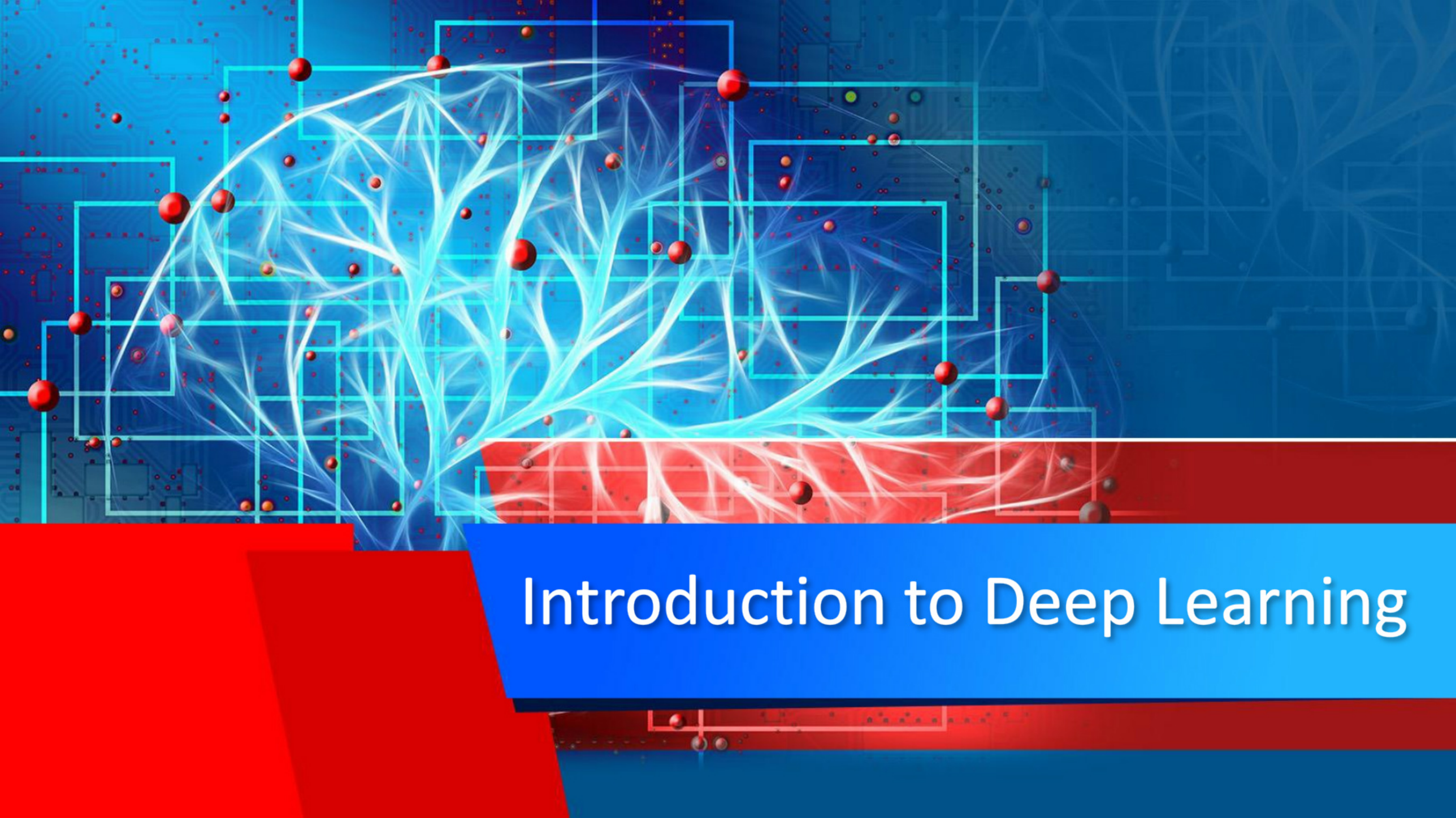


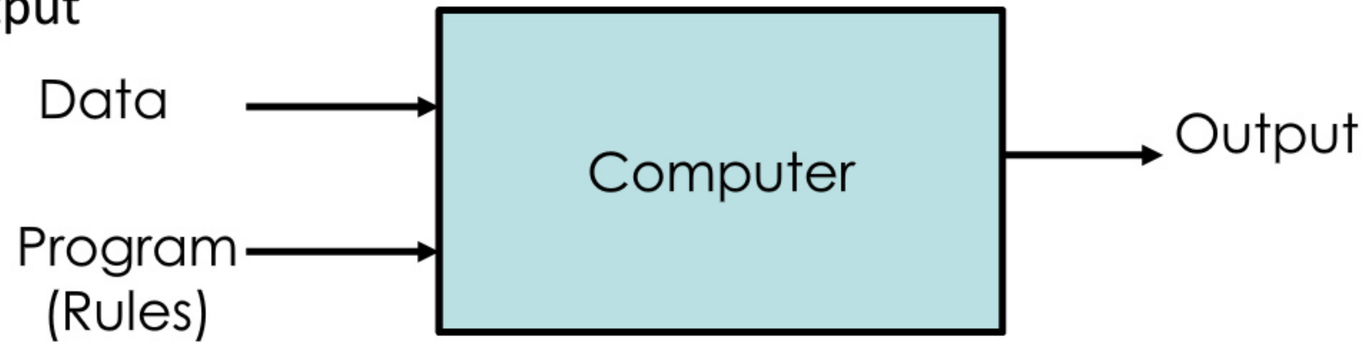
Introduction to Deep Learning

What is Machine Learning?

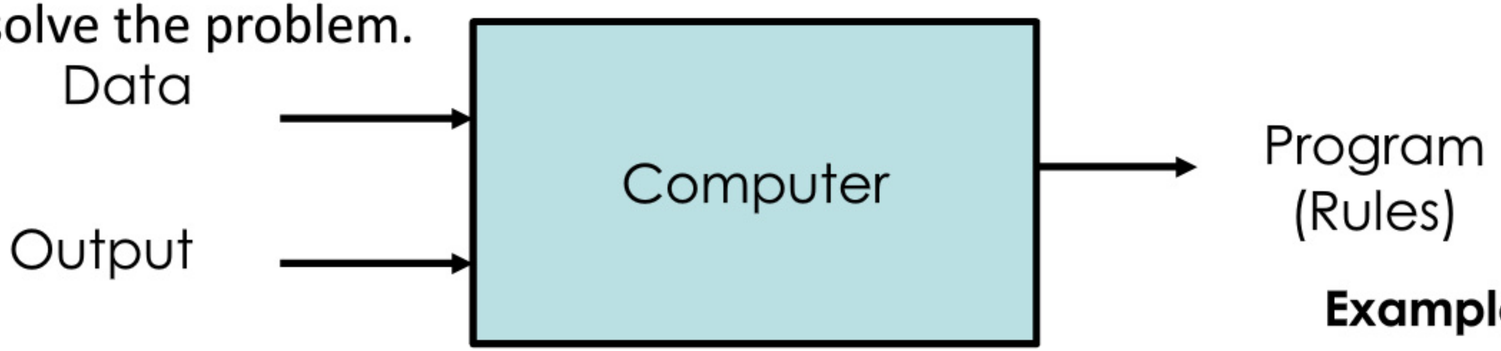
> “Field of study that gives computers the ability to learn without being explicitly programmed.”
>
> — Arthur Samuel

Programming vs ML

- **Traditional Programming**
  - You have the data and you are a programmer writing a program on the computer to produce the output



- **Machine Learning**
  - You have the data and desired output and would like to train the computer to produce the program that solve the problem.



Example: Sorting numbers

Slide Credit: Pedro Domingos, Tom Mitchel, Tom Dietterich

Activity Recognition

```
if(speed<4){
    status=WALKING;
}
```

```
if(speed<4){
    status=WALKING;
} else {
    status=RUNNING;
}
```

```
if(speed<4){
    status=WALKING;
} else if(speed<12){
    status=RUNNING;
} else {
    status=BIKING;
}
```

```
// Uh oh
```

Activity Recognition using ML

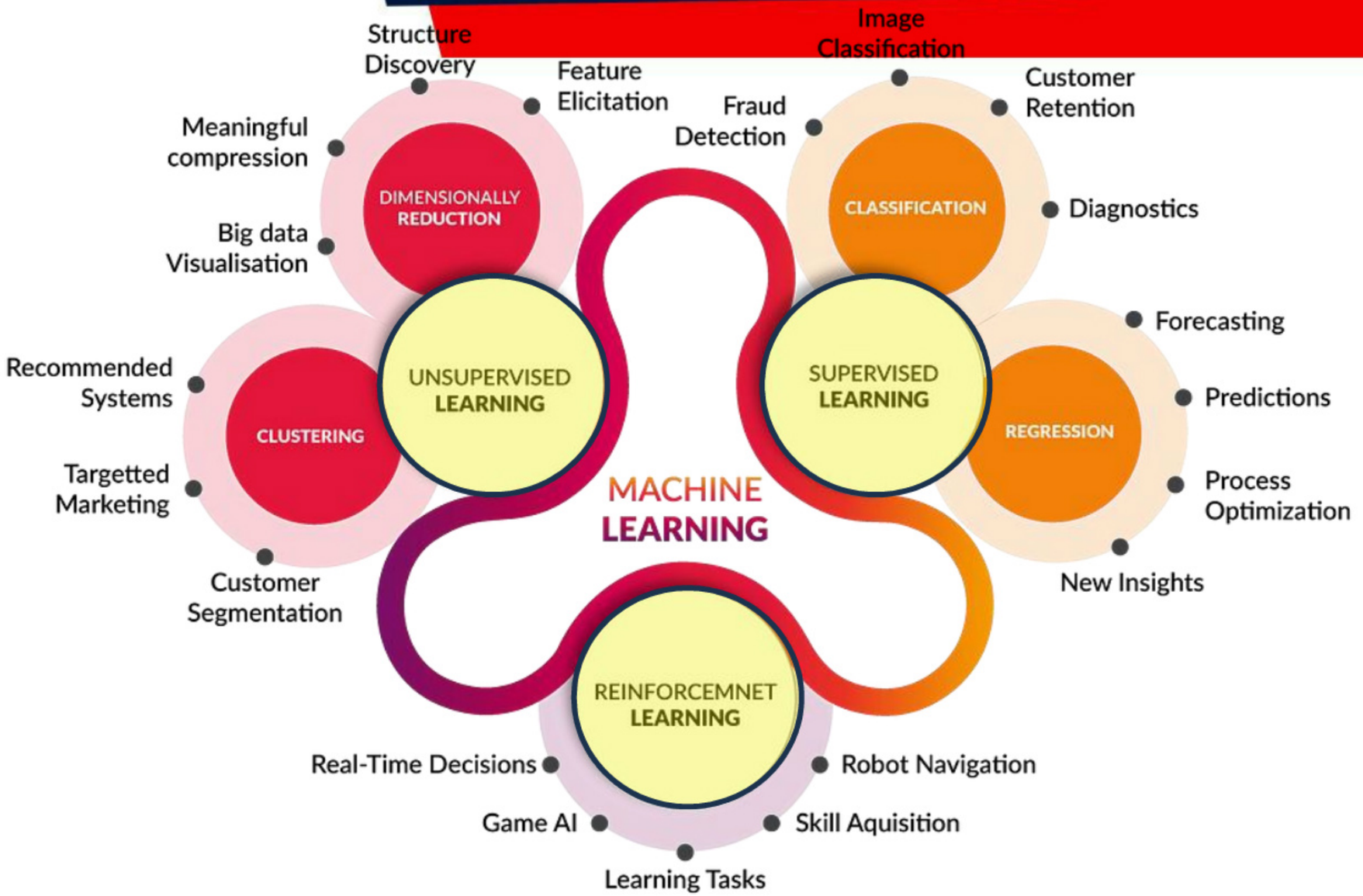
0101001010100101
0101001010101001
0111010100101010
0101010010101001
010100101010
Label =
WALKING

1010100101001010
1010101010010010
0100010010011111
0101011111010100
100111101011
Label =
RUNNING

1001010011111010
1011101010111010
1011101010101111
0101010111111110
001111010101
Label =
BIKING

1111111111010011
1010011111010111
1101010101110101
0101011101010101
010100111110
Label = GOLFING
(Sort of)

Types of Machine learning



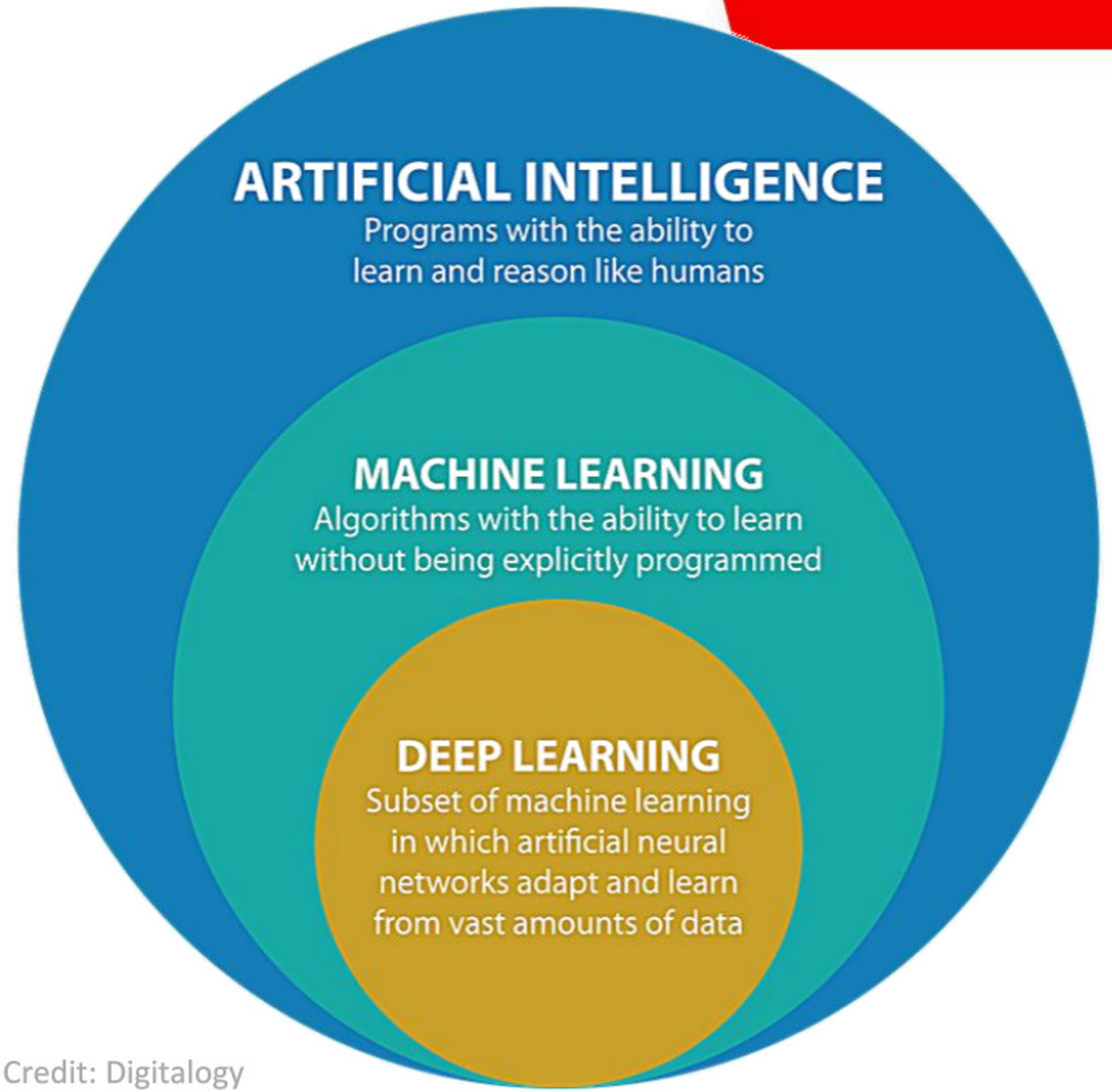
Types of Machine learning

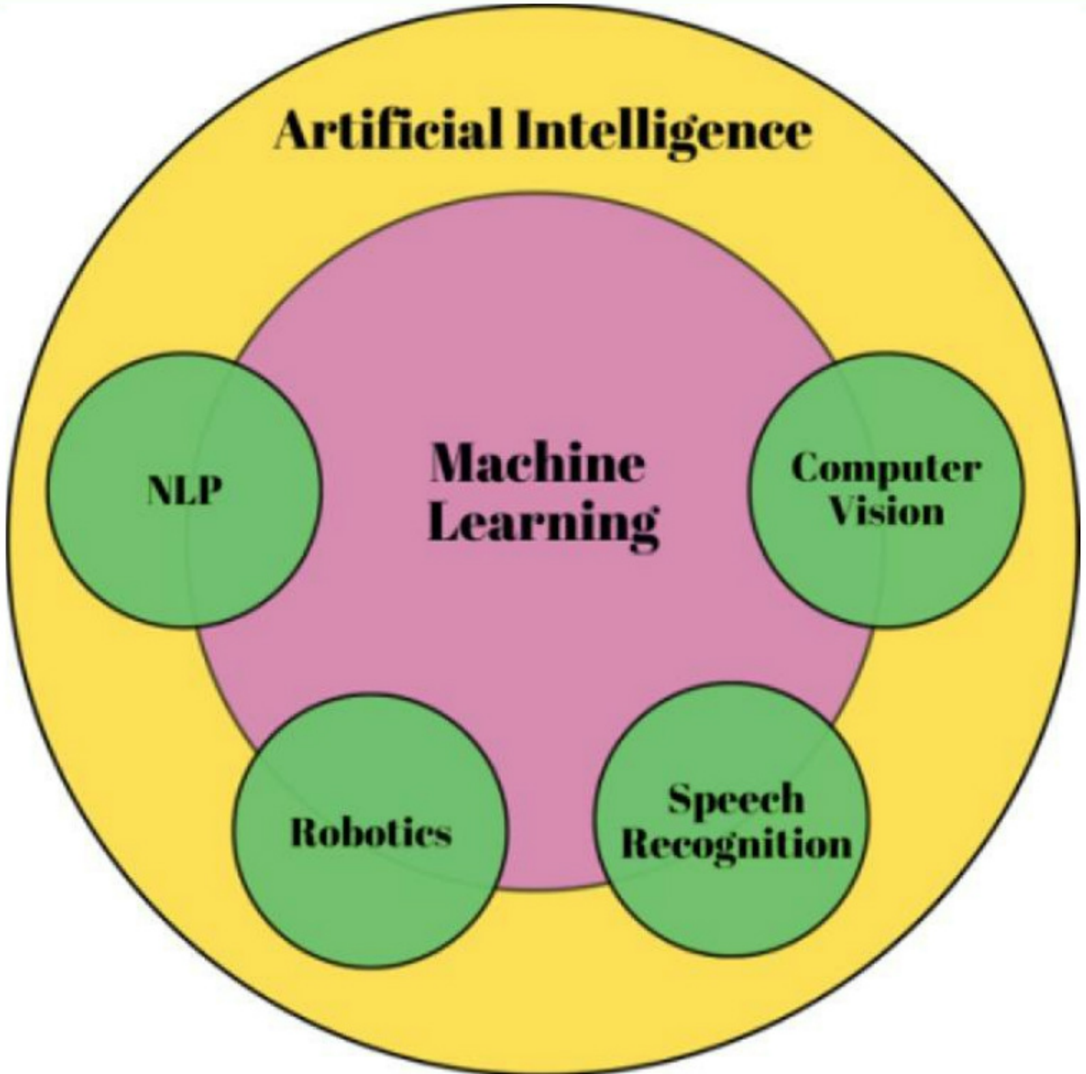
- **Supervised learning**: Training data includes desired outputs
- The goal is to predict an “output” y from an “input” x
- Classification and Regression

- **Unsupervised learning**: Training data does not include desired outputs. instead, looking for interesting patterns or structure in the data

- **Reinforcement learning**: Rewards from sequence of actions, tries to learn to maximize the reward signal

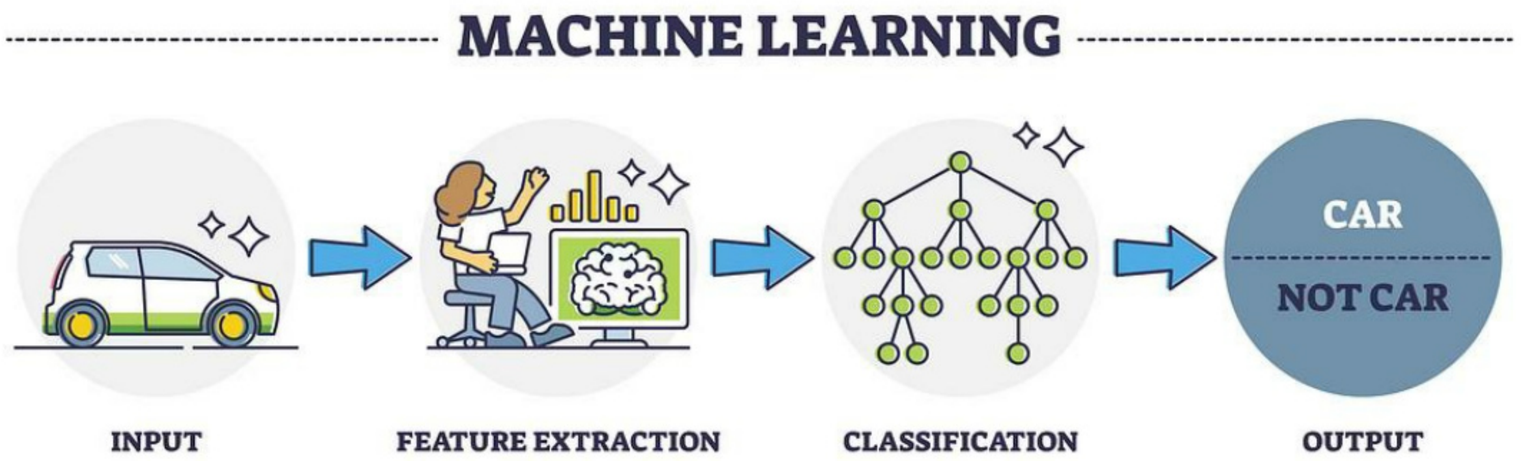
AI, ML, and DL

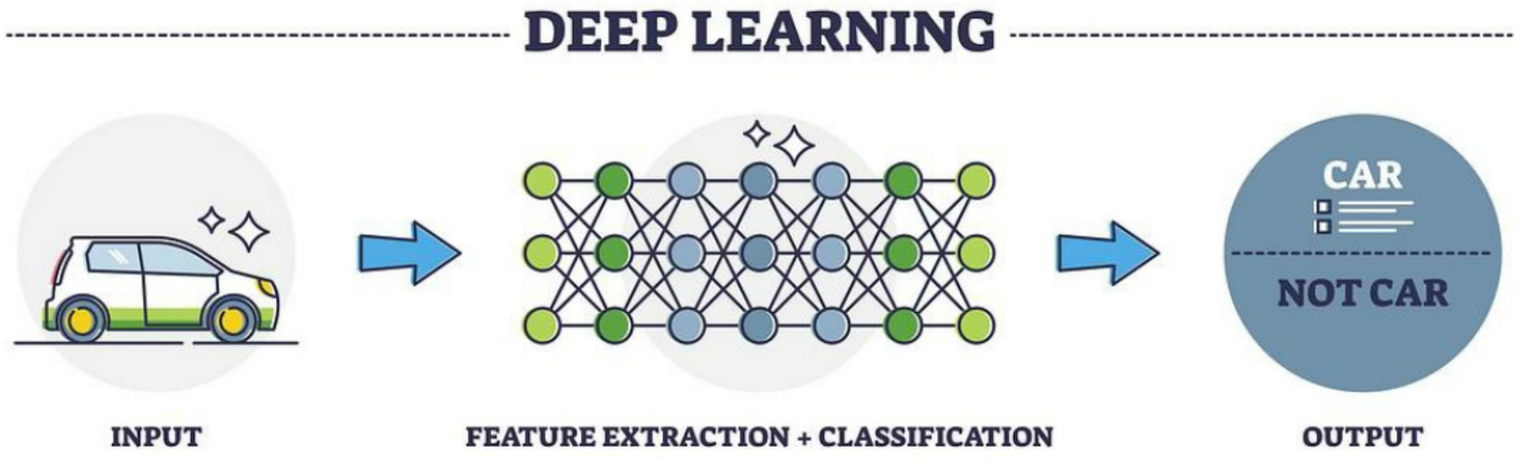




Slide Credit: Digitalogy

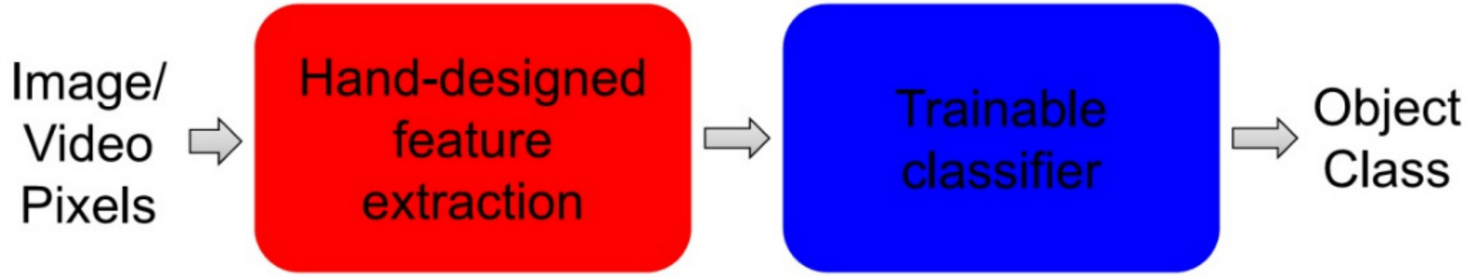
Machine learning vs Deep learning



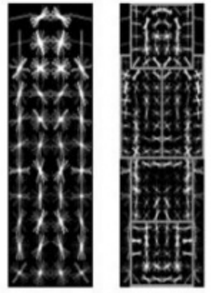


Shallow learning

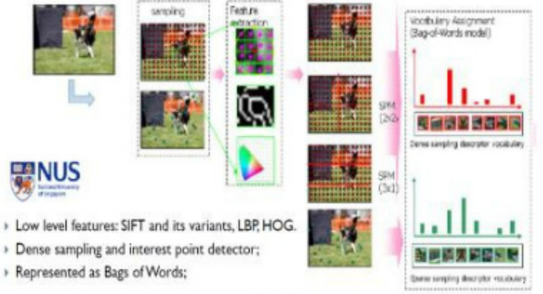
"Shallow" architecture



- Feature extraction is a manual process that requires domain knowledge of the data that we are learning from.
- Trainable classifier is often generic (e.g. SVM)
- Features are key to recent progress in recognition.
- Multitude of hand-designed features currently in use
  - • SIFT, HOG, ……



Felzenszwalb, Girshick, McAllester and Ramanan, PAMI 2007

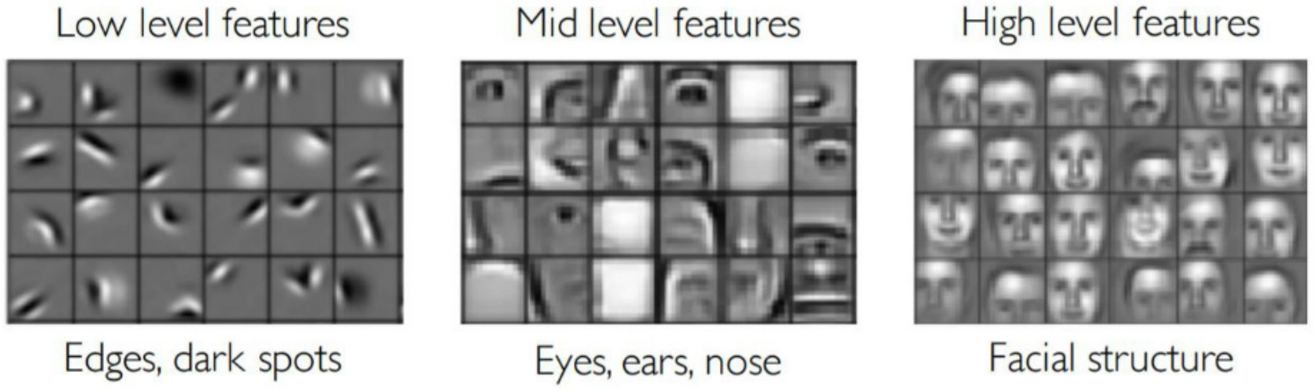


Yan & Huang
(Winner of PASCAL 2010 classification competition)

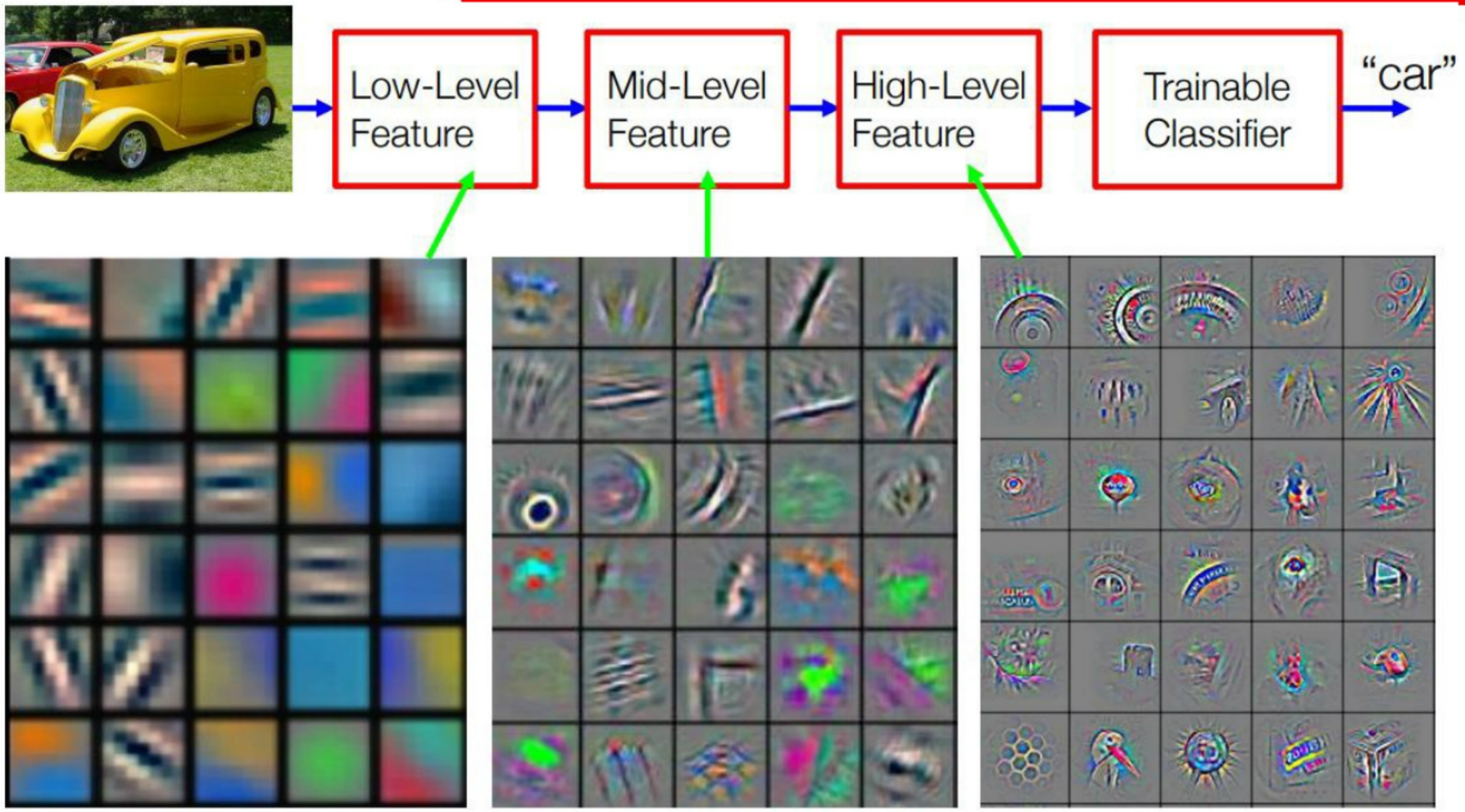
Deep learning

Deep learning: "Deep" architecture

- Deep learning methods have in common:
  - refers to a set of algorithms or models that are composed of multiple processing layers
  - Learn underlying features directly from data.
  - Have multiple layers that extract features from the output of previous layer
  - Learn feature hierarchies representations in each layer.

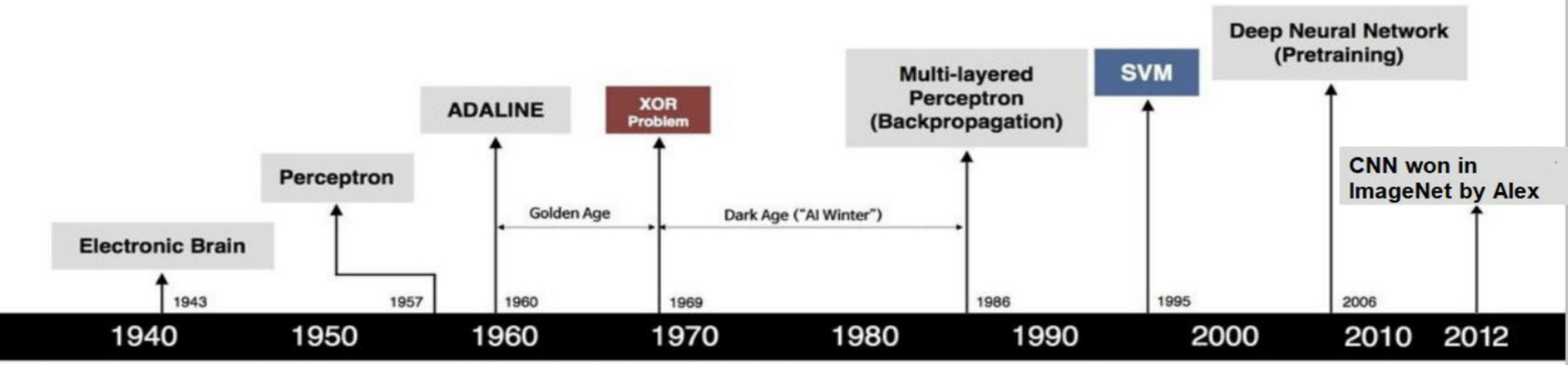


Deep Learning = Hierarchical Compositionality



M.D. Zeiler and R. Fergus, "Visualizing and Understanding Convolutional Networks", In ECCV 2014

Brief history of Deep Learning



S. McCulloch - W. Pitts

F. Rosenblatt

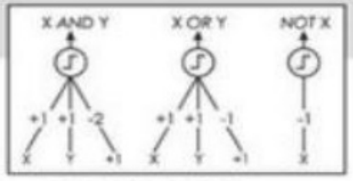
B. Widrow - M. Hoff

M. Minsky - S. Papert

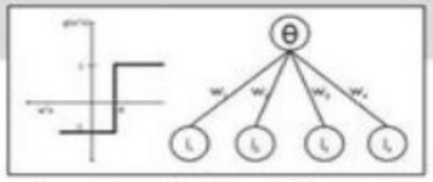
D. Rumelhart - G. Hinton - R. Wiliams

V. Vapnik - C. Cortes

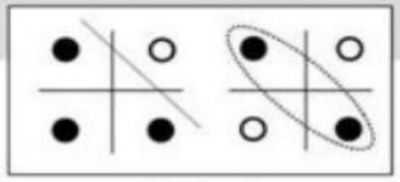
G. Hinton - S. Ruslan



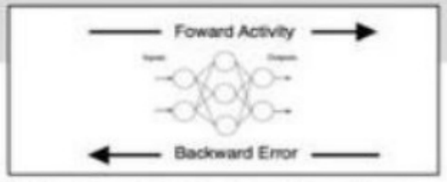
- Adjustable Weights
- Weights are not Learned



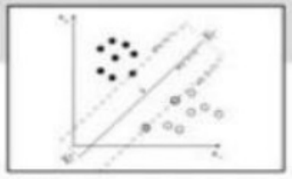
- Learnable Weights and Threshold



- XOR Problem



- Solution to nonlinearly separable problems
- Big computation, local optima and overfitting



- Limitations of learning prior knowledge
- Kernel function: Human Intervention

- Hierarchical feature Learning

Brief history of Deep Learning

- **1940s: The Birth of Artificial Neurons**
  - **1943: Warren McCulloch** and **Walter Pitts** proposed the first mathematical model of an artificial neuron. Their work laid the foundation for neural networks by mimicking the way biological neurons process information.
- **1950s: The Perceptron**
  - **1957: Frank Rosenblatt** developed the **Perceptron**, the first neural network model capable of learning from data. It was a significant breakthrough but limited in its ability to solve complex problems.
- **1960s: Early Development and Challenges**
  - **1960: Bernard Widrow** and **Marcian Hoff** introduced the **ADALINE** (Adaptive Linear Neuron) and **MADALINE** models, which were among the first to be used in practical applications like adaptive filters.
  - **1969: Marvin Minsky** and **Seymour Papert** published *"Perceptrons,"* which highlighted the limitations of single-layer neural networks in solving the XOR problem, leading to a decline in neural network research—a period often referred to as the "AI Winter."
- **1980s: The Backpropagation Revolution**
  - **1986: David Rumelhart**, **Geoffrey Hinton**, and **Ronald J. Williams** rediscovered and popularized the **backpropagation** algorithm, which allowed multi-layered networks (often called deep networks) to be trained effectively. This rekindled interest in neural networks.
- **1990s: SVMs and Renewed Interest**
  - **1995: Vladimir Vapnik** and **Corinna Cortes** introduced **Support Vector Machines (SVMs)**, which became popular due to their effectiveness in classification tasks. Neural networks faced competition from SVMs and other machine learning methods during this time.
- **2000s: The Deep Learning Era Begins**
  - **2006: Geoffrey Hinton** and his colleagues introduced **pretraining** techniques using deep belief networks, which made training deep neural networks more feasible and effective. This marked the beginning of the deep learning era.

Brief history of Deep Learning

AlexNet

Rise of Convolutional Networks (CNNs)

Generative Adversarial Networks (GANs)

ResNet

AlphaGo

Transformer

BERT

GPT-2, GPT-3
The rise of Large Language Models

ViT

ChatGPT

GPT-4

2012 2013-2014 2015 2016 2017 2018 2019-2020 2022 2023

Ian Goodfellow

Vaswani, Niki Parmar

OpenAI

ChatGPT VS Bard

Brief history of Deep Learning

- **2010s: Breakthroughs in Deep Learning**
  - **2012: AlexNet**, developed by **Alex Krizhevsky**, **Ilya Sutskever**, and **Geoffrey Hinton**, won the ImageNet competition by a large margin, showcasing the power of Convolutional Neural Networks (CNNs) and sparking widespread interest in deep learning.
- 2014: Generative Adversarial Networks (GANs), introduced by Ian Goodfellow, revolutionized generative modeling, enabling the creation of realistic images, videos, and more.
- 2015: Residual Networks (ResNet) by He et al. introduced the concept of skip connections, enabling the training of ultra-deep networks (over 100 layers) without the vanishing gradient problem.
- 2016: AlphaGo: Google DeepMind's AlphaGo defeated the world champion Go player, marking a milestone in deep reinforcement learning. AlphaGo combined deep neural networks with reinforcement learning
- 2017: **Transformers**, introduced by Vaswani et al. Google Brain, revolutionized natural language processing (NLP), leading to models like BERT and GPT that have set new benchmarks in various NLP tasks. Unlike RNNs, Transformers use self-attention mechanisms to process input sequences in parallel, enabling faster and more effective training.
- 2018 **BERT** (Bidirectional Encoder Representations from Transformers), introduced by Google, transformed NLP tasks by applying pre-training followed by fine-tuning on specific tasks. Use Transfer learning.
- **2020: The Rise of Large-Scale Models**
  - OpenAI introduced GPT-2, a large-scale language model trained to generate coherent and fluent text, pushing the boundaries of natural language understanding and generation. This sparked widespread interest in large-scale language models, showcasing the power of unsupervised learning on massive datasets.
  - OpenAI's GPT-3 (175 billion parameters) became the largest and most powerful language model at the time, demonstrating unprecedented capabilities in text generation, translation, summarization, and even code writing.
  - GPT-3.5 was introduced in late 2022 as an upgraded version of GPT-3, with enhanced conversational capabilities. This version powered the initial ChatGPT product,
  - ChatGPT is a fine-tuned version of GPT-3, specifically optimized for conversation.
  - Training: ChatGPT was fine-tuned using Reinforcement Learning from Human Feedback (RLHF). Human AI trainers provided conversations and rated responses, which allowed the model to learn not just from static data but from interactions and preferences.
  - Multimodal capabilities: GPT-4 introduced multimodal abilities. March 14, 2023. This means it can accept both text and images as inputs, making it capable of tasks such as image captioning, understanding complex visual scenes, and integrating text with images for more contextual understanding.

Why Now?

- Neural Networks date back decades, so why the dominance?

1. Big Data

- Larger Datasets
- Easier Collection & Storage

IMAGENET

2. Hardware

- Graphics Processing Units (GPUs)
- Massively Parallelizable

3. Software

- Improved Techniques
- New Models
- Toolboxes

TensorFlow

ImageNet Challenge

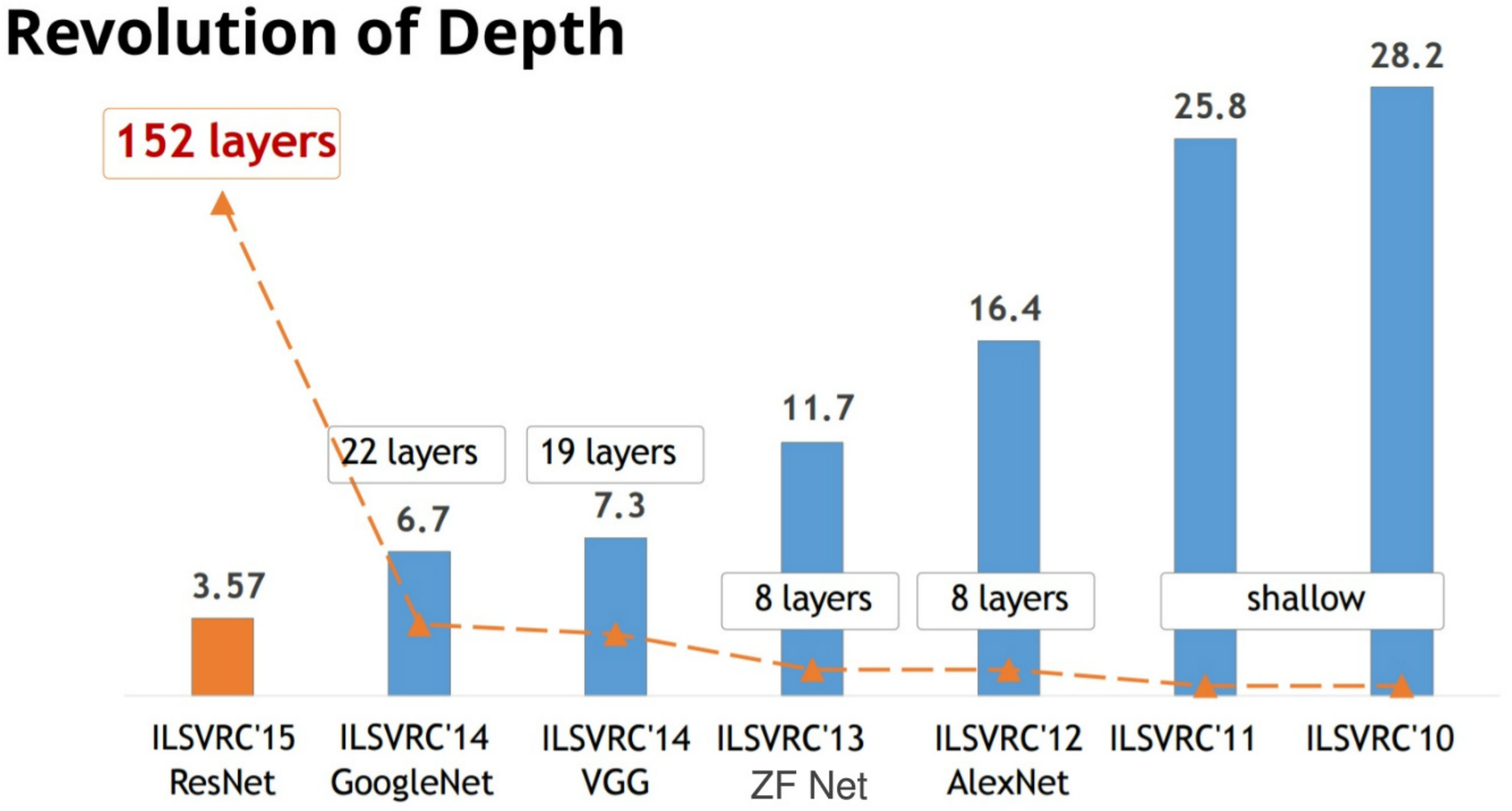
IMAGENET

[Deng et al. CVPR 2009]

ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

- ~14 million labeled images, 20k classes
- Images gathered from Internet
- Human labels via Amazon Turk
- Challenge: 1.2 million training in 1000 classes
- Classification task: produce a list of object categories present in image. 1000 categories.

ImageNet Challenge & Deep learning



AlphaGo seals 4-1 victory over Go grandmaster Lee Sedol

DeepMind's artificial intelligence astonishes fans to defeat human opponent and offers evidence computer software has mastered a major challenge

The world's top Go player, Lee Sedol, lost the final game of the Google DeepMind challenge match. Photograph: Yonhap/Reuters

Google DeepMind's AlphaGo program triumphed in its final game against South Korean Go grandmaster Lee Sedol to win the series 4-1, providing further evidence of the landmark achievement for an artificial intelligence program.