



GLOBAL T



UNIÓN EUROPEA



FONDO EUROPEO DE
DESARROLLO REGIONAL

"Una manera de hacer Europa"

Desarrollo de infraestructuras IoT de altas prestaciones contra el cambio climático basadas en inteligencia artificial

Retos-colaboración (RTC2019-007159-5)

José María Cecilia
Universitat Politècnica de València
(UPV)



Agenda

El proyecto GLOBALoT [1]

El contexto científico [20]

Uso de GPUs remotas para una ejecución eficiente en el edge [10]

Uso de drones inteligentes para la ayuda a los servicios de rescate en emergencias climáticas [8]

Colaboración Universidad-Empresa

El proyecto GLOBALoT.

GLOBALoT. Consortio Universidad- Empresa

GLOBALoT call Retos
Colaboración 2019.

(911.556,05€ = Préstamo:
512.201,21€ + Subvención
292.458,35€)



nutricontrol
Automatic Fertigation & Climate Control



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UCAM
UNIVERSIDAD CATÓLICA
DE MURCIA

El contexto científico

Industria 4.0 & Sociedad 5.0

Nuevo status quo

Convergencia de negocio, procesos y estándares de gobernanza como Industria 4.0 y Sociedad 5.0



Producción en masa



Personalización masiva



Comprar



Leasing



Pay as You Go

Pago por adelantado



Pago por uso



Manual



Autónomo

X *IoT enables industrial and socioeconomic process transformation*

AI+IoT

AIoT brings sensors, machines, cloud computing, analytics and people together to improve productivity and efficiency



Manufacturing



Agriculture



Policy-makers



*Natural disaster
Management & prevention*

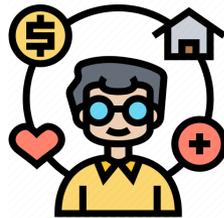
Resultados para las empresas- La principal motivación detrás de la digitalización



Nuevos servicios y modelos de negocio



Productos que mejoran con el tiempo



Mejora de relaciones con los clientes



Incremento de eficiencia



Toma de decisiones inteligentes



Disciplina basada en datos

Crecimiento de los ingresos

Los datos de IoT conducen a crecimiento de negocio

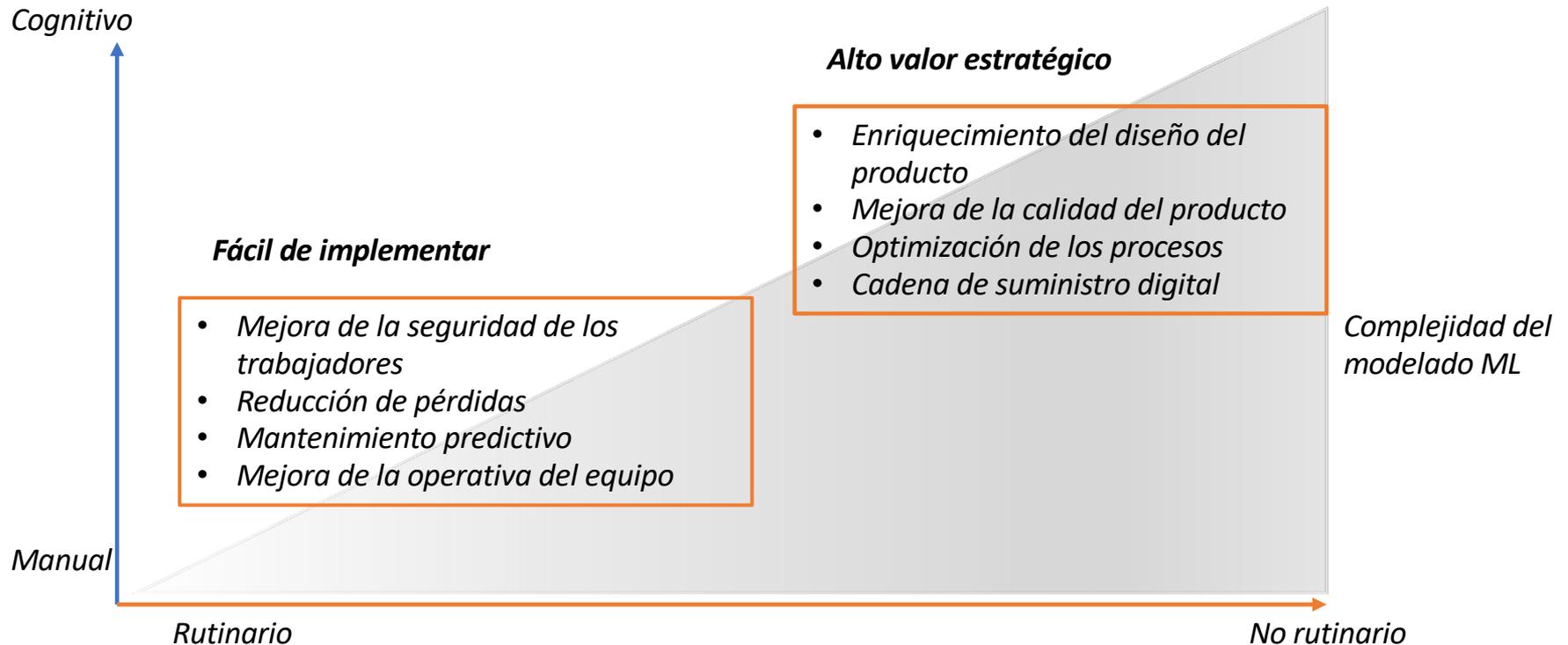
Eficiencia operacional

Los datos de IoT reducen los costes operacionales

*No body just buys “AI”... or “IoT”
.... they seek **bussiness outcomes***

Innovación gracias a la intersección AI/ML e IoT

Progresión hacia la concienciación industrial de la IA



ML en datos generados por el IoT es complicado

ML is all about data

Los datos generados por las máquinas muchas veces son liosos, no representan la realidad por no hablar de los fallos....

IoT devices deben reportar los datos en formatos simples para ser flexibles para la abstracción de funciones

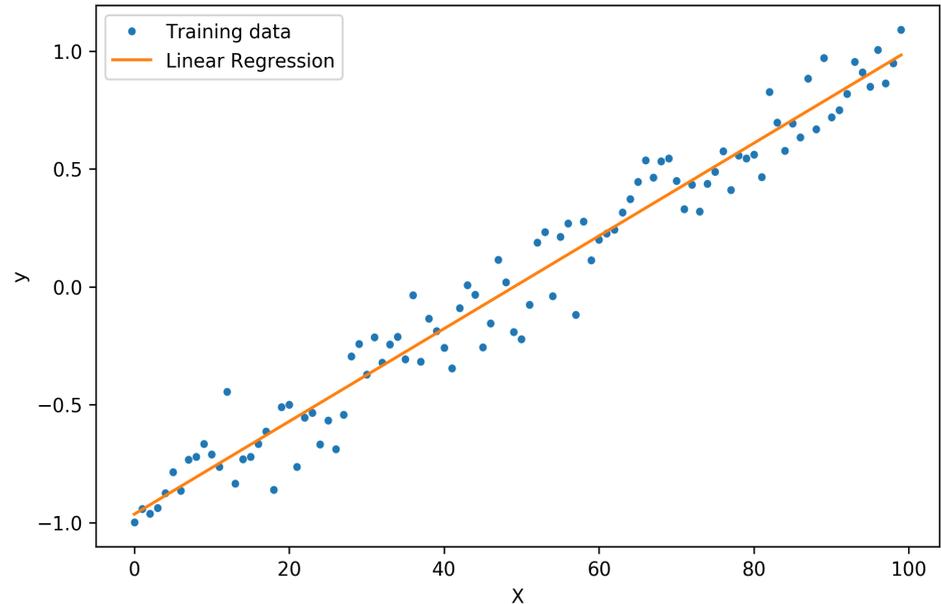
- ***Falta de etiquetas, contexto y relaciones:*** Los sensores sólo mandan datos simples (float, int). Se debe construir el contexto de esos datos para añadirlo.
- ***Datos de mala calidad e integridad:*** El IoT tiene recursos limitados y en algunos casos de mala calidad o en situaciones muy desafiantes tecnológicamente.
- ***Distinguir desviación de variación:*** Se necesita un análisis de datos avanzado para distinguir causalidad de casualidad.
- ***Datos con muchas dimensiones:*** Se refiere al número de parámetros independientes en el análisis. Cuantas más variables los métodos de ML son computacionalmente más costosos

Alinear las técnicas ML con un propósito/función

Supervised Machine learning

- Usado comúnmente para alcanzar un objetivo predefinido basado en funciones algorítmicas que conectan la entrada con la salida.
- Dadas las entradas adecuadas, se puede predecir las salidas usando nuevos datos.
- Entre los modelos destaca método de regresión lineal, auto-regresivos (AR)
- Aplicable a series temporales mantenimiento predictivo, temperatura, vibraciones ...

Regresion

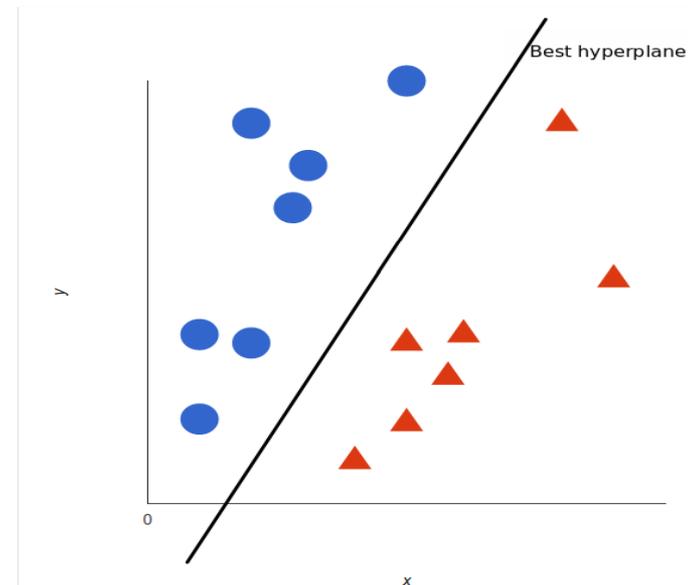


Alinear las técnicas ML con un propósito/función

Supervised Machine learning

- Asume que los datos pertenecen a unas categorías bien definidas.
- Aplicable a modelos basados en contextos donde se quiere determinar la causa (e.g. meteorología, hábitos de consumo)
- Entre los modelos destaca K-NN, SVM, RF, ANNs
- Aplicable a conjuntos de datos con múltiples rangos.

Classificación

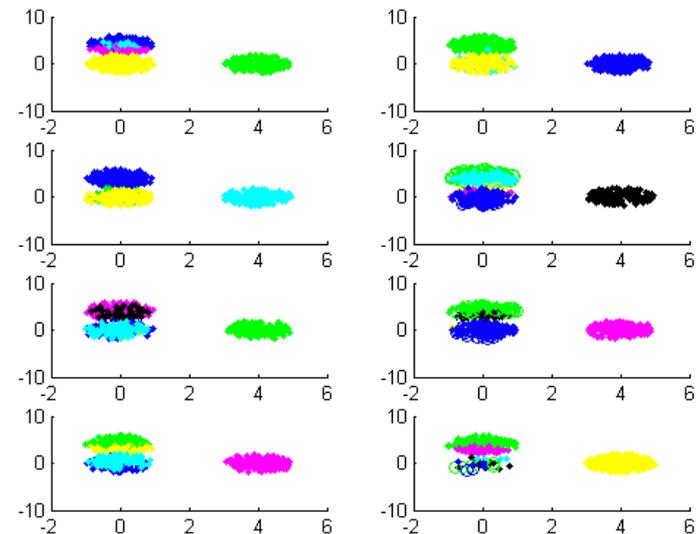


Alinear las técnicas ML con un propósito/función

Unsupervised Machine learning

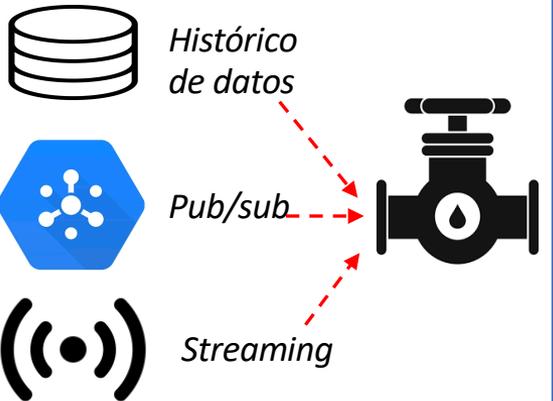
- Asume que no se conocen el resultado final.
- Entre los modelos destaca [Clustering] K-mens, FCM, FM; [Visualización y reducción de la dimensionalidad] PCA, etc.
- Muy efectivo para la mejora de procesos y calidad de producto.
- Complejo de implementar en data sets de alta dimensionalidad.

Clustering



AIoT building blocks

Recogida & Cotejo



Agregados de diferentes fuentes, estructuración y cotejo de datos según ventana temporal. *http, MQTT, API rest.*

Limpieza & Contextualización



Separar señales de ruido, limpieza, enriquecer y preparar datos de IoT. *API notification*

Optimización estructura



Almacenamiento de los datos procesados, análisis de series temporales y reutilización de datos en crudo. *Elastic Search. In-memory database*

Análisis & Visualización



ML models training & inference. *Sklearn, RAPIDS, Tensor Flow ...*

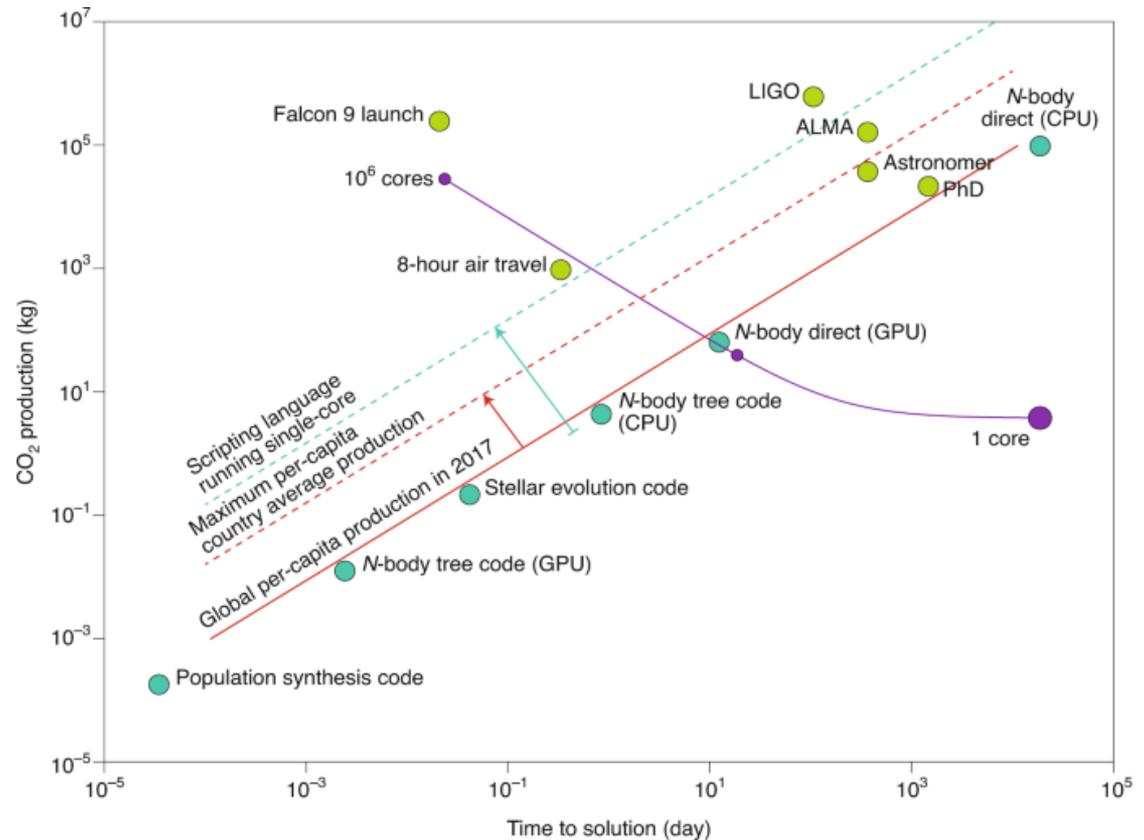
All that glitters is not gold: “*Performance and energy tradeoff*”

Data, Performance and Energy tradeoff



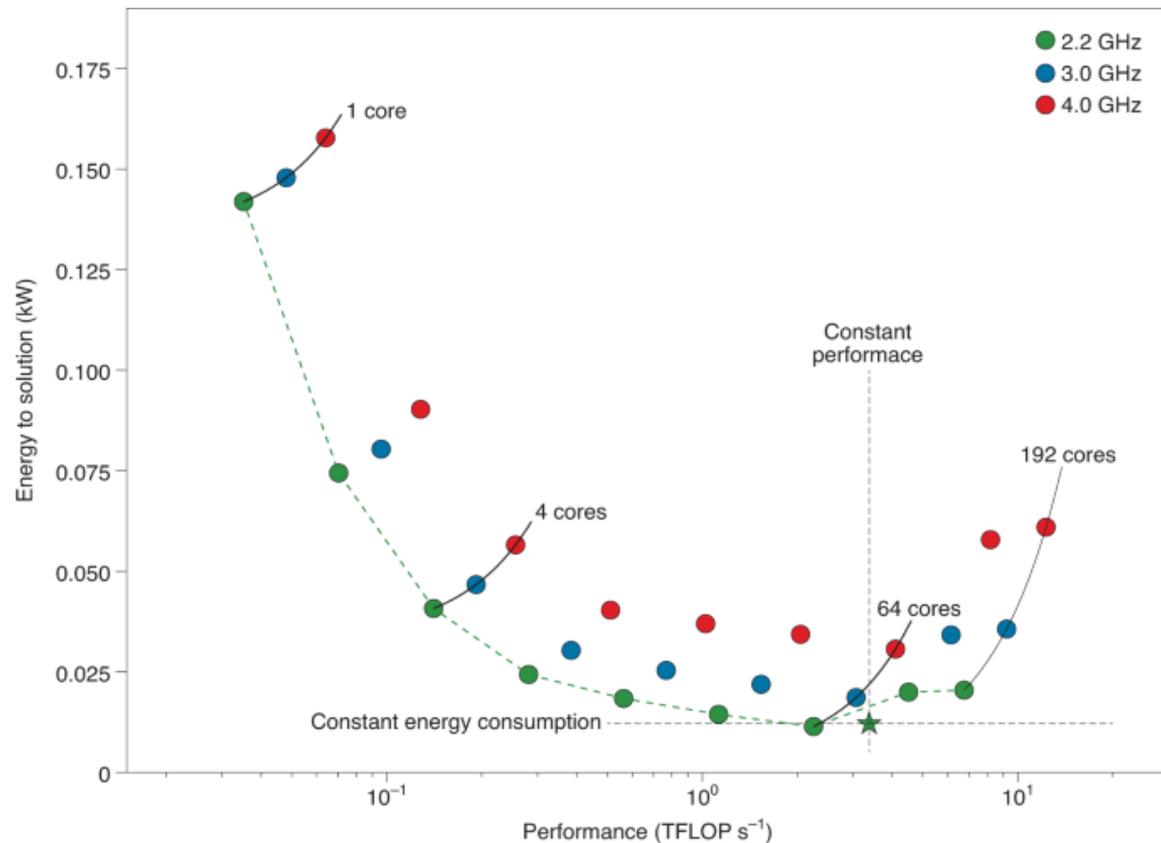
Ecological impact of HPC

Source: Zwart, S. P. (2020).
The ecological impact of high-performance computing in astrophysics. *Nature Astronomy*, 4(9), 819-822.



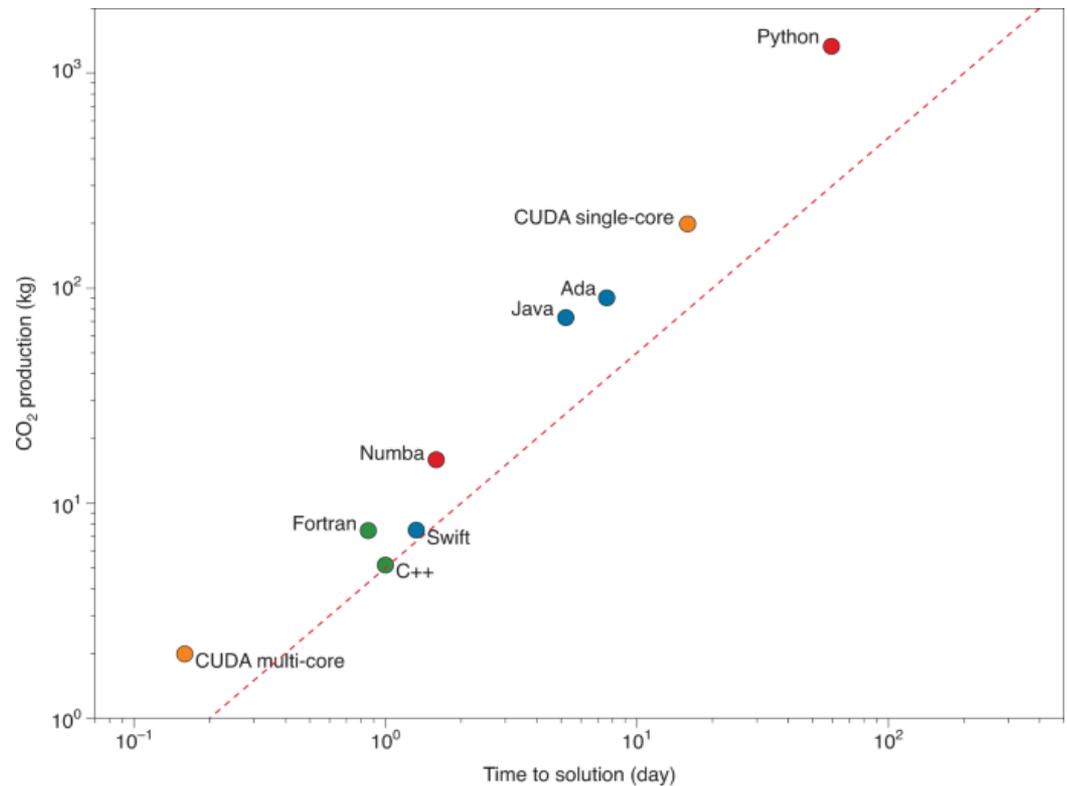
Ecological impact of HPC

Source: Zwart, S. P. (2020).
The ecological impact of high-performance computing in astrophysics. *Nature Astronomy*, 4(9), 819-822.

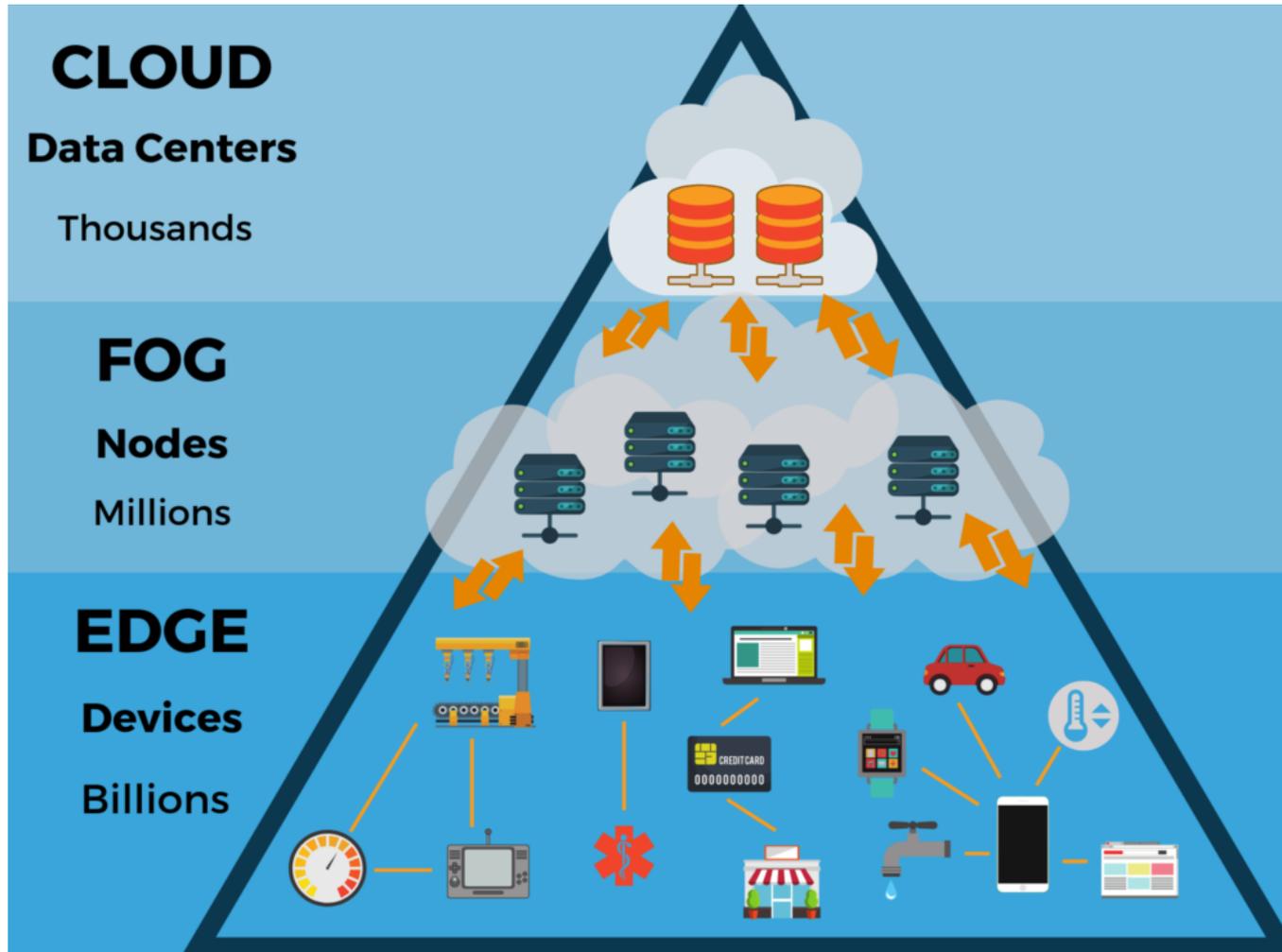


Ecological impact of HPC

- Source: Zwart, S. P. (2020). The ecological impact of high-performance computing in astrophysics. *Nature Astronomy*, 4(9), 819-822.



Edge computing: An energy efficient approach for IoT



...but edge is low-power and low-performance



Nvidia Jetson AGX Xavier:

- K\$
- 10W-15W
- 32 TOPS
- 136.5GB/sec



Coral Google

- 150 \$
- 5-10 W
- 4TOPS



Nvidia Jetson Nano

- 100\$
- 5-10W
- 472GFLOPS
- 25.6GB/s

- ✓ Low Power and Low Performance.
- ✓ Is it enough? -> No, at the moment.
- ✓ Virtualization tech. may help.



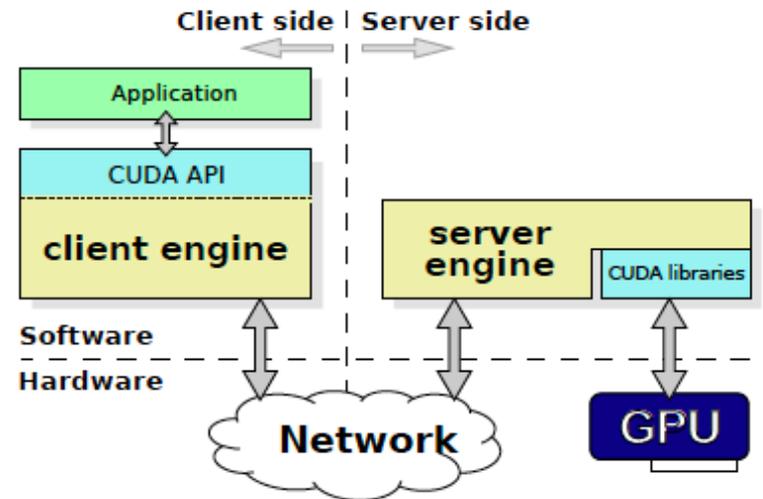
Arduino

- 10\$
- 1W

Using Remote GPU Virtualization Techniques to Enhance Edge Computing Devices

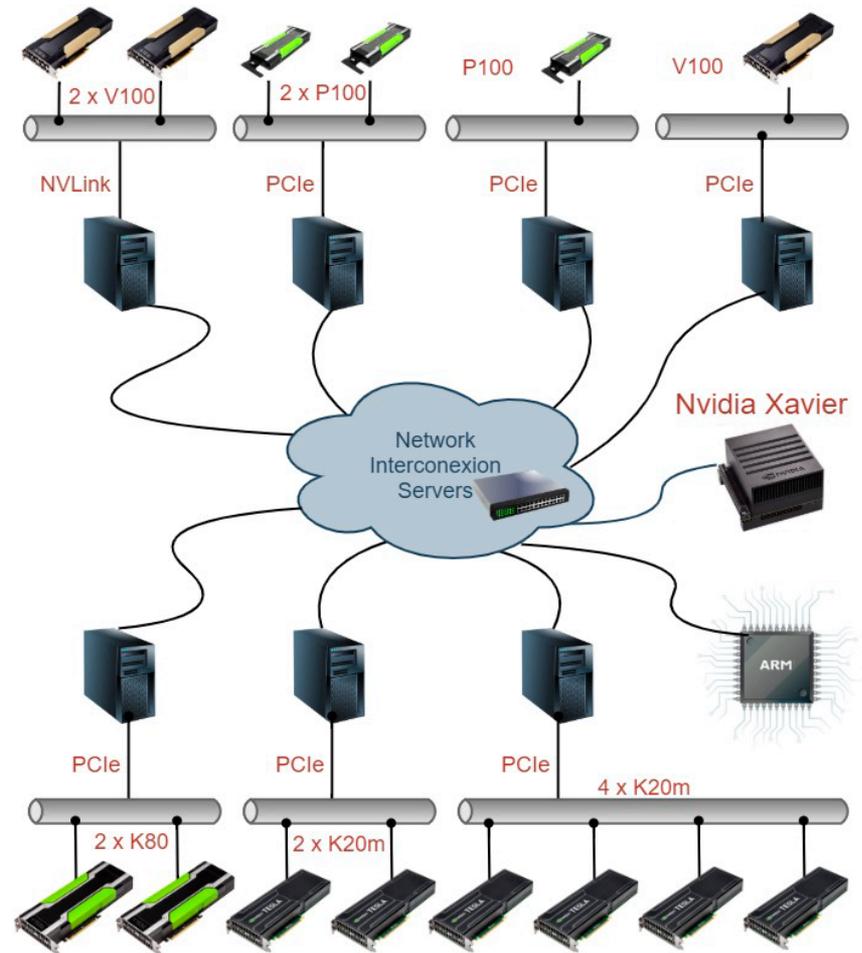
Remote virtualization through rCUDA

- Please visit <http://www.rcuda.net/> for details



Experimental setup

- Gas sensors for home activity monitoring Data Set
- <https://archive.ics.uci.edu/ml/datasets/Gas+sensors+for+home+activity+monitoring>
- 919.438 points.



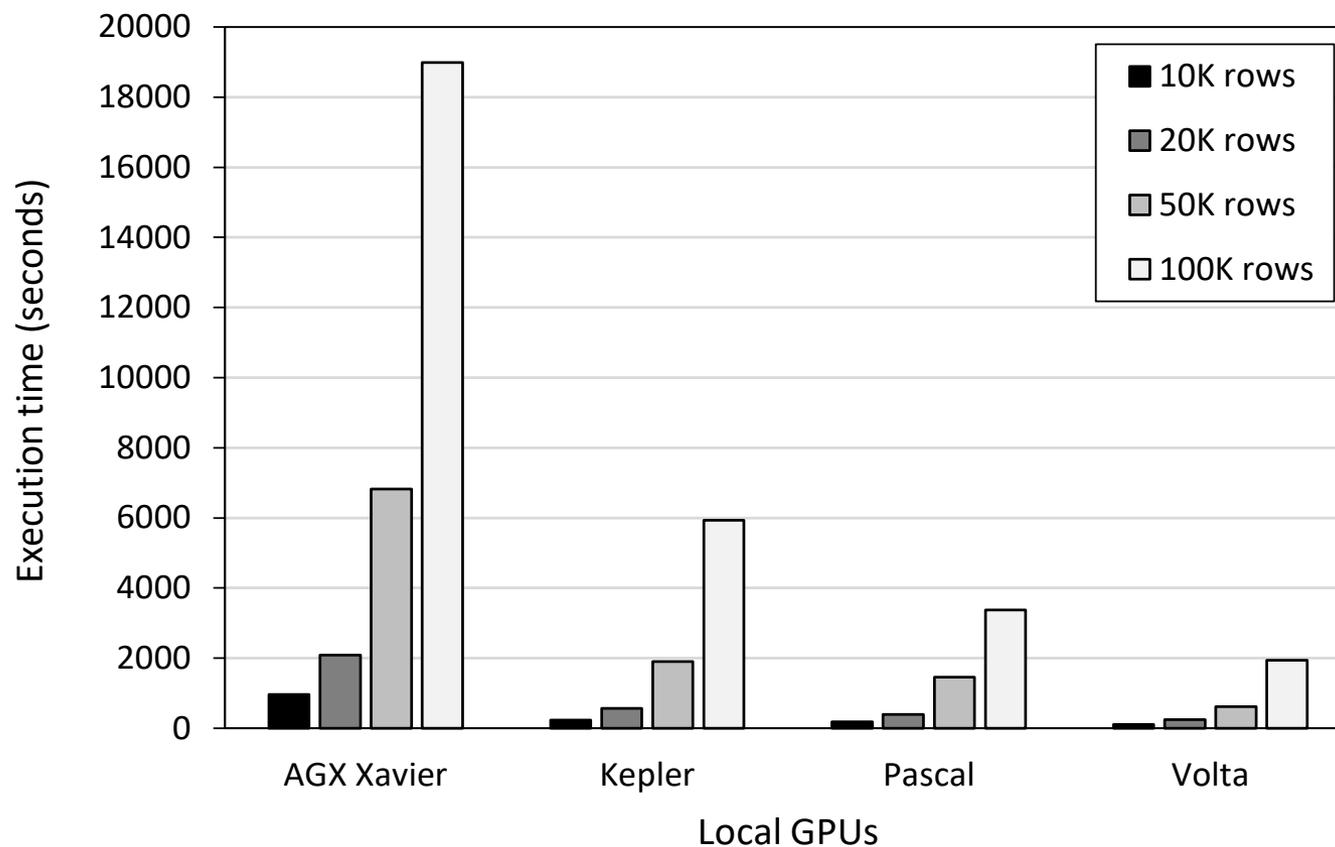
Fuzzy minimals (FM) algorithm

Algorithm 1 Fuzzy Minimals algorithm, where X is the input dataset to be classified, V is the algorithm output that contains the prototypes found by the clustering process. F is the dimension of the vector space.

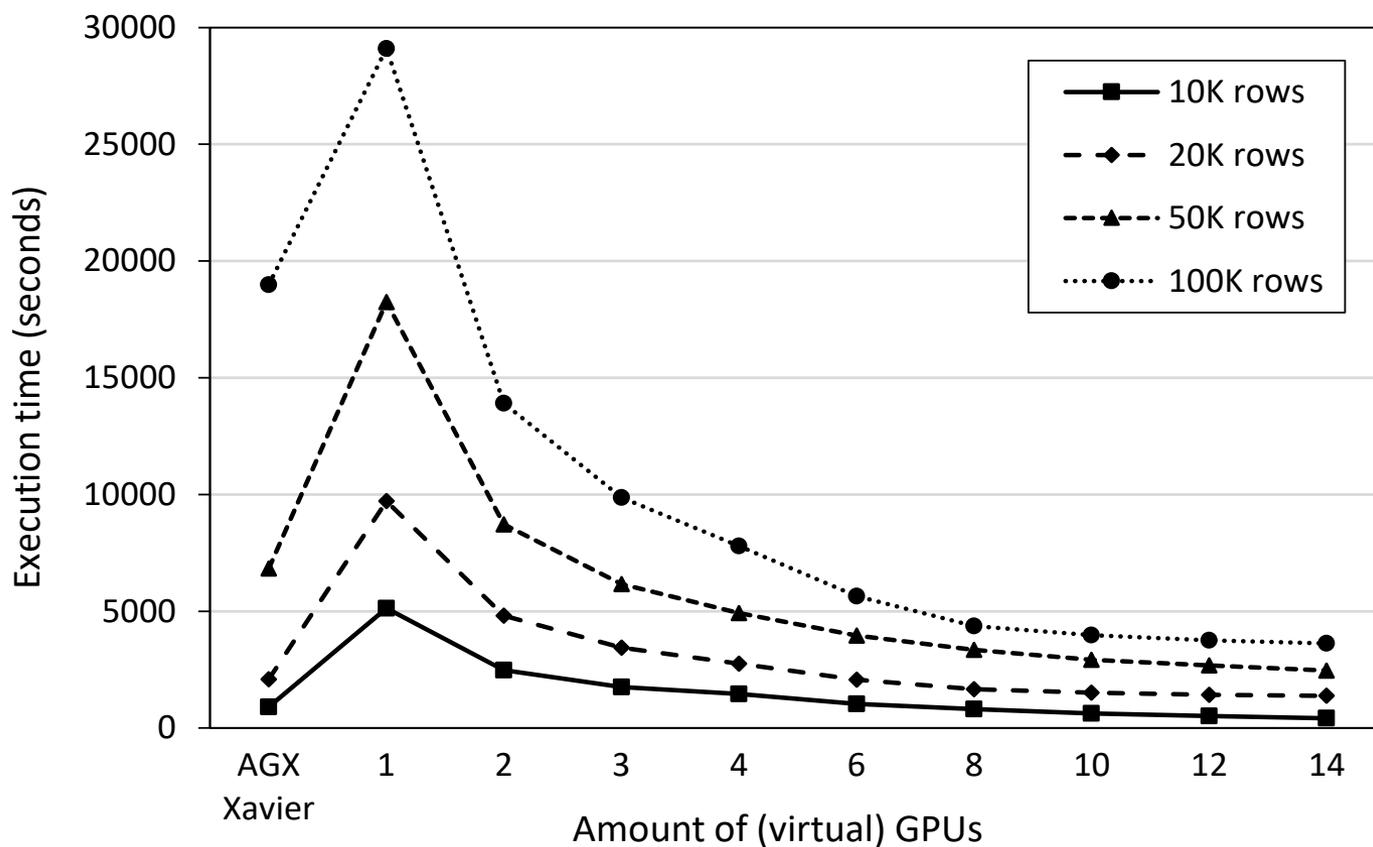
- 1: Choose ε_1 and ε_2 standard parameters.
 - 2: Initialize $V = \{ \} \subset \mathbb{R}^F$.
 - 3: *Load_Dataset*(X)
 - 4: $r = \text{Calculate}_r\text{Factor}(X)$
 - 5: *Calculate_Prototypes*($X, r, \varepsilon_1, \varepsilon_2, V$)
-

$$\frac{\sqrt{|C^{-1}|}}{nr^F} \sum_{x \in X} \frac{1}{1 + r^2 d_{xm}^2} = 1.$$

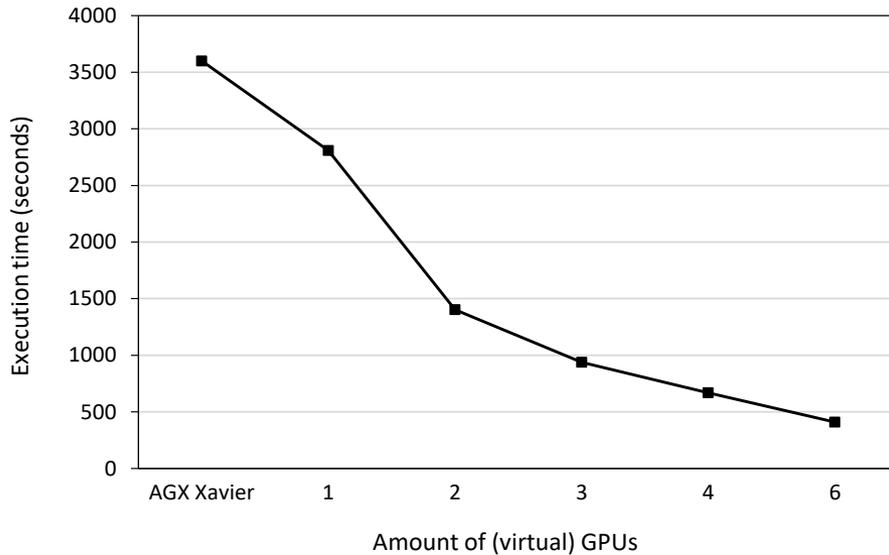
Performance at local GPUs



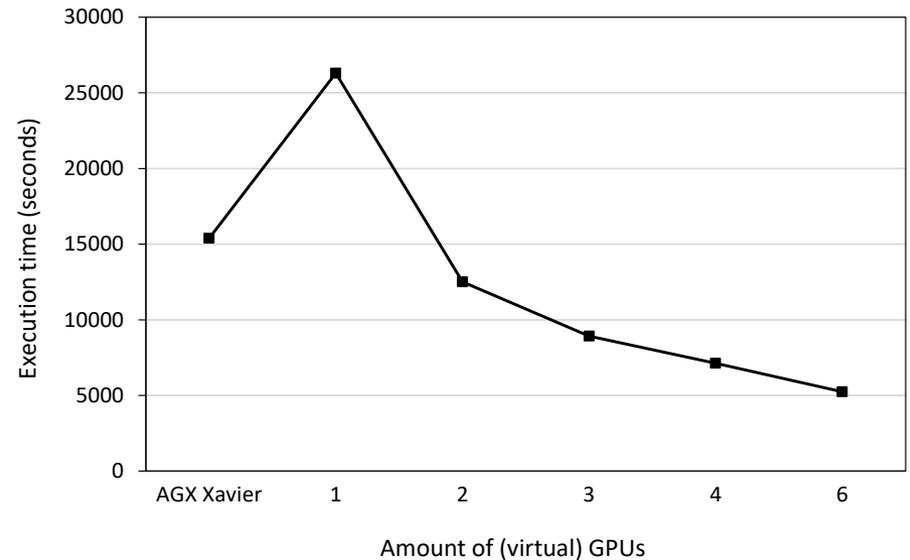
Performance using virtual GPUs



Performance Factor r and prototypes

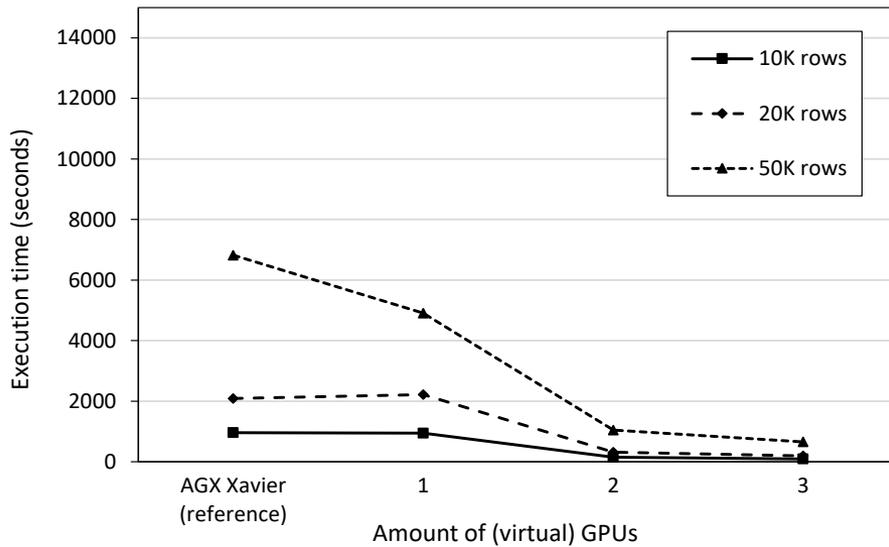


Execution time of *factor r*

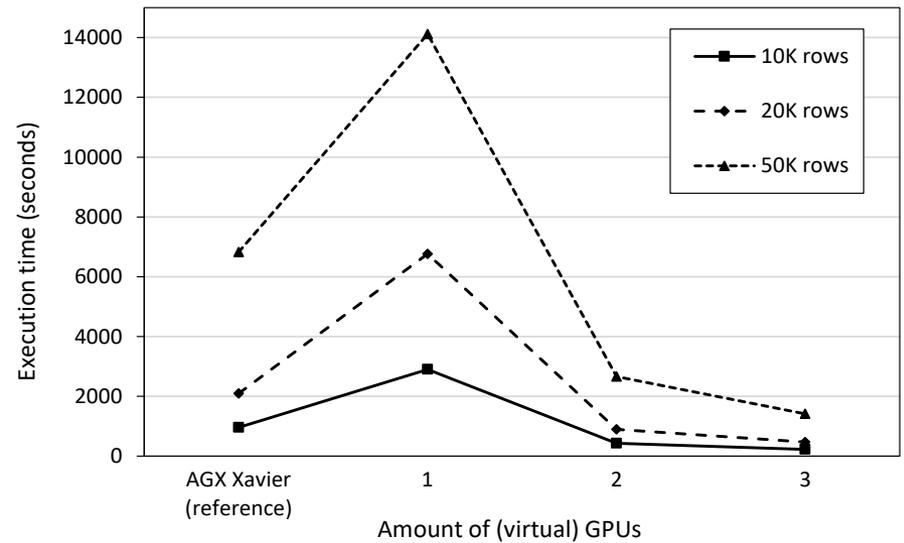


Execution time of prototypes calculation

Communication impact: Ethernet and InfiniBand

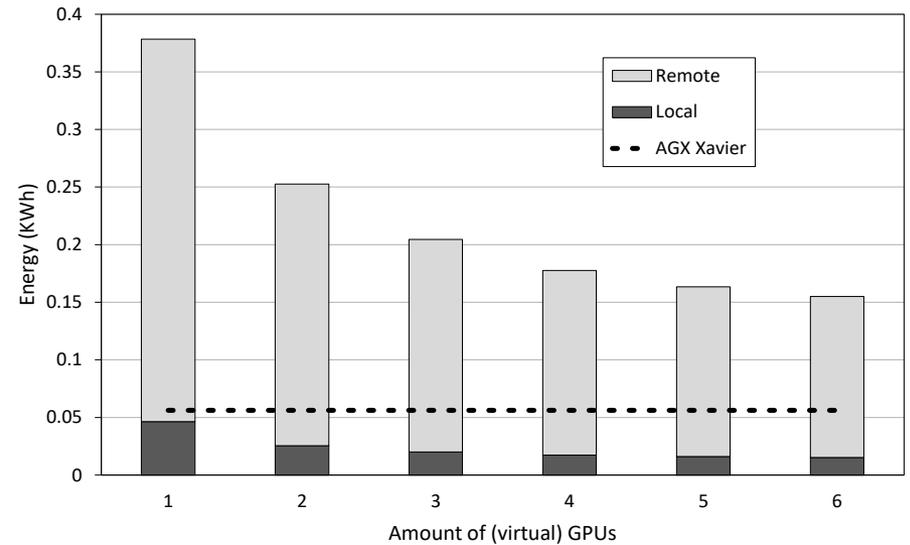
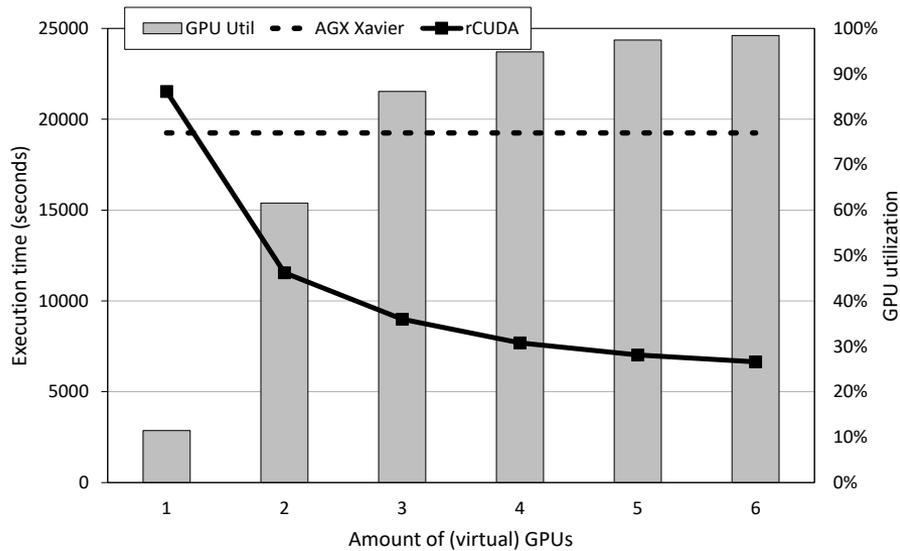


ARM-based node with InfiniBand

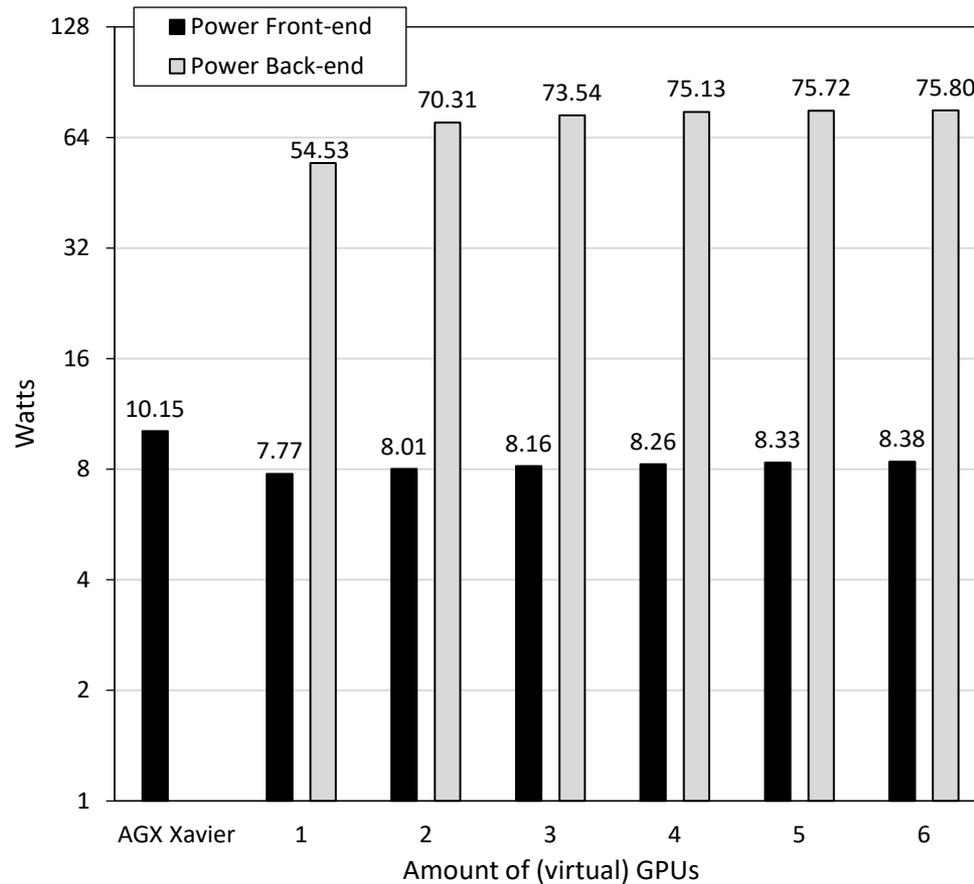


ARM-based node with Ethernet

Execution time & Energy consumption



Average power consumption (Front-end Vs. Back-end)



Conclusiones

- ✓ Remote **virtualization of GPUs** can be a good solution to increase computing power without increasing the power budget of edge devices.
- ✓ We report a performance gain of up to **3.1x speed-up** factor by just using a single remote GPU running multiple virtual GPU instances on it.
- ✓ The power consumption of the edge device is reduced by up to 30%, obtaining up to **80% of energy savings** by delegating the GPU workload to the back-end as its GPU can remain off.

AI-enabled autonomous drones for fast climate change crisis assessment



Phases of disaster management







Autonomous AI swarm drone to help decision makers



Reducing the number of images...

- ... To be processed by humans → Early response and decision making
- ... to decrease the communication bandwidth requirements

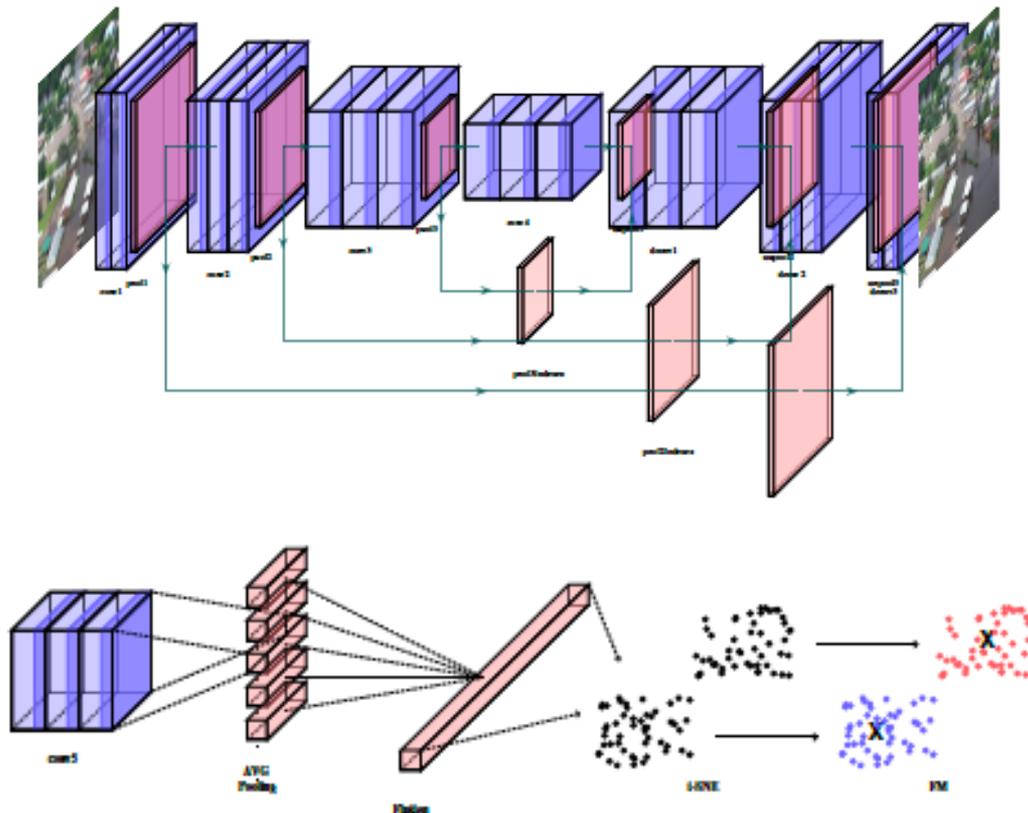
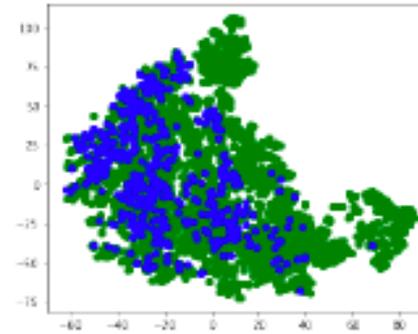


Fig. 2: AI-pipeline for the latent space clustering extracted from the autoencoder.



Fig. 3: Classes within the AIDER dataset.

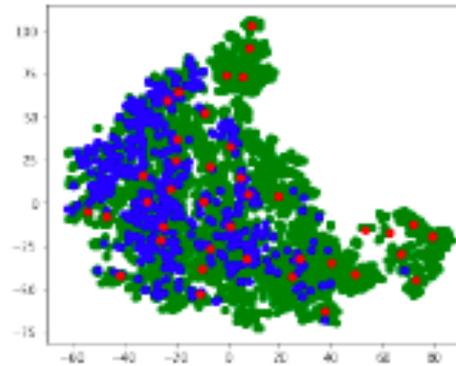


(a) Data points for flooding images (blue), and others (green)

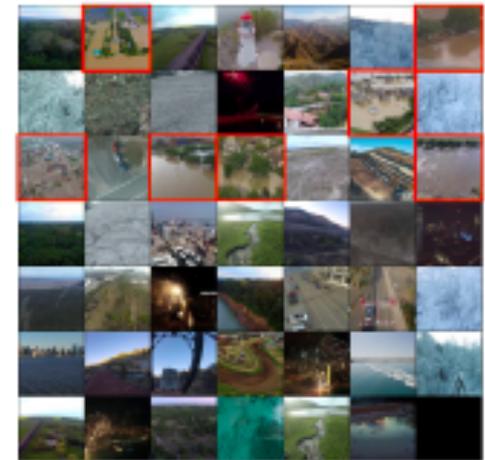


(b) Images corresponding to data points.

Fig. 4: Data points and images after running the t-SNE algorithm.



(a) Data points of flooding images (blue), others (green), and prototypes obtained by the FM algorithm (red).



(b) Images closer to the prototype found by FM.

Fig. 5: Data points and images after running the FM algorithm.

TABLE I: Specification of the various GPU platforms used in our experiments.

	Pedra	Jetson AGX Xavier	Jetson TX2	Jetson Nano
CPU	Intel Silver 4216	NVIDIA Carmel ARM v8.2	ARMv8	ARM Cortex-A57 MPCore
2xGPU (NVIDIA)	GeForce RTX 2080 Ti	Volta	Pascal	Maxwell
Memory [Gib]	12 DDR5	32 LPDDR4x	8 LPDDR4	4 LPDDR4
Size [mm]	N/A	105 x 105	50 x 87	70 x 45
Weight [g]	N/A	280	85	61
Energy consumption [W]	N/A	10-30	7.5	3-5

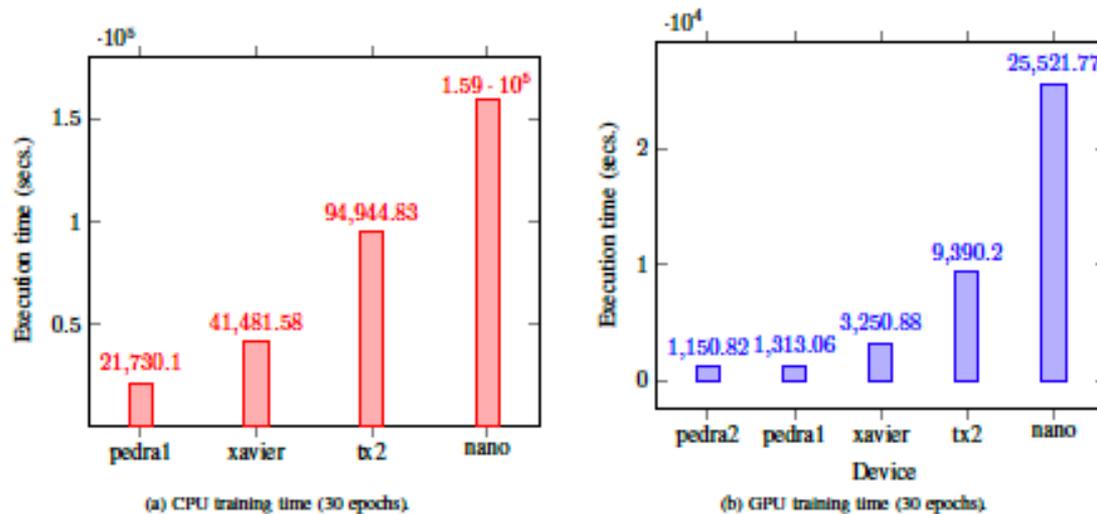


Fig. 7: Execution time for the auto-encoder's training stage.

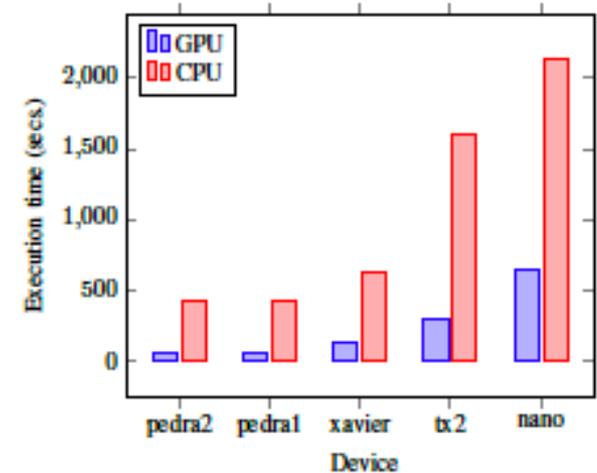


Fig. 8: Inference comparison of GPU/CPU for the entire dataset.

Performance evaluation

Conclusions

- ✓ **Autonomous UAVs** could play a "key role" in addressing the consequences of climate change.
- ✓ The intersection between **AI and edge computing** is undoubtedly the answer today, allowing to make autonomous drones a useful tool under different emergency situations.
- ✓ An AI-based pipeline to be run on edge computing platforms to enable efficient processing of drone images of natural disasters.
- ✓ Our results reveal that **edge computing is a compelling alternative** for processing these heavy computational workloads.



GOBIERNO
DE ESPAÑA

MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

GLOBAL T

<http://www.grc.upv.es/>

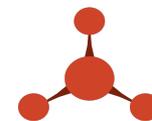
Departamento DISCA. Edificio 1G.

Camí de Vera S/N C.P. 46022.

Valencia

José María Cecilia

jmcecilia@disca.upv.es



SIE

30 *Años*