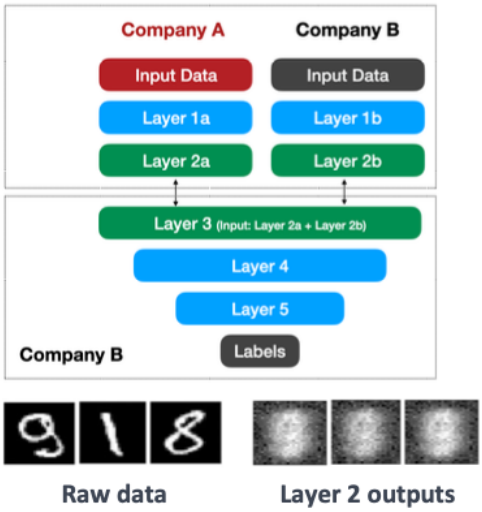
# Plataforma de Inteligencia Artificial para entrenar redes neuronales manteniendo la privacidad de los datos.

**Acuratio Europe SL**

El proceso de transformación digital al que se están sometiendo todas las industrias está convirtiendo los datos en el principal activo de muchas empresas. Las exigencias de los ciudadanos, las nuevas regulaciones como el Reglamento General de Protección de Datos (RGPD) y la necesidad de confidencialidad de empresas y organizaciones. Está demandado un cambio en la forma en la que se almacenan, tratan y comparten los datos (Gobierno del Dato). Esto unido a la creciente demanda de productos personalizados, que requieren conocer los gustos de los usuarios, ha creado una necesidad en el mercado de tecnologías que preserven la privacidad de los datos a la vez que cumplan las expectativas de los consumidores. Las llamadas Privacy-enhancing technologies.

De esta necesidad nace Acuratio, que en 2017 publica la primera implementación de Federated Learning en open-source. Una técnica que permite entrenar modelos de Machine Learning, moviendo los modelos a donde están los datos y así preservando la privacidad de los datos con varias técnicas como differential privacy y protocolos de agregación segura. Desde esa primera implementación Acuratio ha trabajado en proporcionar al mercado una plataforma para que las organizaciones puedan entrenar redes neuronales manteniendo la privacidad de los datos.

Entre otras innovaciones Acuratio en colaboración con el MIT Media Lab sentó las bases del llamado Vertical Federated Learning con Redes Neuronales. La cual es una técnica que permite a dos o más organizaciones con una base de usuarios en común entrenar un modelo conjunto con los datos que cada una conoce de los individuos.

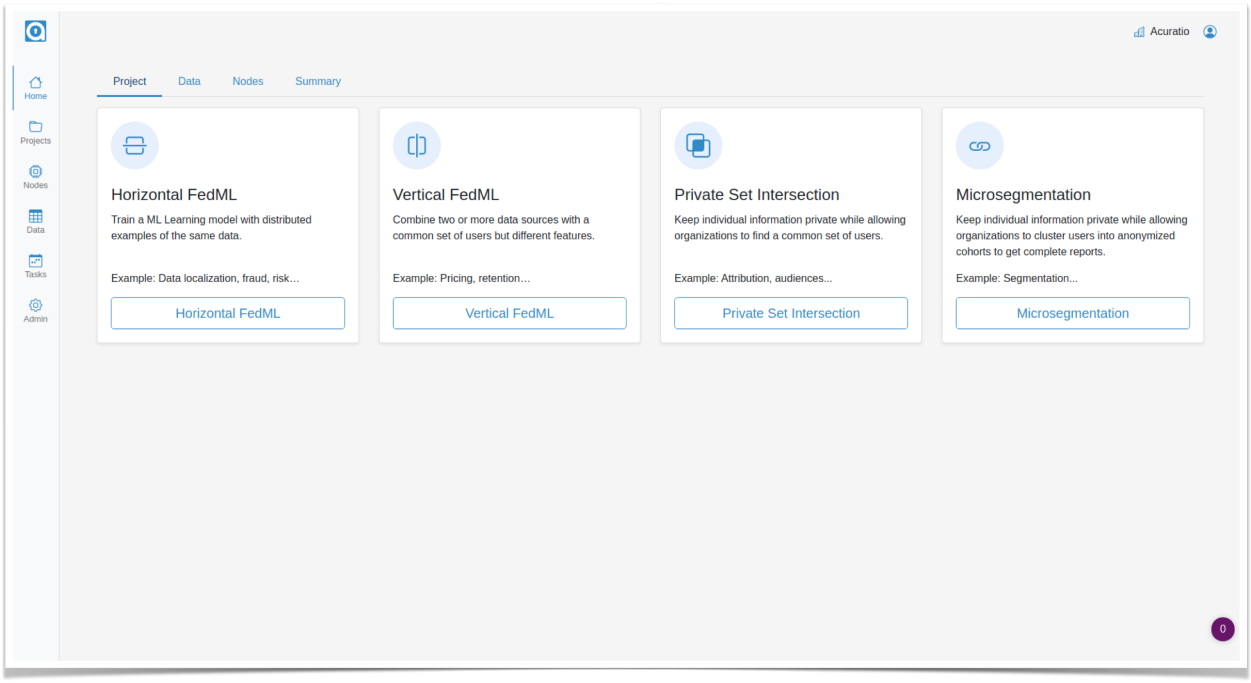



Esta técnica y la plataforma que actualmente comercializa Acuratio entre las entidades financieras más sólidas de Europa, permite a estas organizaciones controlar mejor sus datos, auditar su uso y preservar su valor cuando son compartidos con otras organizaciones.

Acuratio da respuesta a dos retos a los que están haciendo frente las empresas:

- Gobierno del Dato: Muchas empresas tienen políticas o bien muy laxas o estrictas, a la hora de acceder a los datos. Por ejemplo, el departamento de innovación de un banco puede tardar hasta 3 meses en obtener datos de operaciones fraudulentas para probar un nuevo modelo. Lo contrario ocurre en otras empresas en las que toda la organización tiene acceso a datos personalmente identificables de los clientes.
- Colaboraciones: Actualmente cuando por ejemplo un banco, una aseguradora o un operador de telefonía quiere mejorar sus modelos de riesgo o fraude, o bien envía los datos en un USB o bien dependen de una tercera empresa que hace de intermediario. Con nuestra plataforma por primera vez pueden colaborar directamente, a menor coste y reduciendo el proceso de meses a días.

La plataforma de Acuratio se comercializa a través de licencias y actualmente 2 Bancos, 3 Aseguradoras y un operador de Telefonía hacen uso de ellas para distintas operaciones analíticas en las que obtienen resultados agregados sin comprometer la privacidad de los usuarios y cumpliendo con total garantía con las regulaciones vigentes.




*Adecuación al premio*: El equipo de Acuratio lleva trabajando desde el 2017 en las Privacy-enhancing technologies para que **no tenga que haber un compromiso entre privacidad y utilidad de los datos**. La privacidad es un derecho fundamental que hay que proteger y a la vez proporcionar a las empresas europeas herramientas para personalizar sus ofertas y competir con empresas extranjeras.

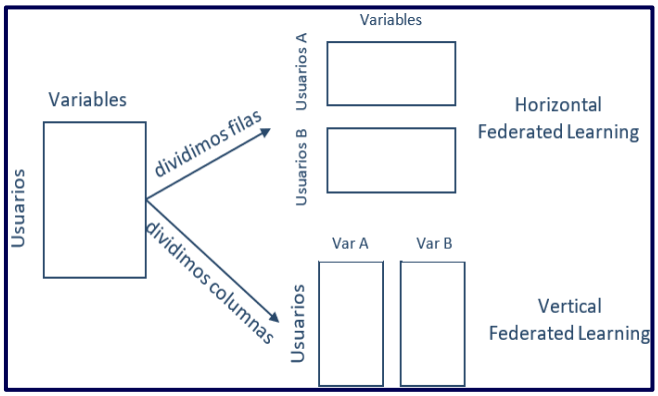
Acuratio actualmente es un emprendimiento totalmente sostenible y con perspectivas de crecer y doblar personal el año próximo. Seguiremos con nuestra estrategia comercial en Europa e iniciaremos un acercamiento al mercado latinoamericano a la vez que continuaremos haciendo labor comercial con en USA, donde recientemente adquirimos nuestro primer cliente.

Plataformas como la nuestra e investigación básica como la que realizamos con MIT Media Lab son ejemplos para que proyectos tan ambiciosos como GAIA-X u otras iniciativas lleguen a buen puerto. Los espacios de datos federados son el futuro de las colaboraciones entre empresas, hospitales, entidades públicas y países.

El acceso distribuido y privado a imagen médica, historiales clínicos y datos genéticos supondrá un antes y un después para la realización de estudios o la creación de modelos para la medicina personalizada. En ese sentido seguiremos trabajando en adaptar nuestra plataforma para el tratamiento de estos datos especialmente sensibles. El año pasado empezamos a colaborar con el CIC bioGUNE - Centro de Investigación Cooperativa en Biociencias (Donostia) en el tratamiento de datos de genéticos.

## Capacidades de la Plataforma

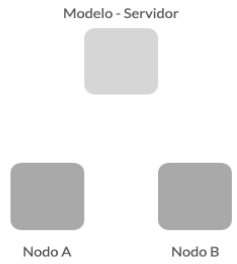
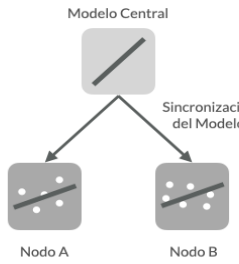
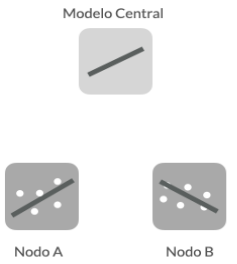
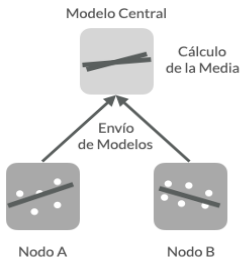
La plataforma de Acuratio está diseñada para entrenar modelos de datos distribuidos de dos tipologías (partiendo de la premisa de compartición de información entre 2 entidades o silos):




**Horizontal FML:** que se da cuando ambas entidades / silos tienen el mismo tipo de variables, pero cada una tiene una base de usuarios distinta.
**Vertical FML:** cuando ambas entidades / silos compartan una base de usuarios en común, pero cada una conoce características/variables distintas de los usuarios.
Este sistema es capaz de aprender de forma automática, integrándose en los desarrollos y adaptándose a los cambios de entorno cada vez que este sea alimentado con nuevos datos.
De esta manera, se entrenan los modelos de datos en local, es decir, en lugar de mover los datos a un servidor central, se moverán los algoritmos a dónde se encuentran los datos, manteniendo la privacidad de los datos en todo momento. Además, para incrementar el nivel de privacidad y seguridad, se aplicarán técnicas de criptografía para aumentar la seguridad.

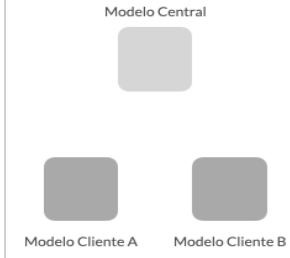
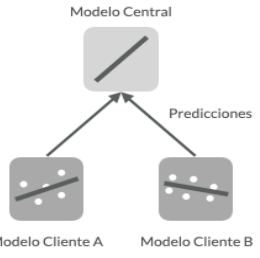
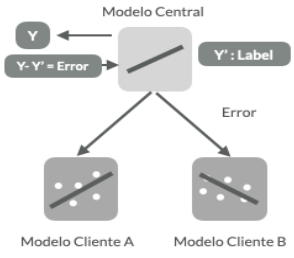
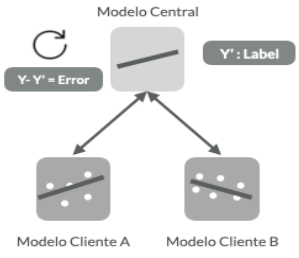
La plataforma permite (a elección) aplicar técnicas de k-anonimato y/o differential privacy a los datos antes de ponerlos a disposición de terceros. Esto añade una capa extra de privacidad para el proceso pero que no es necesaria normalmente, dado que el propio funcionamiento ordinario del software ya produce una anonimización completa de los datos. Sin embargo, a demanda, la solución permite incorporar esta capa extra de privacidad.
Además, se pueden aplicar 2 metodologías de FML en función de cómo estén distribuidos los datos:

**Horizontal FML**: los datos nunca saldrán del control de los propietarios, solo se mandarán los modelos comprimidos y entrenados.

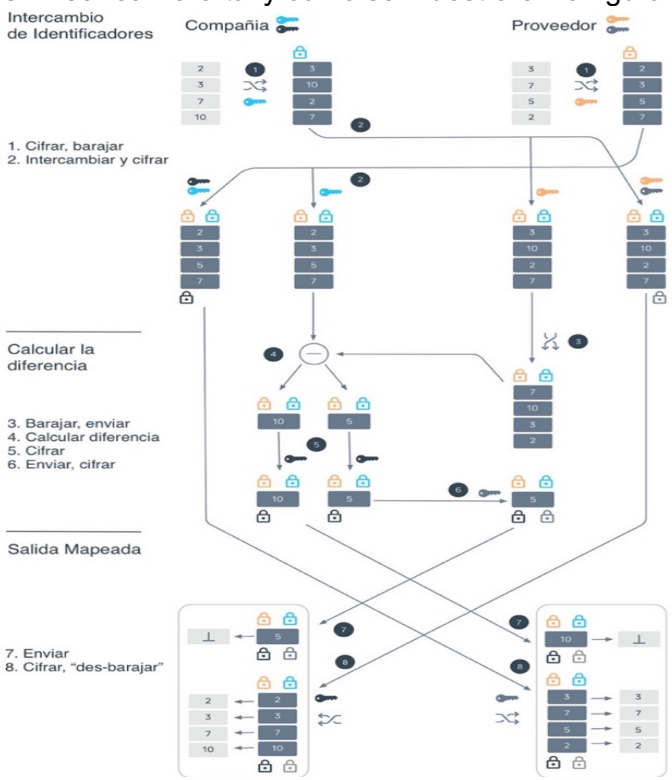
| Paso 1 | Paso 2 | Paso 3 | Paso 4 |
|---|---|---|---|
|  |  |  |  |
| El servidor central elige un modelo estadístico para entrenar. | El servidor central envía el modelo a los distintos nodos que tienen los datos (Workers) | Los nodos entrenan el modelo localmente con sus datos. | Los nodos envían su modelo entrenando a servidor central, que recoge todos los modelos y hace la media de los mismos. |

Para hacer la media se enviarán estos pesos con un protocolo de agregación segura que genere ruido simétrico que cada participante añade a sus pesos y luego se cancela al hacer la suma. Por lo que si alguien intercepta la comunicación solo tendría los pesos ofuscados. Además, se podrán comprimir los modelos, lo que añadirá indirectamente más dificultad para la interpretación de los resultados para alguien ajeno al protocolo, y en consecuencia, mayor nivel de ofuscación de los datos.

**Vertical FML:** el proceso será distinto al anterior, como paso previo, será necesario encontrar un identificador (ID) Común para que permita la correlación de datos (imagen de la derecha), y por lo tanto será necesaria la generación de un ID privado.

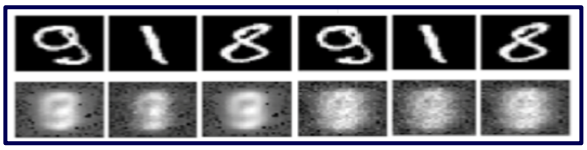
| Paso 1 | Paso 2 | Paso 3 | Paso 4 |
|---|---|---|---|
|  |  |  |  |
| Se diseñan los modelos para cada parte de los datos. Y se ordenan los usuarios comunes para comenzar el entrenamiento. | Comienza el entrenamiento sobre los usuarios comunes, cada modelo envía sus resultados. | El modelo central concatena la predicciones individuales y las toma como parámetros de entrada, hace una predicción y comprueba si hay un error. Finalmente modifica los modelos corrigiendo el error. | El proceso continua hasta que el error es mínimo para todos los usuarios sobre los que se ha entrenado el modelo. Y todas las partes del modelo están optimizadas para el objetivo buscado. |

El intercambio de estos ID se realizará tal y como se muestra en la figura de la siguiente figura:

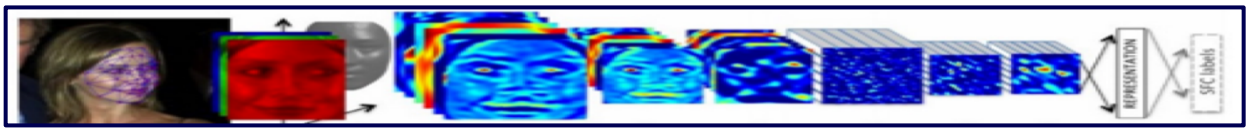


A continuación, se listan las características técnicas y funcionales más notables del Vertical FML:

- **Generación de ID común**, por parte de ambos clientes, el cual será totalmente privado, a partir del cual el otro cliente no podrá re-identificar a los usuarios. A partir del mismo, se podrá realizar el entrenamiento.
- **Distance Correlation**, capa de privacidad que consistirá en añadir un término a la función de coste que maximiza la ofuscación al mismo tiempo que minimizará la pérdida de información.
- **Corrección de modelos**, los datos una vez ofuscados, pasarán por las capas de las redes neuronales que procesan en distintas fases el funcionamiento del modelo de cada entidad, realizándose en la sede de cada una. Al final del proceso de análisis, el resultado comparará con la predicción que tiene una de ellas en su modelo, comprobando el índice de error y realizando la consiguiente corrección.



- **Mejora de resultados iniciales**, a cada entidad se le devolverá una mejora del modelo predictivo que tenían en origen, manteniendo la privacidad de los datos.
- **Inversión de parámetros**, durante el proceso de entrenamiento para que, la otra parte, pueda construir un decodificador que le permita inferior los datos a partir de la información recibida.




- **Retroalimentación**, tras el entrenamiento del algoritmo, se permitirá la entrada de nuevos registros a los que se les aplicará el modelo en su versión resultante del análisis realizado mediante las redes neuronales.

Periódicamente, se deberá realizar un nuevo entrenamiento del algoritmo, debido a la incorporación en cartera de un número de clientes entre la última valoración y la actual, que incorpore la comparación con los resultados obtenidos de la aplicación del algoritmo desde el último entrenamiento, y que permita aprender y obtener en todo caso parámetros de mejora suficientemente buenos.

A modo de resumen, los aspectos diferenciales más significativos de la plataforma son las siguientes:

- **Modelado de datos**: mejora continua del modelado de datos debido a que los algoritmos se entrenarán de manera continuada a medida que haya nuevos datos.
- **Privacidad**: se moverá el algoritmo al dato, por lo que el dato se mantendrá privado en todo momento.
- **Agnosticismo**: se tratará de una plataforma con un alto grado de adaptabilidad y flexibilidad a diferentes sectores y tipología de almacenamientos de datos.
- **Registro**: en todo momento quedará registrado qué dato se ha tratado, quién ha accedido a la información, cómo se ha accedido a la información, etc. por lo que permitirá auditar todo lo que tenga que ver con el propio acceso a la información y al tratamiento de los datos.

Cabe destacar que el equipo de Acuratio es el inventor de esta técnica para redes neuronales profundas y ha publicado varios papers en colaboración con el MIT Media Lab como este: https://arxiv.org/abs/2008.04137