AGENCIA NACIONAL DE INVESTIGACIÓN E INNOVACIÓN

# Informe final publicable de proyecto Anonimización de datos basada en redes generativas antagónicas

**Código de proyecto ANII: FSDA_1_2018_1_154419**

**26/11/2021**

**MAYR, Franz** (Investigador)

**VISCA ZANONI, Ramiro** (Investigador)

**YOVINE, Sergio** (Responsable Técnico - Científico)

UNIVERSIDAD ORT. FACULTAD DE INGENIERÍA (Institución Proponente)

**Resumen del proyecto**

El intercambio de información, ya sea en forma de datos brutos o de modelos entrenados usando aprendizaje automático, debe garantizar niveles adecuados de privacidad. Esta cuestión no es sólo técnica, sino también jurídica, ya que existen leyes que definen el derecho a la privacidad. En este proyecto estudiamos un escenario en el que varias organizaciones (públicas y/o privadas) comparten datos y modelos entrenados con datos privados de cada organización (o de sus pacientes, usuarios, clientes, etc.). El resultado del proyecto fue la propuesta de una solución que consiste en un esquema de privacidad diferencial de tipo mixto que compone un mecanismo centralizado y otro local. El primero es la disponibilización pública de un ensemble de modelos a través de un curador confiable que aplica ruido aleatorio a las predicciones, protegiendo así los datos de las organizaciones participantes en dicho ensemble de modelos. El segundo permite a un tercero realizar consultas al ensemble protegiendo su datos a través de un curador confiable que aplica un mecanismo de privacidad diferencial local que agrega ruido aleatorio a los datos enviados en la consulta. Esta técnica fue implementada y evaluada experimentalmente con éxito en aplicaciones en ciberseguridad y salud.

**Ciencias Naturales y Exactas / Ciencias de la Computación e Información / Ciencias de la Computación / Aprendizaje automático**

**Palabras clave: Anonimización de datos / Aprendizaje profundo / Redes generativas antagónicas /**

## Introducción

Impulsados por el crecimiento de los datos disponibles y la potencia de cálculo, los avances en el campo de la inteligencia artificial están dando lugar a mejoras significativas en la capacidad de resolver una variedad de tareas con la ayuda de artefactos inteligentes impulsados por algoritmos de aprendizaje automático. Este es el caso en dominios críticos como salud y seguridad, donde se trabaja activamente en el desarrollo de algoritmos cada vez más precisos para abordar problemas como diagnóstico de enfermedades [1] y detección de intrusiones [2,3,4,5].

Sin embargo, la oportunidad de construir sistemas predictivos inteligentes conlleva grandes desafíos. Por lo general, para aprender modelos predictivos se necesitan cantidades considerables de datos de entrenamiento que permitan alcanzar niveles de exactitud satisfactorios, pero este requisito no puede ser cumplido por una sola organización. Esta carencia podría superarse si las organizaciones compartieran los datos brutos o los modelos predictivos entrenados con dichos datos.

Como ejemplo en esta línea, en la última década las ONG y los institutos de investigación han impulsado la publicación de datos públicos abiertos [6]. En el caso de Europa, por ejemplo, el acceso a los datos públicos está legislado por la Directiva (UE) 2019/1024 sobre datos abiertos y reutilización de la información del sector público [7]. En Uruguay, la iniciativa de llevar adelante políticas públicas de datos abiertos está legislada en la Ley Sobre el Derecho de Acceso a la Información Pública (Ley Nº 18.381), cuya finalidad es fomentar y prescribir la disponibilización de los datos producidos, obtenidos, en poder y/o bajo control de organismos públicos.

Ciertamente, la compartición de datos no está solamente motivada por la legislación sobre datos abiertos, sino también por la necesidad de ponerlos a libre disposición de los agentes públicos y privados que tienen la capacidad técnica de utilizar estos datos para la innovación productiva y científica [8].

Sin embargo, a pesar de las ventajas de compartirlos, en la mayoría de los casos los datos no pueden publicarse ni transferirse fácilmente [6,9]. De hecho, la mayoría de los datos recogidos por las organizaciones, ya sean públicas o privadas, contienen información sobre los ciudadanos, clientes, usuarios o pacientes que son los verdaderos propietarios de esos datos, como por ejemplo, el documento de identidad, las contraseñas y los números de la seguridad social, de las cuentas bancarias y de las tarjetas de crédito. Evidentemente, independientemente del valor que pueda tener para una organización o para terceros externos, esos datos no les pertenecen. En Uruguay, los datos disponibilizados están sujetos al cumplimiento de la legislación vigente sobre protección de la privacidad de los datos contemplada en la Ley Nº 18.331 Protección de Datos Personales y acción de Habeas Data.

Para comprender el contexto, consideremos el siguiente caso de uso de inteligencia artificial como servicio, en el que los datos de un grupo de individuos (pacientes, usuarios, clientes, etc.), reales propietarios de los mismos, se almacenan en la base de datos de una organización (mutualista, empresa de comercio electrónico, banco, etc.) de confianza que permite a un tercero, proveedor de servicios de análisis de datos, consultar su base de datos. Sin las medidas de protección adecuadas, estas consultas, aunque no sean mal intencionadas, pueden revelar datos sensibles del propietario, como la

identidad de la persona. Para evitar esta pérdida de privacidad, deben tomarse las medidas adecuadas al permitir el acceso a la base de datos de una organización, más allá del uso de cualquier técnica de anonimización que elimine la información de identificación personal [10]. Esto se debe a que se han descrito varios ataques capaces de reidentificar a los individuos aún habiendo sido anonimizados [11,12,13,14]. Además, aunque los datos sensibles no sean publicados, estos pueden quedar expuestos al permitir que terceros consulten los modelos predictivos entrenados con ellos mediante los llamados ataques de inversión de modelos [15].
En suma, el intercambio de información, ya sea en forma de datos brutos o de modelos entrenados usando aprendizaje automático, debe garantizar niveles adecuados de privacidad. Esta cuestión no es sólo técnica, sino también jurídica, ya que, como dijimos, existen leyes relativas a la privacidad.
Por lo expuesto, queda claro que es esencial contar con mecanismos para proteger la información privada contenida en los datos que se ponen a disposición de terceros. Dichos mecanismos deben aplicarse independientemente de la forma en que se compartan los datos, ya sea publicando un conjunto de datos o permitiendo que interesados externos consulten una base de datos. Además, los mecanismos de protección de datos deben ser capaces de conservar suficiente información útil para resolver las tareas para las cuales se los necesita [17].
En consecuencia, en este proyecto estudiamos un escenario en el que varias organizaciones (públicas y/o privadas) comparten datos y modelos entrenados con datos privados de cada organización (o de sus pacientes, usuarios, clientes, etc.). En este contexto, Papernot y otros [18,19] habían propuesto un mecanismo denominado "Private Aggregation of Teacher Ensembles" (PATE). Este consiste en construir un conjunto (ensemble) de modelos que añade ruido aleatorio a las salidas de los predictores antes de agregarlos. PATE garantiza la privacidad de los datos de las organizaciones que participan en el ensemble mediante la implementación de un mecanismo de privacidad diferencial [20]. Sin embargo, no proporciona ninguna protección a las organizaciones que consultan el ensemble con con sus propios datos (privados) con el objetivo de obtener una respuesta para tomar una decisión (de negocio, diagnóstico, etc.) o para entrenar su propio modelo predictivo.
Para resolver este problema, propusimos una solución que consiste en un esquema de privacidad diferencial de tipo mixto que compone un mecanismo centralizado y otro local. El primero, basado en PATE, permite la disponibilización pública de un ensemble de modelos a través de un curador confiable que aplica ruido aleatorio a las predicciones, protegiendo así los datos de las organizaciones participantes. El segundo permite a un tercero realizar consultas al ensemble protegiendo su datos a través de un curador confiable que aplica un mecanismo de privacidad diferencial local que agrega ruido aleatorio a los datos enviados en la consulta. Esta técnica fue implementada y evaluada experimentalmente con éxito en aplicaciones en ciberseguridad y salud.

**Metodología/diseño del estudio**

El proyecto estuvo motivado inicialmente por la necesidad de encontrar herramientas de privatización de datos secuenciales, en particular los registros (logs) de sistemas de seguridad que contienen información sensible de los usuarios, para que puedan ser liberados con el objetivo de llevar a cabo experimentos predictivos de ciberseguridad, concretamente el entrenamiento de redes neuronales dedicadas a la detección y prevención de ciberataques [2, 3, 4, 5]. Durante el desarrollo del proyecto surgió también el interés de aplicar y validar los resultados en datos secuenciales en el ámbito de la salud, como por ejemplo electrocardiogramas [21, 22].
La metodología aplicada para el desarrollo del proyecto se centró en el estudio de diferentes mecanismos basados en privacidad diferencial como forma de definir procesos o desplegar herramientas que permitieran un acceso a los datos para la liberación de los mismos o la realización de cualquier consulta que se realice sobre ellos. Para cada enfoque se analizó la pérdida de información con respecto al uso de datos considerados sensibles y la utilidad de la liberación de datos una vez privatizados, en particular, para desarrollar modelos de aprendizaje automático en las áreas de interés antes mencionadas.
Concretamente, en este proyecto se investigaron dos enfoques. El primero consiste en generar una nueva base de datos sintética construida a partir de la privatización de los datos originales vía la aplicación de mecanismos de privacidad diferencial. El segundo apunta a explorar una forma de proteger los modelos de aprendizaje automático mientras se protegen los datos originales utilizados para crear dichos modelos.
En cuanto al desarrollo de software, se llevaron adelante dos líneas de trabajo. Una de ellas es enfocó en la implementación de una herramienta para la generación de logs. La otra, tuvo por objetivo diseñar e implementar herramientas de software para automatizar la validación experimental de cada uno de los enfoques estudiados. La necesidad de dichas herramientas se debe a la complejidad que conlleva transformar una base de datos en otra base de datos evaluando diferentes mecanismos de privatización, con diferentes parámetros de privacidad, que se suman a los requeridos para el entrenamiento de los modelos generativos y predictivos, así como evaluar las métricas de privacidad y

utilidad predictiva apropiadas en cada caso.

**Resultados, análisis y discusión**

El principio director detrás del primer enfoque, orientado a generar una nueva base de datos sintética, consiste en mantener la representación de las secuencias y generar nuevas secuencias en sustitución de las originales. Para esto, se investigó una variante novedosa de privacidad generativa adversaria (GAP) en la que cada secuencia es considerada como una base de datos, lo que corresponde a un enfoque local desde el punto de vista de la privacidad diferencial. Para la generación se usaron modelos de aprendizaje automático profundo basados en redes neuronales recurrentes (RNN) y se llevaron adelante dos líneas de investigación. La primera consistió en entrenar RNN con una capa de "embedding" cuyo objetivo es aprender similitudes entre los datos que forman parte de las secuencias. Así, esta relación de similitud actúa como función de utilidad de un mecanismo de privacidad diferencial llamado mecanismo exponencial. A continuación, se genera una base de datos sintética aplicando este mecanismo exponencial para cada dato de la secuencia de forma individual y utilizando sólo una vez la secuencia original completa. La segunda línea de trabajo consistió en entrenar RNN para generar secuencias a partir de una distribución de probabilidad, de tal manera que cada dato de la secuencia se elige siguiendo la distribución aprendida durante el entrenamiento. En este caso, se investigaron diferentes maneras de integrar mecanismos de privacidad diferencial en la generación de las secuencias sintéticas a partir de las originales. De todas las ideas exploradas, la que mejores resultados arrojó, en cuanto al balance entre privacidad y de capacidad predictiva, es una versión del segundo enfoque que consiste en entrenar una RNN generativa conjuntamente con un modelo predictivo dedicado a la tarea específica, por ejemplo, una clasificación binaria (normal vs anormal) de las secuencias. En este método se observó que no sólo aumenta la exactitud del aprendizaje realizado sobre los datos sintéticos a medida que aumenta el presupuesto de privacidad diferencial, sino que también aumenta la distorsión o distancia de Hamming, que calcula una puntuación de coincidencia para dos secuencias de datos como el número de posiciones en las que difieren. La exploración más profunda de este escenario queda para trabajo futuro. En efecto, si bien los resultados observados experimentalmente son muy interesantes ya que el modelo aprendió a generar secuencias normales y anormales que son muy diferentes a las originales pero que mantienen altos niveles de utilidad predictiva, se consiguieron mejores resultados con el segundo enfoque que presentamos a continuación.

El caso de uso planteado como punto de partida para la segunda línea de investigación consiste en que varias organizaciones comparten datos y modelos entrenados con datos privados de cada organización. Para resolver este problema, propusimos una solución que consiste en un esquema de privacidad diferencial de tipo mixto que compone un mecanismo centralizado y otro local. Como mecanismo centralizado, se tomó como base el denominado "Private Aggregation of Teacher Ensembles" (PATE) [18,19], que consiste en construir un ensemble de modelos a través de un mecanismo de agregación que perturba aleatoriamente las predicciones de los modelos constituyentes antes de combinarlos para producir la predicción final. Así, PATE logra garantizar la privacidad de los datos de las organizaciones que participan en el ensemble permitiendo obtener buenos resultados predictivos cuando el conjunto de participantes es suficientemente grande. Sin embargo, PATE no está pensado para proporcionar protección alguna a las organizaciones que envían consultas al ensemble para obtener sus predicciones (por ejemplo, con el objetivo de entrenar su propio modelo con sus propios datos). El mecanismo centralizado está basado en PATE y permite la disponibilización pública de un ensemble de modelos a través de un curador confiable que aplica ruido aleatorio a las predicciones, protegiendo así los datos de las organizaciones participantes. El mecanismo local permite a un tercero realizar consultas al ensemble protegiendo su datos a través de un curador confiable que aplica privacidad diferencial perturbando aleatoriamente los datos enviados en la consulta. Desde el punto de vista de la utilidad, la propuesta está fundada en que varios trabajos científicos han mostrado experimentalmente que los modelos de ensemble son robustos al ruido en los datos [23, 24]. Esta técnica fue implementada y los resultados experimentales obtenidos en aplicaciones en ciberseguridad y salud permitieron validar exitosamente esta hipótesis en tanto y en cuanto se observaron pérdidas de exactitud predictiva poco significativas.

Tangencialmente, los experimentos realizados con logs de WAF y electrocardiogramas ayudaron a descubrir que la denominada "pérdida de privacidad dependiente de los datos" propuesta en [18] puede estar sujeta a una alta varianza. Hasta donde sabemos, este fenómeno no había sido reportado previamente en otros trabajos de investigación y abre la puerta a futuras investigaciones.

Todos los experimentos realizados y resultados obtenidos en ambas líneas de investigación están detallados en la tesis de Maestría de Ramiro Visca [25], así como la herramienta DP-GEM diseñada con el objetivo de automatizar la validación experimental del enfoque generativo. La herramienta desarrollada para validar experimentalmente el segundo enfoque está detallada en la tesis de Ingeniería en Sistemas de Sebastián Sosa [26]. El soporte teórico y los resultados obtenidos con el segundo enfoque fueron publicados en [27].

Además de estos dos enfoques centrales, se desarrollaron paralelamente otros trabajos que sirvieron como soporte a las dos líneas centrales y se exploraron otras pistas que contribuyeron a hacer un mapa del estado del arte y abrir caminos que pensamos continuar recorriendo en el futuro. Entre los primeros mencionaremos aquí el desarrollo de modelos predictivos con redes neuronales para WAF y de un framework para la generación automática de logs para el entrenamiento de modelos de aprendizaje automático. Entre los segundos, cabe mencionar la realización de una prueba de concepto basada en el framework de OpenMined para modelos de Machine Learning, que combina privacidad diferencial y encriptación homomórfica, una forma de encriptación que permite realizar cálculos sobre sus datos encriptados sin desencriptarlos primero, y la reconstrucción de sesiones a partir de web logs, lo que permitiría, en un trabajo de investigación futuro, aplicar mecanismos de privacidad y de detección de ataques basados en sesiones. Estos resultados, así como los datasets utilizados y código desarrollado también forman parte de los productos alcanzados (detallados en la lista de producciones y disponibles en REDI).

## Conclusiones y recomendaciones

La conclusión del proyecto es que es factible conciliar privacidad y utilidad predictiva en el caso de datos secuenciales definiendo apropiadamente el escenario de uso. Como sugerencia de aplicación en Uruguay de los resultados obtenidos, se propuso una instancia particular del esquema general diseñado en este proyecto y publicado en [27]. Este caso de uso es ilustrado de manera gráfica en el artículo de divulgación "Privacidad diferencial, o cuando la confusión viene a protegernos" (ver lista de producciones del proyecto). En palabras, planteamos un escenario en el que AGESIC cumple el rol de curador confiable, tanto para el ensemble de modelos predictivos (enfoque global) como para el tercero que envía consultas al ensemble (enfoque local). A modo de ejemplo, supongamos que varias organizaciones (organismos públicos, mutualistas, etc.) ponen a disposición de AGESIC sus modelos predictivos, obtenidos mediante el uso de técnicas de aprendizaje automático a partir de sus datos privados. De esta manera, aceptan que AGESIC pueda eventualmente tener acceso a sus datos sensibles (ya sea propios o de sus clientes, usuarios, socios, pacientes, etc.), pero con el compromiso de protegerlos frente a consultas hechas a los modelos por terceros. AGESIC cumple este compromiso mediante la disponibilización de consultas a esos modelos a través del mecanismo de privacidad diferencial provisto por PATE. Así, pone a resguardo los datos sensibles de las organizaciones participantes en el ensemble. Por otra parte, una organización externa que desea consultar el ensemble disponibilizado por AGESIC, ya sea para construir su propio modelo de IA o para obtener una respuesta que le permita tomar una decisión de negocios, detectar un fraude, bloquear un ataque o realizar un diagnóstico médico, puede utilizar a la propia AGESIC como su curador confiable. En este caso, AGESIC es responsable de aplicar el mecanismo de privacidad diferencial a los datos enviados por esta organización para la consulta, antes de ejecutar la consulta en el modelo disponibilizado. De esta manera, la privacidad de los datos sensibles de la organización consultante está garantizada frente a cualquier uso que puedan dar las organizaciones que los reciben. De esta manera, nos aseguramos que la privacidad de la información de todas las organizaciones participantes (tanto consultadas como consultantes) no sea vulnerada, dentro de los parámetros de privacidad acordados entre las organizaciones y AGESIC. La profundización de esta propuesta será objeto de investigaciones futuras.

**Referencias bibliográficas**


1. Iqbal, M.J.; Javed, Z.; Sadia, H.; Qureshi, I.A.; Irshad, A.; Ahmed, R.; Malik, K.; Raza, S.; Abbas, A.; Pezzani, R.; et al. Clinical applications of artificial intelligence and machine learning in cancer diagnosis: Looking into the future. Cancer Cell Int. 2021, 21, 270.
2. Kim, J.; Kim, J.; Thi Thu, H.L.; Kim, H. Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection. In Proceedings of the 2016 International Conference on Platform Technology and Service (PlatCon), Jeju, Korea, 15–17 February 2016; pp. 1–5.
3. Bontemps, L.; Cao, V.L.; McDermott, J.; Le-Khac, N. Collective Anomaly Detection Based on Long Short-Term Memory Recurrent Neural Networks. In International Conference on Future Data and Security Engineering; Dang, T.K., Wagner, R.R., Küng, J., Thoai, N., Takizawa, M., Neuhold, E.J., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2016; Volume 10018, pp. 141–152.
4. Thi, N.N.; Cao, V.L.; Le-Khac, N. One-Class Collective Anomaly Detection Based on LSTM-RNNs. Trans. Large Scale Data Knowl. Centered Syst. 2017, 36, 73–85.
5. Yin, C.; Zhu, Y.; Fei, J.; He, X. A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks. IEEE Access 2017, 5, 21954–21961.
6. Ruijer, E.; Détienne, F.; Baker, M.; Groff, J.; Meijer, A. The Politics of Open Government Data: Understanding Organizational Responses to Pressure for More Transparency. Am. Rev. Public Adm. 2020, 50, 260–274.
7. Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-Use of Public Sector Information. Available online: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32019L1024 (accessed on 5 August 2021).
8. Irvin, J.; Rajpurkar, P.; Ko, M.; Yu, Y.; Ciurea-Ilcus, S.; Chute, C.; Marklund, H.; Haghgoo, B.; Ball, R.L.; Shpanskaya, K.S.; et al. CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison. In Proceedings of the 33rd AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; pp. 590–597.
9. Gruschka, N.; Mavroeidis, V.; Vishi, K.; Jensen, M. Privacy Issues and Data Protection in Big Data: A Case Study Analysis under GDPR. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 5027–5033.
10. Rocher, L.; Hendrickx, J.; de Montjoye, Y. Estimating the success of re-identifications in incomplete datasets using generative models. Nat. Commun. 2019, 10, 3069.
11. Harmanci, A.; Gerstein, M. Quantification of private information leakage from phenotype-genotype data: Linking attacks. Nat. Methods 2016, 13, 251–256.
12. Narayanan, A.; Shmatikov, V. Robust de-anonymization of large sparse datasets. In Proceedings of the 2008 IEEE Symposium on Security and Privacy (sp 2008), Oakland, CA, USA, 18–21 May 2008; pp. 111–125.
13. Sweeney, L.; Abu, A.; Winn, J. Identifying participants in the personal genome project by name (a re-identification experiment). arXiv 2013, arXiv:1304.7605.
14. De Montjoye, Y.A.; Hidalgo, C.A.; Verleysen, M.; Blondel, V.D. Unique in the crowd: The privacy bounds of human mobility. Sci. Rep. 2013, 3, 1376.
15. Fredrikson, M.; Jha, S.; Ristenpart, T. Model Inversion Attacks that Exploit Confidence Information and Basic Countermeasures. In Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, Denver, CO, USA, 12–16 October 2015.
16. General Data Protection Regulation. Available online: https://gdpr-info.eu/ (accessed on 10 May 2021).
17. Chen, B.; Kifer, D.; LeFevre, K.; Machanavajjhala, A. Privacy-Preserving Data Publishing. Found. Trends Databases 2009, 2, 1–167.
18. Papernot, N.; Abadi, M.; Erlingsson, U.; Goodfellow, I.; Talwar, K. Semi-supervised knowledge transfer for deep learning from private training data. arXiv 2016, arXiv:1610.05755.
19. Papernot, N.; Song, S.; Mironov, I.; Raghunathan, A.; Talwar, K.; Erlingsson, Ú. Scalable private learning with pate. arXiv 2018, arXiv:1802.08908.
20. Dwork, C.; Roth, A. The Algorithmic Foundations of Differential Privacy. Found. Trends Theor. Comput. Sci. 2014, 9, 211–407.
21. Kachuee, M.; Fazeli, S.; Sarrafzadeh, M. ECG Heartbeat Classification: A Deep Transferable Representation. In Proceedings of the 2018 IEEE International Conference on Healthcare Informatics (ICHI), New York, NY, USA, 4–7 June 2018; pp. 443–444.

22. Moody, G.; Mark, R. The impact of the MIT-BIH Arrhythmia Database. IEEE Eng. Med. Biol. Mag. 2001, 20, 45–50.
23. Melville, P.; Shah, N.; Mihalkova, L.; Mooney, R.J. Experiments on ensembles with missing and noisy data. In International Workshop on Multiple Classifier Systems; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2004; Volume 3077, pp. 293–302.
24. Strauss, T.; Hanselmann, M.; Junginger, A.; Ulmer, H. Ensemble Methods as a Defense to Adversarial Perturbations against Deep Neural Networks. arXiv 2017, arXiv:1709.03423.
25. Visca, R. Estudio de modelos de privacidad de datos. Tesis de Maestría, Universidad ORT Uruguay, 2021. https://hdl.handle.net/20.500.12381/463
26. Sosa, S. Application of private aggregation of teacher ensembles framework for malicious web request detection. Trabajo final de Ingeniería en Sistemas, Universidad ORT Uruguay, 2021. https://hdl.handle.net/20.500.12381/459
27. Yovine, S.; Mayr, F.; Sosa, S.; Visca, R. An Assessment of the Application of Private Aggregation of Ensemble Models to Sensible Data. Mach. Learn. Knowl. Extr. 2021, 3, 788-801.https://hdl.handle.net/20.500.12381/456