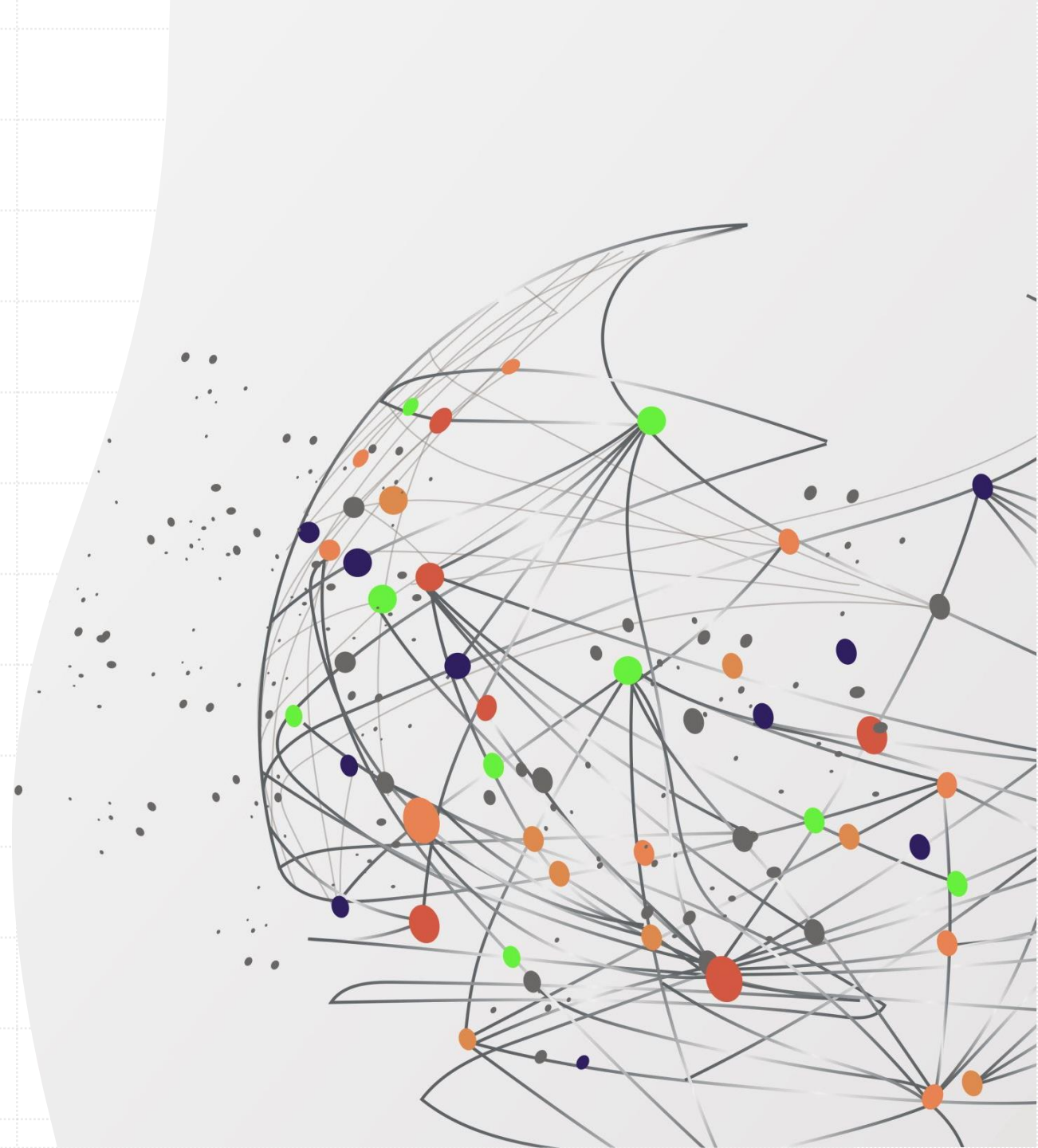


El aprendizaje automático en sistemas embebidos

Por Ignacio Latre Ayen (735913@unizar.es)

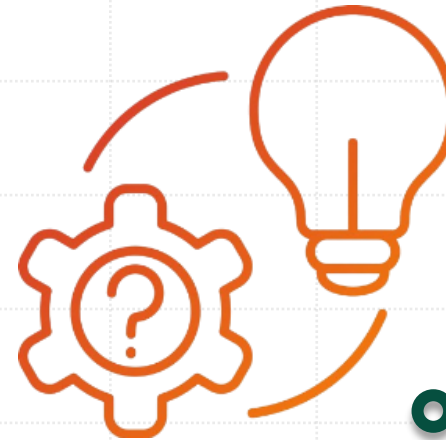
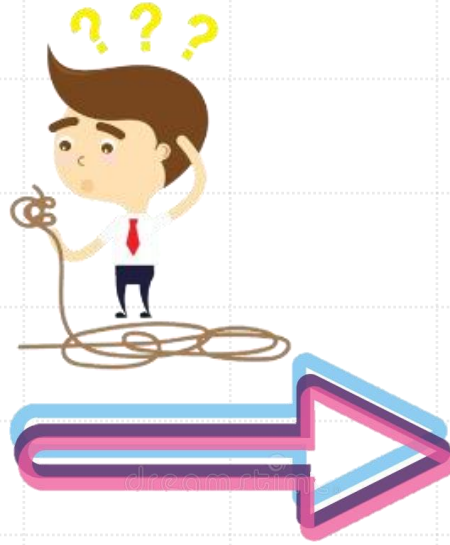
2024



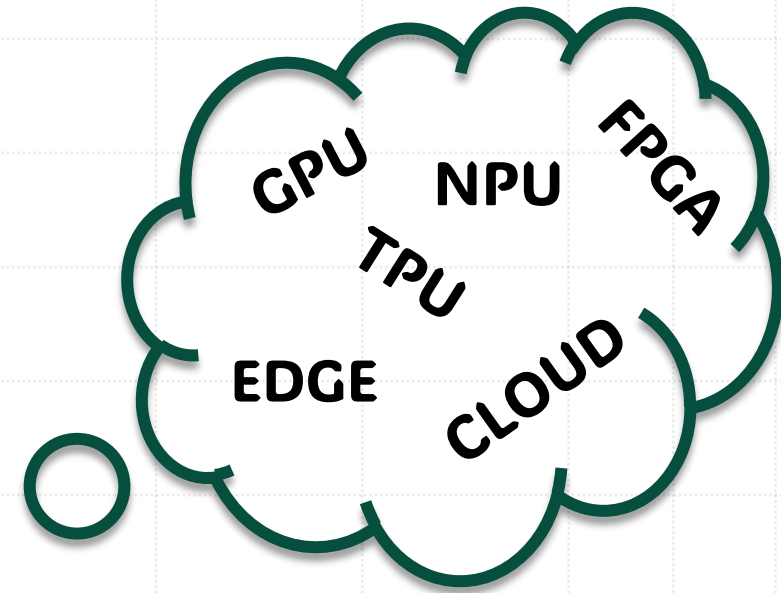
Introducción



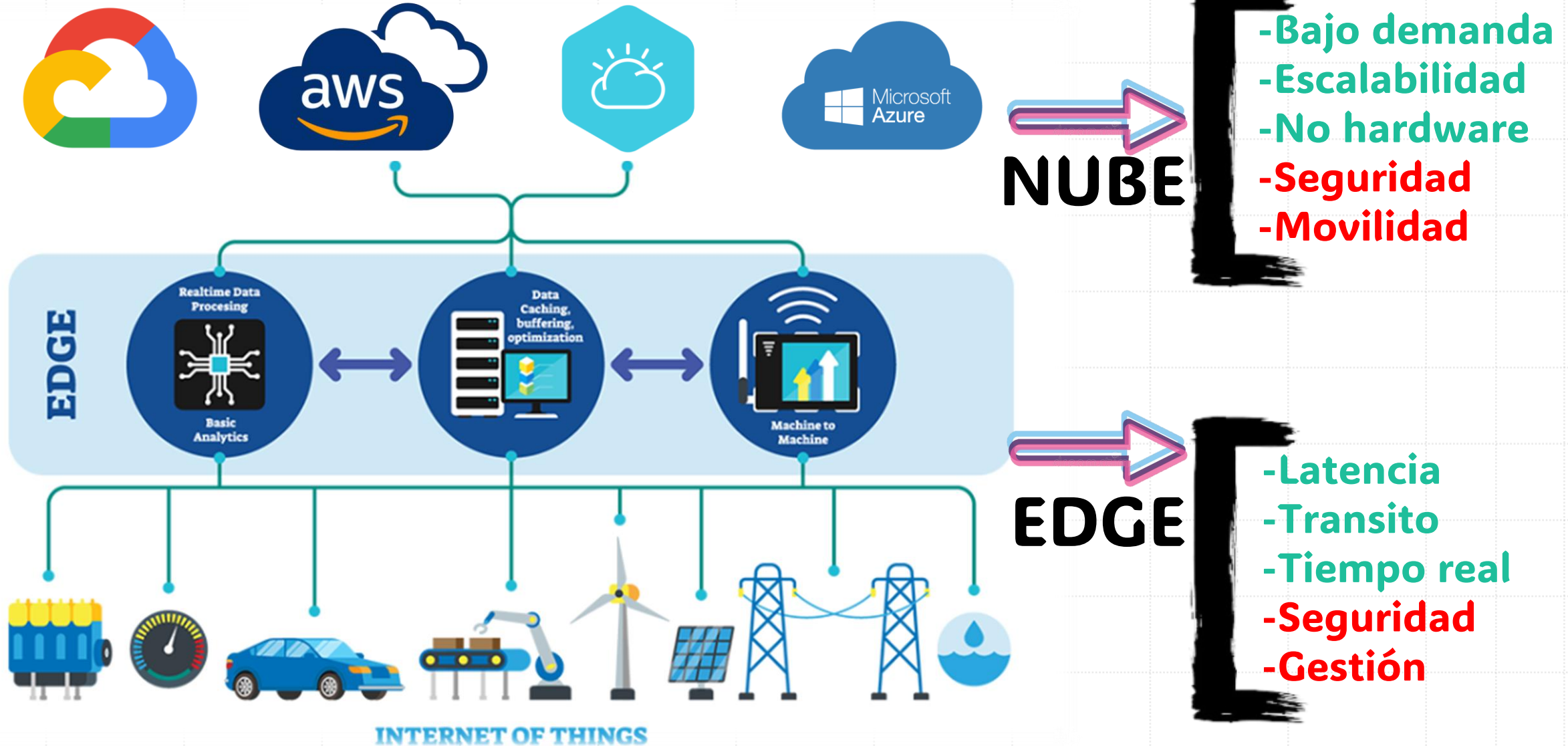
¿Por qué ML?,
¿Por qué embebido?



¿Qué soluciones
hay?

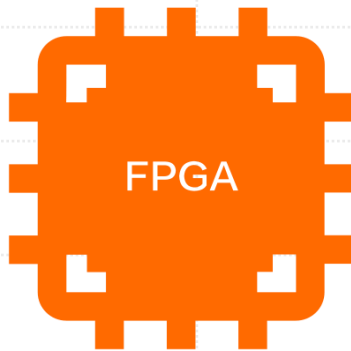


Fuera del dispositivo

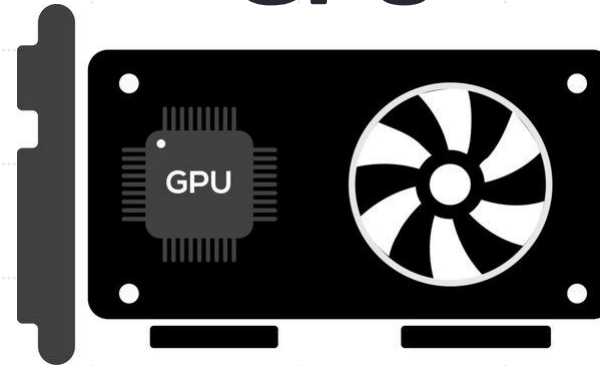


En el dispositivo (no embebido)

FPGA



GPU



TPU



Coma flotante



Disponibilidad



XILINX®



NVIDIA



Google



Consumo

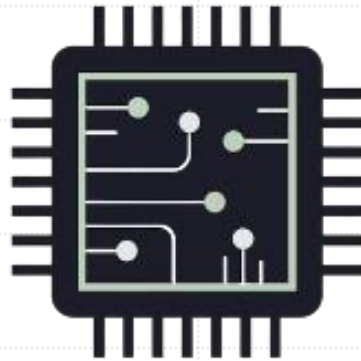


Paralelismo



En el dispositivo (embebido)

ASIC



Velocidad



Eficiencia



Tamaño



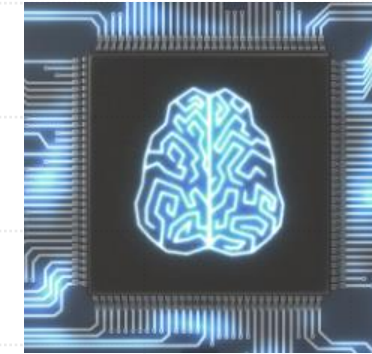
Paralelismo



Precio



NPU



Optimización



¿Y sí no es posible utilizar hardware especializado?

**Coma
fija**

**Sustituir el formado
de coma flotante a
coma fija**

**Pesos
reducidos**

**Usar las
arquitecturas
adecuadas y evitar
algoritmos muy
pesados como las
redes neuronales
profundas**

**Compr
esión**

- Técnicas de poda
- Técnicas de cuantización
- Técnica de destilación
- Técnicas de regularización

Conclusiones

¿Qué método de implementación es mejor?



Pues depende....

Consumo

Cantidad

Algoritmos

Precio

Tamaño

Conectividad

