Aprendizaje distribuido para mejorar los procesos industriales manteniendo la privacidad de los datos

M.A. Melgarejo Aragón[a,*], A. Pérez Núñez (PhD)[a], A. Casas García-Minguillán (PhD)[a]

DOI: 10.17979/ja-cea.2024.45.10976

Introducción

El proyecto FUTCAN, coordinado por el Centro Tecnológico CTC, desarrolla una metodología de aprendizaje distribuido para mejorar los procesos industriales manteniendo la privacidad de los datos. Esta metodología se basa en el Aprendizaje Federado (AF), permitiendo entrenar modelos de inteligencia artificial (IA) utilizando datos de múltiples clientes sin necesidad de compartir información confidencial. En lugar de intercambiar datos sensibles, los clientes solo comparten los pesos de los modelos de IA, asegurando que la información sensible permanezca segura dentro del dominio de cada cliente.

Esta estrategia no solo optimiza el rendimiento de los modelos de manera individual, sino que también mejora la inteligencia artificial en el sector industrial en su conjunto. El estudio incluye simulaciones con arquitecturas distribuidas utilizando datasets bien conocidos, como EMNIST y CIFAR-10. Los resultados muestran que este enfoque es prometedor para entornos industriales, permitiendo aprovechar grandes cantidades de datos sin comprometer la seguridad.

Desafíos actuales:

- Reglamento General de Protección de Datos (GDPR) [1].
- EU AI Act: primera regulación de la IA [2].
- Redes neuronales cada vez más complejas en el estado del arte [3].

En este estudio se propone un mediador en el servidor central de la arquitectura de AF para asegurar la compatibilidad y calidad de los datos mediante la supervisión de anomalías/divergencia de los modelos de IA compartidos y el seguimiento de regulaciones y procedimientos como las normas ISO.

Metodología

Algoritmo 1. Arquitectura de aprendizaje federado con servidor central.

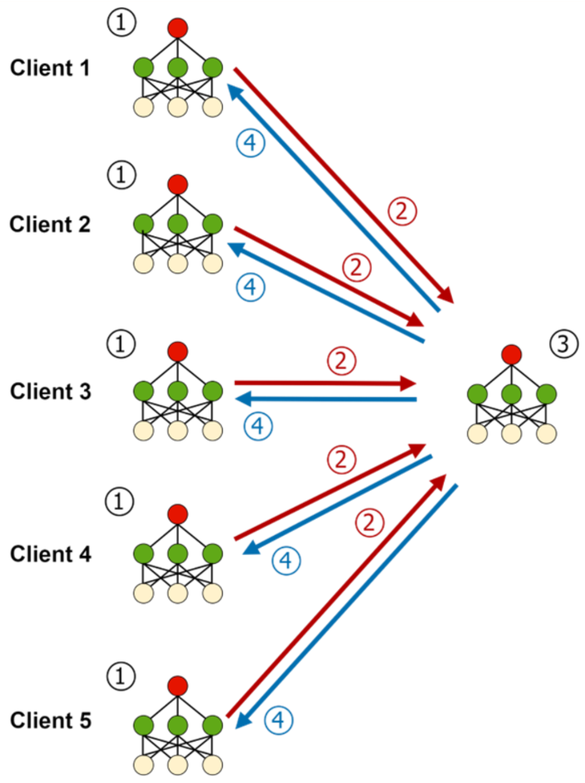
Input: A dictionary *data_dict* containing the data for all clients with train/test splits, the pre-trained model *model*, the number of communication rounds *n_times*.
Output: Final aggregated model weights *save_weights*.
Create a list save_weights to store the weights from all rounds.
Create a list n_client containing the number of training data points per client using *data_dict*.
Create a parameter *initial_weights* with the *model*'s initial weights.
FOR each round from 1 to *n_times*:
Create an empty list *weights_round* to store the weights from the clients in the current round.
FOR each client *i* in *data_dict*:
Assign *initial_weights* to *model*.
Retrieve the training data of client *i*, *X_train* and *y_train*.
Train *model* with *X_train* and *y_train*.
Obtain the trained weights *weights_client* from *model*.
Store the weights *weights_client* of the client *i* in *weights_round*.
END FOR
Create *avg_weights* with the function *ave_weights(n_client, weights_round).*
Update *initial_weights* with *avg_weights* for the next round.
Store the aggregated model weights *avg_weights* in *save_weights*.
END FOR

Algoritmo 2. Función, *ave_weights,* de agregación de los pesos de los modelos de cada cliente.

Input: A list *n_client* containing the number of training data points per client, and another list *weights_round* with the weights from each client.
Output: Aggregated model weights *ave_weights*.
Initialize global weights *ave_weights* with null values.
i = 0
FOR each *n* in *n_client*:
Create *rec_weight* with the weights *weights_round*[*i*] from the client.
Scale *rec_weight* according to *n*.
Normalize *rec_weight* by the total sum of *n_client*.
Iteratively accumulate *rec_weight* into *ave_weights* by summing element by element.
i = *i* + 1
END FOR

Agradecimientos

Este material forma parte del proyecto: "*Posicionamiento Estratégico de Especialización Inteligente y Sostenible para las Factorías del FUTuro de CANtabria* (**FUTCAN**)". Esta entidad ha recibido una ayuda cofinanciada por el Fondo Europeo de Desarrollo Regional a través del Programa Operativo **FEDER 2021-2027** de Cantabria por medio de la línea de subvenciones "Ayudas a proyectos de investigación con alto potencial industrial de agentes tecnológicos de excelencia para la competitividad industrial **TCNIC**".



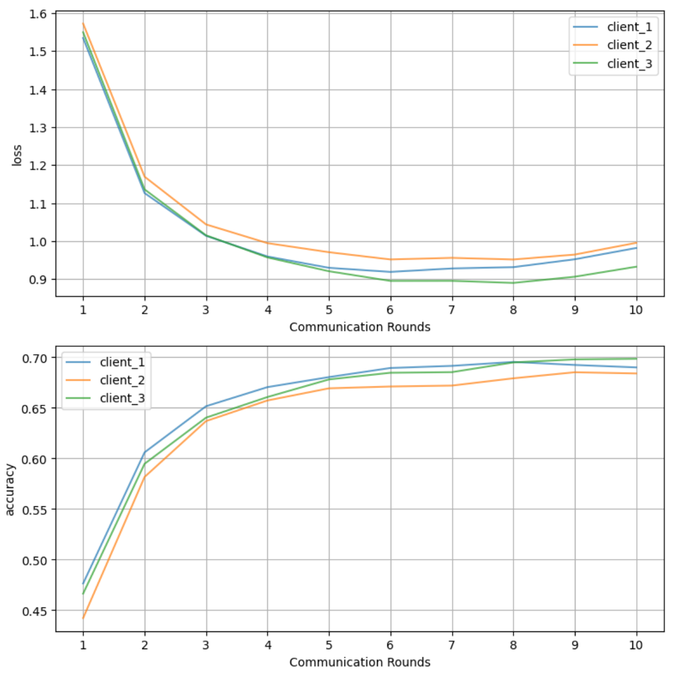
① Entrenamiento en cada cliente.
② Extracción de pesos de los modelos.
③ Agregación de los pesos de los modelos en un modelo global.
④ Los pesos del modelo global se envían a cada cliente.
⑤ Se repiten los pasos anteriores para más rondas.

Figura 1. Arquitectura de Aprendizaje Federado con servidor central y cinco clientes sintéticos [4].

Resultados

Caso de uso:

- Dataset **CIFAR-10** [5] de **clasificación de imágenes** ampliamente usado en Machine Learning.
- 10 clases divididas en **50000/10000** imágenes **train/test**.
- **Tres clientes sintéticos y 10 rondas de comunicación**.
- Se consideró una **red neuronal sencilla** común.



✓ Mejora considerable para los modelos de todos los clientes.
✓ Convergencia en pocas rondas.
✓ Solo se comparten los pesos de los modelos de red neuronal.
✓ Los datos no salen del servidor local y permanecen a salvo.

Figura 2. Evolución de las métricas de pérdida y precisión en el conjunto de test para el dataset CIFAR-10.

Conclusiones

- **Mejora del rendimiento predictivo** de modelos de Machine Learning mediante aprendizaje distribuido.
- Garantía de **privacidad y seguridad** de datos industriales sensibles a través del Aprendizaje Federado.
- **Escalabilidad eficiente del entrenamiento distribuido** mediante la utilización de múltiples nodos.
- **Resultados prometedores** en simulaciones con datasets reales EMNIST y CIFAR-10.
- **Necesidad de estándares** comunes para asegurar la compatibilidad y calidad en arquitecturas distribuidas.
- Esta forma de colaboración no solo permitiría **mejorar procesos industriales individuales**, sino que ayudarían a **mejorar la eficiencia del sector en su conjunto**.

Referencias

[1] Unión Europea. (2016). Reglamento General de Protección de Datos (GDPR). Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016.
URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj.
[2] Parlamento Europeo, 2021. The EU's AI regulation. Think Tank. URL: https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2021)698792.
[3] Tang, Z., Shi, S., Chu, X., Wang, W., Li, B., 2020. Communication-efficient distributed deep learning: A comprehensive survey.
[4] Sáinz-Pardo Díaz, J. and López García, A., 2023. Study of the performance and scalability of federated learning for medical imaging with intermittent clients. Neurocomputing, 518:142-154.
[5] TensorFlow, 2023. CIFAR-10 dataset in tensorflow datasets. URL: https://www.tensorflow.org/datasets/catalog/cifar10.

[a]Fundación Centro Tecnológico de Componentes
Parque Científico y Tecnológico de Cantabria (PCTCAN)
Isabel Torres, 1 . 39011 – Santander (SPAIN)
Tel.: +34 942 76 69 76 / info@ctcomponentes.com
*Correspondencia: mmelgarejo@centrotecnologicoctc.com

Mejor Centro Tecnológico 2024 (AEE)

Cofinanciado por la Unión Europea