
# WIND POWER PRODUCTION FORECASTING USING ANT COLONY OPTIMIZATION AND EXTREME LEARNING MACHINES

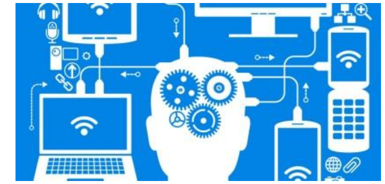
**MARIA CARRILLO[1], JAVIER DEL SER[1,2,3], MIREN NEKANE BILBAO[1], CRISTINA PERFECTO[1], DAVID CAMACHO[4]**

1: UNIVERSITY OF THE BASQUE COUNTRY EHU/UPV, BILBAO (BIZKAIA, SPAIN)
2: TECNALIA RESEARCH & INNOVATION, ZAMUDIO (BIZKAIA, SPAIN)
3: BASQUE CENTER FOR APPLIED MATHEMATICS (BCAM), BILBAO (BIZKAIA, SPAIN)
4: UNIVERSIDAD AUTÓNOMA DE MADRID, MADRID (MADRID, SPAIN)

# Outline of the Presentation

- Introduction and motivation
- Research hypothesis
- Contribution and Objectives
- Problem Formulation
- Proposed Scheme
- Experimental results and discussion
- Conclusions
- Future Research Lines

# INTRODUCTION: DIGITALIZATION OF THE ENERGY GRID (1/2)

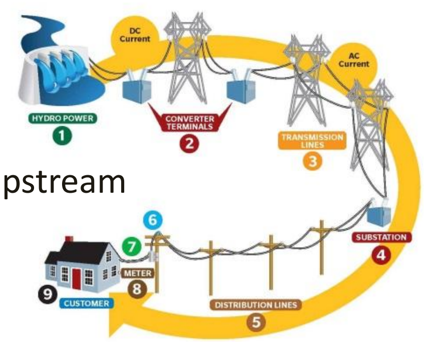


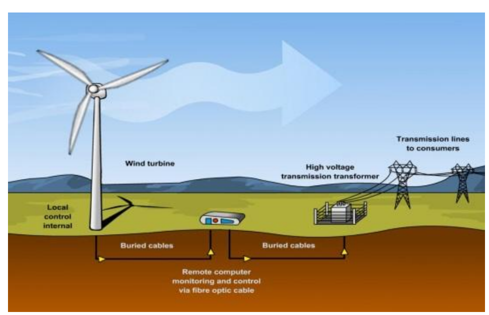
## Digitalization of the energy grid

- Thanks to the deployment of the ICT-powered infrastructure
- Bidirectional information flows from the grid to the operator, supervisor or customer:
  - Demand side management
  - Fraud detection
  - Improved energy efficiency in buildings



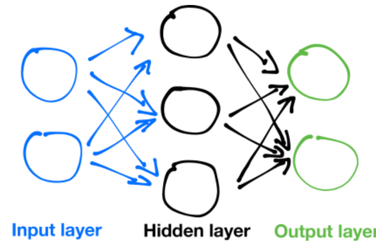
- Particular byproduct: rich data substrate of energy consumptions
  - Match overall generation to consumption
  - Detect abnormal patterns in the consumed energy traces (tampering)
- Renewable energy sources
  - Maximization of the installation productivity
  - Pattern characterization of the produced energy towards its injection upstream



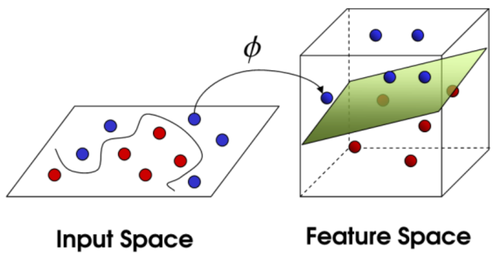



# INTRODUCTION: PREDICTIVE MODELS (2/2)

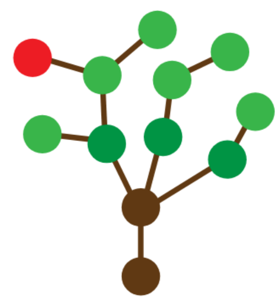
- Naïve Machine Learning approaches
  - –Neural Networks



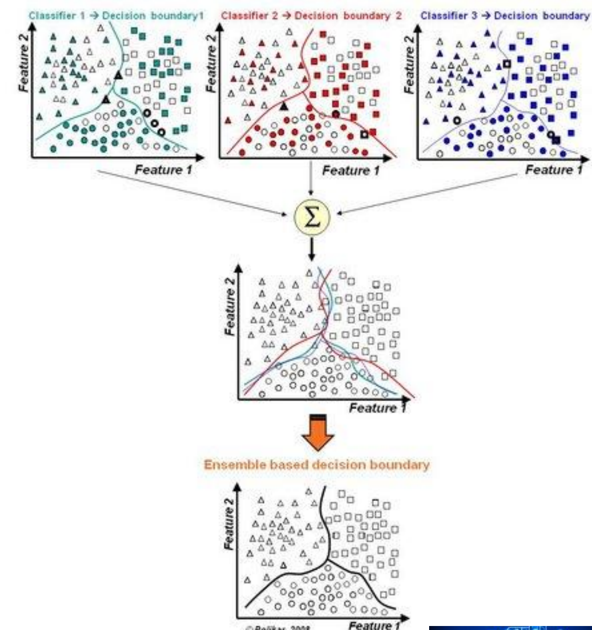

  - –Support Vector Machines




  - –Decision Tree Regressors



  - –Model ensembles




  - –Deep Learners

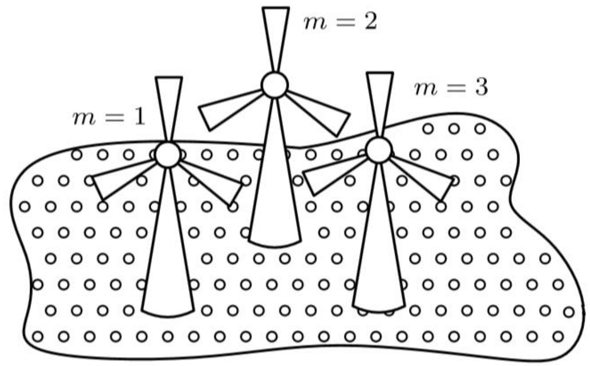
- Hybridization of { nature-inspired heuristics / machine-learning models }

Good for very **big datasets**

# CONTRIBUTION AND OBJECTIVES

This manuscript proposes a hybrid ACO-ELM model to

1) Predict the **total instantaneous power Pt [W]** at time t when being at t - $\Delta_t$
2) Finding an **optimal feature subset** such that the generalization performance of the **predictive model $M_\theta$** is maximized
3) Extending the considered data with a $\Delta_x$ sized window (power values and meteorological variables) before the instant t - $\Delta_t$ at hand




DATASET:

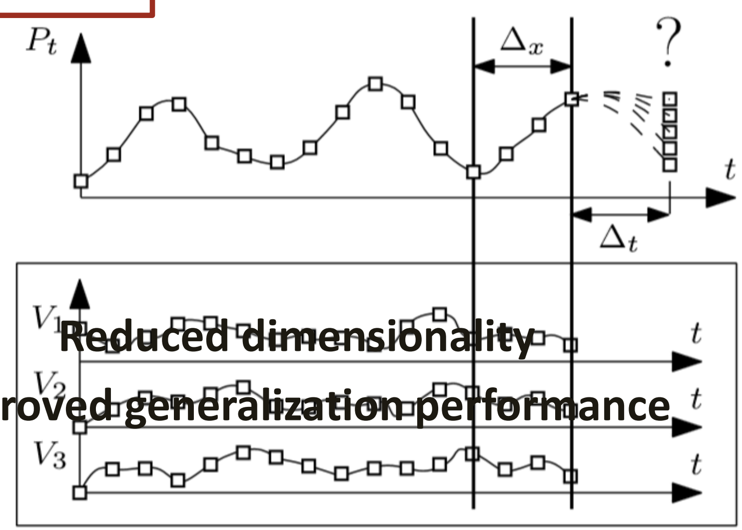
- Total instantaneous power $P_t$[W] at time t
- V meteorological variables obtained at different points
  1. Wind vector modulus
  2. Wind direction
  3. U/V components
  4. Temperature

  by NWP model

- **Results** for two farms in Spain

**HYPOTHESIS**

Proposed hybrid scheme excels at constructing datasets of




**Reduced dimensionality**

**Improved generalization performance**

# Problem Formulation

$$\mathbf{X}_t \doteq \{\mathbf{X}_t^{\Diamond}, \mathbf{X}_{t-1}^{\Diamond}, \ldots, \mathbf{X}_{t-\Delta_X}^{\Diamond}, P_t, P_{t-1}, \ldots, P_{t-\Delta_X}\}$$

Where:

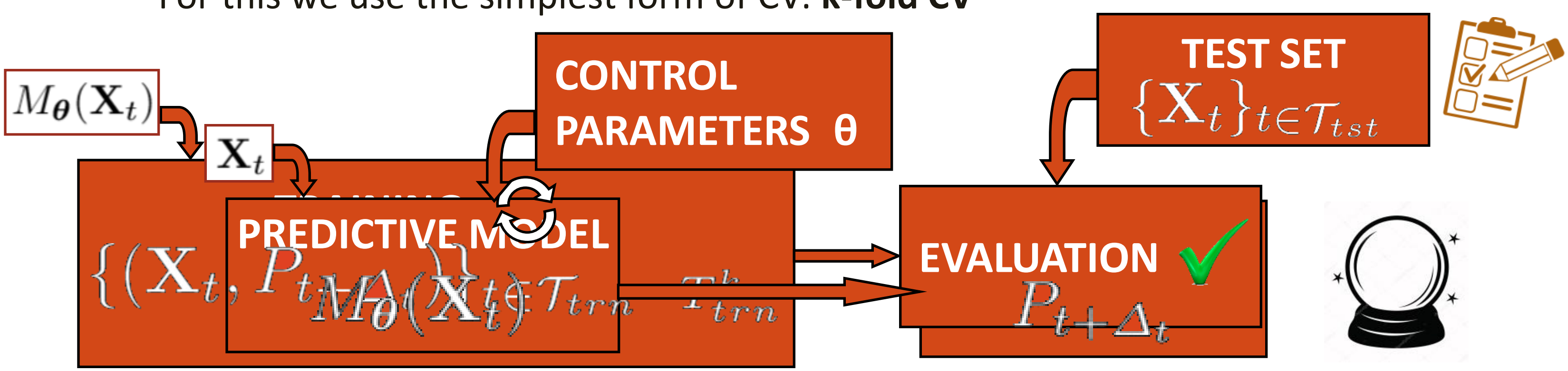
$$\mathbf{X}_t^{\Diamond} \doteq \{\mathbf{X}_t^{\Diamond,\mathbf{p}_1}, \mathbf{X}_t^{\Diamond,\mathbf{p}_2}, \ldots, \mathbf{X}_t^{\Diamond,\mathbf{p}_P}\}; \quad \mathbf{p} \in \{\mathbf{p}_1, \ldots, \mathbf{p}_P\}; \; \mathbf{p}_i \in \mathbb{R}^2$$

And: $P_{t+\Delta_t} = M_{\boldsymbol{\theta}}(\mathbf{X}_t)$ such that

$$\underset{\mathcal{X}' \subseteq \mathcal{X}}{\text{Maximize}} \; \widehat{\varphi}(\mathcal{X}') \doteq \frac{1}{K}\sum_{k=1}^{K} \varphi(k, \mathcal{X}'); \qquad \varphi(k, \mathcal{X}) \in \mathbb{R}^+$$

- Feature selection problem that can be formulated as an optimization problem
  - Fitness metric quantifies the generalization performance of the model when processing unseen data instances
  - For this we use the simplest form of CV: **k-fold CV**

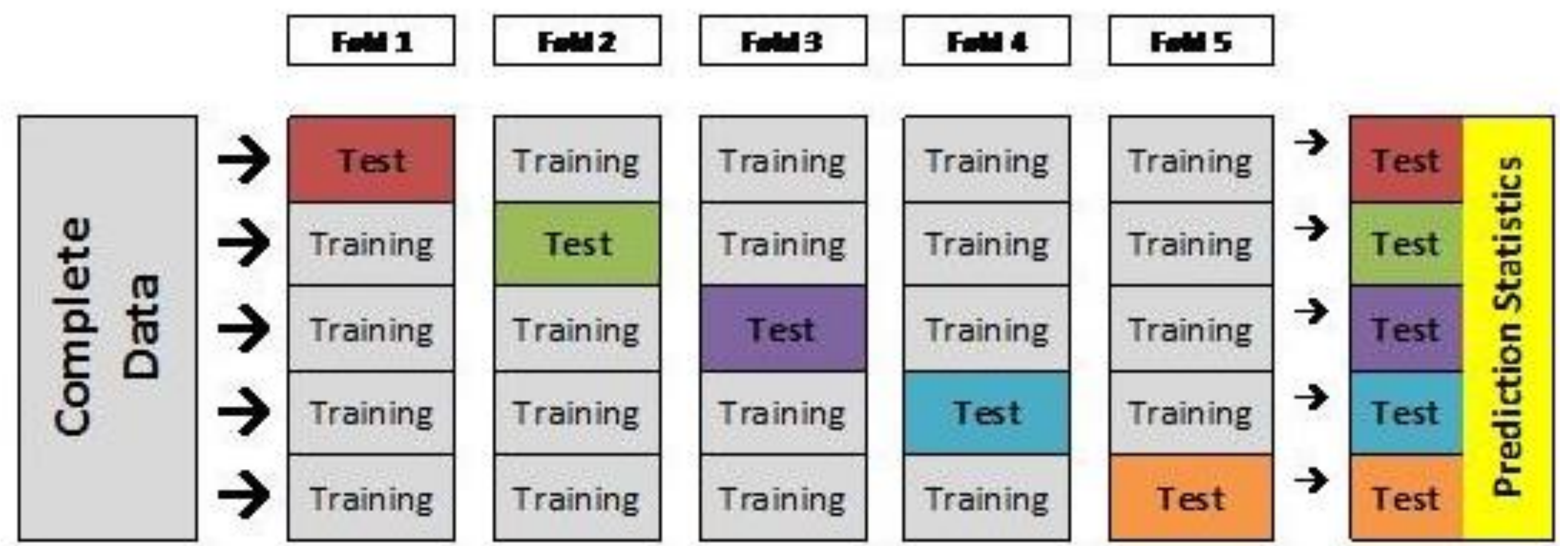
How will the predictive model generalize to an independent set?



# INGREDIENT #1: k-fold CV

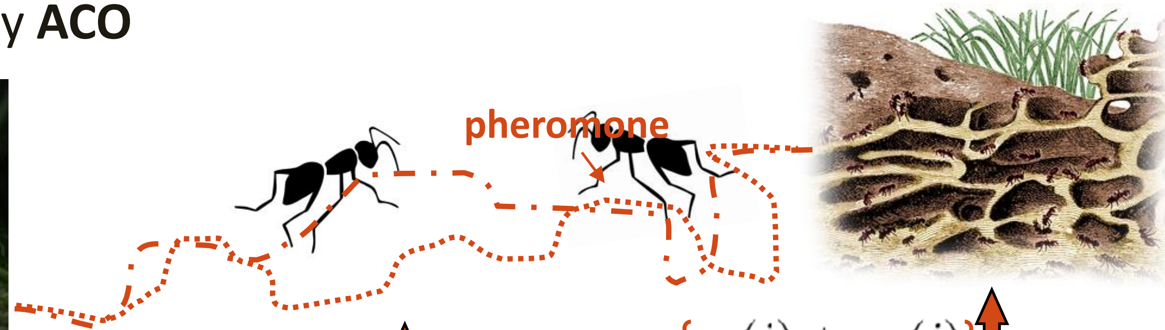
How will the predictive model generalize to an independent set?

1) The original sample is randomly partitioned into *k* equal sized subsamples.
2) One of the k subsamples: **validation data** for testing the model.
3) The remaining k - 1 subsamples: **training data**.
4) Repeat k times (k folds), with each of the k subsamples used exactly once as the **validation data**.
5) The k results from the folds can then be averaged to produce a single estimation.

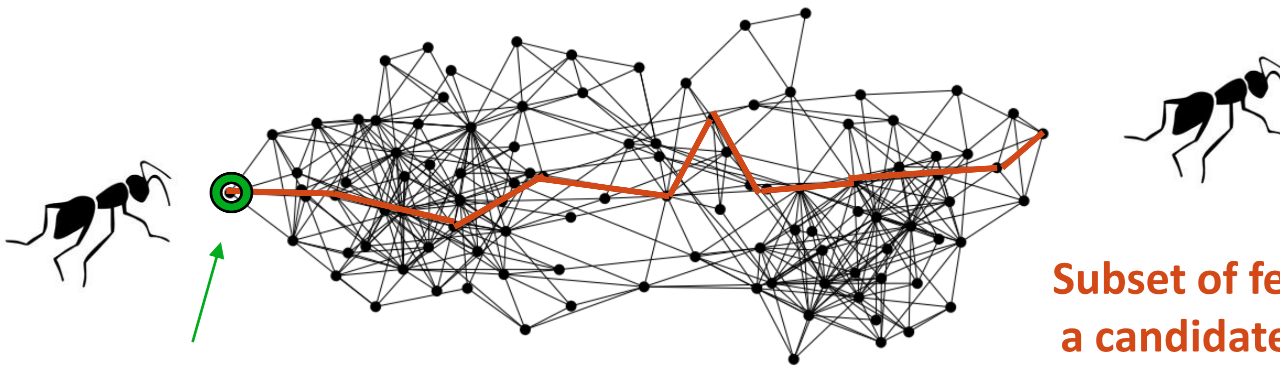


# INGREDIENT #2: Ant Colony Optimization (ACO)

- Feature selection by **ACO**



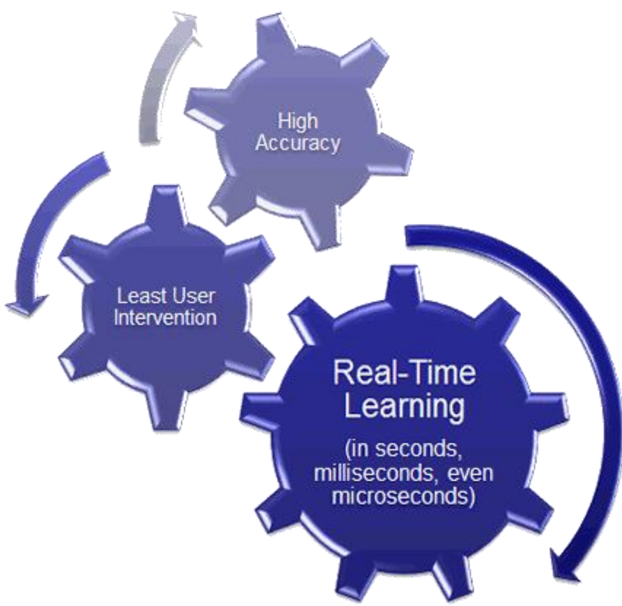

$$p^{a}_{uw}(i) = \frac{[\tau_u(i) + \tau_w(i)]}{\sum_{w' \in \mathcal{U}/u,w} \tau_{uw'}(i)}$$



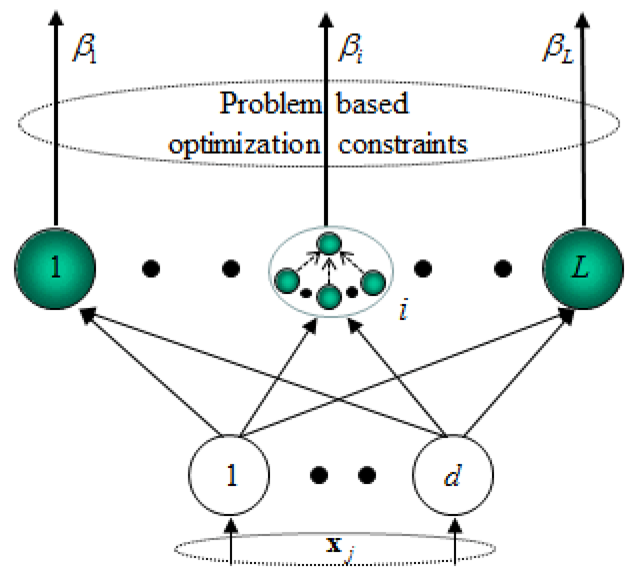

- Movement guided probabilistically by the **pheromone** deployed
- Probability that the ants led to this node

# INGREDIENT #3: Extreme Learning Machines (ELM)

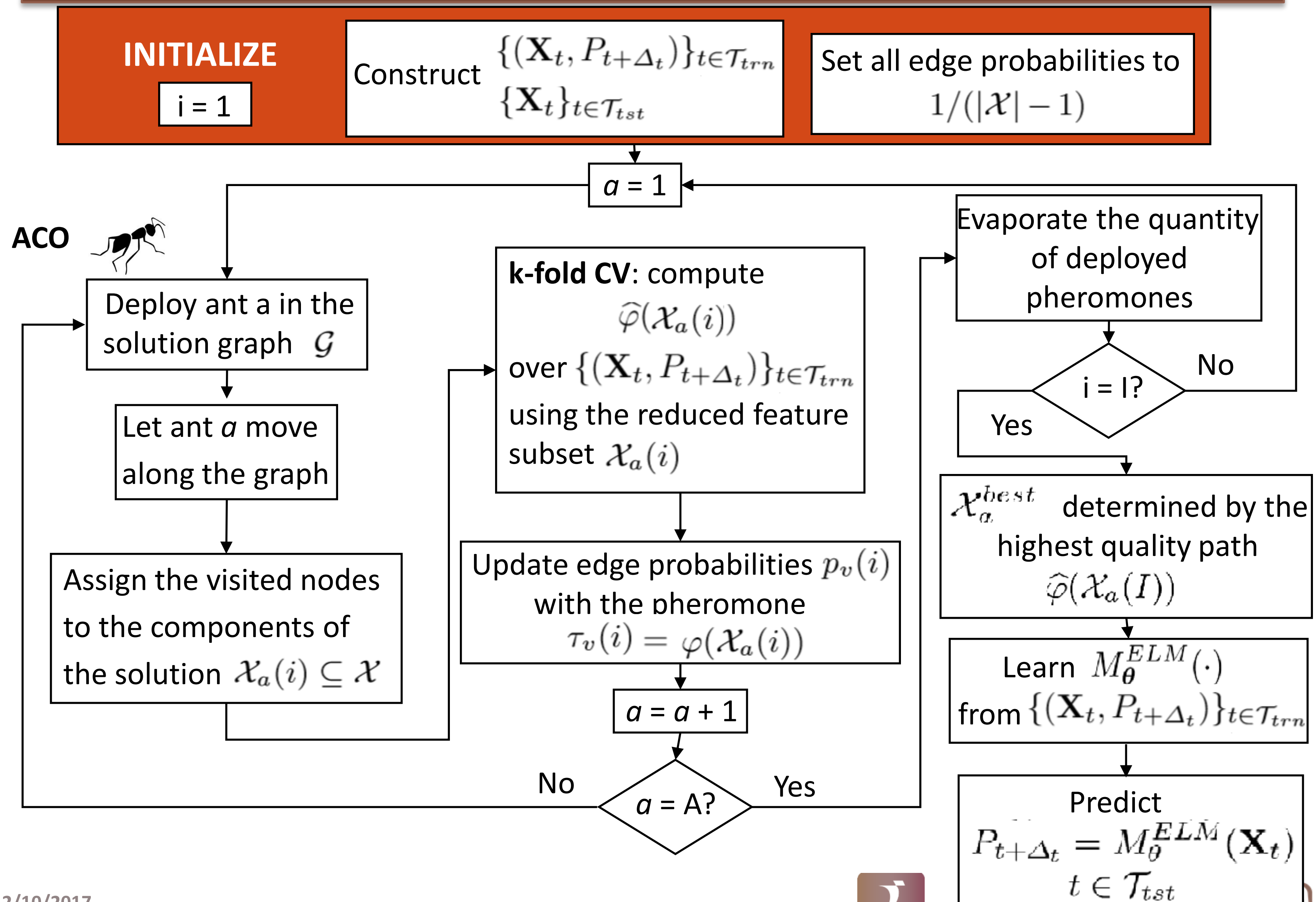
- Low complexity variant of neural networks characterized by a fast training and learning process




- Most significant characteristic:
  - Can be carried out by randomly setting the weights of the underlying neural network, and then taking the pseudo-inverse of the hidden-layer output matrix.
    - Clustering
    - Regression
    - Classification



# The Recipe: Proposed Algorithm






tecnalia
Universidad del País Vasco
Euskal Herriko Unibertsitatea
Campus de Excelencia Internacional
basque center for applied mathematics

# EXPERIMENTAL RESULTS AND DISCUSSION

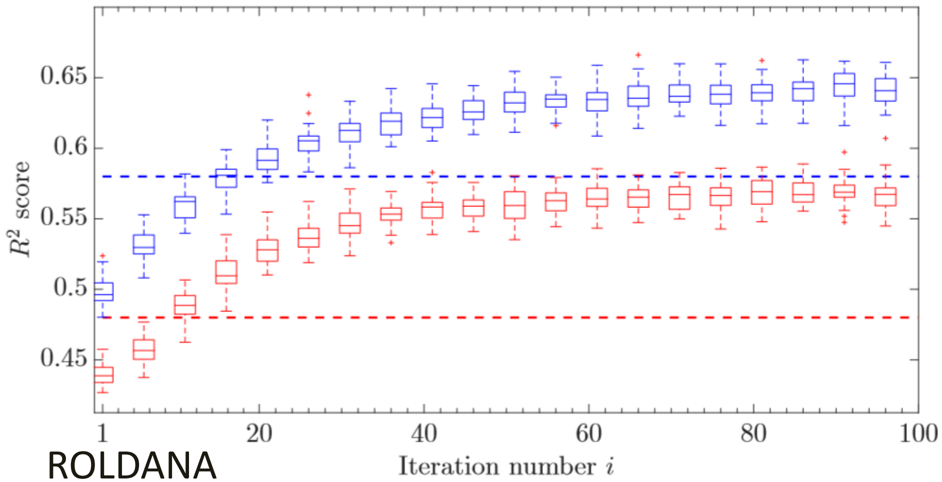
- ROLDANA parameters
  - M = 22 turbines
  - Total nominal power: 36,740 KW
- DATA
  - From January 2013 to October 2015
  - Time step between wind power measurements: 1h
- Variables (NWP model):
  - Temperature (V1)
  - Wind module (V2)
  - Wind U/V components (V3 and V4)
- Rectangular grid of P = 45 points

- FARO FARELO parameters
  - M = 18 turbines
  - Total nominal power: 30,060 KW
- Used values
  - Δt = 1 time steps
  - ΔX = 2 past values of every feature
  - T45*5*2 = 450 possible features/scenario
  - A = 100 ants
  - ELM with 200 hidden neurons
  - I = 100 iterations
  - Predictive performance: coefficient of determination or $R^2$ score.

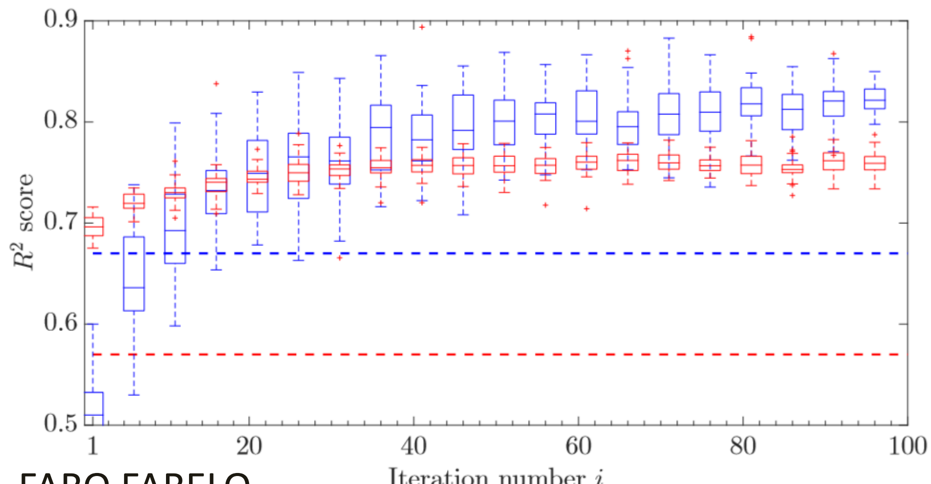
- 20 independent experiments per every simulated scenario have been run in order to account for the ACO algorithm to be stochastic.

# EXPERIMENTAL RESULTS AND DISCUSSION

- Does the prediction perform better when using ACO wrapper?




ROLDANA




FARO FARELO

**$R^2$ convergence plots of the proposed ACO-ELM model**

Horizontal bold dashed lines correspond to the $R^2$ score when no feature selection is made

**Convergence curves for an alternative ACO-Nearest Neighbors model**

In both simulated scenarios the feature selection process provides a predictive gain with respect to the case when no feature selection is made

(in the order of 0.1 in the $R^2$ scale)

# Conclusions and Future Research Lines

- Wind power production forecasting has been tackled by a hybrid predictive model that combines ELMs and ACOs
  - Main design principle: possible input features to the model = nodes of a solution graph
  - This is efficiently explored by using ant colonies guided by a fitness equal to the cross-validated prediction score
  - ELM: light optimization procedure of the overall model due to the renowned low complexity training process of this particular class of supervised learners
  - Validated with real data recorded in two different wind farms located in Spain characterized by very distinct wind patterns
  - The performance enhancement obtained is promising ($R^2$ increases of up to 0.1)
- Future research lines:

1) Accelerate the convergence properties of the ACO wrapper by adding heuristic information to the pheromone calculation
2) How to reflect the collinearity between nodes u and v in the pheromone calculation expression
   - So as to avoid transitions between nodes (features) when they are strongly correlated to each other.
3) Other swarm heuristics will be also under active investigation as alternative feature selection wrappers
4) And many-many other ideas to come!

# THANK YOU!

## WIND POWER PRODUCTION FORECASTING USING ANT COLONY OPTIMIZATION AND EXTREME LEARNING MACHINES

***MARIA CARRILLO BARRENETXEA***
***mcarrillo007@ikasle.ehu.eus***

*This work has been co-funded by the following research projects: EphemeCH (TIN2014-56494-C4-4-P) from the Spanish Ministry of Economy and Competitivity, under the European Regional Development Fund FEDER, and by the Basque Government under its ELKARTEK research program (BID3ABI project).*