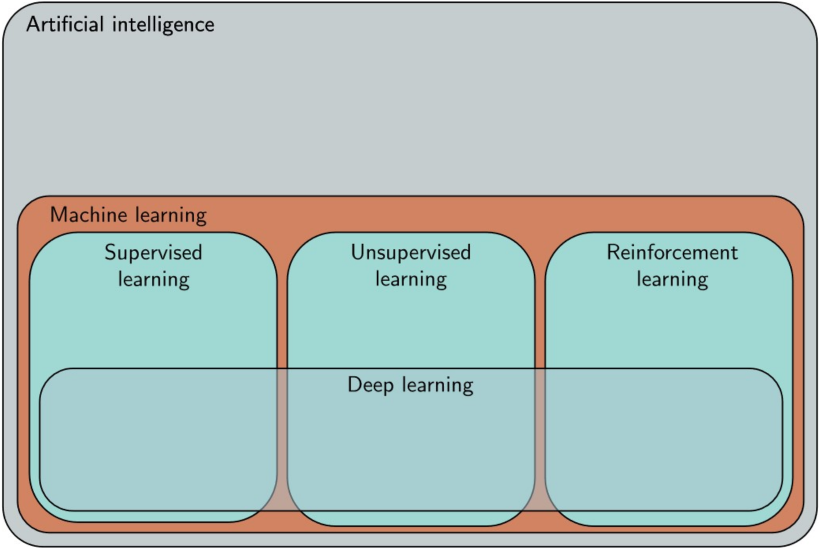
Faculty of Sciences of Tetouan

Introduction to Deep Learning

1. Introduction

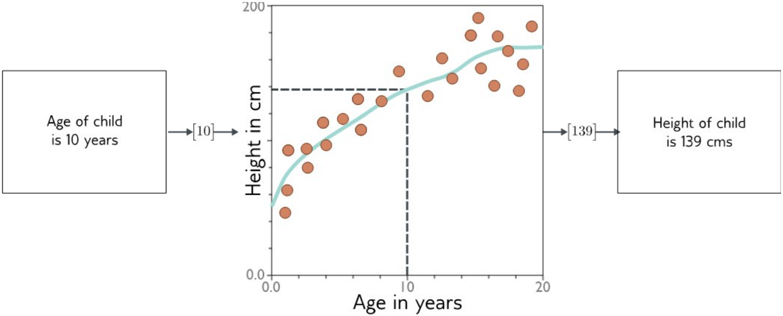
Prof. Monir EL ANNAS

Classical Machine Learning vs. Deep Learning



Types of Learning

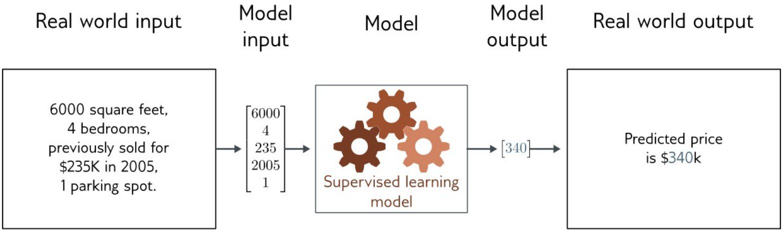
Supervised learning case



- Define a mapping from input to output
- Learn this mapping from paired input/output data examples
- Deep neural networks are just a very flexible family of equations that can represent an extremely broad family of relationships between input and output

Types of Learning

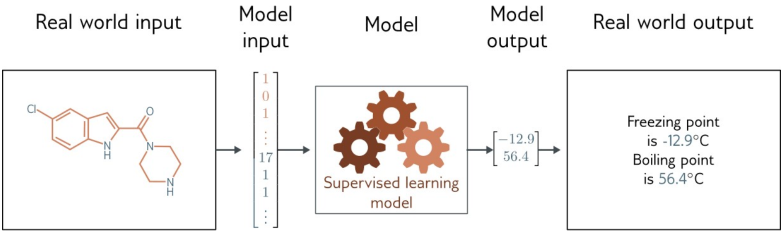
Supervised learning case: Regression



- Univariate regression problem (one output, real value)
- Fully connected network

Types of Learning

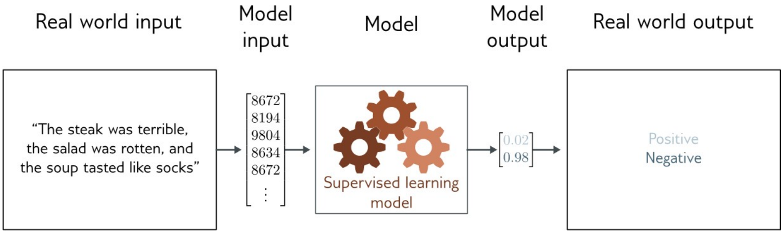
Supervised learning case: Graph regression



- Multivariate regression problem (>1 output, real value)
- Graph neural network

Types of Learning

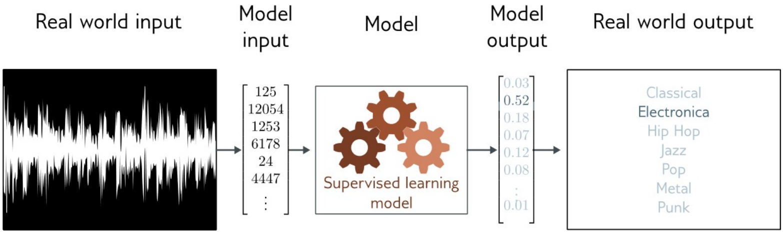
Supervised learning case: Binary classification problem



- Binary classification problem (two discrete classes)
- Transformer network

Types of Learning

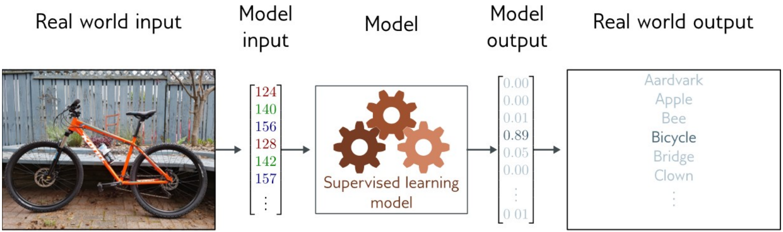
Supervised learning case: Multiclass classification problem



- Multiclass classification problem (discrete classes, >2 possible values)
- Recurrent neural network (RNN)

Types of Learning

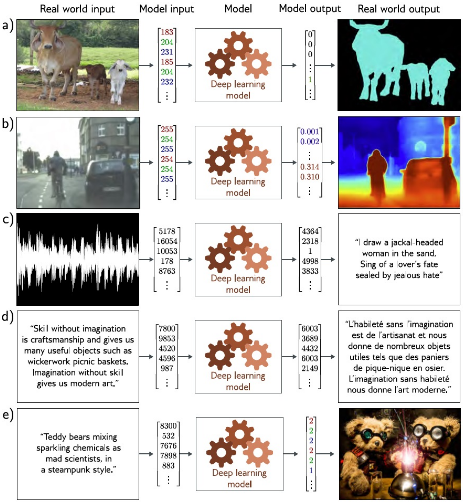
Supervised learning case : Multiclass classification problem



- Multiclass classification problem (discrete classes, >2 possible classes)
- Convolutional network

Types of Learning

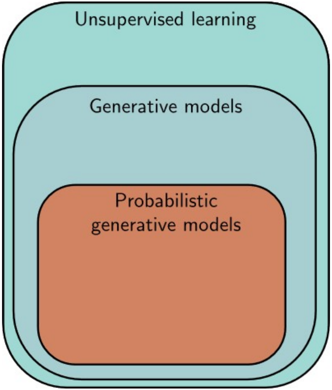
Supervised learning case: More complex examples



Types of Learning

Unsupervised Learning case

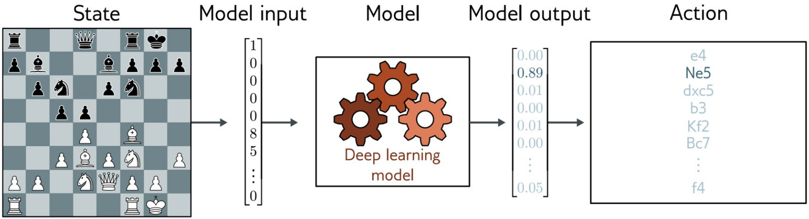
- Learning about a dataset without labels (e.g., clustering)
- Generative models can create examples (e.g., generative adversarial networks)
- PGMs learn distribution over data (e.g., diffusion models)



Types of Learning

Reinforcement learning case

- A set of states: States are valid states of the chess board
- A set of actions: Actions at a given time are valid possible moves
- A set of rewards: Positive rewards for taking pieces, negative rewards for losing them
- Goal: take actions to change the state so that you receive rewards



History and Present of (Deep) Neural Networks

Why now

I. Big Data

- Larger Datasets
- Easier Collection & Storage

II. Hardware

- Graphics Processing Units (GPUs)
- Massively Parallelizable

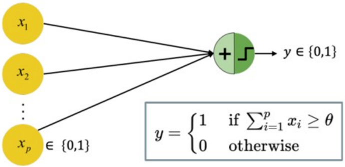
III. Software

- Improved Techniques
- New Models
- Toolboxes

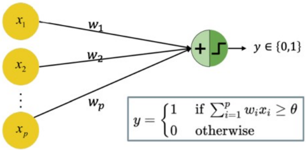
History and Present of (Deep) Neural Networks

The Beginning

- **1943:** the first artificial neuron, the "Threshold Logic Unit (TLU)", was proposed by Warren McCulloch & Walter Pitts
- The model is limited to binary inputs
- It fires/outputs $+1$ if the input exceeds a certain threshold θ
- The weights are not adjustable, so learning could only be achieved by changing the threshold θ



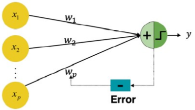
- **1957:** the perceptron was invented by Frank Rosenblatt
- The inputs are not restricted to be binary
- The weights are adjustable and can be learned by learning algorithms
- As for the TLU, the threshold is adjustable based on the classification result and decision boundaries are linear



History and Present of (Deep) Neural Networks

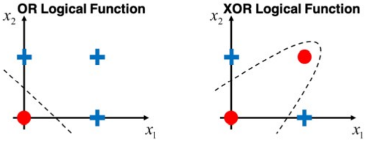
First AI Winter

- 1960: Adaptive Linear Neuron (ADALINE) was invented by Bernard Widrow & Ted Hoff; weights are now adjustable according to the weighted sum of input, yielding a numeric error instead of just misclassification.
- 1965: group method of data handling (also known as polynomial neural networks) by Alexey Ivakhnenko. The first learning algorithms for supervised deep feedforward multilayer perceptrons.



- 1969: the first "AI Winter" kicked in.

  Marvin Minsky & Seymour Papert proved that a perceptron cannot solve the XOR-Problem (linear separability)
  Less funding led to standstill in AI / DL research



History and Present of (Deep) Neural Networks

Second AI Winter

1985: Multilayer perceptron with backpropagation

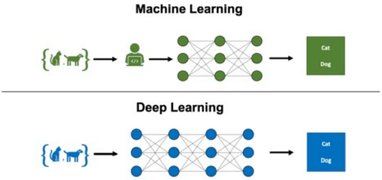
- Invented by David Rumelhart, Geoffrey Hinton, and Ronald Williams
- Efficiently compute derivatives of composite functions
- Backpropagation was developed already in 1970 by Linnainmaa

1985: The second "AI Winter"

- Overly optimistic expectations concerning potential of AI / DL
- The phrase "AI" reached a pseudoscience status
- Kernel machines and graphical models both achieved good results on many important tasks
- Some fundamental mathematical difficulties in modeling long sequences were identified.

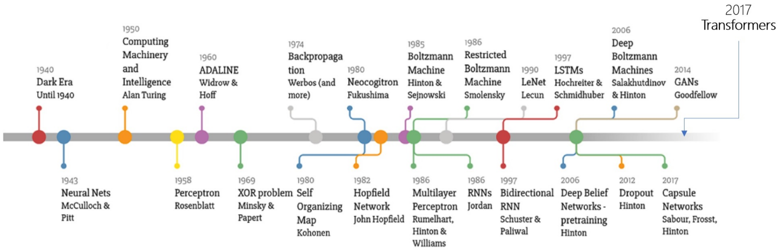
2006: Age of deep neural networks began

- Geoffrey Hinton showed that a deep belief network could be efficiently trained using greedy layer-wise pretraining.
- This wave of research popularized the term deep learning to emphasize that researchers were now able to train deeper neural networks than had been possible before.
- At this time, deep neural networks outperformed competing AI systems based on other ML technologies as well as hand-designed functionality.



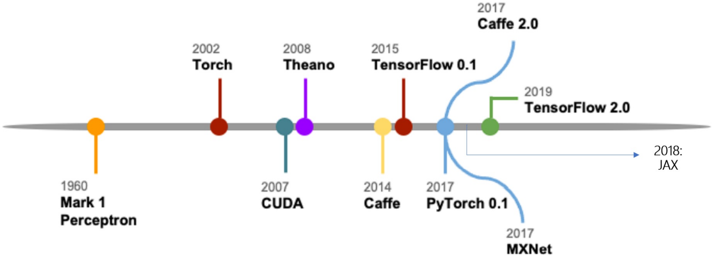
History and Present of (Deep) Neural Networks

Timeline of Algorithms



History and Present of (Deep) Neural Networks

Timeline of Tools



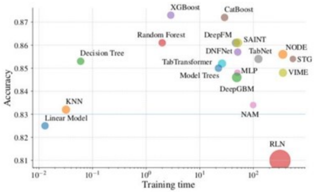
When is Deep Learning Useful? (1/3)

Deep learning can be extremely valuable if the data has these properties:

- It is high dimensional.
- Each single feature itself is not very informative but only a combination of them is.
- Large amounts of training data are available.

For tabular data, deep learning is therefore rarely the correct model choice.

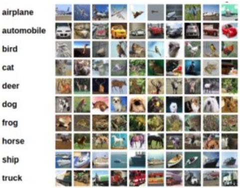
- Without extensive tuning, models like random forests or gradient boosting will outperform deep learning most of the time.
- One exception is data with categorical features with many levels.



Borisov, V. et al. Deep neural networks and tabular data: A survey. arXiv [cs.LG] (2021)

When is Deep Learning Useful? (2/3)

- One promising use case for deep learning are tasks based on images as they are characterized by:
  - High dimensionality: a color image with 255×255 (3 Colors) pixels already has 195075 features.
  - Informativeness: a single pixel is not meaningful but only a combination of pixels is.
  - Training data: depending on the desired application, huge amounts of data are available.

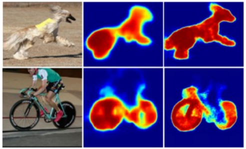


Credit: Alex Krizhevsky (2009)

Image classification: predict a single label for each image

Credit: Kaiming He (2017)

Object detection: generate bounding boxes for each instance



Credit: Hyeonwoo Noh (2015)

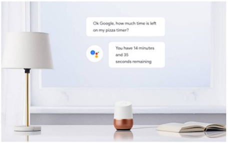
Instance segmentation: partition the image into segments

When is Deep Learning Useful? (3/3)

- Another promising use case for deep learning are tasks based on text as it is characterized by:
  - High dimensionality: each word can be a single feature.
  - Informativeness: a single word does not provide much context.
  - Training data: huge amounts of text data available.

Sentiment analysis: systematically identify the emotional and subjective information in texts

Machine translation: predict likelihood of a sequence of words, typically modeling entire sentences in a single integrated model



Speech recognition & generation: Extract features from audio data for downstream tasks, e.g. to classify emotions in speech