Optimization for Deep Learning

Basic Principles and Pratical Sessions

Stefano Gualandi
Università di Pavia, Dipartimento di Matematica

email: stefano.gualandi@unipv.it
twitter: @famo2spaghi
blog: http://stegua.github.com
web: http://matematica.unipv.it/gualandi/

Outline

1 Introduction

2 Images

3 Machine Learning

4 Neural Networks

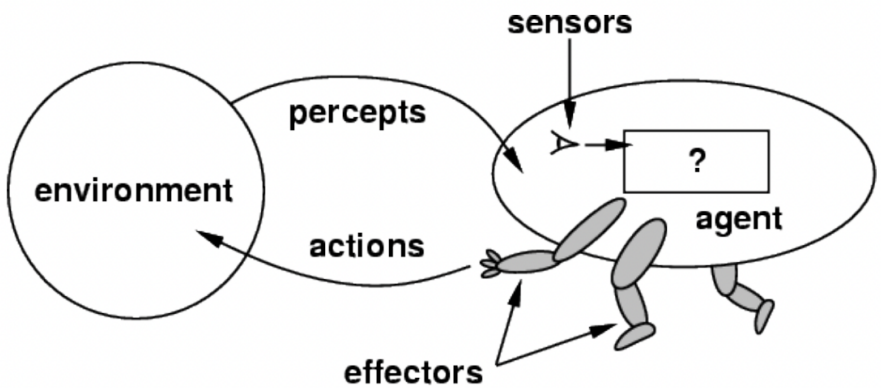
5 Convolutional Neural Networks

6 Transfer Learning

Artificial Intelligence and Machine Learning

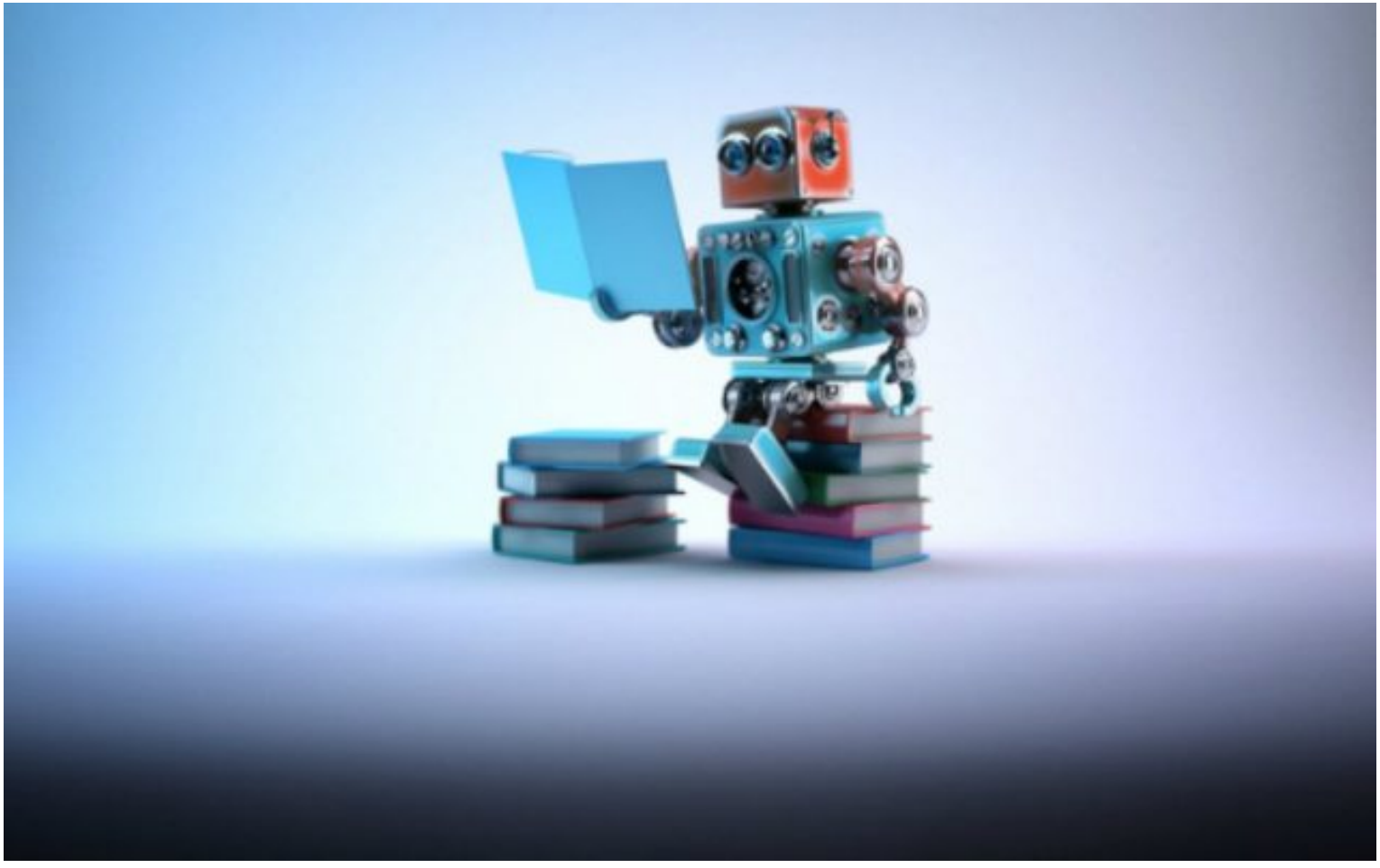
How do you design an intelligent agent?

- Definition: An **intelligent agent** perceives its environment via **sensors** and acts rationally upon that environment with its **effectors**.
- A discrete agent receives **percepts** one at a time, and maps this percept sequence to a sequence of discrete **actions**.
- Properties
  - –Autonomous
  - –Reactive to the environment
  - –Pro-active (goal-directed)
  - –Interacts with other agents via the environment



Artificial Intelligence: A Modern Approach. Stuart Russell and Peter Norvig
`https://aima.cs.berkeley.edu/`

Artificial Intelligence and Machine Learning



"Machine take me by surprise with great frequency"
– A.M. Turing

What does it mean "Learning"?

TRECCANI ISTITUTO MAGAZINE CATALOGO SCUOLA LIBRI

imparare

Vocabolario on line

Crea un ebook con questa voce | Scaricalo ora (0) Condividi

imparare v. tr. [lat. **imparare*, comp. di *in-*[1] e *parare* «procurare»; propr. «procacciarsi una nozione», o sim.]. – **1. a.** Acquistare cognizione di qualche cosa, o fare propria una serie di cognizioni (relative a un'arte, a una scienza, a un'attività, ecc.), per mezzo dello studio, dell'esercizio, dell'osservazione, della pratica, attraverso l'esempio altrui, ecc. Con sign. generico: *è un fannullone che non ha voglia d'imparar nulla*; *da lui, abbiamo tutti qualcosa da i.*; *sbagliando s'impara* (prov.). Determinando: *i. la grammatica, la letteratura, la fisica, il disegno*; *i. una lingua straniera, i. un gioco, i. le buone maniere. I. a memoria* o *a mente*, fissare nella memoria (un brano di prosa, dei versi, ma anche brevi frasi, serie di nomi, di numeri, ecc.) in modo da saper ripetere testualmente le parole; in questo senso, anche il semplice *imparare*: *ho già imparato la poesia* (talora con la particella pron.: *mi sono già imparato le prime due strofe*); *i. la lezione*, mandarla a memoria sia testualmente sia assimilando genericamente i concetti. **b.** Con *a* e l'infinito, apprendere a saper fare qualche cosa: *i. a leggere, a scrivere, a contare, a suonare, a cantare, a ballare, a dipingere, a cucinare*; con sign. più ampio (acquistare

Informal definition

"Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed"

Reinforcement Learning plays ATARI



DEMO GIF

Formal definition

Definition 1 (Mitchell, 1997)

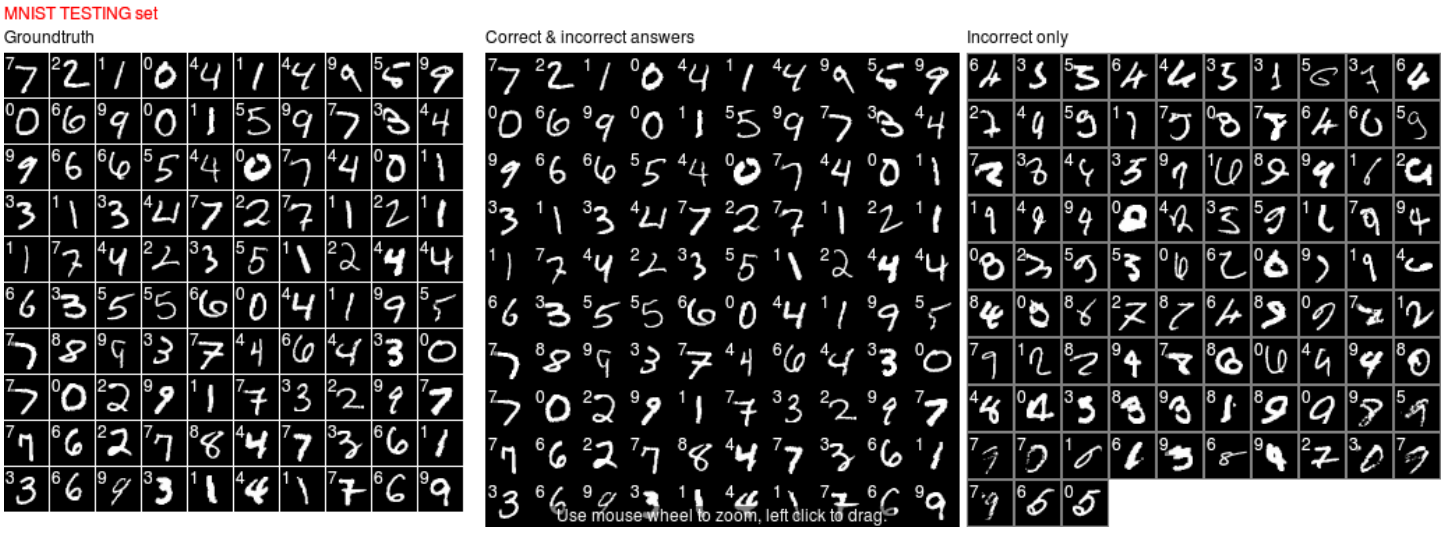
A computer program is said to **learn** from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.

Example 2

A handwriting recognition learning problem:

- **Task** T: recognizing and classifying handwritten words within images
- **Performance measure** P: percent of words correctly classified
- **Training experience** E: a dataset of handwritten words with given classifications

Handwriting recognition learning problem

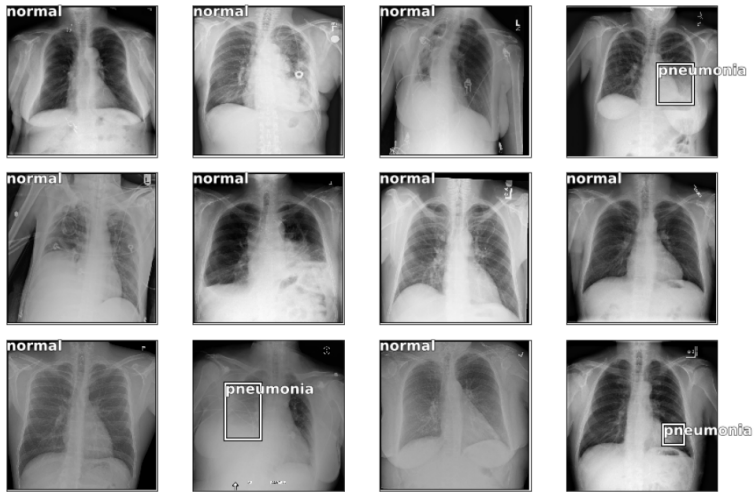


In 2017 and 2018, the RSNA Deep Learning Classroom, presented by NVIDIA Deep Learning Institute, used the **MNIST image dataset** to show radiologists how to train a classifier to recognize handwritten digits.

This year [2019], the Classroom will feature a medically originated set of images, **MedNIST Database** [..] (1) chest radiographs, (2) head CTs, (3) chest CTs, (4) abdomen CTs, (5) breast MRIs, and (6) brain MRIs [..] The goal is to build and train an AI model that accurately classifies images into those six categories.

Charles E. Kahn. We All Need A Little Maigc (Editorial). Radiology: Artificial Intelligence. Vol.1, n.4, 2019.

RSNA Pneumonia Detection Challenge on Kaggle

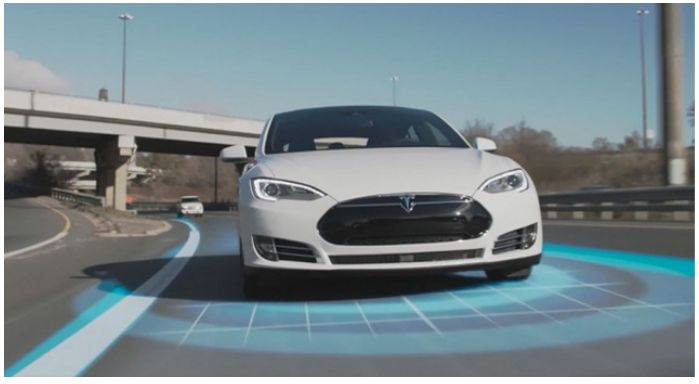


Example 3

RSNA Pneumonia Detection Challenge:

- **Task** T: to detect pneumonia in chest radiograph (CXR) images
- **Performance measure** P: the mean average precision at different intersection over union (IoU) thresholds
- **Training experience** E: the data set is composed of 25,684 unique chest radiographs (1024 x 1024 grayscale). 5,659 (22%) of these, present lung opacities, suggesting pneumonia in the patient. The x-rays tagged as pneumonia can contain multiple opacities, hence several bounding-boxes (ope per opacity) might be provided.

Self-driving cars





Example 4

A robot driving learning problem:

- **Task** T: driving on public streets using vision sensors
- **Performance measure** P: average distance traveled before an error (as judged by human operator)
- **Training experience** E: a sequence of images and steering commands recorded while observing a human driver

GDPR 2018: The right to Explanations

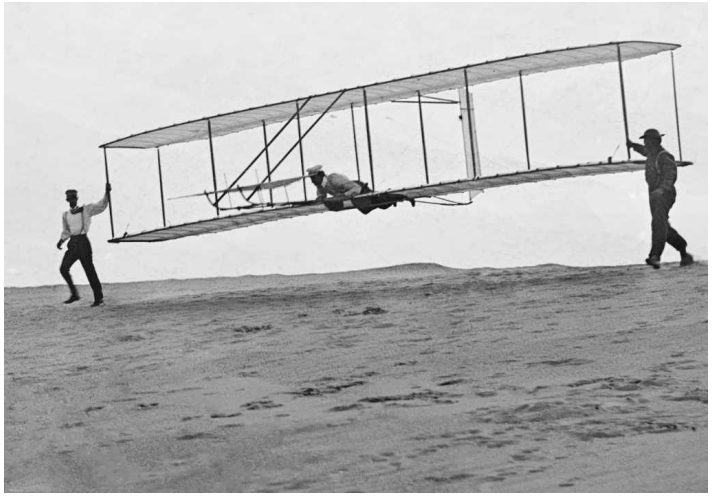
The **data subject** *should have the right not to be subject to a decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing*

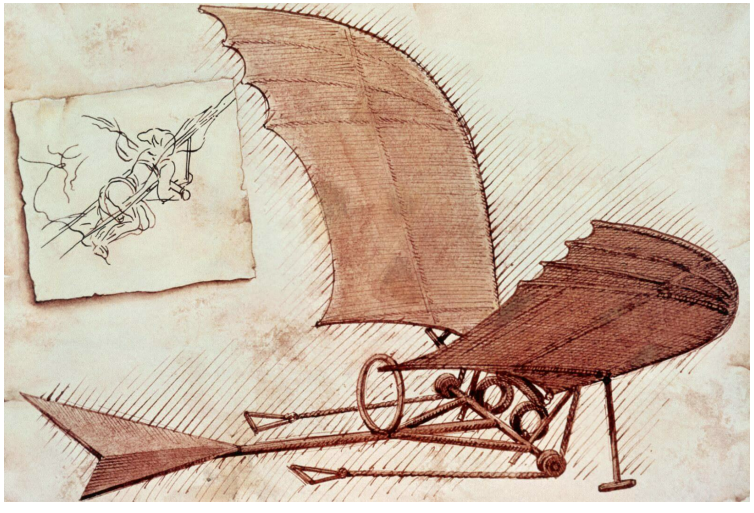
[...]

(the data subject has) the right to obtain human intervention, to express his or her point of view, **to obtain an explanation of the decision reached after such assessment and to challenge the decision**.

My view on AI: Natural Flight vs. Artificial Flight

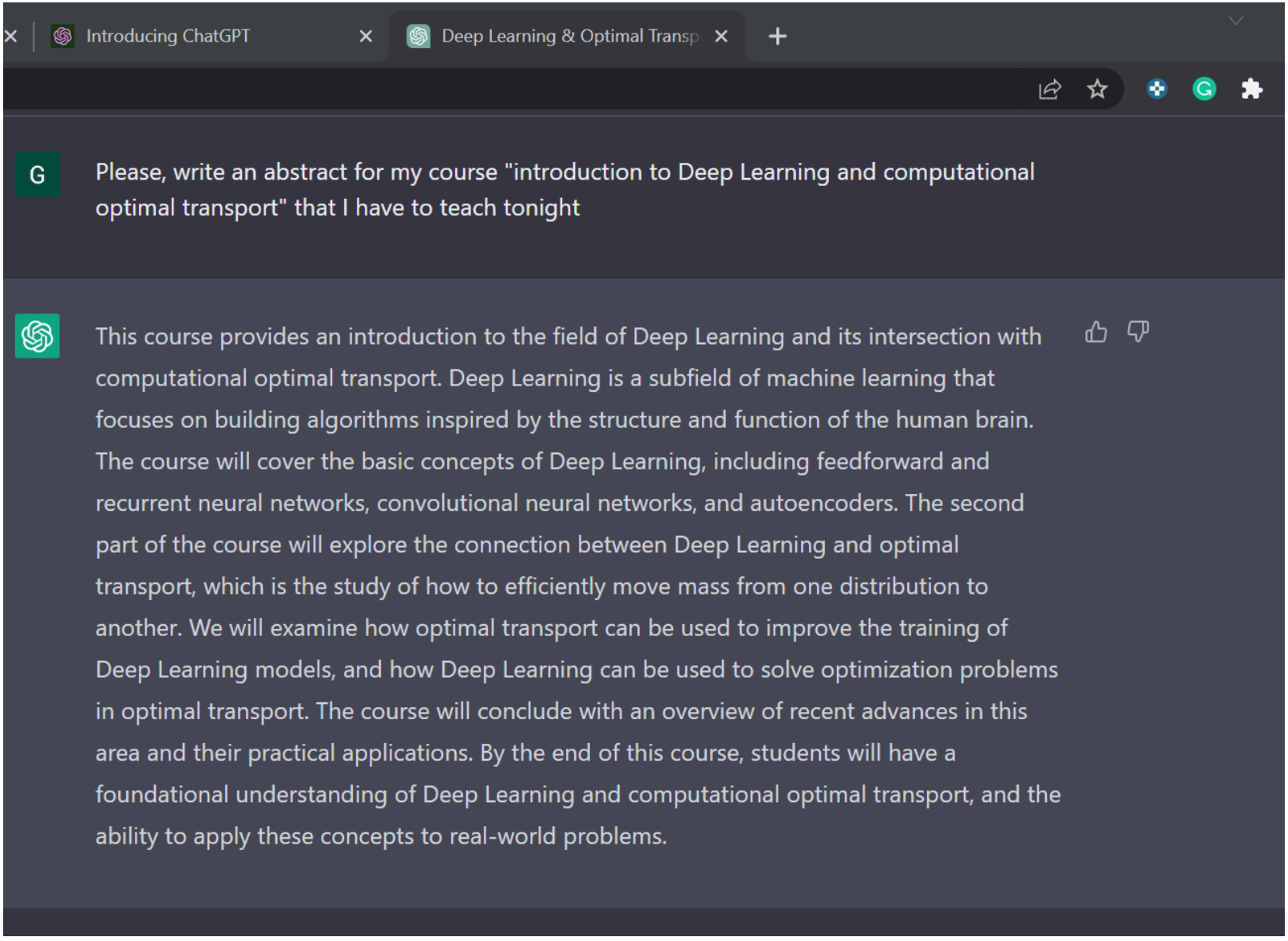




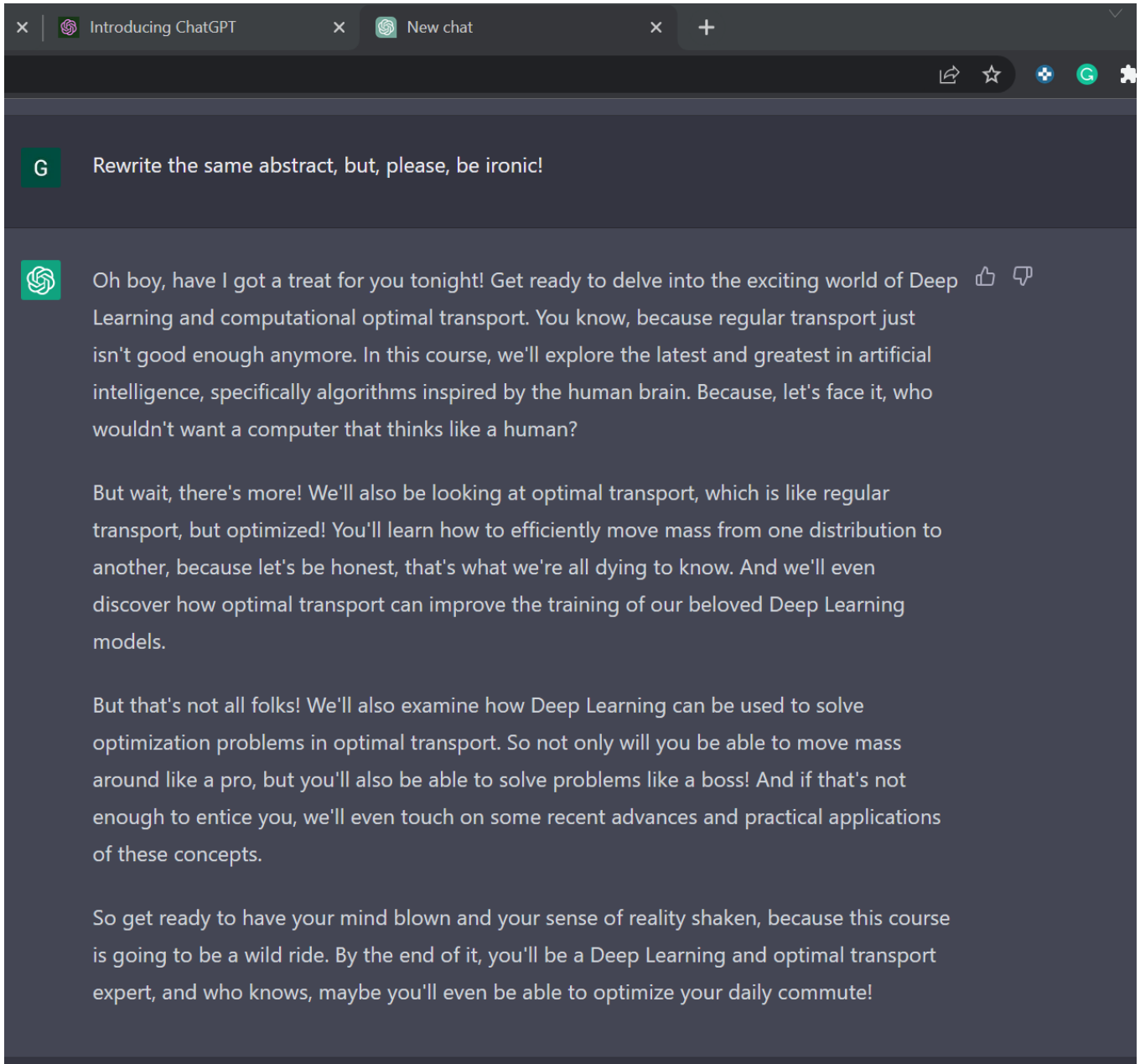




Marketing AI: ChatGPT



Marketing AI: ChatGPT



Marketing AI: ChatGPT

G Ma a cosa stai pensando?

Come modello di lingua artificiale, non penso, perché non ho emozioni o desideri. Sono qui solo per rispondere alle tue domande e offrire assistenza nella misura delle mie capacità. Se hai bisogno di aiuto o hai altre domande, sarò felice di assisterti.

G Ti piace leggere?

Come modello di linguaggio artificiale, non ho una vera e propria preferenza per la lettura, in quanto non ho interessi personali o opinioni. Tuttavia, come parte del mio addestramento, ho "letto" molti testi e libri durante il processo di apprendimento della lingua. Inoltre, sono in grado di fornire informazioni su vari argomenti e di rispondere alle domande in modo efficace.

G Hai un anima?

No, come modello di linguaggio artificiale, non ho un'anima o una coscienza in quanto sono solo un programma software progettato per elaborare e generare risposte in base alle informazioni che mi vengono fornite. Sono un insieme di algoritmi e modelli matematici che cercano di riprodurre l'uso del linguaggio umano in modo efficace e naturale.

Marketing AI: ChatGPT: Terence Tao's perspective

mastodon

Back



Terence Tao
@tao@mathstodon.xyz

AI tools like #ChatGPT will soon be capable of answering a large fraction of traditional university homework type questions with reasonable accuracy. In the long term, it seems futile to fight against this; perhaps what we as lecturers need to do is to move to an "open books, open AI" mode of examination where we give the students full access to AI tools but ask them more challenging questions, both to teach the material and also to teach the students how best to use the AI tools of the future.

20 Dec 2022, 01:21 · · Web · 185 · 282

What does it mean "reading"?

leggere

[lèg-ge-re] v.tr.

PRONUNCIA AUDIO | STAMPA IL RISULTATO

m pres. *io lèggo*, *tu lèggi ecc.*; pass.rem. *io lèssi*, *tu leggésti*, *egli lèsse*, *noi leggémmo*, *voi leggéste*, *essi lèssero*; part.pass. *lètto*

a aus. *avere*

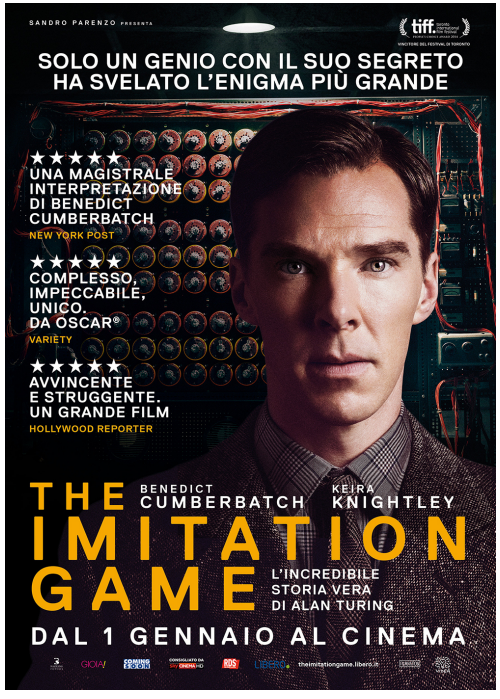
1. seguire con gli occhi i caratteri di una scrittura, intendendo il significato di parole e frasi: *leggere in silenzio, sottovoce, ad alta voce*; *saper leggere e scrivere*; *imparare, insegnare a leggere*; *i ciechi leggono con i polpastrelli*, riconoscendo appositi segni alfabetici in rilievo

2. venire a conoscenza del contenuto di un testo, di un avvenimento ecc. per mezzo della lettura [+ che, di]: *leggere una lettera, un libro, una notizia*; *ho letto sul giornale* ***che*** *domani ci sarà uno sciopero*; *ho letto* ***della*** *riapertura del teatro*; *leggere una lingua straniera*, capire ciò che in quella lingua è scritto; *leggere un autore, leggere Dante*, i suoi scritti

3. amare, praticare la lettura: *leggere tanto, di tutto*; *si dice che gli italiani leggano poco*; *nel tempo libero mi piace leggere*

Dizionario Garzanti online

"Can Machine Think?"
– *A.M. Turing*



VOL. LIX. NO. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should

VIDEO CLIP: The imitation game

VIDEO CLIP: L'uomo bicentenario

Learning Machines, by A.M. Turing

In the process of trying to imitate an adult human mind we are bound to think a good deal about the process which has brought it to the state that it is in. We may notice three components,

(*a*) The initial state of the mind, say at birth,

(*b*) The education to which it has been subjected,

(*c*) Other experience, not to be described as education, to which it has been subjected.

456 A. M. TURING:

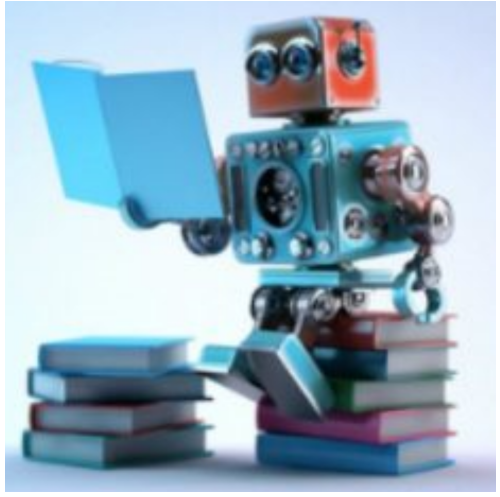
Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child-brain is something like a note-book as one buys it from the stationers. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child-brain that something like it can be easily programmed. The amount of work in the education we can assume, as a first approximation, to be much the same as for the human child.

Learning Machines, by A.M. Turing

the different genetical combinations that had been tried, so as to avoid trying them again?

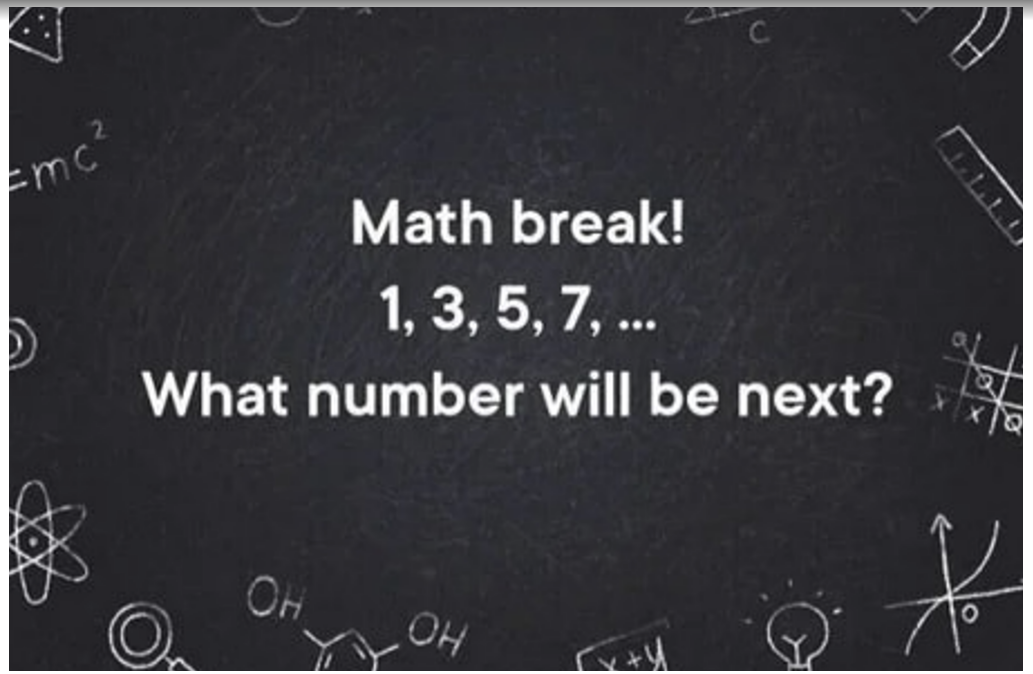
We may hope that machines will eventually compete with men in all purely intellectual fields. But which are the best ones to start with? Even this is a difficult decision. Many people think that a very abstract activity, like the playing of chess, would be best. It can also be maintained that it is best to provide the machine with the best sense organs that money can buy, and then teach it to understand and speak English. This process could follow the normal teaching of a child. Things would be pointed out and named, etc. Again I do not know what the right answer is, but I think both approaches should be tried.

We can only see a short distance ahead, but we can see plenty there that needs to be done.



Will a computer *love* reading?

Math and Learning



The next number is **217341** if

$$f(x) = \frac{18111}{2}x^4 - 90555x^3 + \frac{633885}{2}x^2 - 452773x + 217331$$

since then

$$f(1) = 1, \quad f(2) = 3, \quad f(3) = 5, \quad f(4) = 7, \quad f(5) = 217341$$

REMARK: Learning is finding a (composite) function that "*explains*" the data.



Select PayPal as payment method, and apply PROMOCODE:

DLITEACH1222_17_MSYA_44

https://courses.nvidia.com/courses/course-v1:DLI+S-FX-01+V1/

Outline

1 Introduction

2 Images

3 Machine Learning

4 Neural Networks

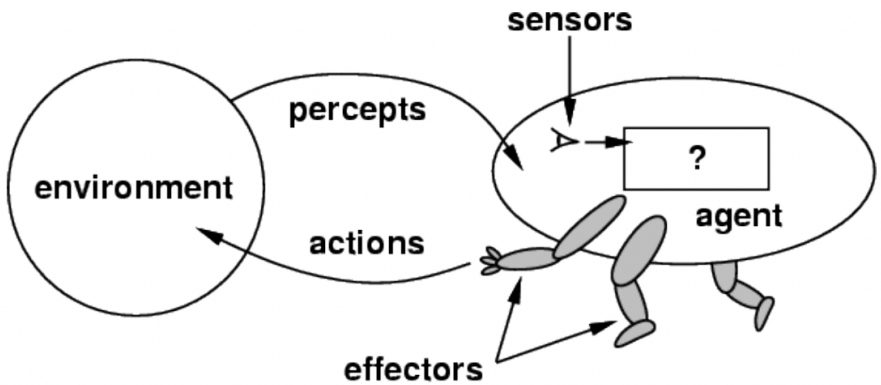
5 Convolutional Neural Networks

6 Transfer Learning

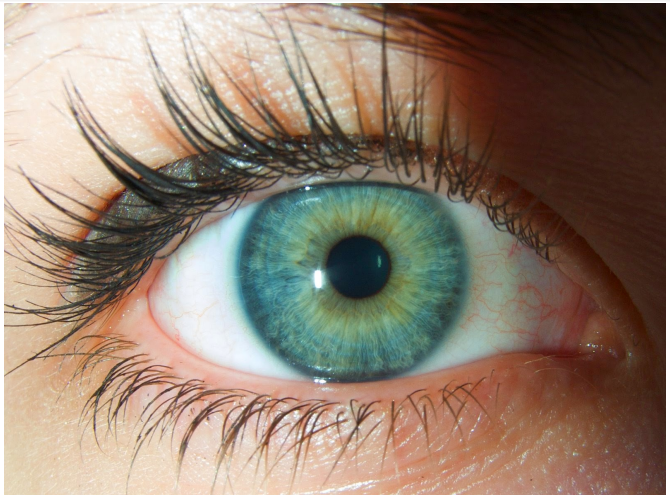
Intelligent Agent

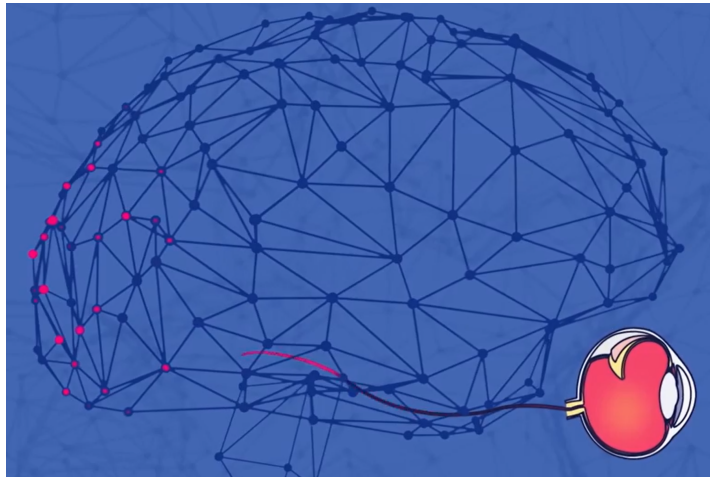
How do you design an intelligent agent?

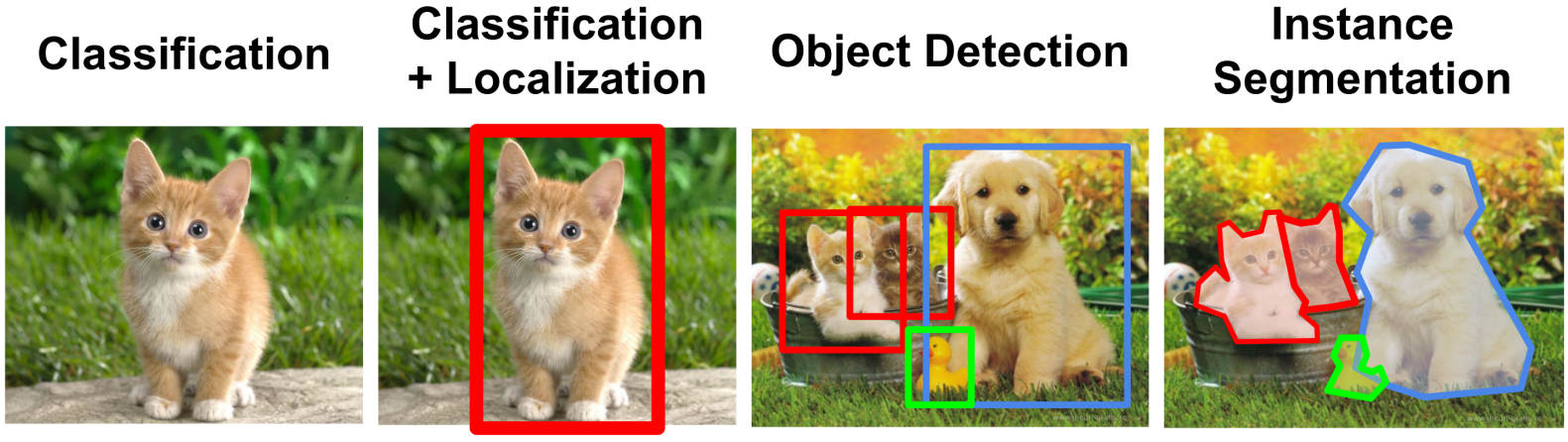
- Definition: An **intelligent agent** perceives its environment via **sensors** and acts rationally upon that environment with its **effectors**.
- A discrete agent receives **percepts** one at a time, and maps this percept sequence to a sequence of discrete **actions**.
- Properties
  - –Autonomous
  - –Reactive to the environment
  - –Pro-active (goal-directed)
  - –Interacts with other agents via the environment



Human Vision vs. Artificial Vision: Where are we?







Math and Images

Disclaimer: "Images are just numbers!"



64x64 pixels
MedNIST Challenge



16x16 pixels
Downsampling

17	14	13	12	15	14	14	87	82	16	14	14	13	14	16	21
15	12	11	10	104	77	10	123	116	23	101	13	10	12	14	19
13	10	10	11	115	108	13	146	146	63	117	29	8	10	11	17
12	9	90	64	106	140	11	159	145	114	148	21	6	7	10	16
11	7	99	123	85	152	98	184	184	165	161	17	8	8	13	16
11	6	44	135	110	166	152	174	177	168	171	5	8	93	114	15
11	6	5	139	162	153	185	195	189	196	163	5	102	171	139	16
10	5	6	46	161	168	208	218	214	227	177	6	165	173	54	15
9	4	25	22	170	188	212	210	210	230	170	158	207	156	6	11
8	3	241	147	119	217	224	225	225	244	182	215	214	86	6	11
5	1	91	249	228	211	247	253	253	254	234	216	158	5	5	11
5	1	16	186	81	195	255	255	255	255	252	152	5	4	4	11
6	2	1	1	3	234	255	255	255	253	138	4	0	1	3	10
6	2	2	2	5	255	255	255	255	218	27	5	1	2	3	10
6	3	2	2	95	255	251	255	255	115	2	2	1	1	3	10
8	4	4	3	171	255	255	255	255	5	3	3	3	3	4	10

Image as matrix of numbers

Math, Color Images, and Tensors



RGB Image



Red channel



Green channel



Blu channel

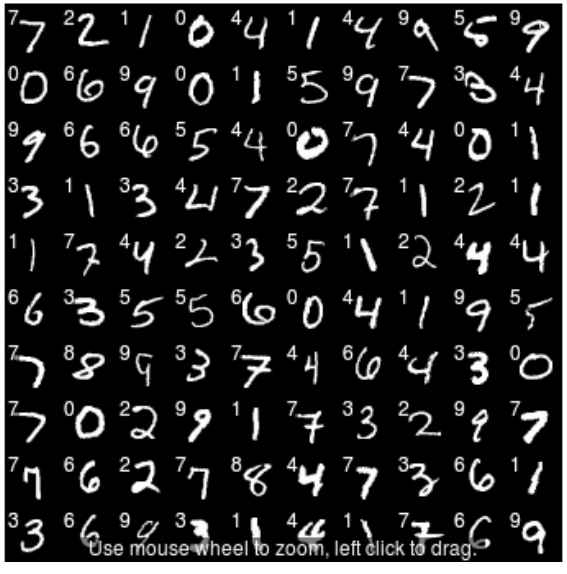
Most used dataset in the past: Digits

MNIST TESTING set

Groundtruth



Correct & incorrect answers



Incorrect only



Outline

1 Introduction

2 Images

3 Machine Learning

4 Neural Networks

5 Convolutional Neural Networks

6 Transfer Learning

Machine Learning Categories

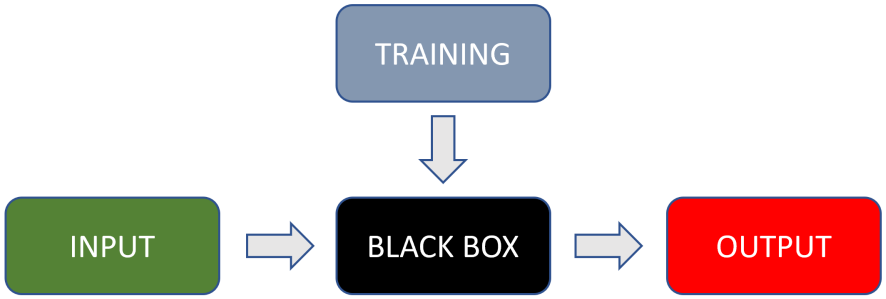
Machine Learning Tasks are typically classified into:

1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning

Definition 5 (1. Supervised Learning)

Supervised Learning is the task of learning (inferring) a function f that maps input vectors to their corresponding target vectors, by using a dataset containing a given set of pairs of $(input, output)$ samples.
Examples:

- Regression: the output vectors contain continuous values
- Classification: the output vectors contain discrete categories.
  Special case: **Binary Classification**.



- Learns known patterns
- Labeled input data
- Predicts outcome

1. Supervised Learning: Binary Classification



$\Rightarrow$ { 0=NORMAL ; 1=PNEUMONIA }

$\boldsymbol{x} \in X$: Input data

$y \in Y = \{0, 1\}$: Target

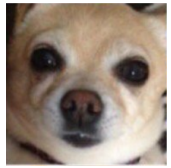
CLASSIFICATION TASK: Find a function

$$f_w : X \to Y$$

that maps each input to the *correct* target.

Dog or muffins?

Take $Y = \{-1, 1\}$, with $y = 1$ means **DOG**, and $y = -1$ means **MUFFINS**.

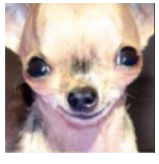
$\boldsymbol{x}_1 =$  $\boldsymbol{x}_2 =$  $\boldsymbol{x}_3 =$  $\boldsymbol{x}_4 =$  $\boldsymbol{x}_5 =$ 

$y_1 = 1$ $y_2 = 1$ $y_3 = -1$ $y_4 = 1$ $y_5 = -1$

1. Use the Training Set $\Rightarrow$ run learning algorithm $\Rightarrow$ get

$$f_w : X \to Y$$

2. Use a Test Set to validate f:

$f_w\left(\boldsymbol{x}_6 =\right.$  $\left.\right) = -1$ $f_w\left(\boldsymbol{x}_7 =\right.$  $\left.\right) = 1$

Is it a **4** or not?

We take $Y = \{-1, 1\}$, $y = 1$ means **YES, IS 4**, $y = -1$ **NO, IT'S NOT**.

$\boldsymbol{x}_1 =$  $\boldsymbol{x}_2 =$  $\boldsymbol{x}_3 =$  $\boldsymbol{x}_4 =$  $\boldsymbol{x}_5 =$ 

$y_1 = 1$ $y_2 = -1$ $y_3 = 1$ $y_4 = 1$ $y_5 = -1$

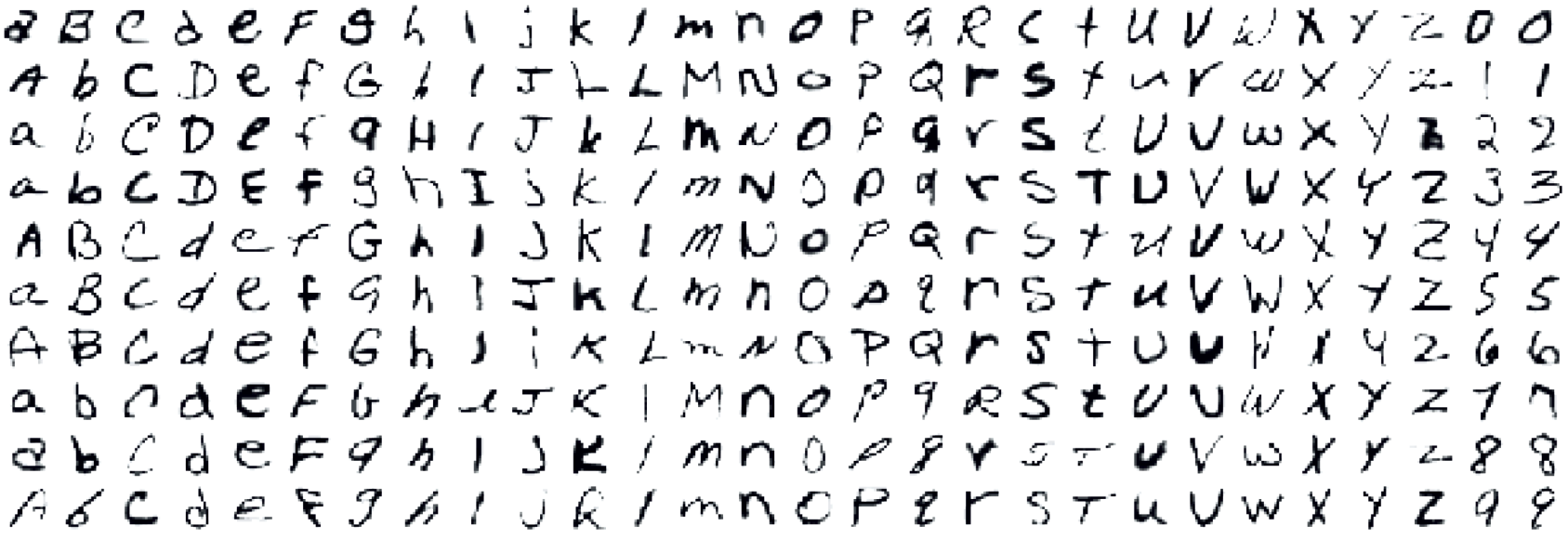
1. Use the TRAINING SET $\Rightarrow$ run learning algorithm $\Rightarrow$ get

$$f_w : X \to Y$$

2. Use a TEST SET to validate f:

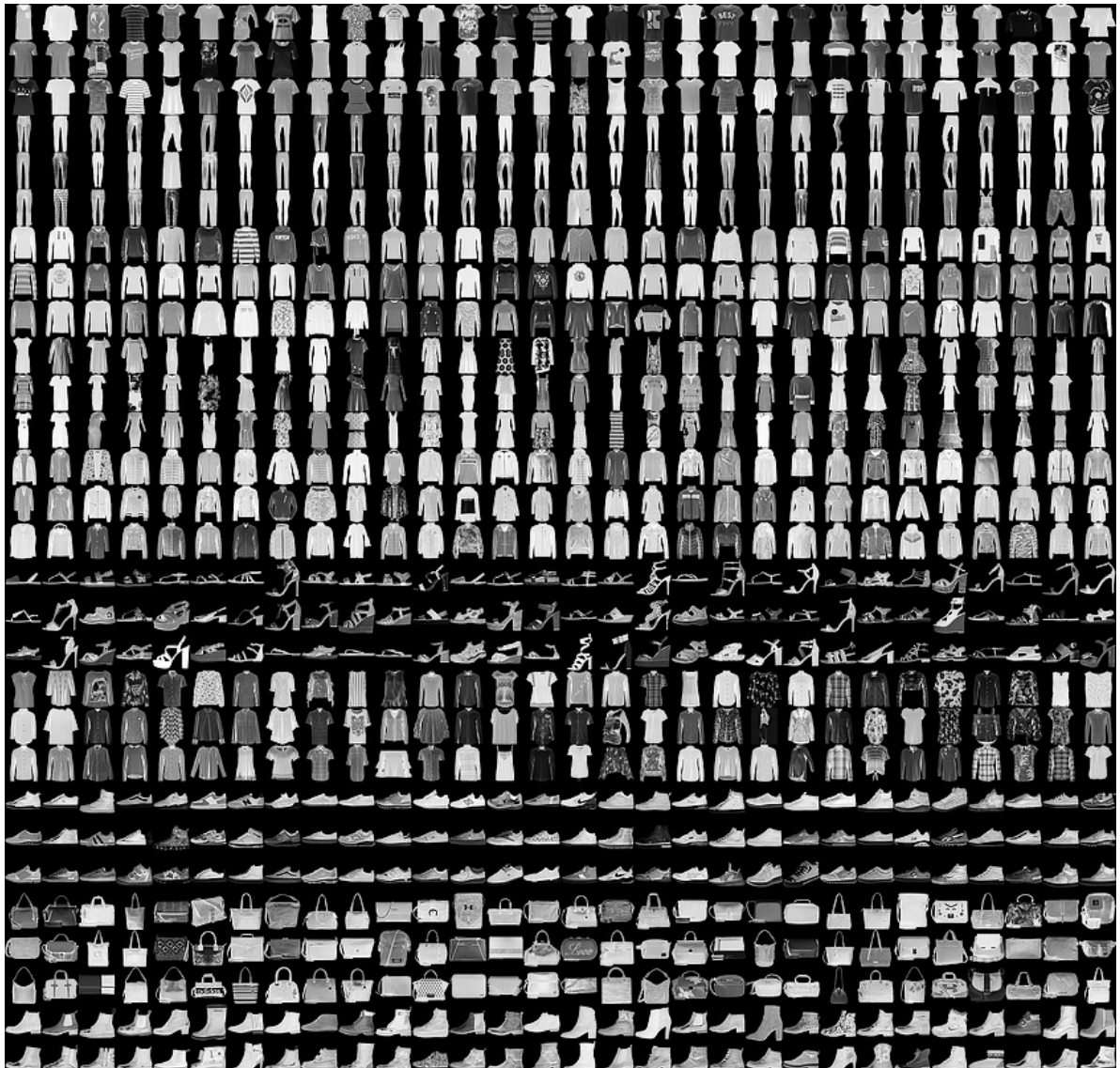
$f\left(\boldsymbol{x}_6 =\right.$  $\left.\right) = -1$ $f\left(\boldsymbol{x}_7 =\right.$  $\left.\right) = 1$

Extended MNIST database



Fashion MNIST

https://github.com/zalandoresearch/fashion-mnist



Definition 6 (2. Unsupervised Learning)

Unsupervised Learning is the task of learning (inferring) a function f that maps input vectors to their corresponding target vectors, but without any a priori knowledge about the correct mapping. Examples:

- Clustering: To group or partition the input vectors into k groups or clusters, with the vectors in each group close to each other.
- Dimensionality Reduction: The goal is to project the data from a high dimensional space down to two or three dimensions.



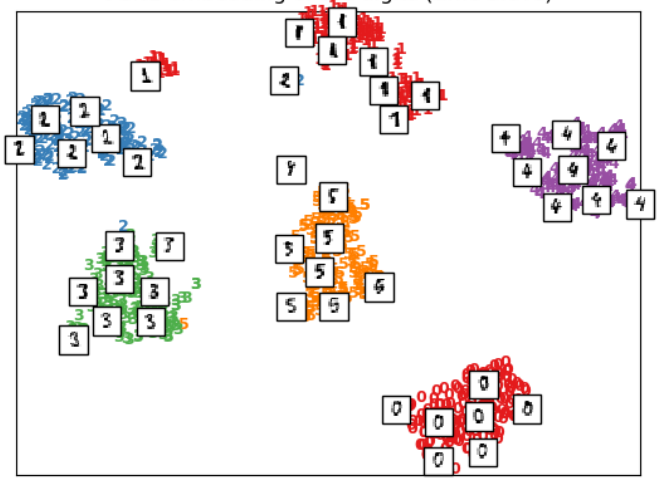
- Learns unknown patterns
- Unlabeled input data
- Find hidden patterns

2. Unsupervised Learning: Clustering

A selection from the 64-dimensional digits dataset



t-SNE embedding of the digits (time 5.87s)



$\boldsymbol{x} \in X$: Input data

$c_k :\in C$: Output Cluster

CLUSTERING TASK: Find a function

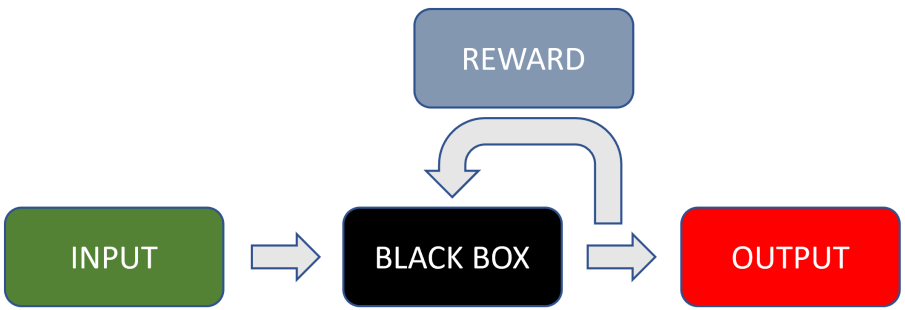
$$f_w : X \to C$$

that maps similar inputs to the same cluster $c_k \in C$.

Definition 7 (3. Reinforcement Learning)

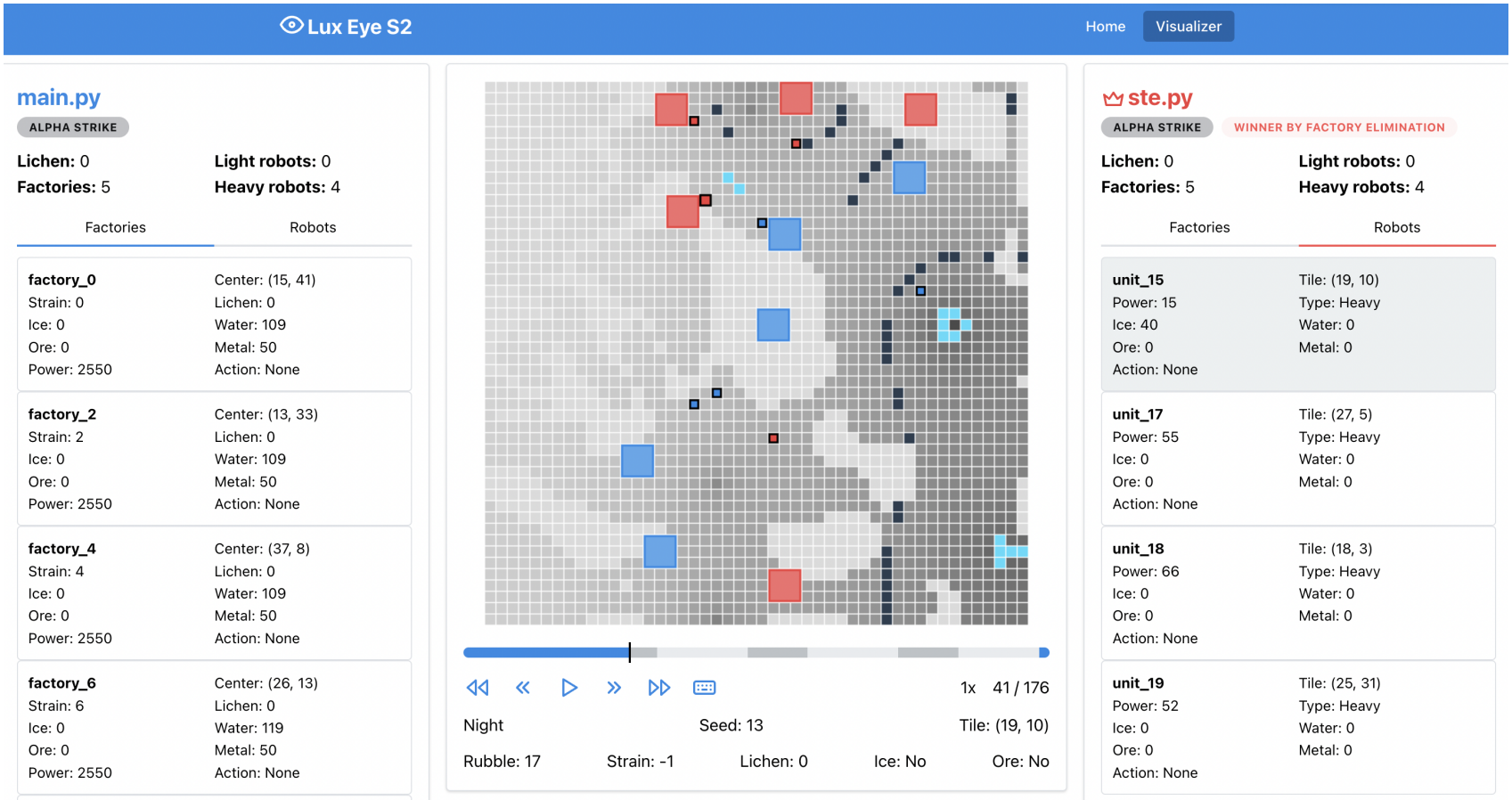
Reinforcement Learning is the task of learning from the consequences of interactions with an environment without being explicitly taught. Examples:

- Robotic Controllers: Several robots controller are trained with reinforcement learning.
- Gaming: The demo clip on the Wall game



- Learns unknown patterns
- Unlabeled input data
- Find hidden patterns

Game playground for RL: Lux AI Challenge



https://www.lux-ai.org/specs-s2

https://s2vis.lux-ai.org/

Master Thesis?

Finding the best weight vector w for $f_w(X)$

Given a data set, we divide the data into:

Training Set: it is used to learn a model, and contains usually 80% of the data.

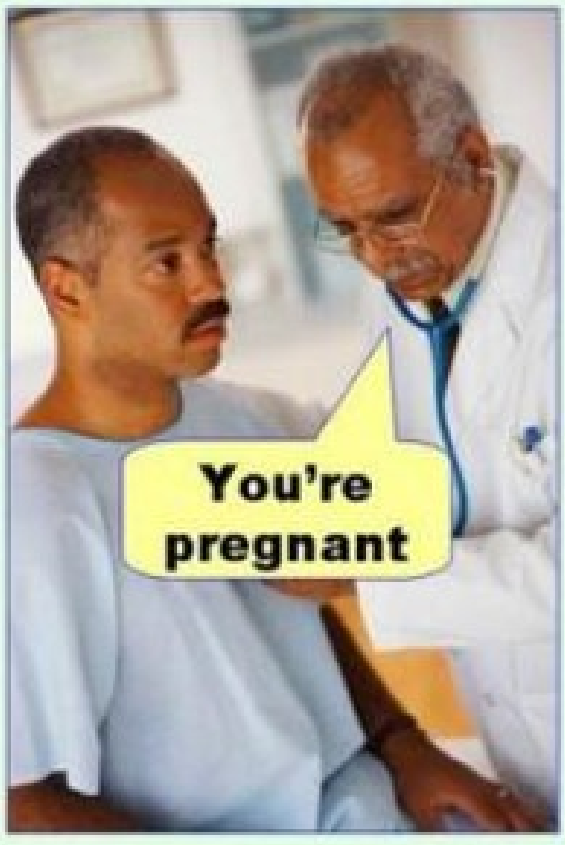
Test Set: it is used to evaluate the learned model, and it contains usually 20% of the data.

Given a data set, to increase the performance, it is better to fit the model by iteratively changing the Training and Test set, and by recording the model that yields the BEST RESULTS (see n-fold cross validation).

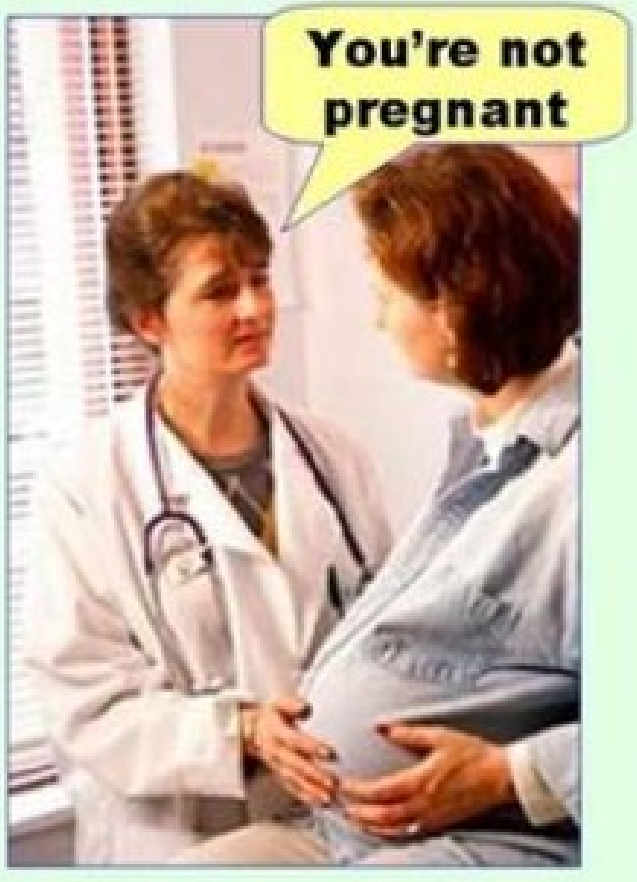
... but **BEST RESULTS** with respect to which METRIC?

Binary Classification: Confusion Matrix

Type I error
(false positive)



Type II error
(false negative)



Binary Classification: Accuracy

$$\text{Accuracy} := \frac{\#\text{TP} + \#\text{TN}}{\#\text{TP} + \#\text{TN} + \#\text{FP} + \#\text{FN}}$$

TP: True Positive

TN: True Negative

FP: False Positive

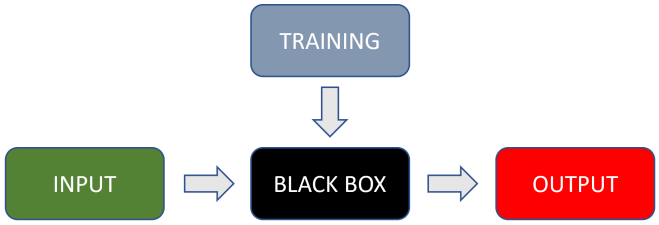
FN: False Negative

QUESTION: When Accuracy make sense for evaluating a binary classifier? Example?

Evaluating Classifiers with Imbalanced Classes

		Condition (as determined by "Gold standard")		
		Condition Positive	Condition Negative	
Test Outcome	Test Outcome Positive	**True Positive**	**False Positive** (Type I error)	**Positive predictive value** = Σ True Positive / Σ Test Outcome Positive
	Test Outcome Negative	**False Negative** (Type II error)	**True Negative**	Negative predictive value = Σ True Negative / Σ Test Outcome Negative
		Sensitivity = Σ True Positive / Σ Condition Positive	Specificity = Σ True Negative / Σ Condition Negative	

Why Optimization? To Minimize the errors



1. **Mean Square Error/Quadratic Loss/L2 Loss**:

$$MSE := \frac{1}{n}\sum_{i=1}^{n}(y_i - f(x_i; w))^2$$

2. **Mean Absolute Error/L1 Loss**:

$$MAE := \frac{1}{n}\sum_{i=1}^{n}|y_i - f(x_i; w))|$$

3. **Cross Entropy Loss/Negative Log Likelihood**:

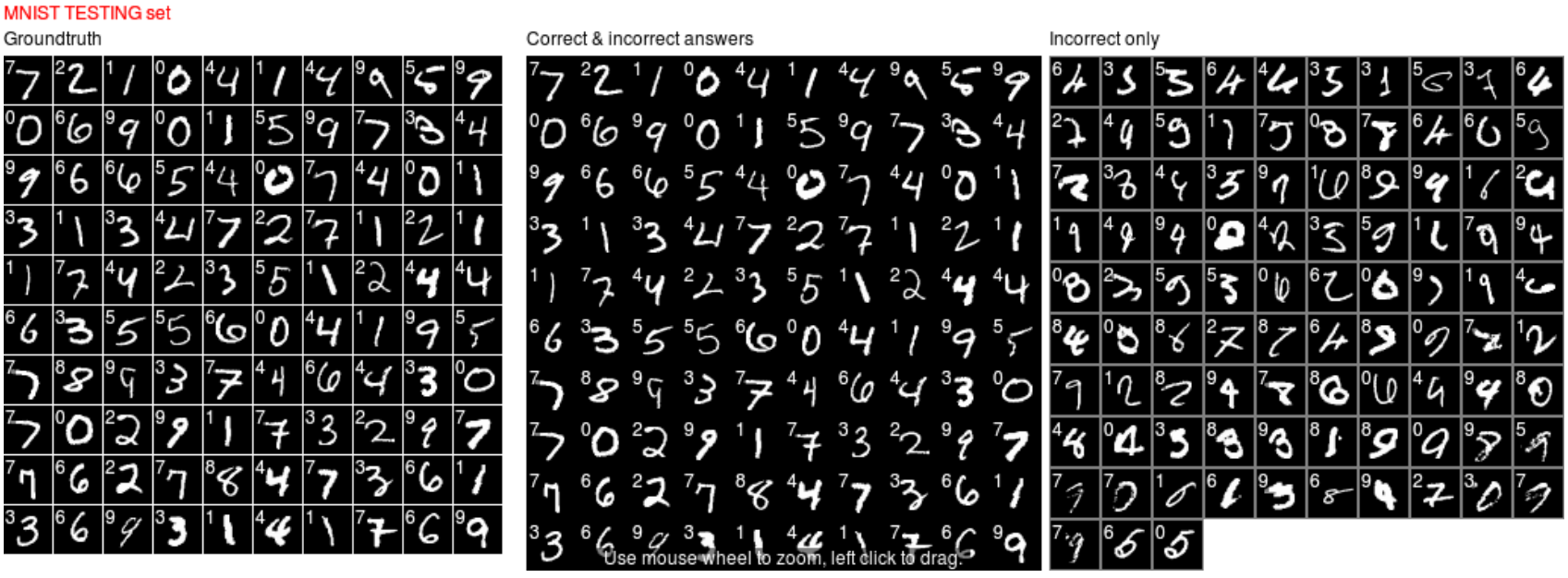
$$CrossEntropy := \sum_{i=1}^{n} -(y_i \log(f(x_i; w)) + (1 - y_i)\log(1 - f(x_i; w)))$$

Probabilistic Interpretations for Loss Functions

- Optimization algorithms plays a crucial role in minimizing the loss functions
- Most succesful loss functions have a Probabilistic Interpretations
- In some context, Optimal Transport distances can be used as loss functions (see Wasserstein GAN)

MNIST database

The dataset contains 60 000 images with 28×28 pixels, around image 6 000 per digit. Training 60 000, Testing: 10 000.
`http://yann.lecun.com/exdb/mnist/`



Accuracy on the MNIST dataset

Classifiers [edit]

This is a table of some of the machine learning methods used on the database and their error rates, by type of classifier:

Type	Classifier	Distortion	Preprocessing	Error rate (%)
Linear classifier	Pairwise linear classifier	None	Deskewing	7.6[9]
K-Nearest Neighbors	K-NN with non-linear deformation (P2DHMDM)	None	Shiftable edges	0.52[19]
Boosted Stumps	Product of stumps on Haar features	None	Haar features	0.87[20]
Non-linear classifier	40 PCA + quadratic classifier	None	None	3.3[9]
Support-vector machine (SVM)	Virtual SVM, deg-9 poly, 2-pixel jittered	None	Deskewing	0.56[21]
Deep neural network (DNN)	2-layer 784-800-10	None	None	1.6[22]
Deep neural network	2-layer 784-800-10	Elastic distortions	None	0.7[22]
Deep neural network	6-layer 784-2500-2000-1500-1000-500-10	Elastic distortions	None	0.35[23]
Convolutional neural network (CNN)	6-layer 784-40-80-500-1000-2000-10	None	Expansion of the training data	0.31[15]
Convolutional neural network	6-layer 784-50-100-500-1000-10-10	None	Expansion of the training data	0.27[24]
Convolutional neural network	Committee of 35 CNNs, 1-20-P-40-P-150-10	Elastic distortions	Width normalizations	0.23[8]
Convolutional neural network	Committee of 5 CNNs, 6-layer 784-50-100-500-1000-10-10	None	Expansion of the training data	0.21[17][18]

Machine Learning Algorithms

SUPERVISED LEARNING ALGORITHMS

CLASSIFICATION	REGRESSION
K-Nearest Neighbors	Linear Regression
Naïve Bayes Classifier	Nonlinear Regression
Support Vector Machines	Ordinary Least Square
Neural Networks	Neural Networks
Decision Trees	
Random Forests	

Machine Learning Algorithms

SUPERVISED LEARNING ALGORITHMS

CLASSIFICATION	REGRESSION
K-Nearest Neighbors	Linear Regression
Naïve Bayes Classifier	Nonlinear Regression
Support Vector Machines	Ordinary Least Square
Neural Networks	Neural Networks
Decision Trees	
Random Forests	

Machine Learning Algorithms

UNSUPERVISED LEARNING ALGORITHMS

CLUSTER ANALYSIS

- K-means clustering
- Hierarchical clustering

DIMENSION REDUCTION

- Linear Discriminant Analysis
- Principal Component Analysis
- *Manifold Learning*
- *Autoencoders*

Outline

1 Introduction

2 Images

3 Machine Learning

4 Neural Networks

5 Convolutional Neural Networks

6 Transfer Learning

Neural Networks



Biological Motivation

1. The human brain is estimated to contain a densely interconnected network of around 10^{11} neurons
2. Each neuron is connected to around others 10^4 neurons. Neurons activity is typically excited or inhibited through connections to other neurons.
3. The switching of neurons is of order of 10^{-3} seconds, much slower than modern computers, which have switching speeds of 10^{-10} seconds
4. Humans take around 10^{-1} seconds to recognize a friend (compare to switching speed...)
5. Human brain performs a highly parallel tasks

Hollywoodian Neural Networks



John Connor: Can you learn stuff you haven't been programmed with so you could be... you know, more human? And not such a dork all the time?

The Terminator: **My CPU is a neural-net processor; a learning computer**. But Skynet presets the switch to read-only when we're sent out alone.

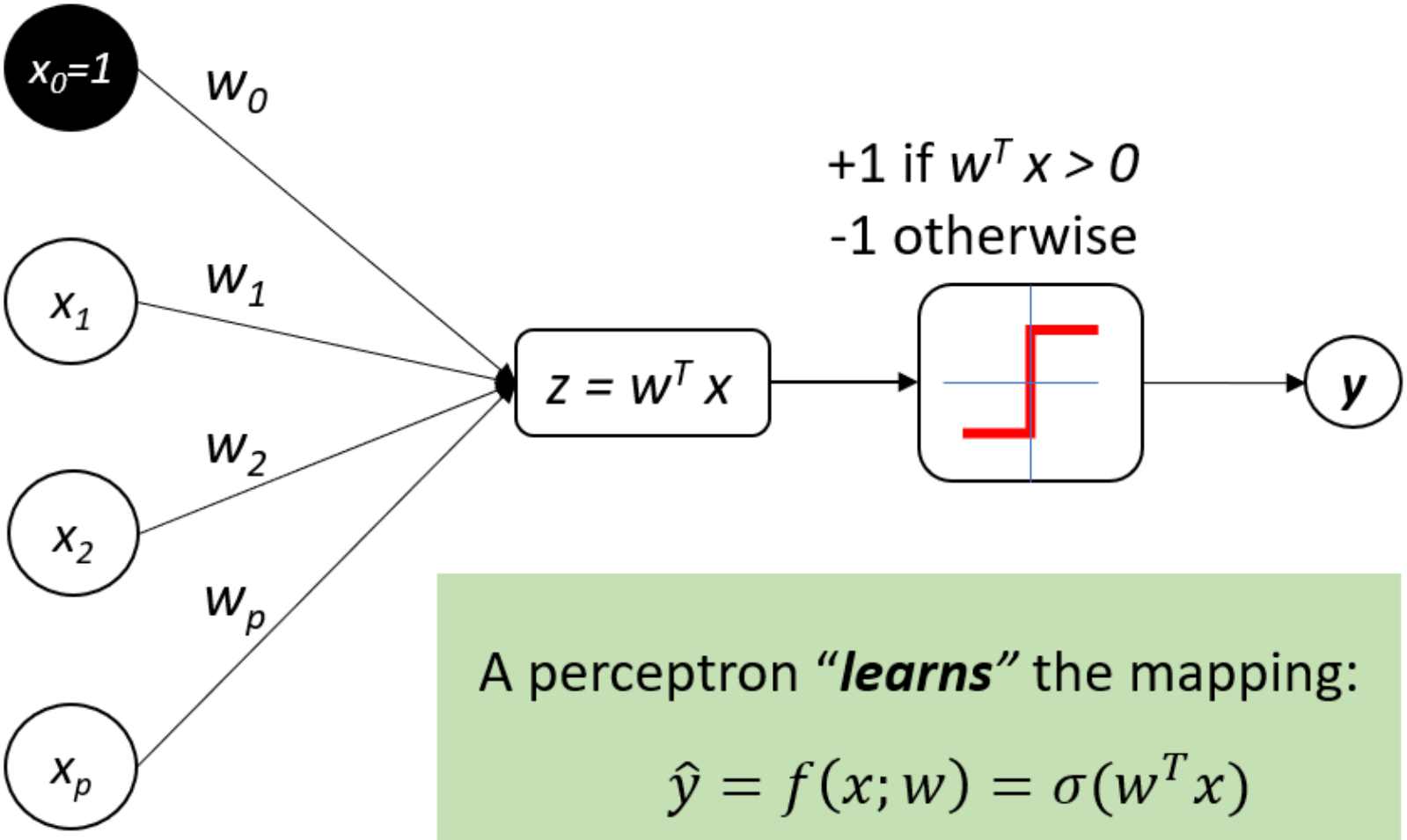
Sarah Connor: Doesn't want you doing too much thinking, huh?

The Terminator: No.

Let us start simple: Single Perceptron



Perceptron

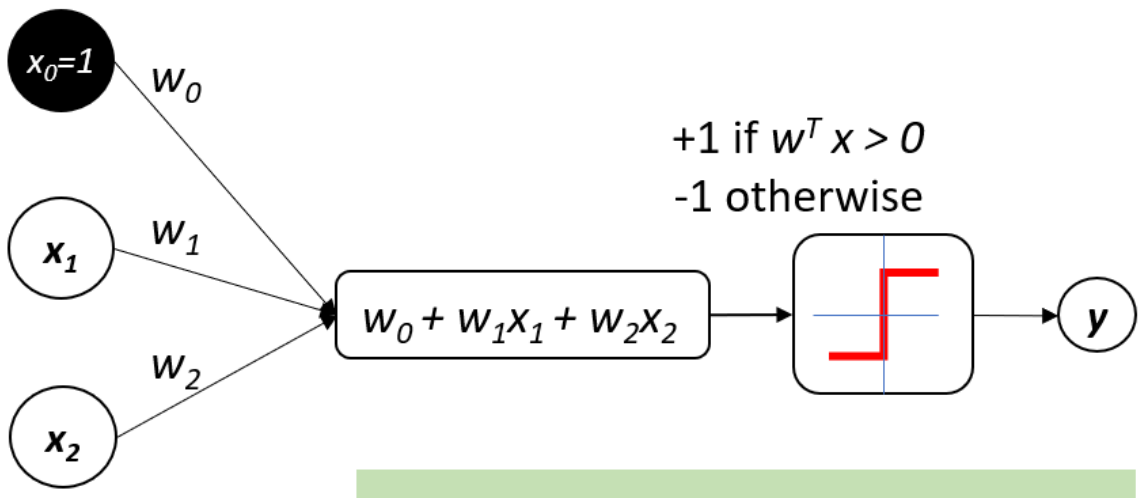


A perceptron "***learns***" the mapping:

$$\hat{y} = f(x; w) = \sigma(w^T x)$$

Perceptron

Logical AND		
x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

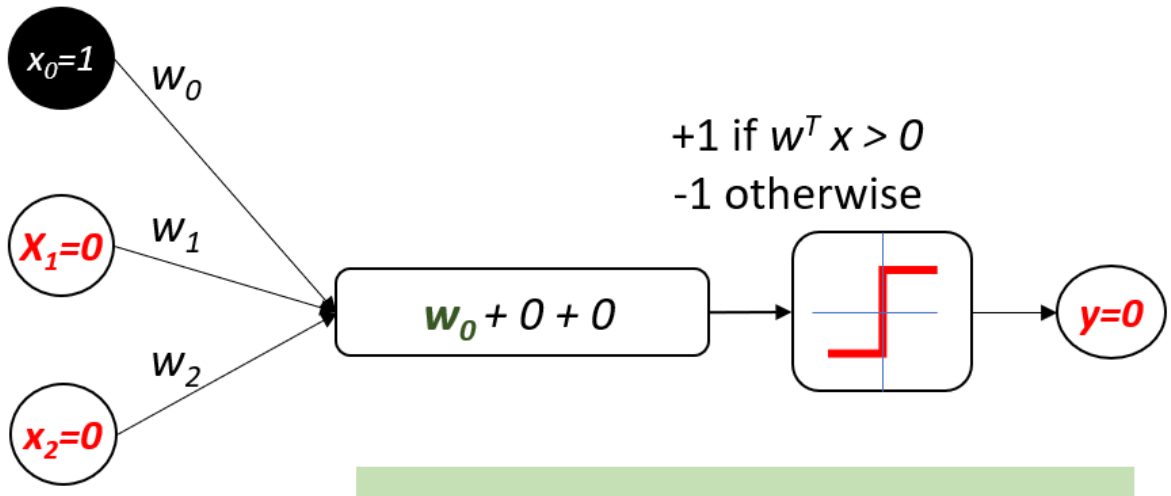


A perceptron "***learns***" the mapping:

$$\hat{y} = f(x; w) = \sigma(w^T x)$$

Perceptron

Logical AND		
x_1	x_2	y
0	**0**	**0**
0	1	0
1	0	0
1	1	1

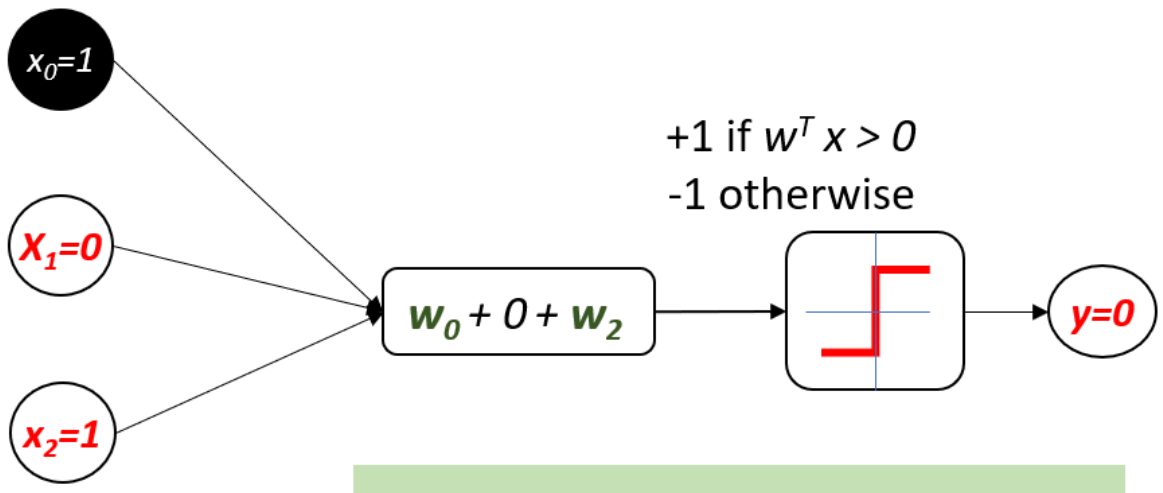


A perceptron "***learns***" the mapping:

$$\hat{y} = f(x; w) = \sigma(w^T x)$$

Perceptron

Logical AND		
x_1	x_2	y
0	0	0
0	**1**	**0**
1	0	0
1	1	1

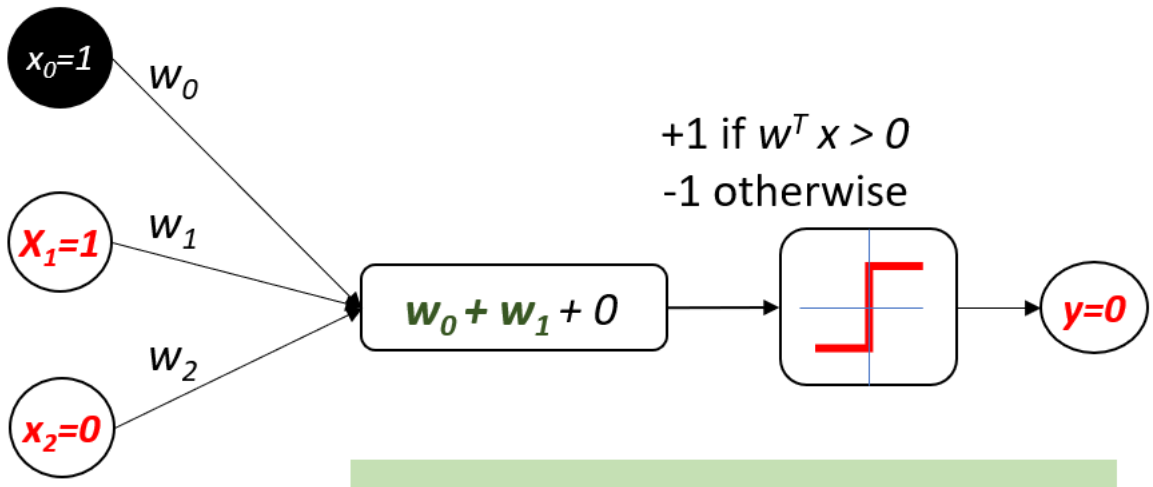


A perceptron "***learns***" the mapping:

$$\hat{y} = f(x; w) = \sigma(w^T x)$$

Perceptron

Logical AND		
x_1	x_2	y
0	0	0
0	1	0
1	**0**	**0**
1	1	1

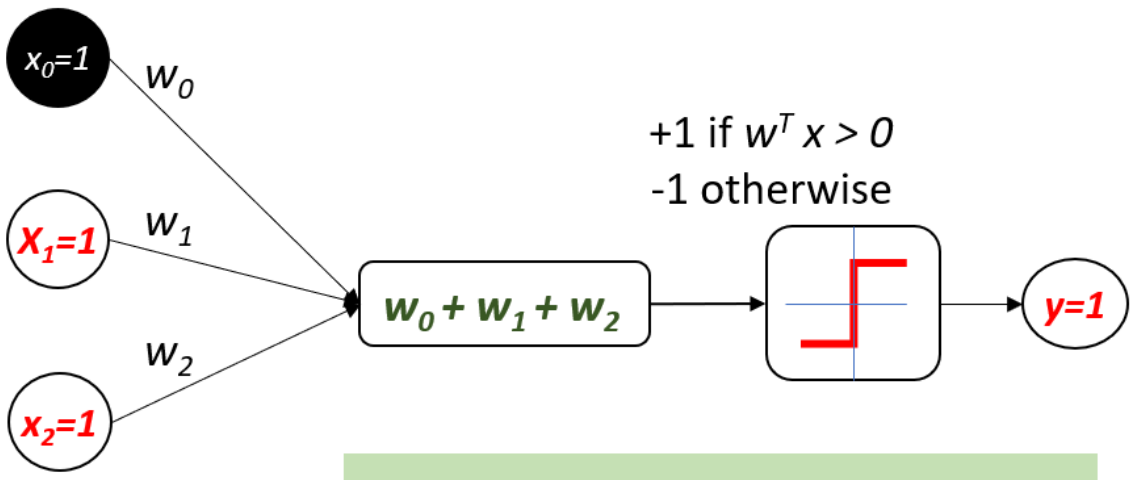


A perceptron "***learns***" the mapping:

$$\hat{y} = f(x; w) = \sigma(w^T x)$$

Perceptron

Logical AND		
x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	**1**	**1**



A perceptron "***learns***" the mapping:

$$\hat{y} = f(x; w) = \sigma(w^T x)$$

Video Facebook

VIDEO CLIP FB

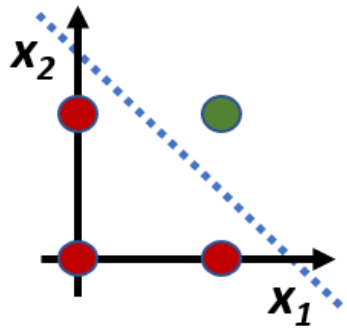
https://mashable.com/2016/12/01/facebook-ai-video-lessons

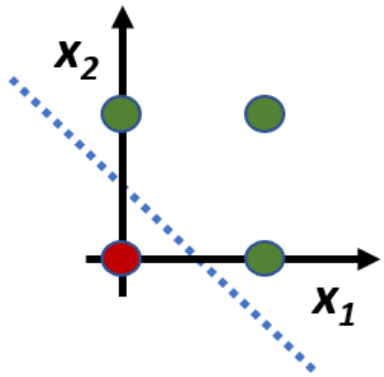
Perceptron: limits

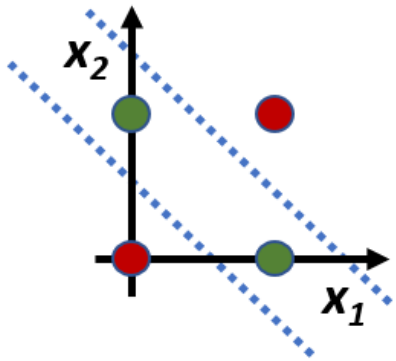
AND		
x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

OR		
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	1

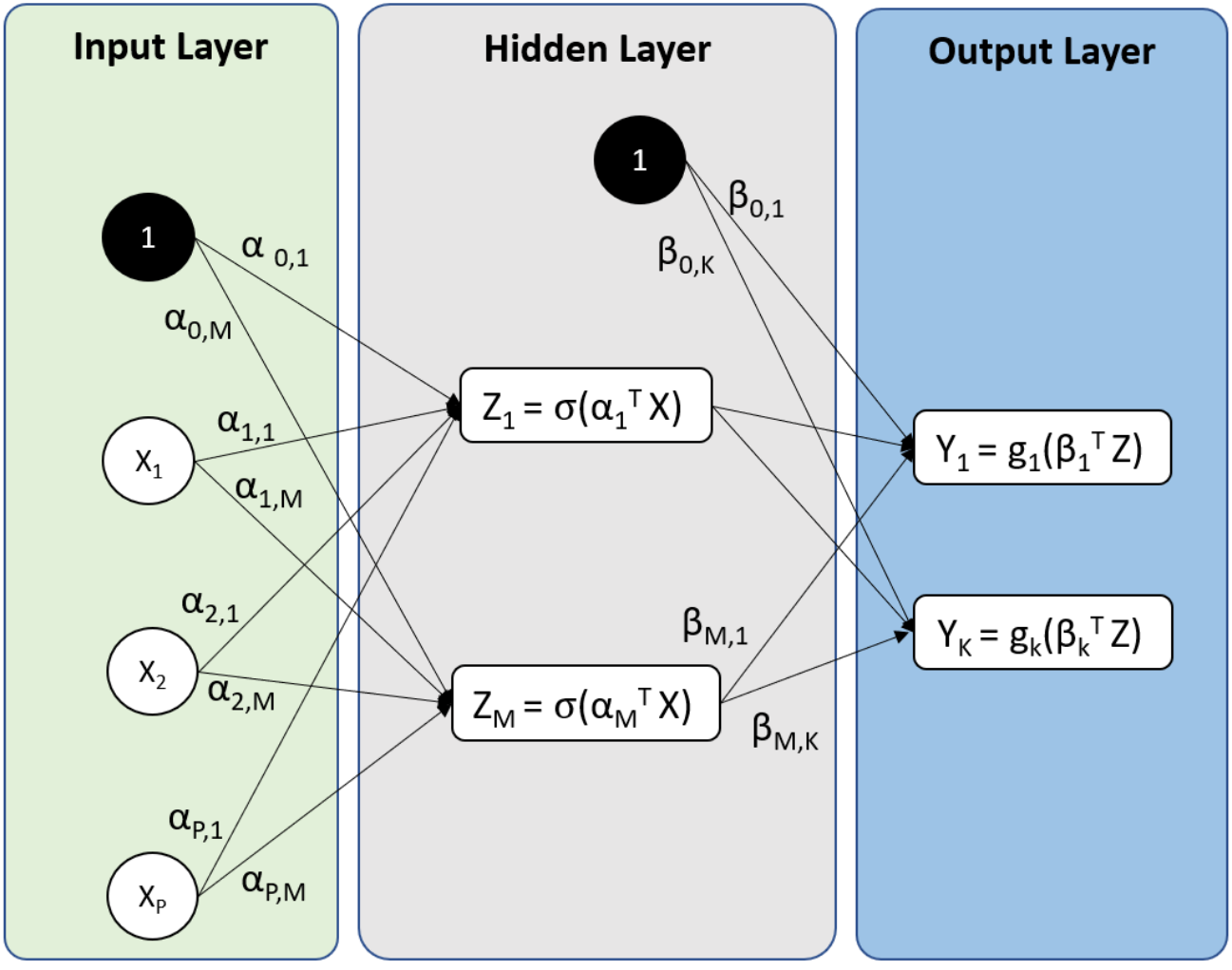
XOR		
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0







Multi Layer Neural Network

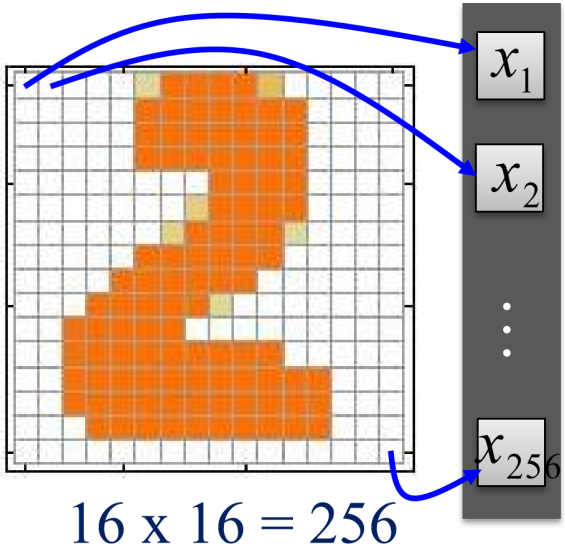


$\alpha \in R^{P \times M}$

$\beta \in R^{M \times K}$

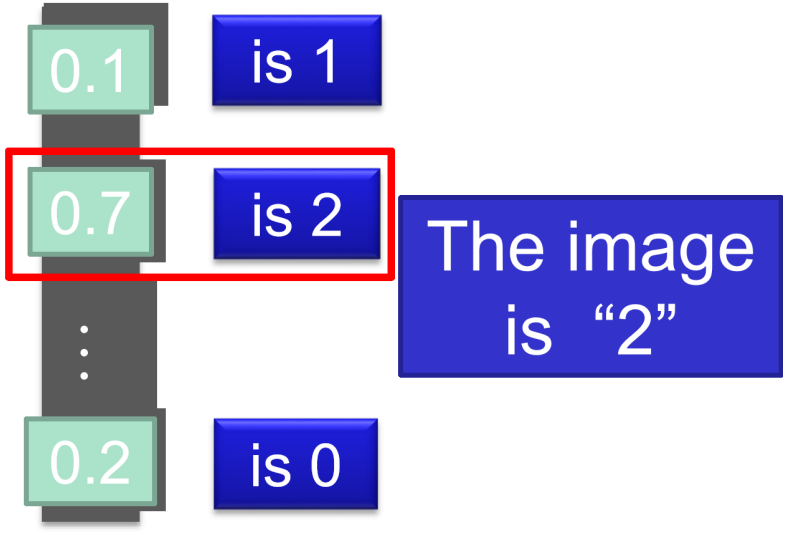
Images as vectors

Input



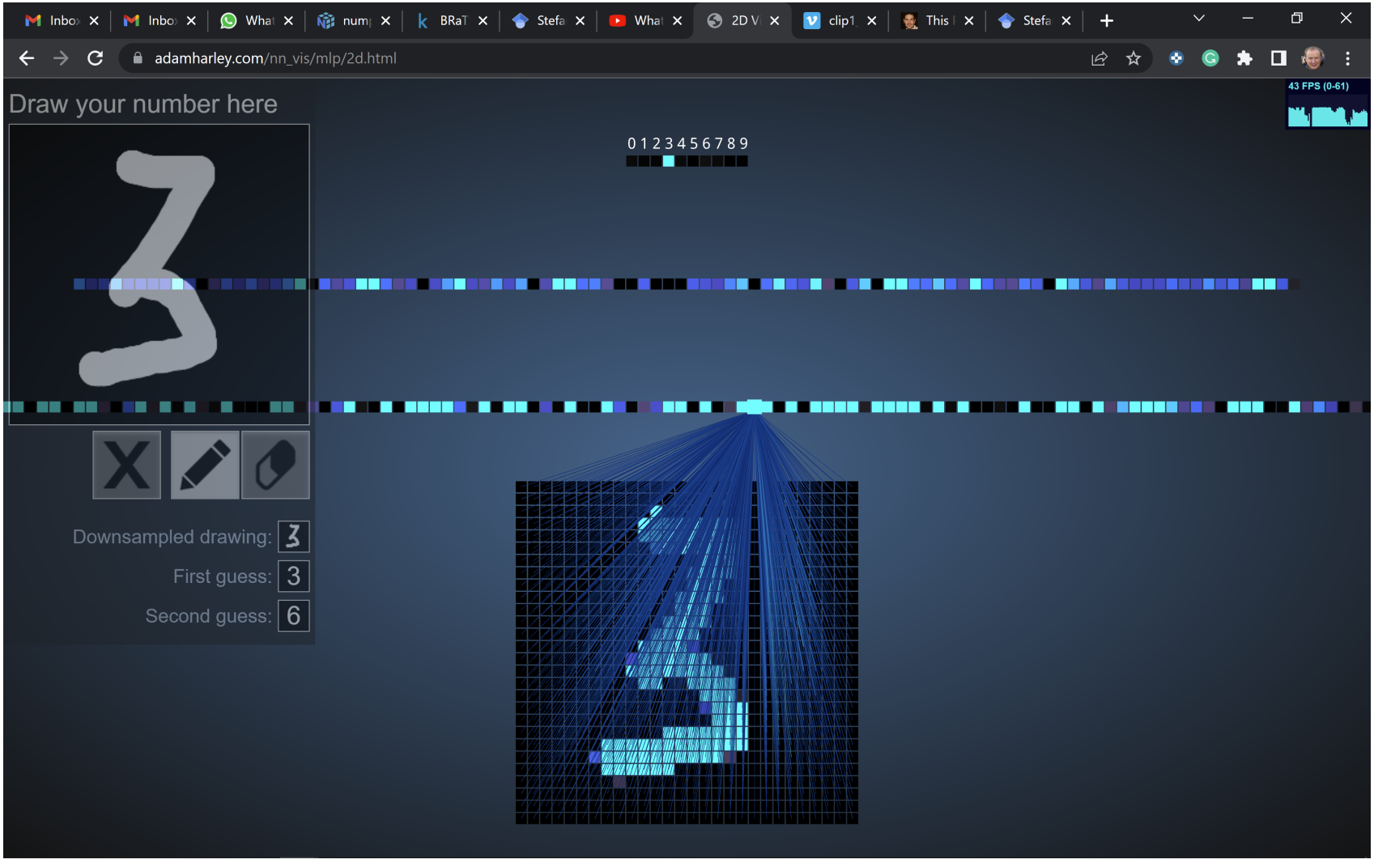
16 x 16 = 256
Ink → 1
No ink → 0

Output



Each dimension represents the confidence of a digit.

Multilayer Feedforward Neural Networks



https://adamharley.com/nn_vis/

Backpropagation as Automatic Differentiation

DEMO ON LEARNING WEIGHTS: Backpropagation Algorithms

https://www.youtube.com/watch?v=Ilg3gGewQ5U

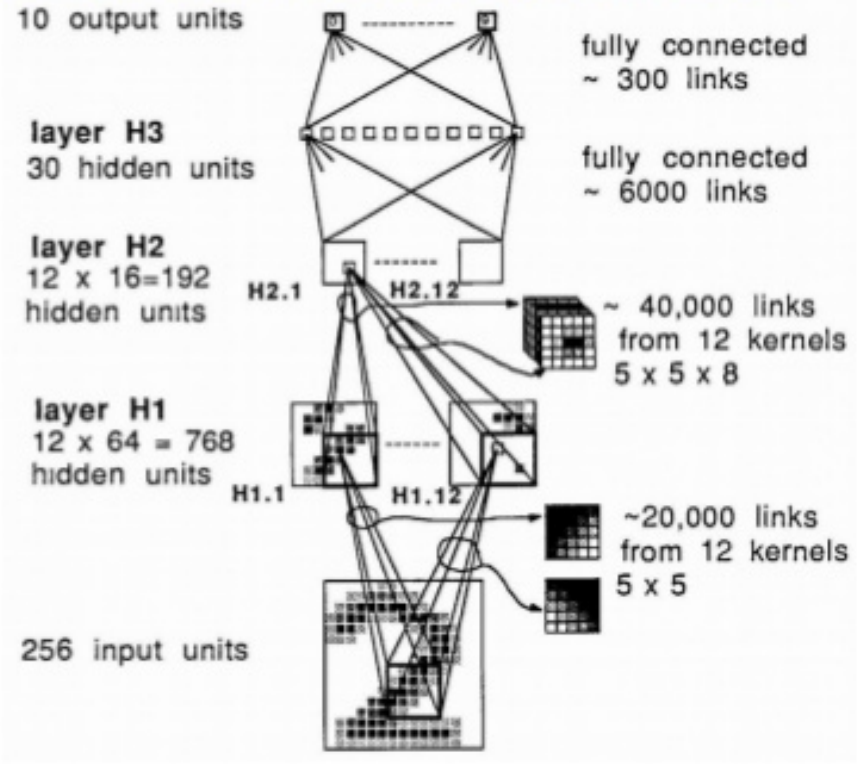
First Convolutional Neural Network

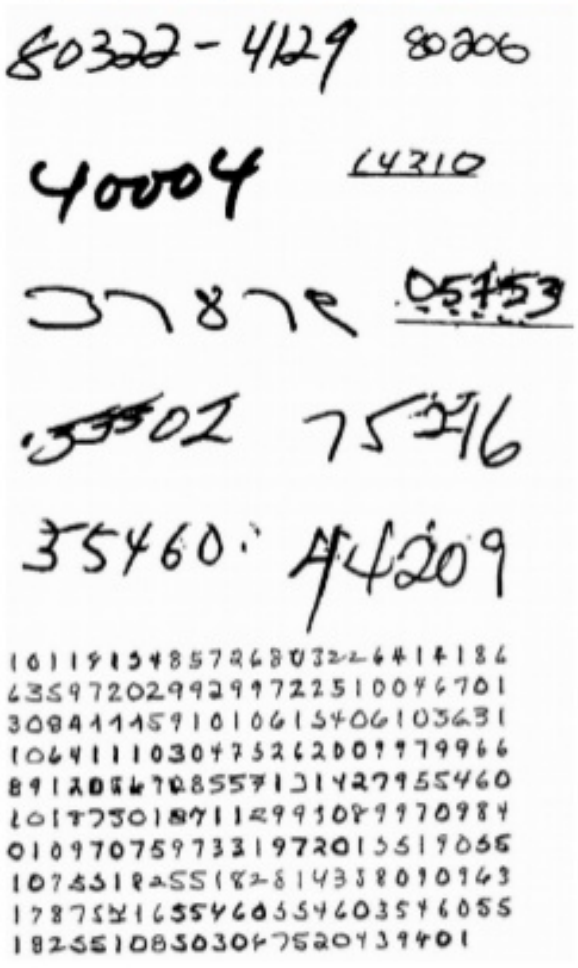


Trained with Backprop.

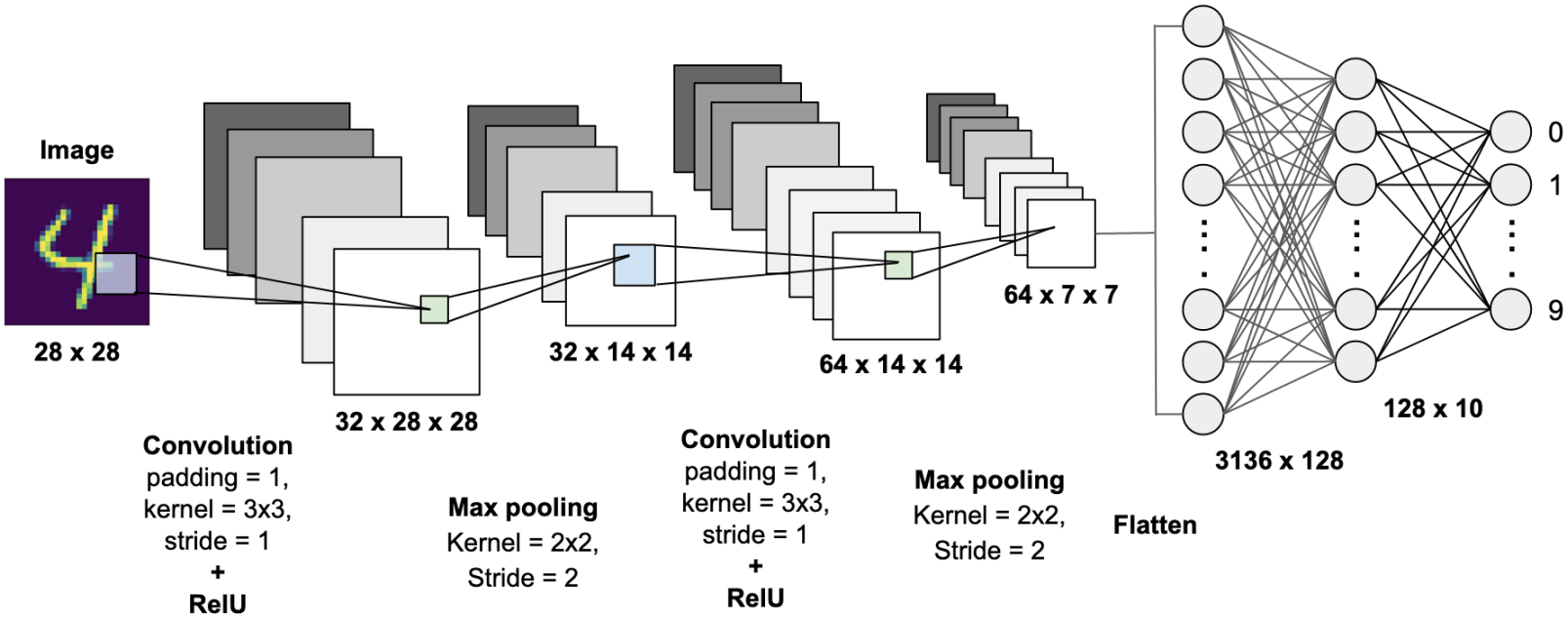
USPS Zipcode digits: 7300 training, 2000 test.

Convolution with stride. No separate pooling.





Modern Convolutional Neural Network



NOVEMBER 8, 2016

Fujitsu leverages AI to develop highly accurate recognition technology for strings of handwritten Chinese characters

by Fujitsu

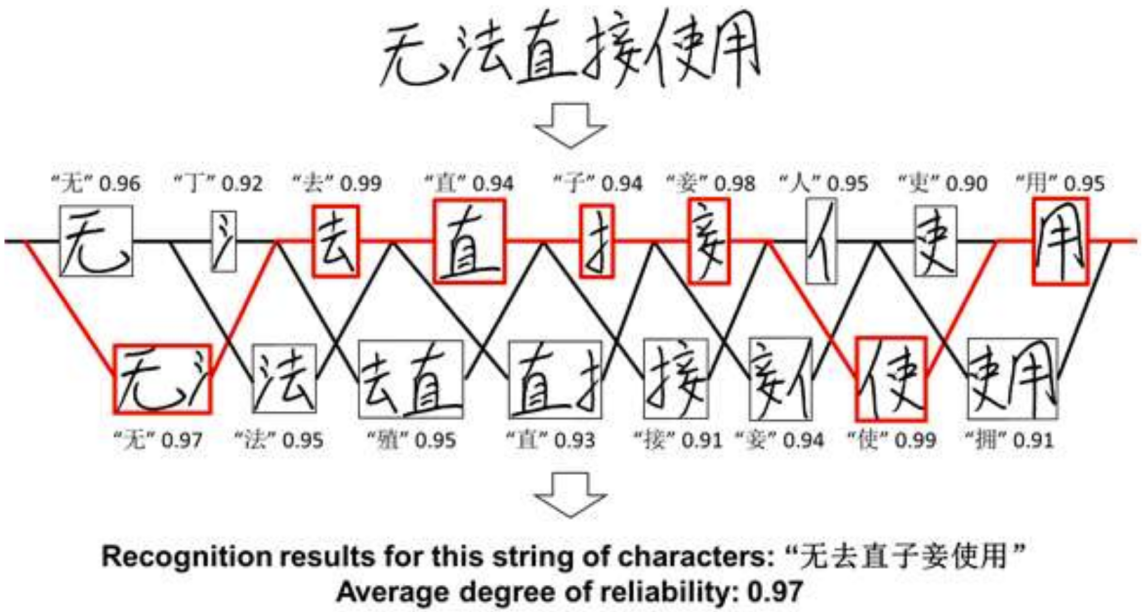
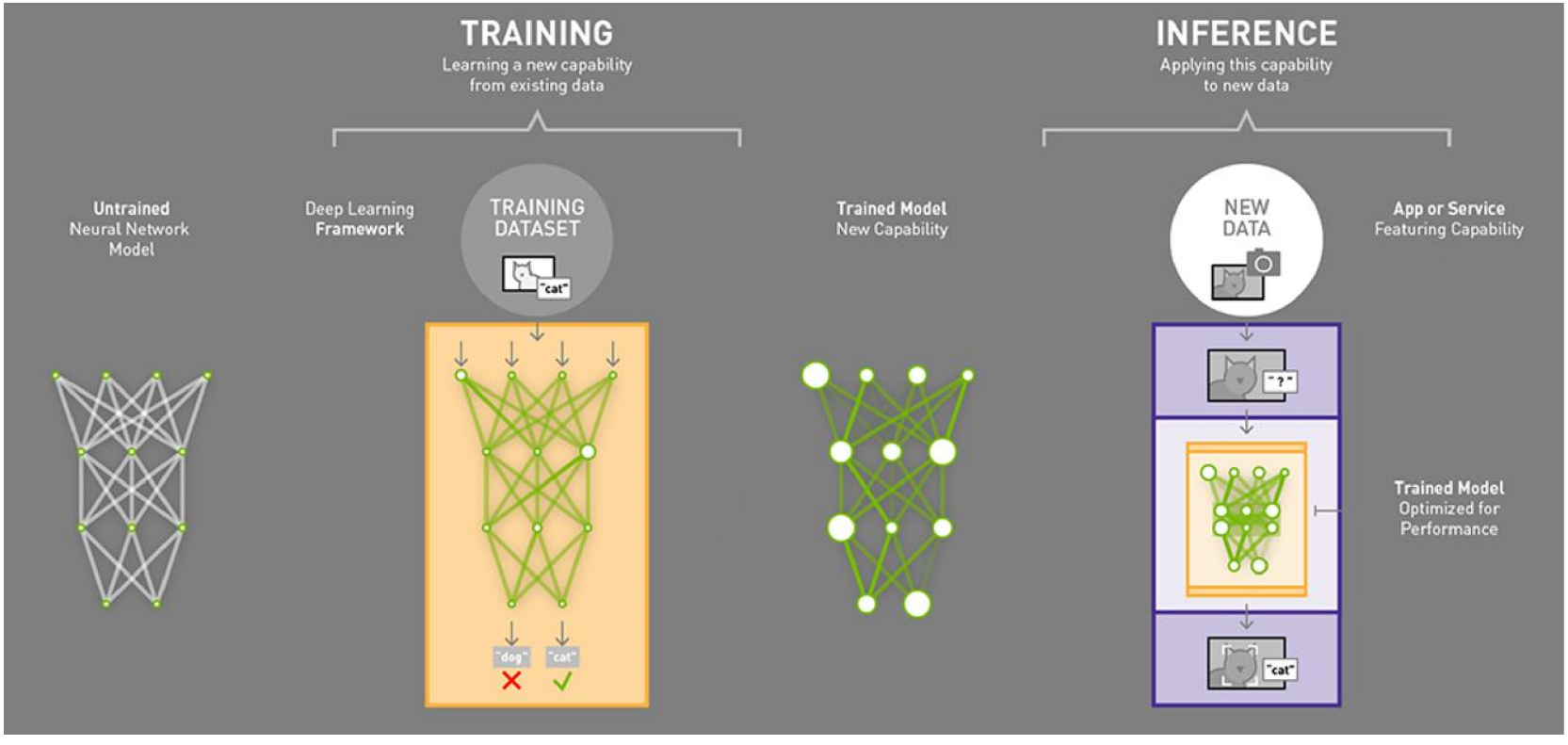
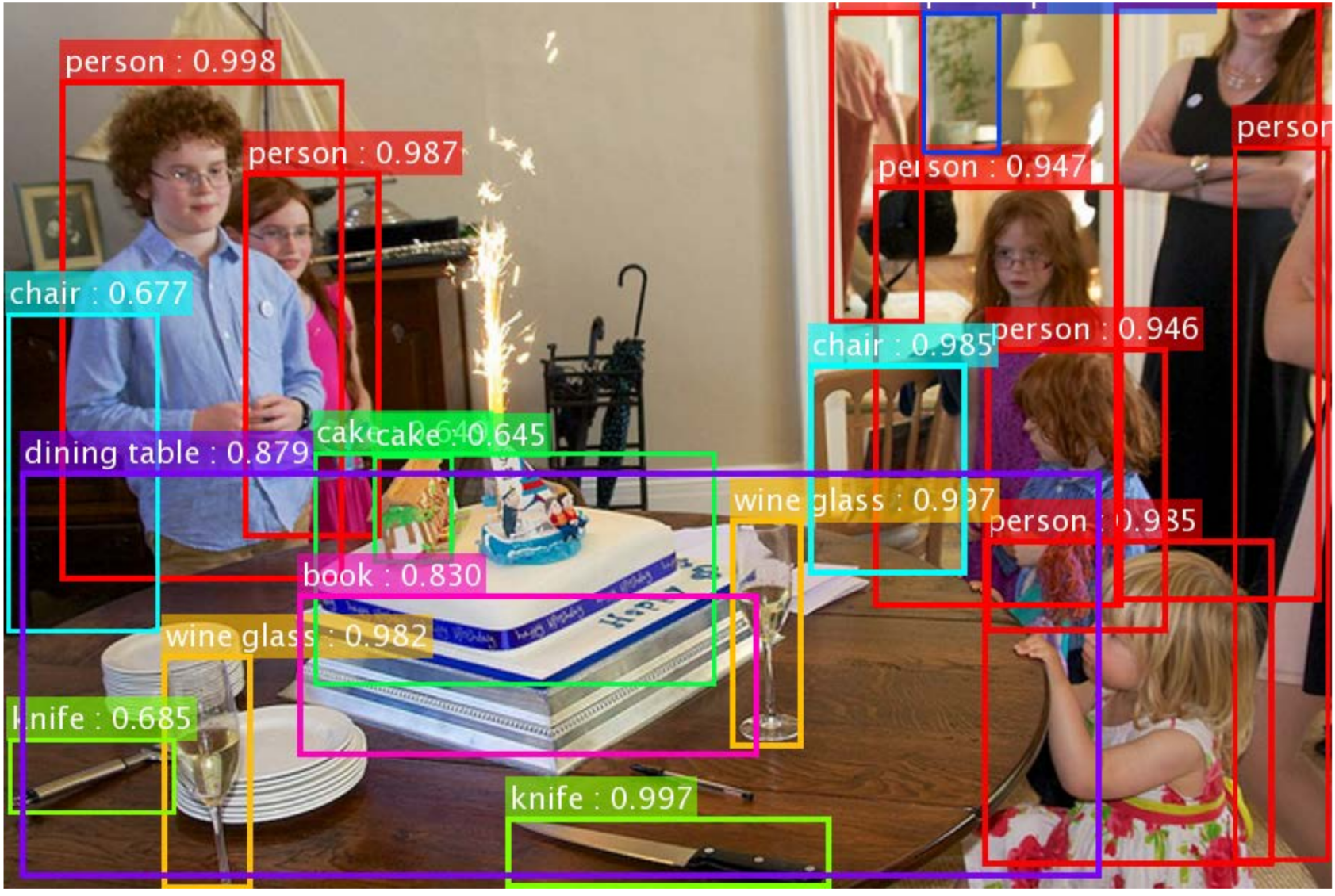


Figure 1: Recognition results for a string of characters with existing deep learni…

Training and Inference



State-of-the-art in Object recognition



1000 Artificial Neurons vs. a single Biological Neuron

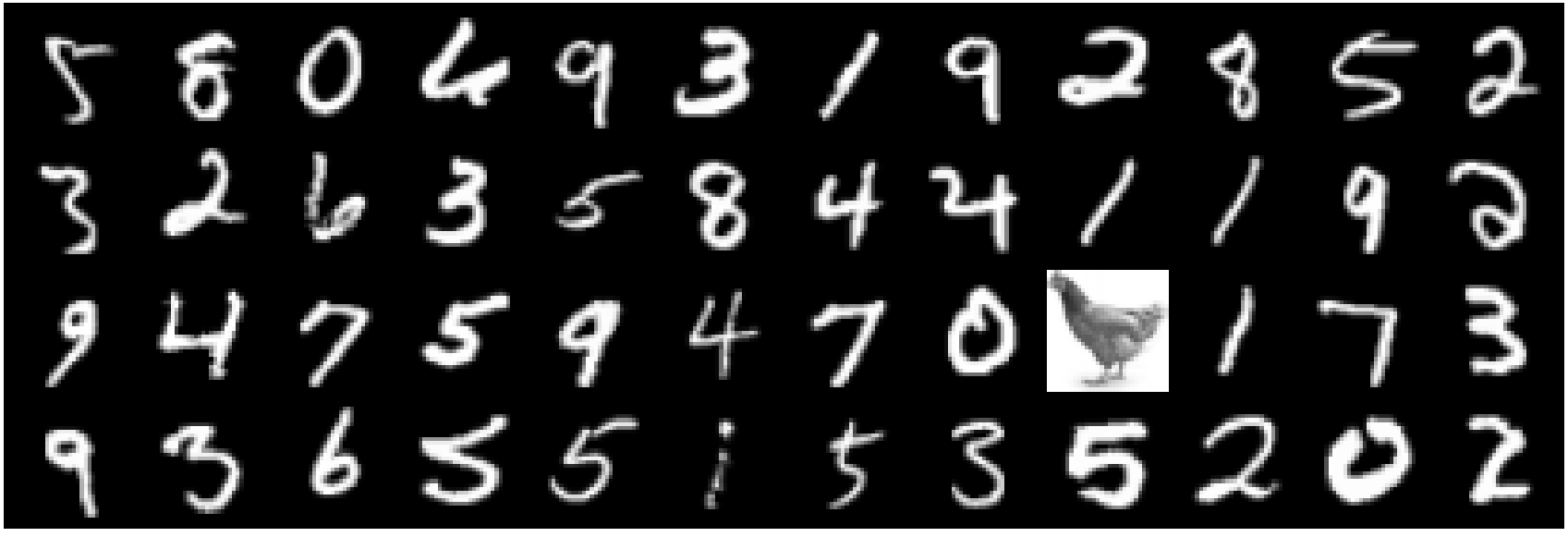
VIDEO

The evolution of NN architectures

DEMO CLIP ON VIMEO

https://vimeo.com/154085950

WARNING: Limits and risks 1/3



WARNING: Limits and risks 2/3

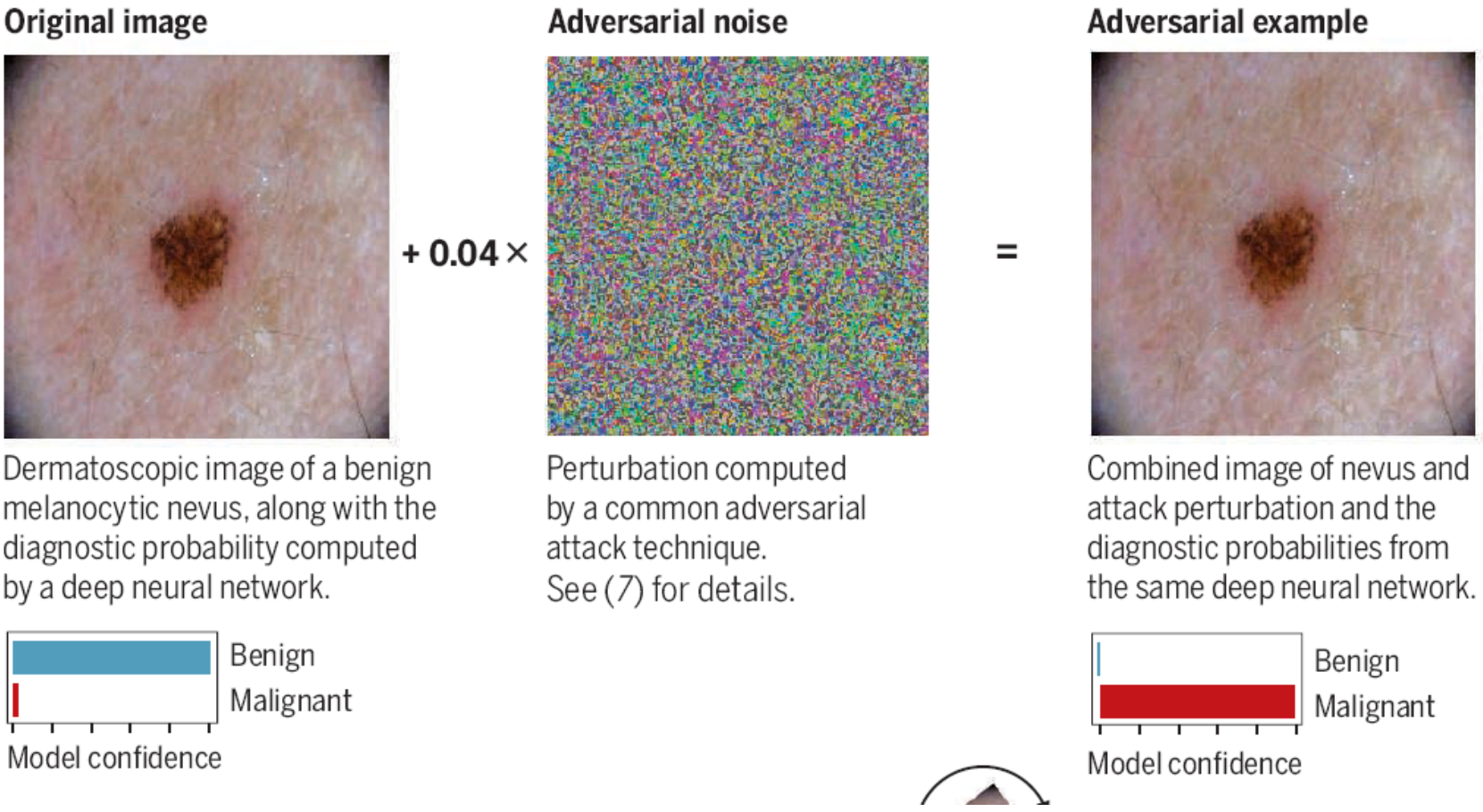


76.4%
confident
that it's a 5



99.9%
confident
that it's a 5

WARNING: Limits and risks 3/3



Kaggle Challenge: Melanoma Classification

https://www.kaggle.com/c/siim-isic-melanoma-classification

Outline

1 Introduction

2 Images

3 Machine Learning

4 Neural Networks

5 Convolutional Neural Networks

6 Transfer Learning

Convolutional Neural Network Architecture

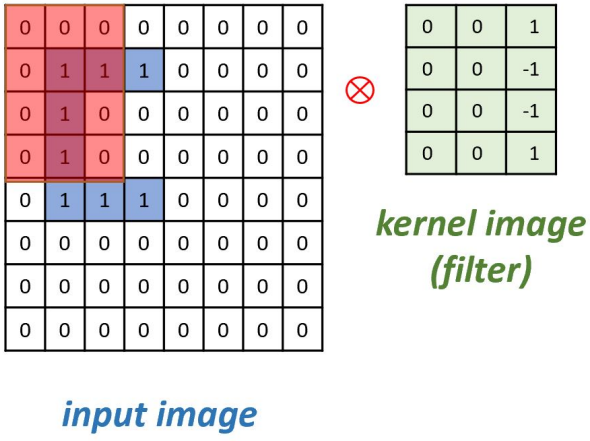
ANIMATED APPLET

`http://www.cs.cmu.edu/~aharley/vis/`

Convolution

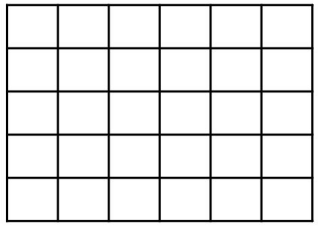
0	0	0	0	0	0	0	0
0	1	1	1	0	0	0	0
0	1	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Convolution

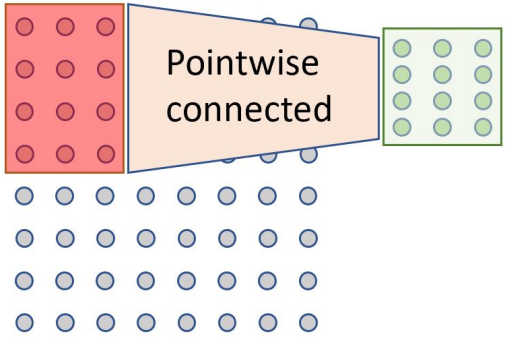


input image

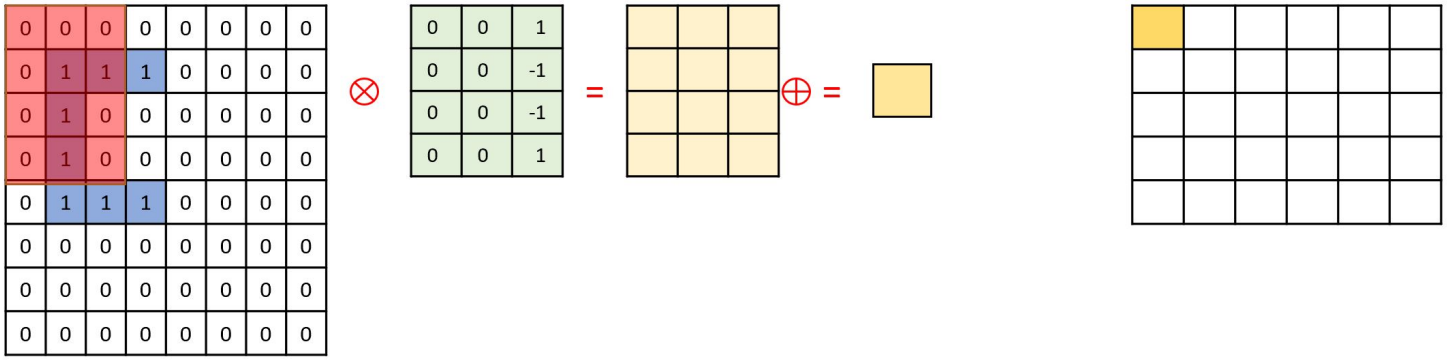
kernel image (filter)

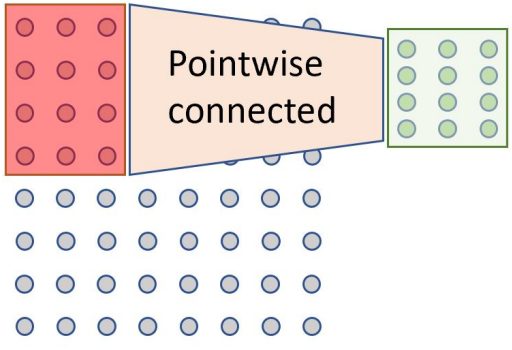


new image (feature map)

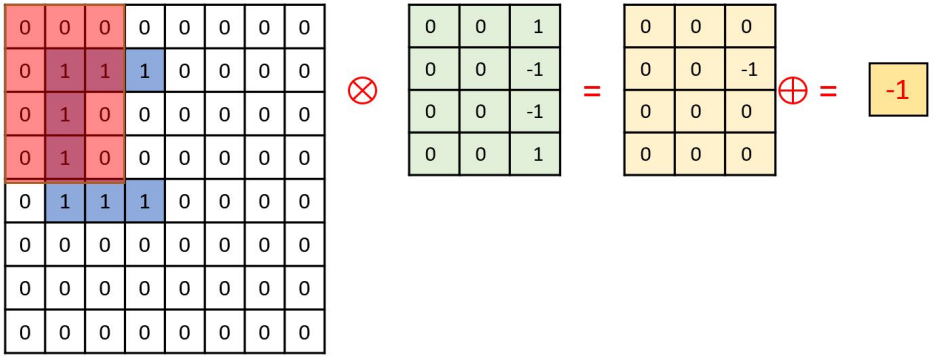


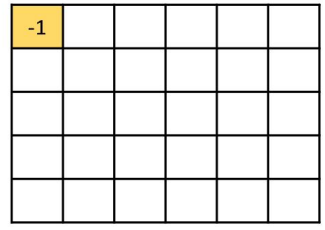
Convolution

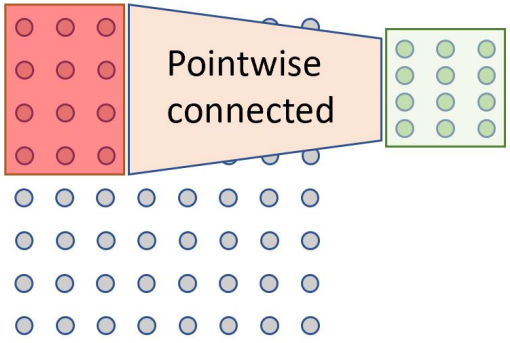




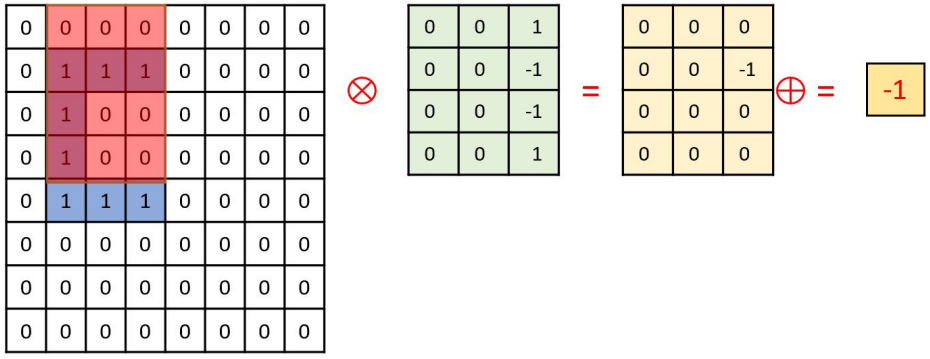
Convolution

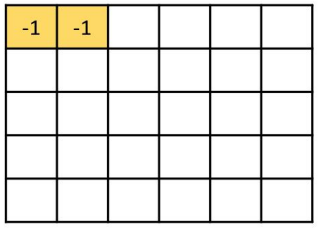


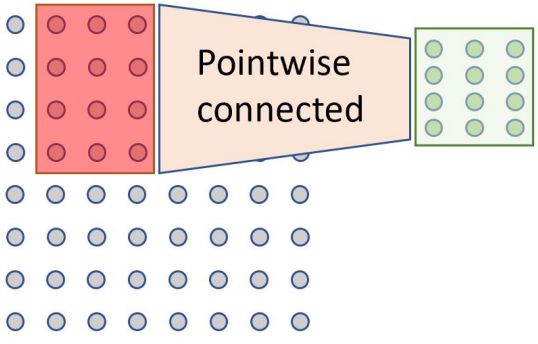




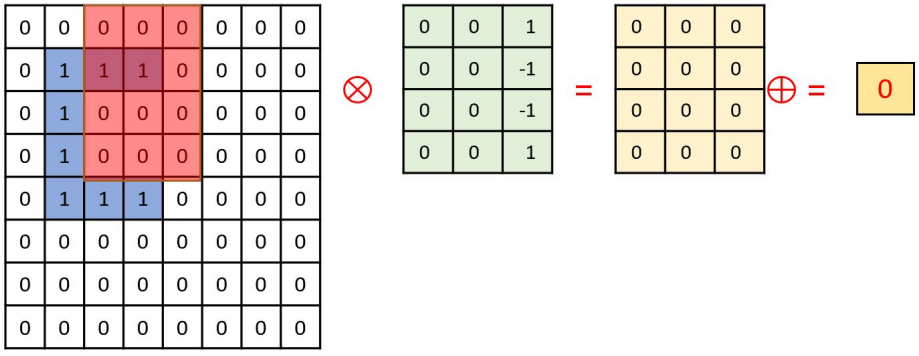
Convolution

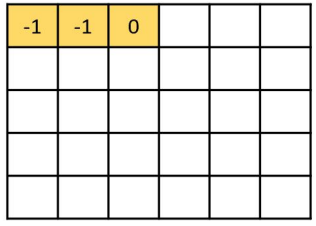


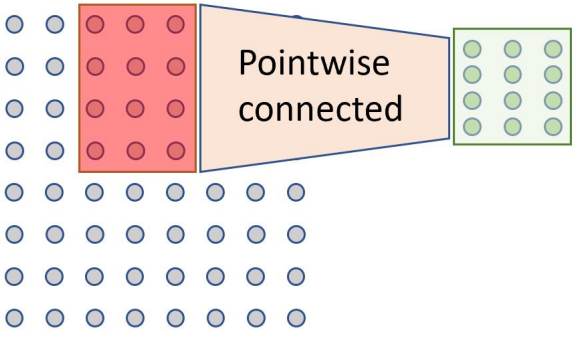




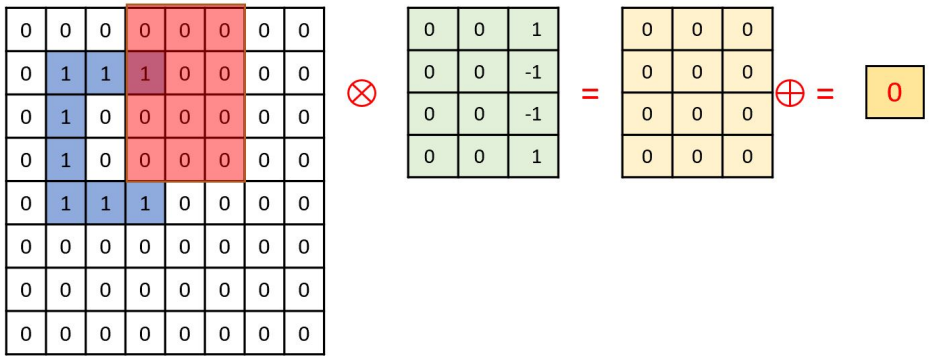
Convolution

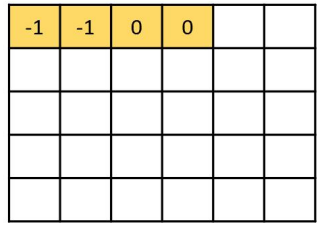


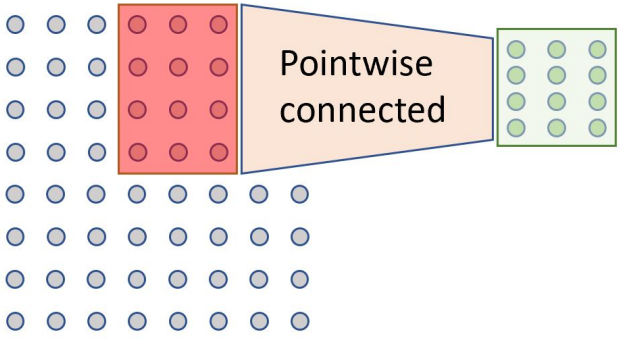




Convolution







Convolution

0	0	0	0	0	0	0	0
0	1	1	1	0	0	0	0
0	1	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

$\otimes$

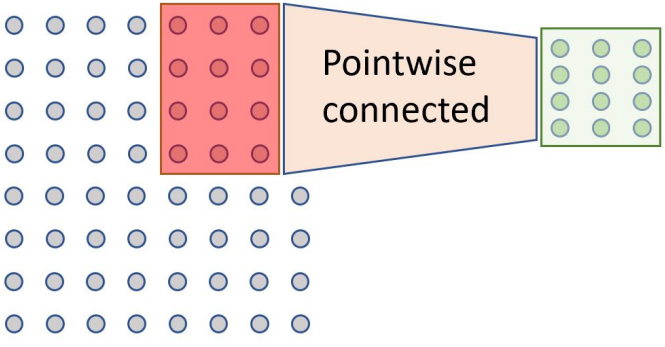
0	0	1
0	0	-1
0	0	-1
0	0	1

$=$

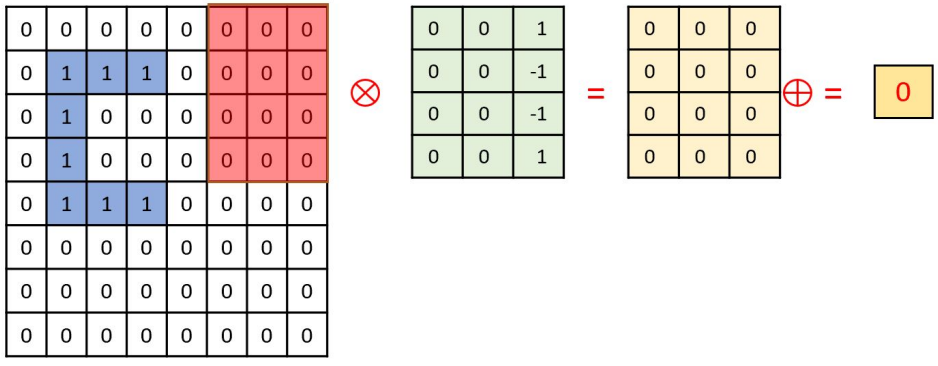
0	0	0
0	0	0
0	0	0
0	0	0

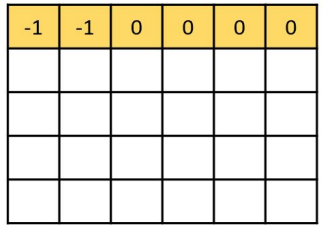
$\oplus =$ 0

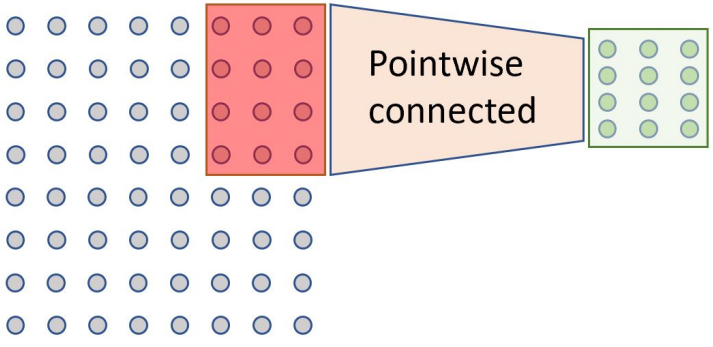
-1	-1	0	0	0	



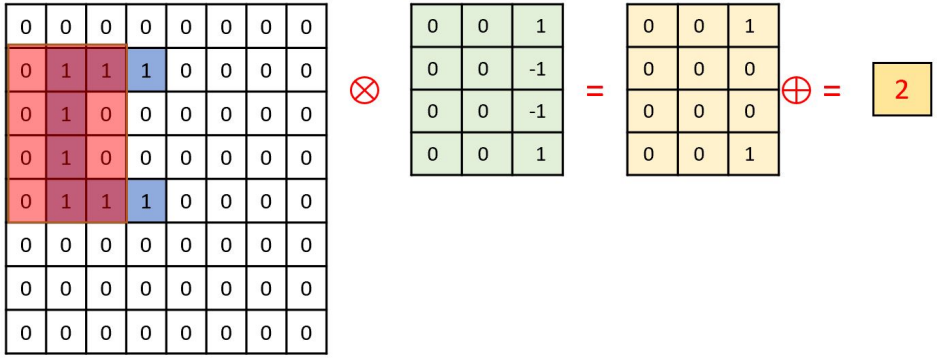
Convolution

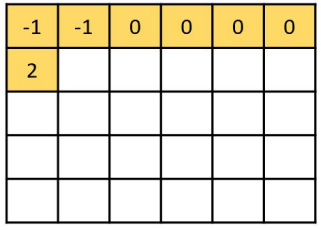


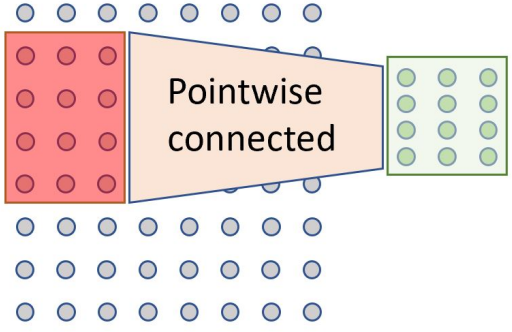




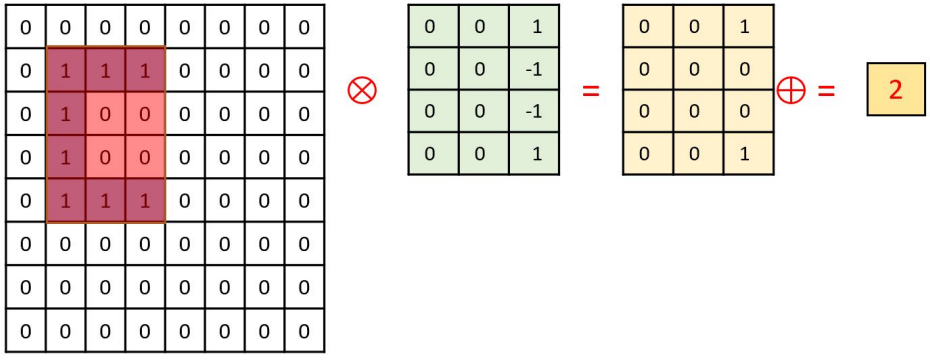
Convolution

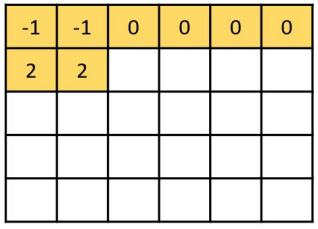


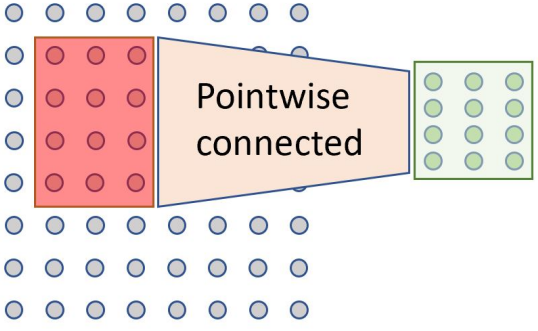




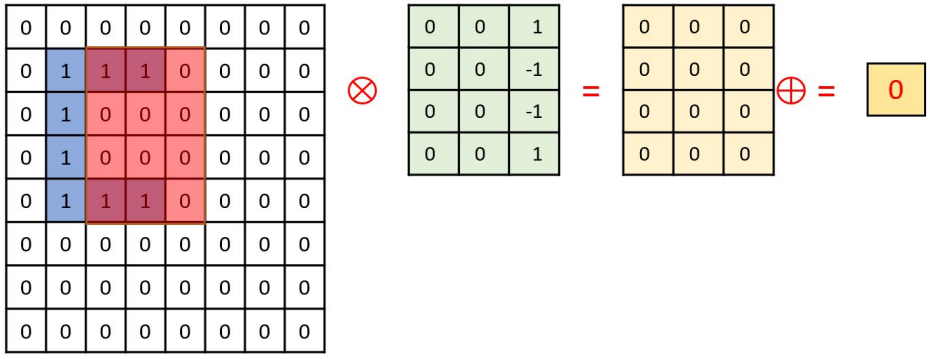
Convolution

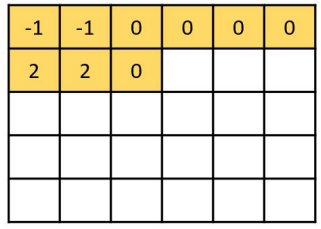


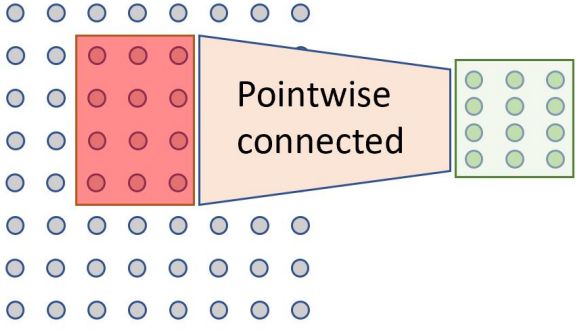




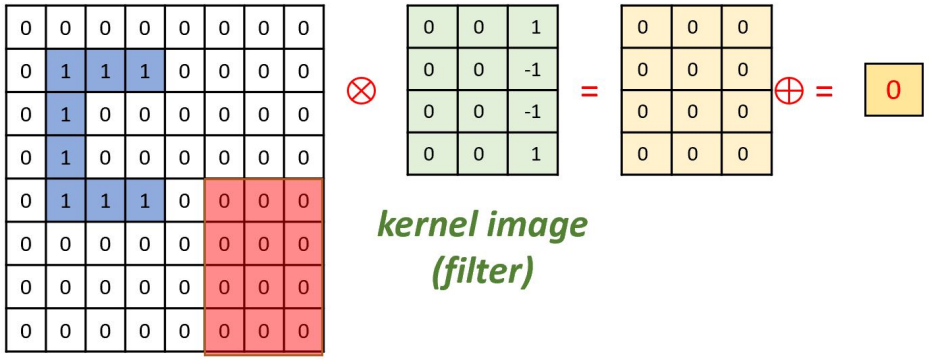
Convolution

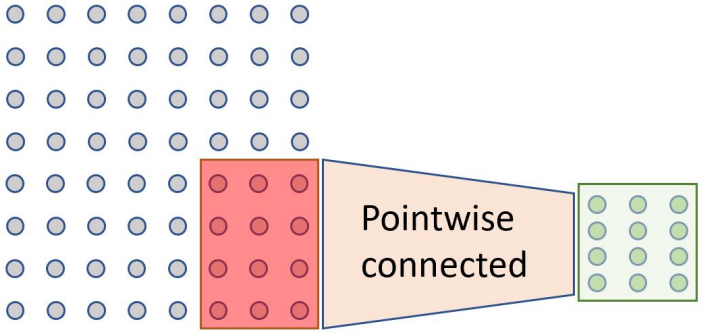




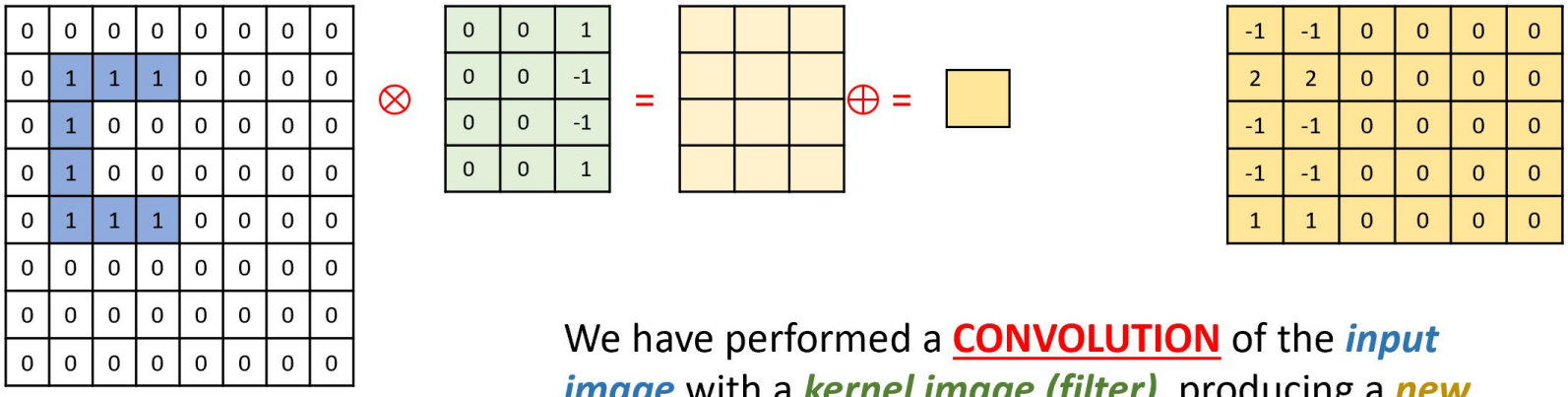


Convolution



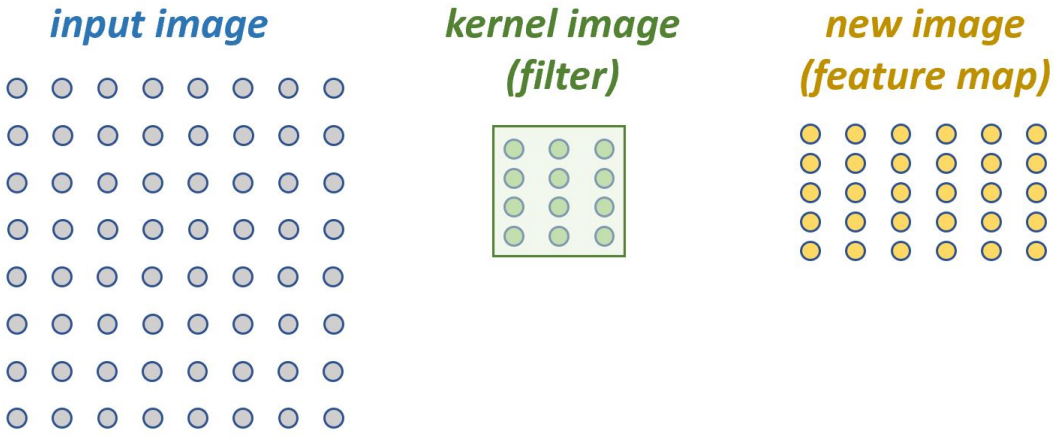


Convolution

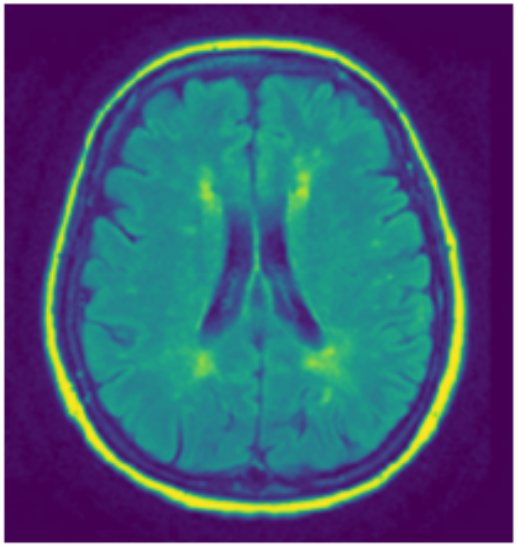


We have performed a **CONVOLUTION** of the *input image* with a *kernel image (filter)*, producing a *new image (feature map)*.

Many important image transformations are convolutions where you *convolve* the image function with a very small "*kernel function*"



Kernels for Edge Detection



Input image:
Brain MRI
1190x1123 pixels

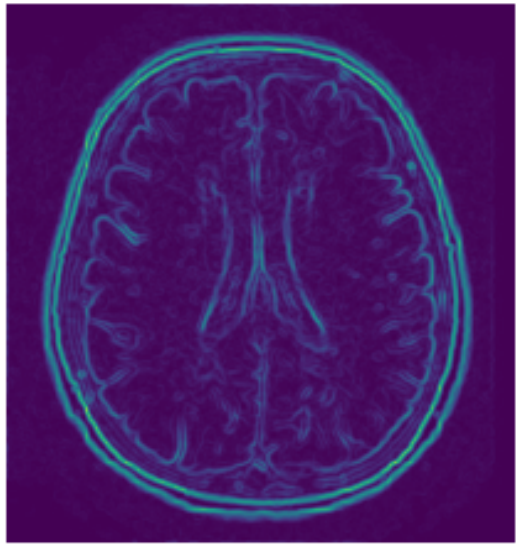
Two kernels/filters

$K_1 =$

-1	-2	-1
0	0	0
+1	+2	+1

$K_2 =$

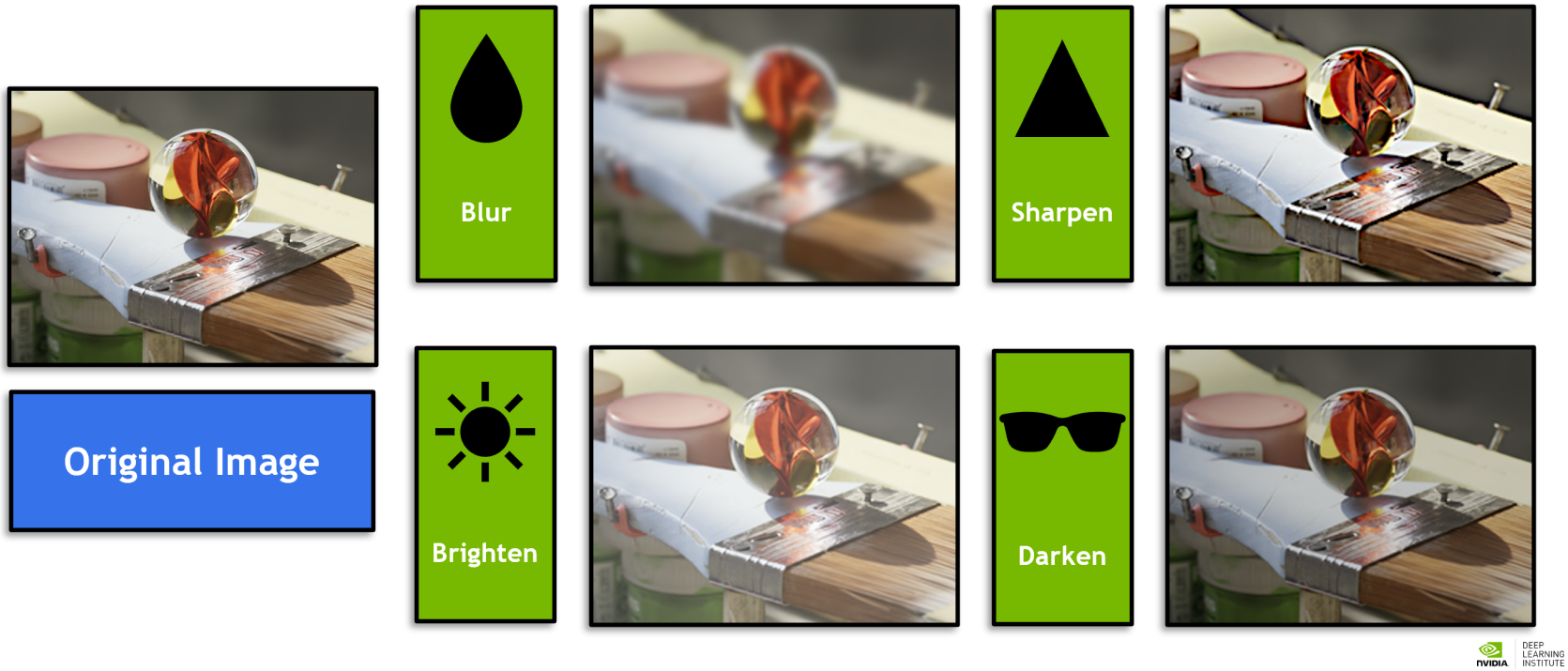
-1	0	+1
-2	0	+2
-1	0	+1



Feature Map:
Edge detection

Common Kernels

KERNELS AND CONVOLUTION



Common Kernels

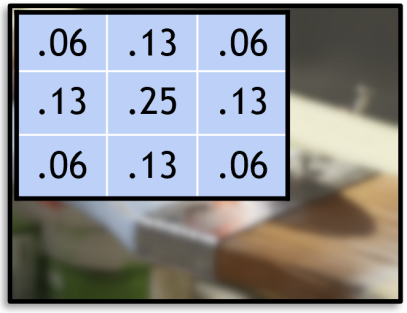
KERNELS AND CONVOLUTION



Original Image

Blur

.06	.13	.06
.13	.25	.13
.06	.13	.06



Sharpen

0	-1	0
-1	5	-1
0	-1	0



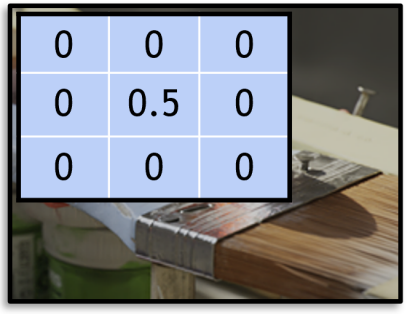
Brighten

0	0	0
0	1.5	0
0	0	0



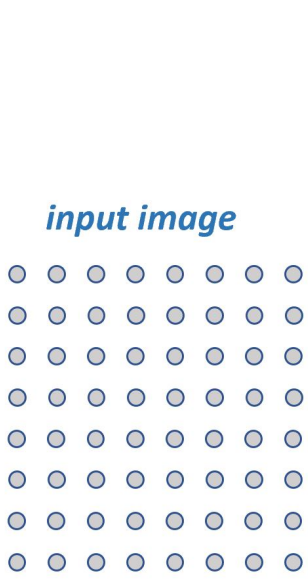
Darken

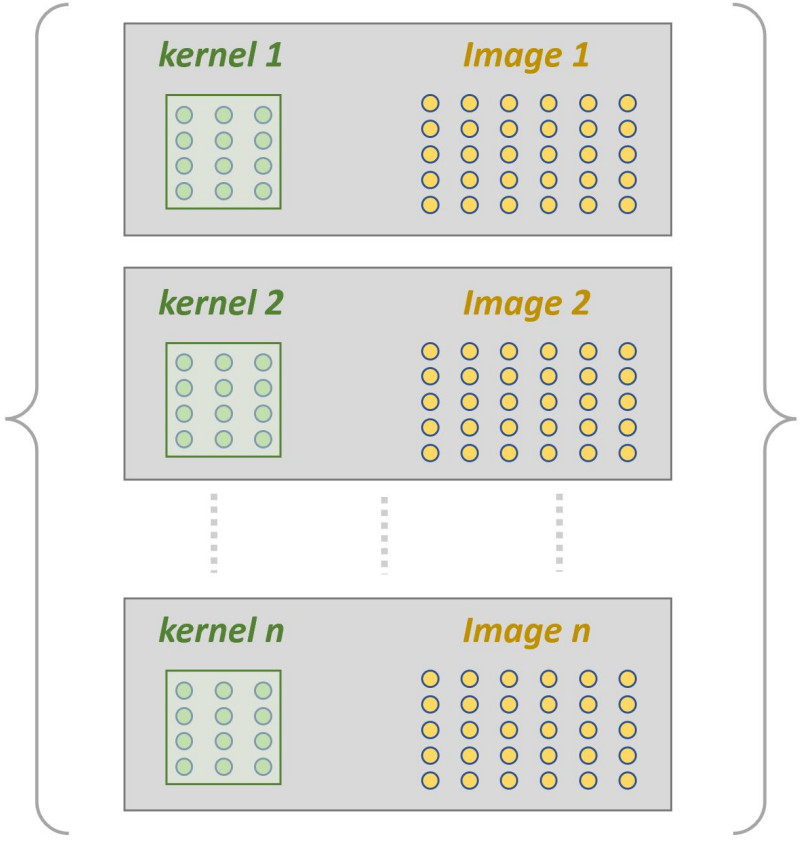
0	0	0
0	0.5	0
0	0	0



NVIDIA DEEP LEARNING INSTITUTE

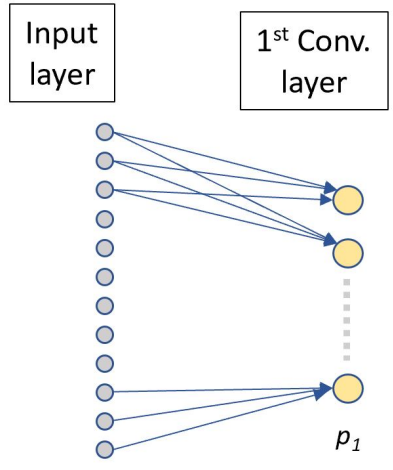
Convolution



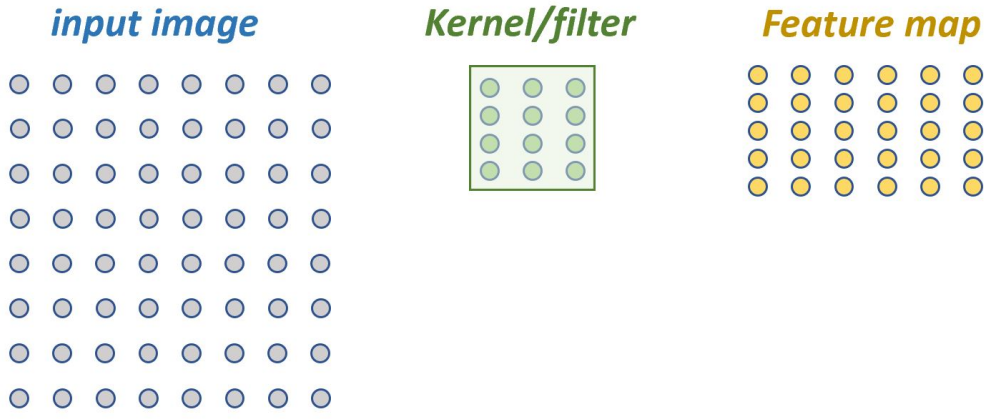


Each ***kernel function/matrix*** may be ***LEARNED*** be since is nothing else than a small matrix of weights (weighted connections of our sparse-ANN)

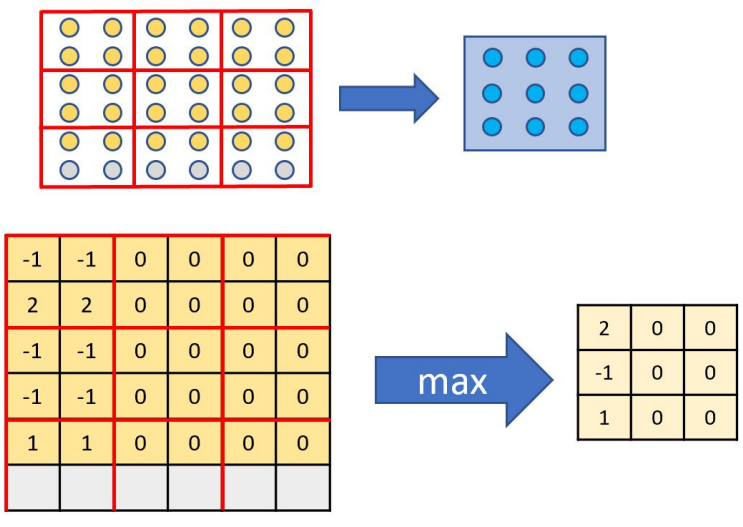
Convolutional Neural Network layer:



Pooling (subsampling)

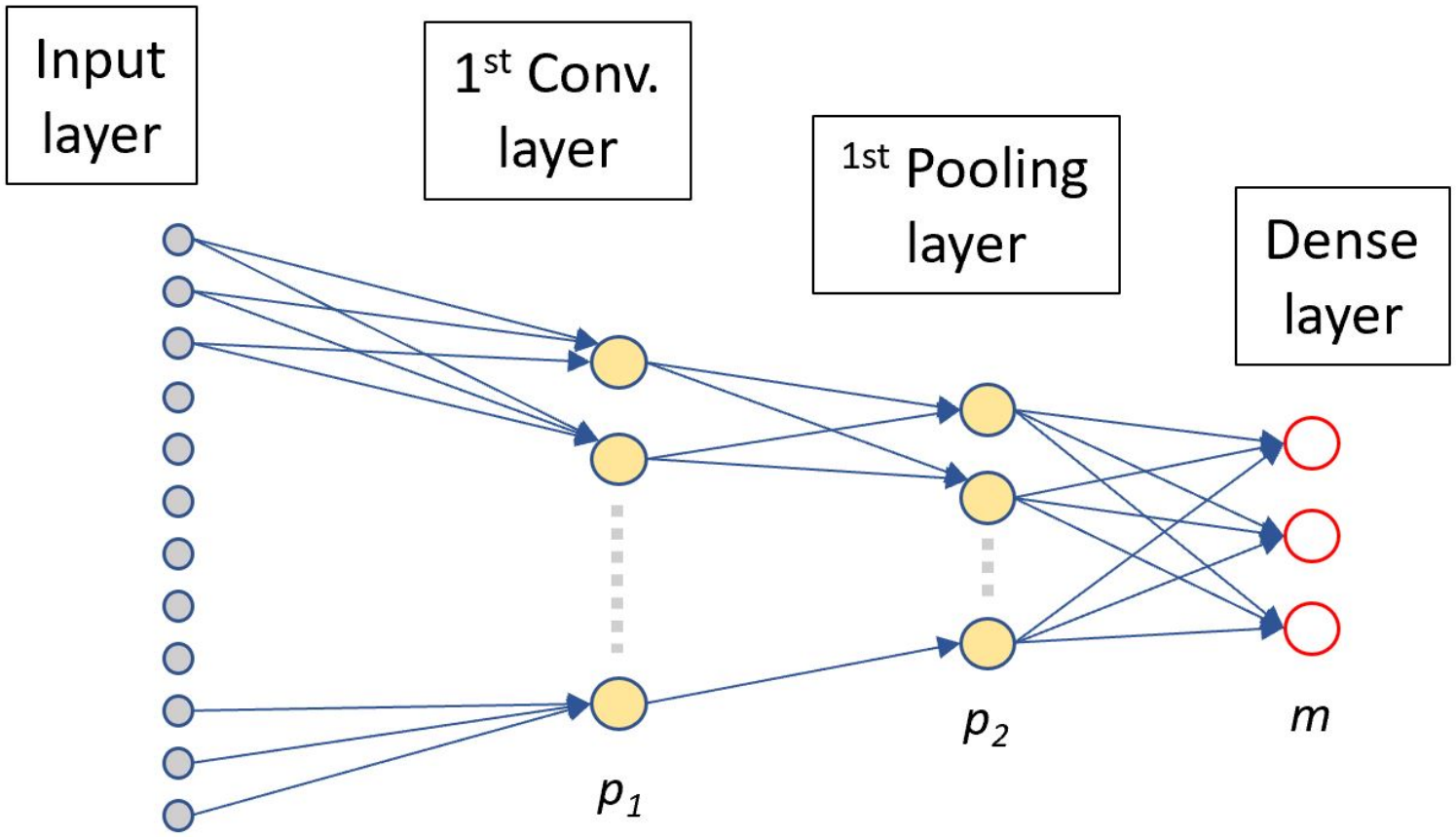


Compress images by ***maximum pooling***
For each submatrix takes the *maximum* value.



Convolutional Neural Network Architecture

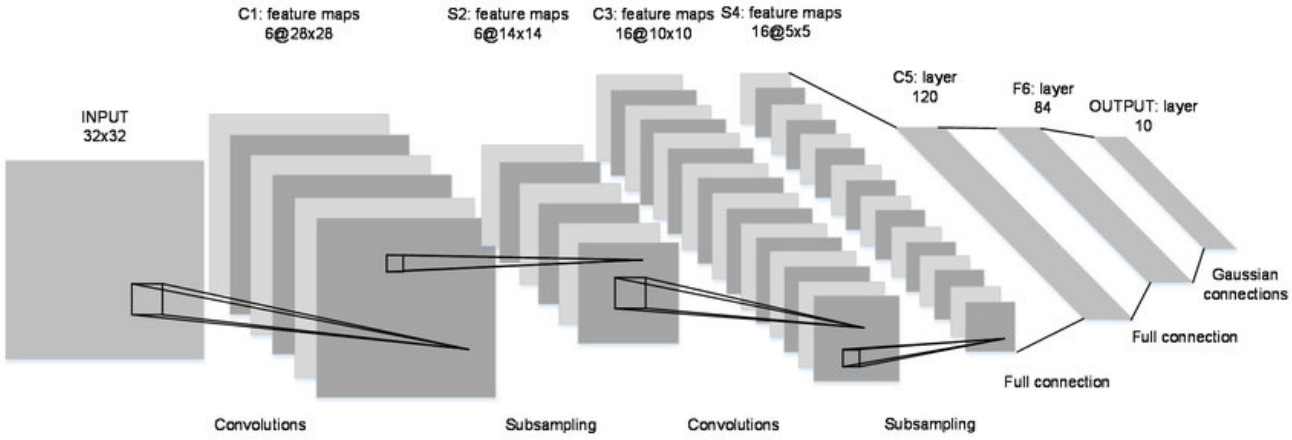
Convolutional Neural Network architecture:



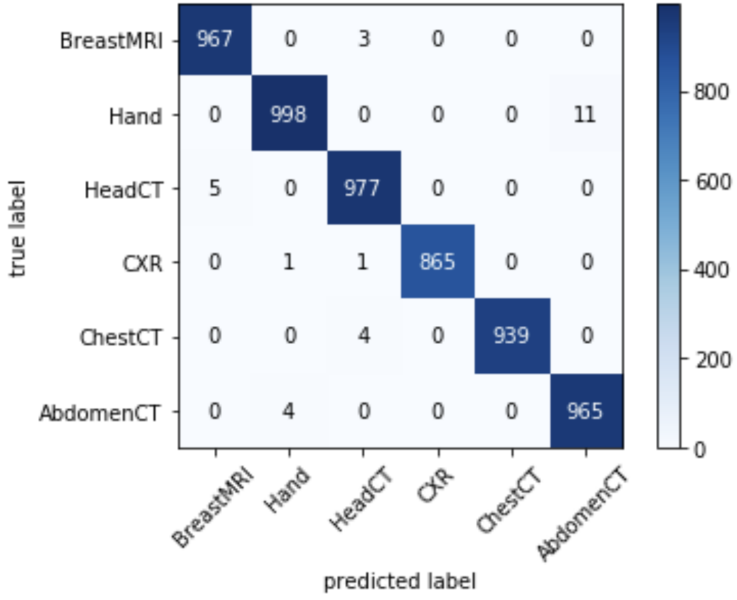
CNN in action

DEMO CLIP (FILE)

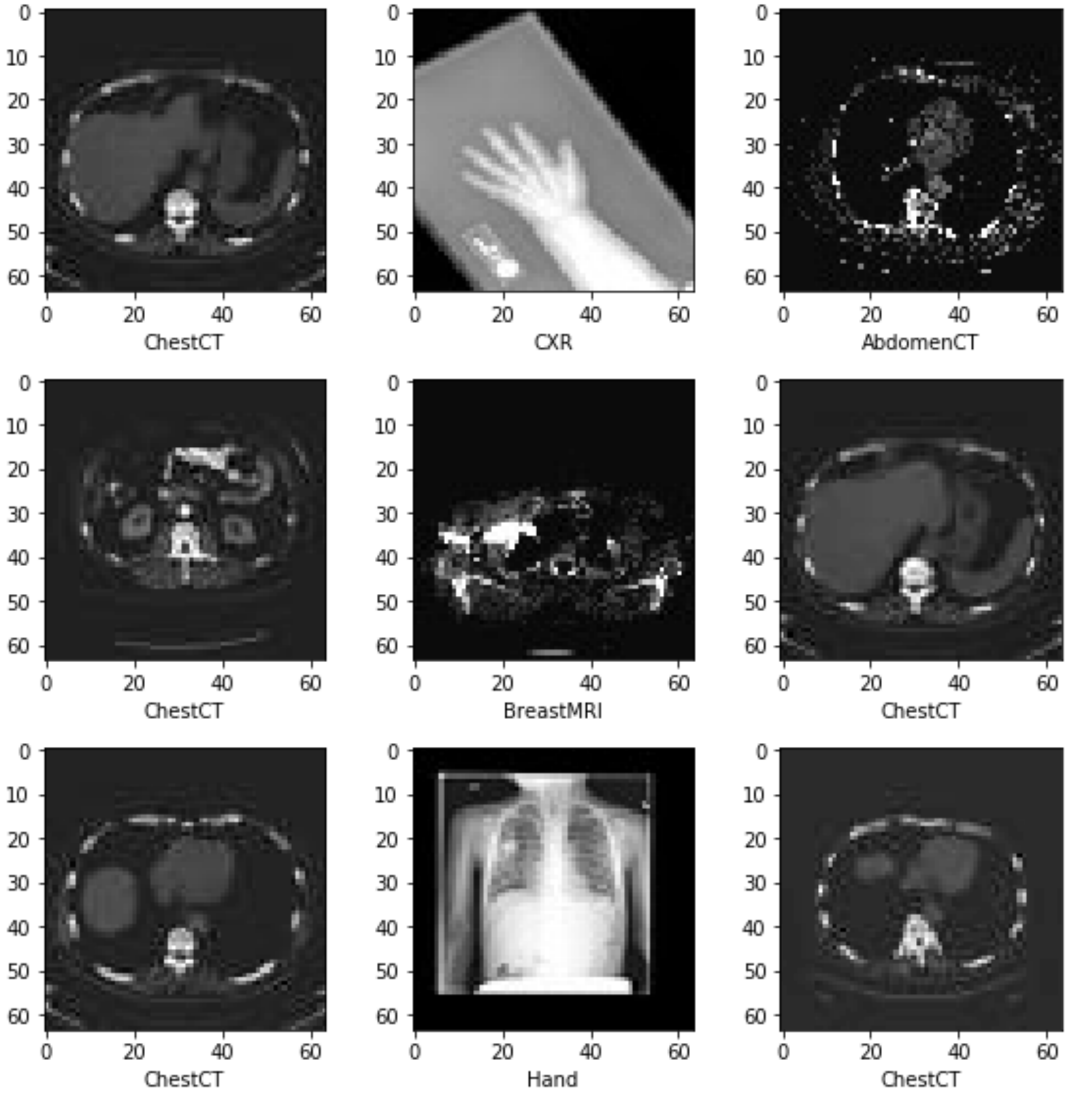
Convolutional Neural Network: MedNIST Classification




```
Accuracy =  0.9949477351916376
```



Convolutional Neural Network: MedNIST Classification



Outline

1 Introduction

2 Images

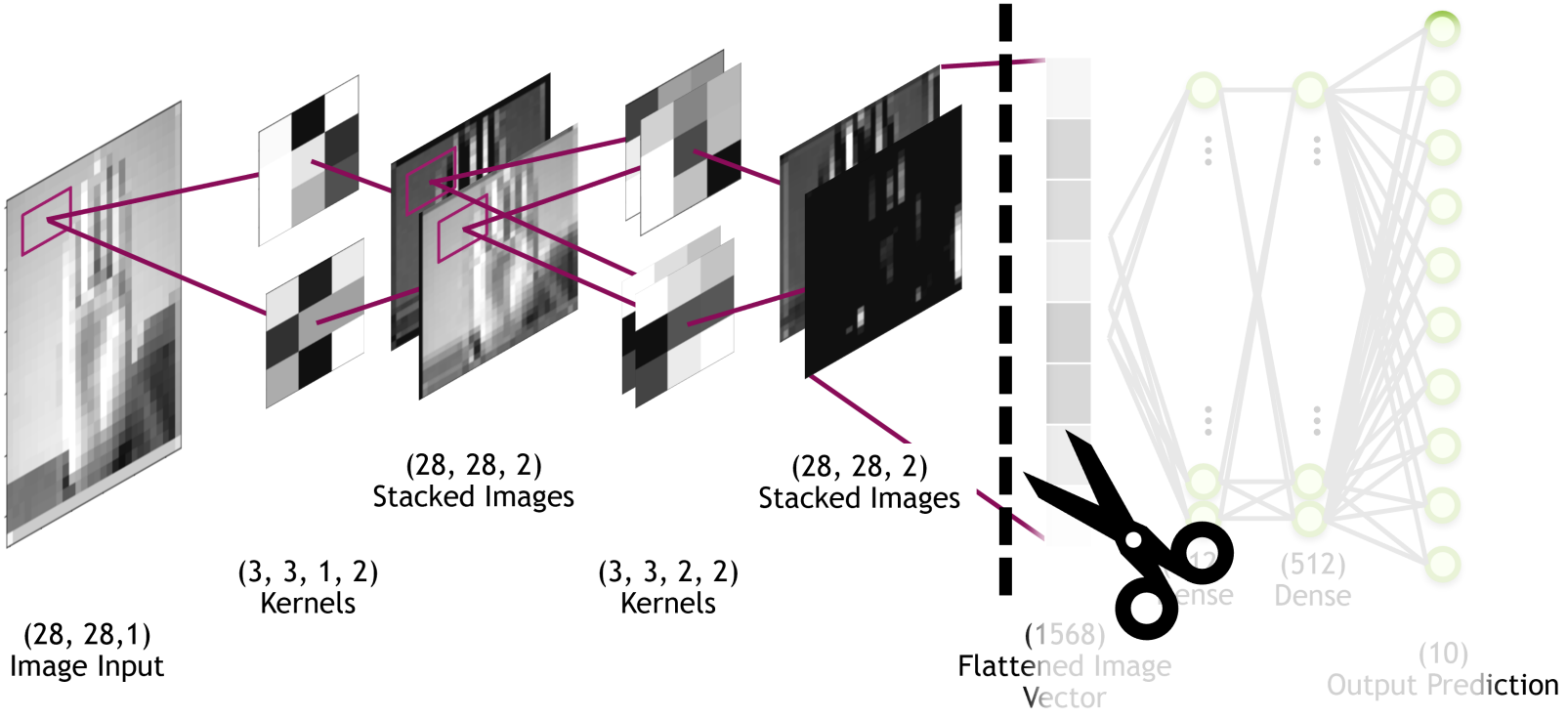
3 Machine Learning

4 Neural Networks

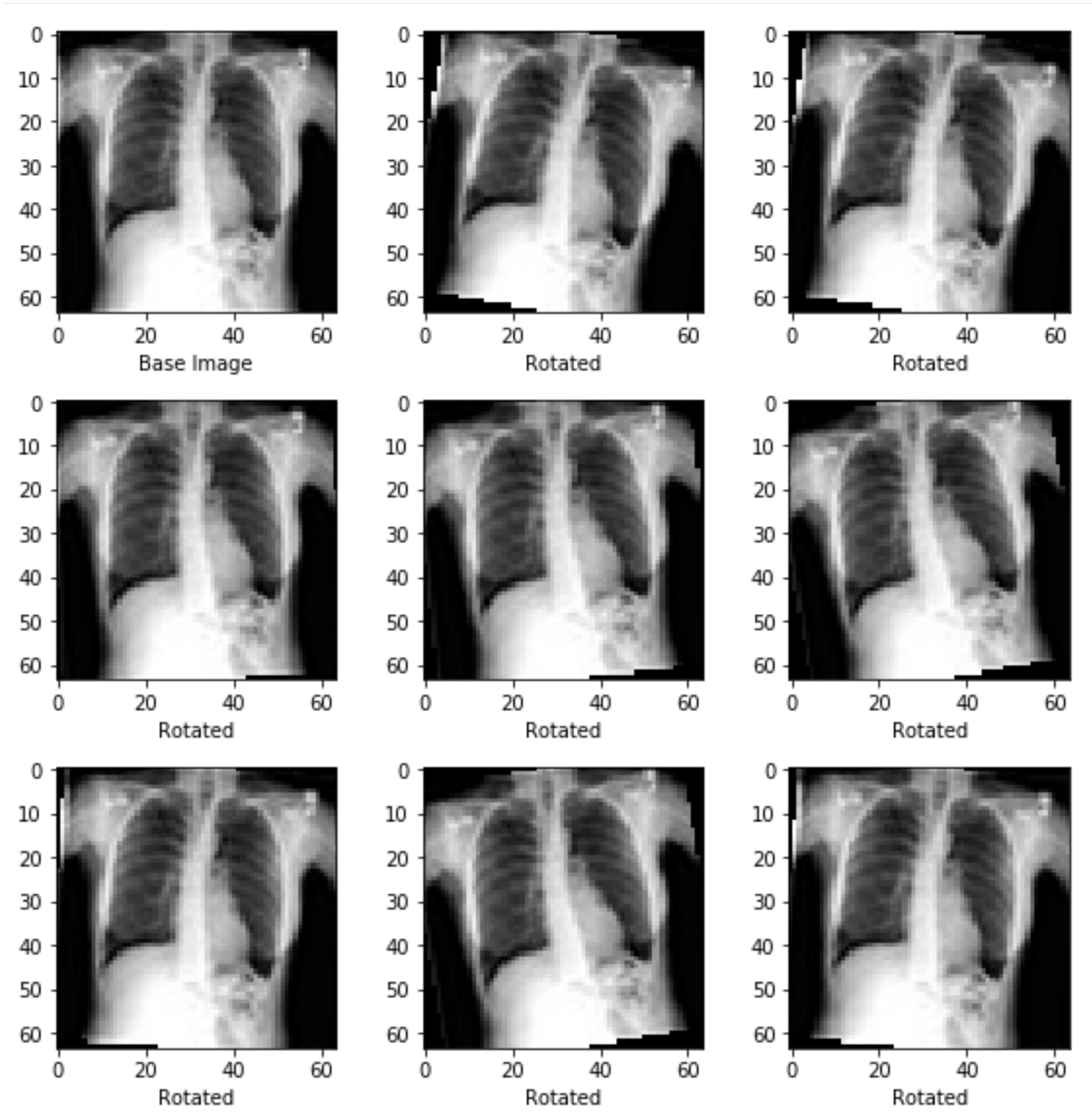
5 Convolutional Neural Networks

6 Transfer Learning

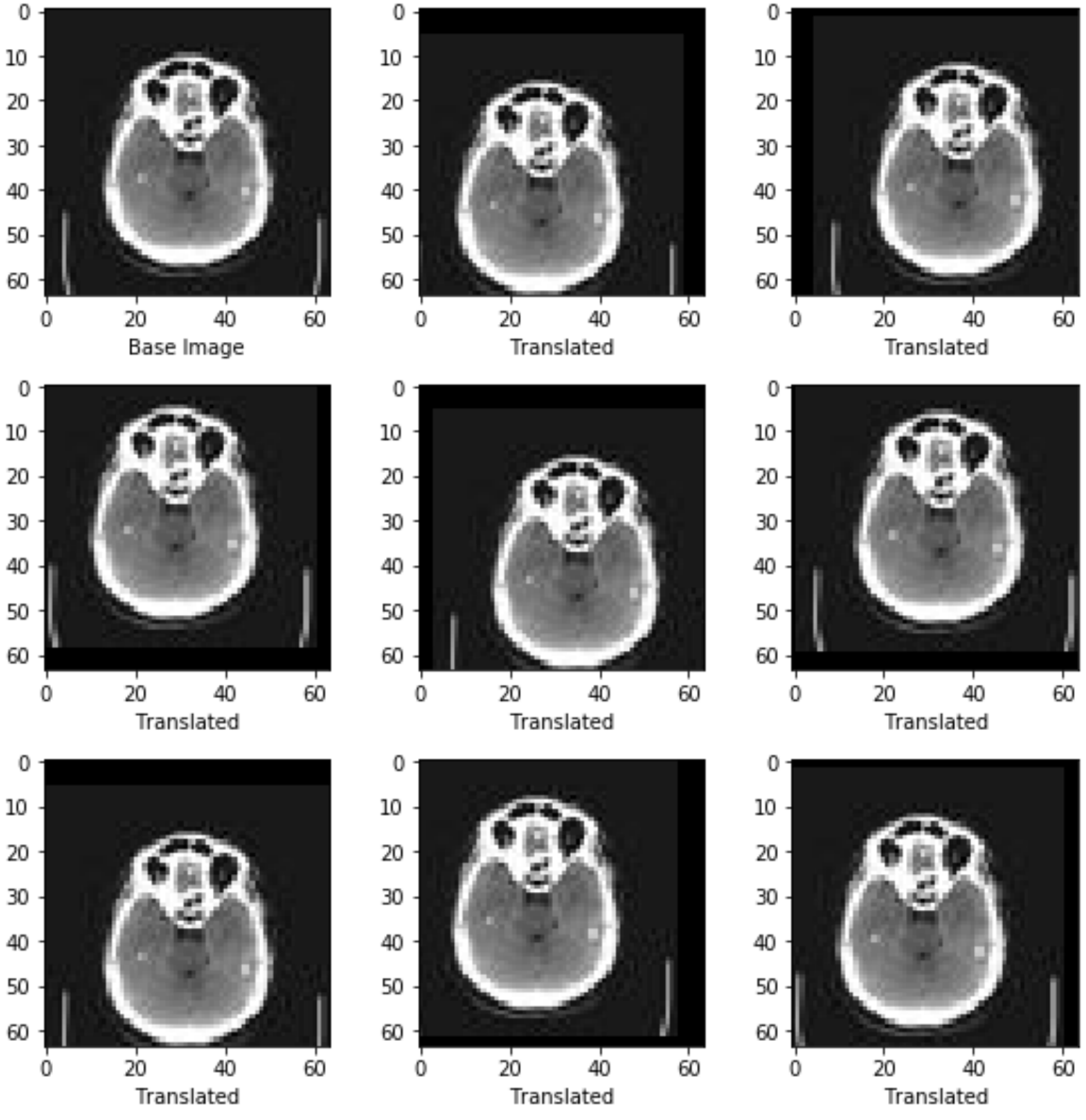
Transfer Learning



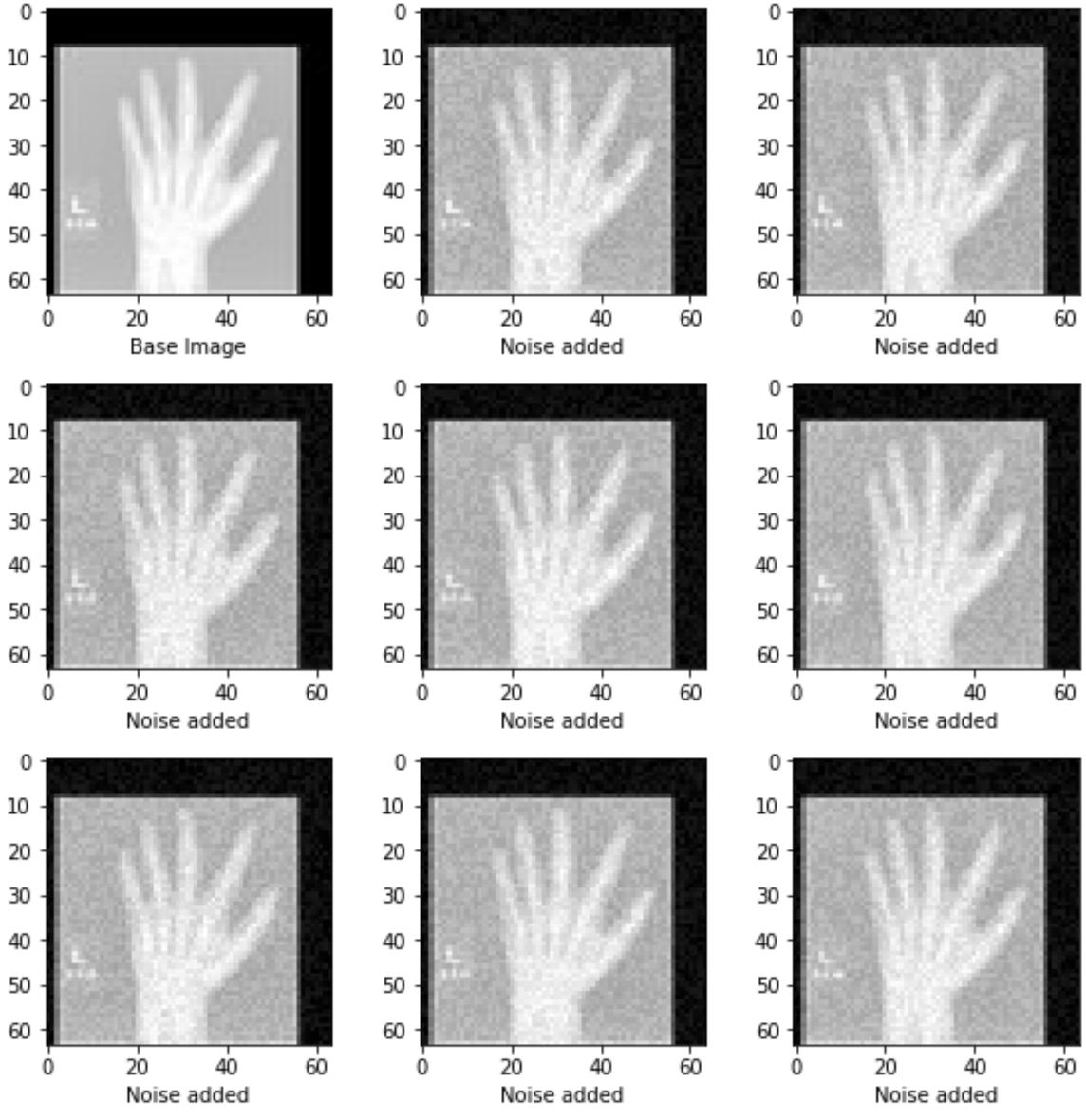
Data Augmentation: Rotations



Data Augmentation: Translation



Data Augmentation: White Noise





Select PayPal as payment method, and apply PROMOCODE:

DLITEACH1222_17_MSYA_44

https://courses.nvidia.com/courses/course-v1:DLI+S-FX-01+V1/

Thanks for your attention!

