

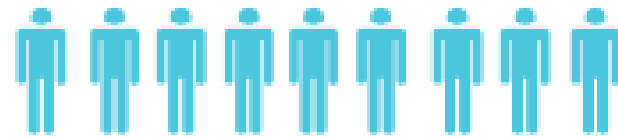


Data Analytics in Precision Agriculture

SARITH DIVAKAR M | [GITHUB.COM/SARITHDM](https://github.com/SARITHDM)

Population growth and Urbanization

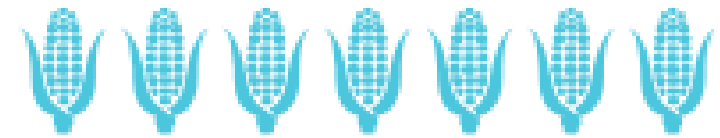
POPULATION
GROWTH =
HIGHER DEMAND
FOR FOOD



10 billion

world population in 2050

=



70%

More food to be produced by farmers

URBANIZATION
DRIVES CHANGE IN
CONSUMPTION
PATTERN



36.4 kg

processed food and meat annual
per capita meat consumption
1997-1999

→



45.3 kg

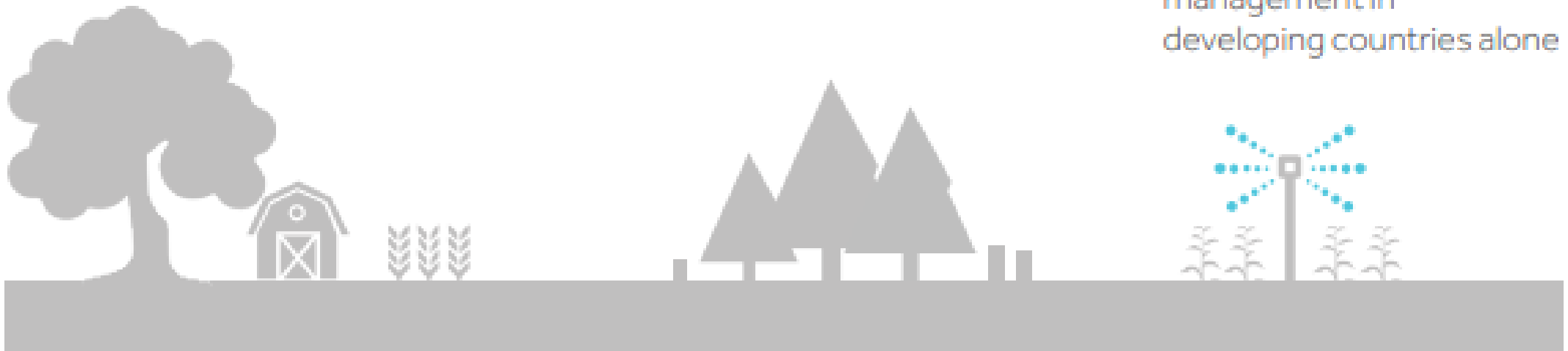
processed food and meat annual
per capita meat consumption
2030

Natural Resources

25% of all farmland is already rated as highly degraded

~80% global deforestation driven by agricultural concerns

\$1 trillion investment necessary until 2050 for irrigation water management in developing countries alone

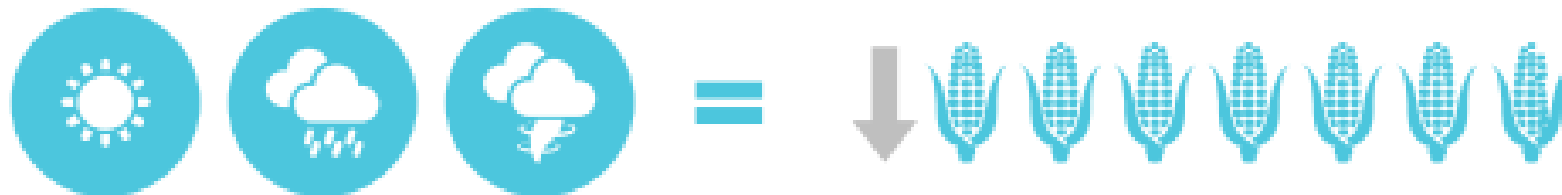


Climate Change

GREENHOUSE GAS EMISSIONS

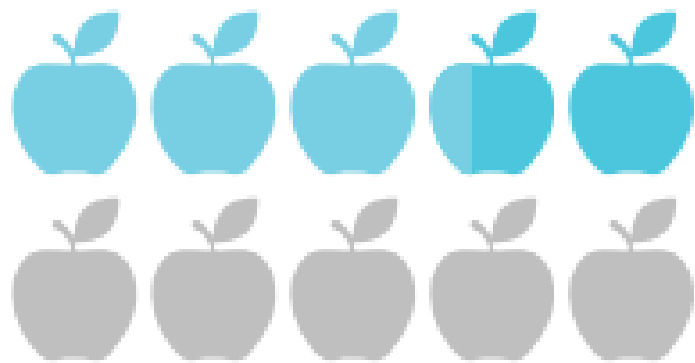


VARIABILITY OF PRECIPITATION REDUCE CROP YIELDS



Rise in the frequency of droughts and floods, all of which tend to reduce crop yields

Food Waste



between

33%-50% =

of all food produced
globally is never eaten



25%

of all fresh water
consumption globally

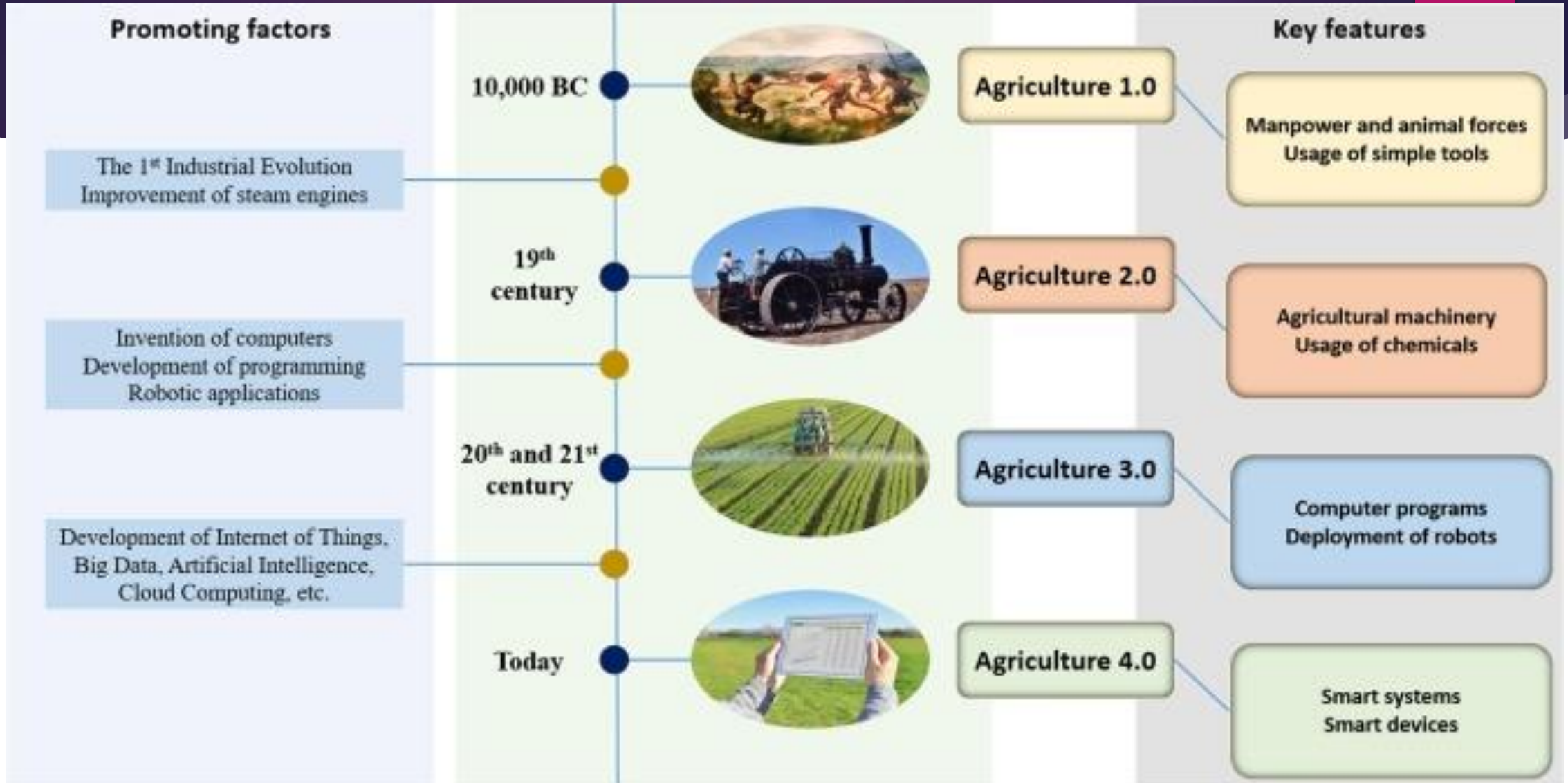
3rd

largest emitter of greenhouse gases after **China** and the **US**, if food waste were a country

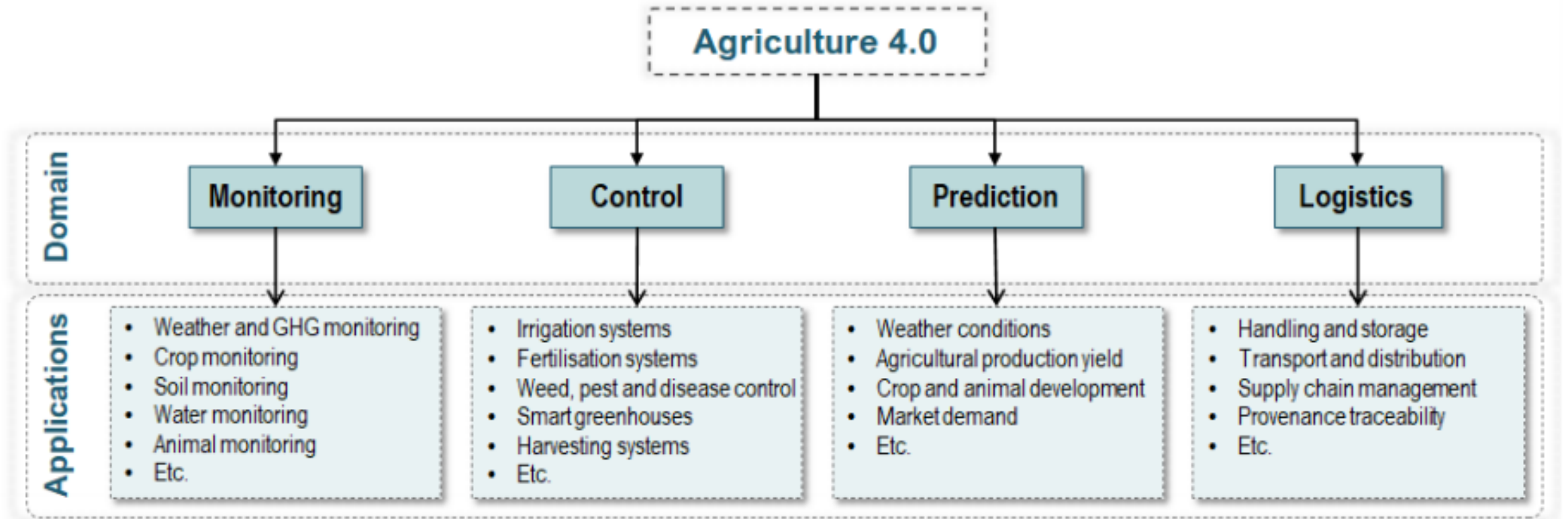
Solution: Agriculture 4.0



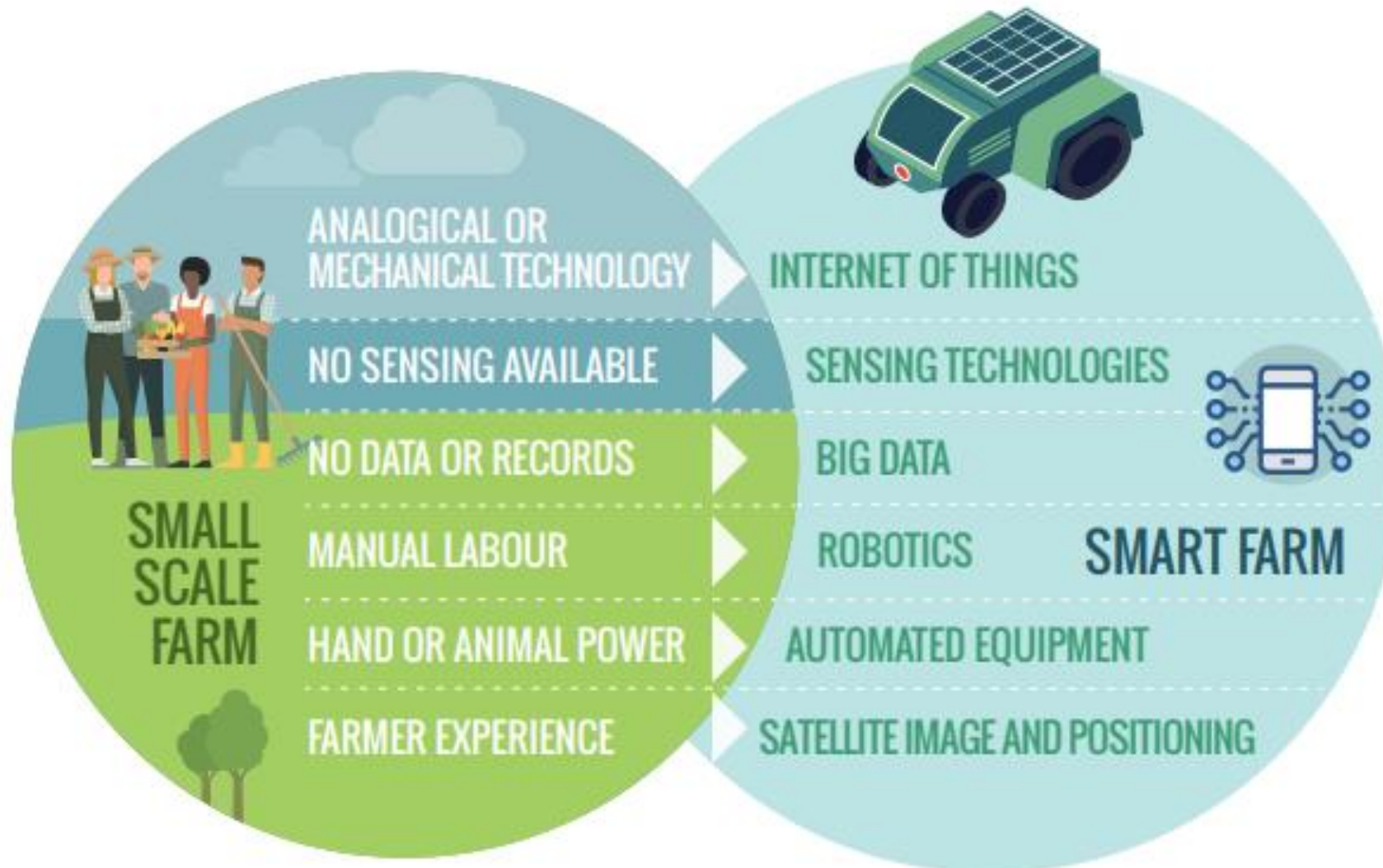
Agriculture 4.0

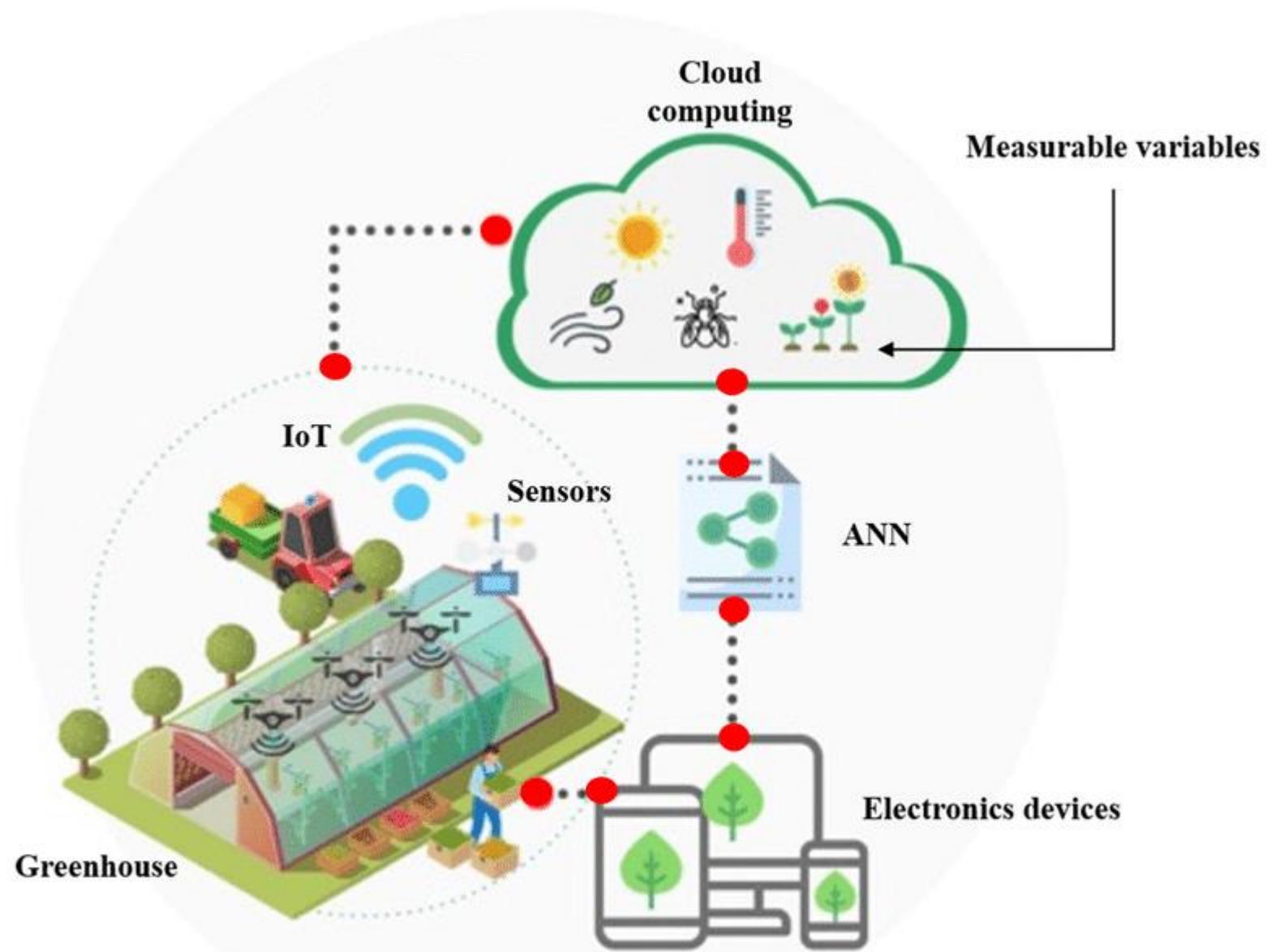


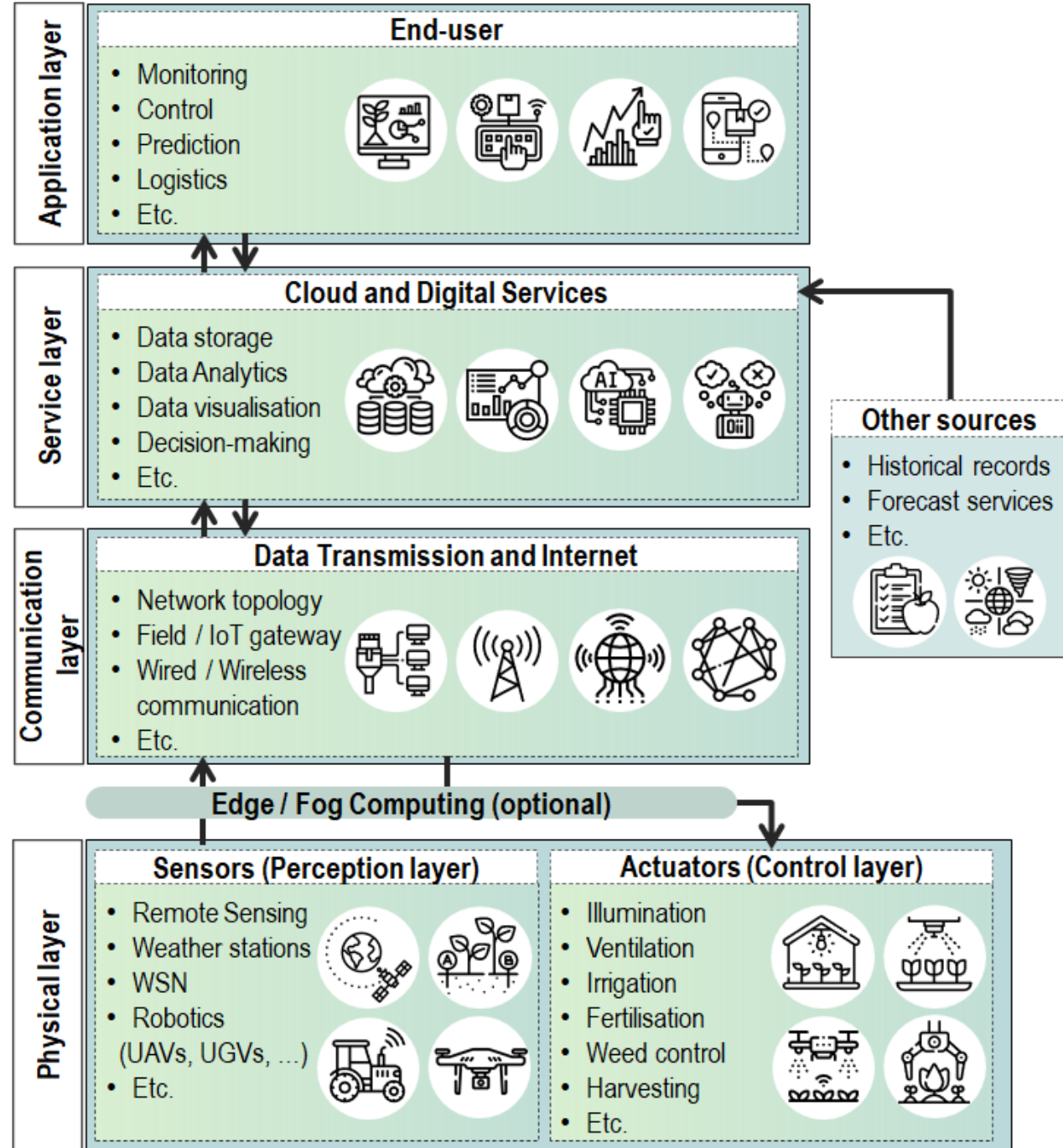
Applications

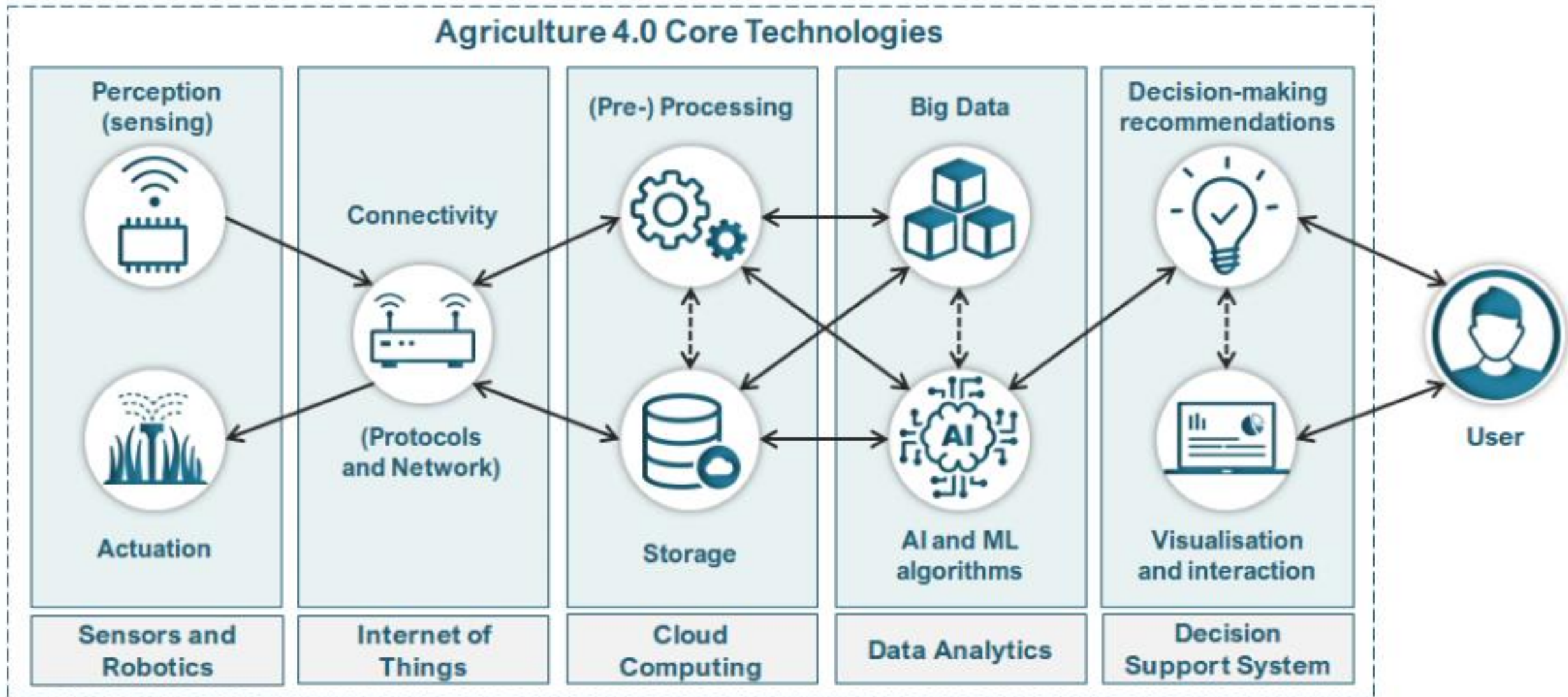


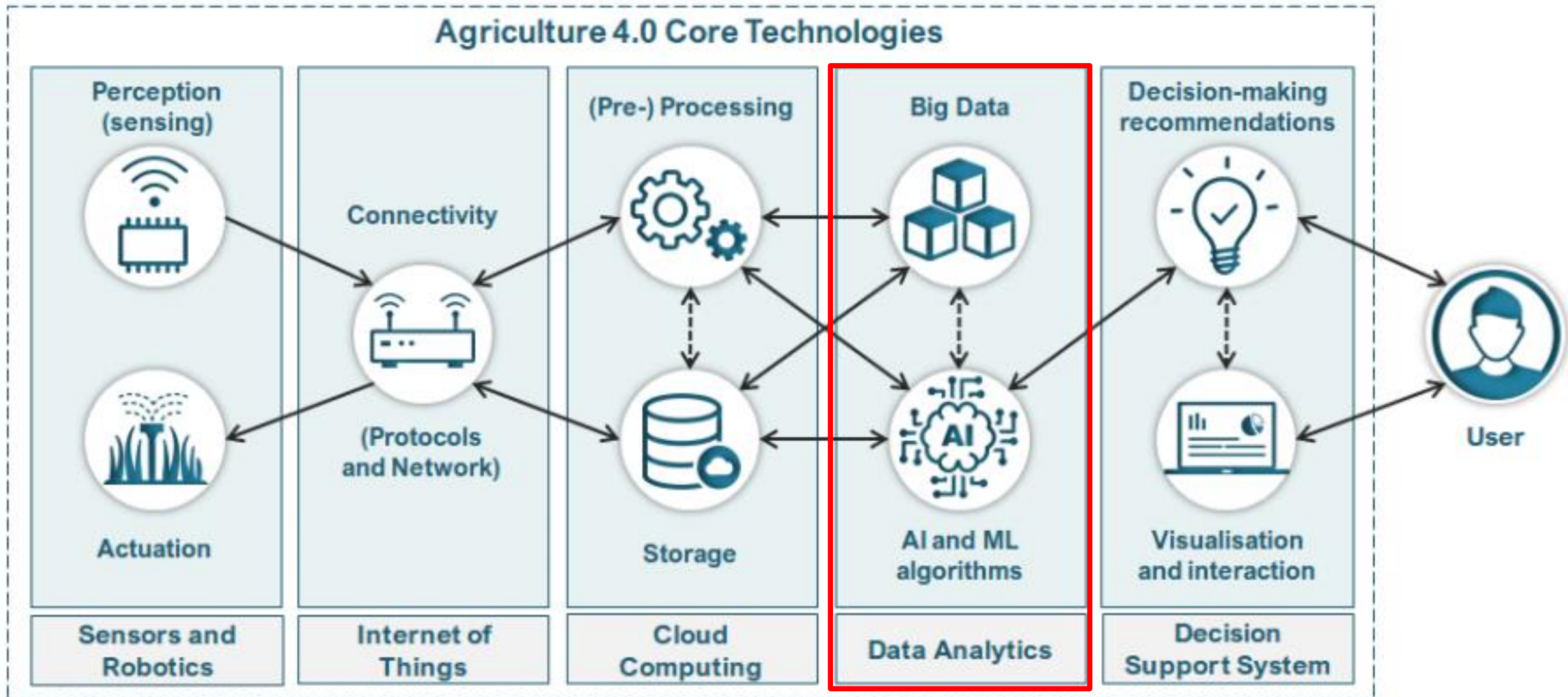
Precision Agriculture/Smart Farming











AI/ML Algorithms

An algorithmic way of **making sense** (learning) from data.

Learning means Improving with Experience at some Task

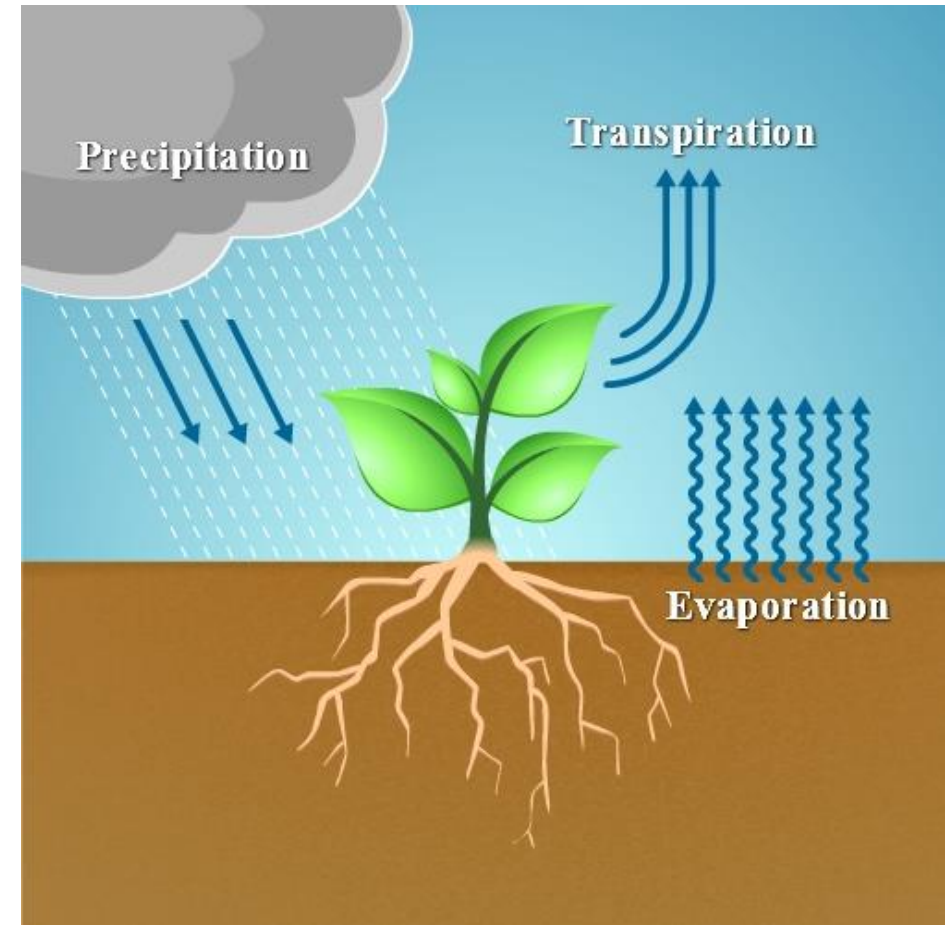
Tom M. Mitchell: A computer program is said to learn from Experience E with respect to some Task T and some Performance measure P, if its performance on T, as measured by P, improves with experience E.

T: Estimate Evapotranspiration rate

E: Learning from Evapotranspiration dataset

P: Accuracy of the Evapotranspiration rate

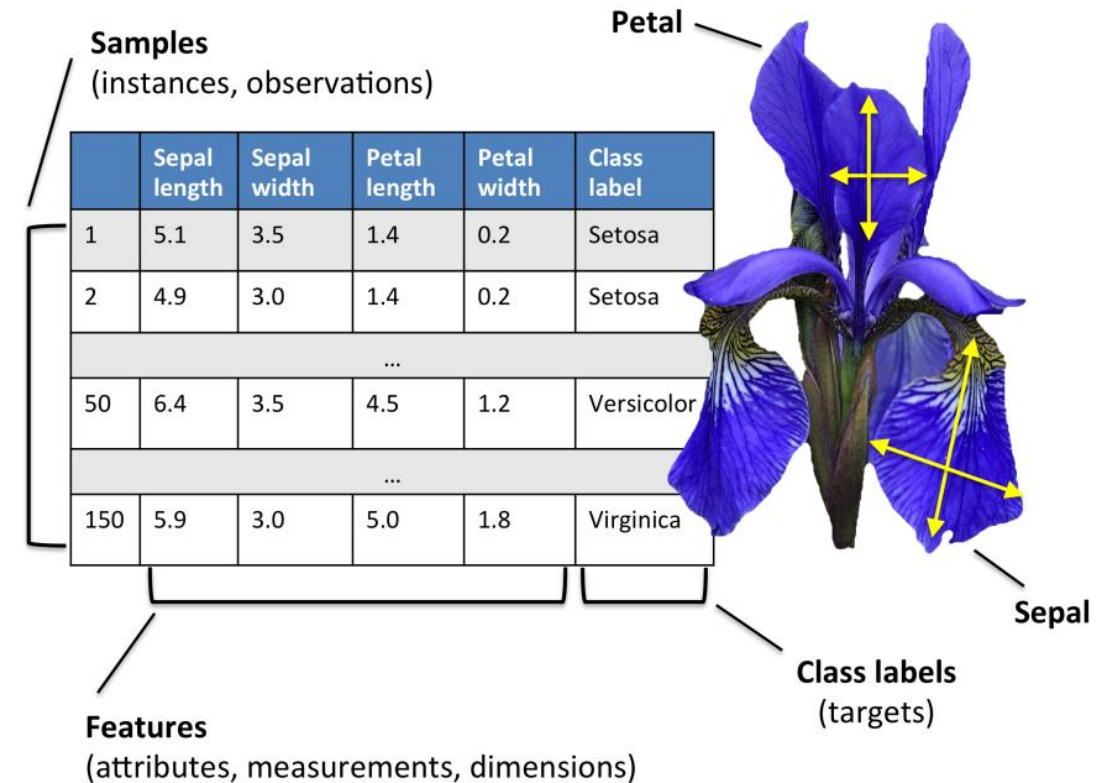
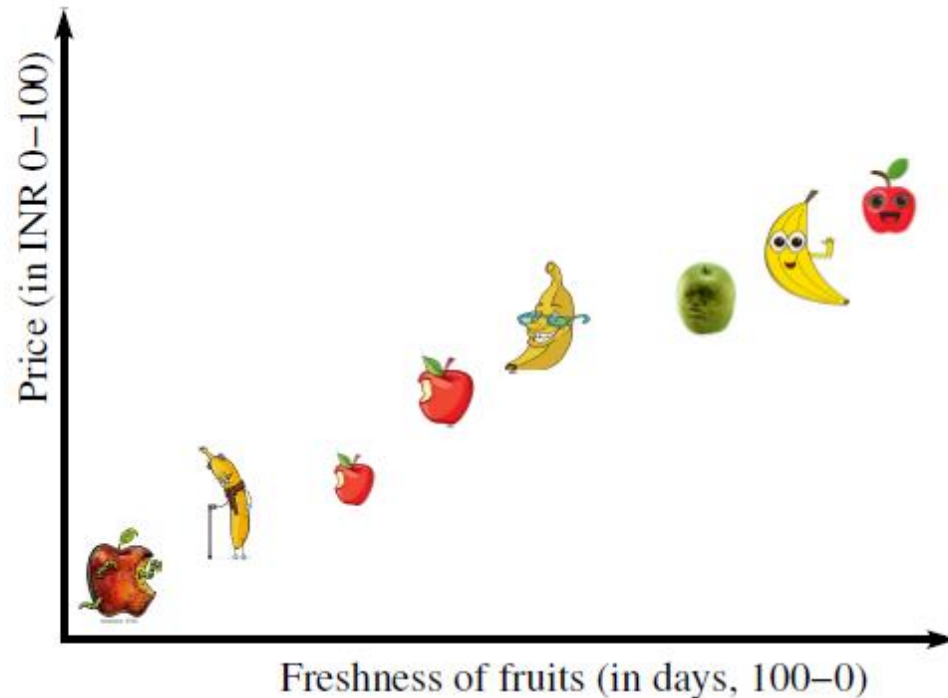
Decision: Irrigation Scheduling



Predictive Analytics/ Supervised learning

Build an analytical model predicting a target measure of interest

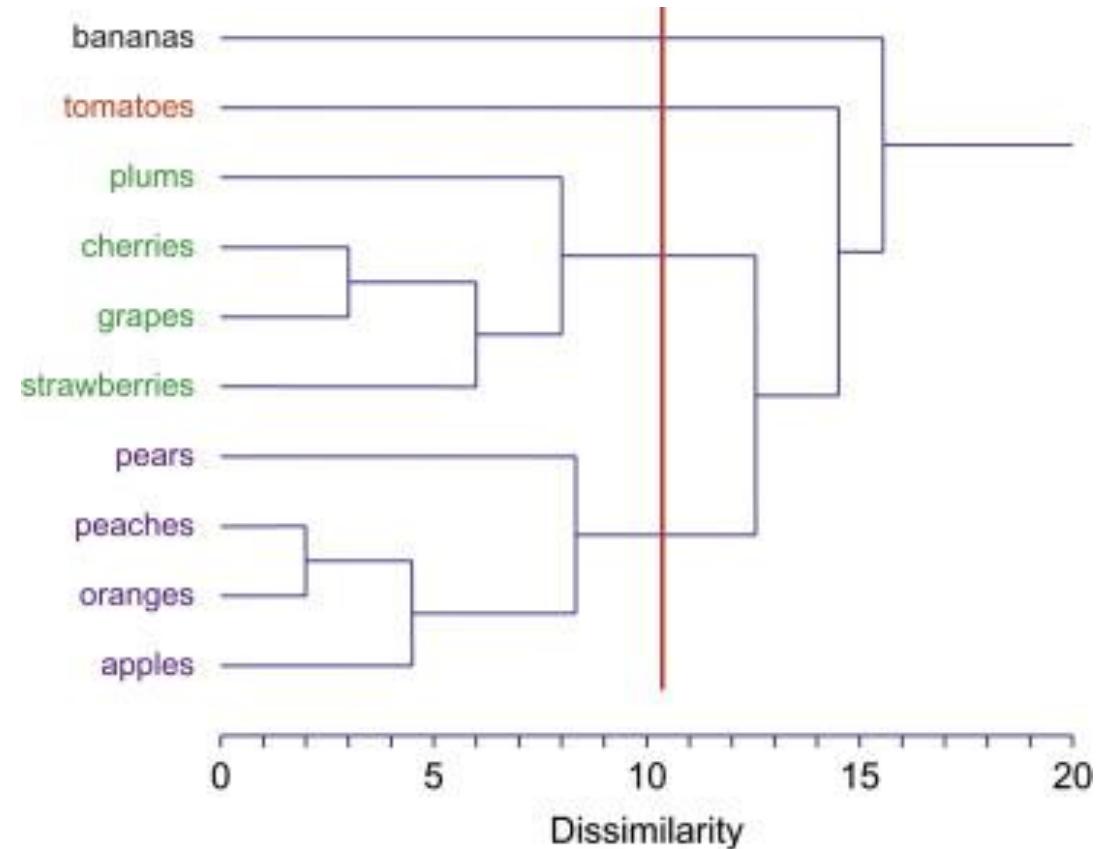
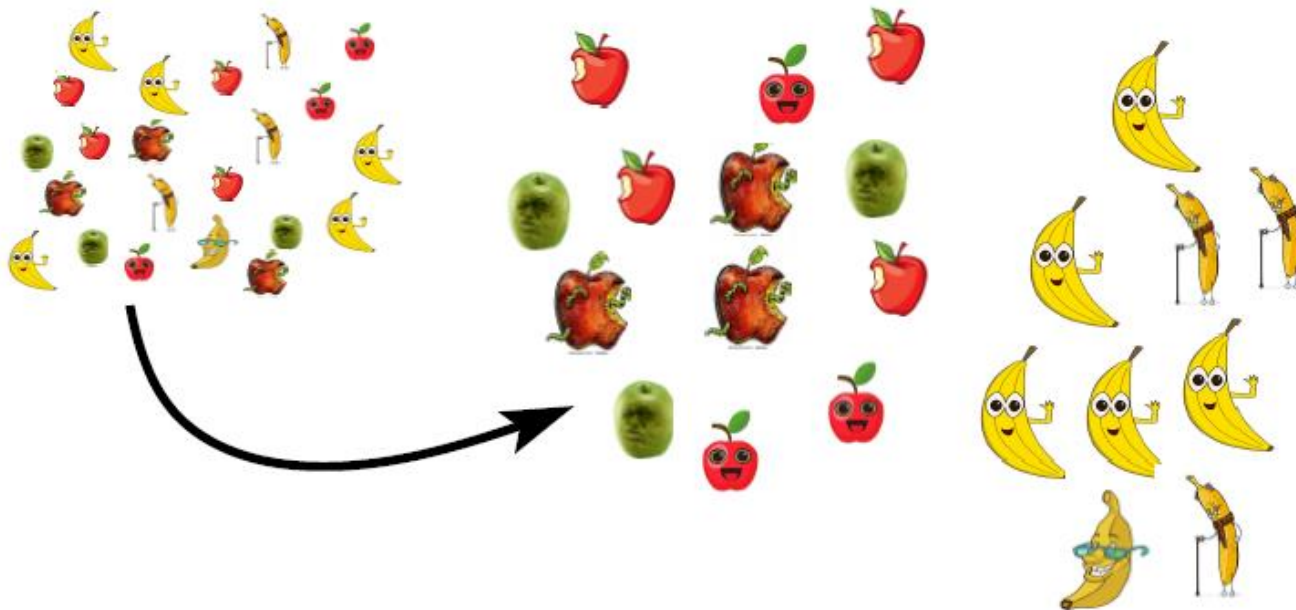
Example: Predict Price based on freshness



Descriptive Analytics/Unsupervised learning

Describe patterns from the data

Example: Group fruits into two groups



Deep Learning

Neural Networks are considered universal function approximators

They can compute and learn any function

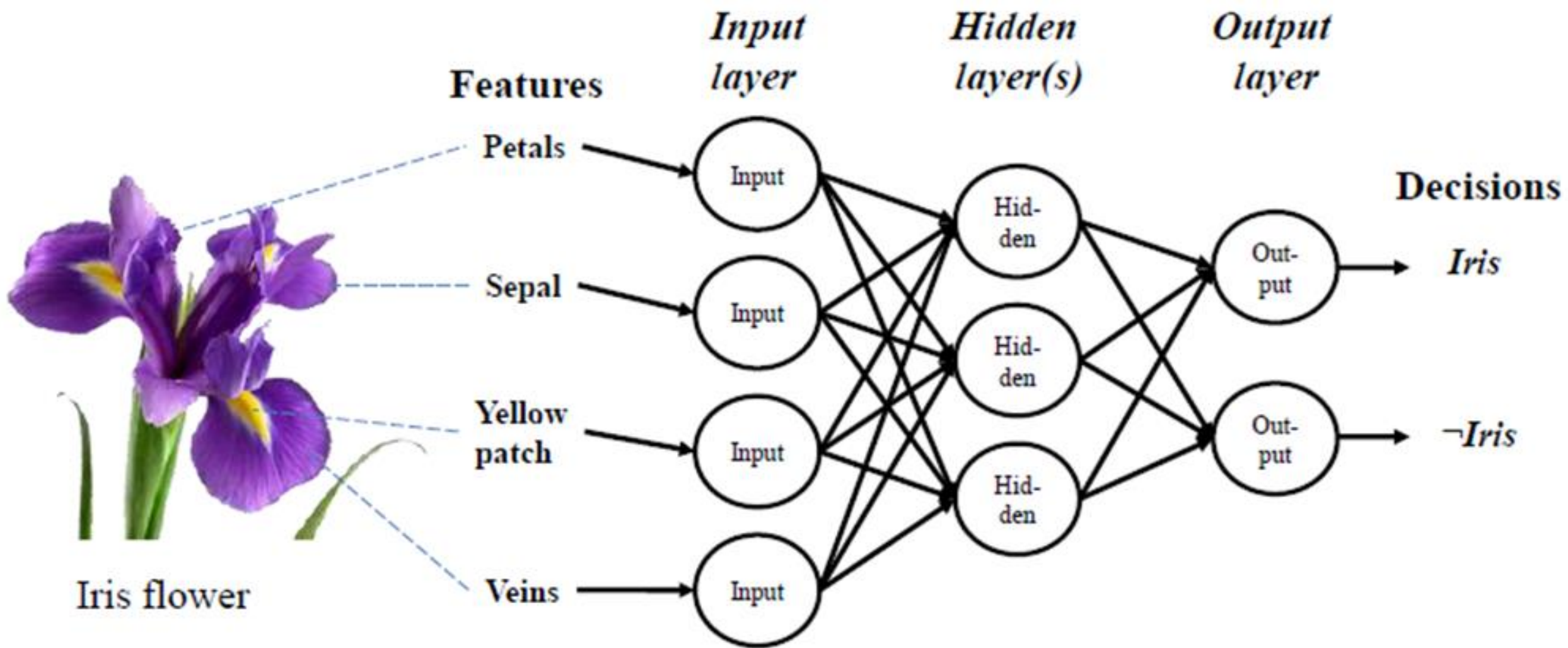
The neuron count has risen over the years to express more complex models.

More parameters efficiently learned using high computing powers.

More connections means that our networks have more parameters to optimize, and this required the explosion in computing power that occurred over the past 20 years.

Core components

- Parameters
- Layers
- Activation functions
- Loss functions
- Optimization methods
- Hyperparameters

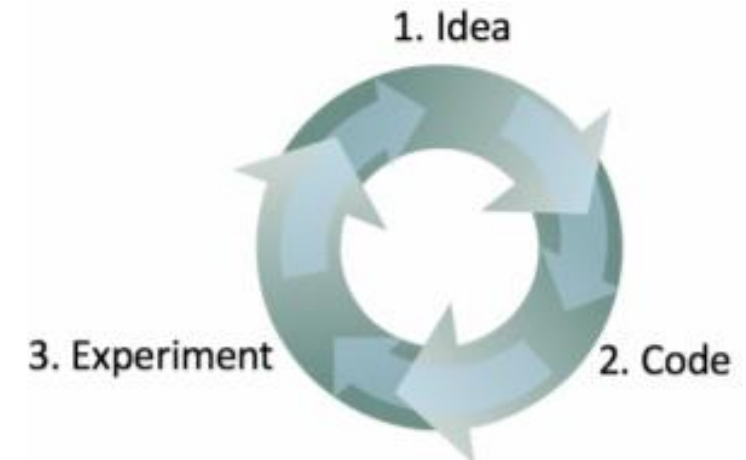


ML/DL is an iterative process: Don't expect it to work first time

Andrew Ng (Co-Founder of Coursera and a former head of Google Brain) states that his approach to building machine learning/deep learning software is threefold:

1. Start off with an idea
2. Implement the idea in code
3. Carry out an experiment to conclude how well the idea worked

The faster you can go around this loop, the quicker progress will be made!



How to solve
an agricultural
problem?

Is it a ML task? Are you sure ML is the best solution?

- Hard: X is independent of Y : $X, Y=?$
- Easy: X is a set with limited variations. Configure $Y=F(X)$

Appropriate ML scenario?

- Supervised learning
- Unsupervised learning

Appropriate model?

- Data size (small data -> linear model, large data -> consider non-linear)
- Imbalanced data (special treatment of the minority class required)

Enough training data?

- Investigate how precision improves with more data

Model overly complicated?

- Start simple first, increase complexity and evaluate performance
- Avoid overfitting to training set

Feature quality

- Have you identified all useful features?
- Use domain knowledge of an expert to start
- Include any feature that could be found and investigate model performance

Feature engineering

- The best strategy to improve performance and reveal important input
- Encode features, normalize [0:1], combine features

Combine models

- If multiple models have similar performance there is a chance of improvement
- Use one model for one subset of data and another model for the other

Model Validation

- Use appropriate performance indicator (Accuracy, Precision, Recall, F1, RMSE, MSE, etc.)
- How well does the model describe data? (AUC)
- Data typically divided into Training and Validation
- Evaluated accuracy on disjoint dataset (other than training dataset)
- Tune model hyper parameters (i.e. number of iterations)

Use Cases



The Smart
Agriculture market
is expected to
reach \$18.45
Billion in 2022, at
a CAGR of 13.8%
- Business Intelligence

DRONES

Health assessment,
irrigation, crop monitoring,
crop spraying, planting,
and soil and field analysis

Precision Farming

With IoT, all data from
different sensors is
accessible to the
agriculturist on their
mobile phones

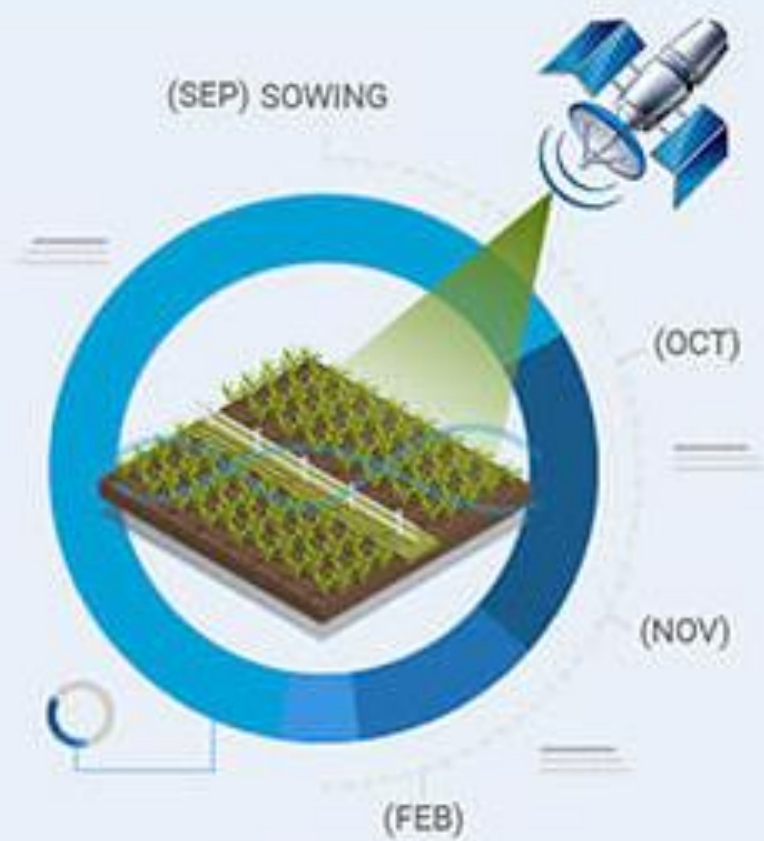
Soil Management

Analyze soil status,
temperature and
humidity

Livestock Management

Monitor livestock productivity and health

Water Management with Automated Irrigation



**Satellite
Imagery**

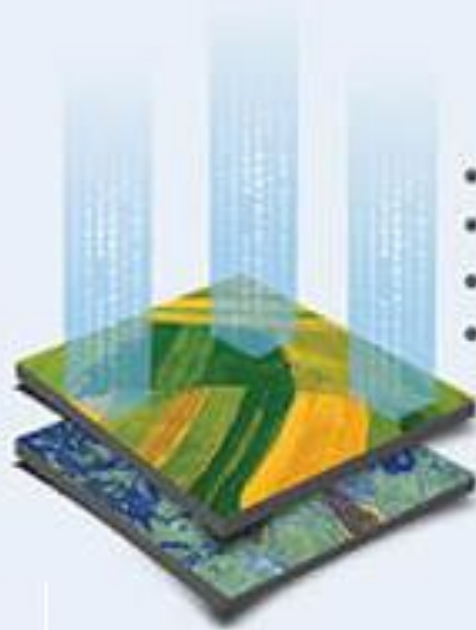


**Weather
Data**



**SmartFarm
Geo tagging labels**

- NDVI
- LAI
- SAVI
- EVI
- NDWI

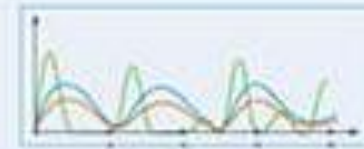


- Precipitation
- Temperature
- Humidity
- Evaporation

Feature Engineering



CROP 1



CROP 2



CROP 3

Crop Signatures



Crop Model 1

Crop Model 2

Yield Model 1

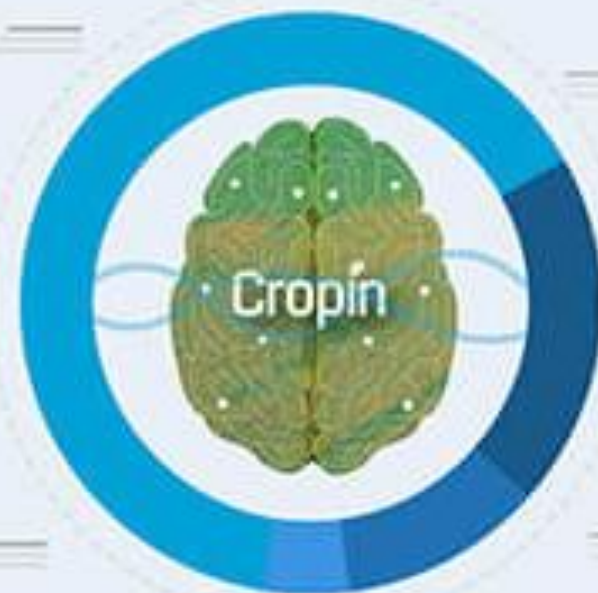
Yield Model 2

Stress Model 1

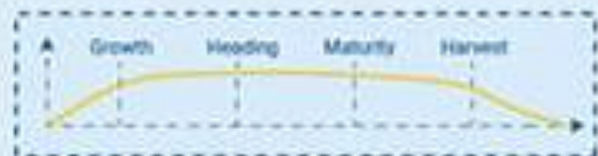
Stress Model 2



AI & ML Prediction



Model
Training



BUSINESS USE CASE



Agri Input
Companies

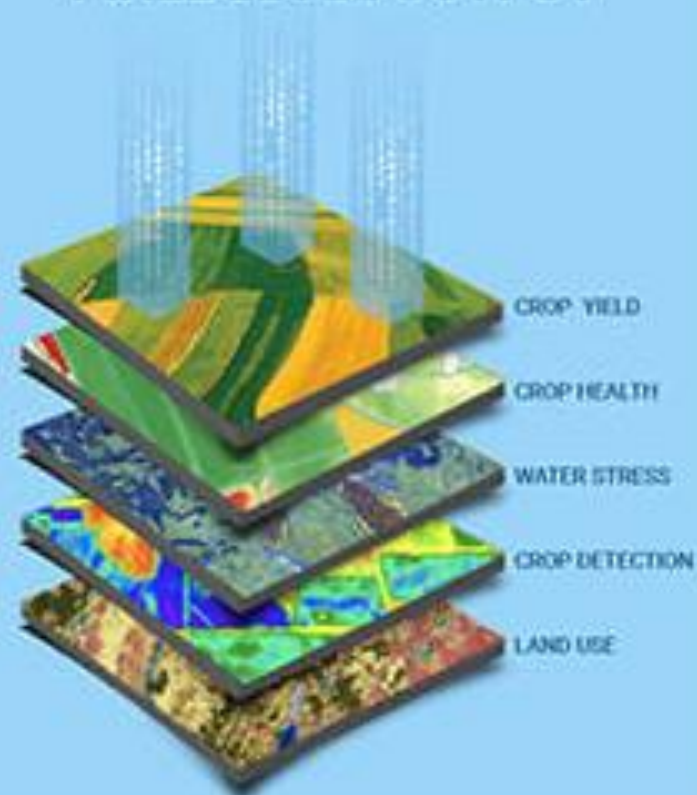


BFSI



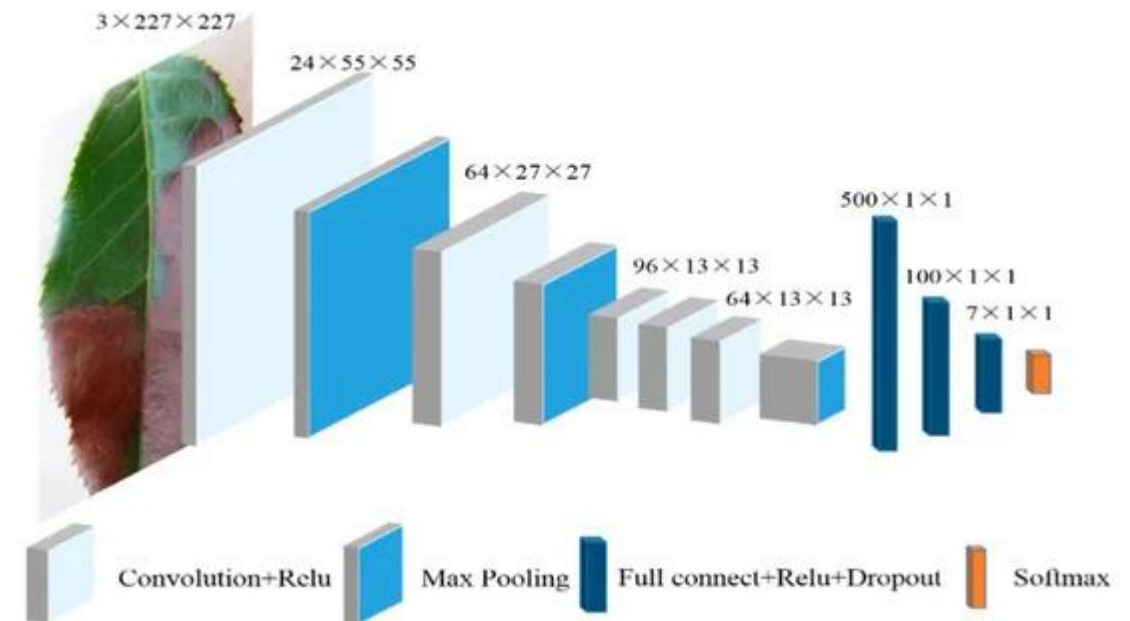
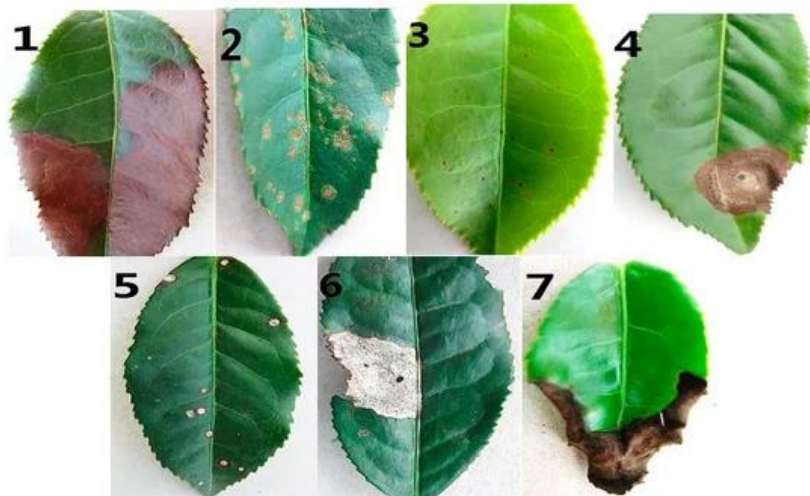
Farming
Companies

PIXEL LEVEL OUTPUT

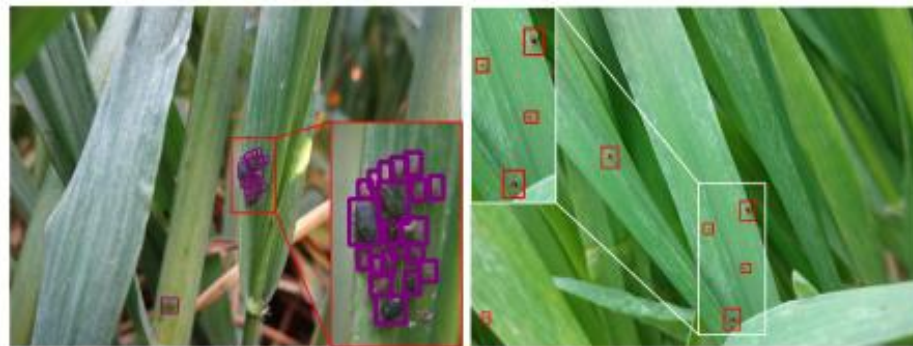


Case Studies

Leaf disease detection/ Identification of weeds - CNN



Pest Population Counting- RetinaNet



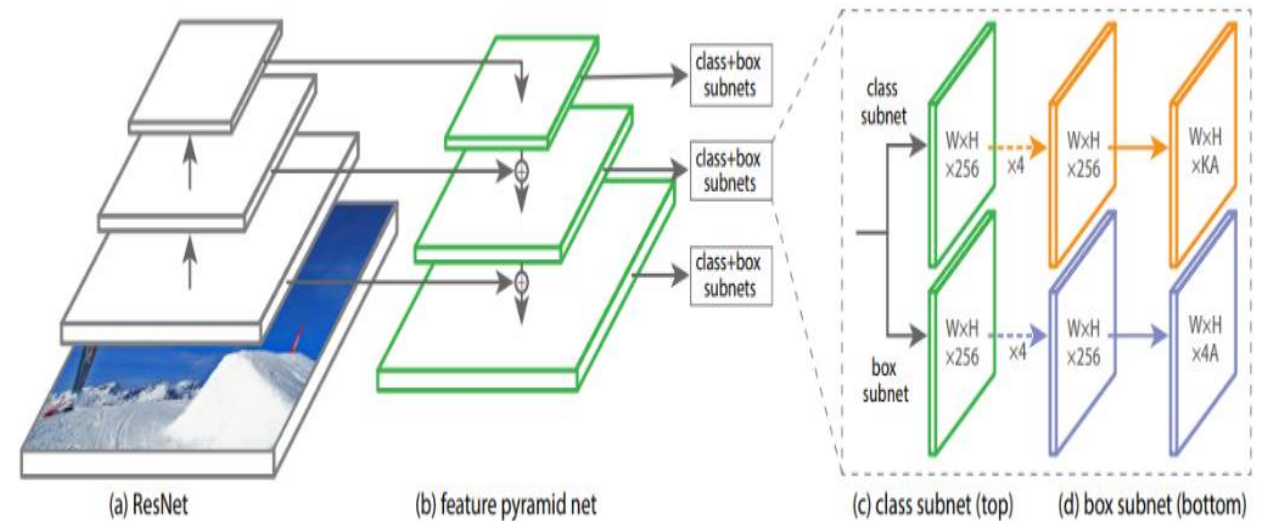
(a) Dense distribution of pests

(b) Sparse distribution of pests



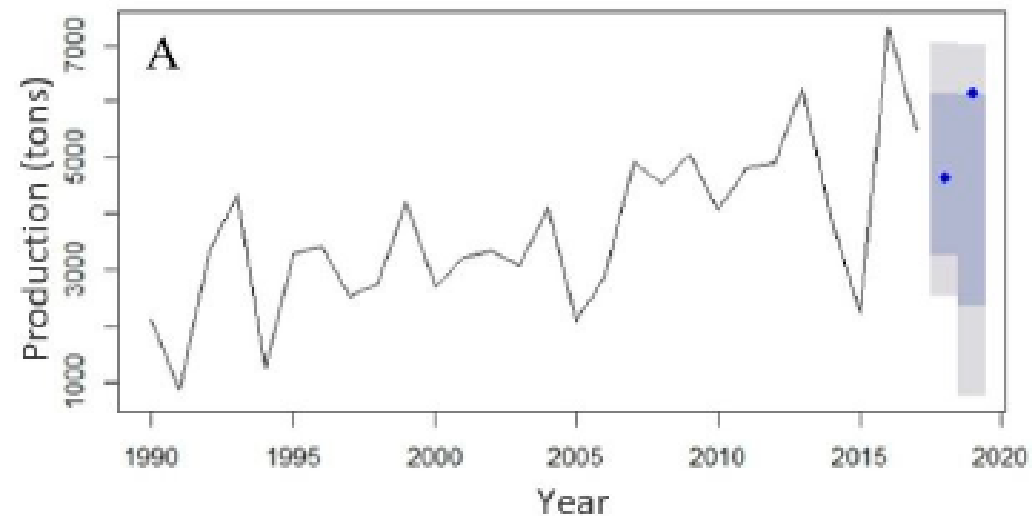
(c) Illumination variations

(d) Background clutter

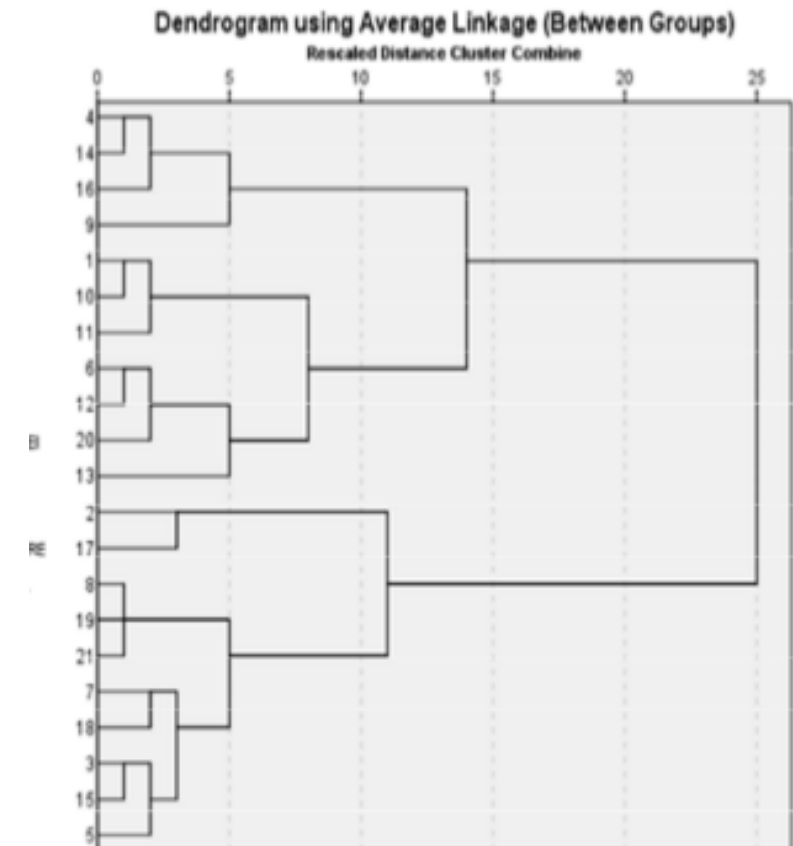
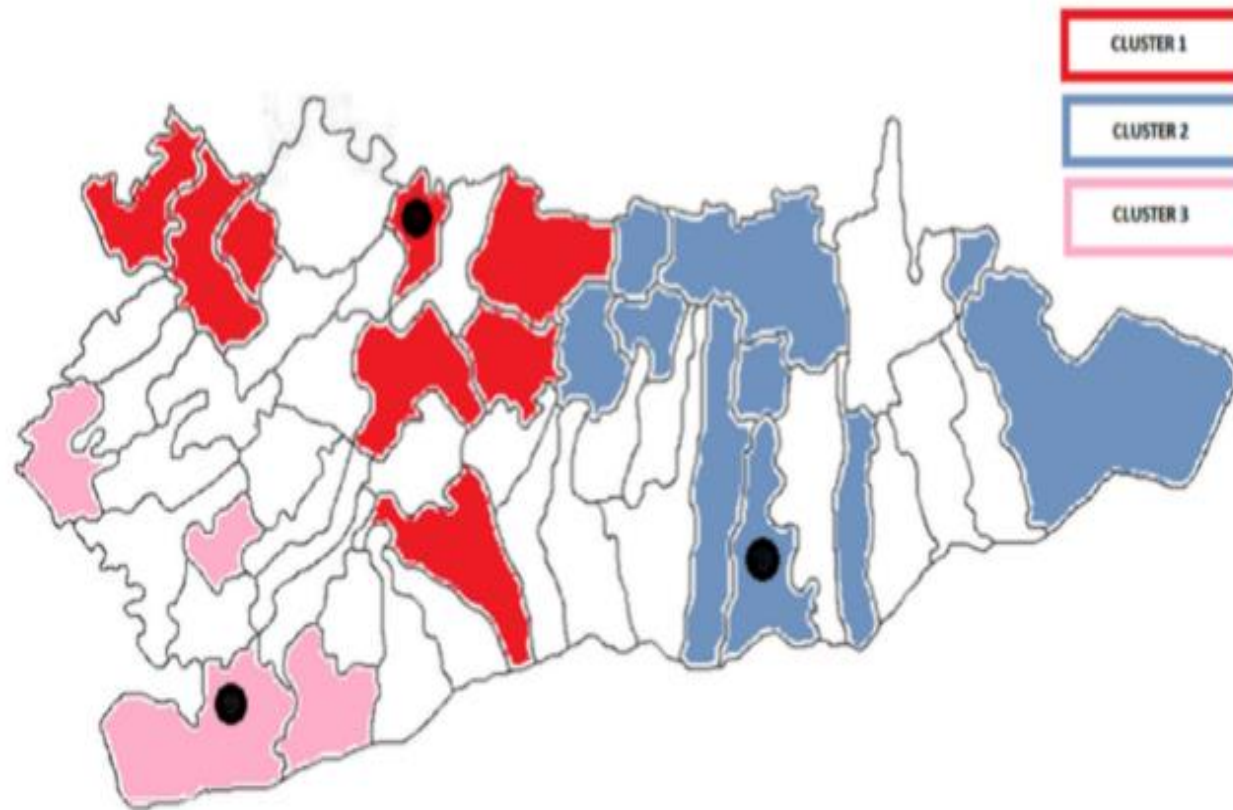


Crop Yield Prediction

Data Names	Abbreviation	Sources
Normalized Difference Vegetation Index	NDVI	MODIS (MOD13Q1)
Potential Evapotranspiration	PET	Climate Research Unit
Precipitation	PRE	Climate Research Unit
Minimum Temperature	TMN	Climate Research Unit
Maximum Temperature	TMX	Climate Research Unit
Soil Moisture	SM	European Space Agency
Size of land cultivated for maize production	Land	Department of Agriculture, Forestry and Fisheries



Organic Farming Patterns Analysis Based on Clustering Methods



Hands-on Session

Plant disease detection

Dataset:

- ▶ <https://www.kaggle.com/emmarex/plantdisease>

Code:

- ▶ <https://www.kaggle.com/sarithdivakar/plant-disease-detection-using-keras>