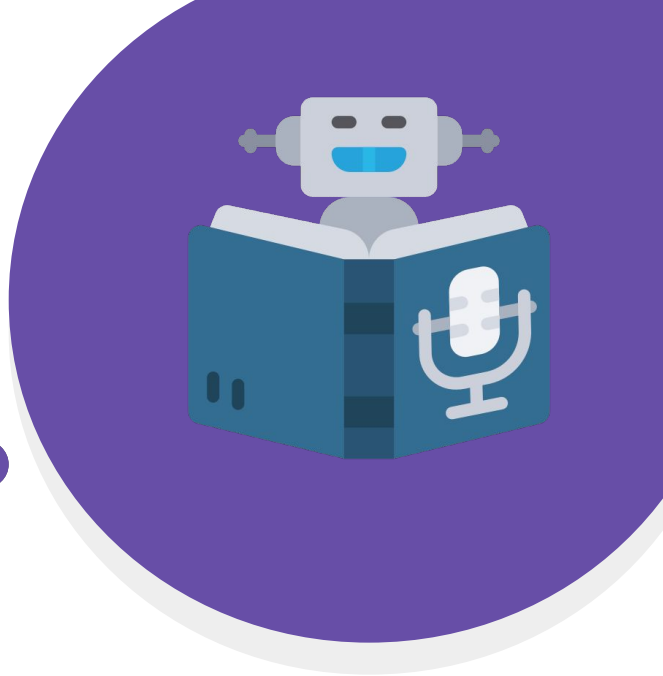
Introduction to Machine Learning



Nisal Mihiranga

Facilitator

Nisal Mihiranga

Microsoft MCT

Areas of Interest & Expertise:
AI, Technology, Science, Teaching, Consulting, Mentoring

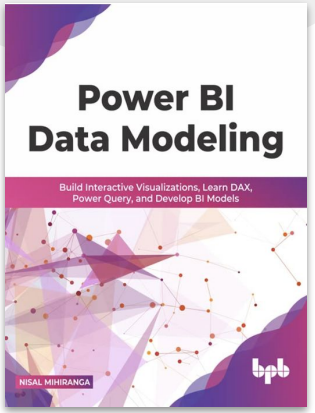
Experience:
Head of AI and Data Science, Architect at
Zone24x7 pvt Ltd
Corporate Trainer

12 Years of Industry exposure to Data Engineering, Data Science and Business Intelligence

Credentials:
M.Sc in Data Science

B.Sc in Information Technology

Microsoft Certified Trainer



Curriculum

PDU
PROFESSIONAL DEVELOPMENT UNIT

Week	Module
Week 1	Python for Machine Learning
Week 2	**Introduction to Machine Learning**
Week 3	Data Transformation and Analysis
Week 4	Regression Analysis
Week 5	Classification, KNN, DT, SVM, Ensemble Systems
Week 6	Clustering Algorithms
Week 7	Neural Networks
Week 8	MLOPS, Machine Learning in Cloud

Agenda

PDU
PROFESSIONAL DEVELOPMENT UNIT

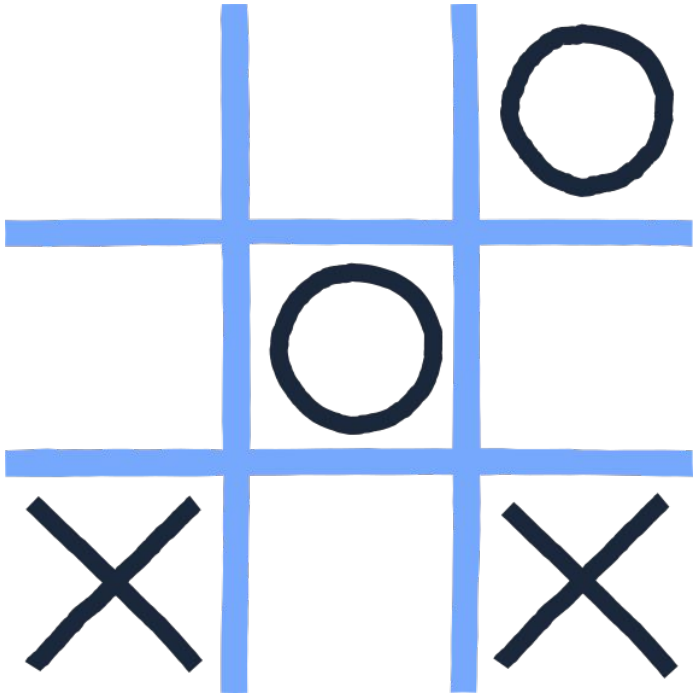
Week	2nd Week
Day	7th of Sep
Duration	4hrs

TIME	TOPIC & ACTIVITY	COMPLETED-
60 Mins	Intro to ML	Yes
60 Mins	Types of ML	Yes
20 Mins	Break	
45 Mins	ML Life cycle	Yes
30 Mins	Frameworks and getting started to develop	Yes
5 Mins	Q&A Session on lesson learn	Yes

Introduction to Machine Learning

Tic Tac Toe Game

PDU
PROFESSIONAL DEVELOPMENT UNIT



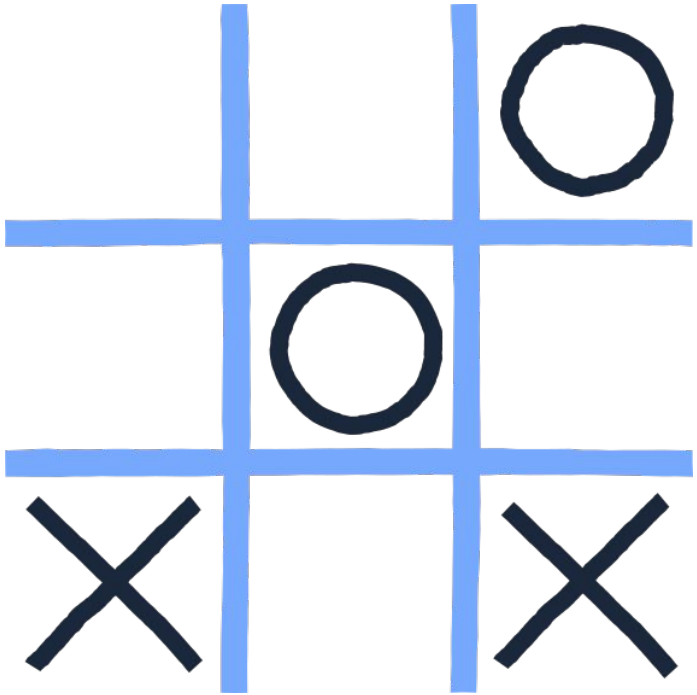
Move 1:
Place an X in a corner.

Move 2:
IF the other player did not place an O in the opposite corner
THEN place an X in the opposite corner to move 1.
ELSE place an X in a free corner.

Move 3:
IF there are 2 Xs and a space in a line
THEN place an X in the free space on that line.
ELSE IF there are 2 Os and a space in a line
THEN place an X in that space.
ELSE place an X in a free corner.

Move 4:
IF there are 2 Xs and a space in a line
THEN place an X in the free space on that line.
ELSE IF there are 2 Os and a space in a line
THEN place an X in that space.
ELSE place an X in a free corner.

Move 5:
Place an X in the free space.



```
def make_move_1():
    return "Place X in a corner"

def make_move_2(opponent_move):
    if opponent_move != "opposite corner":
        return "Place X in the opposite corner to move 1"
    else:
        return "Place X in a free corner"

def make_move_3(board_state):
    if "2 Xs in a line" in board_state:
        return "Place X in the free space on the line"
    elif "2 Os in a line" in board_state:
        return "Place X in that space to block"
    else:
        return "Place X in a free corner"

def make_move_4(board_state):
    if "2 Xs in a line" in board_state:
        return "Place X in the free space on the line"
    elif "2 Os in a line" in board_state:
        return "Place X in that space to block"
    else:
        return "Place X in a free corner"

def make_move_5():
    return "Place X in the free space"

def play_game():
    print(make_move_1())
    opponent_move = "side"
    print(make_move_2(opponent_move))
    board_state = ["no win or block available"]
    print(make_move_3(board_state))
    print(make_move_4(board_state))
    print(make_move_5())

play_game()
```

Key points in Rule-based

PDU
PROFESSIONAL DEVELOPMENT UNIT

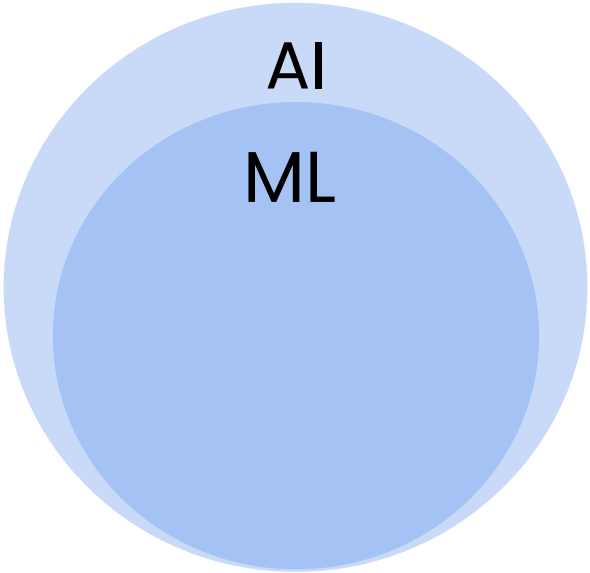
Fixed Logic: This follows strict rules that do not change, regardless of how many games it plays. There is no learning or adaptation.
Predictability: The system always behaves in the same way given the same conditions.
Manual Setup: All rules must be predefined by humans, which means the system will not improve unless we explicitly add more complex rules.

The system cannot adapt or learn. It works well for structured, deterministic tasks like Tic-Tac-Toe, but as the complexity of the problem increases (e.g., Chess or Go), writing explicit rules becomes extremely difficult.

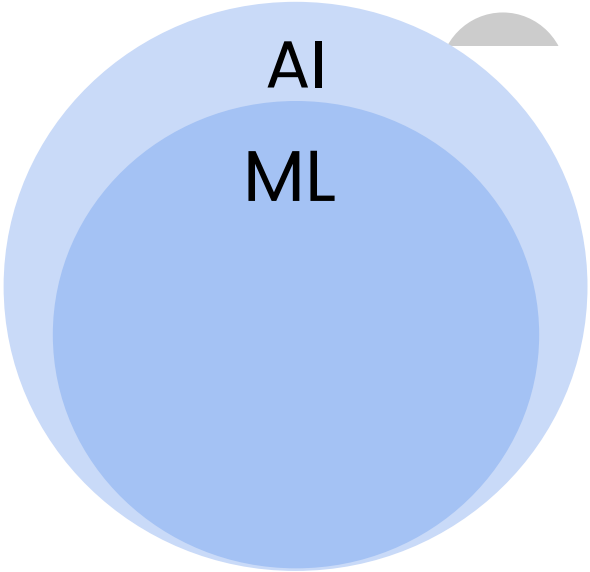
What is Machine Learning

PDU
PROFESSIONAL DEVELOPMENT UNIT

ML is a subset of AI that focuses on algorithms that allow machines to **learn from data** and improve their performance over time without being explicitly programmed.



What is Machine Learning

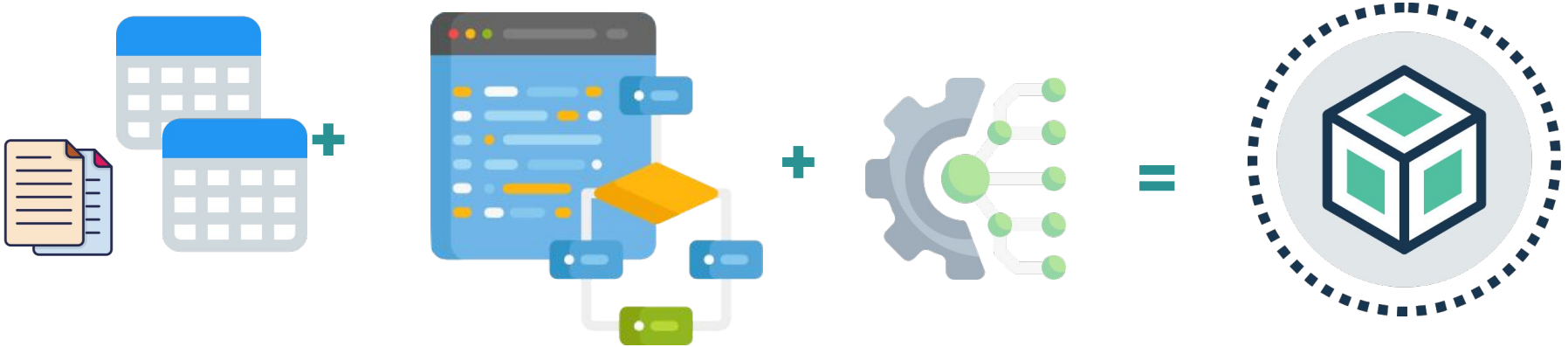


Unlike rule-based systems, machine learning (ML) algorithms allow a computer to learn patterns from data rather than follow predefined rules. This behavior allow machine learning to solve complex problems which involves training on historical data.

What is a ML Model

PDU
PROFESSIONAL DEVELOPMENT UNIT

A ML Model is a representation of a real-world context.
Train on historical Data



Data + Algorithms + Training = ML Model

Types of Data

PDU
PROFESSIONAL DEVELOPMENT UNIT

Structured Data

SQL Databases,
Excel File Data

Semi-structured Data

XML/JSON Data

Unstructured Data

Video, Audio, Image
Data, Documents

Activity

Classroom Activity

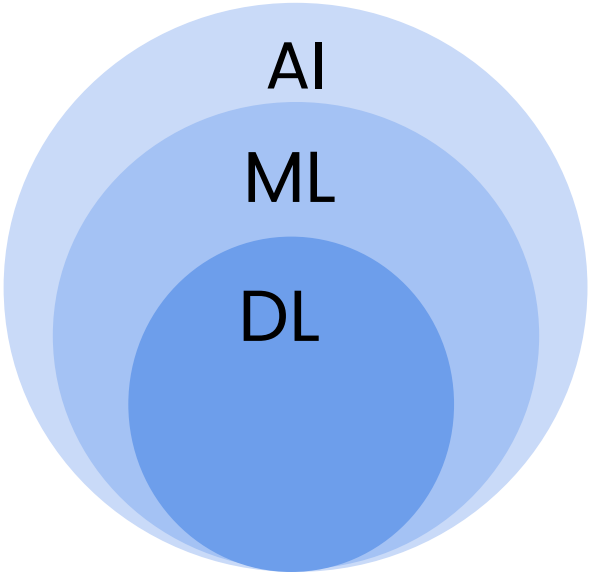
Duration: *10 mins*

1. Discuss real-world examples for different data types

What is Deep Learning

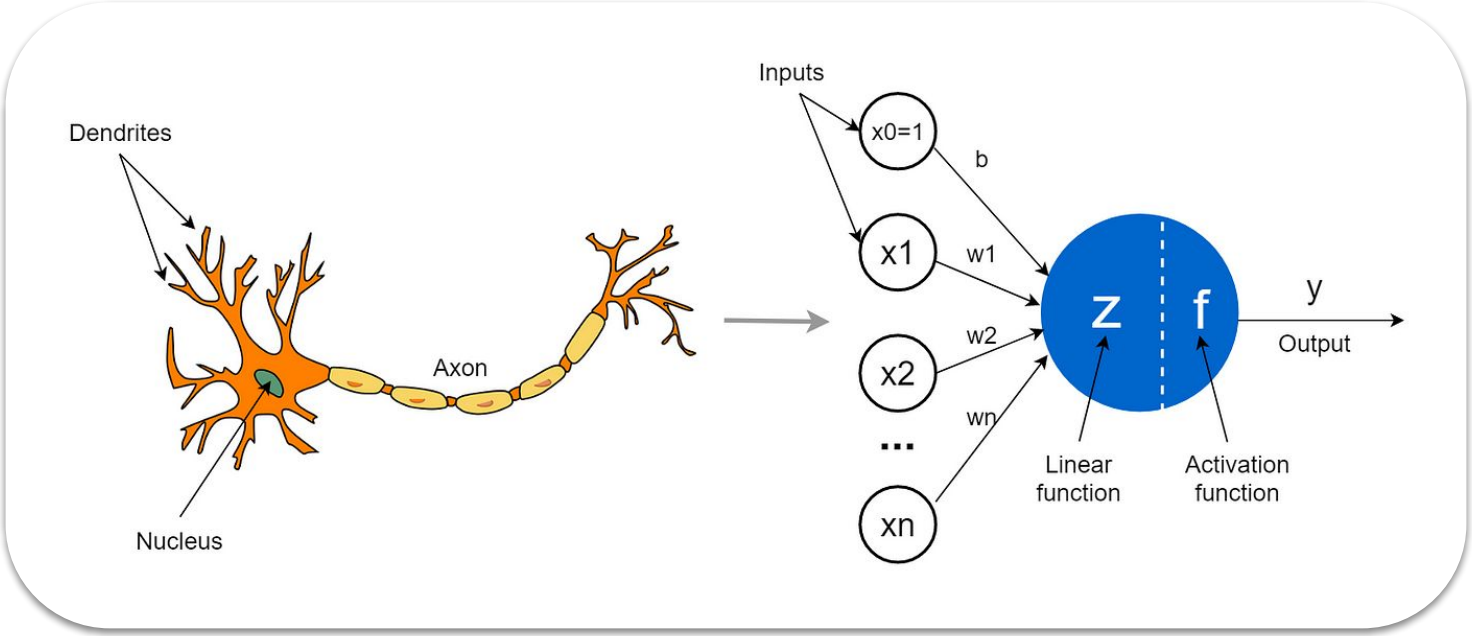
PDU
PROFESSIONAL DEVELOPMENT UNIT

DL is a specialized subset of ML that uses neural networks with many layers (hence “deep”) to analyze various factors of data. These neural networks attempt to simulate the behavior of the human brain, enabling it to “learn” from large amounts of data.



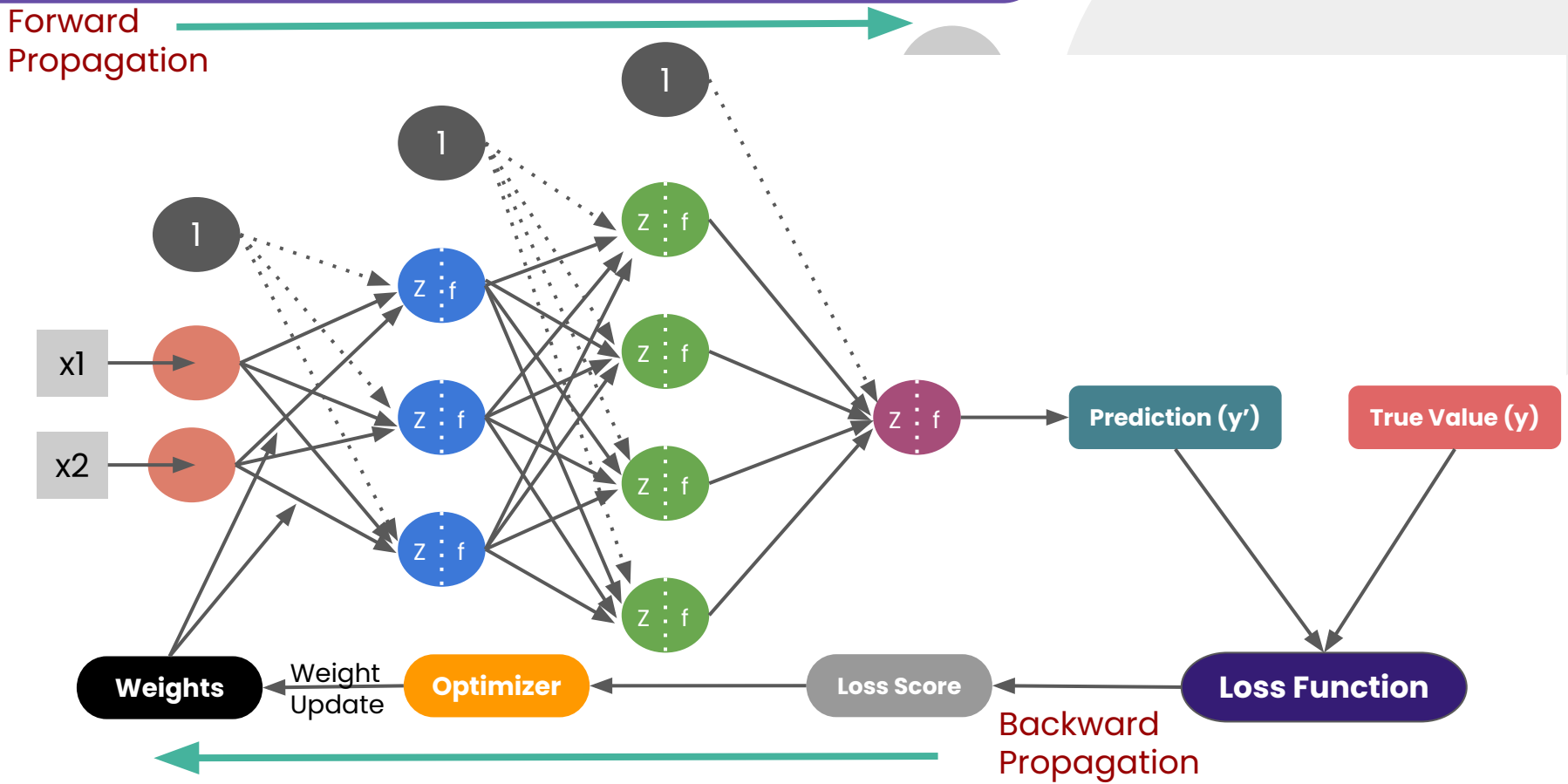
Neural Networks

- Deep learning primarily uses neural networks, which are computational models inspired by the human brain



Structure of a Neural Net

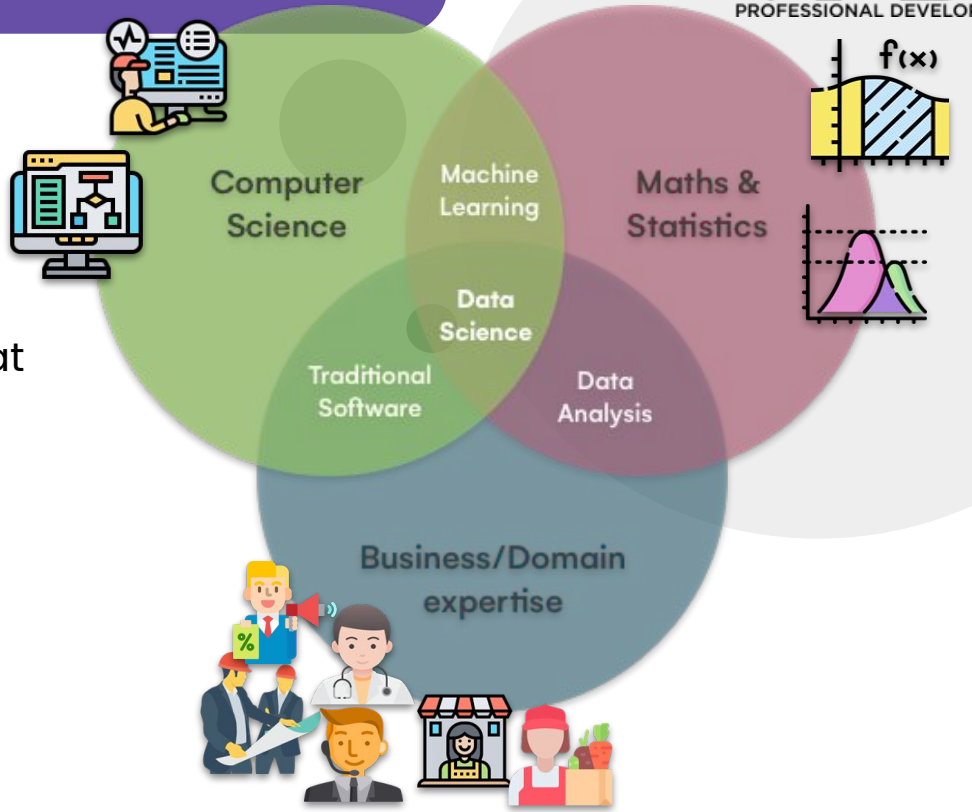
PDU
PROFESSIONAL DEVELOPMENT UNIT



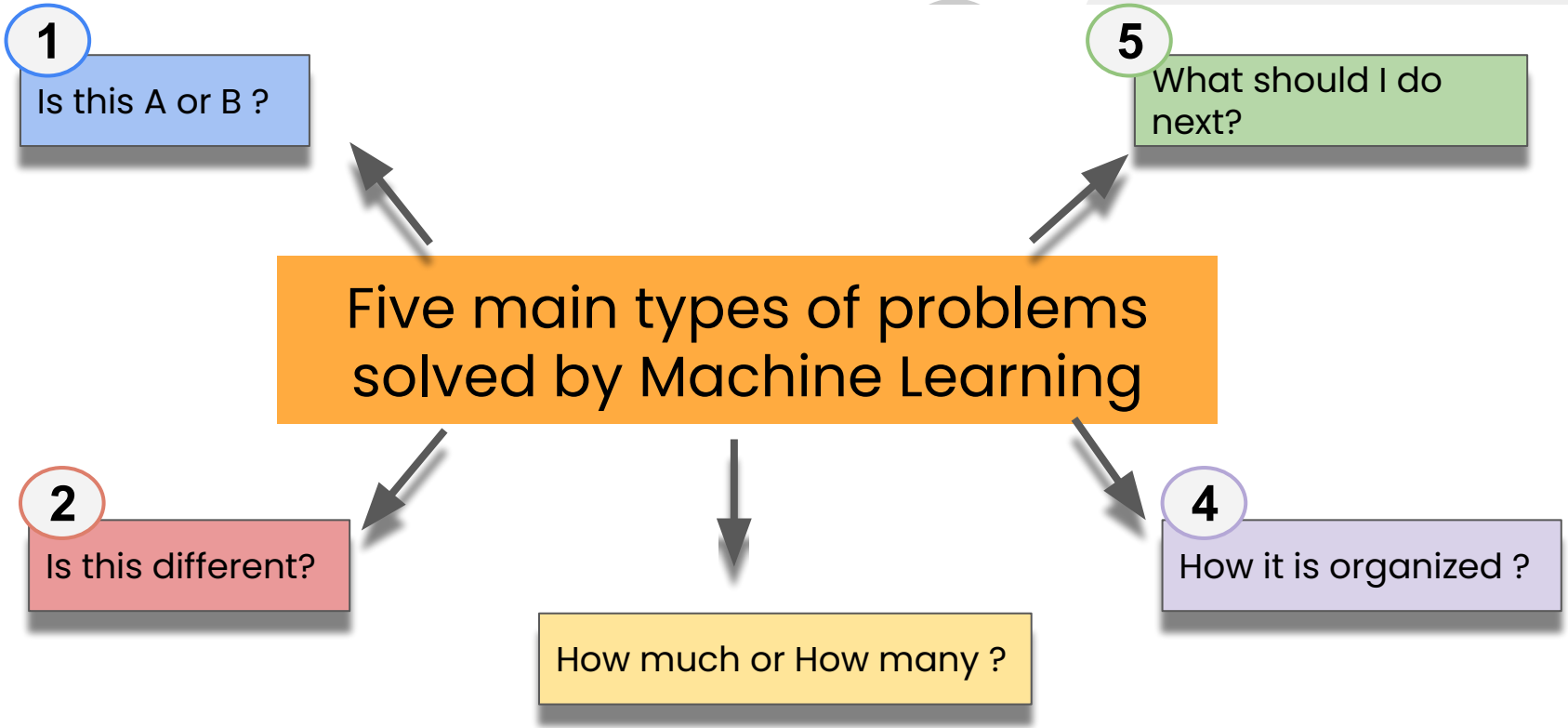
What Is Data Science?

PDU
PROFESSIONAL DEVELOPMENT UNIT

Data science is the field of study that combines domain expertise, computer science programming skills, and knowledge of mathematics and statistics to extract meaningful insights from data.



Main Types of ML Problems ?

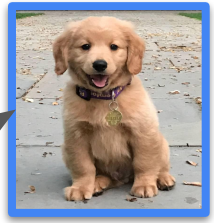


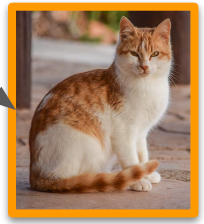
Classification

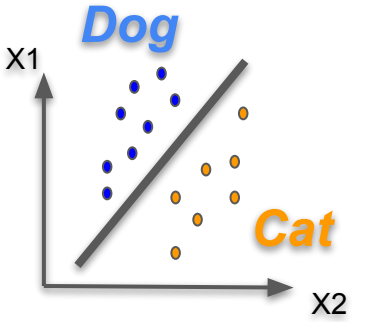
1

Is this A or B ?

Classify Cat ? or Dog ?







Supervised learning

Activity

Classroom Activity

Duration: *10 mins*

1. Discuss real world use cases of Classification

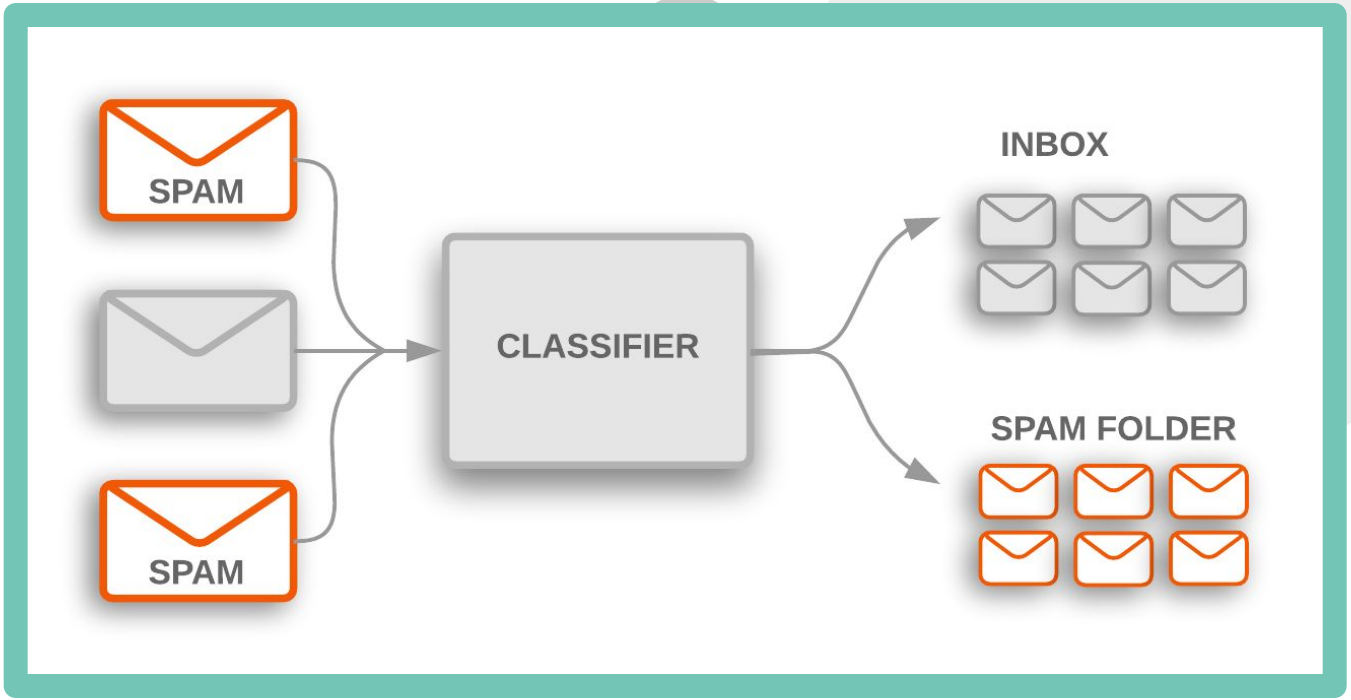
Classification

Classify Spam Emails

PDU
PROFESSIONAL DEVELOPMENT UNIT

1

Is this A or B ?

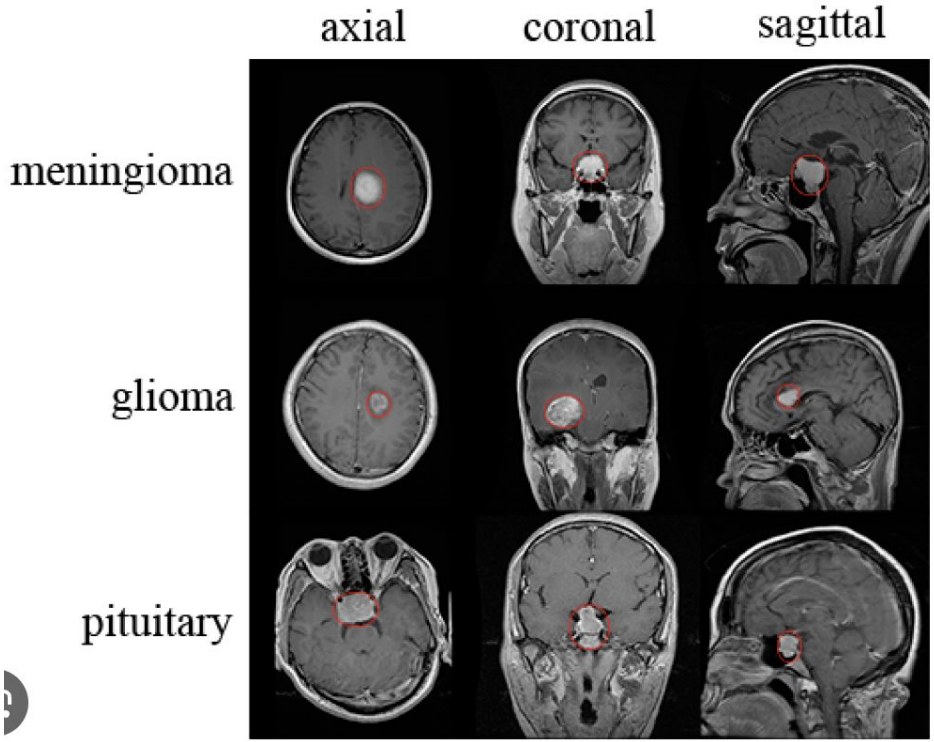


Classification

1

Is this A or B ?

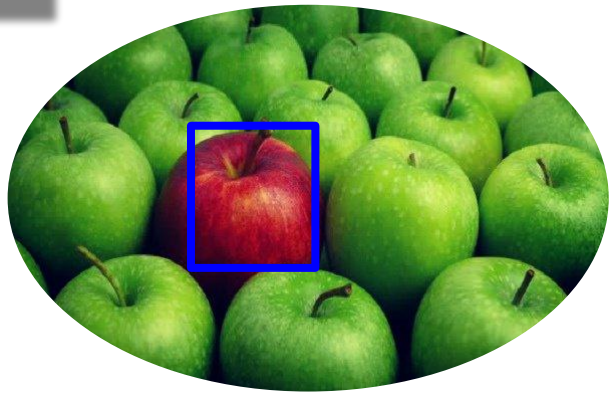
Classify brain tumors from MRI



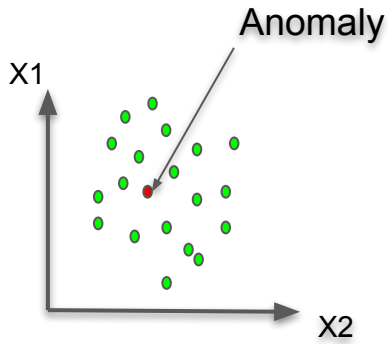
Anomaly Detection

2

Is this different?



Detect the red apple in the box of green apples



Both Supervised and unsupervised learning

Activity

Classroom Activity

Duration: *10 mins*

1. Discuss real world use cases of Anomaly Detection

Anomaly Detection

2

Is this different?

Regression Analysis

3

How much or How many ?

How much will this cost ?

Activity

Classroom Activity

Duration: *10 mins*

1. Discuss real world use cases of Regression Analysis

Regression

Stock Price Prediction

PDU
PROFESSIONAL DEVELOPMENT UNIT

3

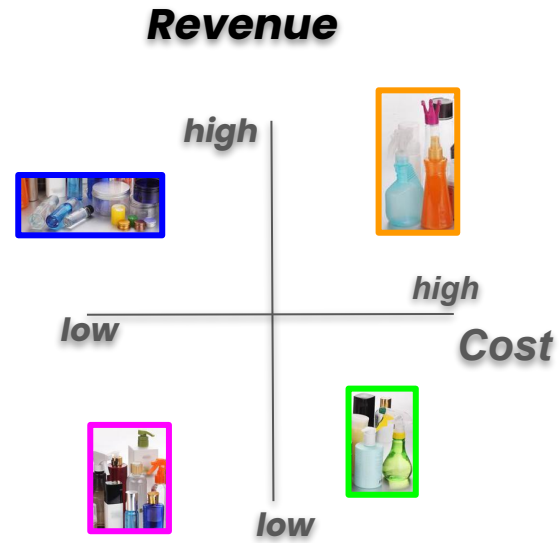
How much or How many ?

Clustering

4

How it is organized ?

How is products portfolio organized ?



Unsupervised learning

Activity

Classroom Activity

Duration: *10 mins*

1. Discuss real world use cases of Cluster Analysis

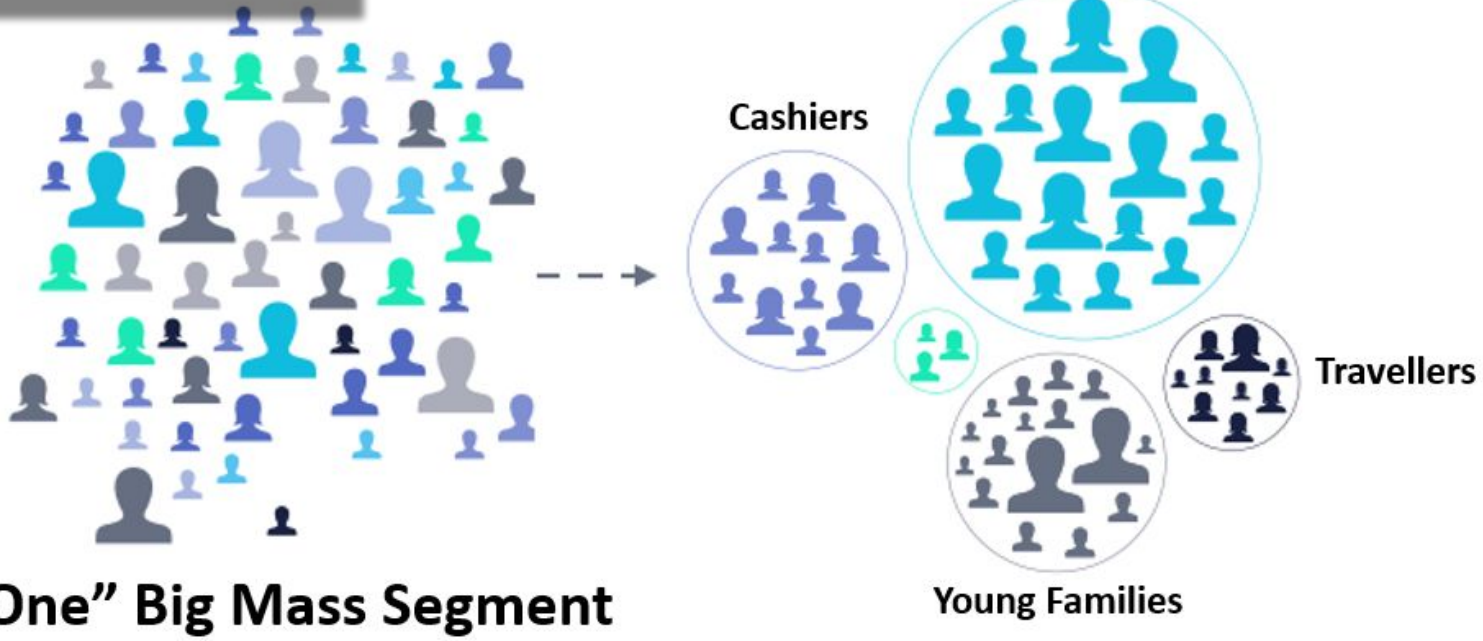
Clustering

Segment customers and target marketing

PDU
PROFESSIONAL DEVELOPMENT UNIT

4

How it is organized ?



Reinforcement Learning

PDU
PROFESSIONAL DEVELOPMENT UNIT

RL is a type of machine learning where an agent learns to make decisions by interacting with an environment. The agent receives rewards or penalties (positive or negative feedback) based on its actions and aims to maximize cumulative reward over time.

Reinforcement Learning

5

What should I do next?

PDU
PROFESSIONAL DEVELOPMENT UNIT

Command

Bring me the ball

Action

Reward

No toys

Learn from mistakes

Activity

Classroom Activity

Duration: *5 mins*

1. Discuss real world use cases of Reinforcement Learning

Reinforcement Learning

5

What should I do next?

Training a self-driving car to navigate complex environments and make safe driving decisions.

20 mins Break

PDU
PROFESSIONAL DEVELOPMENT UNIT

Developing Machine Learning Models

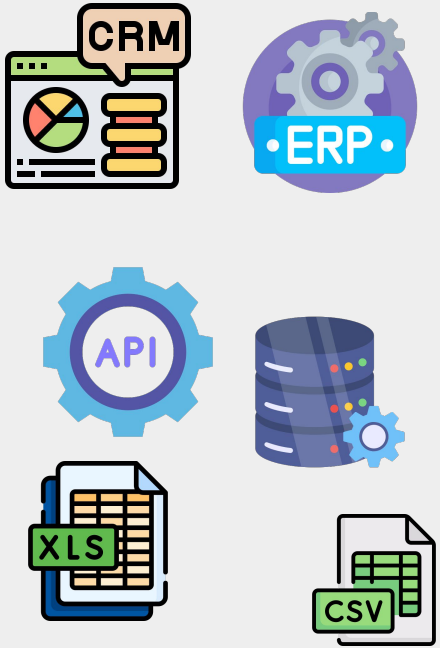
Libraries for Machine Learning

PDU
PROFESSIONAL DEVELOPMENT UNIT

Data Sources for ML

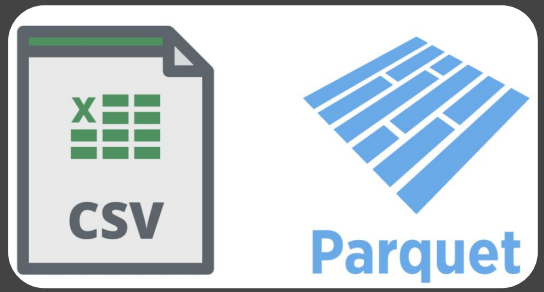
PDU
PROFESSIONAL DEVELOPMENT UNIT

Operational Data



ETL

Data for Machine Learning



Data Splitting

PDU
PROFESSIONAL DEVELOPMENT UNIT

Data Splitting is a fundamental concept in machine learning used to evaluate the performance of a model. The process involves dividing a dataset into two main subsets: the training set and the testing set.

Training set

Used to train the machine learning model. The model learns the relationships between the input features and the target variable from this subset.

Testing set

Used to evaluate the performance of the trained model. It provides an unbiased assessment of how well the model generalizes to new, unseen data.

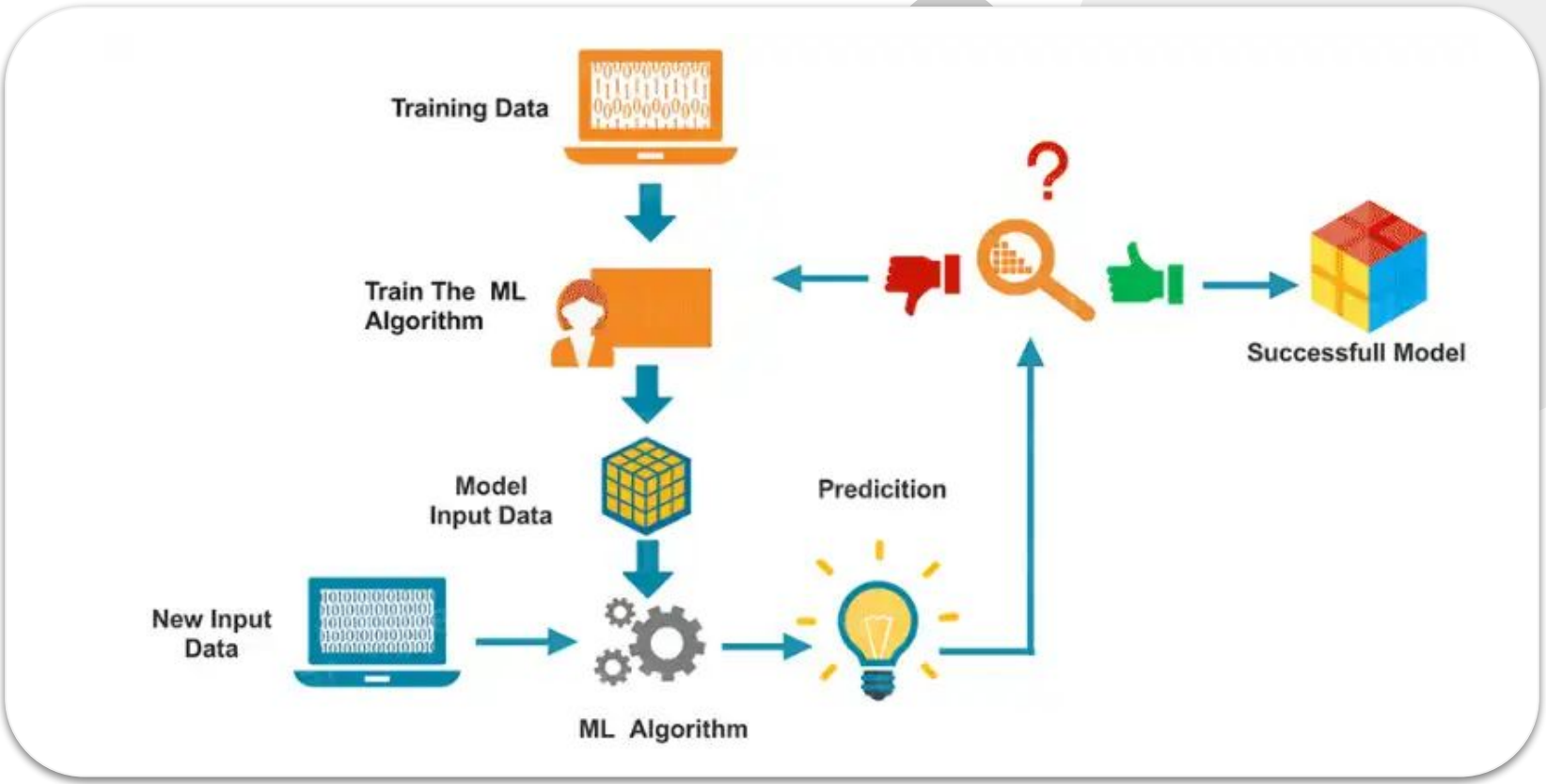
Data Splitting

PDU
PROFESSIONAL DEVELOPMENT UNIT

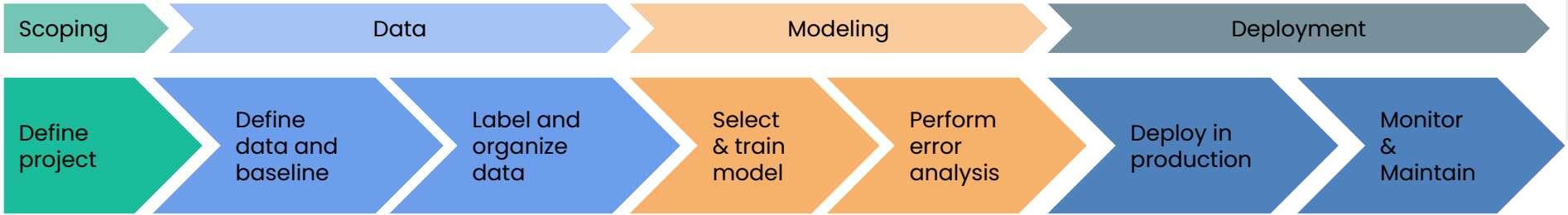
Data Splitting is a fundamental concept in machine learning used to evaluate the performance of a model. The process involves dividing a dataset into two main subsets: the training set and the testing set.

80% Training / 20% Testing is the most widely used split ratio, suitable for many machine learning tasks. It works well when you have a reasonably large dataset. **70%/30%** ratio used when you want to allocate a larger portion of the data for testing, especially in cases where ensuring good model evaluation is critical or when you have a smaller dataset.

How It Works


Training Data
Train The ML Algorithm
Model Input Data
New Input Data
ML Algorithm
Predicition
Successfull Model

Machine Learning Life Cycle



Q&A

PDU
PROFESSIONAL DEVELOPMENT UNIT

Thank You

Nisal Mihiranga
Head of AI & DS at **Zone24x7**
Consultant, Trainer
M: +94 71 726 3044