Introduction to Artificial Intelligence and Machine Learning

School of Engineering
Nanyang Polytechnic

Trainer

Foo Yong Wee, Ph.D.

Manager, Computational Intelligence and Analytics Group

School of Engineering

Nanyang Polytechnic

Email: foo_yong_wee@nyp.edu.sg

Tel: 6550 0962

Introduction to Artificial Intelligence and Machine Learning

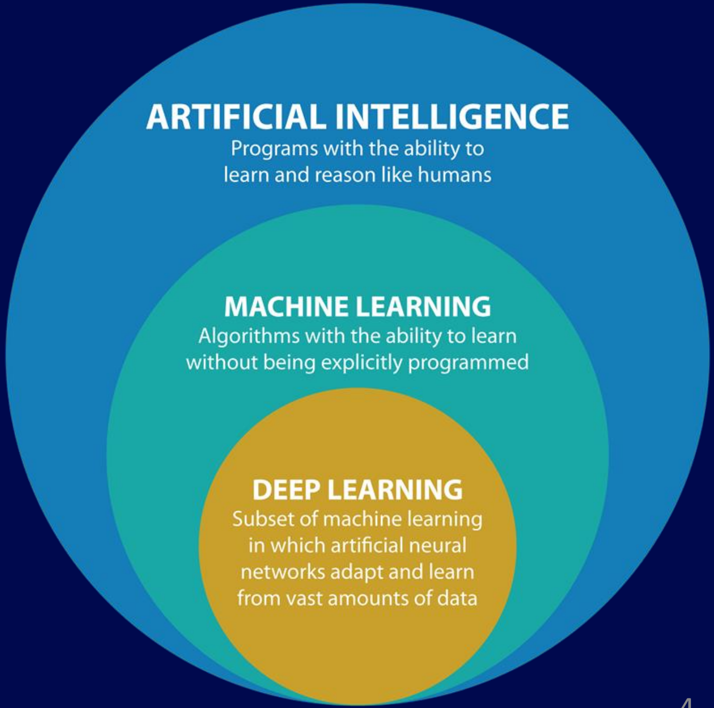
Outline

1 What is Artificial Intelligence (AI)

2 What is Machine Learning (ML)

3 What is Deep Learning (DL)

4 Applications of AI and ML



1 What is AI

Definition of AI

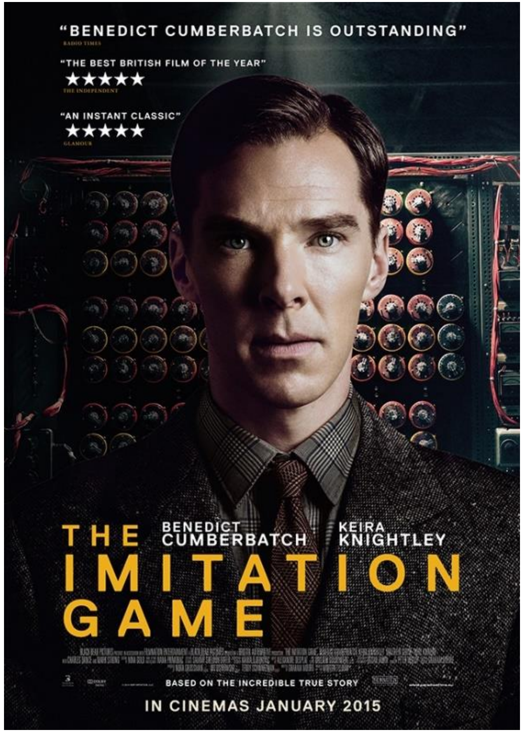
Artificial Intelligence is the ability of a program or machine to think and behave like a human.

Turing Test

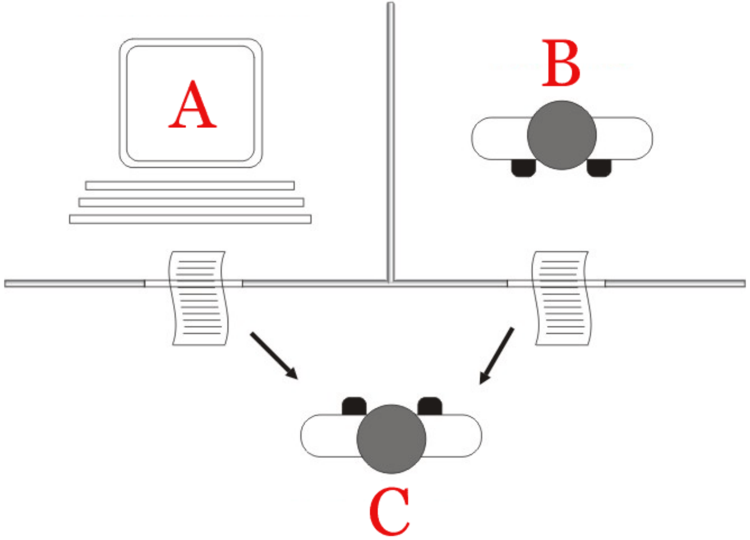
A machine is said to pass the Turing Test when it exhibits behaviour that is indistinguishable from that of a human.

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

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

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

Watson won Jeopardy game

Deep Blue beat world chess champion

Google BERT

AlphaGo won Go champion

ChatGPT

Generative AI

1943	1956	The golden years	1974	First AI winter	1980	AI boom	1987	Second AI winter	1993	1997	2011	AI renaissance	2016	2020	AI revolution ……	2023 …

1943

Artificial neurons (McCullouch and Pitts)

1956

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*

1980

1993

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

2011

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

AI Resurgence

Computing performance
& Cloud

Big data

IOT

Internet technologies
& Social Media

+

AI / Machine Learning

Advanced data analytics

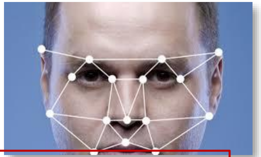
Deep learning

Deep neural networks

BIG DATA, HIGH-PERFORMANCE COMPUTING AND CLOUD, HAS TRANSFORMED AI

AI is All Around Us

Image search



Facial recognition

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

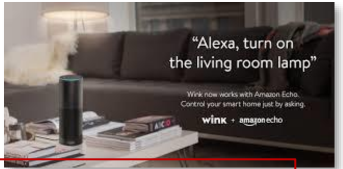
Google BERT

ChatGPT

Drones, UAV

Smart IoT devices

Self-driving vehicles



Virtual smart assistants

Generative AI

Recommendation engines

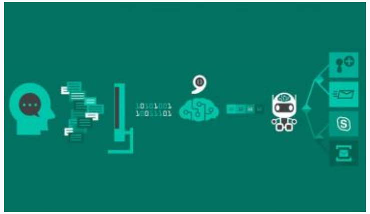
Spam filters

Chatbots

Robo-advisors

Language translation

AI Technologies



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

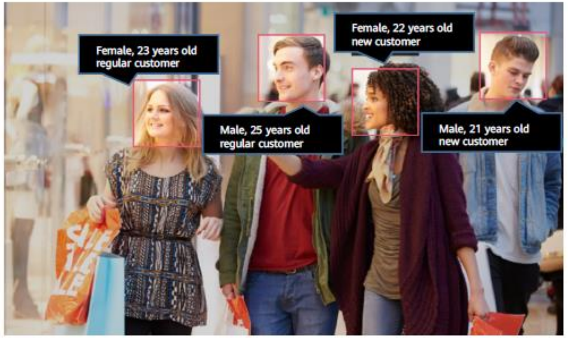
Speech processing: a general term for various processing technologies used to research the voicing process, statistical features of speech signals, speech recognition, machine-based speech synthesis, and speech perception

Natural language processing (NLP): is the use computer technologies to understand and process human language for a range of applications, including machine translation, sentiment analysis, and text analytics.

Generative AI generate new, unique data samples that are similar to the examples it was trained on such as images.

Language Models a deep learning Generative model that can understand the context and nuances of natural language text.

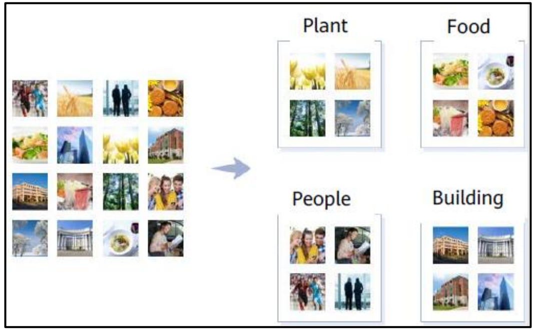
Computer Vision



Traffic Analysis

Electronic Attendance

Common applications include image classification, target detection, image segmentation, target tracking, optical character recognition (OCR), and facial recognition.



Smart Album

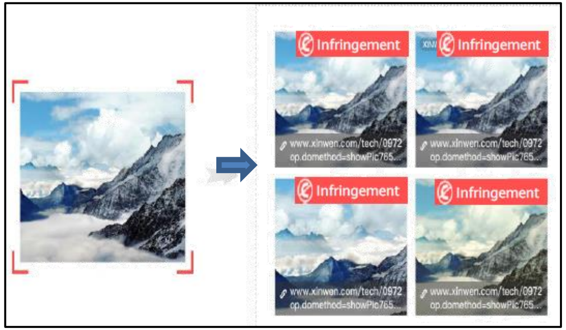
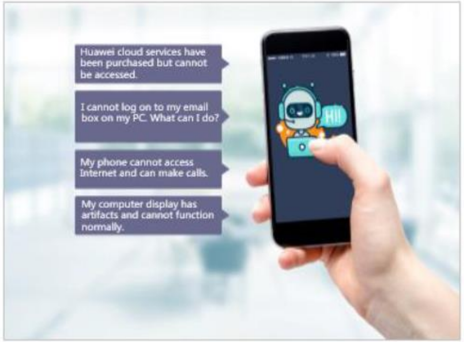


Image Search

Authentication

Action Analysis

Voice/Speech Processing



Question Answering Bot (QABot)

Voice Navigation

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

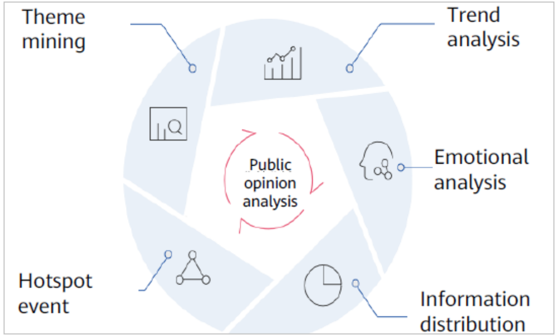
Intelligent Education

Other applications:

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

Real-time Conference Records

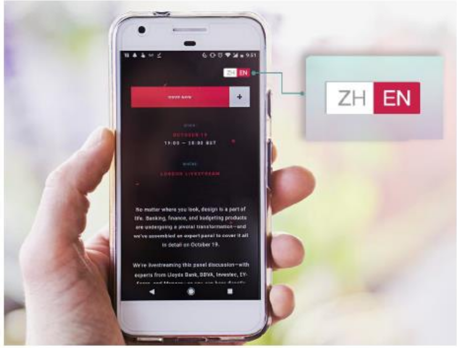
Natural Language Processing



Public Opinion Analysis


Evaluation Analysis

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



Machine Translation

Other applications:

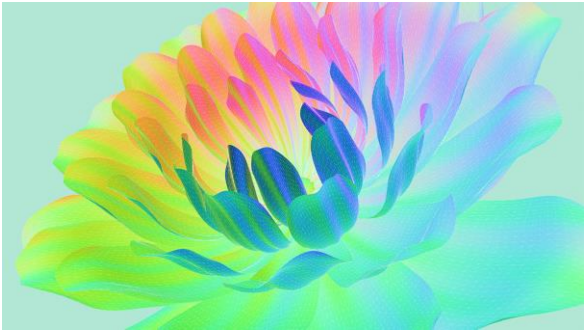
- Knowledge graph
- Intelligent copywriting
- Video subtitle
- ...

Text Classification

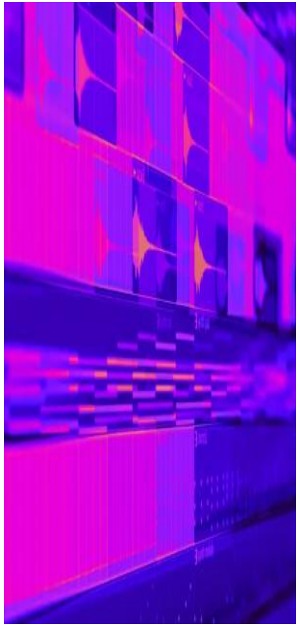
Generative AI

Video Generation
Google's Imagen video (https://imagen.research.google/video/) can produce short high-definition video clips using a similar technique to the one it uses to make images. (Photo by Google AI)

Data Augmentation

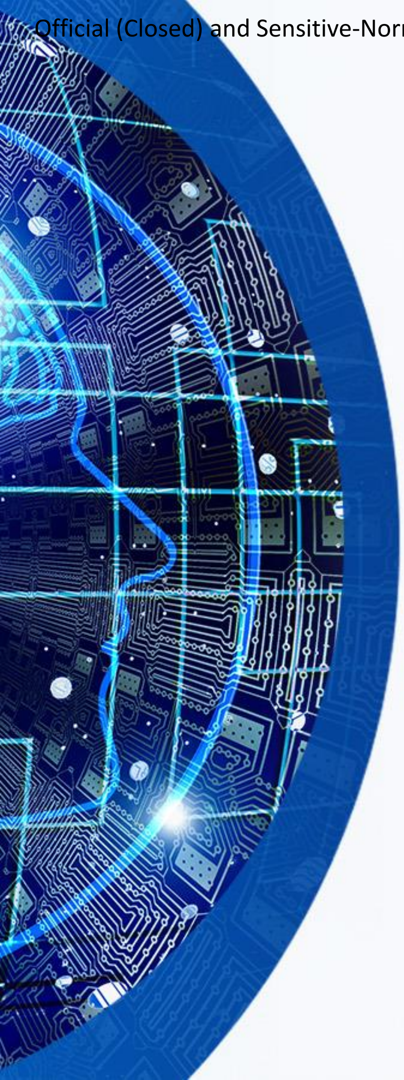


Art Generation

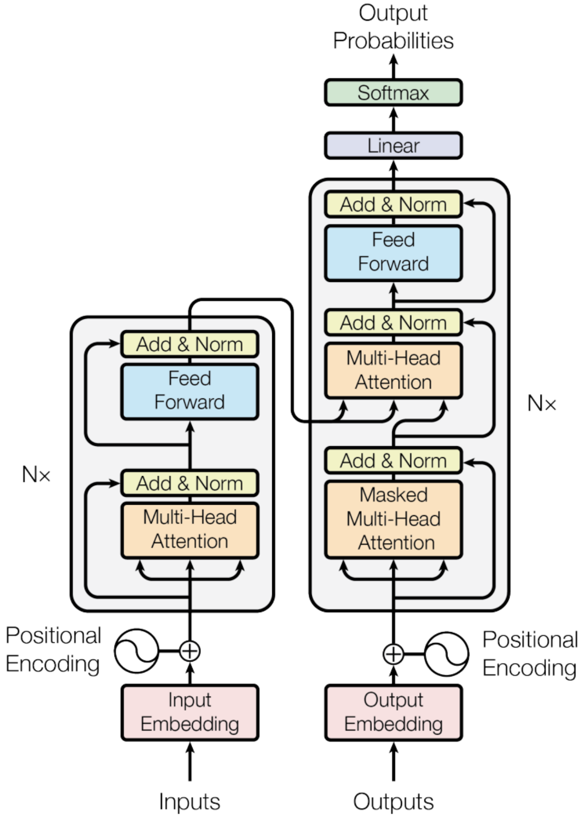


AI-Generated Music

Generative AI and Large Language Models (LLMs)

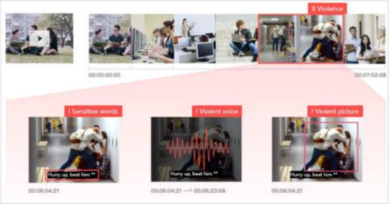


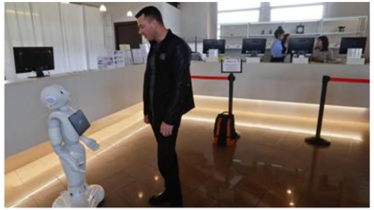
- Text Generation
- Summarization
- Coding
- Image Generation
- Translation
- Chatbots

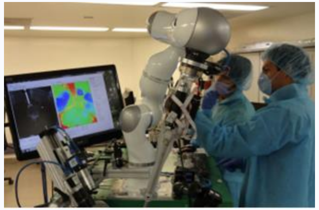


Question

- How can AI technology revolutionize different industries, and what are some specific examples of its applications in these industries?



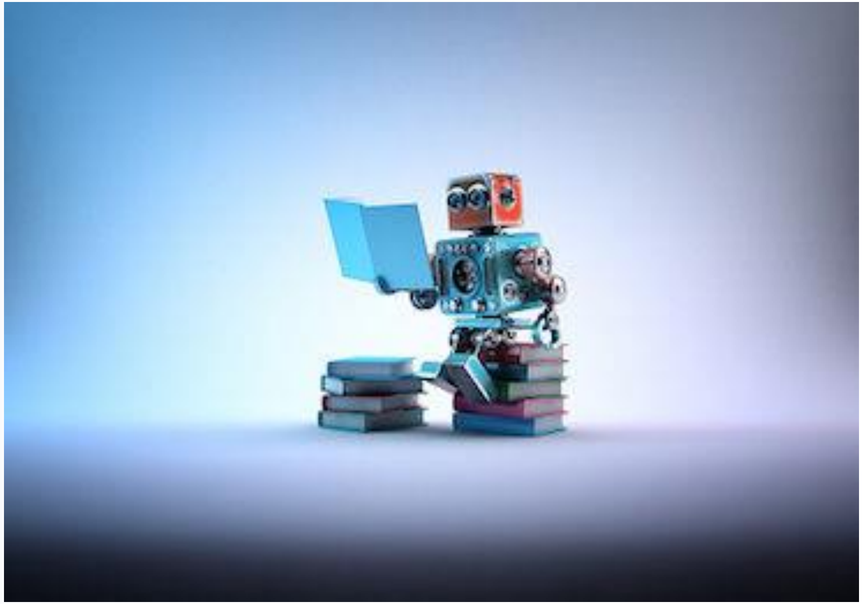




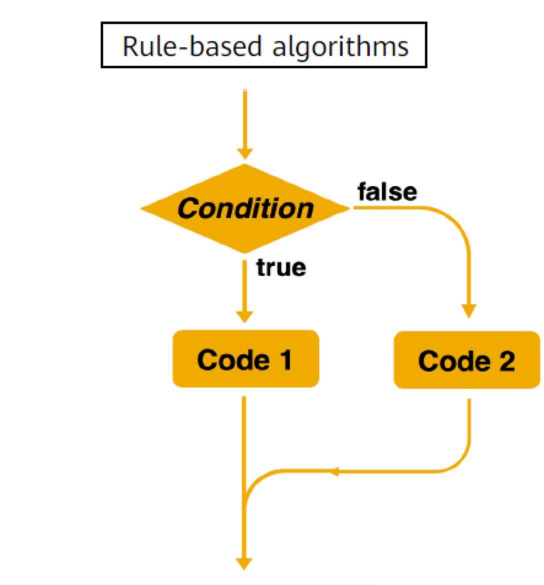
End of Chapter 1

2 What is Machine Learning

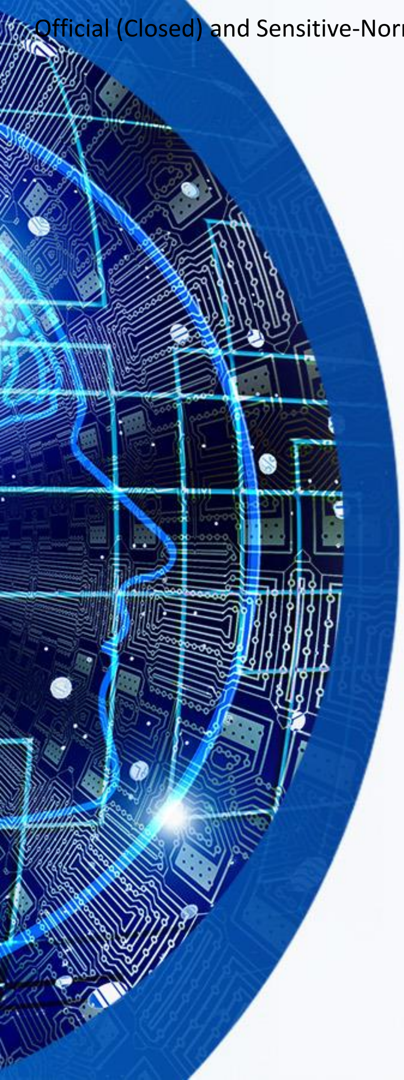
Can a machine learn?



Classical Rule-Based Algorithms



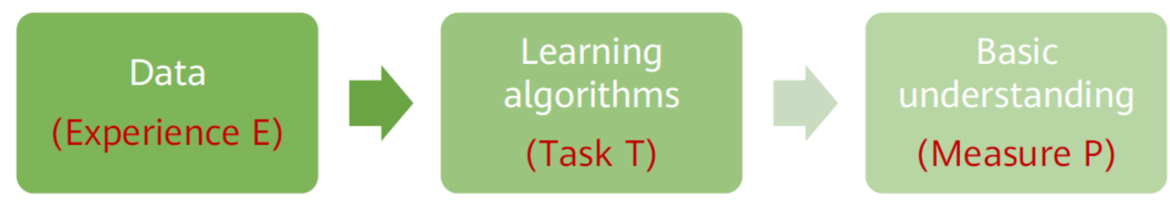
- Explicit programming is used to solve problems
- Rules can be manually specified

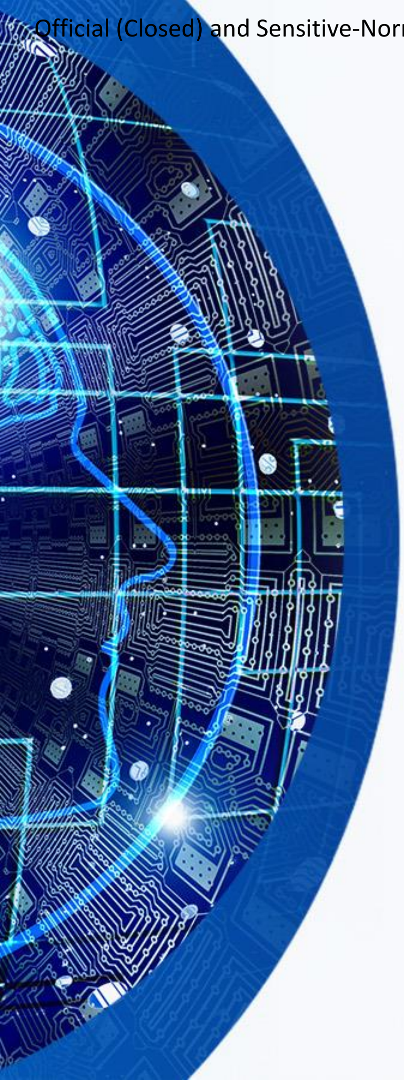


Definition of learning

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.”

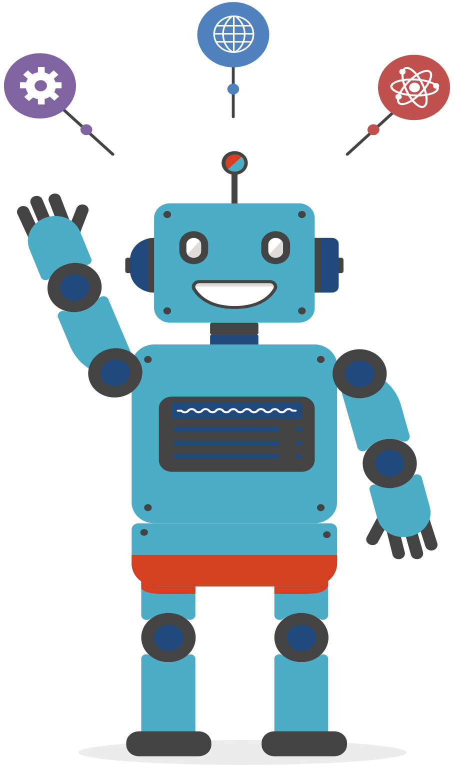
– Tom Michell

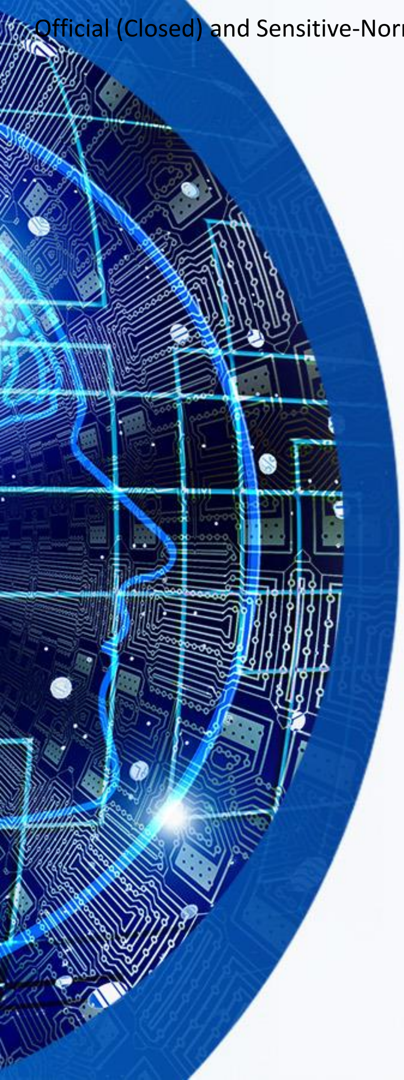




Machine Learning

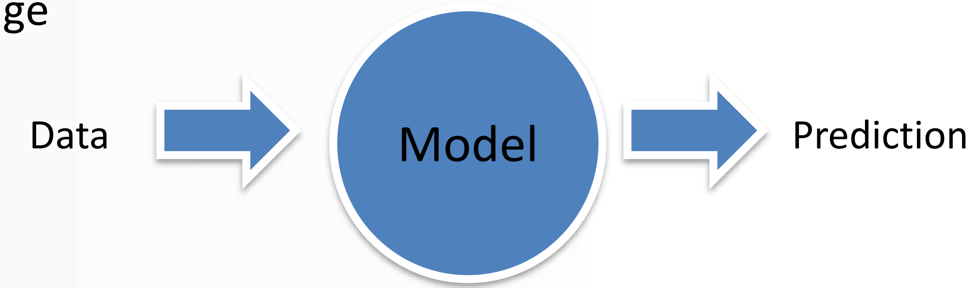
Machine learning allows computers to learn and infer from data

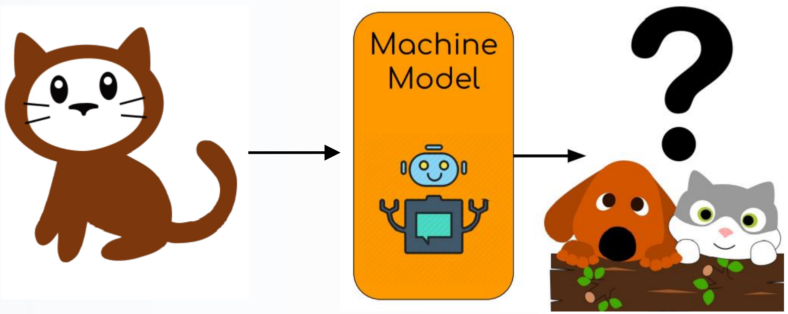




Machine Learning Model

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





Machine Learning Vocabulary

Feature

Target

Sepal length	Sepal width	Petal length	Petal width	Species
6.7	3.0	5.2	2.3	Virginica
6.4	2.8	5.6	2.1	Virginica
4.6	3.4	1.4	0.3	Setosa
6.9	3.1	4.9	1.5	Versicolor
4.4	2.9	1.4	0.2	Setosa
4.8	3.0	1.4	0.1	Setosa
5.9	3.0	5.1	1.8	Virginica
5.4	3.9	1.3	0.4	Setosa
4.9	3.0	1.4	0.2	Setosa
5.4	3.4	1.7	0.2	Setosa

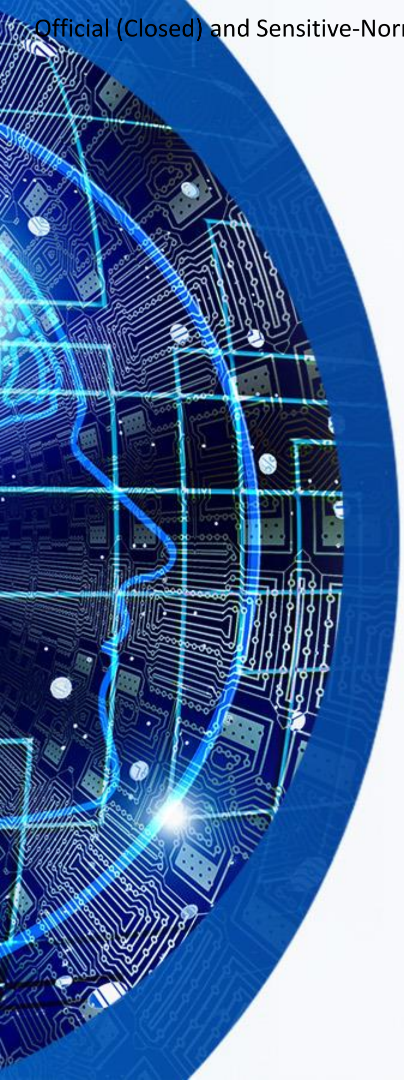
Label

Example

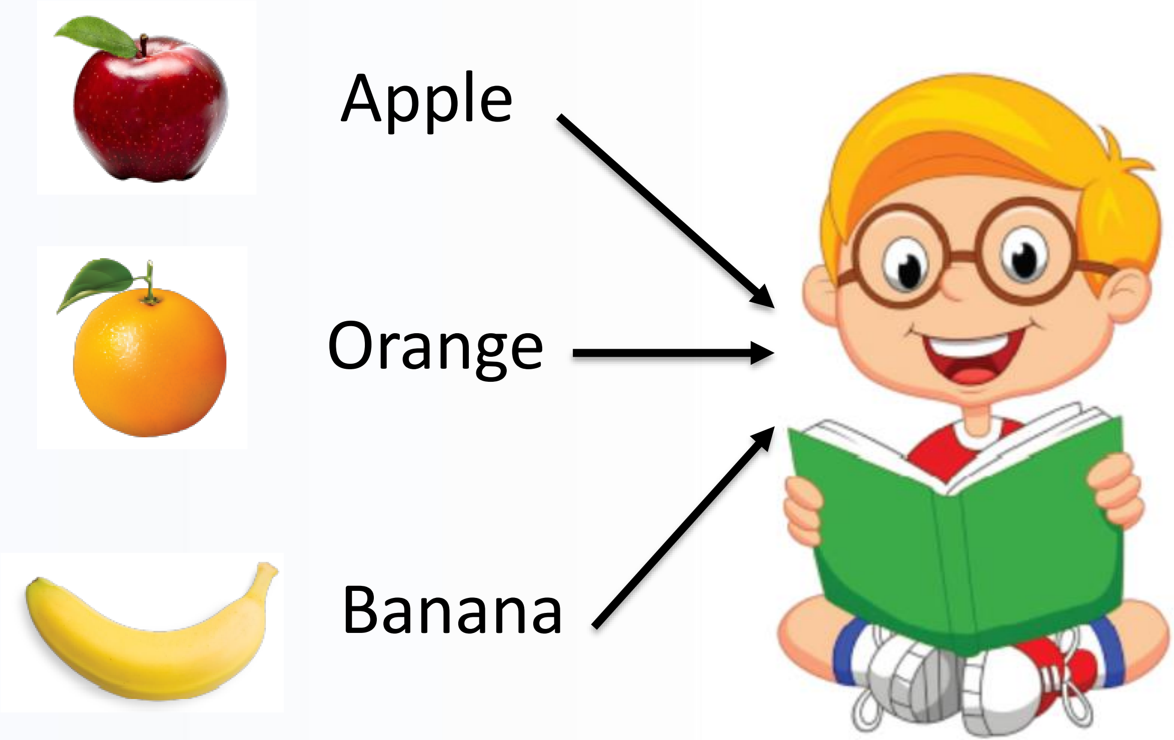
- **Target**: Predicted category or value of the data (column to predict)
- **Features**: properties of the data used for prediction (non-target columns)
- **Example**: a single data point within the data (one row)
- **Label**: the target value for a single data point (for category), also known as a response value (for numeric)

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

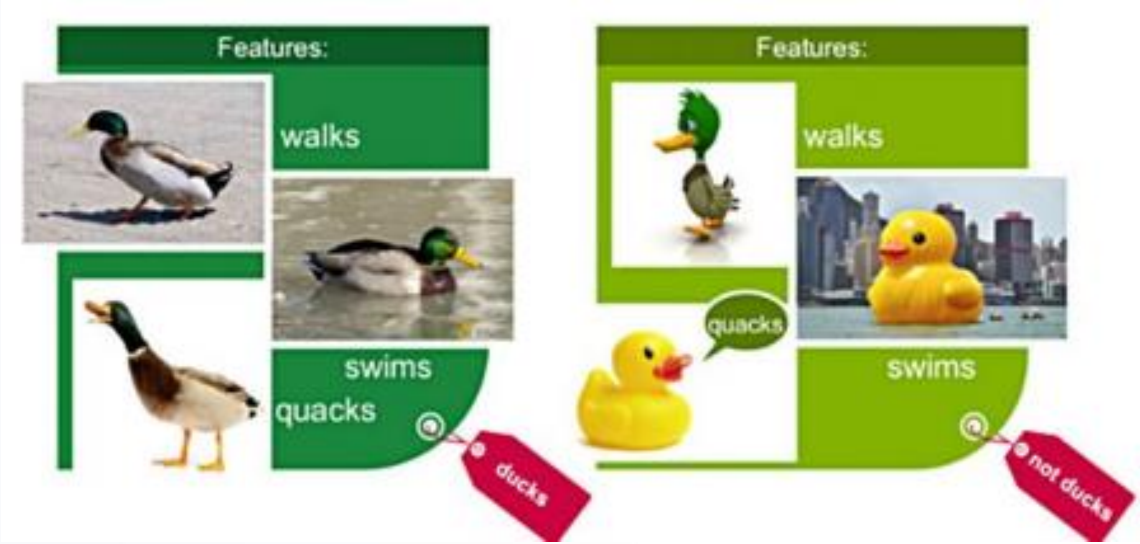
Not Duck

→ Supervised Learning → Predictive Model

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

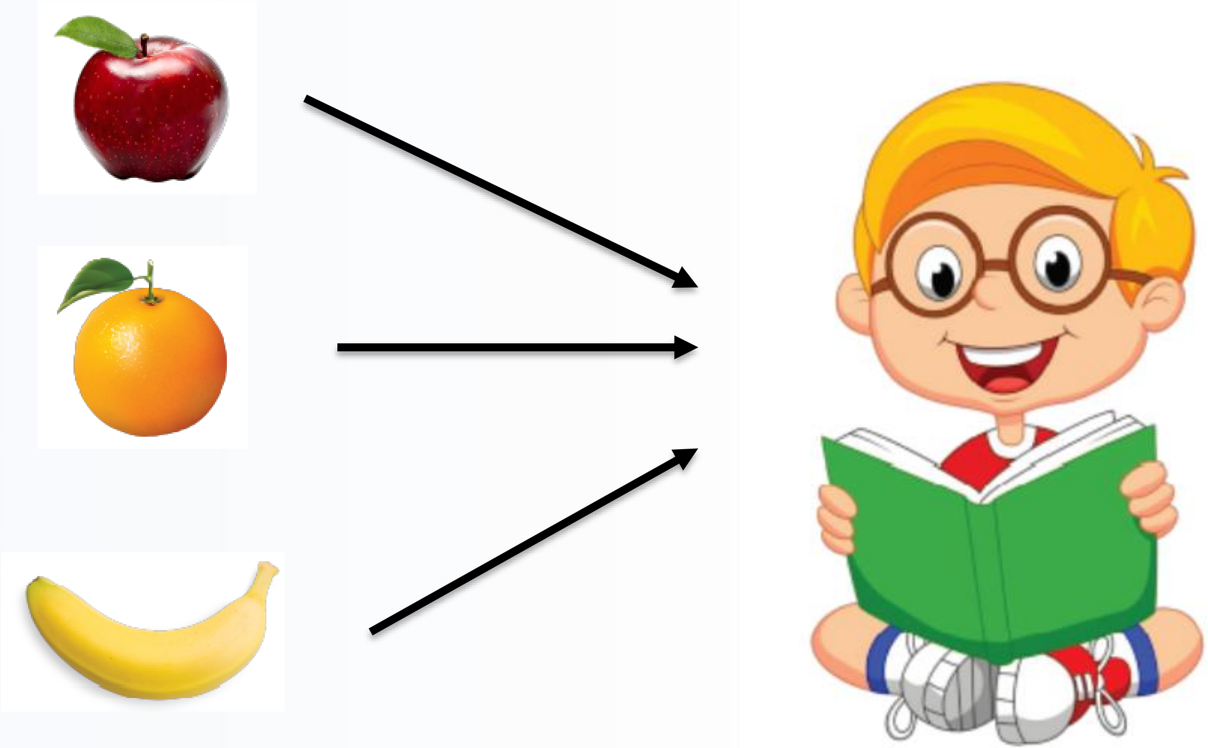
→ Predictive Model → Duck

Supervised Learning

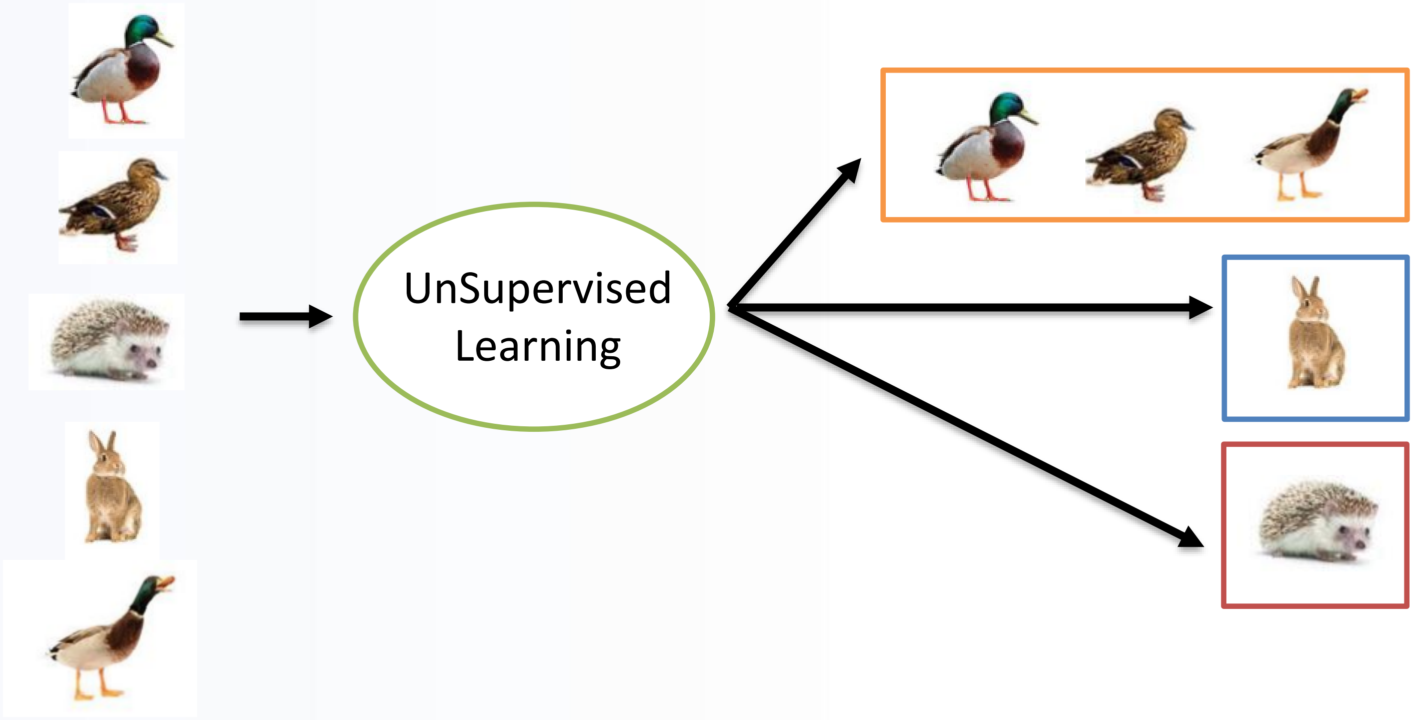


- Uses labelled training data to improve programs future actions. Reproduces known knowledge. Learn by example approach.
- Supervised learning needs to be given examples of what is “good” and what is “bad”

UnSupervised Learning



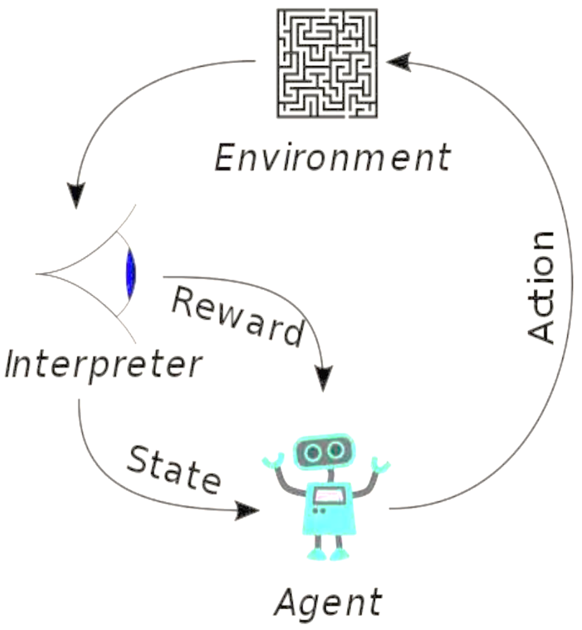
UnSupervised Learning



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

Reinforcement Learning

- Close to human learning
- Algorithm (agent) learns a policy of how to act in a given environment
- Every action has some impact in the environment, and the environment provides rewards that guides the learning algorithm
- Reinforcement Learning is learning how to act in order to maximize a numerical reward



Type of Supervised Learning

Regression

Outcome is continuous (numerical)

Classification

Outcome is a category

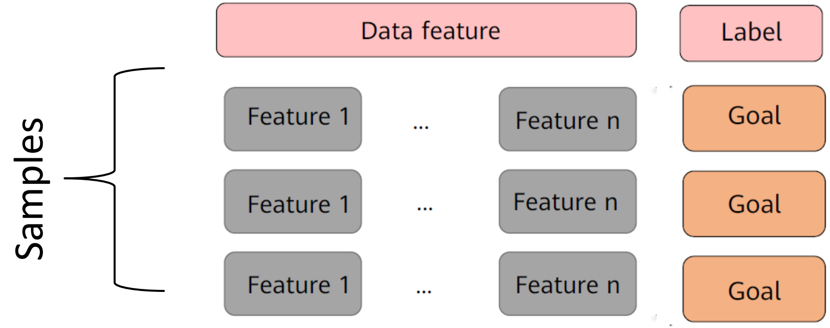
- Binary
- Multiclass

Supervised Learning Overview

Data with answer + Model → Fit → Model

Data without answer + Model → Predict → Predicted Answers

Supervised Learning



Weather	Temperature	Wind Speed	Enjoy Sports
Sunny	Warm	Strong	Yes
Rainy	Cold	Fair	No
Sunny	Cold	Weak	Yes

Features →

Samples

Weather	Temperature	Wind Speed
Sunny	Warm	Strong
Rainy	Cold	Fair
Sunny	Cold	Weak

Supervised Learning Algorithm

Enjoy Sports
Yes
No
Yes

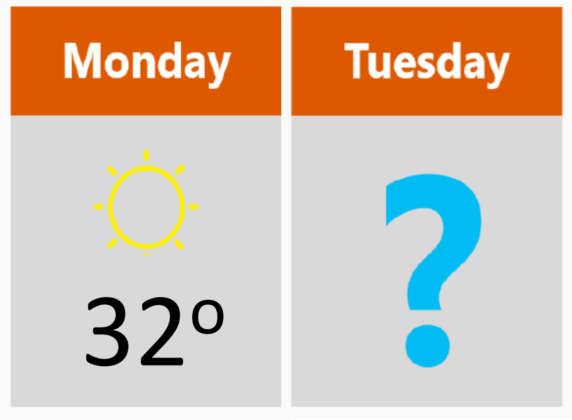
← Class

Predictions

35

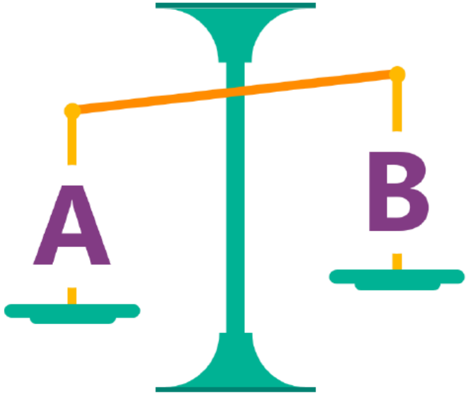
Supervised Learning - Regression

- Regression: To discover the dependencies in the sample dataset by mapping the relationships between the feature values and the response.
- Regression Questions:
  - How much will I benefit from the stock next week?
  - What's the temperature on Tuesday?

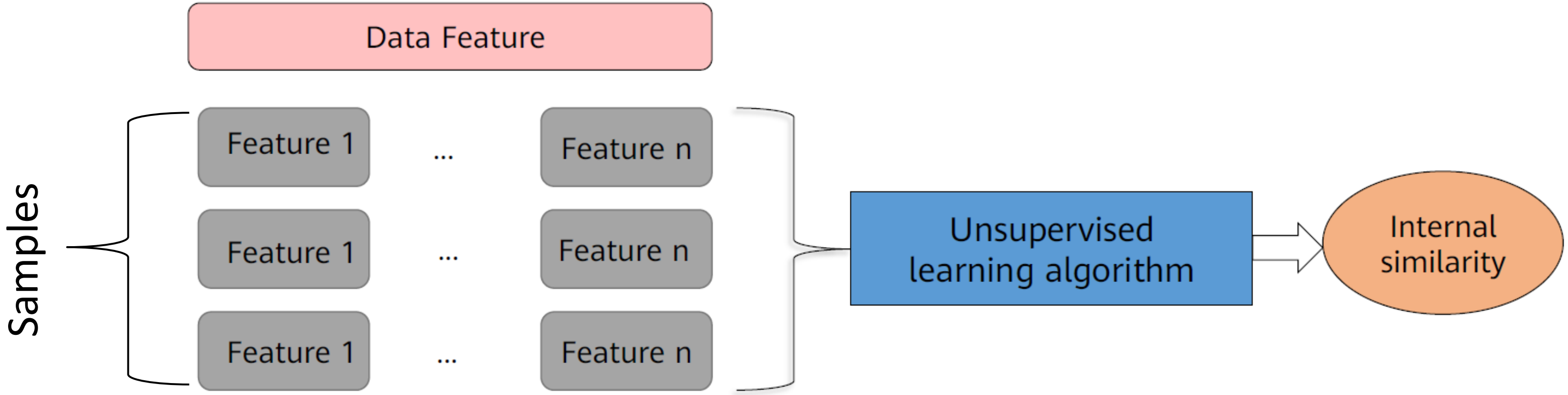


Supervised Learning - Classification

- Classification: maps samples in a sample dataset to a specified category by using a classification model.
- Classification Questions:
  - Will there be a traffic jam on XX road during the morning rush hour tomorrow?
  - Which method is more attractive to customers: $10 voucher or 25% off?



Unsupervised Learning



Features →

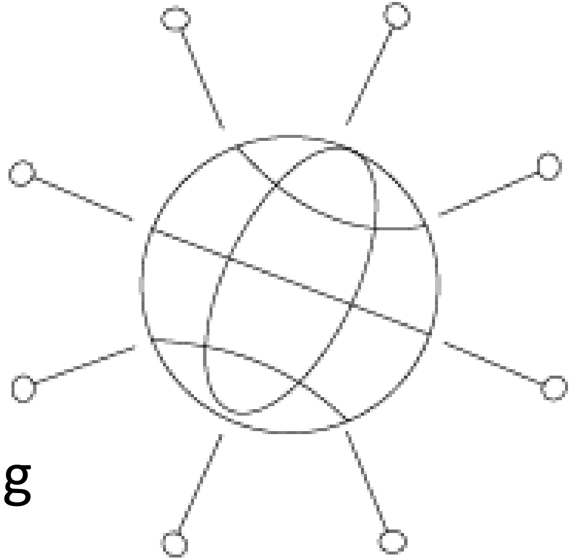
Samples

Monthly Consumption	Commodity	Consumption Time
1000–2000	Badminton racket	6:00–12:00
500–1000	Basketball	18:00–24:00
1000–2000	Game console	00:00–6:00

Category
Cluster 1
Cluster 2

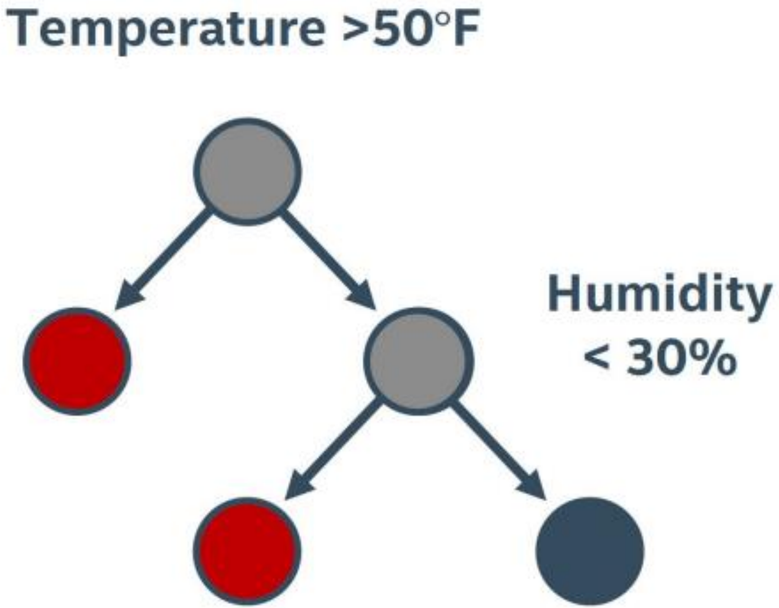
Examples of Classification Algorithms

- K Nearest Neighbors
- Logistic Regression
- Support Vector Machine
- Gaussian Naive Bayes
- Decision Tree
- Ensemble Methods
  - Random Forest, Gradient Boosting



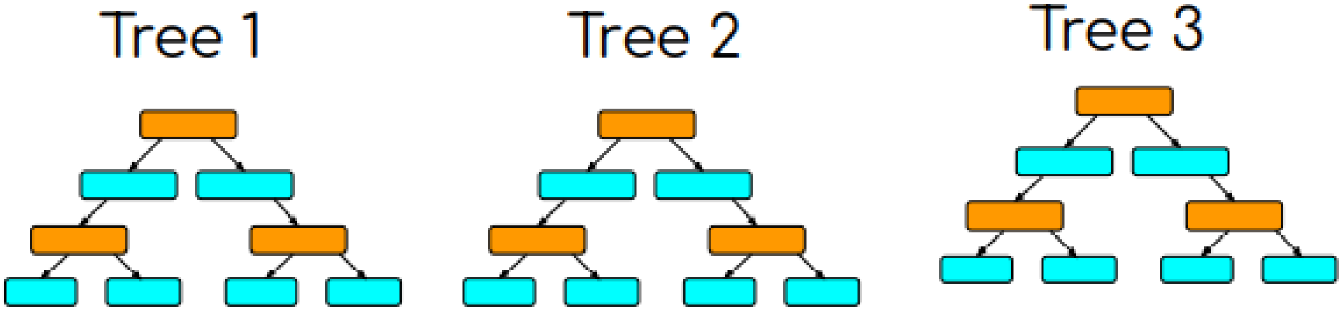
Example: Decision Tree

- Decision tree is easy to interpret and implement
- Heterogeneous input data allowed, preprocessing required
- However, decision trees tend to overfit
- Pruning helps reduce variance to a point. Often not significant for model to generalize well



Example: Random Forest

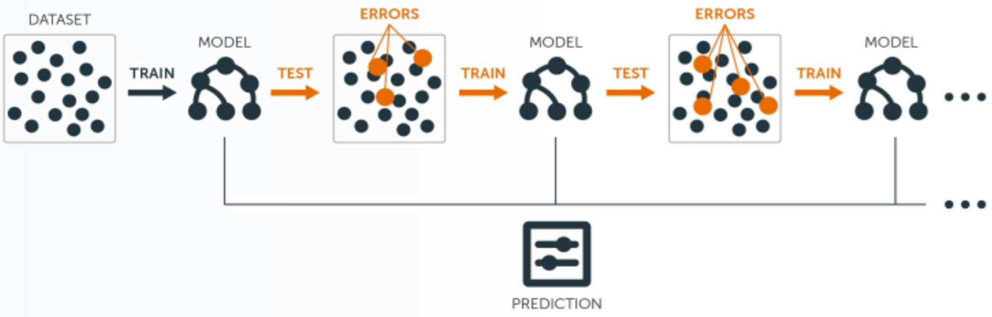
Random Forest classifier is a bagging ensemble method based on lots of decision trees with random selection of subsets of training samples.



The result is based on the majority votes from all the decision trees

Example: Gradient Boosting

- Gradient Boosting is a ensemble boosting method that "boosting" many weak predictive models into a strong one, in the form of ensemble of weak models.
- Boosting utilizes different loss functions At each stage, the margin is determined for each point. Margin is positive for correctly classified points and negative for misclassifications. Value of loss function is calculated from margin



Binary Classification

- In many predictive analysis, we are interested in YES/NO analysis such as spam/not-spam, health/not-healthy etc.
- In this case, we can form a confusion matrix of 2 columns and 2 rows

Type II Error

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Type I Error

True Positive: You predicted positive and it's true

True Negative: You predicted negative and it's true

False Positive: (Type 1 Error) You predicted positive and it's false

False Negative: (Type 2 Error) You predicted negative and it's false.

Confusion Matrix

Several measurements are used to assess the performance of a classifier. These are: Accuracy, Precision, Recall, and F1 score.

	Predicted Positive	Predicted Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

Accuracy = $\frac{TP + TN}{TP + FN + FP + TN}$

(is the percentage (x100) of correct hits from all classes – positive and negative)

Recall = $\frac{TP}{TP + FN}$

(is the percentage (x100) of correct hits from all the positive classes)

Precision = $\frac{TP}{TP + FP}$

(is the percentage (x100) of correct hits from all the positive predictions)

F1 Score = $\frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}$

takes both false positive and false negatives into account

AUC and ROC Curve in Machine Learning

- Receiver Operating Characteristics (ROC) curve provides a graphical representation of the effectiveness of the binary classification model
- It plots the true positive rate (TPR) vs the false positive rate (FPR) at different classification thresholds
- Area Under the Curve (AUC) represents the area under the ROC curve

$$\textbf{TPR} = \frac{TP}{TP + FN}$$

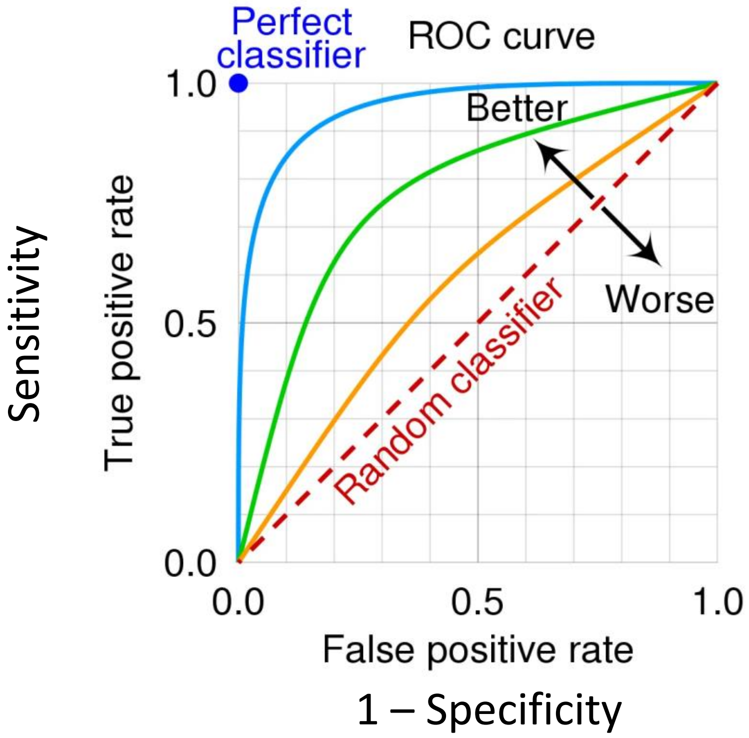
True Positive Rate / Sensitivity / Recall

$$\textbf{FPR} = \frac{FP}{TN + FP}$$

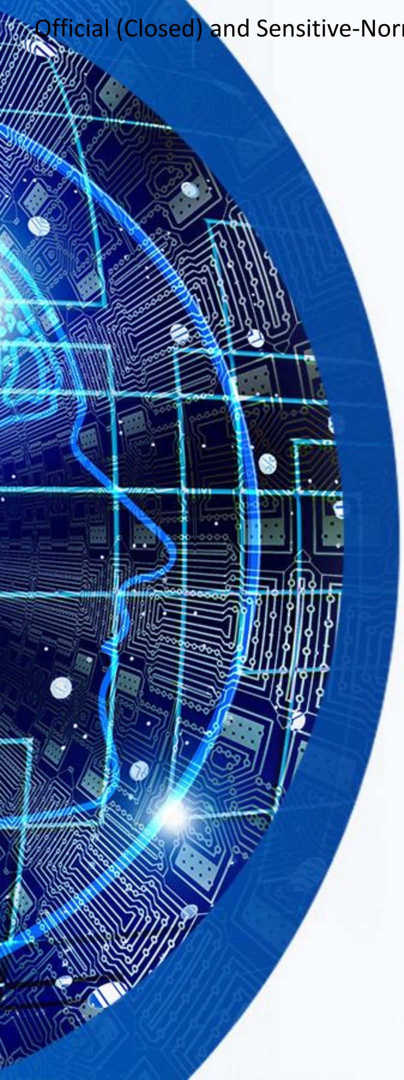
Also define as (1 – Specificity)

AUC-ROC

- ROC-AUC tries to measure and ranks the performance of classifiers
- AUC can help you decide which model is better



https://en.wikipedia.org/wiki/Receiver_operating_characteristic



Regression Algorithms

- Linear Regression (Most Common)
- Ridge Regression
- Lasso Regression
- Elastic Net Regression

What is Linear Regression

- Linear regression is a statistical method of finding the relationship between independent and dependent variables
- In linear regression, we are trying to obtain the line of best fit
- Two common methods to achieve this is:
  - Ordinary Least Square
  - Gradient Descent

Loss Function and Mean Square Error

- Mean Square Error (MSE) is the common loss function to measure how good is the linear regression model
- Machine Learning aims to minimize the MSE to find the best linear regression model or the “line of best fit”
- In other words, we want to identify a slope-intercept, in the equation, *y = mx + b* (hence the word linear)

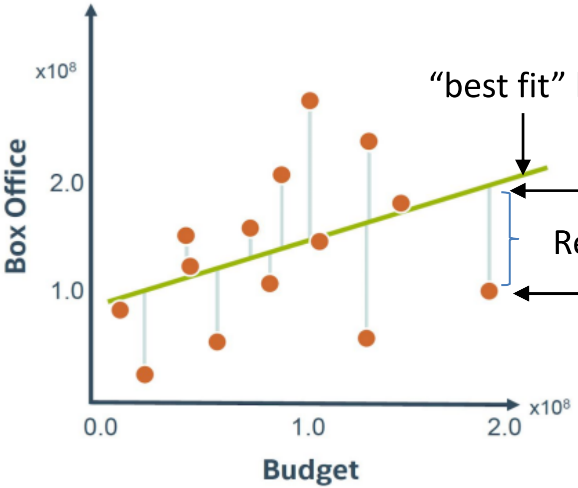
Parameters or weights

Sample size

$$\frac{1}{m}\sum_{i=1}^{m}\left(\left(\beta_0 + \beta_1 x_{obs}^{(i)}\right) - y_{obs}^{(i)}\right)^2$$

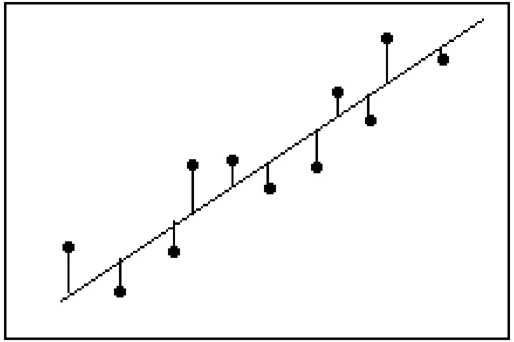
Actual/target/label

Feature or predictor



Assessing the Goodness-of-Fit

- After you have fit a linear model using regression analysis, you need to determine how well the model fits the data.
  - R-squared (or Coefficient of Determination)
    - $R^2 = 1 -$ (Sum of Square Error / Total Sum of Square)
    - $0 \leq R^2 \leq 1$
    - If $R^2 = 0$, model is not useful, if $R^2 = 1$, then model fits all points perfectly or the dependent variable(s) is explained by the independent variables
    - A rule of thumb is $R^2 \geq 0.7$ is a good fit

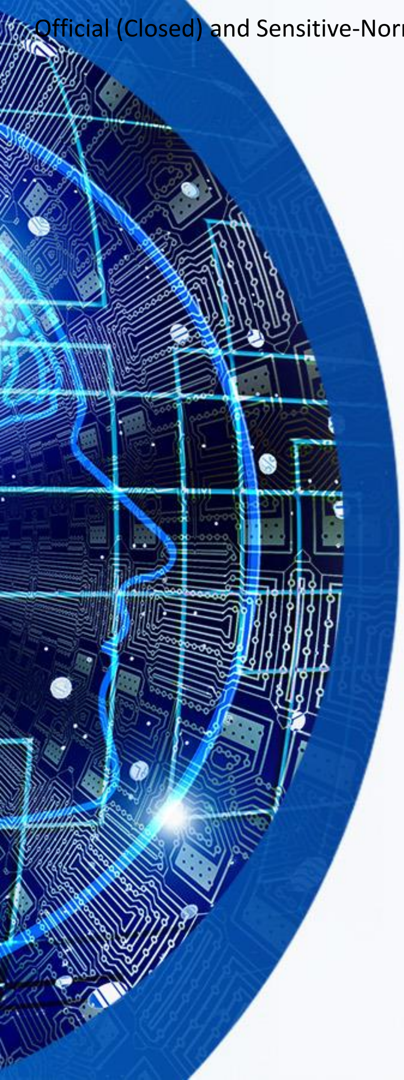


Regularization

- Fitting the training data too closely is called overfitting and not fitting it well enough is called underfitting
- Overfitting is a problem in machine learning, and we want to discourage the model from learning overly complex patterns in the data
- Regularization in machine learning is a technique to prevent the model from overfitting by adding extra information, or penalty term, to the model’s loss function

Common Regularization Techniques

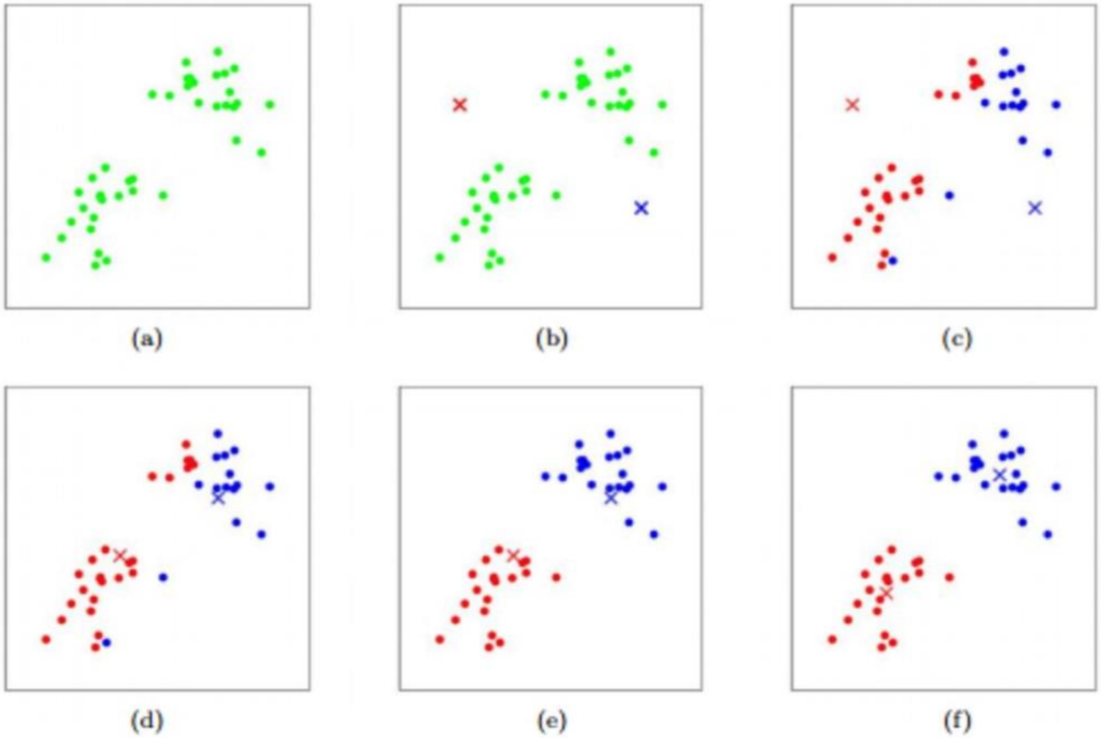
- L1 Regularization (Lasso): This technique adds the absolute values of the coefficients as a penalty term. It encourages sparsity in the model, making some features have exactly zero coefficients. It's useful for feature selection.
- L2 Regularization (Ridge): L2 regularization adds the squared values of the coefficients as a penalty term. It doesn't force coefficients to become exactly zero but helps in reducing the magnitude of all coefficients, which can lead to a simpler model.
- Elastic Net Regularization: Elastic Net combines both L1 and L2 regularization. It balances between feature selection (L1) and coefficient shrinkage (L2).



Clustering Algorithms

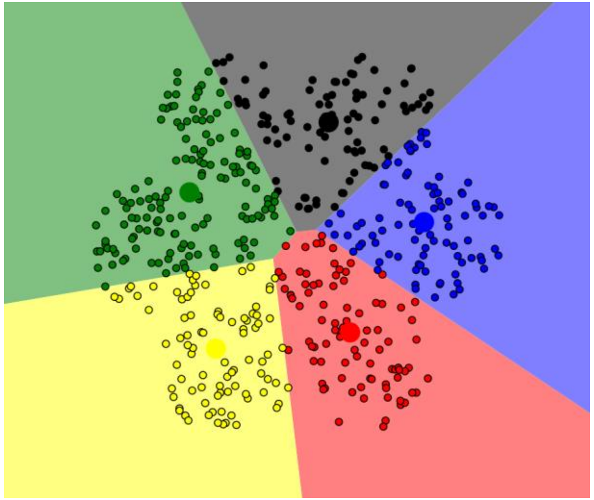
- K-Means Clustering
- Hierarchical Agglomerative Clustering

K-Mean Algorithm



K-Mean Algorithm Simulation

Click on the images below to start k-mean clustering simulation



https://www.naftaliharris.com/blog/visualizing-k-means-clustering/

3 What is Deep Learning

History of Deep Learning Milestones and Tools

DL Milestones

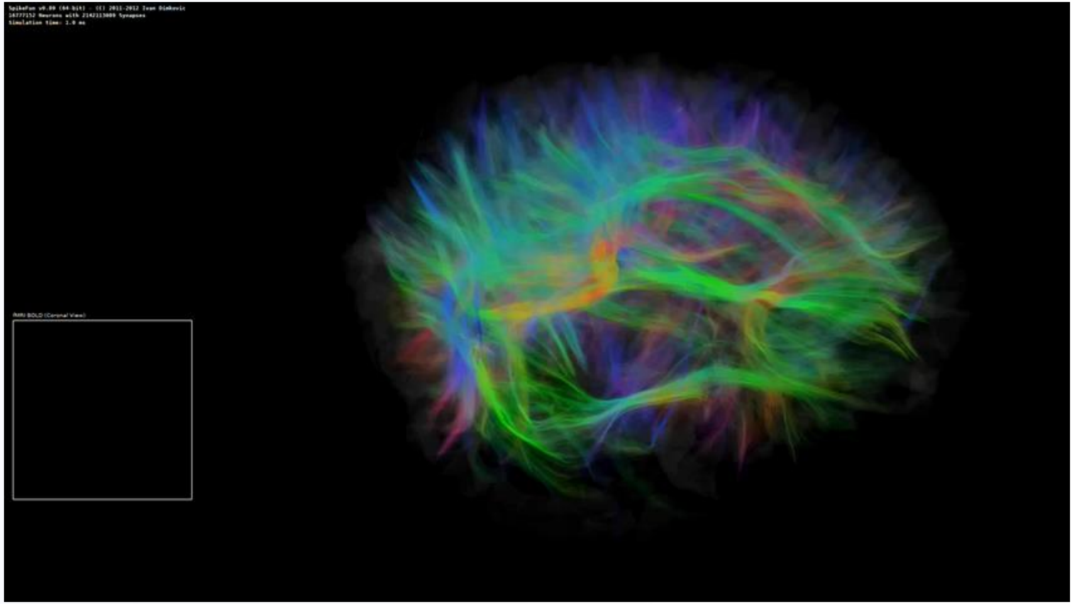
- 1943: Neural networks
- 1957: Perceptron
- 1974-86: Backpropagation, RBM, RNN
- 1989-98: CNN, MNIST, LSTM, Bidirectional RNN
- 2006: “Deep Learning”, DBN
- 2009: ImageNet
- 2012: AlexNet, Dropout
- 2014: GANs
- 2014: DeepFace
- 2015: RestNet-152
- 2016: AlphaGo
- 2017: AlphaZero, Capsule Networks
- 2018: Google BERT, OpenAI GPT
- 2019: GPT-2
- 2020: GPT-3
- 2021: GPT-3.5

DL Tools

- Mark 1 Perceptron–1960
- Torch –2002
- CUDA –2007
- Theano –2008
- Caffe –2014
- DistBelief–2011
- TensorFlow 0.1 –2015
- PyTorch0.1 –2017
- TensorFlow 1.0 –2017
- PyTorch1.0 –2017
- TensorFlow 2.0 –2019
- PyTorch1.10 - 2021
- TensorFlow 2.7 –2021

Full list of acronym can be found in the reference slides

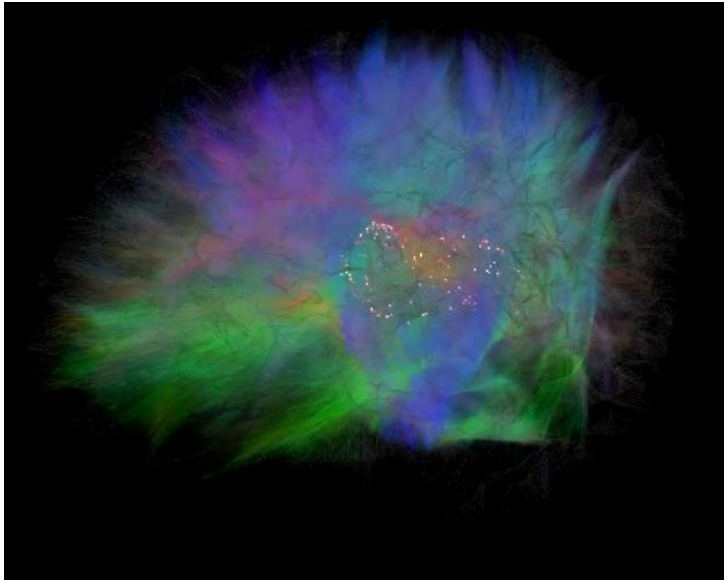
The Human Brain and the Neurons



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

Biological and Artificial Neural Networks



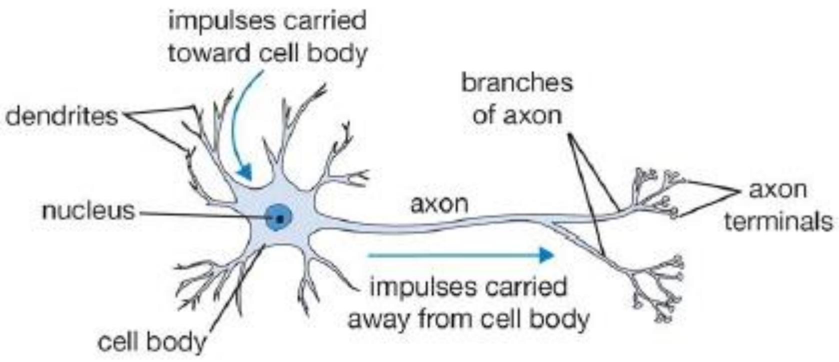
Human Brain

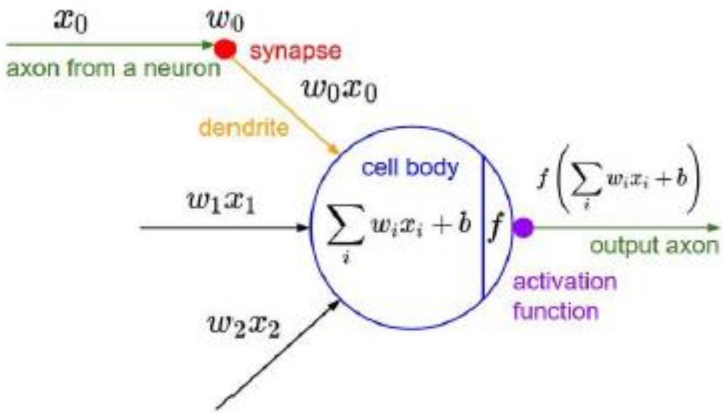
- 100 billion neurons and 1,000 trillion synapses

Artificial Neural Network

- ResNet-152: 60 million synapses
- ChatGPT-2: 50 billion neurons
- ChatGPT-3: 60-80 billion neurons and around 100 trillion synapses

Neuron: Biological Inspiration for Computation

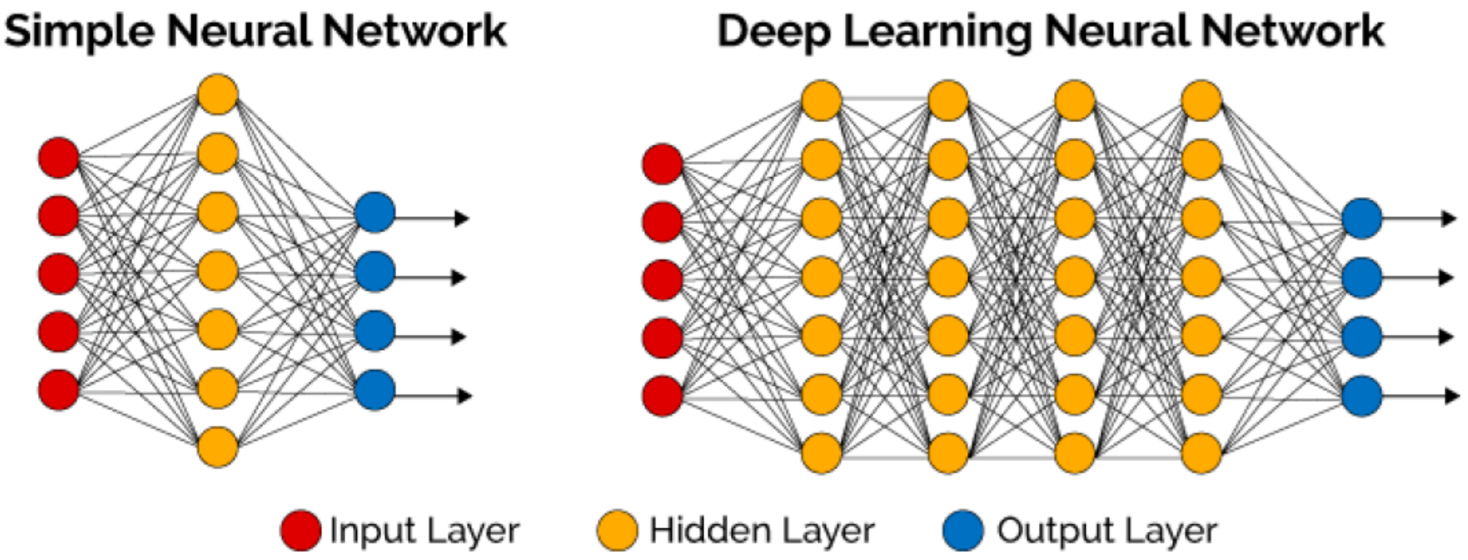




Neuron: computational building block for the brain

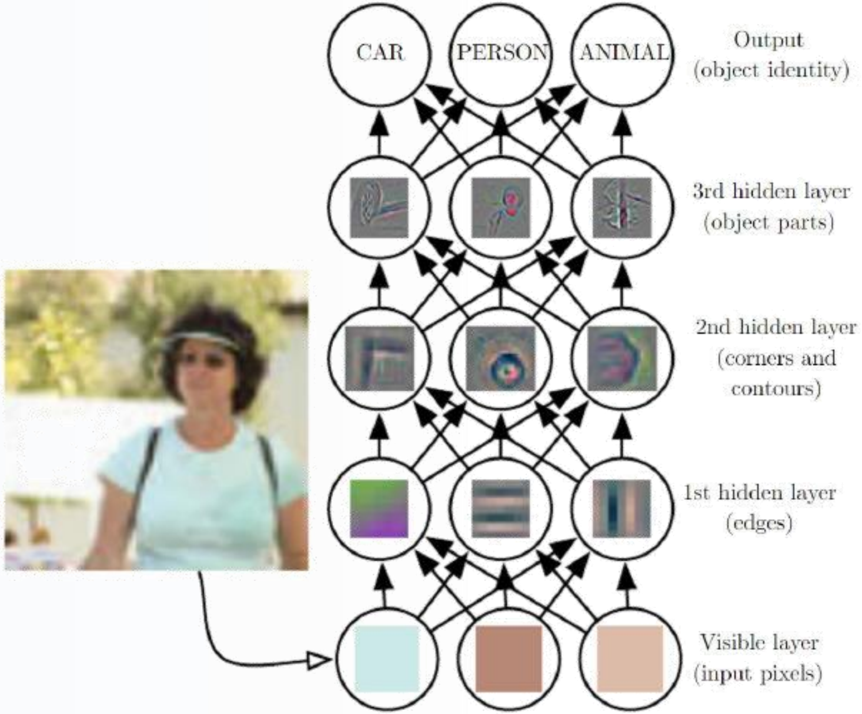
(Artificial) Neuron: computational building block for the “neural network”

Combing Neurons in Hidden Layers: Power to Approximate

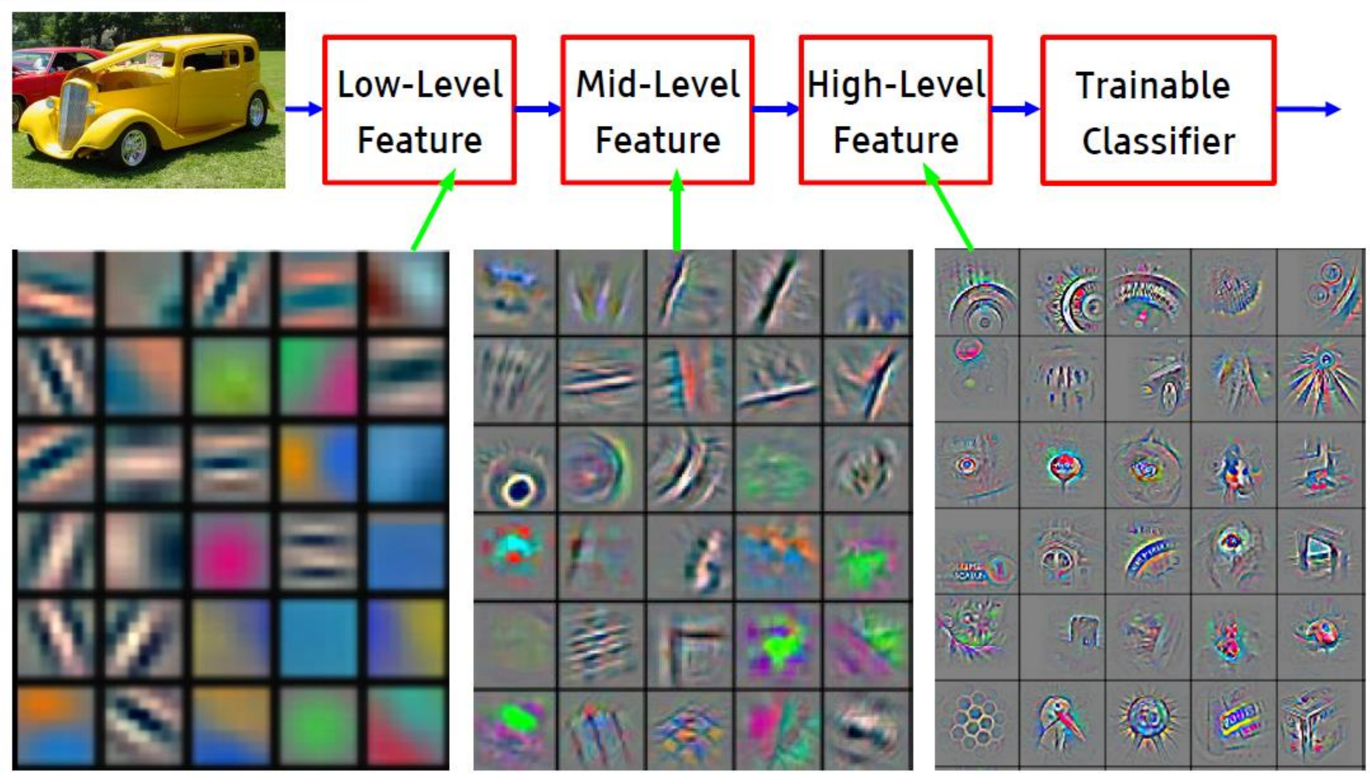


Universality: For any arbitrary function *f(x)*, there exists a neural network that closely approximate it for any input *x.*

Deep Learning is Representation Learning (aka feature learning)

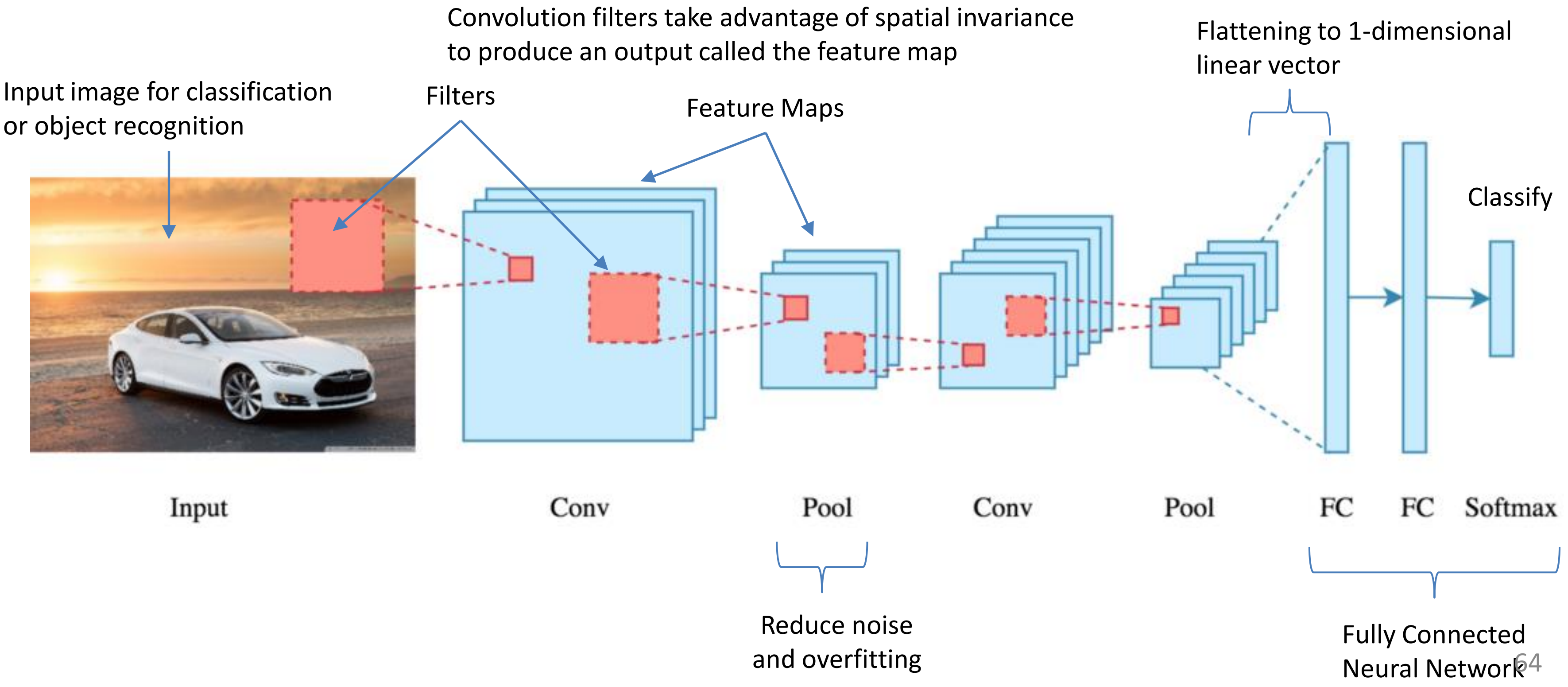


Deep Learning = Learning Hierarchical Representations

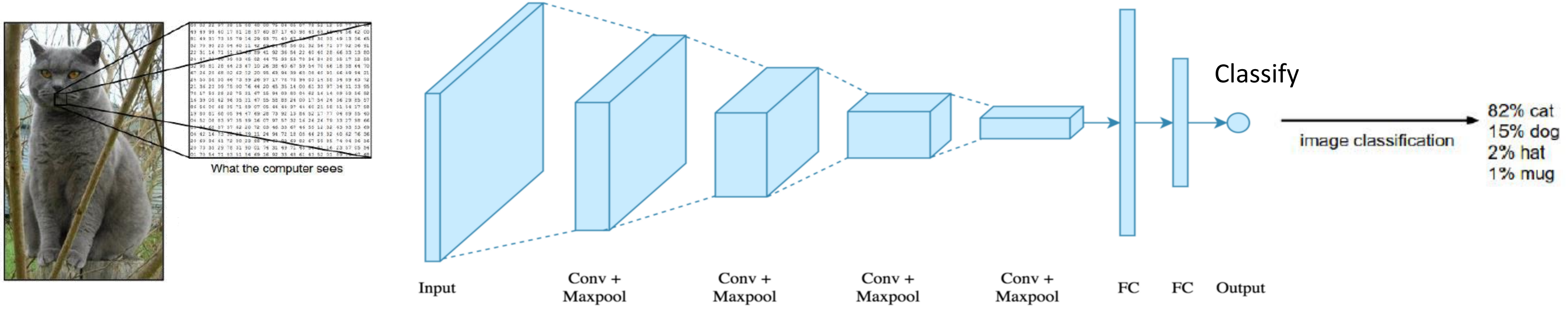


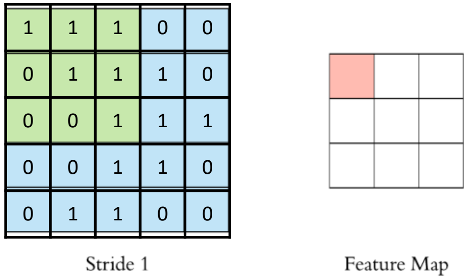
Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

Convolutional Neural Network Architecture



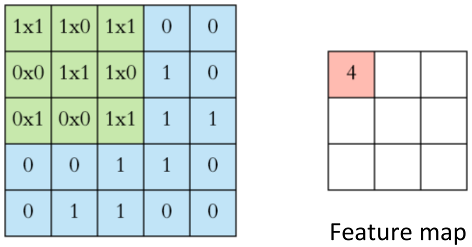
CNN Implementation





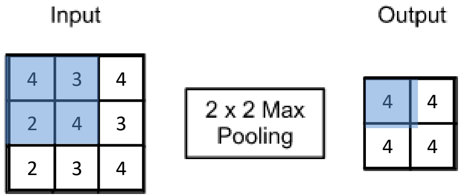
Filter

Convolution filters slides through the image



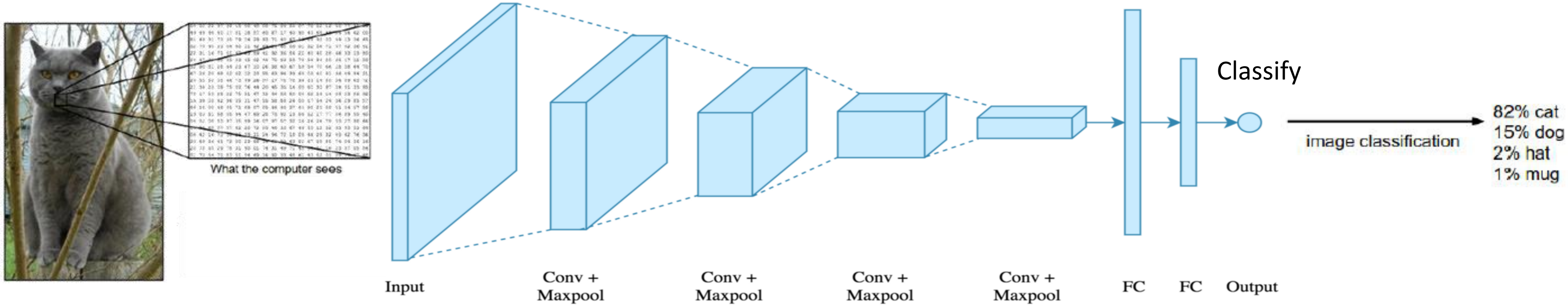
The convolution process

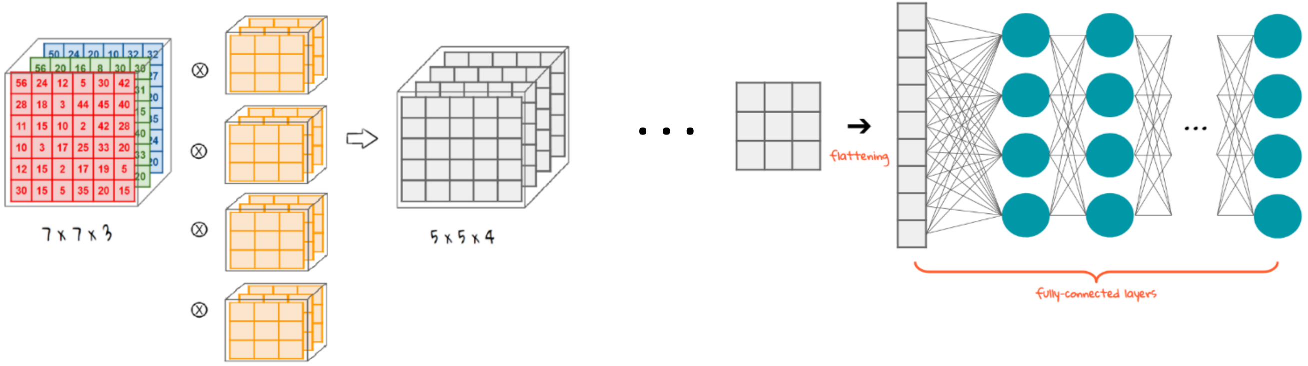
Feature map



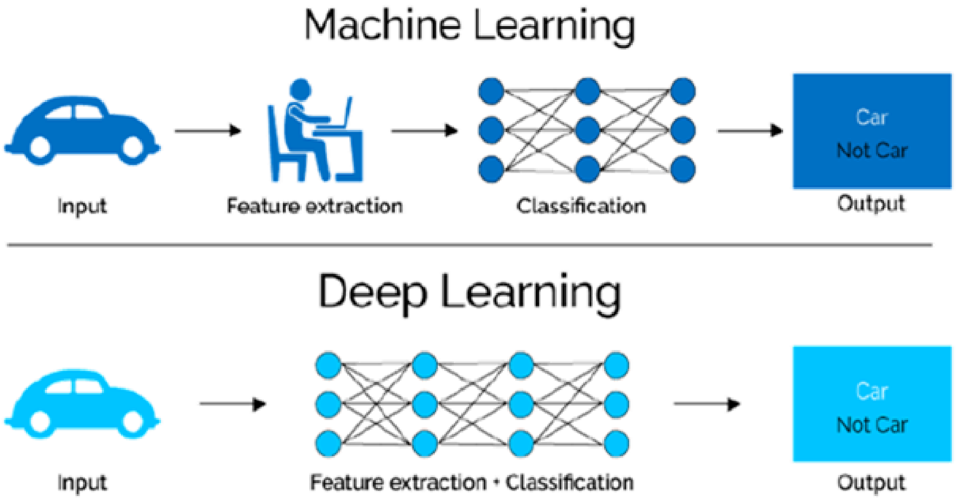
The pooling process

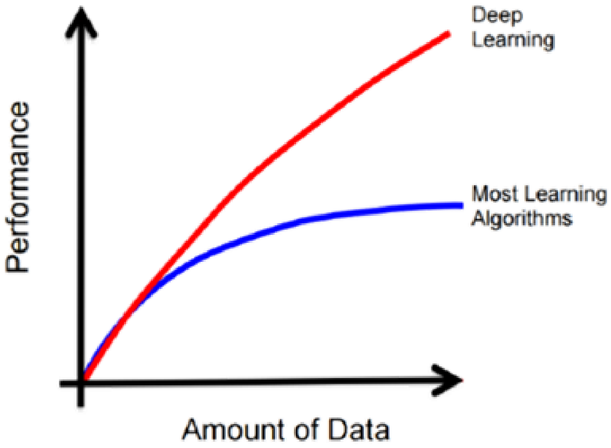
CNN Implementation (Cont’d)





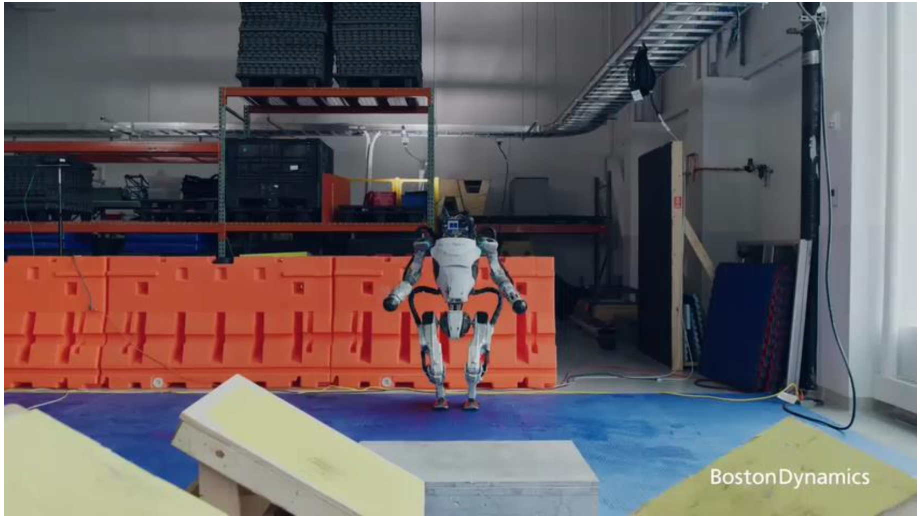
Why Deep Learning? Scalable ML

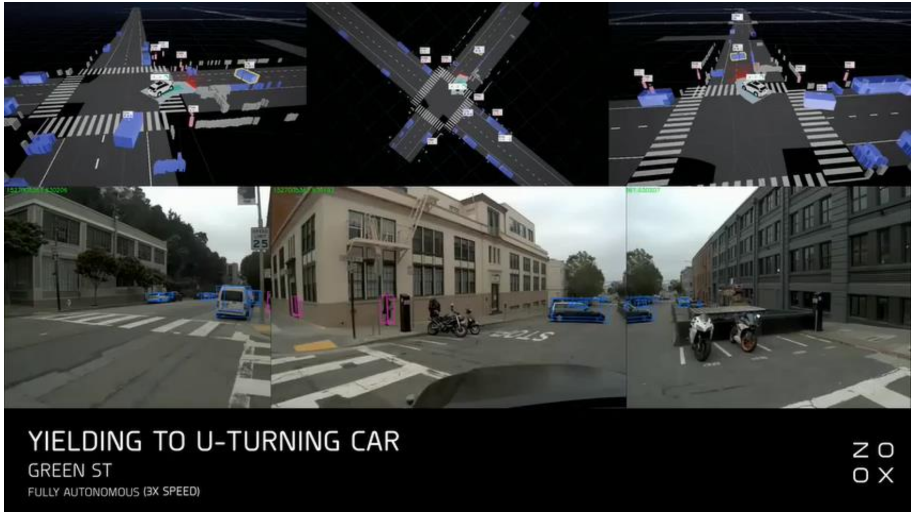




4 Applications of AI and ML

Why not Deep Learning? Real-World Applications



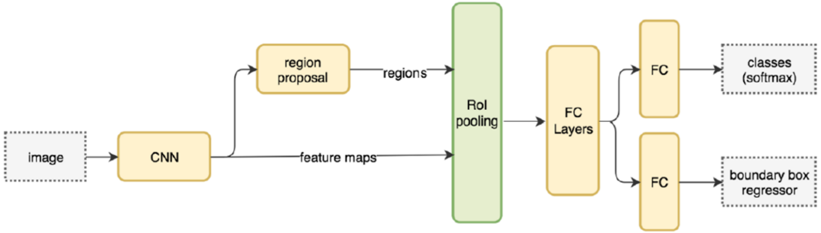


https://www.youtube.com/watch?v=tF4DML7FIWk

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

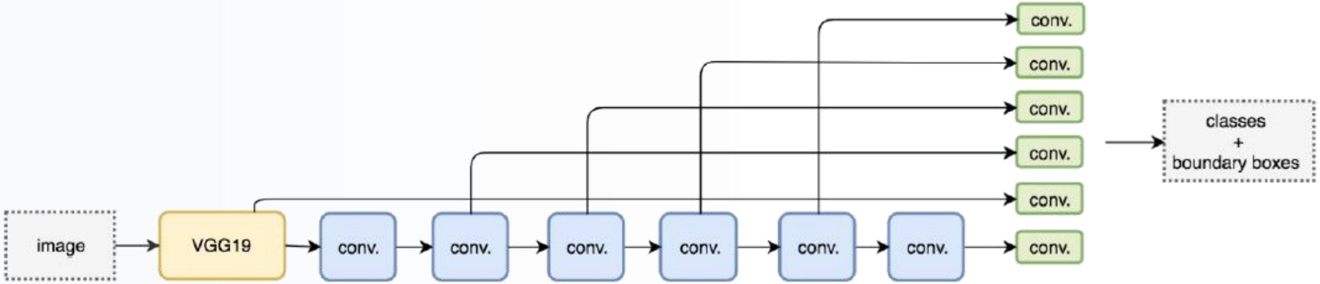
Image Detection / Localization with Region-Based Methods (Faster R-CNN)

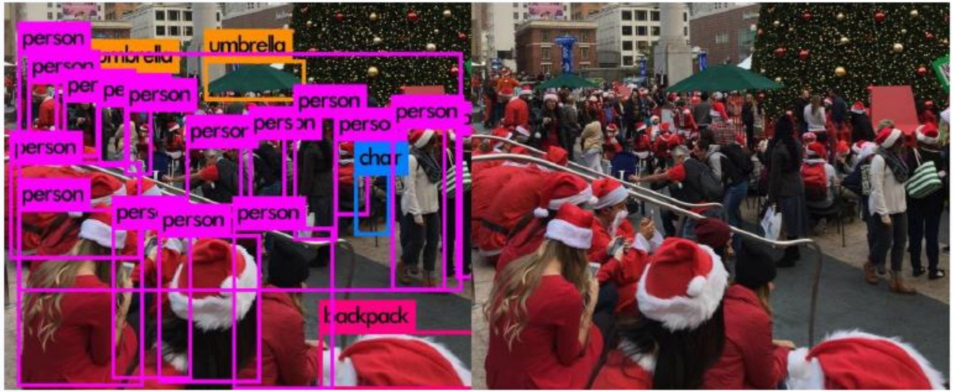
```
ROIs = region_proposal(image)
for ROI in ROIs
    patch = get_patch(image, ROI)
    results = detector(patch)
```



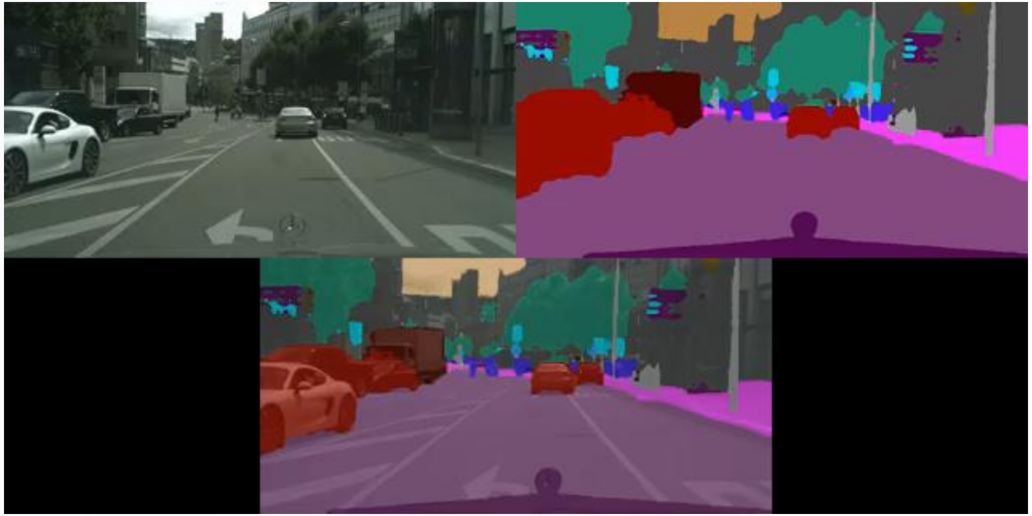
https://www.youtube.com/watch?v=zrHfvVdvmz4&t=1s

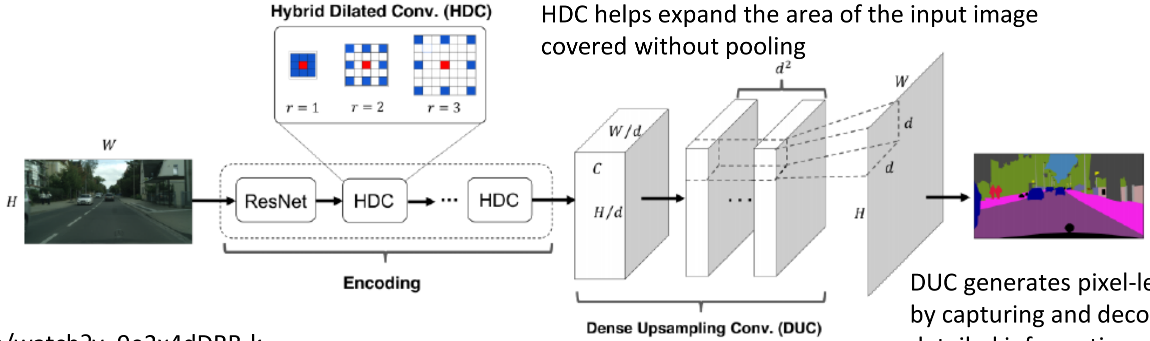
Image Detection / Localization with Single Shot Detection Methods (SSD)





Semantic Segmentation



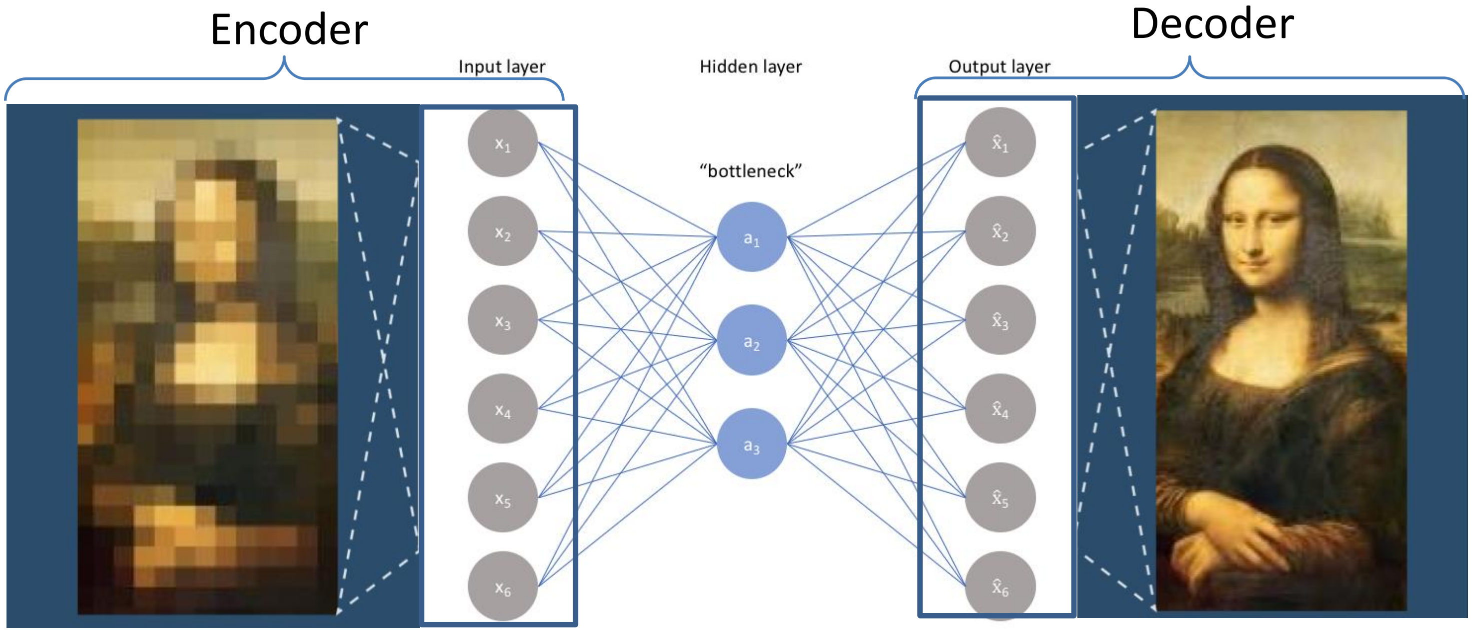


HDC helps expand the area of the input image covered without pooling

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

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

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.

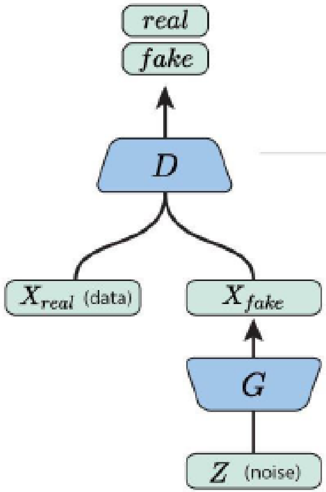
Decoder upsampling the convolutional blocks and reconstructs the data back from its encoded form from the bottleneck's output.

Image Enhancement using AI

- Let's Enhance - https://letsenhance.io/
- Topaz Labs - https://topazlabs.com/
- Vance AI - https://vanceai.com/
- DeepArt.io - https://deepart.io/
- Remini - https://www.remini.ai/

Generative Adversarial Networks (GANs)

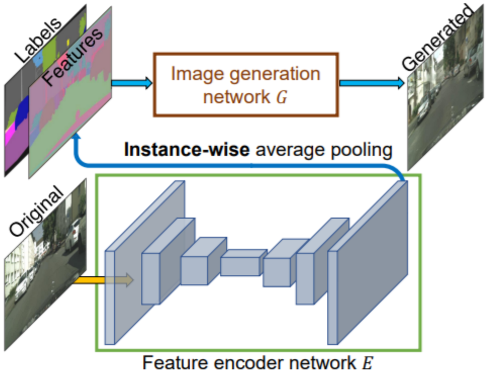
Generative Adversarial Networks (GANs) are a way to make a generative model by having two neural networks compete with each other.



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.

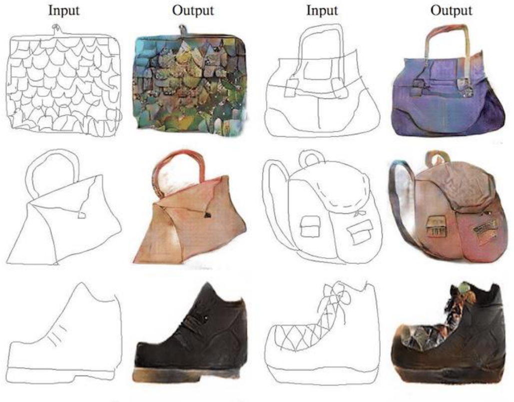
Generate Realistic Photographs

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

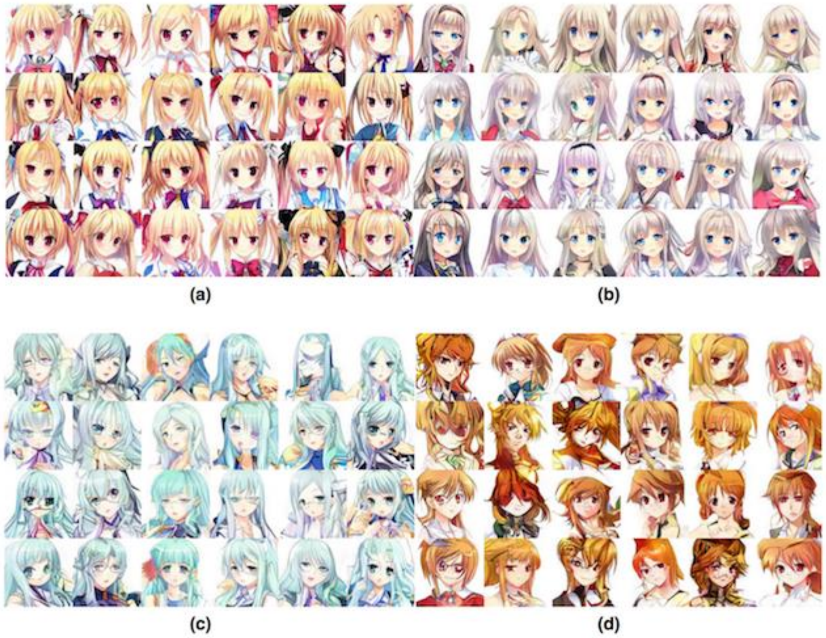
Translation of sketches to color photographs



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

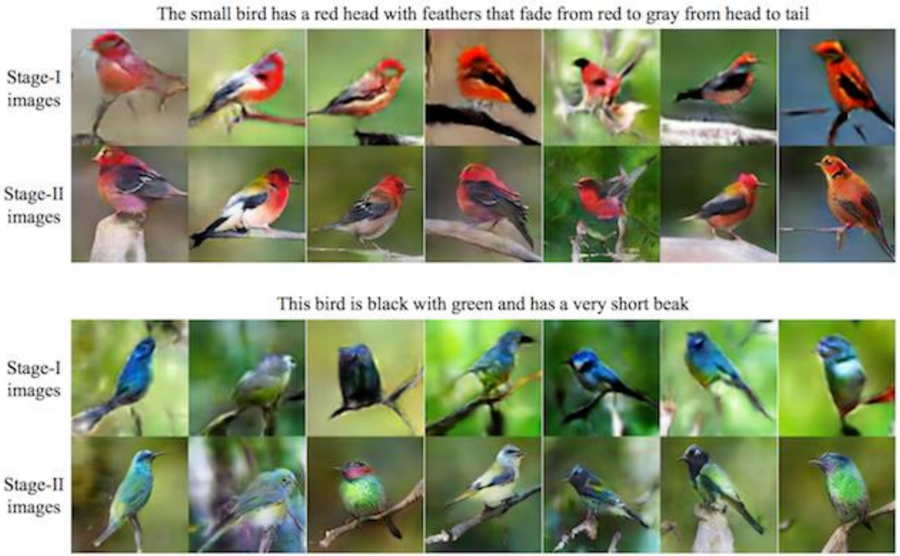
Applications of Generative Adversarial Networks (GANs)

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.

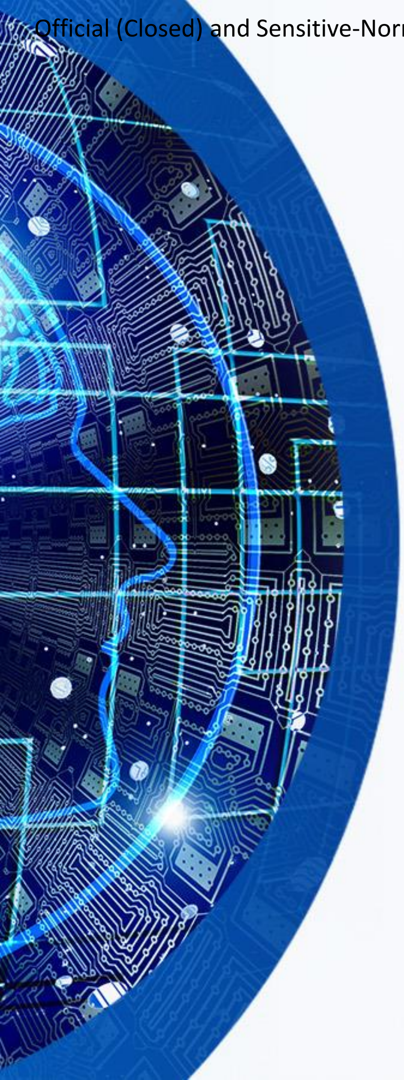
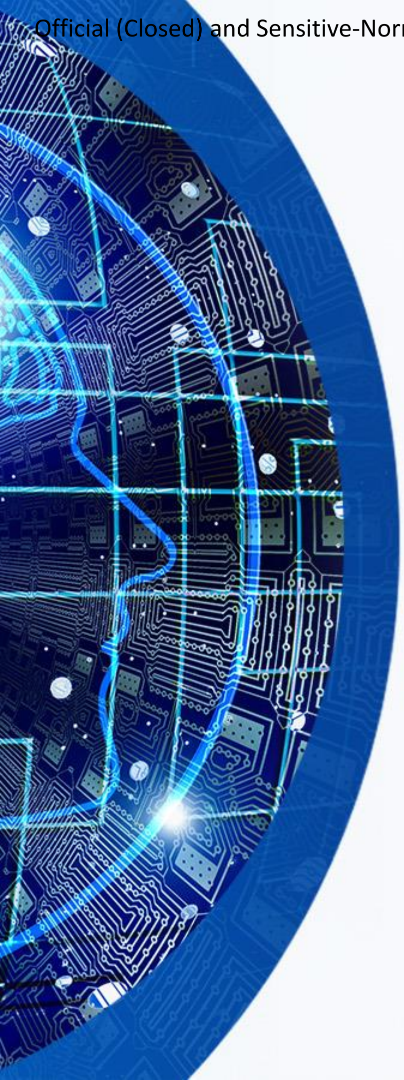


Image Generation using AI

- ArtBreeder - https://www.artbreeder.com/
- This Person Does Not Exist - https://thispersondoesnotexist.com/
- Generated Photos - https://generated.photos/
- Deep Dream Generator - https://deepdreamgenerator.com/
- Runway ML - https://runwayml.com/

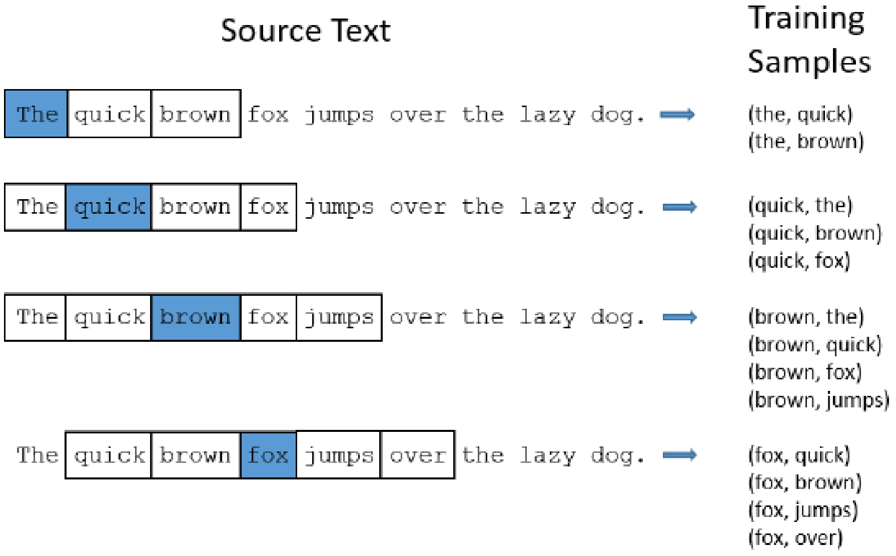


Music/Lyrics Generation using AI

- https://www.ampermusic.com/
- OpenAI Jukebox - https://openai.com/blog/jukebox/

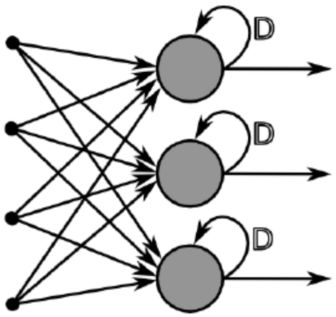
Word Embeddings (Word2Vec)

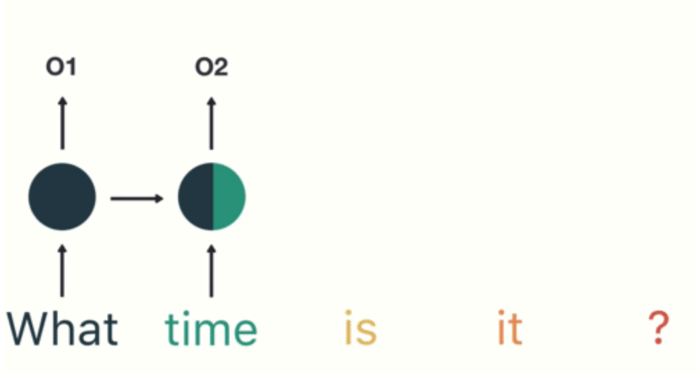
- Word embedding is capturing context of a word in a document, semantic and syntactic similarity, relation with other words
- Word2Vec is a method to construct an embedding. Another way to put it is, they are vector representations of a particular word.
- It is used to predict the source context words (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

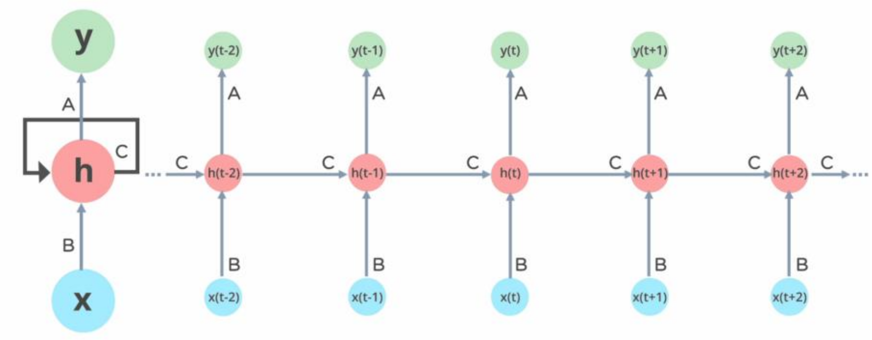


Skip Gram Model

Recurrent Neural Networks (RNNs)







Applications

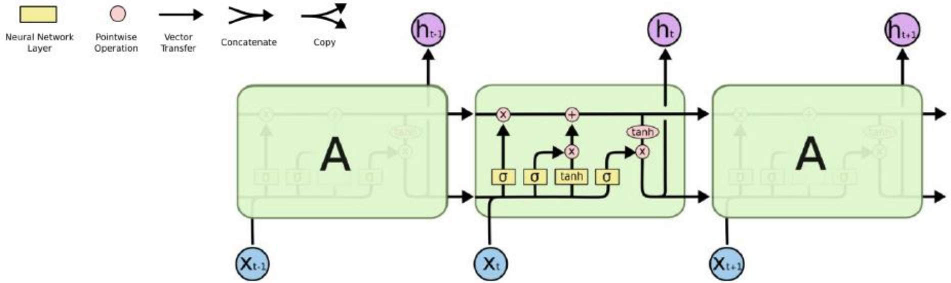
- Sequence Data
- Text
- Speech
- Audio
- Video
- Generation

Long Term Dependency

Short-term dependence:
Bob is eating an apple

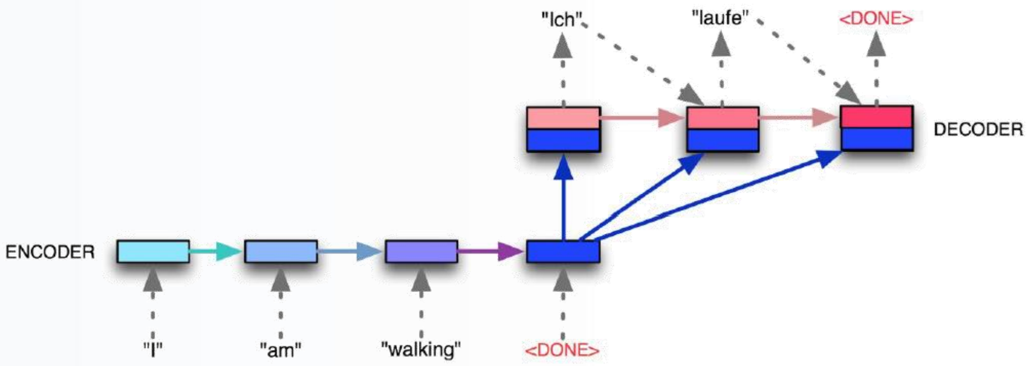
Context

Long-term dependence:
Bob likes apples. He is hungry and decided to have a snack. So now he is eating an apple.

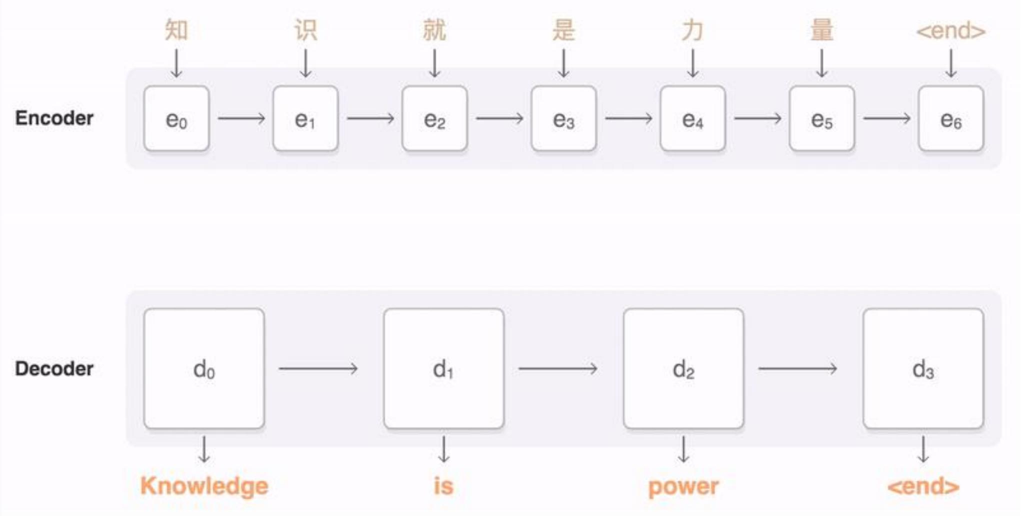


Long Short-Term Memory (LSTM) Networks:
Pick What to Forget and What To Remember

Encoder-Decoder Architecture



Machine Language Translation

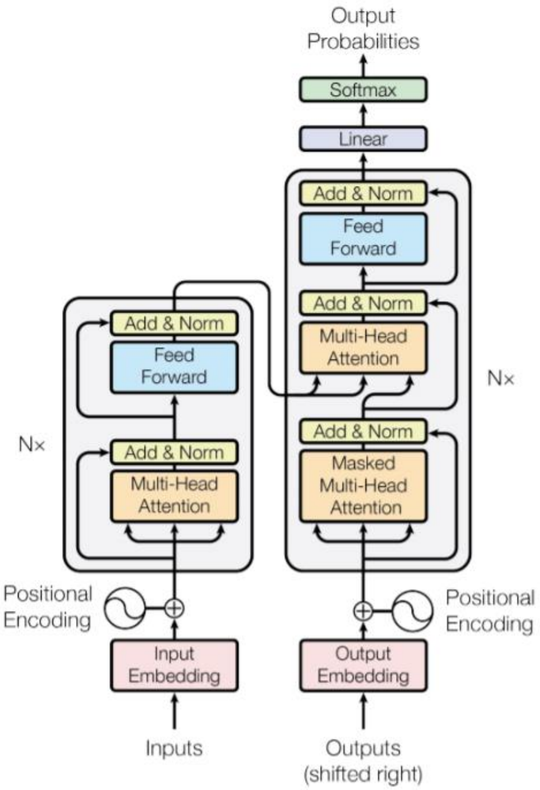


Language Model – Generative Pre-Train Transformer (GPT)

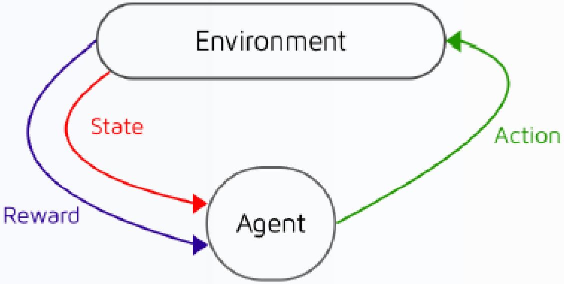
1. Input embedding
2. Multi-head self-attention
3. Feed-forward network
4. Normalization and residual connections

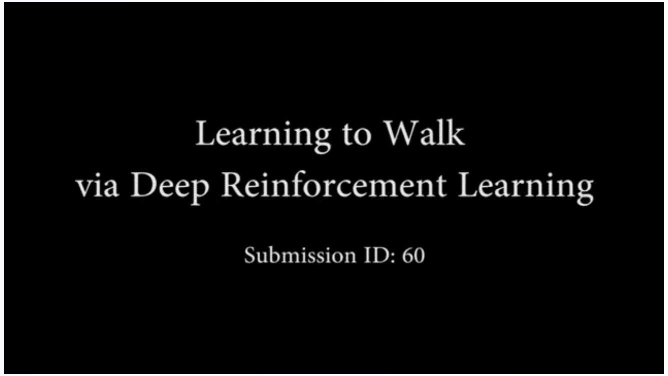
Ashish Vaswani et. al. 2017. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 6000–6010.

https://arxiv.org/abs/1706.03762



Deep Reinforcement Learning

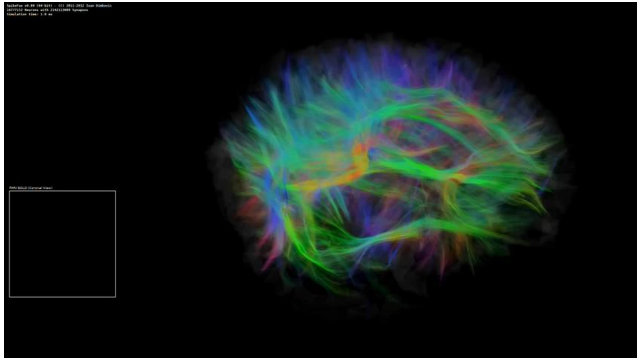


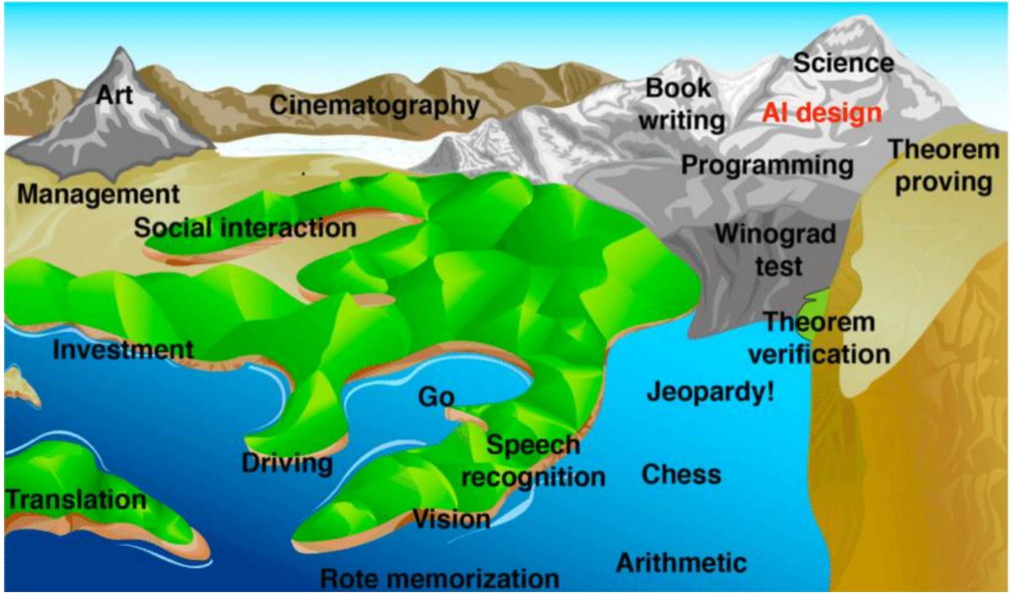


https://www.youtube.com/watch?v=n2gE7n11h1Y

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
- 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 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.

Quiz III

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

A. Machine Learning

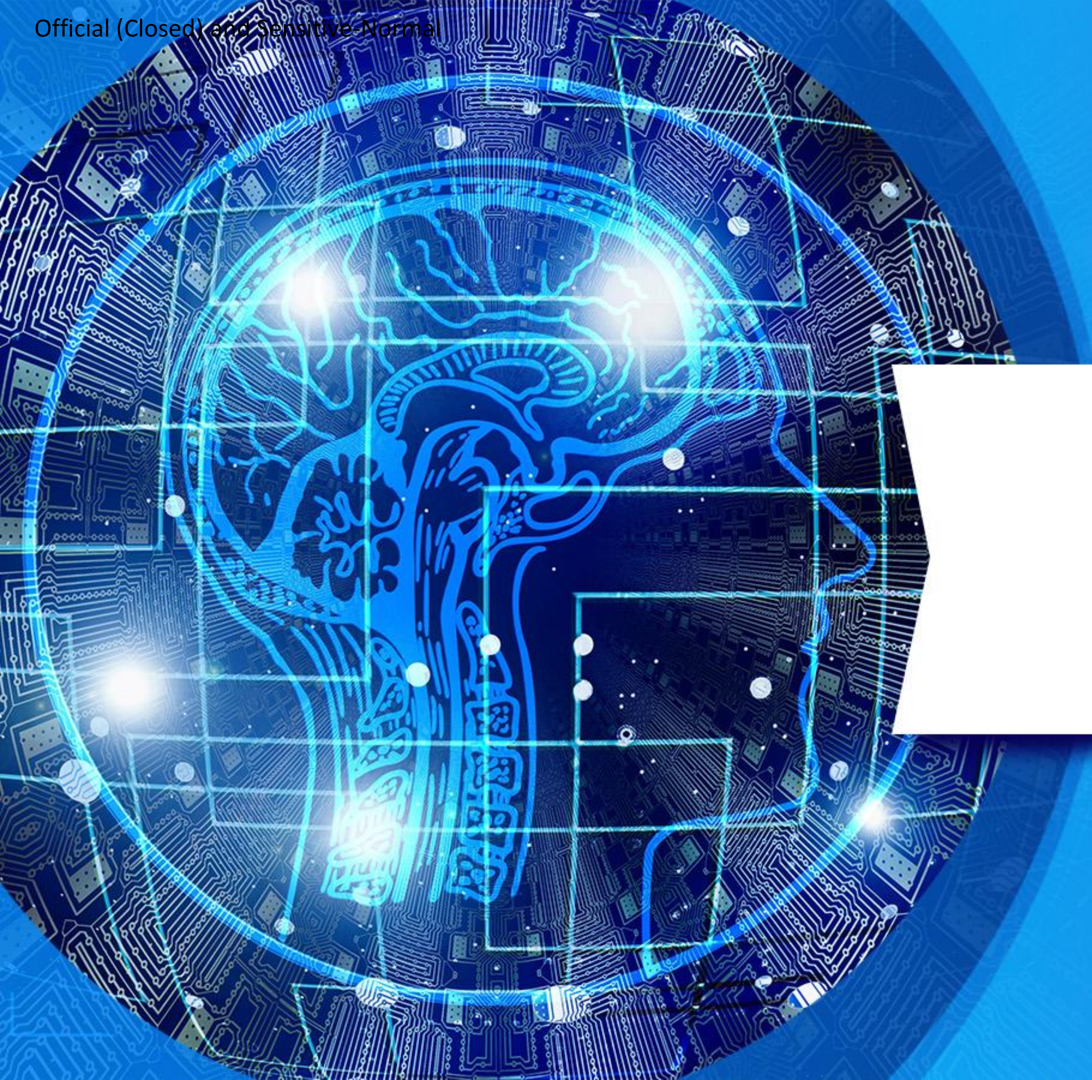
B. Supervised or Unsupervised

C. Representation Learning

D. Artificial Neural Networks

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



The End

URL: [Qualtrics Survey | Qualtrics Experience Management]

COURSE RUN ID: 793880

Course Title:	**INTRODUCTION TO AI & MACHINE LEARNING**
Course Start Date:	5 Oct 2023
Name of Trainer(s):	Dr Foo Yong Wee

URL: https://forms.gle/NBo6VvmeTrqqxvKV8

Course Title: **INTRODUCTION TO AI & MACHINE LEARNING (CE1574-230001)**

Course Start Date: 5 Oct 2023
Name of Trainer(s): Dr Foo Yong Wee