Aprendizaje federado aplicado a la ciberseguridad. Retos y oportunidades

José Luis Martínez
Universidad de Castilla–La Mancha

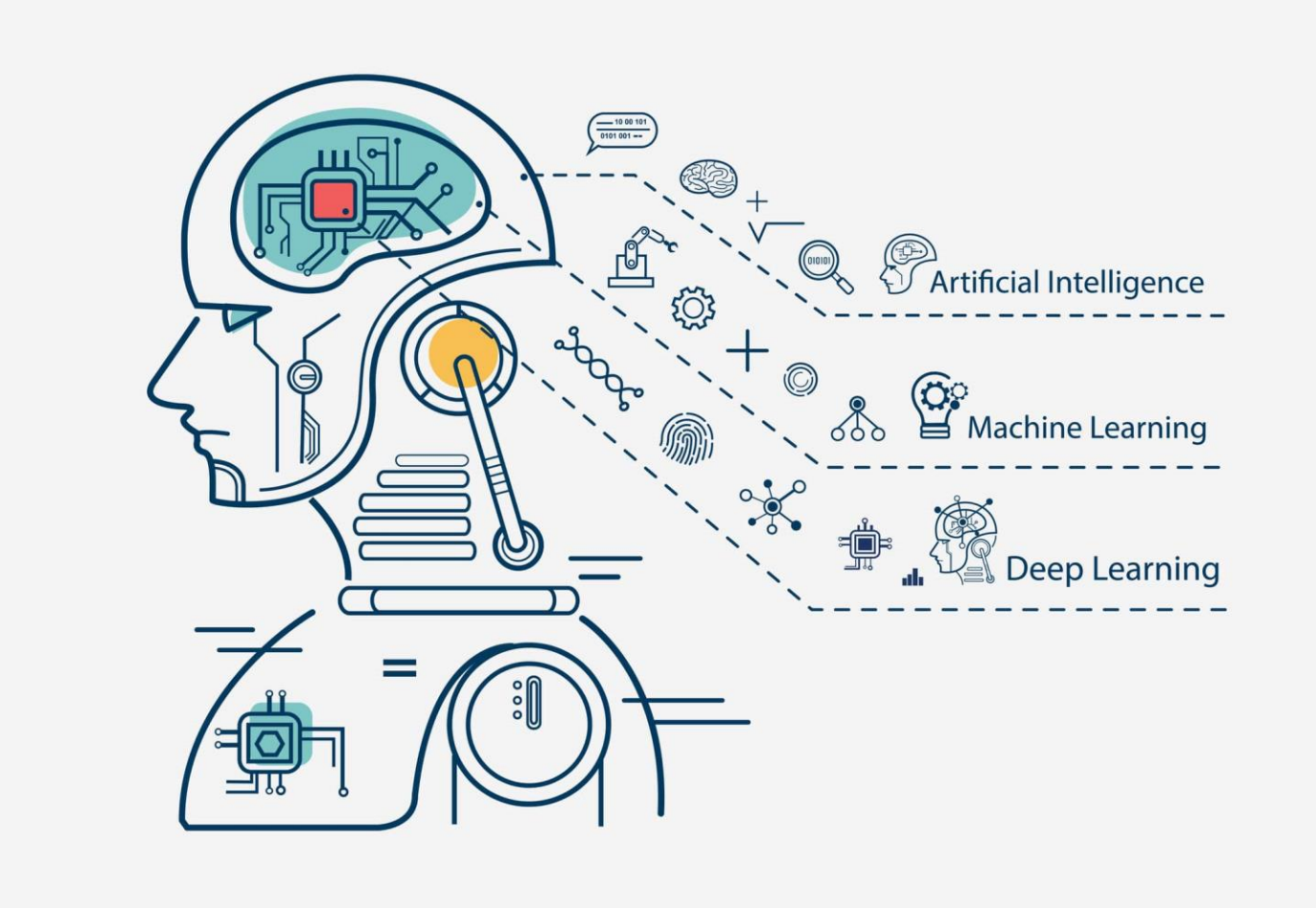
Breve presentación

- Afiliación & Formación:
  - Catedrático de Universidad en la UCLM
  - Ingeniero Informático y Doctor en Informática en el 2005 y 2009 por la UCLM
- Perfil Docente:
  - 15 años de experiencia docente
  - Director Master en Formación Permanente en Ciberseguridad
- Perfil Investigador:
  - Investigación en Malware, amenazas y forense en IoT
  - Perito forense judicial y Consultor + formador
- Certificaciones:
  - Instructor CCNA Security y CCNP Switch CISCO
  - CSNA de Stormshield
  - Instructor Certified Ethical Hacking (CEH) de EC Council
  - Instructor Computer Hacking Forensic Investigator (CHFI) de EC Council
- Linkedin:
  - https://www.linkedin.com/in/josé-luis-martínez-martínez-060679120

Índice

- Presentación
- Introducción & Motivación & Escenario
- Federated Learning
- Retos
- Retos relacionados con la seguridad
- Direcciones Futuras
- Conclusiones


Data Breach
Cyber Attack
Protection Failed
System Safety Compromised

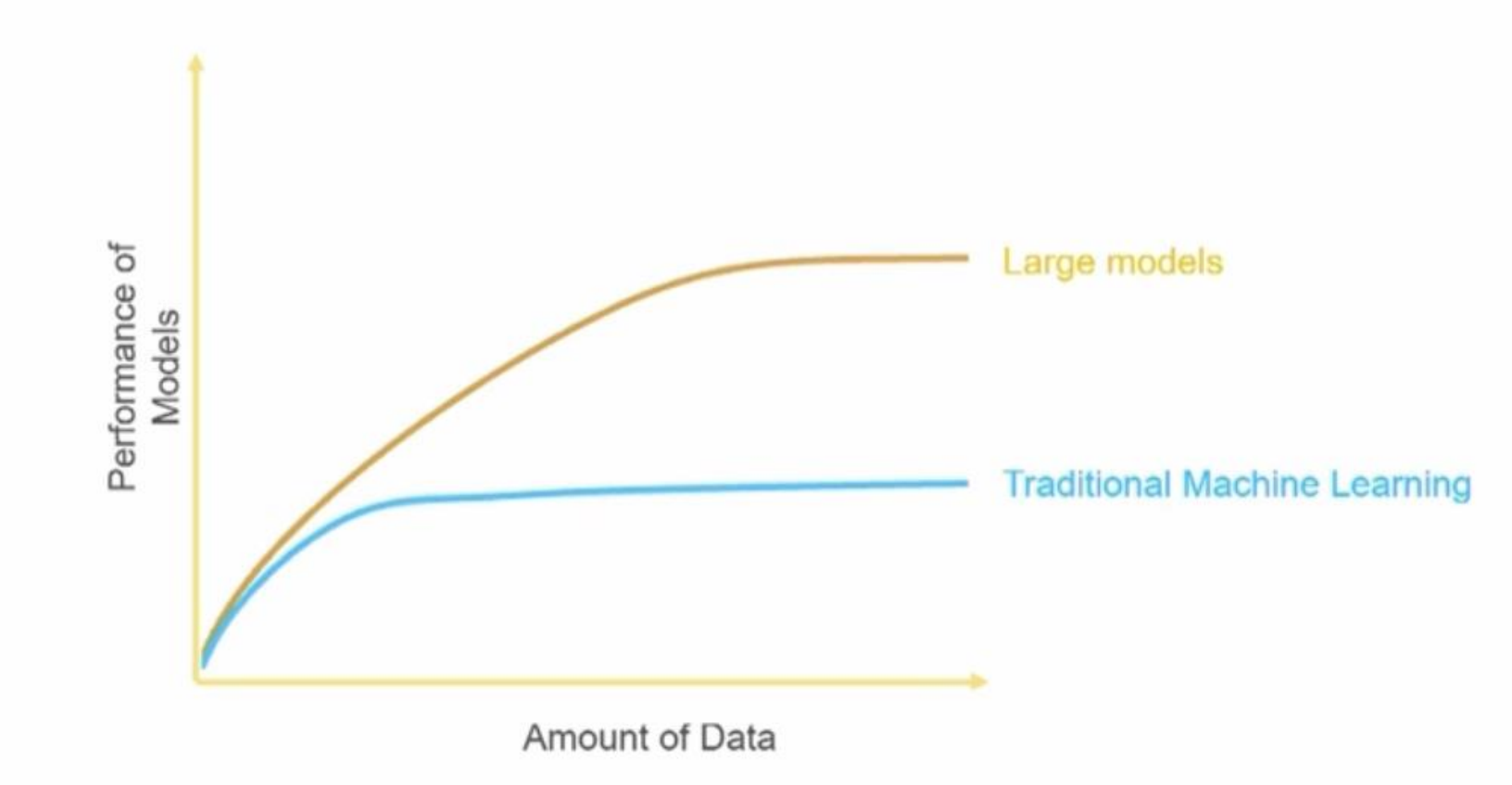


Fuente:https://atriainnovation.com/blog/el-machine-learning-en-la-industria/

Fuente: https://www.sevensquaretech.com/best-llm-models/

FEDERATED MACHINE LEARNING

Motivación

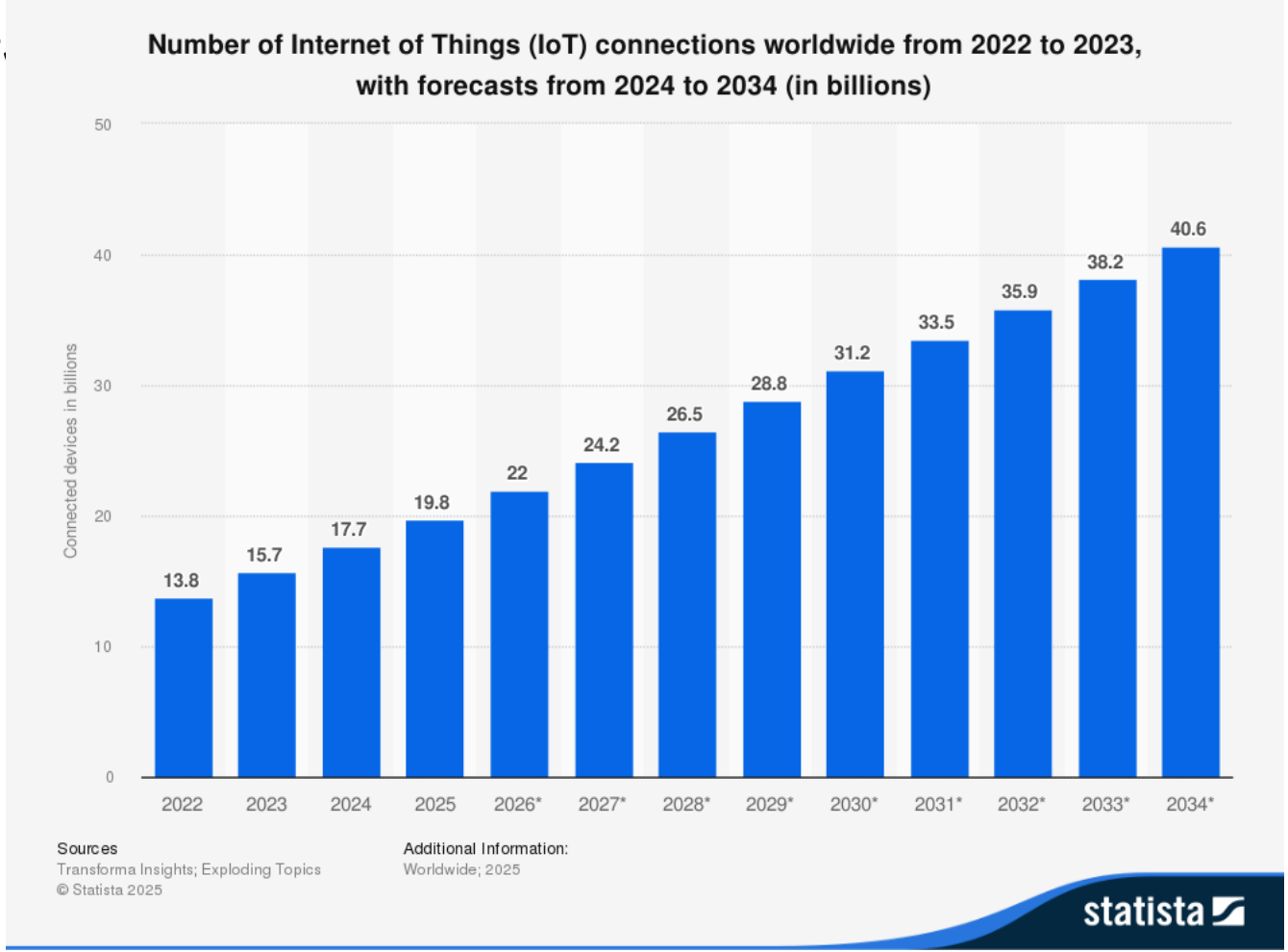

Performance of Models
Large models
Traditional Machine Learning
Amount of Data

Motivación

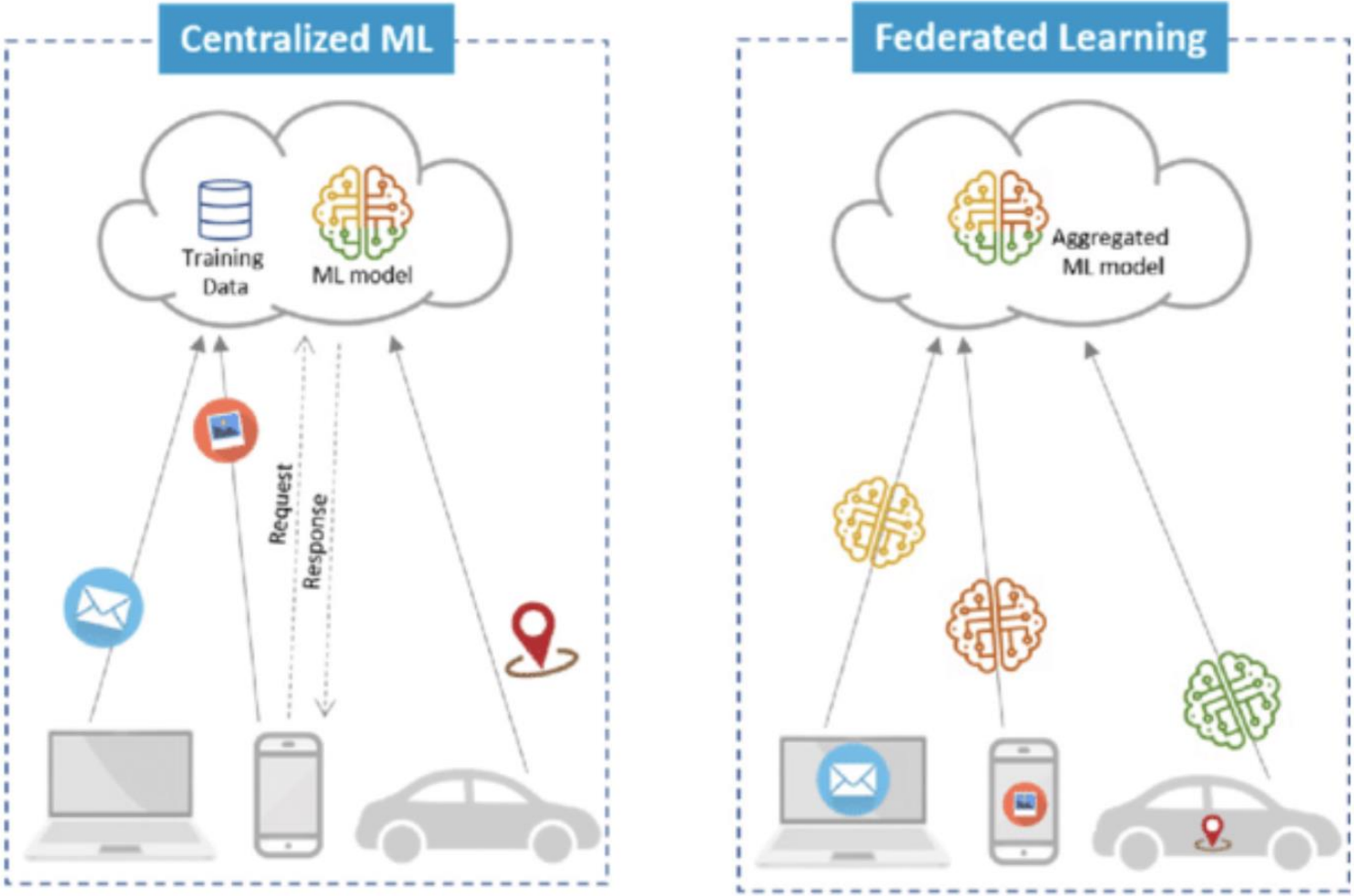
- Los datos….
- Problema-1:
  - …están protegidos legalmente (HIPAA, GDPR).
  - …son confidenciales.
  - …son demasiado valiosos para compartirlos.
- Problema-2:
  - …son demasiado grandes para transmitirlos.
  - …es ineficiente transmitirlos en bruto en algunos contextos….

Health Insurance Portability and Accountability Act

Moti



Fuente: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/

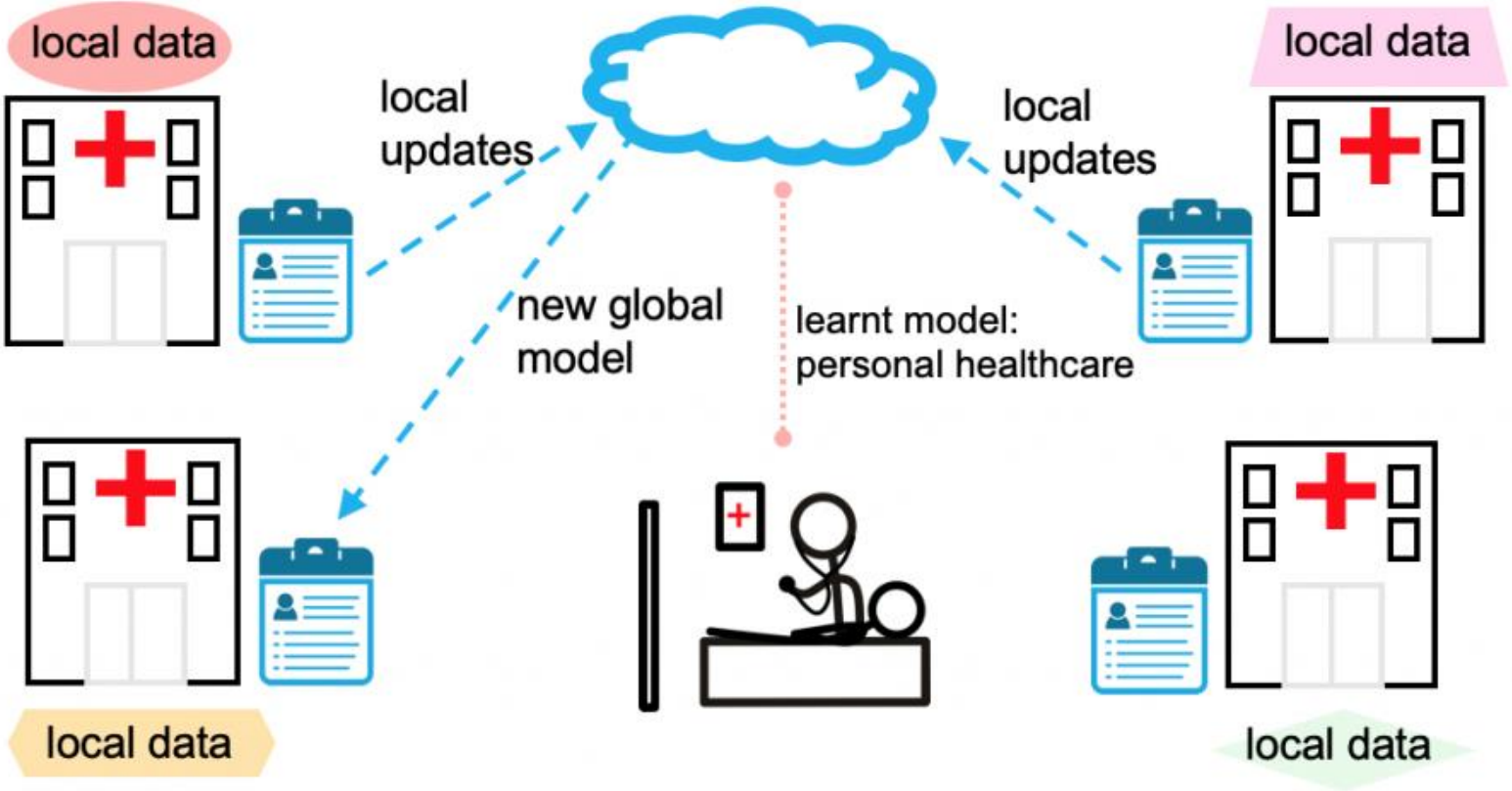

Centralized ML
Training Data
ML model
Request
Response
Federated Learning
Aggregated ML model

Motivación

- La adquisición de los datos es distribuida
  - Privacidad de éstos
  - Normativas y reglamentos
  - Geolocalización
- Eficiencia energética en la transmisión
- Eficiencia en el uso del ancho de banda

[REF] Abdulrahman, Sawsan & Tout, Hanine & Ould-Slimane, Hakima & Mourad, Azzam & Talhi, Chamseddine & Guizani, Mohsen. (2020). A Survey on Federated Learning: The Journey From Centralized to Distributed On-Site Learning and Beyond. IEEE Internet of Things Journal. PP. 10.1109/JIOT.2020.3030072.

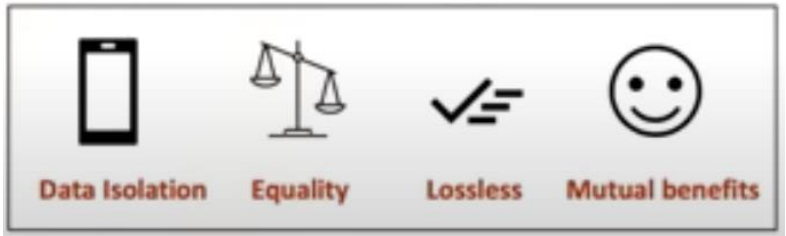
Escenario

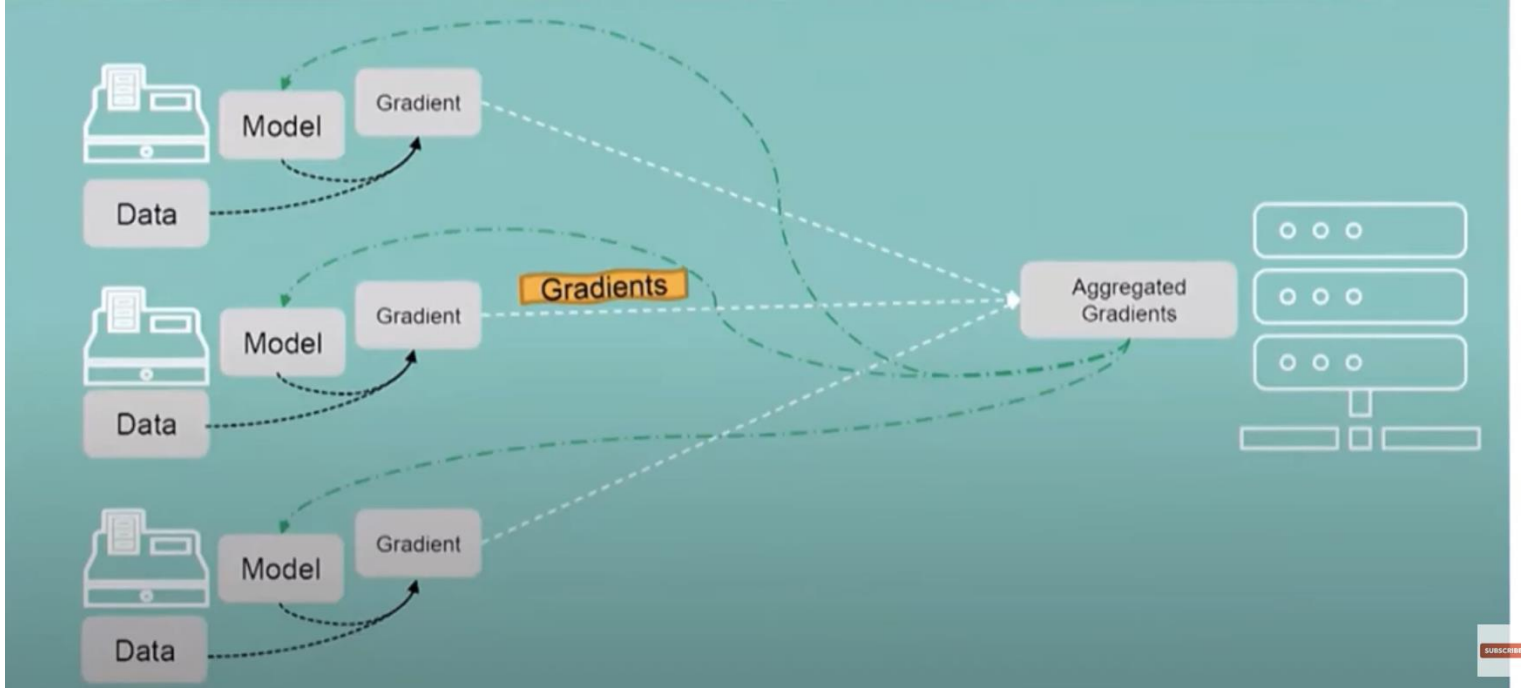


Fuente: https://blog.ml.cmu.edu/2019/11/12/federated-learning-challenges-methods-and-future-directions/

Federated Learning

- Datos son recogidos desde diferentes fuentes y se entrena un modelo ML/DL local.
  - Igualdad: Todos los participantes contribuyen
- Cada nodo NO tiene acceso a los datos de otra fuente
  - Sin pérdidas. El rendimiento es el mismo si todos los participantes enviaran los datos
- Cada nodo **transmite los parámetros de su modelo entrenado** al servidor central
  - Aislamiento de datos. **Datos en crudo no son transmitidos**
- El servidor central actualiza un modelo general con los datos recibidos y devuelve a cada nodo el modelo actualizado
  - Beneficio Mutuo: Todos los nodos se benefician de la federación

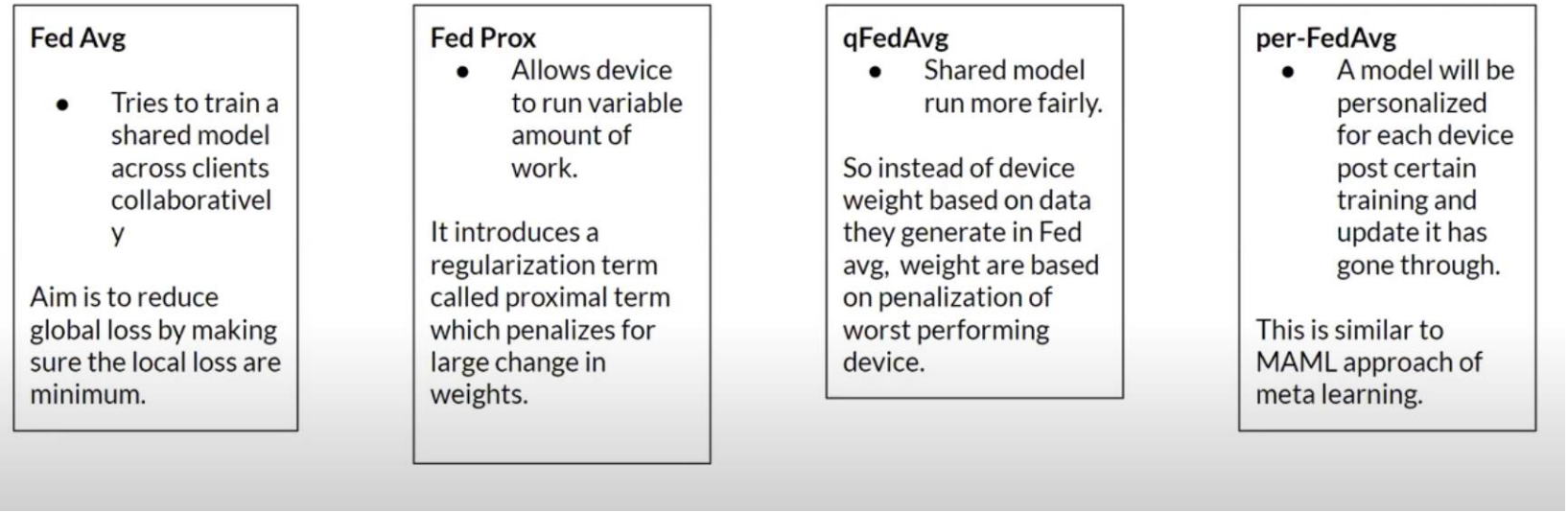




- El modelo es la red neuronal, regresor, clasificador, etc. que se entrena de forma colaborativa entre los clientes de la federación
- Los pesos son los valores actuales del modelo.
- El gradiente es la dirección y magnitud de cambio que indican cómo ajustar los parámetros del modelo para reducir la función de pérdida.

Federated Learning

- Modelos de agregación.
  - Las actualizaciones del modelo son enviadas de vuelta al servidor central y se actualiza el modelo general a partir de los modelos locales

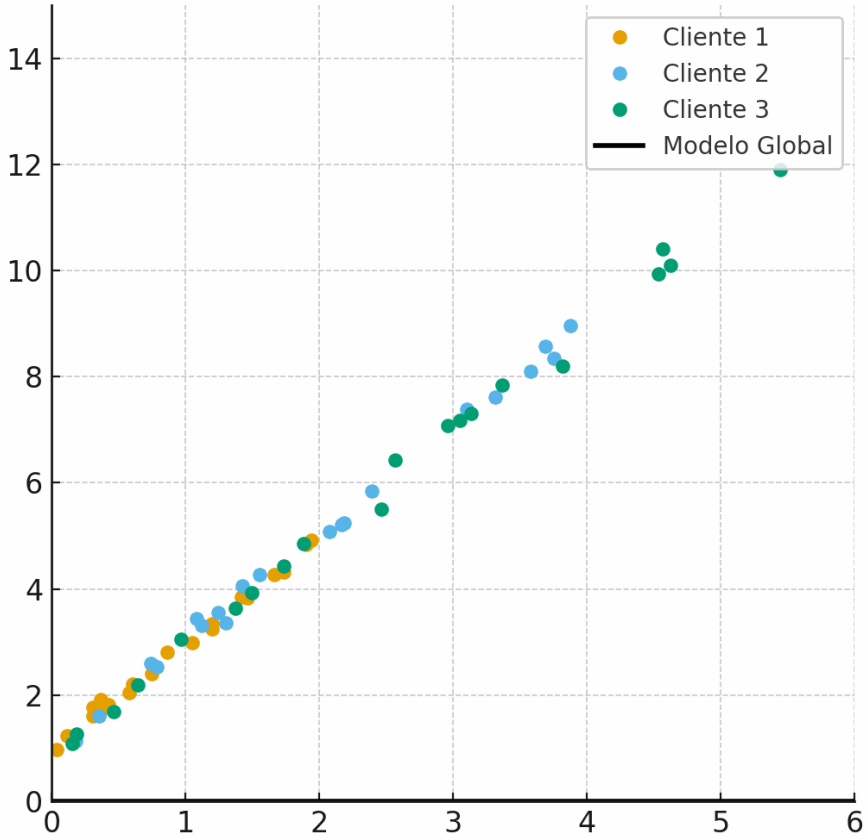


Federated Learning

- Supongamos que queremos entrenar un modelo muy simple:
  - $y = w \cdot x$
  - donde w es el parámetro que queremos aprender
- Hay 3 clientes (nodos) que entrenan localmente:
  - Los clientes tienen datos
    - (x,y)={(1,2),(2,4),….}
- El servidor inicializa el modelo con un peso arbitrario, digamos:
  - $w_0 = 0$
- Cada cliente ajusta localmente el parámetro w con sus datos usando, por ejemplo, una sola iteración de descenso de gradiente (SGD).
  - La función de pérdida (MSE) para cada cliente es:
    - $L(w)=1/N \sum (y-w \cdot x)^2$

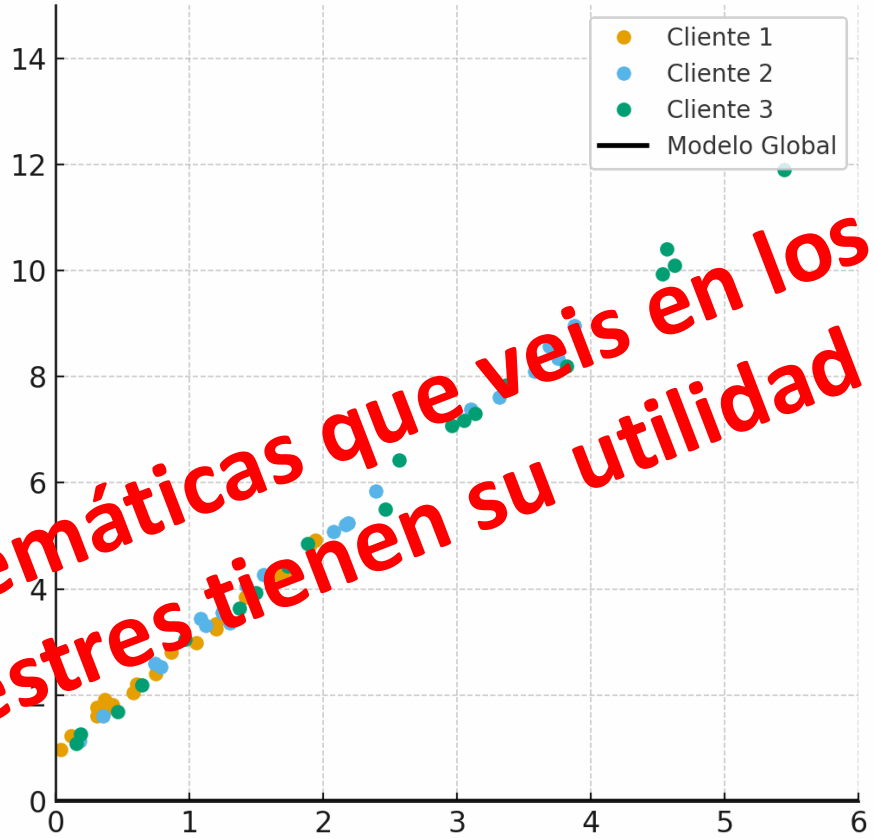
Federate

- Supongamos c e:
  - $y = w \cdot x$
  - donde w es el
- Hay 3 clientes
  - Los clientes tie
    - (x,y)={(1,2),(2,
- El servidor inic nos:
  - $w_0 = 0$
- Cada cliente a s usando, por ejemplo, una s .
  - La función de
    - L(w)=1/N ∑ (y-



Federate

- Supongamos c e:
  - $y = w \cdot x$
  - donde w es el
- Hay 3 clientes
  - Los clientes tie
    - (x,y)={(1,2),(2,
- El servidor inic nos:
  - $w_0 = 0$
- Cada cliente a s usando, por ejemplo, una s .
  - La función de
    - L(w)=1/N ∑ (y-



Las matemáticas que veis en los primeros semestres tienen su utilidad practica

Federated Learning

- Diferencias entre FL y *distributed learning*
  - Nodos son heterogéneos (multimodal) en términos de capacidad de cómputo, ancho de banda, potencia, autonomía, memoria, etc
  - El coste de comunicación es mayor que el de computación
  - Los datos almacenados en cada nodo son **no-IID** (independiente e idénticamente distribuidos)
  - La cantidad de datos no es balanceada

[REF] McMahan, H. B., Eider Moore, Daniel Ramage, Seth Hampson and Blaise Agüera y Arcas. “Communication-Efficient Learning of Deep Networks from Decentralized Data.” International Conference on Artificial Intelligence and Statistics (2016).

Federated Learning

- Beneficios:
  - Poca Latencia
  - Buena predicción de modelos
  - Bajo Consumo
  - Privacidad
- Aplicaciones:
  - Monitor de tráfico
  - Temperatura / monitor
  - Recomendaciones de localizaciones
  - Monitor de salud
  - Actividad
  - Coches autónomos
  - Plantas manufacturadas

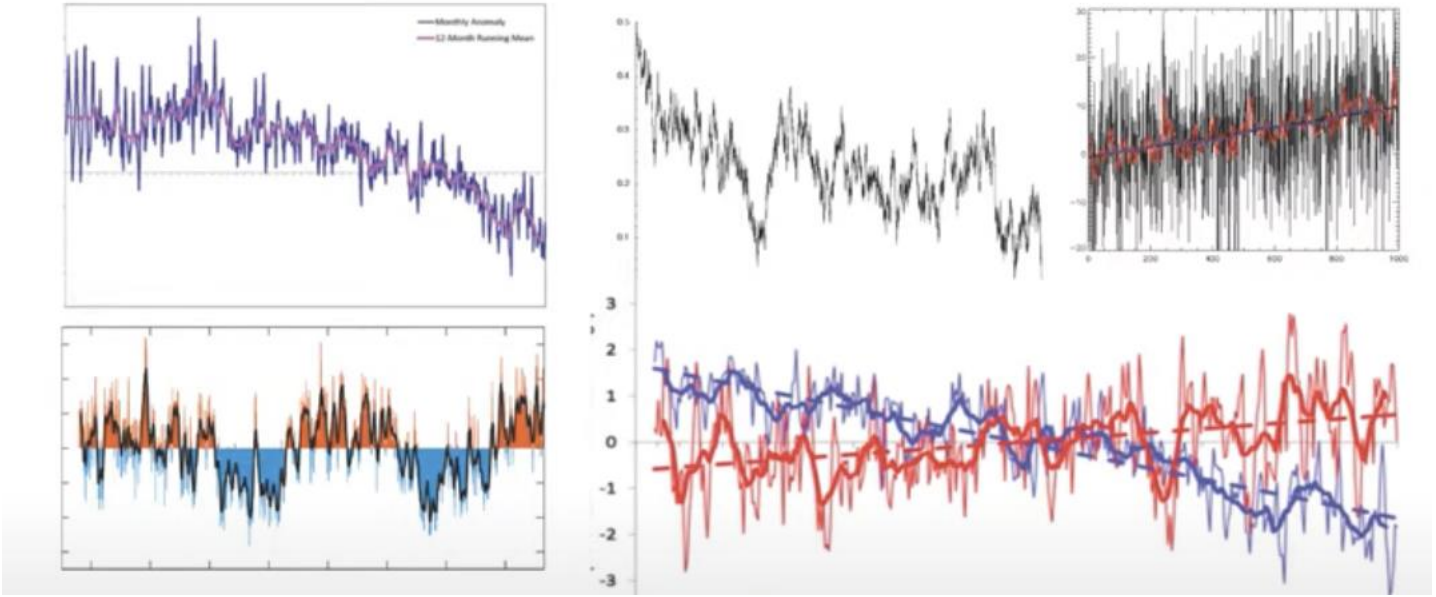
¿pero esto no era un congreso de ciberseguridad?

Federated Learning

[REF] Pinto, Rui P., Bruno MC Silva, and Pedro RM Inácio. "Federated learning for anomaly detection on Internet of Medical Things: A survey." Internet of Things (2025): 101677.

- **Anomaly Detection for Cyber Security using Federated Learning**

La idea es en una ristra de datos y eventos y el objetivo es encontrar anomalías en ellos



Federated Learnir

SPRINGER NATURE Link

Find a journal | Publish with us | Track your research

Home > The Journal of Supercomputing > Article

A TabPFN-based intrusion de for the industrial internet of t

Open access | Published: 30 May 2024

Volume 80, pages 20080–20117, (2024) Cite this article

Download PDF

You have full access to this open access artic

Sergio Ruiz-Villafranca, José Roldán-Gómez, Juan Manuel Castelo Gó & José Luis Martinez

View PDF | Download full issue

ELSEVIER

Future Generation Computer Systems

Volume 166, May 2025, 107707

WFE-Tab: Overcoming limitations of TabPFN in IIoT-MEC environments with a weighted fusion ensemble-TabPFN model for improved IDS performance

Sergio Ruiz-Villafranca [a], José Roldán-Gómez [b 1], Javier Carrillo-Mondéjar [b 1], José Luis Martinez [a 1], Carlos H. Gañán [c 1]

Show more

Add to Mendeley | Share | Cite

https://doi.org/10.1016/j.future.2025.107707

Get rights and content

Retos Generales

[REF] Wang, Cong, Bin Hu, and Hongyi Wu. "Energy minimization for federated asynchronous learning on battery-powered mobile devices via application co-running." *2022 IEEE 42nd international conference on distributed computing systems (ICDCS)*. IEEE, 2022.

- Maximizar autonomías de las baterías
  - Dispositivos limitados en memoria y procesamiento
- Desigual data frecuency - no-iid - *imbalance*
  - Heterogeneidad de datos: los datos no siguen la misma distribución entre clientes y la cantidad y frecuencia de datos disponibles varía entre clientes.
- Tiempo de convergencia del modelo

[REF] Wang, Hongyi, et al. "Federated learning with matched averaging." arXiv preprint arXiv:2002.06440 (2020).

Strategies
for improving communication efficiency." arXiv preprint arXiv:1610.05492 (2016).

- Modelo
  - Algori

14:45-15:00, Paper We-S3-T13.4

Adaptive Federated Learning-Based Architecture for Intrusion Detection in IoT/IIoT Environments

García-Sáez, Luis Miguel — University of Castilla-La Mancha
Ruiz-Villafranca, Sergio — University of Castilla-La Mancha
Roldán-Gómez, José — University of Zaragoza
Carrillo-Mondéjar, Javier — University of Zaragoza
Martínez, José Luis — University of Castilla-La Mancha

Keywords: AIoT, Cloud, IoT, and Robotics Integration, Machine Learning

Abstract: The rapid expansion and growth of Internet of Things (IoT) and Industrial Internet of Things (IIoT) environments has led to an increase in the number of attacks and risks in these environments. This presents new cybersecurity challenges that require more advanced intrusion detection systems (IDS). However, IDS based on centralised Machine Learning (ML) face problems of scalability, latency, and privacy. In this context, Federated Learning (FL) offers a decentralised approach that allows multiple nodes to train models collaboratively without exposing sensitive data. This work presents a federated IDS tailored for IoT/IIoT environments and introduces FedWLA, an aggregation strategy that dynamically weights updates according to the quality and uncertainty of local data. The proposed architecture is evaluated through different IoT/IIoT traffic datasets orientated to cybersecurity and widely used in these environments. It shows comparable and even superior performance to centralised methods, with an average F1-Score ranging between 0.98 - 0.99 for the tests performed. Moreover, the proposed FedWLA strategy consistently outperforms other federated aggregation approaches, such as FedAvg and FedProx, particularly in heterogeneous scenarios. These results demonstrate the capability and potential of FL in intrusion detection, effectively leveraging the scalability and privacy advantages it offers.

Retos Generales

[REF] Konečný, Jakub, et al. "Federated learning: Strategies for improving communication efficiency." arXiv preprint arXiv:1610.05492 (2016).

- Modelos de agregación federada
  - Algoritmos clásicos de ML, federarlos (ML, deeplearning, etc)
- Reducir comunicación: cuantificando o comprimiendo
- Derivados de conector IoT a internet…objeto de los atacantes

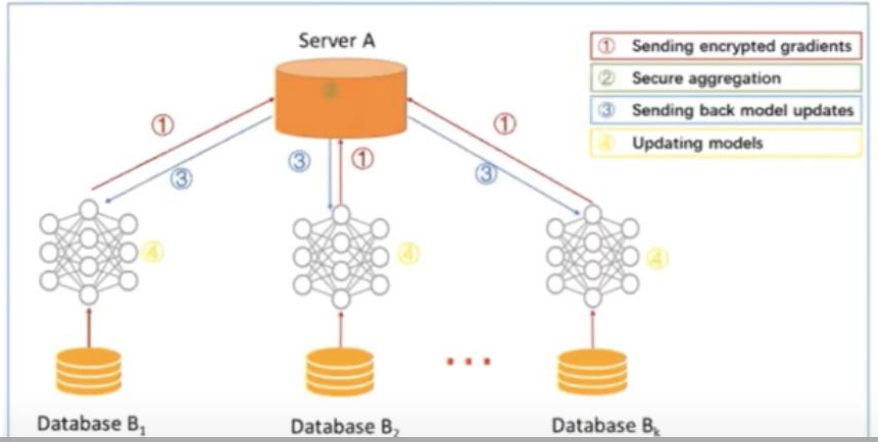
[REF] M. Hao, H. Li, X. Luo, G. Xu, H. Yang and S. Liu, "Efficient and Privacy-Enhanced Federated Learning for Industrial Artificial Intelligence," in IEEE Transactions on Industrial Informatics, vol. 16, no. 10, pp. 6532-6542, Oct. 2020,

Retos relacionados con la seguridad

- Ataques de inferencia
- Un nodo puede *leakear* información sensible del modelo asociado a sus datos
- Posible solución:
  - Cifrado homomorfico
  - Differencial privacy

[REF] Behnia, Rouzbeh, et al. "Efficient secure aggregation for privacy-preserving federated machine learning." 2024 Annual Computer Security Applications Conference (ACSAC). IEEE, 2024.

- Cifrado Homomórfico
  - permite realizar operaciones directamente sobre datos cifrados.
  - El resultado de esas operaciones es igual a si las hubieras hecho sobre los datos en claro.



[REF] Yang, Zhaoxiong et al. “FPGA-Based Hardware Accelerator of Homomorphic Encryption for Efficient Federated Learning.” ArXiv abs/2007.10560 (2020)

- Research Directions:
  - Eficiencia computacional. El cifrado es mucho más lento que la computación en claro.
    - Re-diseñar algoritmos más ligeros y que haga uso hardware especializado (GPU, FPGA, ASIC)
  - Tamaño de los datos cifrados. Los cifrados homomórficos generan mensajes mucho más grandes
    - Mejorar los esquemas de compresión y reducir la sobrecarga de los ciphertexts

[REF] Jin, Weizhao, et al. "FedML-HE: An efficient homomorphic-encryption-based privacy-preserving federated learning system." arXiv preprint arXiv:2303.10837 (2023)

Retos relacionados con la seguridad

[REF] Correia, Pedro, et al. "Federated Learning: An approach with Hybrid Homomorphic Encryption." *arXiv preprint arXiv:2509.03427* (2025)

- El cifrado soporta bien sumas y multiplicaciones, pero no tanto operaciones no lineales (sigmoide, ReLU, softmax), lo que introduce pérdida de precisión.
  - Desarrollar técnicas que permitan más variedad de operaciones sobre datos cifrados sin perder exactitud.

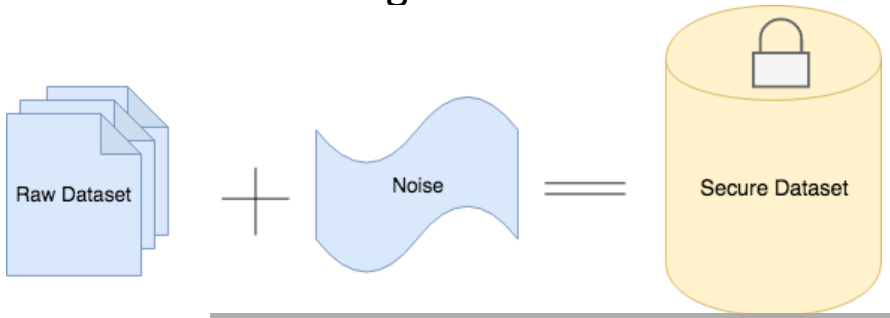
- Profundidad de los circuitos. Estos sistemas tienen un límite en la cantidad de operaciones encadenadas (profundidad). Hay que “renovar” el cifrado (bootstrapping)
  - Reducir o eliminar la necesidad de bootstrapping para hacer viable entrenar modelos profundos.

[REF] Brutzkus, Alon, Ran Gilad-Bachrach, and Oren Elisha. "Low latency privacy preserving inference." *International Conference on Machine Learning*. PMLR, 2019.

[REF] Dowlin, Nathan et al. “CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy.” International Conference on Machine Learning (2016).

Retos relacionados con la seguridad

- Differential Privacy
  - Asegura que los datos de un cliente no se puedan reconstruir ni identificar a partir de lo que comparte.
  - Con Differential Privacy, se protege esa información aplicando ruido matemático controlado a las actualizaciones.
  - Compatible con otras técnicas de seguridad como cifrado homomórfico o secure aggregation.



[REF] McMahan, H. B., Eider Moore, Daniel Ramage, Seth Hampson and Blaise Agüera y Arcas. “Communication-Efficient Learning of Deep Networks from Decentralized Data.” International Conference on Artificial Intelligence and Statistics (2016).

Retos relacionad

[REF] Fu, Jie, Zhili Chen, and Xiao Han. "Adap dp-fl: Differentially private federated learning with adaptive noise." 2022 IEEE international conference on trust, security and privacy in computing and communications (TrustCom). IEEE, 2022.

- Retos a resolver en Differential Privacy
  - Precisión Vs. Privacidad. Agregar ruido degrada la utilidad del modelo.
    - diseñar mecanismos de ruido más inteligentes (ej. dependientes de la sensibilidad de cada cliente o capa)

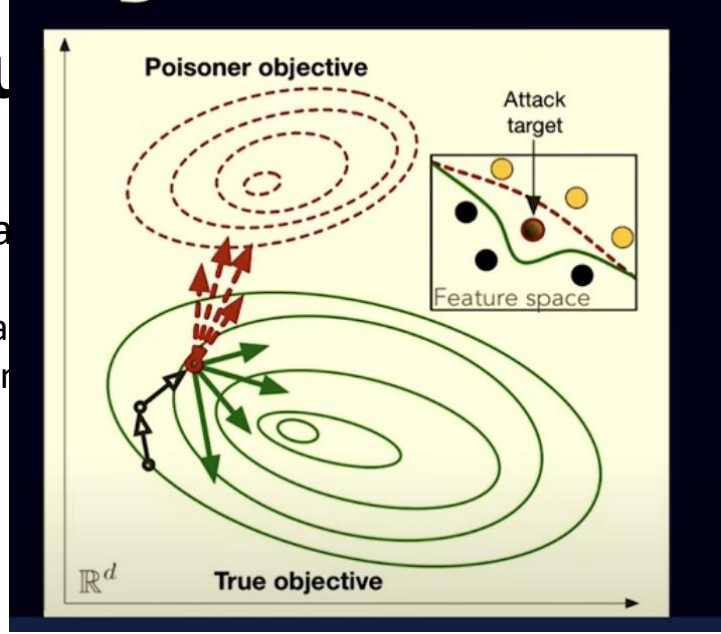
[REF] Debao Wang, Shaopeng Guan, FedFR-ADP: Adaptive differential privacy with feedback regulation for robust model performance in federated learning,Information Fusion,Volume 116,2025,

  - Heterogeneidad de datos y clientes. El ruido tiene un efecto desigual según el volumen de datos y la distribución de datos (non-IID).
    - Crear mecanismos de DP que se adapten dinámicamente al tamaño y distribución de los datos por cliente.
  - Combinación con otros mecanismos de privacidad como encriptación homomorfica
    - Estimar si el ruido se añade antes o después de la agregación / cifrado

[REF] Sébert, Arnaud Grivet, et al. "Protecting data from all parties: Combining FHE and DP in federated learning." arXiv preprint arXiv:2205.04330 (2022).

Retos relacionados con la segu

- Model Poisonnng
- Un atacante puede manipular esos gradientes o pesos para
  - Degradar la precisión global
  - Backdoor attack: definir trigger: clasificar siempre como “A” cua
  - Sesgar el modelo: hacer que favorezca ciertas salidas (ej. mar para promocionar un producto concreto).



- Solución:
  - Agregación robusta: propuestas que reducen el impacto de actualizaciones anómalas.
    - los outliers se descartan, y el modelo global se actualiza con las contribuciones fiables.
  - Análisis de anomalías: detectar clientes que envían gradientes sospechosos.
  - Differential Privacy combinada con robust aggregation: añade ruido y evita que un cliente malicioso domine la actualización.

Retos relacionados con la seguridad

[REF] Li, Youpeng, et al. "FedCAP: Robust Federated Learning via Customized Aggregation and Personalization." 2024 Annual Computer Security Applications Conference (ACSAC). IEEE, 2024.

- Retos a resolver en la agregación robusta:
  - Escalabilidad y coste computacional. No escala bien cuando hay miles o millones de dispositivos.
  - Heterogeneidad de datos (non-IID). Una actualización legítima puede parecer un ataque si el cliente tiene datos muy distintos.
    - distinguir entre “cliente malicioso” y “cliente honesto con datos muy heterogéneos”.

[REF] Lian, Z., Zhang, C., Nan, K., Su, C.. SPoiL: Sybil-Based Untargeted Data Poisoning Attacks in Federated Learning. NSS 2023. Lecture Notes in Computer Science, vol 13983. Springer.

  - Ataques adaptativos y colusión. Los ataques no son aislados; si varios clientes maliciosos colaboran (Sybil o collusion attacks), pueden enviar actualizaciones coordinadas que engañen incluso a métodos robustos.
    - diseñar mecanismos resistentes a ataques coordinados y dinámicos.
  - Compatibilidad con técnicas de privacidad (DP, cifrado homomorfico, secure aggregation).
    - Problema: Se requieren ver las actualizaciones en claro.
    - hacer agregación robusta sobre datos cifrados o privados, sin comprometer la seguridad ni la privacidad.

[REF] Lian, Z., Zhang, C., Nan, K., Su, C.. SPoiL: Sybil-Based Untargeted Data Poisoning Attacks in Federated Learning. NSS 2023. Lecture Notes in Computer Science, vol 13983. Springer.

Retos relacionados con la seguridad

- Ataques de tipo Advesarial network
- Varios clientes distintos (legítimos o comprometidos) cooperan entre sí de forma coordinada para manipular el modelo global.
- Sybil / Collusion / Byzantine
  - Sybyl: Un atacante crea múltiples **identidades falsas** para aumentar su peso de influencia en la agregación global
  - Collusion: varios clientes se ponen de acuerdo (**pueden ser reales o Sybils**) para engañar al sistema.
  - Byzantine: comportamiento arbitrario y malicioso de un cliente, sin necesidad de coordinarse ni multiplicarse.

- Objetivos:
  - Sesgar el modelo global hacia un comportamiento concreto (data poisoning).
  - Insertar un backdoor (ejemplo: que el modelo reconozca una clase incorrectamente bajo cierta condición).
  - Bloquear o ralentizar el entrenamiento

Retos relacionados con la seguridad

- Sc

Home > ACM Journals > ACM Transactions on Internet Technology > Just Accepted > Poisoning-Resilient Federated Learning for MEC-IoT Environments Using Blockchain

RESEARCH-ARTICLE | FREE ACCESS

Just Accepted

Poisoning-Resilient Federated Learning for MEC-IoT Environments Using Blockchain

Authors: Luis Miguel García-Sáez, Sergio Ruiz-Villafranca, José Roldán-Gómez, Javier Carrillo-Mondéjar, José Luis Martínez Authors Info & Claims

ACM Transactions on Internet Technology • Accepted on 08 September 2025 • https://doi.org/10.1145/3767740

Online AM: 13 September 2025 Publication History

Check for updates

Direcciones futuras…

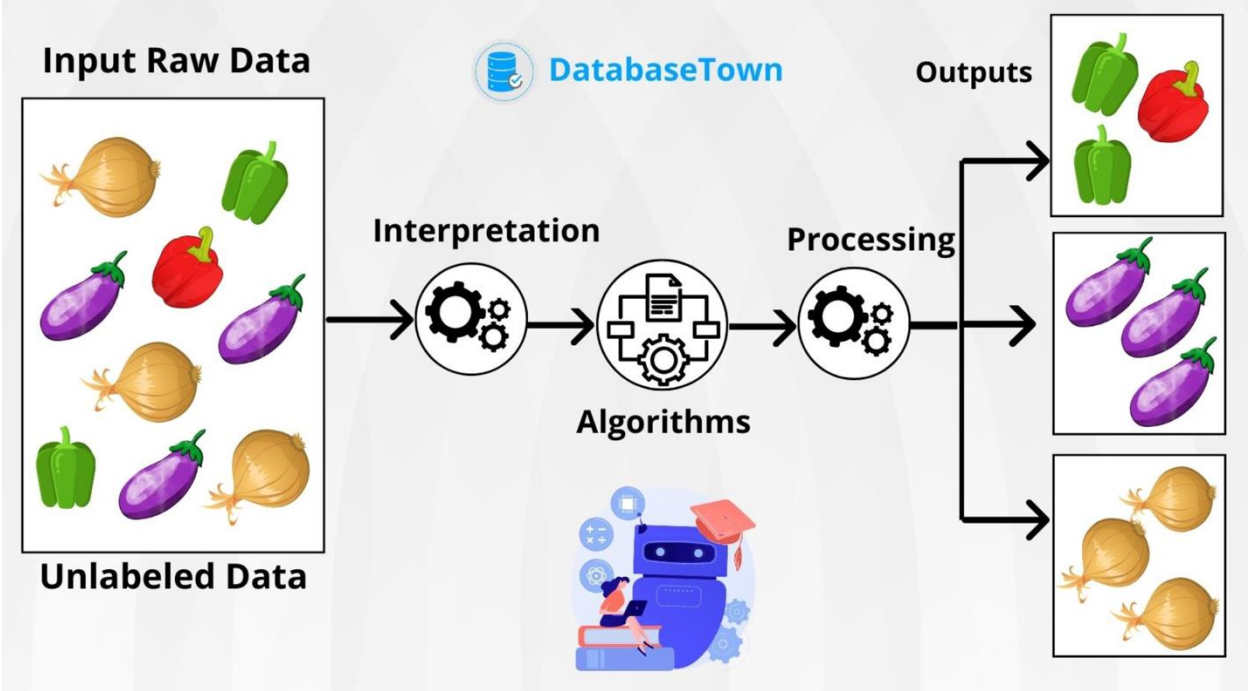
- Adaptative Learning per client
  - Cada cliente tiene datos y capacidades distintas.
  - La idea es ajustar dinámicamente el entrenamiento local a las características de cada cliente en lugar de aplicar la misma estrategia a todos.
    - Adaptive Learning Rate. Ajustar la tasa de aprendizaje local según la calidad de los datos
    - Adaptive Local Epochs**.** Decidir cuántas épocas entrena cada cliente.
    - Algunos clientes participan en todas las rondas, otros solo de vez en cuando (según disponibilidad o energía).

[REF] Kim, Junhyung Lyle, et al. "Adaptive federated learning with auto-tuned clients." arXiv preprint arXiv:2306.11201 (2023).

Direcciones futuras...

[REF] Anwar, Ahmed, et al. "Fedad-bench: A unified benchmark for federated unsupervised anomaly detection in tabular data." 2024 2nd International Conference on Federated Learning Technologies and Applications (FLTA). IEEE, 2024.

- Aprendizaje Federado **No Supervisado**
  - No hay clase. Técnicas:
    - Clustering
    - Autoencoders
    - Detección de anomalías
    - Reducción de dimensión



Direcciones futuras…

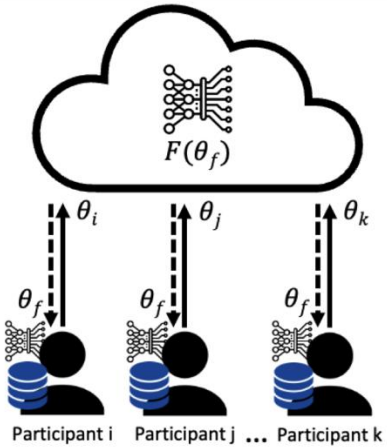
- Aprendizaje Federado **No Supervisado**
  - Retos:
    - Heterogeneidad de datos (non-IID y desequilibrio)
      - Cada cliente puede tener distribuciones muy diferentes de datos sin etiquetar.
      - Los espacios latentes aprendidos pueden no estar alineados entre clientes
    - Resiliencia a ataques. En ausencia de etiquetas, detectar clientes maliciosos (Byzantine, Sybil, poisoning) es mucho más difícil.
    - Integrar differential privacy, homomorphic encryption, secure aggregation sin añadir demasiado coste computacional o pérdida de utilidad

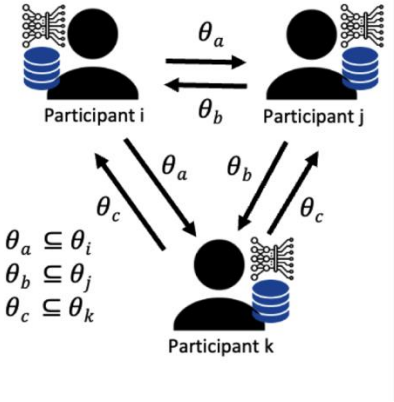
[REF] Li, Yiwei, et al. "Differentially private federated clustering over non-IID data." IEEE Internet of Things Journal 11.4 (2023): 6705-6721.

Direcciones futuras.

[REF] Yuan, Liangqi, et al. "Decentralized federated learning: A survey and perspective." IEEE Internet of Things Journal 11.21 (2024): 34617-34638.

- Aprendizaje Federado **No centralizado**
  - Los clientes peer-to-peer
  - Modelo de agregación global emerge a través de consensos distribuidos
- Principal Ventaja:
  - **Mayor robustez:** no hay un único punto de fallo (ni de ataque).
  - **Privacidad más fuerte:** menos riesgo de fuga de información
  - Mejoras en escalabilidad y latencia





Direcciones futuras…

- Retos:
  - Sincronización y convergencia: garantizar que el modelo converge
  - Topología dinámica: los nodos pueden entrar o salir de la red, complicando el consenso.
  - Coste de comunicación: el P2P puede generar más mensajes
  - Seguridad
    - Ataques Sybil o colusión pueden ser más peligrosos si no hay un servidor que verifique.
    - Agregación robusta descentralizada: diseñar mecanismos resistentes a nodos maliciosos sin servidor central.

[REF] Bornstein, Marco, et al. "SWIFT: Rapid decentralized federated learning via wait-free model communication." arXiv preprint arXiv:2210.14026 (2022).

[REF] Liu, Yuze et al. “Towards Heterogeneity-Aware and Energy-Efficient Topology Optimization for Decentralized Federated Learning in Edge Environment.” *ArXiv* abs/2508.08278 (2025): n. pag.

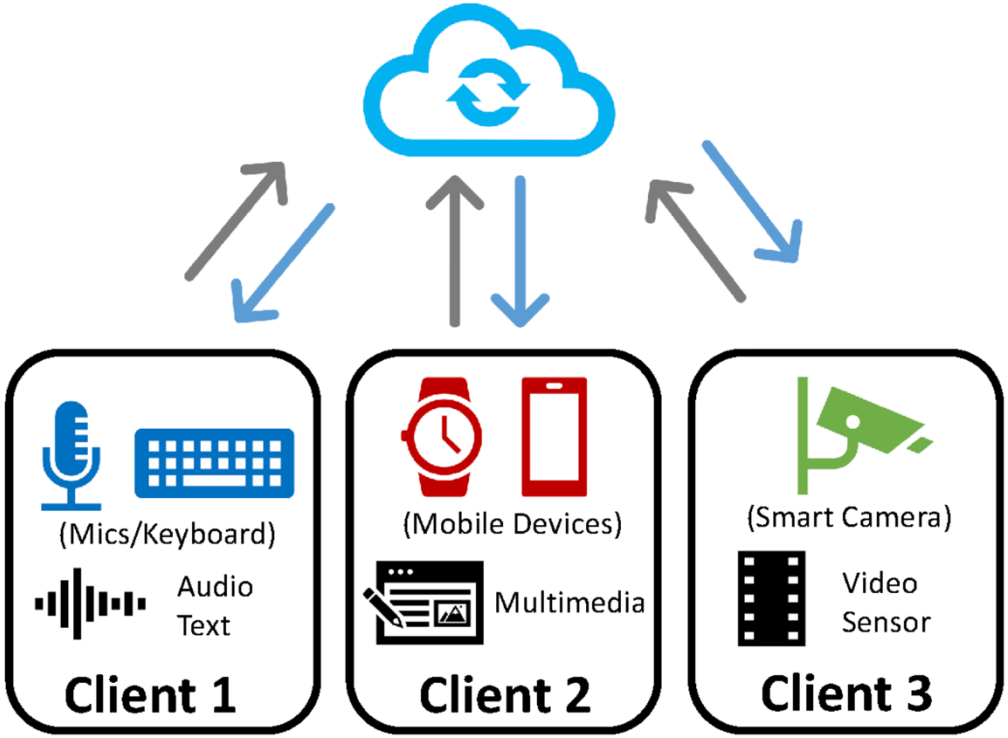
[REF] Wang, He, and Yuejie Chi. "Communication-efficient federated optimization over semi-decentralized networks." *IEEE Transactions on Signal and Information Processing over Networks* (2025).

[REF] Feng, Chao, et al. "Sentinel: An aggregation function to secure decentralized federated learning." *arXiv preprint arXiv:2310.08097* (2023).

Direcciones futuras

[REF] Yongheng Deng, Ningxin He, Xinyi Li, Fan Wu, Yaoxue Zhang, and Ju Ren. 2025. Cross-Modal Federated Learning among Unimodal Devices. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 9, 3, Article 78 (September 2025), 26 pages.

- Multimodal
  - Múltiples tipos de datos (modalidades)
  - Retos:
    - Heterogeneidad de las Modalidades (datos y clientes)
    - Incongruencia de datos
    - Falta de Modalidades.
      - Esto genera un problema de datos incompletos en el entrenamiento



[REF] Bao, Guangyin, et al. "Multimodal federated learning with missing modality via prototype mask and contrast." *arXiv preprint arXiv:2312.13508* (2023).

Conclusiones

- Aprendizaje Federado
- Aplicado a la detección de amenazas
- Retos a resolver:
  - Federado
  - Ataques a los sistemas federados
  - Direcciones futuras

https://github.com/joseluispplu/TallerFederatedLearning

Muchas Gracias

Aprendizaje federado aplicado a la ciberseguridad. Retos y oportunidades

José Luis Martínez
Universidad de Castilla–La Mancha