AI for Good



You, Your Research and Beyond.

Anthony, Faustine

PhD researcher (Ghent-Tecnico Lisboa, Portugal)

sambaiga sambaiga@gmail.com

sambaiga.github.io/sambaiga/





Research Area

- Machine learning and signal processing for smart-energy systems with a focus on Non-Intrusive Load Monitoring (NILM), Community Energy and Smart Grid.
- Artificial Intelligence for Earth Observations(EO) data with a focus on sustainable energy, agriculture and natural resources management.
- Artificial Intelligence for Industrial 4.0 with focus on preventive maintenance, machine vision and streaming analytics and MLOPs.
- Robust deep learning for industrial applications with a focus on a learning algorithm that allows uncertainty quantification, is robust to adversarial attacks and allows self-supervised or semi-supervised learning.
- Data management with a focus on specification, strategies, best practises and guideline for machine learning training dataset (TDs) development.

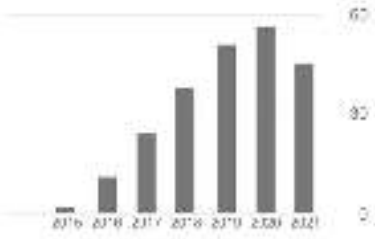
Publications



Senior Industrial Analytics Researcher, Irish Manufacturing Research
Verified email at lmrc.ie
Machine learning Signal processing Deep learning

TITLE	CITED BY	YEAR
Watt's up at Home? Smart Meter Data Analytics from a Consumer-Centric Perspective B Völker, A Beinhauer, A Faustine, L Pereira Energies 14 (8), 719		2021
UNet-NILM: A Deep Neural Network for Multi-tasks Appliances State Detection and Power Estimation in NILM A Faustine, L Pereira, H Bouchiba, S Kulkarni Proceedings of the 5th International Workshop on Non-Intrusive Load ...	7	2020
Adaptive weighted recurrence graphs for appliance recognition in non-intrusive load monitoring A Faustine, L Pereira, C Klemenjak IEEE Transactions on Smart Grid 12 (1), 398-406	6	2020
Leveraging Machine Learning for Sustainable and Self-sufficient Energy Communities A Faustine, L Pereira, D Ngendya, L Bandiziwa		2020
Multi-Label Learning for Appliance Recognition in NILM Using Fryze-Current Decomposition and Convolutional Neural Network A Faustine, L Pereira Energies 13 (16), 4154	6	2020
Improved appliance classification in non-intrusive load monitoring using weighted recurrence graph and convolutional neural networks A Faustine, L Pereira Energies 13 (13), 3374	10	2020
On metrics to assess the transferability of machine learning models in non-intrusive load monitoring C Klemenjak, A Faustine, S Makonin, W Elmenreich arXiv preprint arXiv:1912.06200	3	2019
A survey on non-intrusive load monitoring methodies and techniques for energy disaggregation problem A Faustine, NH Mvungi, S Kaijage, K Mkocha arXiv preprint arXiv:1703.00785	100	2017
Open source cellular technologies for cost effective cellular connectivity in rural areas F Anthony, M Gabriel, B Shao International Journal of Computer Applications 148 (15)	9	2016
Wireless sensor networks for water quality monitoring and control within lake victoria basin: prototype development A Faustine, AN Mvuma, HJ Mongi, MC Gabriel, AJ Tenge, SB Kucel	70	2014

	All	Since 2016
Citations	226	226
h-index	6	6
i10-index	4	4



Public access VIEW ALL

0 articles	3 articles
not available	available

Based on funding mandates

Co-authors EDIT

- Lucas Pereira — IT, LARSyS, Instituto Superior T...
- Prof. Nerey mvungi — Professor of Computer Systems...
- Hector Mongi — The University of Dodoma
- Christoph Klemenjak — Research Associate @ University...

Technical Report

- A State-of-the-Art and Literature Review of Machine Learning for Green Energy, *CeADAR, University College Dublin, January 2021*
- AI-ready Earth Observation Training Dataset (AIREO) Best Practice Guidelines, *CeADAR, Irish Centre for High-End Computing (ICHEC) and European Space Agency March 2021*
- AI-ready Earth Observation Training Dataset (AIREO) State-of-the-Art Review Report. *CeADAR, Irish Centre for High-End Computing (ICHEC) and European Space Agency; November 2020*
- A State-of-the-Art and Literature Review of Artificial Intelligence for Earth Observation Data, *CeADAR, University College Dublin, May 2020*
- Artificial Intelligence for Earth Observation Final Report.*CeADAR, University College Dublin, September 2020*
- Machine learning solution for Event Detection in a Smart-home environment using data from sensors. *IDLab, Ghent University, March 2018*

Talks

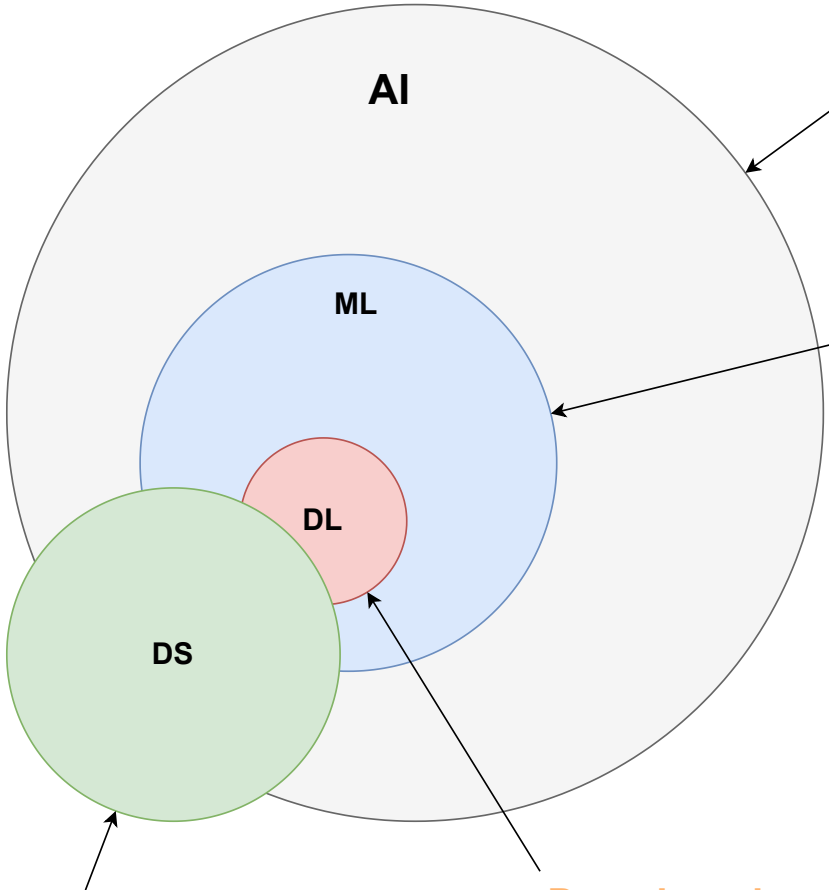
Conference Presentation

- **Leveraging Machine learning for Sustainable and Self-sufficient Energy Communities**
  A. Faustine, L. Pereira, D. Ngondya, L. Benabbou
  NeurIPS 2020 Workshop Tackling Climate Change with Machine Learning (2021) slides link video
- **UNet-NILM A Deep Neural Network for Multi-tasks Appliances State Detection and Power Estimation in NILM**
  A. Faustine, L. Pereira H, Bousbiat and S. Kulkarni
  5th International Workshop on Non-Intrusive Load Monitoring (co-located with ACM BuildSys 2020 and jointly organized with the EU NILM Workshop) (2020) slides link video
- **Improved Appliance Classification in NILM using Recurrence Plots and Convolutional Neural Networks**
  A. Faustine, D .Deschrijver and T. Dhaene
  EU NILM Workshop, Thessaloniki, Greece (2019) slides link video
- **Convolutional Neural Network for Appliance Classification in NILM**
  A. Faustine
  Data Science Africa, Arusha, Tanzania (2017) slides link video

Invited Talks and Workshop

- **Deep Neural Networks (DNN) with Energy-Based Learning**
  A. Faustine
  CeADAR Technical Talk (2020) slides link video
- **Machine Learning for Energy Disaggregation**
  A. Faustine
  CeADAR Technical Talk (2020) slides link video

AI defined



Artificial Intelligence

- the science and engineering of making intelligent machines.
- machines that can learn, at least somewhat as human beings do.

Machine Learning

- A subset of AI studying how computer agents can improve their perception, knowledge, thinking, or actions based on experience or data.

Supervised-ML

- A computer learns from input-output pairs in order to predict the output of the new input.

Unsupervised-ML

- Discover the pattern and learn the structure of unlabeled data.

Reinforcement-ML

- Agent learn action sequences that optimize its total rewards

Data science

- Understanding or making sense of data e.g visualization, exploratory analysis, data products etc

Deep learning

- A subset of ML that use cascade multi-layer (artificial) neural networks of non-linear preprocessing for feature learning and pattern recognition.

Figure 1: AI defined

AI-research area

- Self-supervised learning.
- Reinforcement learning.
- Beyond prediction and classification $\Rightarrow$ ML-driven optimisation etc
- Robust and trusthworth AI algorithms (biasness, fairness, anomaly detection, out of distribution detection etc
- Streaming analytics $\Rightarrow$ incremental or online learning, drift detection etc
- Data-centric AI
- AI for Climate Change and SDG.

Where can you contribute

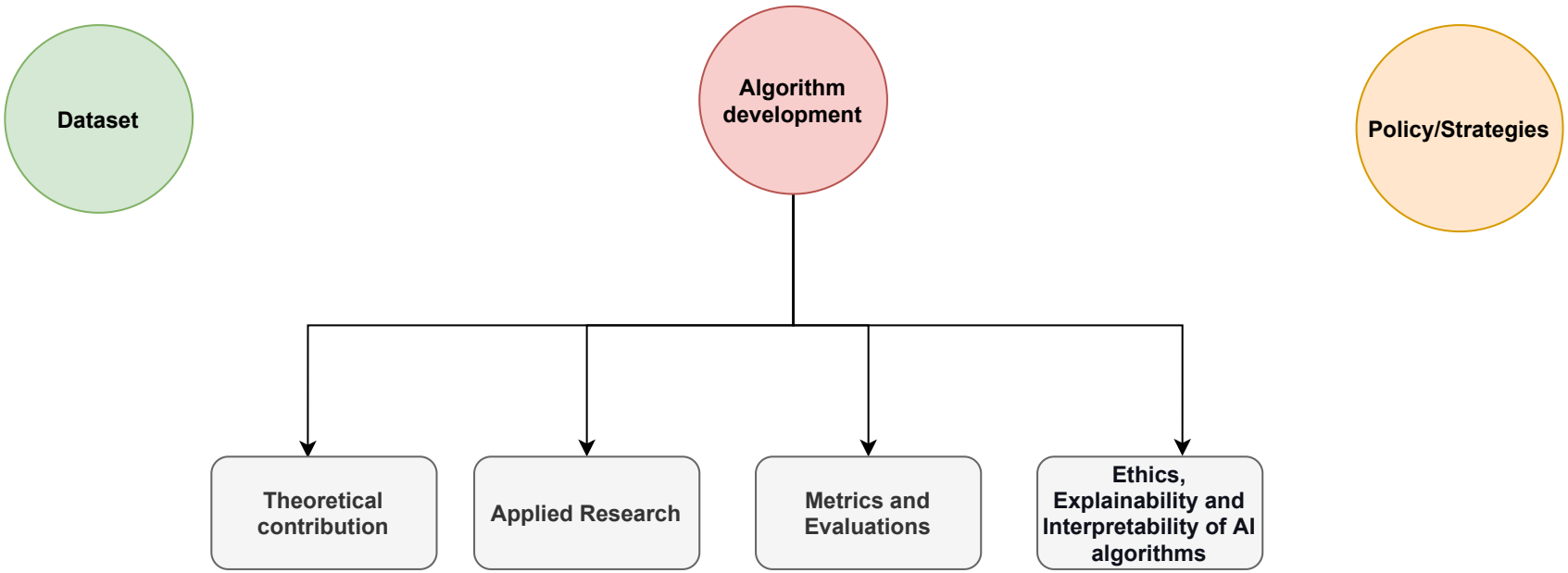
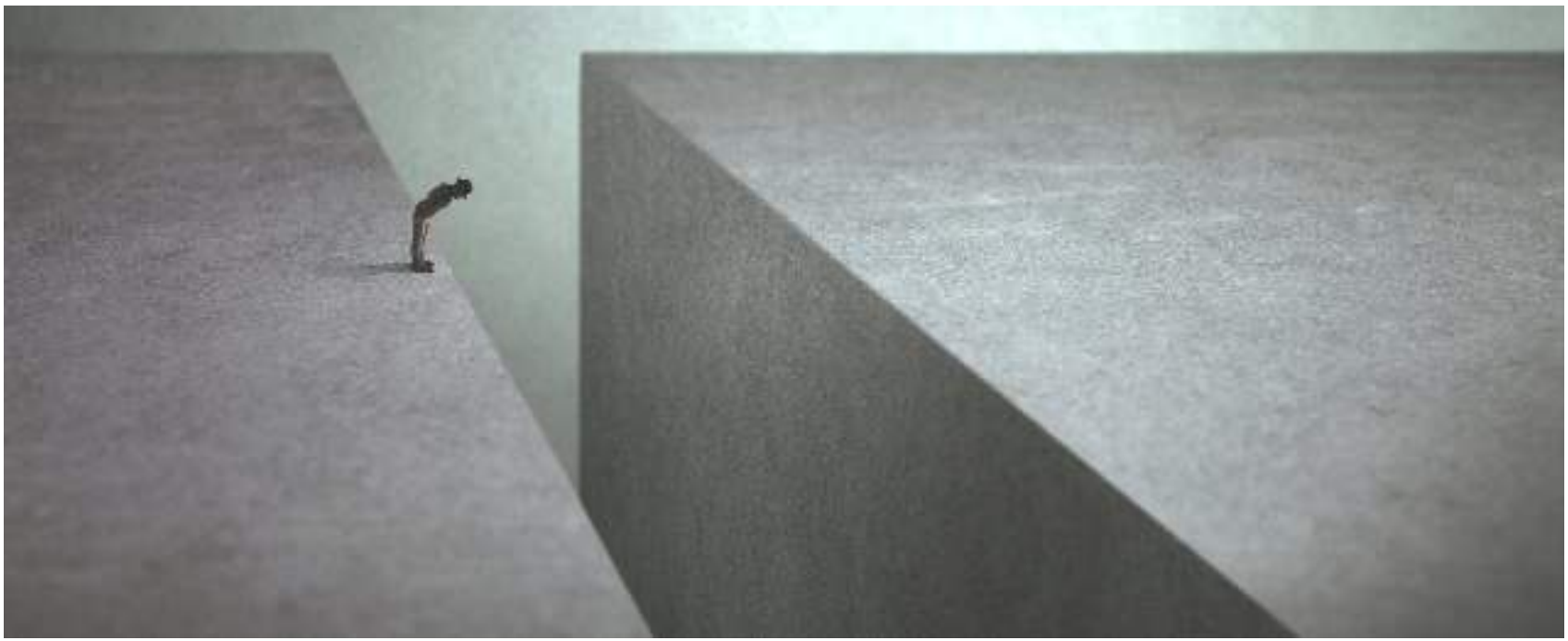


Figure 2: Contributions

The Gap



Read, Connect, Reflect

- Critically read several papers.
- Take note of every-paper you read.
- Reflect on what research is useful and impactful.
- Look beyond your immediate horizon.
- Seek advice from experienced researchers.
- Collaborate with other people working on similar topics

```
Take the time, on a regular basis, to ask the larger questions,
and not stay immersed in the sea of detail where almost everyone
stays almost all of the time
```

Choosing research problem to work on

Idea-Driven vs Goal-Driven Research

1. Idea-driven ⇒ As you read a paper showing how to do X, you have an idea of how to do X even better.
2. Goal-driven ⇒ develop a vision of some new ML/AI capabilities you'd like to achieve, and solve problems that bring you closer to that goal.

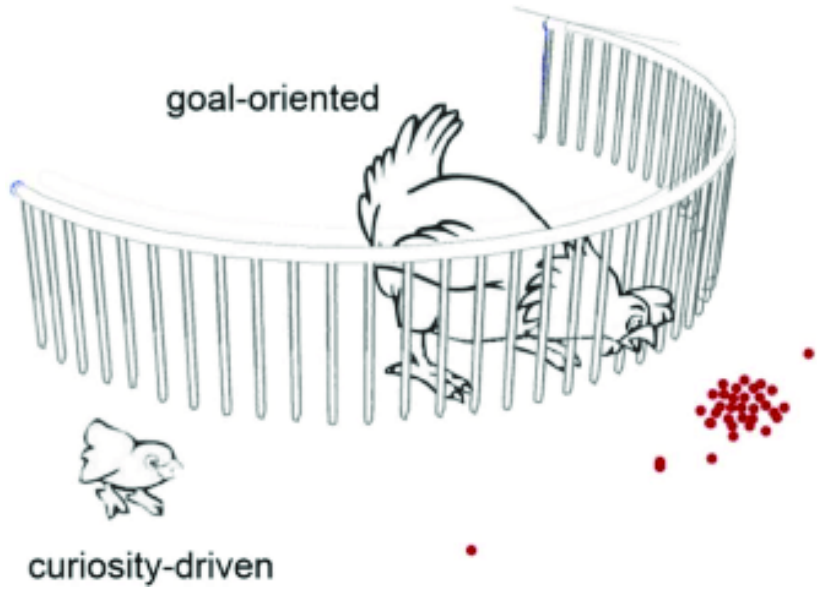


Figure 3: Caption

`Brains are nice to have, but many people who seem not to have great IQs have done great things.`

Develop important research skills

1. Theoretical and practical understanding of your area of research.
2. Learn and understand how to use your microscope
3. Improve your writing, presentation and networking skills.



Figure 4

`Luck favors the prepared mind. You prepare yourself to succeed or not, as you choose, from moment to moment by the way you live your life.`

Climb incrementally towards your research goals



- To develop potential contribution in research requires consistent $\Rightarrow$ concentrate your efforts on a problem for a long period of time.
- Record your daily ideas and experiments.
- Learn when to switch problems.

Change does not mean progress, but progress requires change.

Establish hypothesis about the problem.



- Identify open questions that need answers ⇒ be ambitious.
- Learn about common methods, datasets and libraries.
- Lay out goals & objectives, constraints, and evaluation criteria.

Identify datasets

Identify data-set(s) to benchmark your solution.

At least one dataset that appeared in related prior work.



Where to find datasets

- Build them.
- Scrape them.
- Find them (contact authors).
- Generate them (artificial data).

Take time to understand your data: exploratory analysis.

Establish baseline

Define your baseline:Any publishable performance with simplest approach.







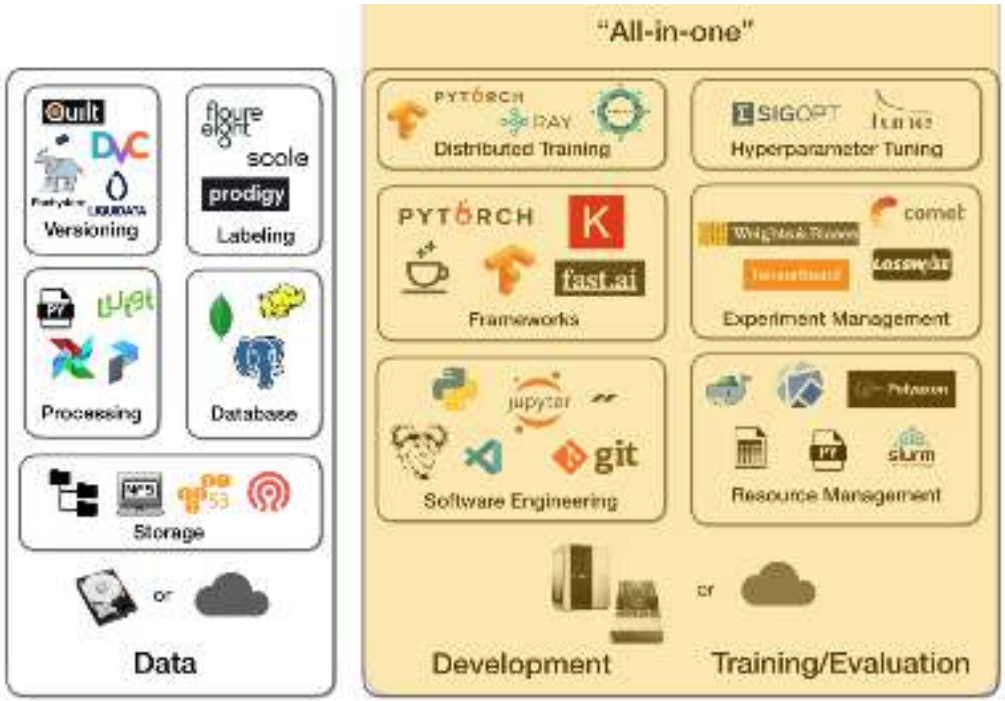
Write code quickly

- data-pipeline.
- Training-evalution-pipeline.
- Analysis-pipeline.

First get a baseline running $\Rightarrow$ this is good research practise.

Run Experiments

Track experiments: Take notes of what each experiment was meant to test.



- Use recommended best practices for managing and monitoring ML experiments.
- Version code, data and ML experiments.
- Use existing frameworks and tools.

Figure 5: credit:Full stack DL

More references: CS329: Versioning & experiment tracking

Experiment Evaluation

When the cook tastes the soup that is formative; when the customer testes that is summative.

Formative evaluation

- They guide further investigations.
- Compare design option A to B, tune hyper-parameters etc.

Summative evaluation

- compare your approach to previous approaches,
- compare different major variants of your approach.

Don't save all your qualitative evaluation for the summative evaluation.

Experiment Evaluation

Follow prior work precisely in how to choose and implement main evaluation metric.



Quantitative evaluation

- Show metric as many variants of your model as you can.
- Test for statistical significance (for highly variable models or small difference performance).
- If your results are not significant. say so and explain what you found.

Experiment Evaluation

Conduct a thorough analysis



Qualitative-evaluation

- convince reader for your hypothesis ⇒ look to prior work to get started.
- Show examples of system output.
- Plot how your model performance varies with the amount of data.
- Present error analysis.

Your story in the paper

- Identify the core contribution ⇒ main story (selling point) of your paper.
- Lay out the structure of your paper ⇒ ensure good flow of ideas.
  1. X (+define X if not obvious) is an important problem
  2. The core challenges are this and that.
  3. Previous work on X has addressed these with Y but the problems with this are Z.
  4. In this work we do W.
  5. This has the following appealing properties and our experiments show this and that.
- The language ⇒ use formal language acceptable in your domain.
- Get feedback prior to submission deadline

Define your life, career and research goal

Pick the goals you believe are worth striving for.

- Why are you doing that PhD?
- What next after your PhD?
- What impact do you want to create?

```
Yes, it is nice to end up where you wanted to be, but the person
you are when you get there is far more important.
```

Networking and Personal branding

- Attend conferences
- Give talks
- Be active on professional social network (e.g Linkedin, google scholar, twitter etc).
- Maintain your personal site

Develop other soft skills beyond your research

- Communication and Personal development skills.
- Grants and Research Proposal Writing.
- Project management and leadership skills.
- Team work and other technical skills in your domain.

Physical and mental health

- The PhD does not define the real you.
- You can still survive without your PhD.

Conclusion

The desire for excellence is an essential feature for doing great work. Without such a goal you will tend to wander like a drunken sailor.

References