

Pathways to Become an AI/DS Expert

Dr. Sophea PRUM

Chief of Business Development Officer

ELITE VIGOUR

Contact Us

sopheaprum@elitevigour.com

My 20 years life on a single slide



2002: Bac II, Cambodia
2008: Engineer degree, Cambodia
2009: Master degree, France
2013: PhD Degree, France



2015: Researcher, Malaysia



2015: Lecturer Malaysia



Cambodia ICT Engineer
Award 2020



2020: Head of Product, Cambodia



2021: Co-founder Elite

Industrial Revolution 4.0



1st Revolution
End of 18th century

Steam and Water Power



2nd Revolution
Early 20th century

Electrical Power



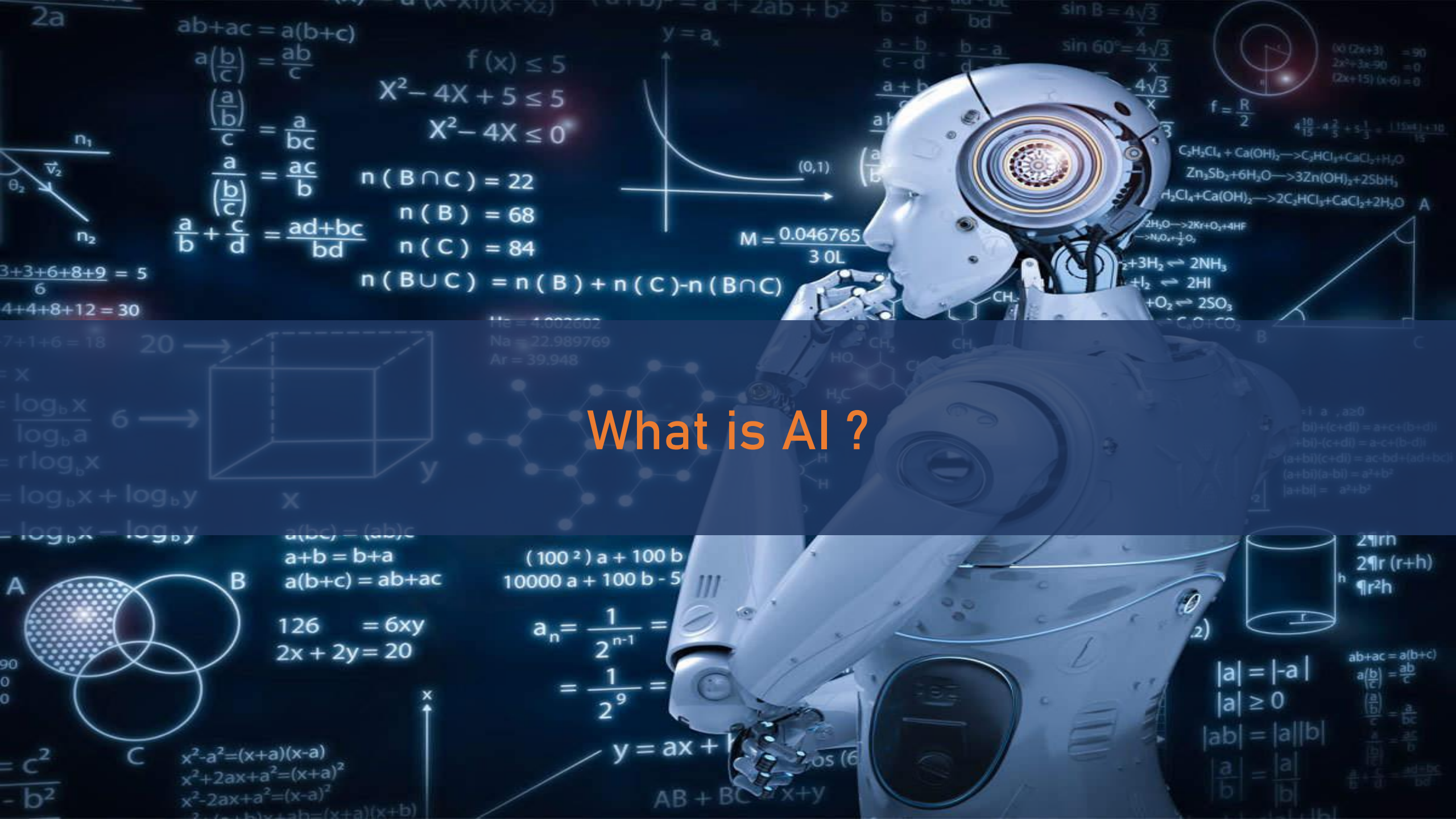
3rd Revolution
Late 20th century

Automated Production,
IT System and
Robotic



4th Revolution
Today

Autonomous Decision
using AI solution(s)



What is AI ?

How do you unlock your phone?



Password

Is it an AI based system?

Unlikely



Fingerprint

Is it an AI based system?

Yes

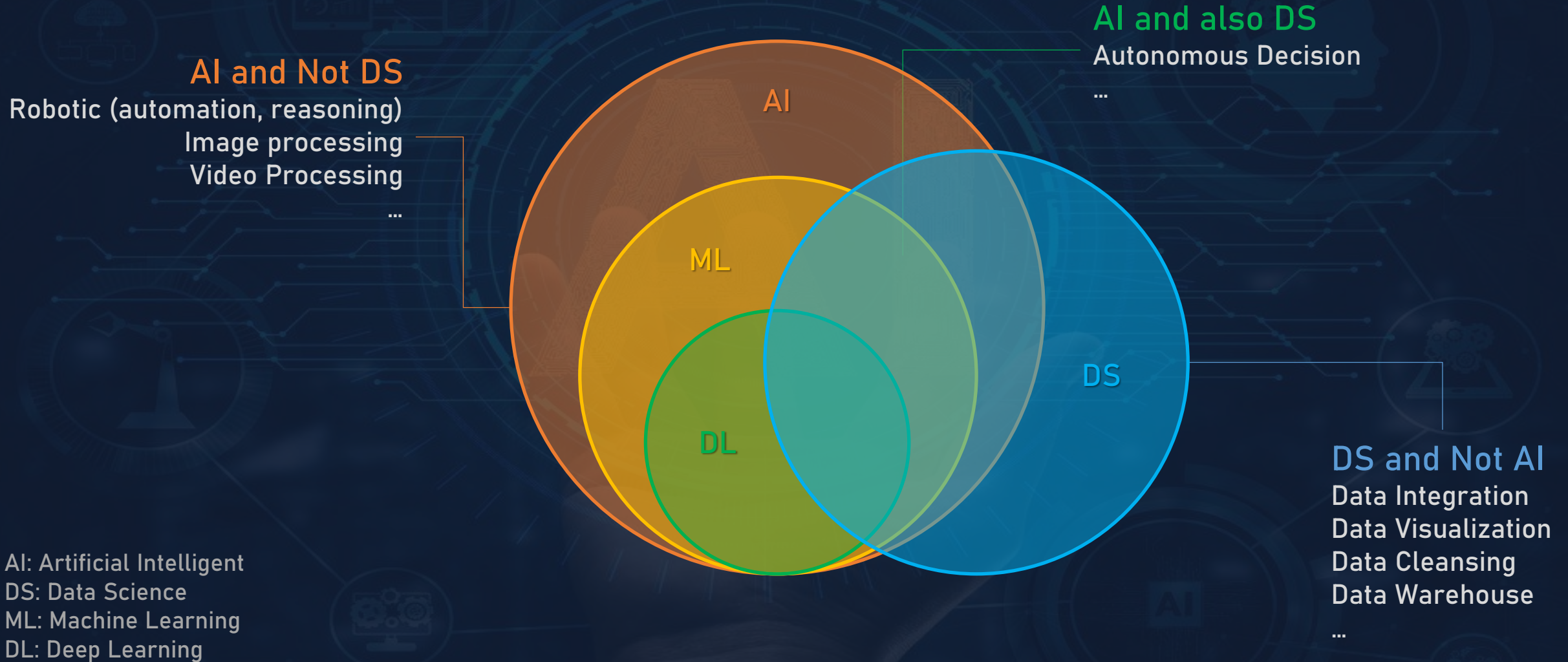


Face Recognition

Is it an AI based system?

Yes

AI vs. DS vs. ML vs. DL ?



5. Programming Skill

Tools/Libraries/Frameworks/Algorithms + implementation

4. Critical Thinking and Analysis

How to evaluate
ML result ?

Able to understand
"why" ?

How to
process your
data ?

3. Machine learning

Advantages/Disadvant
ages of each ML

ML
Algorithm/Concept

Feature Extraction
vs. Feature Learning

Data Integration,
Data Transformation
Data Cleansing
Data Normalization

1. Domain knowledge

2. Data Preparation

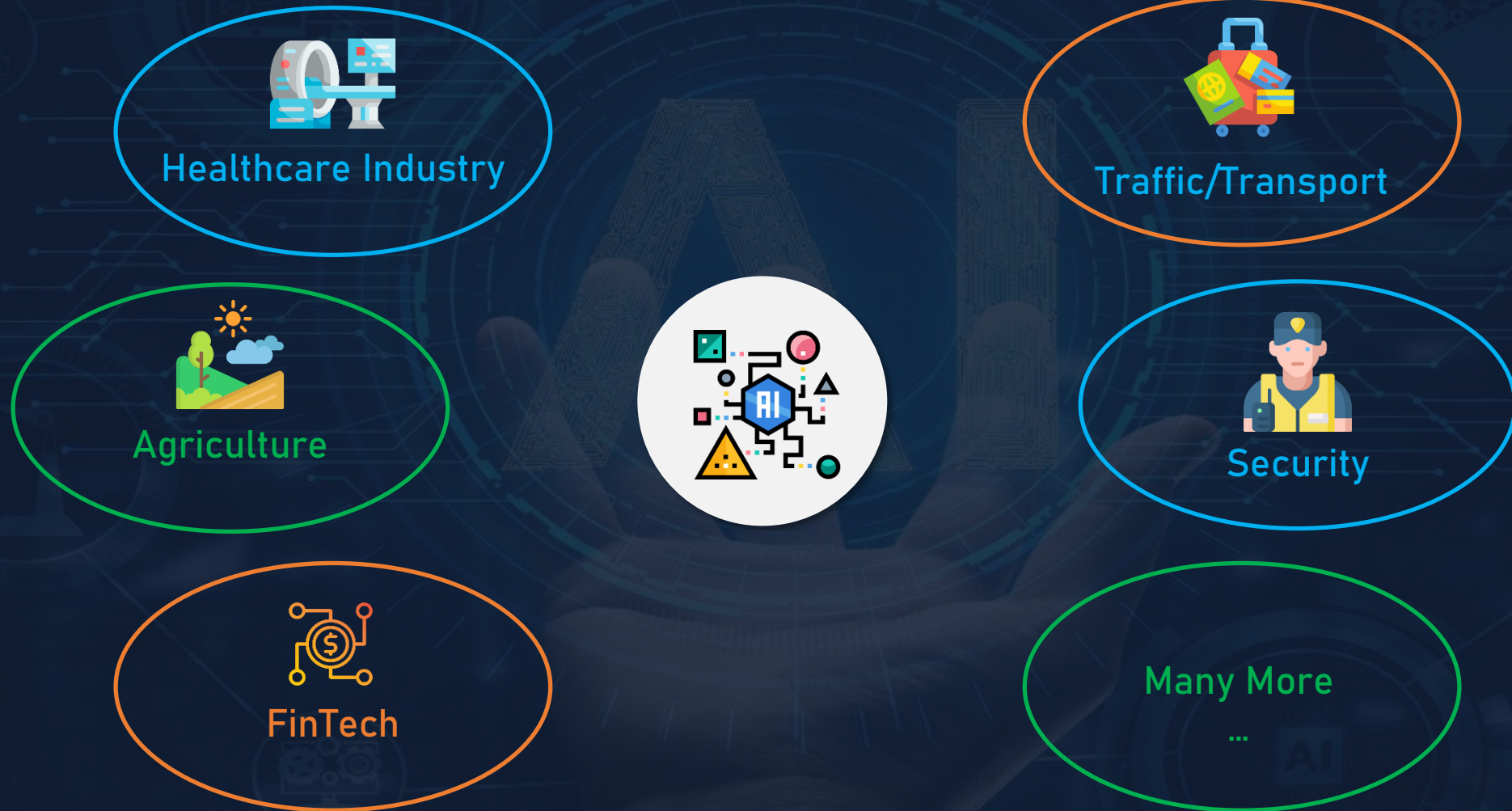
What are the
challenges ?

How do you want to
overcome the
challenges ?

What data do you
have ?

Or how can you
collect the data you
need ?

1/5 Domain Knowledge



1/5 Domain Knowledge



Business Perspective
(NON ICT)

Challenges

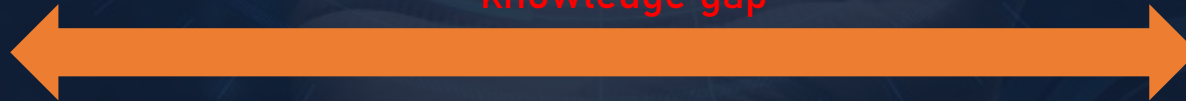
Tech Perspective
(ICT)

Education Institutions play crucial role here

- What are the challenges?
- How can you overcome the challenges?
- What data do you have?
- Or how can you collect the data you need?

- What are the challenges?
- How can you overcome the challenges?
- What data do you have?
- Or how can you collect the data you need?

Challenge?
knowledge gap



Solution?
Bridge the knowledge gap

But How?

2/5 Data Preparation



Objective: to get/generate

- Sufficient Data,
- Close to the reality (as close as possible)
- Clean Data

You Get What You Feed

2/5 Data Preparation

- What data do you have ?
- How can you get the data that you need ?



1. Data Integration

Integrate Data From
Different sources



2. Data Transformation

Changing the format,
structure, or values
to the desire one



3. Data Cleansing

Detecting and correcting
(or removing) corrupt or
inaccurate records



4. Data Normalization

Adjusting values measured on
different scales to a notionally
common scale

2/5 Data Preparation

1. Wine Quality Dataset

Different scale

Normalization or Standardization

| Fixed acidity | Volatile acidity | Citric acid | Residual sugar | Chlorides | Free sulfur dioxide | Total sulfur dioxide | Density | pH | Sulphates | Alcohol | Class |
|---------------|------------------|-------------|----------------|-----------|---------------------|----------------------|---------|------|-----------|---------|-------|
| 7 | 0.27 | 0.36 | 20.7 | 0.045 | 45 | 170 | 1.001 | 3 | 0.45 | 8.8 | 1 |
| 6.3 | 0.3 | 0.34 | 1.6 | 0.049 | 14 | 132 | 0.994 | 3.3 | 0.49 | 9.5 | 1 |
| 8.1 | 0.28 | 0.4 | 6.9 | 0.05 | 30 | 97 | 0.9951 | 3.26 | 0.44 | 10.1 | 2 |
| 7.2 | 0.23 | 0.32 | 8.5 | 0.058 | Null | 186 | 0.9956 | 3.19 | 0.4 | 9.9 | 2 |
| 7.2 | 0.23 | 0.32 | 8.5 | 0.058 | 47 | Null | 0.9956 | 3.19 | 0.4 | 9.9 | 3 |

Incorrect Value:

- Type of Data: Structured data
- Number of records: 4898
- Number of attributes: 11
- Number of classes: 10

1. Remove, or
2. Correct ?

3/5 Machine Learning



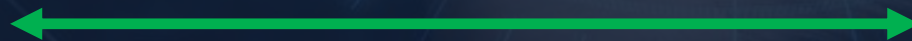
The dream to create AI system that can create another AI system

Emulating the decision-making ability of a human expert

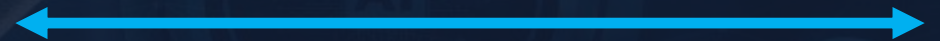
Widely available of Machine Learning algorithms and tools

Adaptive model that can learn with smaller dataset using adaptive learning strategy

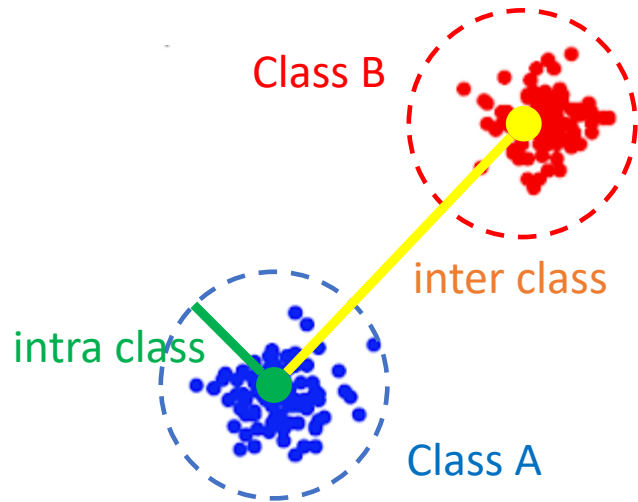
Conventional Machine Learning Algorithms



Deep Learning Algorithms



3/5 Machine Learning

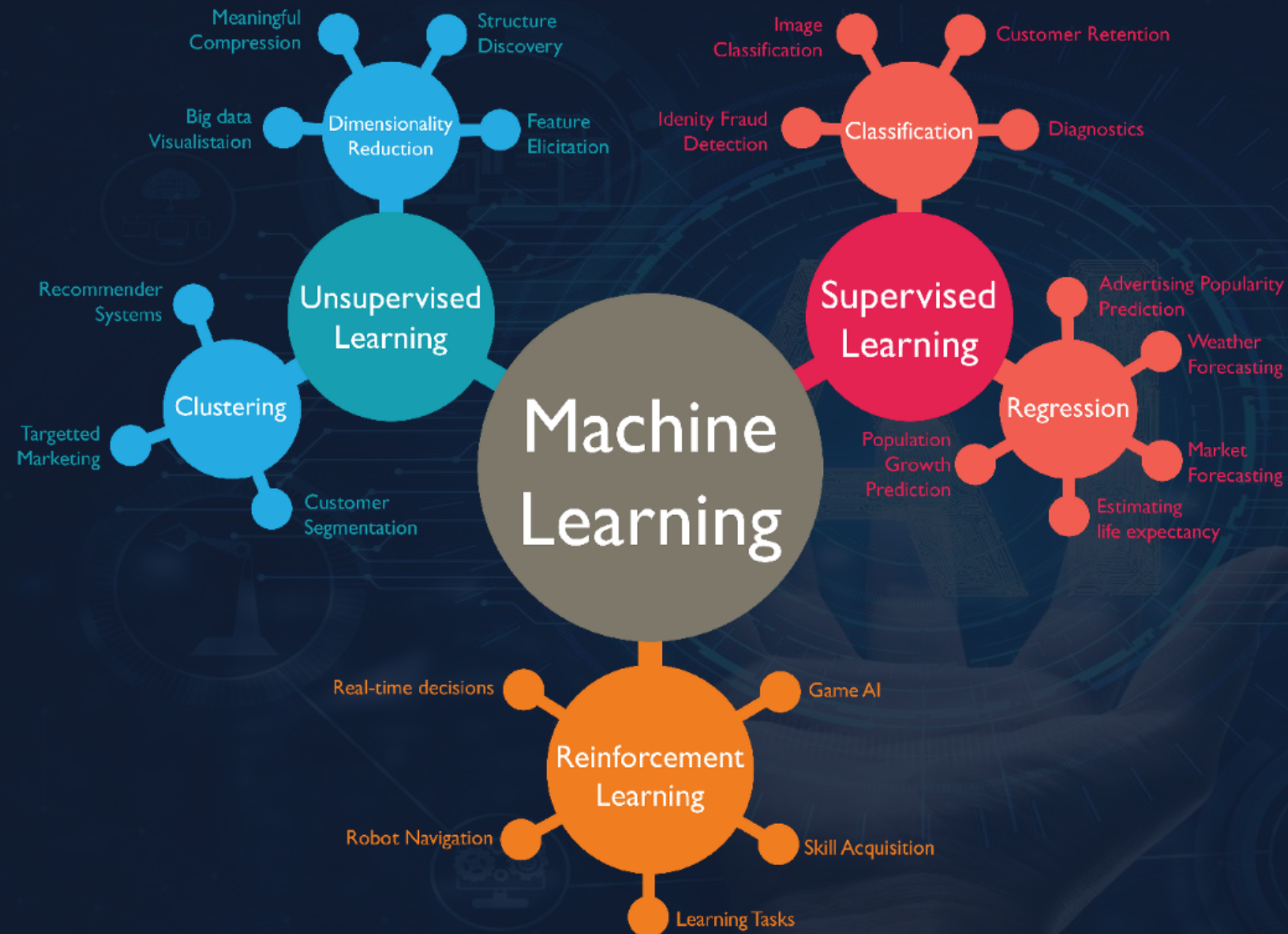


Goal of Machine Learning

- Minimize the distance of the intra class samples
- Maximize the distance of the inter class samples

➤ What you have to know about ML?

- ML Algorithm/Concept
- Advantages/Disadvantages of each ML
- How to process your data?



Dimensionality Reduction:
Principal Component Analysis (PCA)
Feature Selection
Linear Discriminant Analysis (LDA)
...

Classification:
Neural Network
SVM
K-Nearest Neighbor
CNN
RNN/LSTM/BLSTM
...

Algorithms:
Decision Tree/Random Forest Regressor
Linear Regression
Logistic Regression
...

Clustering:
Mean Shift
K-Medoids
K-means
...

3/5 Machine Learning

- Applying ML without
 - Understanding ML algorithm/concept
 - Understanding advantages/disadvantages of the ML
 - Understanding the nature of your data



4/5 Critical Thinking and Analysis

- How to evaluate ML results ?
- Able to understand “Why it works” and “Why it does not work” ?



Data Benchmark

Objective

- Create a standard dataset to evaluate your algorithm(s)
- Close to the reality (“as close as possible”)



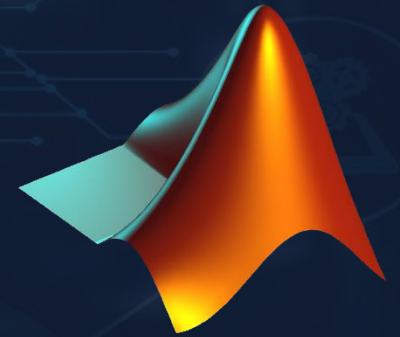
Evaluation Matrices

Algorithms

- Confusion matrix
- F1 Score
- AUC (Area Under ROC Curve)
- ...

5/5 Programming Skill

Being a full options researcher



5/5 Programming Skill



Tools for data analysis and predictive models



Library (C/C++ and Python) : computer vision and machine learning



Framework (Python, CUDA, C++, C) : Machine Learning, NLP, Computer Vision



Library (C/C++ and Python) : Machine Learning, AI, Focus on Deep Learning



Library (Python) : Act as interface for Tensorflow. Later on, it support Ms Cognitive, Theano and PlaidML

Many More ...

Summary

To expertise AI:

1. Domain Knowledge
2. Data Preparation
3. Machine Learning
4. Critical Thinking & Analysis
5. Programming skill

- ☐ AI is fun
- ☐ Start from small things
- ☐ AI theory and mathematic can be a bit complex → play to learn
- ☐ How to be great in AI ?

1.
Getting
Start

2.
Work
Hard

3.
Work
Harder

It is 'Almost' impossible to be an AI expert,
but we can learn and take part of the journey
because AI is the future



ස්තූතියි!
Thank You!