Machine Learning and AI in Process Systems Engineering

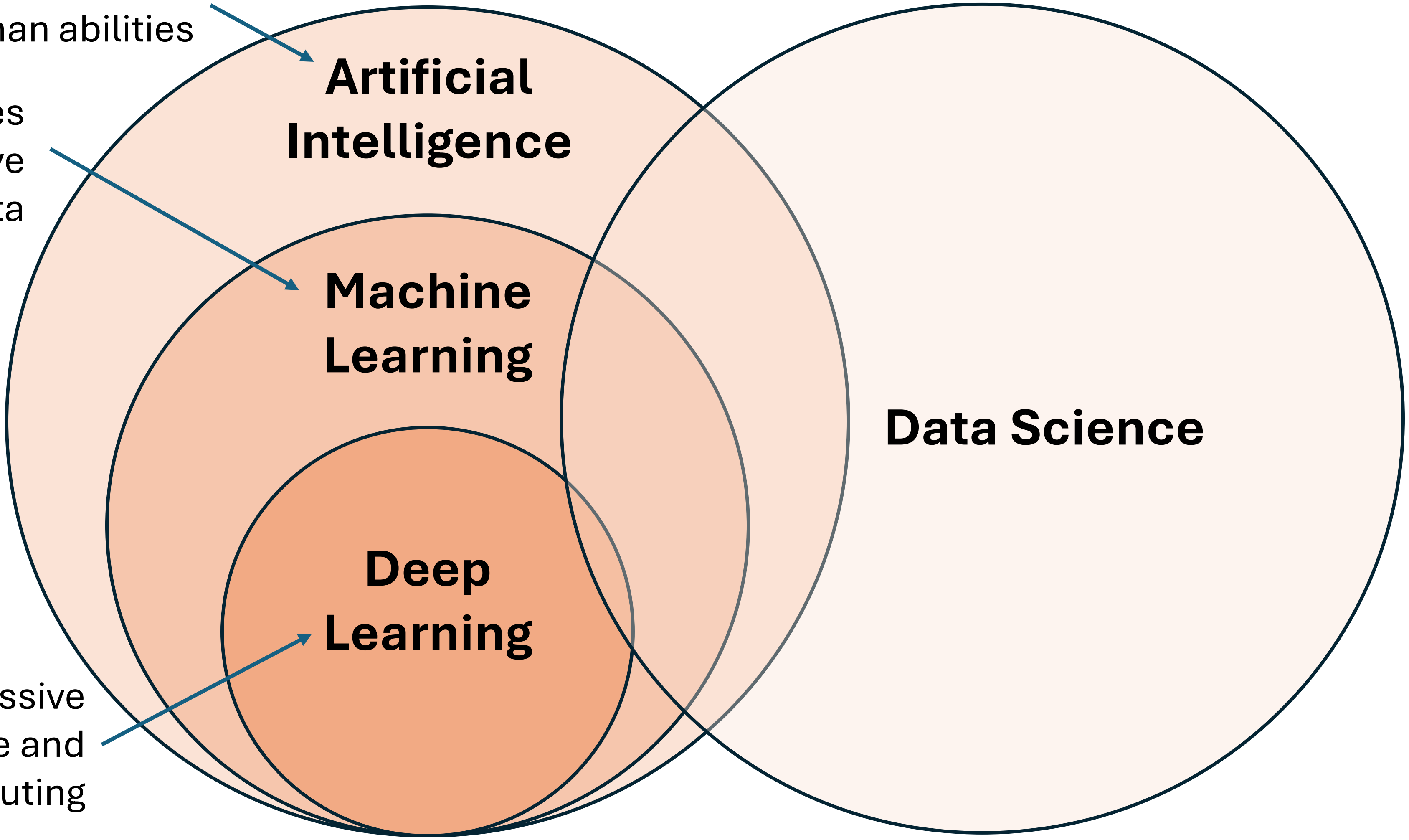
Karl Ezra Pilario, Ph.D.
University Scientist 1
Visiting Professor, *National University of Singapore*
Associate Professor, *University of the Philippines Diliman*

What is AI / ML / DS?

Have machines mimic
human abilities

Have machines
learn and improve
from data

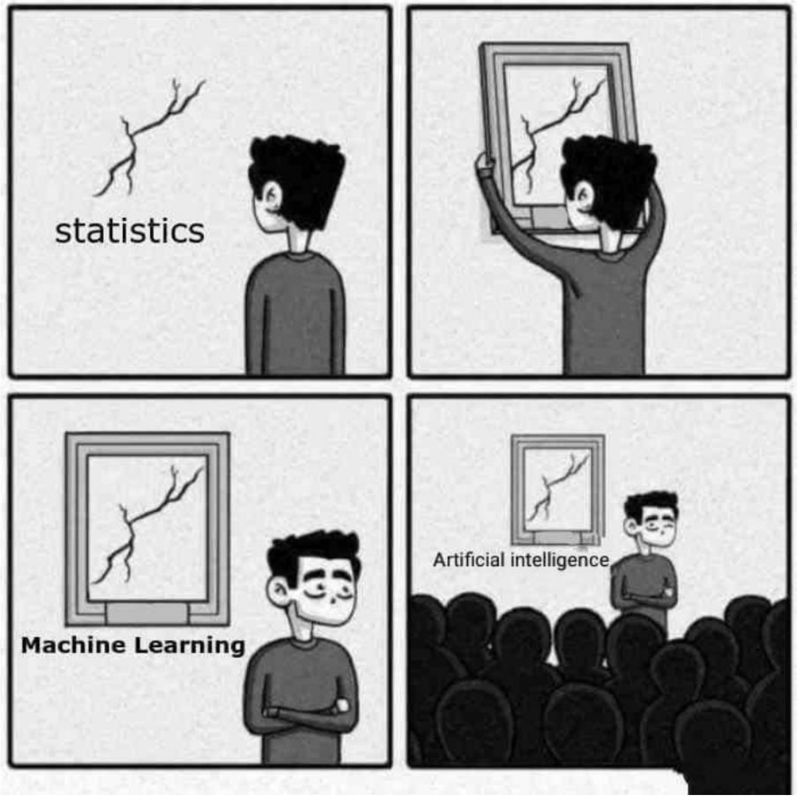
Building
intelligent
systems
and
automation

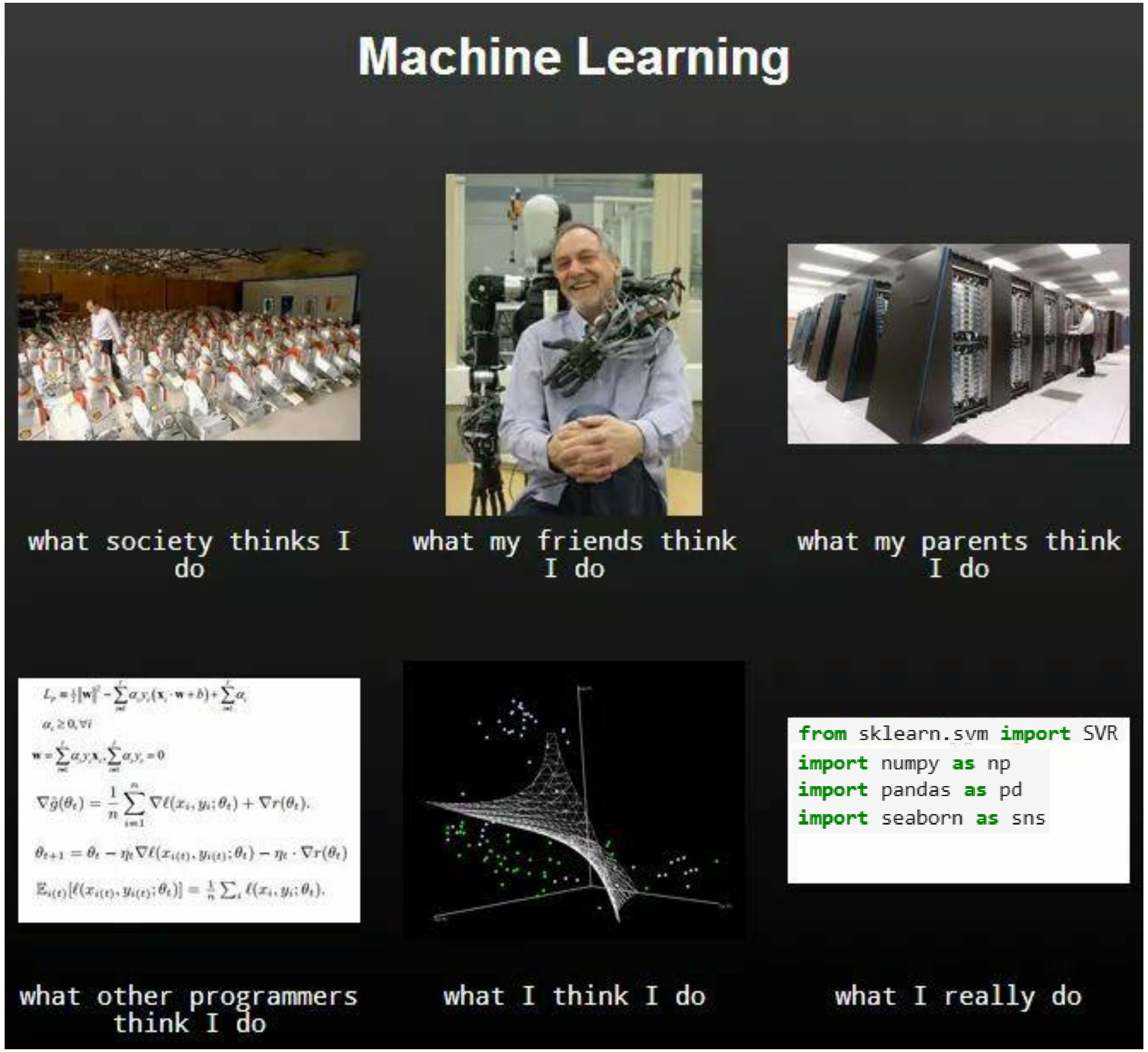


ML but with massive
architecture and
computing

Leveraging
data to gain
insights and
aid human
decision-
making

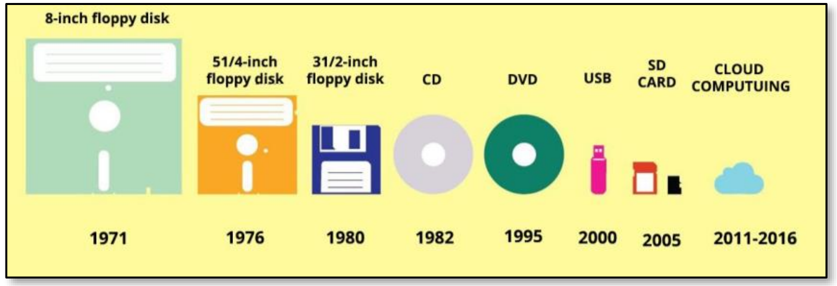
What is AI / ML / DS?

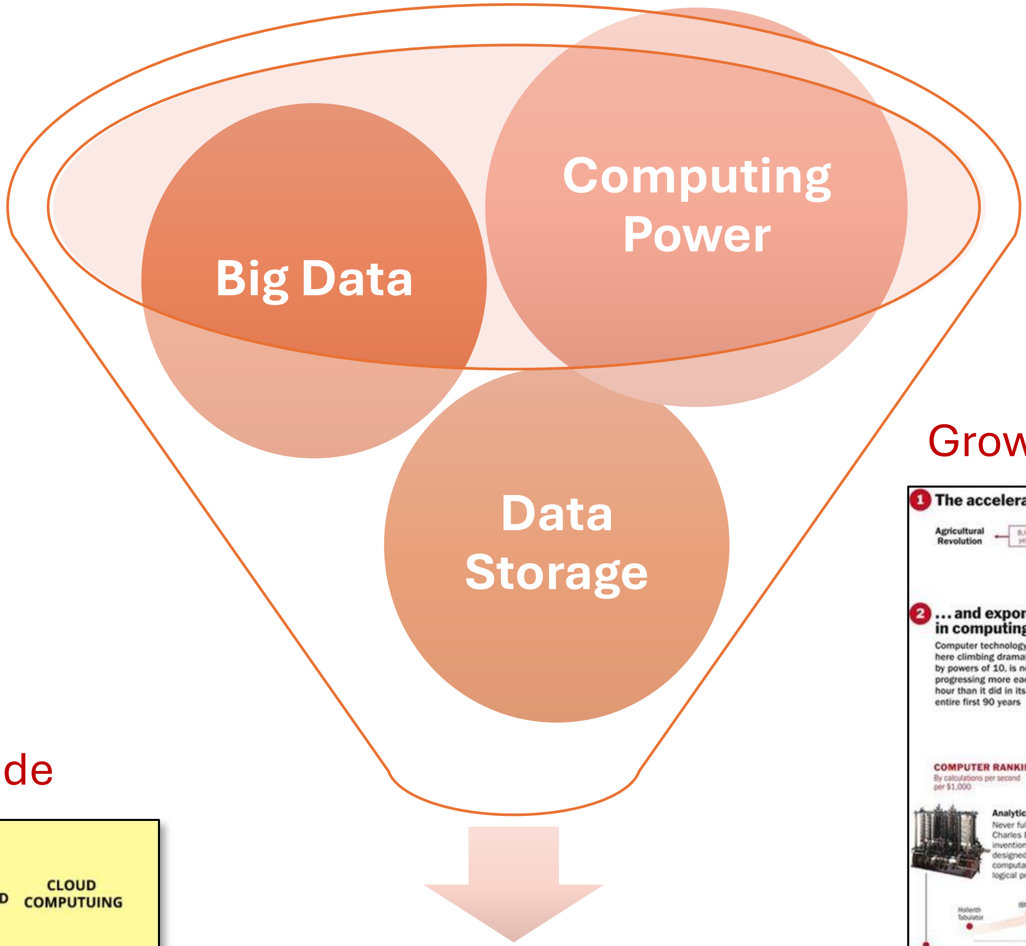




Why did it take off only now?

Growth in **Data Storage** worldwide

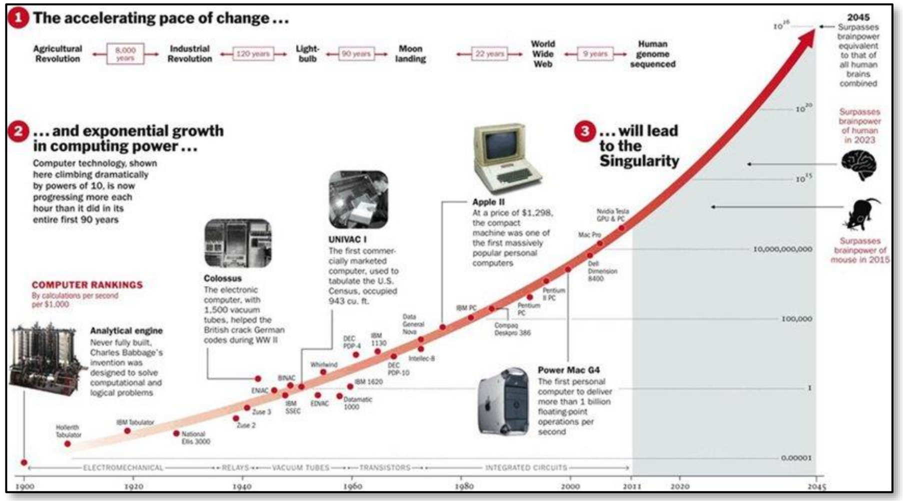




Machine Learning +

Practical Applications

Growth in **Computing Power** worldwide

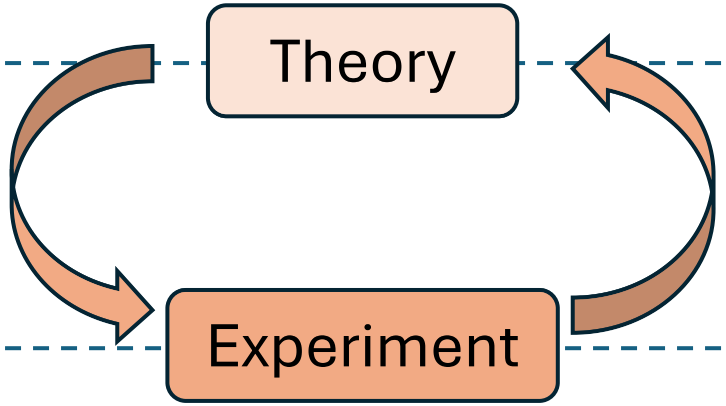


Why use AI / ML in the first place?

Scientific Method

1. Make a hypothesis

4. Revise



2. Test the hypothesis

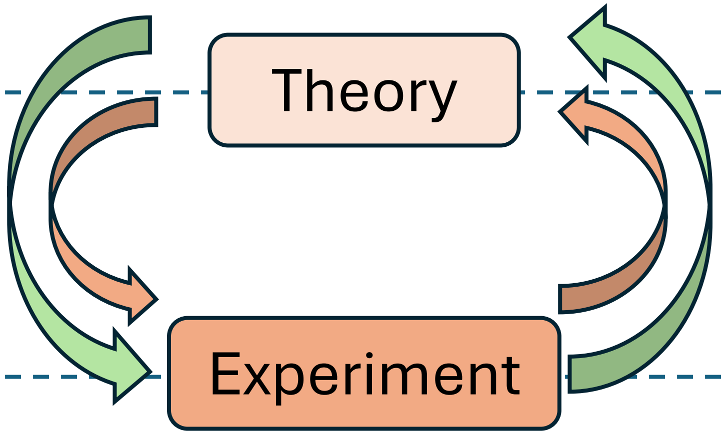
3. Gather data

AI for Scientific Discovery

- *AI is now being used to accelerate discoveries in many fields.*

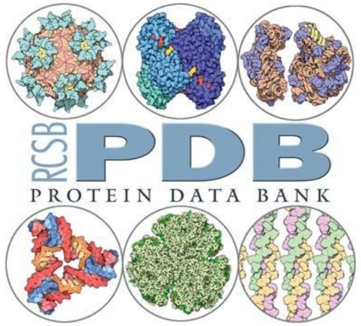
3. Suggest a new model

2. Discover new patterns with AI

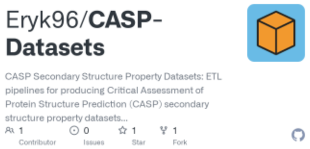


4. Test the models

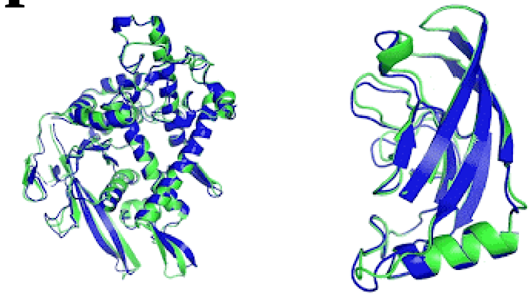
1. Gather data



CASP (Critical Assessment of Protein Structure Prediction)

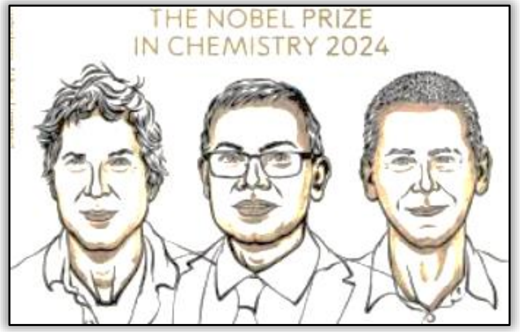


AlphaFold



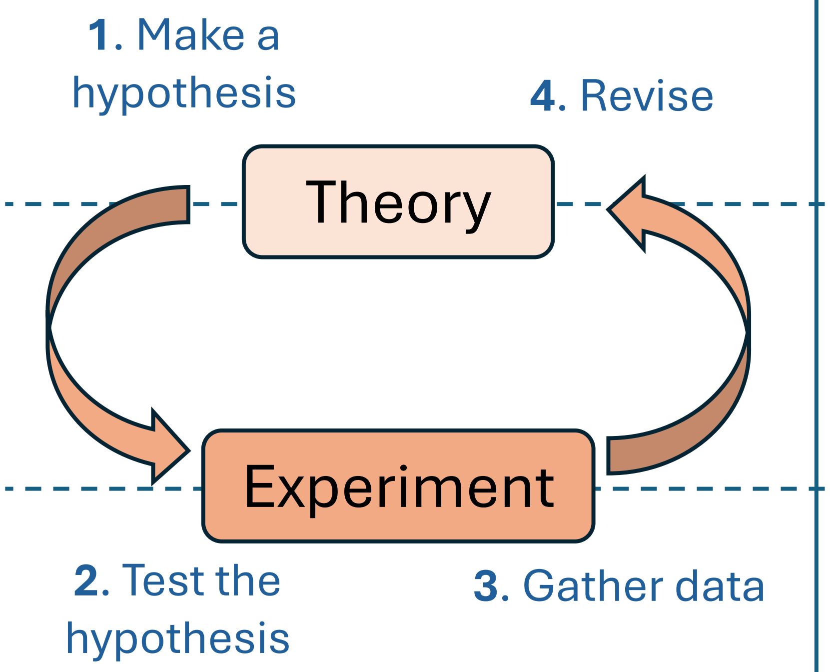
T1037 / 6vr4
90.7 GDT
(RNA polymerase domain)

T1049 / 6y4f
93.3 GDT
(adhesin tip)



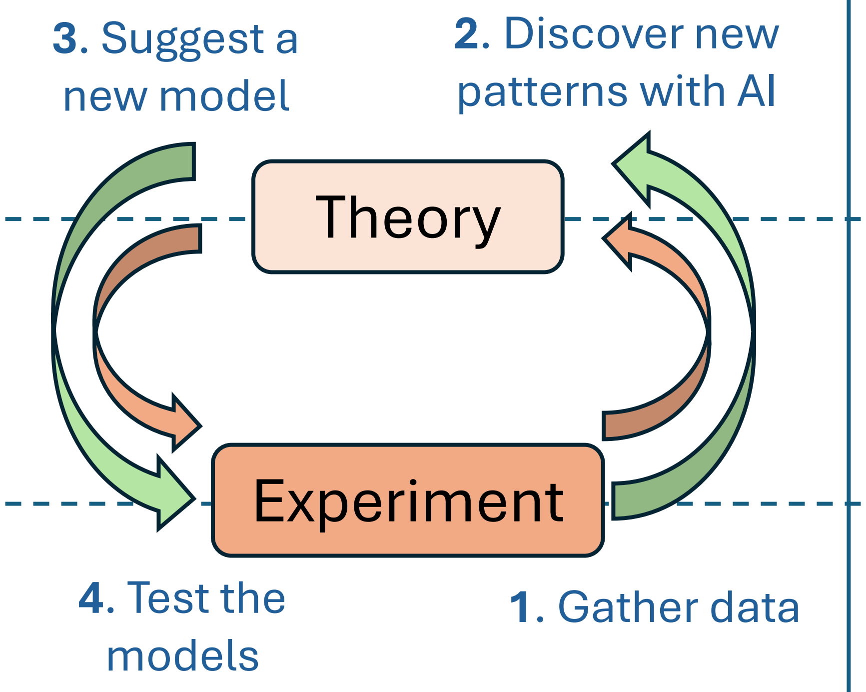
Why use AI / ML in the first place?

Scientific Method



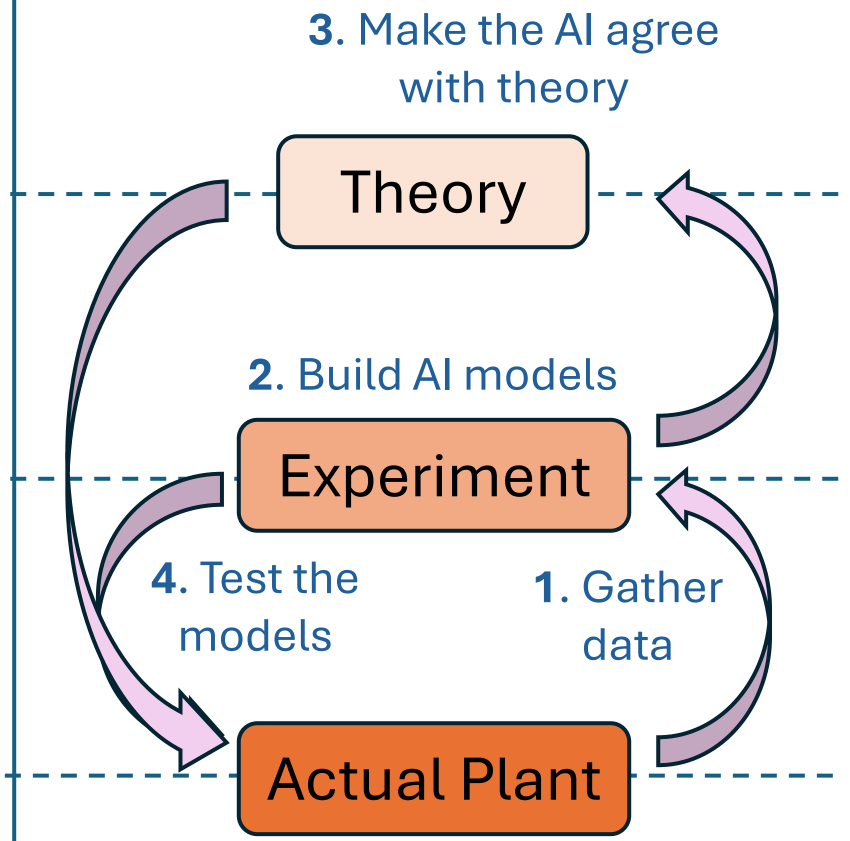
AI for Scientific Discovery

- *AI is now being used to accelerate discoveries in many fields.*



AI for Engineering Practice

- *AI models can be deployed with or without theoretical backing.*



*or any point of application

Why use AI / ML in the first place?

Some points to consider:

1. Working with the first-principles equations of an entire plant is *too difficult*.

- All parameters must be known (rxn const, heat coeff., etc.).
- Theory should relate *‘00s of variables* in *‘00s equations*.
- Computational cost of solving dynamic ODEs / PDEs is high.

2. The real-world is *noisy*.

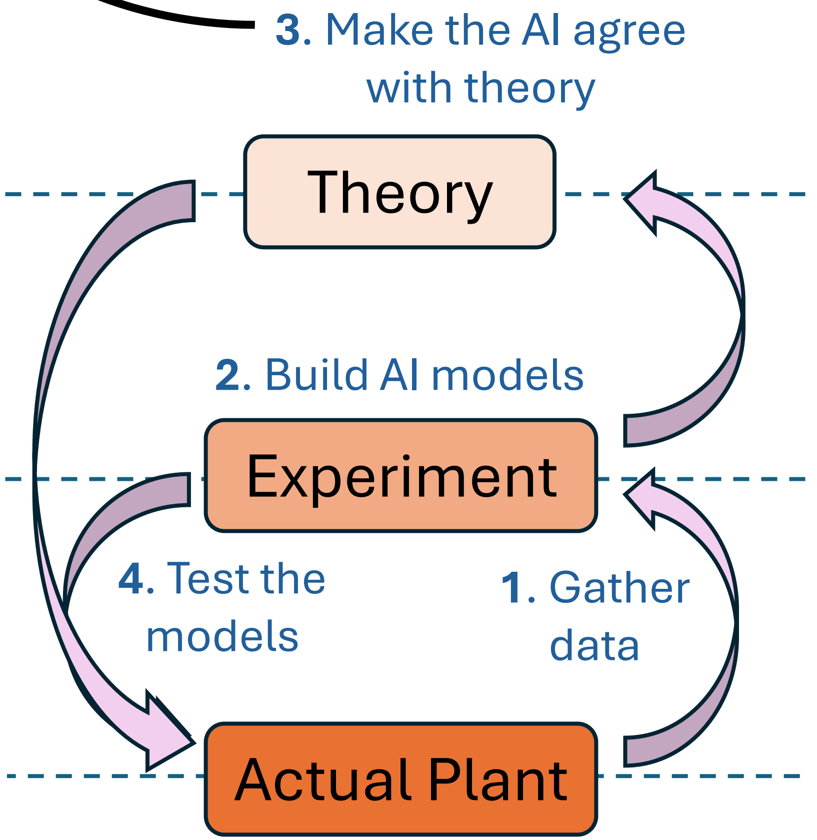
- Theoretical assumptions don’t hold in practice.
- Plant simulators (that employ theory) are too perfect.

3. Theory is *not available yet* for some phenomena.

- We still don’t understand some aspects of:
    - Multi-component phase change
    - Multi-phase flows for ≥ 3 phases (+ solids)
    - Co-crystallization in pharmaceutical plants
    - Protein-protein interactions
    - Catalyst deactivation, Battery degradation, etc.

AI for Engineering Practice

- *AI models can be deployed with or without theoretical backing.*



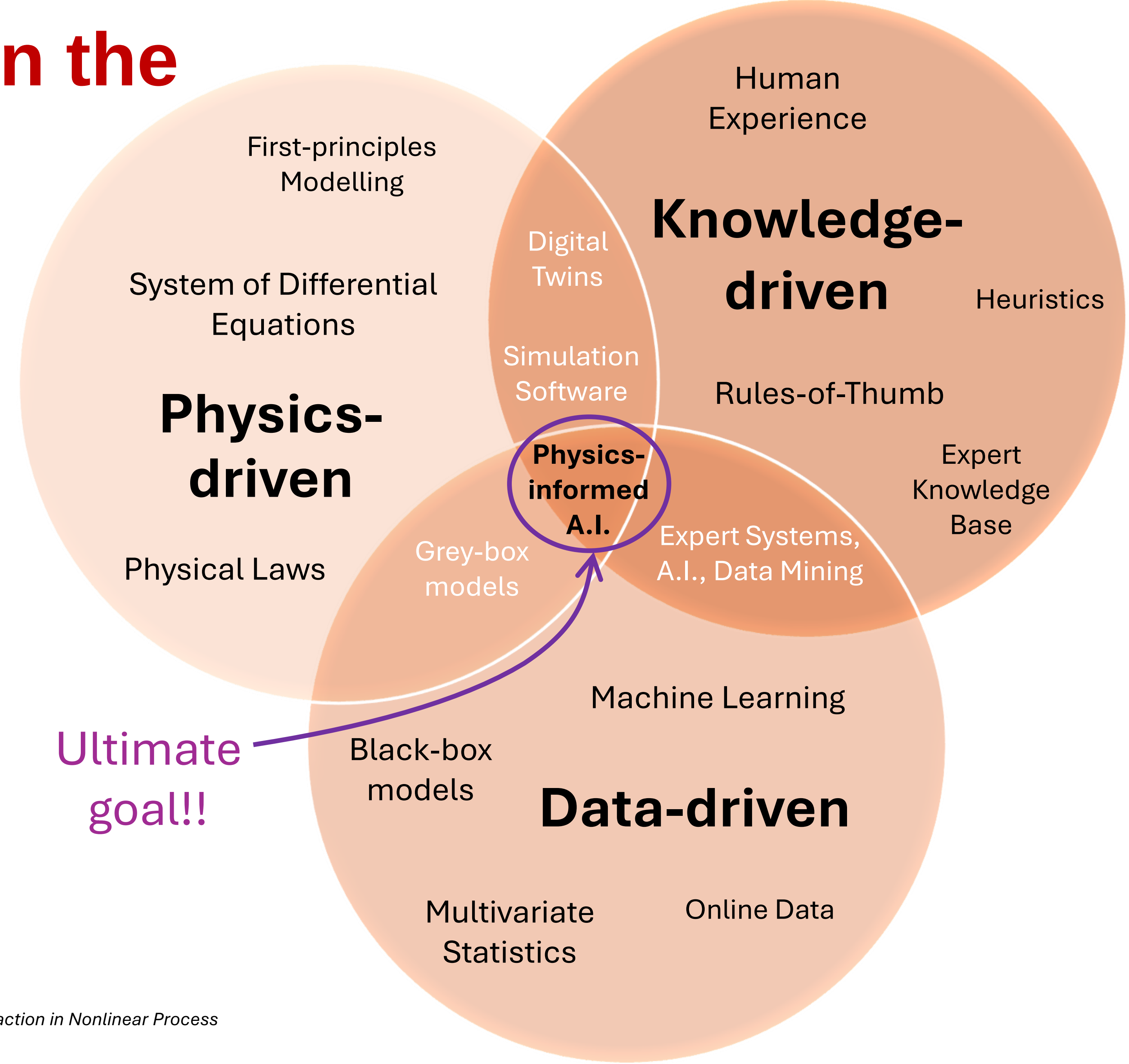
*or any point of application

Why use AI / ML in the first place?

Three approaches to engineering problems:

1. Physics-driven Methods
2. Knowledge-driven Methods
3. Data-driven Methods

Machine learning is a **data-driven approach**.



Reference: Pilario et al. (2020), *A Review of Kernel Methods for Feature Extraction in Nonlinear Process Monitoring*. MDPI: Processes, https://doi.org/10.3390/pr8010024

What is Process Systems Engineering?

Process Control and Monitoring

Process Design and Synthesis

Process Automation

Process Optimization

Process Engineering **+** **Systems Thinking**

Predictive Maintenance

Process Simulation

Process Operations

Computational Tools

Multi-scale

Impacts

Net Zero

From molecules to plants to supply chains

Sustainability

Integration

Synthesis

Process Systems Engineering Lab @ UPD

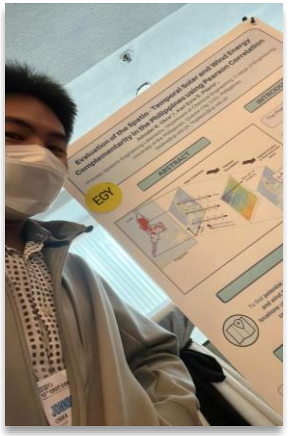
Mission

PSEL aims to provide **algorithmic, data-driven, and sustainable engineering solutions** to both process industries and allied problems in energy, water, materials, and the environment.

The lab leverages **computational tools** by integrating **machine learning, data analytics, and artificial intelligence with chemical engineering principles** to aid process systems engineering activities such as design and simulation, modeling, control and monitoring, optimization, and predictive maintenance.

Vision

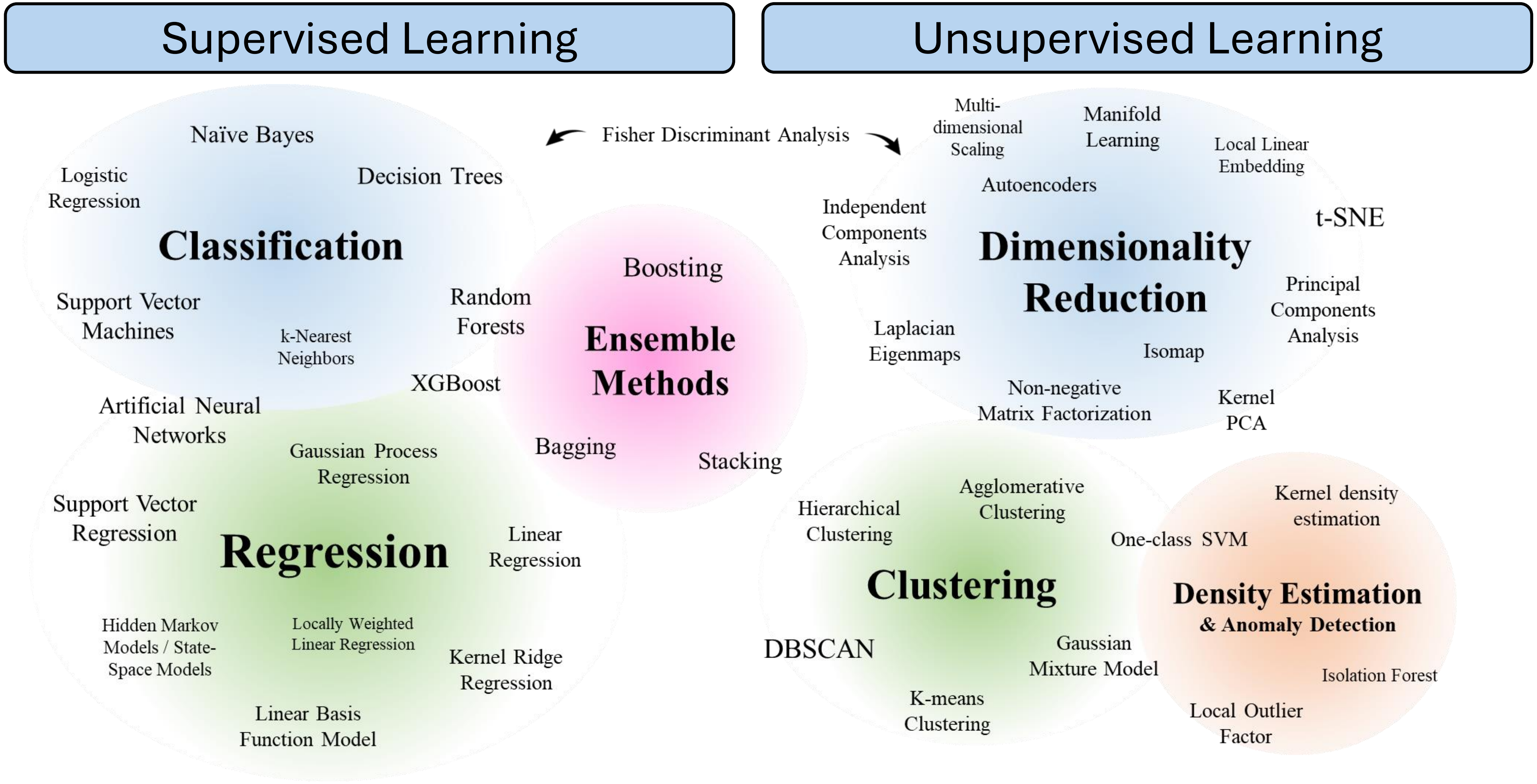
In the long term, PSEL aspires to be a leading research group that houses the best talent in **computer-aided chemical engineering** in the Philippines, while being a **trusted partner** of industry practitioners and researchers **locally and globally**.



Machine Learning and AI
in Process Systems Engineering:
Personal Experiences in Research

- AI in Process Control and Monitoring
- AI in Process Design and Synthesis
- AI in Energy Systems
- AI in Food Industry
- AI in Environment

ML Models and Algorithms



Reference: Pilario et al. (2020), *A Review of Kernel Methods for Feature Extraction in Nonlinear Process Monitoring.* MDPI: Processes, https://doi.org/10.3390/pr8010024

ML Models and Algorithms

Supervised Learning

These are images of dogs.

These are images of cars.

Now, what is this an image of?

Unsupervised Learning

Here are some images…

Is there an image that does not belong?

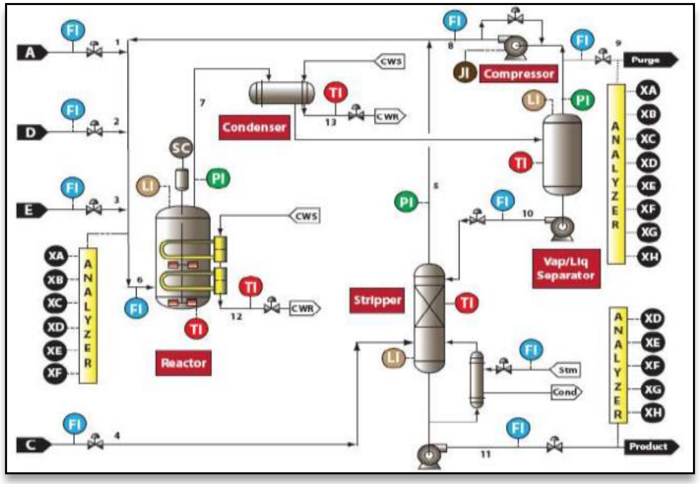
Are there images with similar patterns?

Reference: Pilario et al. (2020), *A Review of Kernel Methods for Feature Extraction in Nonlinear Process Monitoring.* MDPI: Processes, https://doi.org/10.3390/pr8010024

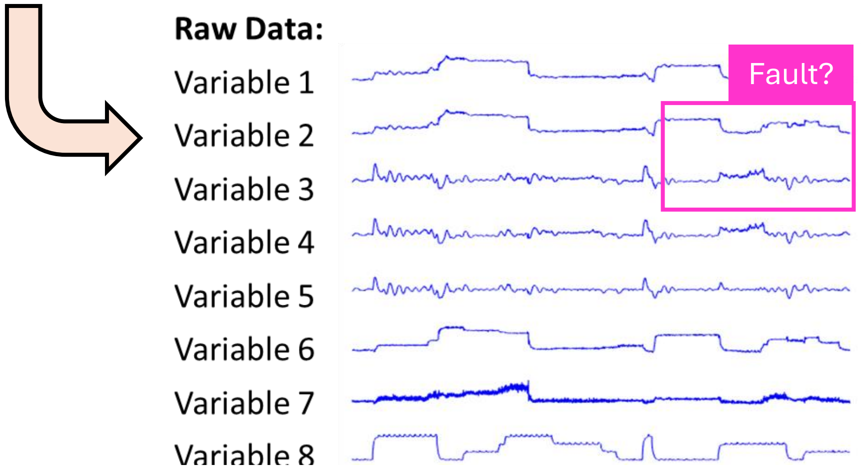
Process Control and Monitoring

AI models can **detect, diagnose, and prognose faults** in industrial plants.

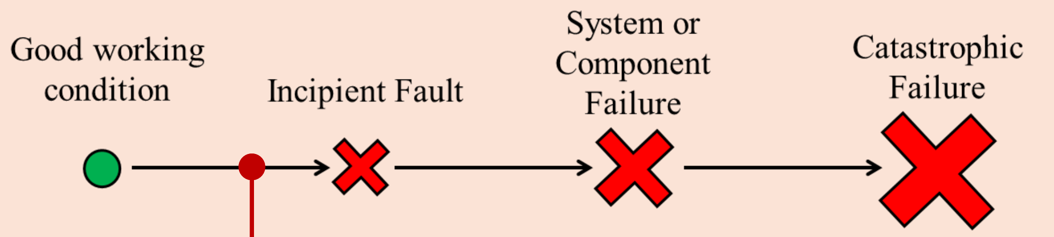
Industrial Plants



- Hundreds of process variables and KPIs.
- Different types of faults can occur at any time: *leakage, blockage, catalyst decay, fouling, loss of power…*
- Maintenance cost is high.



What is Process Monitoring?



Fault Detection

Did a fault occur?

Fault Diagnosis

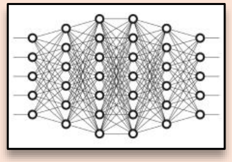
What is the root cause?
How severe is it?

Fault Prognosis

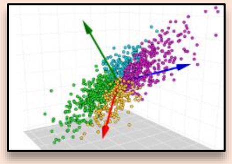
How long can we continue operating before failure occurs?

Process Control and Monitoring

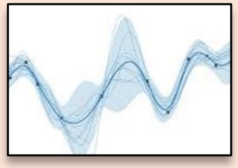
AI models can **detect, diagnose, and prognose faults** in industrial plants.



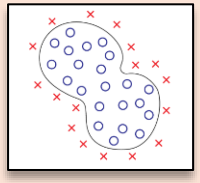
Deep Learning



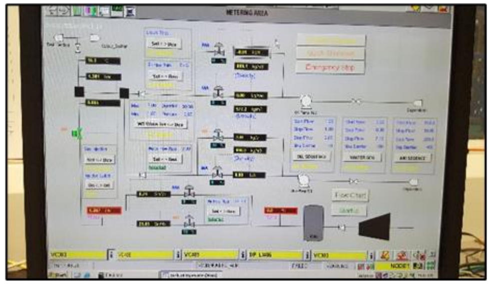
Dimensionality Reduction



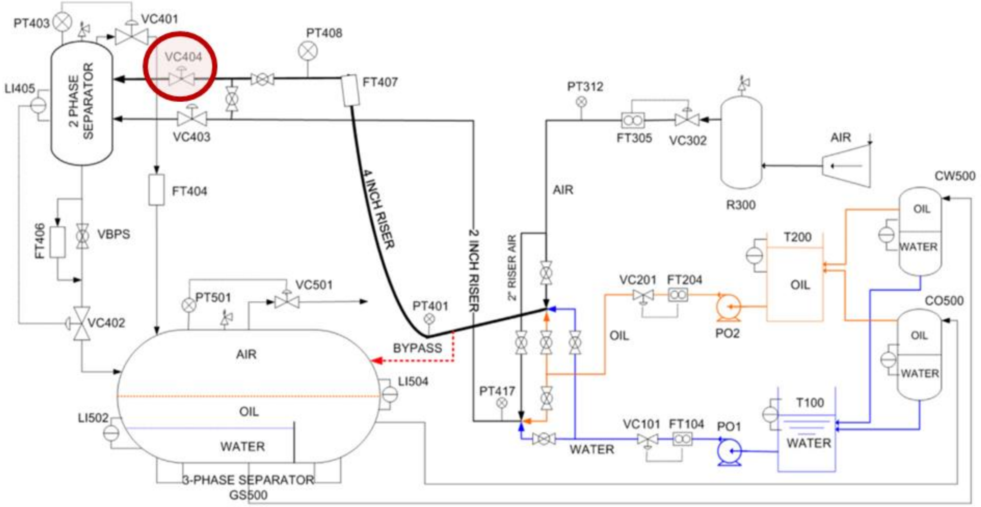
Forecasting



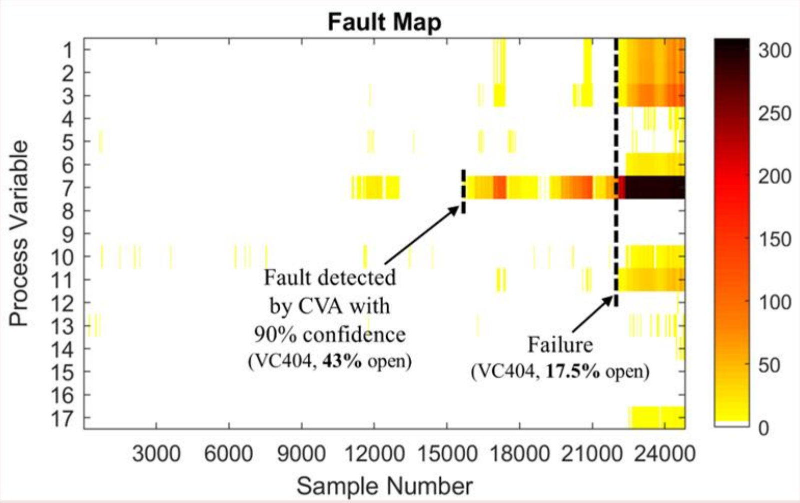
Anomaly Detection



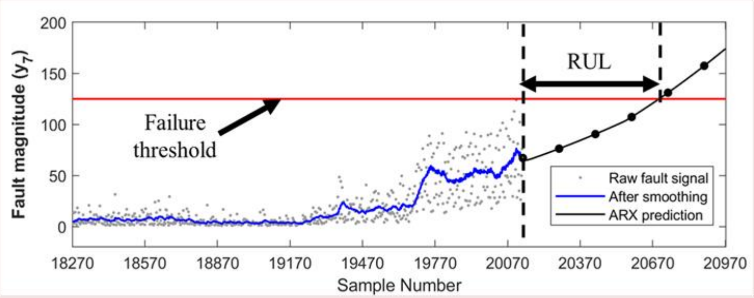
Cranfield Multiphase Flow Rig
Cranfield, United Kingdom



Fault Detection and Diagnosis



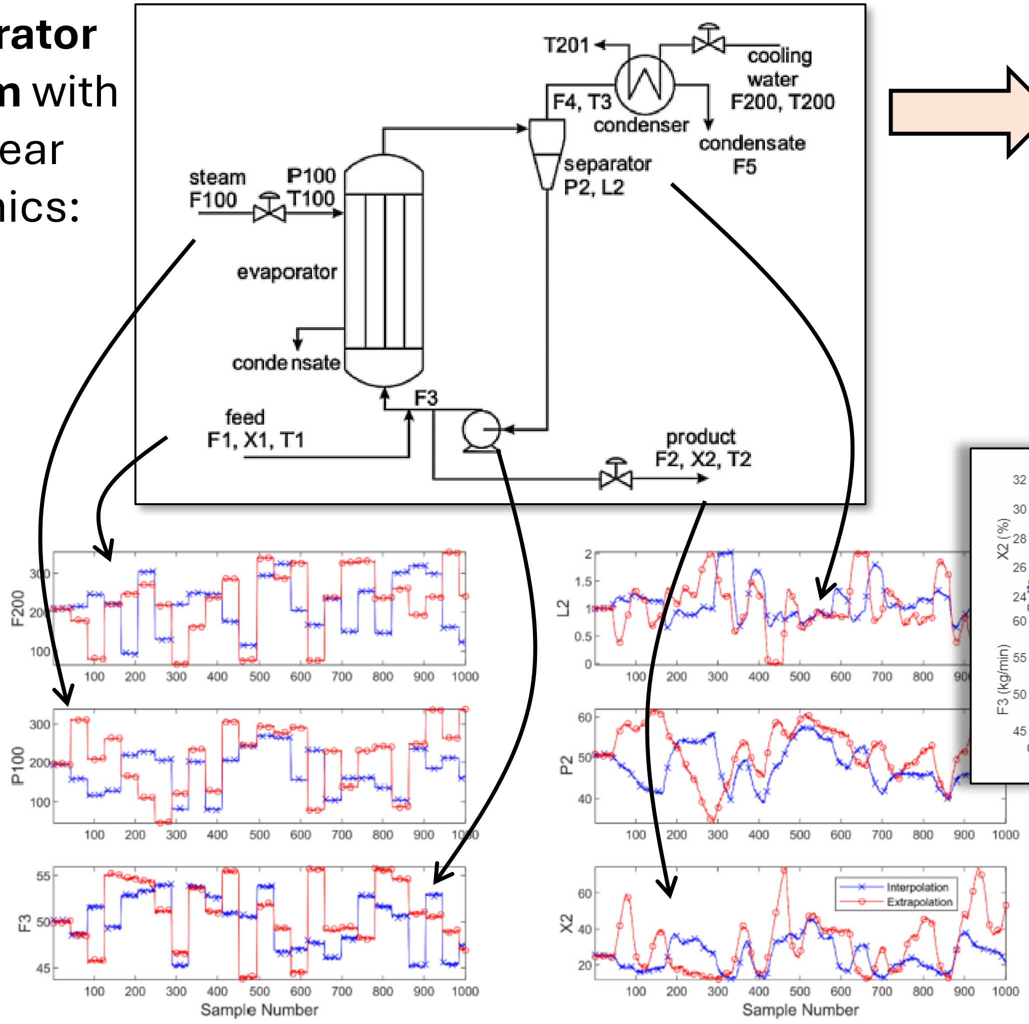
Fault Prognosis

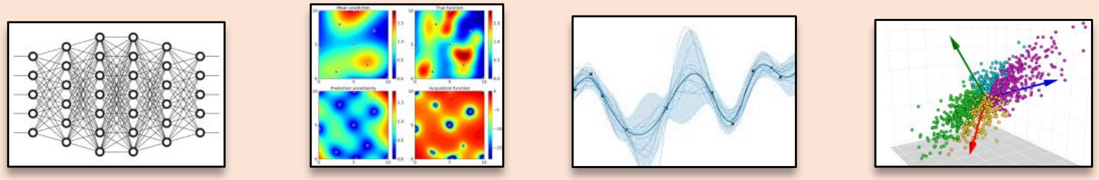


Process Control and Monitoring

AI surrogate models *accelerate* computation in **model predictive control**.

Evaporator System with Nonlinear Dynamics:



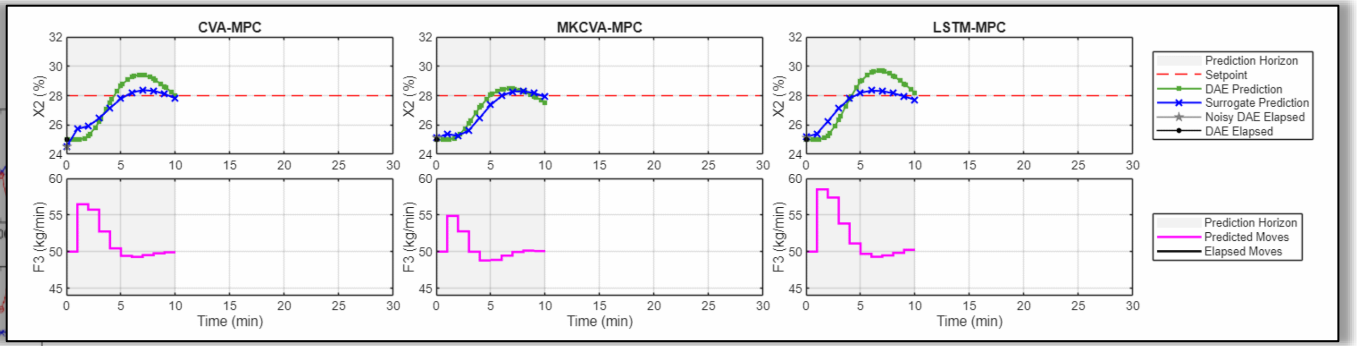


Deep Learning

Bayesian Optimization

Forecasting

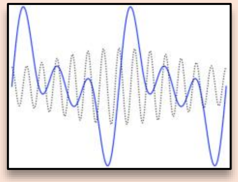
Dimensionality Reduction



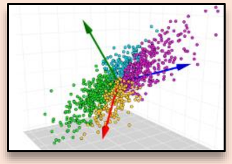
- Faster computation (from *minutes* down to *seconds*)
- Better control for setpoint tracking.
- No need for mass-energy balances and ODEs.

Process Control and Monitoring

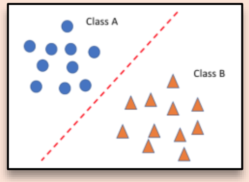
AI models can **predict hard-to-measure variables such as flow regimes.**



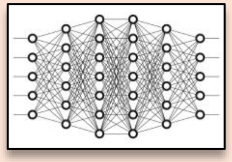
Signal Processing



Dimensionality Reduction



Classification



Deep Learning

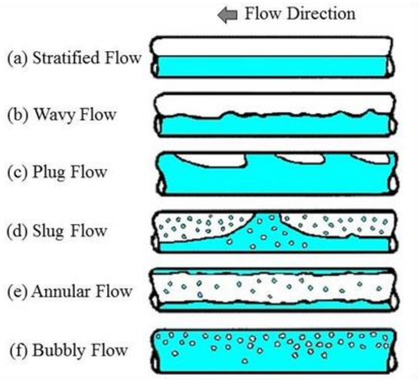
Two-Phase Flow Identification

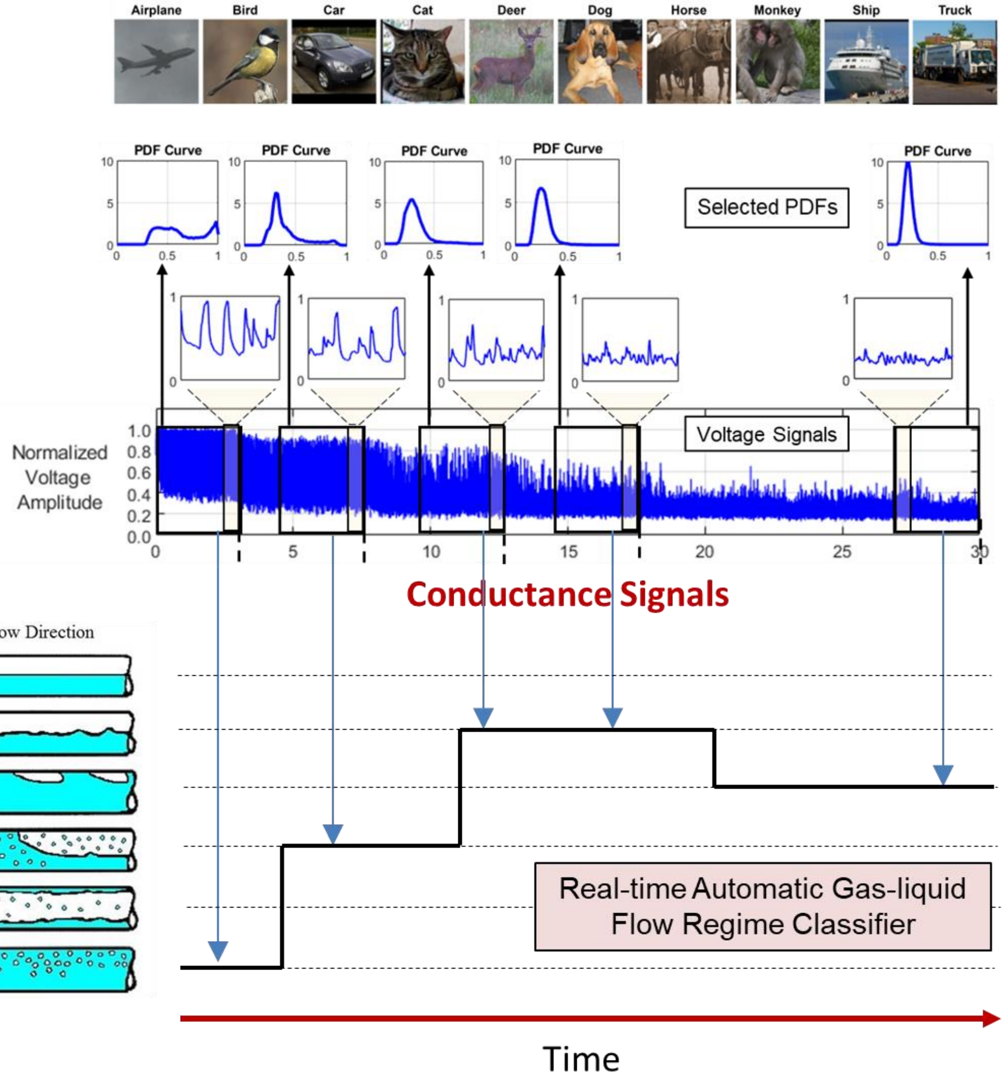
Gas-liquid Flow

Cranfield Multiphase Flow Rig
Cranfield, United Kingdom

Problem:

- How to identify the governing type of flow inside an opaque pipe?





Process Design and Synthesis

AI models can **recommend optimal materials, process designs for carbon capture**.

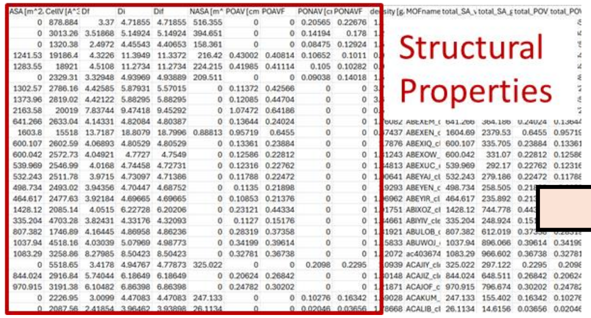
CoRE-MOF Database
- Computation-Ready Experimental (CoRE)
- Chung et al. (2014, 2019)
- ~14,000 **experimental** MOFs

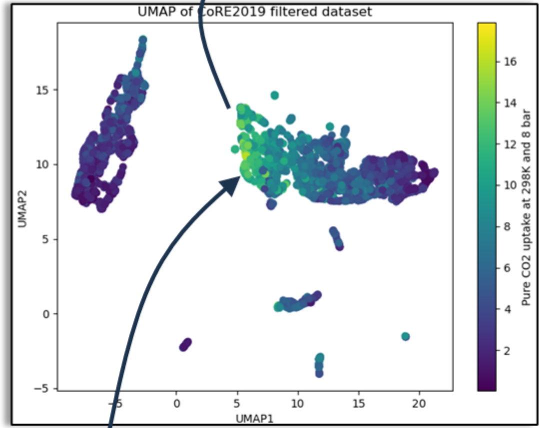
- Metal-organic Frameworks (MOFs) are promising for adsorbing CO_2.
- There are >100,000s of MOFs in literature.
- Which one is most selective to CO_2 and most cost-effective to operate?

Remove MOFs with rare metals

4,319 MOFs

MOF-1
MOF-2
MOF-3
…
…
…



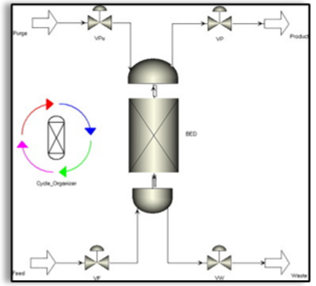


Sample a MOF

Fit a suitable CO_2 and N_2 isotherm

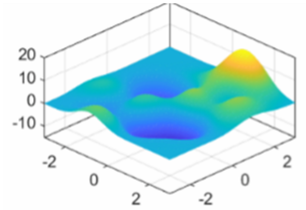
Freundlich, Langmuir, Sips, Toth, Linear

Input the MOF and its isotherms into *Aspen Adsorption Suite*

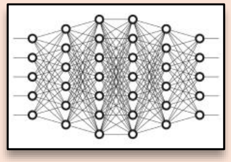


Calculate energy requirements, CO_2 purity, CO_2/N_2 selectivity

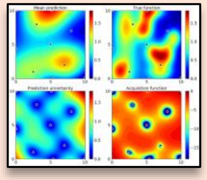
Use an AI recommender system to suggest a next MOF



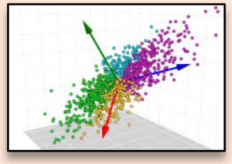
Find the nearest MOF to the one recommended by AI.



Regression



Bayesian Optimization



Dimensionality Reduction

AI in Energy

AI models are good for **long-term forecasting of renewable energy resources**.

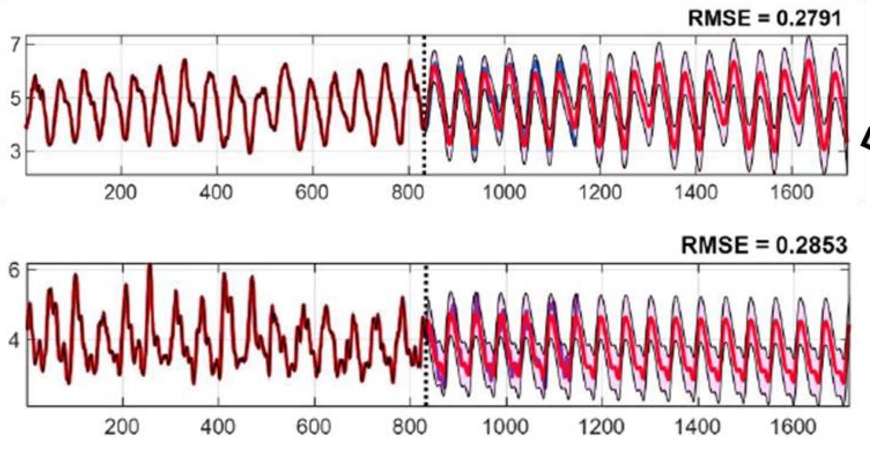
Forecasting and Complementarity Analysis of Philippine Solar & Wind Energy

Problem: Can we predict the *future* solar and wind energy output per site in the Philippines?

Training Data | Test Data (Forecast)

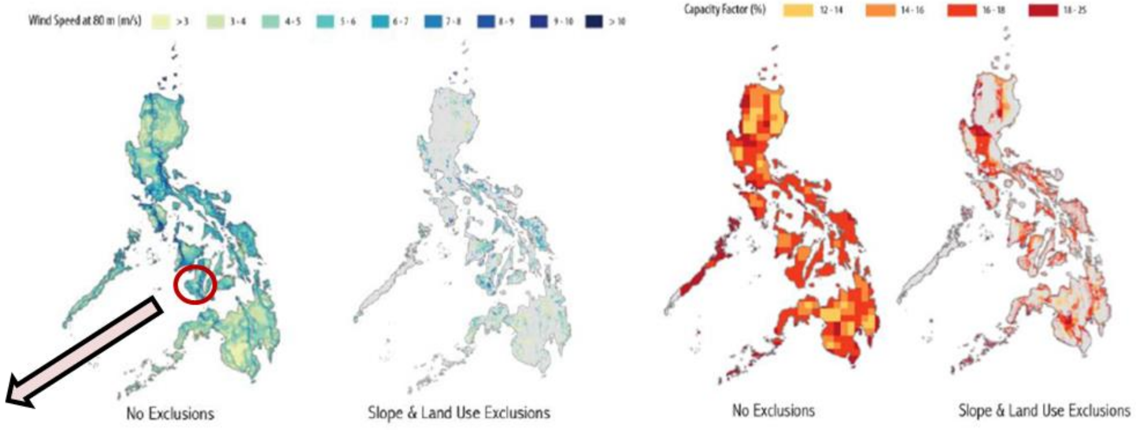
Solar power

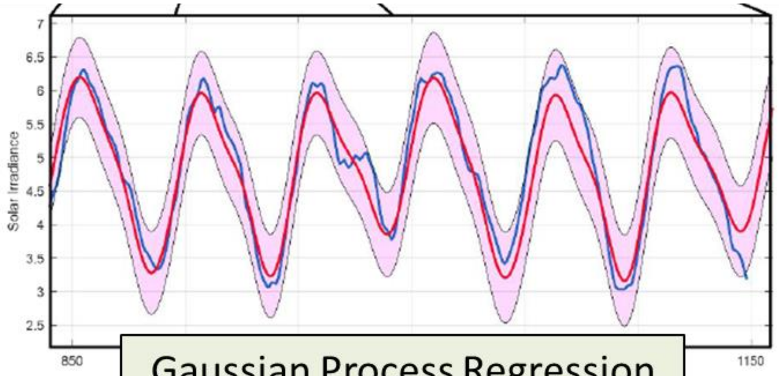
Wind power



Weekly Average Values

Wind Resource Map | Solar Resource Map





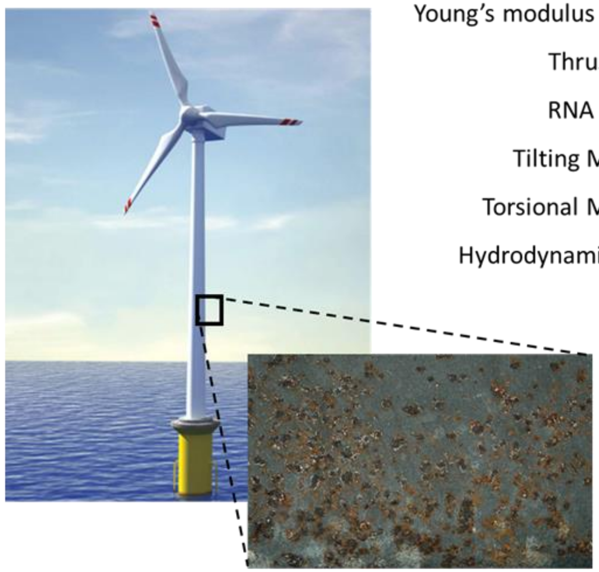
Gaussian Process Regression

AI in Energy

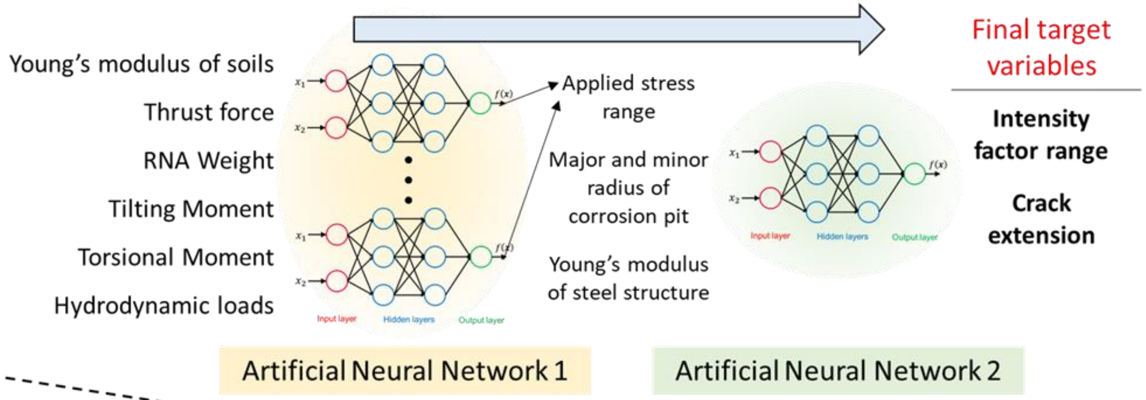
AI models for developed for prognosis can be used in **wind turbines**.

Reliability Analysis of Offshore Wind Turbines

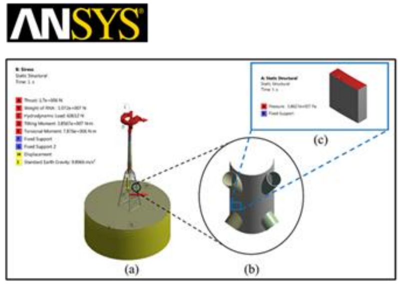
- **Problem:** Pitting corrosion in offshore wind turbine monopiles.
- **Question:** How to calculate the Remaining Useful Life of the wind turbine?
- **Solution:** Neural networks help predict crack extensions from input information.



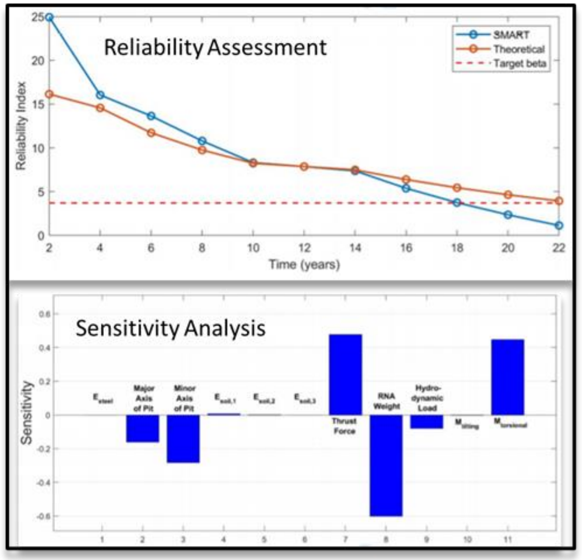
Pitting Corrosion



- **Data Set:** 300 Finite-Element Analysis (FEA) simulations of the structure, using ANSYS, while varying stochastic variables.
- **Structure:** 10 neurons each hidden layer + sigmoid activation
- Training set, 70%; Validation set, 15%; Testing set, 15%.



Result: 18 yrs service life

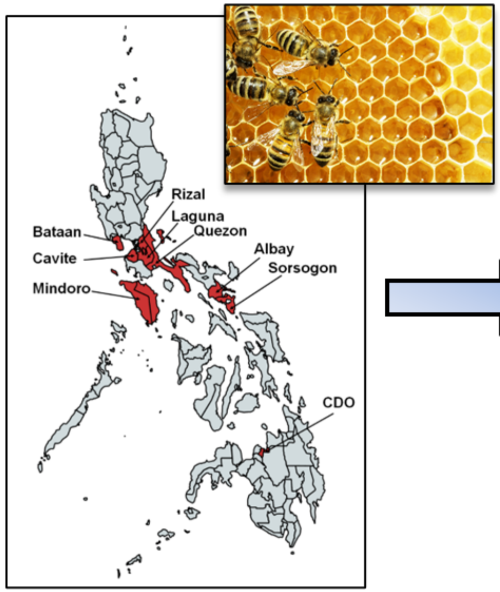


AI in the Food Industry

AI models can be used to **authenticate and differentiate food substances.**

Geographical Differentiation of *Bee Honey* using Unsupervised Machine Learning

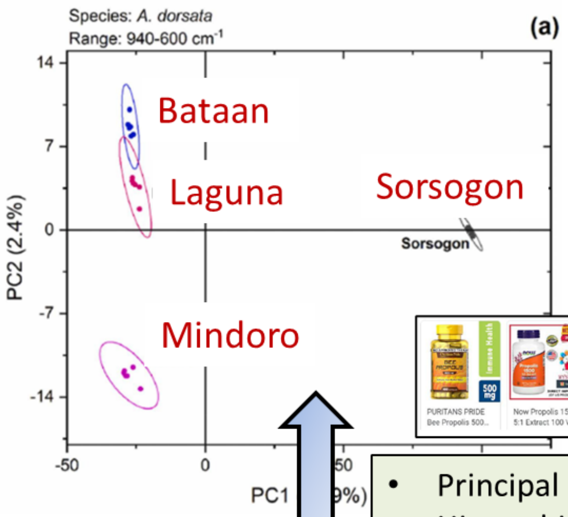
Problem: Can we figure out the geographical source of honey based on its composition?

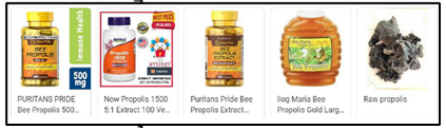


Chemometric Analysis of Philippine Honey

GCMS
(Gas chromatograph – Mass Spectrometer)

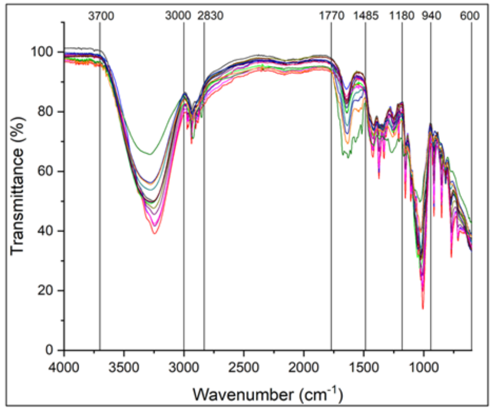
FTIR
(Fourier Transform Infrared Spectroscopy)





Uses:
- Honey authentication
- Bee species differentiation
- Mapping honey properties

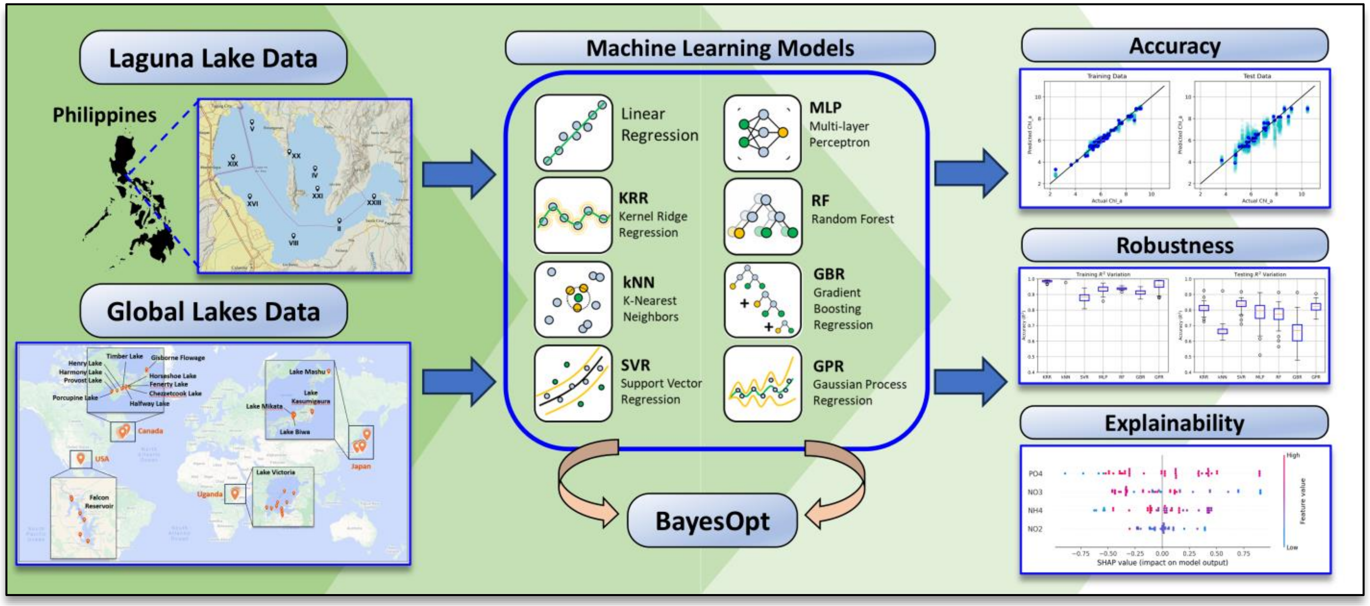
- Principal Components Analysis
- Hierarchical Clustering

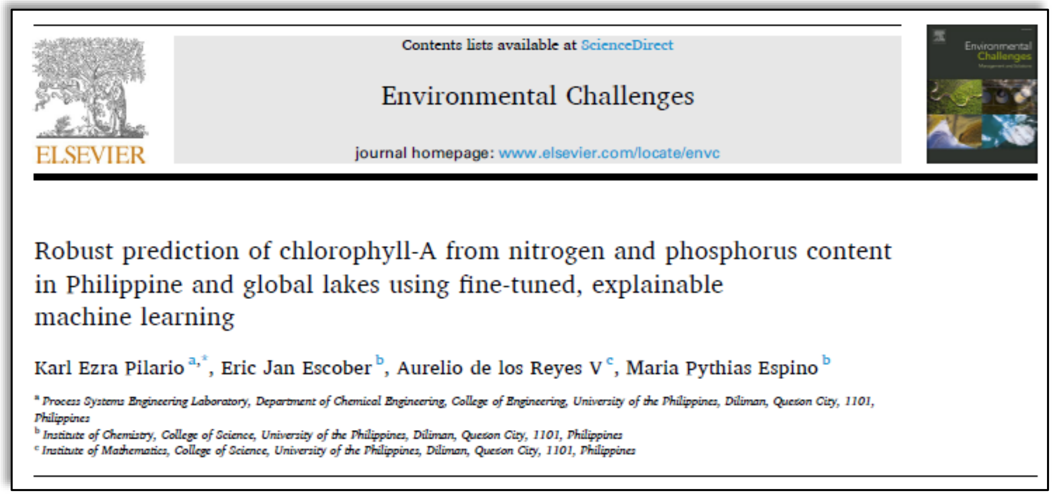


- Honey Composition

AI for the Environment

AI models can **predict water quality parameters and detect algal blooms.**





Contents lists available at ScienceDirect

Environmental Challenges

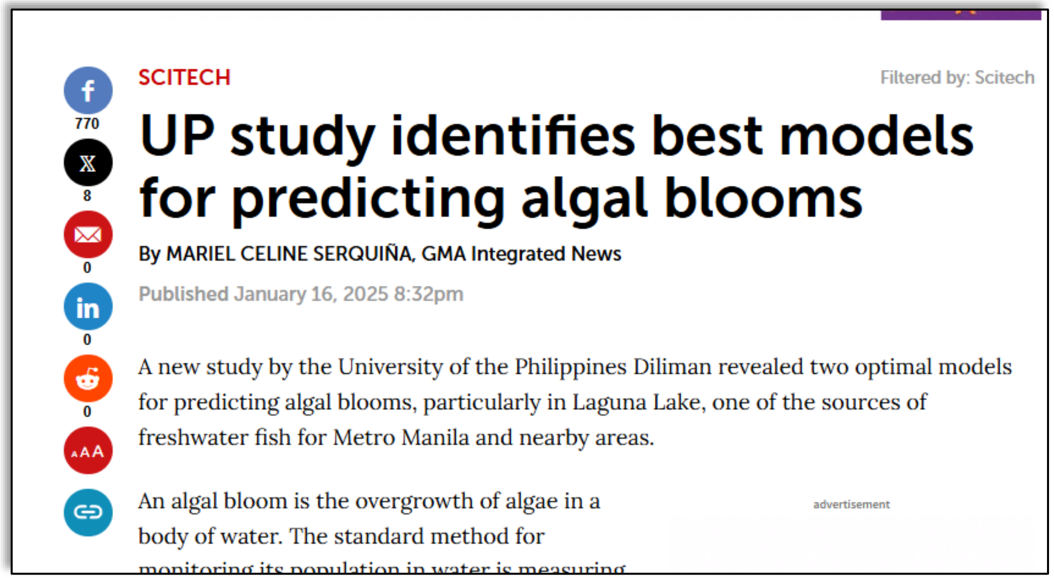
journal homepage: www.elsevier.com/locate/envc

ELSEVIER

Robust prediction of chlorophyll-A from nitrogen and phosphorus content in Philippine and global lakes using fine-tuned, explainable machine learning

Karl Ezra Pilario [a,*], Eric Jan Escober [b], Aurelio de los Reyes V [c], Maria Pythias Espino [b]

[a] Process Systems Engineering Laboratory, Department of Chemical Engineering, College of Engineering, University of the Philippines, Diliman, Quezon City, 1101, Philippines
[b] Institute of Chemistry, College of Science, University of the Philippines, Diliman, Quezon City, 1101, Philippines
[c] Institute of Mathematics, College of Science, University of the Philippines, Diliman, Quezon City, 1101, Philippines



SCITECH

Filtered by: Scitech

UP study identifies best models for predicting algal blooms

By MARIEL CELINE SERQUIÑA, GMA Integrated News

Published January 16, 2025 8:32pm

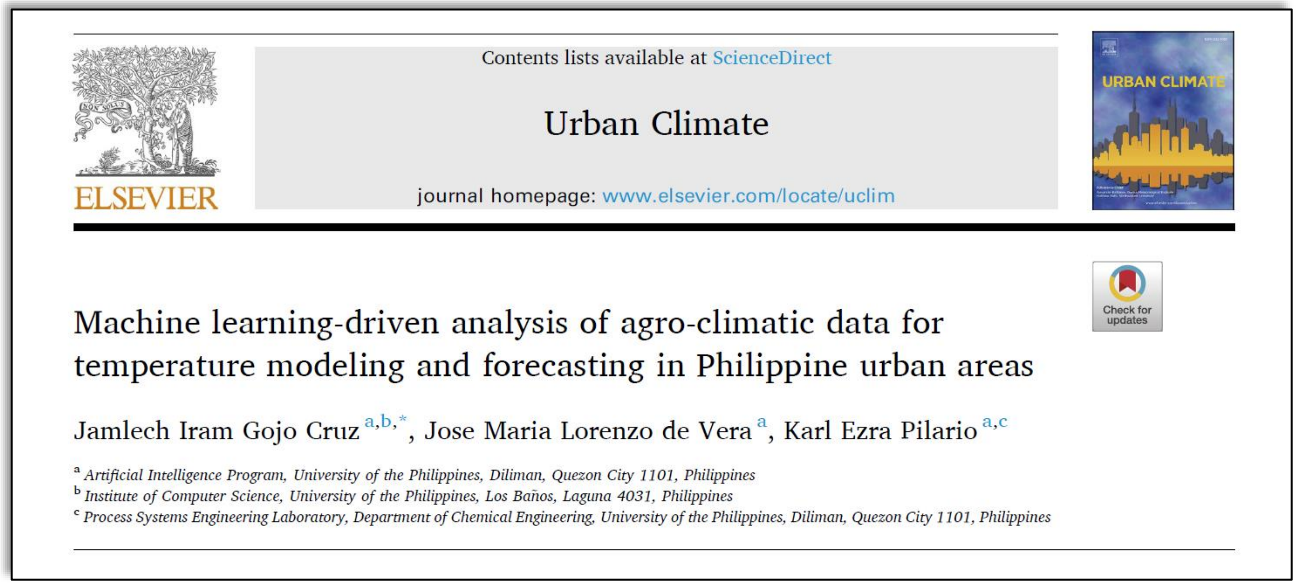
A new study by the University of the Philippines Diliman revealed two optimal models for predicting algal blooms, particularly in Laguna Lake, one of the sources of freshwater fish for Metro Manila and nearby areas.

An algal bloom is the overgrowth of algae in a body of water. The standard method for

- Best ML models were found for predicting algal blooms in **Laguna Lake** and **global lakes**.
- **P-content** in lakes is the best predictor of algal blooms.

AI for the Environment

AI models can **be used to forecast climate variables in the long-term.**


Contents lists available at ScienceDirect

Urban Climate

journal homepage: www.elsevier.com/locate/uclim

ELSEVIER

URBAN CLIMATE

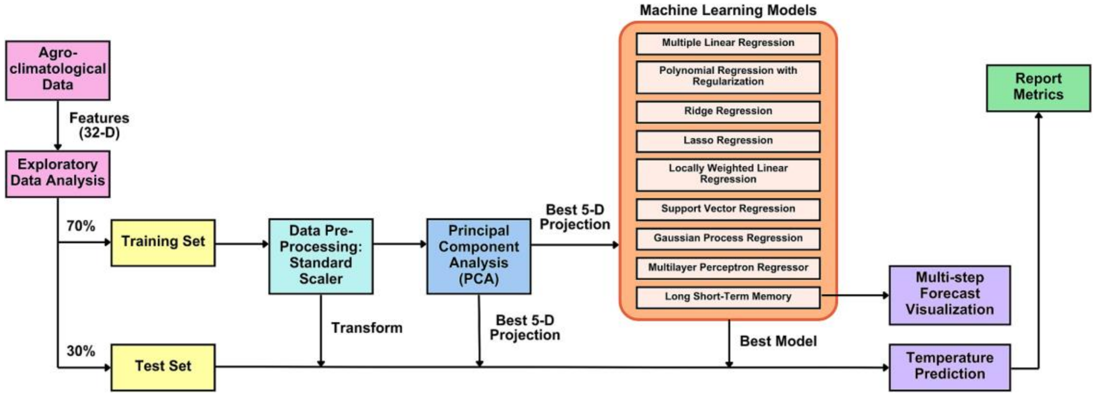
Check for updates

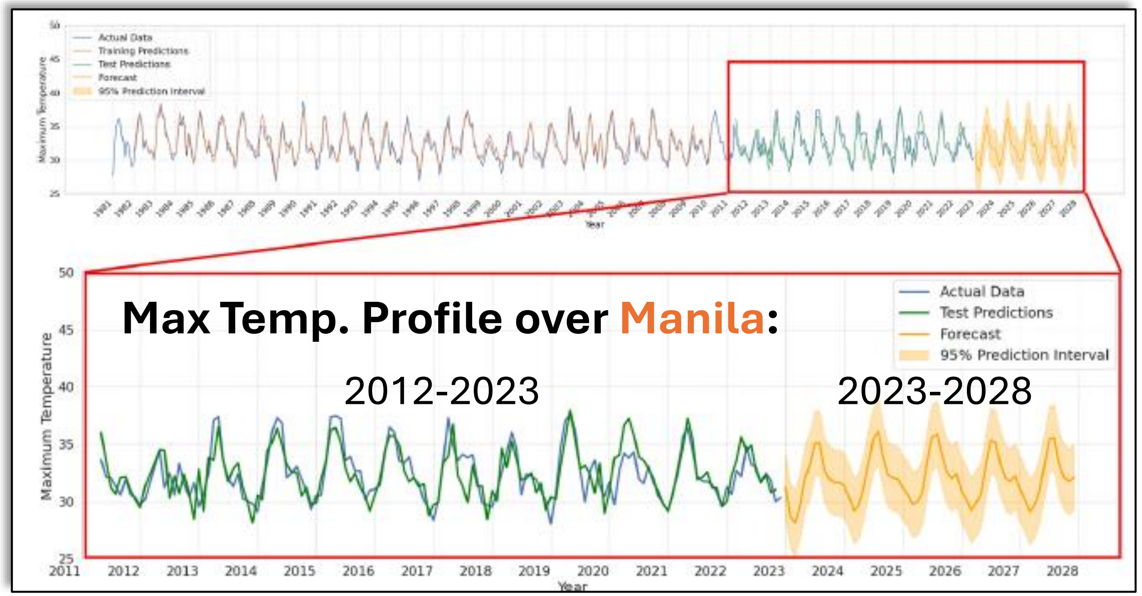
Machine learning-driven analysis of agro-climatic data for temperature modeling and forecasting in Philippine urban areas

Jamlech Iram Gojo Cruz [a,b,*], Jose Maria Lorenzo de Vera [a], Karl Ezra Pilario [a,c]

[a] Artificial Intelligence Program, University of the Philippines, Diliman, Quezon City 1101, Philippines
[b] Institute of Computer Science, University of the Philippines, Los Baños, Laguna 4031, Philippines
[c] Process Systems Engineering Laboratory, Department of Chemical Engineering, University of the Philippines, Diliman, Quezon City 1101, Philippines

- The max temperature in Manila is 80% strongly associated with other climate variables: precipitation, humidity, etc.
- LSTM predicts max temp. for 2025-2027 as **35°C ± 2.8°C**.





Publications from PSEL@UPD

AI in Process Control and Monitoring

Fault Detection, Diagnosis, Prognosis

- **Pilario and Cao (2018).** doi.org/10.1109/TII.2018.2810822
- **Pilario et al. (2019).** doi.org/10.1016/j.compchemeng.2018.12.027
- **Pilario et al. (2019).** doi.org/10.23919/IConAC.2019.8895249
- **Pilario et al. (2019).** doi.org/10.1016/B978-0-12-818634-3.50200-9
- **Pilario et al. (2020).** doi.org/10.3390/pr8010024
- **Jang et al. (2023).** doi.org/10.1109/TII.2023.3240601

Flow Regime Identification

- **Nnabuife et al. (2019).** doi.org/10.1016/j.flowmeasinst.2019.05.002
- **Eyo et al. (2021).** doi.org/10.1109/TCYB.2019.2910257
- **Roxas et al. (2022).** doi.org/10.1016/j.dche.2022.100024
- **Khan et al. (2023).** doi.org/10.1108/HFF-09-2023-0526
- **Khan et al. (2024).** doi.org/10.1016/j.enganabound.2024.03.006

Process Control and System Identification

- **Pilario et al. (2021).** doi.org/10.1109/TIE.2020.2996142
- **Madayag, Pilario (2023).** doi.org/10.1016/B978-0-443-15274-0.50250-X
- **Pilario and Wu.** *Under Review*

AI in Energy Systems

- **Shittu et al. (2020).** doi.org/10.1002/we.2542
- **Pilario et al. (2022).** doi.org/10.1016/j.dche.2022.100036
- **Pilario et al. (2022).** doi.org/10.3390/su14063253
- **Gnetchejo et al. (2024).** doi.org/10.1080/01430750.2023.2281611

AI in Materials Informatics

- **Hipolito et al. (2022).** doi.org/10.1016/j.commatsci.2023.112607
- **Hipolito et al. (2023).** doi.org/10.14886/jvss.2023.0_2Fp08
- **Tomacruz et al. (2022).** doi.org/10.3303/CET2294122
- **Ordillo et al. (2024).** doi.org/10.1088/1361-651X/ad53ee

AI for Environment

- **Bhagat et al. (2023).** doi.org/10.1016/j.jclepro.2022.135522
- **Pilario et al. (2024).** doi.org/10.1016/j.envc.2024.101056
- **Gojo Cruz et al. (2025).** doi.org/10.1016/j.uclim.2025.102339

AI in the Food Industry

- **Grabato et al. (2022).** doi.org/10.1016/j.jfca.2022.104853
- **Pilario et al. (2022).** doi.org/10.1016/j.eswa.2021.115938

AI in ChemE Education

- **Pilario (2024).** doi.org/10.1016/j.dche.2024.100163

Publications from PSEL@UPD

AI in Process Co

Github: https://github.com/kspilario/MLxChE

rmatics

Fault Detection, Diagnos

- **Pilario and Cao (201**
- **Pilario et al. (2019).** d
- **Pilario et al. (2019).** d
- **Pilario et al. (2019).** d
- **Pilario et al. (2020).** d
- **Jang et al. (2023).** do

2023.112607
2Fp08
2
3ee

Flow Regime Identificati

- **Nnabuife et al. (2019**
- **Eyo et al. (2021).** doi
- **Roxas et al. (2022).** d
- **Khan et al. (2023).** d
- **Khan et al. (2024).** d

5522
56

Process Control and Sys

- **Pilario et al. (2021).** d
- **Madayag and Pilario**
- **Pilario and Wu.** *Under Review*

53
38

Digital Chemical Engineering 11 (2024) 100163

Contents lists available at ScienceDirect

Digital Chemical Engineering

journal homepage: www.elsevier.com/locate/dche

ELSEVIER

Original Article

Teaching classical machine learning as a graduate-level course in chemical engineering: An algorithmic approach

Check for updates

Karl Ezra Pilario

Process Systems Engineering Laboratory, Department of Chemical Engineering, University of the Philippines, Diliman, Quezon City, 1101, Philippines

AI in Energy Systems

- **Shittu et al. (2020).** doi.org/10.1002/we.2542
- **Pilario et al. (2022).** doi.org/10.1016/j.dche.2022.100036
- **Pilario et al. (2022).** doi.org/10.3390/su14063253
- **Gnetchejo et al. (2024).** doi.org/10.1080/01430750.2023.2281611

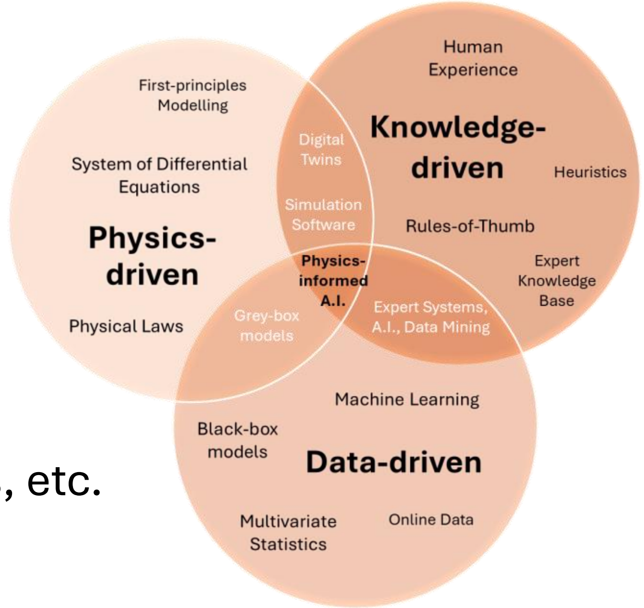
AI in ChemE Education

- **Pilario (2024).** doi.org/10.1016/j.dche.2024.100163

Ideas for Future Research

1. More ideas on physics-informed AI + Human-machine interactions

- How to constrain AI models to be consistent with theory?
- How to reduce computational cost of AI model training and deployment?
- How to make AI work with less data? Collaborate better with humans?
- AI explainability and transparency issues
- Uncertainty quantification in deep learning models



2. Other applications of AI in PSE

- AI for Sustainable Development Goals (SDGs)!!
- We need more AI in materials discovery: new catalysts, adsorbents, etc.
- Applications of Generative AI, e.g. Large Language Models in PSE
- Automatic / Real-time design of experiments

3. Increasing the Technology Readiness Level of AI in Practice

- Addressing issues in AI deployment, i.e. requires inter-disciplinary teams
- AI adoption is still slow as it entails system-wide changes
- Need to increase public awareness of AI and its pitfalls
- How to leverage trustworthy AI for data-driven public policy-making?

Some Lessons Learned

1. **In industry, Big Data does not always equal Good Models.**
    - Big data may contain missing values, outliers, redundant data.
    - The data should represent a large range of scenarios. Otherwise, we learn nothing.
    - We still need simulators (theory) + human inputs (knowledge) to augment real data.

2. **The size and quality of training data dictates what model is best.**
    - Neural networks are not always the best model. They are data-hungry.
    - A finely-tuned shallow model is better than deep learning for small data.
    - No free lunch: No single model is consistently best for all case studies.
    - Always do hyper-parameter tuning when comparing models.

3. **More efforts should be done to make AI trustworthy in practice.**
    - Always apply explainable AI, especially if the model is completely black-box.
    - Always report the right metrics to stakeholders, especially uncertainty.
    - Stakeholders are more interested in instances when the AI makes mistakes.

Thank you!

Karl Ezra S. Pilario
Associate Professor
Department of Chemical Engineering
University of the Philippines, Diliman

Email: kspilario@up.edu.ph

Acknowledgment

DOST-ERDT

MAYNILAD

FLUOR®