# CS396: Selected CS2

(Deep Learning for visual recognition)

**Spring 2022**

**Dr. Wessam EL-Behaidy**
Associate Professor, Computer Science Department,
Faculty of Computers and Artificial Intelligence,
Helwan University.

# Course Logistics

# Course Objectives

- To introduce students to recent developments in neural network (aka “deep learning”) approaches that have greatly advanced the performance of visual recognition systems.
- To develop skills of using deep learning architectures with a focus on learning end-to-end models for these tasks, particularly image classification, and object detection.

# Resources

**Lectures (Course slides) are based on** Stanford course : Convolutional Neural Networks for Visual Recognition (CS231n): http://cs231n.stanford.edu/index.html

**Practical Labs based on:**
Francois Chollet, “ Deep Learning with Python”, Manning Publications, 2017.

# Topics to Cover

1- Introduction to deep learning
2- Review Neural Network, Gradient Descent, …
3- CNN Layers
4- Different activation functions (Sigmoid, Relu,Tanh, ...)
5- Batch Normalization and its variants
6- Different optimizers (SGD, SGD+Momentum, RMSProp, …)
7- Regularization functions (L1,L2, dropout, data augmentation)
8- CNN architectures (AlexNet, VGG, Resnet,…)
9- Transfer learning, fine tuning and ensemble models
10- Object detection and segmentation
11- GAN Generative model
12- Visualizing and understanding
13- Performance measures (Confusion matrix, AUC, ROC)

# Provisional Grading Distribution

| | |
|---|---|
| Midterm | 15% |
| Lab | 5% |
| Project | 20% |
| Final Exam | 60% |

# Join our Team Class

To communicate with us, join our team class using this code:

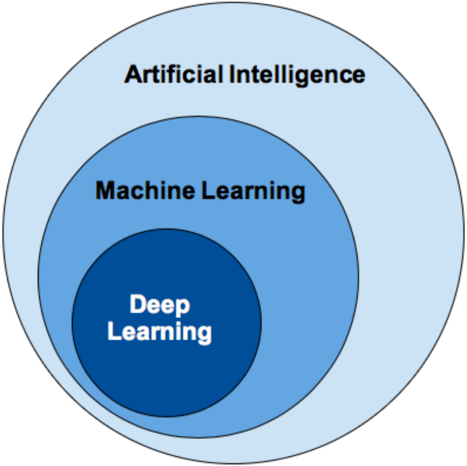
m77wt0u

# Lecture 1

Introduction to Deep Learning

# What is Deep Learning?

Deep Learning is a subset of Machine Learning that uses mathematical functions to map the input to the output. These functions can extract patterns from the data, which enables them to form a relationship between the input and the output. This is known as **learning,** and the process of learning is called **training**.




https://www.sumologic.com/wp-content/uploads/compare_AI_ML_DL.png

# History of Deep Learning Ideas and Milestones

You might be surprised to know that the history of deep learning dates back to the 1940s. Indeed, deep learning has not appeared overnight, rather it has evolved slowly and gradually over seven decades.

1943: Mathematical model of Neuron
1957-71: Perceptron, Multilayer NN
1960-68: Backpropagation
1979-98: CNN, MNIST
2006: "Deep Learning", DBN
2009: ImageNet
2012: AlexNet
2014: GANs
2016: AlphaGo
2017-19: Transformers

# History of Deep Learning Ideas and Milestones

You might be surprised to know that the history of deep learning dates back to the 1940s. Indeed, deep learning has not appeared overnight, rather it has evolved slowly and gradually over seven decades.

1943: Mathematical model of Neuron
1957-71: Perceptron, Multilayer NN

**1943**: McCulloch Pitts Neuron is the mathematical model of a biological neuron, with no learning mechanism.
**1957:** Frank Rosenblatt Creates "Perceptron" that had true learning capabilities to do binary classification on its own.
**1965-71:** Ivakhnenko creates the first multi-layer perceptron and he is often considered as the **father of deep learning**.

# History of Deep Learning Ideas and Milestones

You might be surprised to know that the history of deep learning dates back to the 1940s. Indeed, deep learning has not appeared overnight, rather it has evolved slowly and gradually over seven decades.

## 1960-86: Backpropagation (BP)

**1960:** The first-ever version of the continuous BP model in context to Control Theory.
**1962:** BP model that uses simple derivative chain rule.
**1970:** BP is computer coded.
**1982:** Proposal For Backpropagation In ANN
**1986:** Geoffrey Hinton et al. shows the successful implementation of BP in ANN.

# History of Deep Learning Ideas and Milestones

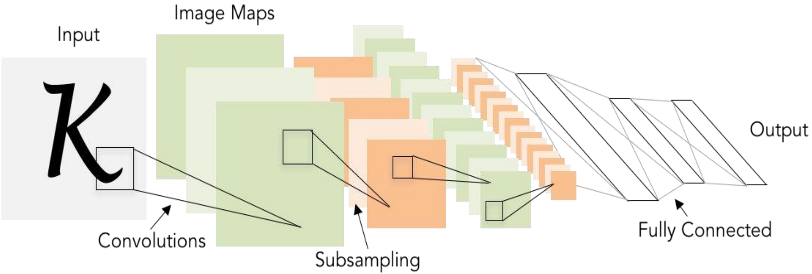
You might be surprised to know that the history of deep learning dates back to the 1940s. Indeed, deep learning has not appeared overnight, rather it has evolved slowly and gradually over seven decades.

## 1980-98: CNN, MNIST

**1980:** Neocognitron – First CNN Architecture could recognize visual patterns such as handwritten characters.
**1991:** Vanishing Gradient Problem Appears.
**1997:** The Milestone Of LSTM
**1998**: Yann LeCun -CNN Using Backpropagation to recognize handwritten digits. A breakthrough moment as it lays the foundation of modern computer vision using deep learning.

# LeNet-5 – A Classic CNN Architecture

**Gradient-based learning applied to document recognition**
*[LeCun, Bottou, Bengio, Haffner 1998]*




LeNet-5

# History of Deep Learning Ideas and Milestones

You might be surprised to know that the history of deep learning dates back to the 1940s. Indeed, deep learning has not appeared overnight, rather it has evolved slowly and gradually over seven decades.

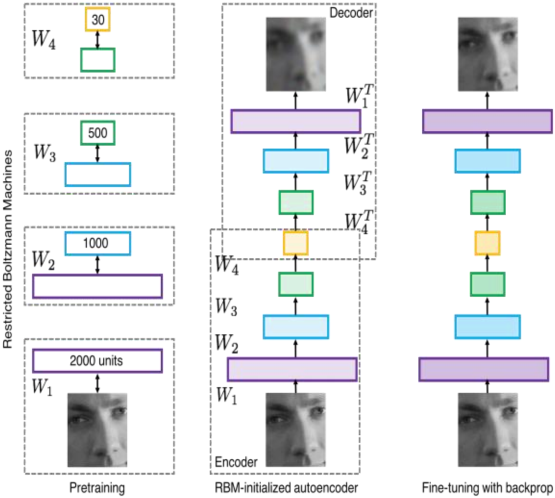
2006: “Deep Learning”, DBN
2009: ImageNet

**2006:** Deep Belief Network (DBF)
**2008:** GPU Revolution Begins- Andrew NG’s group in Stanford starts the use of GPUs for training Deep Neural Networks.
**2009:** Fei-Fei Li a professor at Stanford– ImageNet is launched
**2011:** Yoshua Bengio shows that ReLU activation function can avoid vanishing gradient problem.

# Hinton and Salakhutdinov 2006

[Hinton and Salakhutdinov 2006]

Reinvigorated research in Deep Learning



# History of Deep Learning Ideas and Milestones

You might be surprised to know that the history of deep learning dates back to the 1940s. Indeed, deep learning has not appeared overnight, rather it has evolved slowly and gradually over seven decades.

2006: “Deep Learning”, DBN

2009: ImageNet

**2006:** Deep Belief Network (DBF)

**2008:** GPU Revolution Begins- Andrew NG’s group in Stanford starts the use of GPUs for training Deep Neural Networks.

**2009:** Fei-Fei Li a professor at Stanford– ImageNet is launched

**2011:** Yoshua Bengio shows that ReLU activation function can avoid vanishing gradient problem.

# Appearance of Large Datasets

IMAGENET

www.image-net.org

**22K** categories and **14M** images

- Animals
  - Bird
  - Fish
  - Mammal
  - Invertebrate
- Plants
  - Tree
  - Flower
- Food
- Materials
- Structures
- Artifact
  - Tools
  - Appliances
  - Structures
- Person
- Scenes
  - Indoor
  - Geological Formations
- Sport Activities

Deng, Dong, Socher, Li, Li, & Fei-Fei, 2009

# IMAGENET LSVRC Challenge

IMAGENET Large Scale Visual Recognition Challenge

The Image Classification Challenge:
1,000 object classes
1,431,167 images

Output:
Scale
T-shirt
Steel drum
Drumstick
Mud turtle

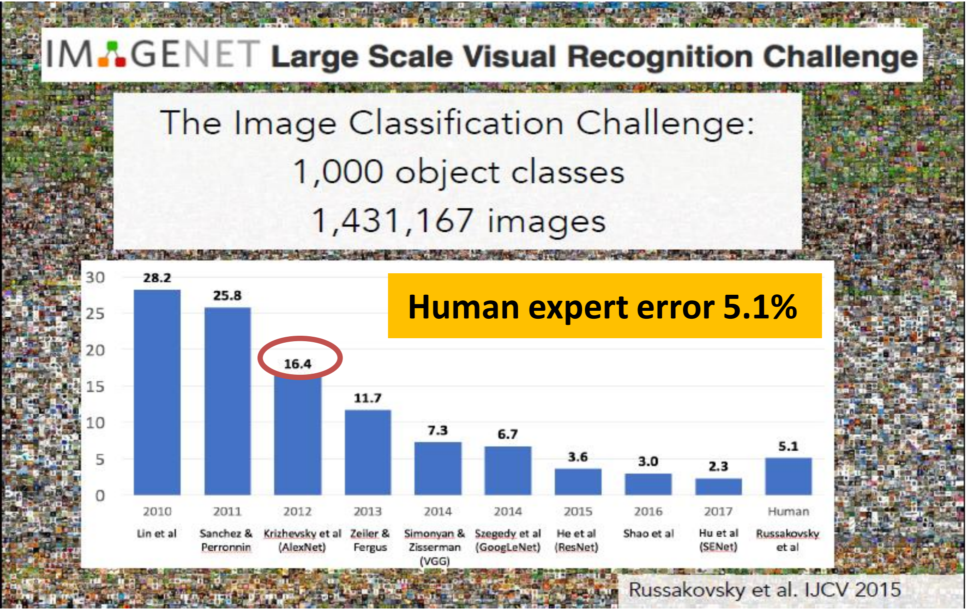
✔

Output:
Scale
T-shirt
Giant panda
Drumstick
Mud turtle

✗

Russakovsky et al. IJCV 2015

# IMAGENET LSVRC Challenge results



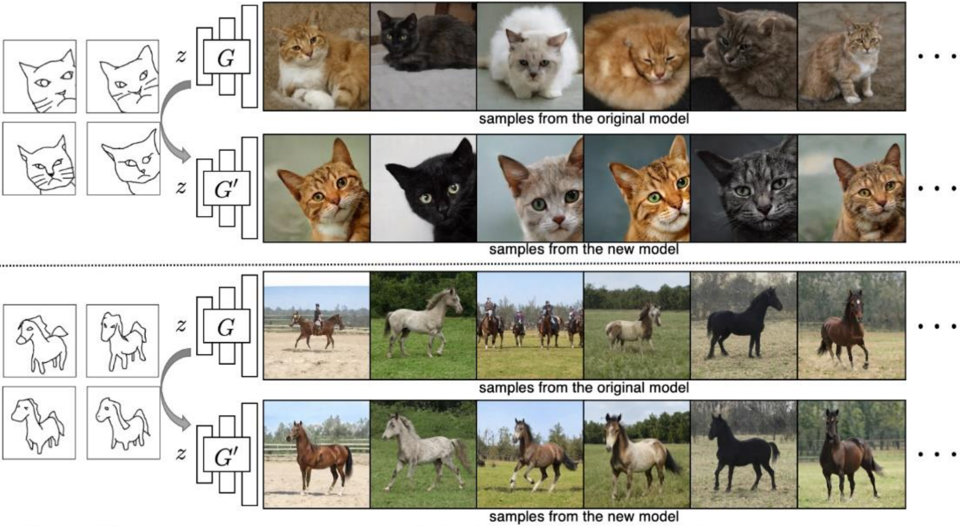
# History of Deep Learning Ideas and Milestones

You might be surprised to know that the history of deep learning dates back to the 1940s. Indeed, deep learning has not appeared overnight, rather it has evolved slowly and gradually over seven decades.

2012: AlexNet
2014: GANs
2016: AlphaGo
2017-19: Transformers

**2012:** AlexNet Starts Deep Learning Boom-- Alex Krizhevsky, wins Imagenet's image classification contest with accuracy of 84%.
**2014:** The Birth Of GANs--Ian Goodfellow with GANs opens whole new doors of application of deep learning in fashion, art, science due to its ability to synthesize real-like data.
**2016:** AlphaGo Beats Human
**2017:** Transformers were introduced

# GAN Sketching




User sketches

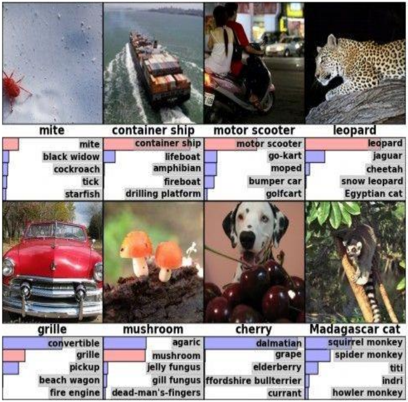
Customizing a GAN using human sketches

https://peterwang512.github.io/GANSketching/

# Ingredients for Deep Learning

# Fast-forward to today: ConvNets are everywhere

Classification




Retrieval

Figures copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

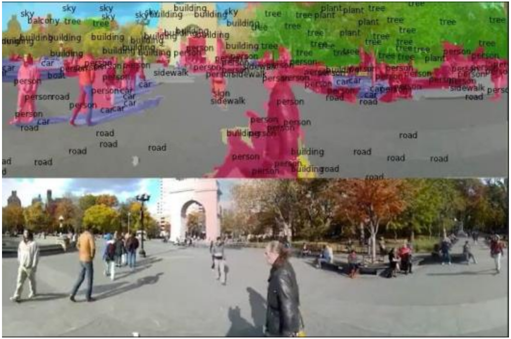
# Fast-forward to today: ConvNets are everywhere

Detection


Figures copyright Shaoqing Ren, Kaiming He, Ross Girschick, Jian Sun, 2015. Reproduced with permission.

*[Faster R-CNN: Ren, He, Girshick, Sun 2015]*

Segmentation



Figures copyright Clement Farabet, 2012. Reproduced with permission.

*[Farabet et al., 2012]*

# Fast-forward to today: ConvNets are everywhere

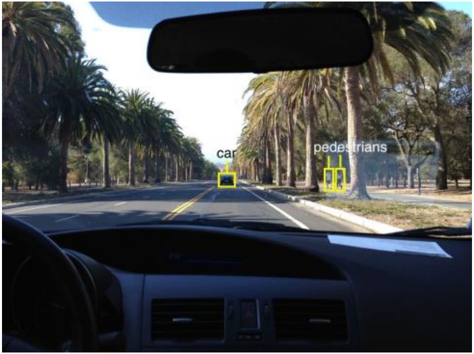


Photo by Lane McIntosh. Copyright CS231n 2017.

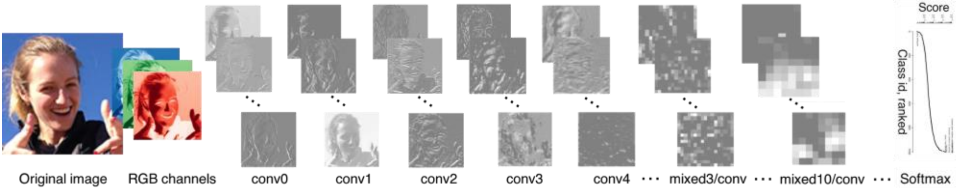
self-driving cars

This image by GBPublic_PR is licensed under CC-BY 2.0

NVIDIA Tesla line
(these are the GPUs on rye01.stanford.edu)

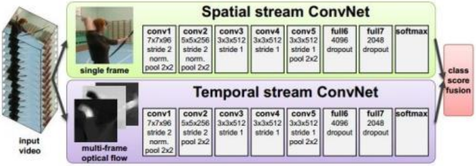
Note that for embedded systems a typical setup would involve NVIDIA Tegras, with integrated GPU and ARM-based CPU cores.

# Fast-forward to today: ConvNets are everywhere




*[Taigman et al. 2014]*

Activations of [inception-v3 architecture] [Szegedy et al. 2015] to image of Emma McIntosh, used with permission. Figure and architecture not from Taigman et al. 2014.




*[Simonyan et al. 2014]*

Figures copyright Simonyan et al., 2014. Reproduced with permission.

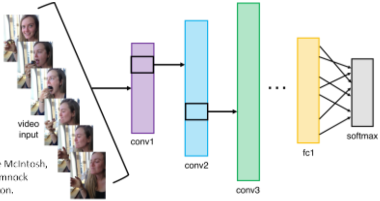
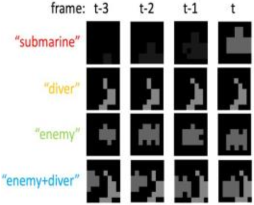


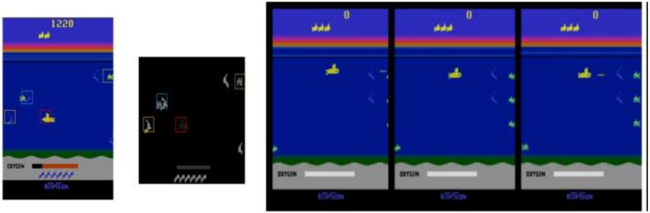

Illustration by Lane McIntosh, photos of Katie Cumnock used with permission.

# Fast-forward to today: ConvNets are everywhere

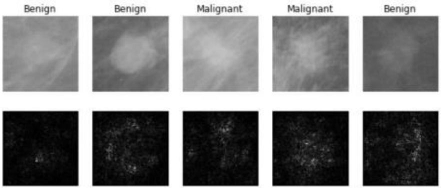
Images are examples of pose estimation, not actually from Toshev & Szegedy 2014. Copyright Lane McIntosh.

*[Toshev, Szegedy 2014]*



# Fast-forward to today: ConvNets are everywhere




*[Levy et al. 2016]*

Figure copyright Levy et al. 2016.
Reproduced with permission.

*[Dieleman et al. 2014]*

From left to right: public domain by NASA, usage permitted by ESA/Hubble, public domain by NASA, and public domain.


*[Sermanet et al. 2011]*
*[Ciresan et al.]*

Photos by Lane McIntosh.
Copyright CS231n 2017.

# Fast-forward to today: ConvNets are everywhere

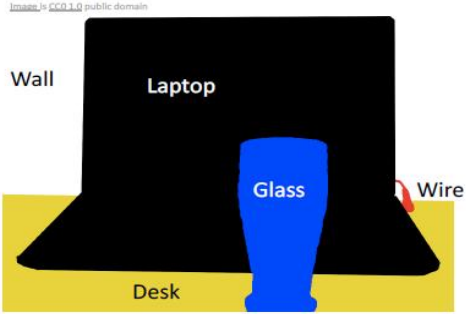
This image by Christin Khan is in the public domain and originally came from the U.S. NOAA.

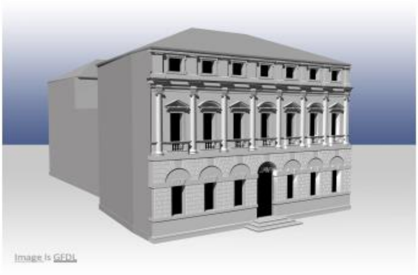
*Whale recognition, Kaggle Challenge*

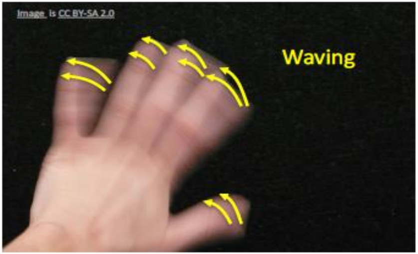
Photo and figure by Lane McIntosh; not actual example from Mnih and Hinton, 2010 paper.

*Mnih and Hinton, 2010*

# Fast-forward to today: ConvNets are everywhere




# Fast-forward to today: ConvNets are everywhere

Minor errors

Somewhat related

*A white teddy bear sitting in the grass*

*A man in a baseball uniform throwing a ball*

*A woman is holding a cat in her hand*

*A man riding a wave on top of a surfboard*

*A cat sitting on a suitcase on the floor*

*A woman standing on a beach holding a surfboard*

## Image Captioning

*[Vinyals et al., 2015]*
*[Karpathy and Fei-Fei, 2015]*

All images are CC0 Public domain:
https://pixabay.com/en/luggage-antique-cat-1643010/
https://pixabay.com/en/teddy-plush-bears-cute-teddy-bear-1623436/
https://pixabay.com/en/surf-wave-summer-sport-litoral-1668716/
https://pixabay.com/en/woman-female-model-portrait-adult-983967/
https://pixabay.com/en/handstand-lake-meditation-496008/
https://pixabay.com/en/baseball-player-shortstop-infield-1045263/

Captions generated by Justin Johnson using Neuraltalk2

# Types of Neural networks (NNs)

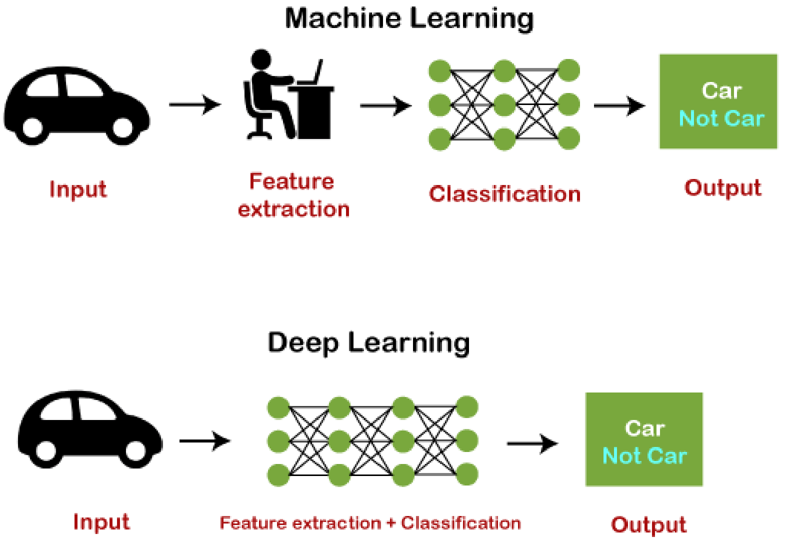
1- Shallow Neural Networks: ordinary fully connected NN.
2- Convolutional Neural Networks or CNNs are primarily used for modeling spatial data such as 2D or 3D images and videos. Their tasks are related to computer vision (image classification or object detection ) or image processing.
3- Recurrent Neural Networks or RNN are primarily used to model sequential data, such as text, audio, or any type of data that represents sequence or time. They are often used in tasks related to natural language processing (NLP).

# Types of Neural networks (NNs)

4- Generative adversarial networks or GANs are frameworks that are used for the tasks related to unsupervised learning. This type of network essentially learns the structure of the data, and patterns in a way that it can be used to generate new examples, similar to that of the original dataset.

5- Transformers are the new class deep learning model that is used mostly for the tasks related to modeling sequential data, like that in NLP. Recently, they are also being applied in computer vision tasks and they are proving to be quite more effective than the traditional CNNs.

# Machine Learning Vs. Deep Learning



# Machine Learning Vs. Deep Learning

| | Traditional ML | Deep Learning |
|---|---|---|
| **Data Dependency** | It can work with a small amount of data. Adv. | It highly depends on a large amount of data for good performance. |
| **Execution time** | It takes less time to train the model than deep learning, but a long-time duration to test the model. | It takes a long time to train the model, but less time to test the model. Adv. |

# Machine Learning Vs. Deep Learning

| | Traditional ML | Deep Learning |
|---|---|---|
| **Hardware Dependencies** | They can work on low-end machines. Adv. | They need GPU's and hence the high-end machine. |
| **Feature Engineering** | They need a step of feature extraction by the expert, and then it proceeds further. | It does not need to develop the feature extractor for each problem; instead, it tries to learn high-level features from the data on its own. Adv. |

# Machine Learning Vs. Deep Learning

| | Traditional ML | Deep Learning |
|---|---|---|
| **Problem-solving approach** | It breaks the problem in sub-parts, and after solving each part, produces the final result. | It takes input for a given problem, and produce the end result. Hence it follows the end-to-end approach. Adv. |
| **Interpretation of result** | The interpretation of the result for a given problem is easy; it means why this result occur, what was the process. Adv. | The interpretation of the result for a given problem is very difficult; we can not find why this result occurred, and the reasoning of this result. |

# Machine Learning Vs. Deep Learning

| | Traditional ML | Deep Learning |
|---|---|---|
| **Type of data** | Machine learning models mostly require data in a structured form. | Deep Learning models can work with both structured and unstructured data as they rely on the layers of the ANNs. Adv. |
| **Suitable for** | Machine learning models are suitable for solving simple or bit-complex problems. | Deep learning models are suitable for solving complex problems. Adv. |

# Resource of this lecture

Stanford course : Convolutional Neural Networks for Visual Recognition (CS231n):
http://cs231n.stanford.edu/index.html

A Gentle Introduction to Deep Learning:
https://www.v7labs.com/blog/deep-learning-guide#what-is-deep-learning

Brief History of Deep Learning from 1943-2019:
https://machinelearningknowledge.ai/brief-history-of-deep-learning/

https://www.javatpoint.com/machine-learning-vs-deep-learning

Thanks