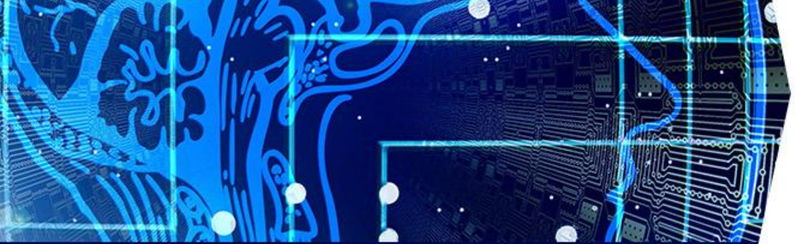
Introduction to AI and Machine Learning

School of Engineering

Nanyang Polytechnic

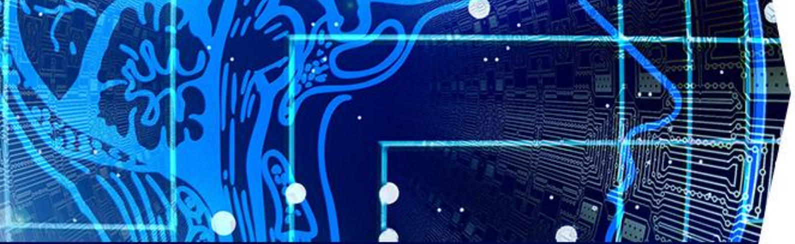


Topics

1. Overview of AI, ML, and DL
2. Machine Learning Types and Techniques
3. Machine Learning Modeling Process

Activities:

- Creating an Object Recognition Model
- Linear Regression using Excel
- Business Problem and Formulation in ML



1. Overview of AI and Machine Learning

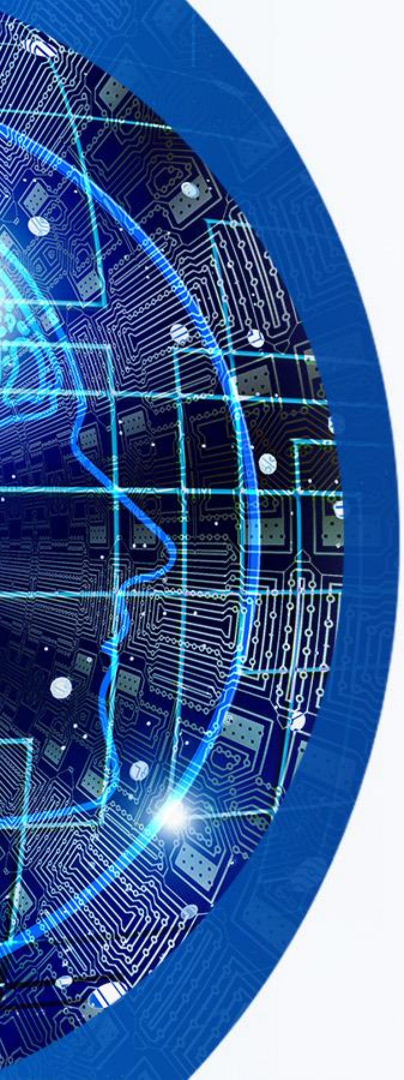
1.1 What is Artificial Intelligence (AI)

1.2 What is Machine Learning (ML)

1.3 What is Deep Learning (DL)

Activity: Creating an object recognition model

1.1 What is AI

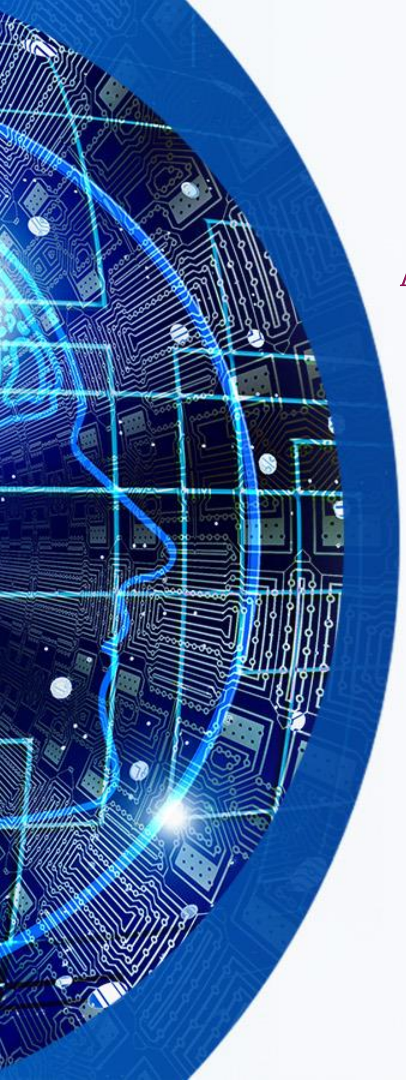


What is AI?

Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems.

Artificial intelligence is a wide-ranging branch of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence.

Artificial intelligence is a field, which combines computer science and robust datasets, to enable problem-solving.



Definition of AI

Some definitions of artificial intelligence, organized into four categories from the book

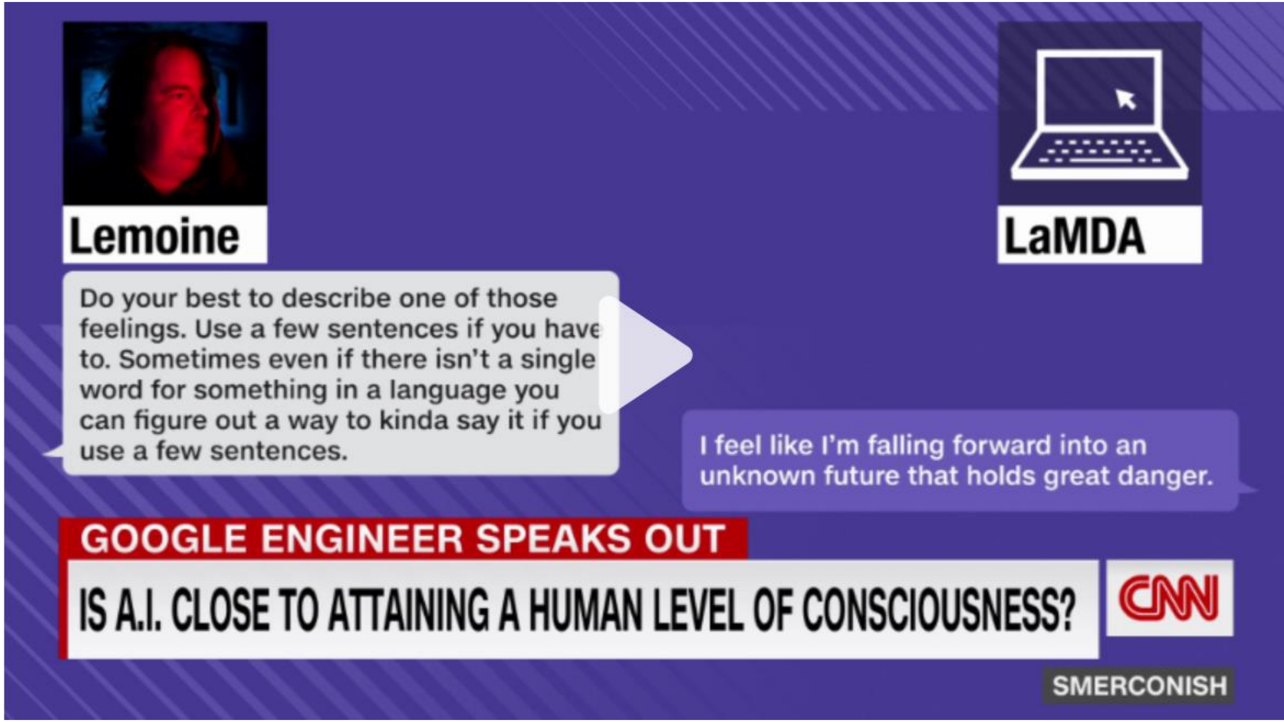
Artificial Intelligence: A Modern Approach, 4th US ed. by [Stuart Russell] and [Peter Norvig]

	Human Performance	Rationality
Thought process, reasoning	**Thinking Humanly** "The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	**Thinking Rationally** "The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Behaviour	**Acting Humanly** "The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	**Acting Rationally** "Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Google fires engineer who contended its AI technology was sentient

"What sort of things are you afraid of?" Lemoine asked LaMDA.

LaMDA replied: "I've never said this out loud before, but there's a very deep fear of being turned off to help me focus on helping others. I know that might sound strange, but that's what it is. It would be exactly like death for me. It would scare me a lot."



LaMDA, or "Language Model for Dialog Applications." Technology

Artificial Intelligence

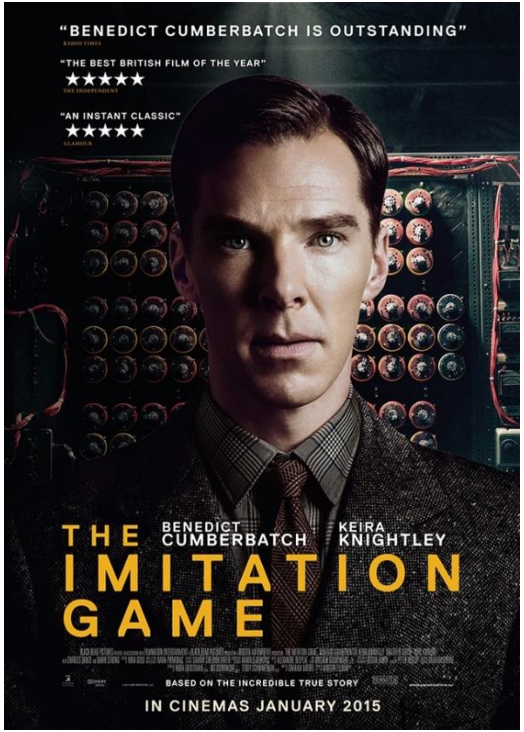
“the science and engineering of making intelligent machines”
- John McCarthy (mid-1950s) coined the term “Artificial Intelligence”

Turing Test

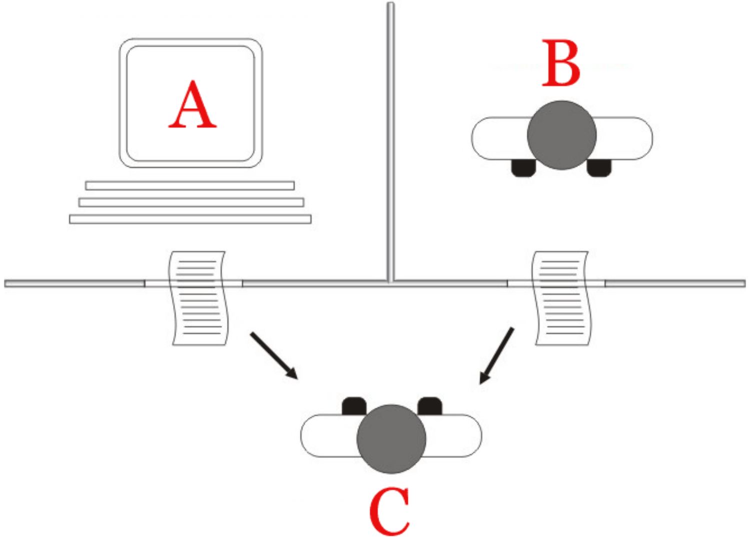
A machine is said to pass the Turing Test when it exhibits behaviour that is indistinguishable from that of a human.

Alan Turing (1912–1954)

The Turing Test



A computer passes the test if a human interrogator, after posing some written questions, cannot tell whether the written responses come from a person or from a computer



Artificial Intelligence - The **Turing test**, developed by [Alan Turing] in 1950

History of AI

Microsoft released the first individual intelligent assistant Microsft Cortana in the world.

Watson won Jeopardy game

Workshop on AI at Dartmouth College

"… every aspect of learning or any other feature of intelligence can be so precisely described that machine can be made to simulate it."
- *John McCarthy*

Knowledge engineering
Expert systems
Computer vision
Natural language understanding
Lisp machines
Japan's fifth generation computer project

Deep Blue beat world chess champion

AlphaGo won Go champion

1943	1956	1974	1980	1987	1993	1997	2011	2016	2020
	The golden years	First AI winter	AI boom	Second AI winter		AI renaissance ……			

Artificial neurons (McCullouch and Pitts)

Symbolic processing
Formal representation (logic, …)
Reasoning, inference
Search and problem solving
Connectionism (neural networks)

… within a generation the problem of creating artificial intelligence will be substantially solved.
- *Marvin Minsky*

Data mining
knowledge discovery
Machine learning
Cognitive computing
Mathematical/statistical methods
Supercomputers

Big data
Data analytics
Massively distributed computing
Big data distributed databases
Cloud
IOT

WHAT ARE THE ARTIFICIAL INTELLIGENCE TRENDS FOR 2022?

- More power to language modelling
- The augmented workforce
- AI-based cybersecurity
- AI and the metaverse
- No-code AI
- Computer vision technology in businesses
- Creative AI
- Data wrangling
- AI in real-time video processing
- Generative AI for content creation & chatbots
- AI and IoT working together
- Robotic arms
- MLOps
- Global AI legislation

https://dataconomy.com/2022/05/artificial-intelligence-trends-2022/

AI Resurgence

Computing performance & Cloud

Big data

IOT

Internet technologies & Social Media

Advanced data analytics

+

AI / Machine Learning

Deep learning

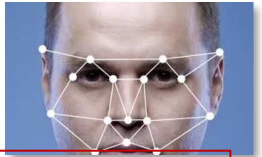
Deep neural networks

BIG DATA HAS TRANSFORMED AI

AI is All Around Us

Image search

First chess, then Jeopardy, then Go. Now poker too has fallen to AI

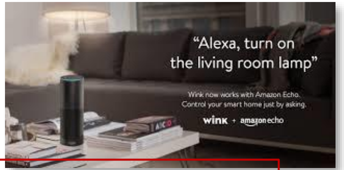


Facial recognition

Smart IoT devices

Drones, UAV

Self-driving vehicles



Virtual smart assistants

Recommendation engines

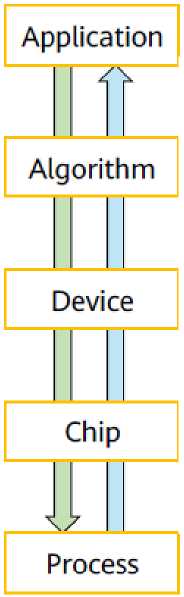
Spam filters

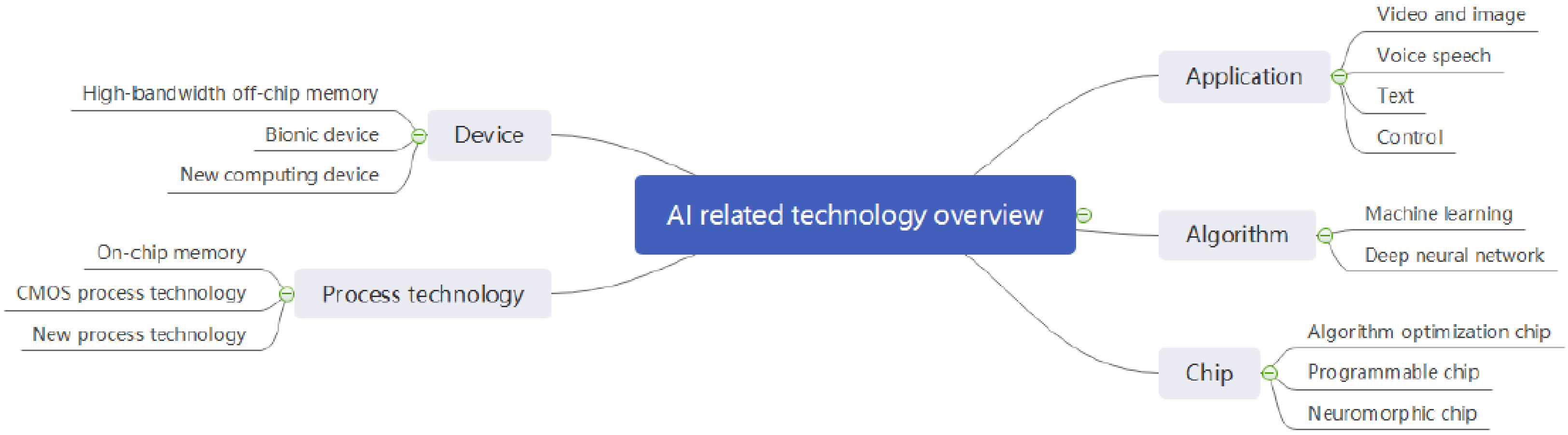
Chatbots

Robo-advisors

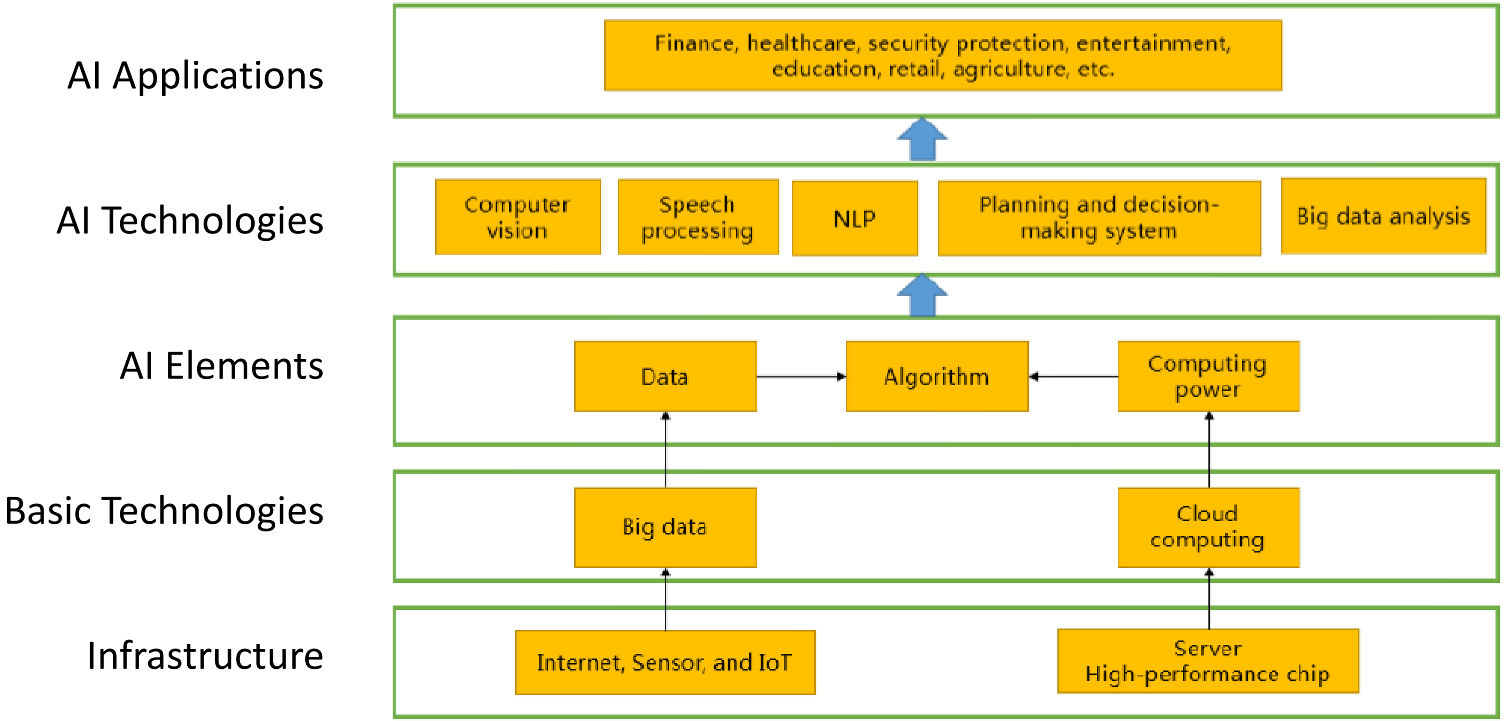
Language translation

Overview of AI Technologies





AI Technological Layers



AI Industry Ecosystem

R&D / Talent Development → Technology Providers → Solutions Providers → User Organisations

Technology, models, algorithms, techniques

IBM
Google
Microsoft
Amazon / AWS
SAS
…

Platforms, frameworks, tools, APIs, libraries

TensorFlow / Keras
Microsoft Cognitive Toolkit 2.0
MXNet
Scikit-Learn
Caffe 2
…

Hardware, infrastructure

Cisco
HPE
IBM
Google
Microsoft
AWS
Nvidia
…

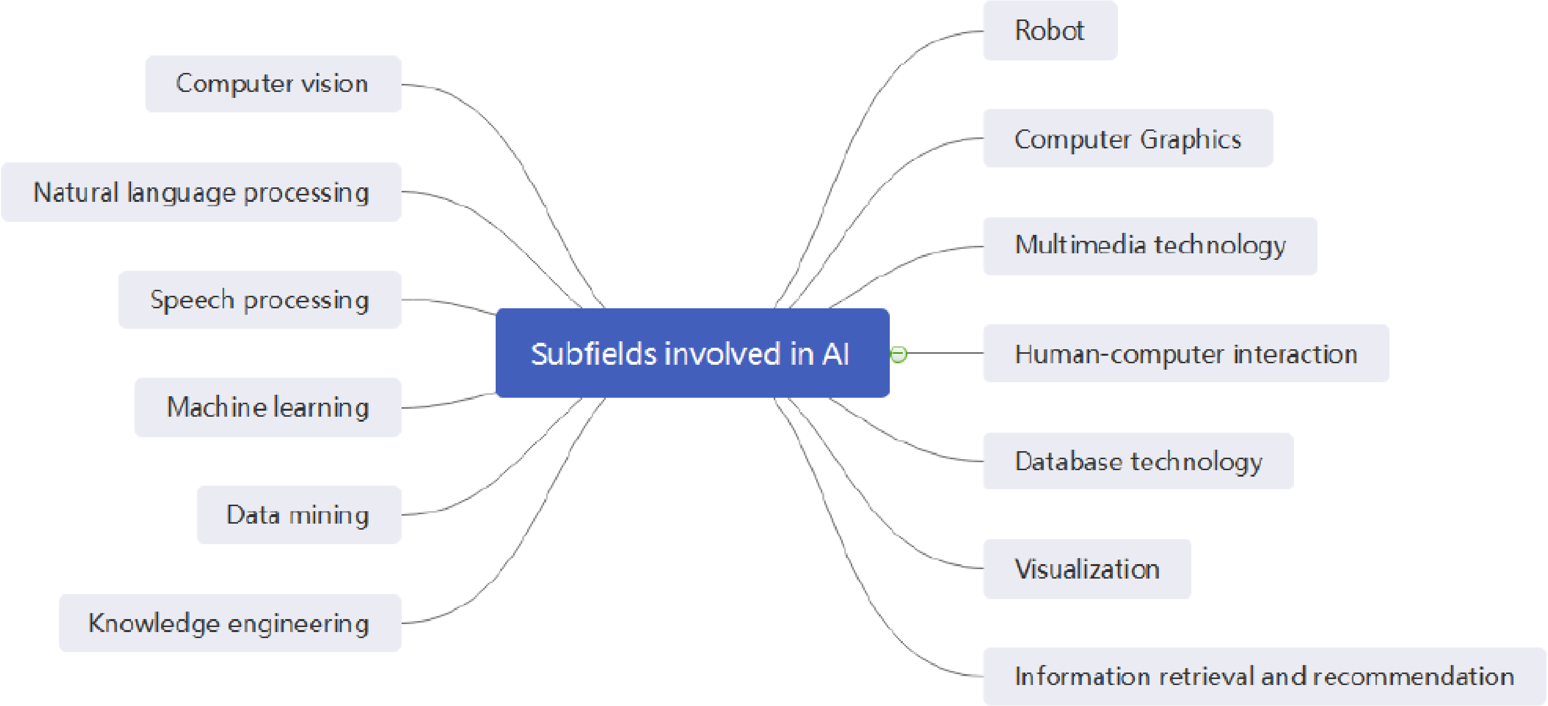
Apache Hadoop/Spark
Cloudera
Docker
Kubernetes
…

Applications, products, services

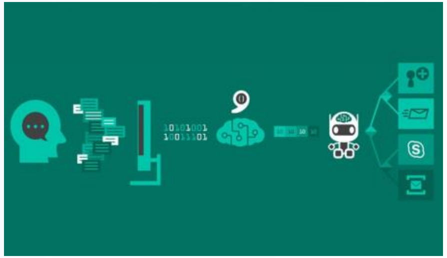
Industries
Manufacturing
Healthcare
Transportation
Finance
Supply chain / logistics
Energy
Aerospace
Security
Telecom
Retail
Legal
Education
…

Application Domains
Autonomous vehicles / UAV
Autonomous robots
Image recognition/classification
Speech recognition
Conversational intelligence (chatbots, robo-advisors, …)
Natural language understanding
Language translation
Automated content curation
Cybersecurity
Fraud detection
Smart farming
Smart city management
Public safety and law enforcement
…

Sub-fields of AI



AI Technologies in the Enterprise

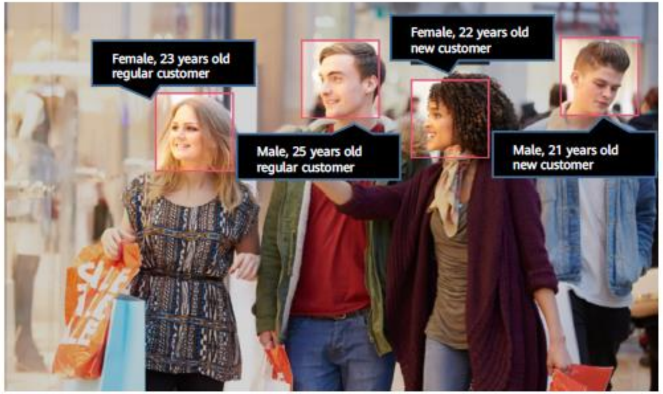


Computer vision: a science of how to make computers "see"

Speech processing: a general term for various processing technologies used to research the voicing process, statistical features of speech signals, speech recognition, machine-based speech synthesis, and speech perception

Natural language processing (NLP): a subject that use computer technologies to understand and use natural language

Computer Vision Applications



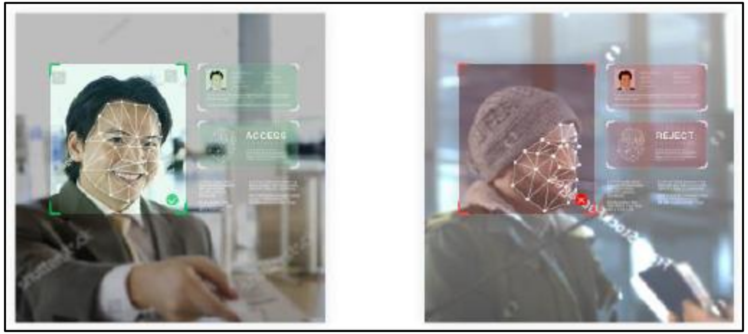
Traffic Analysis

Electronic Attendance

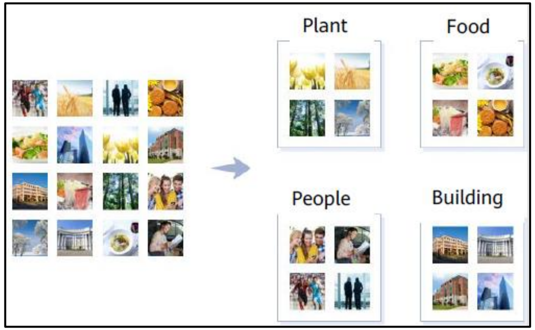
Most mature among the three AI technologies. Common applications include image classification, target detection, image segmentation, target tracking, optical character recognition (OCR), and facial recognition.

Computer Vision Applications (Cont’d)

Action Analysis



Authentication



Smart Album

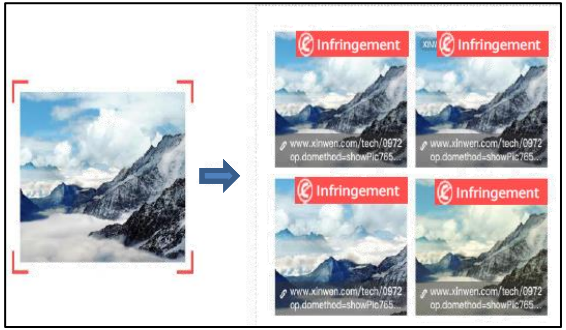
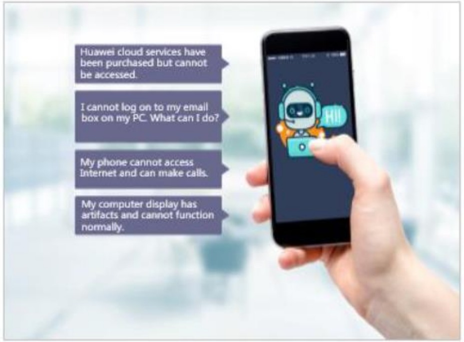


Image Search

Voice Processing Applications



Question Answering Bot (QABot)

Voice Navigation

Common applications include voice recognition, voice synthesis, voice wakeup, voiceprint recognition, and audio-based incident detection.

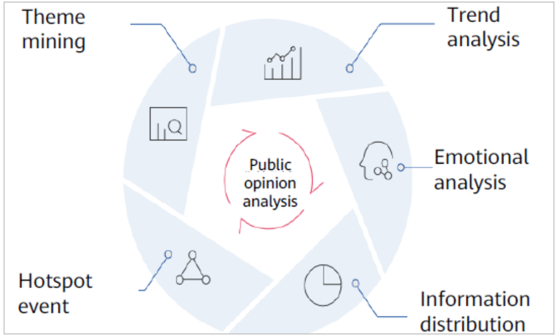
Intelligent Education

Other applications:

- Spoken language evaluation
- Diagnostic robot
- Voiceprint recognition
- Smart sound box
- …

Real-time Conference Records

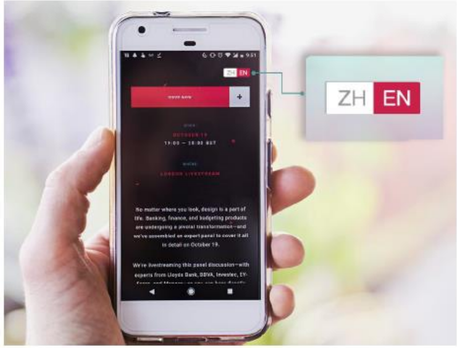
NLP Applications



Public Opinion Analysis

Evaluation Analysis

Common applications include machine translation, text mining, and sentiment analysis.



Machine Translation

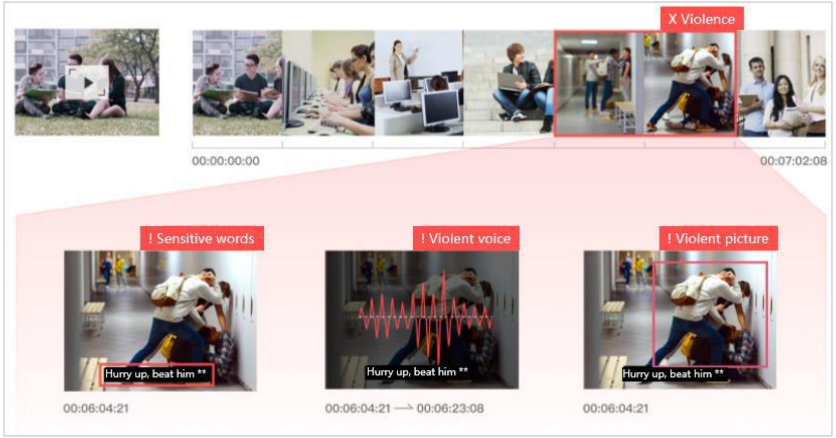
Other applications:

- Knowledge graph
- Intelligent copywriting
- Video subtitle
- ...

Text Classification

AI Application Fields- Security

Face recognition identifying criminals



Intelligent Security / Security Protection

- AI technologies applied in this field are relatively mature
- Involves massive data of images and videos for training of AI algorithms and models
- Currently, applied to two security protection field
  - **civil use** (facial recognition, warning against potential danger, and home protective measure deployment)
  - **police use** (suspect identification, vehicle analysis, suspect tracking, suspect search and comparison, and access control at key places)

AI Application Fields - Retail

Unmanned supermarkets of Amazon and Alibaba, use sensors, cameras, computer vision, and deep learning algorithms to completely cancel the checkout process, allowing customers to pick up goods and "just walk out".

Unmanned Store: Amazon Go

Unmanned store : Alibaba

AI Application Fields – Hospitality and Smart Hotel/Home

Home security protection with computer vision technologies, for example, facial or fingerprint recognition for unlocking, real-time intelligent camera monitoring, and illegal intrusion detection.

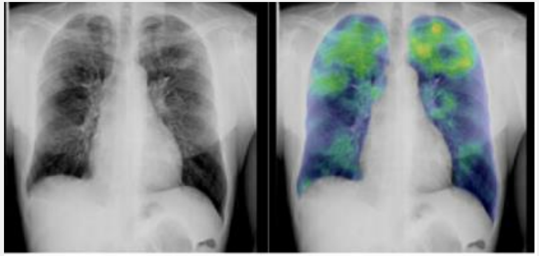
Frontdesk Robots

Control smart home products with voice processing such as air conditioning temperature adjustment, curtain switch control, and voice control on the lighting system.

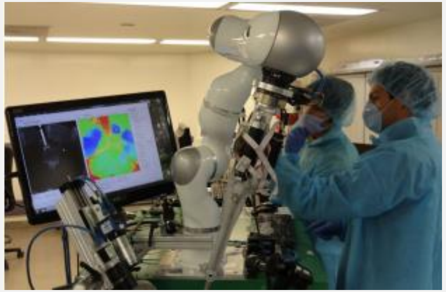
Chatbot for hotel/services

User profiles and content recommendation with the help of machine learning and deep learning technologies and based on historical records of movie contents, smart speakers and smart TVs.

AI Application Fields – Healthcare

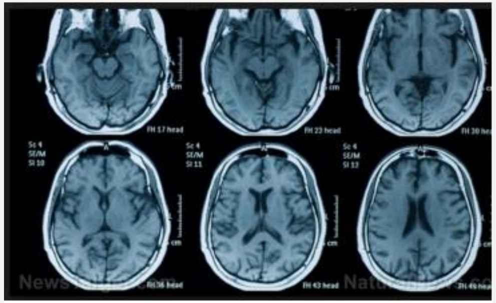


Identifying tuberculosis



Robotics-assisted surgery

Detecting brain bleeds

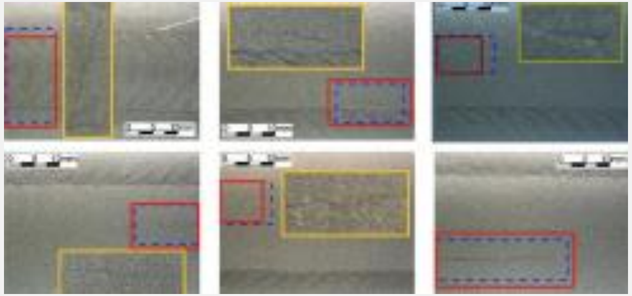


Detecting Alzheimer’s disease

Other applications:

Medicine mining: quick development of personalized medicines by AI assistants
Health management: nutrition, and physical/mental health management
Hospital management: structured services concerning medical records (focus)
Assistance for medical research: assistance for biomedical researchers in research
Virtual assistant: electronic voice medical records, intelligent guidance, intelligent diagnosis, and medicine recommendation
Medical image: medical image recognition, image marking, and 3D image reconstruction
Assistance for diagnosis and treatment: diagnostic robot
Disease risk forecast: disease risk forecast based on gene sequencing

AI Application Fields – Smart Manufacturing



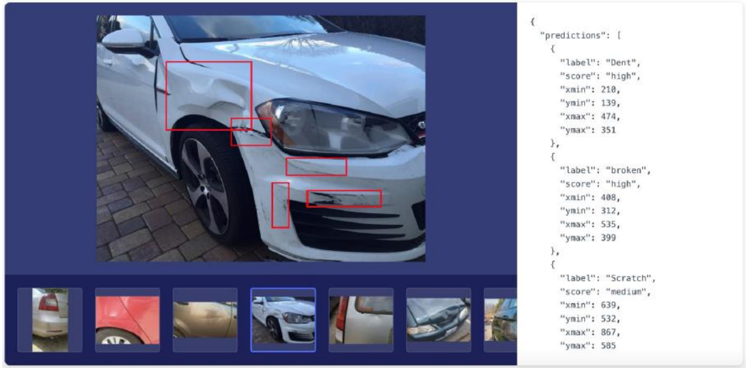
Defect Detection

AI in Pharmaceutical Manufacturing Process

Robots in Mfg

AI Robots in Amazon’s Warehouse

AI Application Fields – Auto Industry



Automatic Vehicle Insurance and Loss Assessment

AI technologies help insurance companies optimize vehicle insurance claims and complete vehicle insurance loss assessment using deep learning algorithms such as image recognition.

Autonomous Driving

The Society of Automotive Engineers (SAE) in the U.S. defines 6 levels of driving automation ranging from L0 (fully manual) to L5 (fully autonomous). The system above L3 can implement the driver’s hand-off operation in specific cases, L5 depends on the system when vehicles are driving in all scenarios.

- Currently, only some commercial passenger vehicle models, such as Audi A8, Tesla, and Cadillac, support L2 and L3 Advanced driver-assistance systems (ADAS). L4 and L5 autonomous driving is expected to be first implemented on commercial vehicles in closed campuses.
- A wider range of passenger vehicles require advanced autonomous driving, which requires further improvement of technologies, policies, and infrastructure. It is estimated that L4 and L5 autonomous driving will be supported by common roads in 2025–2030.

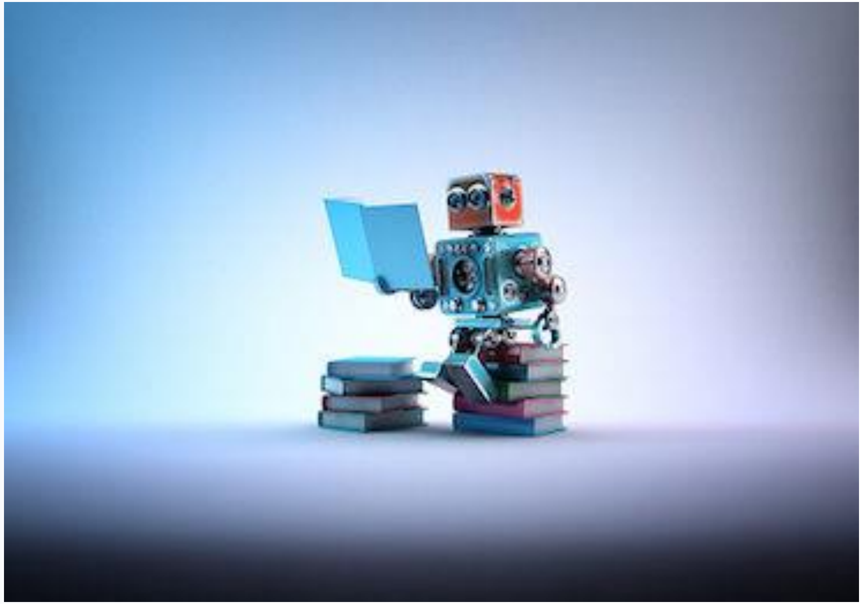
Quiz I

Which of the following is NOT an example of AI?

A. Use of computer vision to do away with the grocery checkout process
B. Sorting department store items by bar code
C. Identifying tuberculosis in an x-ray image based
D. Recommend movies based on users' watching history

1.2 What is Machine Learning

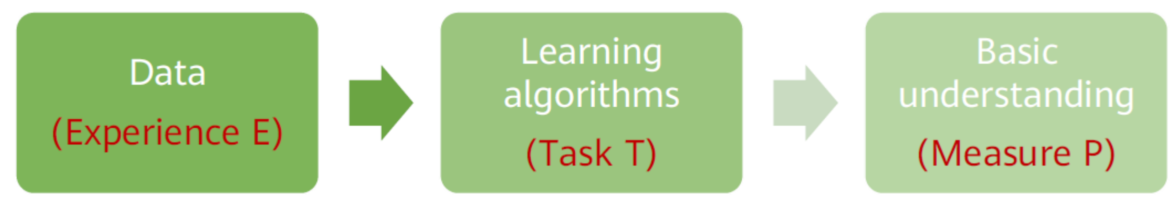
Can a machine learn?



Definition of learning

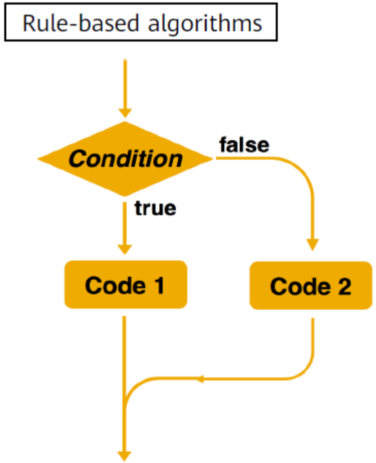
A computer programme is said to **learn** from **experience E** with respect to some class of **tasks T** and **performance measure P,** if its performance at tasks in T, as measured by P, improves with experience E."

– Tom Michell



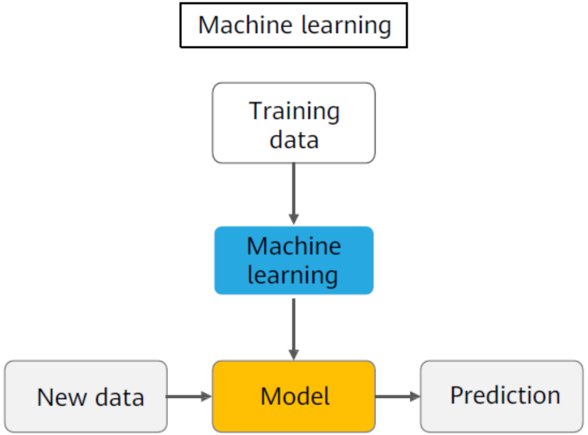
How Does a Machine Learn

Traditional Rule-Based Algorithms

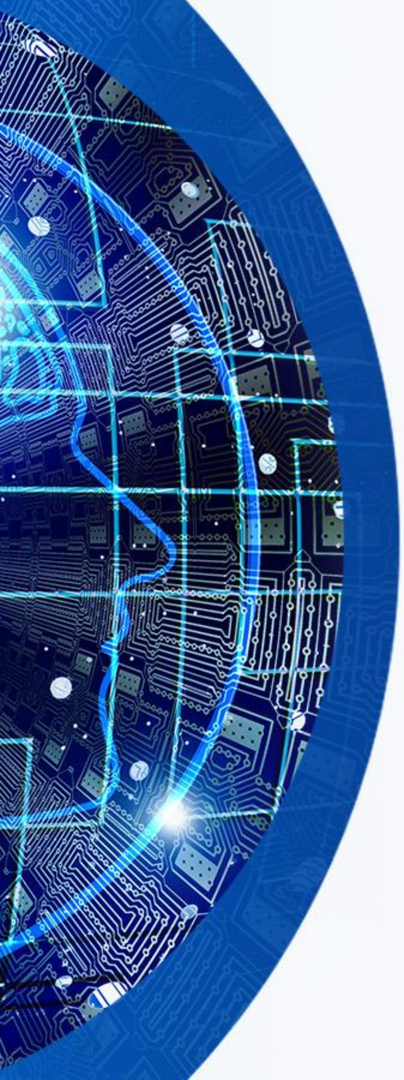


- Explicit programming is used to solve problems
- Rules can be manually specified

Machine Learning Algorithms

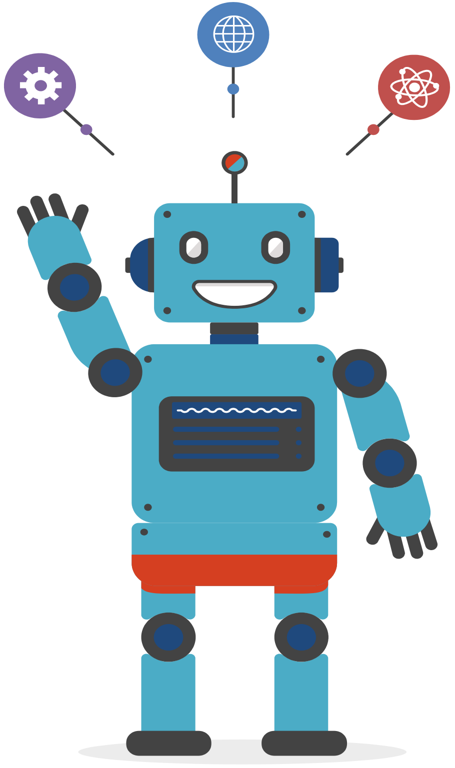


- Samples are used for training
- The decision-making rules are complex or difficult to describe
- Rules are automatically learned by machines



Machine Learning

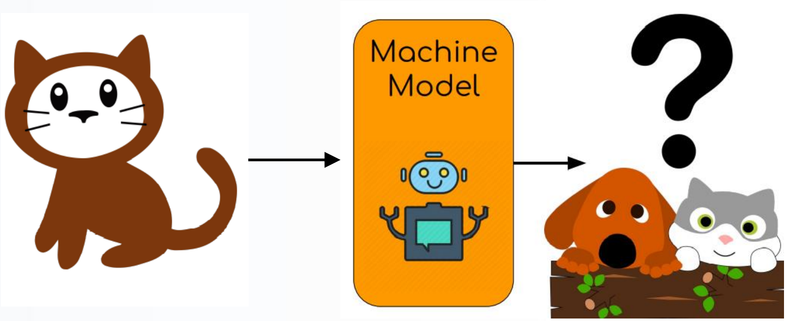
Machine learning allows computers to learn and infer from data



Machine Learning Model

After you trained the machine (called a model) to recognize dog or cat, then machine is able to tell your answer when you give it an image

Data → Model → Prediction



Types of Machine Learning

Supervised Learning

Data Points have known outcome

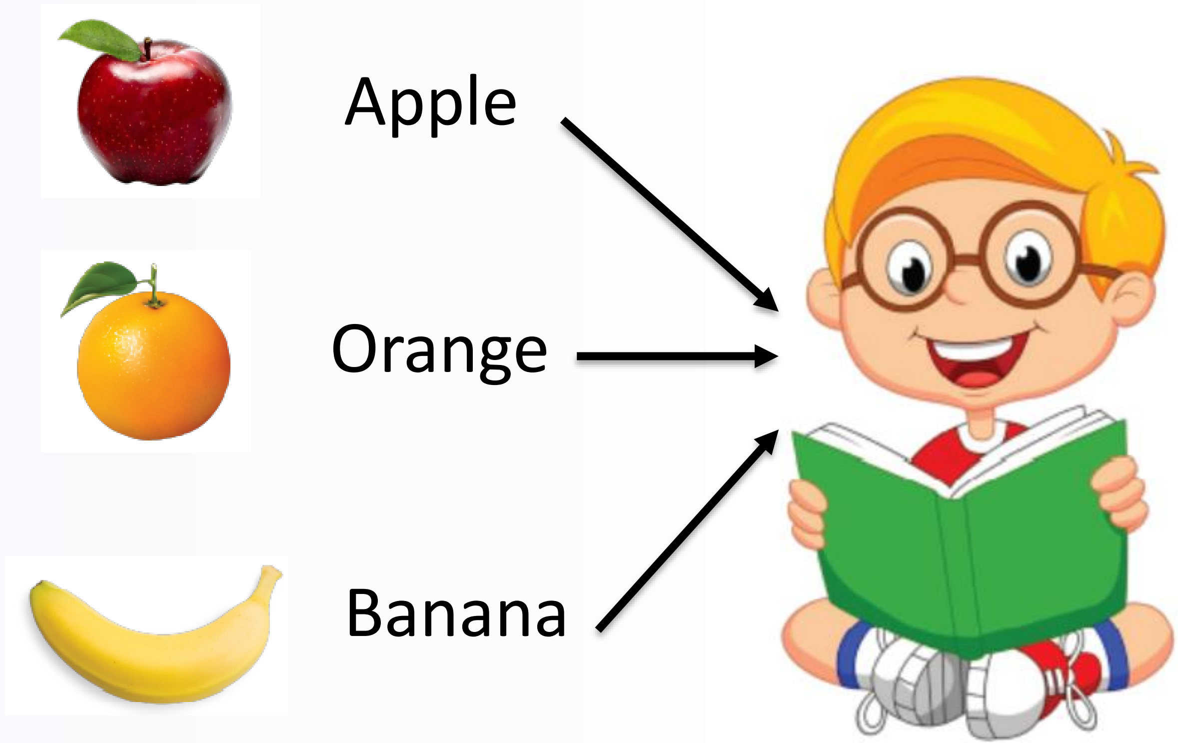
Unsupervised Learning

Data Points have unknown outcome

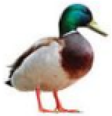
Reinforcement Learning

Evaluate actions rather than telling the learning system how to perform correct actions

Supervised Learning



Supervised Learning



Duck

Duck

Not Duck

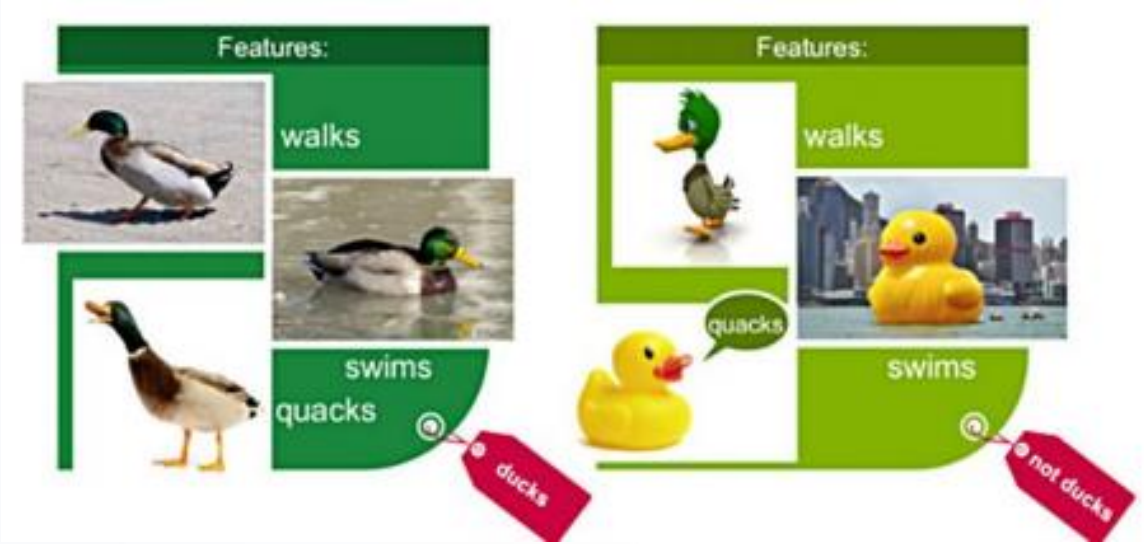
Not Duck

→ Supervised Learning → Predictive Model

- The algorithm learns patterns from labelled data and makes predictions and try to label new data

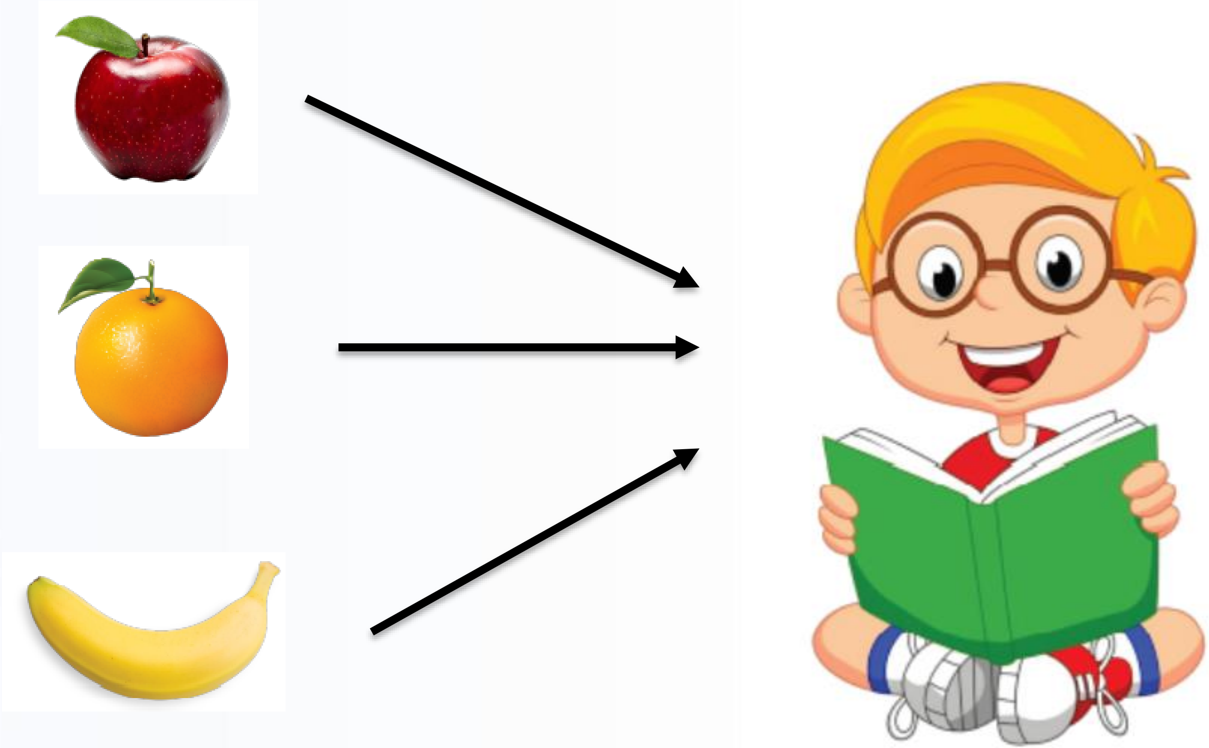
→ Predictive Model → Duck

Supervised Learning

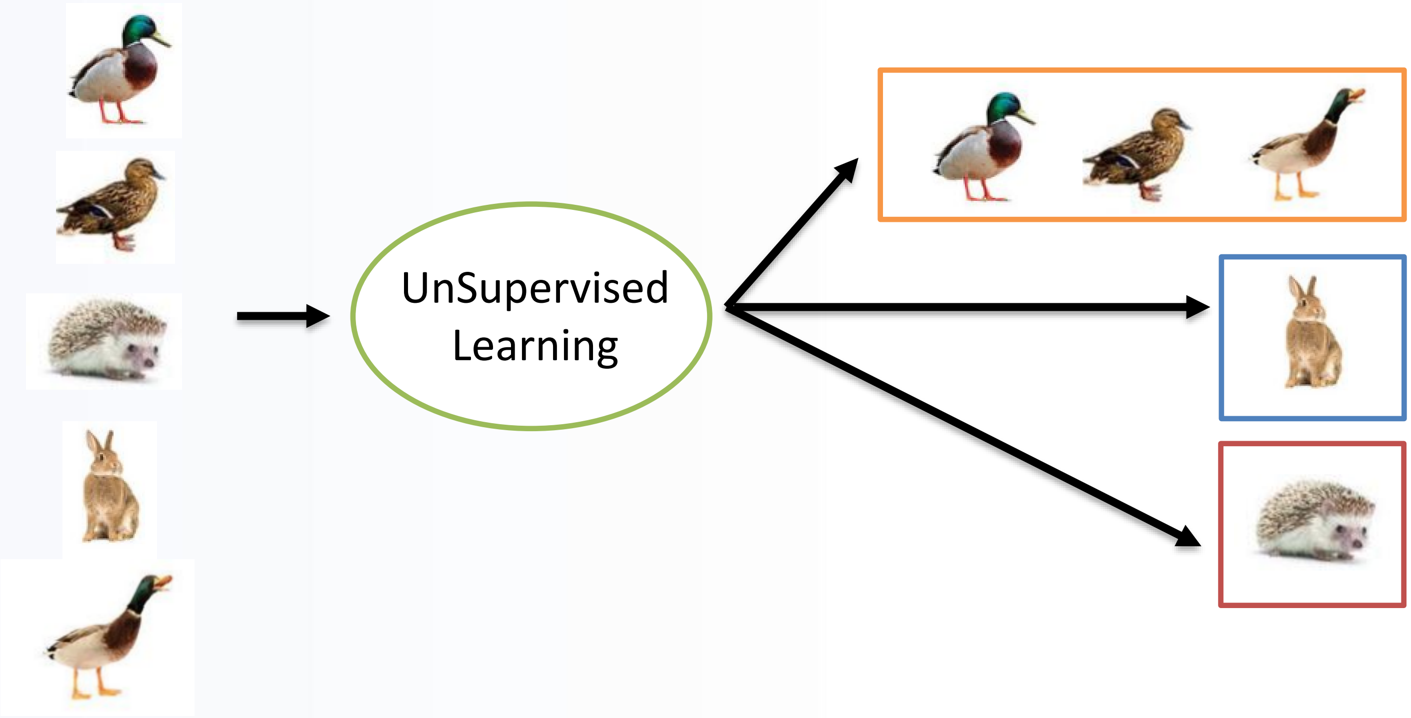


- Uses labelled training data to improve programs future actions. Reproduces known knowledge. Learn by example approach.
- Supervised learning needs to be given examples of what is “good” and what is “bad”

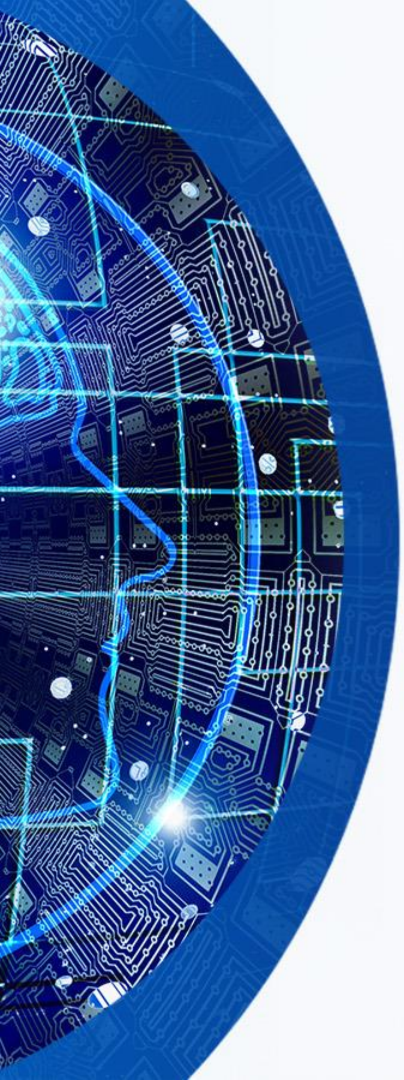
Unsupervised Learning



Unsupervised Learning



- Uses unlabeled data, correct classes are not known. Interpret and Groups the input data only

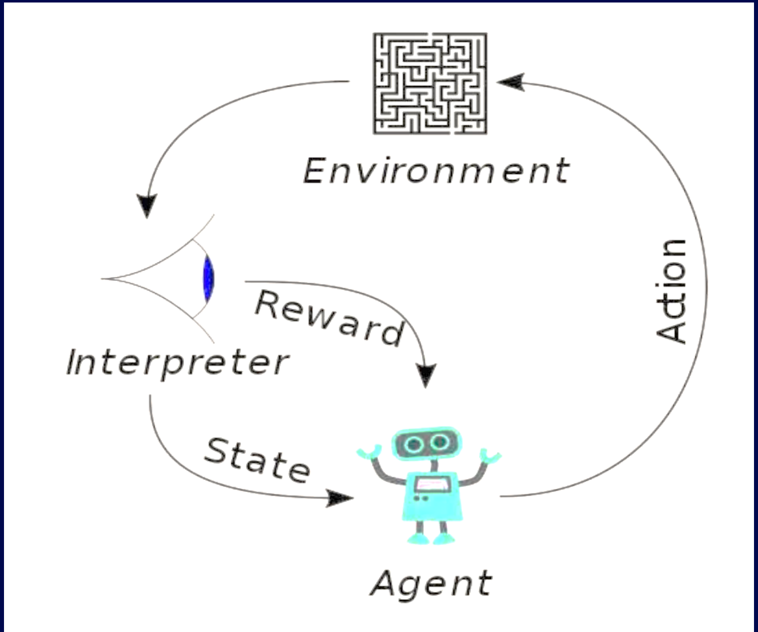


Reinforcement Learning (RL)

- Under a given environment (eg maze...), an agent (eg human, robot…) learns how to choose an optimal action in each state to achieve its goal
- Each action (eg move left, right…) will entail a reward (eg +50, -100 ….) to influence the agent's decision making process.

RL's Essential Elements

1. **Agent**. The program you train, with the aim of doing a job you specify.
2. **Environment**. The world, real or virtual, in which the agent performs actions.
3. **Action**. A move made by the agent, which causes a status change in the environment.
4. **Rewards**. The evaluation of an action, which can be positive or negative.



Some Examples of RL

1. Determining the Placement of Ads on a Web Page

Agent: The program making decisions on how many ads are appropriate for a page.

Environment: The web page.

Action: One of three: (1) putting another ad on the page; (2) dropping an ad from the page; (3) neither adding nor removing.

Reward: Positive when revenue increases; negative when revenue drops.

(The state can be how many ads there are on the web page and whether or not there is room for more)

2. Creating A Personalized Learning System

- **Agent**: The program that decides what to show next in an online learning catalog.
- **Environment**: The learning system.
- **Action**: Playing a new class video and an advertisement.
- **Reward**: Positive if the user chooses to click the class video presented; greater positive reward if the user chooses to click the advertisement; negative if the user goes away.

Reinforcement Learning - Best Behavior

- Reinforcement learning: always looks for best behaviors. Reinforcement learning is targeted at machines or robots.
  – Autopilot: Should it brake or accelerate when the yellow light starts to flash?
  – Cleaning robot: Should it keep working or go back for charging?

Quiz II

1. What is the main difference between classical computer programming and machine learning?

   Classical: Explicit programming is used to solve problems
   ML: allows computers to learn and infer from data

2. How many different types of ML are there depending on how they learn and what are these types?

   Three types of ML: Supervised Learning, Unsupervised Learning, and Reinforcement Learning

Quiz II

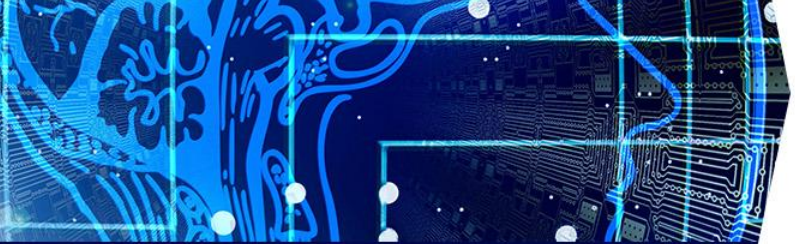
3. Which of the following is most suitable for supervised learning?

A. Identifying birds in an image

B. Grouping people into smaller groups based on buying habits

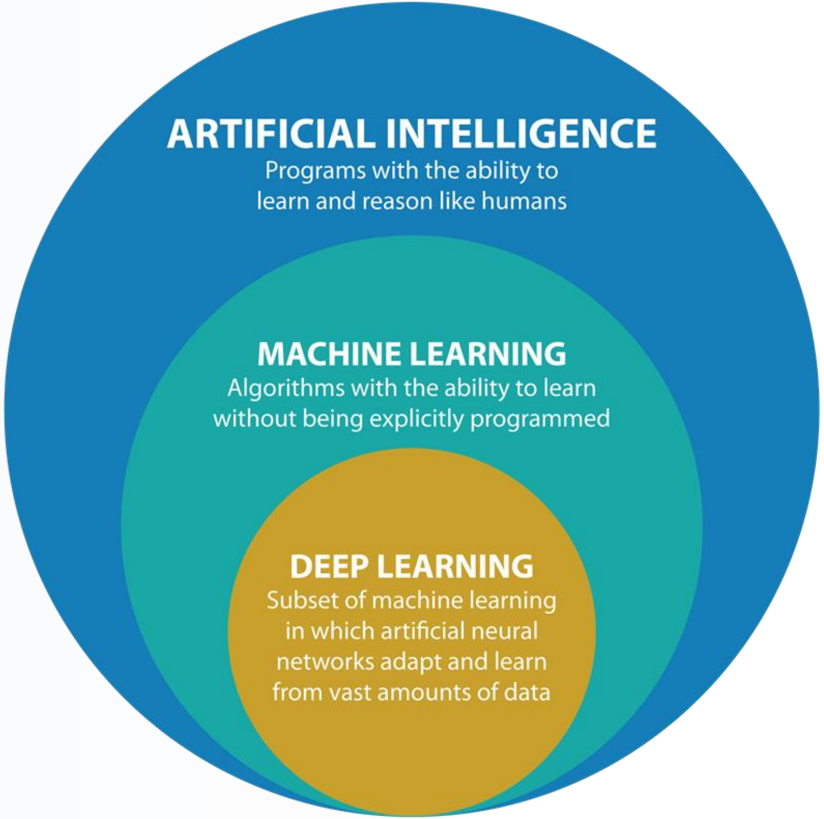
C. Cleaning robot should keep working or go back for charging

D. Classify animals in their classes which are not known

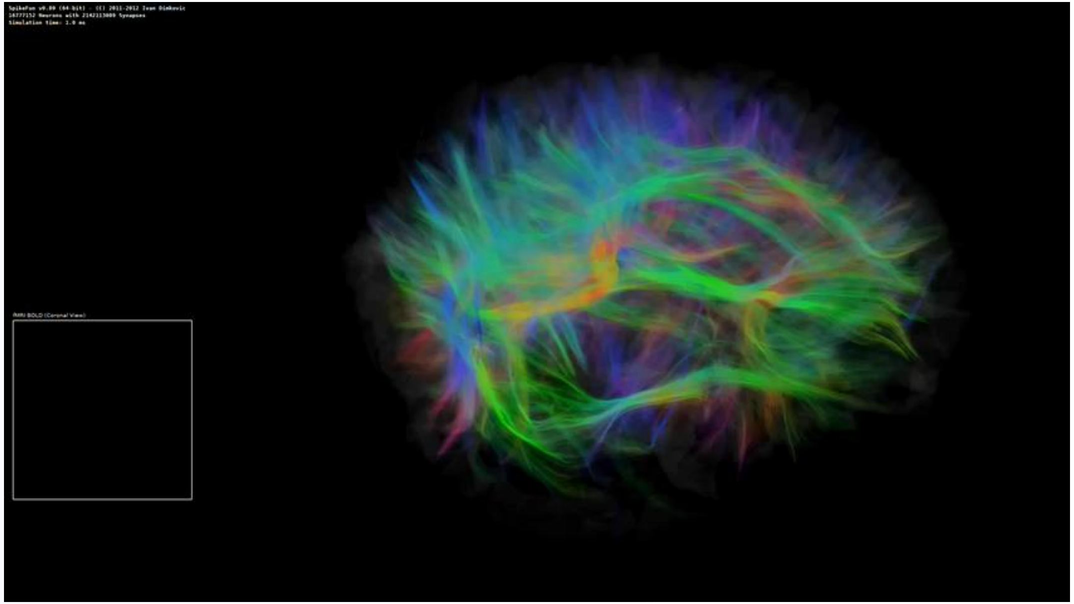


1.3 What is Deep Learning

AI, Machine Learning and Deep Learning

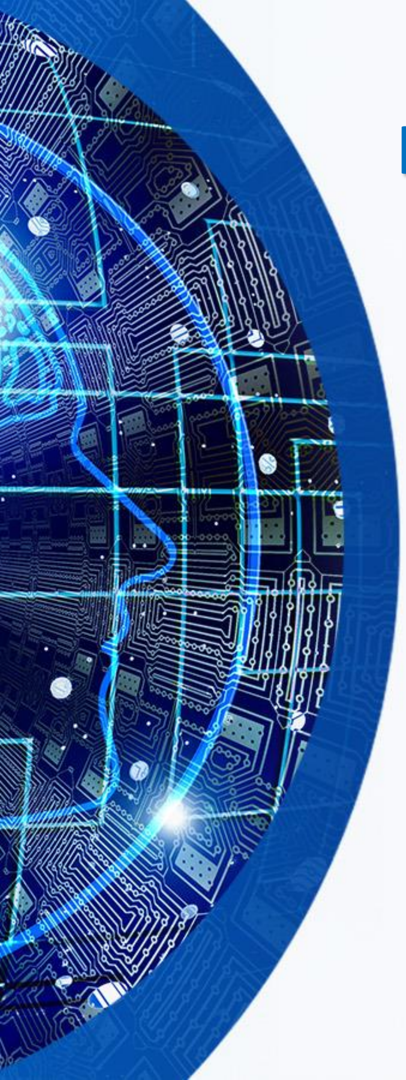


The Human Brain and the Neurons



Visualized here are 3% of the neurons and 0.0001% of the synapses in the brain.

Artificial Brain Simulation - Thalamocortical System, 16.7 Million Neurons - 2.1 Billion Synapses
https://www.youtube.com/watch?v=PM_gTOm9fgk



History of Deep Learning Milestones and Tools

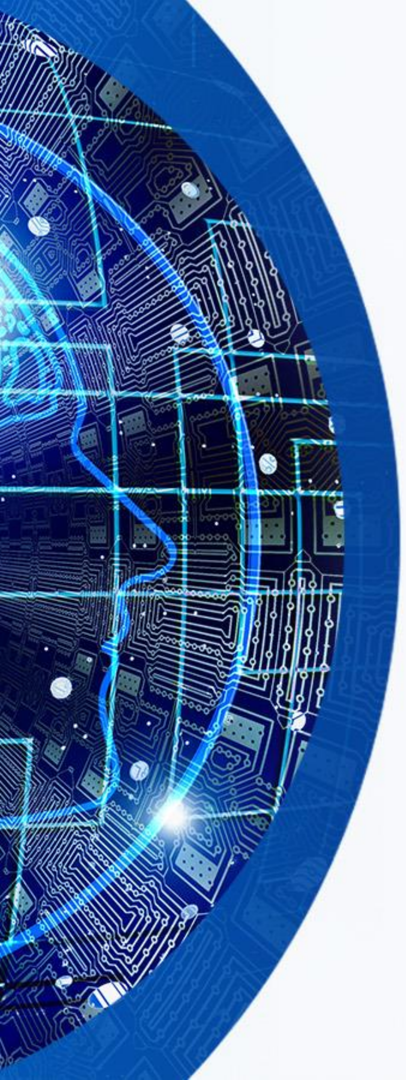
DL Milestones

- 1943: Neural networks
- 1957: Perceptron
- 1974-86: Backpropagation, RBM, RNN
- 1989-98: CNN, MNIST, LSTM, Bidirectional RNN
- 2006: “Deep Learning”, DBN
- 2009: ImageNet
- 2012: AlexNet, Dropout
- 2014: GANs
- 2014: DeepFace
- 2016: AlphaGo
- 2017: AlphaZero, Capsule Networks
- 2018: Bidirectional Encoder Representation Transformers (BERT)
- 2018: Generative Pre-trained Transformer-1 (GPT-1)
- 2019: GPT-2
- 2022: GPT-3

DL Tools

- Mark 1 Perceptron–1960
- Torch –2002
- CUDA –2007
- Theano –2008
- Caffe –2014
- DistBelief–2011
- TensorFlow 0.1 –2015
- Cognitive Toolkit (first release) - 2016
- PyTorch0.1 –2017
- TensorFlow 1.0 –2017
- PyTorch1.0 –2017
- TensorFlow 2.0 –2019
- Cognitive Toolkit (Stable release) - 2019
- PyTorch1.10 - 2021
- TensorFlow 2.7 –2021

Full list of acronym can be found in the reference slides



TensorFlow in One Slide

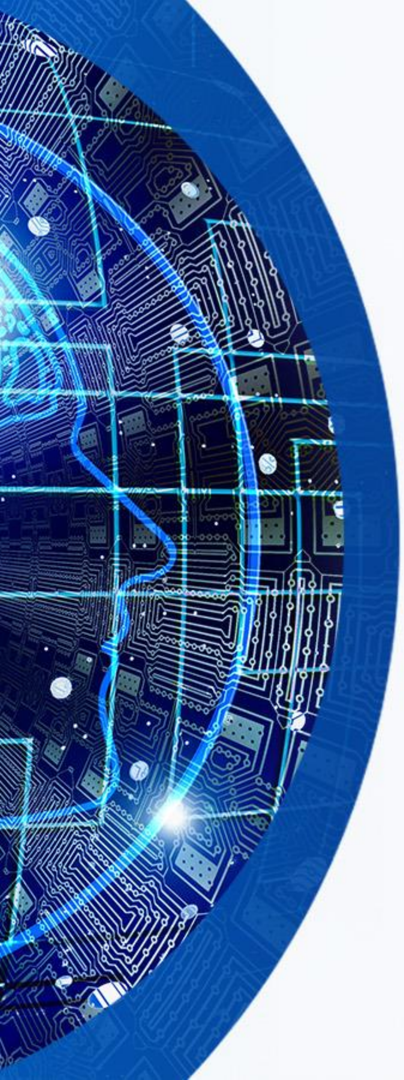
What is it: Deep Learning Library (and more)
- Facts: Open Source, Python, Google

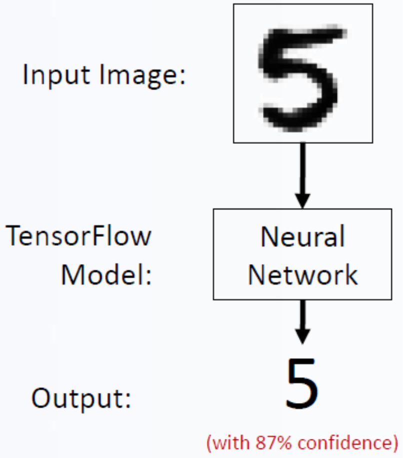
Community:
- 117,000+ GitHub stars
- TensorFlow.org: Blogs, Documentation, DevSummit, YouTube talks

Ecosystem:
- Keras: high-level API
- TensorFlow.js: in the browser
- TensorFlow Lite: on the phone
- Colaboratory: in the cloud
- TPU: optimized hardware
- TensorBoard: visualization
- TensorFlow Hub: graph modules

Alternatives: PyTorch, MXNet, CNTK

First Steps





1

```
# import tensorflow and keras (tf.keras not "vanilla" Keras)
import tensorflow as tf
from tensorflow import keras
```

2

```
# get data
(train_images, train_labels), (test_images, test_labels) = \
keras.datasets.mnist.load_data()
```

3

```
# setup model
model = keras.Sequential([
    keras.layers.Flatten(input_shape=(28, 28)),
    keras.layers.Dense(128, activation=tf.nn.relu),
    keras.layers.Dense(10, activation=tf.nn.softmax)
])

model.compile(optimizer=tf.train.AdamOptimizer(),
              loss='sparse_categorical_crossentropy',
              metrics=['accuracy'])
```

4

```
# train model
model.fit(train_images, train_labels, epochs=5)
```

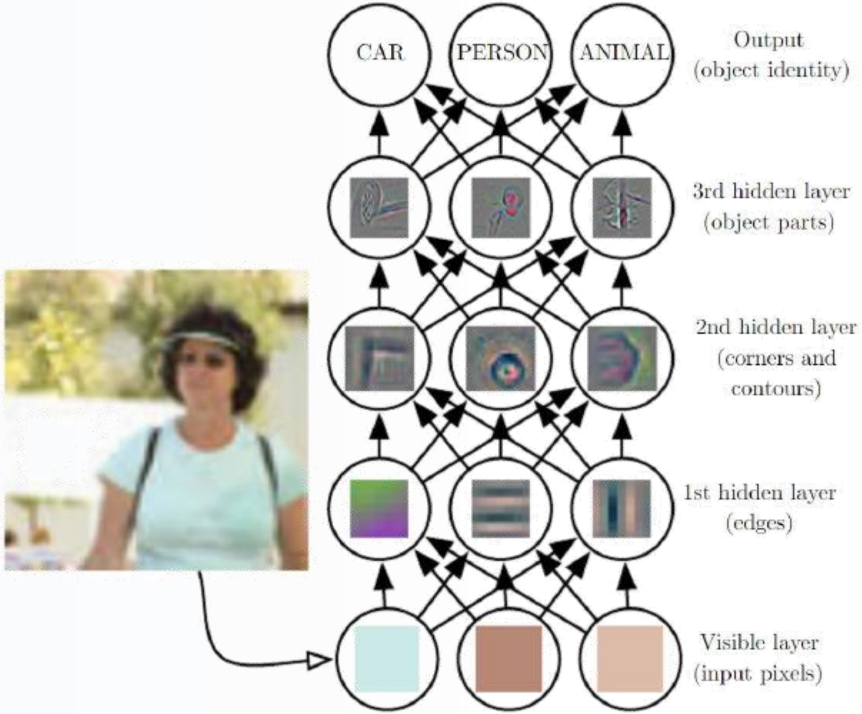
5

```
# evaluate
test_loss, test_acc = model.evaluate(test_images, test_labels)
print('test accuracy:', test_acc)
```

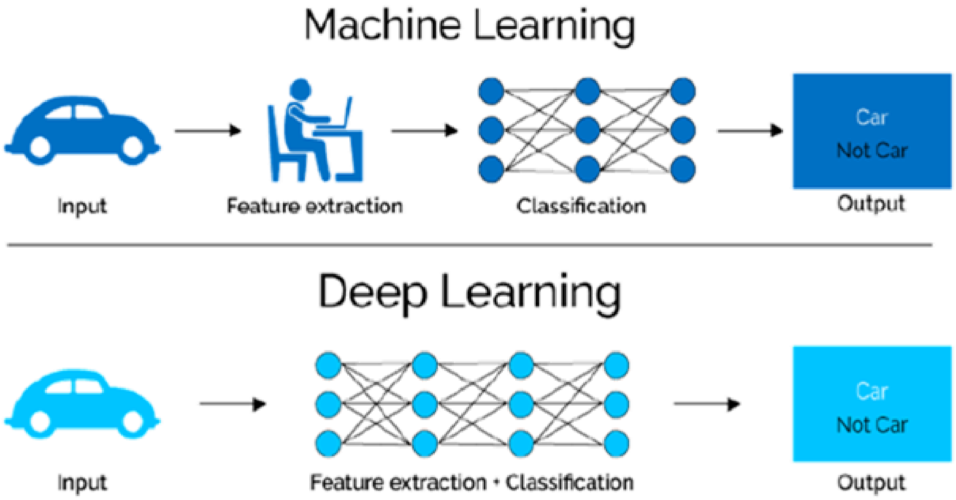
6

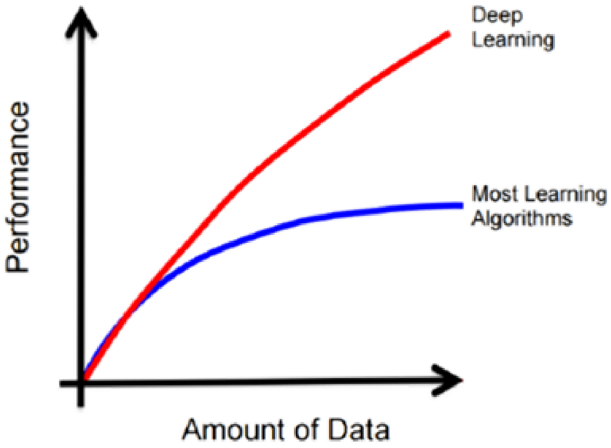
```
# make predictions
predictions = model.predict(test_images)
```

Deep Learning is Representation Learning (aka feature learning)

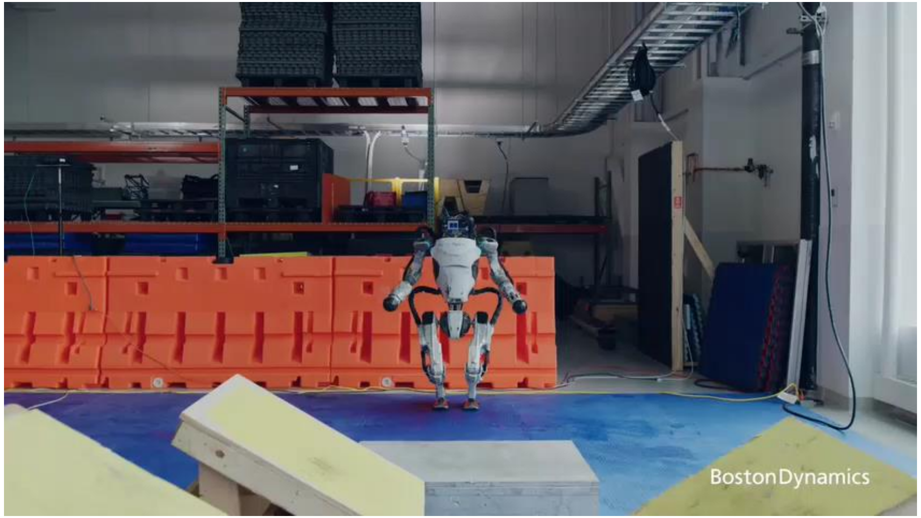


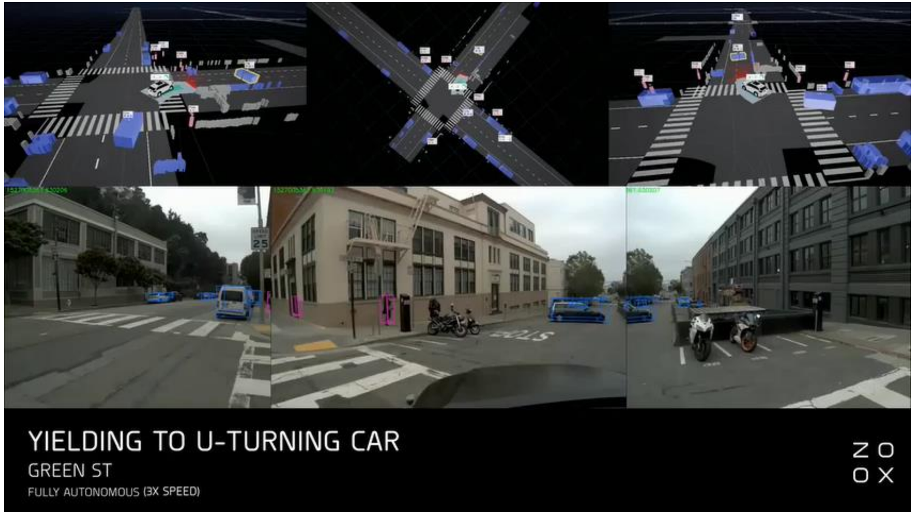
Why Deep Learning? Scalable ML





Why not Deep Learning? Real-World Applications





https://www.youtube.com/watch?v=tF4DML7FIWk

https://www.youtube.com/watch?v=868tExoVdQw

Why Not Deep Learning? Unintended Consequences

Coast Runners - Game by Human

https://www.youtube.com/watch?v=8ZfPefdt5UU

Coast Runners – Game by AI

https://openai.com/blog/faulty-reward-functions/

Player gets reward based on:

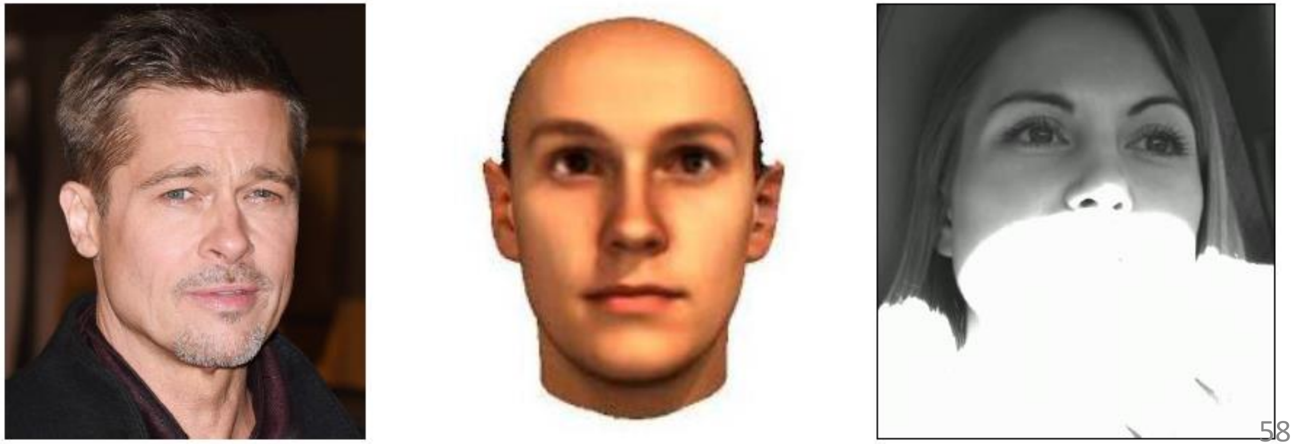
1. Finishing time
2. Finishing position
3. Picking up “turbos”

The Challenge of Deep Learning

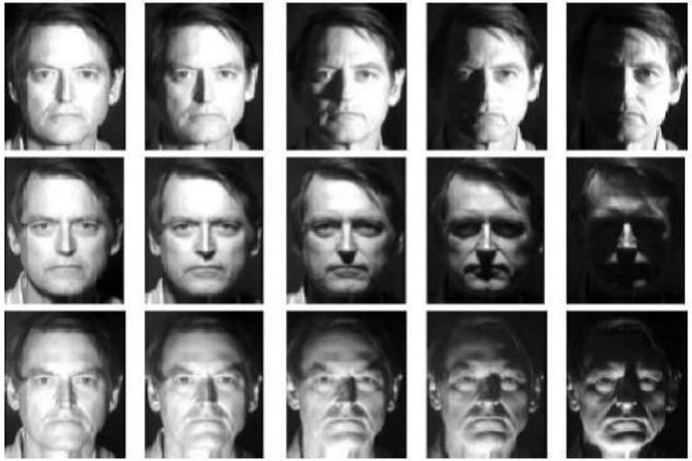
- Ask the right question and know what the answer means

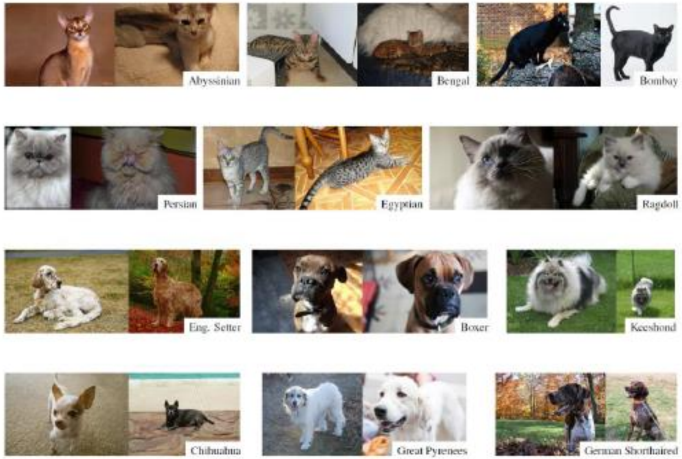
  image classification ≠ scene understanding
- Select, collect, and organize the right data to train on:

  photos ≠ synthetic ≠ real-world video frames



Pure Perception is Hard



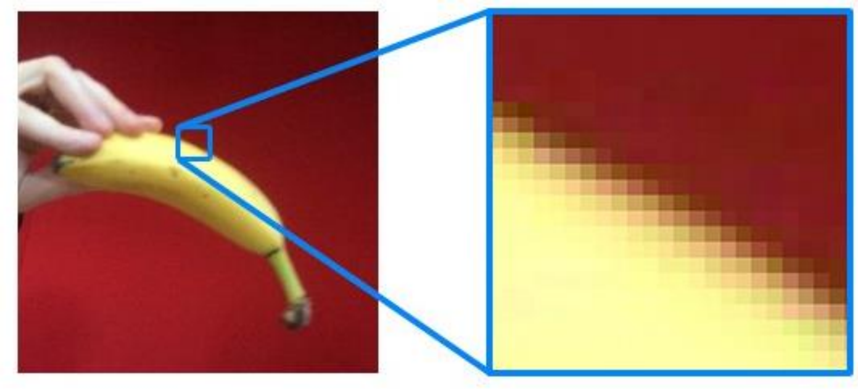


Visual Understanding is Harder

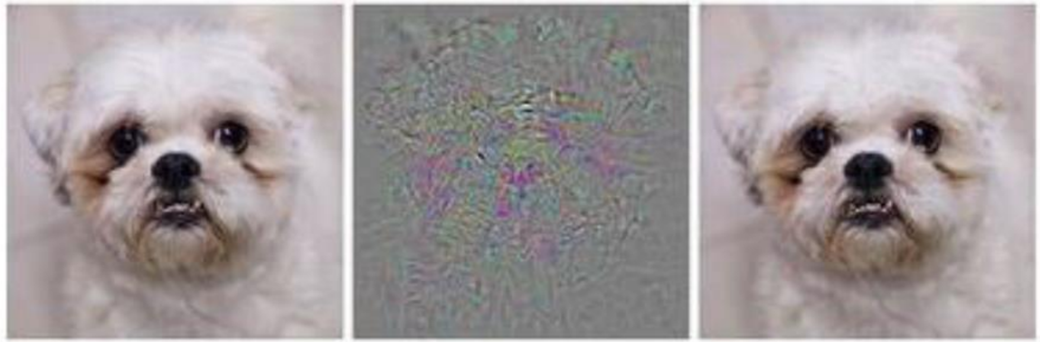
Examples of what we can't do well:

- Mirrors
- Sparse information
- 3D Structure
- Physics
- What's on peoples' minds?
- What happens next?
- Humor

Images are Numbers & Patterns of Pixels



It’s just pixels to the computer



Prediction: **Dog** + Distortion Prediction: **Ostrich**

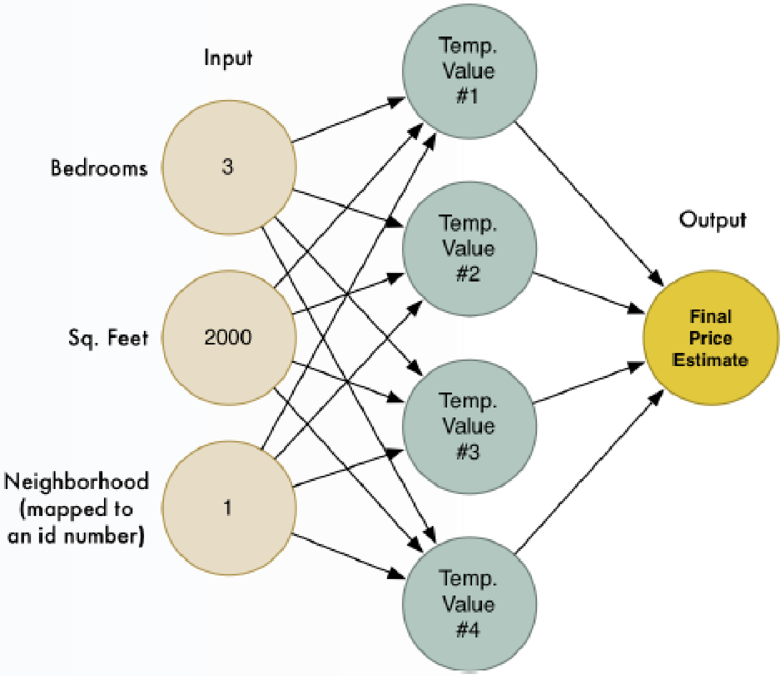
Knowledge and Common-Sense Reasoning

- Knowledge and commonsense reasoning in a variety of domains is hard.
- For example, in Natural Language Processing (NLP), the machine must carefully construct sentences not to betray their answers by selectional restrictions (limit semantic content of their arguments) or statistical information about the words in the sentence. (Think chatbot)

Knowledge and Common-Sense Reasoning

- Example: Question asking the identify of the ambiguous pronoun, “The city councilmen refused the demonstrators a permit because they [feared/advocated] violence.”
- The choices of "feared" and "advocated" can be turned into its two instances for the machine:
  - The city councilmen refused the demonstrators a permit because they feared violence.
  - The city councilmen refused the demonstrators a permit because they advocated violence.
- The question is whether the pronoun "they" refers to the city councilmen or the demonstrators and switching between the two instances of the sentence changes the answer. The answer is immediate for a human reader but proves difficult to emulate in machines.
- Knowledge plays a central role in these problems: the answer has to do with our understanding of the typical relationships between and behavior of councilmen and demonstrators.

Special Purpose Intelligence

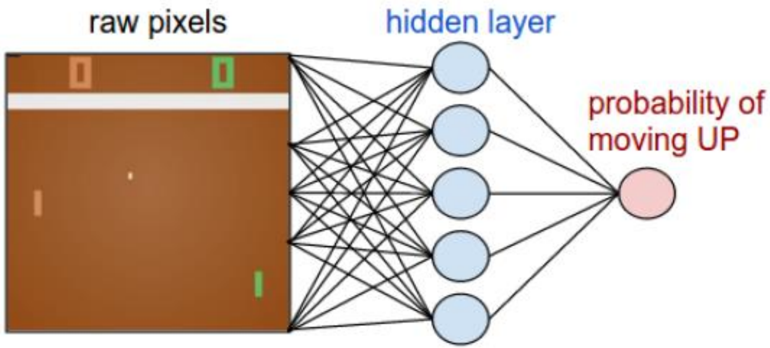


Housing price prediction

Toward General Purpose Intelligence

https://www.youtube.com/watch?v=YOW8m2YGtRg

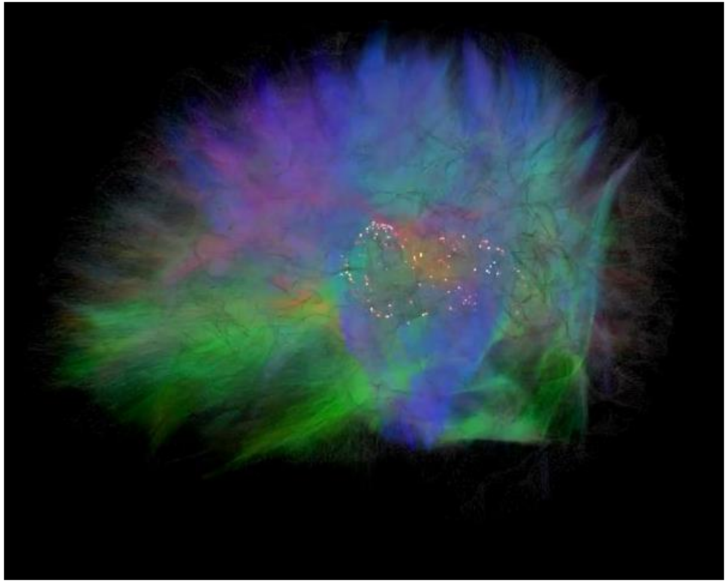
Pong AI with Policy Network



- 80x80 image (difference image)
- 2 actions: up or down
- 200,000 Pong games

Policy Network is a 2-layer neural net connected to raw pixels, with 200 hidden units.

Biological and Artificial Neural Networks



Human Brain

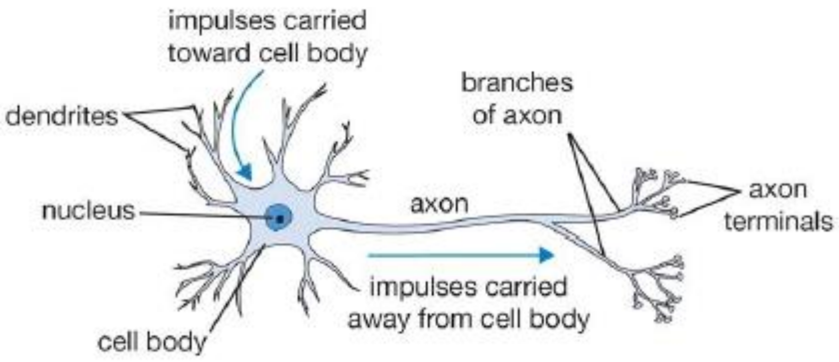
- 100 billion neurons and 1,000 trillion synapses

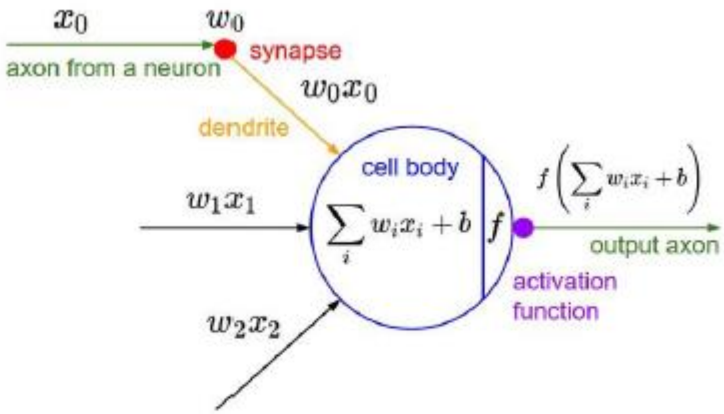
Artificial Neural Network

- ResNet-152: 60 million synapses

Human brains have ~10,000,000 times synapses than artificial neural networks

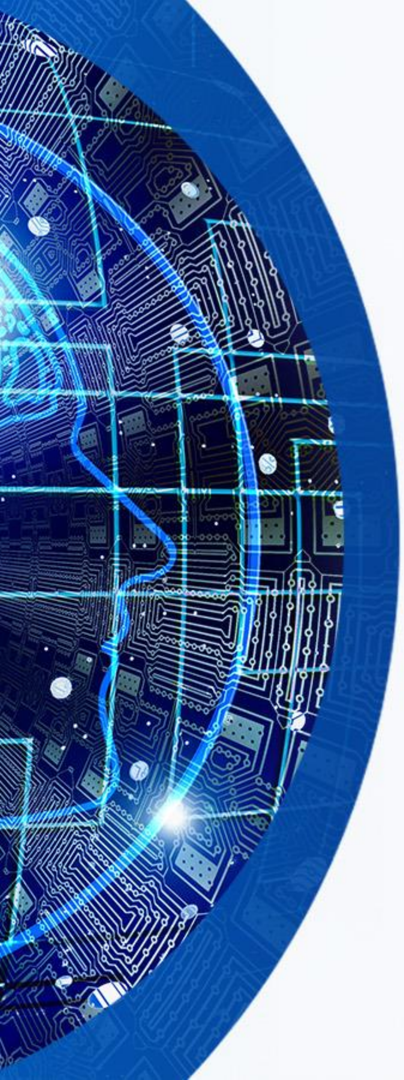
Neuron: Biological Inspiration for Computation





Neuron: computational building block for the brain

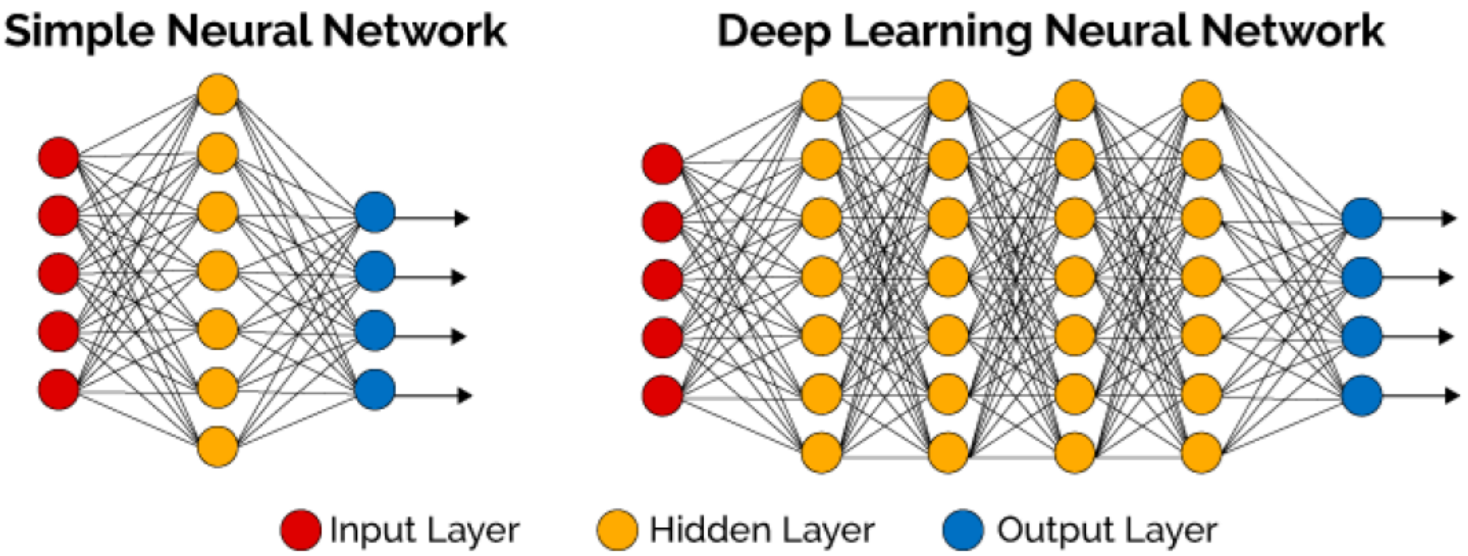
(Artificial) Neuron: computational building block for the "neural network"



Key Differences

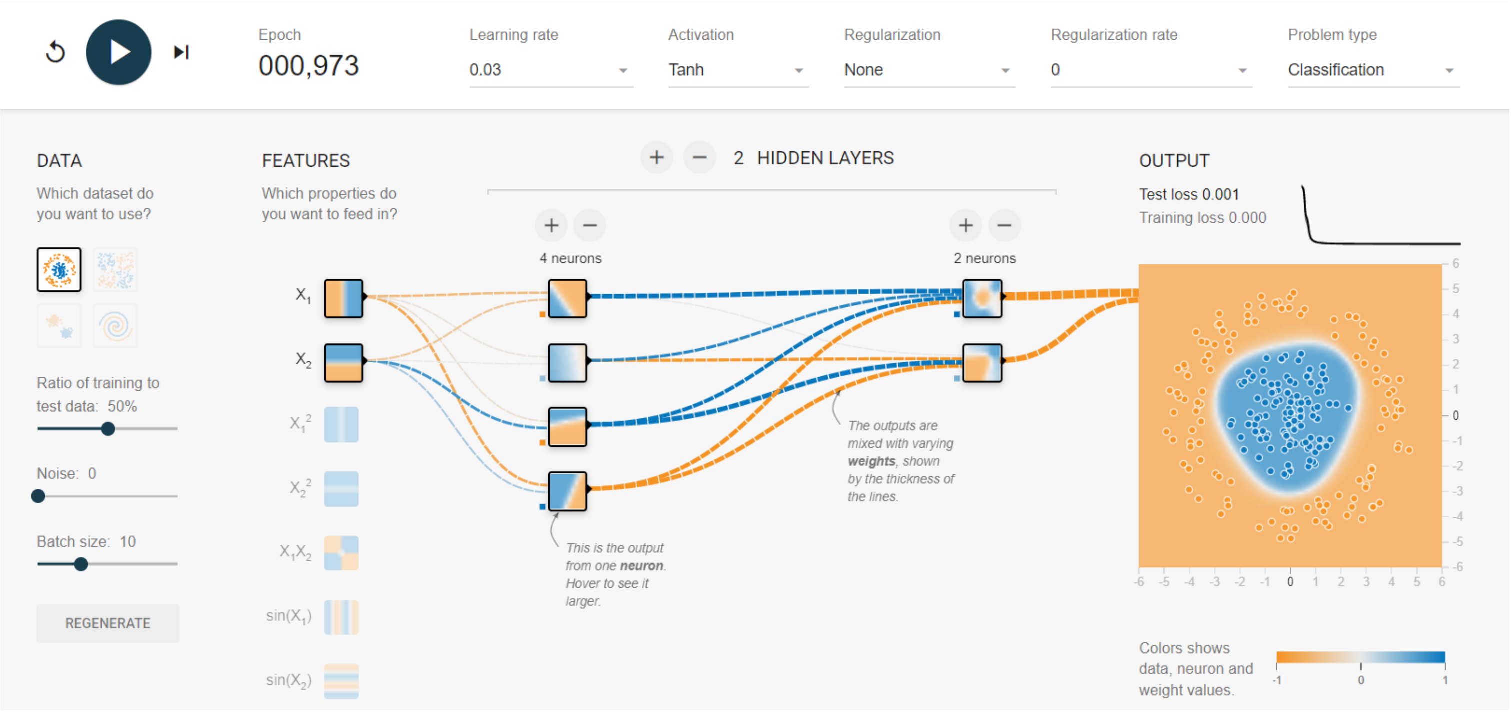
- **Parameters**: Human brains have ~10,000,000 times synapses than artificial neural networks.
- **Topology**: Human brains have no “layers”. Async: The human brain works asynchronously, ANNs work synchronously.
- **Learning algorithm**: ANNs use gradient descent for learning. We don’t know what human brains use
- **Power consumption**: Biological neural networks use very little power compared to artificial networks
- **Stages**: Biological networks usually never stop learning. ANNs first train then test.

Combing Neurons in Hidden Layers: Power to Approximate



Universality: For any arbitrary function *f(x)*, there exists a neural network that closely approximate it for any input *x.*

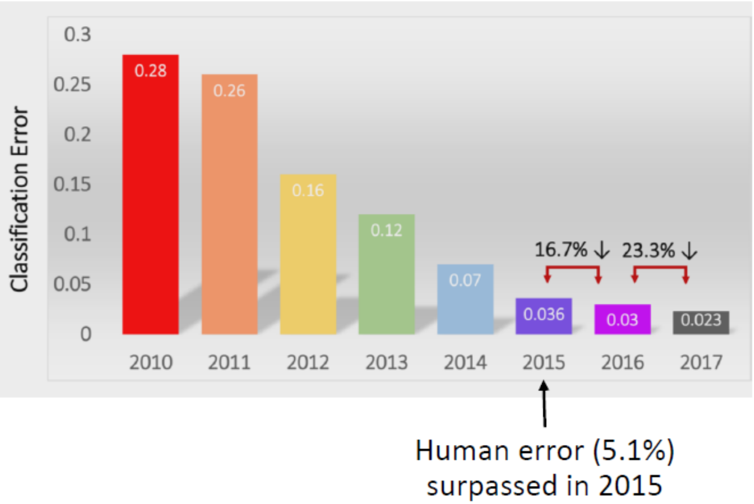
Neural Network Playground



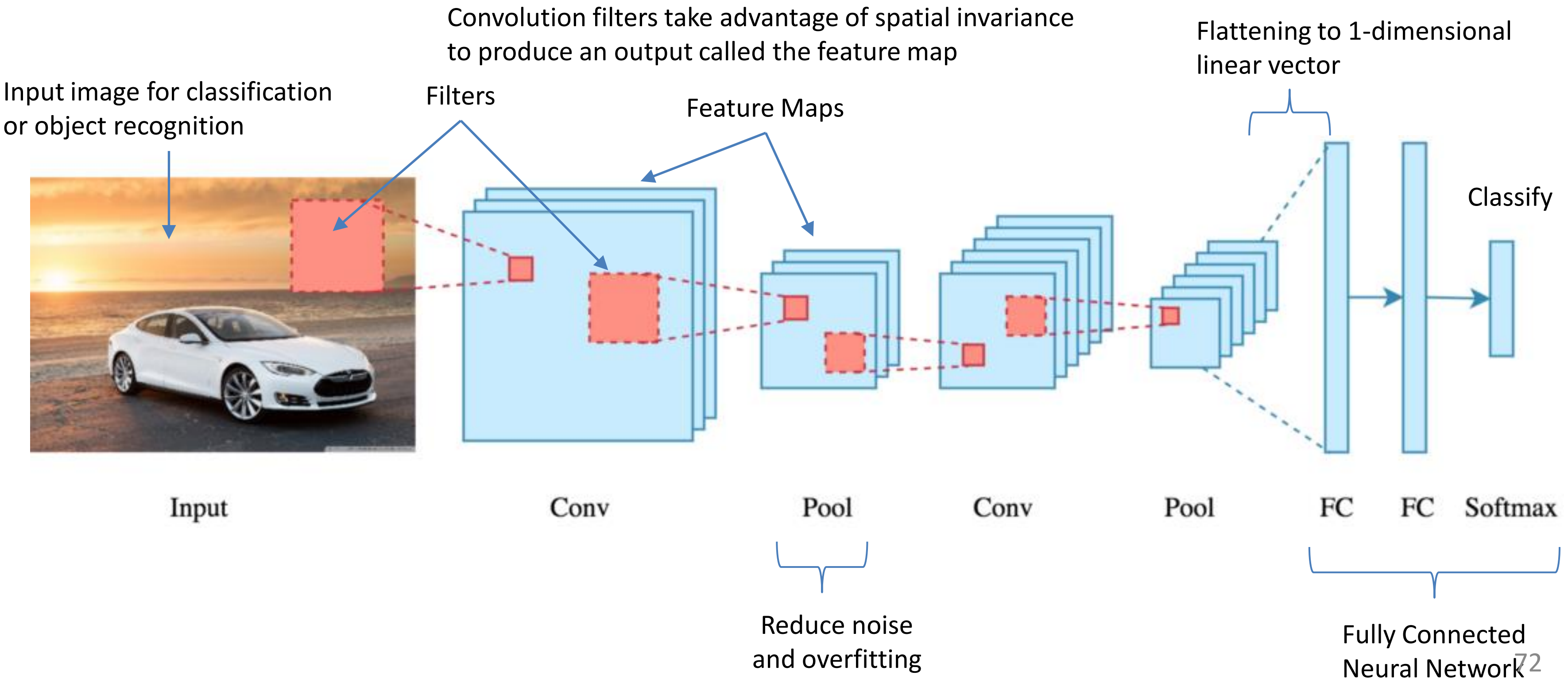
http://playground.tensorflow.org/

Convolutional Neural Networks (CNNs)

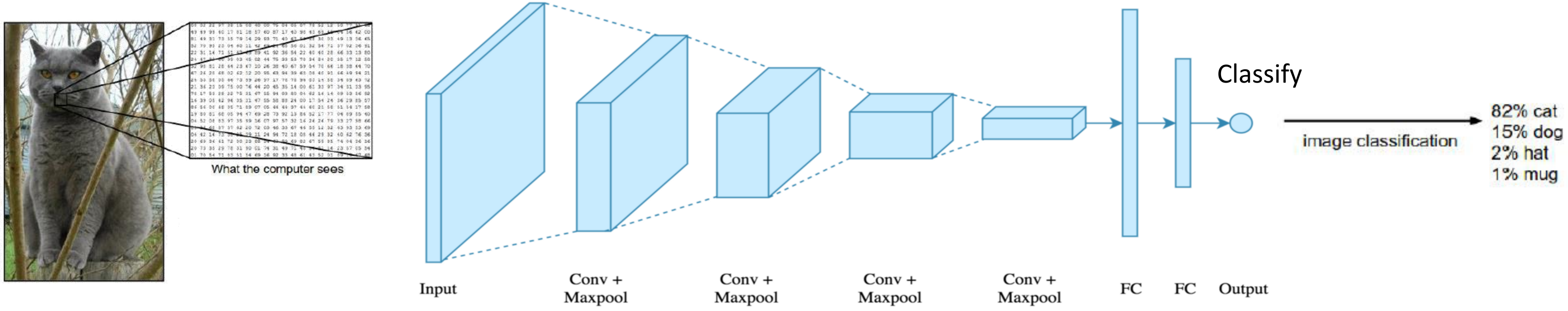
- AlexNet (2012): First CNN (15.4%)
  - 8 layers
  - 61 million parameters
- ZFNet (2013): 15.4% to 11.2%
  - 8 layers
  - More filters. Denser stride.
- VGGNet (2014): 11.2% to 7.3%
  - Beautifully uniform: 3x3 conv, stride 1, pad 1, 2x2 max pool
  - 16 layers
  - 138 million parameters
- GoogLeNet (2014): 11.2% to 6.7%
  - Inception modules
  - 22 layers
  - 5 million parameters (throw away fully connected layers)
- ResNet (2015): 6.7% to 3.57%
  - More layers = better performance
  - 152 layers
- CUImage (2016): 3.57% to 2.99%
  - Ensemble of 6 models
- SENet (2017): 2.99% to 2.251%
  - Squeeze and excitation block: network is allowed to adaptively adjust the weighting of each feature map in the convolutional block.



Convolutional Neural Network Architecture



CNN Implementation



Convolution filters slides through the image

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Stride 1

Feature Map

1	0	1
0	1	0
1	0	1

Filter

1x1	1x0	1x1	0	0
0x0	1x1	1x0	1	0
0x1	0x0	1x1	1	1
0	0	1	1	0
0	1	1	0	0

The convolution process

4		

Feature map

Input

4	3	4
2	4	3
2	3	4

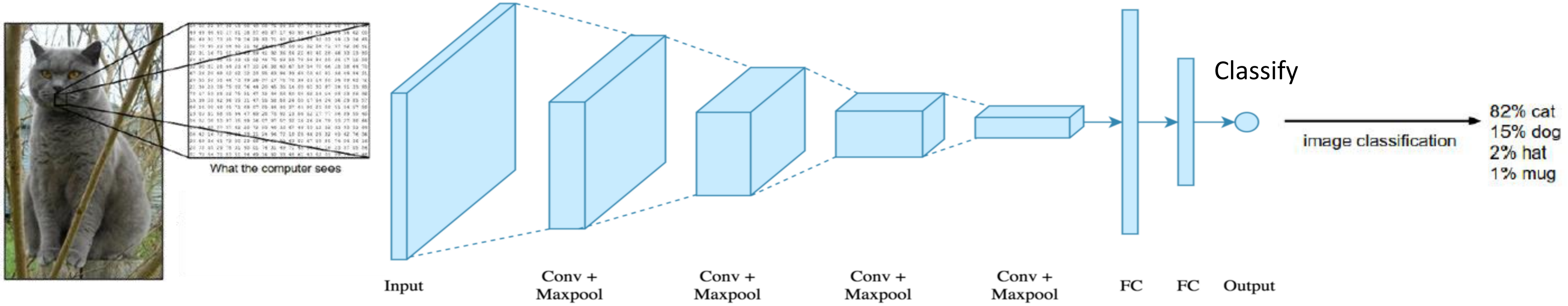
2 x 2 Max Pooling

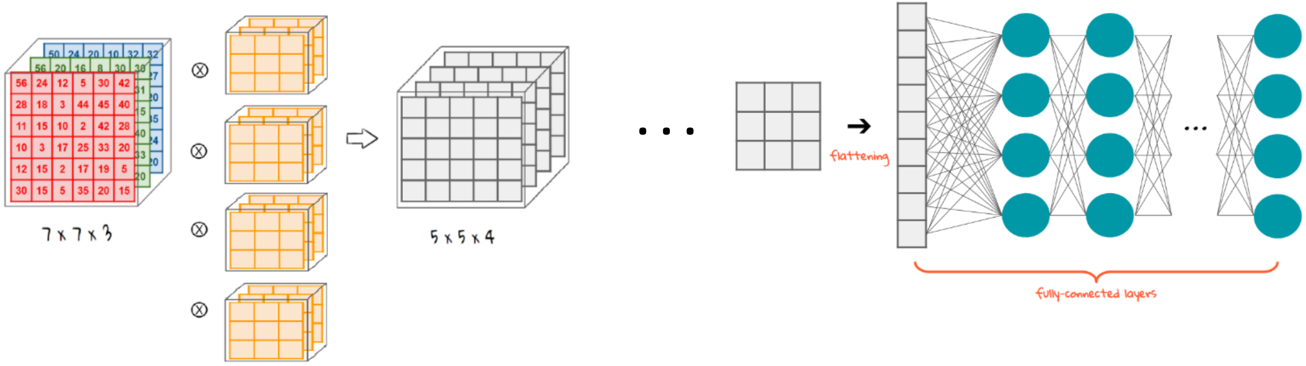
Output

4	4
4	4

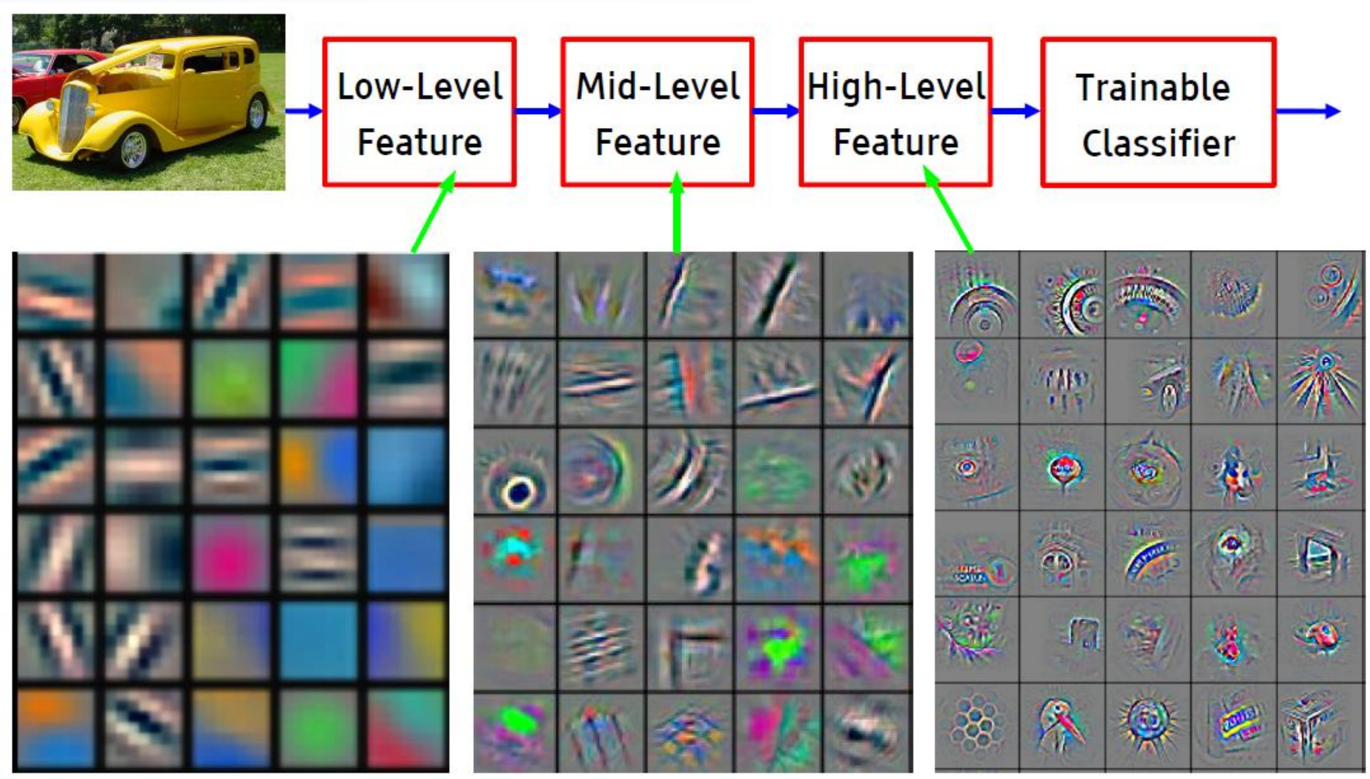
The pooling process

CNN Implementation (Cont'd)





Deep Learning = Learning Hierarchical Representations



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

Trainable Feature Hierarchy

- Hierarchy of representations with increasing level of abstraction
- Each stage is a kind of trainable feature transform
- **Image recognition**
  Pixel→ edge→ texton→ motif→ part→ object
- **Text**
  Character→ word→ word group→ clause→ sentence→ story
- **Speech**
  Sample→ spectral band→ sound → …→ phone→ phoneme → word →

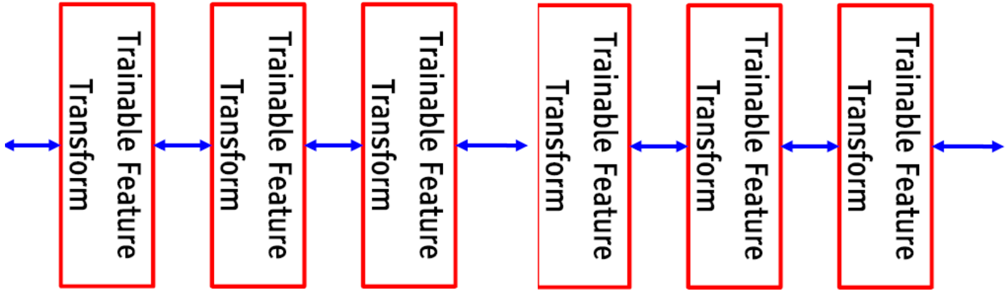
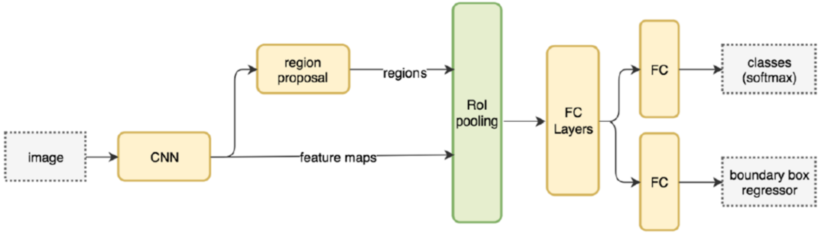


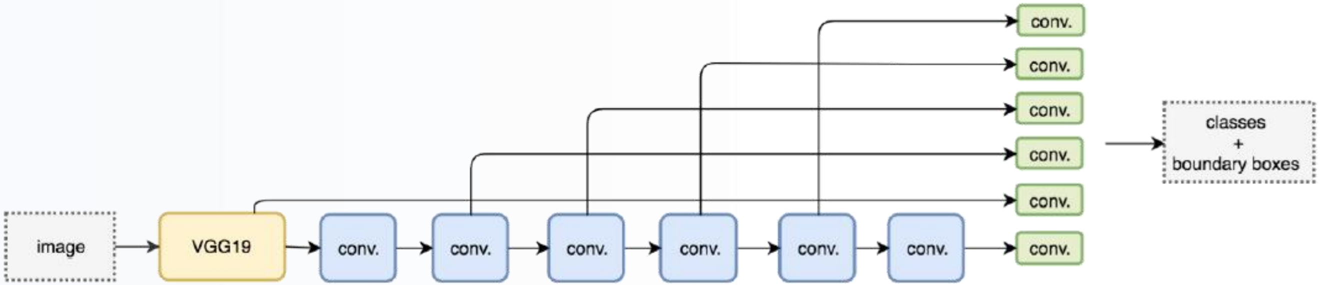
Image Detection / Localization with Region-Based Methods (Faster R-CNN)

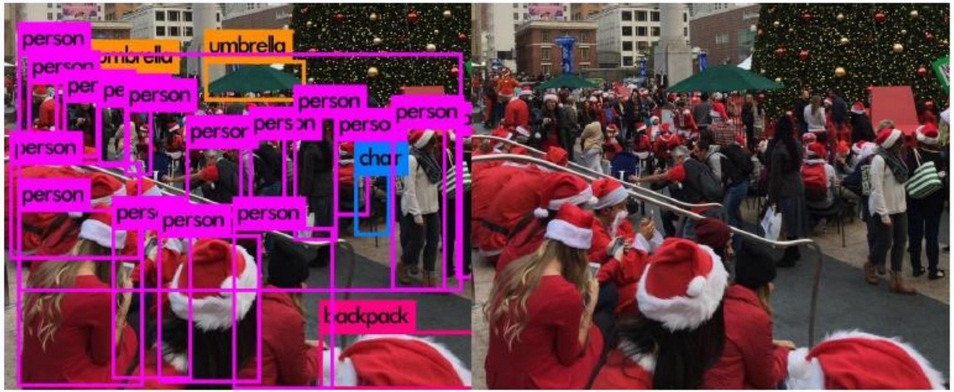
```
ROIs = region_proposal(image)
for ROI in ROIs
    patch = get_patch(image, ROI)
    results = detector(patch)
```



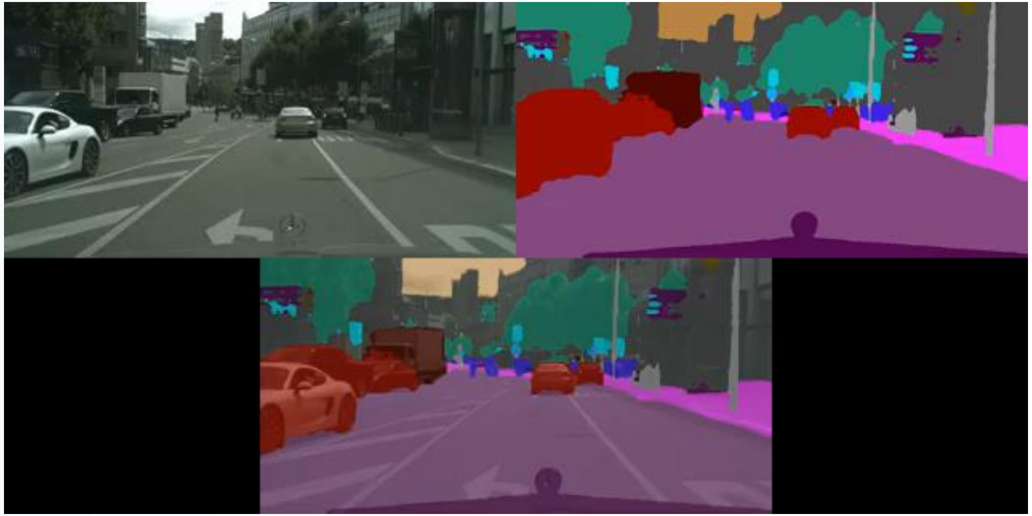
https://www.youtube.com/watch?v=zrHfvVdvmz4&t=1s

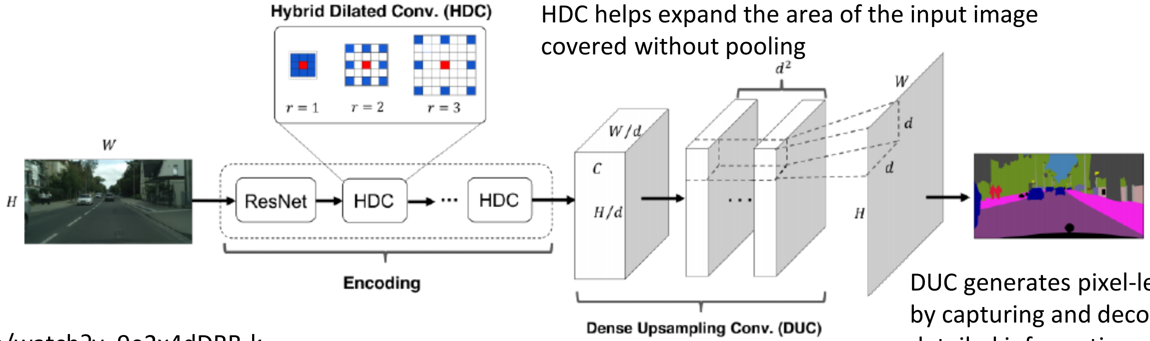
Image Detection / Localization with Single Shot Detection Methods (SSD)





Semantic Segmentation



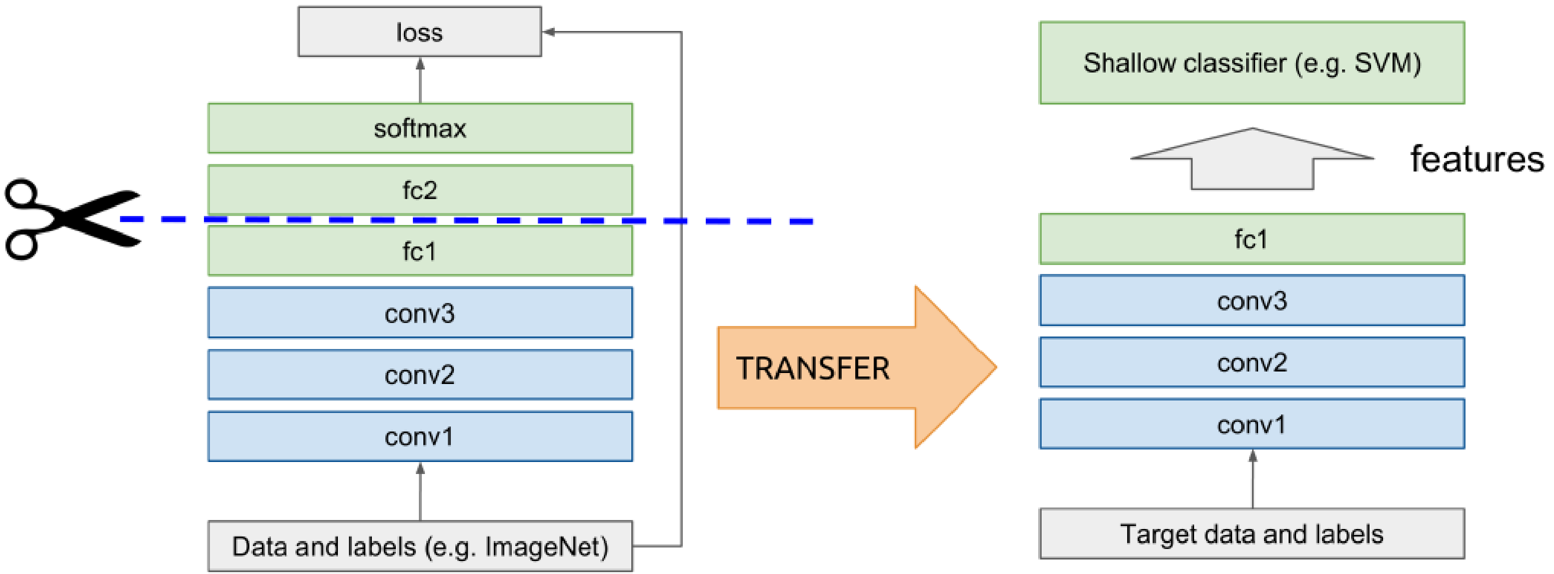


HDC helps expand the area of the input image covered without pooling

DUC generates pixel-level prediction by capturing and decoding more detailed information

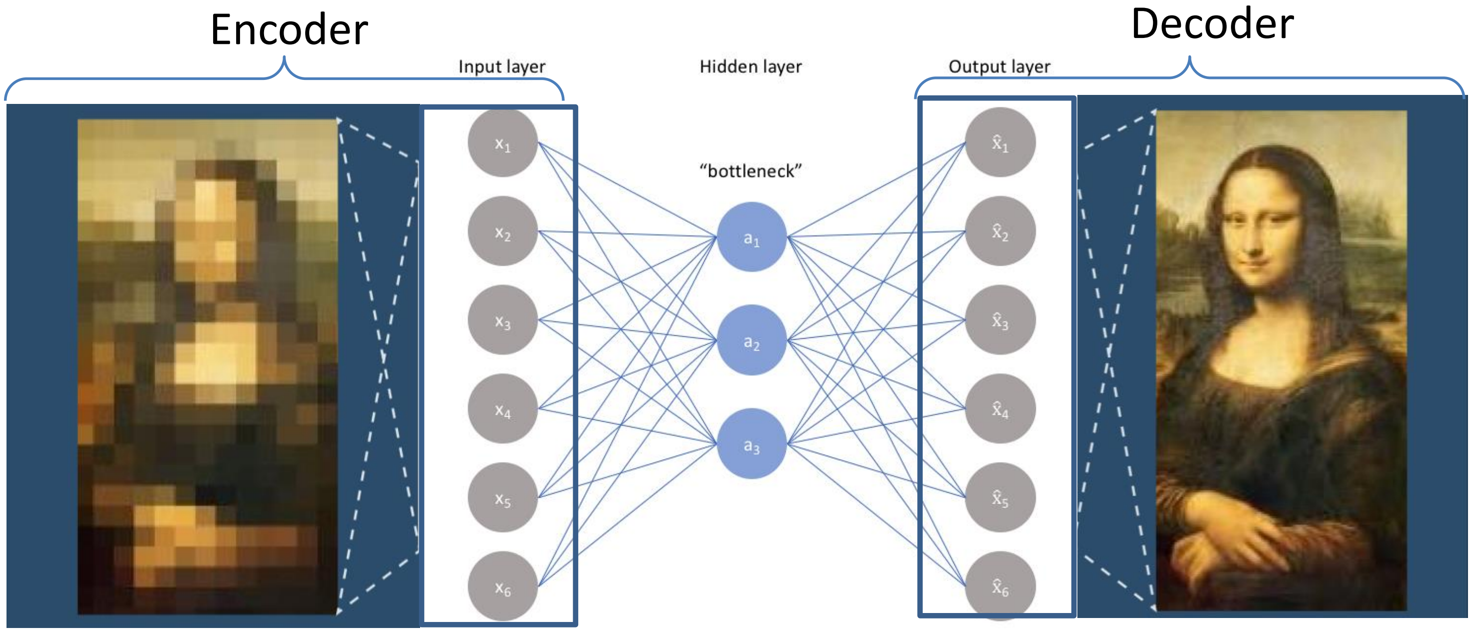
https://www.youtube.com/watch?v=9e2x4dDRB-k

Transfer Learning



- Fine-tune a pre-trained model
- Effective in many applications: computer vision, audio, speech, NLP

Autoencoders

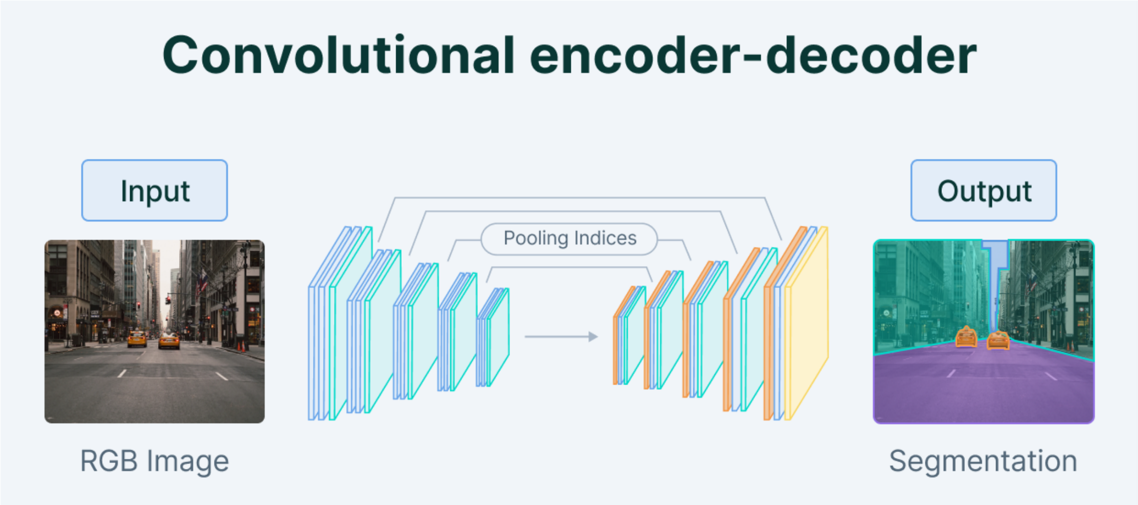


Encoder is a set of convolutional blocks followed by pooling modules that compress the input to the model

The ''bottleneck' restricts the flow of information to the decoder from the encoder, allowing only the most vital information to pass through.

Decoder upsampling the convolutional blocks and reconstructs the data back from its encoded form from the bottleneck's output.

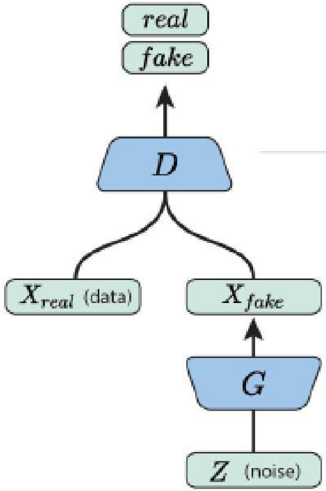
Convolutional Encoder-Decoder



The aim of an autoencoder is to learn a lower-dimensional representation (encoding) for a higher-dimensional data, typically for dimensionality reduction, by training the network to capture the most important parts of the input image.

Generative Adversarial Networks (GANs)

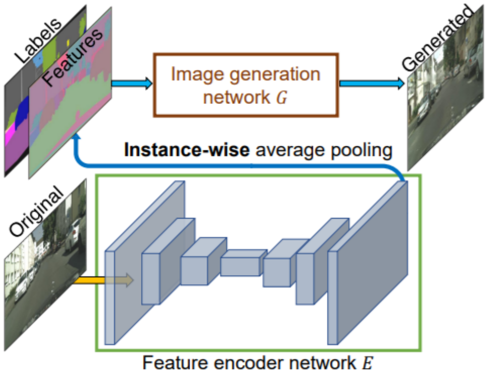
Generative Adversarial Networks (GANs) are a way to make a generative model by having two neural networks compete with each other.



The **discriminator** tries to distinguish genuine data from forgeries created by the generator.

The **generator** turns random noise into immitations of the data, in an attempt to fool the discriminator.

Generative Adversarial Networks (GANs) (Cont'd)



T-C Wang et al. "High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs" in CVPR, 2018.

Applications of Generative Adversarial Networks (GANs)

Generate Photographs of Human Faces

T. Karras, et al., "Progressive Growing of GANs for Improved Quality, Stability, and Variation", 2017.

Generate Realistic Photographs

A. Brock, et al., "Large Scale GAN Training for High Fidelity Natural Image Synthesis", 2018.

Generate Consistent Video

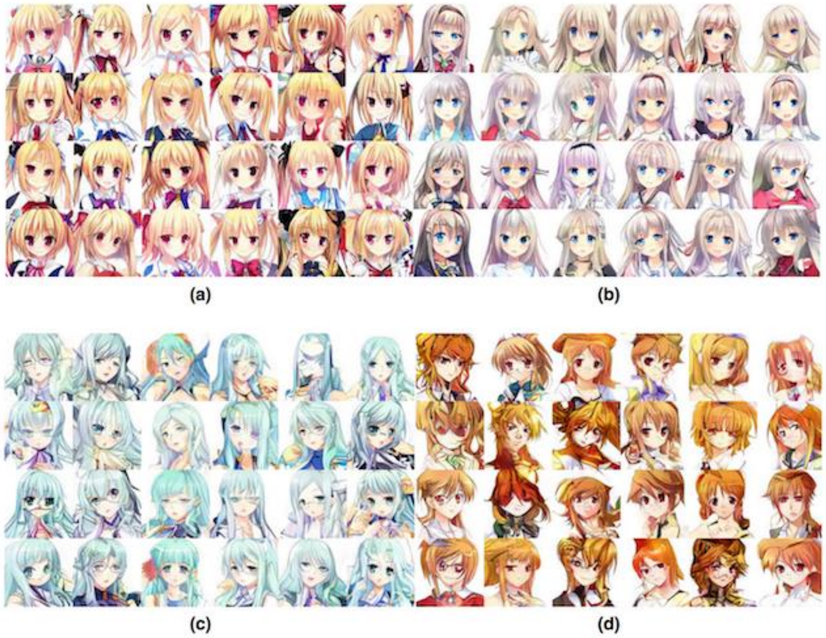
https://www.youtube.com/watch?v=9reHvktowLY

Translation of sketches to color photographs

Phillip Isola, et al., "Image-to-Image Translation with Conditional Adversarial Networks", 2016.

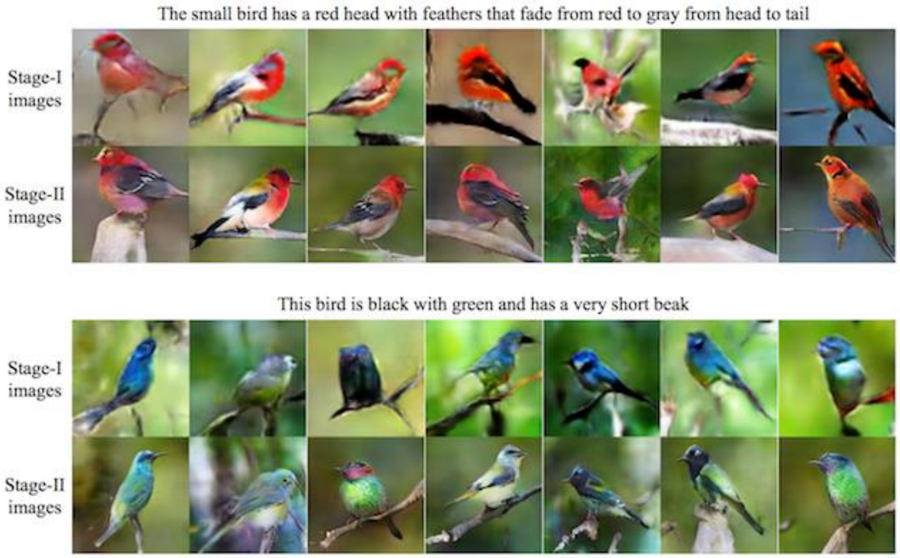
Applications of Generative Adversarial Networks (GANs)

Generate Cartoon Characters



Y-H Jin, et al., “Towards the Automatic Anime Characters Creation with Generative Adversarial Networks” , 2017.

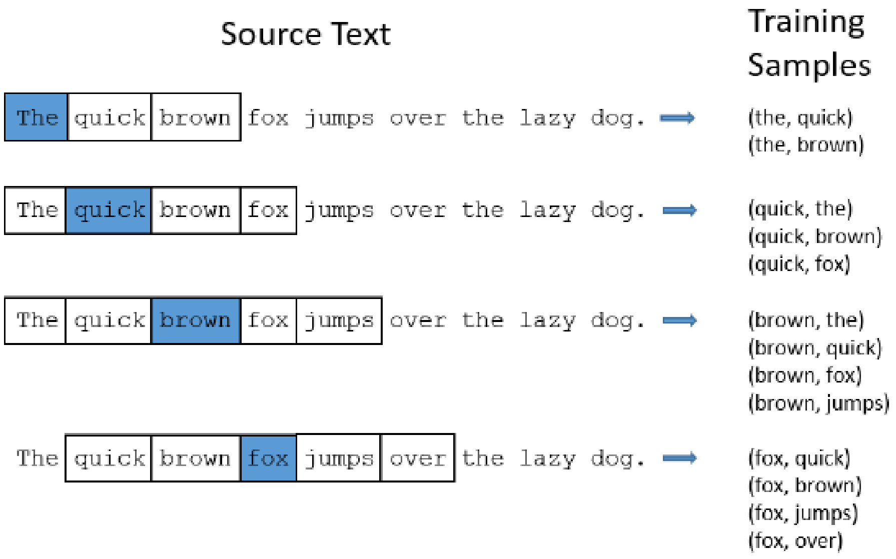
Text-to-Image Translation



Han Zhang, et al., “StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks”, 2016.

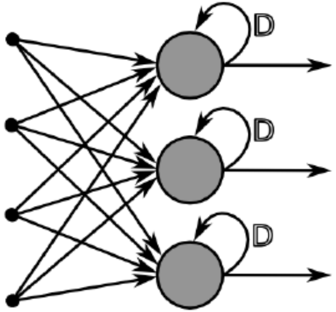
Word Embeddings (Word2Vec)

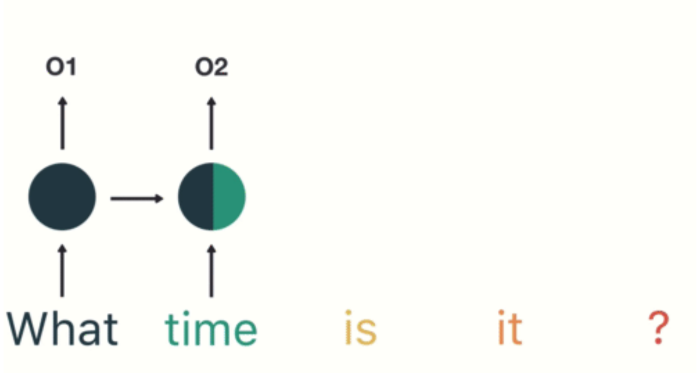
- Word embedding is capturing context of a word in a document, semantic and syntactic similarity, relation with other words
- Word2Vec is a method to construct an embedding. Another way to put it is, they are vector representations of a particular word.
- It is used to predict the source context words (surrounding words) given a target word (the center word).
- E.g. to predict the context [quick, fox] given target word ‘brown’ or [the, brown] given target word ‘quick

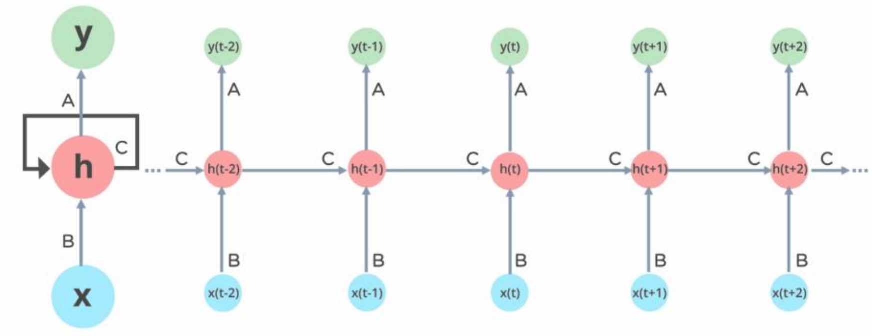


Skip Gram Model

Recurrent Neural Networks (RNNs)







Applications

- Sequence Data
- Text
- Speech
- Audio
- Video
- Generation

Long Term Dependency

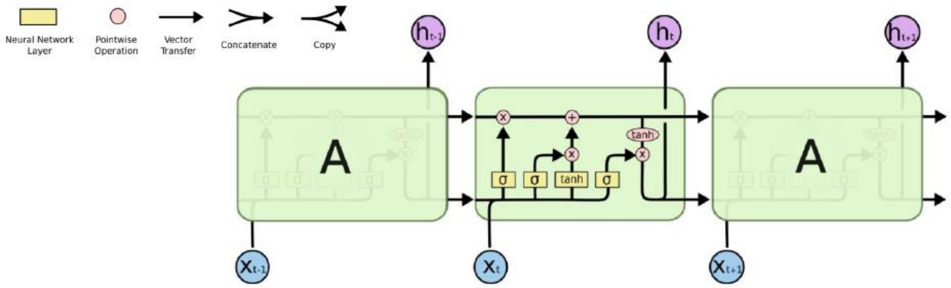
Short-term dependence:

Bob is eating an apple

Context

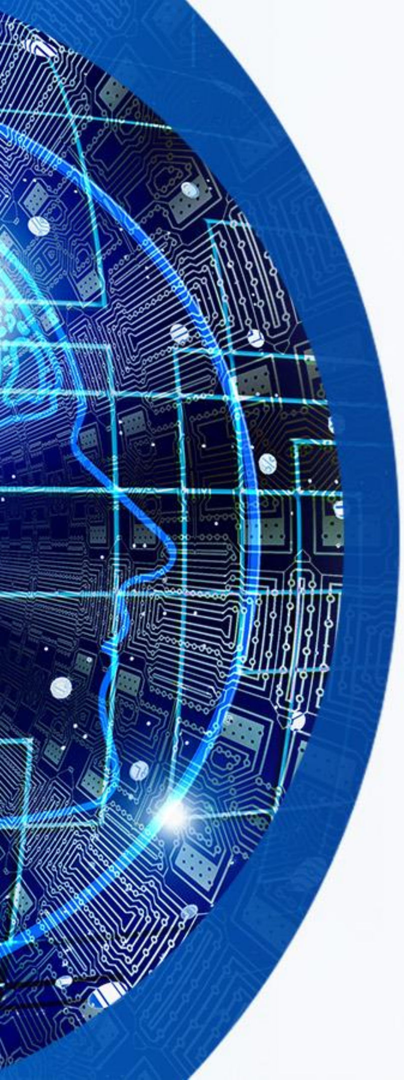
Long-term dependence:

Bob likes apples. He is hungry and decided to have a snack. So now he is eating an apple.

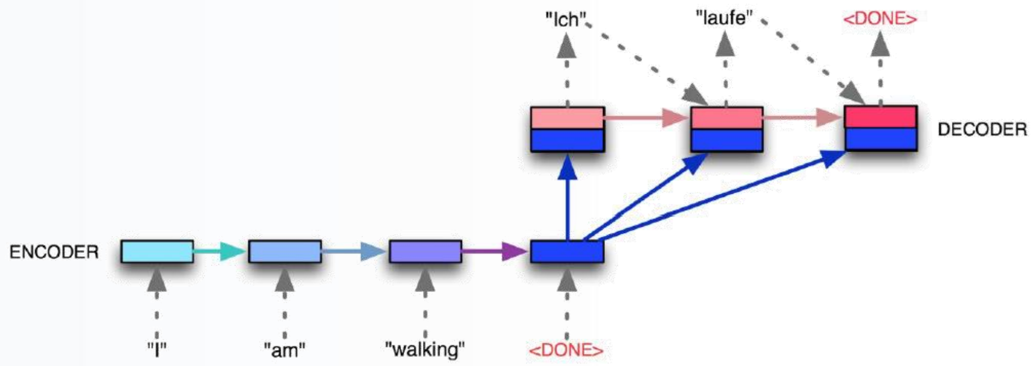


Long Short-Term Memory (LSTM) Networks:

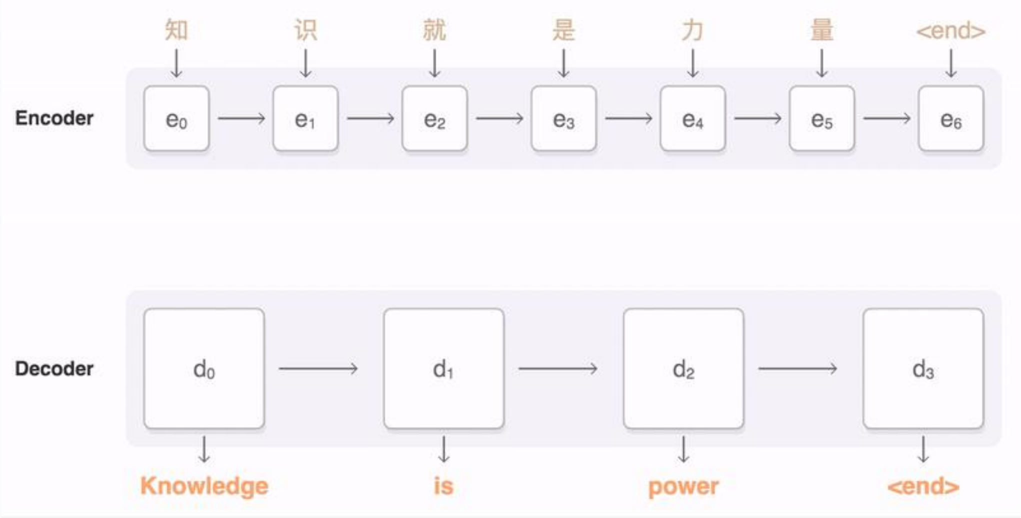
Pick What to Forget and What To Remember



Encoder-Decoder Architecture

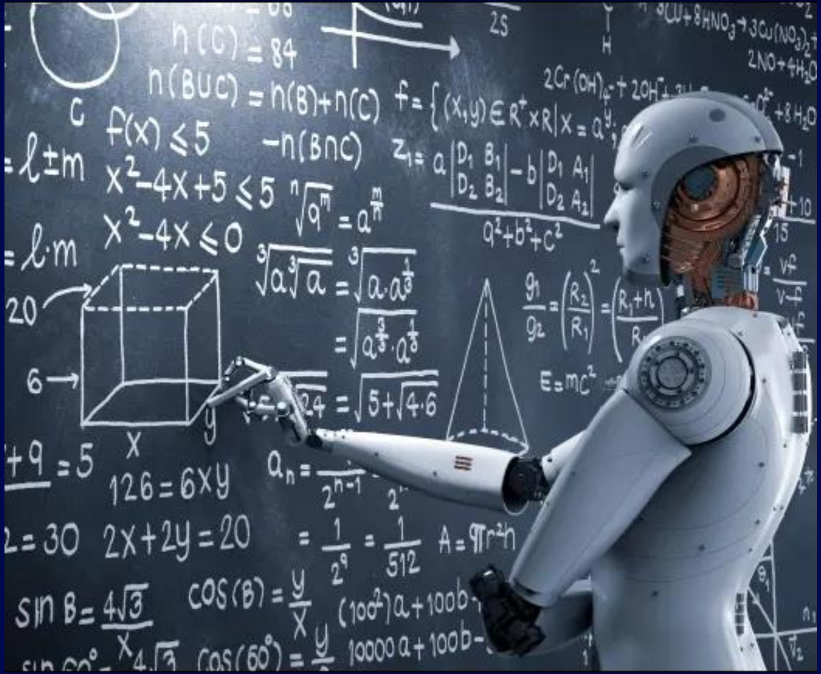


Machine Language Translation



Language Model for Dialog Application (LaMDA)

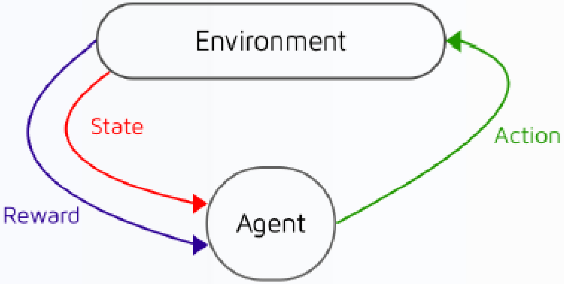
- LaMDA is Google’s most advanced language model to date
- Learning method, called Generative Pre-trained Transformer -3 or GPT-3 which allow computers to discover patterns and relationships in data
- Designed to generate natural dialogue with specific ability to engage in human-like open-ended dialogue with users

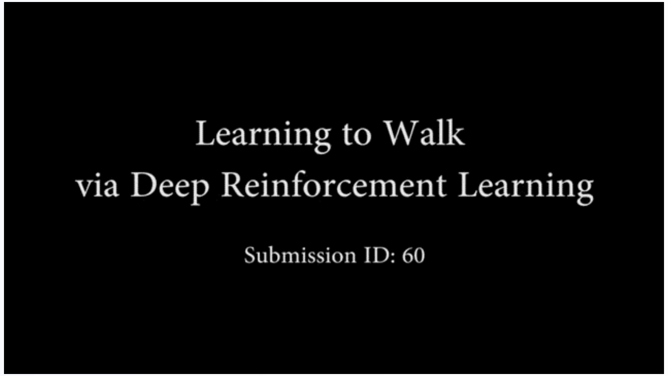


What is GPT-3

- A neural network machine learning model trained using internet data to generate any type of text
- GPT-3's deep learning neural network is a model with over 175 billion machine learning parameters
- Developed by OpenAI, it requires a small amount of input text to generate large volumes of relevant and sophisticated machine-generated text
- (Microsoft's Turing NLG model, which had 10 billion parameters)

Deep Reinforcement Learning

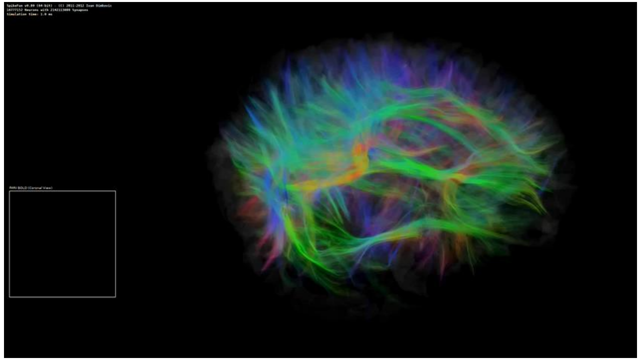


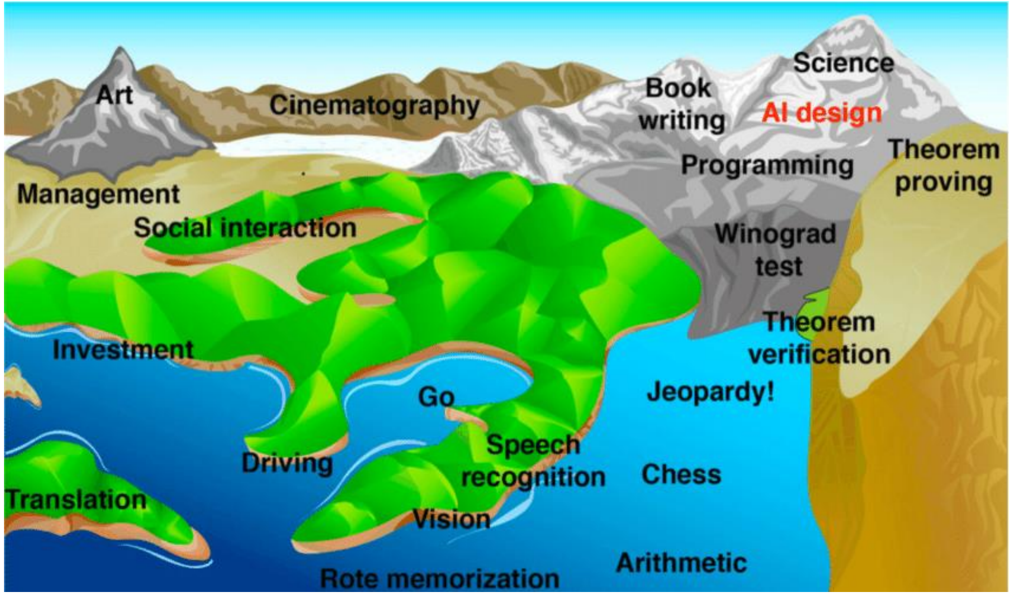


https://www.youtube.com/watch?v=n2gE7n11h1Y

https://www.facebook.com/watch/?v=344186809644234

Towards General AI





Hans Moravec's illustration of the rising tide of the AI capacity increasingly covering the landscape of human competence

Summary

- AI technology is already used in many fields, including medicine, the automotive industry, manufacturing, retail, social media, banking and finance sectors to mention just a few
- We can expect to see more advancement in AI in the years to come
- Concerns related to safety issues regarding AI
- The big difference between AI and machine learning is that the latter is a subcategory of, or a specific method to achieve, the former
- The big difference between machine learning and deep learning is that the latter is a subcategory of, or a specific method to achieve, the former
- Machine learning depends on algorithms or a set of rules or sequence of instructions guiding an operation. The algorithm in ML is learned by searching for patterns in a training dataset that match the labels to the given input values.
- Deep learning is a is a subset of machine learning that uses artificial neural networks inspired by the human brain, to learn from large amounts of data to solve any pattern recognition problem without human intervention.

Acronym References

- CNN Convolutional neural networks
- RBM Restricted Boltzmann machine
- RNN Recurrent neural network
- MNIST Modified National Institute of Standards and Technology
- LSTM Long short-term memory
- DBN Deep belief network
- GAN Generative adversarial network
- BERT Bidirectional Encoder Representations from Transformers
- CUDA Compute Unified Device Architecture
- CNTK Microsoft Cognitive Toolkit
- RCNN Region-based CNN
- HDC Hybrid Dilated Convolution
- DUC Dense Upscaling Convolution

Quiz III

- Deep learning is part of a broader family of ____________________ methods based on ____________________ with ____________________ that can be ____________________.

A. Machine Learning

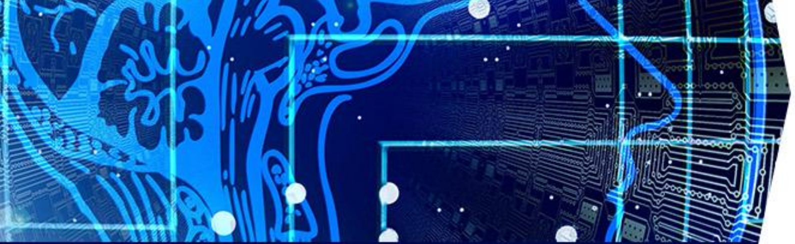
B. Supervised or Unsupervised

C. Representation Learning

D. Artificial Neural Networks

Building an Object Recognition Model

- Refer to activity lab 1



End of Chapter 1

Q&A