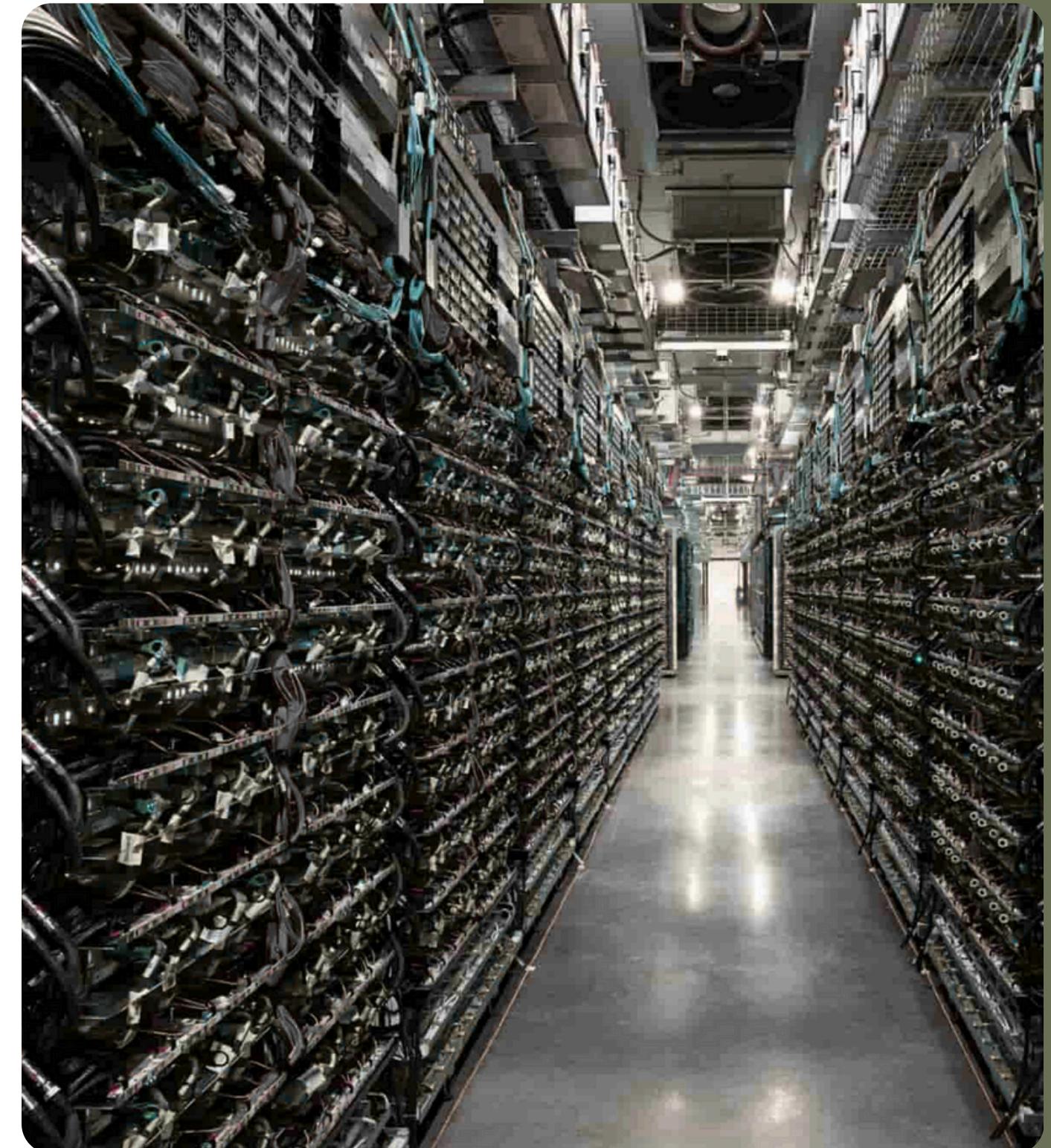


$\text{CO}_2$

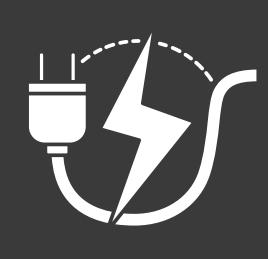
# FROM CLICKS TO CARBON: THE ENVIRONMENTAL TOLL OF RECOMMENDER SYSTEMS

GRUPO 06

BENITO OLIVA, JUSTO SOLIS, DANIEL VERA

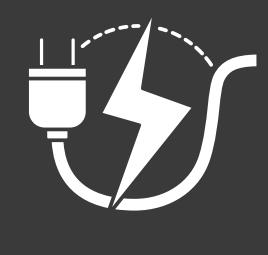


$\text{CO}_2$



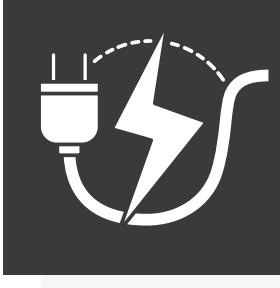
# 01. CONTEXTO

# CONTEXTO DEL PROBLEMA



- Las emisiones de carbono juegan un papel fundamental en el calentamiento global.
- *United Nations Framework Convention on Climate Change* (UNFCCC) tiene como objetivo limitar el aumento de temperatura.
- Implica directamente limitar las emisiones de carbono.

# CONTEXTO DEL PROBLEMA



- Creciente presión por reducir las emisiones de carbono.
- La comunidad de sistemas recomendadores no ha prestado mucha atención al respecto.
- Los algoritmos actuales consumen mayor energía, lo cual implica mayor huella de carbono.

# CONTEXTO DEL PROBLEMA



ItemKNN → MacridVAE

# CONTEXTO DEL PROBLEMA

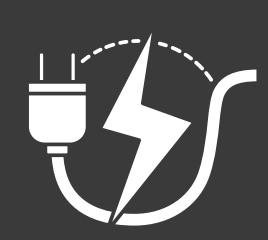


ItemKNN



MacridVAE

- Mayor consumo energético
- Mayor huella de carbono



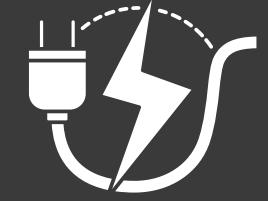
## O2. PROBLEMA DE RECOMENDACION

# PROBLEMA DE RECOMENDACIÓN



- Frente a la crisis climática y la mayor huella de carbono de los sistemas recomendadores actuales.
- Se plantea conocer cuales son los impactos ecológicos de la investigación en el pasado y el presente.

# PROBLEMA DE RECOMENDACIÓN



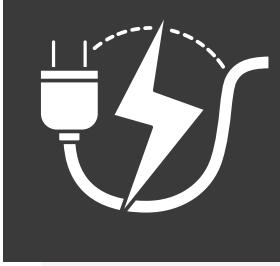
1

¿Cuál es la magnitud del consumo energético de un trabajo moderno de investigación sobre sistemas recomendadores?

2

¿Cuál es la relación entre el consumo de energía y el rendimiento de los algoritmos de los sistemas de recomendación tradicionales y los de aprendizaje profundo?

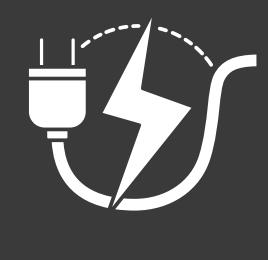
# PROBLEMA DE RECOMENDACIÓN



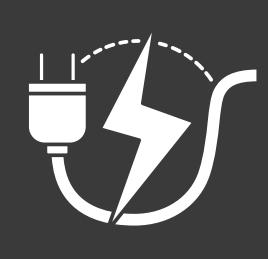
3

¿Cómo ha cambiado la huella de carbono de los experimentos de sistemas de recomendación con la transición de algoritmos tradicionales a algoritmos de aprendizaje profundo?

# PROBLEMA DE RECOMENDACIÓN

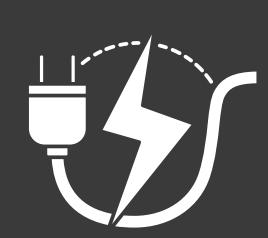


- Para responder las preguntas experimentan utilizando 23 algoritmos sobre 13 datasets.
- Estos se ejecutan en diferentes configuraciones de hardware.
- La elección de algoritmos y datasets se basa en la investigación sobre los RecSys del ACM de 2013 y 2023.



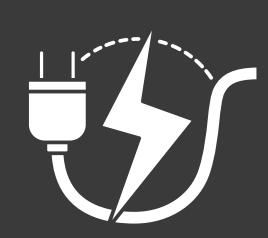
## 03. CONTRIBUCIÓN

# CONTRIBUCIÓN

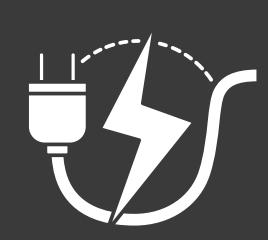


- Contribuye con análisis exhaustivo de las emisiones de carbono de RecSys.
- Evaluando como la evolución histórica (2013 a 2023), el hardware y la ubicación geográfica impactan las emisiones de carbono de las investigaciones.

# CONTRIBUCIÓN

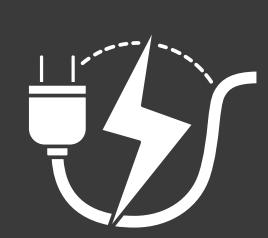


- También, es el primer trabajo en usar medidores físicos para la energía en lugar de software.
- Por último, contribuye al dar cuenta de la creciente emisión de la investigación en RecSys, promoviendo transparencia con el consumo energético y uso de prácticas más sostenibles.



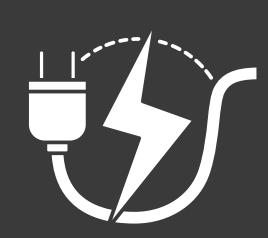
## 4. ESTADO DEL ARTE Y MARCO TEÓRICO

# MARCO TEÓRICO



- CO<sub>2</sub>e (CO<sub>2</sub> equivalentes): Medida estandarizada para expresar el impacto de distintos gases de efecto invernadero.
- kWh: Unidad de energía usada para medir consumo eléctrico en el tiempo.

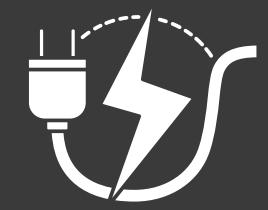
# MARCO TEÓRICO



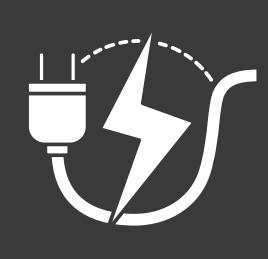
kWh → CO<sub>2</sub>e

- Las emisiones se miden a través de los kWh que gastan en total los algoritmos.
- Es la contaminación de la producción de dicha energía.

# ESTADO DEL ARTE



- Solo un estudio previo analizó la huella de carbono de RecSys. Sin embargo, lo realizó mediante software.
- Existen alternativas energéticamente eficientes a la validación cruzada y estudios sobre consumo de recursos de sistemas automáticos de recomendación.
- No obstante, no hay evaluación detallada ni reproducible como la desarrollada en este paper.



## 5. DETALLE SOLUCIÓN

# ALGORITMOS

2013 VS 2023

- Clásicos: ItemKnn, UserKnn, entre otros.
- Basados en DeepLearning: MacridVae, DGCF, entre otros.

**Table 2: Algorithm implementations used in our experiments.**

Library	Executed on	Algorithm
RecBole	GPU	BPR, DGCF, DMF, ItemKNN, LightGCN, MacridVAE, MultiVAE, NAIS, NCL, NeuMF, NGCF, Popularity, RecVAE, SGL
RecPack	CPU	ItemKNN, NMF, SVD
LensKit	CPU	ImplicitMF, ItemKNN, UserKNN, Popularity, BiasedMF, FunkSVD

# DATASETS

13 con 5 de amazon

5 en el 2013, 12 en el 2023

Dataset Name	Users	Items	Interactions	Density
Amazon2018[39] Books	105,436	151,802	1,724,703	0.0108
Amazon2018[39] CDs-And-Vinyl	71,943	107,546	1,377,008	0.0178
Amazon2018[39] Electronics	62,617	187,288	1,476,535	0.0126
Amazon2018[39] Sports-And-Outdoors	69,781	185,024	1,498,609	0.0116
Amazon2018[39] Toys-And-Games	75,856	192,326	1,686,250	0.0116
Gowalla[12]	64,115	164,532	2,018,421	0.0191
Hetrec-LastFM[8]	1,090	3,646	52,551	1.3223
MovieLens[22] 100K	943	1,349	99,287	7.8049
MovieLens[22] 1M	6,040	3,416	999,611	4.8448
MovieLens[22] Latest-Small	610	3,650	90,274	4.0545
Netflix <sup>5</sup>	11,927	11,934	5,850,559	4.1103
Retailrocket <sup>6</sup>	22,178	17,803	240,938	0.0610
Yelp-2018 <sup>7</sup>	213,170	94,304	3,277,931	0.0163

Year and Prediction Type	Algorithms	Datasets
2013 Rating Prediction	ItemKNN <sup>LK</sup> , UserKNN <sup>LK</sup> , BiasedMF <sup>LK</sup> , FunkSVD <sup>LK</sup>	Movielens-100K, Movielens-1M, Netflix
2013 Top-N Ranking Prediction	ImplicitMF <sup>LK</sup> , ItemKNN <sup>LK</sup> , UserKNN <sup>LK</sup> , Popularity <sup>LK</sup> , ItemKNN <sup>RP</sup> , NMF <sup>RP</sup> , SVD <sup>RP</sup> , BPR <sup>RB</sup> , ItemKNN <sup>RB</sup> , Popularity <sup>RB</sup>	Hetrec-Lastfm, Movielens-100K, Movielens-1M, Gowalla
2023 Top-N Ranking Prediction	ImplicitMF <sup>LK</sup> , ItemKNN <sup>LK</sup> , UserKNN <sup>LK</sup> , Popularity <sup>LK</sup> , ItemKNN <sup>RP</sup> , NMF <sup>RP</sup> , SVD <sup>RP</sup> , BPR <sup>RB</sup> , DGCF <sup>RB</sup> , DMF <sup>RB</sup> , ItemKNN <sup>RB</sup> , LightGCN <sup>RB</sup> , MacridVAE <sup>RB</sup> , MultiVAE <sup>RB</sup> , NAIS <sup>RB</sup> , NCL <sup>RB</sup> , NeuMF <sup>RB</sup> , NGCF <sup>RB</sup> , Popularity <sup>RB</sup> , RecVAE <sup>RB</sup> , SGL <sup>RB</sup>	Gowalla, Hetrec-LastFM, MovieLens: -100K, -1M, -Latest-Small, Amazon2018: -Electronics, -Toys-And-Games, -CDs-And-Vinyl, -Sports-And-Outdoors, -Books, Yelp-2018, Retailrocket

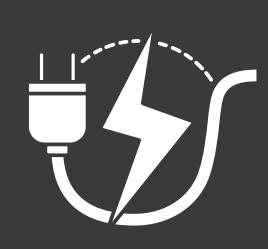
# HARDWARE

**Table 4: Hardware used to compare hardware efficiency.**

Computer and Year	CPU	GPU	RAM in GB	Storage in TB
Modern Workstation 2023	Intel Xeon W-2255 @ 3.70 GHz	NVIDIA GeForce RTX 3090	256	2
Mac Studio 2022	M1 Ultra	M1 Ultra	64	1
MacBook Pro 2020	M1	M1	16	1
Legacy Workstation 2013	Intel Core i7-6700K @ 4.00GHz	NVIDIA GeForce GTX 980 Ti	128	1

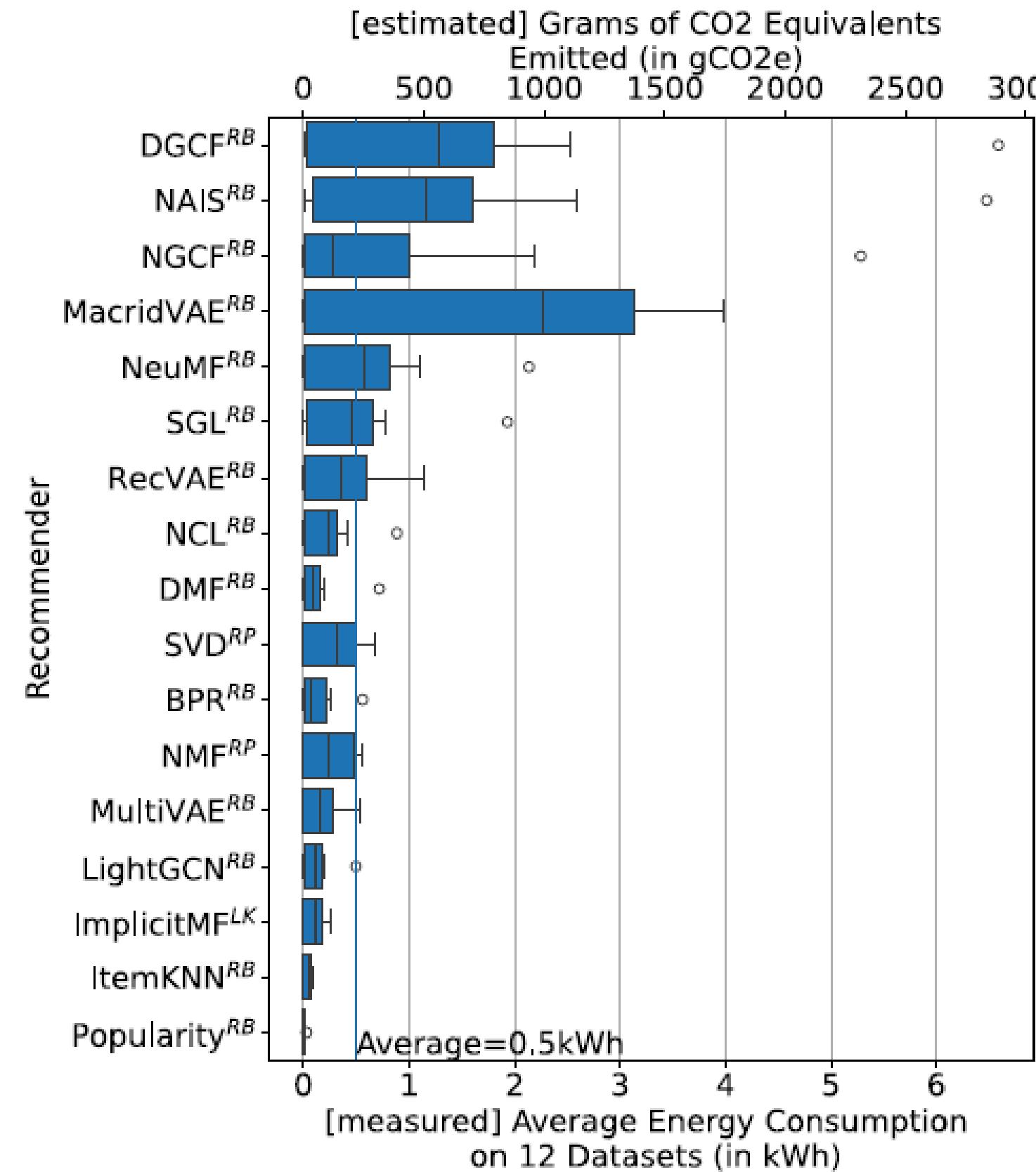
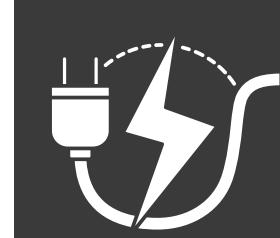
# SUPUESTOS

- Lugar. Calculado en Suecia
- Falta información (código)
- Uso de datos externos (Ember)



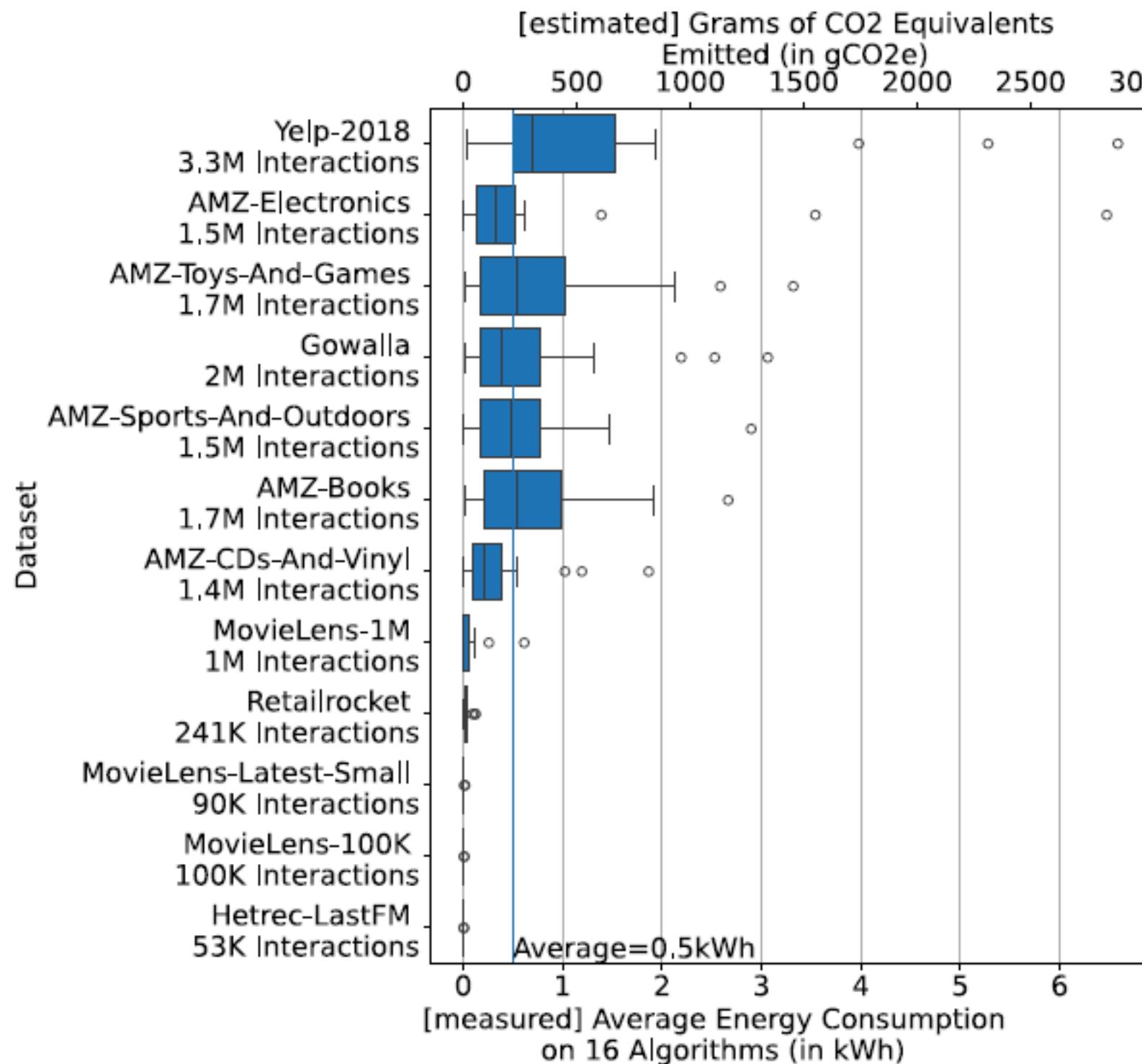
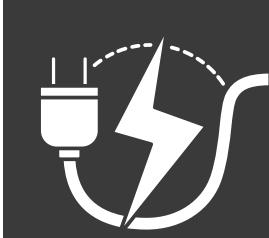
## 6. EVALUACIÓN

# CONSUMO ENERGÉTICO DE ALGORITMOS



- Consumo energético promedio: 0,51 kWh
- *Popularity* e *ItemKNN* con el **menor consumo energético promedio** (0,007 y 0,04 kWh).
- Algoritmos *deep learning* con **bajo consumo de energía**: DMF (0,13 kWh) y LightCGN (0,12 kWh).
- Algoritmos con **mayor consumo de energía**: MacridVAE (1,79 kWh) y DGCF (1,45 kWh).

# CONSUMO ENERGÉTICO EN DATASETS



- Consumo energético de un algoritmo **aumenta** al aplicarlo en datasets con más registros.
  - *Popularity* (0,000036 kWh en *MovieLens-Latest-Small* v/s 0,03 kWh en *Yelp-2018*).
  - DGCF (0,005 kWh en *Hetrec-LastFM* v/s 6,6 kWh en *Yelp-2018*).

# CONSUMO ENERGÉTICO PIPELINE



$$7 \times 3 \times 16 \times 0,51 \text{ kWh} = \underline{\underline{171,36 \text{ kWh}}}$$

- 7 **algoritmos de recomendación**, por cada *pipeline* (promedio).
- Se usaron 3 **datasets** para evaluar los algoritmos (promedio).
- Se realizaron 16 **repeticiones** por algoritmo, para optimizar los hiperparámetros (promedio).
- Promedio consumo energético: **0,51 kWh**.

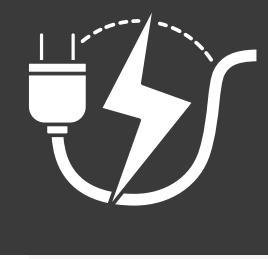
# CONSUMO ENERGÉTICO PAPER 2023



$$171,36 \text{ kWh} \times 40 = \underline{\underline{6854,4 \text{ kWh}}}$$

- Consumo energético *pipeline*: **171,36 kWh**.
- Mediana de los factores estimados de sobrecarga de consumo energético: **40**.

# HUELLA DE CARBONO (PAPER 2023)

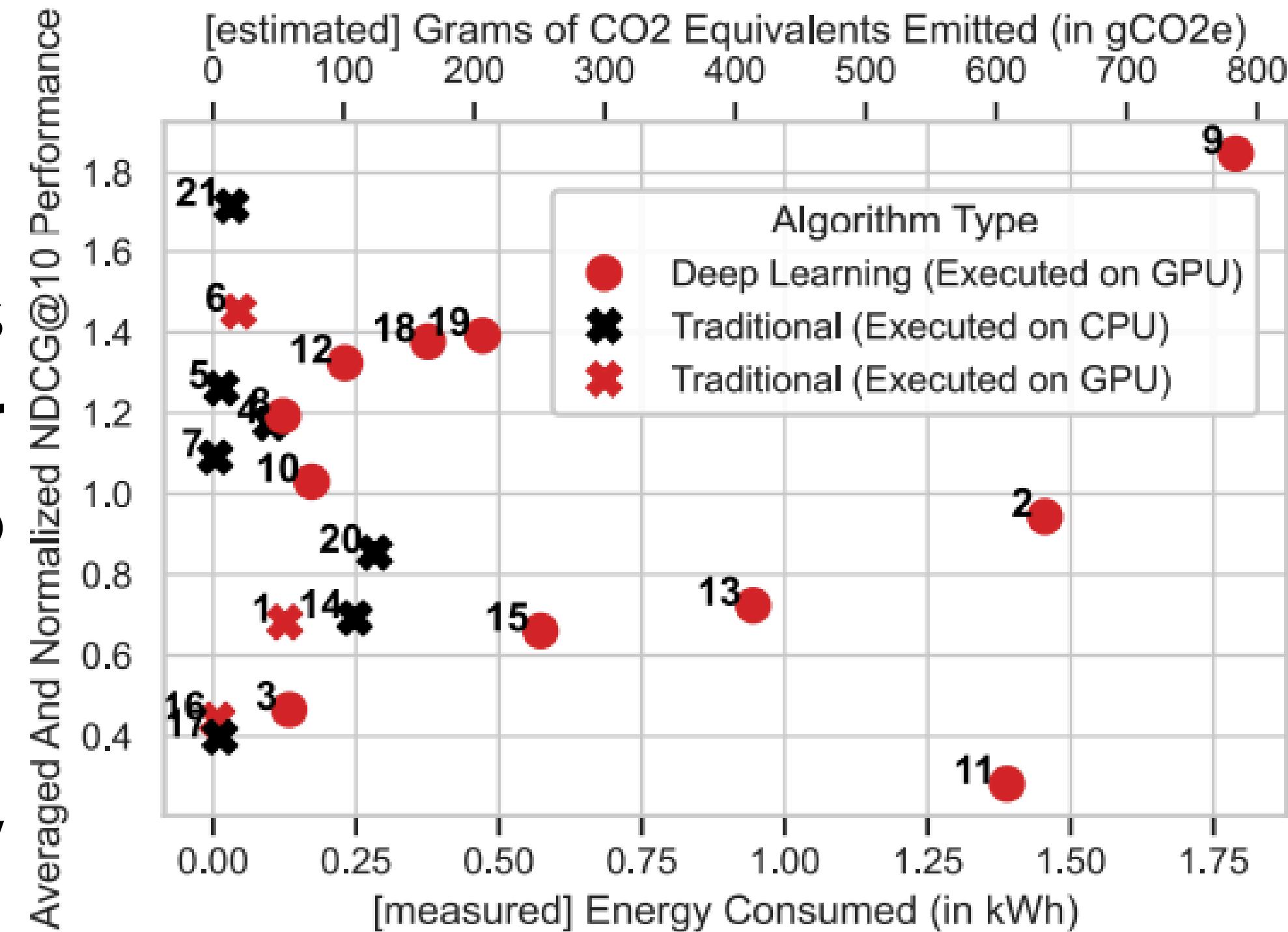


$$6854,4 \text{ kWh} \times 481 \text{ gCO}_2\text{e} \times 269 = \underline{\underline{886,9 \text{ TCO}_2\text{e}}}$$

- Consumo energético paper 2023: **6854,4 kWh**.
- Factor de conversión gCO<sub>2</sub>e a kWh: **481 gCO<sub>2</sub>e** (promedio mundial 2023).
- Envíos a ACM RecSys 2023: **269** (un envío involucra ejecución de experimentos).
- **iEsto es equivalente a 384 vuelos (NY a Melbourne), o lo capturado por un árbol en 80.600 años!**

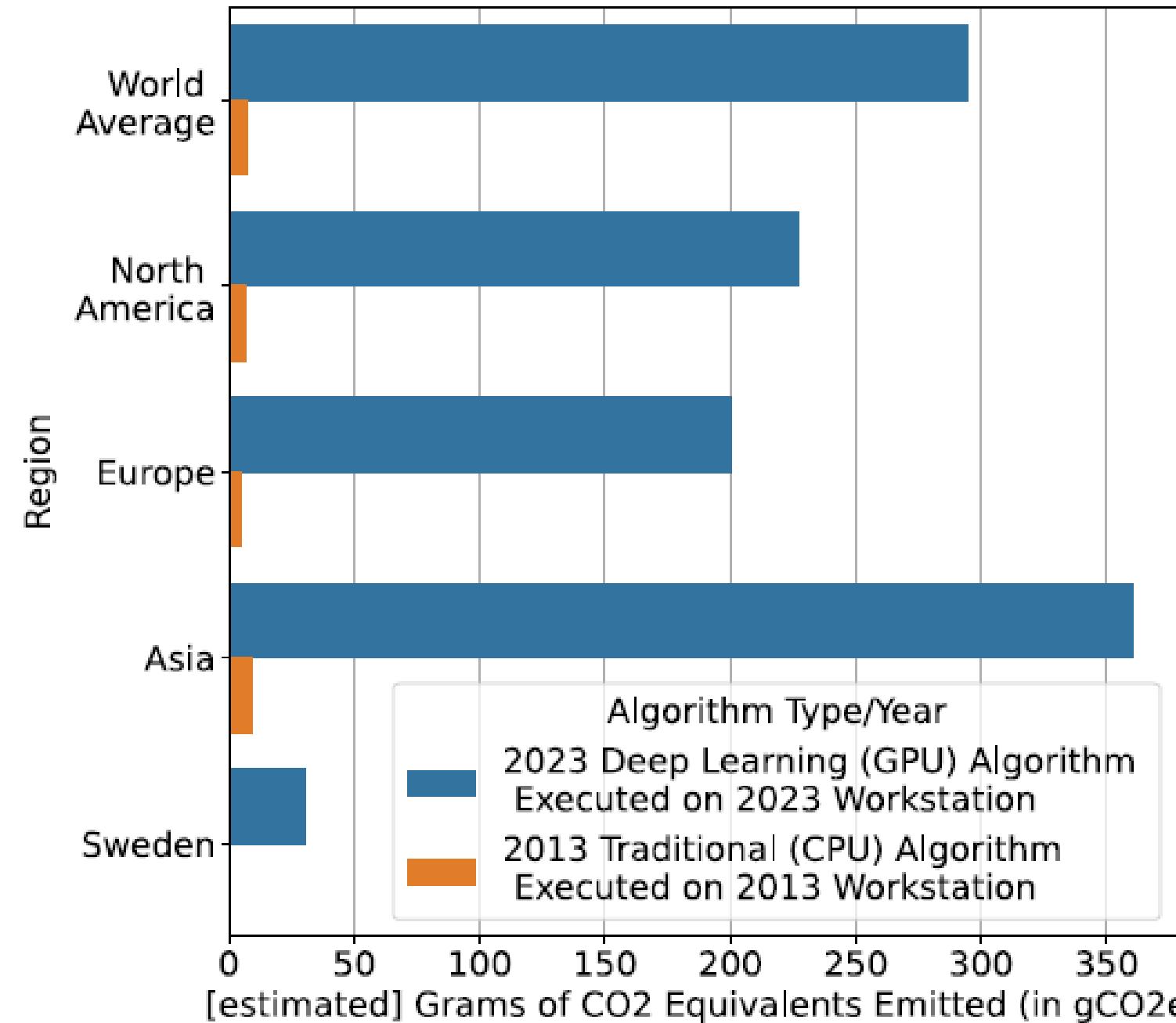
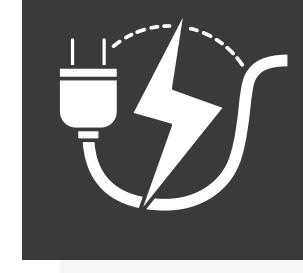
# EVALUACIÓN

- Algoritmos ejecutados en CPU tienen **menor consumo energético** que los ejecutados en GPU.
- **Implementación y complejidad** son factores importantes en el consumo de energía de los algoritmos.



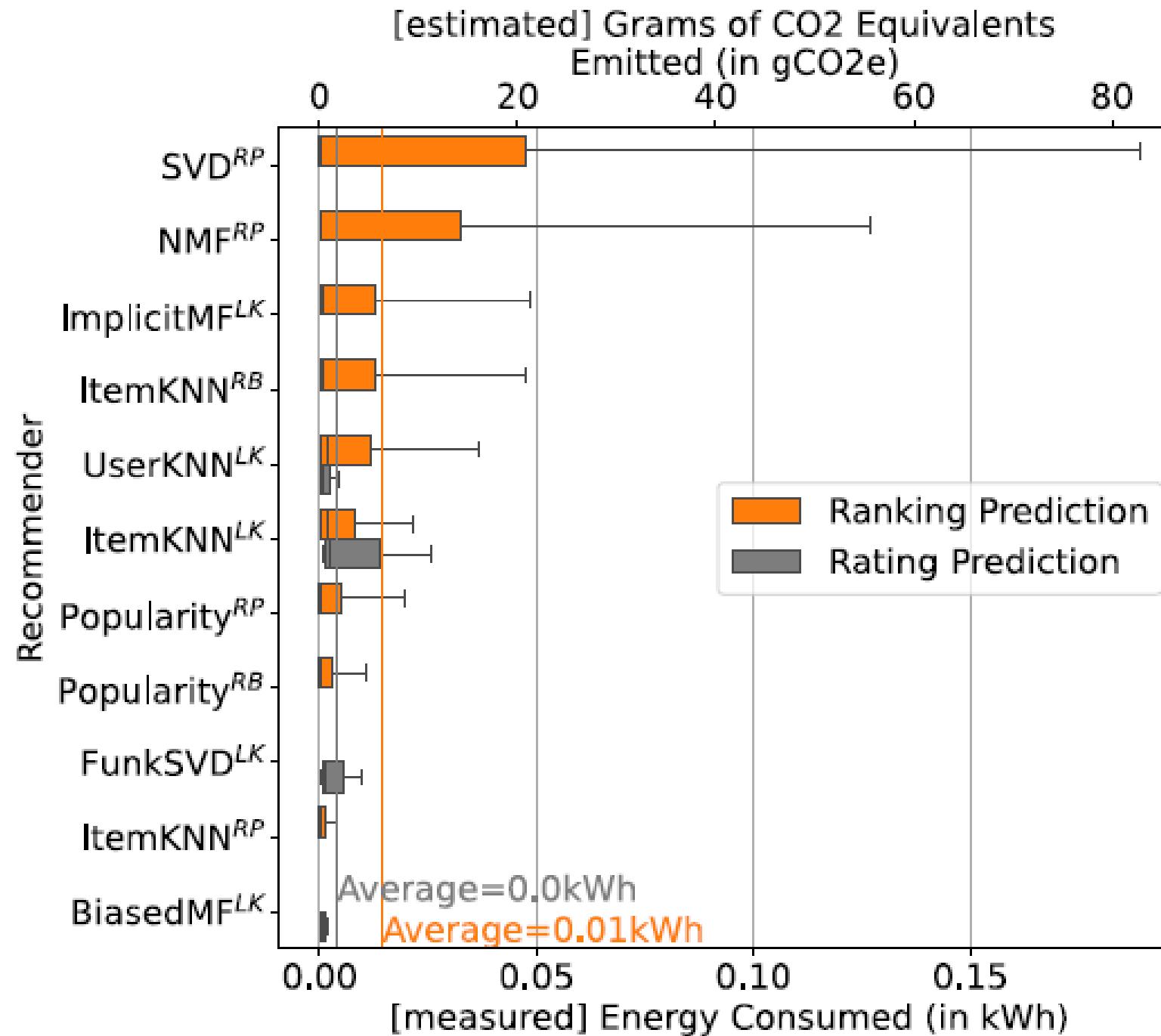
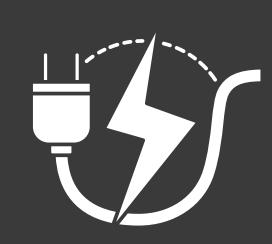
Recommender		
$\times$	$BPR^{RB}$ (1)	$\bullet$ $LightGCN^{RB}$ (8)
$\bullet$	$DGCF^{RB}$ (2)	$\bullet$ $MacridVAE^{RB}$ (9)
$\bullet$	$DMF^{RB}$ (3)	$\bullet$ $MultiVAE^{RB}$ (10)
$\times$	$ImplicitMF^{LK}$ (4)	$\bullet$ $NAIS^{RB}$ (11)
$\times$	$ItemKNN^{LK}$ (5)	$\bullet$ $NCL^{RB}$ (12)
$\times$	$ItemKNN^{RB}$ (6)	$\bullet$ $NGCF^{RB}$ (13)
$\times$	$ItemKNN^{RP}$ (7)	$\times$ $NeuMF^{RB}$ (15)
		$\times$ $Popularity^{RB}$ (16)
		$\times$ $Popularity^{RP}$ (17)
		$\bullet$ $RecVAE^{RB}$ (18)
		$\bullet$ $SGL^{RB}$ (19)
		$\times$ $SVD^{RP}$ (20)
		$\times$ $UserKNN^{LK}$ (21)

# HUELLA DE CARBONO PAPER 2013 V/S 2023



- Aumenta 42 veces la huella de carbono con experimentos de RecSys en 2023 v/s 2013 (7,09 gCO<sub>2</sub>e (2013) v/s 294,90 gCO<sub>2</sub>e (2023)).
- El incremento del uso de energía limpia con mejor hardware no compensa el aumento de emisiones de carbono.

# SOLUCIÓN



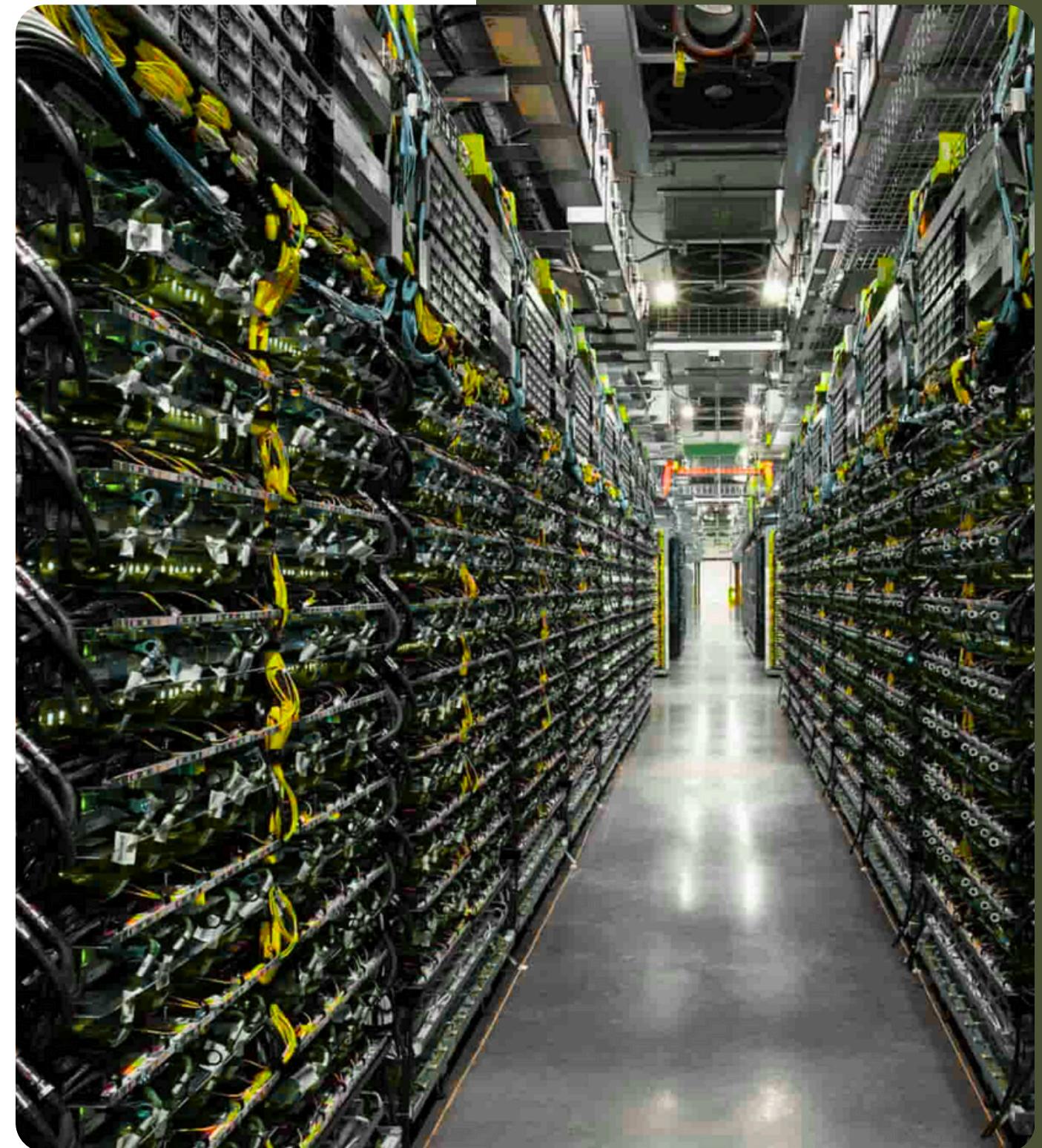
- Ningún algoritmo usa más de 0,2 kWh en su ejecución.
- Uso energético promedio: **0,1 kWh**.

$\text{CO}_2$

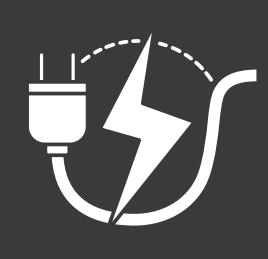
# FROM CLICKS TO CARBON: THE ENVIRONMENTAL TOLL OF RECOMMENDER SYSTEMS

GRUPO 06

BENITO OLIVA, JUSTO SOLIS, DANIEL VERA

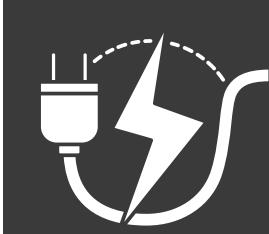


$\text{CO}_2$

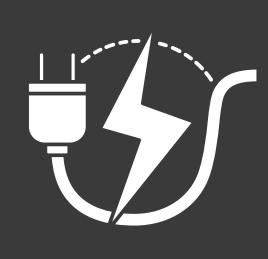


## 8. REFERENCIAS

# REFERENCIAS

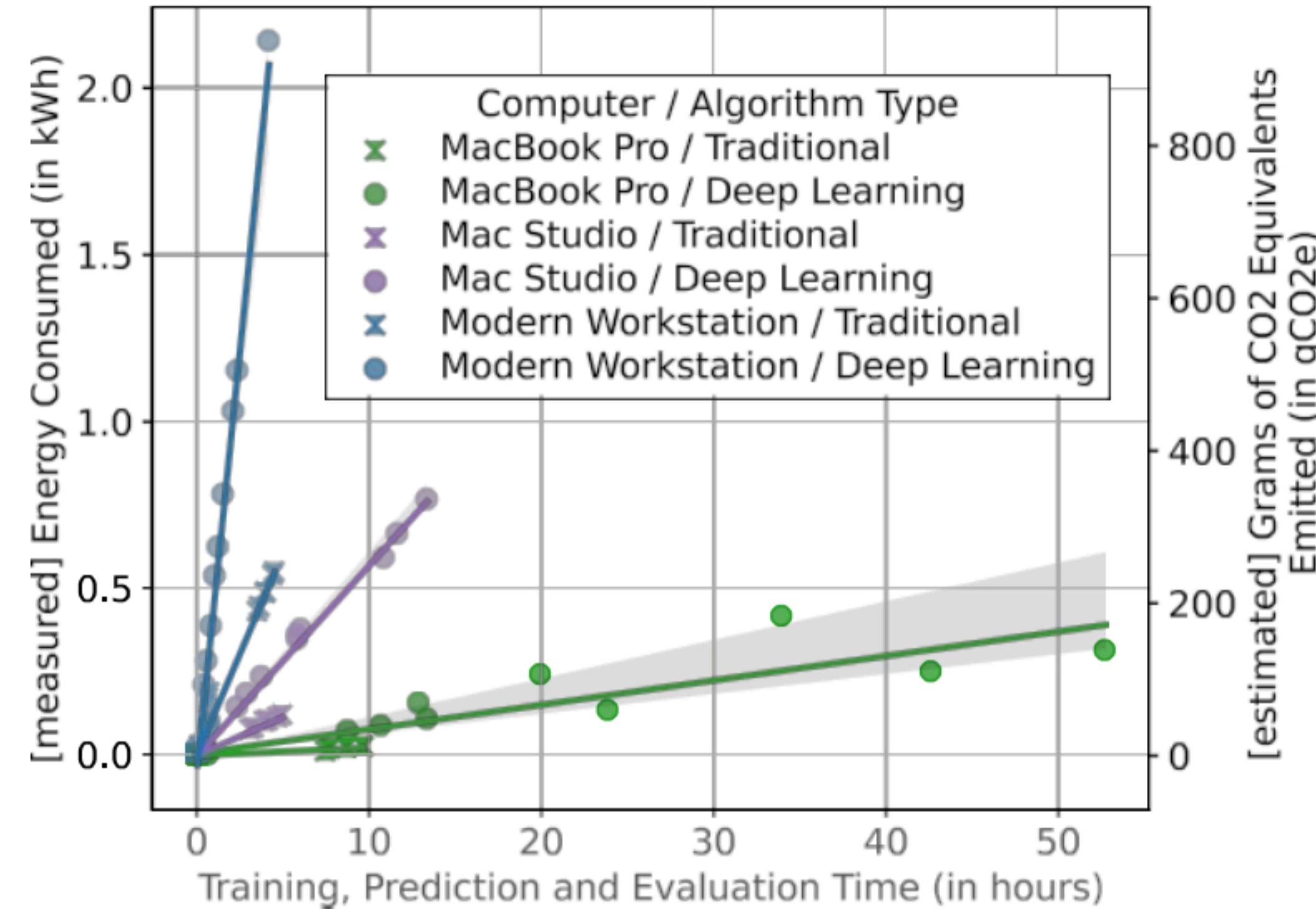
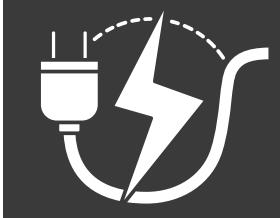


- Alan Said, Joeran Beel, Tobias Vente, Lukas Wegmeth. 2024. From Clicks to Carbon: The Environmental Toll of Recommender Systems. <https://arxiv.org/pdf/2408.08203>
- Joeran Beel, Lukas Wegmeth y Tobias Vente. 2024. e-fold cross-validation: A computing and energy-efficient alternative to k-fold cross-validation with adaptive folds. <https://doi.org/10.31219/osf.io/exw3j>
- Alexander Felfernig, Manfred Wundara, Thi Ngoc Trang Tran, Seda Polat-Erdeniz, Sebastian Lubos, Merfat El Mansi, Damian Garber y Viet Man Le. 2023. Recommender systems for sustainability: overview and research issues. *Frontiers in Big Data* 6 (30 oct. 2023). <https://doi.org/10.3389/fdata.2023.1284511> Copyright de la editorial: Copyright © 2023 Felfernig, Wundara, Tran, Polat-Erdeniz, Lubos, El Mansi, Garber y Le
- Giuseppe Spillo, Allegra De Filippo, Cataldo Musto, Michela Milano y Giovanni Semeraro. 2023. Towards Sustainability-aware Recommender Systems: Analyzing the Trade-off Between Algorithms Performance and Carbon Footprint. En Proceedings of the 17th ACM Conference on Recommender Systems (Singapur, Singapur) (RecSys '23). Association for Computing Machinery, Nueva York, NY, EE.UU., 856-862. <https://doi.org/10.1145/3604915.3608840>
- Michelle Cain, John Lynch, Myles R. Allen, Jan S. Fuglestvedt, David J. Frame y Adrian H Macey. 2019. Improved calculation of warming-equivalent emissions for short-lived climate pollutants. *npj Climate and Atmospheric Science* 2, 1 (Sept. 2019). <https://doi.org/10.1038/s41612-019-0086-4>
- Katharine Ricke, Laurent Drouet, Ken Caldeira y Massimo Tavoni. 2018. Country-level social cost of carbon  
• . *Nature Climate Change* 8, 10 (Sept. 2018), 895-900. <https://doi.org/10.1038/s41558-018-0282-y>
- Showmick Guha Paul, Arpa Saha, Mohammad Shamsul Arefin, Touhid Bhuiyan, Al Amin Biswas, Ahmed Wasif Reza, Naif M. Alotaibi, Salem A. Alyami, y Mohammad Ali Moni. 2023. A Comprehensive Review of Green Computing: Past, Present, and Future Research. *IEEE Access* 11 (2023), 87445-87494. <https://doi.org/10.1109/ACCESS.2023.3304332>
- Yukta Mehta, Rui Xu, Benjamin Lim, Jane Wu y Jerry Gao. 2023. A Review for Green Energy Machine Learning and AI Services. *Energies* 16, 15 (2023), 5718.
- Aimee Van Wynsberghe. 2021. Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics* 1, 3 (2021), 213-218.
- Likang Wu, Zhi Zheng, Zhaopeng Qiu, Hao Wang, Hongchao Gu, Tingjia Shen, Chuan Qin, Chen Zhu, Hengshu Zhu, Qi Liu, Hui Xiong y Enhong Chen. 2023. A Survey on Large Language Models for Recommendation. *arXiv:2305.19860 [cs.IR]*

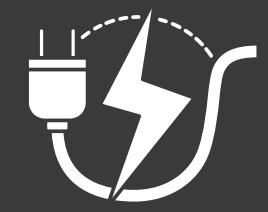


## 9 . ANEXOS

# ANEXOS



# ANEXOS



## Annual CO<sub>2</sub> emissions

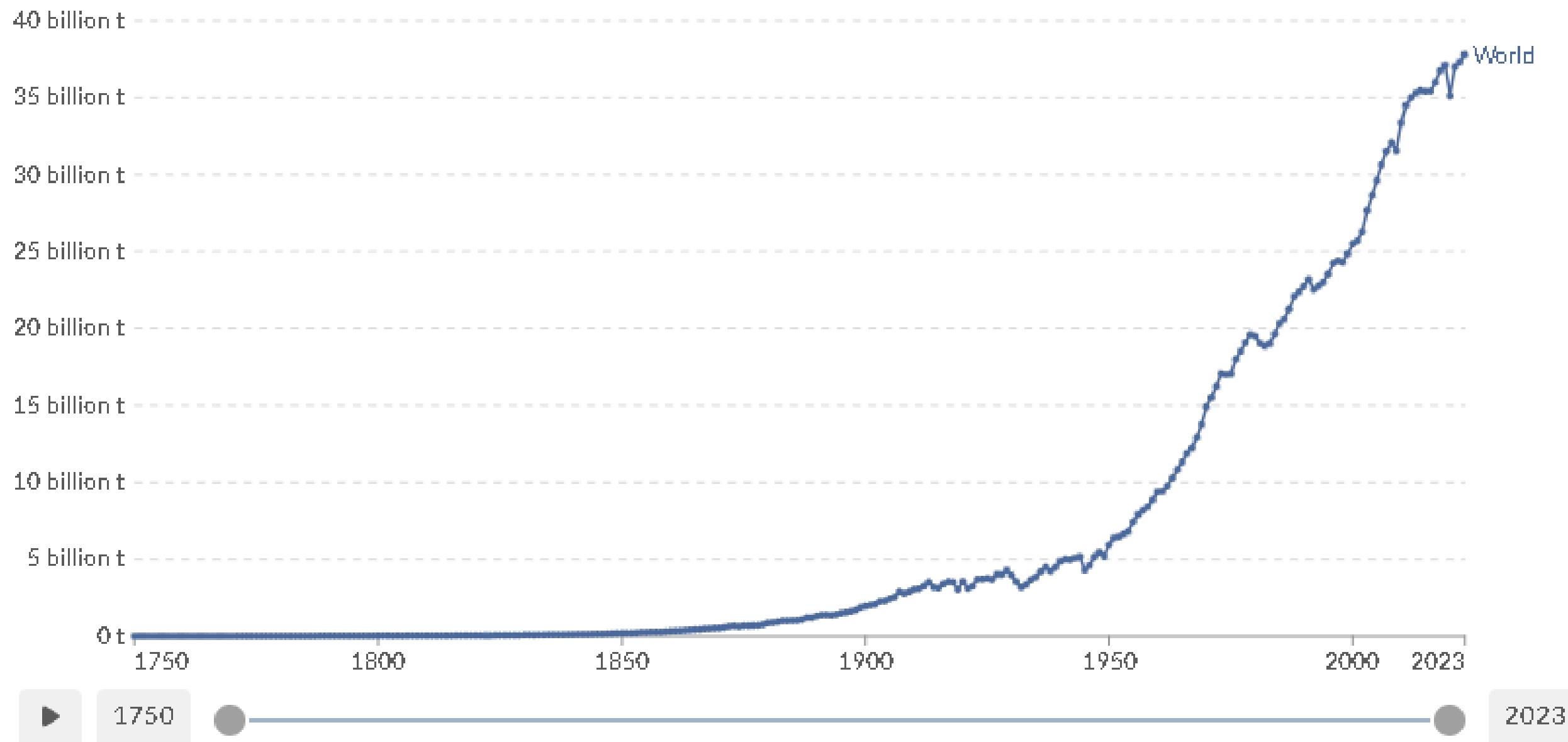
Carbon dioxide (CO<sub>2</sub>) emissions from fossil fuels and industry. Land-use change is not included.

Our World  
in Data

Table Map Chart

Edit countries and regions

Settings



Data source: Global Carbon Budget (2024) – [Learn more about this data](#)  
OurWorldInData.org/co2-and-greenhouse-gas-emissions | CC BY

Download

Share

Enter full-screen