

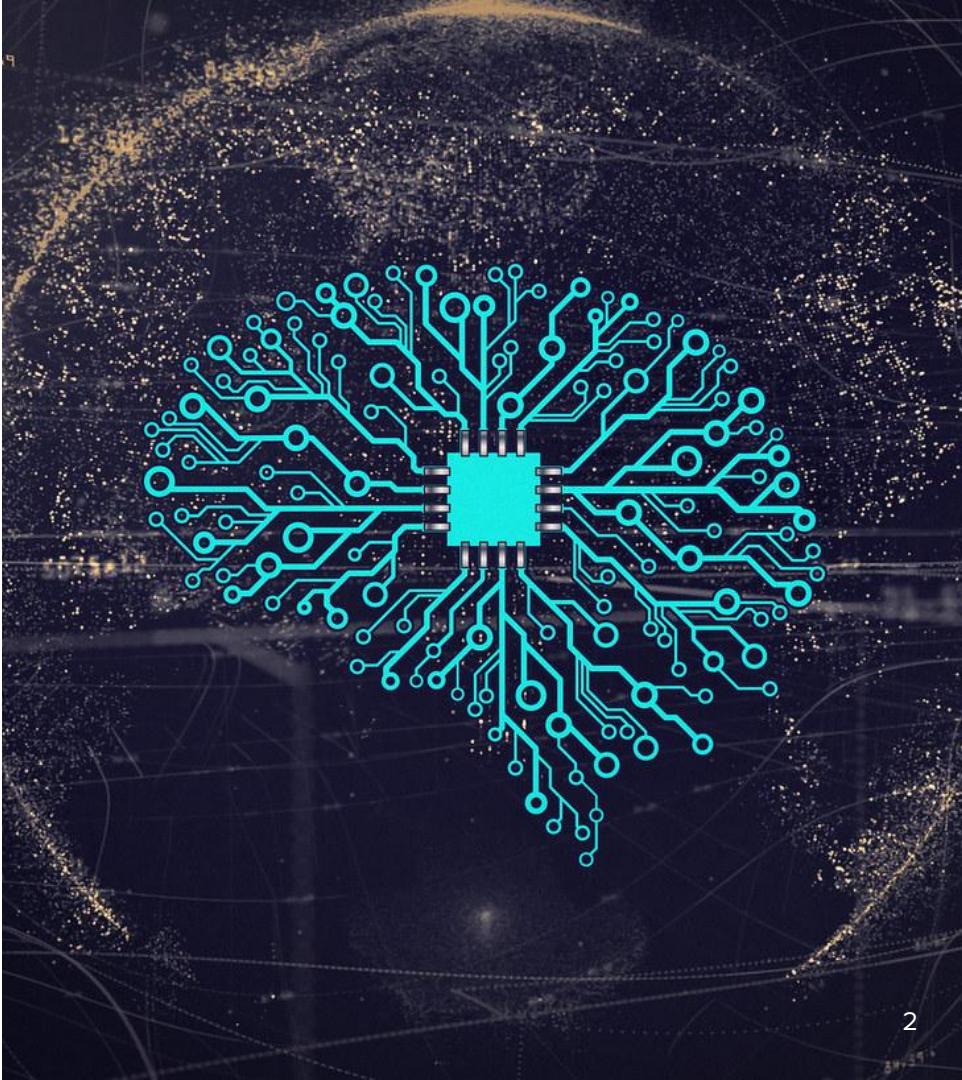
Introduction to Deep Learning

Dr. Paisit Khanarsa

FIBO, Kmutt

Outline

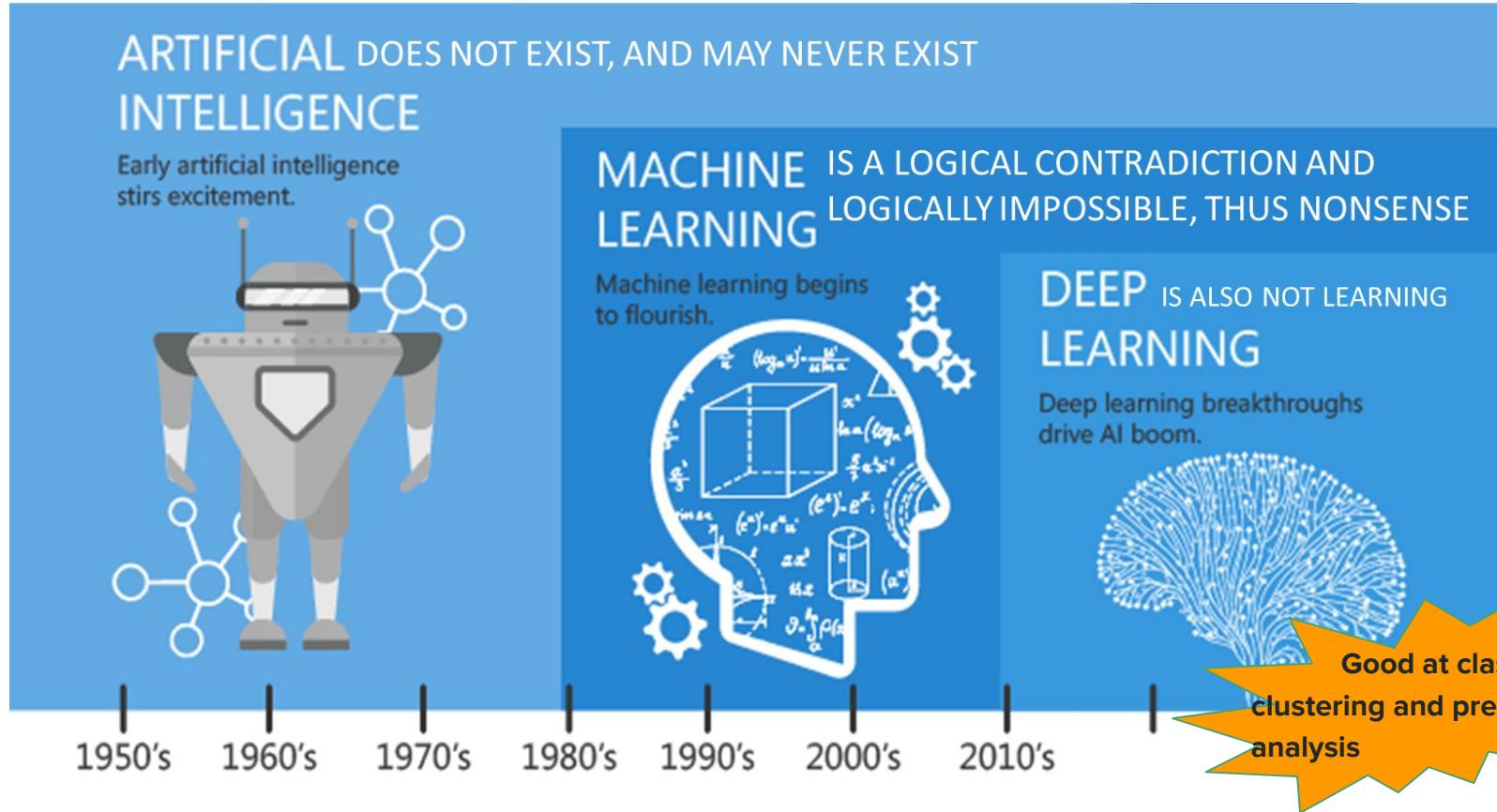
- ❖ What is Deep Learning?
- ❖ Deep Computer Vision
- ❖ Deep Sequence Modeling
- ❖ Deep Learning Development Flow
- ❖ Deep Learning Frameworks and
Interesting Datasets



What is Deep Learning?



Deep Learning Introduction



Deep Learning Introduction

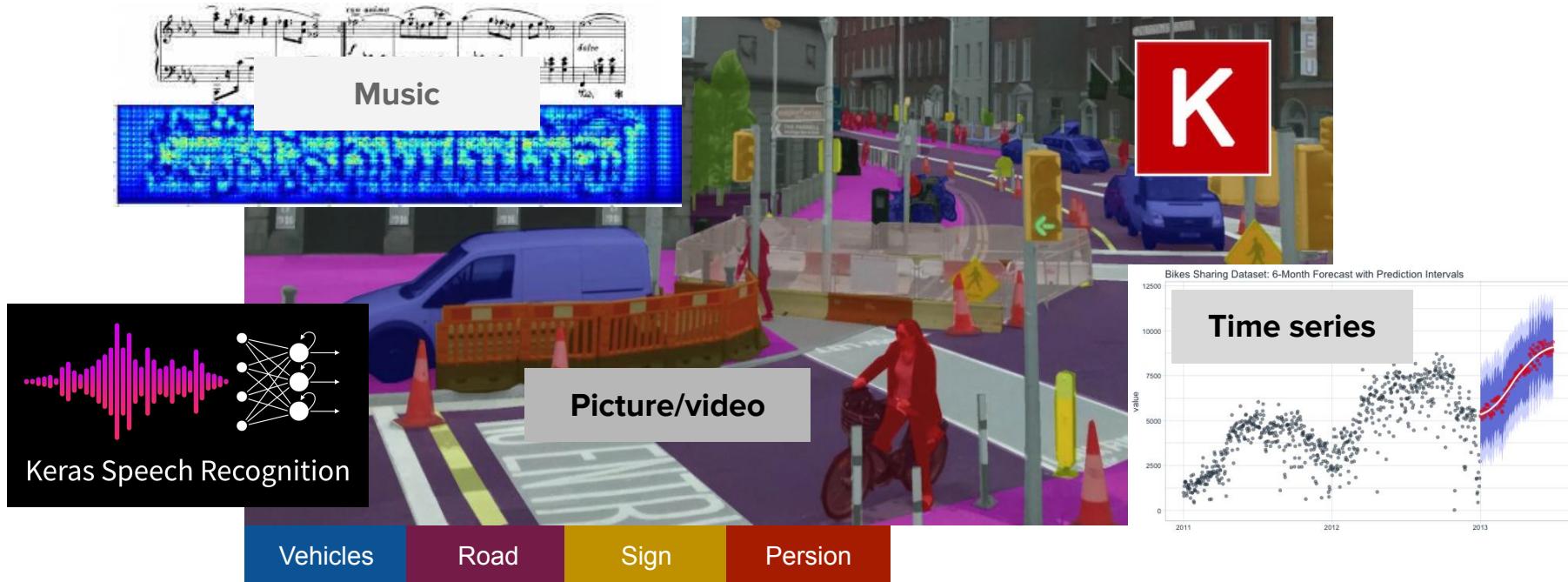
Deep learning is way of classifying, clustering, and predicting things by using a neural network that has been trained on vast amounts of data.



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

Deep Learning Introduction

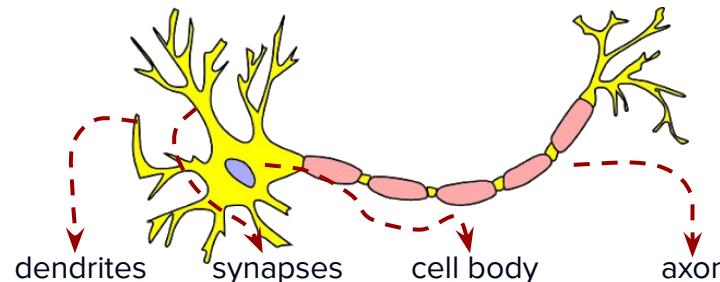
Deep learning is way of classifying, clustering, and predicting things by using a neural network that has been trained on vast amounts of data.



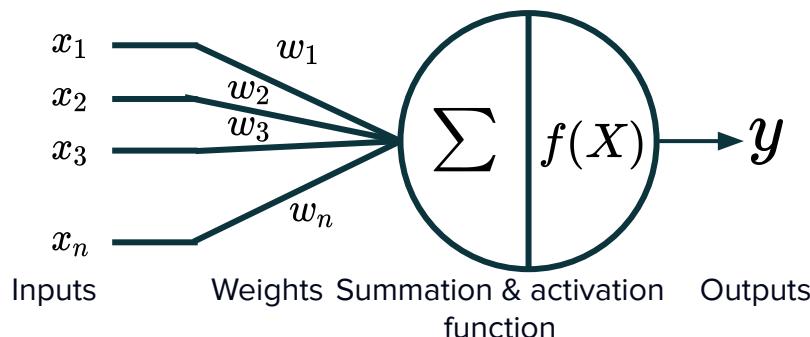
What is deep learning?

- ❖ 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**.

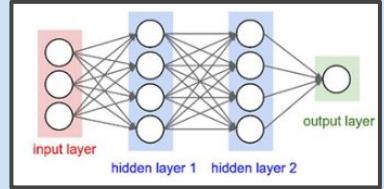
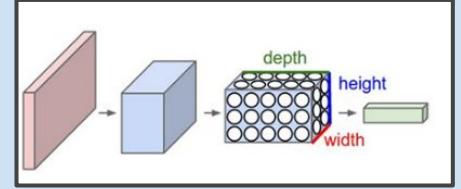
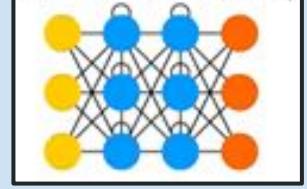
Biological neuron



Artificial neuron



Deep Neural Networks (DNN)

Multi layer perceptron (MLP)	<ul style="list-style-type: none">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.	<p>Fully connected layer</p>  <p>input layer hidden layer 1 hidden layer 2 output layer</p>
Convolutional neural network (CNN) (Deep Computer Vision)	<ul style="list-style-type: none">Feed forward deep neural network.Application for object detection, classification and segmentation.Popular uses for image and video processing.	<p>Feed-forward network</p>  <p>depth height width</p>
Recurrent neural network (RNN) (Deep Sequence Modeling)	<ul style="list-style-type: none">Feed forward deep neural network, extended to feedback connections.Ideal for text, time series and speech analysis.	<p>Cyclical connect</p> 

Deep Computer Vision

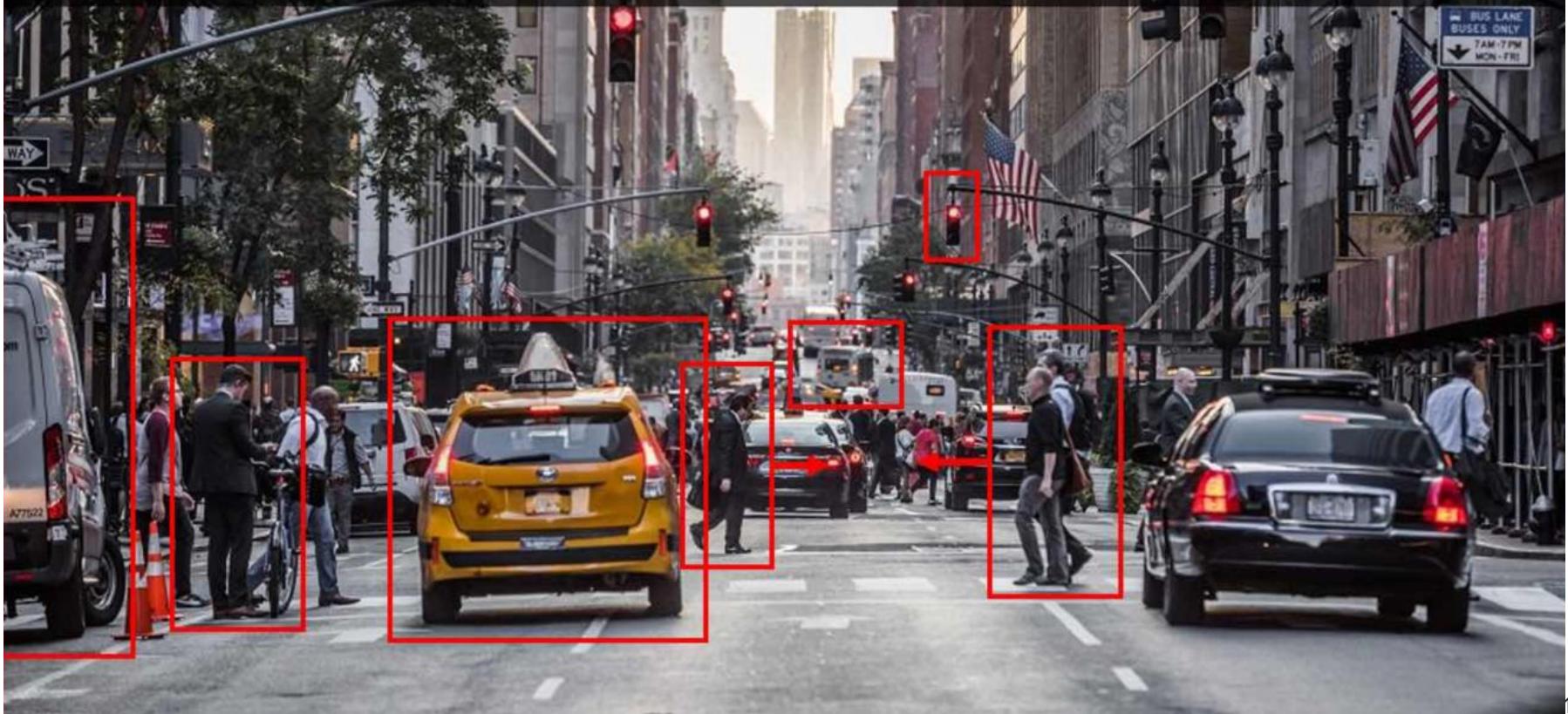
<http://introtodeeplearning.com/>





**“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



To rise and impact of computer vision

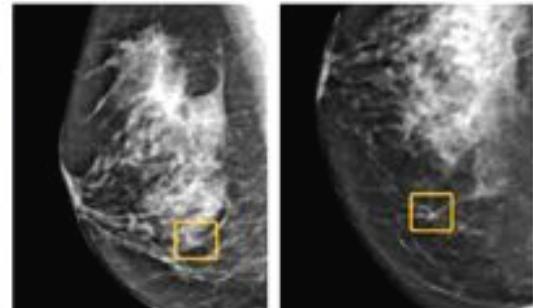
Robotics



Accessibility



Biology & Medicine



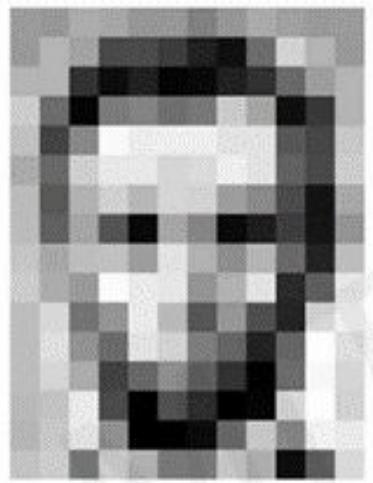
Autonomous driving



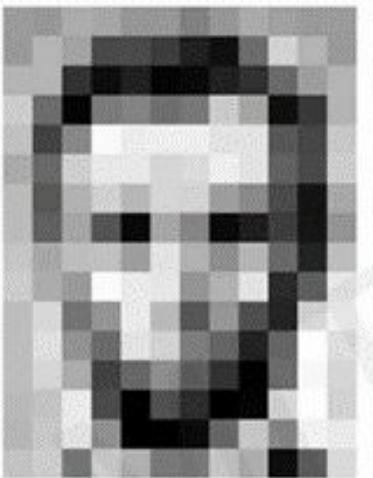
Mobile computing



What computers see



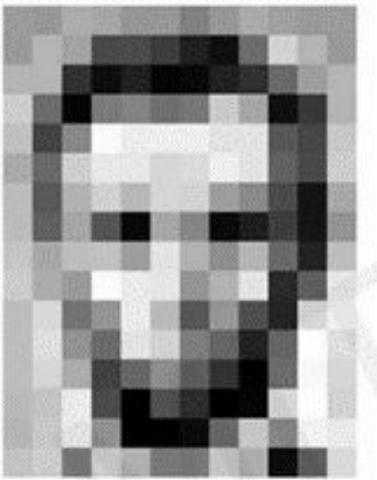
What computers see



197	188	174	148	180	192	129	181	172	143	188	198
180	182	188	74	76	62	38	17	116	234	180	184
180	180	80	14	34	6	10	33	40	196	183	181
204	198	5	128	187	191	180	204	184	16	54	180
194	88	137	261	237	239	239	238	227	87	71	201
172	204	207	233	233	214	235	239	238	88	74	206
188	94	178	209	186	210	211	188	188	76	36	189
188	87	188	54	18	188	758	11	31	42	22	146
199	146	192	189	186	227	179	142	142	226	36	180
208	178	188	282	236	231	148	179	238	43	35	234
190	214	158	146	236	247	88	186	79	38	238	241
196	234	147	166	327	210	227	188	36	191	288	234
196	214	173	54	109	143	36	50	2	236	249	219
187	196	236	75	1	87	47	5	6	217	254	211
189	262	237	149	8	6	12	104	209	129	243	236
196	204	128	207	177	131	129	206	179	18	34	215

What computers see

Images are Numbers



187	188	174	148	186	153	129	161	172	145	188	196
188	182	181	174	25	62	39	57	76	230	180	164
180	180	50	14	34	6	10	33	48	106	180	181
206	198	6	134	187	111	135	204	186	16	54	180
194	48	137	261	237	239	239	236	227	57	71	201
173	204	207	232	232	214	230	239	208	57	74	206
188	60	178	209	186	218	211	188	186	70	35	140
188	87	148	84	16	148	234	31	31	42	29	146
199	148	183	183	183	227	178	145	182	106	26	180
206	178	186	263	236	231	188	178	218	49	39	234
180	216	178	145	236	187	88	186	79	38	218	241
190	234	147	156	227	210	227	186	56	19	258	234
190	214	173	94	183	143	94	96	2	196	249	218
187	196	236	73	1	81	47	9	9	217	264	211
180	202	237	148	6	9	12	104	299	138	343	236
195	206	133	207	177	121	129	200	179	13	96	218

What the computer sees

157	153	174	168	166	182	129	151	172	161	188	196
189	182	149	74	76	62	39	17	110	210	180	184
180	180	98	14	34	6	10	33	48	106	188	183
206	128	6	124	131	111	129	204	186	16	54	186
194	68	131	211	237	238	238	238	237	87	71	201
172	105	257	239	239	214	230	239	239	98	74	206
188	88	179	208	186	216	211	188	139	75	20	168
189	97	146	84	70	168	134	11	31	62	22	148
199	168	141	183	186	227	176	143	182	106	36	180
206	174	189	282	236	231	149	178	238	43	95	234
190	216	116	148	236	187	86	186	79	38	218	241
190	224	147	106	227	218	127	182	36	101	288	234
190	214	173	66	189	143	96	58	2	109	249	218
187	196	236	76	1	81	47	9	6	217	258	211
180	202	237	148	6	9	12	104	200	138	343	236
195	206	133	207	177	121	129	200	179	13	96	218

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

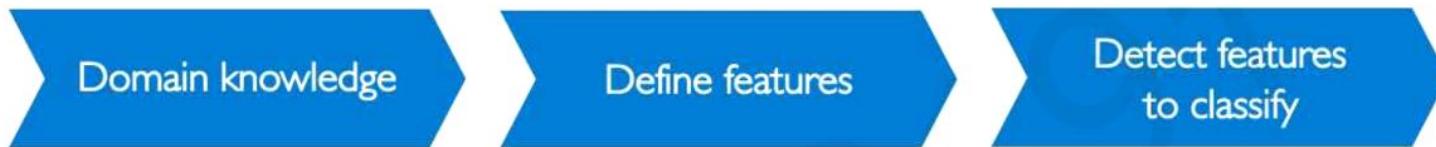


Wheels,
License Plate,
Headlight

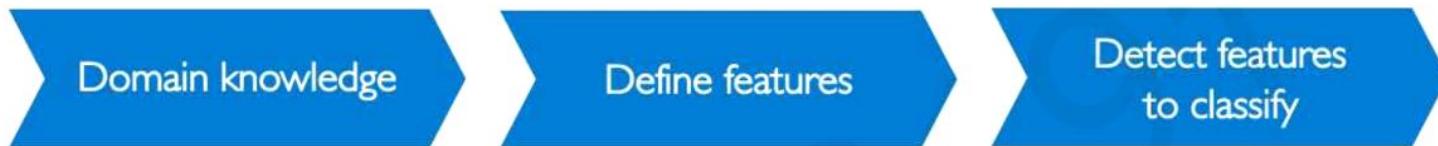


Door, Windows,
Steps

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

Viewpoint variation



Scale variation



Deformation



Occlusion



Illumination conditions



Background clutter

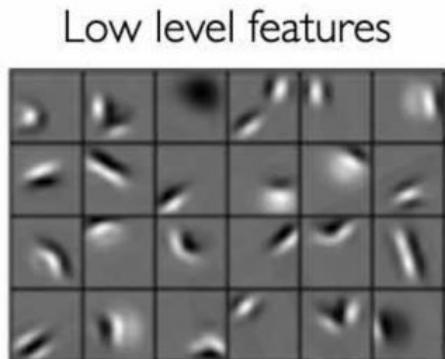


Intra-class variation

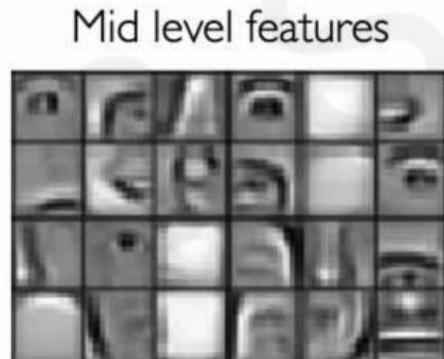


Learning Feature Representations

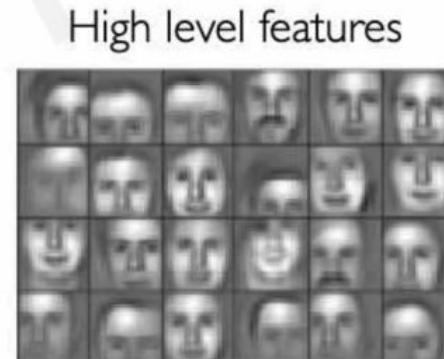
Can we learn a hierarchy of features directly from the data instead of hand engineering?



Edges, dark spots



Eyes, ears, nose

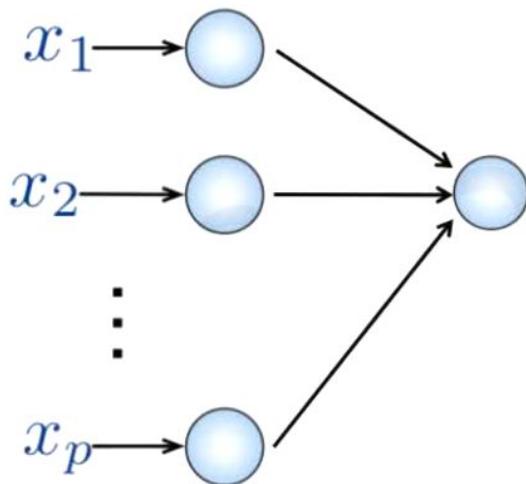


Facial structure

Learning Visual Features

Fully Connected Neural Network

- Input:**
- 2D image
 - Vector of pixel values



Fully Connected:

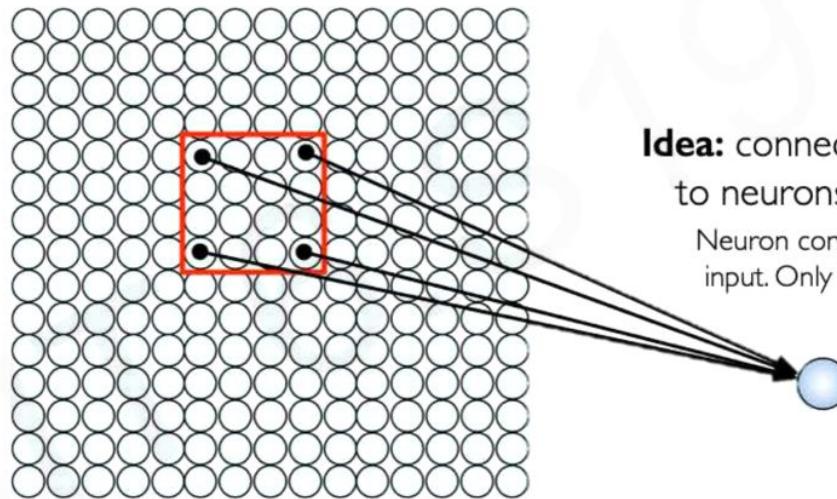
- Connect neuron in hidden layer to all neurons in input layer
- No spatial information!
- And many, many parameters

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

Learning Visual Features

Using Spatial Structure

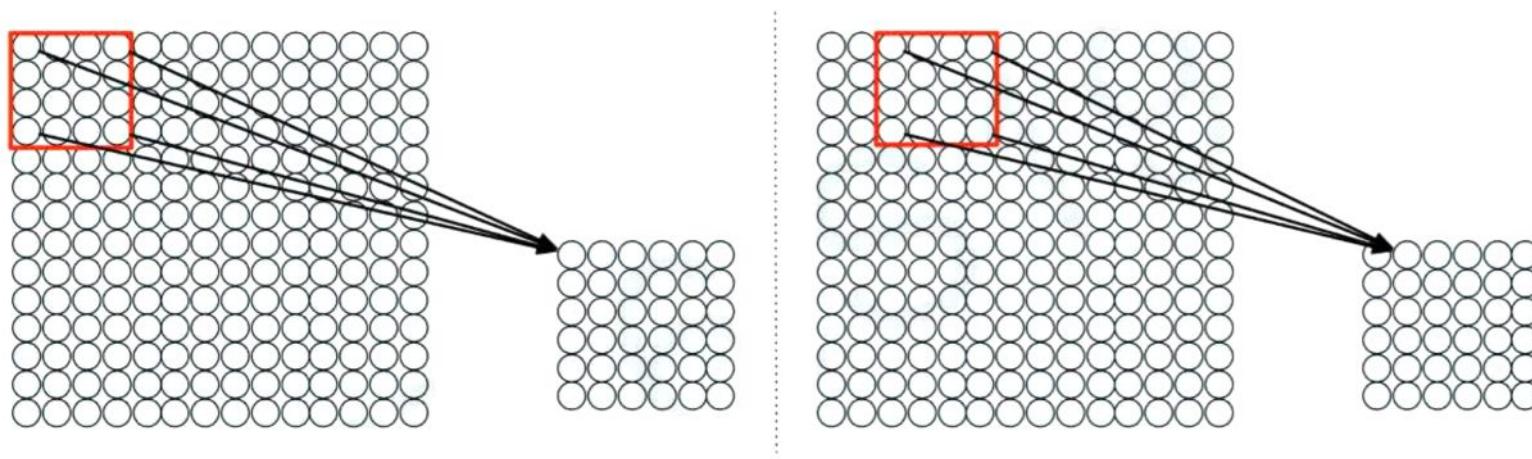
Input: 2D image.
Array of pixel values



Idea: connect patches of input to neurons in hidden layer.
Neuron connected to region of input. Only "sees" these values.

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

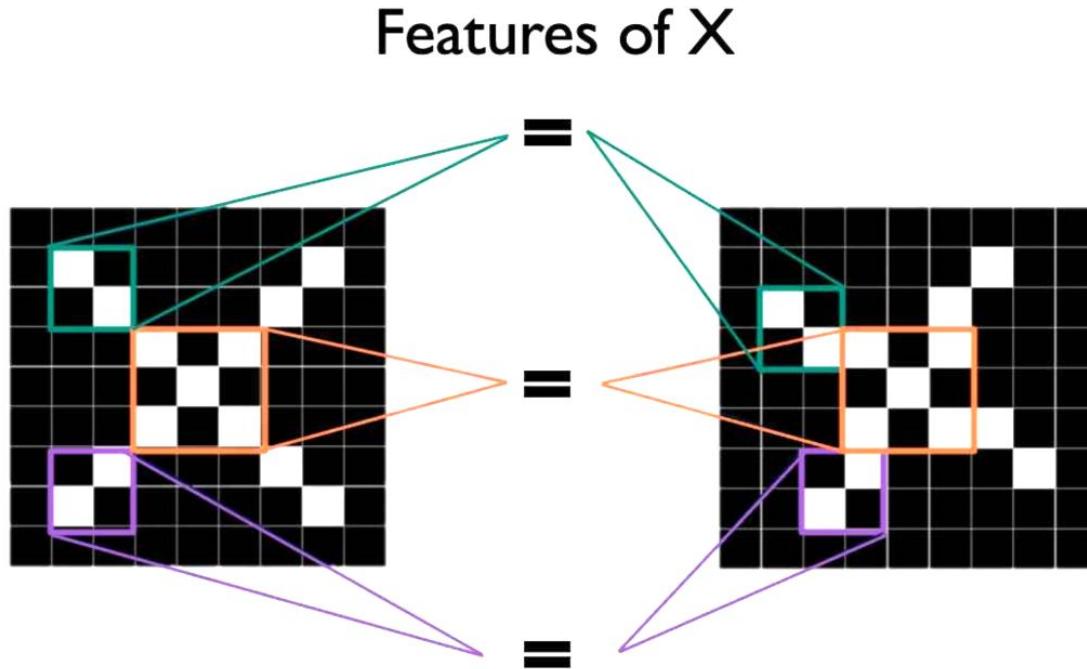
X or X?



-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

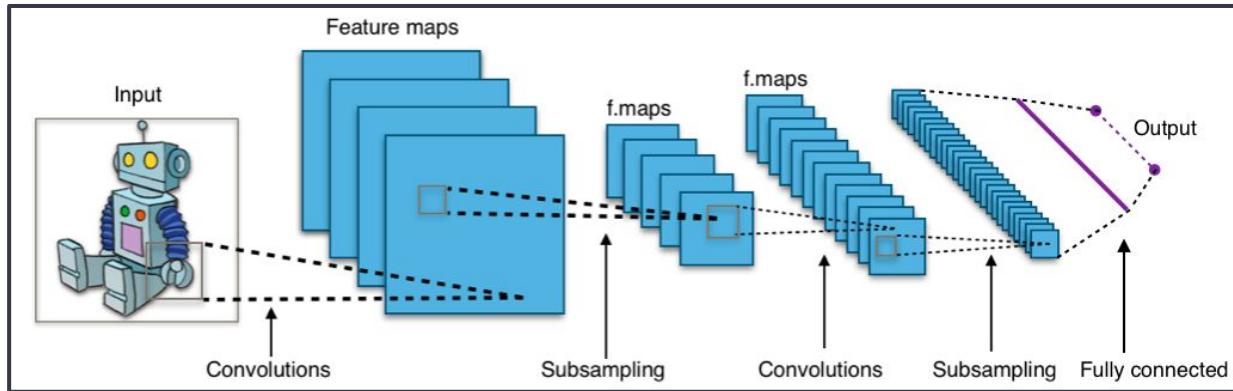
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	1	-1
-1	1	-1	-1	-1	-1	1	-1	-1
-1	-1	1	1	-1	-1	1	-1	-1
-1	-1	-1	-1	1	-1	1	-1	-1
-1	-1	-1	-1	-1	1	-1	-1	-1
-1	-1	-1	1	-1	1	1	-1	-1
-1	-1	-1	1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

Feature Extraction and Convolution: A Case Study



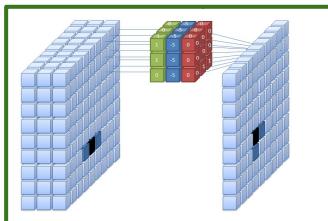
Typical layer involved in Convolutional Neural Networks

Structure of deep learning model

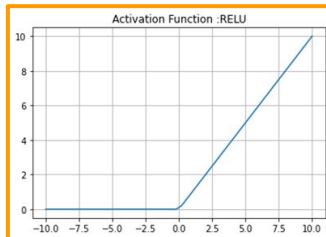


https://medium.com/@penanklihicyrille_16953/introduction-aux-reseaux-de-neurones-de-convolution-b5da87a5321f

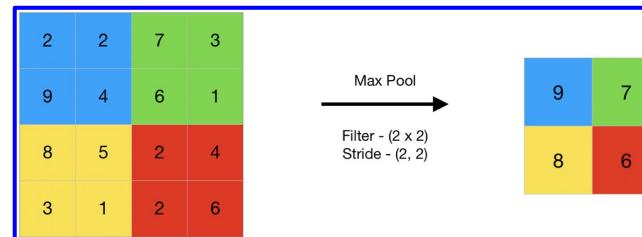
Layers in deep learning model



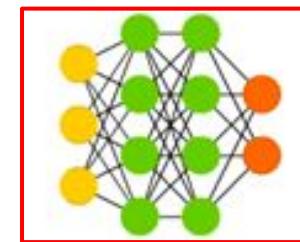
Convolution



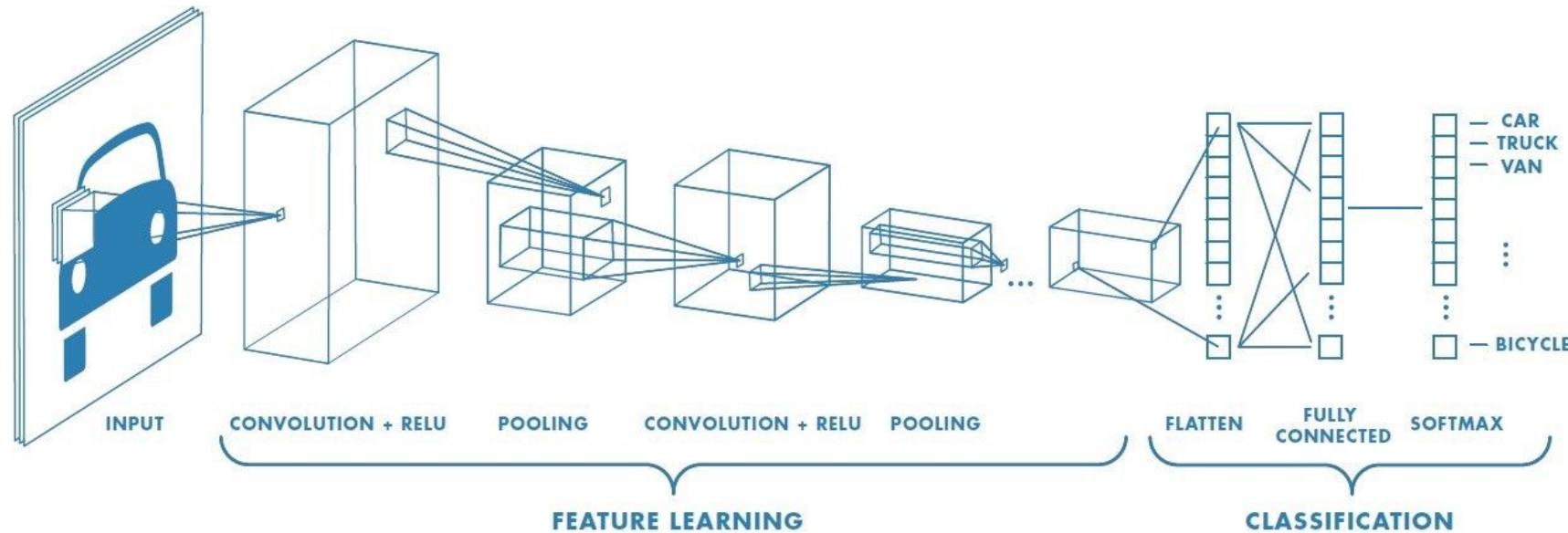
ReLU



Pooling

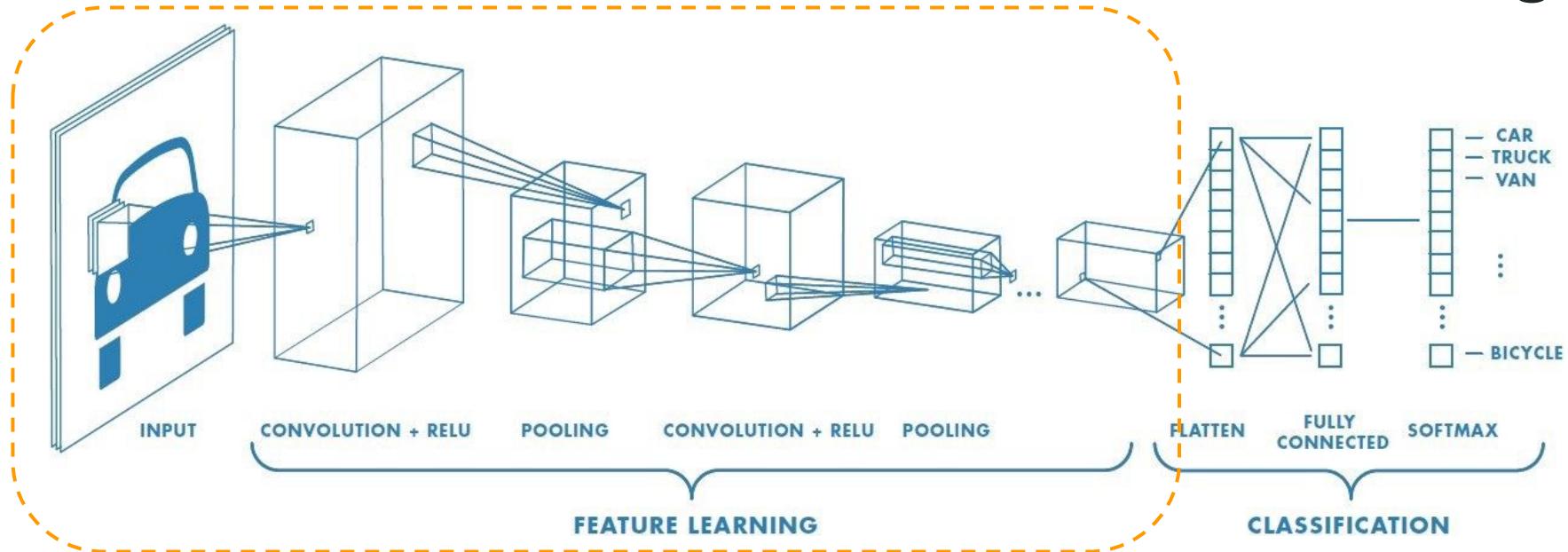


CNNs for Classification: Feature Learning



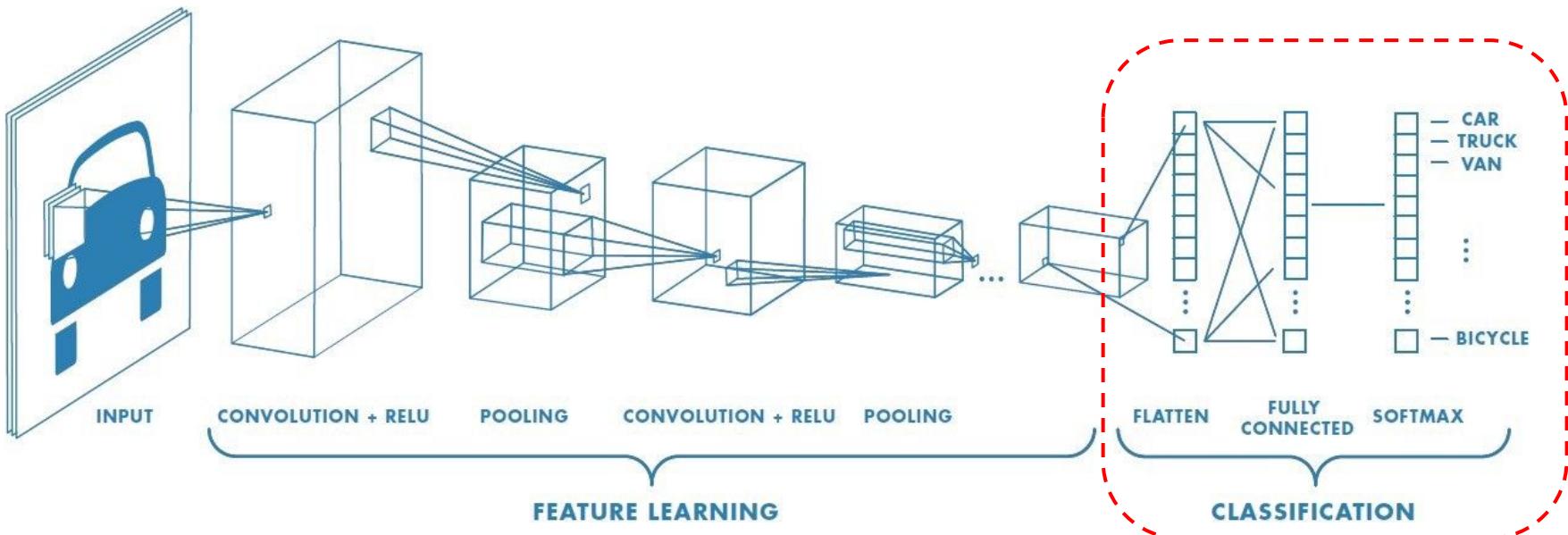
<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

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**

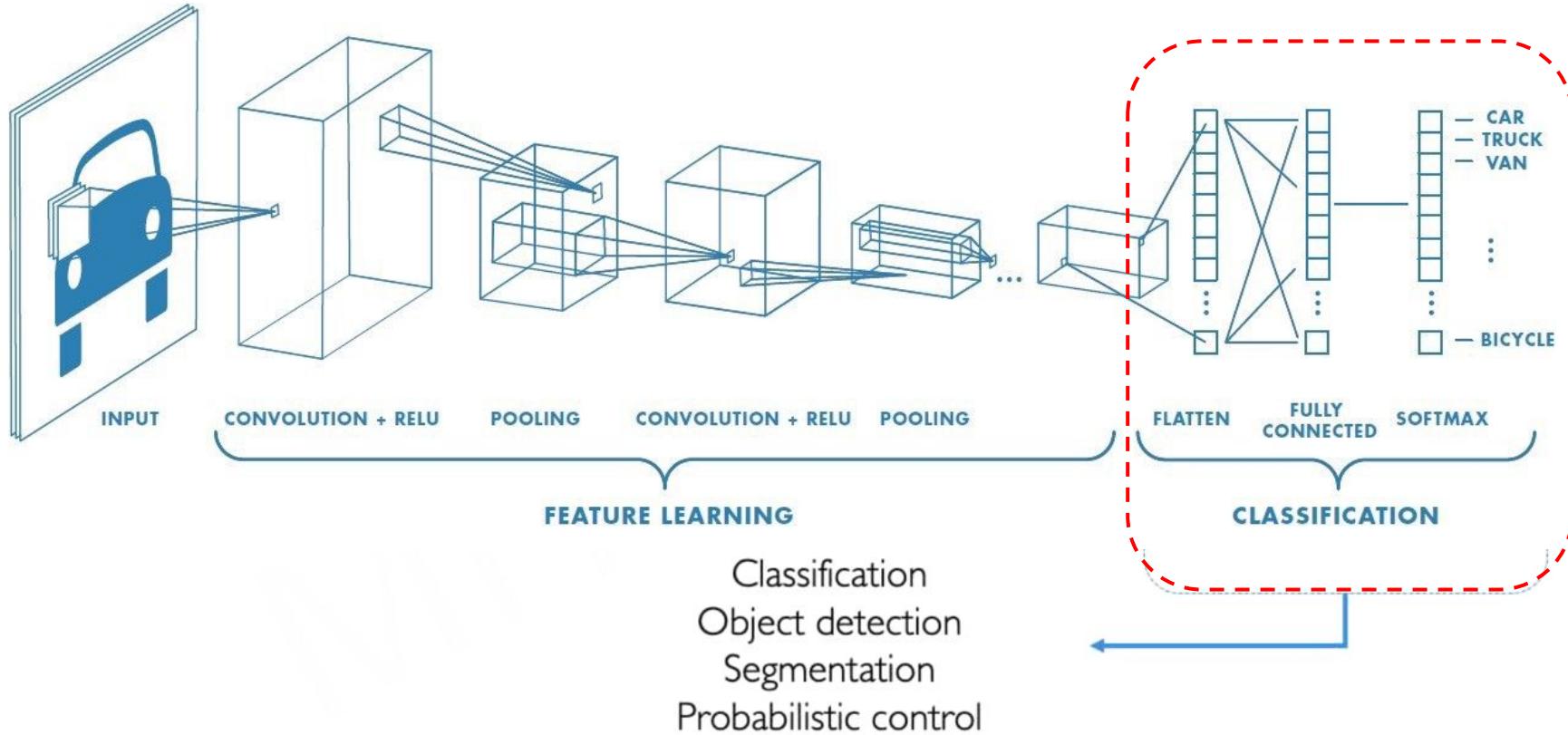
CNNs for Classification: Feature Learning



- 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

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

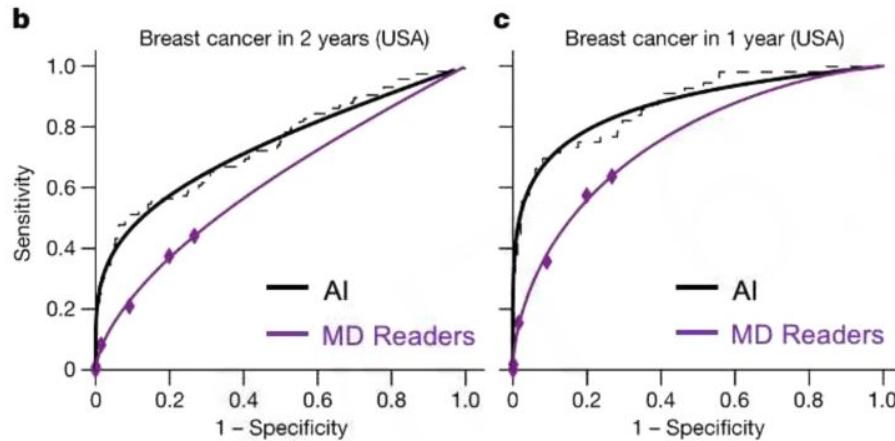
An Architecture for Many Applications



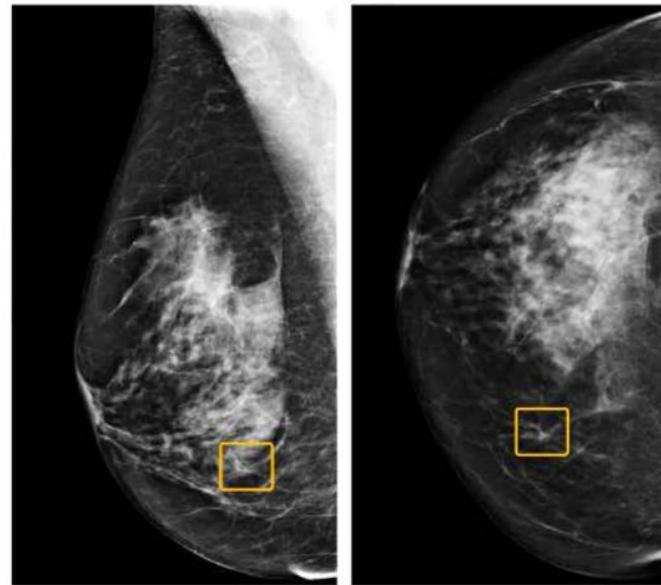
An Architecture for Many Applications

Classification: Breast Cancer Screening

International evaluation of an AI system for breast cancer screening



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

Object Detection



Image X

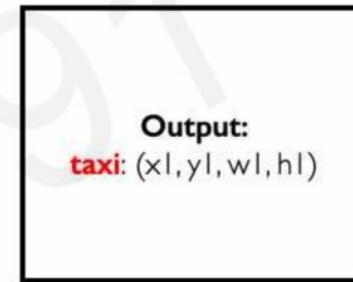
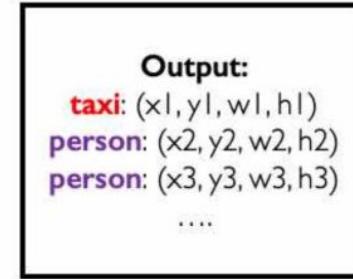
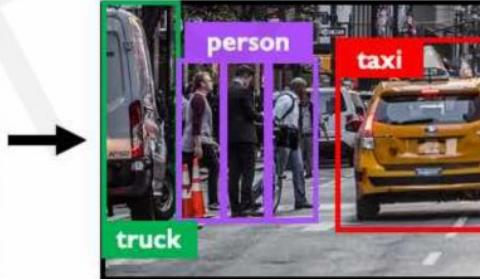


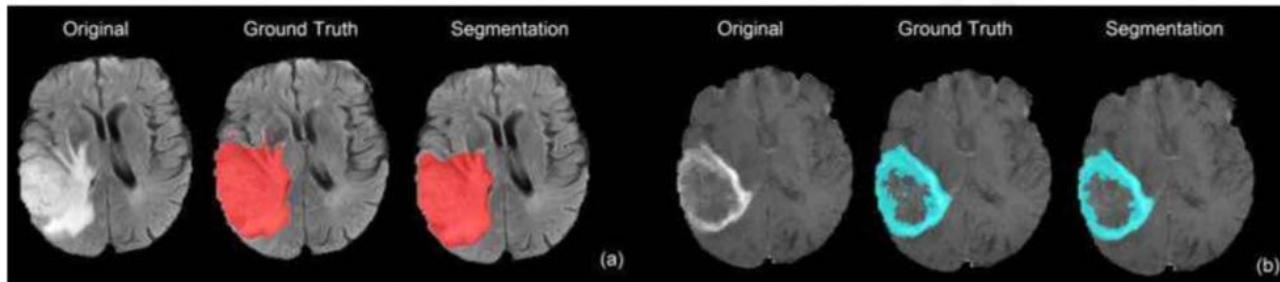
Image X



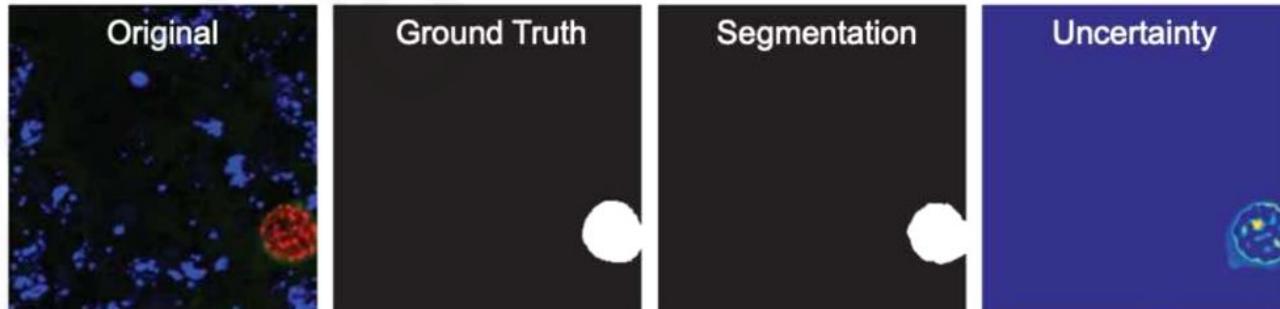
An Architecture for Many Applications

Semantic Segmentation: Biomedical Image Analysis

Brain Tumors
Dong+ *MIUA* 2017.



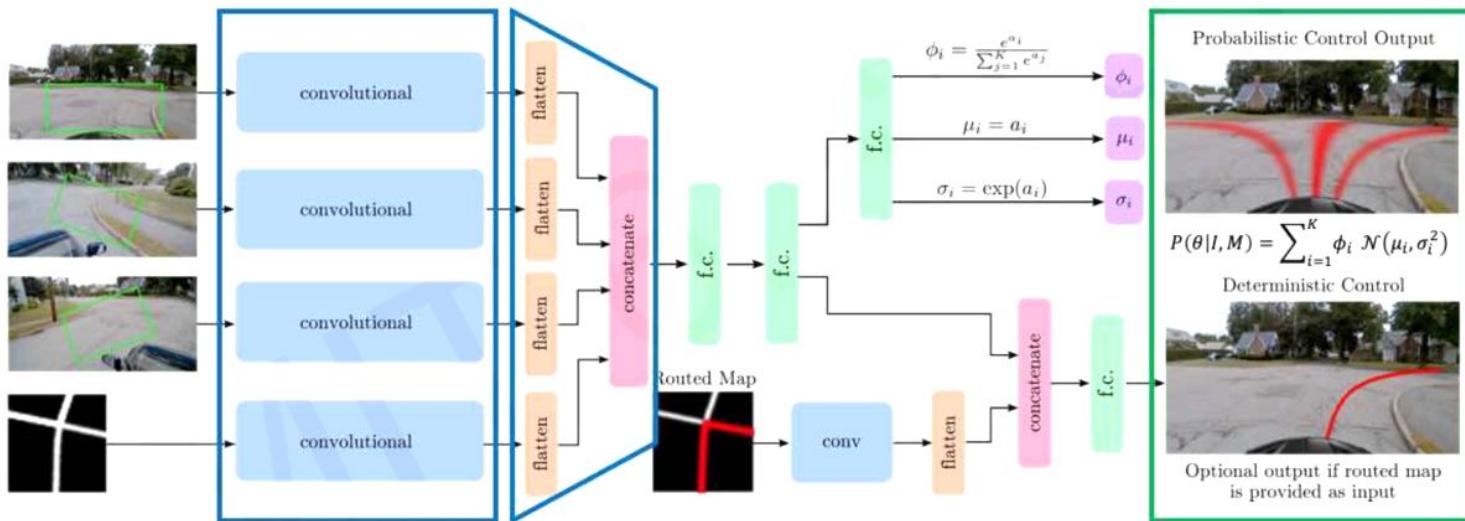
Malaria Infection
Soleimany+ *arXiv* 2019.



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**



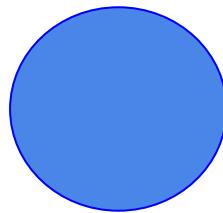
$$L = -\log(P(\theta|I, M))$$

Deep Sequence Modeling

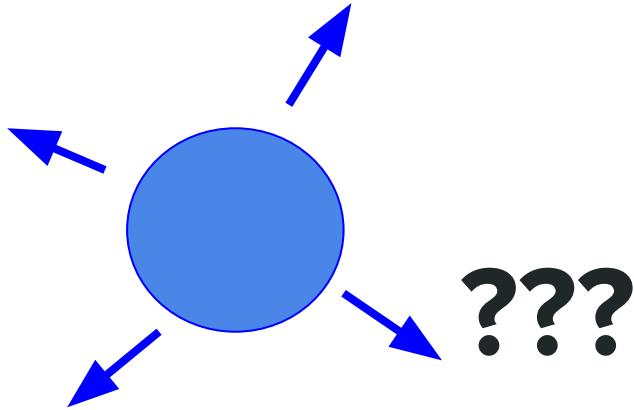
<http://introtodeeplearning.com/>



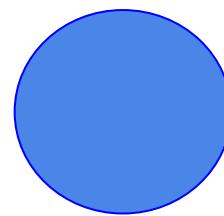
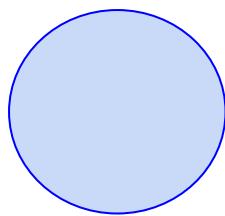
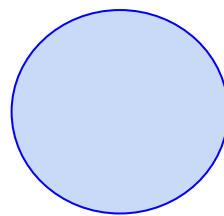
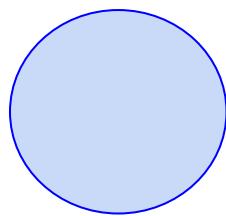
Given an images of a ball, can you predict where it will go next ?



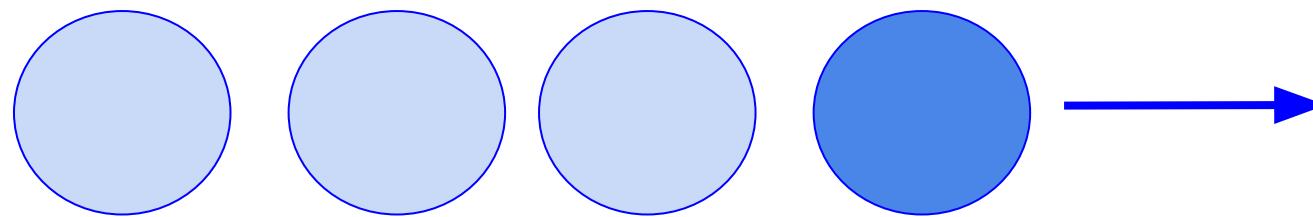
Given an images of a ball, can you predict where it will go next ?



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



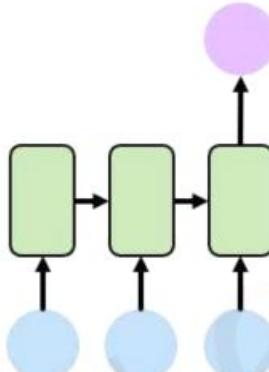
Sequence Modeling Architectures



One to One
Binary Classification

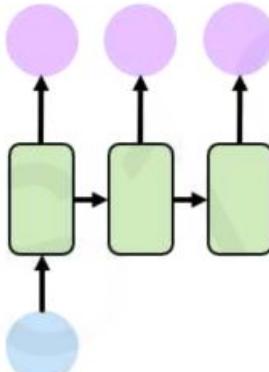


"Will I pass this class?"
Student → Pass?



Many to One
Sentiment Classification

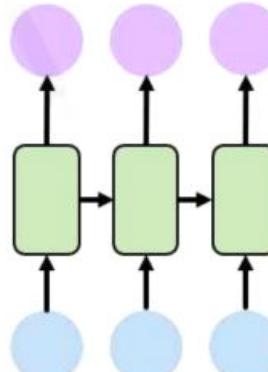
Ivar Hagendoorn
[@IvarHagendoorn](#)
Follow · ...
The @MIT Introduction to #DeepLearning is definitely one of the best courses of its kind currently available online [introtodeeplearning.com](#)



One to Many
Image Captioning



"A baseball player throws a ball."

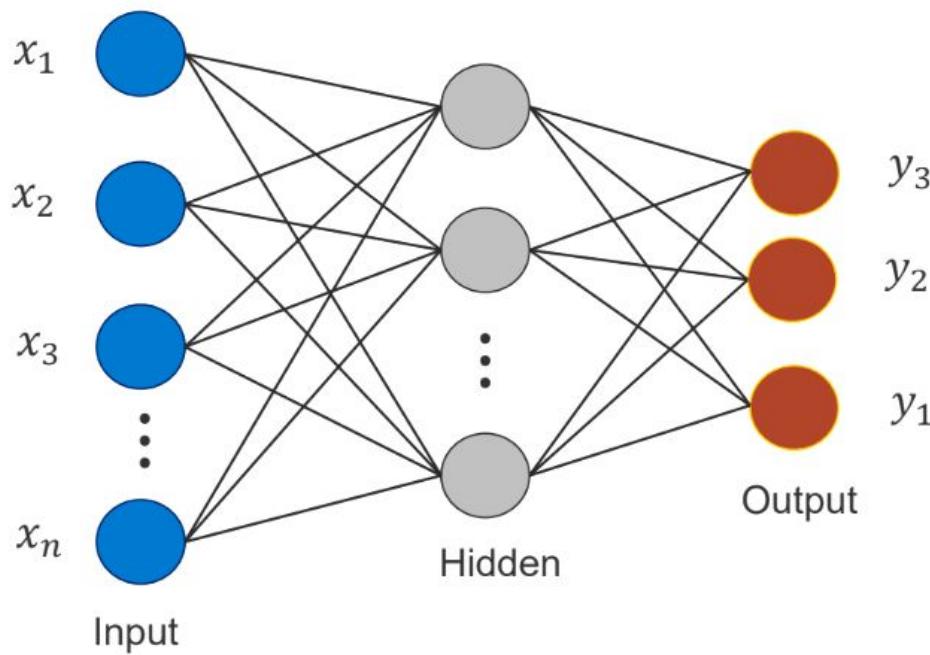


Many to Many
Machine Translation



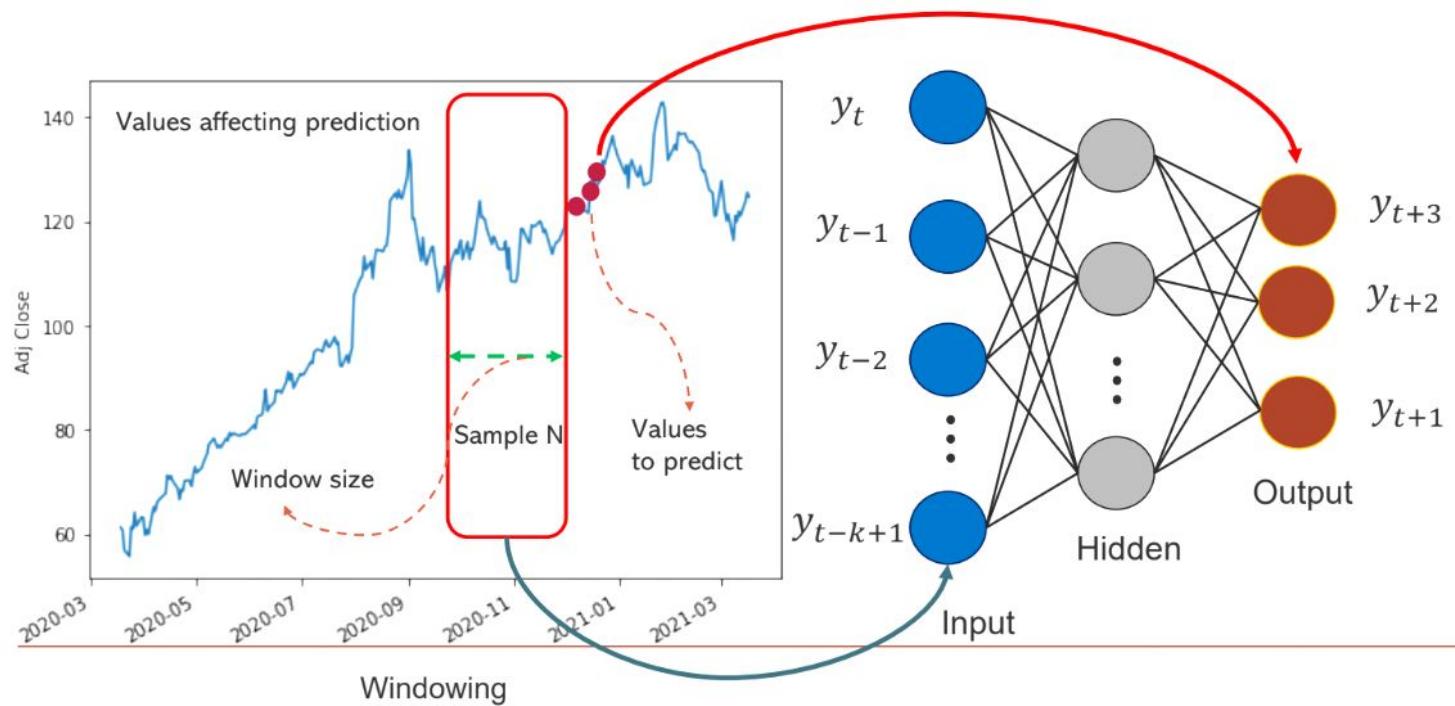
Neurons with Recurrence

Standard Neural Networks



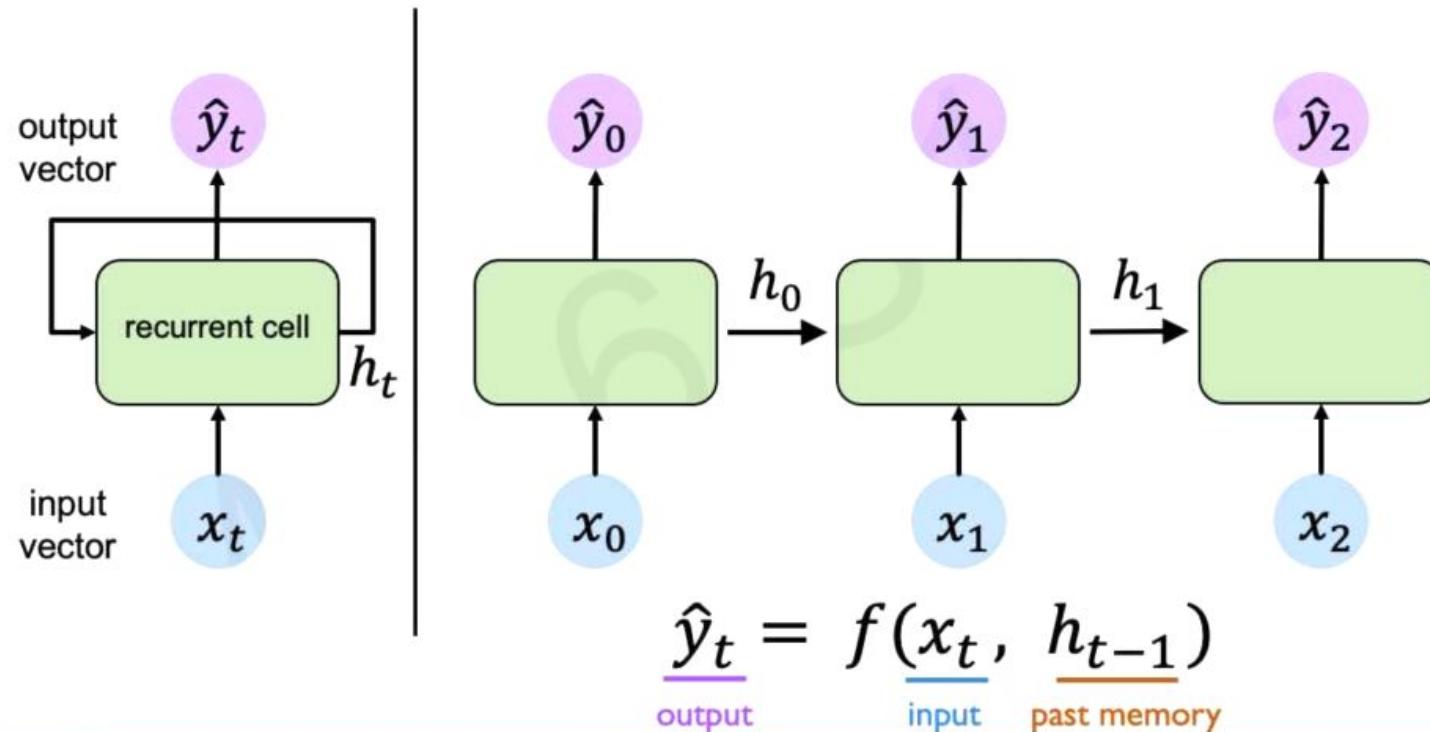
Neurons with Recurrence

Handling Individual Time Steps

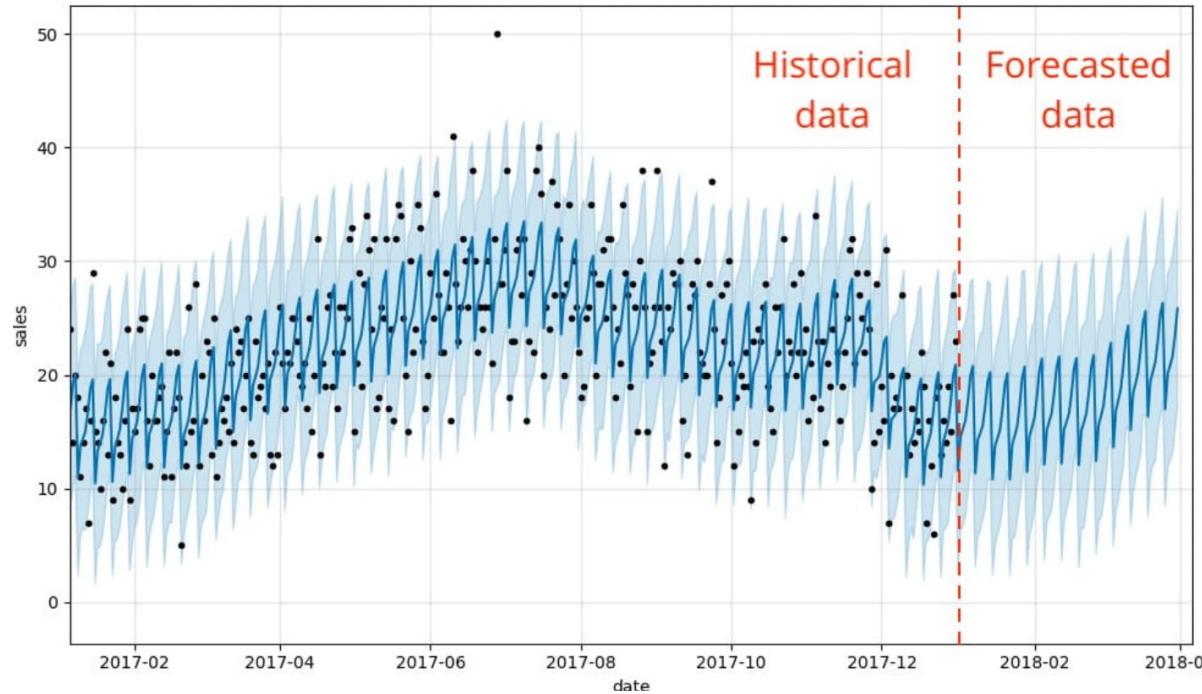


Neurons with Recurrence

Recurrent neural networks (RNNs)



A Sequence Modeling Problems: Time Series Forecasting



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 Problem: Predict the Next Word

“This morning I took my cat for a walk”

given these words

A Sequence Modeling Problems: Natural Language Processing

A Sequence Modeling Problem: Predict the Next Word

“This morning I took my cat for a walk.”

given these words

predict the
next word

A Sequence Modeling Problems: Natural Language Processing

A Sequence Modeling Problem: Predict the Next Word

“This morning I took my cat for a walk.”

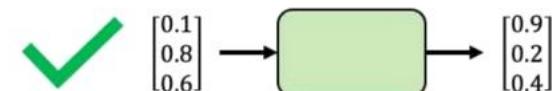
given these words

predict the
next word

Representing Language to a Neural Network



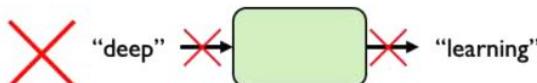
Neural networks cannot interpret words



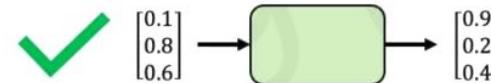
Neural networks require numerical inputs

A Sequence Modeling Problems: Natural Language Processing

Encoding Language for a Neural Network



Neural networks cannot interpret words



Neural networks require numerical inputs

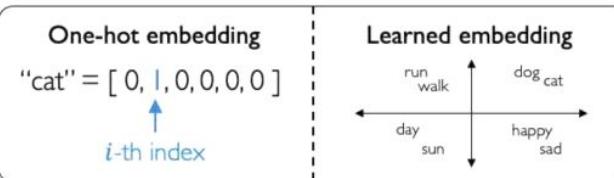
Embedding: transform indexes into a vector of fixed size.

this	cat	for
my	took	
I	walk	
a	morning	

1. Vocabulary:
Corpus of words

a	→ 1
cat	→ 2
...	...
walk	→ N

2. Indexing:
Word to index



3. Embedding:
Index to fixed-sized vector

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



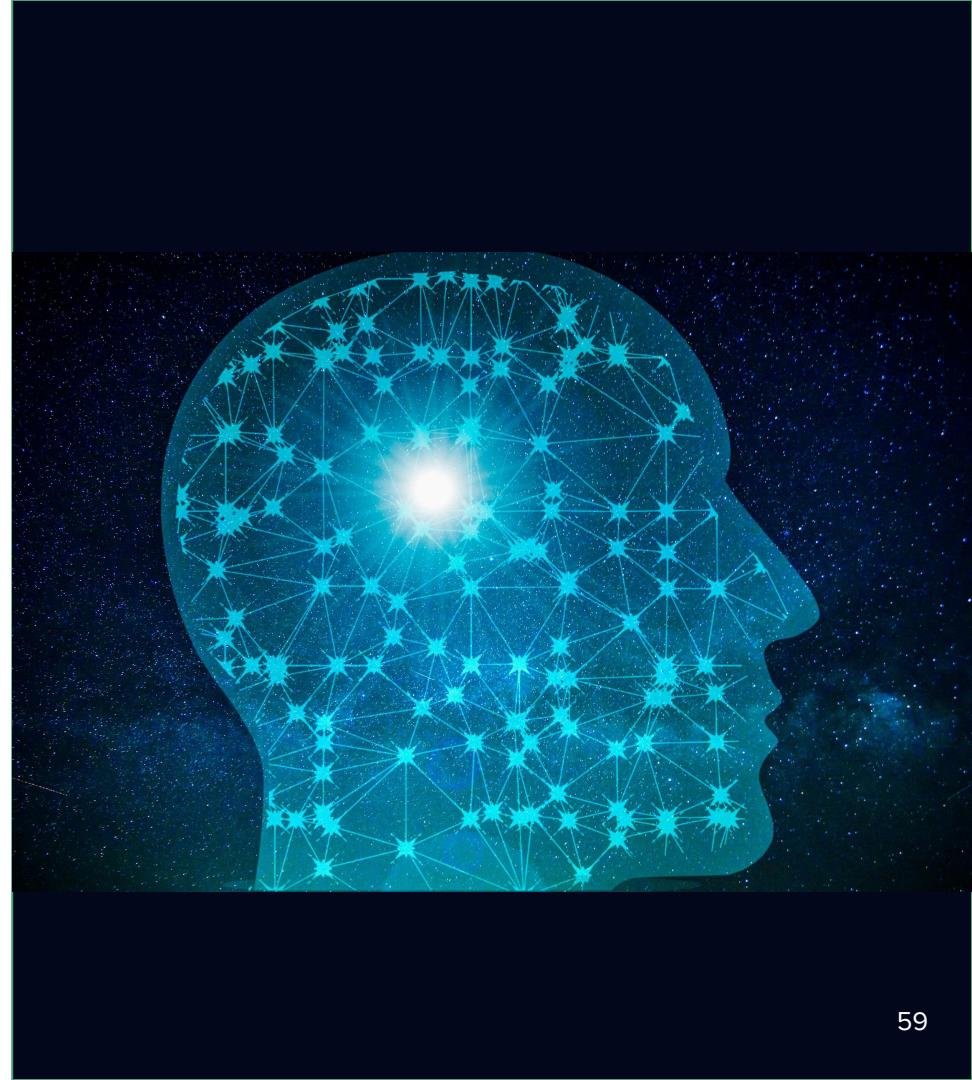
The food was good, not bad at all.

vs.

The food was bad, not good at all.

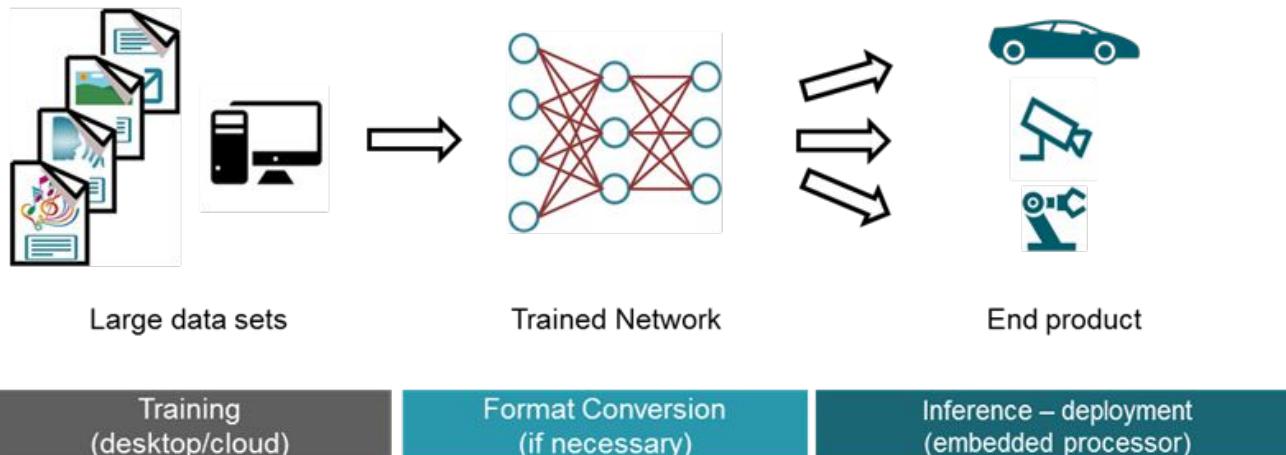


Deep Learning Development Flow



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

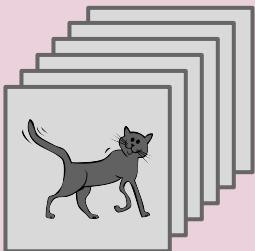


[https://e2e.ti.com/
blogs /b/industry
strength/posts/ti
-brings-deep-lear
ning-inference-to
the-edge](https://e2e.ti.com/blogs/b/industry-strength/posts/ti-brings-deep-learning-inference-to-the-edge)

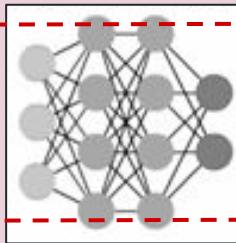
Deep Learning Development Flow

CNN training

Training



forward



Predict

“Dog”

= label ???

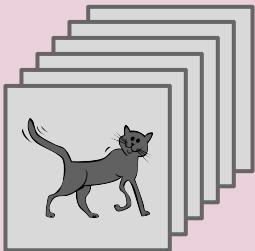
Label

“Cat”

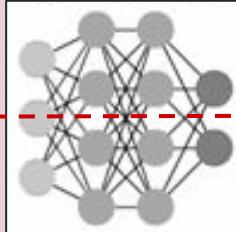
backward

CNN deployment

Interface



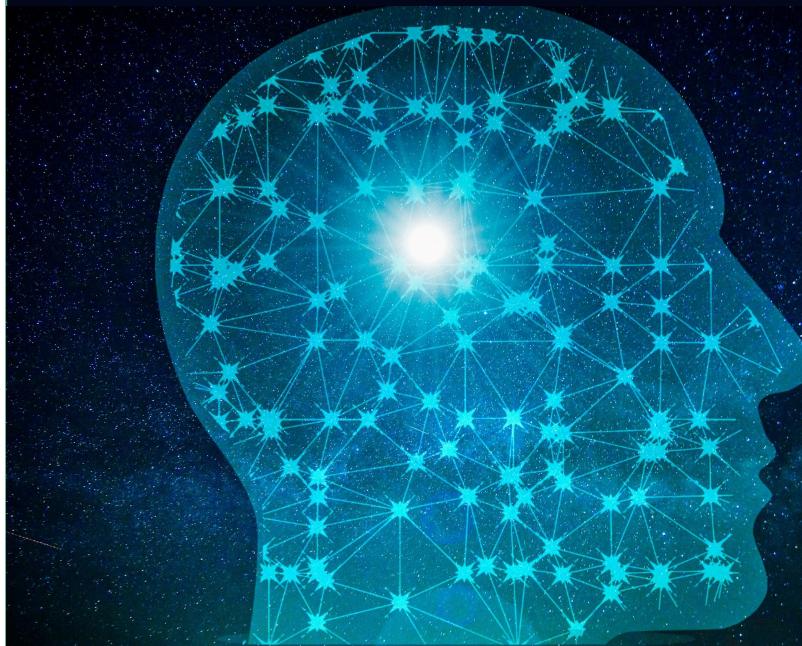
forward



Predict

“Cat”

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 😊

