

Máster Universitario en Ingeniería Informática

SISTEMAS INTELIGENTES PARA LA GESTIÓN EN LA  
EMPRESA

**TEORÍA**  
**Aprendizaje Federado**



**UNIVERSIDAD  
DE GRANADA**



Carlos Santiago Sánchez Muñoz

*Email:* carlossamu7@correo.ugr.es

*DNI:* 75931715K

*22 de junio de 2021*

# Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Aprendizaje Federado</b>	<b>3</b>
2.1. Desafío . . . . .	3
2.2. Desde el ruido hacia la verdad . . . . .	4
2.3. Cifrado homomórfico . . . . .	5
2.4. Funcionamiento . . . . .	5
<b>3. Tipos</b>	<b>7</b>
<b>4. Ventajas/Desventajas</b>	<b>8</b>
4.1. Ventajas . . . . .	8
4.2. Desventajas . . . . .	8
<b>5. Utilidades y logros</b>	<b>9</b>
5.1. Utilidades . . . . .	9
5.2. Aplicación real . . . . .	9
<b>6. Herramientas de uso</b>	<b>10</b>
<b>7. Ejemplo</b>	<b>11</b>
<b>8. Conclusiones</b>	<b>12</b>

## 1. Introducción

El Aprendizaje Profundo significa, en condiciones normales, el acceso a los datos de entrenamiento. El Aprendizaje Profundo, al ser un subcampo del Aprendizaje Automático, se trata de aprender de los datos. Pero a menudo, los datos de los que se aprende son increíblemente personales. Los modelos más significativos interactúan con la información más personal sobre la vida humana y nos dicen cosas sobre nosotros mismos que podrían haber sido difíciles de saber de otra manera. Parafraseando, un modelo de Aprendizaje Profundo puede estudiar miles de vidas para ayudarte a comprender mejor la tuya.

El principal recurso natural para el Aprendizaje Profundo son los datos de entrenamiento. Sin él, el Aprendizaje Profundo no puede aprender; y debido a que los casos de uso más valiosos a menudo interactúan con los conjuntos de datos más personales, el Aprendizaje Profundo encuentra serias dificultades a la hora de poseer esa información tan valiosa.

En 2017, Google publicó un artículo y una publicación de blog muy interesantes que hicieron una mella significativa en este hándicap. Google propuso que no necesitamos centralizar un conjunto de datos para entrenar un modelo sobre él. ¿y si en lugar de llevar todos los datos a un solo lugar, pudiéramos llevar el modelo a los datos? Este es un subcampo nuevo y emocionante del Aprendizaje Automático llamado Aprendizaje Federado. A lo largo de este trabajo se pretende conocer más este campo, entrando en algunos detalles y dejando una idea general bastante completa.

Esta cita describe la situación:

*"Los amigos no espían; la verdadera amistad también se trata de privacidad"*  
Stephen King, Corazones en Atlantis (1999)

La referencia central del trabajo es [2].

## 2. Aprendizaje Federado

Tal y como se ha introducido el Aprendizaje Federado nace para dar una solución al problema de la privacidad en el Aprendizaje Profundo [7].

¿Qué pasa si en lugar de llevar el corpus de datos de entrenamiento a un lugar para entrenar un modelo, pudiera llevar el modelo a los datos donde sea que se generen?

Esta simple inversión es extremadamente importante. En primer lugar, significa que para participar en la cadena de suministro de Aprendizaje Profundo, técnicamente las personas no tienen que enviar sus datos a nadie. Se pueden entrenar modelos valiosos en el cuidado de la salud, la gestión personal y otras áreas sensibles sin necesidad de que nadie revele información sobre sí mismos. En teoría, las personas podrían mantener el control sobre la única copia de sus datos personales (al menos en lo que respecta al Aprendizaje Profundo).

Esta técnica también tendrá un gran impacto en el panorama competitivo del Aprendizaje Profundo en la competencia empresarial y el espíritu empresarial. Las grandes empresas que anteriormente no querían (o no podían, por razones legales) compartir datos sobre sus clientes, pueden potencialmente obtener ingresos a partir de esos datos. Hay algunos dominios problemáticos en los que la sensibilidad y las limitaciones regulatorias que rodean a los datos han sido un obstáculo para el progreso. El cuidado de la salud es un ejemplo en el que los conjuntos de datos a menudo están bien cerrados, lo que dificulta la investigación.

**No es necesario tener acceso a un conjunto de datos para aprender de él.**

La premisa del Aprendizaje Federado es que muchos conjuntos de datos contienen información que es útil para resolver problemas (por ejemplo, identificar el cáncer en una resonancia magnética), pero es difícil acceder a estos conjuntos de datos relevantes en cantidades suficientemente grandes para entrenar un modelo de Aprendizaje Profundo adecuadamente sólido. La principal preocupación es que, aunque el conjunto de datos tiene información suficiente para entrenar un modelo de Aprendizaje Profundo, también tiene información que (presumiblemente) no tiene nada que ver con el aprendizaje de la tarea, pero que podría dañar a alguien si se revelara.

El Aprendizaje Federado se trata de un modelo que entra en un entorno seguro y aprende a resolver un problema sin necesidad de que los datos se muevan a ningún lado.

### 2.1. Desafío

El Aprendizaje Federado tiene dos grandes desafíos, los cuales son peores cuando cada persona en el conjunto de datos de entrenamiento tiene solo un pequeño puñado de ejemplos de entrenamiento. Estos desafíos son el **rendimiento** y la **privacidad**. Resulta que si alguien tiene solo unos pocos ejemplos de entrenamiento (o la mejora del modelo que le envían usa solo algunos ejemplos: un lote de entrenamiento), aún puede aprender bastante sobre los datos. Dadas 10.000 personas (cada una con un poco de datos), pasará la mayor parte de su tiempo enviando el modelo de un lado a otro y no mucho tiempo entrenando (especialmente si el modelo es realmente grande).

Como se basan en una red distribuida, las aplicaciones de Aprendizaje Federado deben abordar los riesgos de posibles ataques o fallas de numerosos trabajadores. Los ataques a un entorno de Aprendizaje Federado pueden adoptar diferentes formas. Un ataque podría originarse en un participante al alterar los datos utilizados para entrenar el modelo, o alterar el modelo en sí, con el potencial de comprometer el modelo global. La falta de confiabilidad de la red, la disponibilidad limitada, la falta de respuesta o el abandono de los trabajadores son problemas más frecuentes en un entorno distribuido, y las implementaciones de Aprendizaje Federado deben diseñarse para ser robustas ante estas amenazas [3].

## 2.2. Desde el ruido hacia la verdad

En esta subsección se va a exponer una paradoja estadística totalmente conectada a este nuevo mundo de aprendizaje.

Supongamos que está realizando una encuesta y quiere preguntar a 100 personas si han cometido un crimen atroz. Por supuesto, todos responderían “No” incluso si les prometiera que no se lo diría. En su lugar, haz que lancen una moneda dos veces (en algún lugar que no puedas ver) y diles que si el primer lanzamiento de la moneda es cara, deben responder honestamente; y si sale cruz, deben responder “Sí” o “No” según el segundo lanzamiento de la moneda.

Dado este escenario, nunca le pides a las personas que te digan si cometieron delitos. Las verdaderas respuestas están ocultas en el ruido aleatorio del primer y segundo lanzamiento de moneda. Si el 60 % de las personas dice “Sí”, puede determinar (usando matemáticas simples) que aproximadamente el 70 % de las personas que encuestaron cometieron delitos atroces (más o menos algunos puntos porcentuales). La idea es que el ruido aleatorio hace plausible que cualquier información que aprenda sobre la persona provenga del ruido en lugar de ella.

### Privacidad diferencial

El nivel de probabilidad de que una respuesta en particular provenga de un ruido aleatorio en lugar de un individuo protege su privacidad dándoles una negación plausible. Esto forma la base para la agregación segura y, de manera más general, gran parte de la privacidad diferencial.

Solo está mirando las estadísticas agregadas en general. Nunca ve la respuesta de nadie directamente; solo ve pares de respuestas o quizás agrupaciones más grandes. Por lo tanto, cuantas más personas pueda agregar antes de agregar ruido, menos ruido tendrá que agregar para ocultarlas y más precisos serán los hallazgos.

En el contexto del Aprendizaje Federado, podríamos agregar un montón de ruido, pero esto perjudicaría el entrenamiento. En su lugar, primero podemos sumar todos los gradientes de todos los participantes de tal manera que nadie pueda ver el gradiente de nadie más que el suyo. La clase de problemas para hacer esto se llama agregación segura, y para hacerlo, necesitará una herramienta más: el cifrado homomórfico.

## 2.3. Cifrado homomórfico

Es posible realizar operaciones aritméticas en valores cifrados. Una de las fronteras de investigación más interesantes es la intersección de la Inteligencia Artificial y la Criptografía. Al frente y al centro de esta emocionante intersección se encuentra una tecnología genial llamada Encriptación Homomórfica. En términos generales, el cifrado homomórfico permite realizar cálculos sobre valores cifrados sin descifrarlos.

Se va realizar un ejemplo de juguete en donde se está interesado en calcular sumas de los valores. No se va a explicar el funcionamiento interno más allá de que mediante claves públicas se va a cifrar y descifrar un número. En esta jerga, un valor cifrado se denomina texto cifrado y un valor no cifrado se denomina texto plano.

Veamos un ejemplo de cifrado homomórfico utilizando la biblioteca phe. (instálese con `pip3 install phe`):

```
import phe

public_key, private_key = phe.generate_paillier_keypair(n_length=1024)

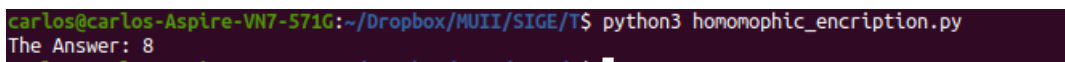
# Se encripta el número 5
x = public_key.encrypt(5)

# Se encripta el número 3
y = public_key.encrypt(3)

# Se suman los dos valores encriptados
z = x + y

# Se desencripta el resultado
z_ = private_key.decrypt(z)
print("The Answer: " + str(z_))
```

El resultado de la ejecución es el siguiente:



```
carlos@carlos-Aspire-VN7-571G:~/Dropbox/MUII/SIGE/TS$ python3 homomophic_encryption.py
The Answer: 8
```

Imagen 1: Ejecución del programa de encriptado

Combinar esta herramienta de encriptado en los modelos de aprendizaje otorga esa privacidad en los datos. Por ejemplo podrían encriptarse los gradientes que se agregan en un modelo de aprendizaje.

## 2.4. Funcionamiento

El Aprendizaje Federado mejora los algoritmos al enviar la versión actual de un algoritmo a dispositivos elegibles. Este modelo del algoritmo luego aprende de los datos privados en los teléfonos de un grupo selecto de usuarios. Cuando finaliza, se envía un resumen del nuevo conocimiento al servidor de la empresa: los datos en sí mismos nunca salen del teléfono [1].

Por seguridad, como ya se ha explicado este conocimiento generalmente se cifra en su camino de regreso al servidor. Para evitar que el servidor pueda averiguar datos individua-

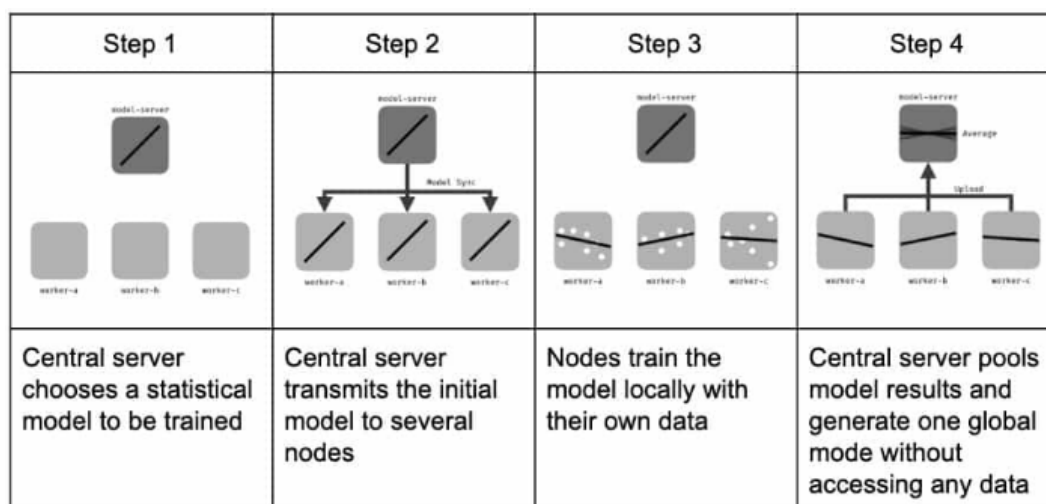


Imagen 2: Funcionamiento del Aprendizaje Federado [1]

les en función del resumen que ha recibido, Google ha desarrollado la agregación segura explicada.

Este protocolo usa criptografía para evitar que el servidor acceda a los resúmenes de información individuales. Bajo este esquema, el servidor solo puede acceder al resumen después de que se haya agregado y promediado con los resultados de cientos o miles de otros usuarios.

Alternativamente, la privacidad diferencial se puede usar para agregar ruido de datos aleatorio al resumen de un individuo, oscureciendo los resultados. Estos datos aleatorios se agregan antes de enviar el resumen al servidor, lo que le da al servidor un resultado lo suficientemente preciso para el entrenamiento algorítmico, sin que se revelen los datos de resumen reales. Esto preserva la privacidad del individuo.

Técnicas como el protocolo de agregación segura y la privacidad diferencial son cruciales para proteger la información del usuario tanto de la organización como de los piratas informáticos. Sin ellos, el Aprendizaje Federado no podría garantizar la privacidad de los usuarios.

Una vez que los resúmenes de información se han enviado de forma segura al servidor, se utilizan para actualizar el algoritmo. El proceso se repite miles de veces, y las versiones de prueba del algoritmo también se envían a varios dispositivos de usuario. Esto permite a las organizaciones evaluar nuevas versiones de algoritmos en datos de usuarios reales. Debido a que el análisis se realiza dentro de los límites de los dispositivos del usuario, los algoritmos se pueden probar sin tener que agrupar los datos del usuario en un servidor central.

Cuando se completan las pruebas, el modelo de algoritmo actualizado se envía a los dispositivos del usuario para reemplazar el antiguo. El algoritmo mejorado se utiliza en sus tareas normales. Si todo ha ido según lo planeado, será más efectivo y preciso para lograr sus resultados.

Finalmente, todo el ciclo se repite una y otra vez:

- El nuevo algoritmo estudia los datos en dispositivos de usuario seleccionados.
- Envía de forma segura resúmenes de estos datos de usuario al servidor.
- Estos datos luego se promedian con los resultados de otros usuarios.
- El algoritmo aprende de esta información, produce actualizaciones y las prueba.
- Se lanza una versión más avanzada del algoritmo a los usuarios.
- Con el tiempo, el algoritmo aprende de los datos del usuario y mejora continuamente, sin tener que almacenar los datos en los servidores de la empresa.

### 3. Tipos

Existen dos grandes categorías del Aprendizaje Federado [6].

El **Aprendizaje Federado Horizontal** y el aprendizaje federado homogéneo pueden hacer frente a los desafíos técnicos y prácticos dividiendo los datos en varias divisiones. El proceso funciona introduciendo conjuntos de datos similares en un espacio comparable. El algoritmo compara las características y los enlaces en consecuencia.

En el **Aprendizaje Federado Vertical**, diferentes conjuntos de datos comparten identificaciones de muestra similares pero espacios de características diferentes. Supongamos que en una ciudad hay dos empresas diferentes. Una es una empresa de comercio electrónico y la otra es un banco. Los conjuntos de usuarios contarán con las personas que viven en la zona para incluir un espacio de usuarios amplio, pero diferente según las tareas y actividades. Así que los conjuntos de datos estarán en espacios diferentes.



## 4. Ventajas/Desventajas

En esta sección se exponen las ventajas y desventajas de este tipo de Aprendizaje [8].

### 4.1. Ventajas

Las ventajas son las siguientes:

- La privacidad de los datos, ya que los modelos locales se añaden a la red y ayudan a un modelo general, pero al no tener que compartir los datos directamente, se garantiza la confidencialidad de cada uno de los nodos y sus datos.
- El proceso de aprendizaje se puede realizar cuando un dispositivo se está cargando, conectado a wifi y no está en uso, minimizando los inconvenientes que enfrentan los usuarios.
- El Aprendizaje Federado también puede ofrecer modelos algorítmicos tanto globales como personalizados.
- El proceso termina transfiriendo menos datos en general que en los modelos de aprendizaje tradicionales.
- La no necesidad de almacenamiento de los datos en un servidor central.
- Se reducen las latencias y el coste de intercambiar datos de forma continua con un servidor [5].

### 4.2. Desventajas

También existen algunas desventajas:

- La implementación de este aprendizaje tiene claras desventajas ya depende de la capacidad del dispositivo local para ser entrenado.
- Disponer de datos etiquetados para el entrenamiento local. Debido a esto, el tiempo de convergencia de estos modelos comparados con el Aprendizaje Automático tradicional es mayor.
- Riesgos de posibles ataques o fallas de nodos.
- Un ataque podría originarse en un participante al alterar los datos utilizados para entrenar el modelo.
- Aprendizaje condicionado a las limitaciones de los dispositivos locales en los que se ejecuta [5]. Se ha menguado con el avance de la tecnología.
- Problemas de conectividad en los nodos.
- Hasta ahora, la falta de herramientas.

## 5. Utilidades y logros

En esta sección se van a abordar usos que ha tenido el Aprendizaje Automático en diferentes campos de conocimiento así como aplicación en casos reales que está teniendo.

### 5.1. Utilidades

El Aprendizaje Federado no fue una gran revolución cuando Google lo inventó en 2017, debido a que no se sabía a qué proyectos reales en desarrollo se podía aplicar tal aprendizaje. Actualmente, se ha descubierto el potencial de este tipo de aprendizajes descentralizados y se está utilizando principalmente en 3 áreas.

#### **Transporte: Conducción autónoma**

Una conducción autónoma de un coche necesita multitud de tecnologías de Aprendizaje Automático para funcionar: Visión para analizar obstáculos, aprendizaje para la adaptación de la velocidad al entorno, etcétera. Debido a la cantidad de vehículos autónomos que se usarían si este proyecto fructuase, y la necesidad imperial de una rápida respuesta, hace que el enfoque tradicional de aprendizaje en la nube genere grandes riesgos de seguridad. El Aprendizaje Federado puede ser una gran solución para limitar este volumen de transferencia de datos y acelerar los procesos de aprendizaje de los coches autónomos.

#### **Industria 4.0: Fabricación Inteligente**

La industria 4.0 tiene una tendencia generalizada a la adopción de técnicas de Aprendizaje Automático, con el objetivo de mejorar la eficiencia y la eficacia de los procesos industriales al tiempo que se garantiza la seguridad. El problema llega con la privacidad de datos confidenciales que tanto las industrias como las empresas manejan. Por ello, el Aprendizaje Federado puede resolver este problema ya que, como se explica anteriormente, no revela ningún dato sensible.

#### **Medicina**

El Aprendizaje Federado ha encontrado en el campo médico y sanitario un gran sector en el que explotar la máximo su potencial, ya que el principal problema de este, era la gran confidencialidad y sensibilidad de los datos que maneja. Con un aprendizaje tradicional en la nube, todos estos datos se verían expuestos de una forma muy clara y con una gran brecha de seguridad de información. Por ello, gracias a este nuevo aprendizaje, se podrá entrenar las redes de los distintos hospitales, para utilizar sus inmensas bases de datos como entrenamiento sin comprometer la seguridad de los mismos. Esto supone un gran avance y se está trabajando intensamente en esta área.

### 5.2. Aplicación real

Varias compañías, como IBM Research, han empezado a trabajar para aplicar el Aprendizaje Federado en la asistencia médica y la salud del mundo real. La start-up Owkin con sede en París (Francia) y respaldada por Google Ventures, también lo está utilizando para predecir la resistencia de los pacientes a diferentes tratamientos y medicamentos, así como sus tasas de supervivencia para ciertas enfermedades. La compañía está hablando con

varios centros de investigación del cáncer en EE. UU. y Europa para aplicar sus datos a estos modelos. Estas conversaciones ya están dando frutos en forma de un próximo trabajo de investigación [4].

Por otro lado está, MELLODDY, un consorcio de descubrimiento de fármacos con sede en el Reino Unido, que está trabajando para demostrar cómo las técnicas de Aprendizaje Federado podrían brindar a los socios farmacéuticos lo mejor de ambos mundos: la capacidad de aprovechar el conjunto de datos de compuestos de fármacos colaborativos más grande del mundo para realizar el entrenamiento de IA sin sacrificar la privacidad de los datos.

Asimismo también existe un proyecto para Google Gboard [1]. El primer despliegue a gran escala de aprendizaje federado en el mundo real fue parte de Aplicación de teclado de Google, Gboard. La compañía tiene como objetivo utilizar la técnica para mejorar las sugerencias de palabras sin comprometer la privacidad del usuario.

## 6. Herramientas de uso

Entre los frameworks más populares, podemos destacar [5]:

- **TensorFlow Federated**, un entorno de código abierto de Google para aprendizaje automático y otros cálculos con datos descentralizados <https://www.tensorflow.org/federated?hl=es-419>.
- **PySyft**, una biblioteca de código abierto construida sobre PyTorch para el aprendizaje profundo cifrado y la privacidad <https://github.com/OpenMined/PySyft>.
- **Federated AI Technology Enabler (FATE)**, un proyecto de código abierto iniciado por el grupo de inteligencia artificial de Webank.
- **Sherpa.ai Federated Learning and Differential Privacy Framework**, framework de código abierto desarrollado para facilitar la investigación y experimentación abierta en el Aprendizaje Federado y la Privacidad Diferencial.

## 7. Ejemplo

Como cúlmen de este inicio en el mundo del Aprendizaje Federado sería interesante estudiar un ejemplo breve de funcionamiento y aplicación del mismo. Se van a dar unas breves pinceladas de un tutorial de [2]. Dicho ejemplo es de Aprendizaje Profundo para detectar *spam* en *mails*.

En primer lugar disponemos de tres usuarios entre los que se reparten los datos.

```
bob = (train_data[0:1000], train_target[0:1000])
alice = (train_data[1000:2000], train_target[1000:2000])
sue = (train_data[2000:], train_target[2000:])
```

El procedimiento de aprendizaje podría ser como sigue:

```
for i in range(3):
    print("Starting Training Round...")
    print("\tStep 1: send the model to Bob")
    bob_model = train(copy.deepcopy(model), bob[0], bob[1], iterations=1)

    print("\n\tStep 2: send the model to Alice")
    alice_model = train(copy.deepcopy(model), alice[0], alice[1], iterations=1)

    print("\n\tStep 3: Send the model to Sue")
    sue_model = train(copy.deepcopy(model), sue[0], sue[1], iterations=1)

    print("\n\tAverage Everyone's New Models")
    model.weight.data = (bob_model.weight.data + alice_model.weight.data
                        + sue_model.weight.data)/3

    print("\t\t% Correct on Test Set: " + str(test(model, test_data, test_target)*100))

    print("\nRepeat!!\n")
```

Por último los resultados:

```
Starting Training Round...
Step 1: send the model to Bob
Loss:0.21908166249699718

.....

Step 3: Send the model to Sue
Loss:0.015368461608470256

Average Everyone's New Models
% Correct on Test Set: 98.8
```

¿Posibles mejoras para este ejemplo? Que la información que los usuarios agregan al modelo esté encriptada. Consultar [2] para ver cómo.

## 8. Conclusiones

El Aprendizaje Federado es uno de los avances más emocionantes en el Aprendizaje Profundo y por consiguiente del Aprendizaje Automático. Creo que el Aprendizaje Federado cambiará el panorama del Aprendizaje Profundo en los próximos años ya que desbloqueará nuevos conjuntos de datos que anteriormente eran demasiado sensibles para trabajar, creando un gran bien social como resultado de estas nuevas oportunidades empresariales disponibles. Esto es parte de una convergencia más amplia y emocionante entre el cifrado y la investigación de Inteligencia Artificial.

Hasta ahora el mayor impedimento del uso de este modelo era la falta de herramientas, marcos de trabajo y otros para estandarizar su uso. A partir de ahora cada vez más datos sensibles podrán ser expuestos a este modelo dándole un lugar más importante en la Inteligencia Artificial. En resumen el Aprendizaje Federado es un paso más hacia la Inteligencia Artificial confidencial.

## 9. Bibliografía

- [1] Dementium. *Aprendizaje federado: ¿es realmente mejor para su privacidad y seguridad?* URL: <https://dementium2.com/seguridad-de-informacion/aprendizaje-federado-es-realmente-mejor-para-su/>.
- [2] O'reilly. *Deep learning on unseen data: introducing federated learning*. URL: [https://www.oreilly.com/library/view/grokking-deep-learning/9781617293702/kindle\\_split\\_023.html](https://www.oreilly.com/library/view/grokking-deep-learning/9781617293702/kindle_split_023.html).
- [3] ICHI PRO. *Aprendizaje federado: ¿por qué y cómo empezar?* URL: <https://ichi.pro/es/aprendizaje-federado-por-que-y-como-empezar-141963632002703>.
- [4] MIT Technology Review. *Aprendizaje federado: la nueva arma de IA para asegurar la privacidad*. URL: <https://www.technologyreview.es//s/11017/aprendizaje-federado-la-nueva-arma-de-ia-para-asegurar-la-privacidad>.
- [5] Paloma Recuero de los Santos. *Aprendizaje federado: IA con privacidad*. URL: <https://empresas.blogthinkbig.com/aprendizaje-federado-ia-con-privacidad/>.
- [6] Data Science. *Aprendizaje federado*. URL: <https://datascience.eu/es/aprendizaje-automatico/aprendizaje-federado/>.
- [7] Andrew W. Trask. *Grokking Deep Learning*.
- [8] Wikipedia. *Aprendizaje Federado*. URL: [https://es.wikipedia.org/wiki/Aprendizaje\\_federado](https://es.wikipedia.org/wiki/Aprendizaje_federado).