Joint Faculty Development Programme, Winter 2019
Deep Learning and Applications, December 9-13, 2019

Introduction to Deep Learning
Artificial Intelligence in the New Era

Aparajita Ojha
Professor, Computer Science and Engineering
PDPM Indian Institute of Information Technology,
Design and Manufacturing, Jabalpur

A Joint Programme by IIITDM Jabalpur, NIT Warangal, MNIT Jaipur and NIT Patna

Evolution of Artificial Intelligence

- Artificial Intelligence is a branch of computer science dealing with the simulation of intelligent behavior in computers. --Merriam Webster Dictionary

- The term Artificial Intelligence began to evolve when a small group of scholars from a variety of fields in 1940s-50s started exploring the possibility of **creating an artificial brain**.

Evolution of Artificial Intelligence

Cybernetics and early neural networks [1940-50s]

Turing's test [1950s]

Nobert Wiener

Dartmouth Conference 1956

AI Emerges

John McCarthy

Alan Turing

Evolution of Artificial Intelligence

Since then, AI has been witnessing rises and falls

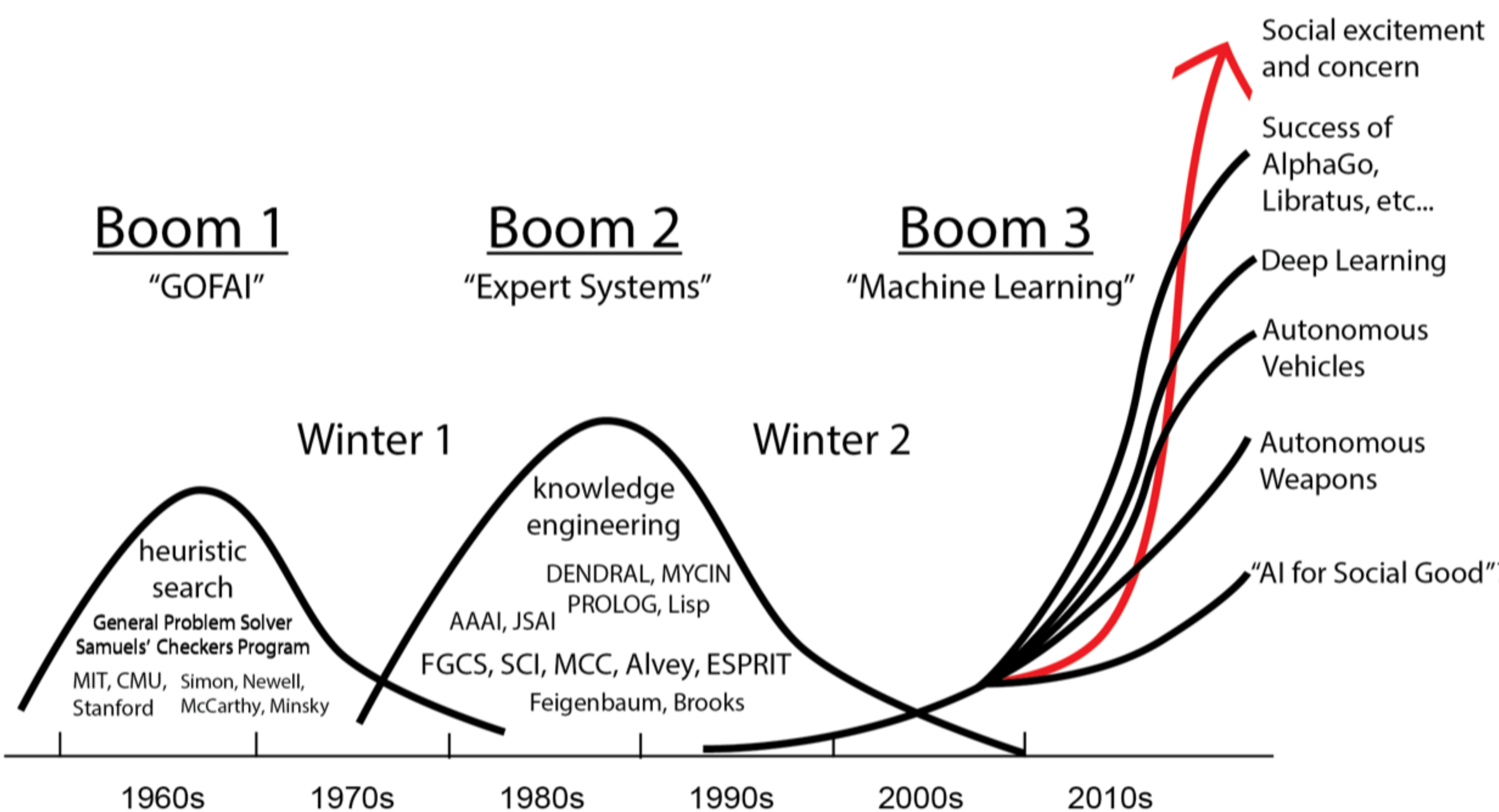
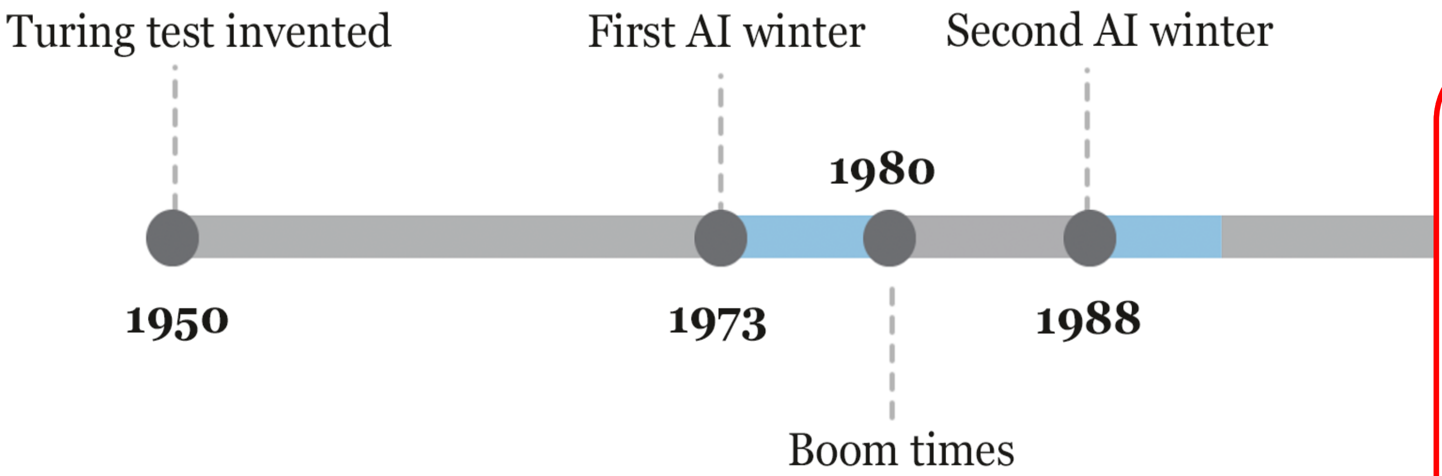


Image source: https://www.technologystories.org/ai-evolution/

Evolution of Artificial Intelligence

Winters in the age of Artificial Intelligence and current resurgence with deep learning



Deep Learning brought the AI to the forefront again

First AI Winter 1973 : The Lighthill report, effect on funding

Second AI Winter 1987: market for specialized AI hardware collapsed.

[Image source: https://towardsdatascience.com/](https://towardsdatascience.com/)

The Era of Deep Learning

Google's AlphaGo AI wins three-match series against the world's best Go player

Posted May 24, 2017 by Jon Russell (@jonrussell)

DESCRIPTION

The first product, GO, offers GPS based, instant mobile broadcasting. Let the world see life how do. Use photo and video to catalogue

WEBSITE

Google's AlphaGo AI has once again made the case that machines are now smarter than

Deep Learning

- The difference in artificial intelligence approaches over the two decades (1997-2017)
  - 1997: The IBM chess computer Deep Blue, was explicitly programmed to win against the grandmaster Garry Kasparov in 1997
  - **2017: AlphaGo was not preprogrammed to play Go.**
  - It learned using a general-purpose algorithm that allowed it to interpret the game's patterns.
- AlphaGo program applied **deep learning**.

Deep Learning Impact

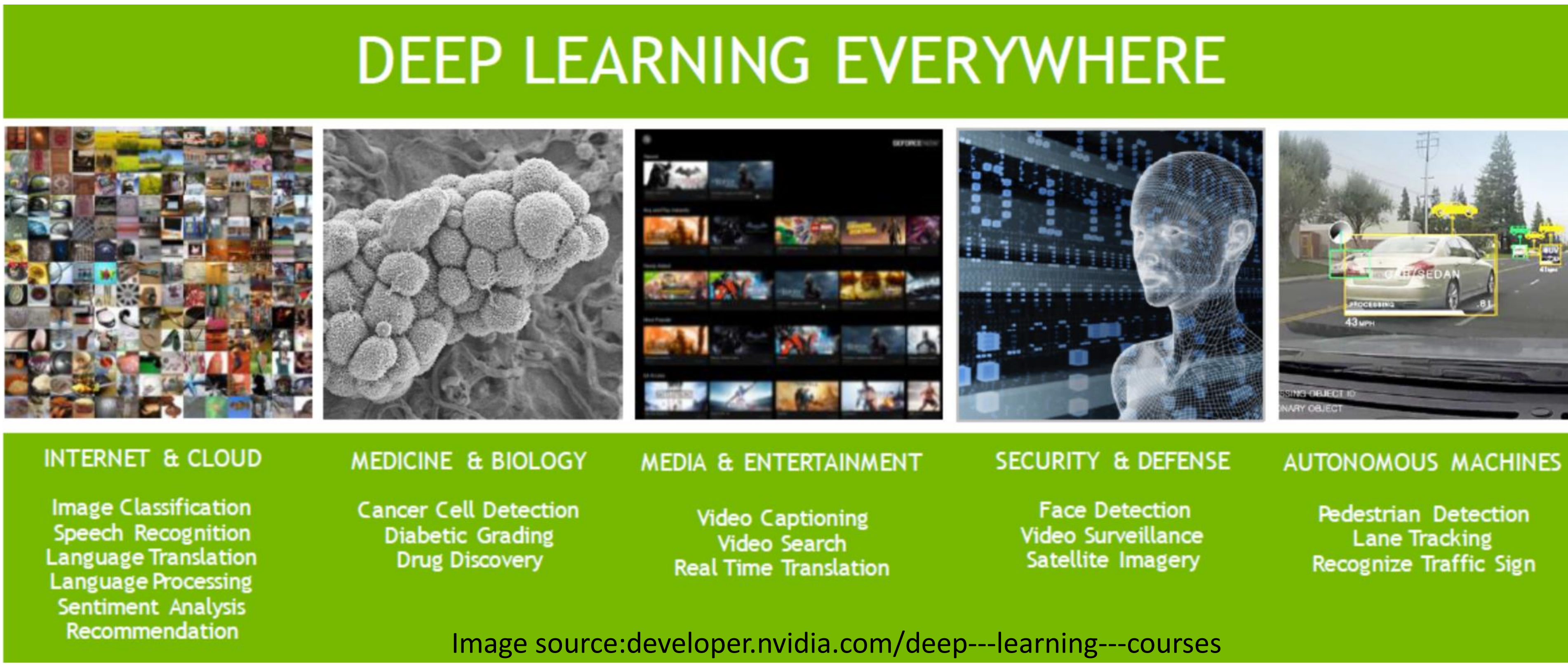


Image source:developer.nvidia.com/deep---learning---courses

Artificial intelligence has penetrated practically every aspect of human life.

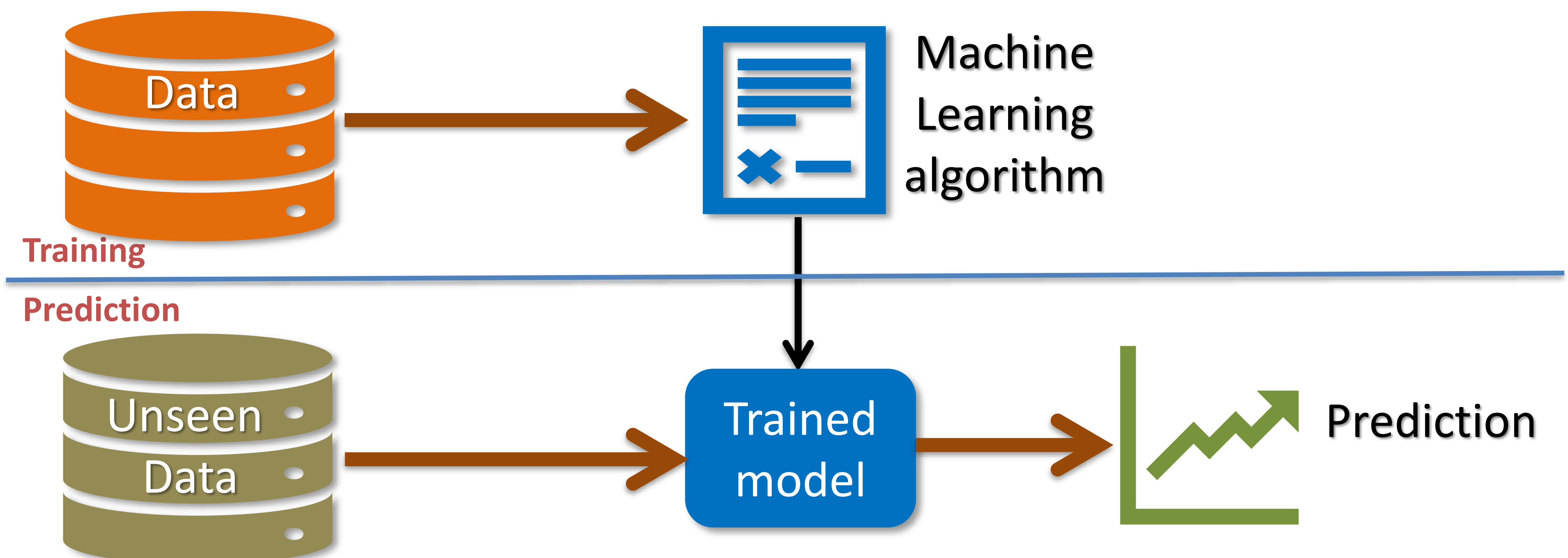
What is Deep Learning ?

- A machine learning technique that allows computers learn and improve their performance with **experience** and **data**.
- Achieves great power and flexibility by learning to represent the world as a nested hierarchy of concepts.
- Each concept defined in relation to simpler concepts.
- More abstract representations computed in terms of less abstract ones.

Source: Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, MIT Press, 2017

Machine Learning

- Machine learning (ML) is a field of computer science that gives computers the ability to **automatically learn without being explicitly programmed.**
- Learning from experience on data to make predictions.



Machine Learning Algorithm

- A machine learning algorithm is an algorithm that is able to learn and **extract patterns** from data.
- **Learning** –
  – A computer program 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.” (*Mitchell , 1997*)

Learning Algorithms…

- General Tasks
  - Classification, Regression, Transcription , Machine Translation etc.
- Performance measures
  - Depends on the type of problem: Examples include –
    - accuracy, error rate etc.
  - Performance is measured on a dataset called test dataset, that is different from the dataset used to train the algorithms.
  - Often difficult to choose a performance measure that corresponds well to the desired behavior of the system.
- Experience
  - Algorithms are termed as supervised learning or unsupervised learning algorithms based on the experience they are allowed to have on datasets.

Learning Algorithms…

- Three broad categories of learning

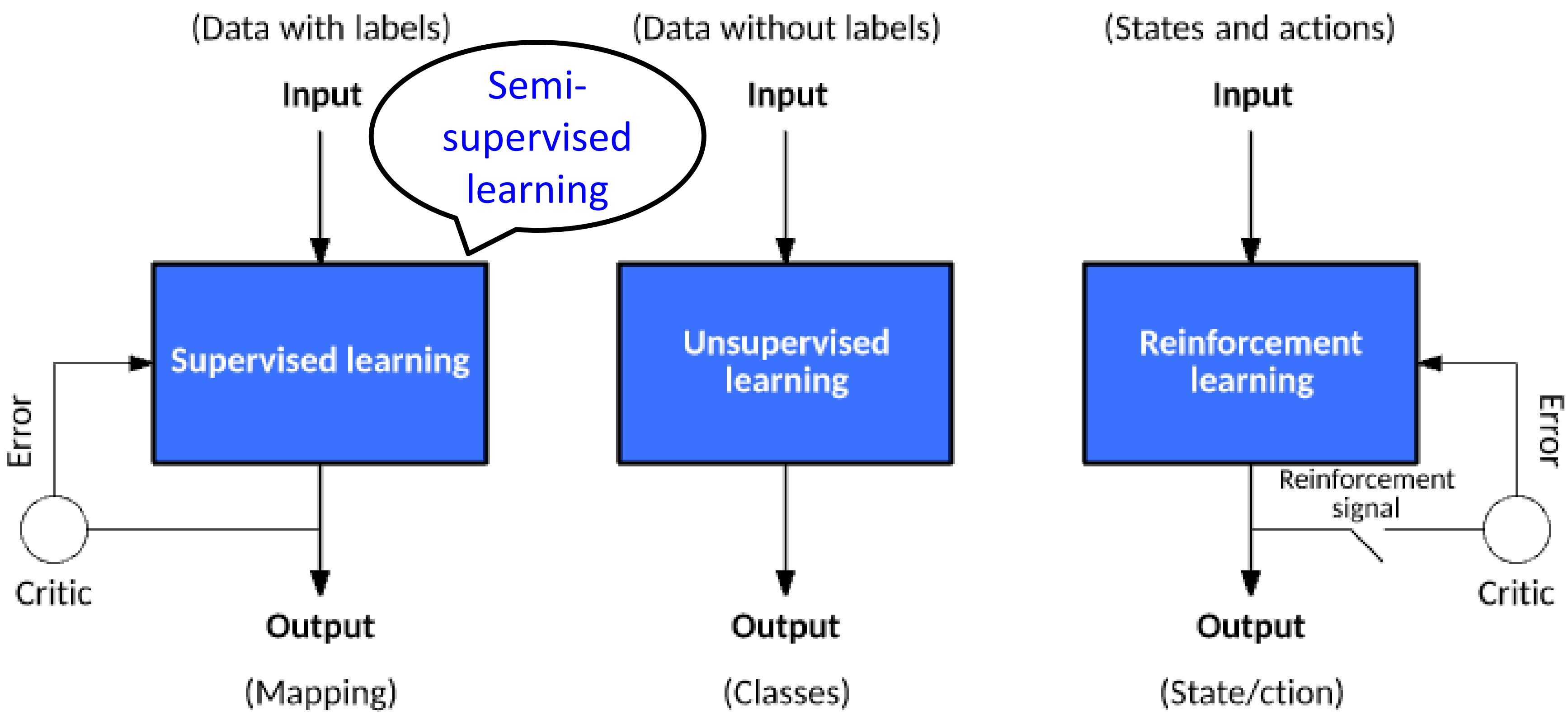
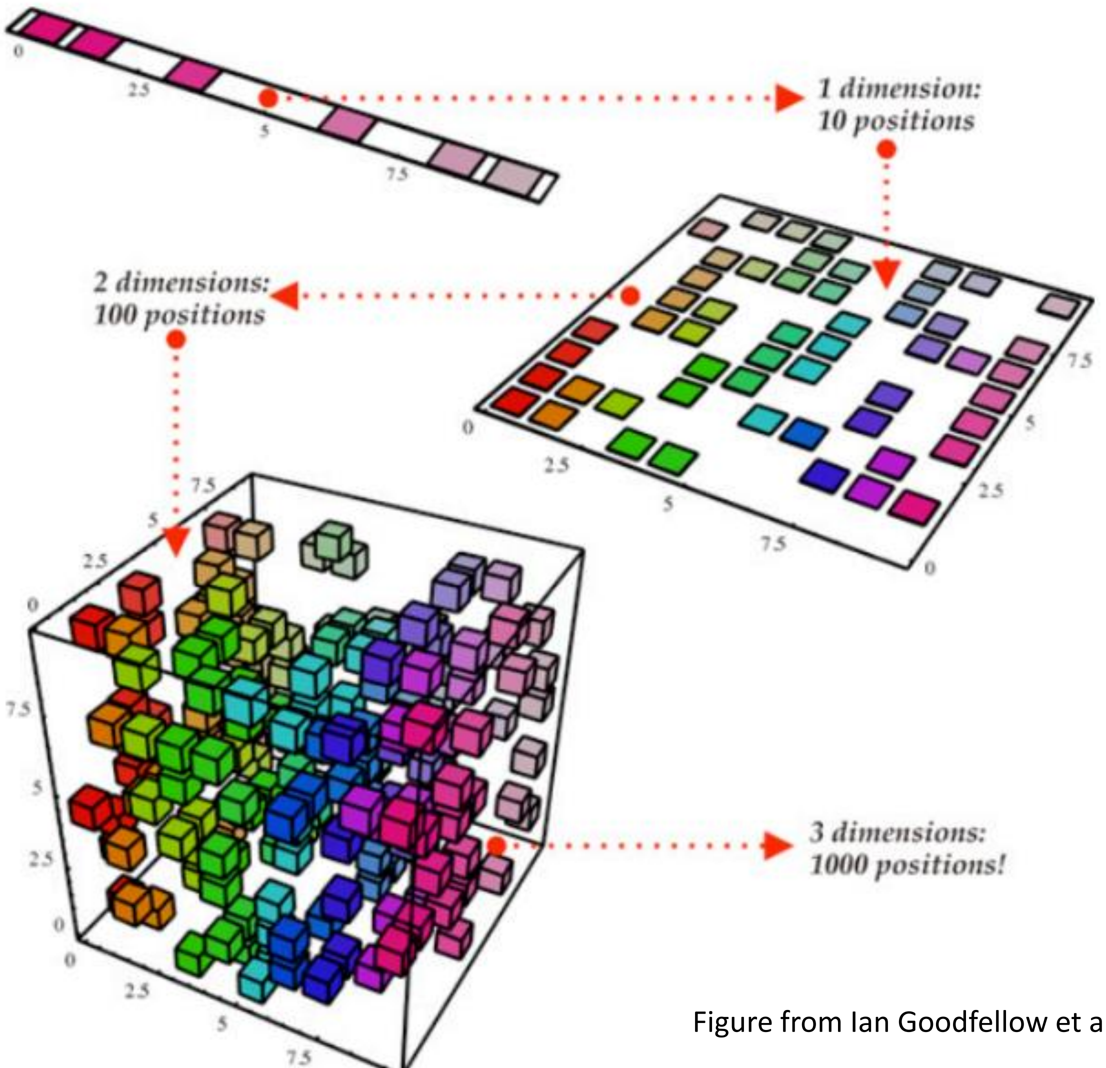


Image source: Unknown

Limitations of Traditional ML Systems



Curse of dimensionality

Many machine learning problems become exceedingly difficult when the number of dimensions in the data is high.

(Hughes/Peaking phenomenon)

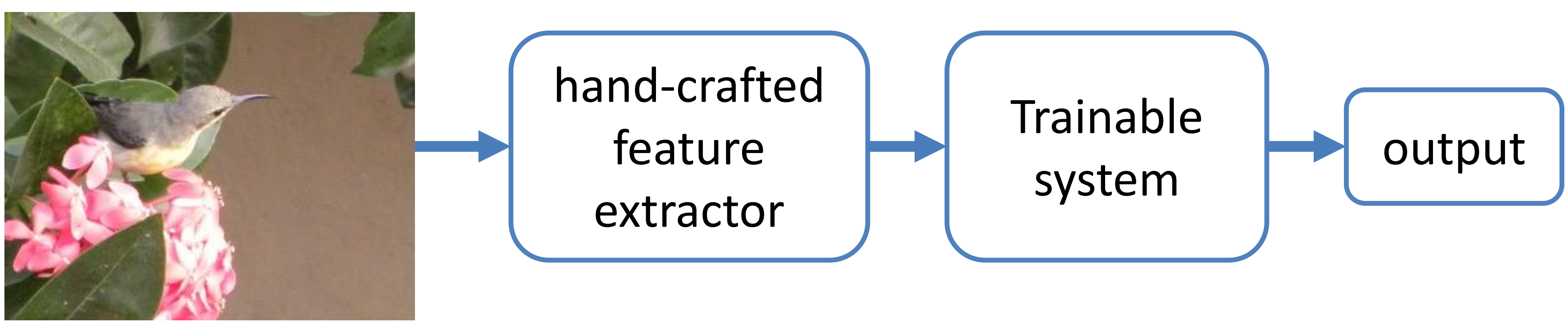
Figure from Ian Goodfellow et al. Book

Curse of Dimensionality

- Hughes phenomenon (or peaking phenomenon)
- With a fixed number of training samples, the predictive power of a classifier or regressor first increases as the number of dimensions/features used is increased but then decreases.
- Examples of ML problems with high dimensional data
  - Speech recognition
  - Object detection in images

Limitations of Traditional ML Systems…

- Traditional machine learning uses manually designed features that are
  - Often over-specified and incomplete.
  - Mostly application-dependent, cannot be generalized easily to other applications.
  - Time consuming in design and validation.
- Developing hand-crafted features is a costly and complex job.



ML vs Deep Learning

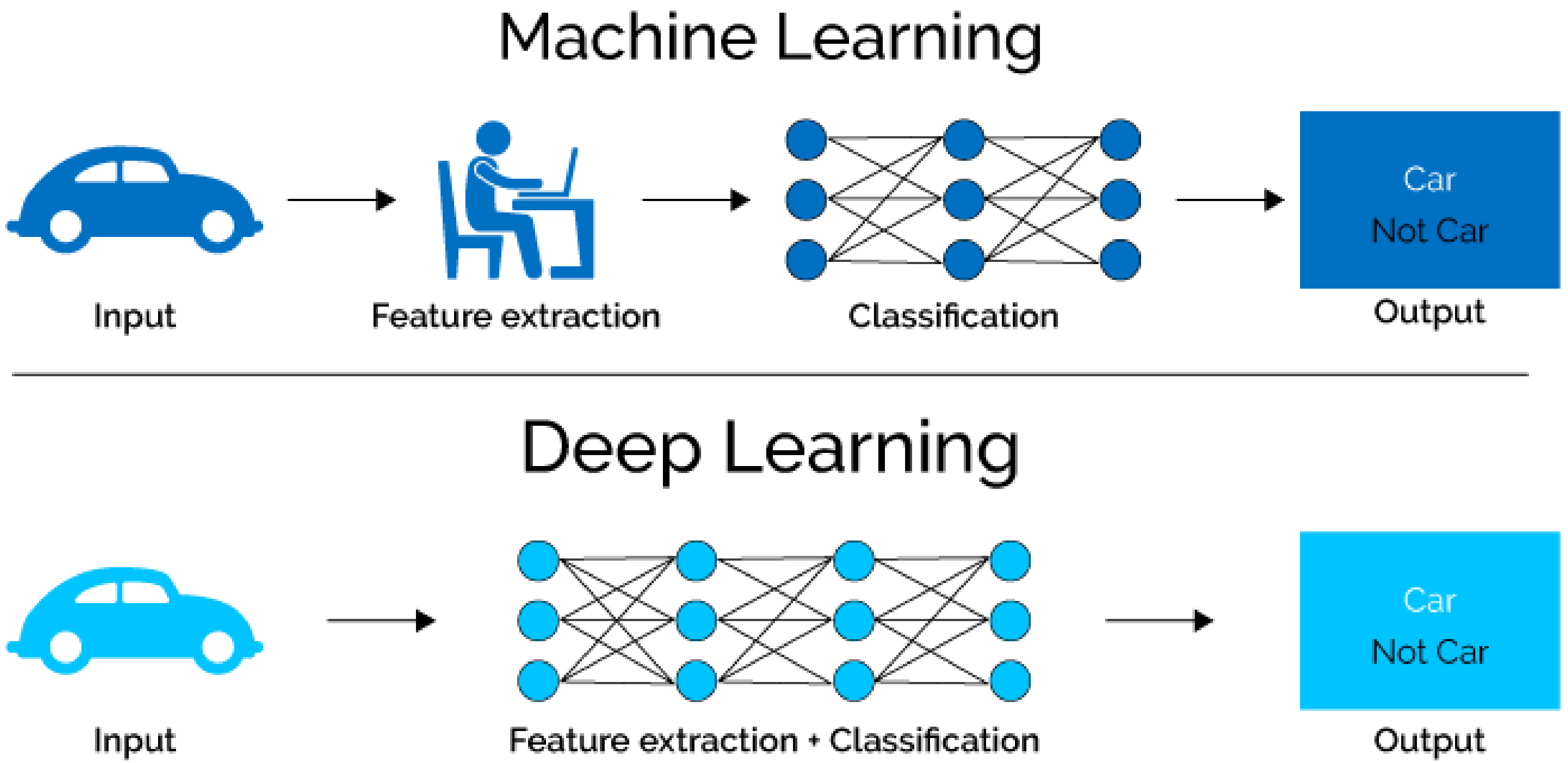


Image source: Unknown

Deep Learning

- Deep learning learns rich hierarchical representations (features) automatically by extracting features at different levels.
- Provides a **flexible** learnable framework.
- Trained algorithms work for **a broad range of problems**.

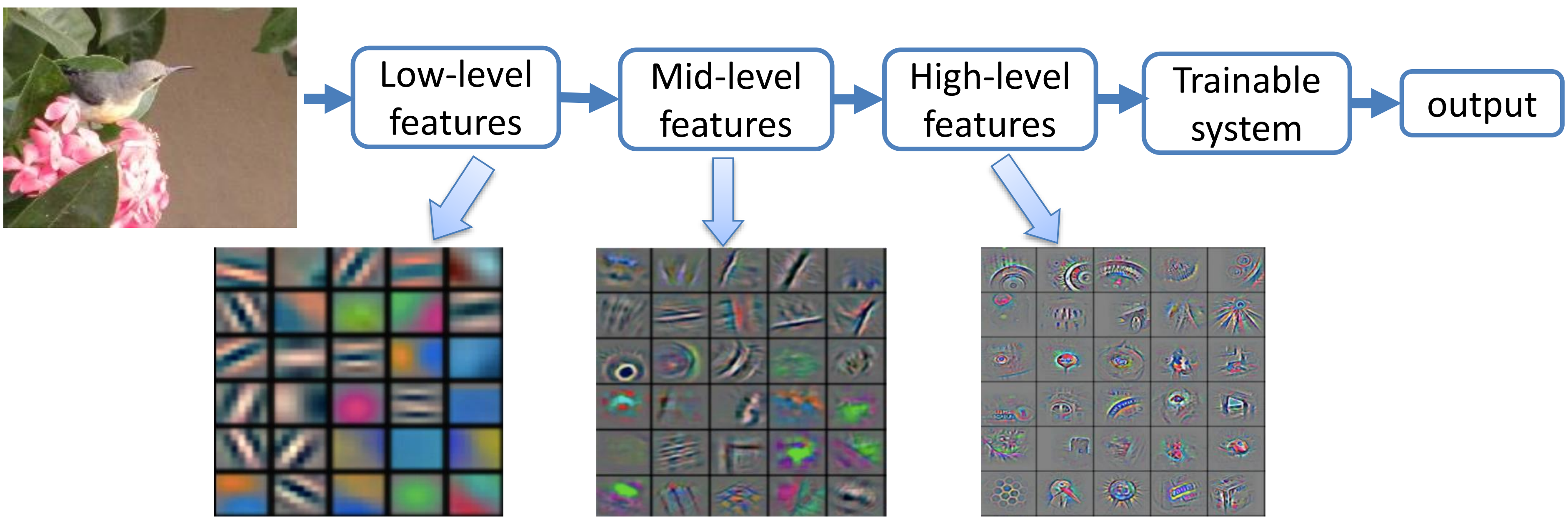


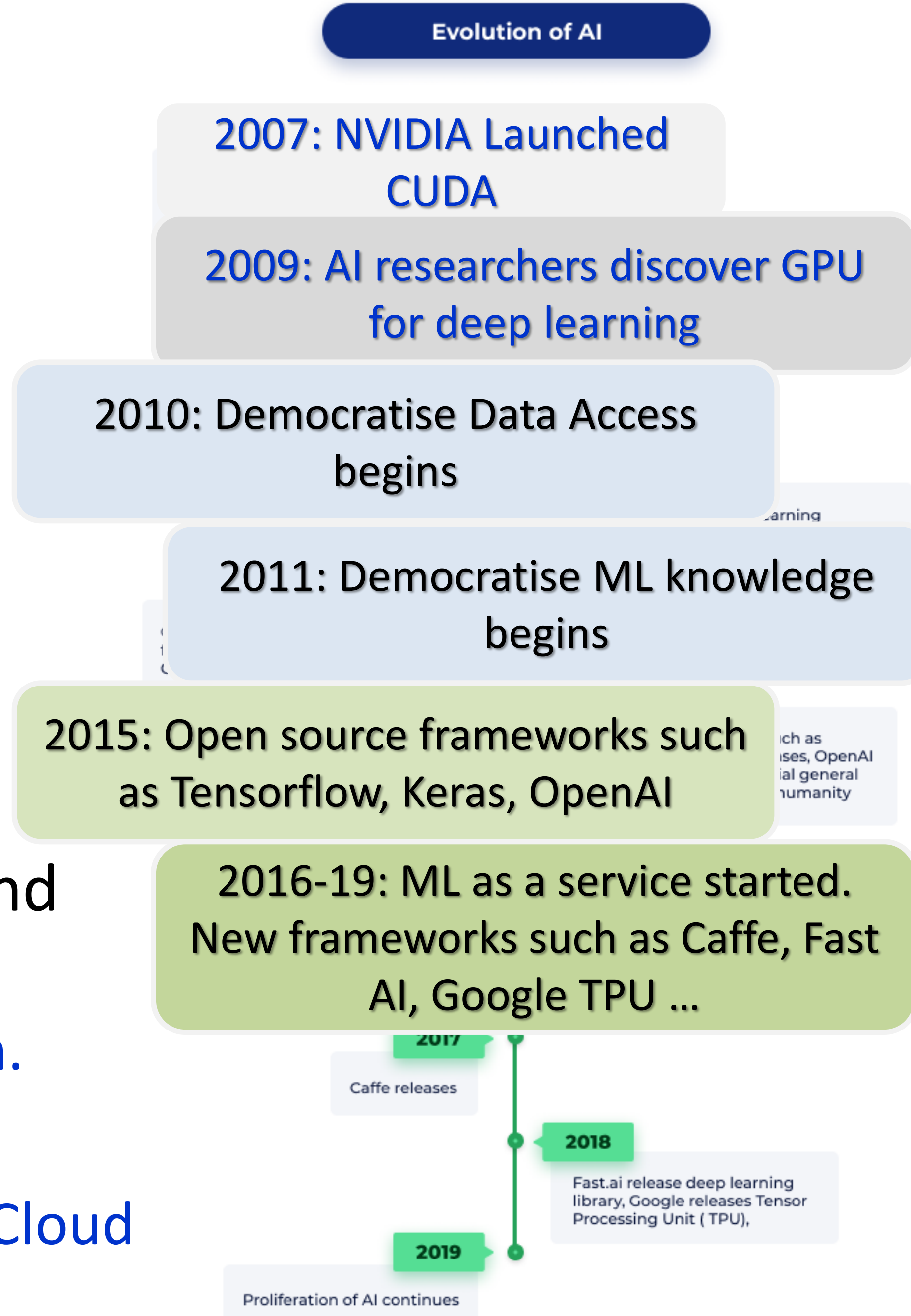
Figure source: Zeiler and Fergus, 2013: Feature visualization of convolution NN on ImageNet database.

Deep Learning Success Mantra

- Increased availability of data
- Significant improvement in computing power
- Increased model size
- **New way of learning representations** (auto-feature extraction)
- Use of Deep Neural Network Architectures that generalize well and are flexible enough to address a broad range of tasks.

AI Technology Advancements with Deep Learning

- Four major catalysts
  - Democratization of AI knowledge.
  - Data and Computing Power
    - Cloud and GPU
  - Proliferation of new tools and frameworks
    - Tensorflow, Keras, PyTorch.
  - AI as a service.
    - Azure AI, AWS AI, Google Cloud AI and IBM Cloud AI.



Source: https://www.lftechnology.com/

Deep Learning Ingredients

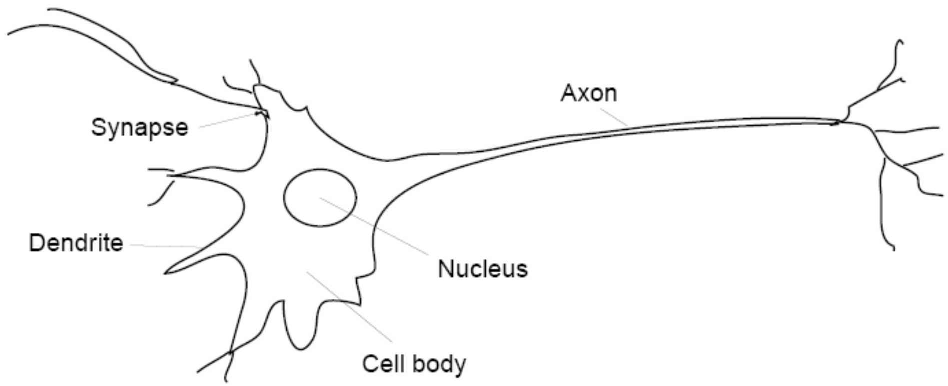
- Artificial Neural Networks (ANN) as a learning machine
- Data
- Algorithms to train ANNs
- Computing resources

Artificial Neural Network (ANN)

- An information processing paradigm inspired by human brain's information processing mechanism.

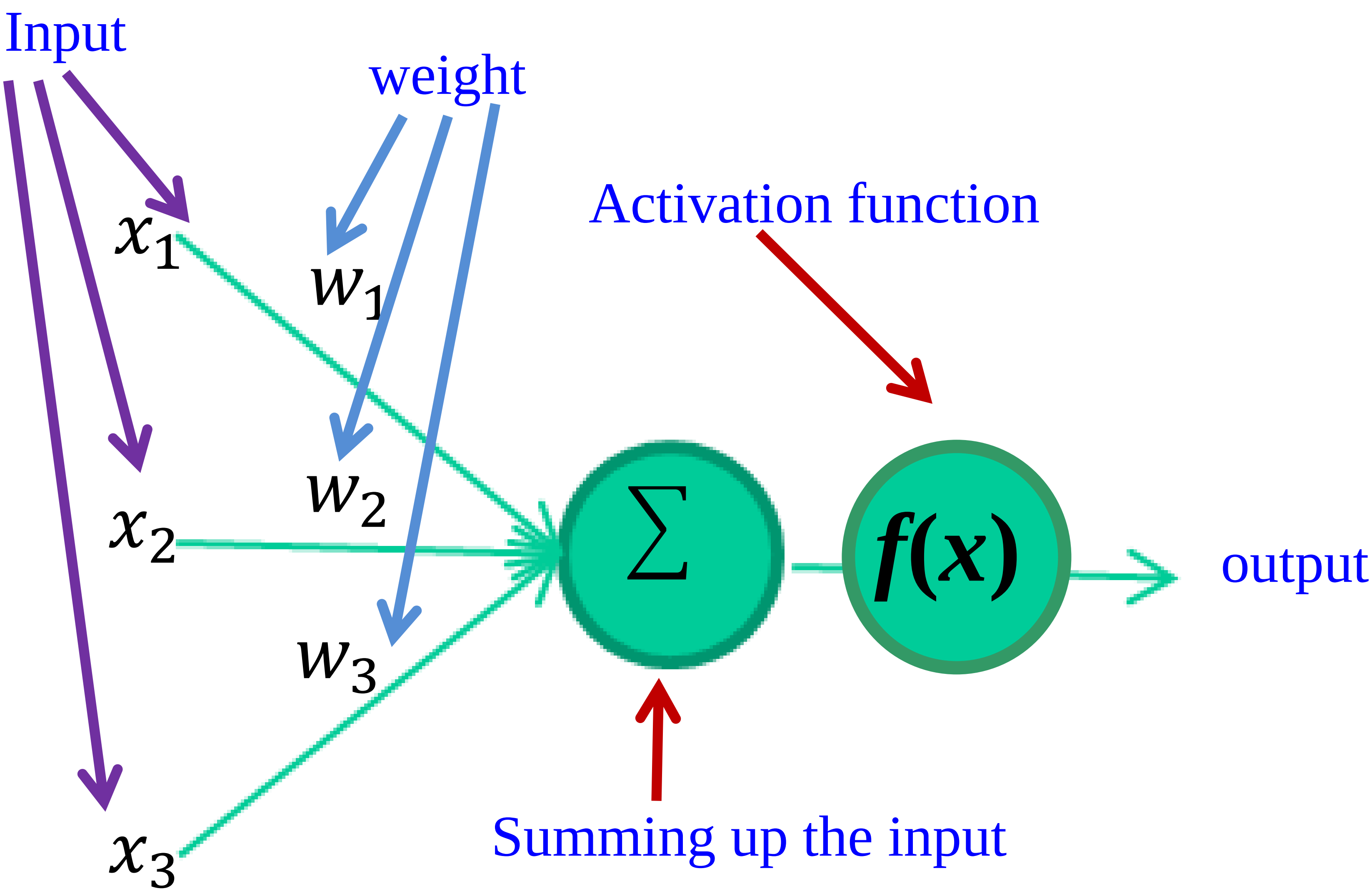
- Composed of a large number of highly interconnected processing elements called neurons working in unison to solve specific problems.
  - ANNs learn by examples like human being.

Neurons in the Brain



- The main purpose of neurons is to receive, analyze and transmit further the information in a form of signals (electric pulses).
- When a neuron sends the information, we say that it “fires”.
- Acting through specialized projections known as dendrites and axons, neurons carry information throughout the neural network.

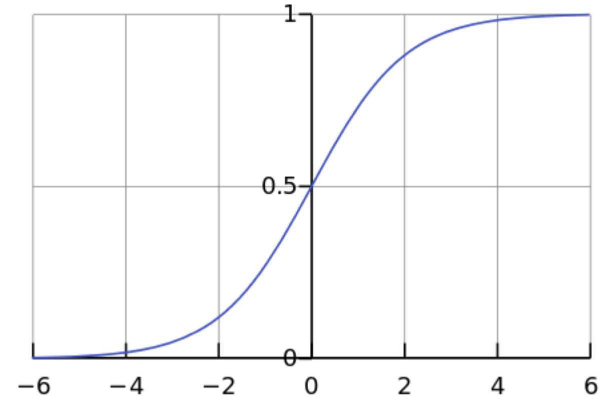
Mathematical Model: Neuron

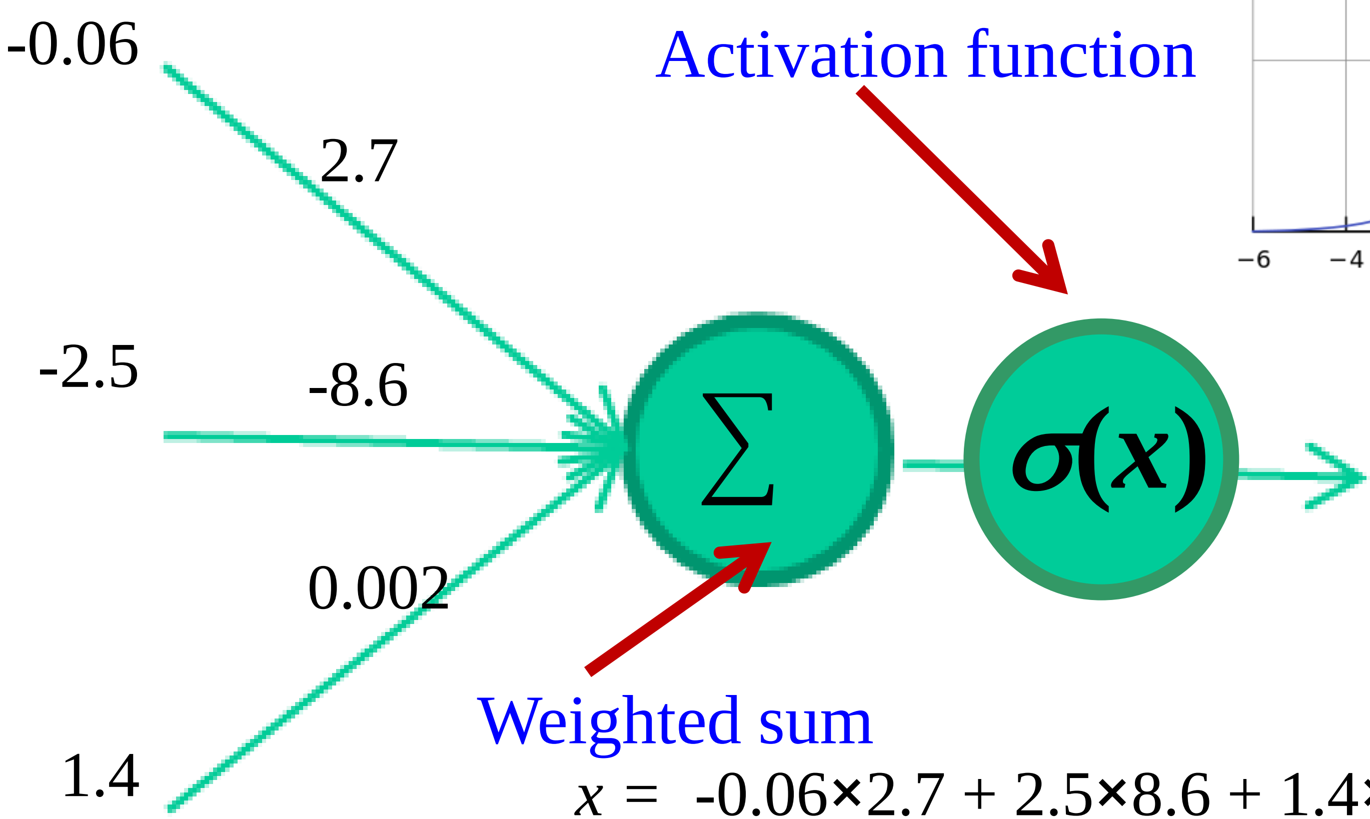


Mathematical Model: Neuron

An Example

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





$$x = -0.06 \times 2.7 + 2.5 \times 8.6 + 1.4 \times 0.002 = 4.0048$$

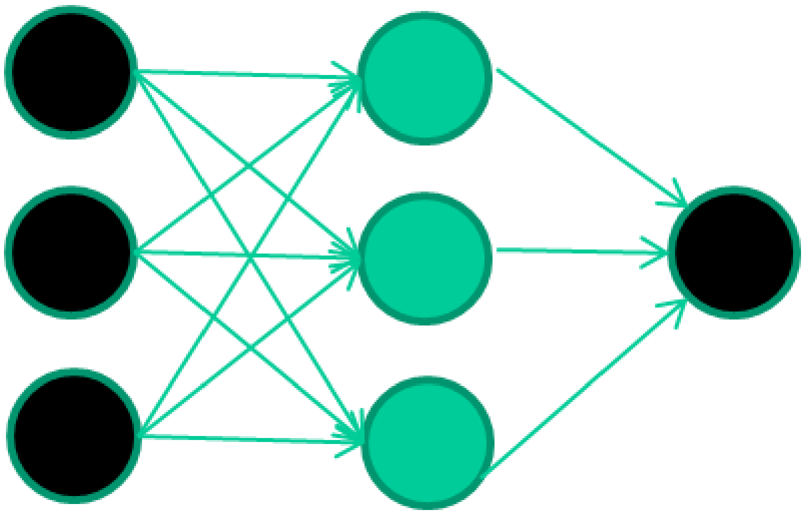
$$\sigma(x) = 0.982$$

Mathematical Model: NN

A dataset

Features			***class***
1.4	2.7	1.9	0
3.8	3.4	3.2	0
6.4	2.8	1.7	1
4.1	0.1	0.2	0

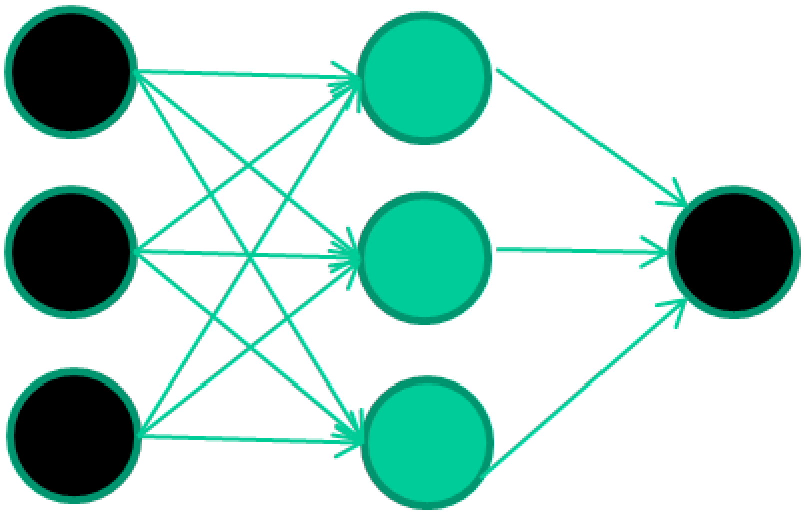
etc …



Mathematical Model: NN…

Train the NN

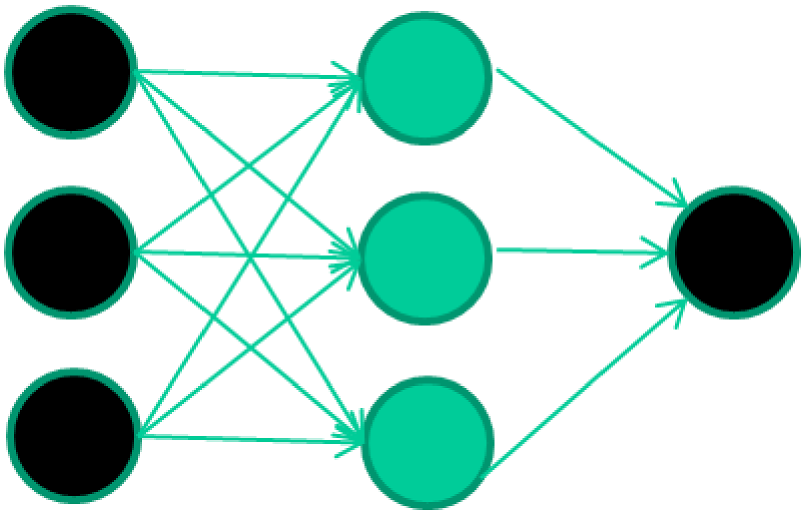
Features			***class***
1.4	2.7	1.9	0
3.8	3.4	3.2	0
6.4	2.8	1.7	1
4.1	0.1	0.2	0
etc …			



Mathematical Model: NN…

Train the NN
Initialise with random weights

Features			***class***
1.4	2.7	1.9	0
3.8	3.4	3.2	0
6.4	2.8	1.7	1
4.1	0.1	0.2	0
etc …			

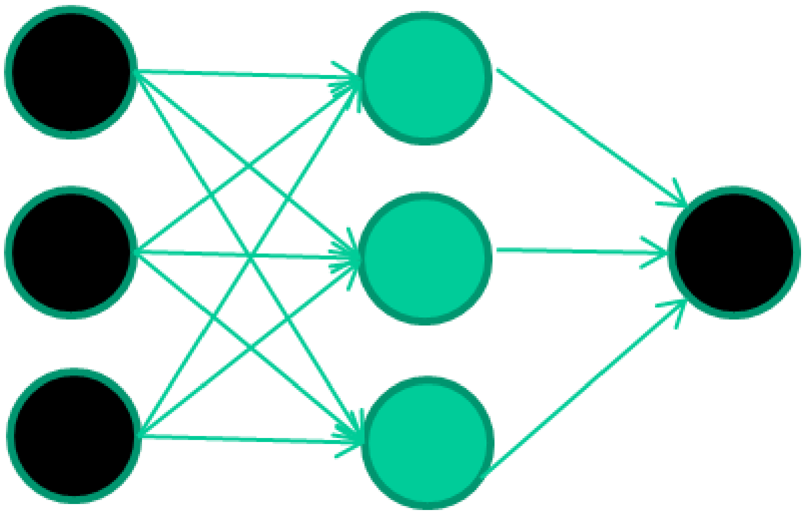


Mathematical Model: NN…

Train the NN
Initialise with random weights
Present a training pattern and feed it to the NN.
Get an output and compare with actual value.

Training Data

Features			*class*
1.4	2.7	1.9	0
3.8	3.4	3.2	0
6.4	2.8	1.7	1
4.1	0.1	0.2	0
etc …			

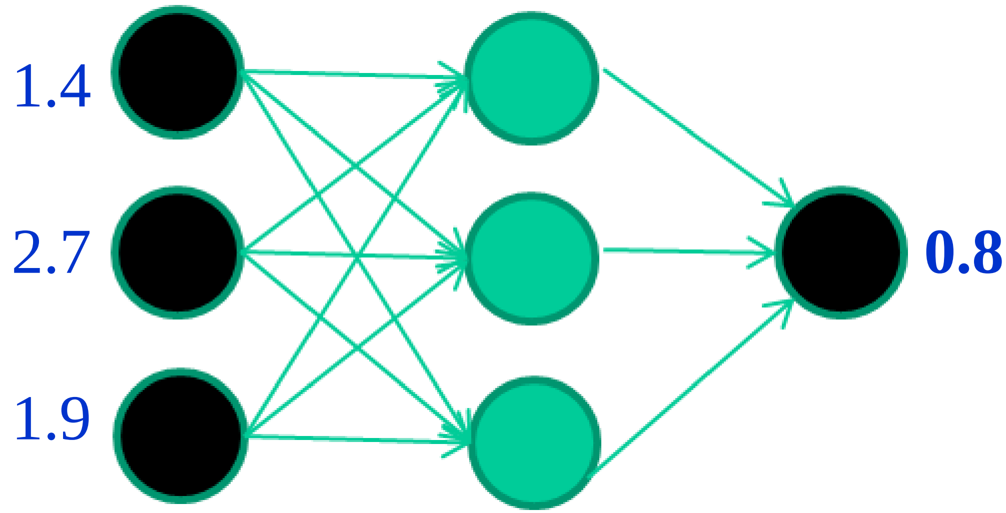


Mathematical Model: NN…

Train the NN
Initialise with random weights
Present a training pattern and feed it to the NN.
Get an output and compare with actual value.

Training Data

Features			***class***
1.4	2.7	1.9	0
3.8	3.4	3.2	0
6.4	2.8	1.7	1
4.1	0.1	0.2	0
etc …			



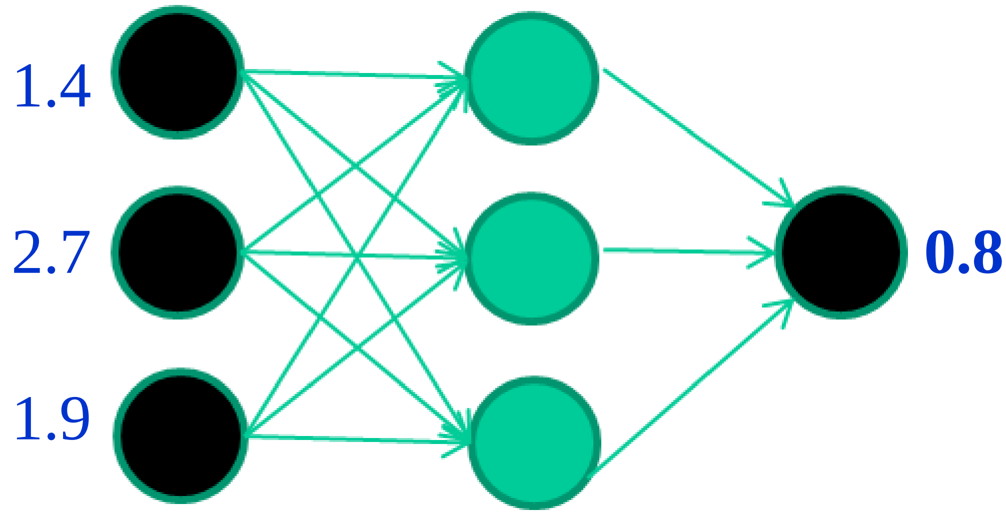
Mathematical Model: NN…

Train the NN
Initialise with random weights
Present a training pattern and feed it to the NN.
Get an output and compare with actual value.

Training Data

Features			*class*
1.4	2.7	1.9	0
3.8	3.4	3.2	0
6.4	2.8	1.7	1
4.1	0.1	0.2	0
etc …			

Expected output



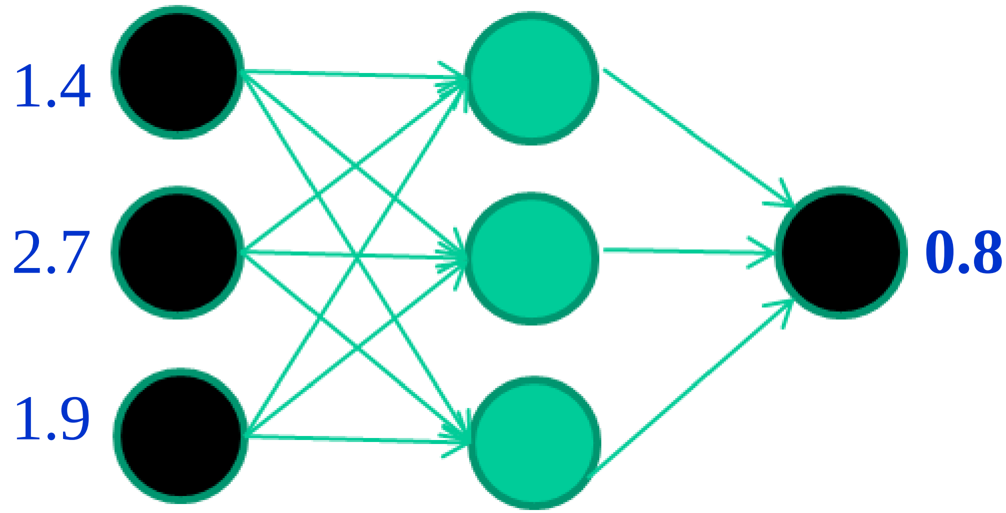
Mathematical Model: NN…

Train the NN
Initialise with random weights
Present a training pattern and feed it to the NN.
Get an output and compare with actual value.

Training Data

Features			***class***
1.4	2.7	1.9	0
3.8	3.4	3.2	0
6.4	2.8	1.7	1
4.1	0.1	0.2	0
etc …			

Expected output



Error = 0.8

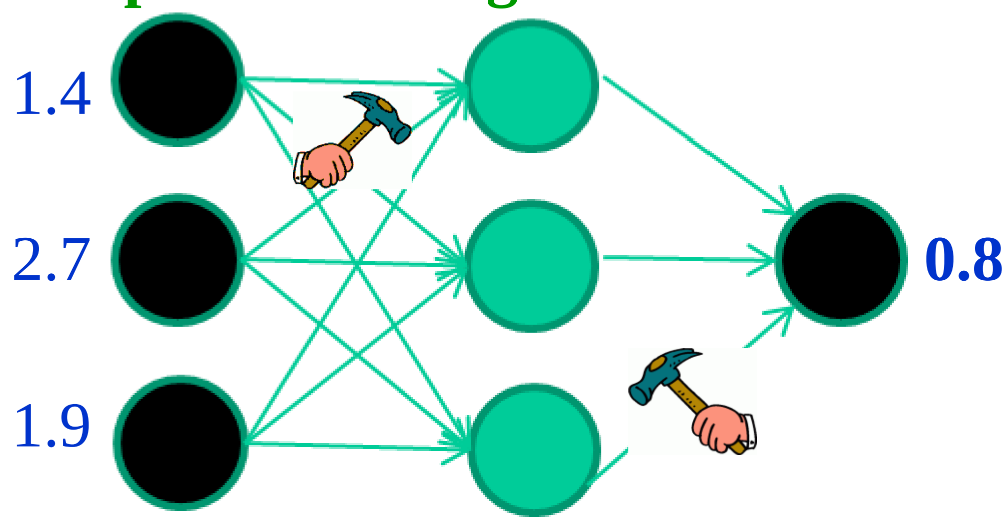
Mathematical Model: NN…

Train the NN
Initialise with random weights
Present a training pattern and feed
Present a training pattern and feed it to the NN.
Get an output and compare with actual value.
Adjust and update the weights.

Training Data

Features			***class***
1.4	2.7	1.9	0
3.8	3.4	3.2	0
6.4	2.8	1.7	1
4.1	0.1	0.2	0
etc …			

Expected output



Error = 0.8

Mathematical Model: NN…

Train the NN

Initialise with random weights

Present a training pattern and feed it to the NN.

Get an output and compare with actual value.

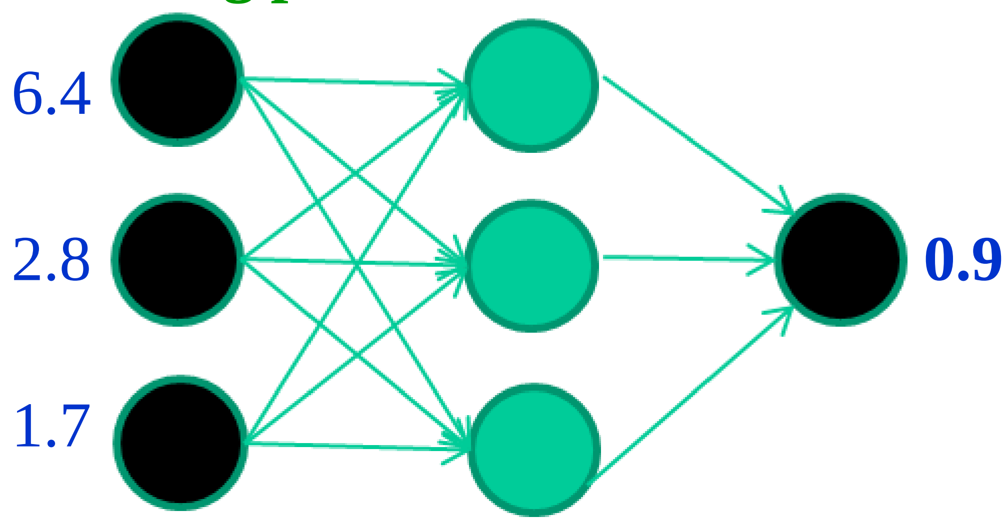
Adjust and update the weights.

Present a training pattern and feed it to the NN

Training Data

Features			***class***
1.4	2.7	1.9	0
3.8	3.4	3.2	0
6.4	2.8	1.7	1
4.1	0.1	0.2	0
etc …			

Training pattern



Error = - 0.1

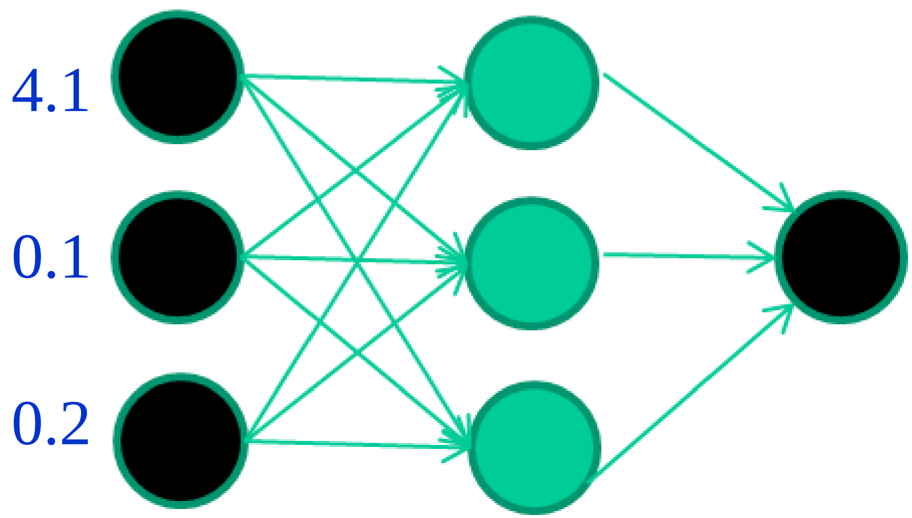
Mathematical Model: NN…

Initialise with random weights
Present a training pattern and feed it to the NN.
Get an output and compare with actual value.
Adjust and update the weights.
Present a training pattern and feed it to the NN.
Repeat the process to reduce the error.

Training Data

Features			***class***
1.4	2.7	1.9	0
3.8	3.4	3.2	0
6.4	2.8	1.7	1
4.1	0.1	0.2	0
etc …			

Next training pattern



Mathematical Model: NN…

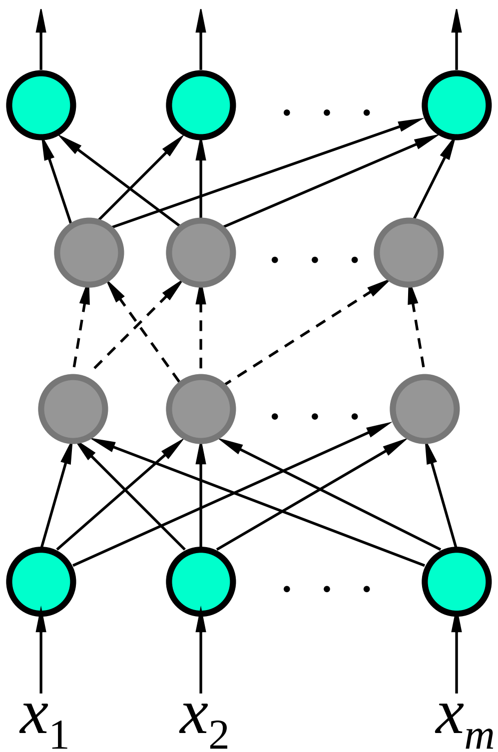
- Weight adjustment process is repeated thousands and thousands of times.
- Each time a random training example is taken, weights are slightly adjusted to tune the system for reducing error of approximation.
- It may not be an efficient adjustment on many other cases.
- But eventually the process of weight adjustments leads to a good enough model for producing an effective classifier.
- It works well in many real applications.

Multilayer Perceptron

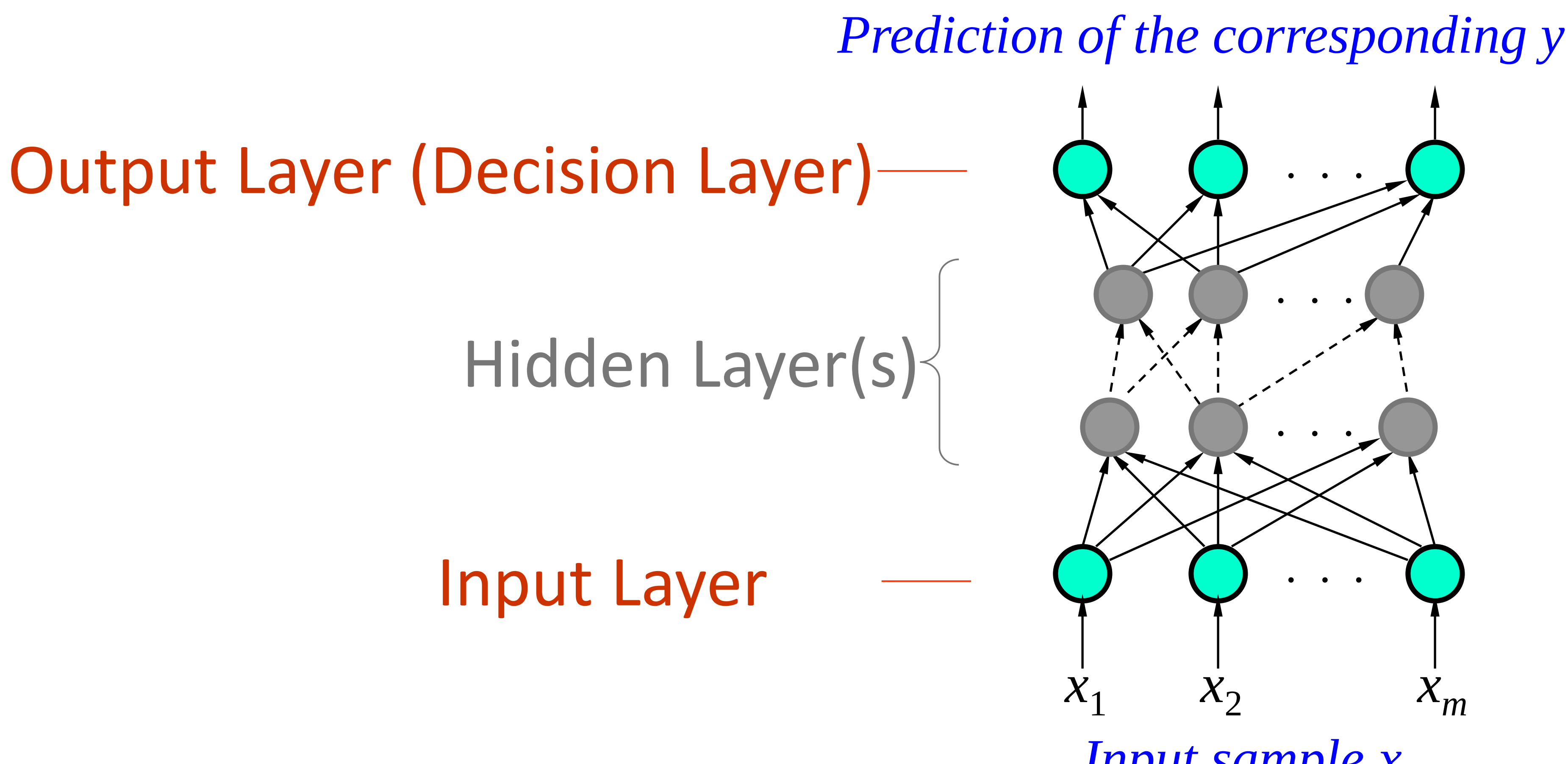
- Universal Approximation Theorem (Cybenko, 1989)
  - Any continuous function over a compact domain can be accurately approximated using a neural network with one hidden layer only.
- If layers increase, accuracy of approximation increases.
- To deal with nonlinearity and for better approximation of more complex functions, multi-layer perceptron models were introduced.
- Let us see how a feed forward NN works with multiple layers of neuron (called units).

Feed Forward NNs

- Graph representation:
  - nodes: neurons
  - arrows: signal flow directions

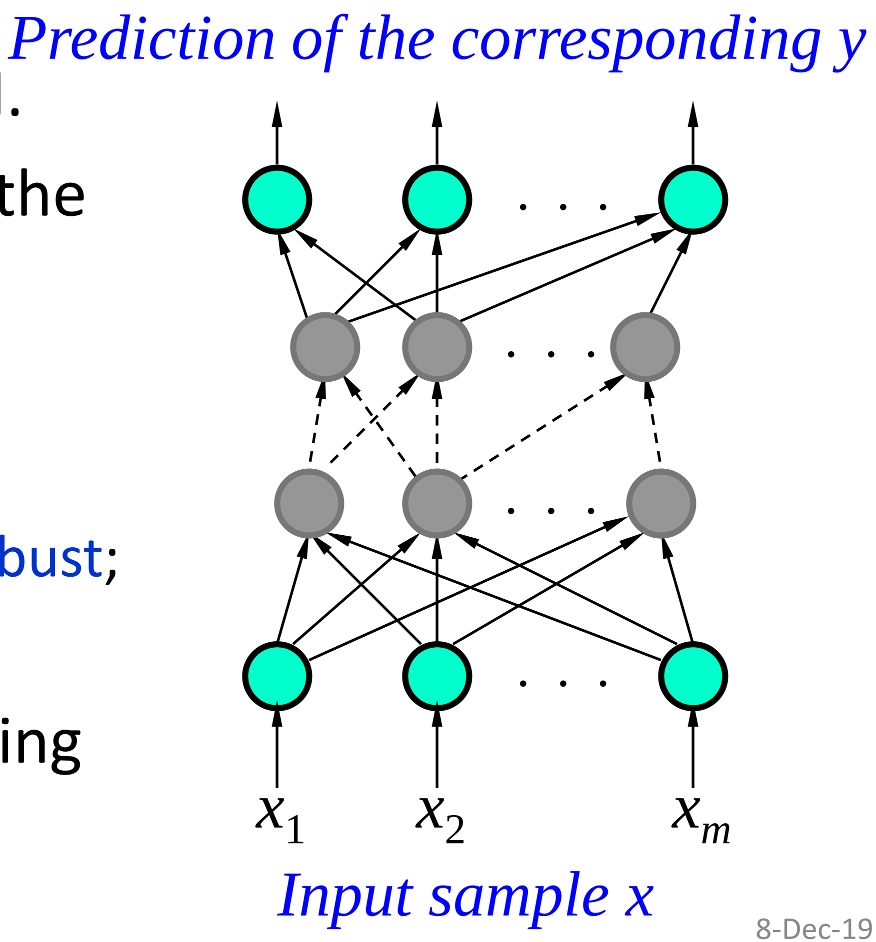


Feed Forward NNs



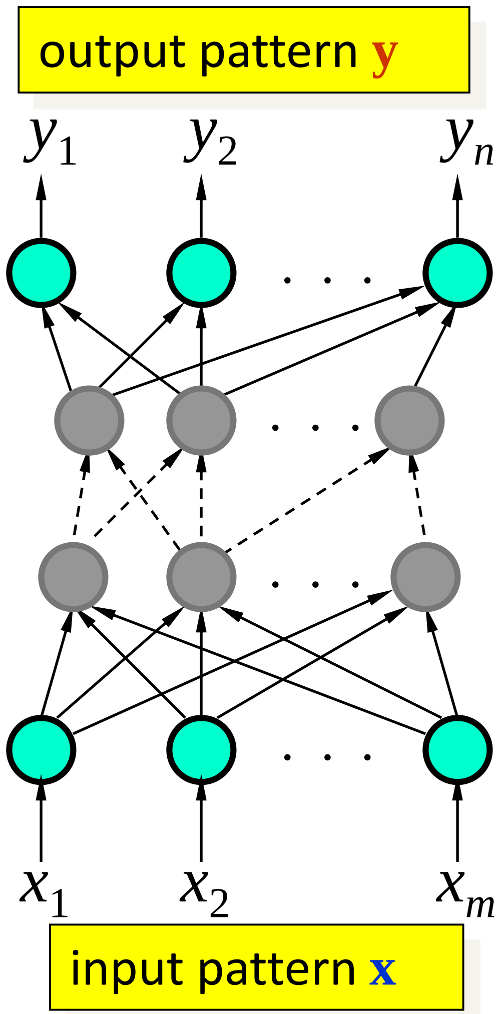
Knowledge and Memory

- The output behavior of a network is determined by the weights.
- Weights – the memory of an NN.
- Knowledge – distributed across the network.
- Large number of nodes
  - increases the storage "capacity";
  - ensures that the knowledge is robust;
  - fault tolerance.
- Store new information by changing weights.



Pattern Classification

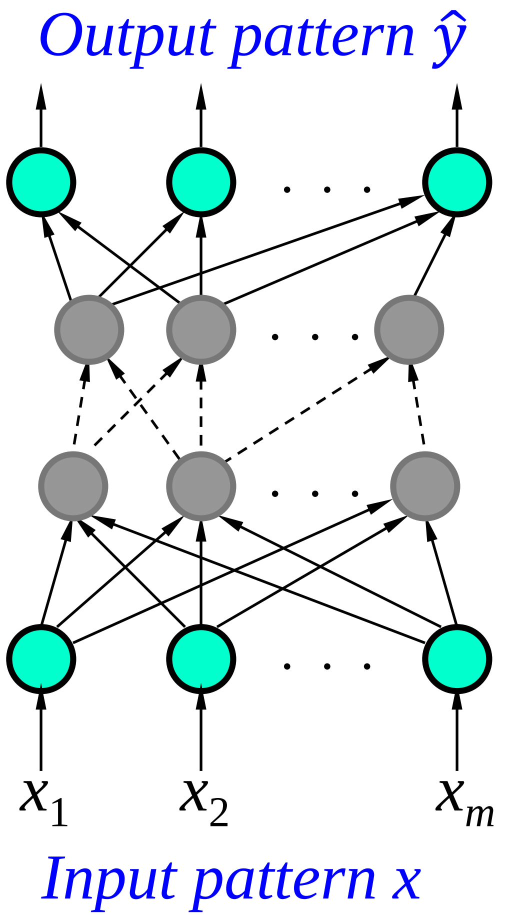
- Function: $\mathbf{x} \rightarrow \mathbf{y}$
- The NN's output is used to distinguish between and recognize different input patterns.
- Different output patterns correspond to particular classes of input patterns.
- Networks with hidden layers can be used for solving more complex problems than just a linear pattern classification.



Generalization

- By proper training, a neural network may produce reasonable answers for input patterns **not seen during training** (generalization).

- Generalization is particularly useful for the analysis of a "**noisy**" data.



Multilayer Perceptron

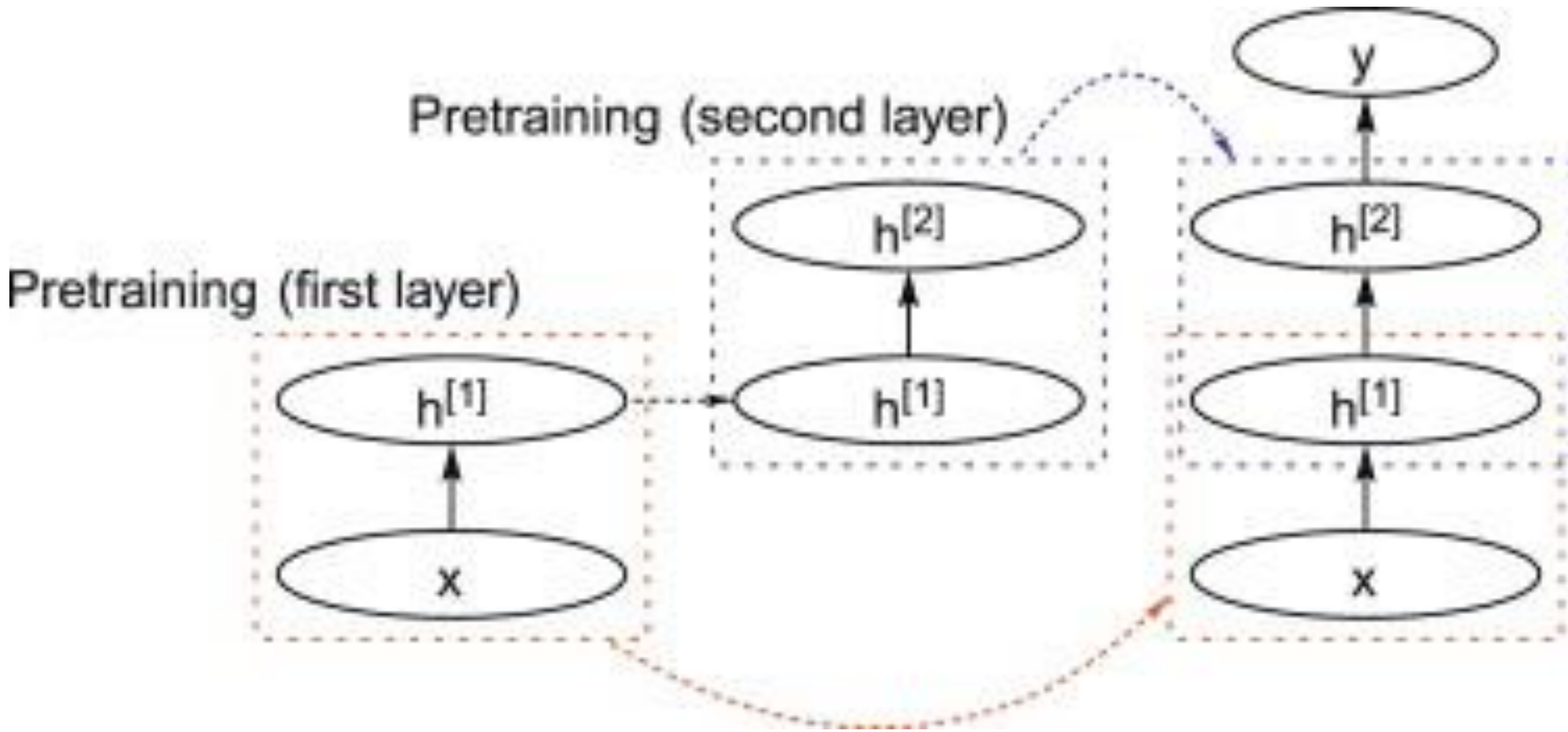
- Theoretically, accuracy of approximation increases if the number of layers increase.
- However, practically it was not feasible.

- Several reasons, like vanishing gradient problem, not enough data, not enough computing power and so on… More in the next sessions.
- Research on improving the learning performance led to several variants of NNs – Restricted Boltzmann Machine, Deep Belief Networks.

Deep Belief Network

- Catalyst in the performance improvement of Neural Networks with several layers.
- A simple way to train deep networks (DBN) to perform better and fast.
- Hinton et al. , 2006.
  - Neural networks composed of multiple layers of hidden units.
  - DBN can learn to probabilistically reconstruct its inputs using unsupervised learning.
  - Layers act as feature detectors.
  - After learning, a DBN is further trained with supervision to perform classification.

Deep Belief Network



DBN led to the development of deep neural networks as a powerful tool for solving many challenging problems like

Translation, object detection in images, speech recognition.

Image source: https://www.sciencedirect.com/topics/engineering/deep-belief-network

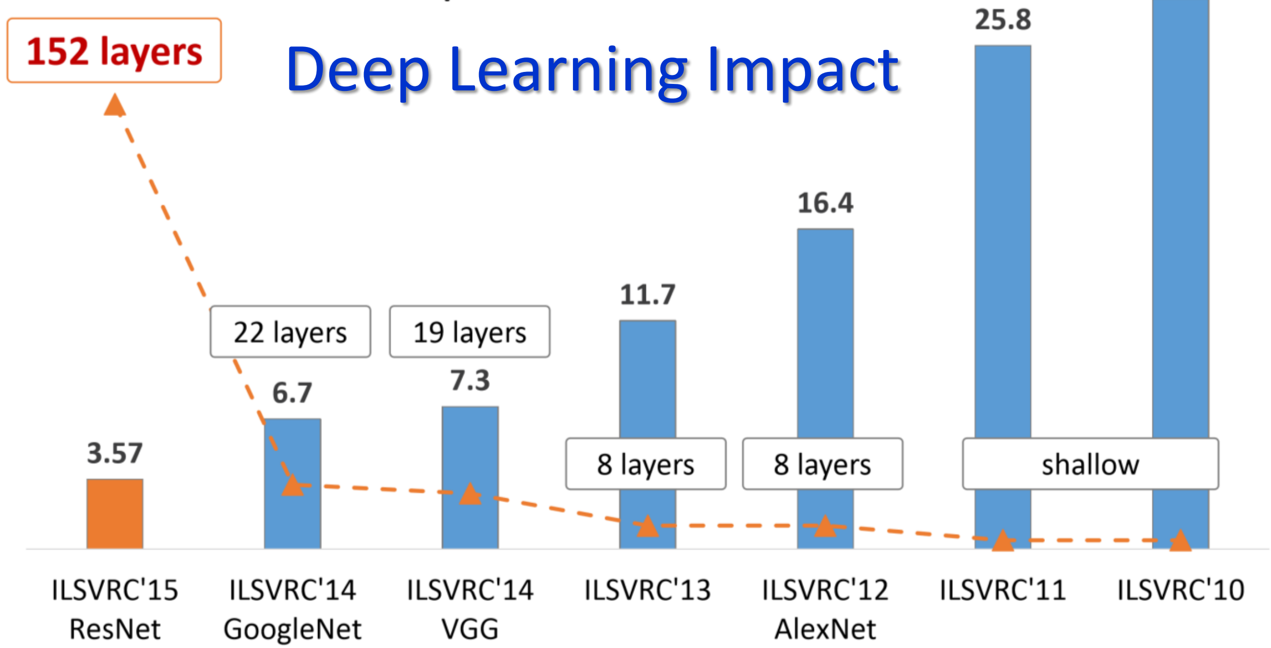
IMAGENET Large Scale Visual Recognition Challenge

The Image Classification Challenge:

1,000 object classes

1,431,167 images

Revolution of Depth



14

ImageNet Classification top-5 error (%)

Reference: Kaiming He, Xiangyu Zhang, Shaoqing Ren, & Jian Sun, 2015 ; Fei Fei Li, Lecture Slides, Introduction to CNN

Deep NN Architectures

- Conventional Neural Networks
- Convolution Neural Networks
- Recurrent Neural Networks
- New concepts are emerging every year
  - Generative Adversarial Networks: 2014
  - Residual Networks: 2015
  - Capsule Networks: 2017
  - U-Net and V-Net and more ....

Deep Learning Limitations

- Unlike human, it needs a large amount of data to learn.
- If the test data's characteristics are not similar to train data, deep learning models do not show any good performance.
  - Deep learning is shallow. The algorithms are good at memorizing, but not good at understanding the characteristics of data.
- Works like a black-box. Millions of parameters don't give much understanding about the system.
  - Which architecture, what type of filters, which hyperparameters are more influential ? Nothing clear…
- Dealing with very high dimensional data is still a problem, not theoretically, but practically.

Some Funny Examples
Wrong Object Detections by DNNs

Left: A man is holding a dog in his hand
Right: A woman is holding a dog in her hand
Image: @SouperSarah

NeuralTalk2: A flock of birds flying in the air
Microsoft Azure: A group of giraffe standing next to a tree
Image: Fred Dunn, https://www.flickr.com/photos/gratapictures - CC-BY-NC

[Source: https://aiweirdness.com/](https://aiweirdness.com/)

Summary

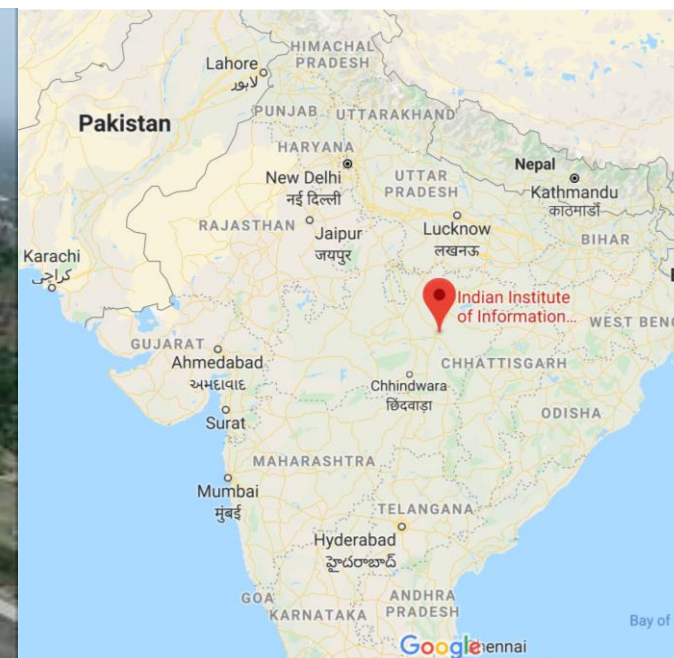
- Artificial Neural Networks are the Basic building Blocks of Present Growth in Artificial Intelligence.
- Deep learning frameworks are based on neural network based machine learning methods.
- Deep learning differs from traditional ML methods.
  - No need to manually selection features learning algorithms.
- Deep Learning Technologies have brought significant improvement in performances of machines in tasks like machine translations, object detection, data hiding etc.
- AI Industry with deep learning technologies is likely to witness exploding growth in the next 4-5 years.

Finally, Our Institute at Jabalpur

PDPM Indian Institute of Information Technology, Design and Manufacturing, Jabalpur

A lush green residential campus on 250 acres with 1500+ students, 3 engineering streams – ME, CES, ECE, Design and natural sciences as well.

State of the art research labs, research oriented programmes led by faculty



Main References

- Ian Goodfellow, Deep Learning, MIT Press, 2016.
- Geoffrey Hinton, 2007 NIPS Tutorial on Deep Belief Nets
- Andrew Ng, Coursera course on Deep Learning (deeplearning.ai)
- Wikipedia and several online image sources mentioned on slides.

Thanks!

Aparajita Ojha
IIITDM Jabalpur
Phone: +91-761-2794221(O)
aojha@iiitdmj.ac.in