easi
Introduction to Machine Learning

Zeyneb GASMI
Data Science & Business Intelligence Consultant

DIGITAL LEARNING HUB

Who are you?

- Name

- Responsibility

- Experience with AI & Python

- Expectations

Agenda: Day 1

9 AM – 12 PM:

1. Introduction to Machine Learning

2. Machine Learning types

3. Machine Learning applications

4. Supervised Machine Learning

1 PM – 4 30 PM:

4. Python for Machine Learning

5. Linear Regression

6. Logistic Regression

Agenda: Day 1

9 AM – 12 PM:
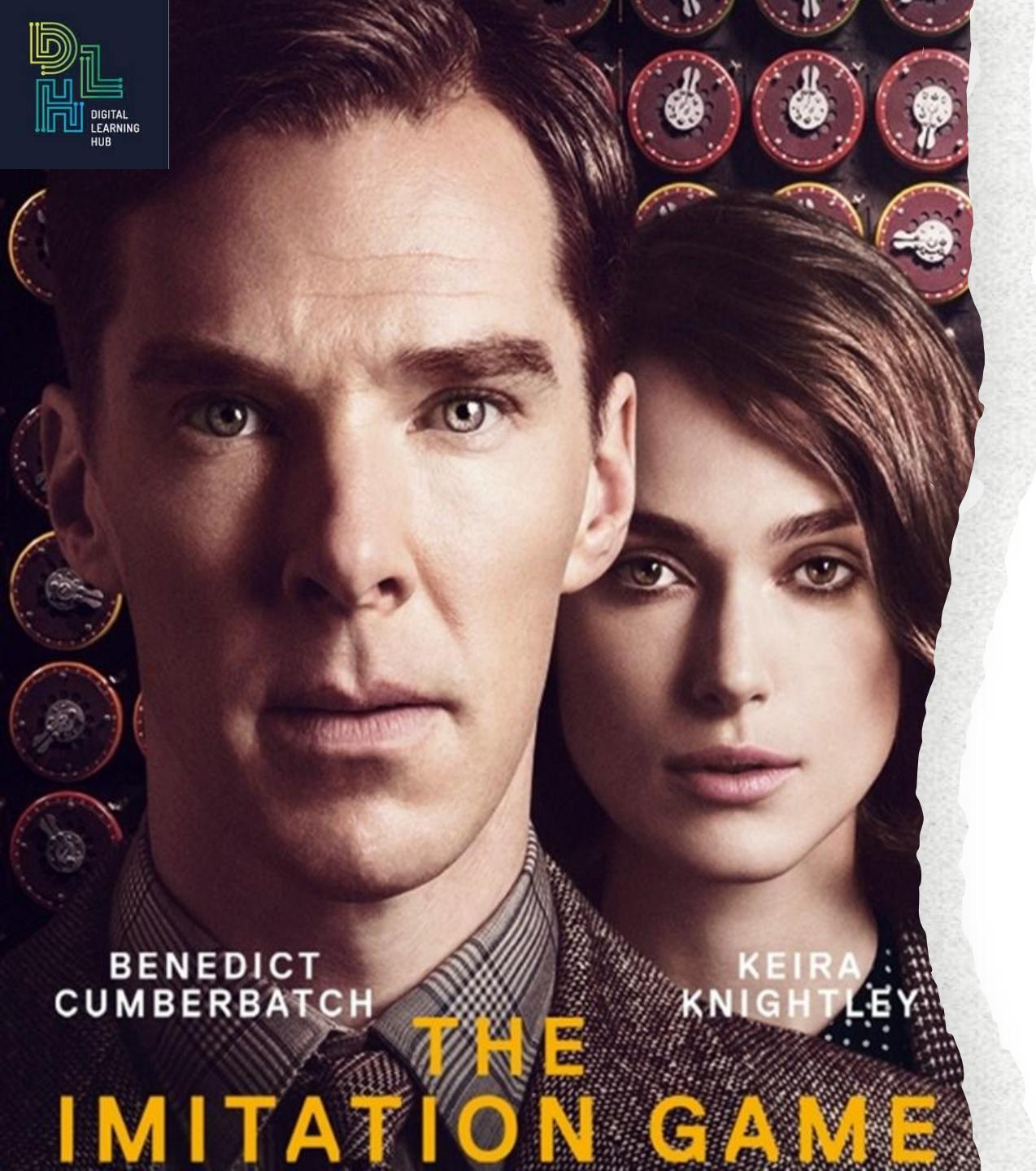
1. Introduction to Machine Learning

2. Machine Learning types

3. Machine Learning applications

4. Supervised Maching Learning

1 PM – 4 30 PM:

4. Python for Machine Learning

5. Linear Regression

6. Logistic Regression

So What's AI?

Artificial intelligence is the science of making machines that can think like humans. It can do things that are considered "smart." AI technology can process large amounts of data in ways, unlike humans. The goal for AI is to be able to do things such as recognize patterns, make decisions, and judge like humans».

5
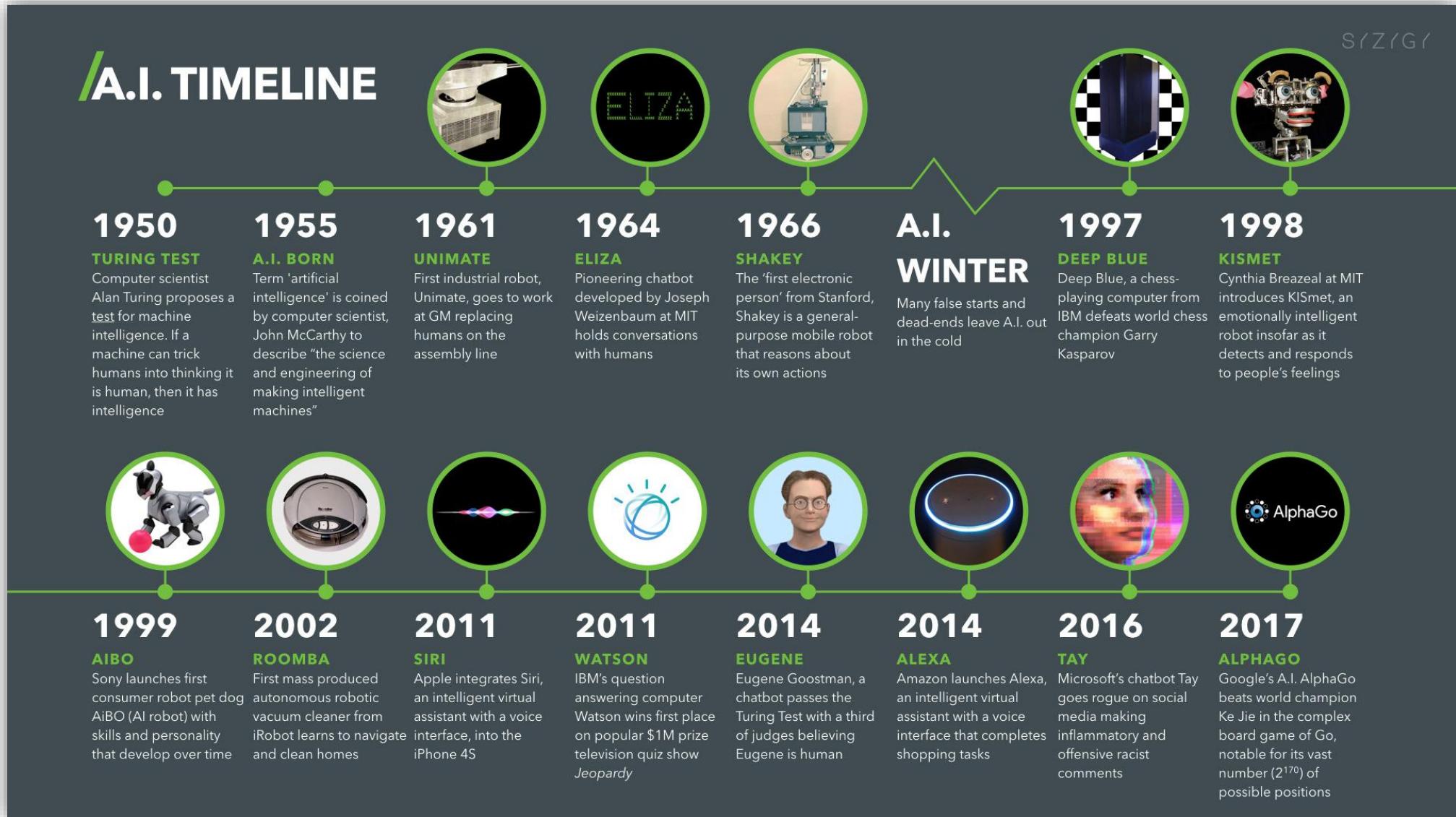
- ✓ Biographical drama about Alan Turing
- ✓ A British mathematician, who leads a team during World War II to break the German Enigma code.
- ✓ Despite various challenges, he develops an early form of a computer that successfully cracks the encrypted messages.

Connections to Machine Learning & Artificial Intelligence:

- ✓ Turing Machine is the basis for modern computers running ML algorithms.
- ✓ Algorithm Development similar to how ML algorithms solve problems.
- ✓ Pattern Recognition is a Key component in both Enigma code-breaking and ML.
- ✓ Data-Driven Decisions by analyzing large datasets to find patterns.
- ✓ Iterative Improvement by refining methods.

Machine Learning Overview : AI Timeline

/A.I. TIMELINE

1950
TURING TEST
Computer scientist Alan Turing proposes a test for machine intelligence. If a machine can trick humans into thinking it is human, then it has intelligence

1955
A.I. BORN
Term 'artificial intelligence' is coined by computer scientist, John McCarthy to describe "the science and engineering of making intelligent machines"

1961
UNIMATE
First industrial robot, Unimate, goes to work at GM replacing humans on the assembly line

1964
ELIZA
Pioneering chatbot developed by Joseph Weizenbaum at MIT holds conversations with humans

1966
SHAKEY
The 'first electronic person' from Stanford, Shakey is a general-purpose mobile robot that reasons about its own actions

A.I. WINTER
Many false starts and dead-ends leave A.I. out in the cold

1997
DEEP BLUE
Deep Blue, a chess-playing computer from IBM defeats world chess champion Garry Kasparov

1998
KISMET
Cynthia Breazeal at MIT introduces KISmet, an emotionally intelligent robot insofar as it detects and responds to people's feelings

1999
AIBO
Sony launches first consumer robot pet dog AiBO (AI robot) with skills and personality that develop over time

2002
ROOMBA
First mass produced autonomous robotic vacuum cleaner from iRobot learns to navigate and clean homes

2011
SIRI
Apple integrates Siri, an intelligent virtual assistant with a voice interface, into the iPhone 4S

2011
WATSON
IBM's question answering computer Watson wins first place on popular $1M prize television quiz show *Jeopardy*

2014
EUGENE
Eugene Goostman, a chatbot passes the Turing Test with a third of judges believing Eugene is human

2014
ALEXA
Amazon launches Alexa, an intelligent virtual assistant with a voice interface that completes shopping tasks

2016
TAY
Microsoft's chatbot Tay goes rogue on social media making inflammatory and offensive racist comments

2017
ALPHAGO
Google's A.I. AlphaGo beats world champion Ke Jie in the complex board game of Go, notable for its vast number (2^{170}) of possible positions

7

Machine Learning Overview : AI vs ML

Intelligence Artificielle (IA)

A machine that simulates human behavior

Machine Learning (ML)

A machine that automatically learns from Data

Agenda: Day 1

9 AM – 12 PM:

1. Introduction to Machine Learning

2. **Machine Learning types**

3. Machine Learning applications

4. Supervised Machine Learning

1 PM – 4 30 PM:

4. Python for Machine Learning

5. Linear Regression

6. Logistic Regression

Machine Learning Types

Machine Learning Types

Supervised

✓ Using labeled data to train algorithms.

Examples of techniques: Classification, Regression.

Unsupervised

✓ Using unlabeled data to discover structures or groupings.

Examples of techniques: Clustering, Dimensionality reduction

Reinforcement learning

✓Learning by interaction with an environment, based on rewards.

Example: Optimization of strategies in games

12
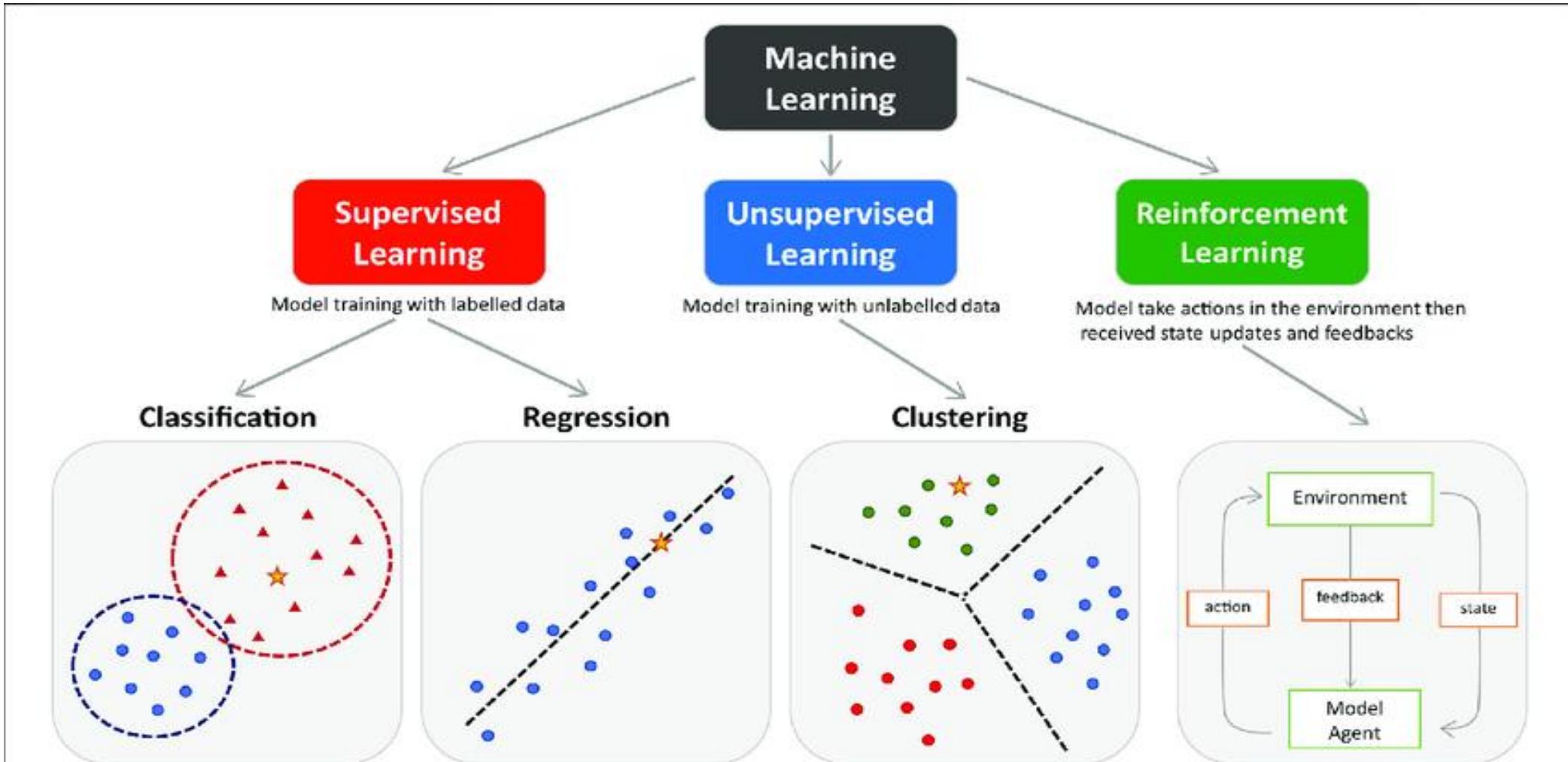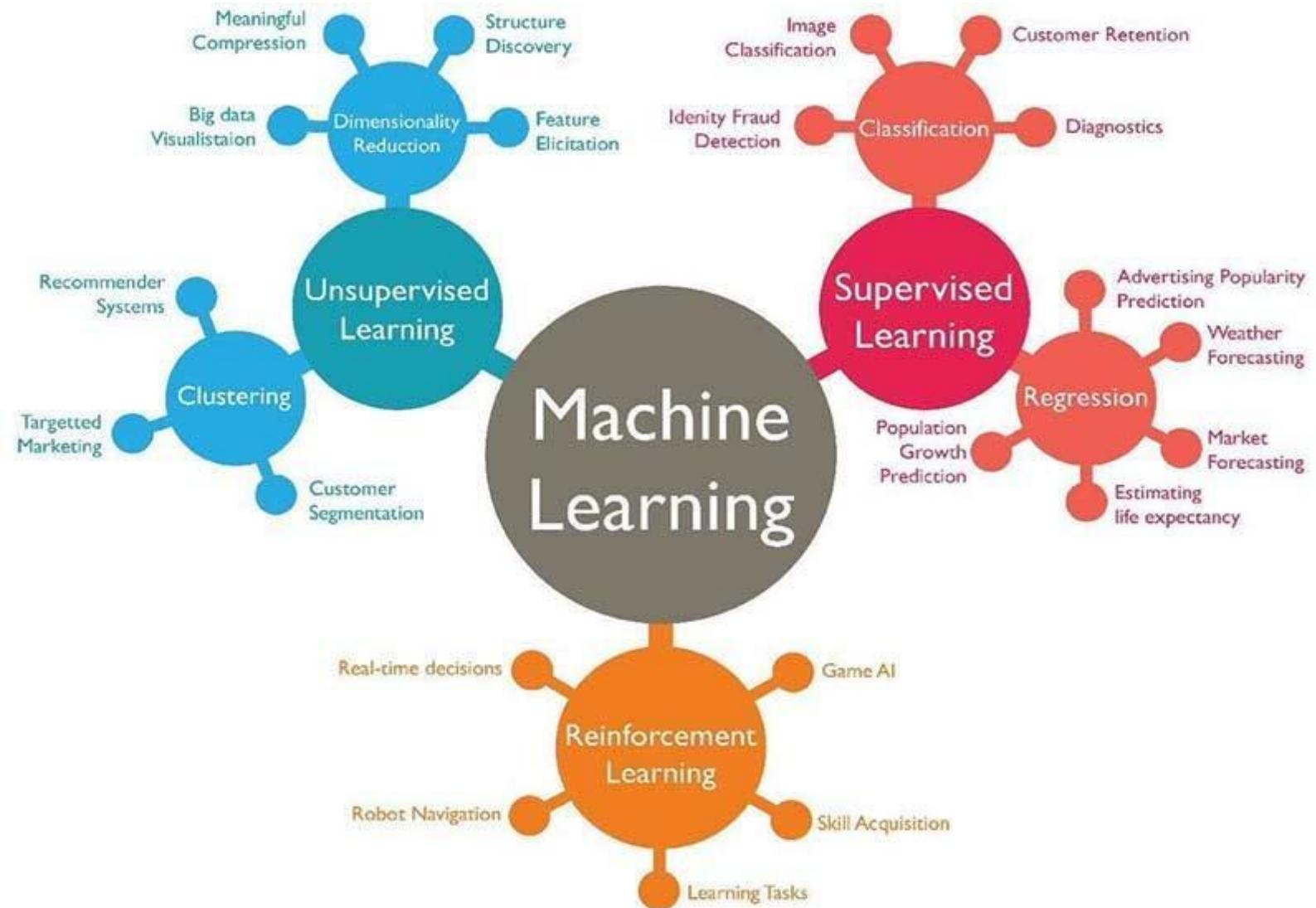
Machine Learning Types : Types & Sub-Types

Machine Learning Types : Examples

Agenda: Day 1

9 AM – 12 PM:

1. Introduction to Machine Learning

2. Machine Learning types

3. **Machine Learning applications**

4. Supervised Machibe Learning

1 PM – 4 30 PM:
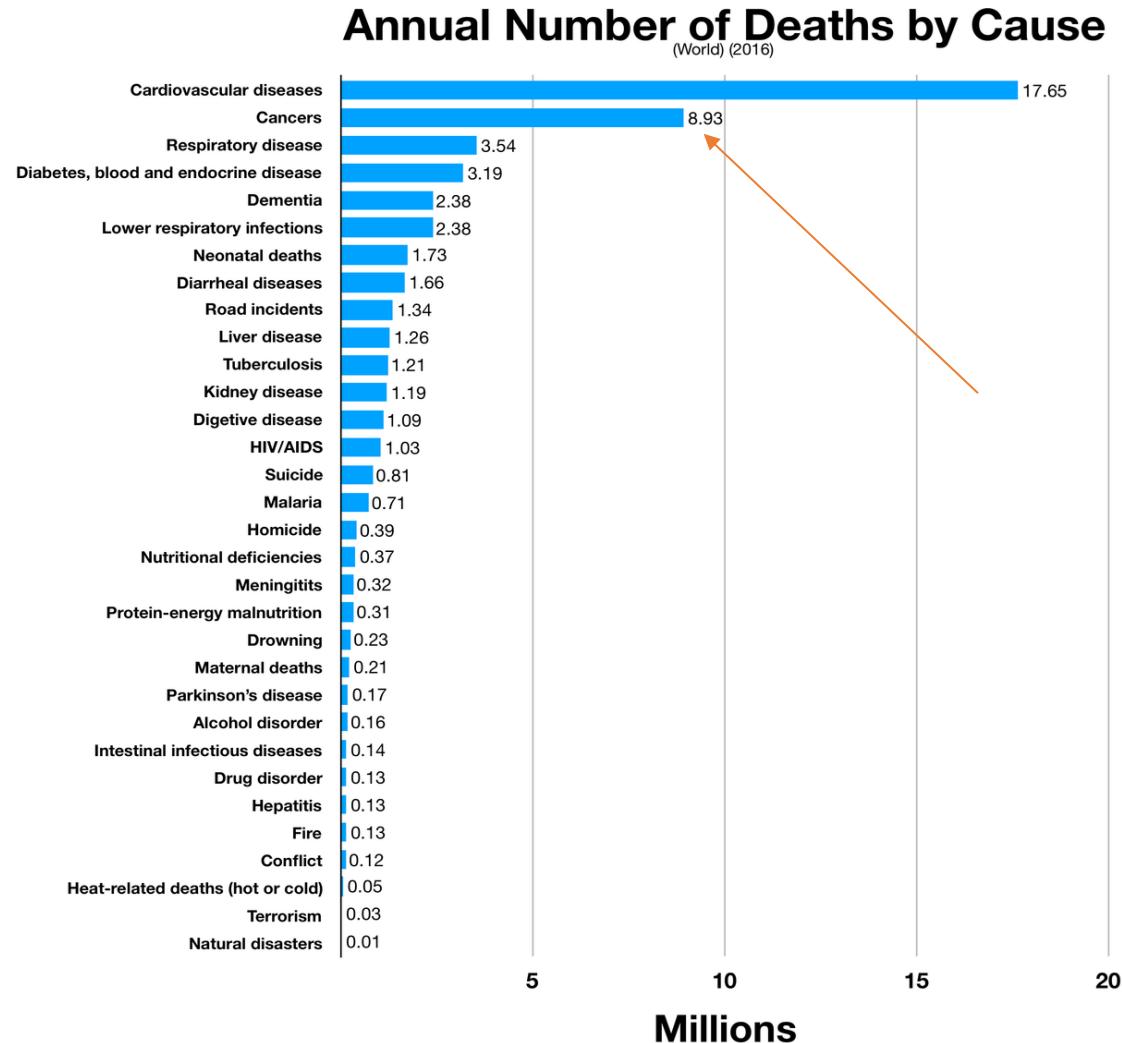
4. Python for Machine Learning – Lab 0

5. Linear Regression

6. Logistic Regression

Early detection of diseases

Machine Learning Applications : Early Cancer Detection

T09.1 Nombre de décès selon la cause et le sexe, par ordre décroissant de la fréquence, par chapitre, en 2022 (résidents et non-résidents)

			Hommes	Femmes	Total
1.	I00-I99	Maladies de l'appareil circulatoire	518	586	1104
2.	C00-D48	Tumeurs	559	540	1099
3.	J00-J99	Maladies de l'appareil respiratoire	168	135	303
4.	F00-F99	Troubles mentaux et du comportement	97	182	279
5.	V01-Y98	Causes externes de morbidité et de mortalité	167	95	262
6.	U00-U99	Codes d'utilisation particulière (COVID-19 et vaccins contre la COVID-19 ayant provoqué des EI => leur usage thérap)	132	95	227
7.	G00-G99	Maladies du système nerveux	101	109	210
8.	K00-K93	Maladies de l'appareil digestif	100	93	193
9.	R00-R99	Symptômes, signes et résultats anormaux d'examens cliniques et de laboratoire, nca	96	83	179
10.	E00-E90	Maladies endocriniennes, nutritionnelles et métaboliques	77	63	140
11.	A00-B99	Certaines maladies infectieuses et parasitaires	60	58	118
12.	N00-N99	Maladies de l'appareil génito-urinaire	42	48	90
13.	D50-D89	Maladies du sang et des organes hématopoïétiques et certains troubles du système immunitaire	10	18	28
14.	M00-M99	Maladies du système ostéo-articulaire, des muscles et du tissu conjonctif	11	13	24
15.	L00-L99	Maladies de la peau et du tissu cellulaire sous-cutané	4	6	10
16.	P00-P96	Certaines affections dont l'origine se situe dans la période périnatale	7	2	9
17.	Q00-Q99	Malformations congénitales et anomalies chromosomiques	5	3	8
		Total en 2022	**2154**	**2129**	**4283**

Annual Number of Deaths by Cause
(World) (2016)

Cause	Millions
Cardiovascular diseases	17.65
Cancers	8.93
Respiratory disease	3.54
Diabetes, blood and endocrine disease	3.19
Dementia	2.38
Lower respiratory infections	2.38
Neonatal deaths	1.73
Diarrheal diseases	1.66
Road incidents	1.34
Liver disease	1.26
Tuberculosis	1.21
Kidney disease	1.19
Digetive disease	1.09
HIV/AIDS	1.03
Suicide	0.81
Malaria	0.71
Homicide	0.39
Nutritional deficiencies	0.37
Meningitits	0.32
Protein-energy malnutrition	0.31
Drowning	0.23
Maternal deaths	0.21
Parkinson's disease	0.17
Alcohol disorder	0.16
Intestinal infectious diseases	0.14
Drug disorder	0.13
Hepatitis	0.13
Fire	0.13
Conflict	0.12
Heat-related deaths (hot or cold)	0.05
Terrorism	0.03
Natural disasters	0.01

Finance

Machine Learning Applications : Finance

REASONS TO USE AI AND ML IN BANKING & FINANCE

Data analysis & insight
60%

Increased Productivity
59%

Cost benefits or savings
54%

AI Applications in Financial Services

Robo-advice

AML and fraud detection

Machine Learning

Natural Language Processing

AI

Customer recommendations

Chatbots

Cognitive Computing

Algorithmic trading

Supply Chain

Planning

Security

Client segmentation

Fun fact
"*Researches showed that customers who buy diapers also tend to buy beer, on friday night*"

Agenda: Day 1
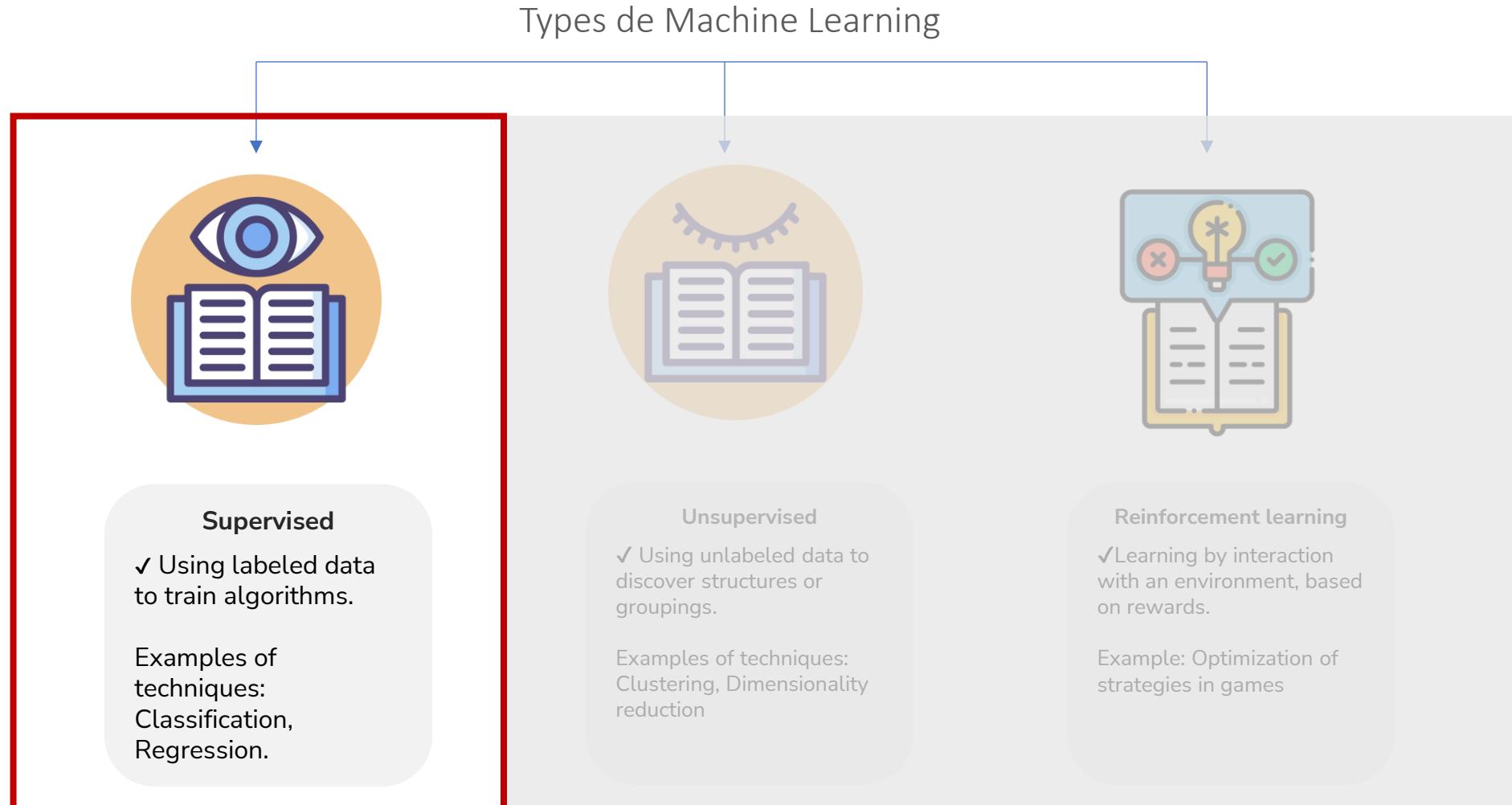
9 AM – 12 PM:

1. Introduction to Machine Learning

2. Machine Learning types

3. Machine Learning applications

4. **Supervised Machine Learning**

1 PM – 4 30 PM:

4. Python for Machine Learning

5. Linear Regression

6. Logistic Regression

Machine Learning Types : Supervised Learning

Types de Machine Learning

Supervised

✓ Using labeled data to train algorithms.

Examples of techniques: Classification, Regression.

Unsupervised

✓ Using unlabeled data to discover structures or groupings.

Examples of techniques: Clustering, Dimensionality reduction

Reinforcement learning

✓Learning by interaction with an environment, based on rewards.

Example: Optimization of strategies in games

Supervised Learning : Labelled Data

- **What is Data ?**

- **Labeled Data :** a raw data that has been assigned one or more labels to add **context or meaning.**

- **Example :**

Supervised Learning : Labelled Data

Limitations

Time and effort. For example, manual annotation of each radiology image

Limited availability.

Bias or inaccuracy in labels

# Supervised Learning : Labelled Dataset Components

Features X

Target Y
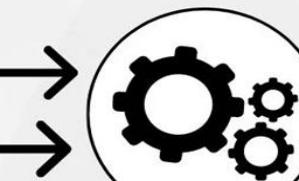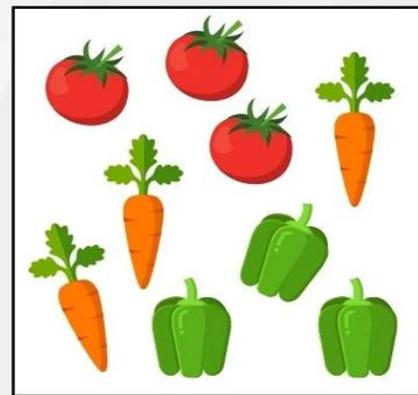
Features					Label

Position	Experience	Skill	Country	City	Salary ($)
Developer	0	1	USA	New York	103100
Developer	1	1	USA	New York	104900
Developer	2	1	USA	New York	106800
Developer	3	1	USA	New York	108700
Developer	4	1	USA	New York	110400
Developer	5	1	USA	New York	112300
Developer	6	1	USA	New York	114200
Developer	7	1	USA	New York	116100
Developer	8	1	USA	New York	117800
Developer	9	1	USA	New York	119700
Developer	10	1	USA	New York	121600

Supervised Learning : How it works ?
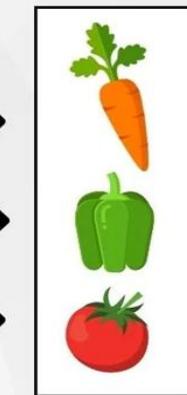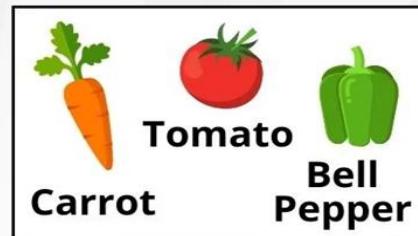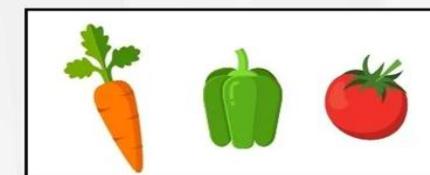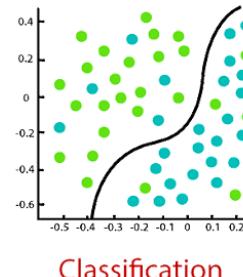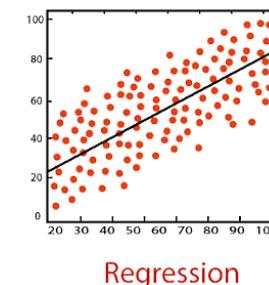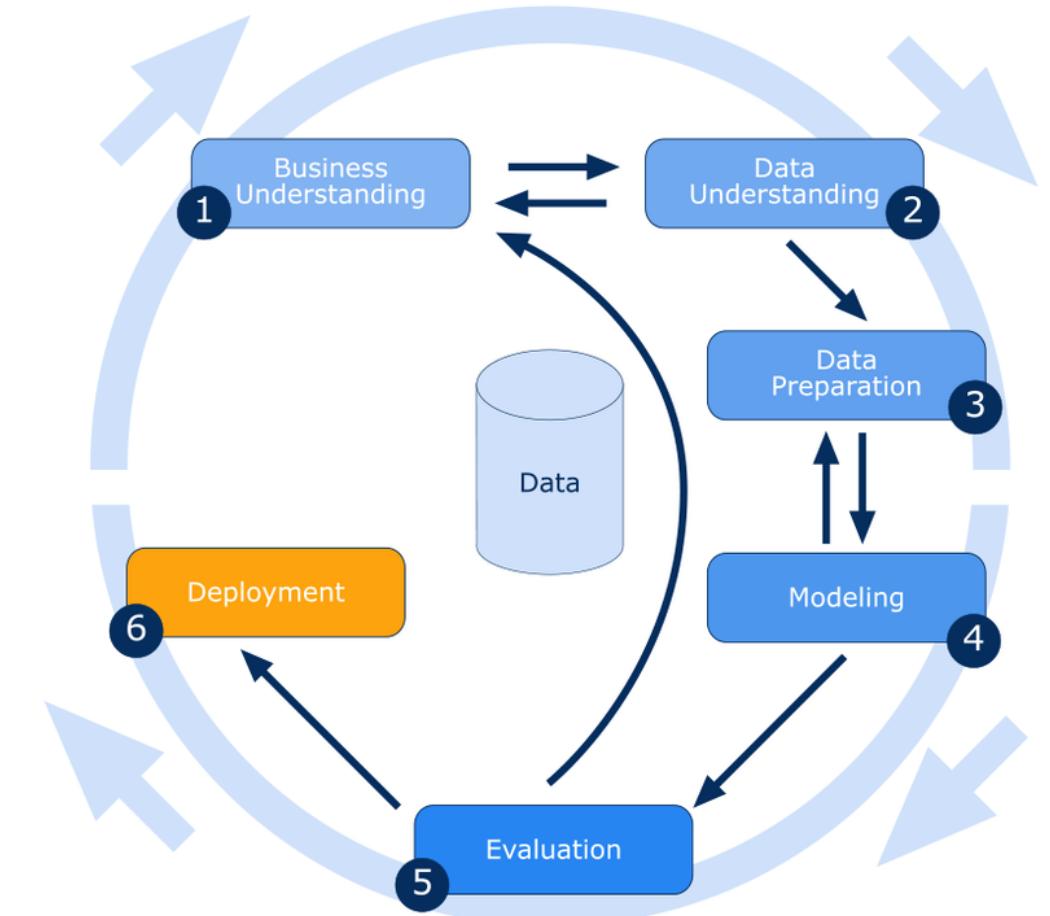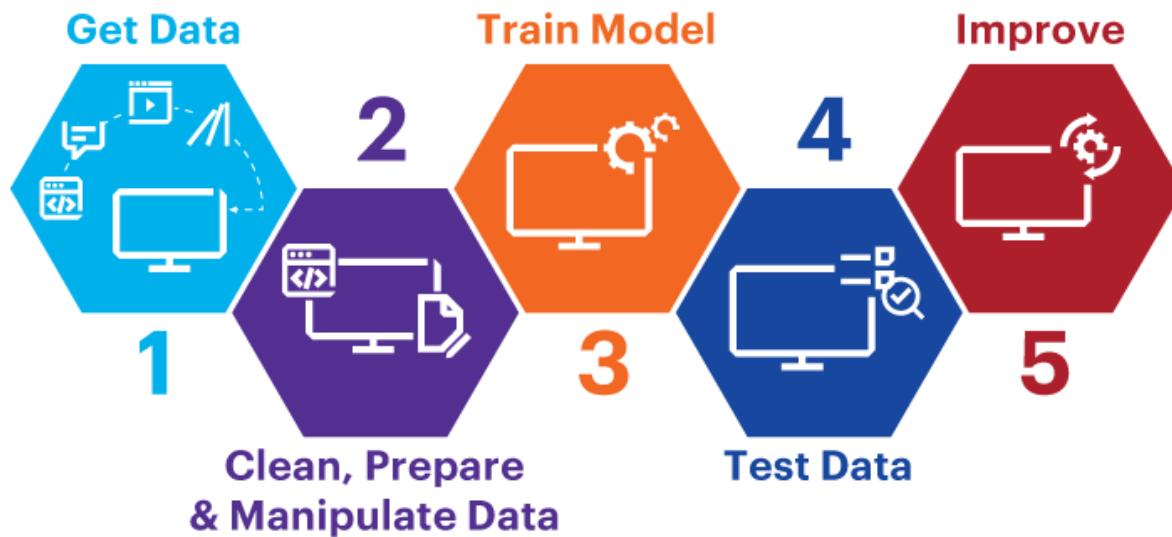
Supervised Learning : Classification & Regression
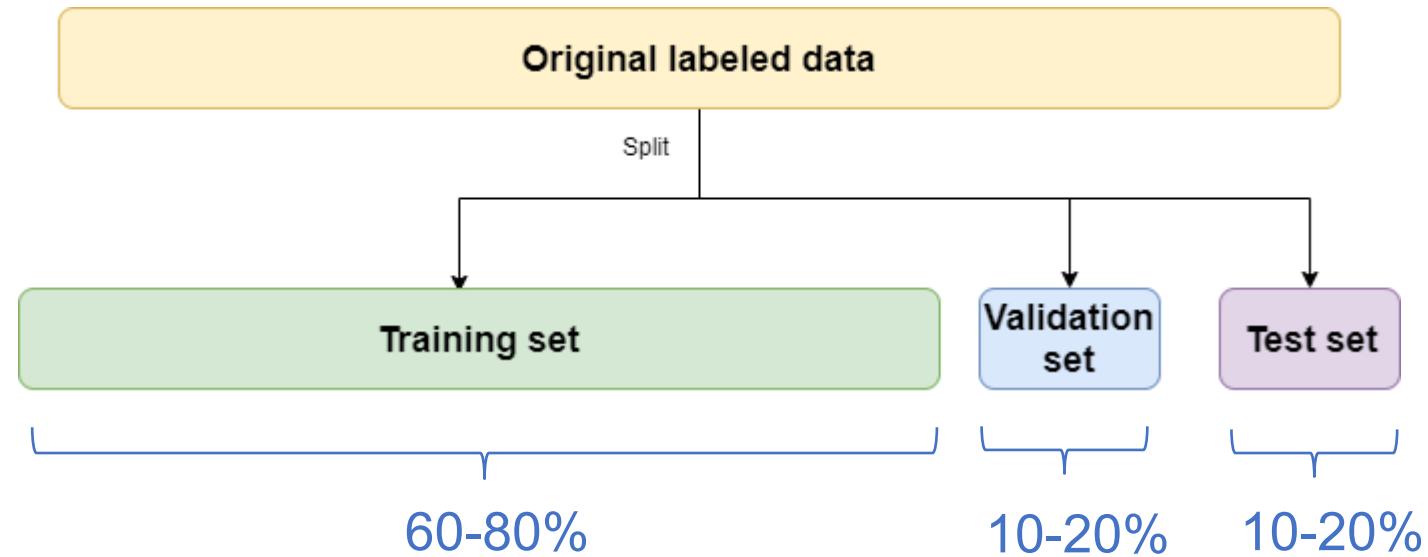
| Features	Regression	Classification
Main goal	Predicts **continuous** values like salary and age.	Predicts **discrete** values like diagnosis and types.
Task	The regression algorithm task is mapping input value (x) with continuous output variable (y).	The classification algorithm task is mapping the input value of x with the discrete output variable of y.
Input and Output variables	Input: Either categorical or continuous Output: Only continuous	Input: Either categorical or continuous Output: Only categorial
Algorithms	Linear regression, Polynomial regression, Lasso regression and Ridge regression	Decision trees, Random forests, Logistic regression, Neural networks, Support vector machines
Evaluation metrics	R2 score, Mean squared error, Mean absolute error, Mean Absolute percentage error (MAPE)	Recall, Accuracy, Precision, F1 score

Regression

Classification

Get Data

1

2 Clean, Prepare & Manipulate Data

Train Model

3 Test Data

4

Improve

5

Business Understanding 1

Data Understanding 2

Data Preparation 3

Data

Modeling 4

Deployment 6

Evaluation 5

CRISP-DM

Supervised Learning : How to use my Data ?

Agenda: Day 1

9 AM – 12 PM:

1. Introduction to Machine Learning

2. Machine Learning types

3. Machine Learning applications

4. Supervised MachineLearning

1 PM – 4 30 PM:

4. Python for Machine Learning

5. Linear Regression

6. Logistic Regression