Machine Learning for Satellite Imaging

Praveen Pankajakshan, PhD

Principal Data Science Manager, Corteva Agriscience

IEEE
HYDERABAD SECTION

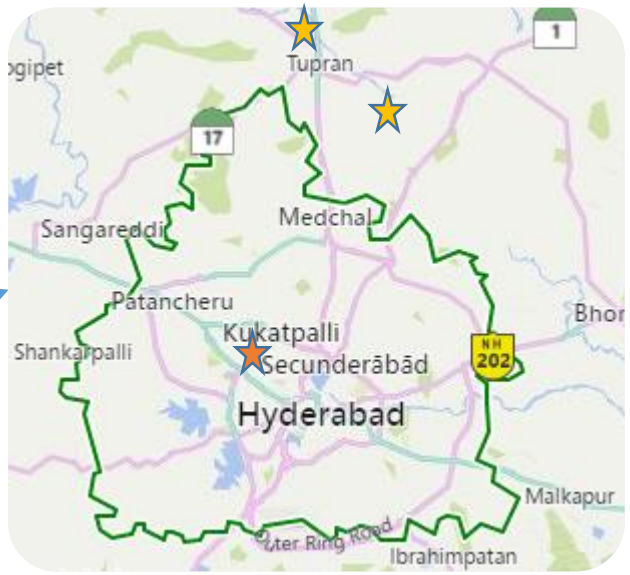
CORTEVA™
agriscience

Introductory Comments

- This presentation is for educational purposes only.
- Information provided is intended to initiate discussions
- This presentation makes considerable use of public sources of information. Every attempt is made to cite the sources.
- Any mention of a company, public-sector organization, individual, or technology is neither an endorsement nor criticism of the stated party or technology


Our Purpose
To enrich the lives of those who produce and those who consume, ensuring progress for generations to come.


SUSTAINABLE DEVELOPMENT GOALS
1 NO POVERTY
2 ZERO HUNGER
3 GOOD HEALTH AND WELL-BEING
4 QUALITY EDUCATION
5 GENDER EQUALITY
6 CLEAN WATER AND SANITATION
7 AFFORDABLE AND CLEAN ENERGY
8 DECENT WORK AND ECONOMIC GROWTH
9 INDUSTRY, INNOVATION AND INFRASTRUCTURE
10 REDUCED INEQUALITIES
11 SUSTAINABLE CITIES AND COMMUNITIES
12 RESPONSIBLE CONSUMPTION AND PRODUCTION
13 CLIMATE ACTION
14 LIFE BELOW WATER
15 LIFE ON LAND
16 PEACE, JUSTICE AND STRONG INSTITUTIONS
17 PARTNERSHIPS FOR THE GOALS
SUSTAINABLE DEVELOPMENT GOALS



Corporate

MCRC

TRC

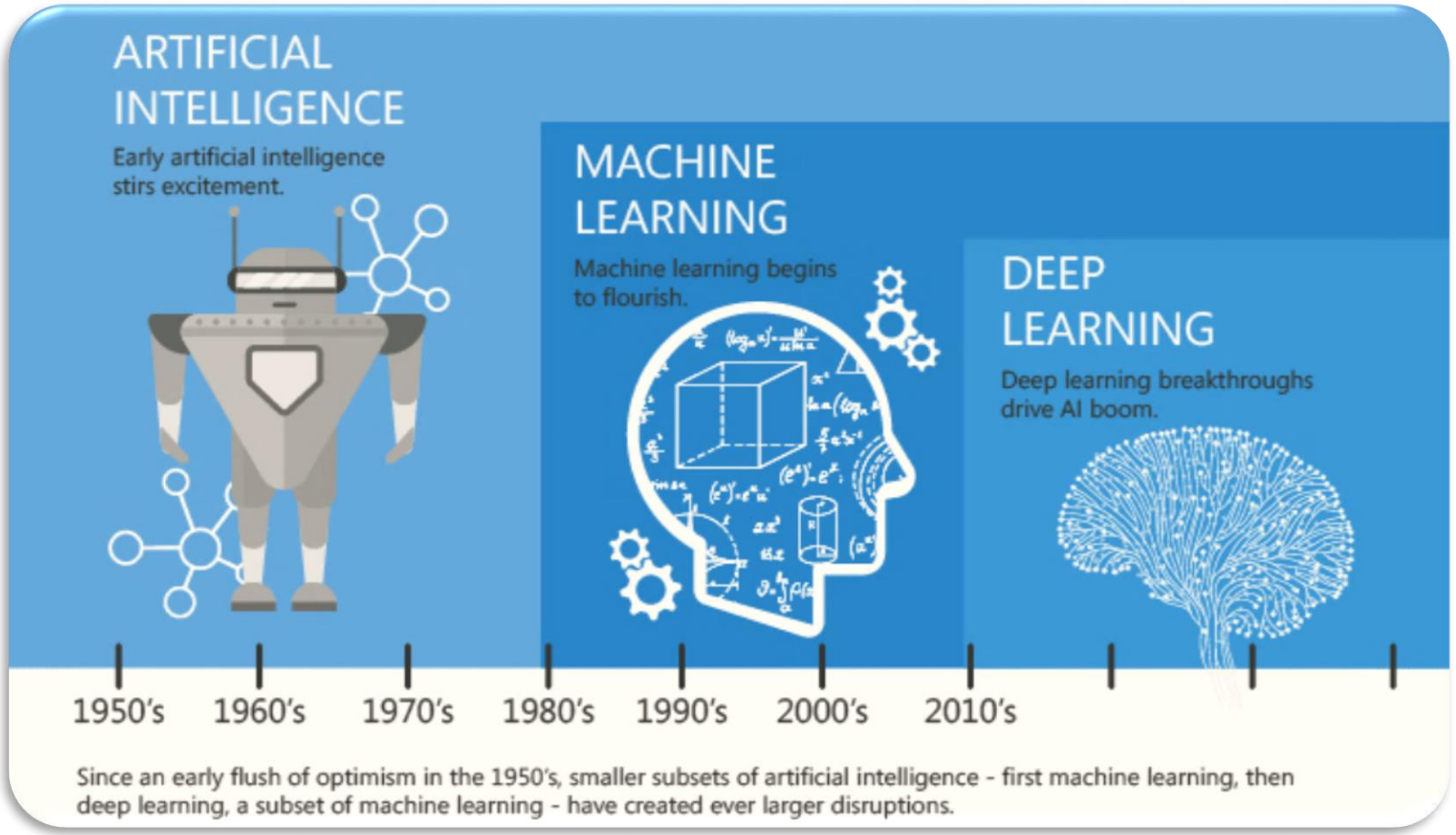
UAV

Satellite Imaging

IoT

Proximal Imaging

$$\text{Performance} \propto f(G, E, M)$$


ARTIFICIAL INTELLIGENCE
Early artificial intelligence stirs excitement.
MACHINE LEARNING
Machine learning begins to flourish.
DEEP LEARNING
Deep learning breakthroughs drive AI boom.
1950's
1960's
1970's
1980's
1990's
2000's
2010's
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.

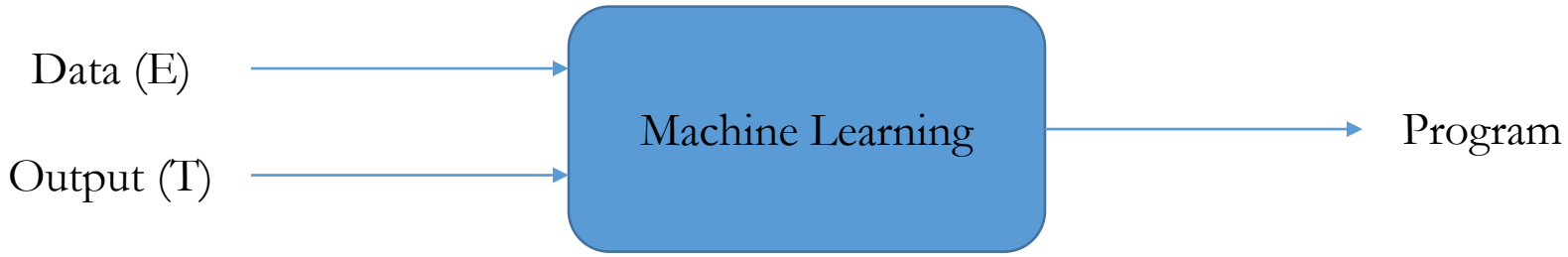
What is Machine Learning?

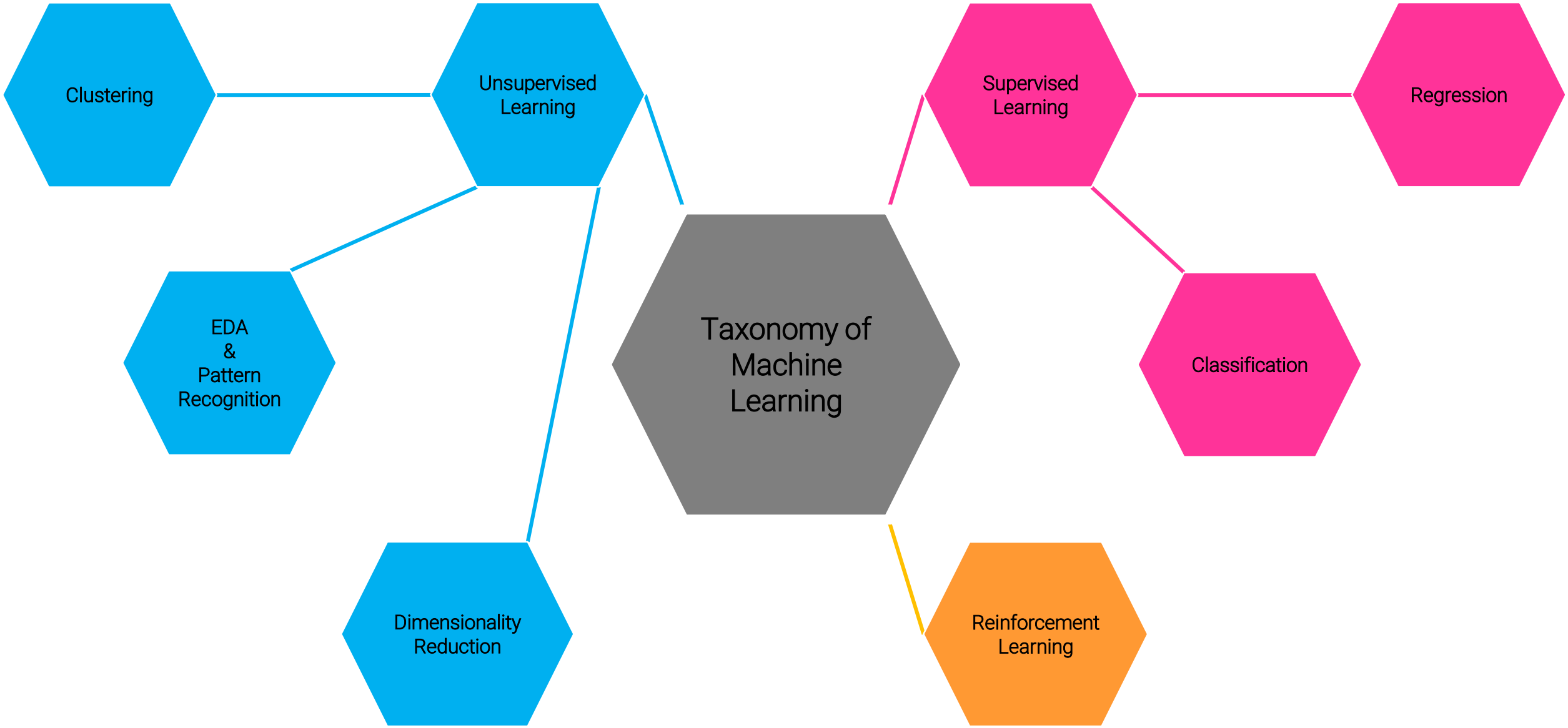
Definition

*A computer program is said to learn from experience **E** with respect to some class of tasks **T** and performance measure **P**, if its performance at tasks T, improves with experience E*

-Mitchell 1997


Data
Program
Traditional Programming
Output


Data (E)
Output (T)
Machine Learning
Program


Clustering
Unsupervised Learning
EDA & Pattern Recognition
Dimensionality Reduction
Taxonomy of Machine Learning
Supervised Learning
Regression
Classification
Reinforcement Learning

Prerequisites for this course

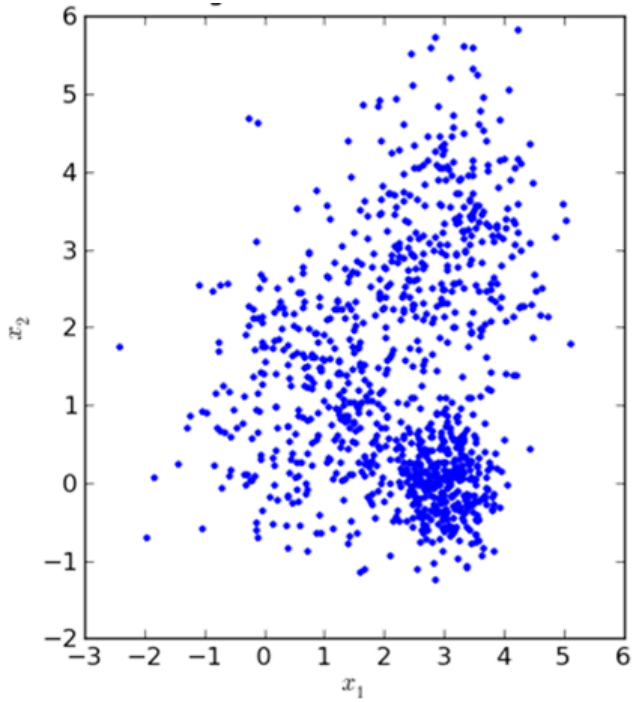
- Linear Algebra
- Basic of image processing
- Python programming
- Google Account

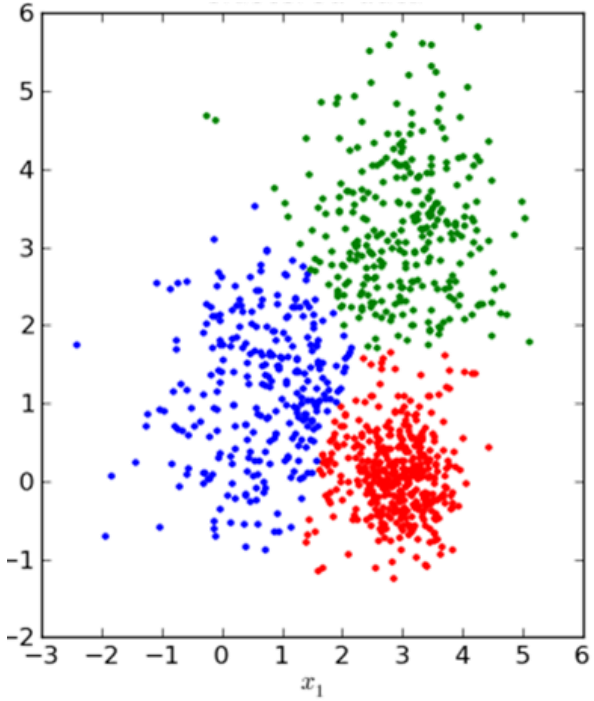
Dataset

http://tiny.cc/91oxhz

Unsupervised Learning

- Dataset (**E**): $\mathrm{X} = (x_1, x_2, \ldots, x_n)$
- Task (**T**): To find the interesting structures in the data (No lables)





K-means

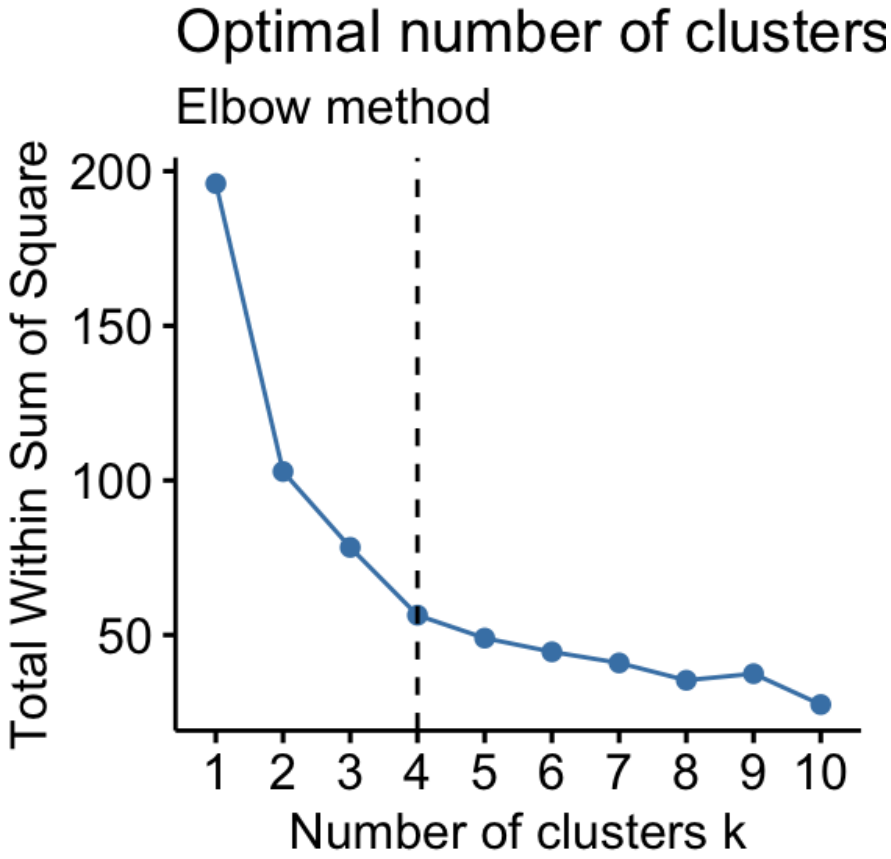
- Given a set of points: X: $(x_1, x_2, \ldots, x_n)$
- Objective function to minimize: $\min_{\{l_c\}_{c=1}^{k}} \sum_{c=1}^{k} \sum_{x \in l_c} \|x - \mu_c\|_2^2$, where $\mu_c = \frac{1}{|l_c|} \sum_{x \in l_c} x$

1. Choose the number of clusters: *k*
2. Assign X to the *k* clusters
3. M: Calculate the mean of each cluster and assign cluster centers to it
4. E: Reassign X to the clusters based on Euclidean distances

$$l^*(i) = argmin_c \, \|x_i - \mu_c\|_2^2$$

B. Kulis, *et al.* "*Revisiting k-means: New Algorithms via Bayesian Nonparametrics*", ICML, 2012.
https://icml.cc/2012/papers/291.pdf

Optimal Cluster Number



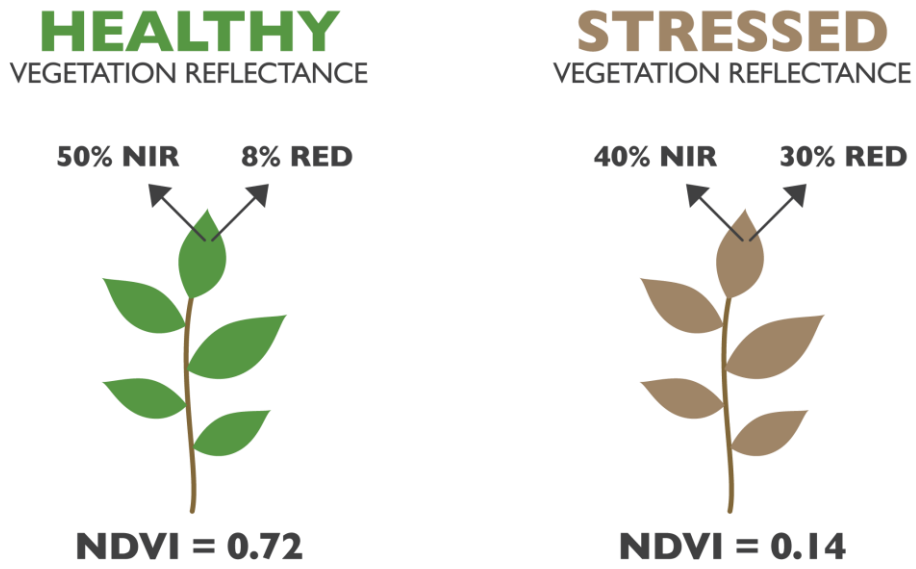
Supervised Learning

Example Object Identification

- X: Pixel values
- **y**: Bounding box

Feature Engineering

- Example: Normalized difference vegetation index (NDVI)

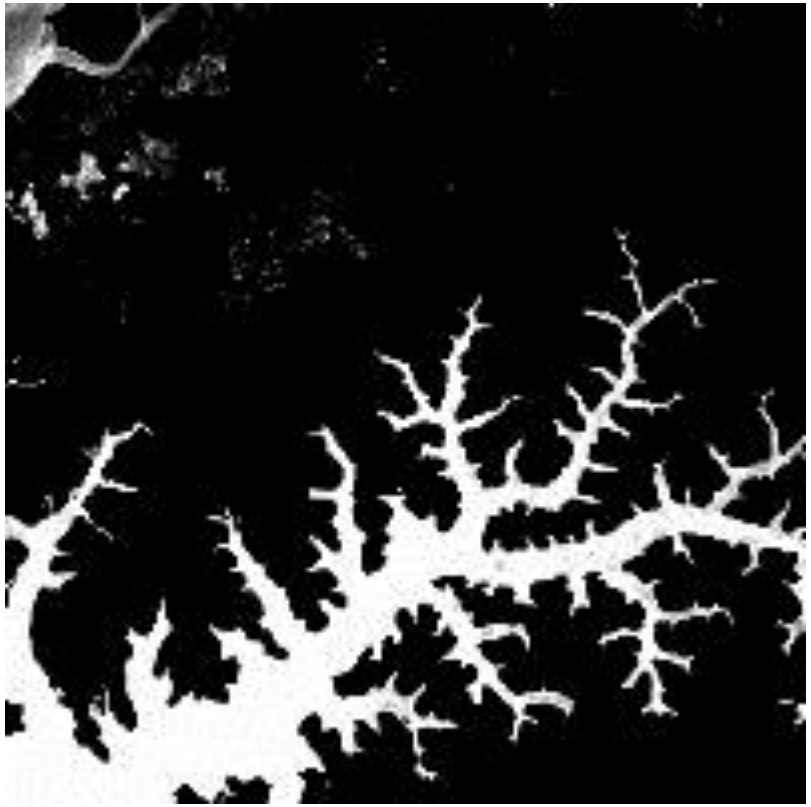


$$NDVI = \frac{NIR - RED}{NIR + RED}$$

Feature Engineering

- Example: Normalized difference Water index (NDWI)

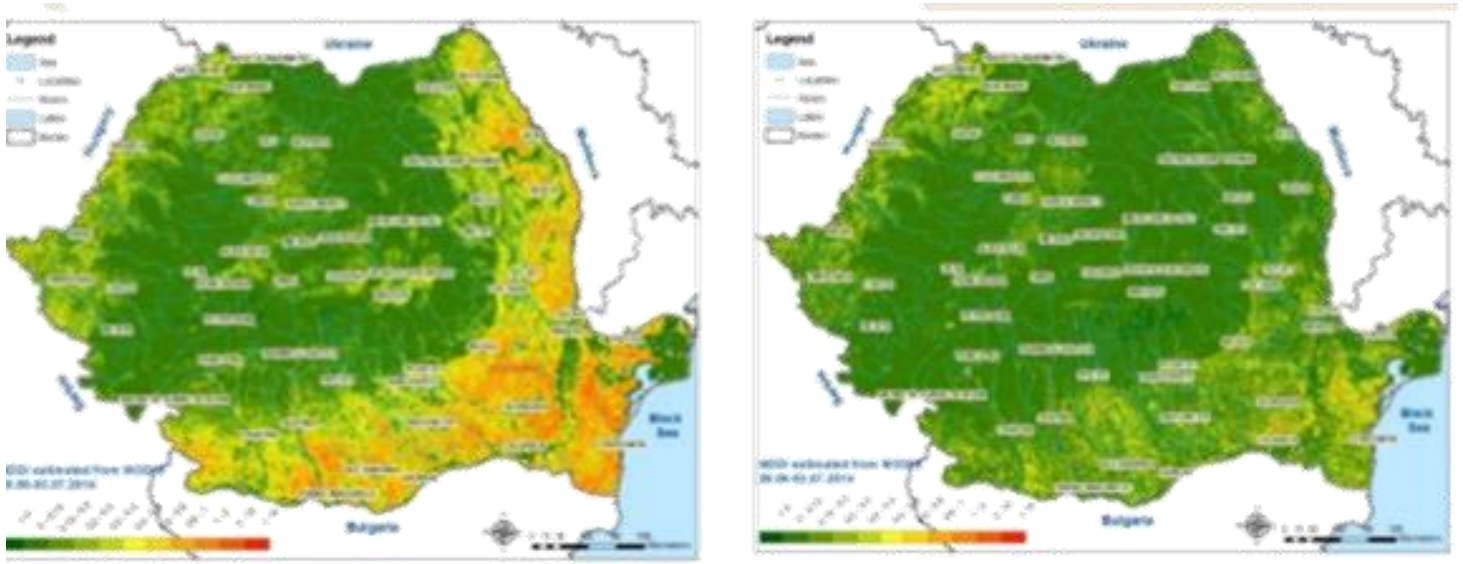
$$NDWI = \frac{B - NIR}{B + NIR}$$



Feature Engineering

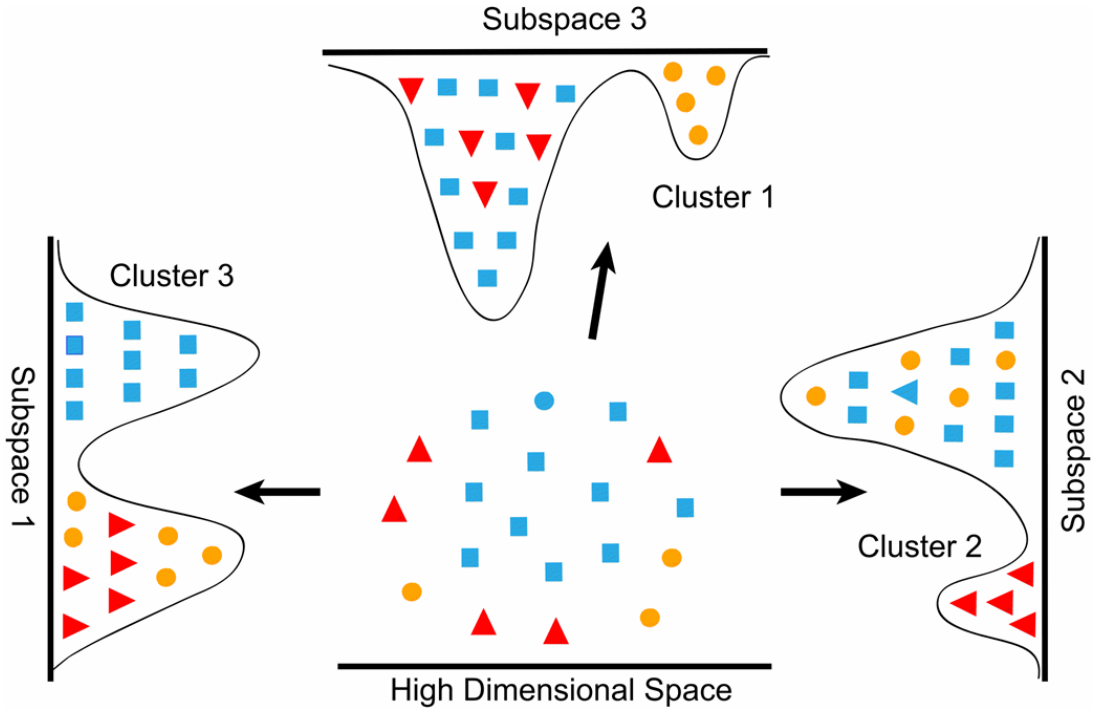
- Example: Normalized Difference Drought Index (NDDI)

$$NDDI = \frac{NDVI - NDWI}{NDVI + NDWI}$$



Dimensionality Reduction

- Reducing the dimensions of the feature space is called Dimensionality Reduction
  - Feature Elimination
  - Feature Extraction



Steps for PCA

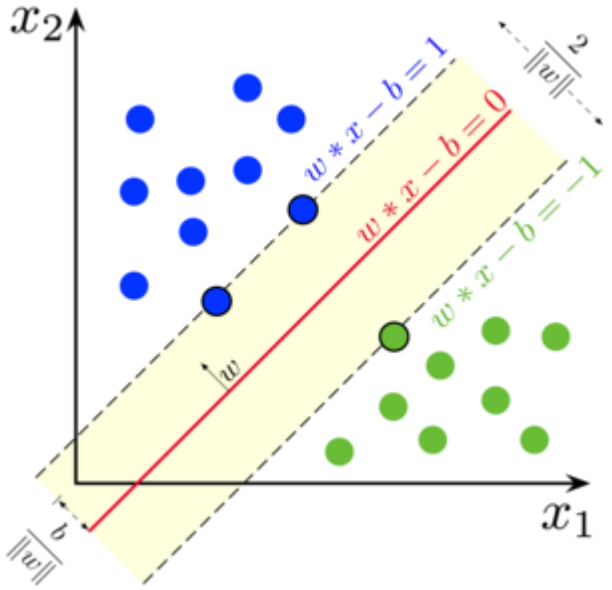
- Find the sample mean

$$\boldsymbol{\mu} = \frac{1}{N}\sum_{i} \boldsymbol{x}_i$$

- Compute the Covariance $C = \frac{1}{N} XX^T = \frac{1}{N}\sum_{i} (\boldsymbol{x}_i - \boldsymbol{\mu})(\boldsymbol{x}_i - \boldsymbol{\mu})^T$
- Find the Eigen values and vectors for C $VC = \Lambda C$
- Arrange them and transform $Y = V^T X$

Supervised Learning

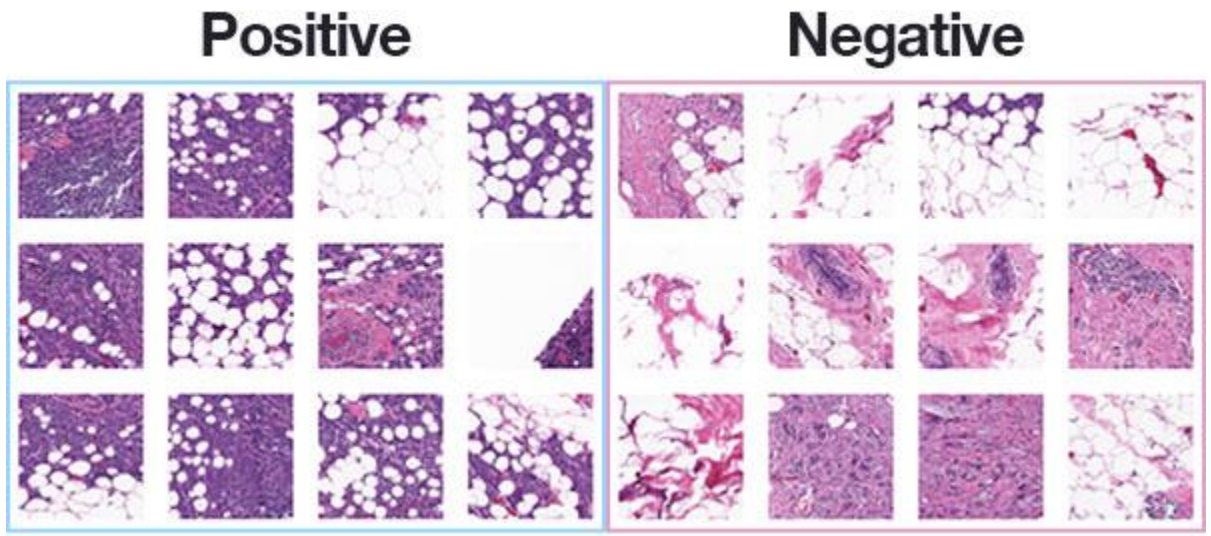
- X: input (E) $\in \Re^M \times \Re^N$
- **y**: Output (T) $\in \Re^M$
- Regression: when Y is a continuous variable
- Classification: when Y is discrete



Supervised Learning

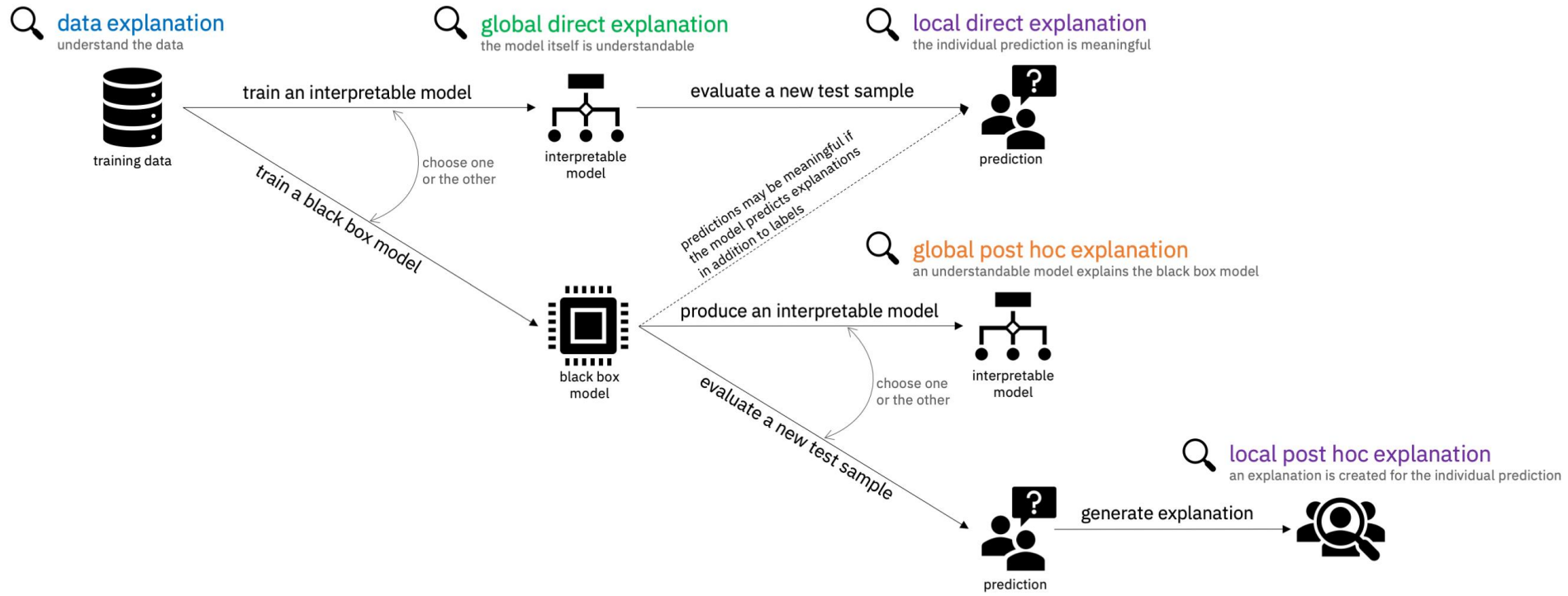
Example Image Classification

- X: Pixel values
- **y**: Class that the images belong to



Human-Centered AI

Explainable Model and Trust

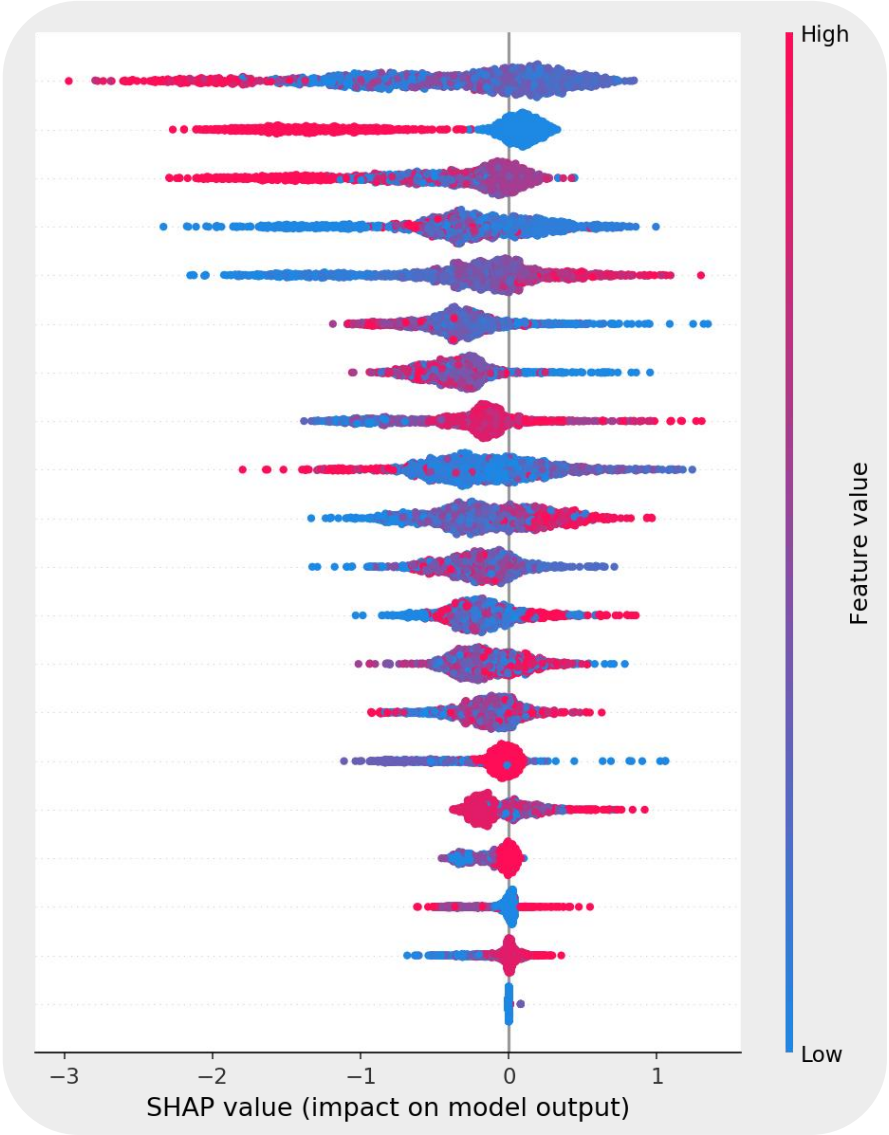


Direct Explainability

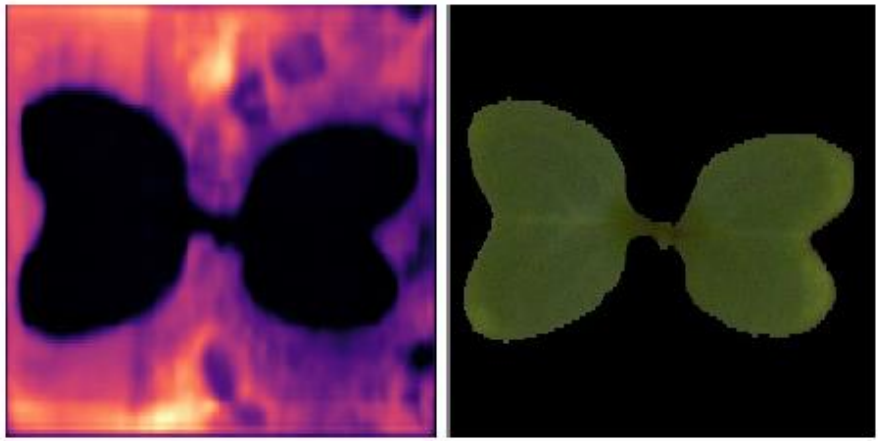
Feature 1

.
.
.
.
.
.
.
.
.
.

Feature 20



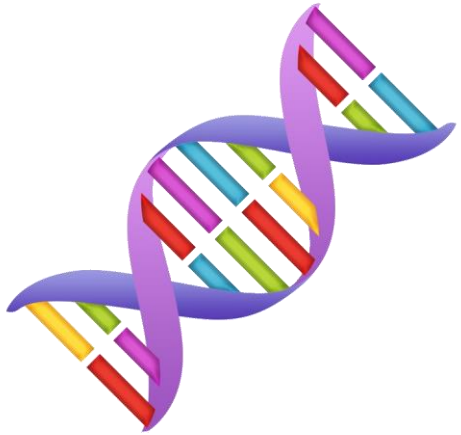
Post-hoc



(a) Checker board pattern

(b) Output activation as a mask on the

Technology



Science

Collaboration

Thank You!

Acknowledgements

- Parmita Ghosh, Sonal Bakiwala, Chaitanyam Potnuru, Harshit Lohani, Vivek Singh Bhadouria, Nitin Kandpal
- Jochen Scheel, Siva Prasad Kumpatla, Anu Swatantran, John Van Hemert, Brent Myers, Manny Ruidiaz, Jeremiah Barr
- Data Science and Informatics members