Introduction to Deep Learning

Jawwad A Shamsi

Dean FAST NUCES

CEO http://syslab.ai & http://iParhai.com

Jawwad.shamsi@nu.edu.pk

About myself

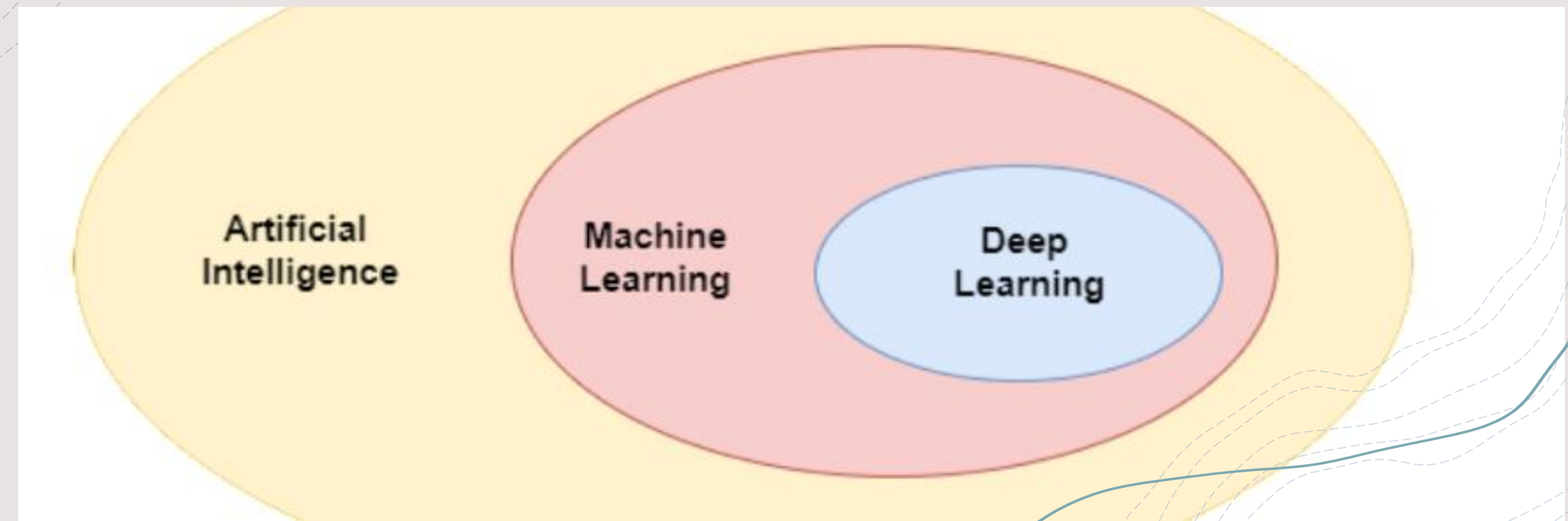
- Dean FAST NUCES
- iParhai (startup)
- Previously
  - Director Campus
  - HoD Computer Science
- Education
  - PhD. Computer Science
    - Wayne State Univ, MI, USA

Topics

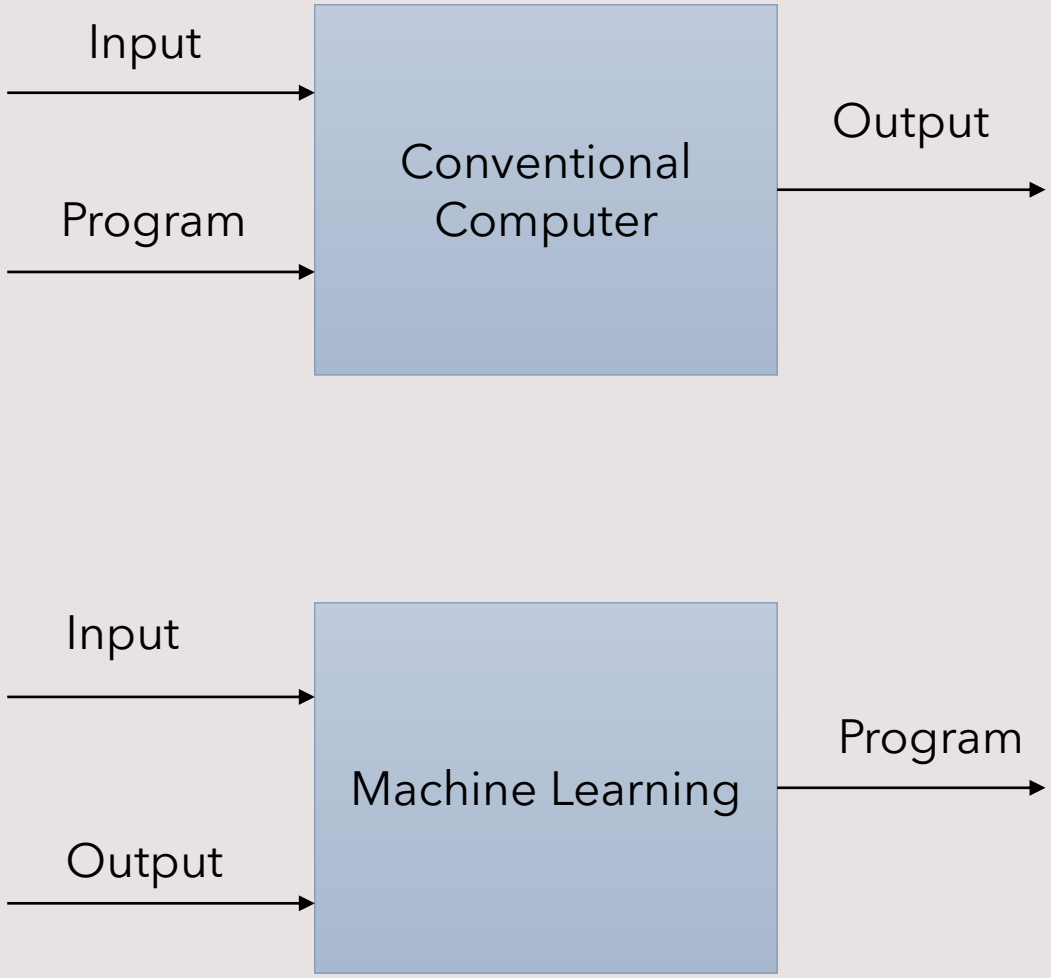
+ Introductory Lecture on Deep Learning
+ AI vs ML DL
+ Features
+ Neural Networks
+ Activation Functions
+ MLP
+ Applications
+ Types of DNNs
+ Hands-on

AI vs ML and DL

- AI: Train Machines so they mimic humans
- ML: A subset of Artificial Intelligence that uses statistical learning algorithms. Has the ability to automatically learn and improve from experiences
- DL: Subset of ML. Feature selection is integrated with classification.



ML vs Programming



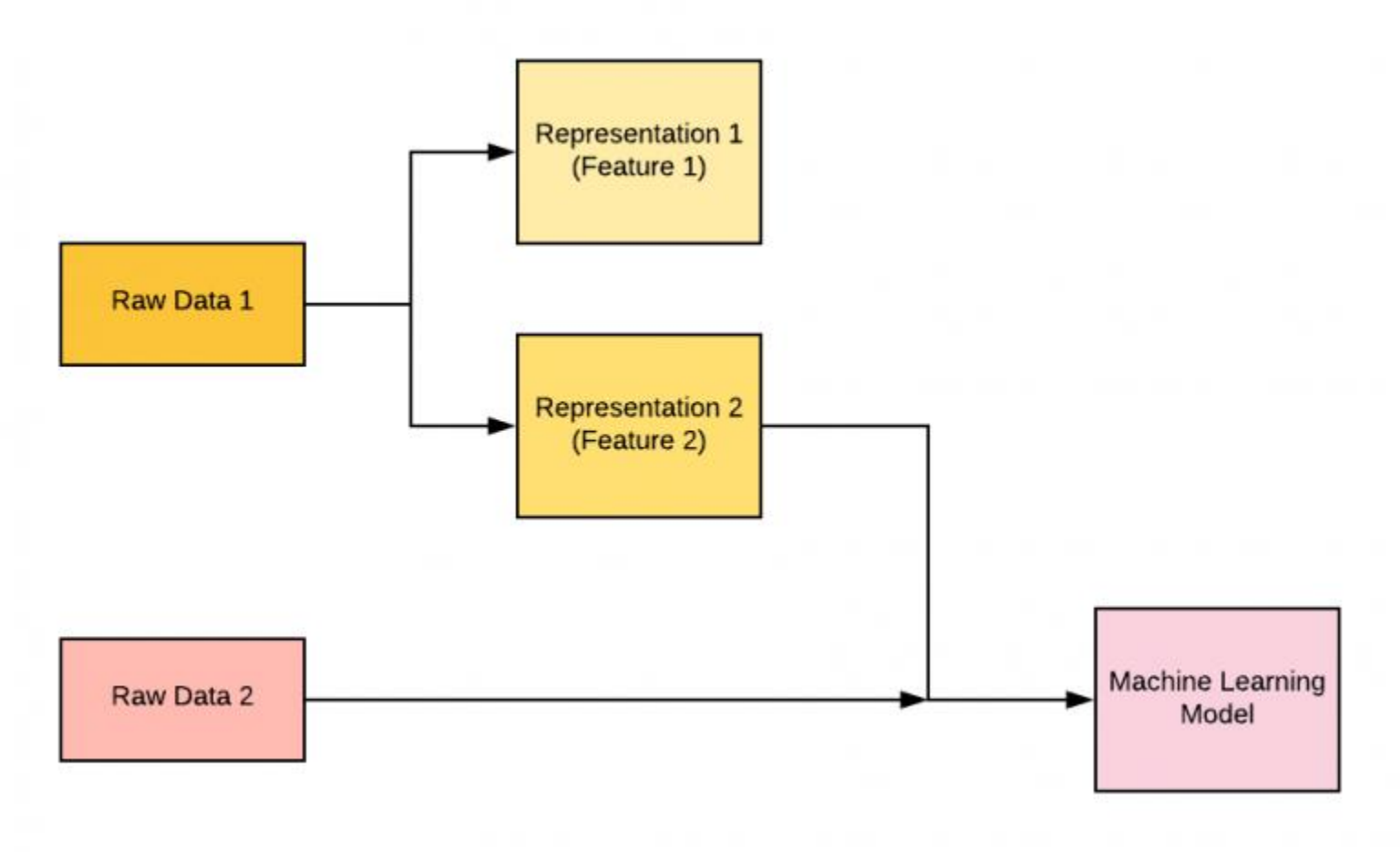
Features

- Set of independent variables
  - Learned to predict the outcome

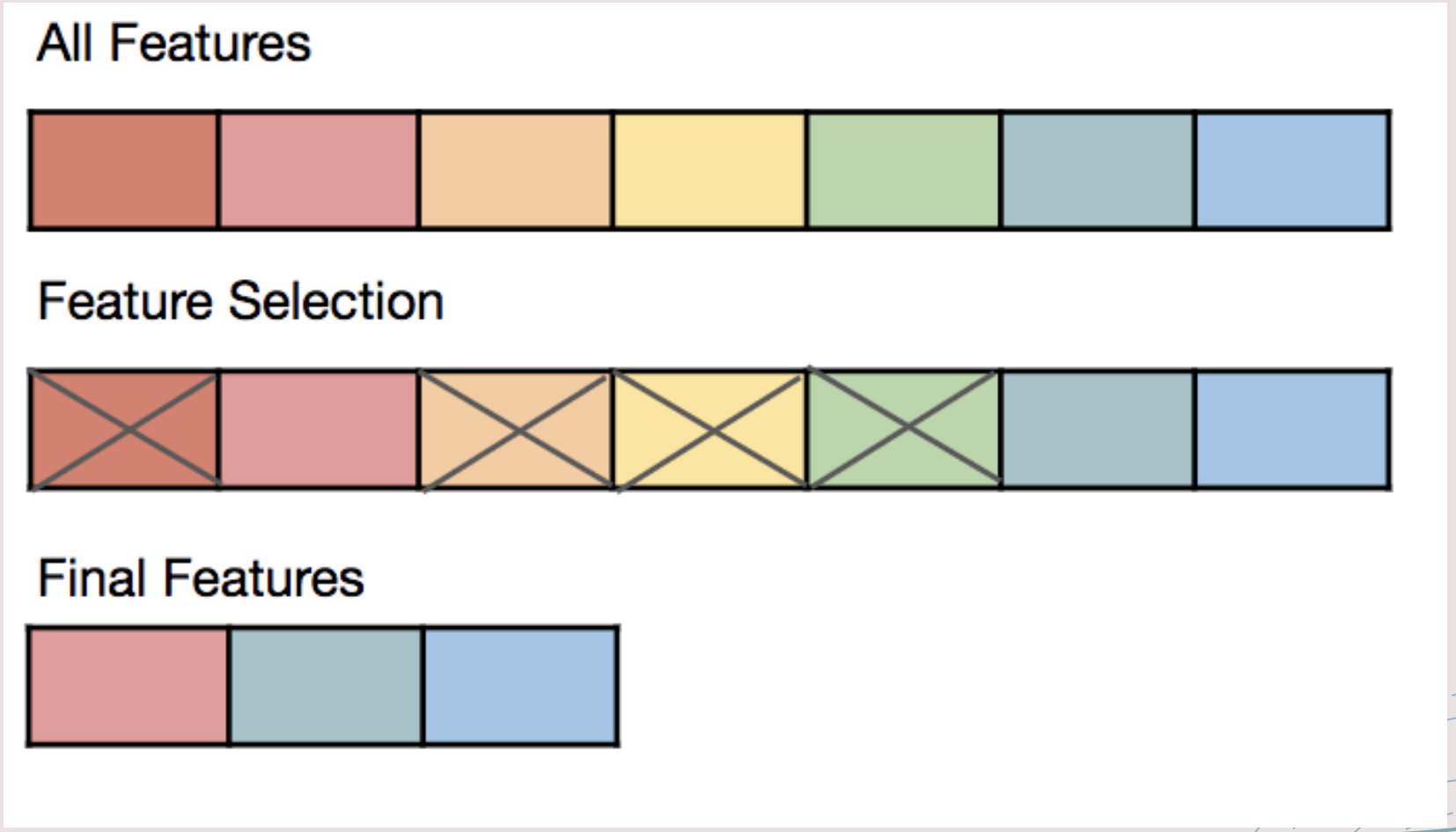
Example of Features

- Predict the Grade ?
  - No. of lectures attended
  - No of Hours of Self Study
  - Previous CGPA

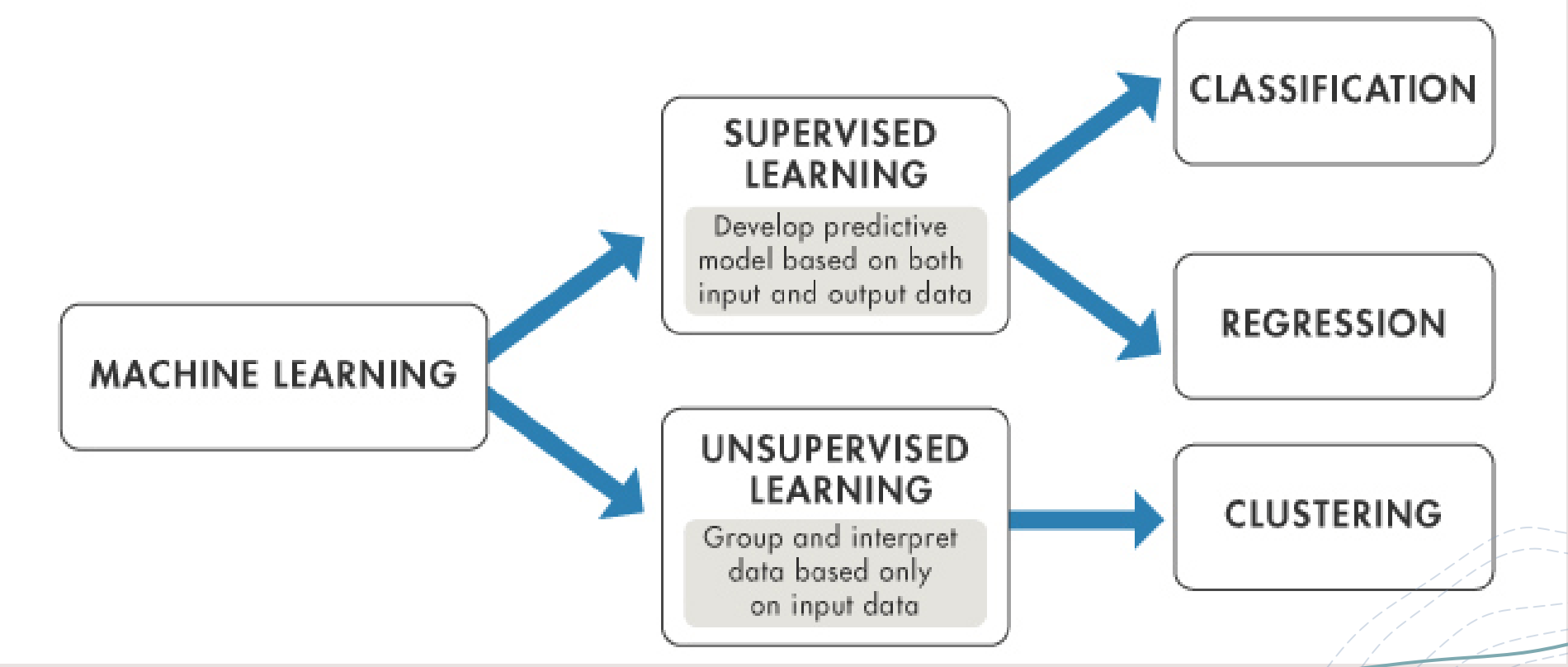
Features


Representation 1
(Feature 1)
Raw Data 1
Representation 2
(Feature 2)
Raw Data 2
Machine Learning
Model

Feature Selection

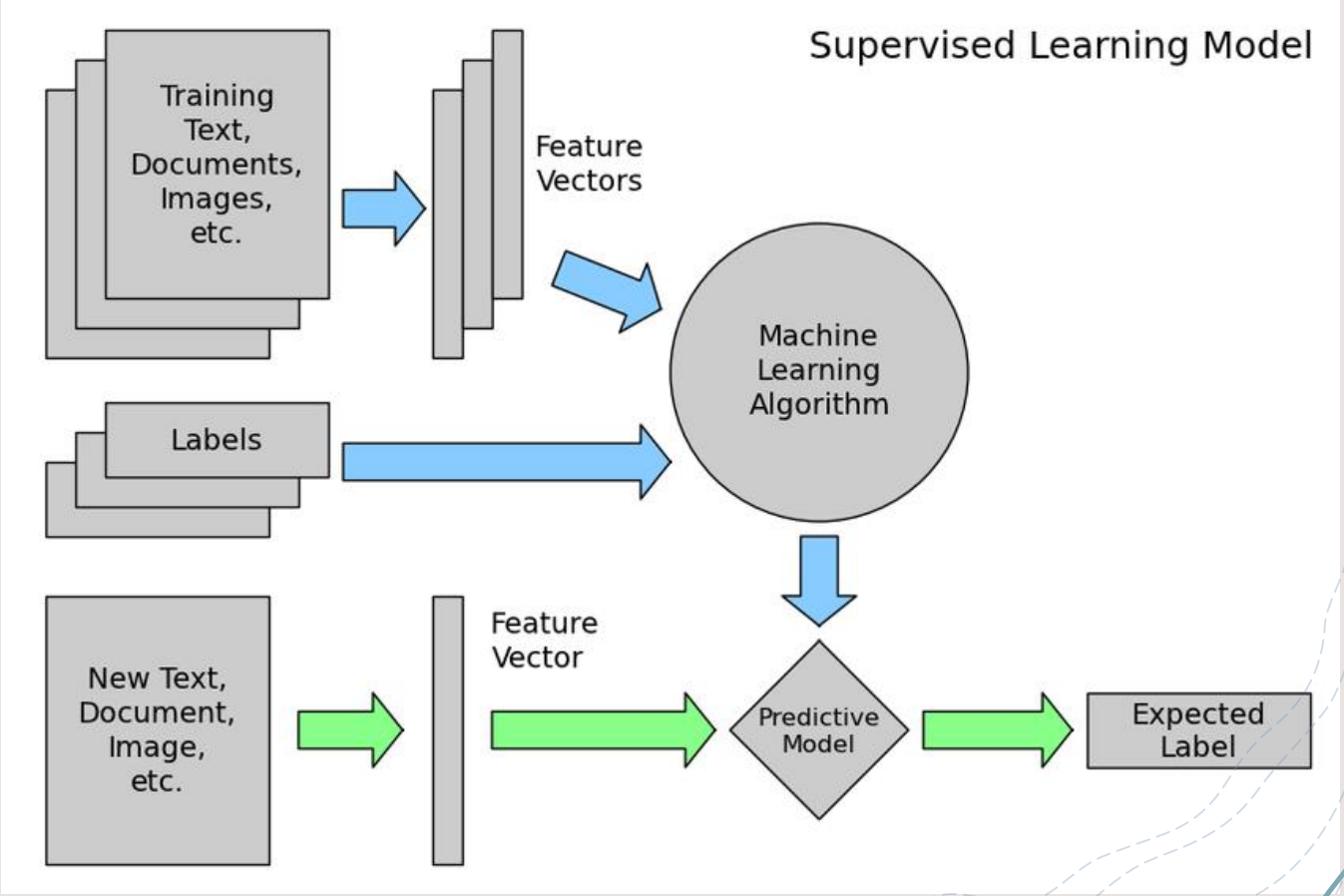


Overview of ML



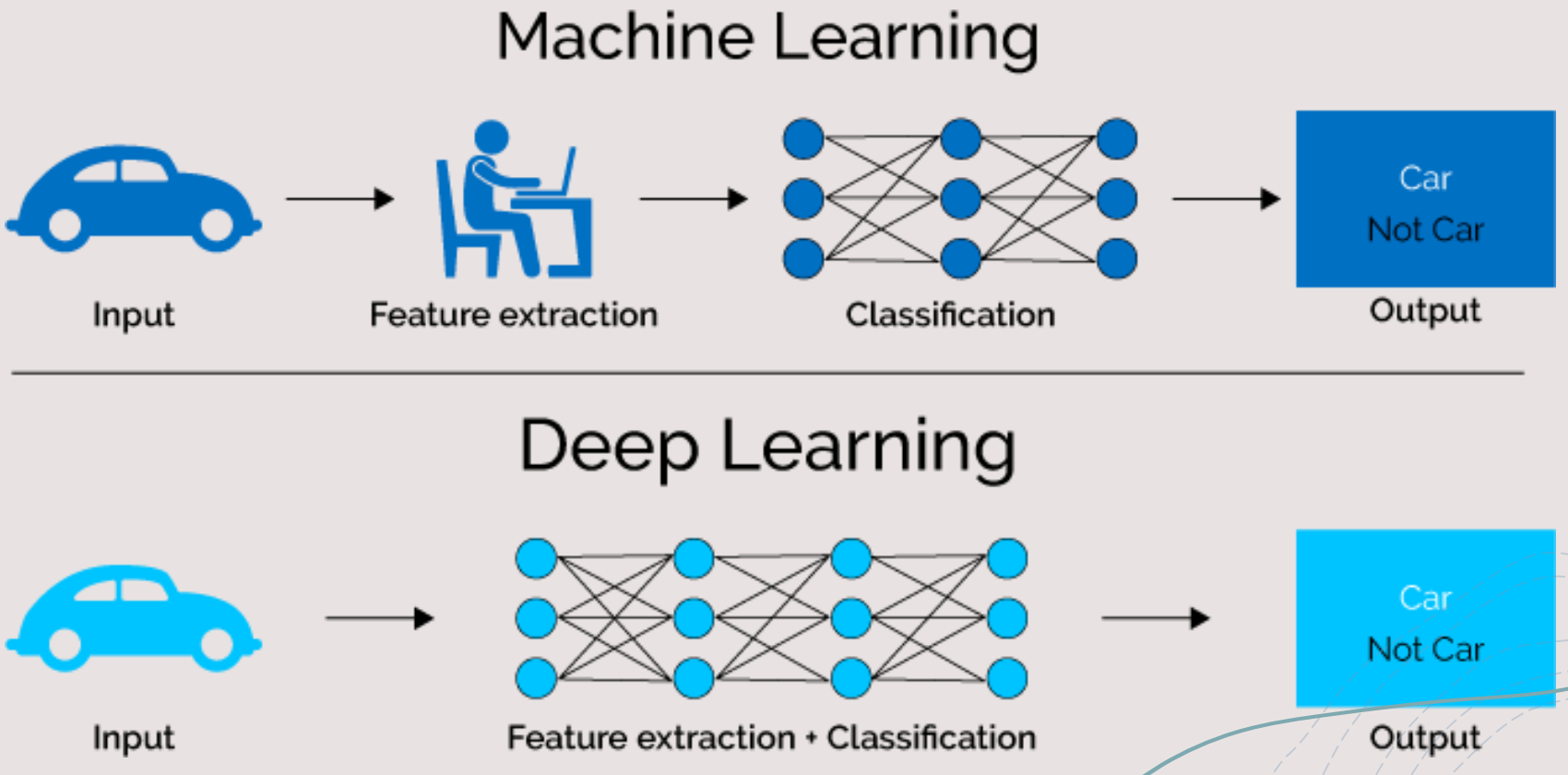
Classification ML Pipeline

- Pre-process data
- Extract Features (Attributes) on the basis of which a label can be predicted
- Develop rules (Train)
- Evaluate rules (Test)
- Iterate between 3 and 4



Deep Learning Pipeline

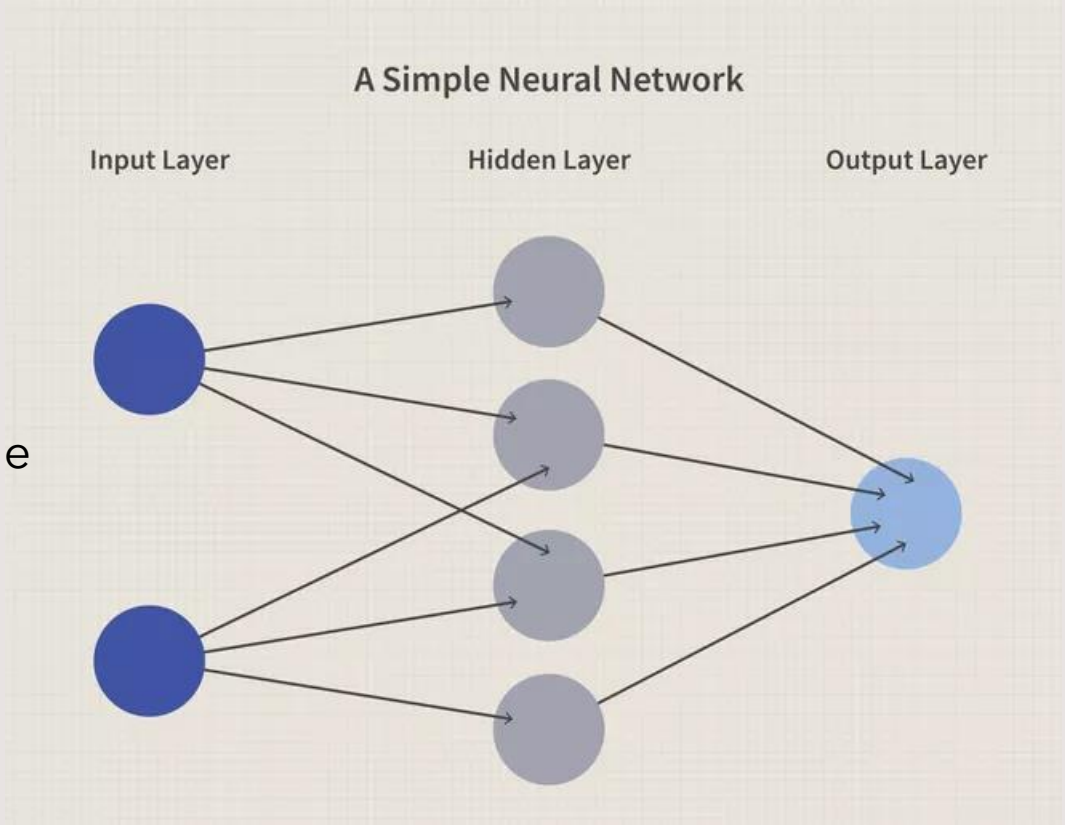
+ In DL, there is no dedicated process of feature extraction. It is bundled with classification



Neural Network

Input layer: Features

Output Layer: Predictor Variable



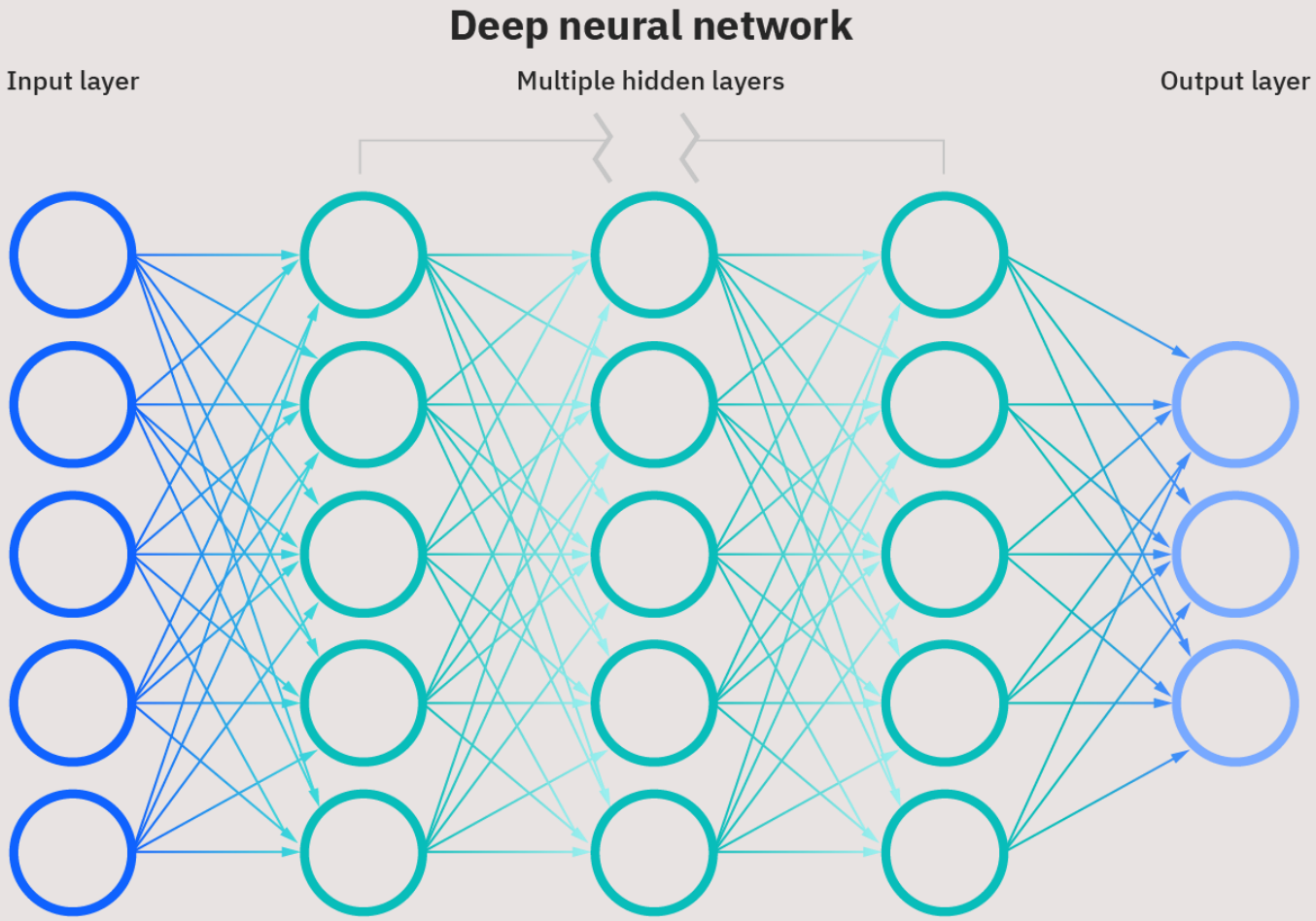
Hidden layer: Non-linear transformations

image source: Investopedia.com

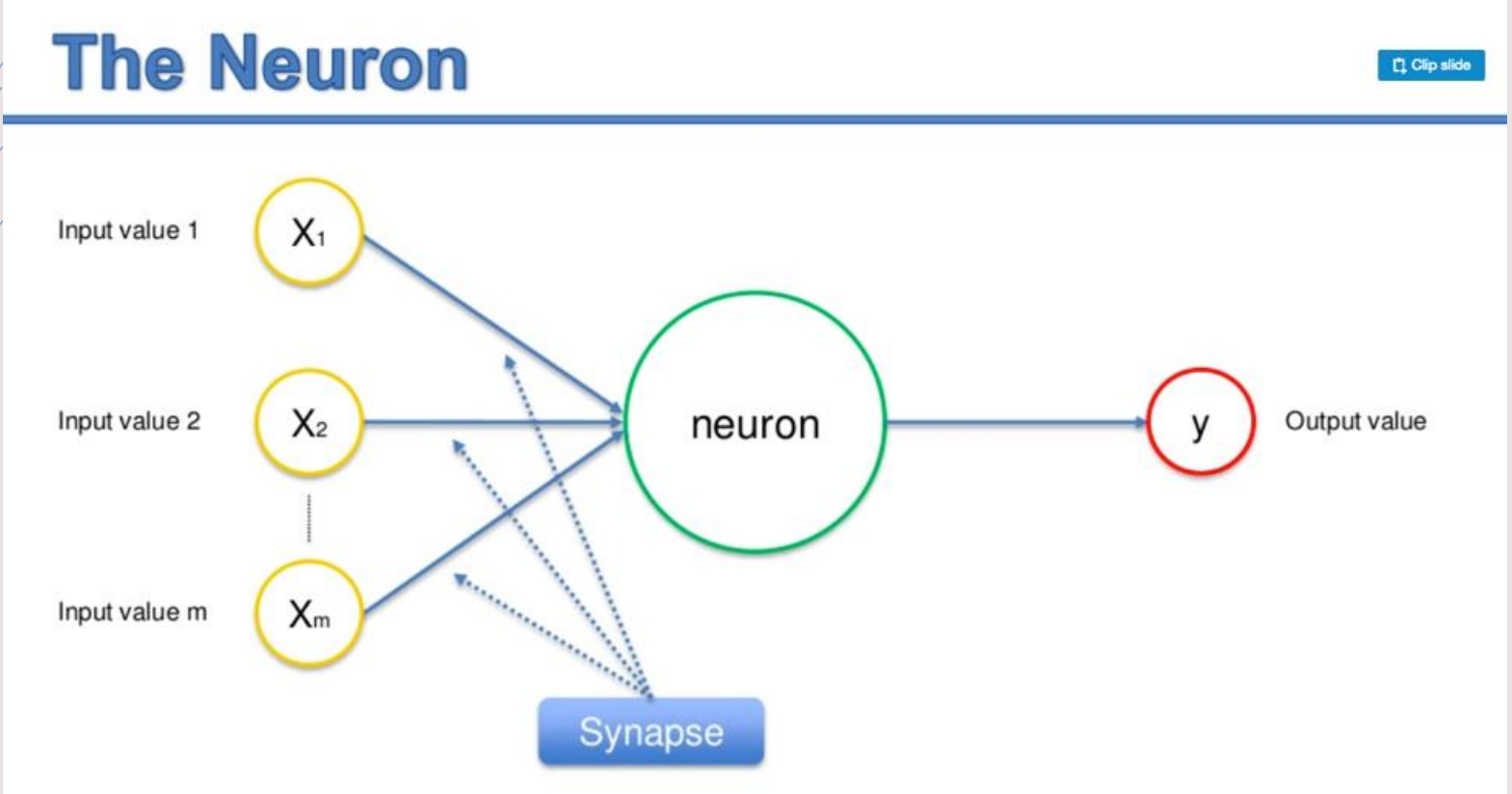
Why Non-Linear Transformations

- Most of the data relationships are non-linear

Deep Neural Network



Neuron



- Neuron is a smallest unit of Computation
- Takes weighted sum of inputs
- Applies activation function
- Computes output

Activation Function

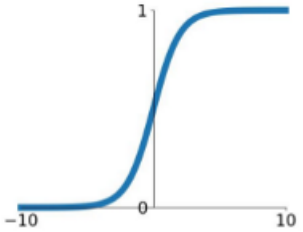
+ Mimics biological Neuron
+ Transforms an input to output

A few Activation Functions

Activation Functions

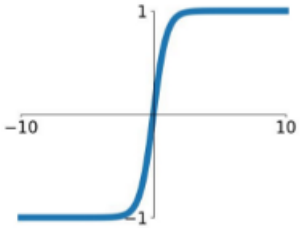
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



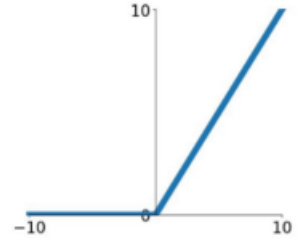
tanh

$$\tanh(x)$$

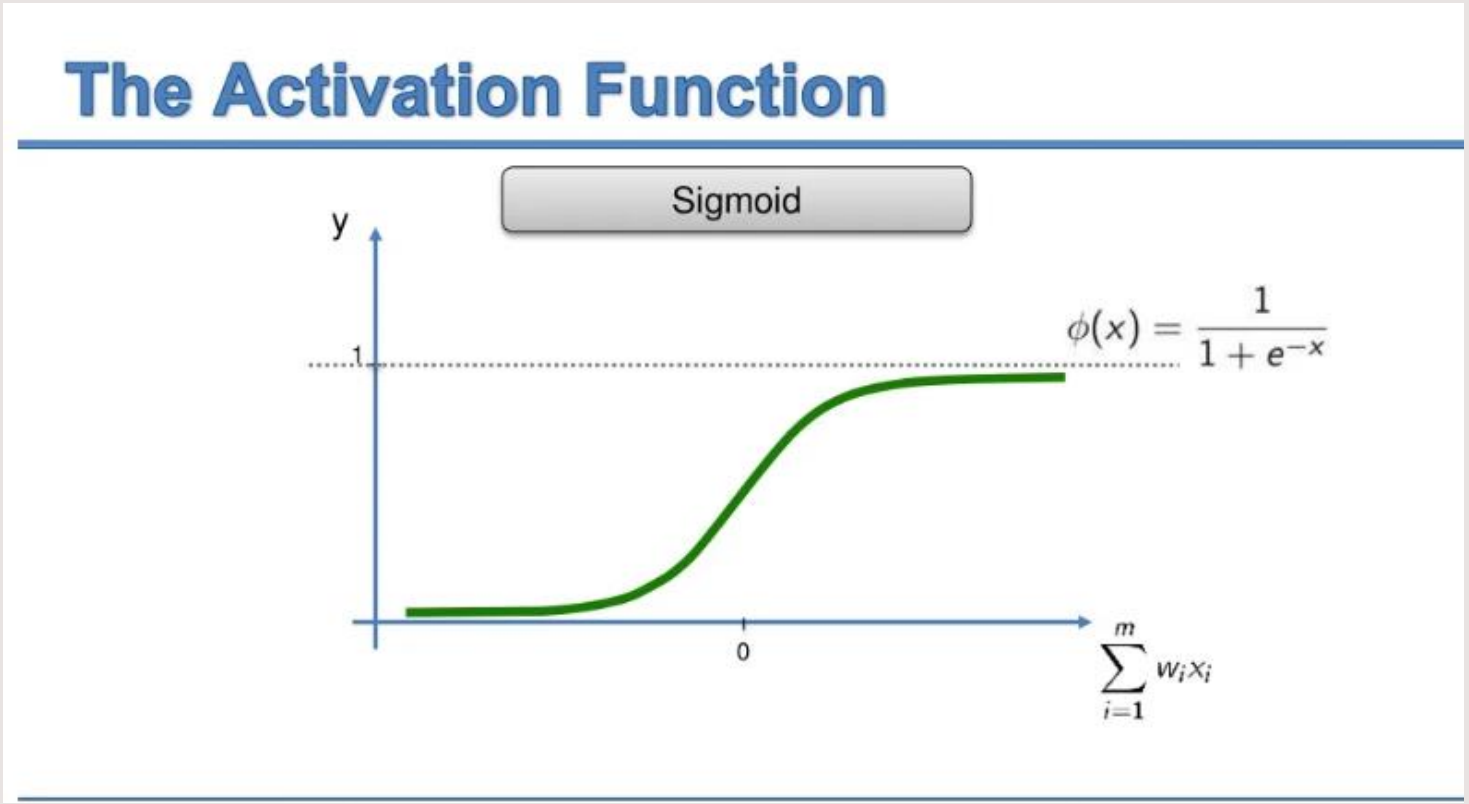


ReLU

$$\max(0, x)$$



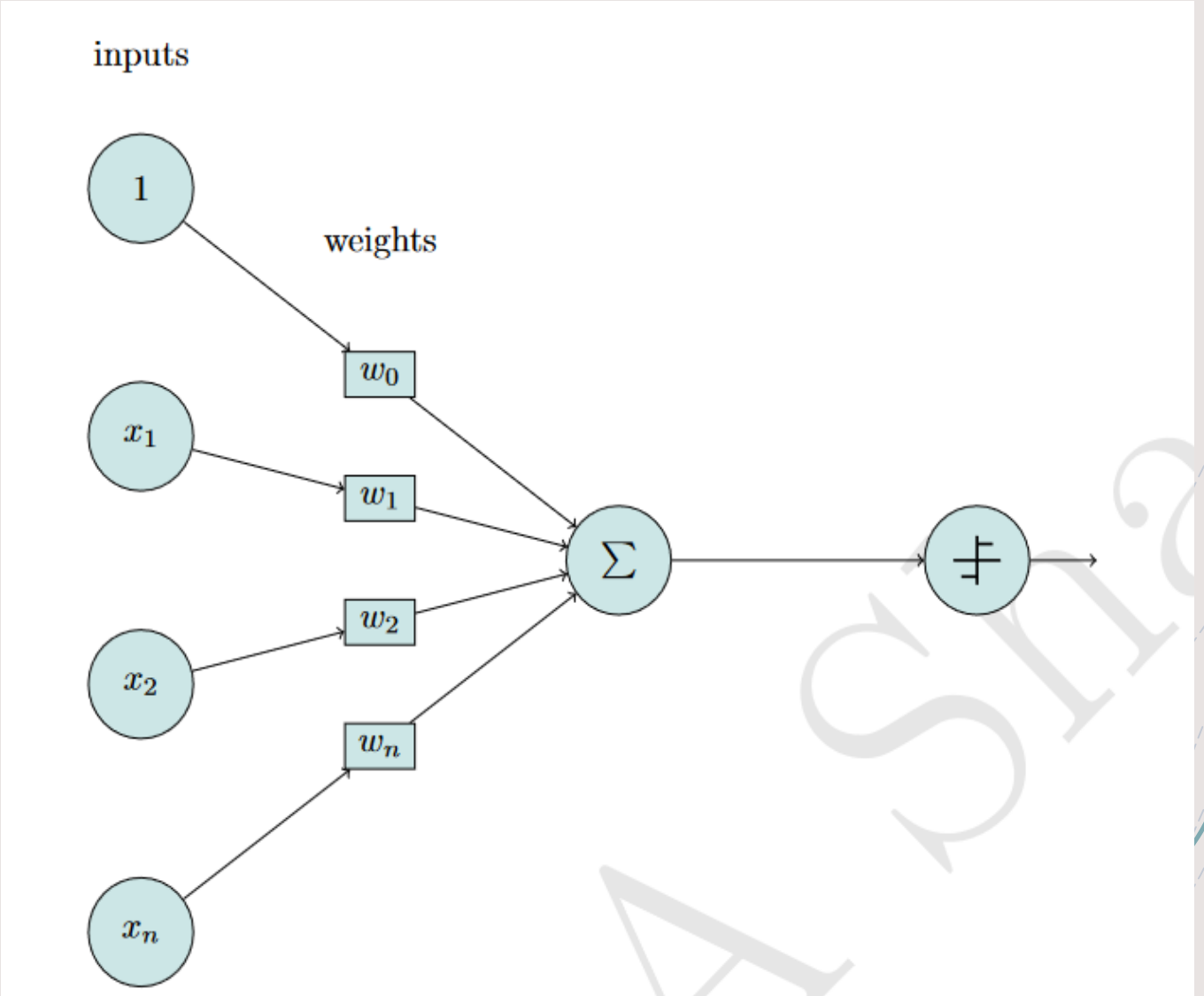
Sigmoid Activation Function



Neuron (Weighted Sum +Activation)

$Z = ActivationFunction(w_1x_1 + w_2x_2 + w_3\ \ x_3 + b)$

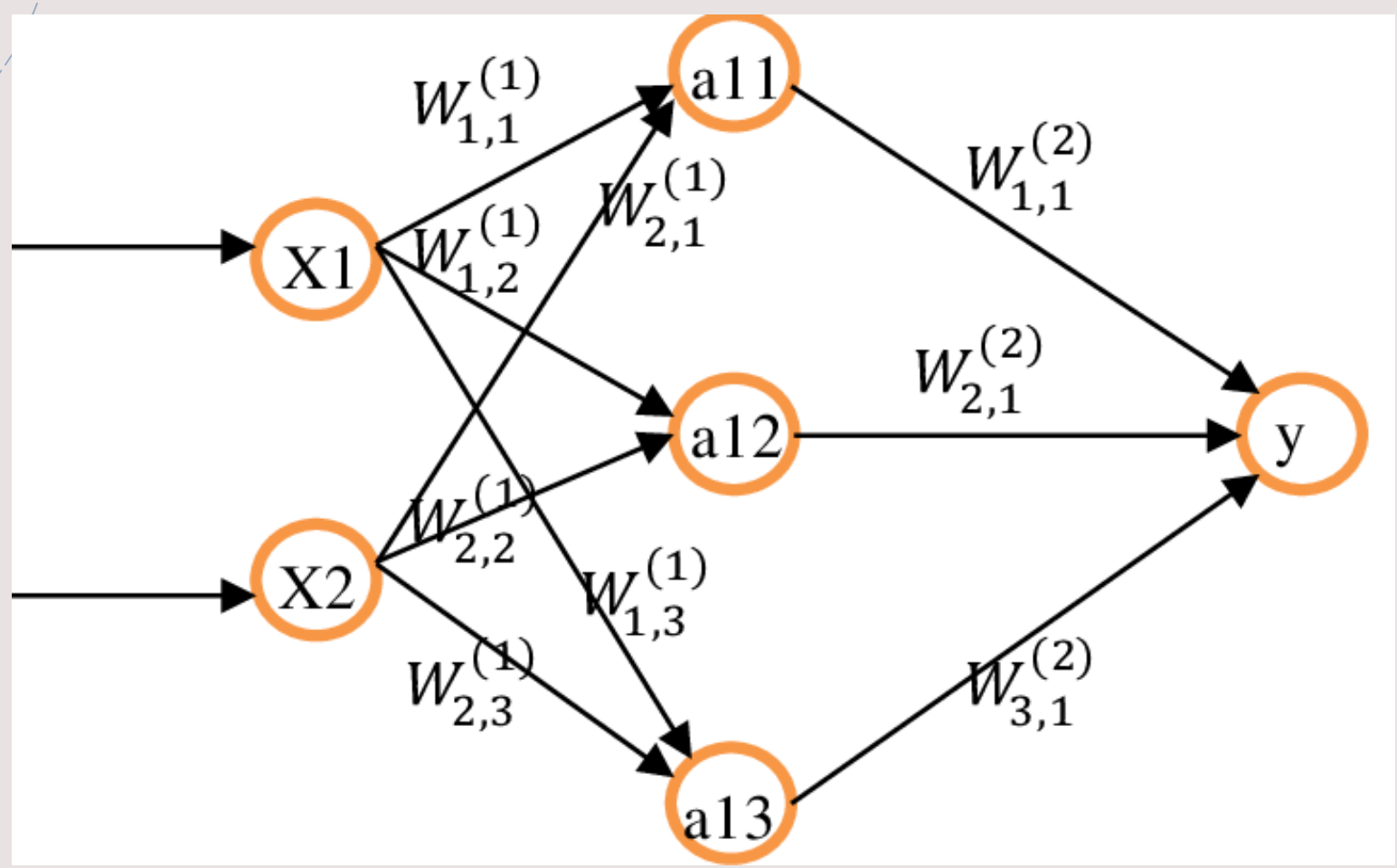
b= $bias = (w_0 1)$



Neuron

1. Takes the input signals
2. Computes the weighted sum of the inputs
3. Applies a step function
4. Checks whether to trigger an output if the threshold exceeds a limit.

Output of DNN



Forward Propagation

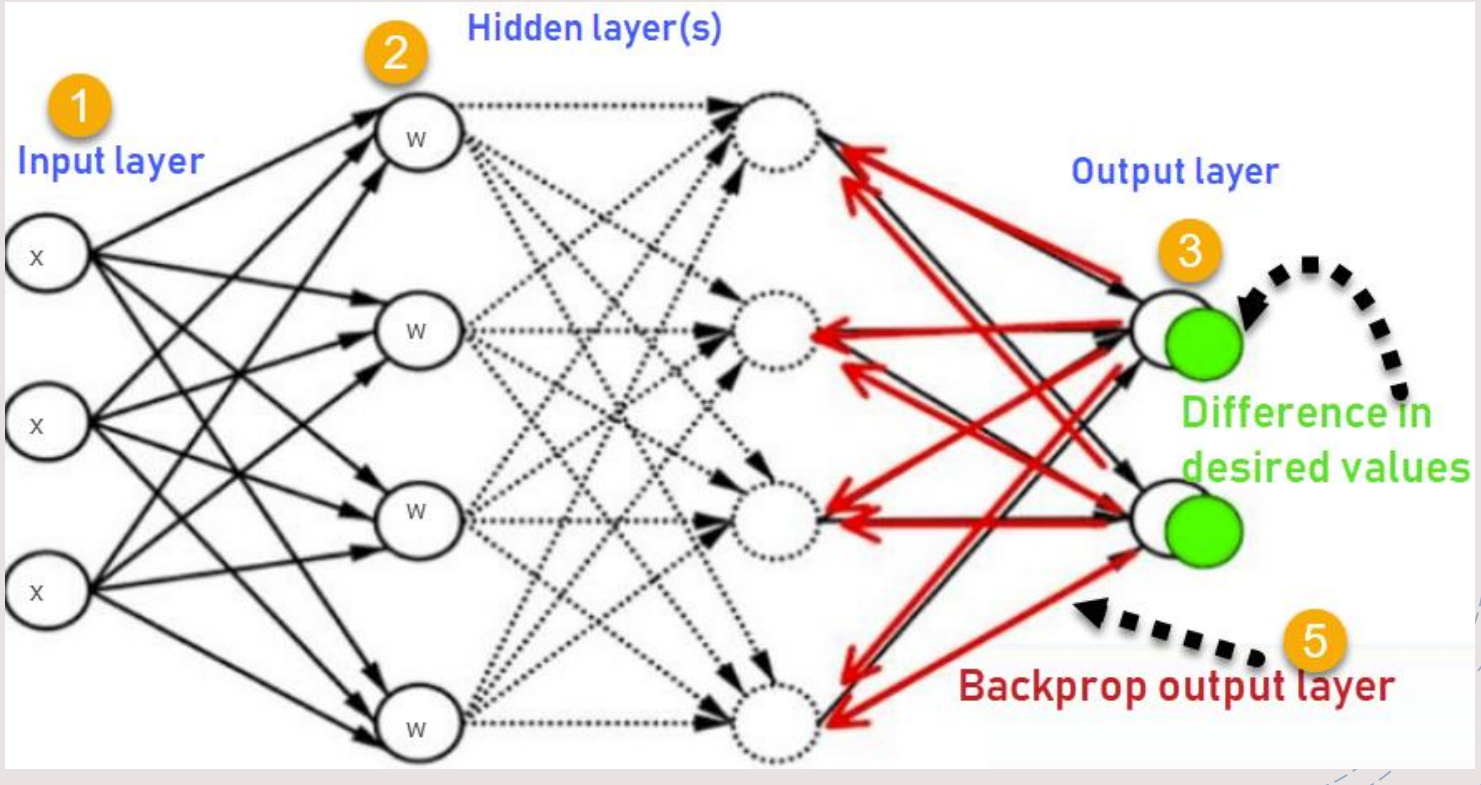
How to get it right

+Prediction is accurate?

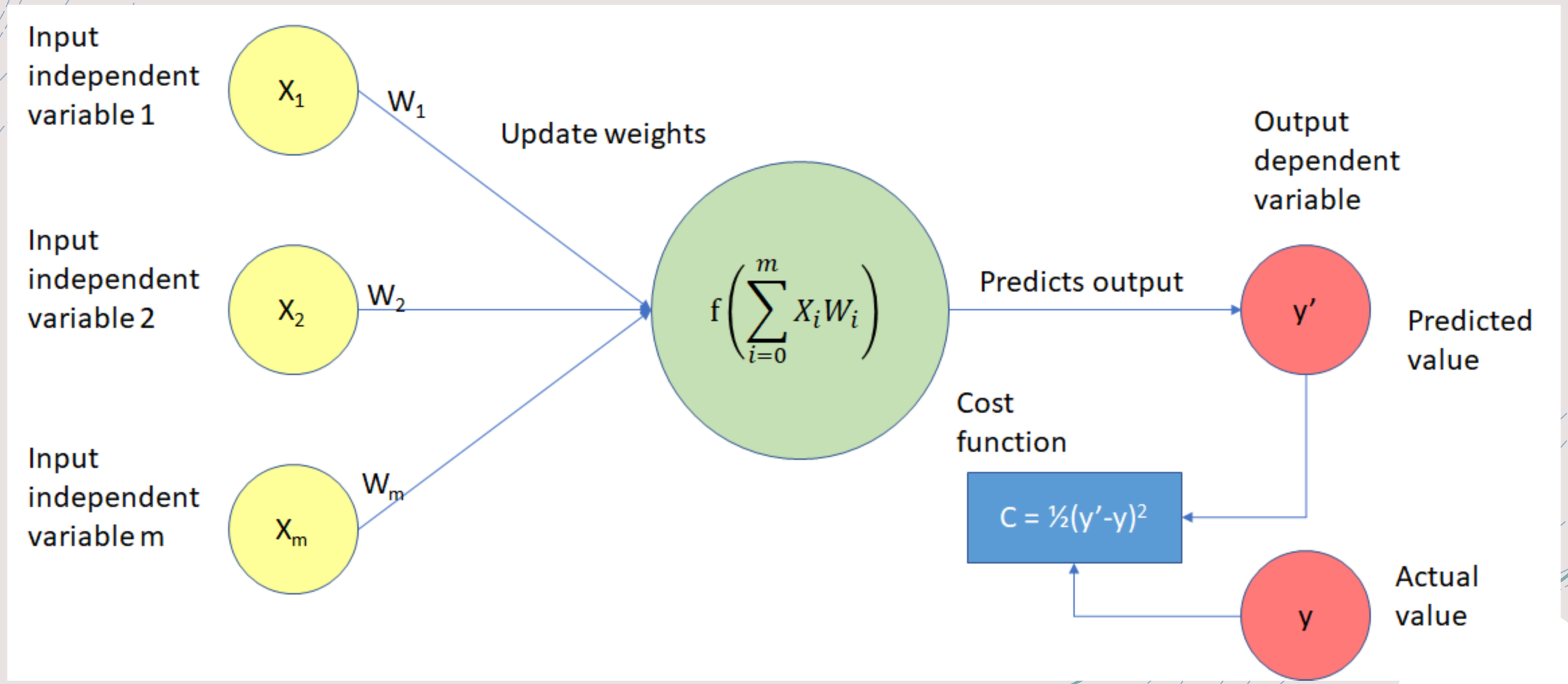
Back Propagation

+Move backwards

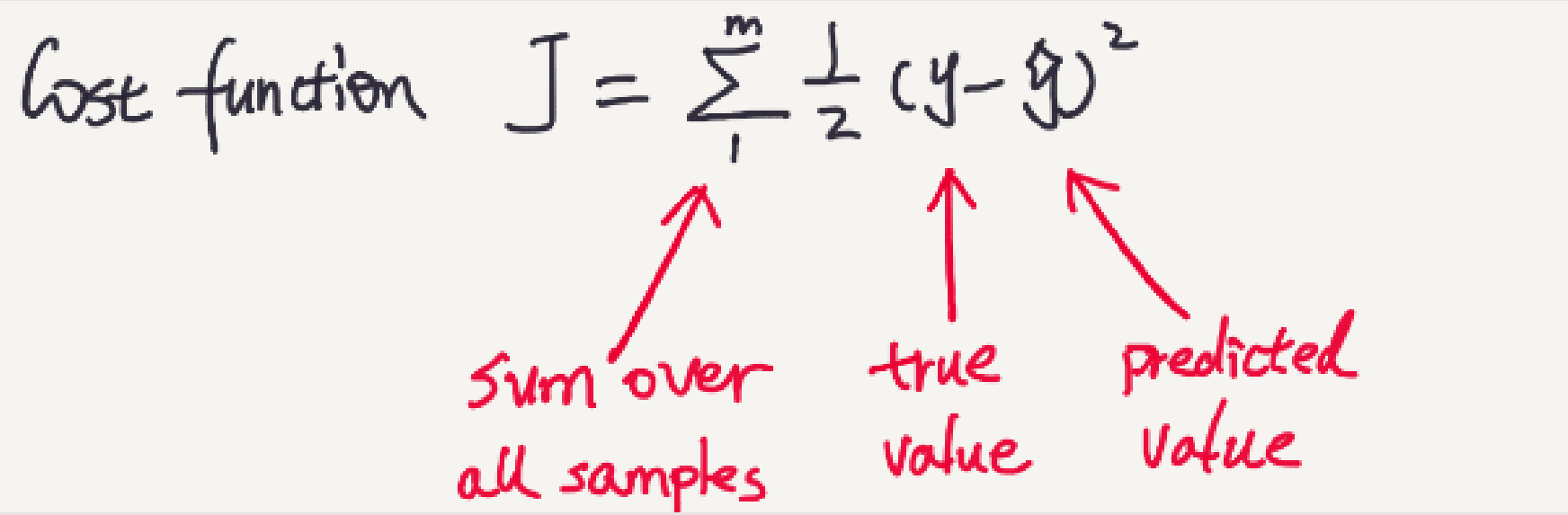
+Adjust weights



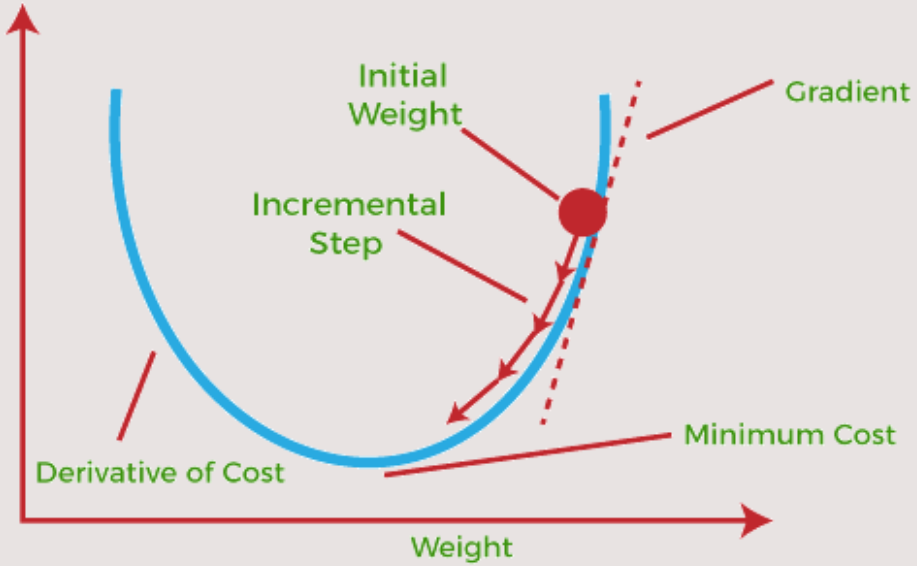
Cost Function



Cost Function



How to minimize the error



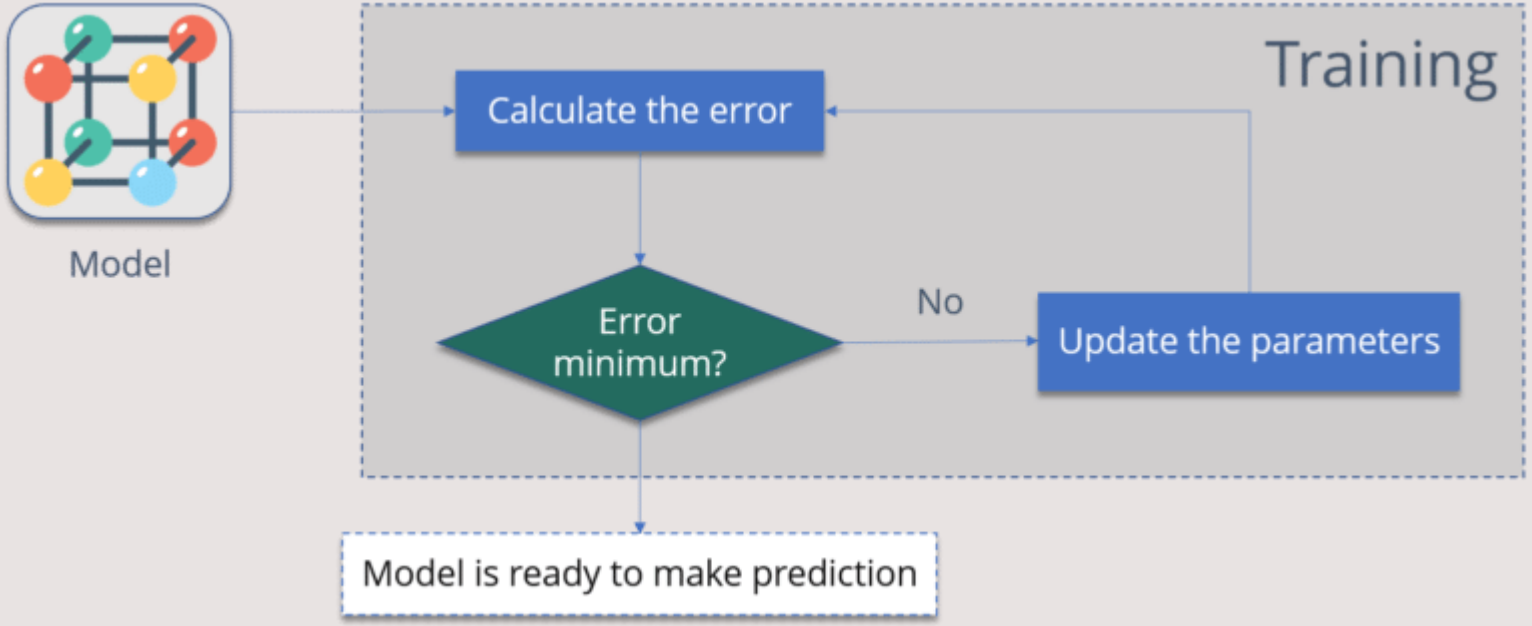
In summary, difference of weights (or weights adjustments) can be mentioned as follows:

$$\Delta W_i = -\alpha \frac{\partial E}{\partial W_i}$$

In the above equation, α gives us the amount of step size; whereas,

$$\frac{\partial E}{\partial W_i}$$

Training


Model
Training
Calculate the error
Error minimum?
No
Update the parameters
Model is ready to make prediction

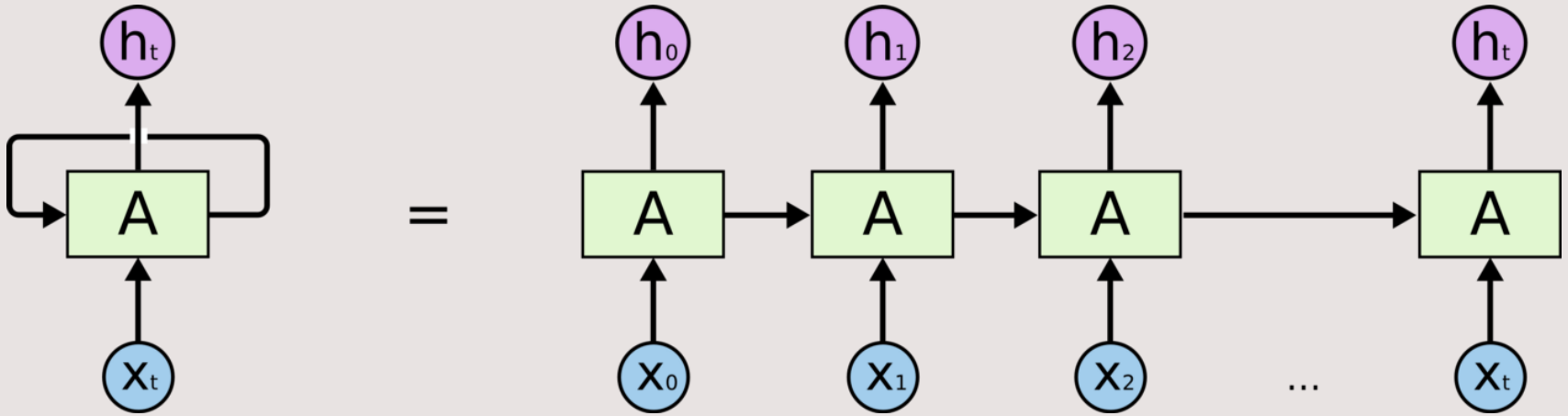
Different Types of Deep Neural Networks and Applications

Application	Examples	Type of DNN
Sequence Learning	Machine Translation Sentiment Analysis Natural Language Processing	RNN LSTM GRU Transformers
Image Processing	Cancer Diagnosis Cat or Dog	Convolutional Neural Network
Object Detection	Search Engines	Modification of CNN (Yolo)
Autonomous Car	Driverless cars Robots	Reinforcement Learning
Deep Fake	Determination of outlier and anomaly. Creation of peer groups of machines and users.	Generative Adversarial Networks
Classification	Malware Other Classification Examples	Multi-Layer Perceptron

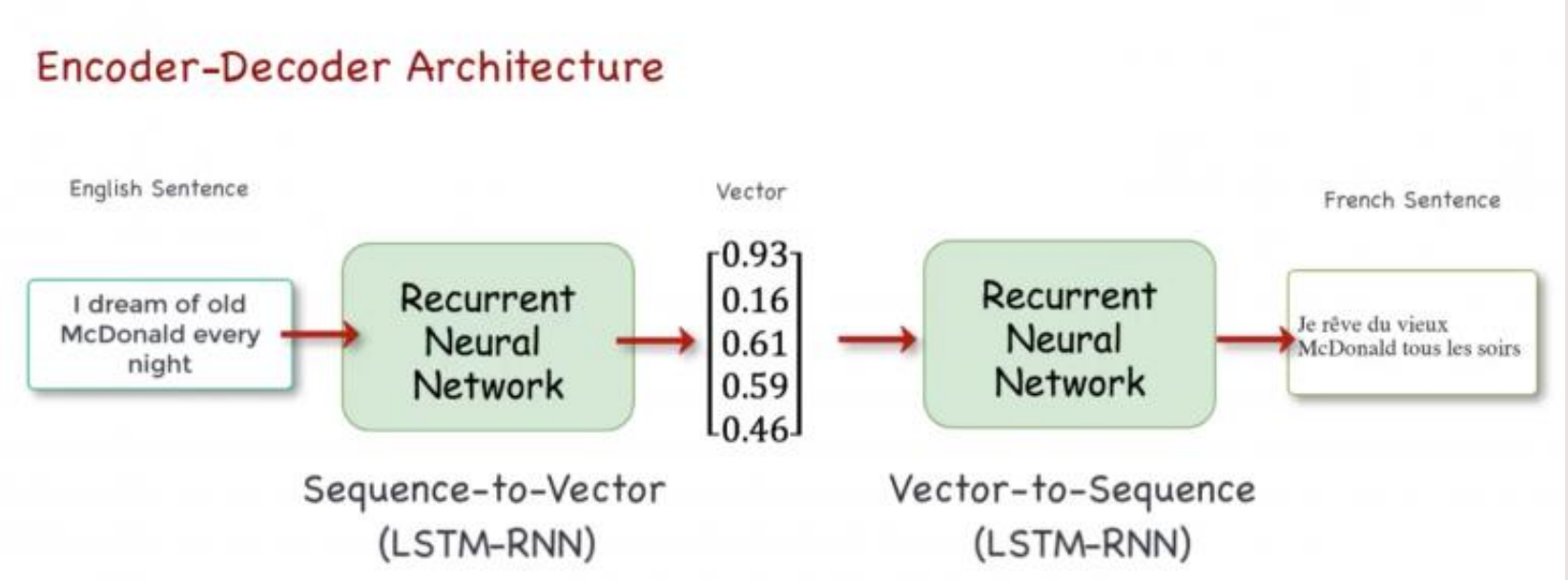
Sentiment Analysis using RNN

+The food was good, not bad at all

+The food was bad, not good at all

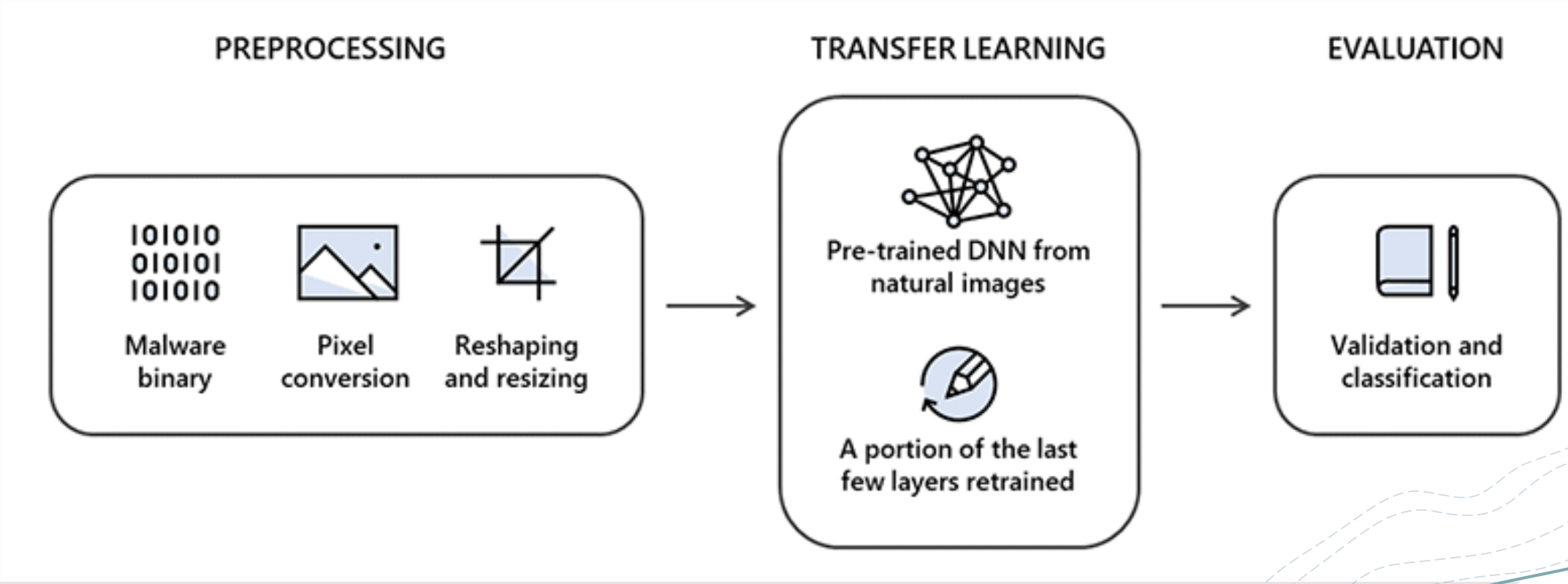


Machine Translation

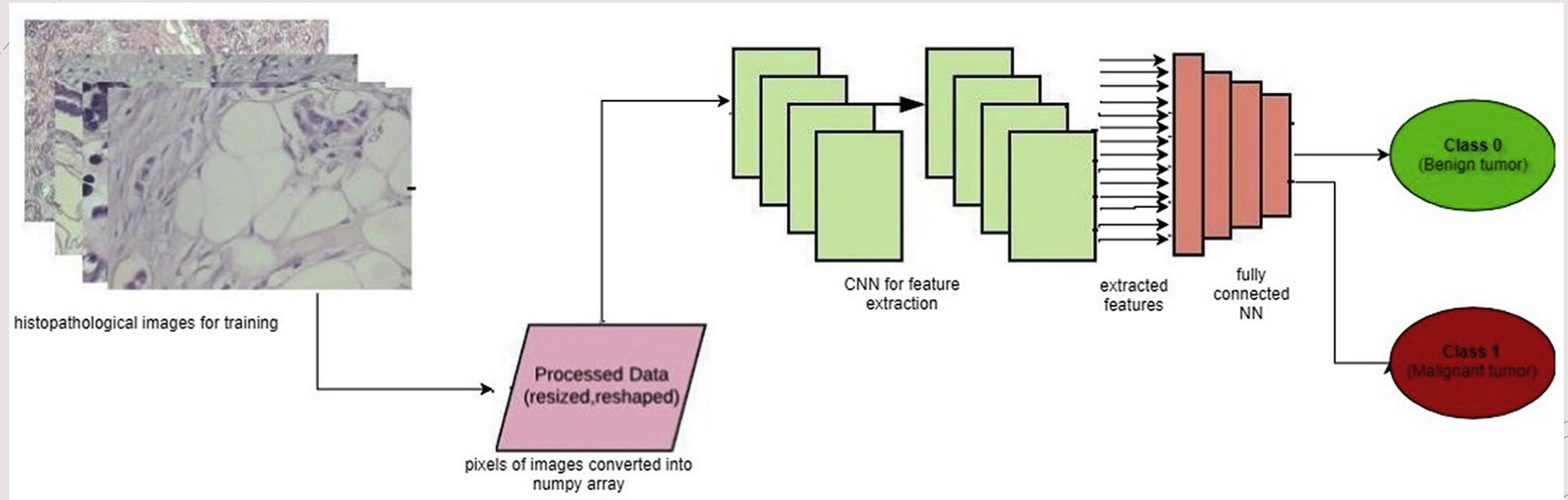


Malware Classification

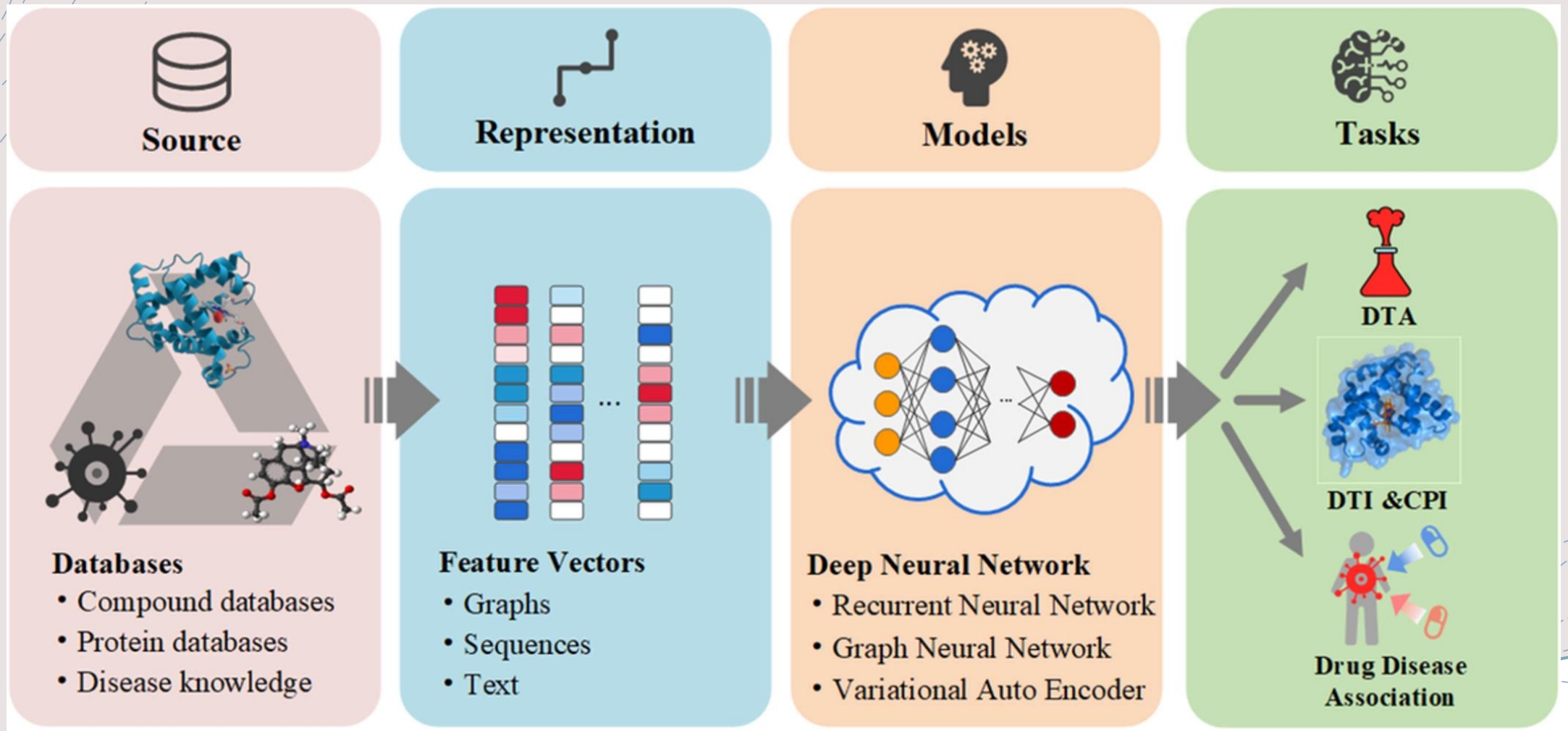
+ Malware files are converted to images
+ Features are extracted using CNN
+ Classify a software as benign or malware



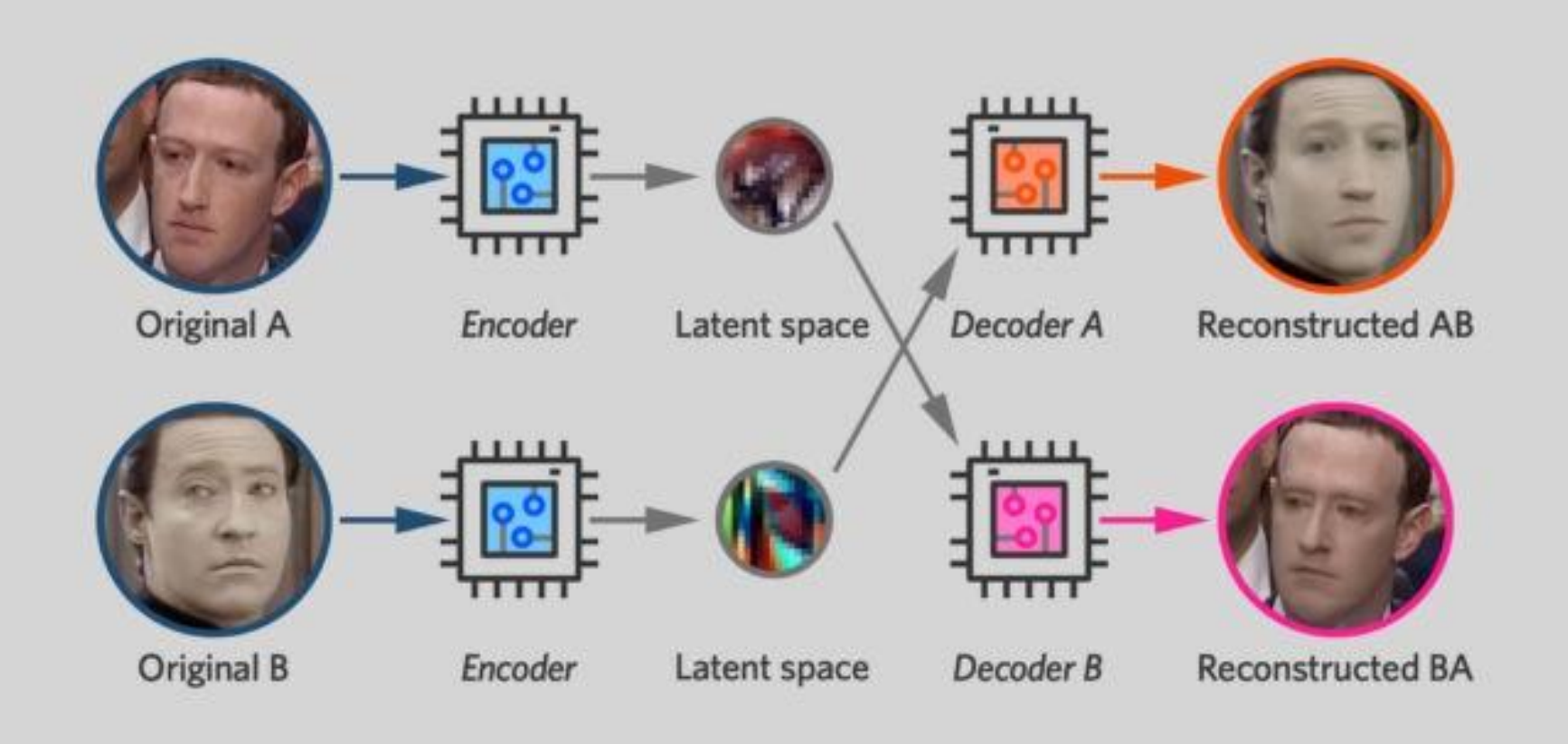
Cancer Diagnosis using CNN



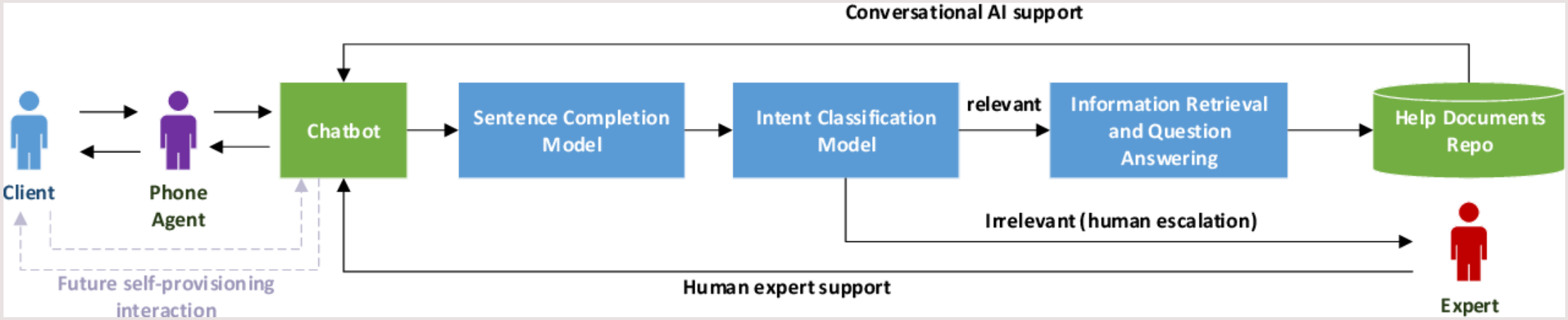
Drug Repurposing



Deepfake


Original A
Encoder
Latent space
Decoder A
Reconstructed AB
Original B
Encoder
Latent space
Decoder B
Reconstructed BA

Chatbot


Conversational AI support
Client
Phone Agent
Chatbot
Sentence Completion Model
Intent Classification Model
relevant
Information Retrieval and Question Answering
Help Documents Repo
Irrelevant (human escalation)
Expert
Human expert support
Future self-provisioning interaction

Conclusion

Hands-on Training

And Gate

https://tinyurl.com/paews

Modify it for OR gate or reduce hidden layers

Questions

Jawwad.shamsi@nu.edu.pk