

Empowering AI through frugality

Amparo Alonso-Betanzos

CITIC-University of A Coruña.
SPAIN



ICPRAM 2025

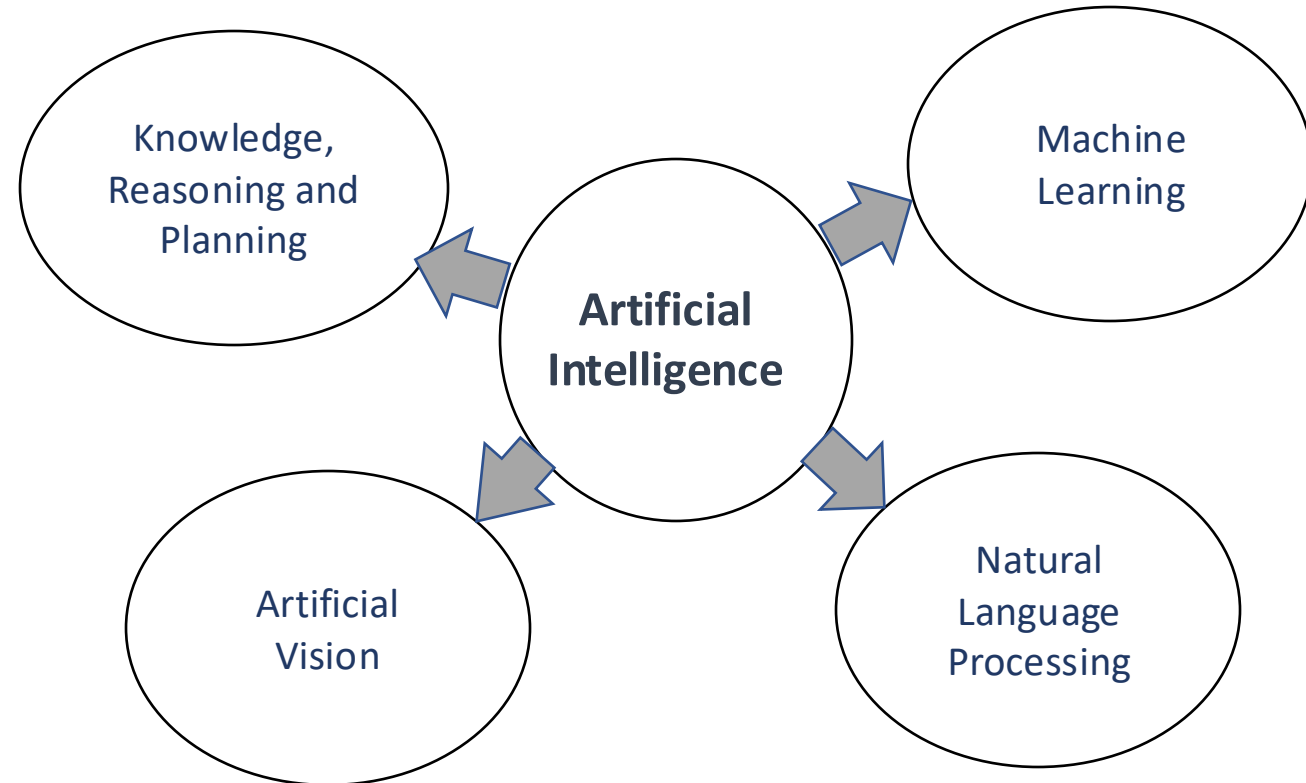
*14th International Conference on Pattern Recognition
Applications and Methods*

Porto, Portugal

23 - 25 February, 2025

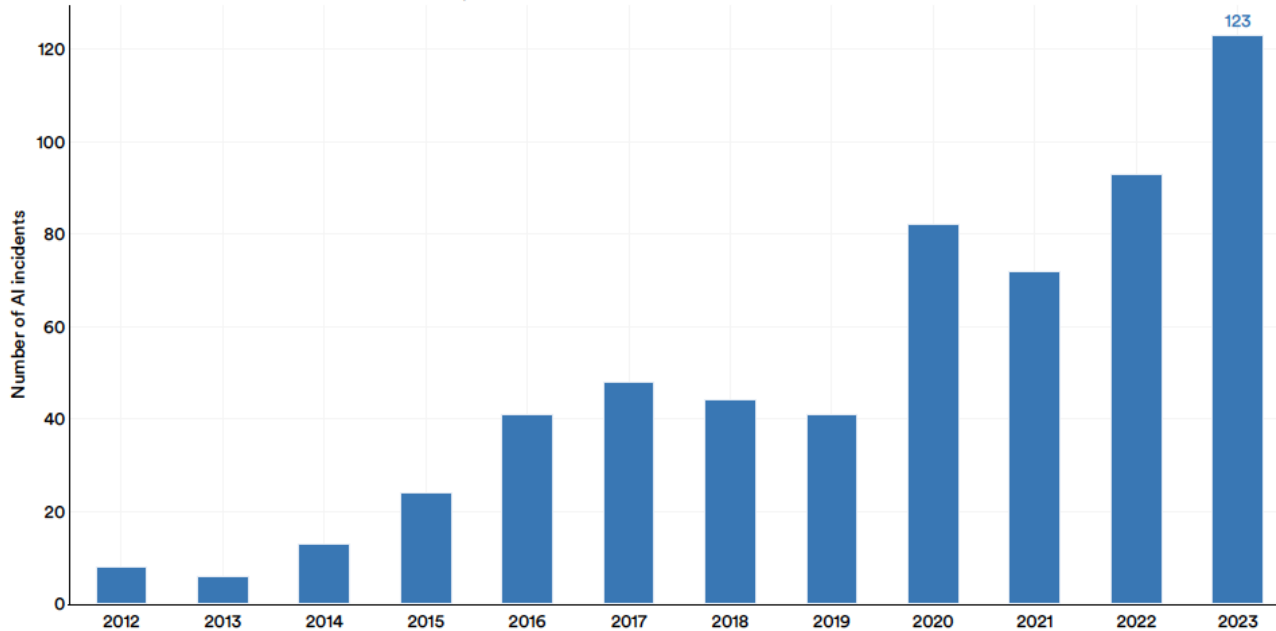


Systems designed by humans which, given a complex goal, act in their environment (physical or digital) by perceiving this environment through data acquisition, interpreting the collected structured or unstructured data, reasoning about the knowledge derived from this data, and deciding the best actions to take.



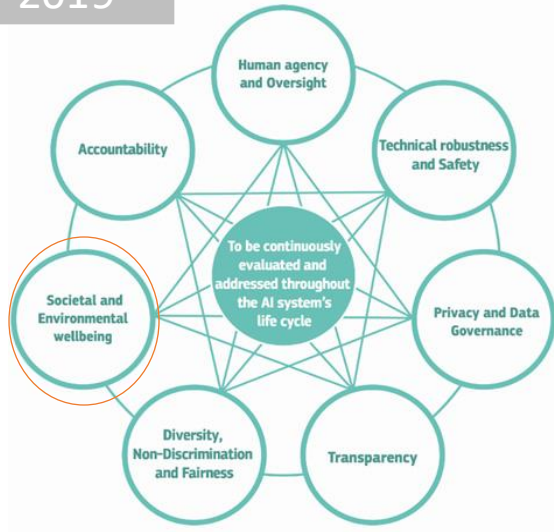
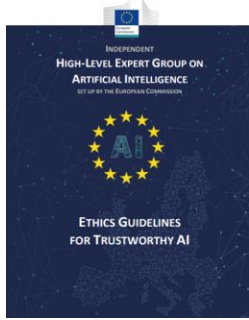
Number of reported AI incidents, 2012–23

Source: AI Incident Database (AIID), 2023 | Chart: 2024 AI Index report



- **Myth:** Technological singularity is the most relevant problem
- **Reality:** Privacy, bias, the possibility of manipulation and exhaustive surveillance, responsibility, **sustainability**, and technological monopoly are real true problems now.
- We need companies and public powers with ethical and moral codes, and more and better informed and critical citizens.

2019

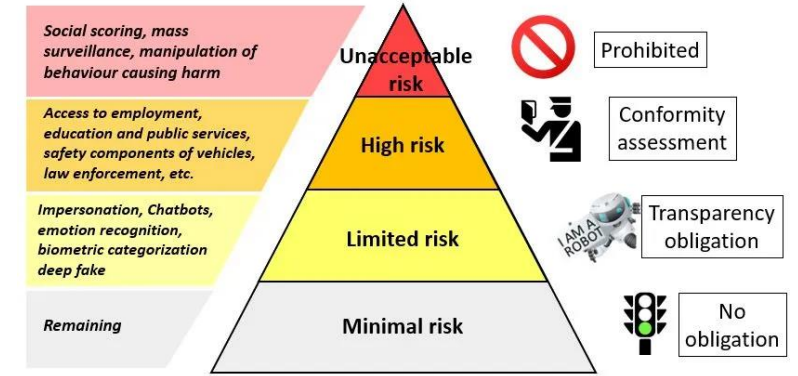


2021



EU Parliament's priority is to make sure that AI systems used in the EU are safe, transparent, traceable, non-discriminatory and environmentally friendly

EU Artificial Intelligence Act: Risk levels



2023



Article 28b(2)(d) required providers of foundation models to adhere to standards for **reducing energy and resource use**, improving **energy efficiency**, and enabling measurement and logging of **environmental impact**

2024



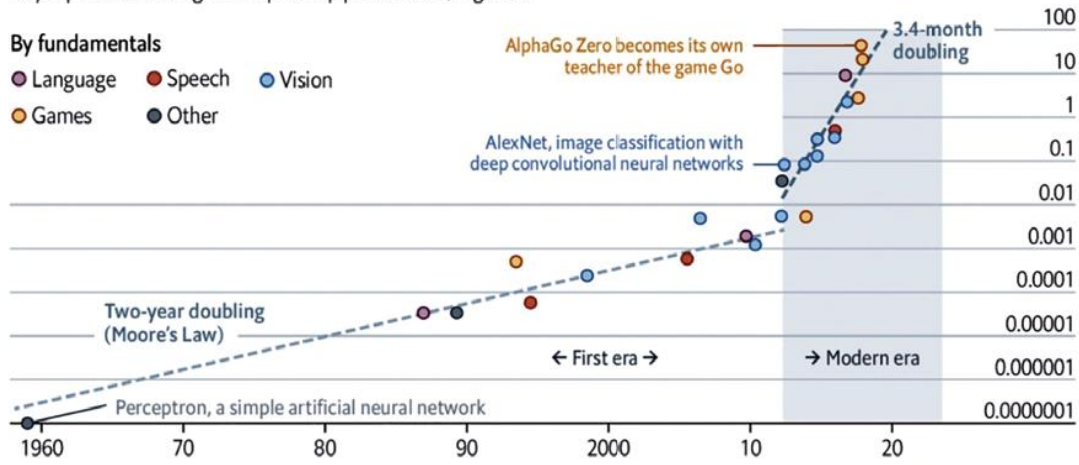
Deep and steep

Computing power used in training AI systems

Days spent calculating at one petaflop per second*, log scale

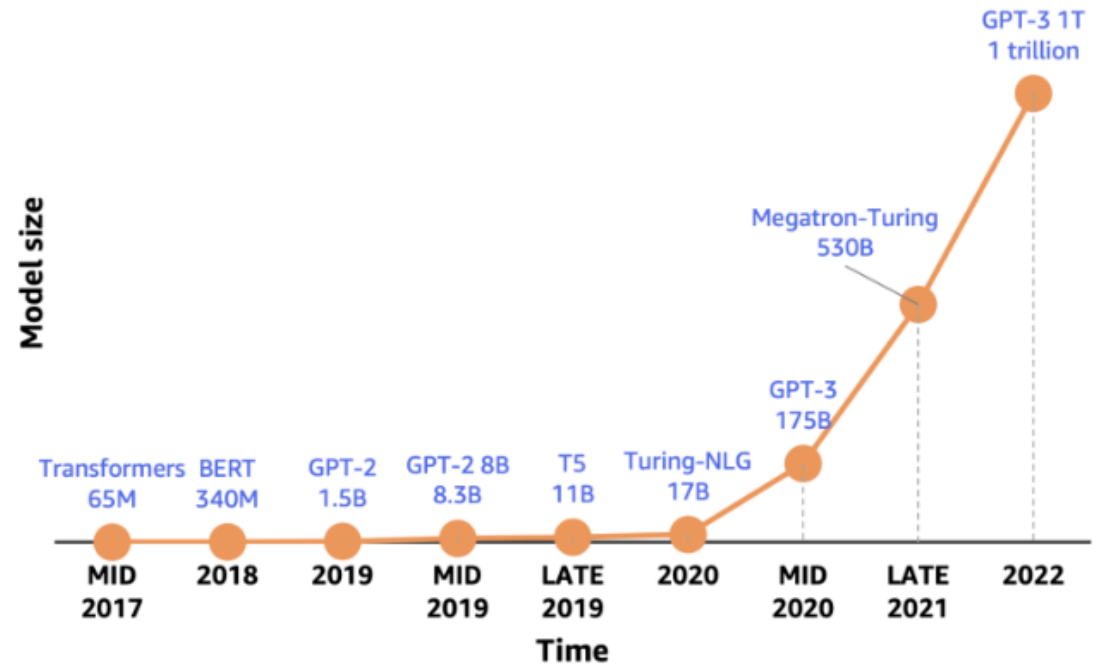
By fundamentals

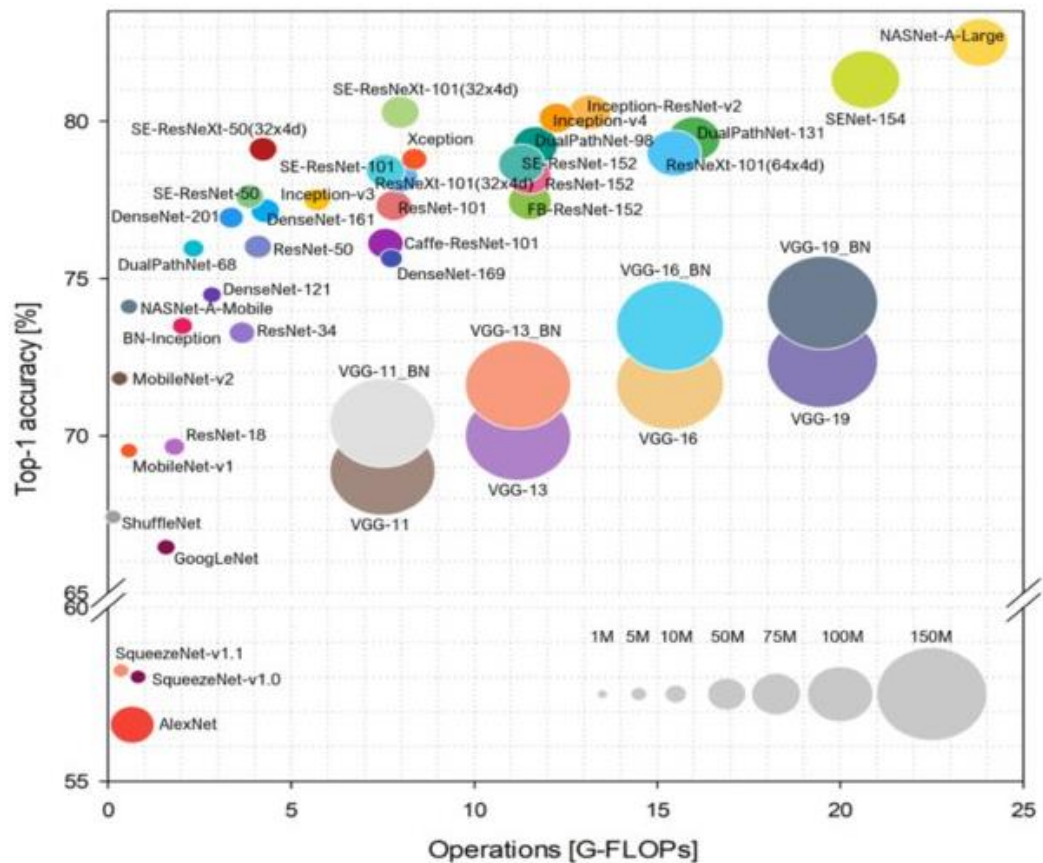
- Language
- Speech
- Vision
- Games
- Other



Source: OpenAI
The Economist

15,000x increase in 5 years





Model	RMSE (Performance)	gCO2 (Emissions)
Random	1,707	N/A
Naive (Dyad Average)	0,958	N/A
MF	0,840	56,449
GC-MC	0,846	243,349
Bayesian SVD++	0,830	358,900
Glocal-K	0,826	2038,492

Annotations: +1% perf (green arrow from MF to GC-MC), x10 CO2 (red arrow from MF to Glocal-K)

Training:

- 1.300.000kWh
- 552 ton. CO2
- 700.000 l. water



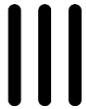
Annual energy consumption of 126 houses in Denmark



700.000 km

Access January 2023:

- 590 milions



Electricity consumption of 175,000 people



CO2 equivalent emissions (tonnes) by select machine learning models and real-life examples, 2020–23

Source: AI Index, 2024; Luccioni et al., 2022; Strubell et al., 2019 | Chart: 2024 AI Index report

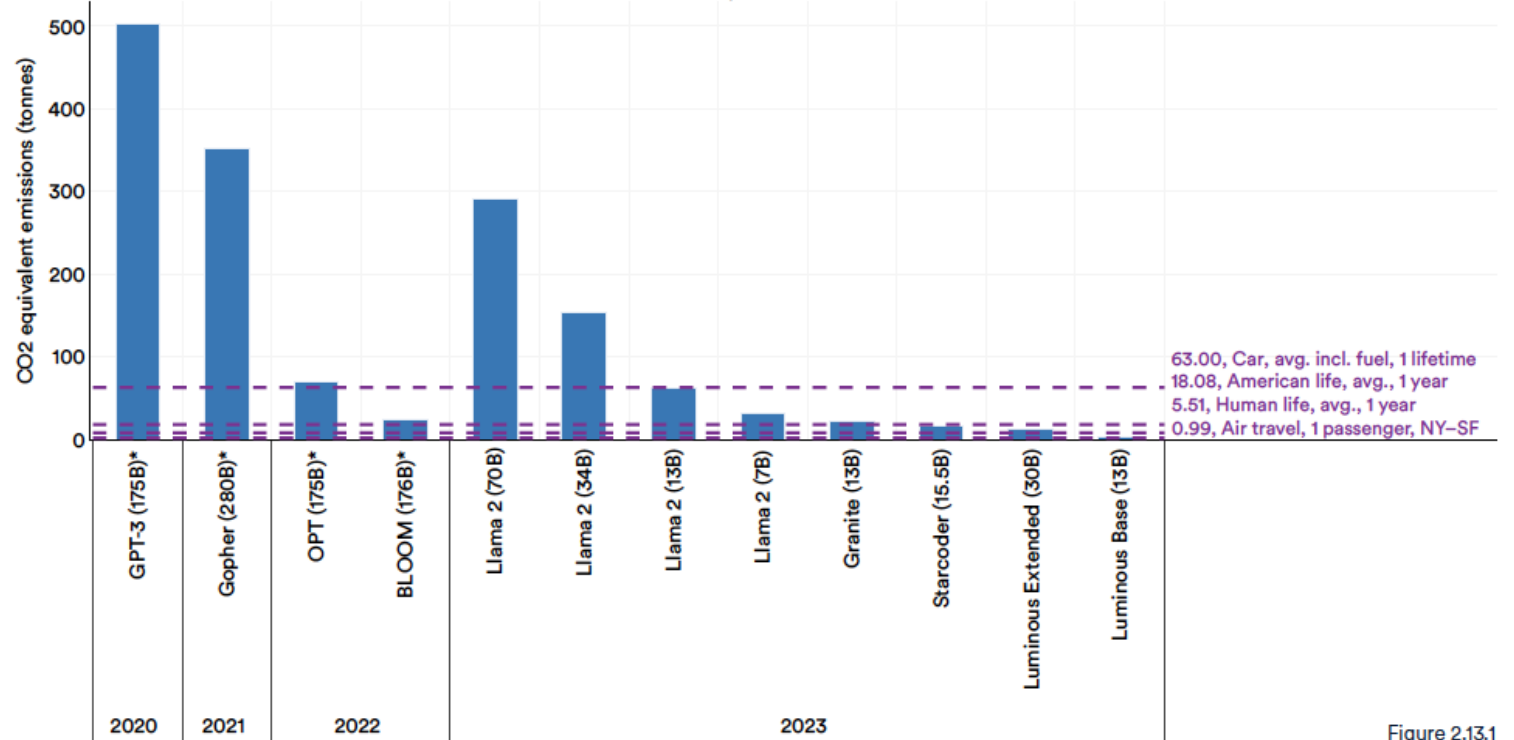


Figure 2.13.1

AI models are devouring energy. Tools to reduce consumption are here, if data centers will adopt.

Amid the race to make AI bigger and better, Lincoln Laboratory is developing ways to reduce power, train efficiently, and make energy use transparent.

SEPTEMBER 22, 2023 | Kylie Foy | Communications & Community Outreach Office



The GPUs to train GPT-3 consumed 1,300 megawatt-hours of electricity

- Use of 1,450 American homes per month

Carbon emissions of the computing industry, with AI at the helm, are greater than those of the aeronautical industry

Google says AI currently accounts for between 10 and 15% of its energy use, 2.3 TWh per year.

Google has used 15.8 billion liters of water to maintain the correct temperature in its data centers.

The world's data centers consume 416TWh

- 3% of global electricity consumption
- 40% more than the UK's annual consumption
- The electricity demand they require will multiply by 15 between now and 2030, reaching 30% of global energy consumption.

If the trend continues, in 2026 AI consumption will equal that of Japan

ENERGY TRANSITION — 16 Oct 2023 | 21:15 UTC

POWER OF AI: Wild predictions of power demand from AI put industry on edge

HIGHLIGHTS

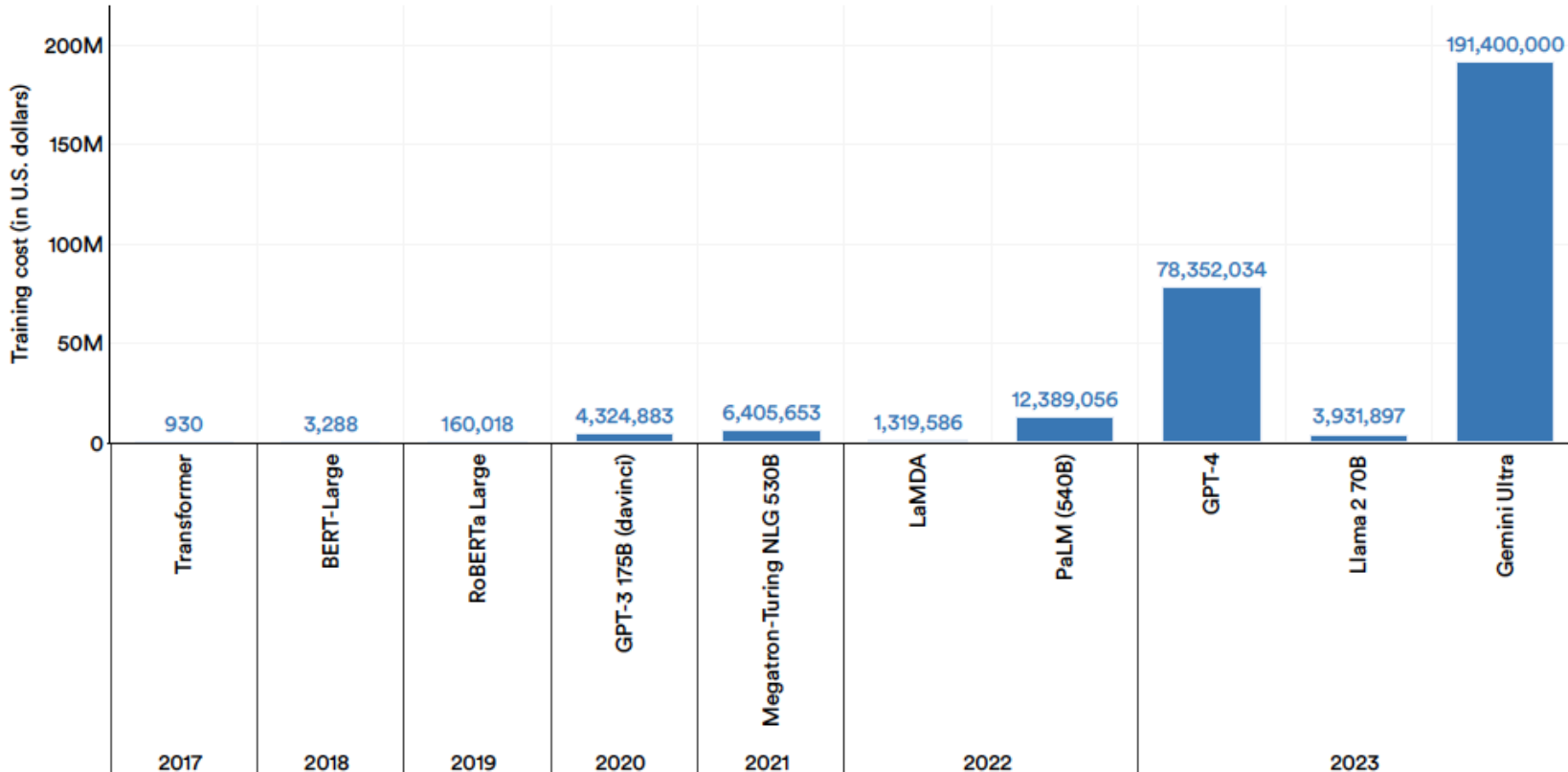
Significant net increase in demand expected

Some applications could reduce demand

MODEL	YEAR	N. PARAMETERS (Billions)	TRAINING COST (Dollars)
GPT-2	2019	1,5	50,000
PaLM	2022	540 (360 times bigger)	8 Milions (160 times more expensive)

Estimated training cost of select AI models, 2017–23

Source: Epoch, 2023 | Chart: 2024 AI Index report



2025



5,6 M dollars

Costs per million output token

Open AI R1- 60\$

DeepSeekR1-2,19\$

Number of notable machine learning models by sector, 2003–23

Source: Epoch, 2023 | Chart: 2024 AI Index report

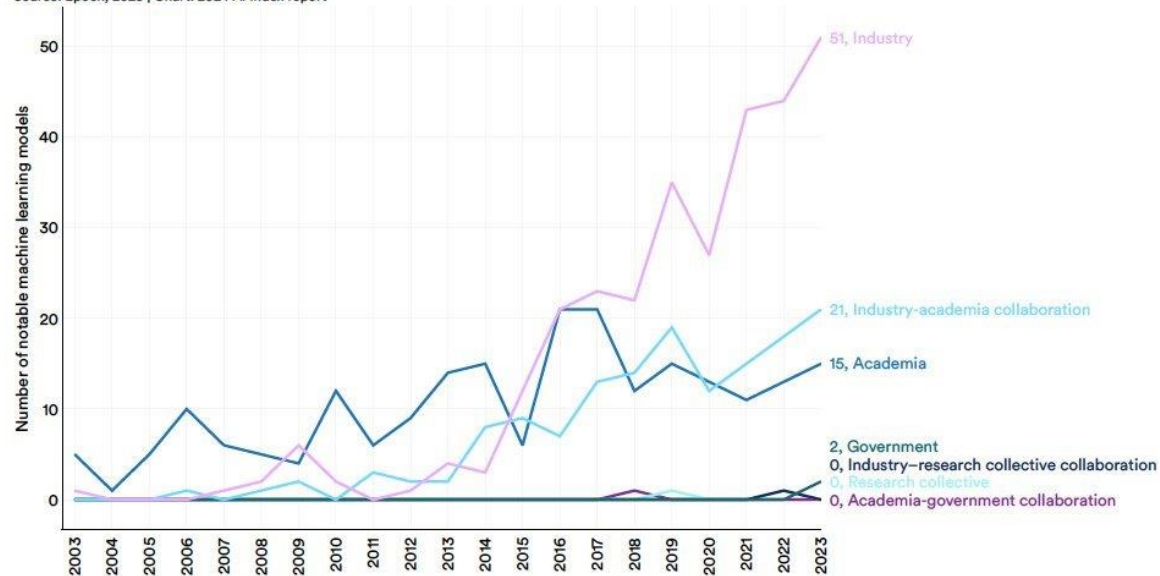
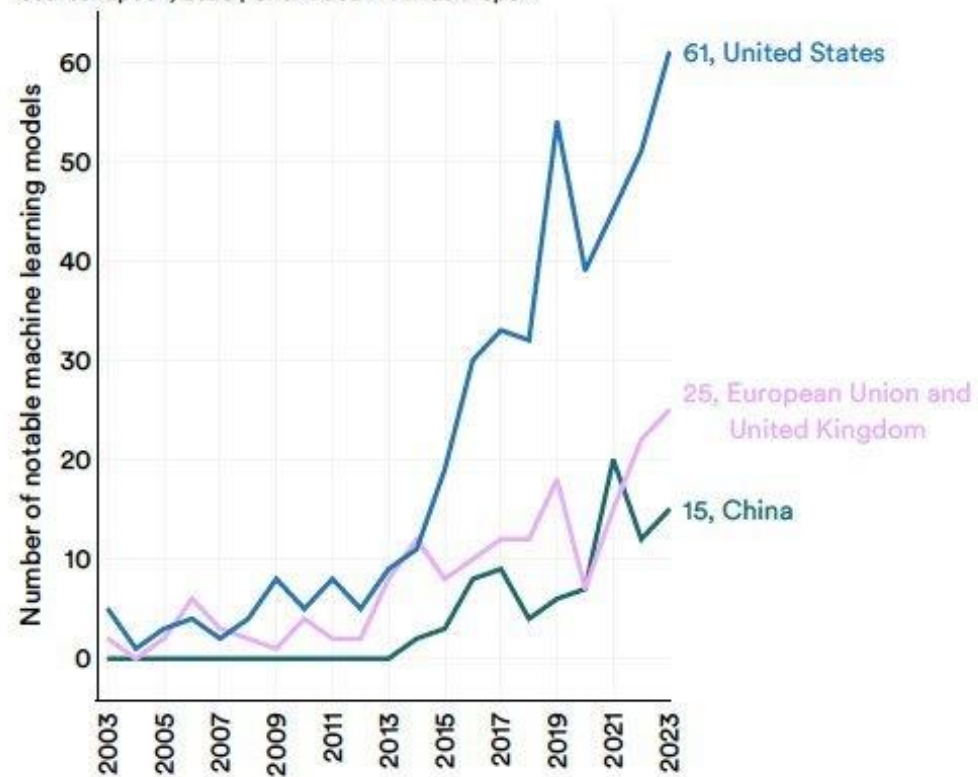
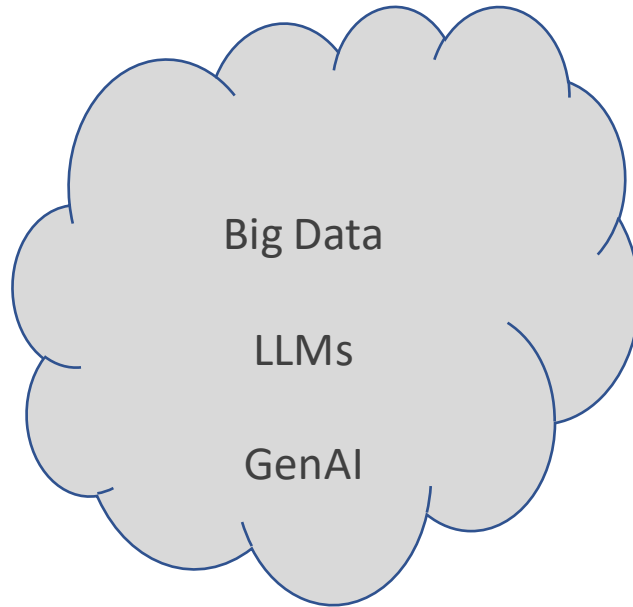
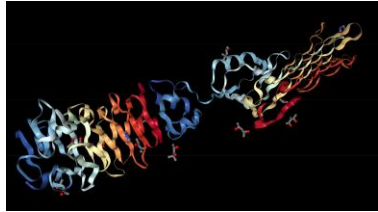


Figure 1.3.1

Number of notable machine learning models by select geographic area, 2003–23

Source: Epoch, 2023 | Chart: 2024 AI Index report





Huge computing overhead

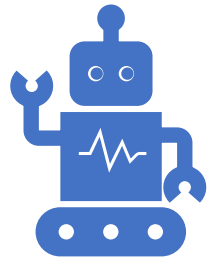
Accuracy as the only objective

Take into account the environmental, social and economic repercussions of AI

Balanced multidimensional assessment

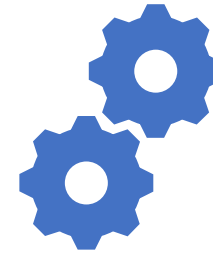
RED AI

GREEN AI



Frugal AI

AI designed to operate efficiently with limited resources



PRINCIPLES

Resource efficiency

Low cost

Simplicity

Sustainability

Democratizes AI access for low-resource environments

Location:	Pennsylvania		

Final Readings			

Average baseline wattage:	1.86 watts		
Average total wattage:	19.42 watts		
Average process wattage:	17.56 watts		
Process duration:	0:00:01		

Energy Data			

Energy mix in Pennsylvania			
Coal:	25.42%		
Oil:	0.17%		
Natural Gas:	31.64%		
Low Carbon:	42.52%		

Emissions			

Effective emission:	4.05e-06 kg CO2		
Equivalent miles driven:	1.66e-12 miles		
Equivalent minutes of 32-inch LCD TV watched:	2.51e-03 minutes		
Percentage of CO2 used in a US household/day:	1.33e-12%		

Assumed Carbon Equivalencies			

Coal:	995.725971 kg CO2/MWh		
Petroleum:	816.6885263 kg CO2/MWh		
Natural gas:	743.8415916 kg CO2/MWh		
Low carbon:	0 kg CO2/MWh		

Emissions Comparison			

Quantities below expressed in kg CO2			
US	Europe	Global minus US/Europe	
Max: Wyoming	9.59e-06 Kosovo	9.85e-06 Mongolia	9.64e-06
Median: Tennessee	4.78e-06 Ukraine	6.88e-06 Korea, South	7.87e-06
Min: Vermont	2.69e-07 Iceland	1.77e-06 Bhutan	1.10e-06

Process used:	1.04e-05 kWh		

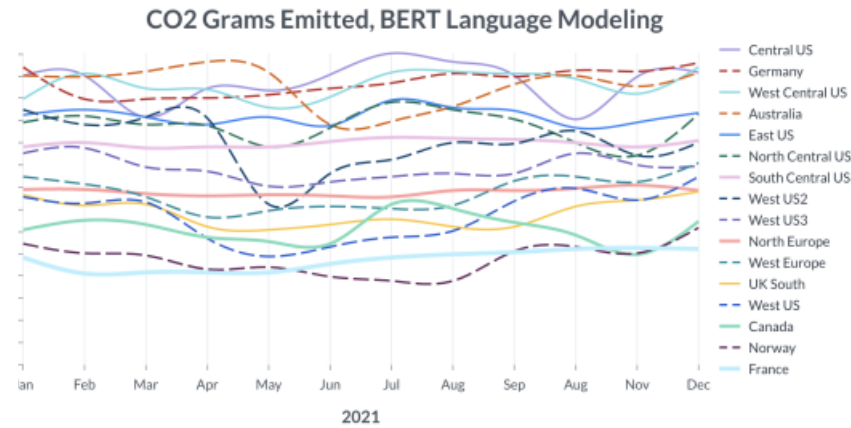
Buscar proyectos Ayuda Patrocinadores Acceder Registrarse

energyusage 0.0.14

pip install energyusage

✓ Versión más reciente
Publicación: 13 dic 2019

Measuring the environmental impact of computation



<https://hiili.org>



ML CO2 Impact Compute Publish Learn Act About

Machine Learning Emissions Calculator

Choose your hardware, runtime and cloud provider to estimate the carbon impact of your research.

This calculator will give you 2 numbers: the **raw** carbon emissions produced and the approximate **offset** carbon emissions. The latter number depends on the grid used by the cloud provider and we are open to update our estimates if anything looks inaccurate or outdated.

Also, keep in mind that the estimate provided below **does not** take datacenter PUE (Power Usage Effectiveness) into account. To do so, you need to find your datacenter's PUE (by asking your computer provider or consulting their documentation) and multiply the quantity of carbon emitted provided below by that number.

Missing a hardware or a region? Open an issue or a PR on GitHub

Hardware type: A100 PCIe 40/80G
Hours Used: 100
Provider: Google Cloud Plat
Region of Compute: asia-east1

COMPUTE

CALCULATING ENERGY AND REPORTING

Buscar proyectos Ayuda Patrocinadores

codecarbon 2.3.4

pip install codecarbon

STRATEGIES



DATA



LEARNING



INFERENCE



DATA

OBJECTIVES:

Reducing acquisition, storage and exploitation costs
Improving data quality

OTHER NEEDS ADDRESSED:

Resource restrictions, privacy,..

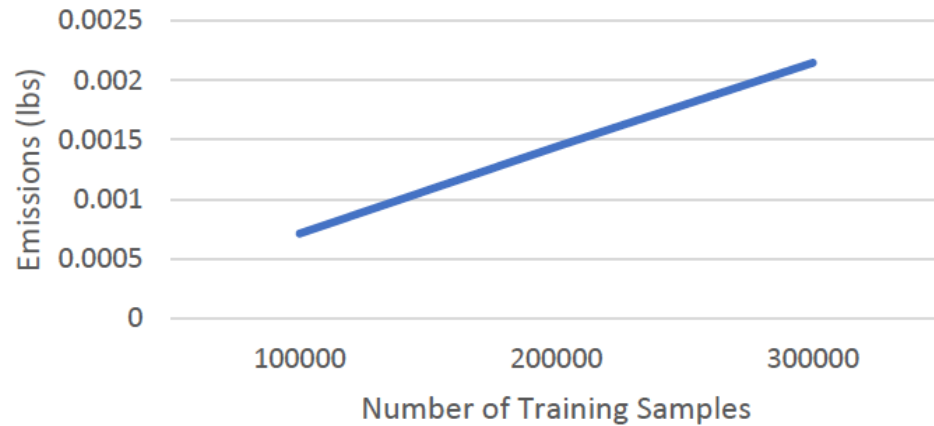
ALTERNATIVES:

Using less data or features
Smaller and well curated datasets
Improve data quality

HOW?

- Feature selection
- Positive Unlabelled Learning (PU Learning)
- Active / Few-shot learning

Influence of Sample Size



FEATURE SELECTION. New measures

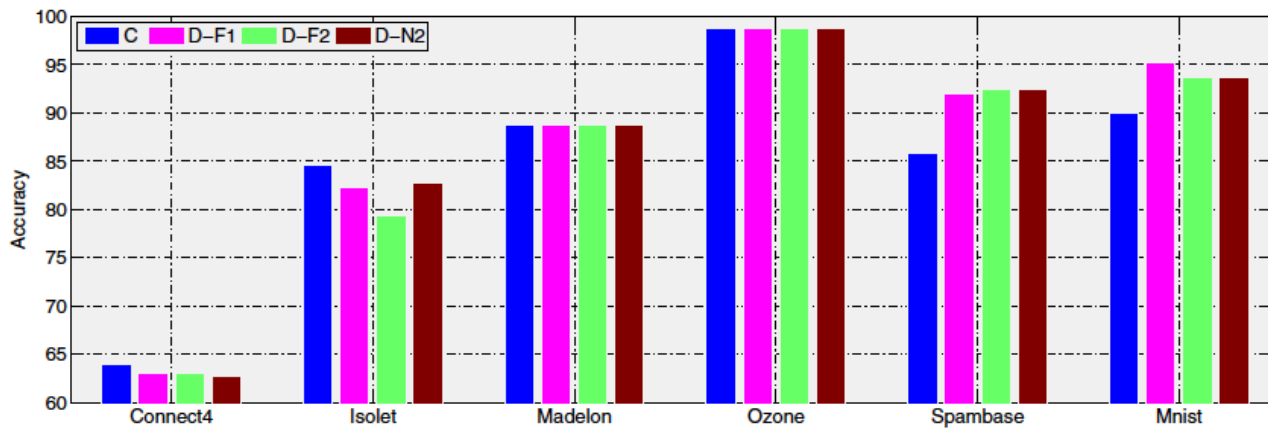
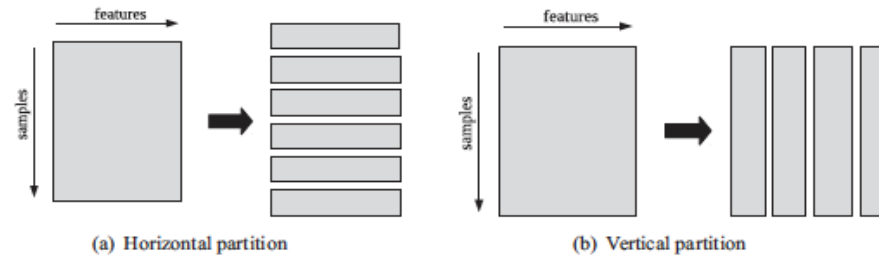


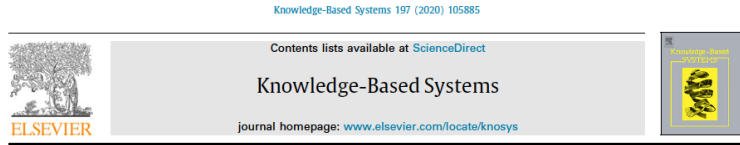
Fig. 4. Comparing the centralized and distributed approaches using horizontal partition in terms of classification accuracy.

Table 6

Maximum runtime (s) for the feature selection methods tested. C stands for centralized approaches, while D refers to the distributed approaches.

		Connect4	Isolet	Madelon	Ozone	Spambase	Mnist	SpeedUp
CFS	C	100	250	36	10	12	1787	5.73
	D	10	77	25	8	6	257	
INT	C	112	196	40	9	13	3145	10.56
	D	11	70	31	8	14	199	
Cons	C	268	245	52	11	14	6163	21.10
	D	10	80	25	6	2	197	
IG	C	97	171	41	9	11	1451	5.30
	D	4	54	29	9	5	235	
ReliefF	C	1680	553	62	14	21	30,413	21.66
	D	11	103	40	8	4	1346	

FEATURE SELECTION. Reduced precision models



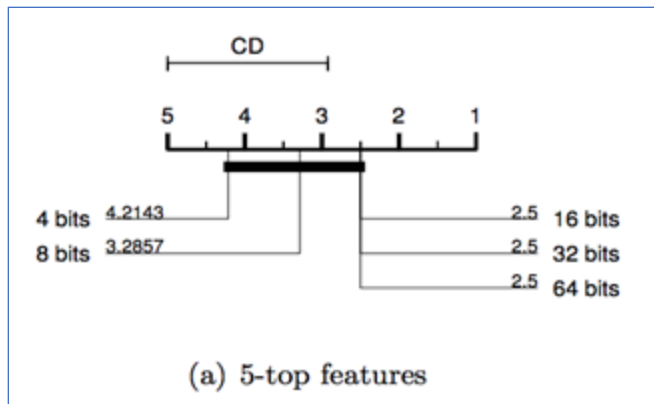
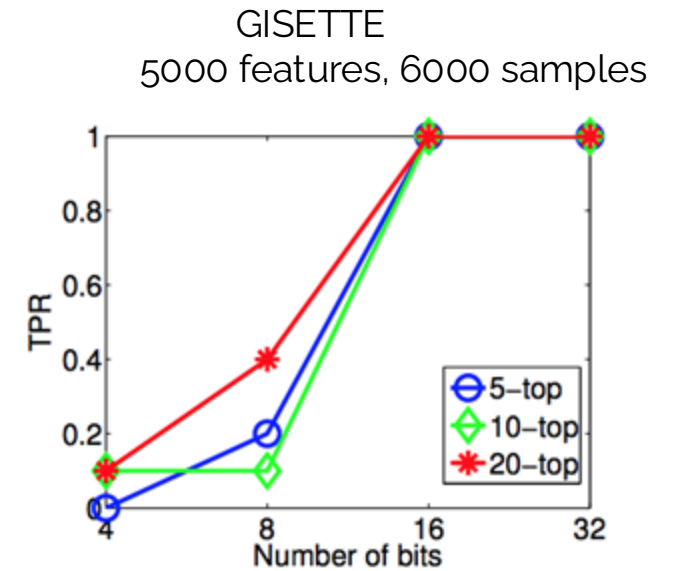
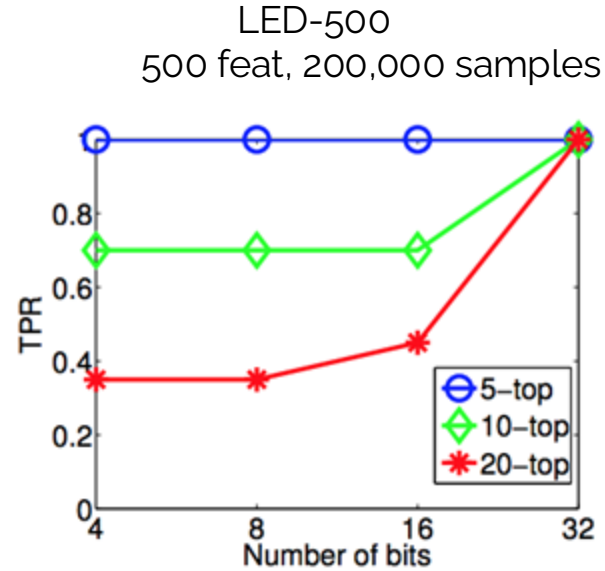
Feature selection with limited bit depth mutual information for portable embedded systems

Laura Morán-Fernández^{a,*}, Konstantinos Sechidis^b, Verónica Bolón-Canedo^a, Amparo Alonso-Betanzos^a, Gavin Brown^b

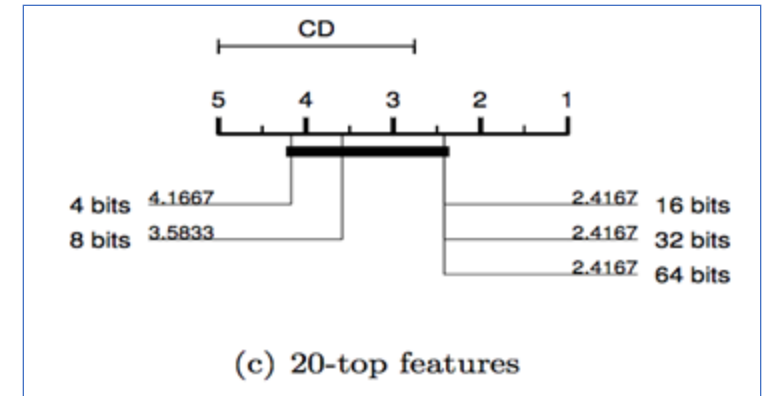
^aCIIC, Universidade da Coruña, A Coruña, Spain
^bSchool of Computer Science, University of Manchester, Manchester, UK



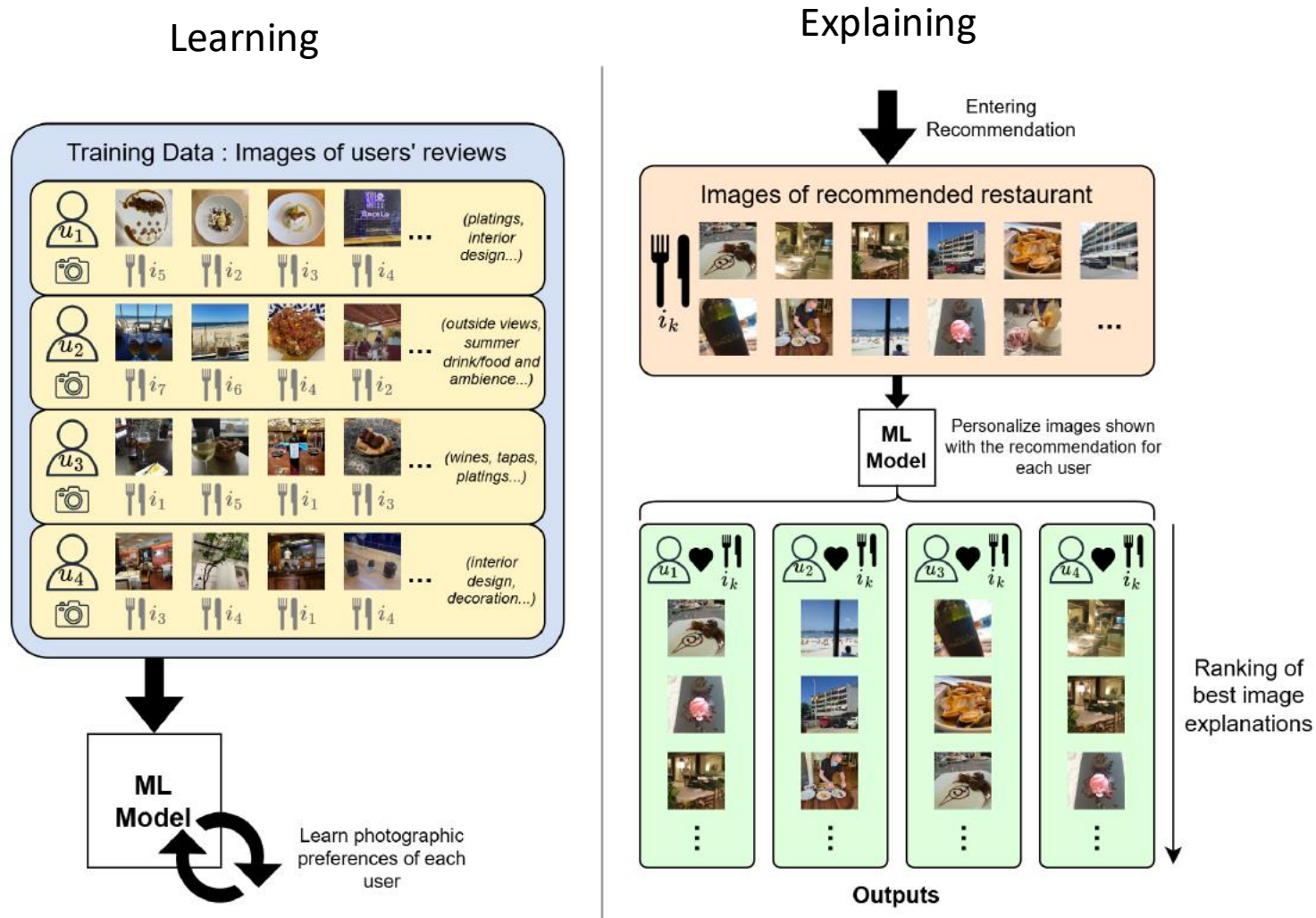
Fixed point in Mutual Information-based algorithms

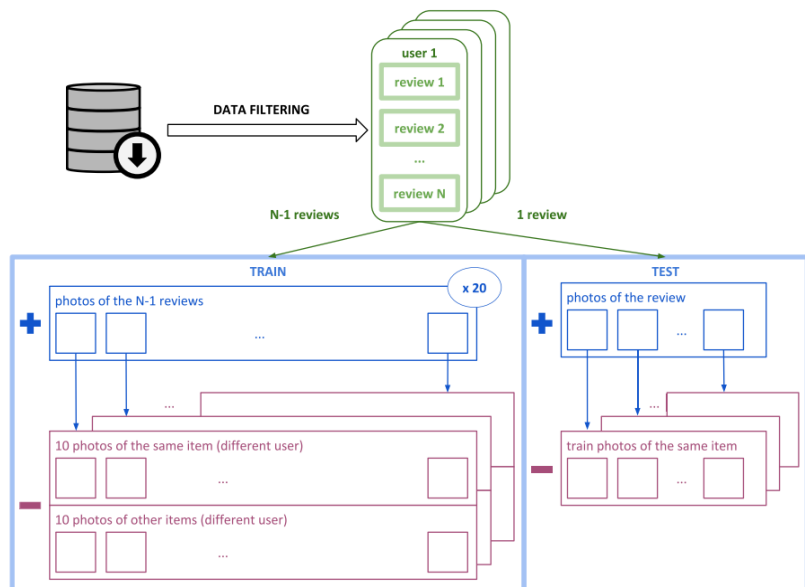


3-NN
 5-fold cross validation
 Friedman test with Nemenyi post-hoc test



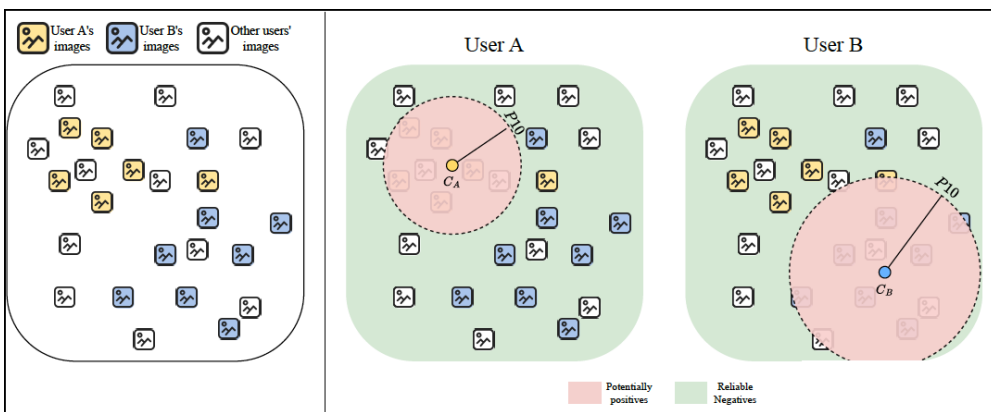
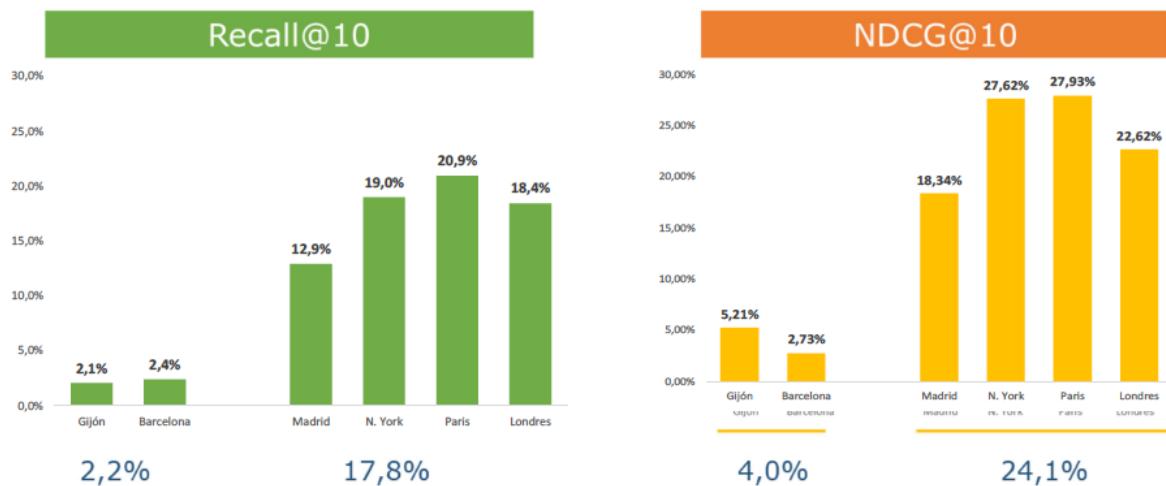
PU Learning : Visual explanation with pre-existent content





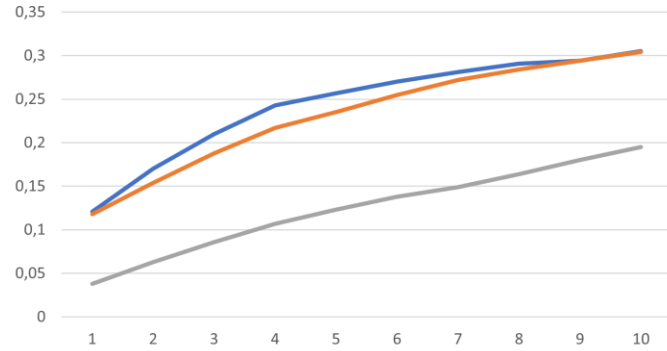
City	Train Partition			Test Partition		
	Users	Restaurants	Images	Users	Restaurants	Images
Gijon	5,139	598	16,302	1,023	346	2,377
Barcelona	33,537	5,881	130,674	8,697	3,211	19,742
Madrid	43,628	6,810	176,763	11,874	3,643	27,142
New York	61,019	7,588	196,315	16,842	4,135	34,826
Paris	61,391	11,982	219,588	15,242	6,345	32,048
London	134,816	13,888	416,356	30,393	8,097	63,442

average percentages of improvement with respect to SOA (ELVis)

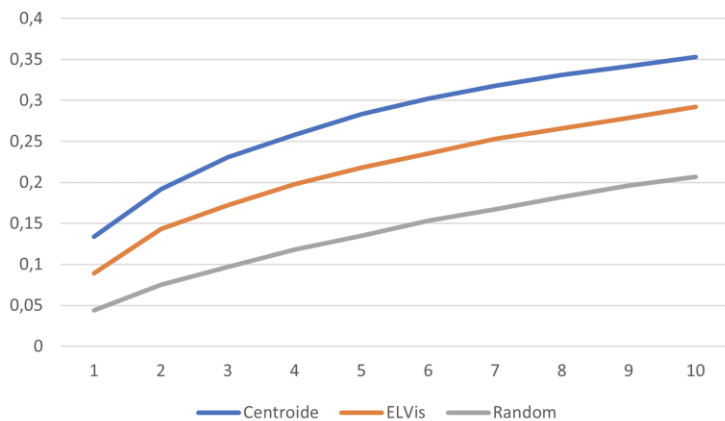


NDCG@10

Gijón



París



Method	Features	Classifier	Performance (avg of 10 runs)		
			AUC-ROC	G. Mean	F1 Score
Original	PathDIP	CAT	0.829	0.717	0.522
		BRF	0.825	0.752	0.450
	GO	CAT	0.832	0.654	0.463
		BRF	0.827	0.755	0.377
PU Learning	PathDIP	CAT	0.829	0.750	0.537†
		BRF	0.815	0.728	0.381
	GO	CAT	0.838†	0.726	0.491
		BRF	0.829	0.763†	0.380

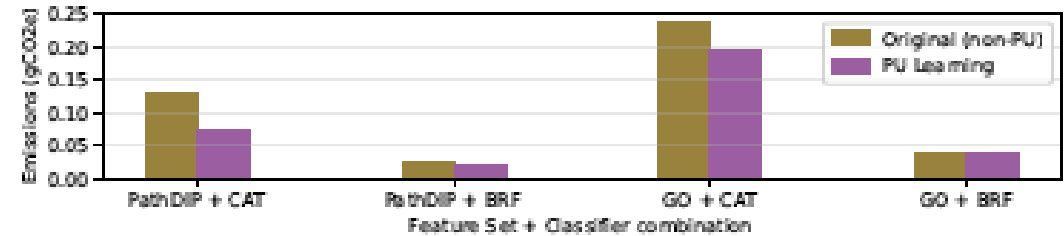


Positive-Unlabelled learning for identifying new candidate Dietary Restriction-related genes among ageing-related genes

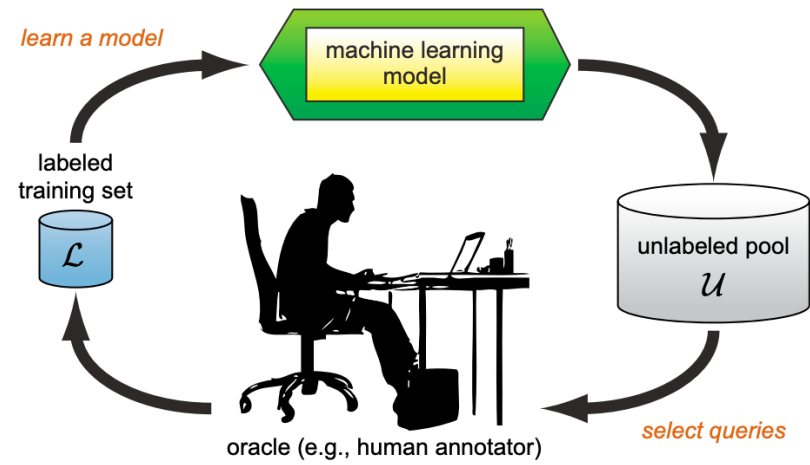
Jorge Paz-Ruza^{a,*}, Alex A. Freitas^b, Amparo Alonso-Betanzos^a, Bertha Guijarro-Berdiñas^a

^a LIDA Group, CITIC, Universidade da Coruña, Campus de Elvella s/n, A Coruña 15071, Spain
^b School of Computing, University of Kent, Canterbury CT2 7FS, United Kingdom

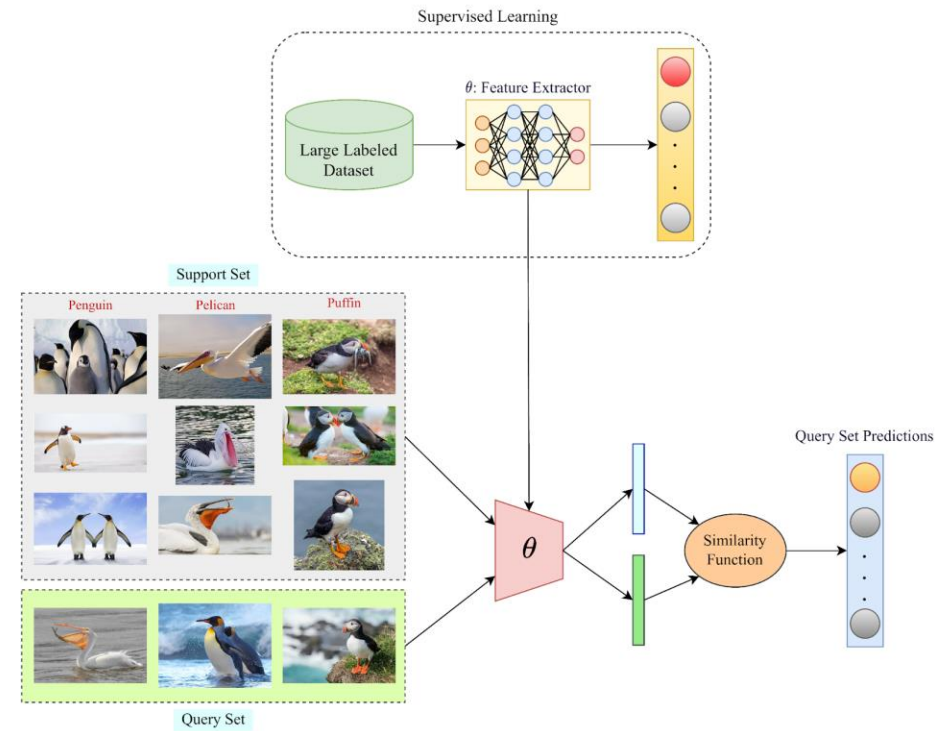
Original non-PU-Learning method		KNN-based PU-Learning method	
Gene	DR-Probability	Gene	DR-Probability
GOT2	0.86	TSC1	0.97
GOT1	0.85	GCLM	0.94
TSC1	0.85	IRS1	0.93
CTH	0.85	PRKAB1	0.92
GCLM	0.82	PRKAB2	0.90
IRS2	0.80	PRKAG1	0.90
SENS2	0.80	IRS2	0.90



ACTIVE LEARNING



FEW-SHOT LEARNING





LEARNING

Network
Architecture
Search (NAS)

Model training

Inference

OBJECTIVES:

To develop more sustainable models

OTHER NEEDS ADDRESSED:

Privacy, Bias, Opacity, Restrictions in resources

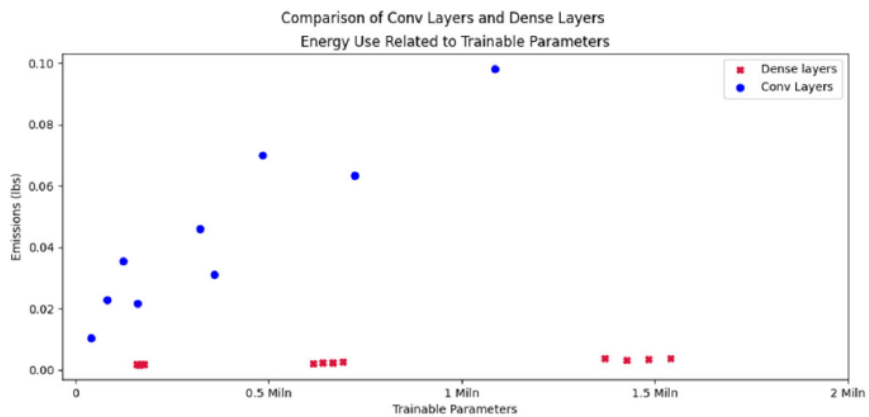
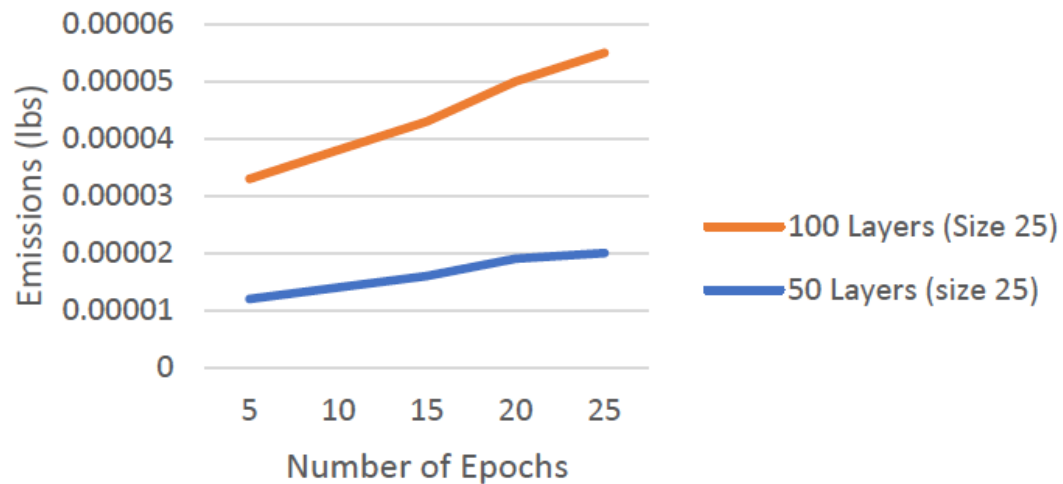
WHAT CAN BE DONE?

Using simpler models, with less parameters and lower hardware needs

HOW?

- Green algorithms and low precision models
- Modular Learning and Long-life learning

Impact of Epochs



Estimating the Sustainability of AI Models Based on Theoretical Models and Experimental Data

Ralf Gitzel, Marie Platenius-Mohr, Andreas Burger

Posted Date: 23 January 2023

Emissions and Network Depth

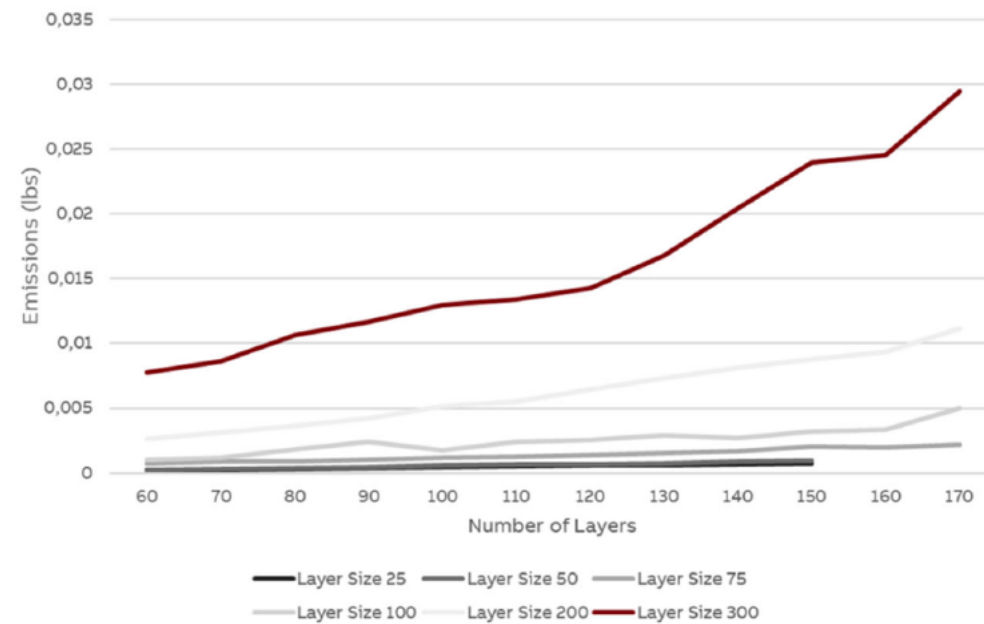


Table 3. Impact of transfer learning on energy consumption.

	RCvD [39]	TLCvD [39]	TLRep	TLMNIST	RMNIST	DMNIST
Accuracy	95%	95%	98%	94.3%	11.3%	98.9%
Energy (kWh)	4.76	0.32	0.337	0.136	0.451	0.005

Xception model trained from scratch

Transfer models

Transfer 6 epochs

Xception from scratch, 20 epochs

Dedicated model from scratch

Classification cats/dogs

Estimating the Sustainability of AI Models Based on Theoretical Models and Experimental Data

Baif Güzal^{*}, Marie Platenius-Mohr, Andreas Burger

Posted Date: 23 January 2023

doi: 10.20944/preprints202301.0406.v1

Keywords: AI, Sustainability, Energy efficiency, Deep learning, Neural networks

Marginal performance gains, exponential increase in CO₂ emissions in training



Model	RMSE (Performance)	gCO2 (Emissions)
Random	1,707	N/A
Naive (Dyad Average)	0,958	N/A
MF	0,840	56,449
GC-MC	0,846	<u>243,349</u>
Bayesian SVD++	<u>0,830</u>	358,900
Glocal-K	0,826	2038,492

+1% perf

x10 CO2

- Personalization is not for free
 - Models should understand, learn and process data from many different individual users.

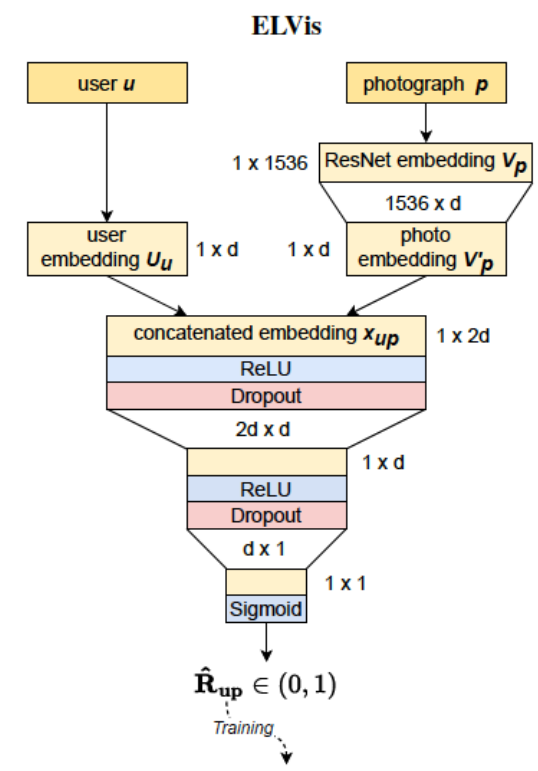
Non-personalised recommendation



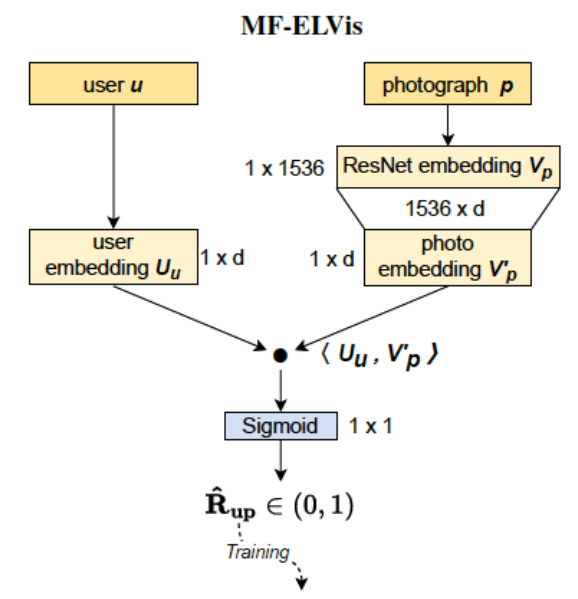
Recommendation personalised for User A



Recommendation personalised for User B



$$\mathcal{L}_{BCE} = -(R_{up} \cdot \log(\hat{R}_{up}) + (1 - R_{up}) \cdot \log(1 - \hat{R}_{up}))$$



$$\mathcal{L}_{BCE} = -(R_{up} \cdot \log(\hat{R}_{up}) + (1 - R_{up}) \cdot \log(1 - \hat{R}_{up}))$$



More sustainable models



Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 207 (2022) 1017–1026

Procedia

Computer Science

www.elsevier.com/locate/procedia

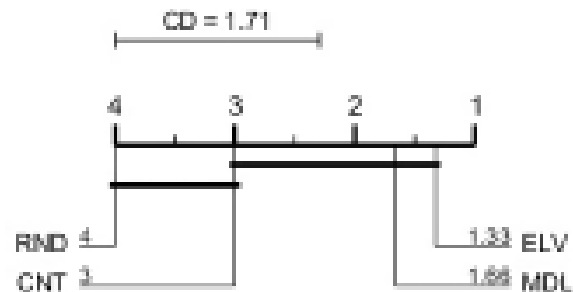
26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Sustainable Personalisation and Explainability in Dyadic Data Systems

Jorge Paz-Ruza^a, Carlos Eiras-Franco^a, Bertha Guijarro-Berdiñas^a, Amparo Alonso-Betanzos^a

^aCIIC, Universidad de Coruña, 15071 A Coruña, Spain

Dataset	Total CO2 emissions (g)		Training Times (s)	
	Deep Learning	Matrix Factorization	ELVis	Proposed Model
Gijón	1.5	0.4	240 ± 4.35	53.90 ± 1.80
Barcelona	9.2	1.8	2465 ± 16.1	696 ± 5.02
Madrid	16.0	3.5	1530 ± 10.4	436 ± 5.49
New York	42.8	7.9	2865 ± 19.5	746 ± 7.31
Paris	57.8	10.0	2940 ± 14.5	786 ± 6.48
London	138.8	19.1	5197 ± 48.9	1578 ± 17.4



	Gijón	
	MRecall@10	MNDCG@10
RND	0.373	0.185
CNT	0.464	0.218
ELVis	<u>0.521</u>	<u>0.262</u>
MF-ELVis	0.538	0.285
	Nueva York	
	MRecall@10	MNDCG@10
RND	0.374	0.168
CNT	0.431	0.217
ELVis	0.553	0.304
MF-ELVis	<u>0.516</u>	<u>0.276</u>
	Londres	
	MRecall@10	MNDCG@10
RND	0.342	0.155
CNT	0.400	0.200
ELVis	<u>0.530</u>	0.293
MF-ELVis	0.531	<u>0.267</u>



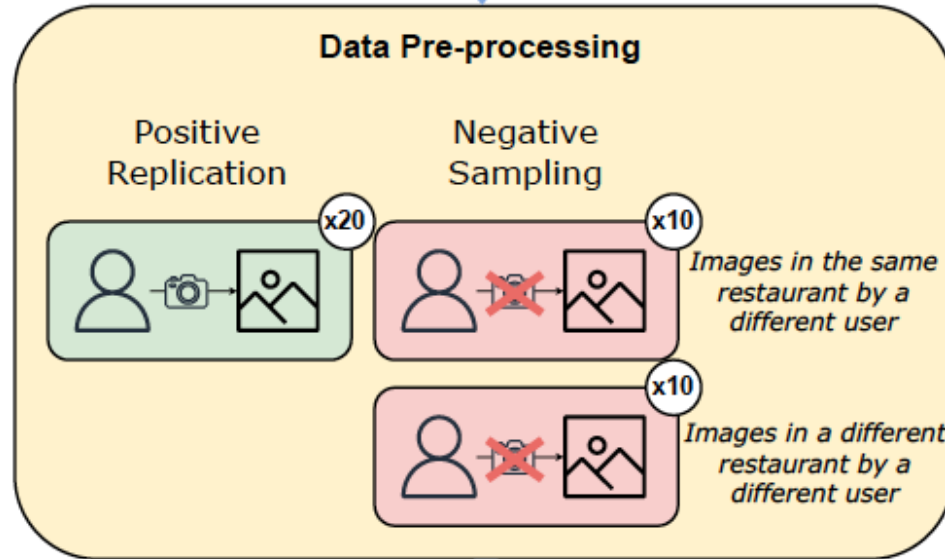
Full Length Article

Sustainable transparency on recommender systems: Bayesian ranking of images for explainability

Jorge Paz-Ruza^a, Amparo Alonso-Betanzos, Bertha Guijarro-Berdiñas, Brais Cancela, Carlos Eiras-Franco

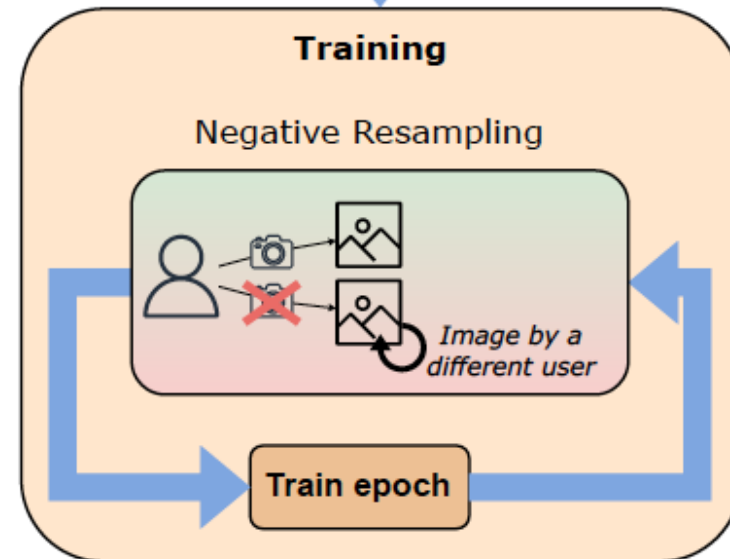
^aUniversidad de Coruña, CITIC, Campus de Bertha s/n, 15006, A Coruña, Spain

ELVIs and MF-ELVIs



Train all epochs

BRIE

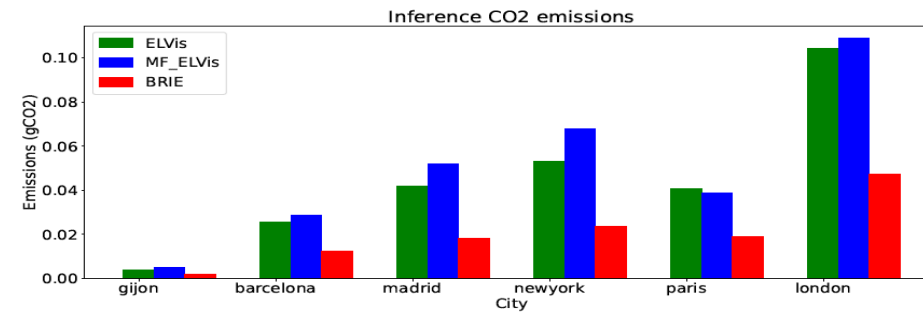
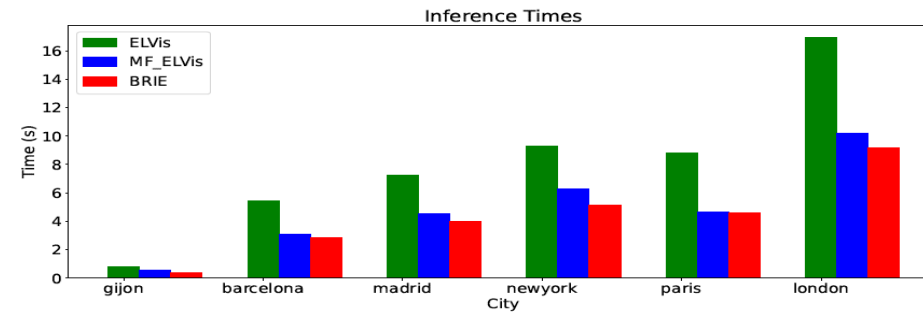


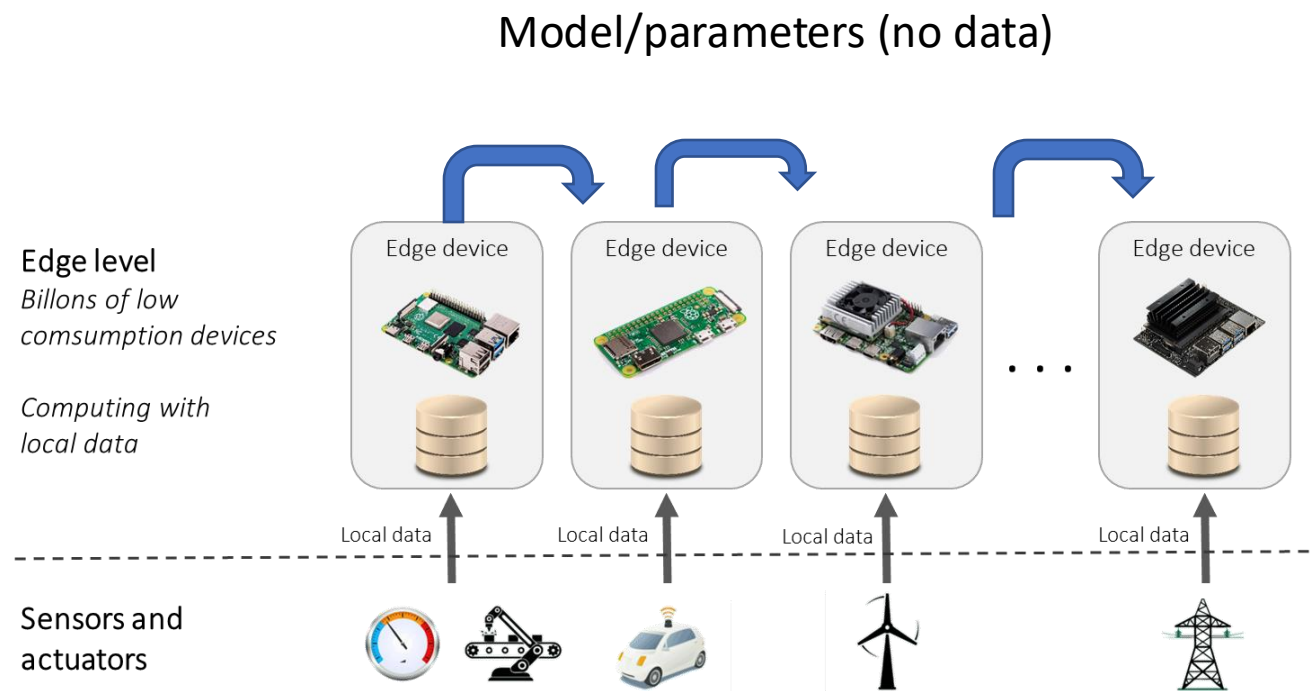
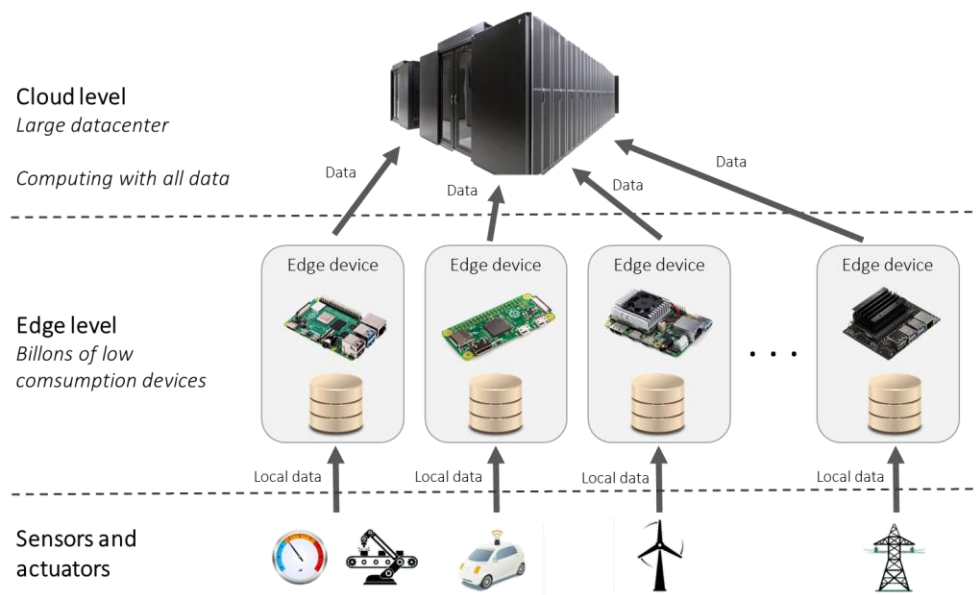
	Gijón			Barcelona			Madrid		
	MRecall@10	MNDCG@10	MAUC	MRecall@10	MNDCG@10	MAUC	MRecall@10	MNDCG@10	MAUC
RND	0.373	0.185	0.487	0.409	0.186	0.502	0.374	0.171	0.499
CNT	0.464	0.218	0.546	0.443	0.219	0.554	0.420	0.203	0.557
ELVis	0.521	0.262	<u>0.596</u>	<u>0.597</u>	<u>0.327</u>	<u>0.631</u>	<u>0.572</u>	<u>0.314</u>	<u>0.638</u>
MF-ELVis	<u>0.538</u>	<u>0.285</u>	0.592	0.557	0.293	0.596	0.528	0.279	0.601
BRIE	0.607	0.333	0.643	0.630	0.368	0.663	0.612	0.348	0.673

	Newyork			Paris			London		
	MRecall@10	MNDCG@10	MAUC	MRecall@10	MNDCG@10	MAUC	MRecall@10	MNDCG@10	MAUC
RND	0.374	0.168	0.502	0.459	0.209	0.502	0.342	0.155	0.500
CNT	0.431	0.217	0.563	0.499	0.245	0.557	0.400	0.200	0.562
ELVis	<u>0.553</u>	<u>0.304</u>	<u>0.637</u>	<u>0.643</u>	<u>0.352</u>	<u>0.630</u>	0.530	<u>0.293</u>	<u>0.629</u>
MF-ELVis	0.516	0.276	0.602	0.606	0.323	0.596	<u>0.531</u>	0.267	0.597
BRIE	0.598	0.341	0.677	0.669	0.391	0.666	0.563	0.318	0.665

GIJÓN

	Performance			Sustainability	
	MRecall@10	MNDCG@10	MAUC	No. of Params	CO ₂ Emiss.
RND	0,373	0,185	0,487	N/A	N/A
CNT	0,464	0,218	0,546	N/A	N/A
ELVis	0,521	0,262	0,592	<u>1.840.128</u>	1,820
MF-ELVis	<u>0,538</u>	<u>0,285</u>	<u>0,596</u>	6.835.200	<u>0,975</u>
BRIE	0,607	0,333	0,643	53.400	0,486





Can Federated Learning Save The Planet?

Xinchi Qiu¹, Titouan Parcollet^{2,1}, Daniel J. Beutel^{1,5}, Taner Topal^{1,5},
 Akhil Mathur³, Nicholas D. Lane¹
¹ University of Cambridge, UK ² Avignon Université, France
³ University College London, UK ⁵ Adap, Germany
 xq227@cam.ac.uk

Country/CO ₂ (g)	V100	K80	V100	K80	FL (IID)		FL (non-IID)	
	<i>PUE = 1.67</i>		<i>PUE = 1.11</i>		1 ep	5 ep	1 ep	5 ep
USA	3.1	6.5	2.1	4.3	2.3	6.5	10.9	8.9
China	5.5	11.5	3.7	7.7	4.1	11.6	19.4	14.2
France	0.4	0.9	0.3	0.6	0.3	0.9	1.6	1.1

Country/CO ₂ (g)	V100	K80	V100	K80	FL	FL
	<i>PUE = 1.67</i>		<i>PUE = 1.11</i>		IID	non-IID
USA	1.6	5.2	1.1	3.5	0.5	1.0
China	2.9	9.2	1.9	6.2	0.9	1.7
France	0.2	0.8	0.2	0.5	0.1	0.1

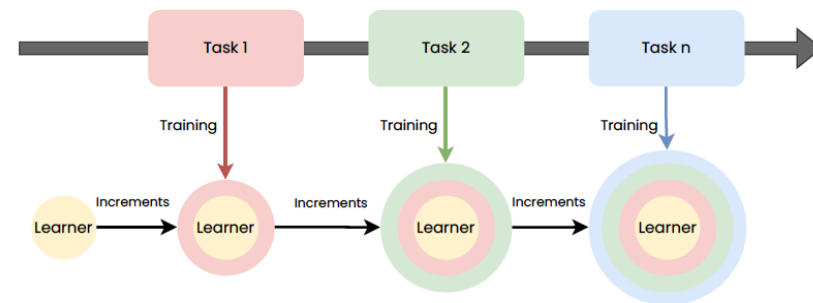
Emissions (gr) for centralized training and federated learning in CIFAR10 and Fashion-MNIST

Ep. Local epochs per clients . IID clients with an equal distribution among all cla

Class-incremental learning

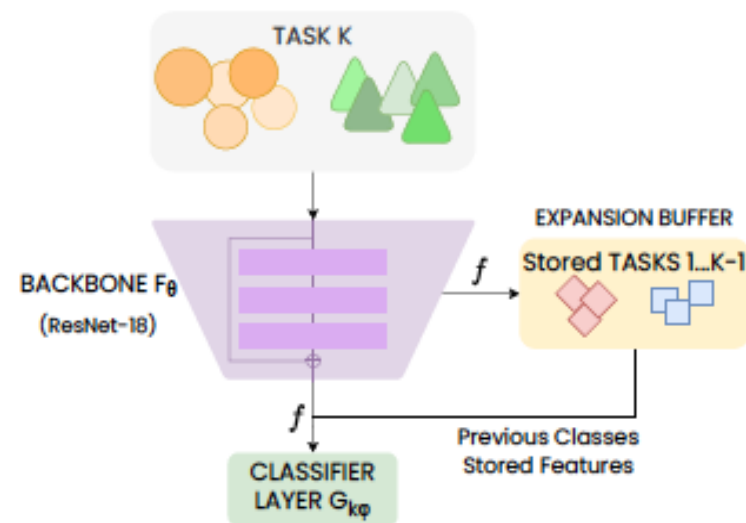
OBJECTIVES

- Competitive accuracy for resource-constrained environments
- Avoid catastrophic forgetting
- Reduce training time and energy-consumption



PROPOSAL

- Single-step optimization process
- Compressed buffer mechanism



Class-incremental learning

TABLE I
COMPARISON OF ENERGY CONSUMPTION, ENVIRONMENTAL IMPACT, AND ACCURACY OF VARIOUS METHODS ON CIFAR-100 WITH A 5-CLASS INCREMENT SETTING.

Model	#P	Duration (min.) ↓	Emissions (kg CO ₂ -eq) ↓	Energy Consumed (kWh) ↓	\bar{A} ↑	A_H ↑
Finetune	0.46	14.17	0.012	0.068	17.58	5.25
BIC	0.46	113.28	0.089	0.512	53.65	26.91
Coil	0.46	309.88	0.180	1.031	59.14	34.61
FOSTER	0.46	75.96	0.075	0.430	63.54	48.43
GEM	0.46	136.72	0.145	0.834	20.87	6.99
ICaRL	0.46	44.94	0.041	0.237	54.58	34.87
PODnet	0.46	62.98	0.056	0.322	49.22	28.41
Replay	0.46	16.33	0.015	0.094	54.49	34.05
RMM-FOSTER	0.46	76.36	0.075	0.429	67.03	51.10
WA	0.46	45.61	0.042	0.239	59.30	42.75
MEMO	7.14	66.80	0.064	0.367	68.49	54.34
DER	9.27	102.44	0.116	0.667	71.34	57.34
Ours	0.46	18.05	0.015	0.085	61.28	47.51

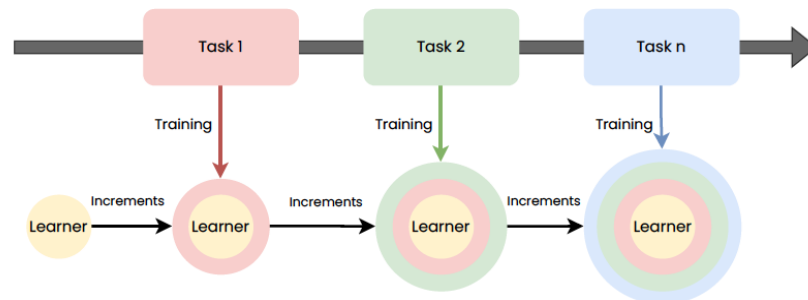
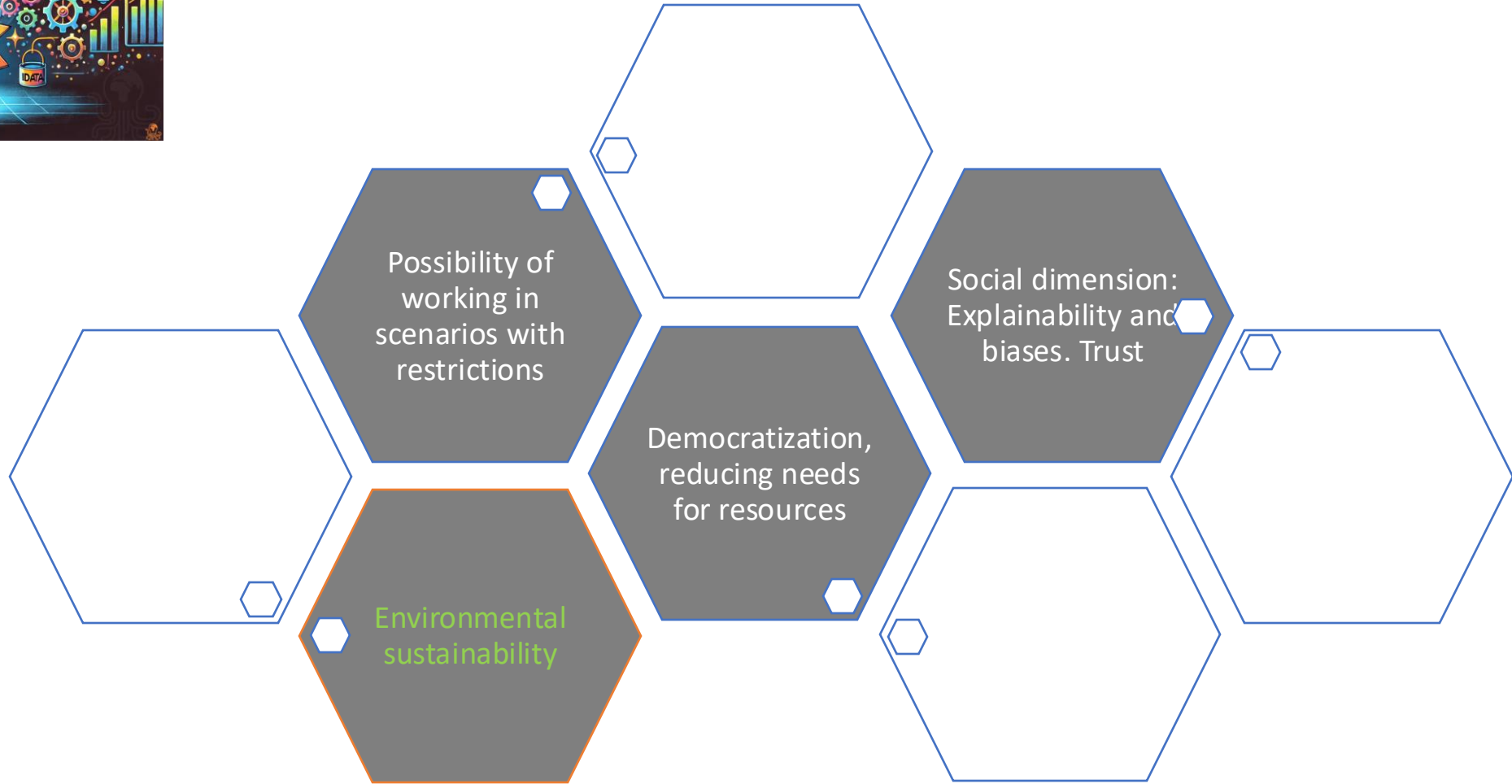
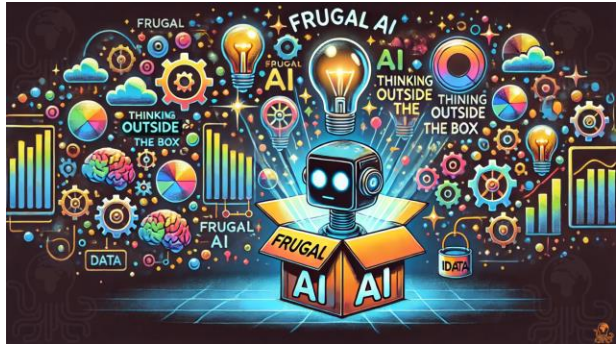


TABLE V
COMPARISON OF ENERGY CONSUMPTION, ENVIRONMENTAL IMPACT, AND ACCURACY OF VARIOUS METHODS ON IMAGENET-100 DATASET WITH A 5-CLASS INCREMENT SETTING.

Model	#P	Duration (min.) ↓	Emissions (kg CO ₂ -eq) ↓	Energy Consumed (kWh) ↓	\bar{A} ↑	A_H ↑
Finetune	11.17	188.93	0.198	1.136	17.60	4.64
Coil	11.17	1812.03	1.323	7.599	57.29	33.60
FOSTER	11.17	953.89	1.119	6.428	66.01	53.50
GEM ³	-	-	-	-	-	-
ICaRL	11.17	536.53	0.599	3.440	54.97	33.06
PODnet	11.17	721.67	0.720	4.138	55.83	37.58
Replay	11.17	226.92	0.227	1.303	56.28	35.38
RMM-FOSTER	11.17	935.67	1.089	6.255	71.73	59.46
WA	11.17	544.00	0.602	3.461	63.21	46.72
MEMO	170.60	758.81	0.847	4.864	68.42	55.52
DER	223.40	1354.40	1.736	9.972	73.88	63.66
Ours	11.17	80.52	0.053	0.304	76.66	68.70



NEXT AI GENERATION

ICPRAM 2025

14th International Conference on Pattern Recognition
Applications and Methods

Porto, Portugal

23 - 25 February



UNIVERSIDADE DA CORUÑA



citic

Amparo Alonso-Betanzos

amparo.alonso.betanzos@udc.es

CITIC-University of A Coruña.

SPAIN