophrenia cases.

However, the server-client architecture is not appropriate for all federated learning applications. NVIDIA FLARE will make federated learning more accessible to a broader range of applications by supporting other architectures. The possible use cases include helping:

- energy corporations analyze seismic and wellbore data,
- Manufacturers optimize industrial processes
- Financial firms improve fraud detection algorithms

The ability to speed federated learning research by open-sourcing NVIDIA FLARE is especially relevant in the healthcare sector, where access to multi-institutional datasets is critical, but patient privacy concerns can hinder data sharing.

NVIDIA FLARE can work with current AI projects, such as the open-source MONAI medical imaging platform. It will also be deployed in the following areas to power federated learning solutions:

1. The American College of Radiography (ACR) has collaborated with NVIDIA on federated learning

research that uses artificial intelligence to radiology images for breast cancer and COVID-19 applications. It wants to make NVIDIA FLARE available through the ACR AI-LAB, a software platform used by society's tens of thousands of members.

2. Flywheel's Flywheel Exchange platform allows users to access and exchange data and biomedical research techniques, manage federated projects for training and research and select their chosen federated learning solution, including NVIDIA FLARE.

3. Taiwan Web Service Corporation offers a GPU-powered MLOps platform that allows users to use NVIDIA FLARE to execute federated learning.

4. The NVIDIA Inception program partner, Rhino Health, has integrated NVIDIA FLARE into its federated learning solution, assisting researchers at Massachusetts General Hospital in developing an AI model that more accurately diagnoses brain aneurysms. In addition, it helps experts at the National Cancer Institute's Early Detection Research Network in developing and validating medical imaging AI models that detect early signs of pancreatic cancer.

By making NVIDIA FLARE open source, researchers and platform developers will have additional options to personalize their federated learning solutions, enabling cutting-edge AI in practically any industry.

Reference: <https://blogs.nvidia.com/blog/2021/11/29/federated-learning-ai-nvidia-flare/>

Platform: <https://developer.nvidia.com/flare>



## ERICSSON AND UPPSALA UNIVERSITY TEAM UP TO RESEARCH AIR QUALITY PREDICTION USING MACHINE LEARNING AND FEDERATED LEARNING

Statistical methods have recently been applied in various sectors, spanning from health care to customer relationship management, to analyze and forecast the behavior of a given event. The goal here is to evaluate the likelihood of an event occurring rather than predict the exact outcome. However, the path is not without bumps; getting access to the data needed to deploy machine learning algorithms is difficult for the following reasons:

- Volume: Transferring such information might be very costly due to network resource constraints.
- Privacy: The data obtained may be sensitive regarding privacy; any procedure that has access to such data is exposed to personal details belonging to distinct individuals.
- Legislation: Data regarding a country's residents cannot be moved outside the country for legal reasons in several countries.

Predictive models, in general, require large amounts of data to perform effectively. Large data sets are expensive to store, and transferring them would significantly strain the network. The only way to solve this problem is to devise a mechanism that allows predictive models to be trained in their raw form without requiring data transfer.

Ericsson is seeking to tackle this issue in partnership with Uppsala University in Sweden. This time it is 'Air Quality Prediction.' The negative consequences of low air quality are well known, and developing a system to predict air quality would be a significant accomplishment. The results can change behavior at all levels, from individual behavior through communities, nations, and even global.

The researchers aspire to create prediction tools that can help figure out what steps can be taken ahead of time to enhance air quality and protect vulnerable groups from its consequences.

The standard strategy for training supervised machine learning models is to deal with centralized data aggregating massive amounts of data at each station. This, however, necessitates the transfer and compilation of vast amounts of raw data. The project's purpose is to move away from the use of centralized data. The researchers looked at federated learning, which allows for a machine learning model to be taught at each station and then federated averaging to merge the models.

This project's scope envisioned a decentralized configuration consisting of many air quality stations, each collecting data for a specific area. They have the processing power to construct a predictive model using data obtained locally and interact with air quality stations elsewhere.

Such a setup does not exist yet; hence measurements obtained by the Swedish Meteorological and Hydrological Institute were used to mimic it (SMHI). The data was divided by weather stations (Stockholm E4/E20 Lilla Essingen, Stockholm Sveavägen 59, Stockholm Hornsgatan 108, and Stockholm Torkel Knutssonsgatan). Although it was a centralized dataset, it resulted in the training of four separate models, which were then combined using federated averaging.

A baseline for comparison is usually needed when validating results. To validate against the federated models in this situation, a high-performing centralized model was created. The same dataset was

investigated using a variety of characteristics and machine learning model architectures.

The models were evaluated based on their accuracy as they were tested simultaneously. The Symmetric Mean Absolute Percentage Error (SMAPE) and Mean Absolute Error (MAE) were employed to conduct the analysis. The researchers could cover a wide range of scenarios and arrive at a high-performing centralized model with these characteristics.

### **RESULT:**

The machine learning model that was trained received ten input features as input. Various models were used to anticipate the next 1, 6, and 24 hours, such as Long Short-Term Memory Networks (LSTM) and Deep Neural Networks (DNNs).

In the centralized scenario, models aimed at predicting the next hour outperformed those aimed at predicting the next day on average. SMAPE scores varied from 0.282 to 0.5214, and MAE scores from 0.22 to 0.47.

In the federated model case, almost similar MAE scores were observed, indicating that decentralized training techniques like federated learning might support the decentralized setup that was initially envisioned.

Techniques like federated learning can help to make the world a more sustainable place to live. They not only make the process of training a machine learning model and managing its lifespan easier, but they also improve the quality of people's lives by predicting air quality. More federated learning and other techniques that contribute to this goal are expected to be used.

Github: <https://github.com/EricssonResearch/damp>

References: <https://www.ericsson.com/en/blog/2021/11/air-quality-prediction-using-machine-learning>

# Federated Learning Research

Series-1

**TO ADVERTISE ON MARKTECHPOST**  
**CONTACT: ASIF@MARKTECHPOST.COM**



Marktechpost, LLC.  
1968 S. Coast Hwy, #1675  
Laguna Beach, CA, 92651  
USA



Asif@marktechpost.com



www.marktechpost.com

# INTRODUCTION TO FEDERATED LEARNING

Large volumes of data are required for training machine learning models. The trained model is run on a cloud server that users can access through various applications such as web search, translation, text production, and picture processing, which is the standard procedure for establishing machine learning applications.

The application must transfer the user's data to the server where the machine learning model is stored every time it wishes to use it, creating privacy, security, and processing issues.

Fortunately, developments in edge AI have allowed sensitive user data to be avoided from being sent to application servers. This current area of study, also known as TinyML, aims to construct machine learning models that fit smartphones and other consumer devices, making on-device inference possible. Even if the device is not connected to the internet, these applications can continue functioning. The on-device inference is more energy-efficient in many applications than transferring data to the cloud.

However, the data is still required to train the models installed on customers' devices. When the entity generating the models already owns the data (e.g., a bank owns its transactions) or the data is public information, this isn't a problem (e.g., Wikipedia or news articles). But, acquiring training data for machine learning models that leverage confidential user information such as emails, chat logs, or personal images poses numerous obstacles.

## Federated Learning

Federated Learning (FL) allows mobile phones to develop a shared prediction model cooperatively while retaining all of the training data on the device, effectively divorcing machine learning from the requirement to store data in the cloud. Bringing model training to the device extends beyond using local models that make predictions on mobile devices.

The foundation for federated learning is a machine learning model in the cloud server. This model has either been trained on a publicly available dataset or has not been trained at all.

Several user devices offer to train the model to the next level. User data relevant to the model's application is stored on these devices, such as chat logs and keystrokes. These devices download the base model while connected to a power outlet and on a wi-fi network (training is a compute-intensive operation that will drain the device's battery if done at an inopportune time). They then use the device's local data to train the model.

After training, the trained model is sent back to the server. The training data is no longer required for inference once the models have been trained to encapsulate the statistical patterns of the data in numerical parameters. As a result, no raw user data is included when the device sends the trained model back to the server. The server updates the basic model with the aggregate parameter values of user-trained models after receiving data from user devices.

The federated learning cycle is repeated numerous times before the model achieves the developers' desired degree of accuracy. When the final model is complete, it can be shared for providing the

on-device interface.

Federated Learning enables smarter models, lower latency, and lower energy use while maintaining privacy. In addition to giving an update to the shared model, the enhanced model on the phone can be used right away, powering experiences tailored to the way users use their phones.

This way, federated learning could revolutionize the way AI models are trained in almost all sectors such as healthcare, automobiles, IoT and FinTech, to name a few.

### **Federated Learning In Healthcare**

The healthcare industry has grown increasingly regulated with the 1996 implementation of HIPAA (Health Insurance Portability and Accountability Act). Organizations have found it extremely challenging to implement new technologies due to the scale and complexity of healthcare regulations. As a result, the healthcare industry's shortage of resources is evident.

FL has the potential to bring AI to the point of care, allowing vast volumes of heterogeneous data from several organizations to be integrated into the model building while adhering to local clinical data regulations.

Larger hospital networks would be better able to collaborate and profit from secure, cross-institutional data access. In addition, smaller community and rural hospitals would benefit from expert-level AI algorithms.

Clinicians would get access to more strong AI algorithms based on data from larger demography of patients for a certain clinical area or from uncommon situations that they would not have encountered locally. They'd also be able to contribute to the algorithm's ongoing training anytime they disagreed with the results.

Thanks to a secure way to learn from more diverse algorithms, healthcare firms could quickly bring cutting-edge inventions to the market. Meanwhile, rather than relying on the restricted supply of free datasets, research institutes would be able to focus their efforts on actual clinical requirements based on a wide range of real-world data.

### **Federated Learning In FinTech**

Whether it's mobile banking, payment apps, or Fintech in general, data security has become a critical concern in the Internet world. Businesses that rely on FinTech confront a number of challenges. These concerns include obtaining clearance and lawful consent, data preservation, and the time and expense of gathering and transporting data across networks.

FL is a distributed and encrypted machine learning method that enables cooperative machine learning training on decentralized data without the need for data transmission between participants. It provides solutions for FinTech, for instance, by looking for data breaches and ATO (Account Takeover) Fraud. It can also analyze credit scores and comprehend a user's digital footprint to prevent fraudulent actions KYC without having to send data to the cloud.

Therefore, FL makes it possible for Fintech to mitigate risks. It develops fresh and inventive techniques for its customers and enterprises and establishes better trust between the two parties.

### **Federated Learning For Autonomous Vehicles**

Furthermore, with real-time data and predictions, federated learning can give a better and safer self-driving car experience. Autonomous vehicles require real-time traffic data and continuous training for better real-time decision making. All of these goals can be met with federated learning, which allows the models to improve over time with input from other vehicles.

Federated Learning necessitates the adoption of new tools and a new way of thinking by machine learning practitioners: model building, training, and evaluation without direct access to or labelling raw data, with communication costs as a limiting factor. For scalability, data protection, and a variety of other reasons, companies like Google, NVIDIA, and other research groups are experimenting with federated learning. The following articles present major highlights of recent advances in machine learning and artificial intelligence when applied with federated learning.

#### References:

- <https://ai.googleblog.com/2017/04/federated-learning-collaborative.html>
- <https://www.analyticsvidhya.com/blog/2021/05/federated-learning-a-beginners-guide/>
- <https://venturebeat.com/2021/08/13/what-is-federated-learning/>
- <https://federated.withgoogle.com/>
- <https://odsc.medium.com/what-is-federated-learning-99c7fc9bc4f5>
- <https://blogs.nvidia.com/blog/2019/10/13/what-is-federated-learning/>
- <https://www.hitechnectar.com/blogs/applications-of-federated-learning/>
- <https://research.aimultiple.com/federated-learning/>