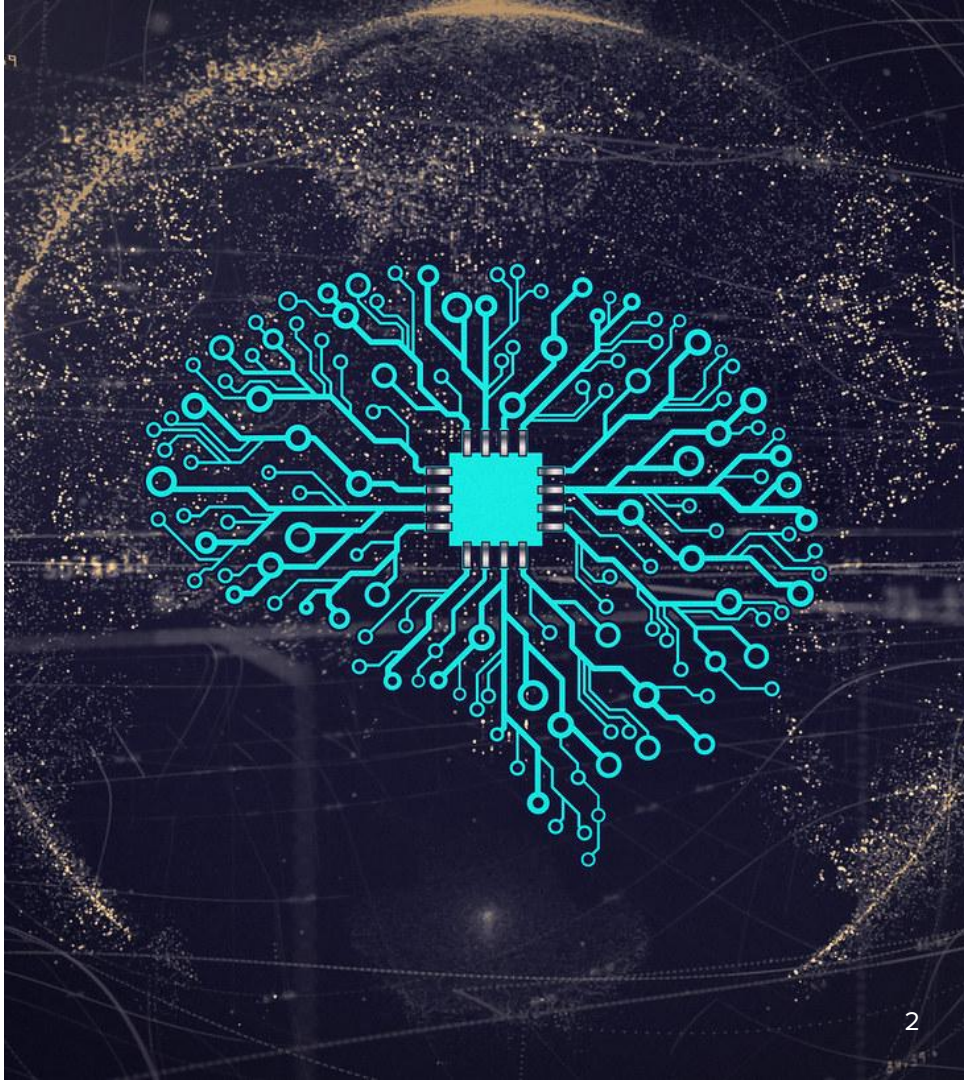
Introduction to Deep Learning

Dr. Paisit Khanarsa

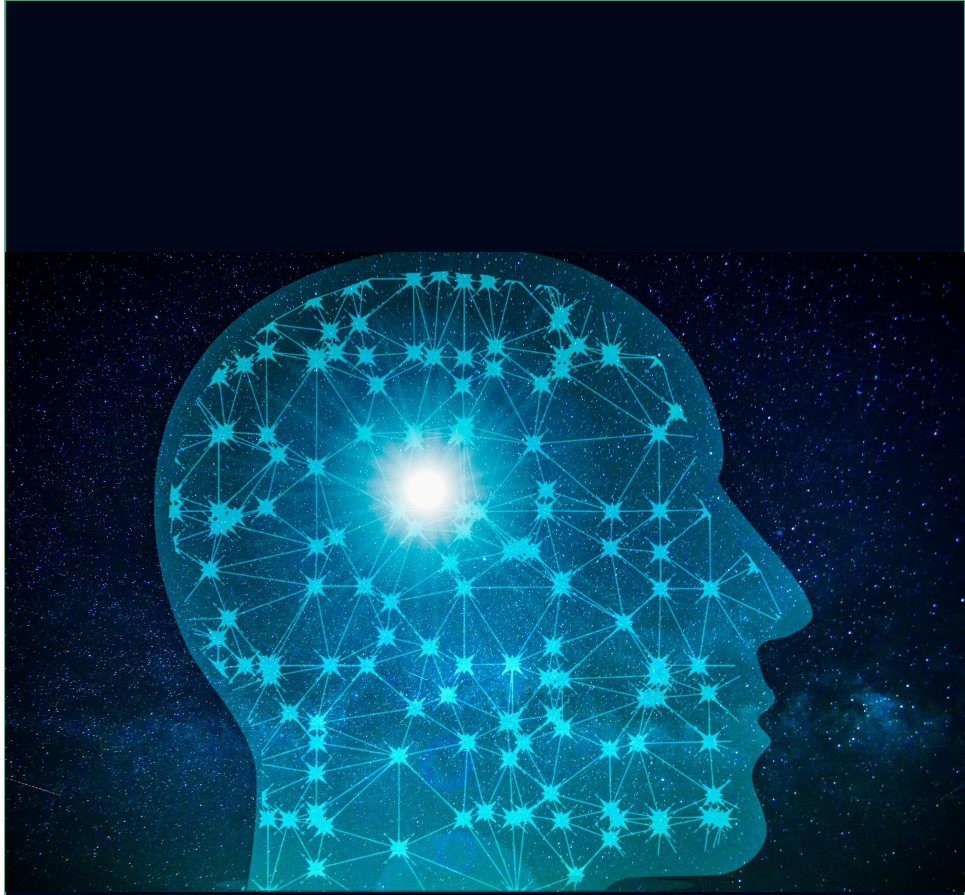
FIBO, Kmutt

Outline

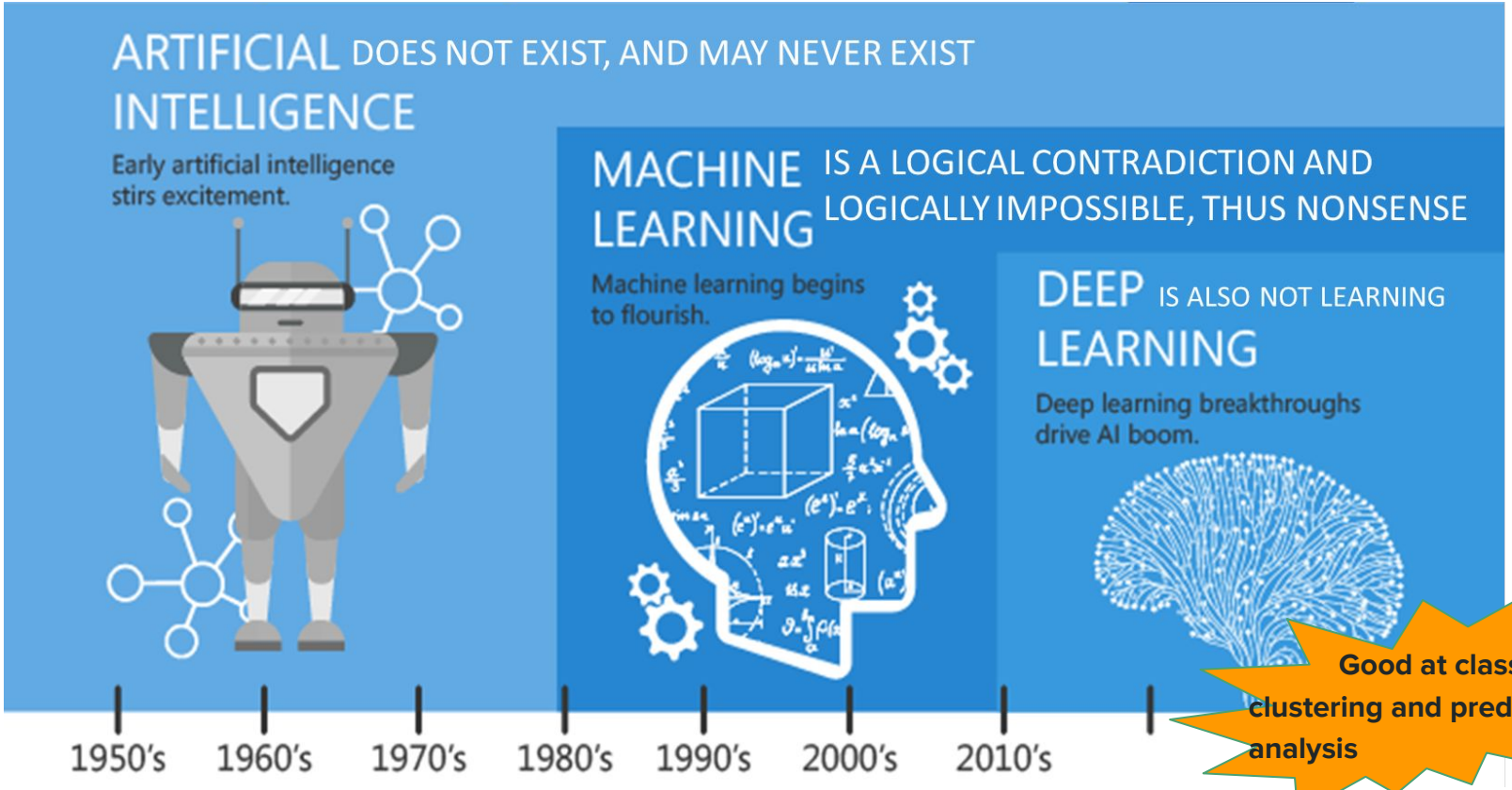
- What is Deep Learning?
- Deep Computer Vision
- Deep Sequence Modeling
- Deep Learning Development Flow
- Deep Learning Frameworks and Interesting Datasets



What is Deep Learning?



Deep Learning Introduction



https://medium.com/@alanb_73111/artificial-intelligence-vs-machine-learning-vs-deep-learning-ai-vs-ml-vs-dl-e6afb7177436

Good at classification, clustering and predictive analysis

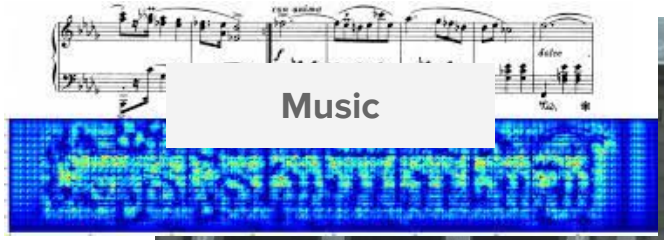
Deep Learning Introduction

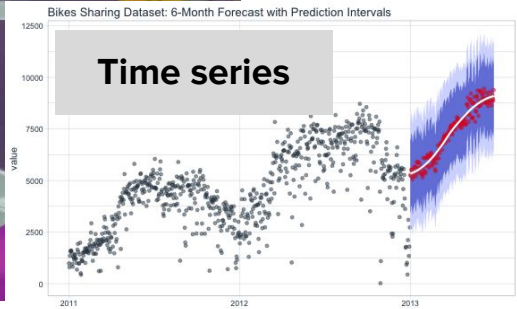
Deep learning is way of classifying, clustering, and predicting things by using a neural network that has been trained on vast amounts of data.

https://divamgupta.com/image-segmentation/2019/06/06/deep-learning-semantic-segmentation-keras.html

Deep Learning Introduction

Deep learning is way of classifying, clustering, and predicting things by using a neural network that has been trained on vast amounts of data.

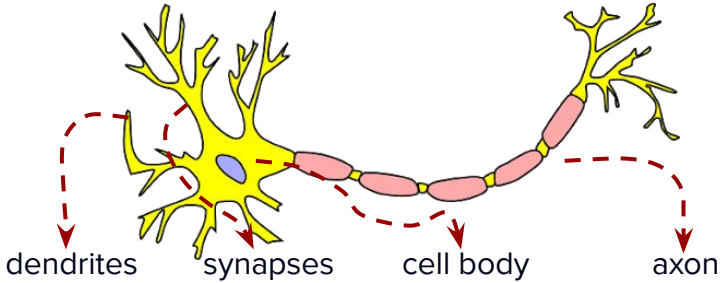




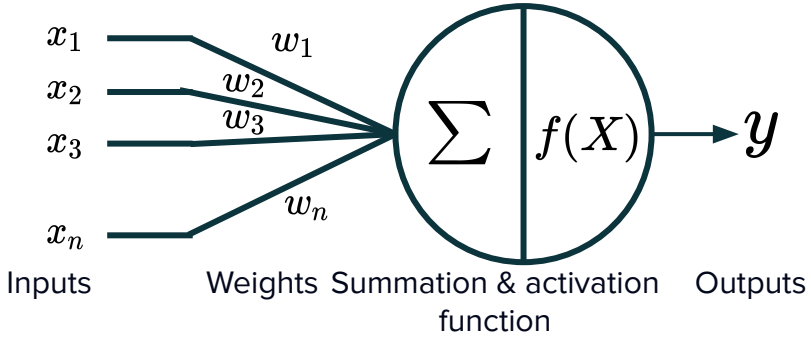
What is deep learning?

- Deep learning is a root in neural networks.
- Neural networks are sets of algorithm, modeled loosely after the human brain, that are designed to **recognize patterns**.

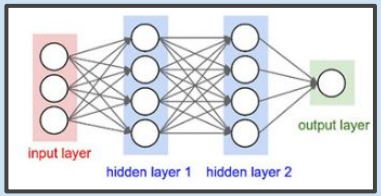
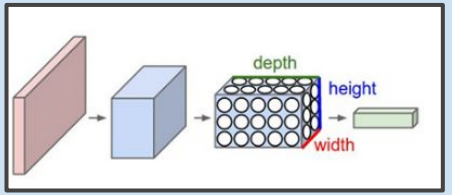
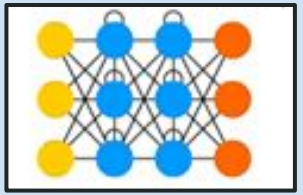
Biological neuron



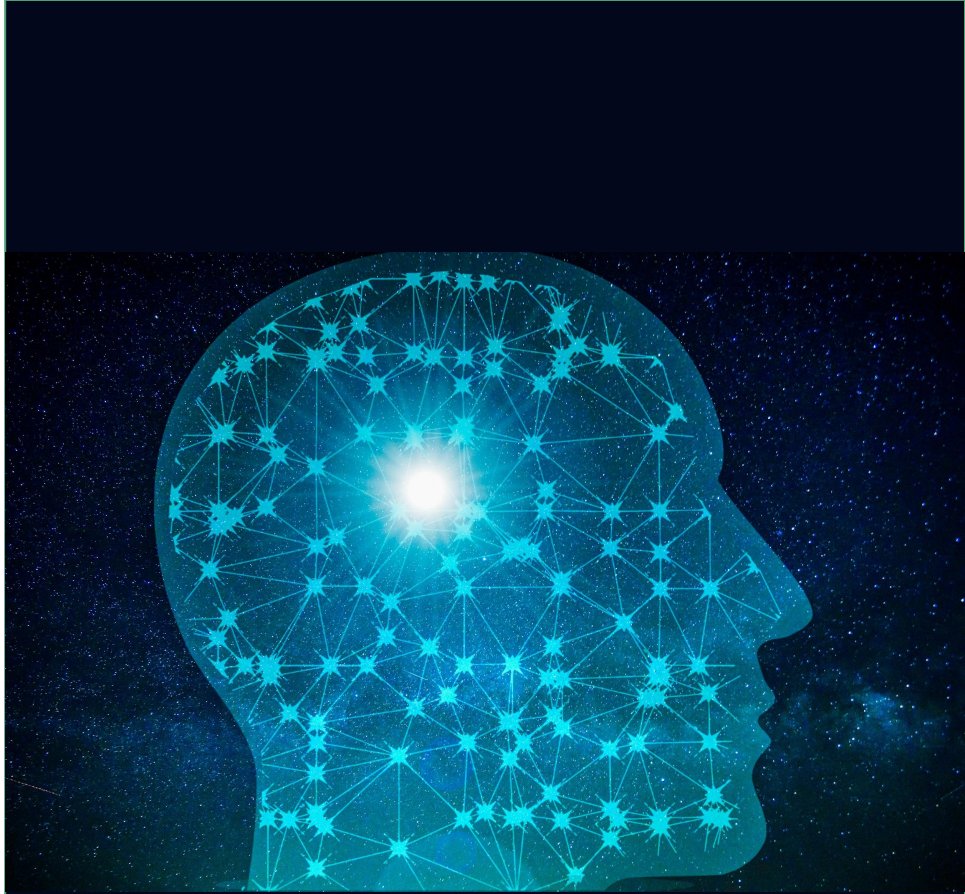
Artificial neuron



Deep Neural Networks (DNN)

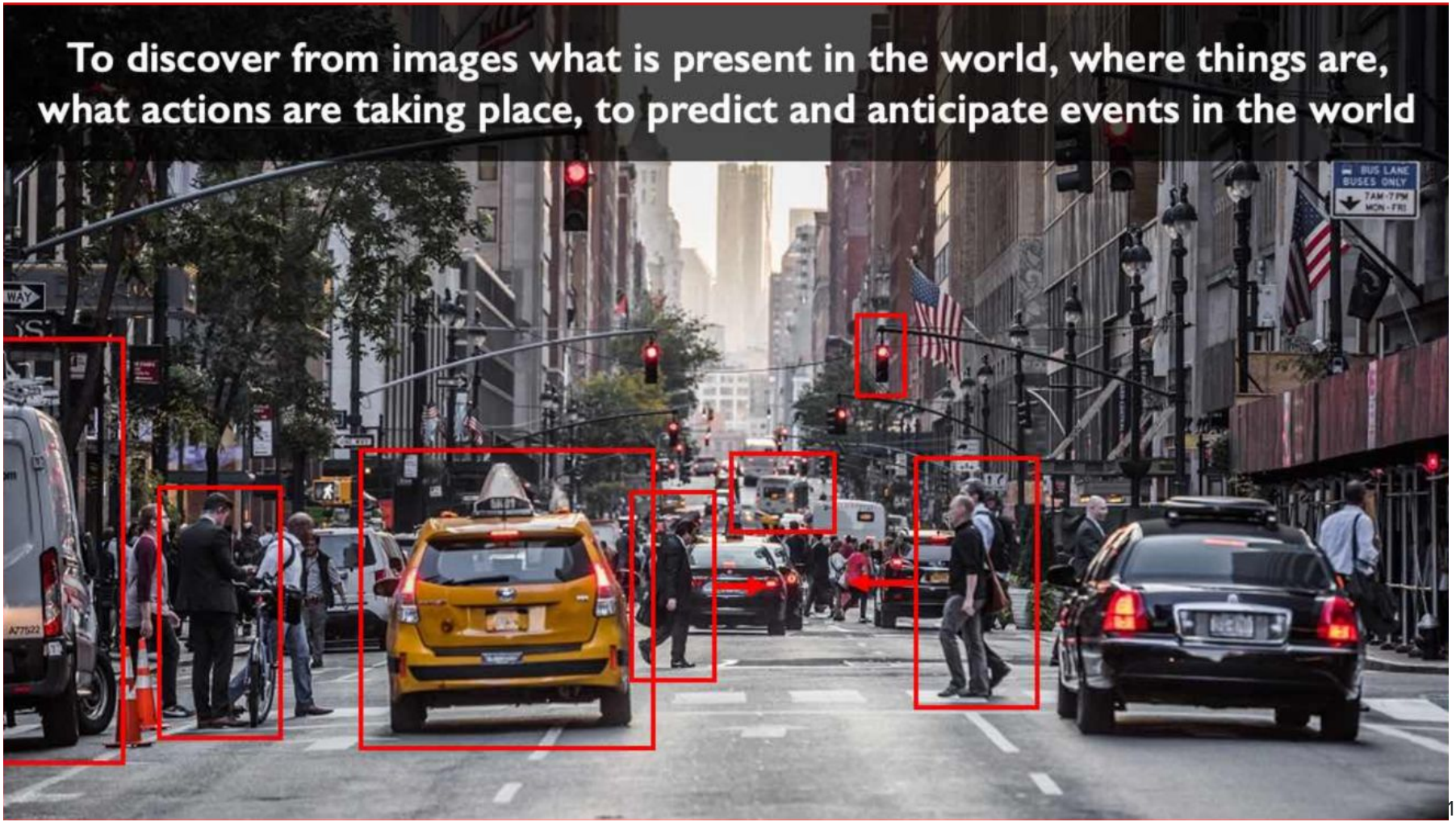
Multi layer perceptron (MLP)	• One of the most traditional types of ML and DL architectures. • Every element of a previous layer is connected to every element of the next layer. • Fell out of favor, in part because it is hard to train.	**Fully connected layer** 
Convolutional neural network (CNN) (Deep Computer Vision)	• Feed forward deep neural network. • Application for object detection, classification and segmentation. • Popular uses for image and video processing.	**Feed-forward network** 
Recurrent neural network (RNN) (Deep Sequence Modeling)	• Feed forward deep neural network, extended to feedback connections. • Ideal for text, time series and speech analysis.	**Cyclical connect** 

Deep Computer Vision



http://introtodeeplearning.com/


“To know what is
where by looking.”

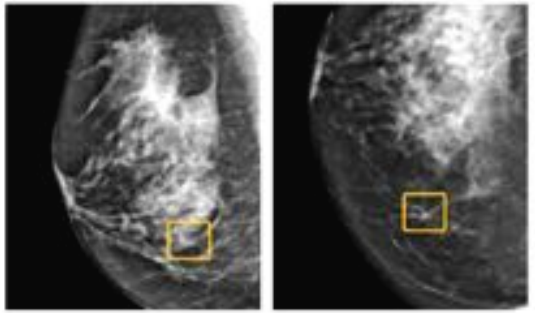

To discover from images what is present in the world, where things are, what actions are taking place, to predict and anticipate events in the world
BUS LANE
BUSES ONLY
7AM-7PM
MON-FRI

To rise and impact of computer vision

Robotics

Accessibility

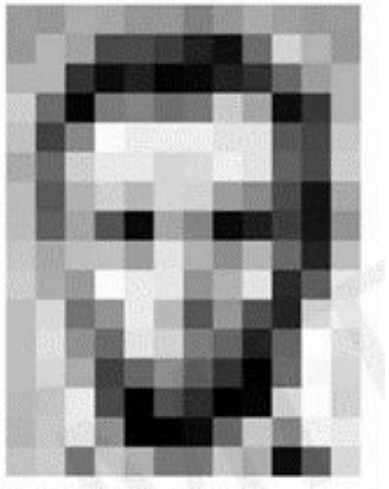
Biology & Medicine



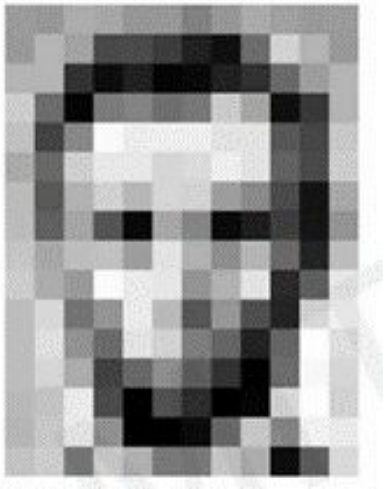
Autonomous driving

Mobile computing

What computers see

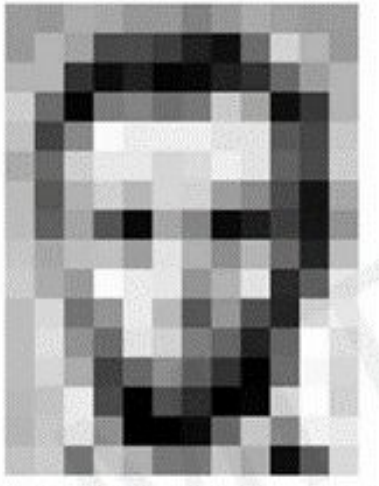


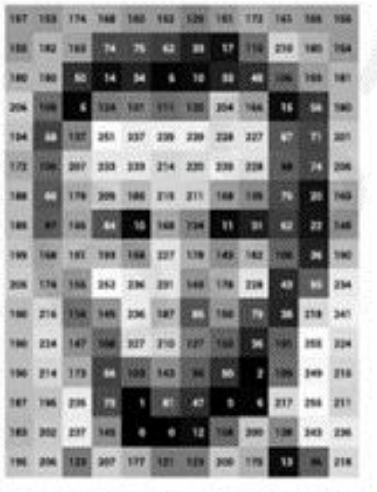
What computers see



What computers see

Images are Numbers





An image is just a matrix of numbers [0,255],
i.e., 1080 x 1080 x 3 for an RGB images

How we identify things ?

How we identify things ?

Let’s identify key features in each image category

How we identify things ?

Let’s identify key features in each image category

Nose, Eyes, Mouth

Wheels, License Plate, Headlight

Door, Windows, Steps

Manual Feature Extraction

Manual Feature Extraction

Problems?

Manual Feature Extraction

Labradoodle
or
Fried chicken

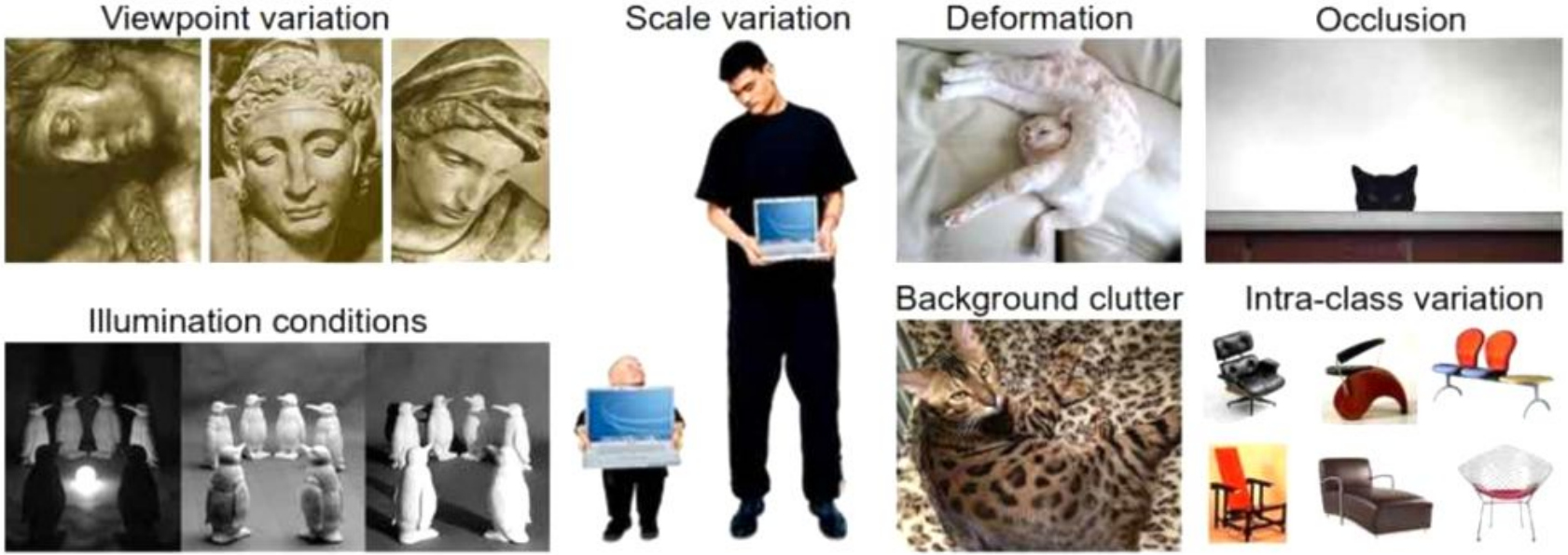
Manual Feature Extraction

Puppy
or
Bagel

Manual Feature Extraction

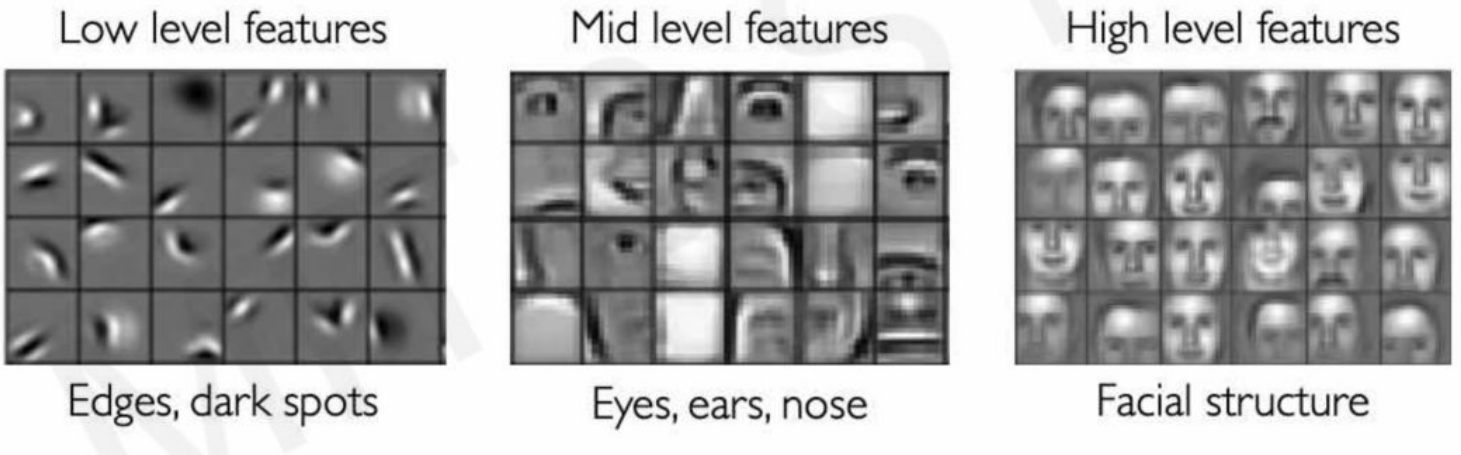
Raw chicken
or
Donald Trump

Manual Feature Extraction



Learning Feature Representations

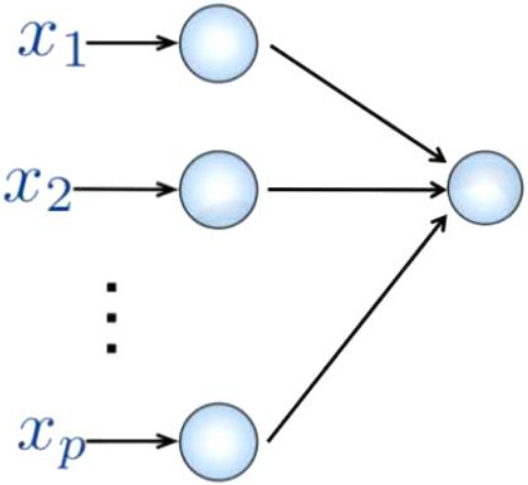
Can we learn a hierarchy of features directly from the data instead of hand engineering?



Learning Visual Features

Fully Connected Neural Network

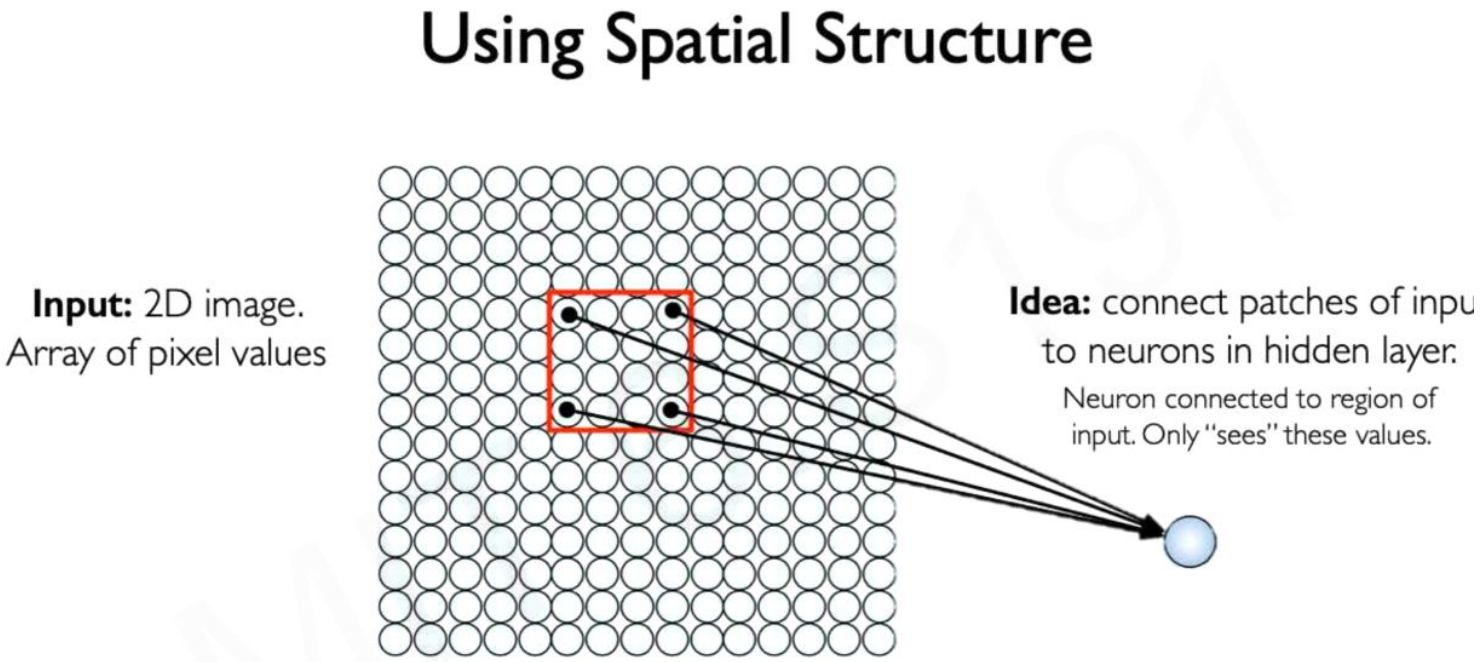
Input:
- 2D image
- Vector of pixel values



Fully Connected:
- Connect neuron in hidden layer to all neurons in input layer
- No spatial information!
- And many, many parameters

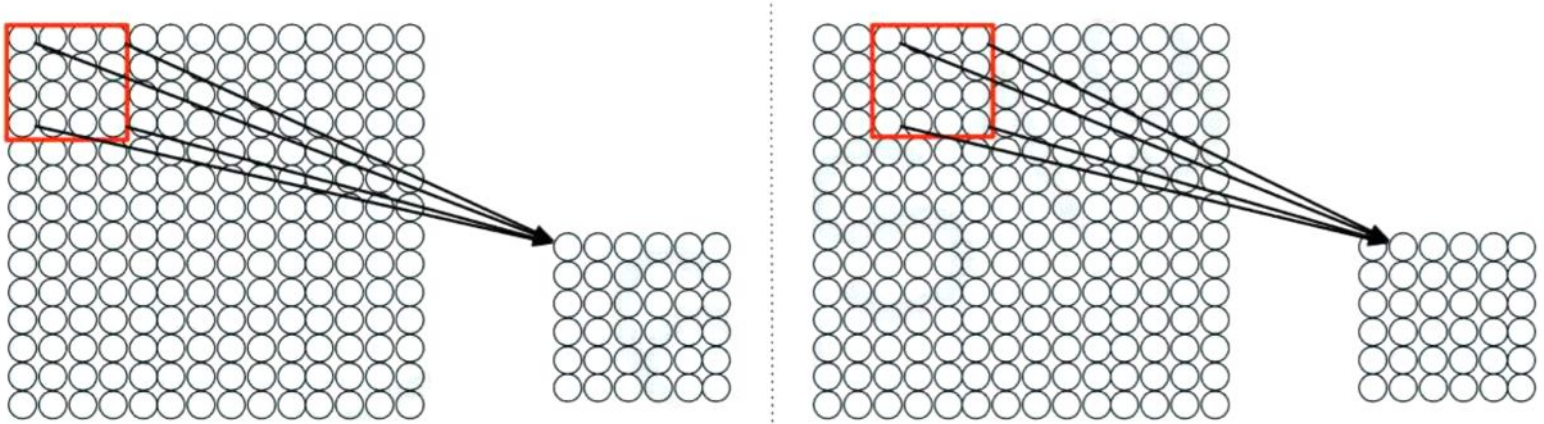
How can we use **spatial structure** in the input to inform the architecture of the network?

Learning Visual Features



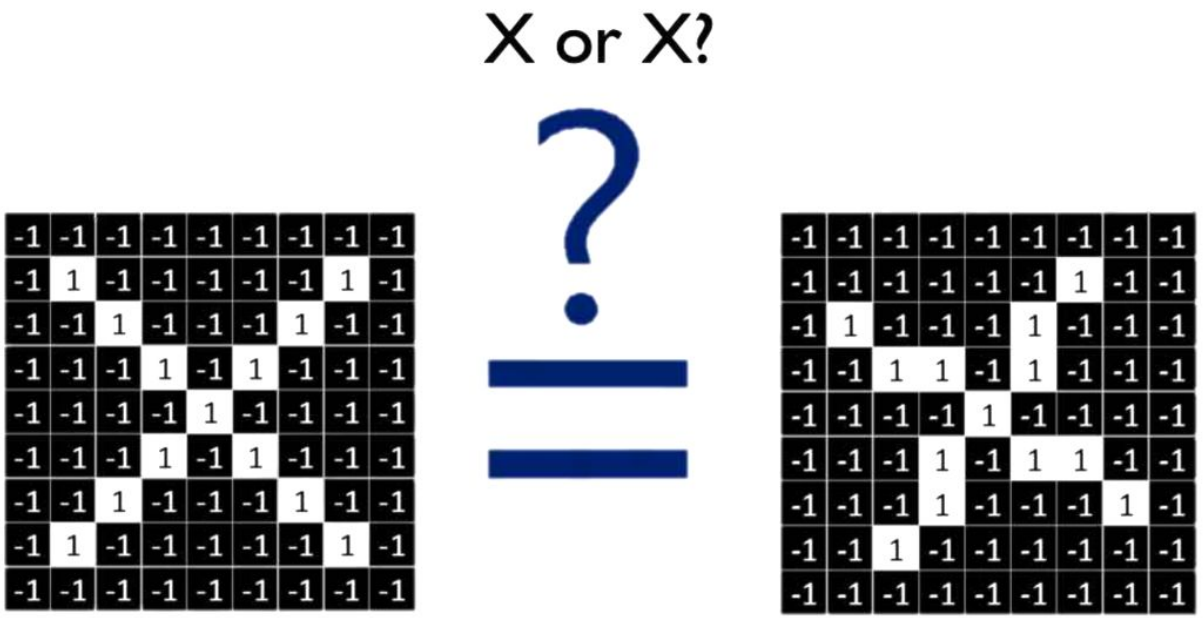
Learning Visual Features

Using Spatial Structure

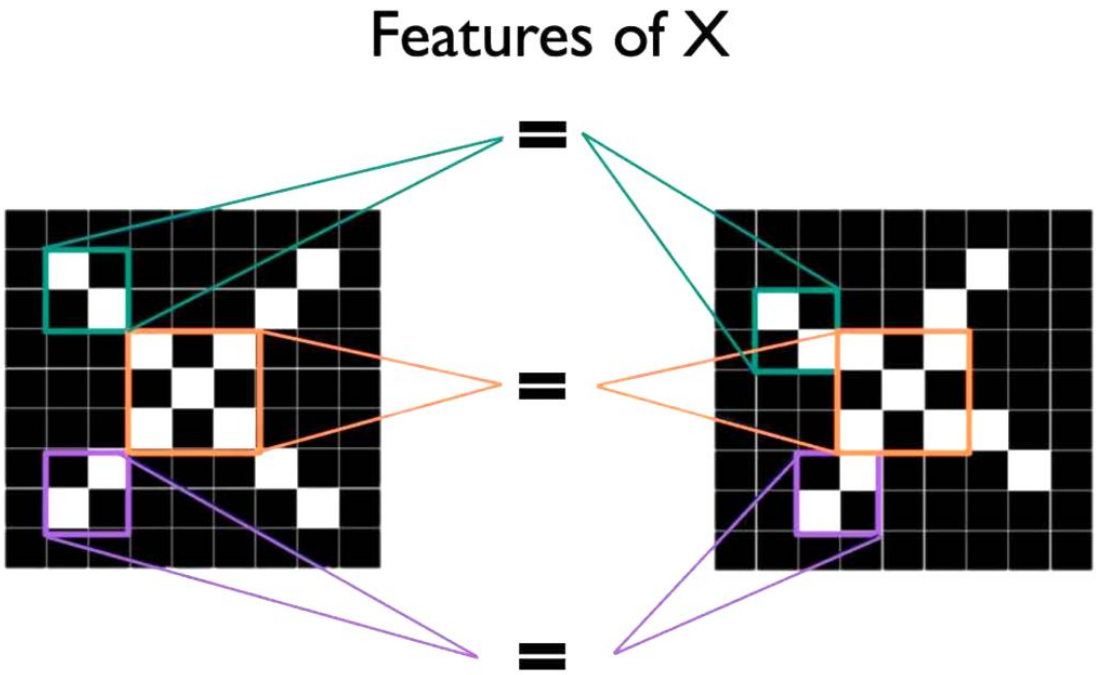


Connect patch in input layer to a single neuron in subsequent layer.
Use a sliding window to define connections.
How can we **weight** *the patch to detect particular features?*

Feature Extraction and Convolution: A Case Study

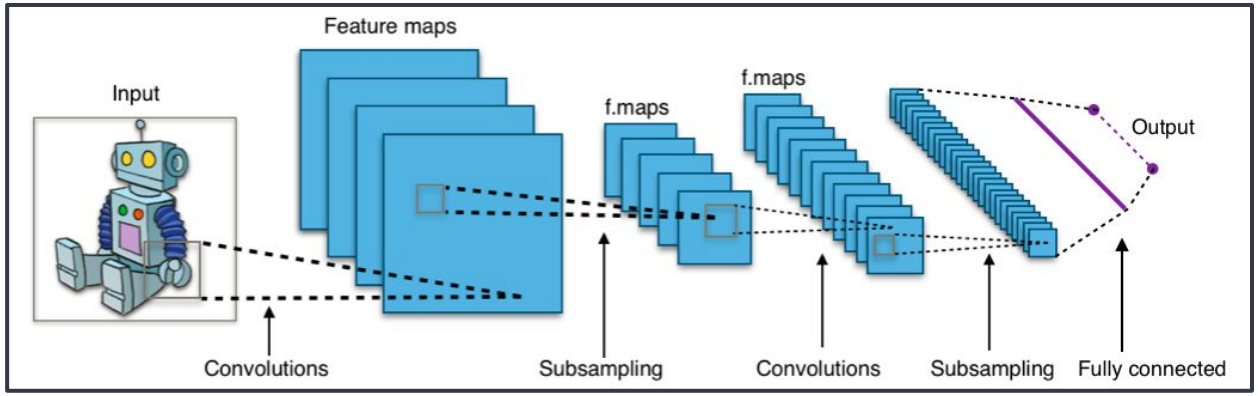


Feature Extraction and Convolution: A Case Study



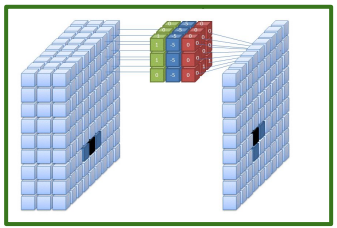
Typical layer involved in Convolutional Neural Networks

Structure of deep learning model

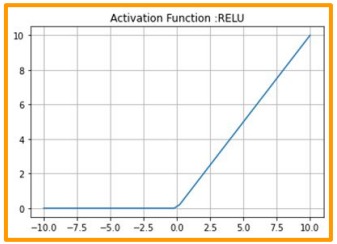


https://medium.com/@penanklihicyrille_16953/introduction-aux-r%C3%A9seaux-de-neurones-de-convolution-b5da87a5321f

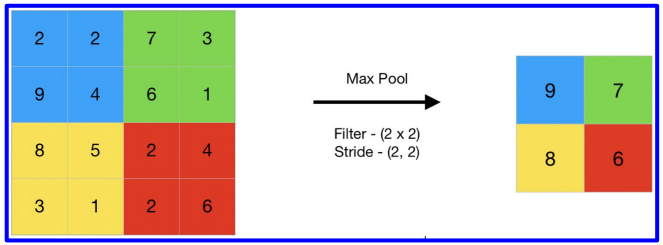
Layers in deep learning model



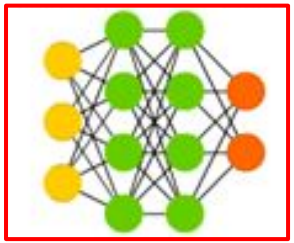
Convolution



ReLu

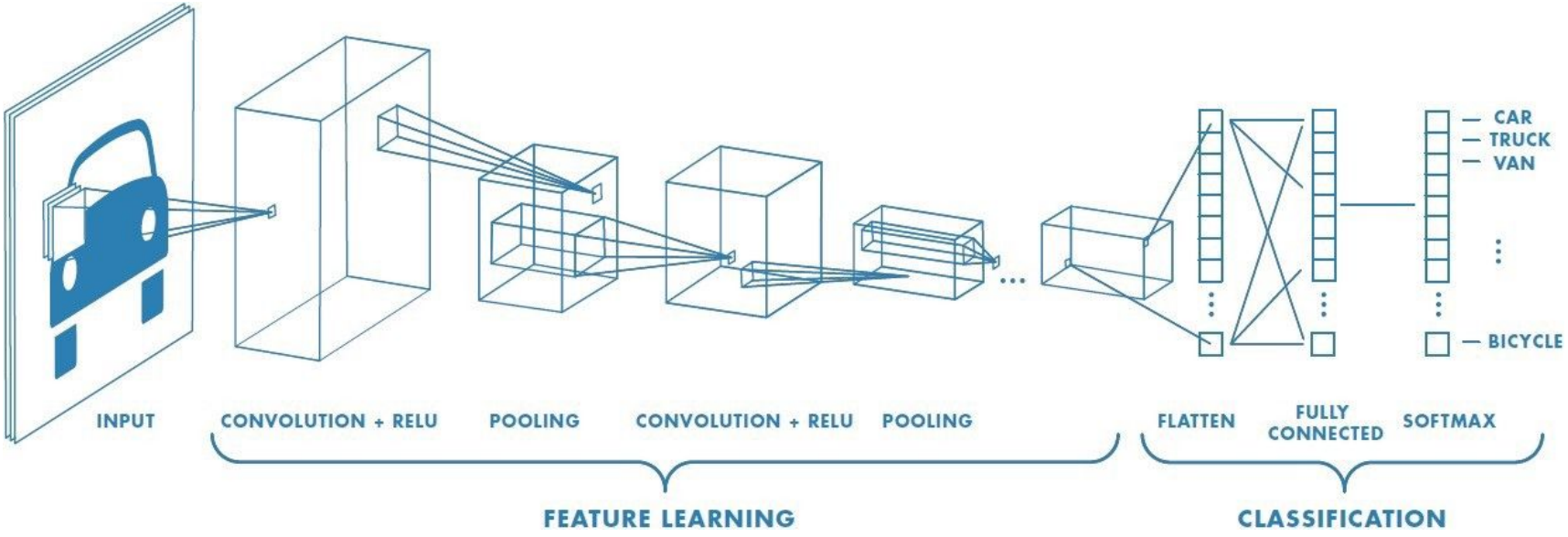


Pooling



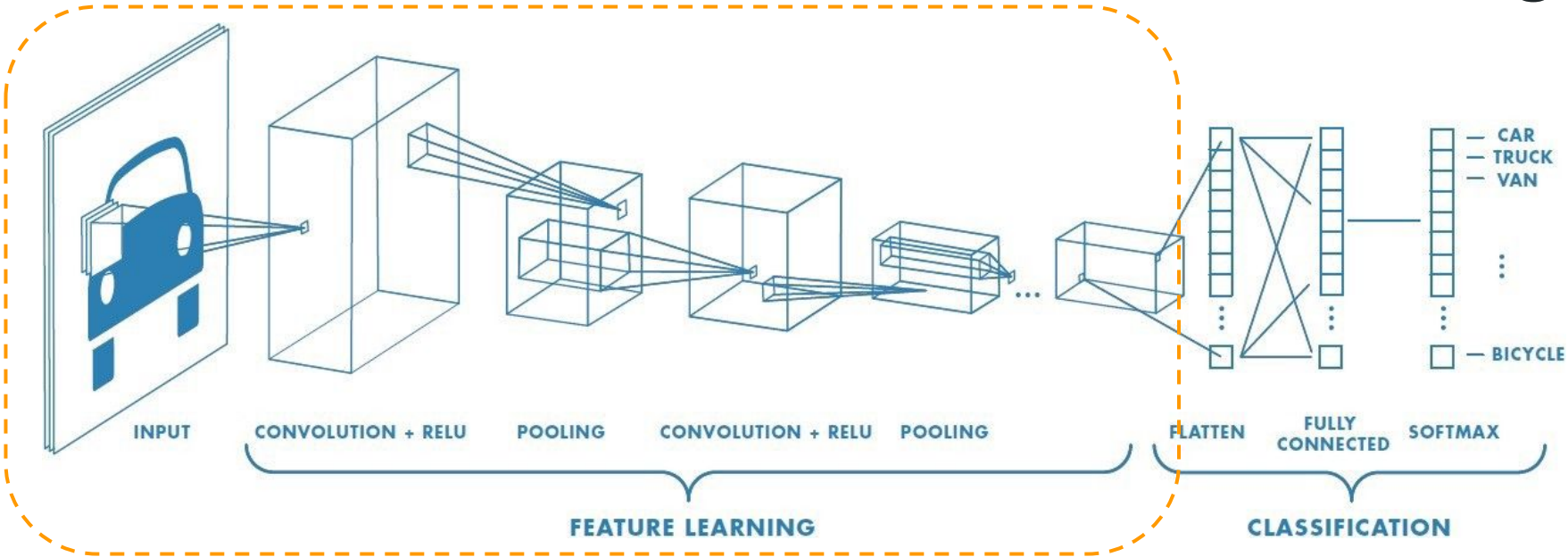
Fully connected layer

CNNs for Classification: Feature Learning



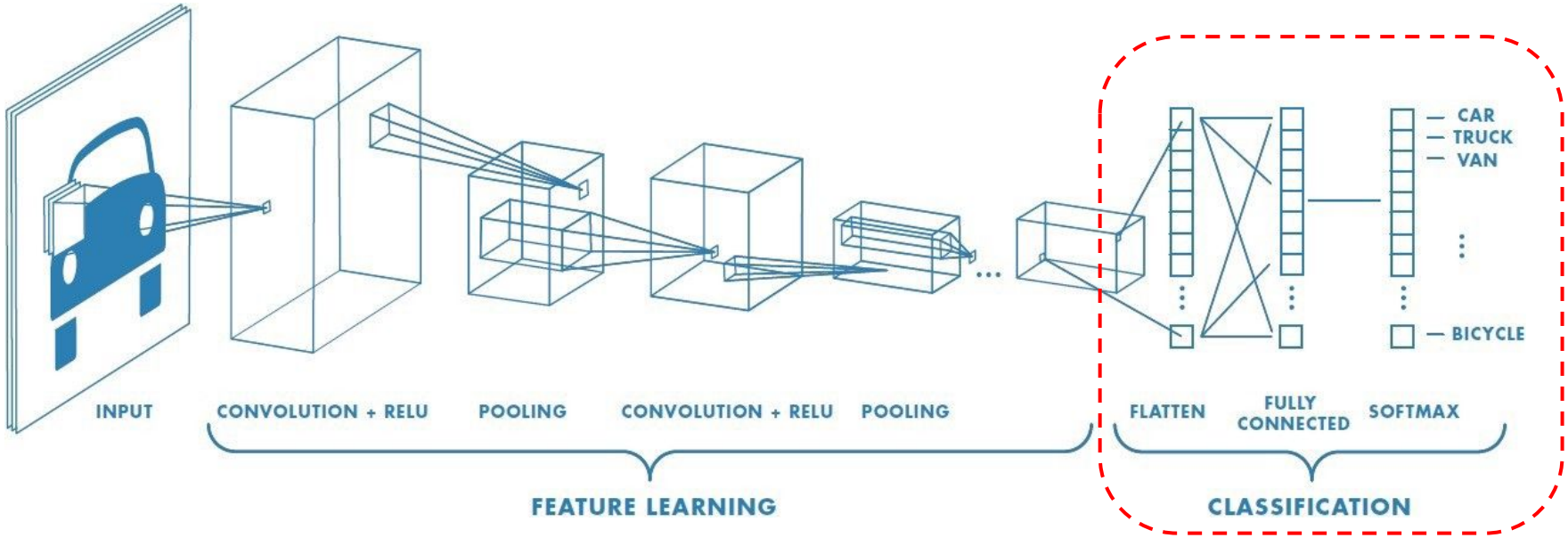
https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

CNNs for Classification: Feature Learning



1. Learn features in input image through **convolution**
2. Introduce **non-linearity** through activation function (real-world data is non-linear!)
3. Reduce dimensionality and preserve spatial invariance with **pooling**

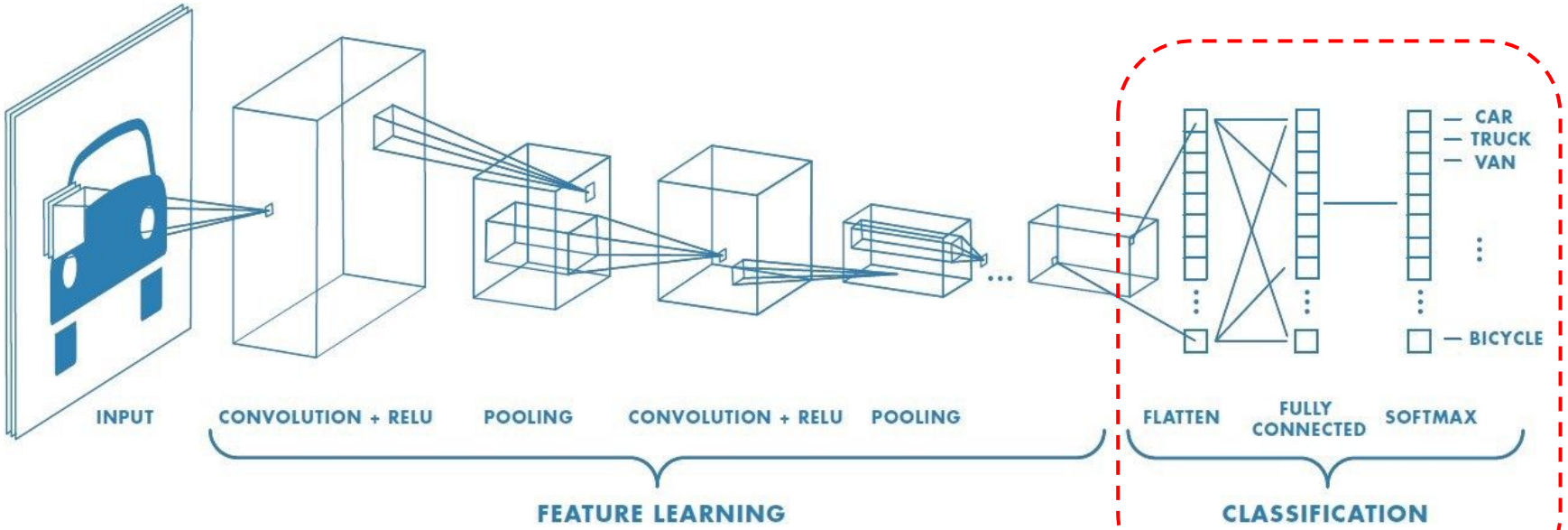
CNNs for Classification: Feature Learning



- CONV and POOL layers output high-level features of input
- Fully connected layer uses these features for classifying input image
- Express output as **probability** of image belonging to a particular class

$$\text{softmax}(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}}$$

An Architecture for Many Applications

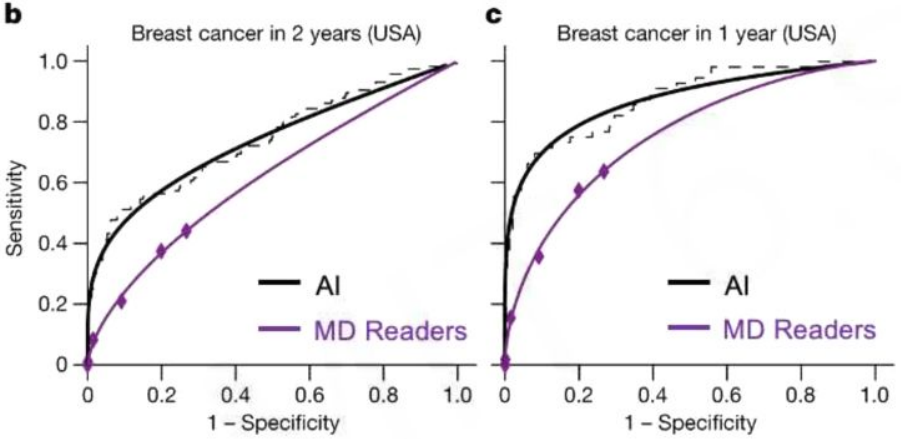


Classification
Object detection
Segmentation
Probabilistic control

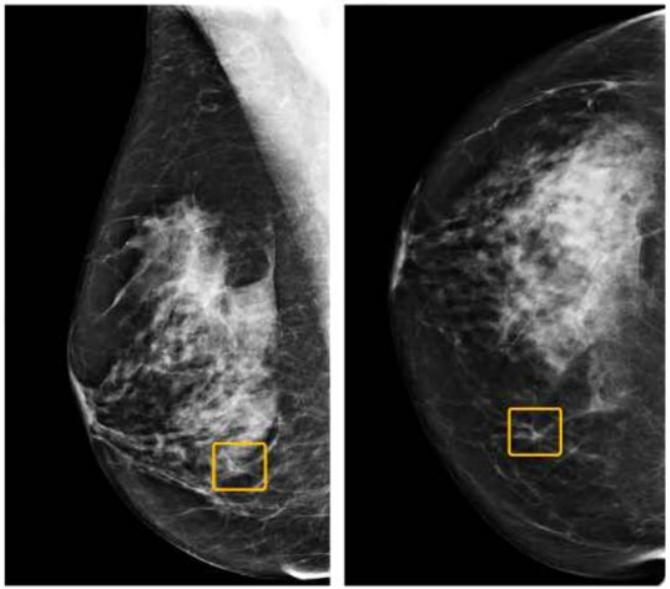
An Architecture for Many Applications

Classification: Breast Cancer Screening

International evaluation of an AI system for breast cancer screening nature

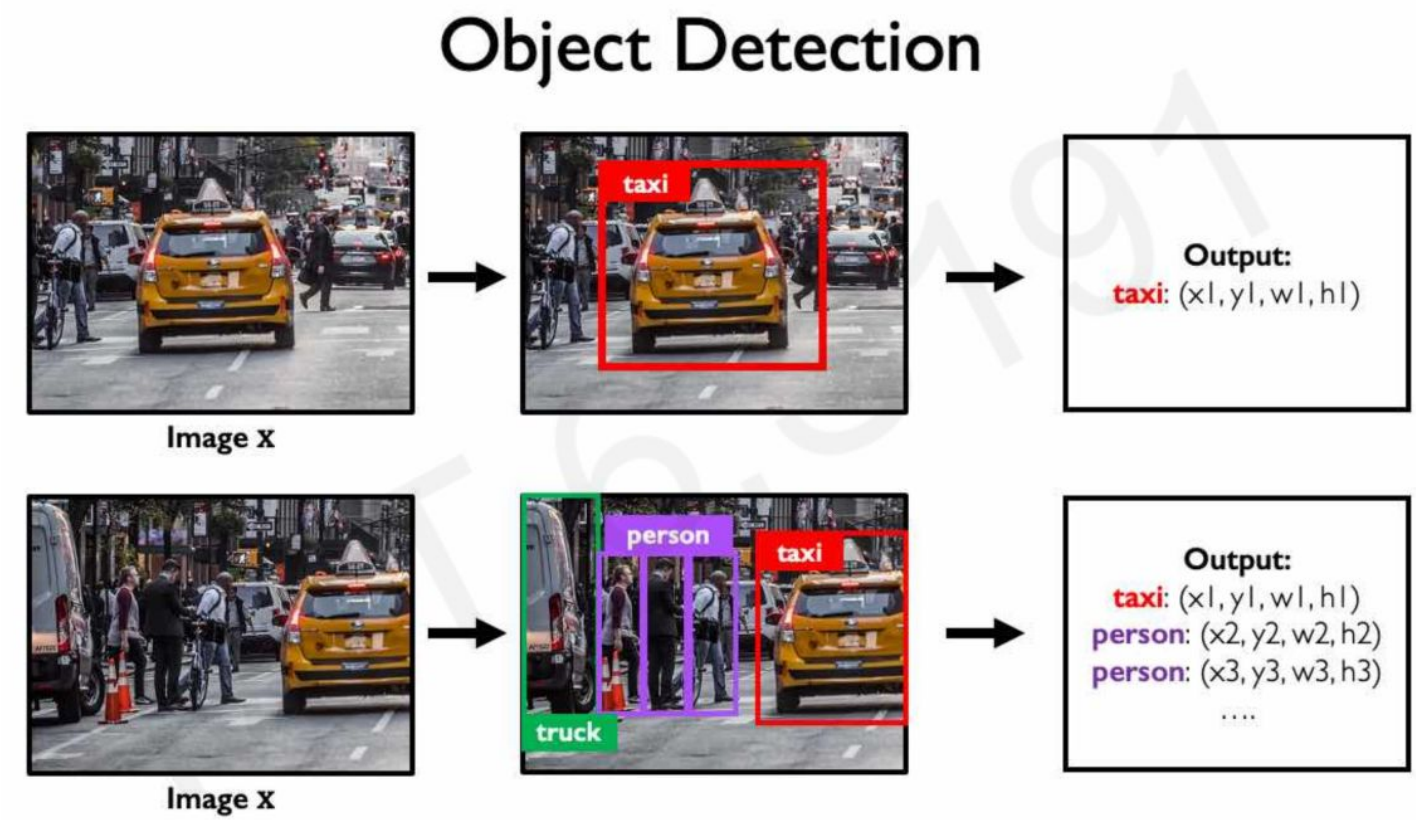


CNN-based system outperformed expert radiologists at detecting breast cancer from mammograms



Breast cancer case missed by radiologist but detected by AI

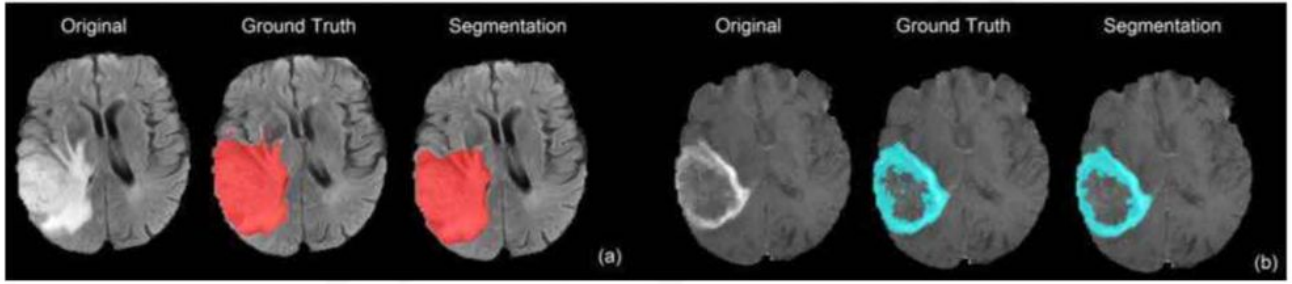
An Architecture for Many Applications



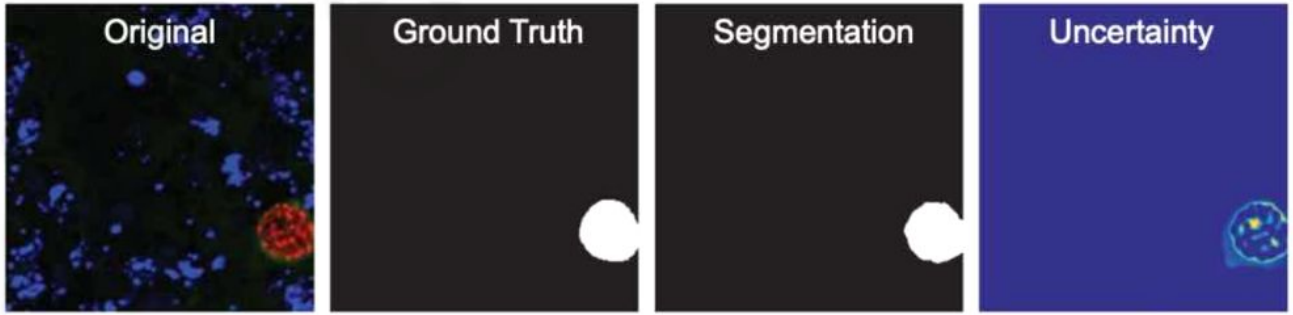
An Architecture for Many Applications

Semantic Segmentation: Biomedical Image Analysis

Brain Tumors
Dong+ *MIUA* 2017.



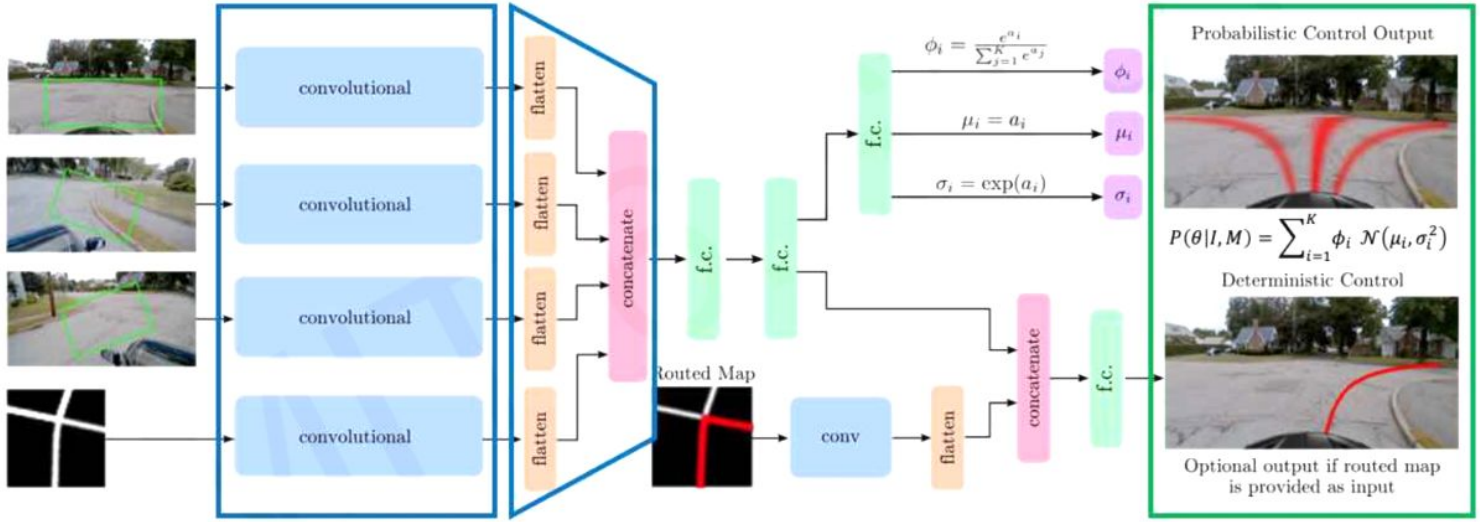
Malaria Infection
Soleimany+ *arXiv* 2019.



An Architecture for Many Applications

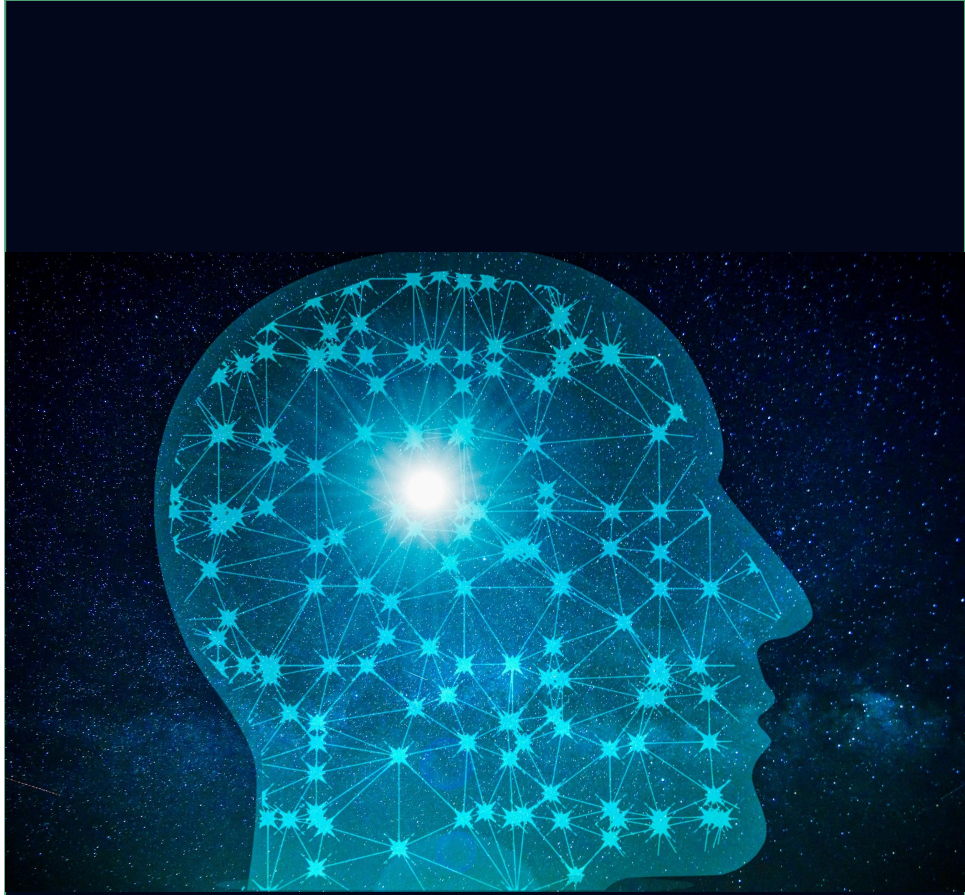
End-to-End Framework for Autonomous Navigation

Entire model is trained end-to-end **without any human labelling or annotations**



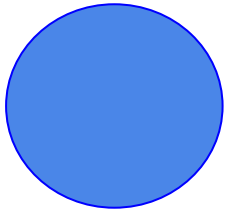
$$L = -\log(P(\theta|I,M))$$

Deep Sequence Modeling

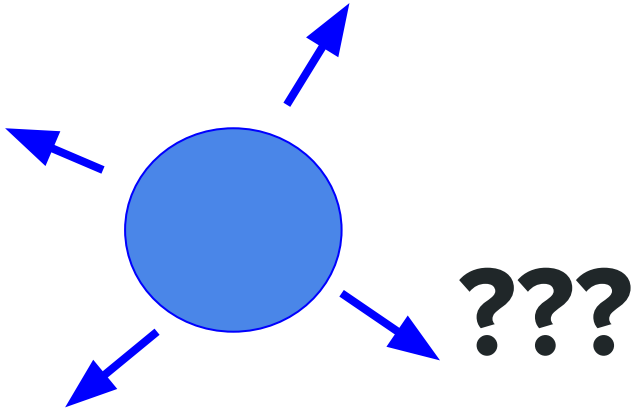


http://introtodeeplearning.com/

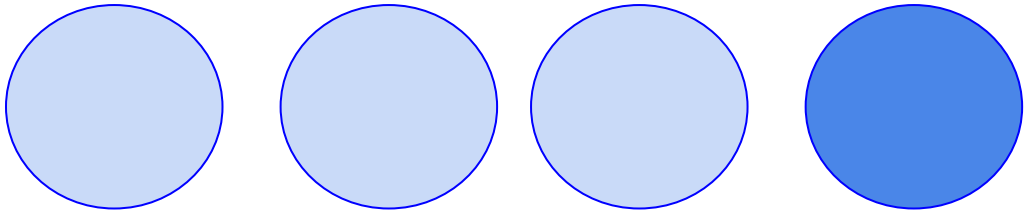
Given an images of a ball, can you predict where it will go next ?



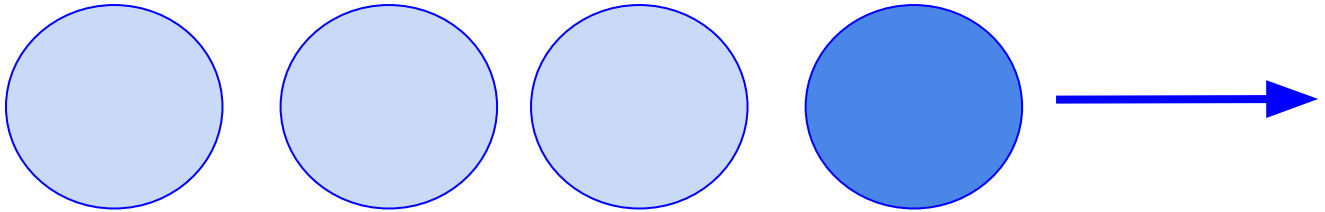
Given an images of a ball, can you predict where it will go next ?



Given an images of a ball, can you predict where it will go next ?



Given an images of a ball, can you predict where it will go next ?

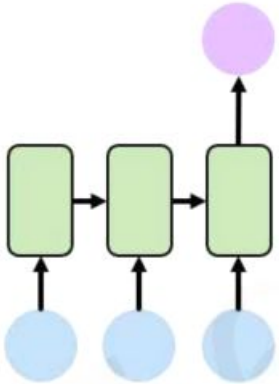


Sequence in the Wild

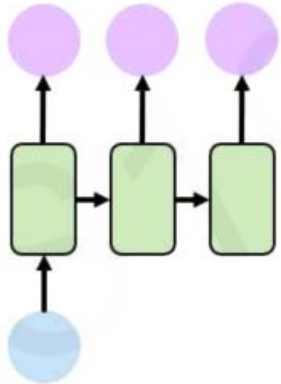
Sequence Modeling Architectures

One to One
Binary Classification

"Will I pass this class?"
Student → Pass?

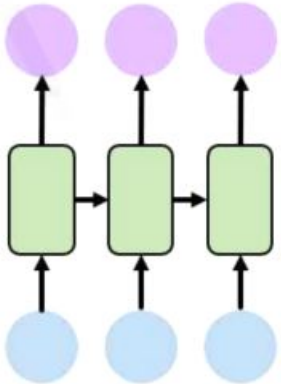


Many to One
Sentiment Classification



One to Many
Image Captioning

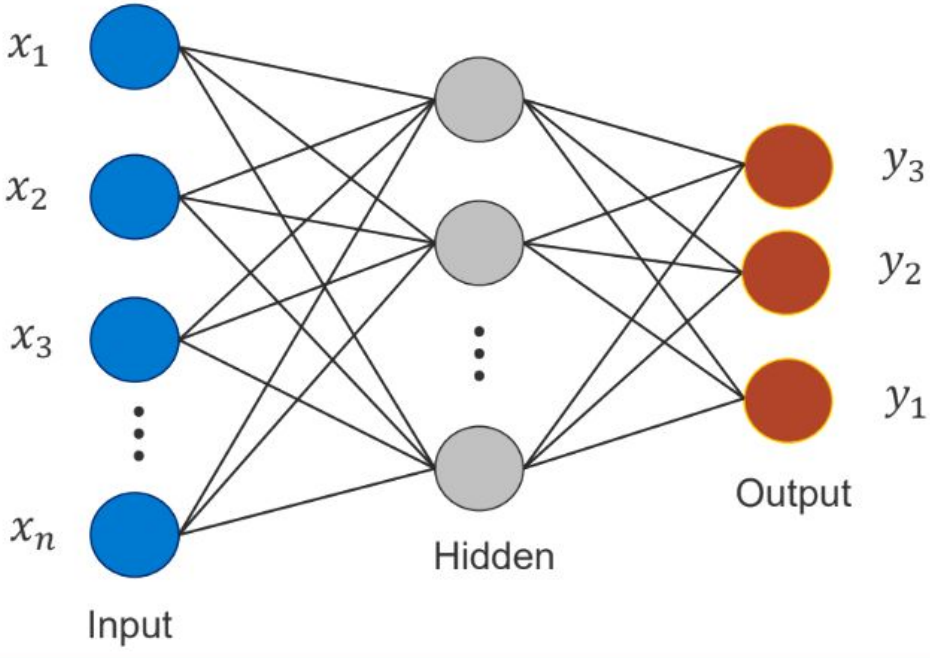
"A baseball player throws a ball."



Many to Many
Machine Translation

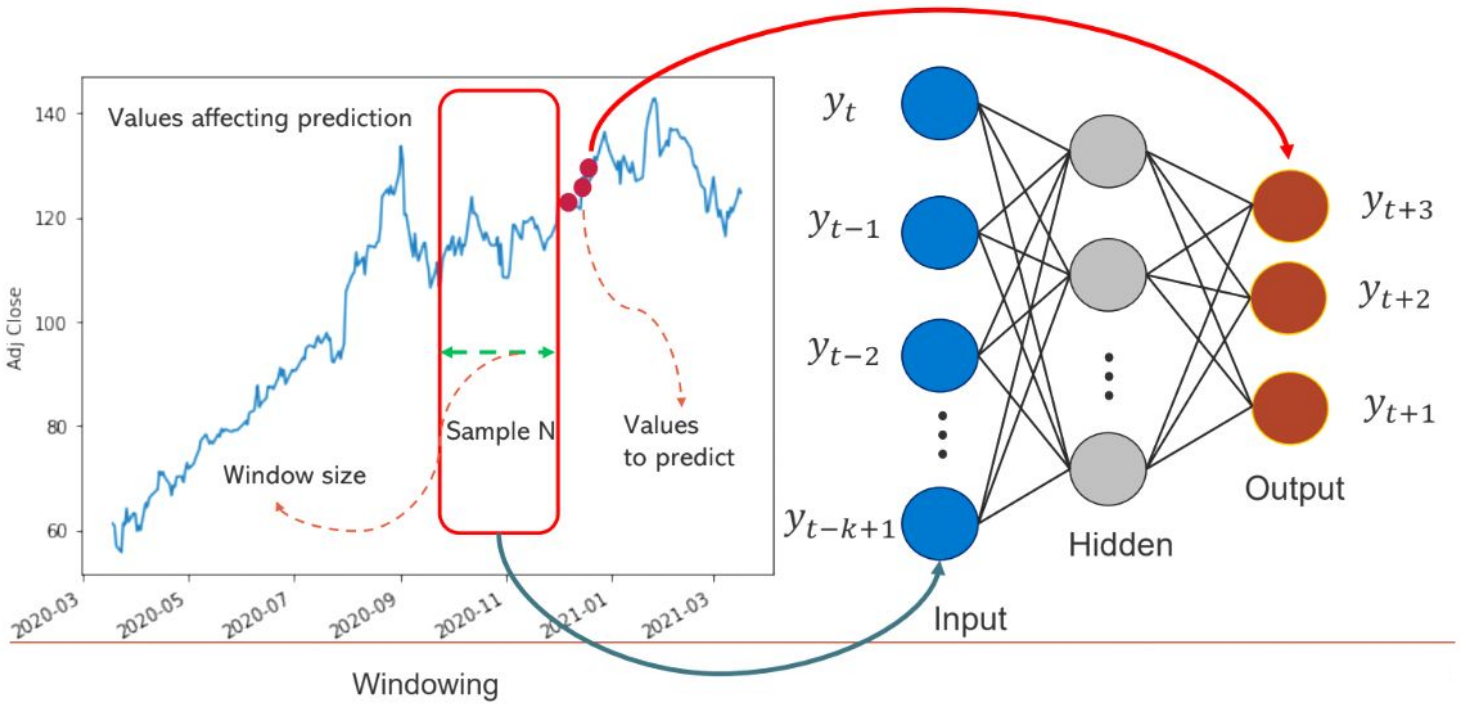
Neurons with Recurrence

Standard Neural Networks



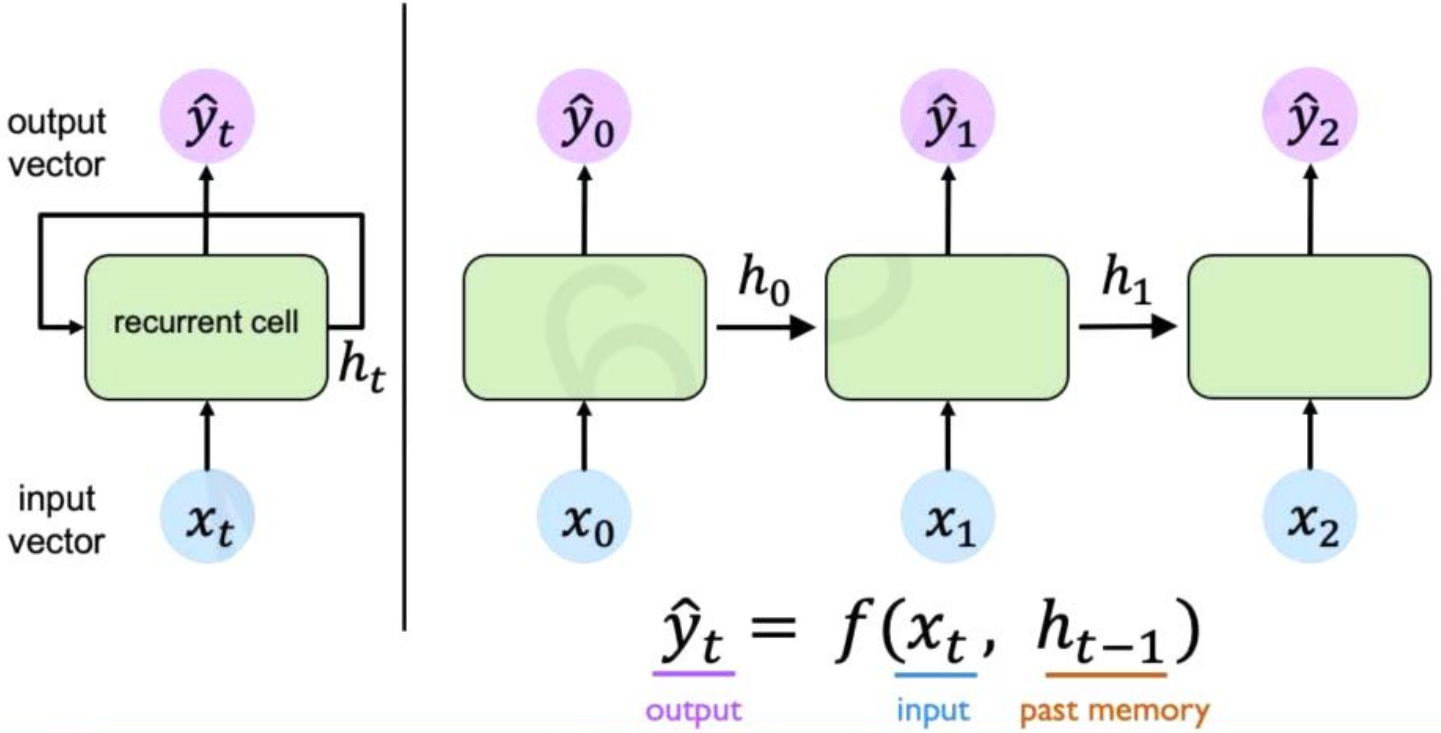
Neurons with Recurrence

Handling Individual Time Steps

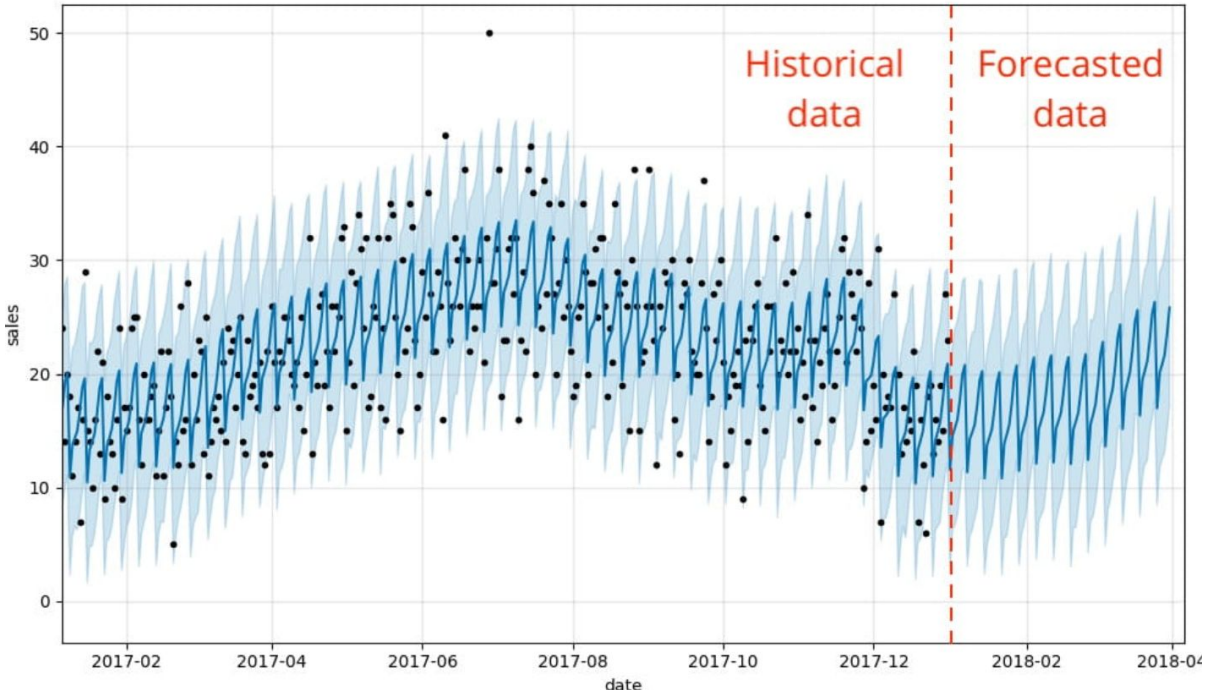


Neurons with Recurrence

Recurrent neural networks (RNNs)



A Sequence Modeling Problems: **Time Series Forecasting**



https://vitalflux.com/different-types-of-time-series-forecasting-models/

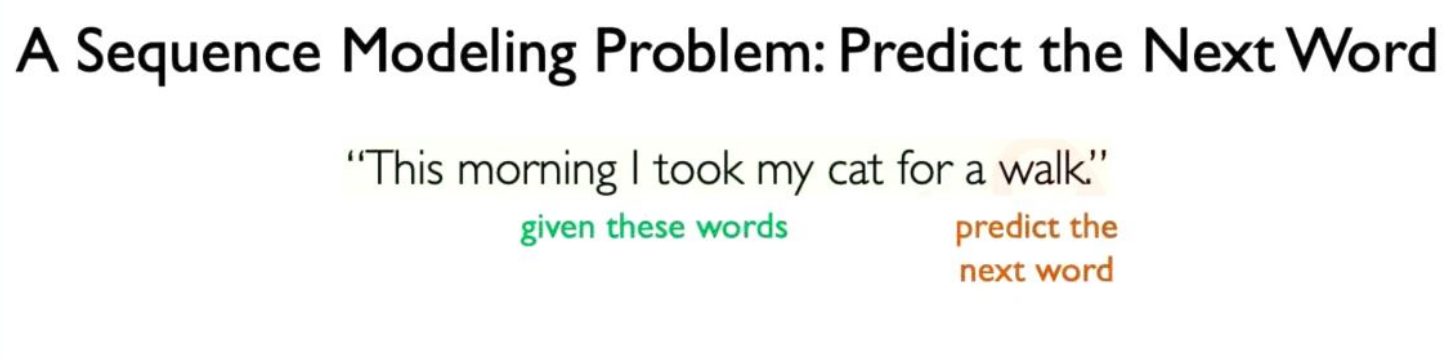
A Sequence Modeling Problems: **Natural Language Processing**

A Sequence Modeling Problem: Predict the Next Word

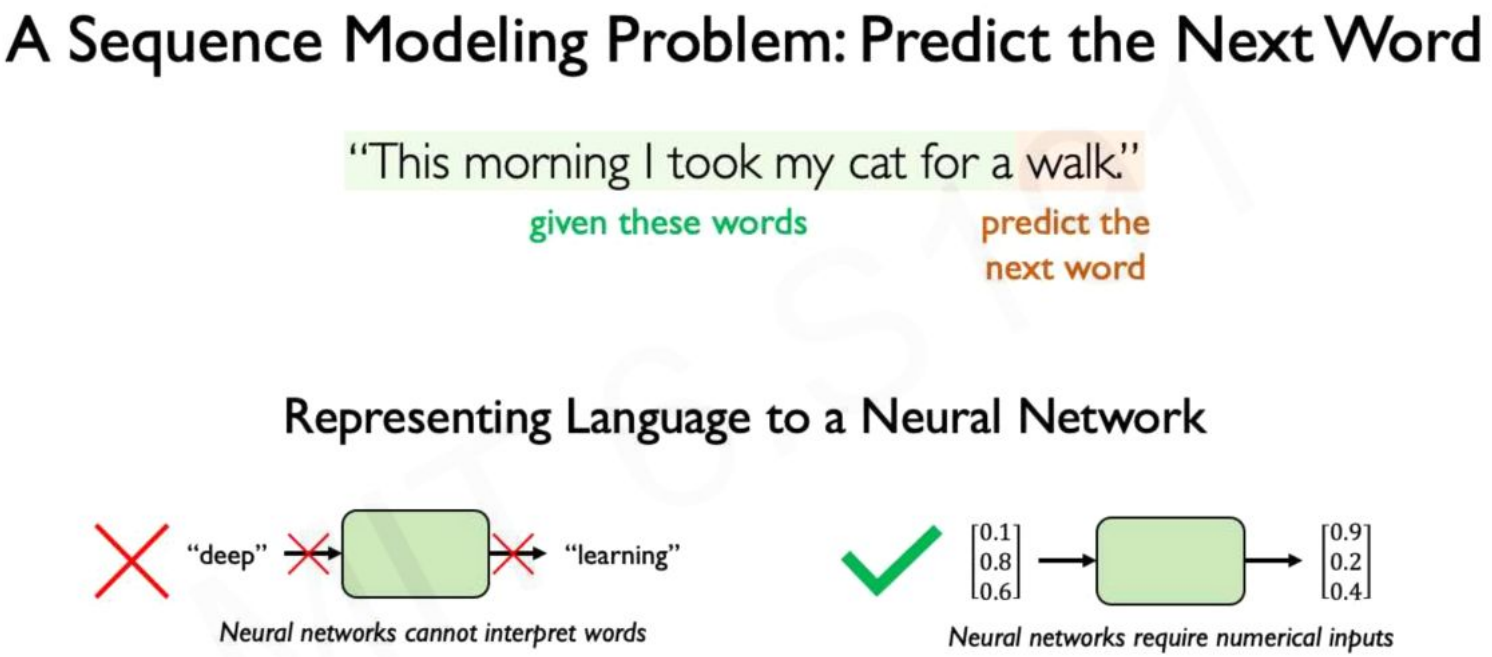
"This morning I took my cat for a walk."

A Sequence Modeling Problems: **Natural Language Processing**

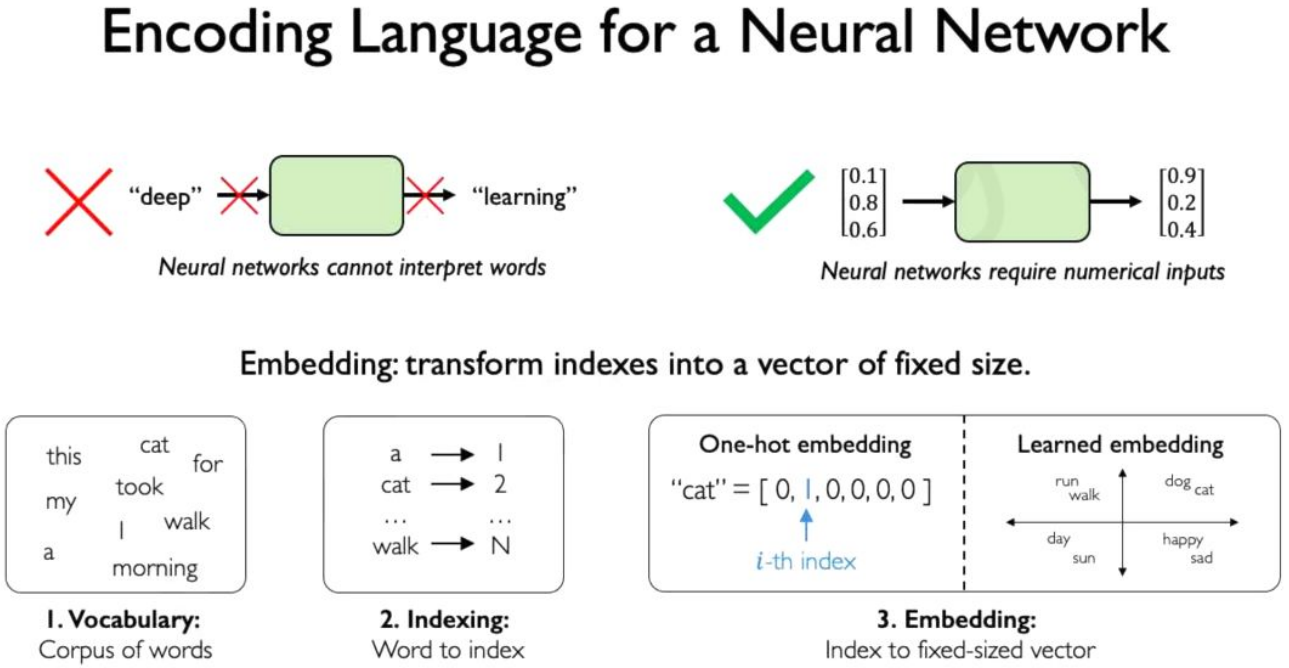
A Sequence Modeling Problems: **Natural Language Processing**



A Sequence Modeling Problems: **Natural Language Processing**



A Sequence Modeling Problems: **Natural Language Processing**



A Sequence Modeling Problems: **Natural Language Processing**

Handle Variable Sequence Lengths

The food was great

vs.

We visited a restaurant for lunch

vs.

We were hungry but cleaned the house before eating

A Sequence Modeling Problems: **Natural Language Processing**

Model Long-Term Dependencies

"**France** is where I grew up, but I now live in Boston. I speak fluent ___."

We need information from **the distant past** to accurately predict the correct word.

A Sequence Modeling Problems: **Natural Language Processing**

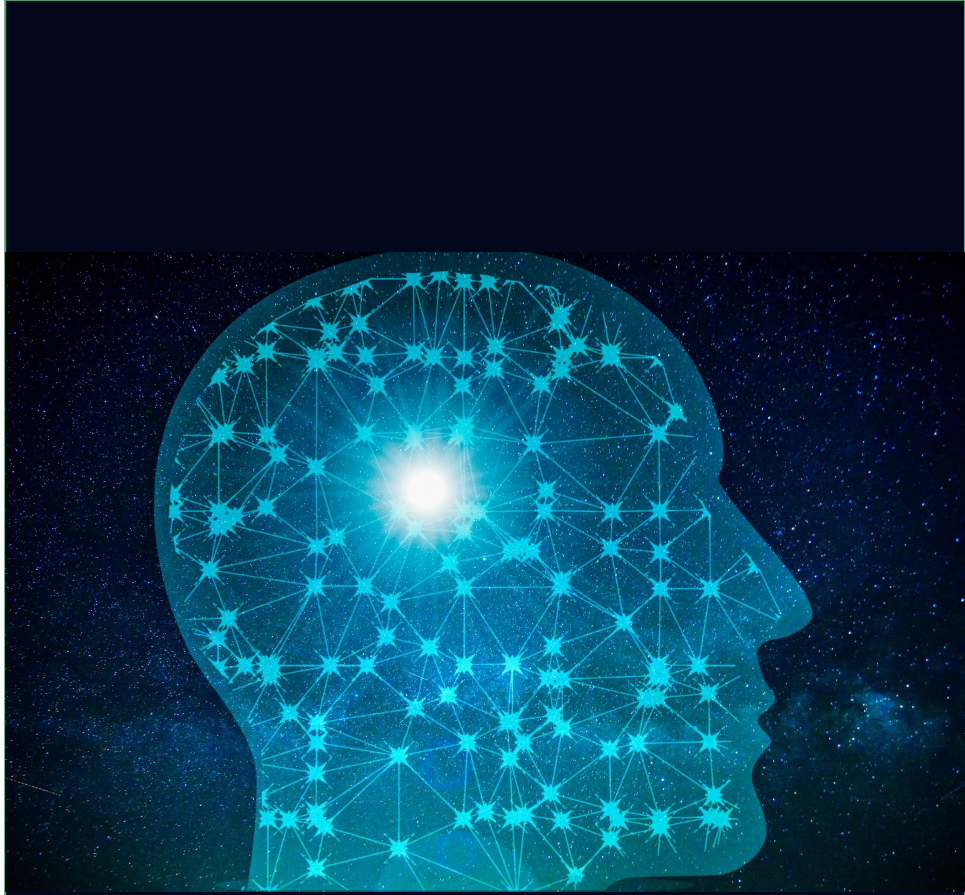
Capture Differences in Sequence Order

The food was good, not bad at all.

vs.

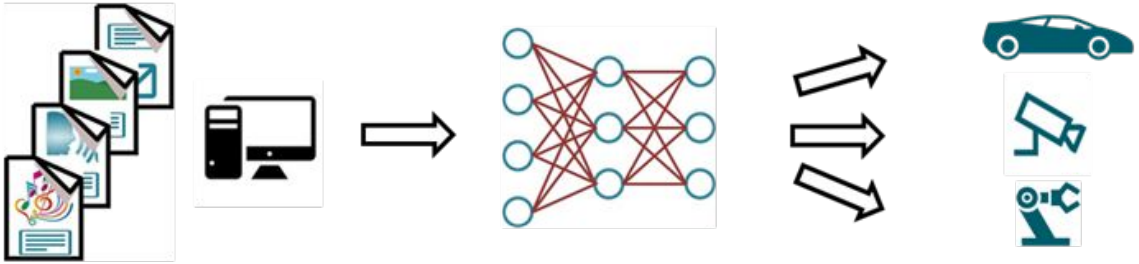
The food was bad, not good at all.

Deep Learning Development Flow



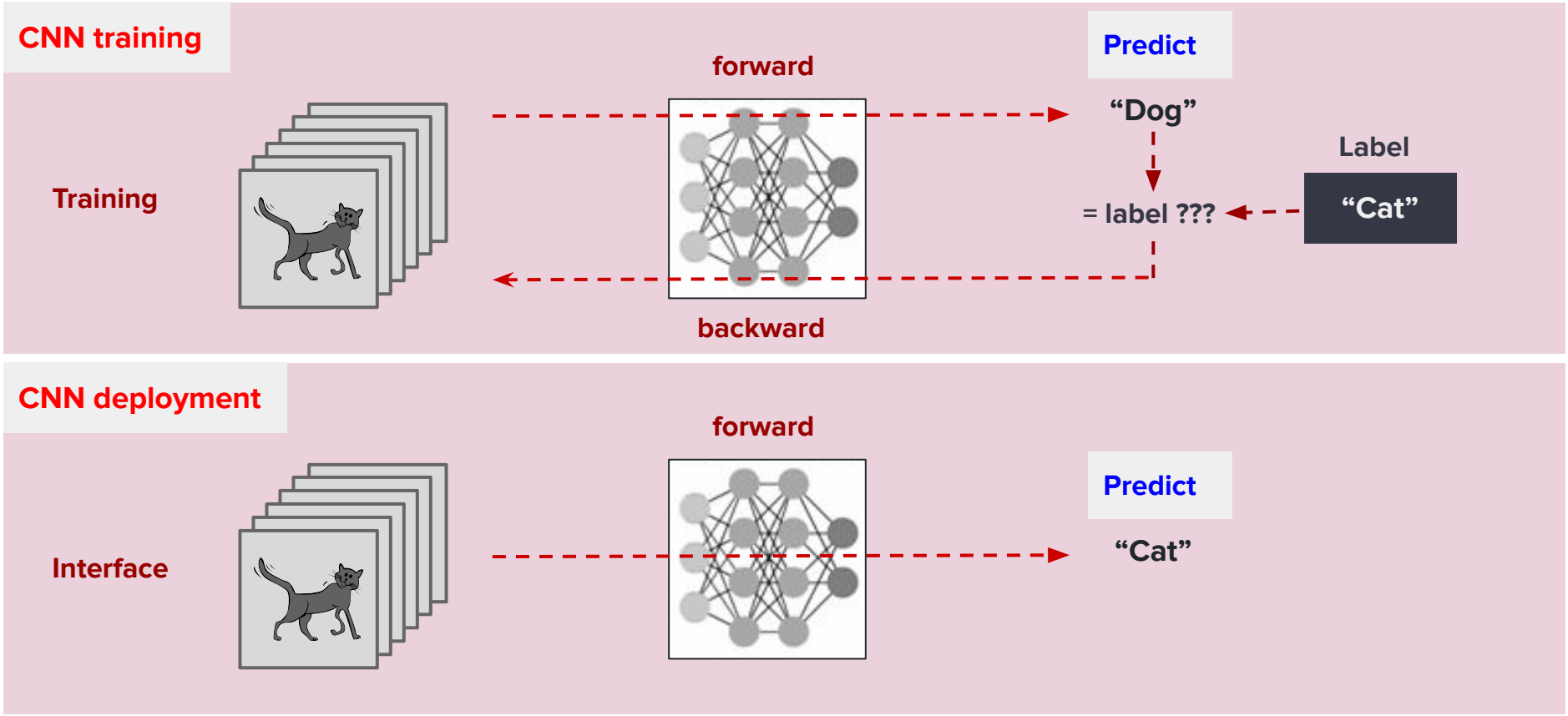
Deep Learning Development Flow

1. Selection of a framework for development
2. Selecting labeled data set of classes to train the network upon
3. Designing initial network model
4. Training the network
5. Saving the parameters and architecture in a binary file
6. Inference

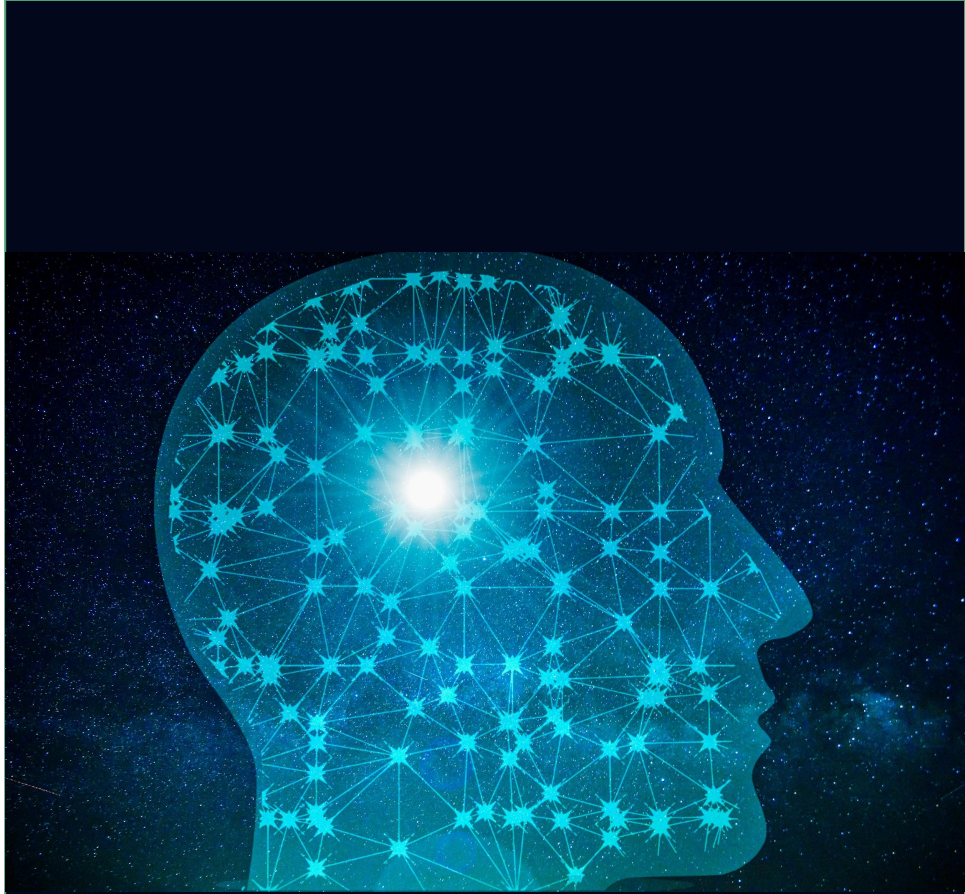


https://e2e.ti.com/blogs_/b/industrial_strength/posts/ti-brings-deep-learning-inference-to-the-edge

Deep Learning Development Flow



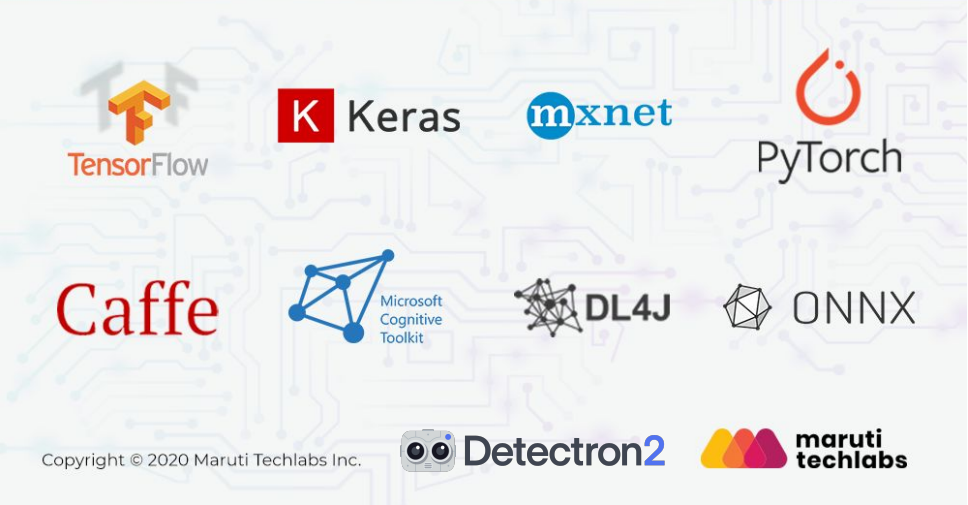
Deep Learning Frameworks and Interesting Datasets



Deep Learning Frameworks

Popular Frameworks

- **TensorFlow:**
  - Developed by Google
  - The most used in deep learning framework
  - Based on Github and forks and Stack Overflow
- **Keras:**
  - There are also interfaces that are wrapped around one or multiple frameworks in TensorFlow.
  - The most well-known and widely-used interface for deep learning.
- **PyTorch and Detectron:**
  - Developed by Facebook.
  - Popular for CNN modeling on object detection



Datasets for Machine Learning and Deep learning Projects

- **50 Public Datasets for ML and DL ---> https://www.ml4devs.com/articles/datasets-for-machine-learning-and-data-science/**
- **Datasets on Kaggle ---> https://www.kaggle.com/datasets**
- **UCI ML Repository ---> https://archive.ics.uci.edu/ml/datasets.php**
- **Harvard Dataverse ---> https://dataverse.harvard.edu/**

Good luck 😉


Uppsala universitet 1852–1916
Studenten, Staden och Samingen
UPPSALA UNIVERSITE