

Jaime Pérez

Sergio Pérez

José M. Moya

Patricia Arroba



THERMAL PREDICTION FOR IMMERSION COOLING DATA CENTERS BASED ON RECURRENT NEURAL NETWORKS

TABLE OF CONTENTS

1. Motivation

2. Objectives
3. State of the art
4. Solution
5. Experiments
6. Conclusions and future directions



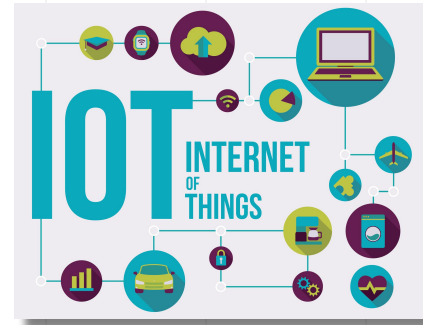
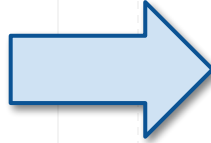
2050

Percentage urban

- 75 or over
- 50 to 75
- 25 to 50
- Less than 25
- No data

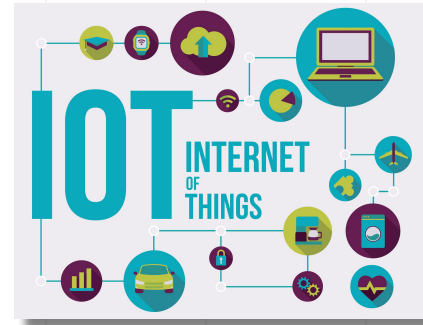
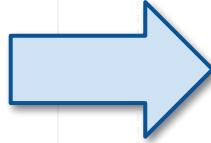
Source: United Nations, UN





100.000.000.000 IoT
Devices in 2030

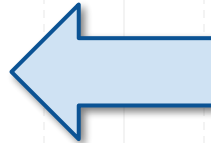




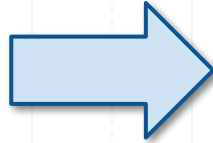
100.000.000.000 IoT
Devices in 2030



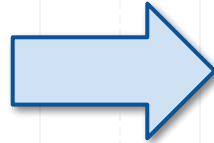
How will we process this
large amount of data?



CLCUD COMPUTING



- Network saturation
- High latencies
- E.g. Self-driving car reaction time at 100 Km/h
 - ↳ Cloud: 2.5 **m** (~100 ms)
 - ↳ Edge: 2.5 **cm** (~10 ms)

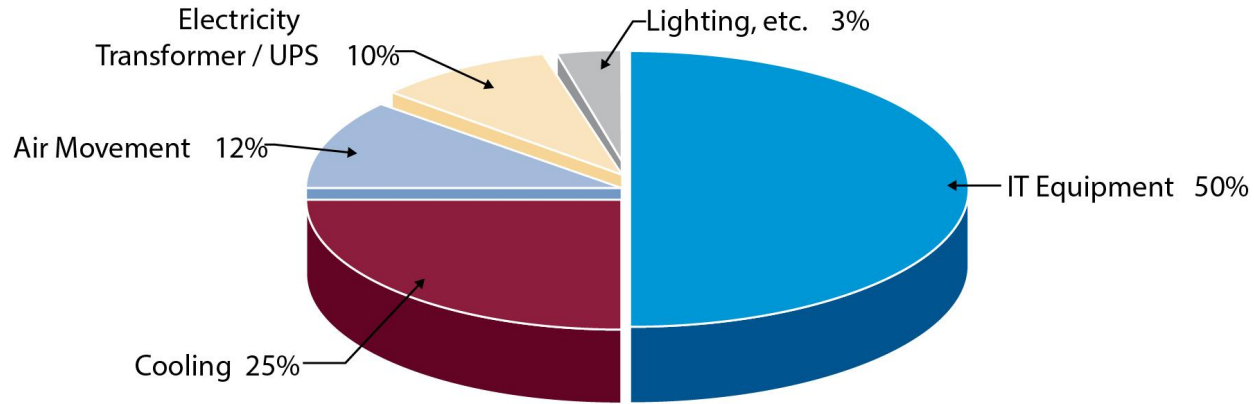


Main challenges:

- Climate-Independent cooling
- Reduce energy consumption
- Reduce area (Power Density)

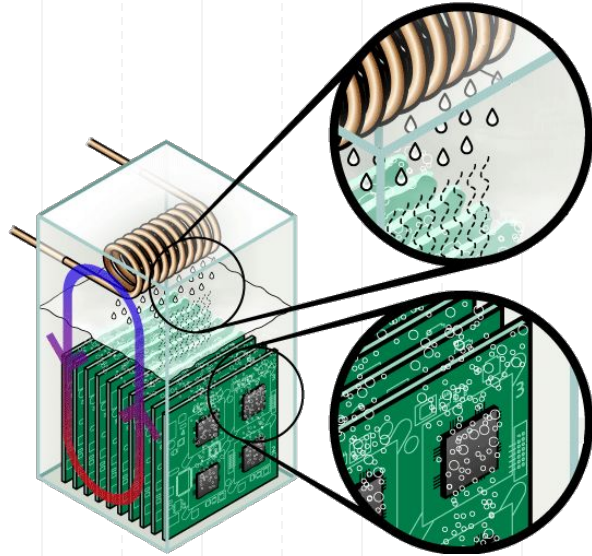
ENERGY BREAKDOWN OF DATA CENTERS

Cooling Energy \approx 40%



Source: EYP Mission Critical Facilities Inc.

TWO-PHASE PASSIVE IMMERSION COOLING IN HYDRO-FLUORO-ETHERS



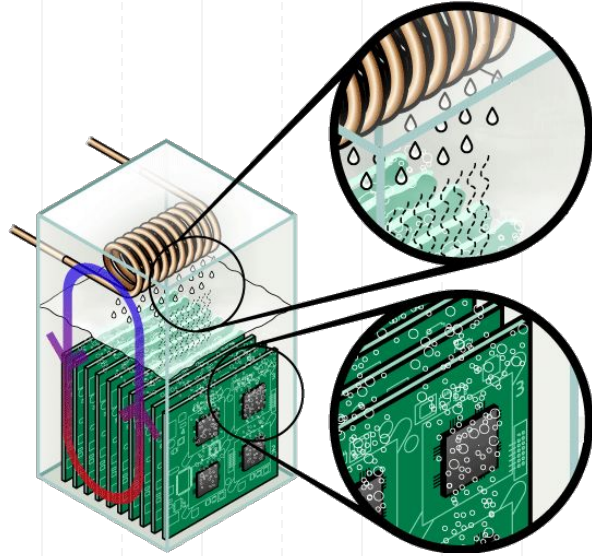
3M
Novec

- Climate independent cooling
- Passive cooling
- High power density

↳ From 40 kW to 250 kW

↳ Reduce area

TWO-PHASE PASSIVE IMMERSION COOLING IN HYDRO-FLUORO-ETHERS



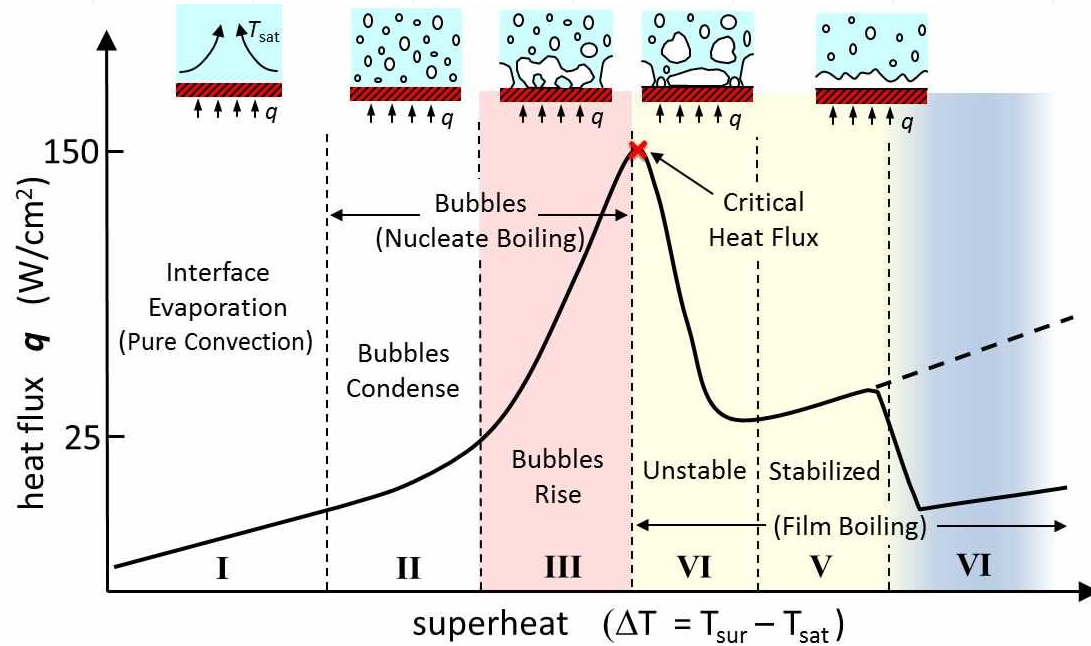
3M
Novec

- How can we protect devices from critical temperatures?
- How can we optimize it?

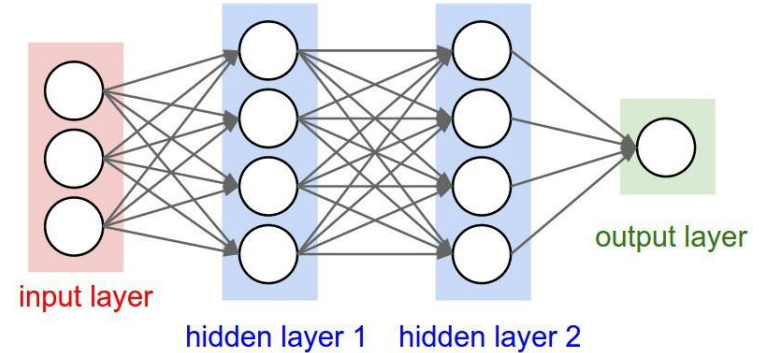
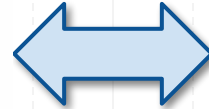
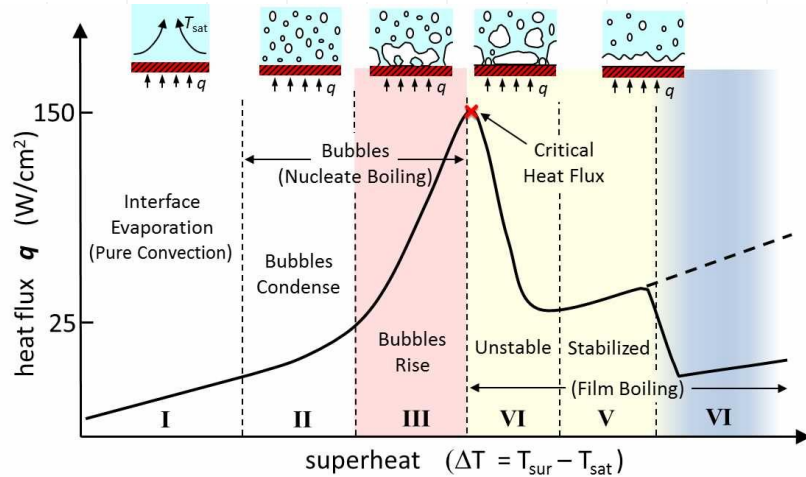
TABLE OF CONTENTS

1. Motivation
2. Objectives
3. State of the art
4. Solution
5. Experiments
6. Conclusions and future directions

EFFICIENCY AND SAFETY



EFFICIENCY AND SAFETY



Predictive models of CPU temperature using neural networks

TABLE OF CONTENTS

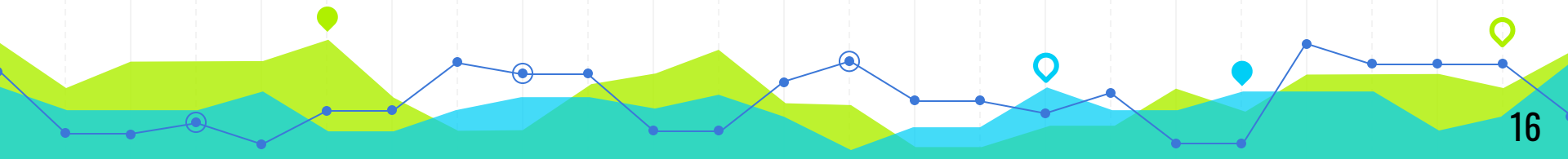
1. Motivation
2. Objectives
- 3. State of the art**
4. Solution
5. Experiments
6. Conclusions and future directions

WHY RECURRENT NEURAL NETWORKS?

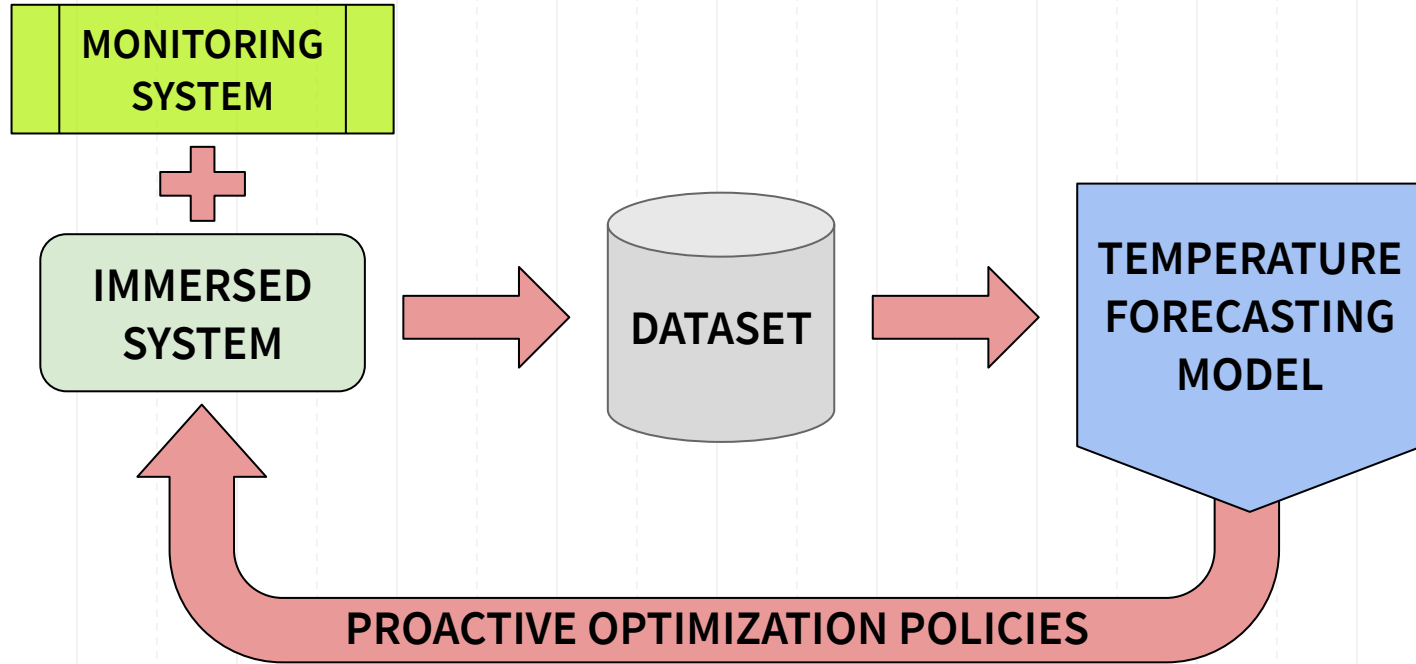
- Able to **learn** complex **non-linear relations** between features
- Native support
 - ↳ Data in **time series** form
 - ↳ **Multiple inputs** and outputs
 - ↳ **Multi-step** forecasting
- Easy deployment with **Tensorflow + Keras**

TABLE OF CONTENTS

1. Motivation
2. Objectives
3. State of the art
- 4. Solution**
5. Experiments
6. Conclusions and future directions

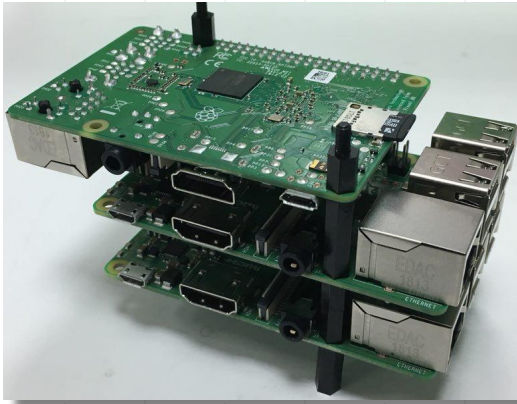


METHODOLOGY



PROTOTYPE

CLUSTERS OF RASPBERRY PI 3 MODEL B+



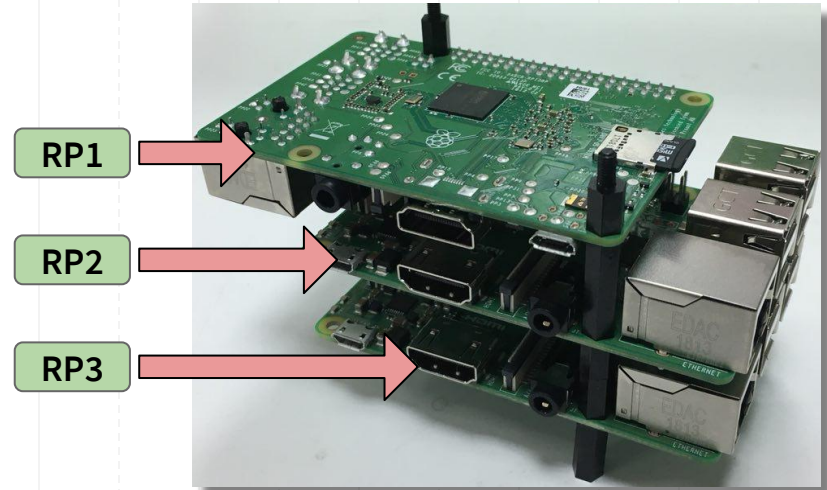
MONITORING SYSTEM

COLLECTD & GRAPHITE

Selected metrics:

- Temperature (**RP1**)
- Temperature (**RP2**)
- CPU Utilization (**RP2**)
- CPU Frequency (**RP2**)
- Temperature (**RP3**)

Workload: Data Analytics



MODEL EVALUATION METRICS

Mean Absolute Error

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - x_j|$$

Standard Deviation

$$STD = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \mu)^2}$$

Coefficient of Determination

$$R^2 = \frac{\sum_{j=1}^n (y_j - x_j)^2}{\sum_{j=1}^n (y_j - \bar{y}_j)^2}$$

Root Mean Square Deviation

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - x_j)^2}$$

TABLE OF CONTENTS

1. Motivation
2. Objectives
3. State of the art
4. Solution
- 5. Experiments**
6. Conclusions and future directions



EXPERIMENTS METHODOLOGY

TENSORFLOW + KERAS

Base Model

- ↳ **Optimizer:** Nadam
- ↳ **Loss Function:** MAE
- ↳ **Time-Step:** 1 min

Optimized Model (Fully Connected, LSTM, GRU)

Set Architecture

- ↳ N. Layers
- ↳ N. Neurons

Optimize Train

- ↳ Batch Size
- ↳ Epochs
- ↳ Activation function

Test

FULLY CONNECTED NETWORK

Neural Architecture	Batch Size	Epochs	MAE \pm STD ($^{\circ}\text{C}$)	RMSD ($^{\circ}\text{C}$)	R ² (%)
3 layers: 16 - 32 - 16	65	80	1.11 \pm 1.14	1.588	71.925
2 layers: 16 - 8	65	80	1.13 \pm 1.08	1.562	72.824
1 layer: 256	65	80	1.09 \pm 0.83	1.371	79.071
1 layer: 256	95	140	0.84 \pm 0.84	1.187	83.897

LSTM NETWORK

Neural Architecture	Batch Size	Epochs	MAE \pm STD ($^{\circ}\text{C}$)	RMSD ($^{\circ}\text{C}$)	R ² (%)
3 layers: 8 - 4 - 2	65	80	0.95 \pm 1.13	1.475	75.782
2 layers: 1 - 1	65	80	1.02 \pm 0.82	1.311	80.871
1 layer: 2	65	80	0.91 \pm 0.89	1.270	82.044
1 layer: 2	95	210	0.71 \pm 0.66	0.969	87.527

GRU NETWORK

Neural Architecture	Batch Size	Epochs	MAE \pm STD ($^{\circ}\text{C}$)	RMSD ($^{\circ}\text{C}$)	R ² (%)
3 layers: 4 - 2 - 1	65	80	0.81 \pm 0.83	1.157	82.211
2 layers: 4 - 2	65	80	0.85 \pm 0.78	1.150	82.429
1 layer: 4	65	80	0.74 \pm 0.68	1.002	86.668
1 layer: 4	95	120	0.75 \pm 0.59	0.957	89.539

GRU NETWORK

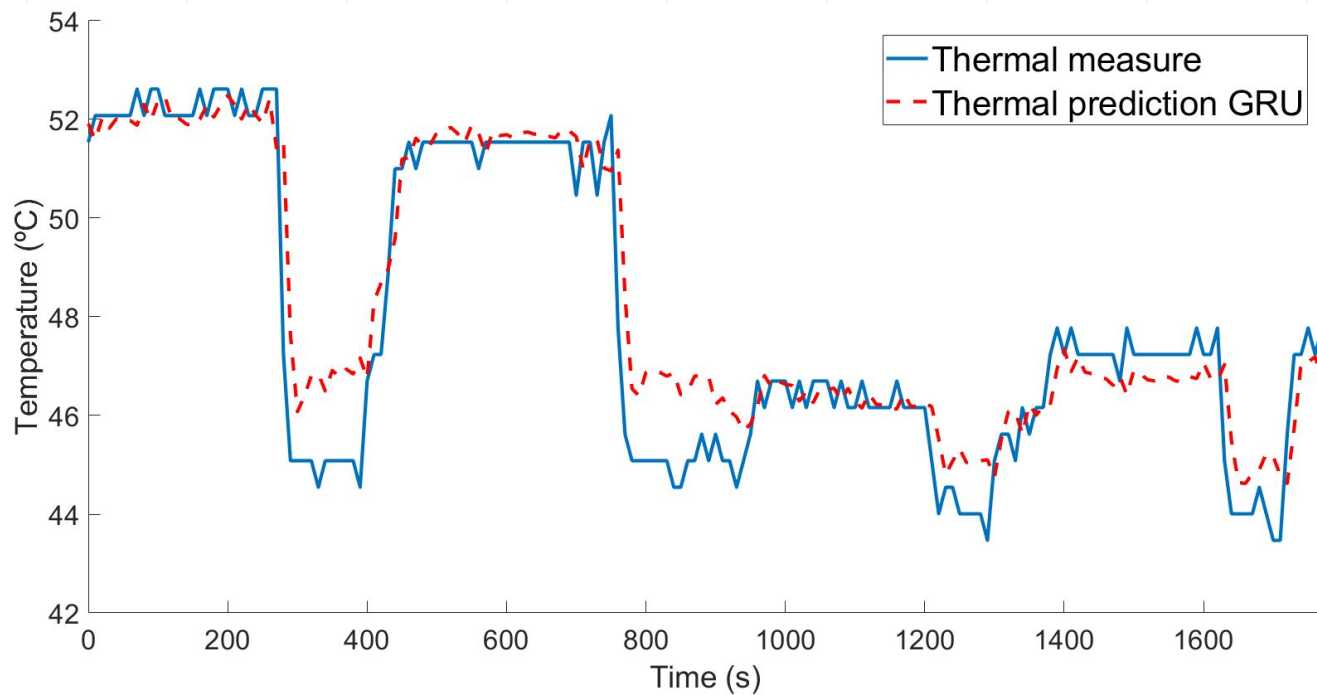


TABLE OF CONTENTS

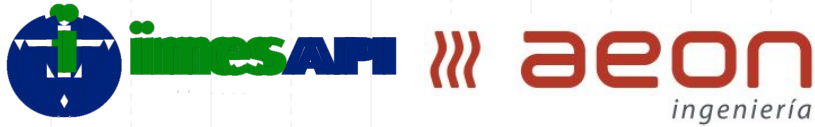
1. Motivation
2. Objectives
3. State of the art
4. Solution
5. Experiments
6. Conclusions and future directions



CONCLUSIONS

- **Temperature forecasting model** (1 min) using GRU neurons
 - ↳ $MAE \pm STD = 0,75 \pm 0,59 \text{ }^{\circ}\text{C}$, $RMSD = 0.957 \text{ }^{\circ}\text{C}$, $R^2 = 89,539 \%$
- Allows the development of automatic and online models
 - ↳ Minimize human assistance
- Solution for the **deployment of Edge Data Centers** in urban areas
 - ↳ Climate independent cooling
 - ↳ Cooling energy $\approx 1\%$
 - ↳ Highly optimized systems

FUTURE DIRECTIONS



THANK YOU FOR YOUR ATTENTION!

Jaime Pérez Sánchez



jaime.perez.sanchez@alumnos.upm.es



www.linkedin.com/in/jaime-perez-sanchez/

THANK YOU FOR YOUR ATTENTION!

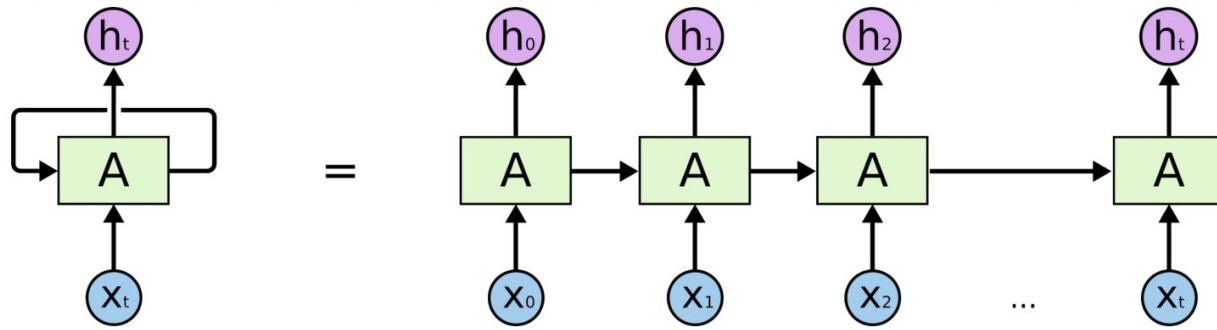
Jaime Pérez Sánchez



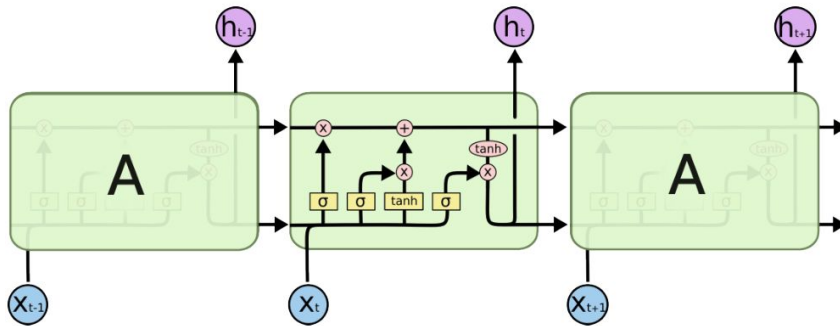
jaime.perez.sanchez@alumnos.upm.es



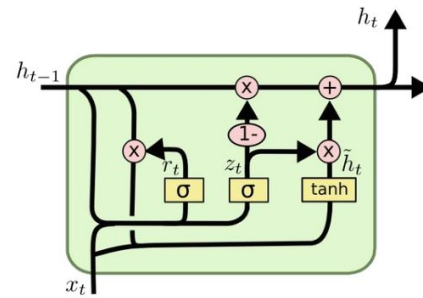
www.linkedin.com/in/jaime-perez-sanchez/



1.a. Simple RNN neuron



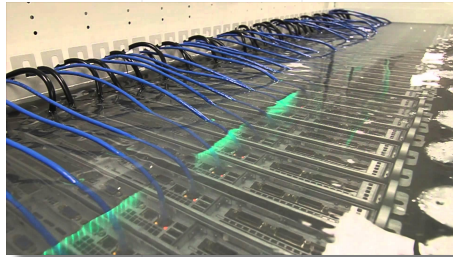
1.b. LSTM neuron



1.c. GRU neuron

IMMERSION COOLING METHODS

One phase immersion



Two phases immersion



ENERGY CONSUMPTION OF DATA CENTERS

- Tianhe-2 (Top 1) > 200.000 MWh / year
- 2020: Top 1 Data Center \approx 1 nuclear plant
- Worldwide:

500 000 Data Centers
Energy: 200.000.000.000 € / year

