



<https://github.com/pvateekul/ieat2026>



**กนอ.**  
การนิคมอุตสาหกรรมแห่งประเทศไทย

หลักสูตรการวิเคราะห์ข้อมูลของระบบฐานข้อมูล  
ขนาดกลาง และ ฐานข้อมูลขนาดใหญ่

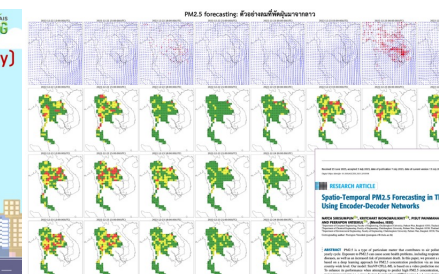
Prof. Peerapon Vateekul, Ph.D.

[Peerapon.v@chula.ac.th](mailto:Peerapon.v@chula.ac.th)

# Prof. Peerapon Vateekul, Ph.D. (Aj.Pop)



- Associate professor at Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Thailand
- **Research: Specialized in applying AI/ML/DL/Gen AI techniques to solve real-world problems across various domains, including healthcare, hydrometeorology, transportation, energy, etc.**
- Parts of NVAITC (NVIDIA AI Technology Center) since 2018
- Senior researcher at AI for Center of Excellence in Digital and AI for Mental Health (AIMET), Faculty of Engineering, Chulalongkorn University
- DATA MIND (Data Analytics Group, Machine Intelligence and Knowledge Discovery Lab)
- In collaboration with many organizations (government & private sectors), e.g., Chulalongkorn hospital.





# วัตถุประสงค์และตารางการเรียนรู้

1. การวิเคราะห์ข้อมูลด้วยเทคนิคปัญญาประดิษฐ์
2. เรียนรู้การเทคนิคการวิเคราะห์ข้อมูลประเภทต่าง ๆ เช่น Supervised & Unsupervised Learning
3. ได้เรียนรู้ผ่านทางการปฏิบัติจริงด้วยเครื่องมือ No-Code / Low-Code เช่น Knime เป็นต้น
4. ได้รู้จักเทคนิคปัญหาประดิษฐ์ขั้นสูงอื่น ๆ เช่น การเรียนรู้เชิงลึก (Deep Learning) และโมเดลภาษาขนาดใหญ่ (Large Language Models: LLM)

## หัวข้อการฝึกอบรม

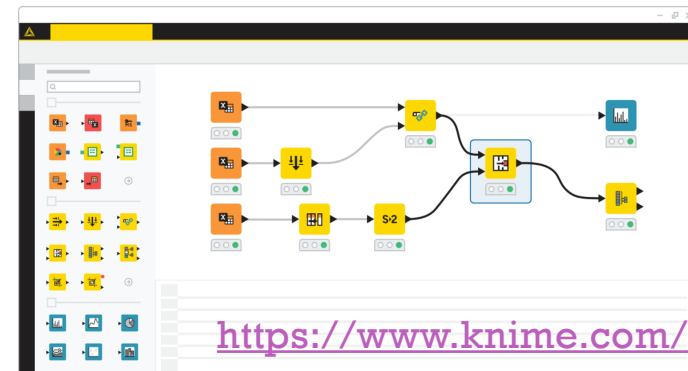
### ช่วงเช้า (09.00 – 12.00 น.)

- 09.00 – 09.30: AI/ML: แนะนำหลักการวิเคราะห์ข้อมูล แนวคิดพื้นฐาน และเครื่องมือที่ใช้
- 09.30 – 10.30: การวิเคราะห์ข้อมูลด้วยเทคนิค Supervised Learning สำหรับโจทย์ Classification
- 10.30 – 10.45: Break
  - 10.30 – 12:00: การวิเคราะห์ข้อมูลด้วยเทคนิค Supervised Learning สำหรับโจทย์ Classification (cont.) + Regression

### วันที่ 1

### ช่วงบ่าย (13.00 – 16.00 น.)

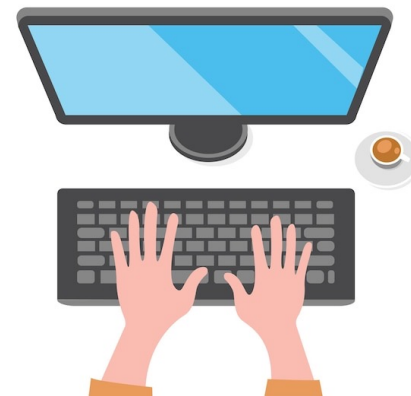
- 13:00 – 14:00: การวิเคราะห์ข้อมูลด้วยเทคนิค Supervised Learning สำหรับโจทย์ Regression (cont.)
- 14:00 – 14.30: การวิเคราะห์ข้อมูลแบบ Unsupervised Learning สำหรับการจัดกลุ่มข้อมูล (Clustering)
- 14.30 – 14.45: Break
- 14.45 – 16:00: เทคนิคการวิเคราะห์ข้อมูลขั้นสูง เช่น Deep Learning และ โมเดลภาษาขนาดใหญ่ (LLM) และถาม-ตอบ



<https://www.knime.com/downloads>

# + Outline

- Slide 1: Introduction
- Slide 2: Classification
- Slide 3: Regression
- Slide 4: Clustering
- Slide 5: Advanced ML
- Slide 7: Conclusion
- Lab 1: Classification
- Lab 2: Regression
- Lab 3: AutoML
- Lab 4: Clusteirng
- Lab 5: Object Detection (Optional)





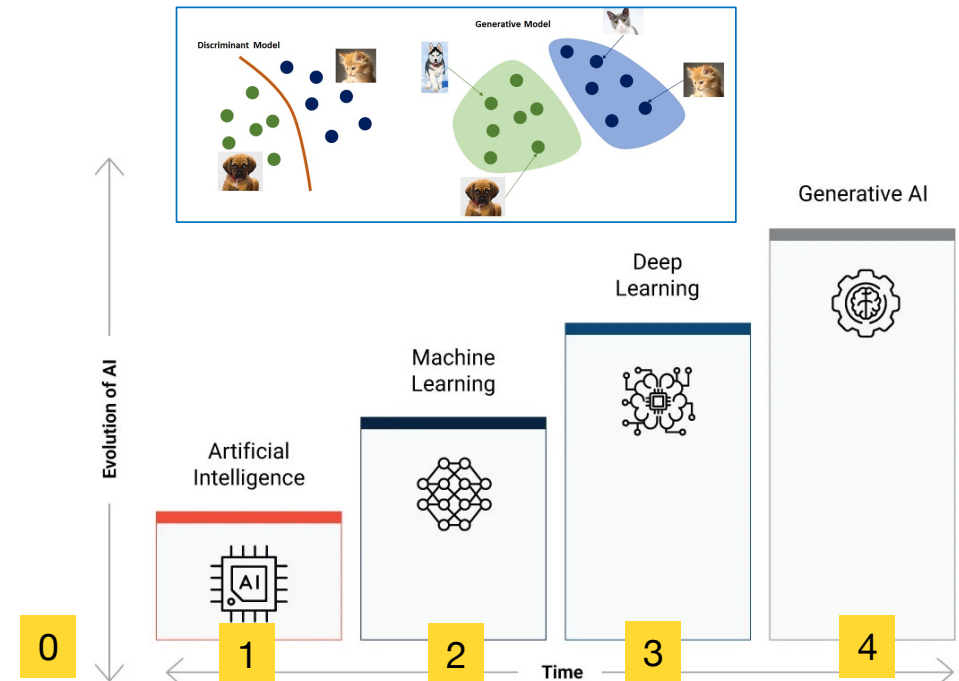
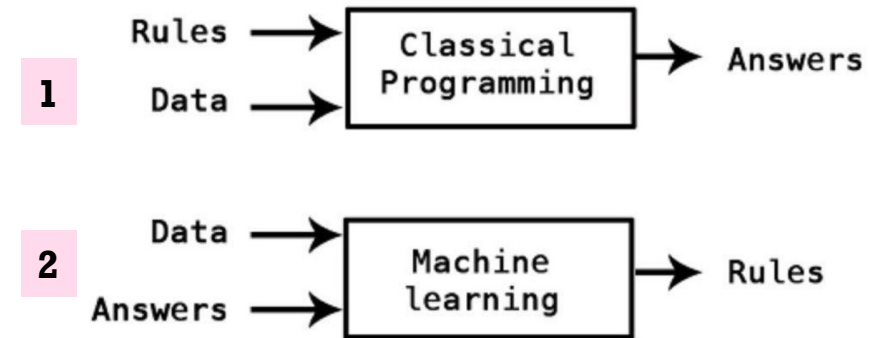
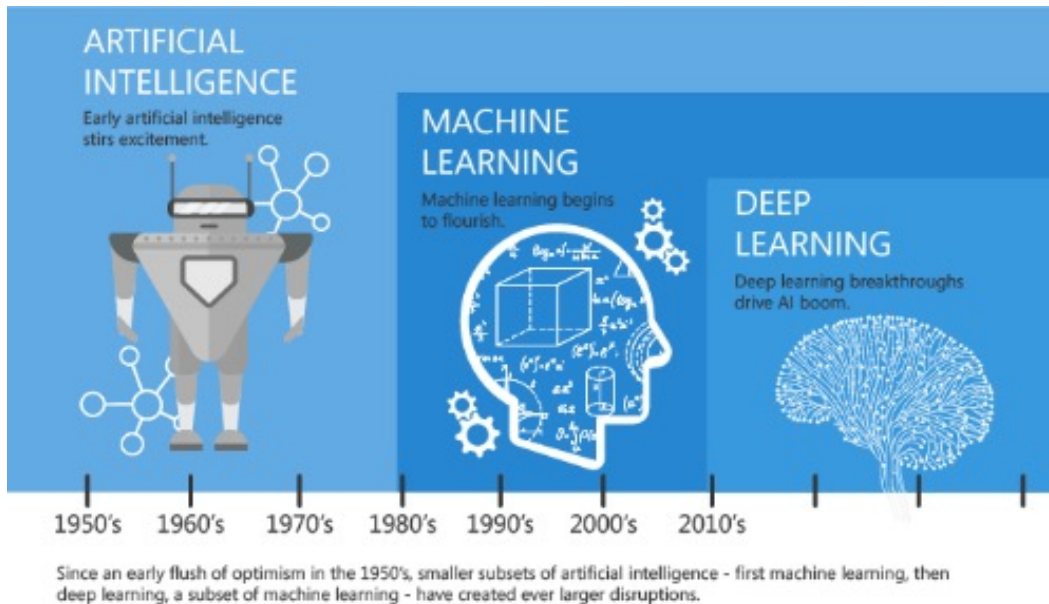
# Introduction

There are many kinds of AI models.



# AI = Automation

- 0) Not AI Solution (not automatic)
- 1) Rule-based AI
- 2) Machine Learning (ML)

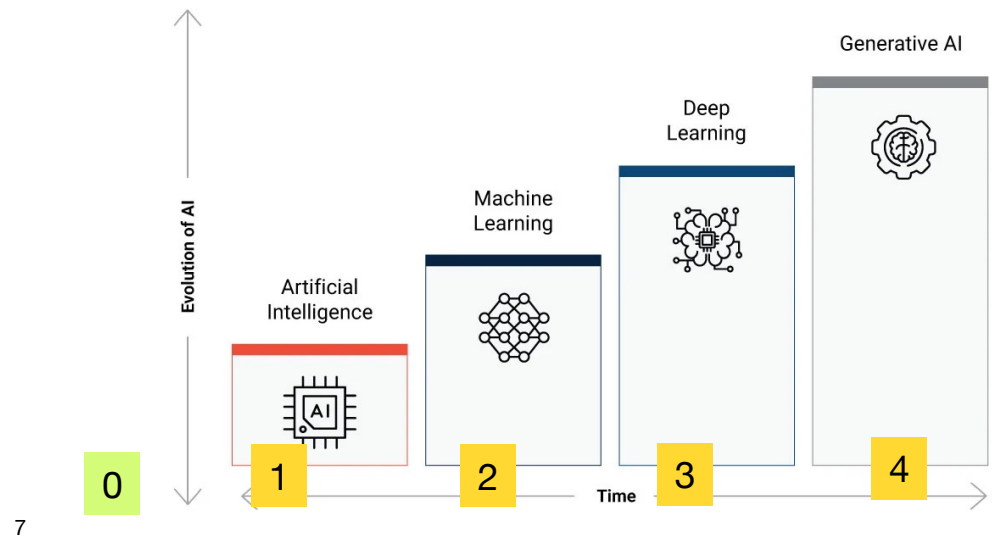
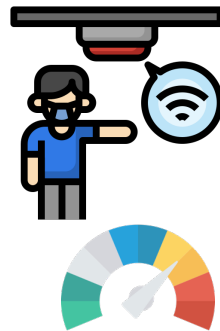
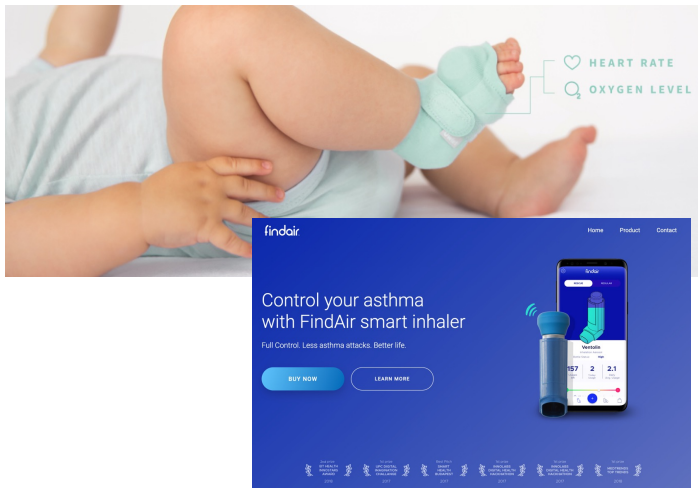


<https://mc.ai/machine-learning-basics-artificial-intelligence-machine-learning-and-deep-learning/>

# Not AI Solution

## Not automatic, not prediction

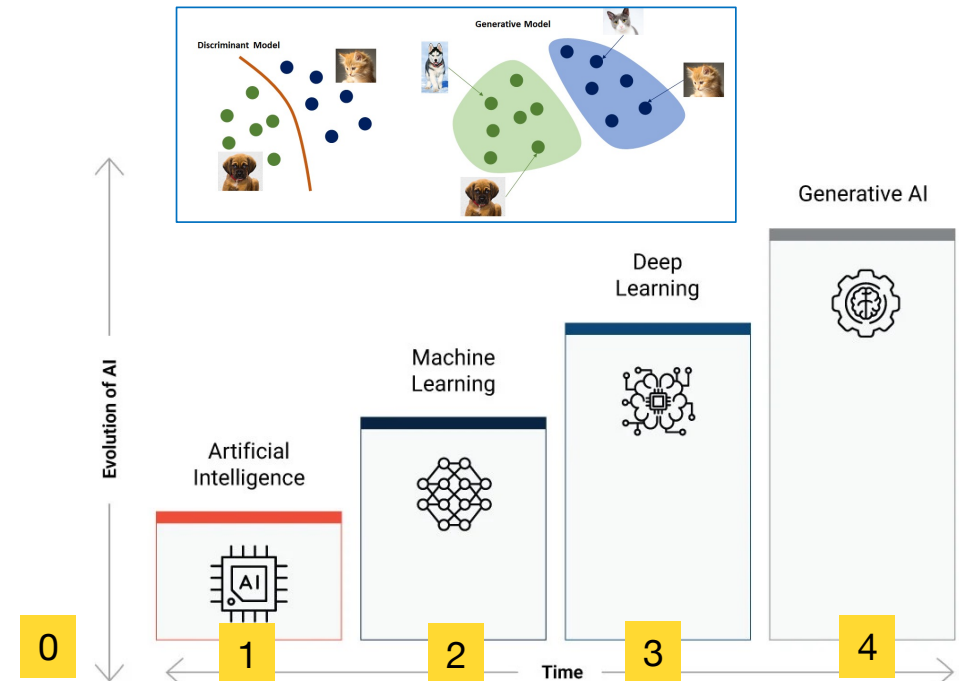
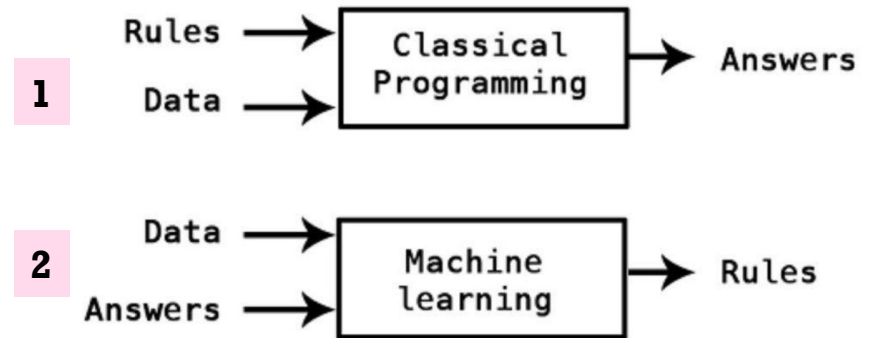
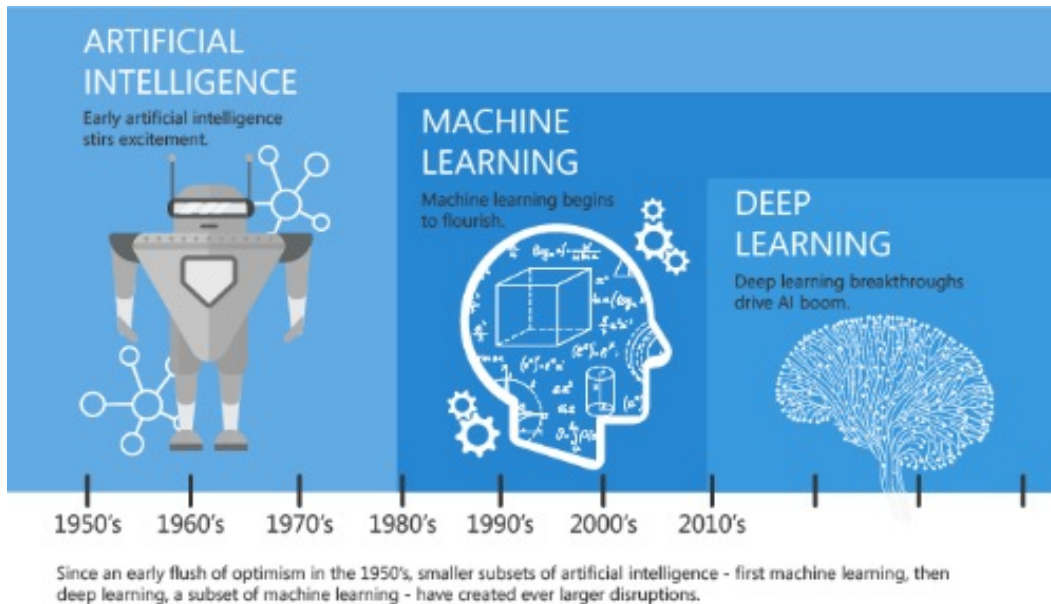
- 1) Information System (e.g., HIS, Web Application)
- 2) Transform data into valuable insights (sensor)
- 3) Transform data into interesting stories





# AI = Automation

- 0) Not AI Solution (not automatic)
- 1) Rule-based AI
- 2) Machine Learning (ML)



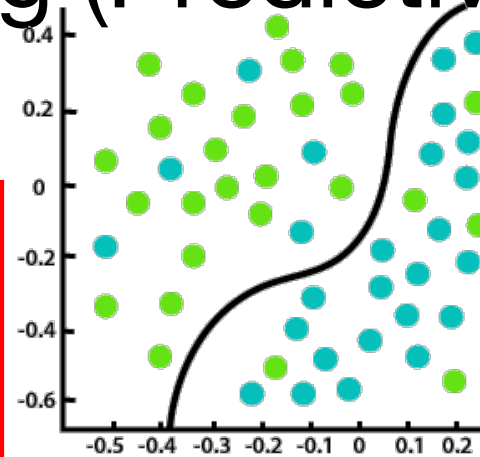
<https://mc.ai/machine-learning-basics-artificial-intelligence-machine-learning-and-deep-learning/>



# Supervised Learning (Predictive Task)

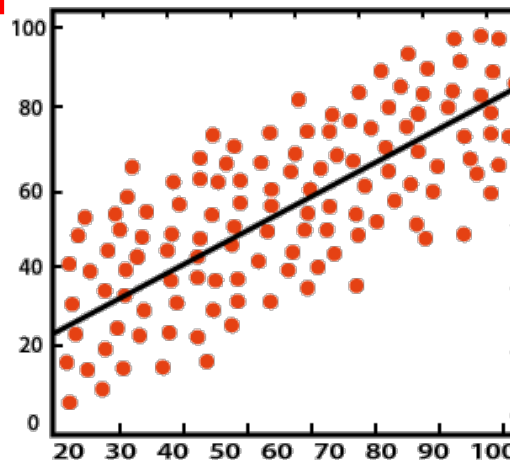
inputs				target
Age	Temp	Gender	Smell	Covid
25	39.0	Female	No	Yes
35	38.9	Female	No	Yes
32	36.5	Male	Yes	No

- Goal: To learn a prediction model mapping from inputs to output.
- Data without label (answer) is meaningless!
- Label should be provided by experts!



- Target is categorical variable.
- Example
- Covid diagnosis (yes/no)
- Disease diagnosis from gait information:
  - 1) Normal,
  - 2) Sick/Knee OA
  - 3) Sick/Parkinson

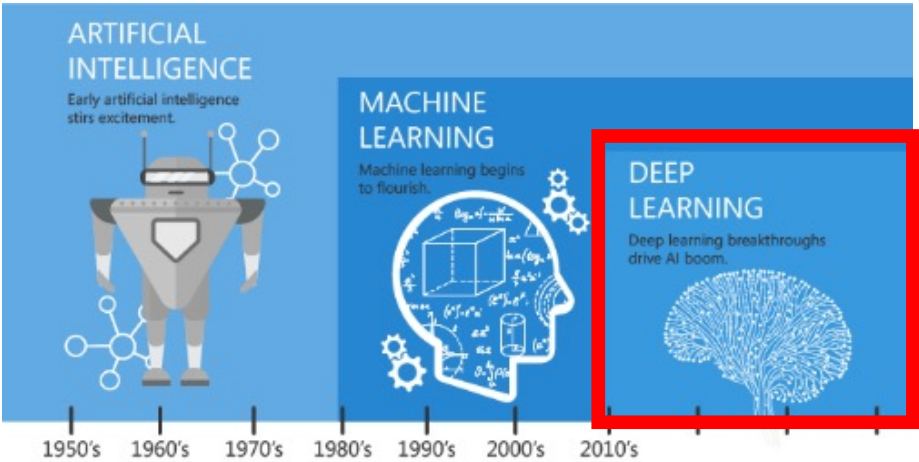
Classification



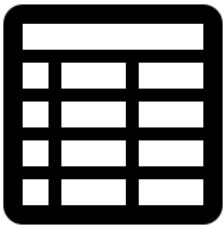
Regression

- Target is numeric variable.
- Example
- PD's state diagnosis from movement data.
- Glucose level prediction from breath particles.

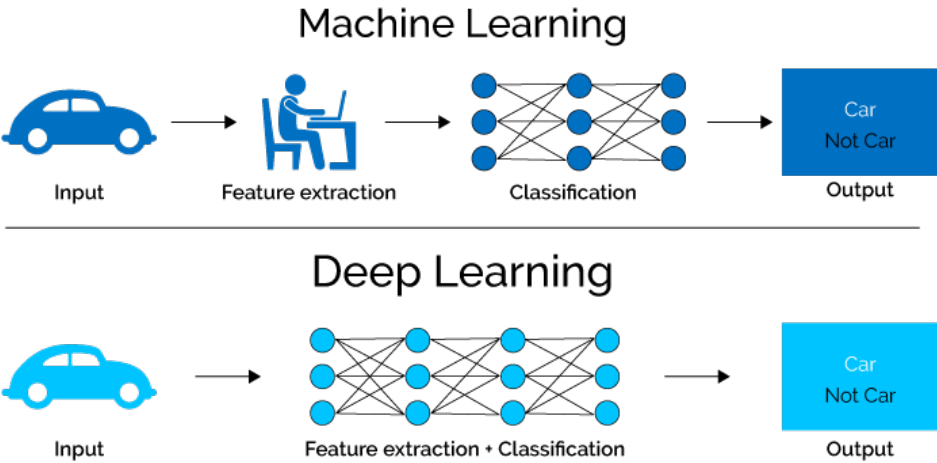
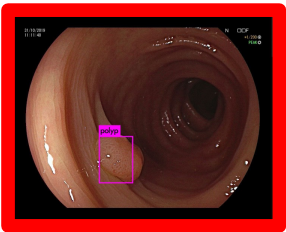
# Arise of Deep Learning



Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.



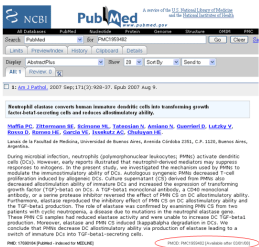
VS



Image, video



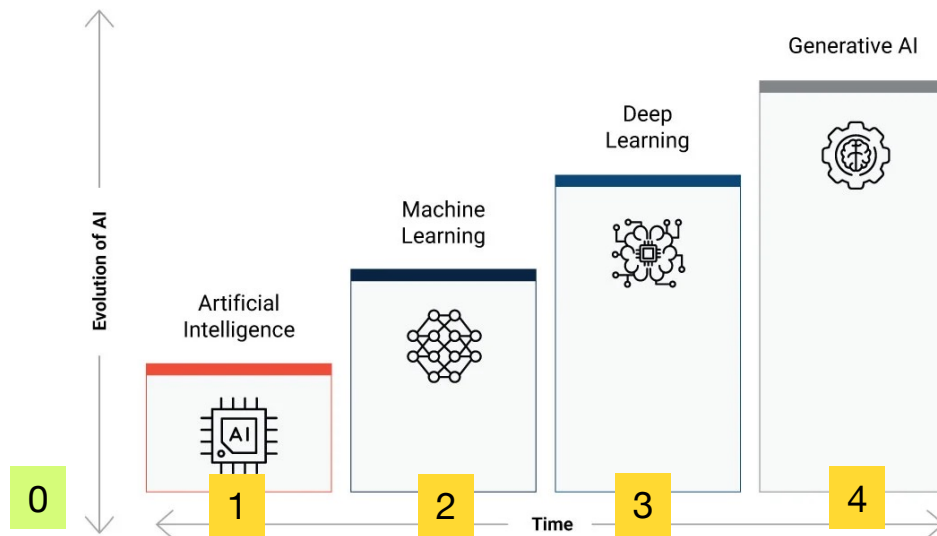
signals, voice



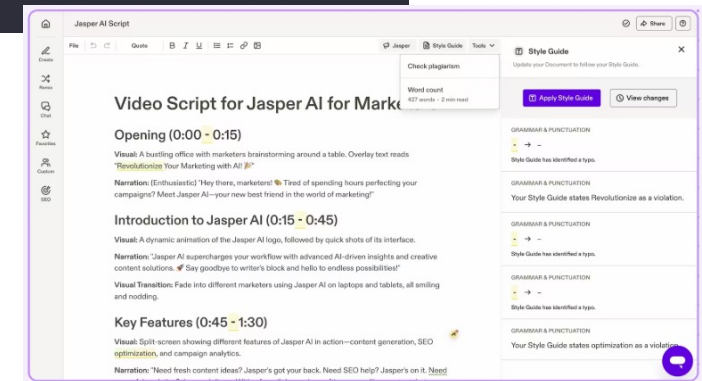
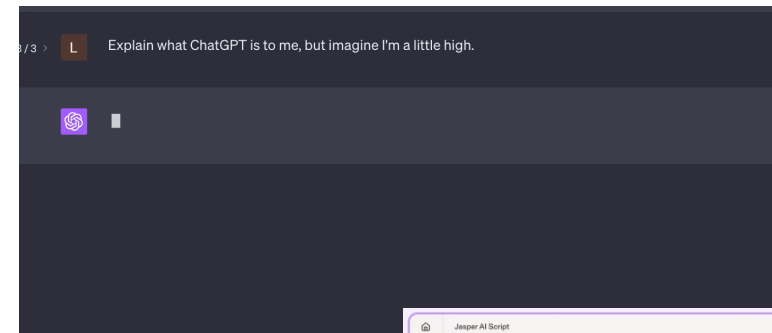
NLP

# Gen AI

Not automatic, not prediction



- 1) Assistant Chatbot
- 2) Text Summarization
- 3) Content Creation
- 4) Data Analytics
- 5) Etc.



# How to implement these AI models?

## Top Python Libraries for Machine Learning



TensorFlow



Keras



Hugging Face  
Transformers



NumPy



PyTorch



XGBoost



Fastai



Matplotlib



Scikit-learn



LightGBM



Pandas

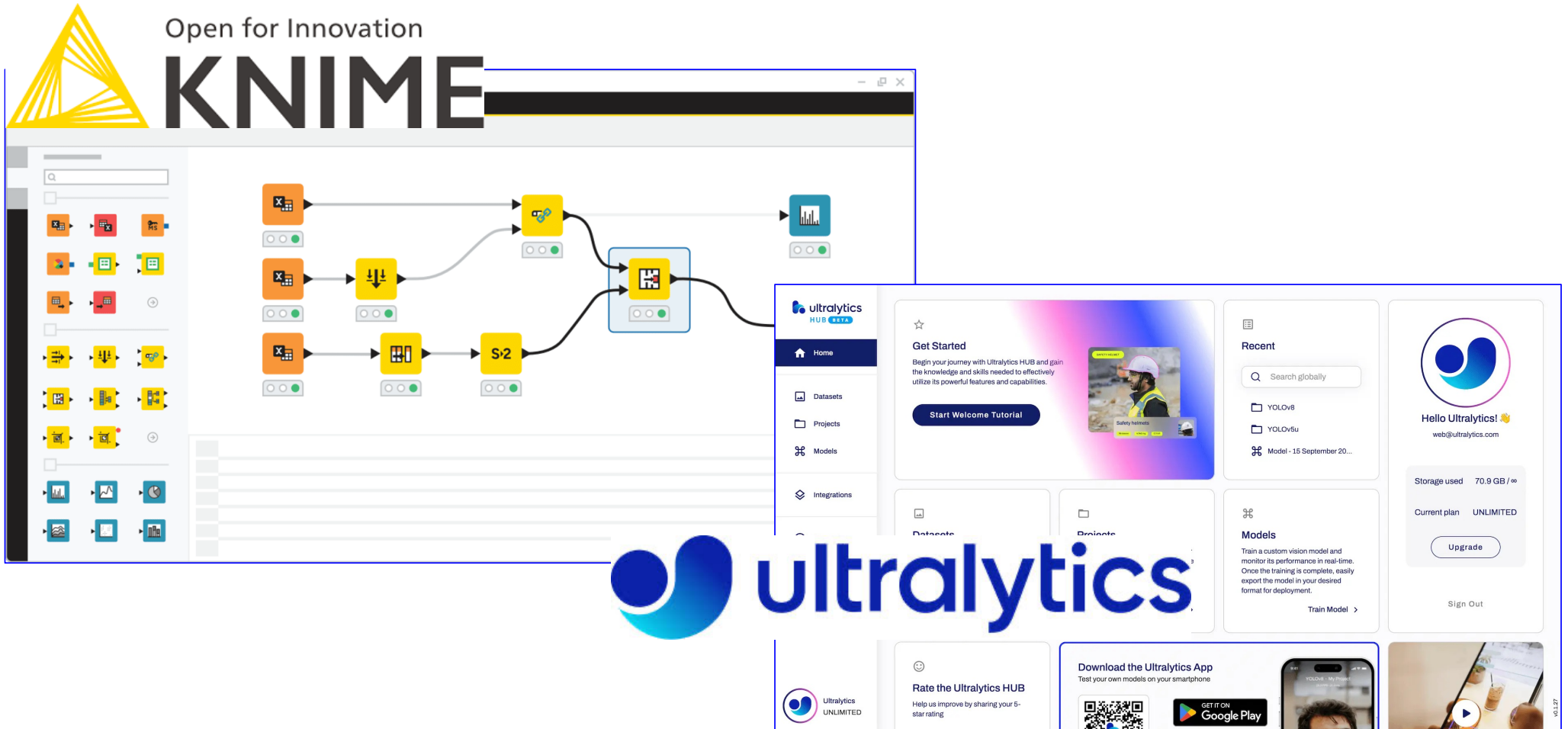


Seaborn



OpenCV

# Low-Code/No-Code Software



+ Thank you  
& any questions