Artificial Hydrocarbon Networks: Chemical Nature Inspiration in Machine Learning

Hiram Ponce

Universidad Panamericana. Facultad de Ingeniería.
Augusto Rodin 498, Ciudad de México, 03920, México.
hponce@up.edu.mx

UNIVERSIDAD PANAMERICANA

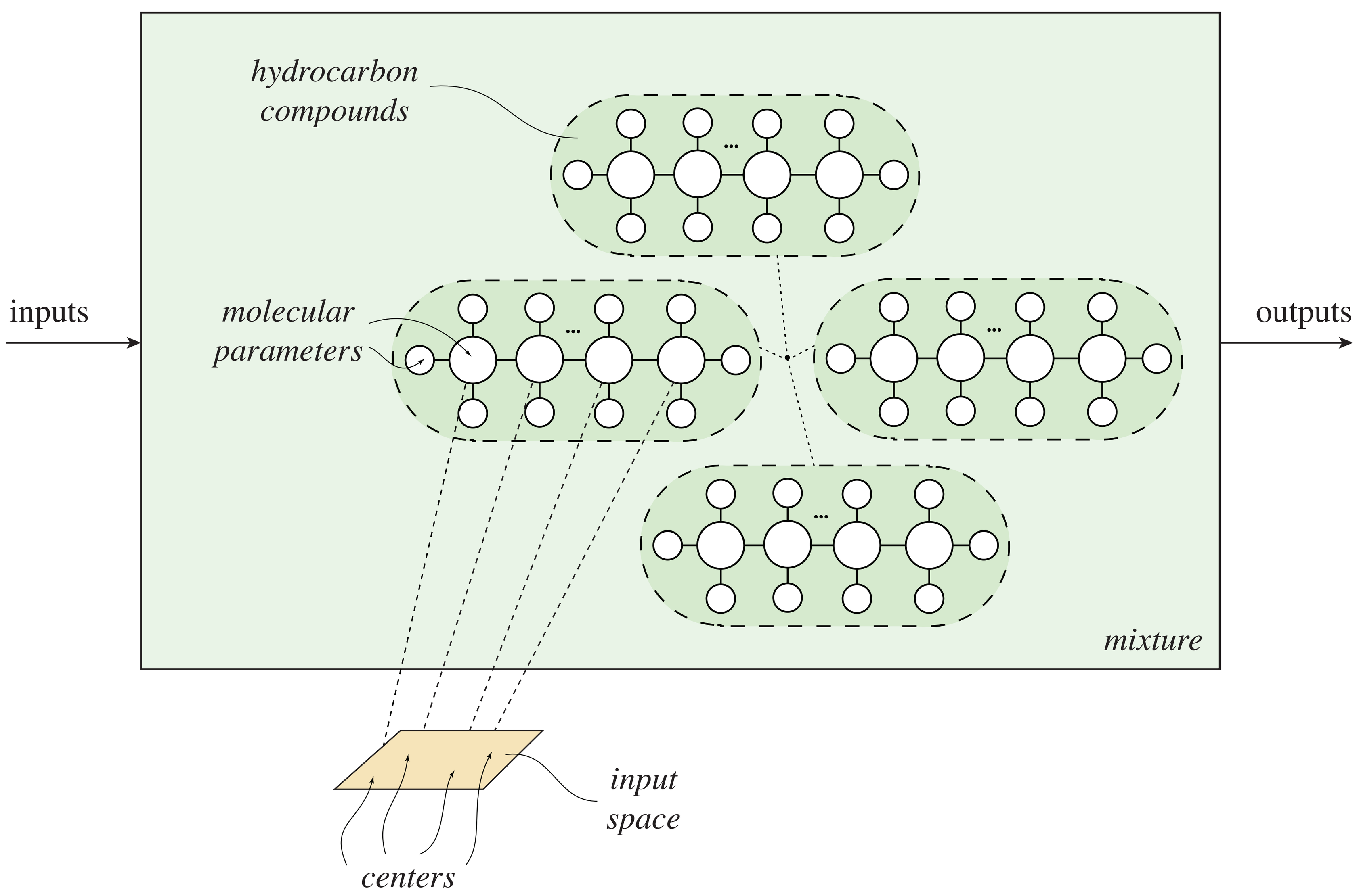
BACKGROUND AND SCOPE

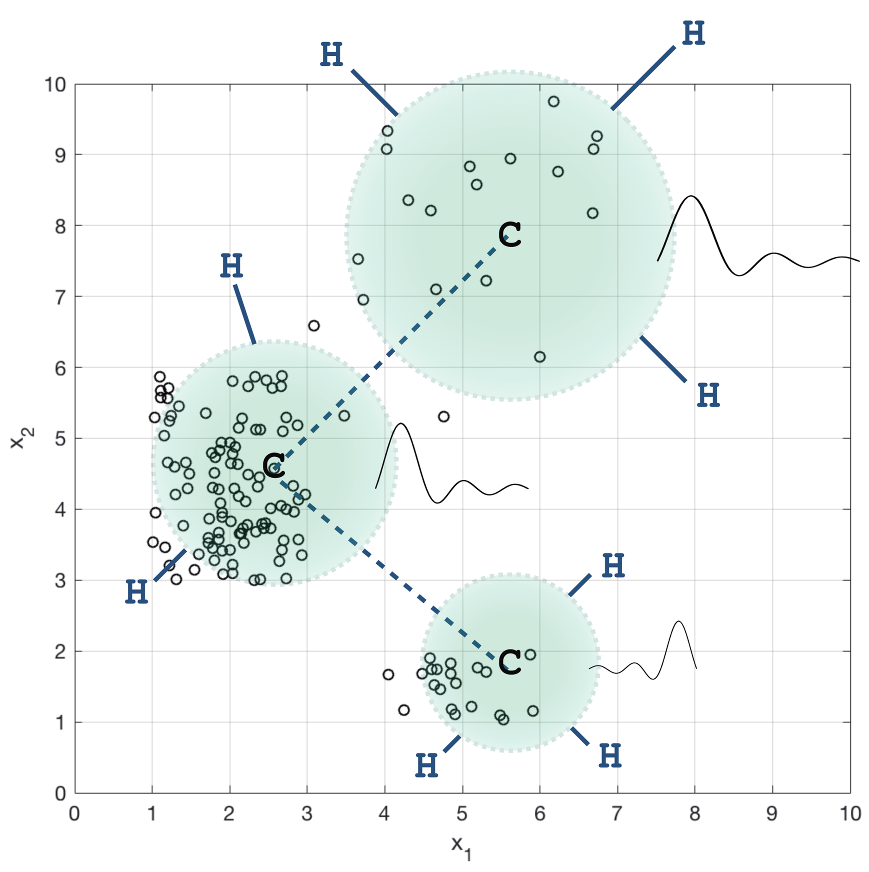
Recently, artificial hydrocarbon networks (AHN), a supervised learning method inspired in the inner structures and mechanisms of chemical compounds, have been proposed as a data-driven approach. AHN have proved to be efficient in predictive power when modeling a data-based problem. However, it stills require more studies on its challenges, issues and applications.

This work aims to discuss challenges and trends of AHN as a data-driven method. It also lays the foundations on AHN for implementing new training algorithms and the way to reveal the chemical nature of data-driven problems.

KEY CONCEPTS OF AHN

The goal of the AHN method is to package information, from a set of instances, in basic units known as molecules. Then, create compounds (nonlinear) and mixtures (linear).





The inspiration in organic compounds to develop a ML method considers three facts observed from nature:

– *Stability*
– *Organization*
– *Multi-functionality*

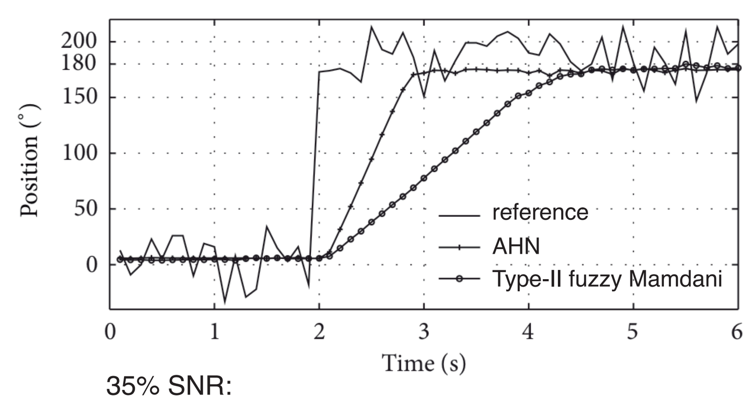
APPLICATIONS & HIGHLIGHTS OF AHN

– Key features in supervised ML.

Features	AHN	DT	NN	NB	kNN	SVM
Accuracy	✓✓✓	✓✓	✓✓✓	✓	✓✓	✓✓✓
Tolerance (missing values)	✓	✓✓✓	✓	✓✓✓	✓	✓✓
Tolerance (noise)	✓✓✓	✓✓	✓✓	✓✓✓	✓	✓✓
Variety of attributes	✓✓	✓✓✓	✓✓	✓✓	✓✓	✓✓
Regression	✓✓✓	✓	✓✓	✓	✓✓	✓✓
Classification	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓	✓✓✓
Interpretability	✓✓	✓✓✓	✓	✓✓✓	✓✓	✓

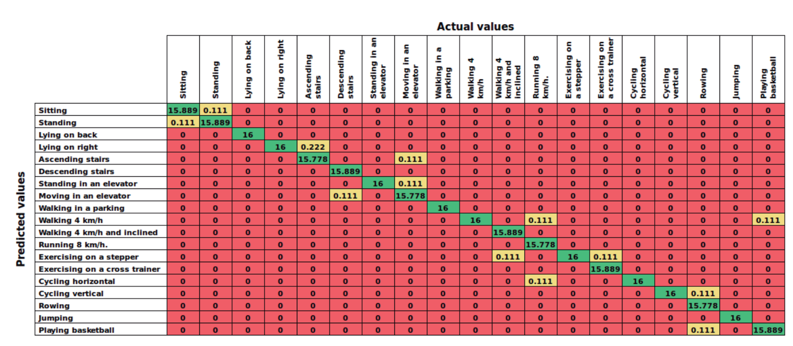
Symbols: low (✓), medium (✓✓) and high (✓✓✓) satisfaction

– Some applications and case studies.



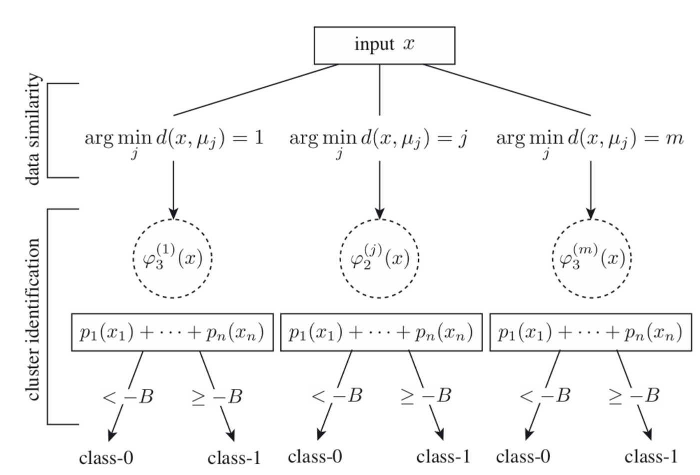
35% SNR:
AHN rise time: 1.0 s, steady-state error: 5.0 %
Type-II rise time: 2.6 s, steady-state error: 5.5 %

Position Control systems in DC motors



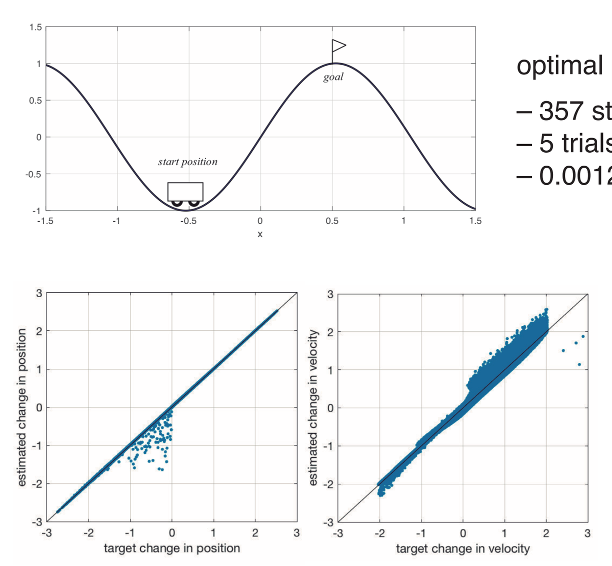
AHN accuracy 99.49(0.44) %
deep-RNN accuracy 99.27(0.16) %
statistical equivalent ($p = 0.109$); AHN 6.98x less parameters

Human activity recognition



if $d(x,0) < 2.63$ **and** $g_1(x) < 0.57$, **then** sample = *benign*, **or**
if $d(x,0) < 2.63$ **and** $g_1(x) \geq 0.57$, **then** sample = *malignant*, **or**
if $d(x,0) \geq 2.63$ **and** $g_2(x) < 1.12$, **then** sample = *benign*, **or**
if $d(x,0) \geq 2.63$ **and** $g_2(x) \geq 1.12$, **then** sample = *malignant*

Interpretability of AHN-based classifier



optimal policy reached on:
– 357 steps
– 5 trials
– 0.0012 RMSE

Reinforcement learning from scratch

TRENDS AND ISSUES OF AHN

– New training algorithms
– Big data processing
– Kernels and relations in molecules
– Hybrid approaches with AHN
– Transfer learning

[1] Ponce, H., Ponce, P., Molina, A. (2013), Artificial Hydrocarbon Networks Fuzzy Inference System, *Mathematical Problems in Engineering* 2013: 531031.
[2] Ponce, H., Miralles, L., Martínez, L., (2016), A Flexible Approach for Human Activity Recognition Using Artificial Hydrocarbon Networks, *Sensors* 16(11): 1715.
[3] Ponce, H., Martínez, L. (2017), Interpretability of Artificial Hydrocarbon Networks for Breast Cancer Classification, *IJCNN*, Anchorage, AK, 3535–2542.
[4] Ponce, H., González, G., Martínez, L., (2018), A Reinforcement Learning Method for Continuous Domains Using Artificial Hydrocarbon Networks, *IJCNN*, Rio de Janeiro, Brazil, 1–6.