Machine Learning in Medicine and Healthcare

João Galego

Siemens Lisbon Tech Hub

Workshop on Biomedical Engineering

14th October 2020

SIEMENS
Ingenuity for life

U LISBOA | UNIVERSIDADE DE LISBOA

IBEB

Workshop on Biomedical Engineering
WBME 12th

Who Am I?

SIEMENS
Ingenuity for life

Academic Background

- MSc Physics
- PgD Forensic Science
- PhD Cognitive Science
  - Brain-Computer Interfaces
  - Machine Learning

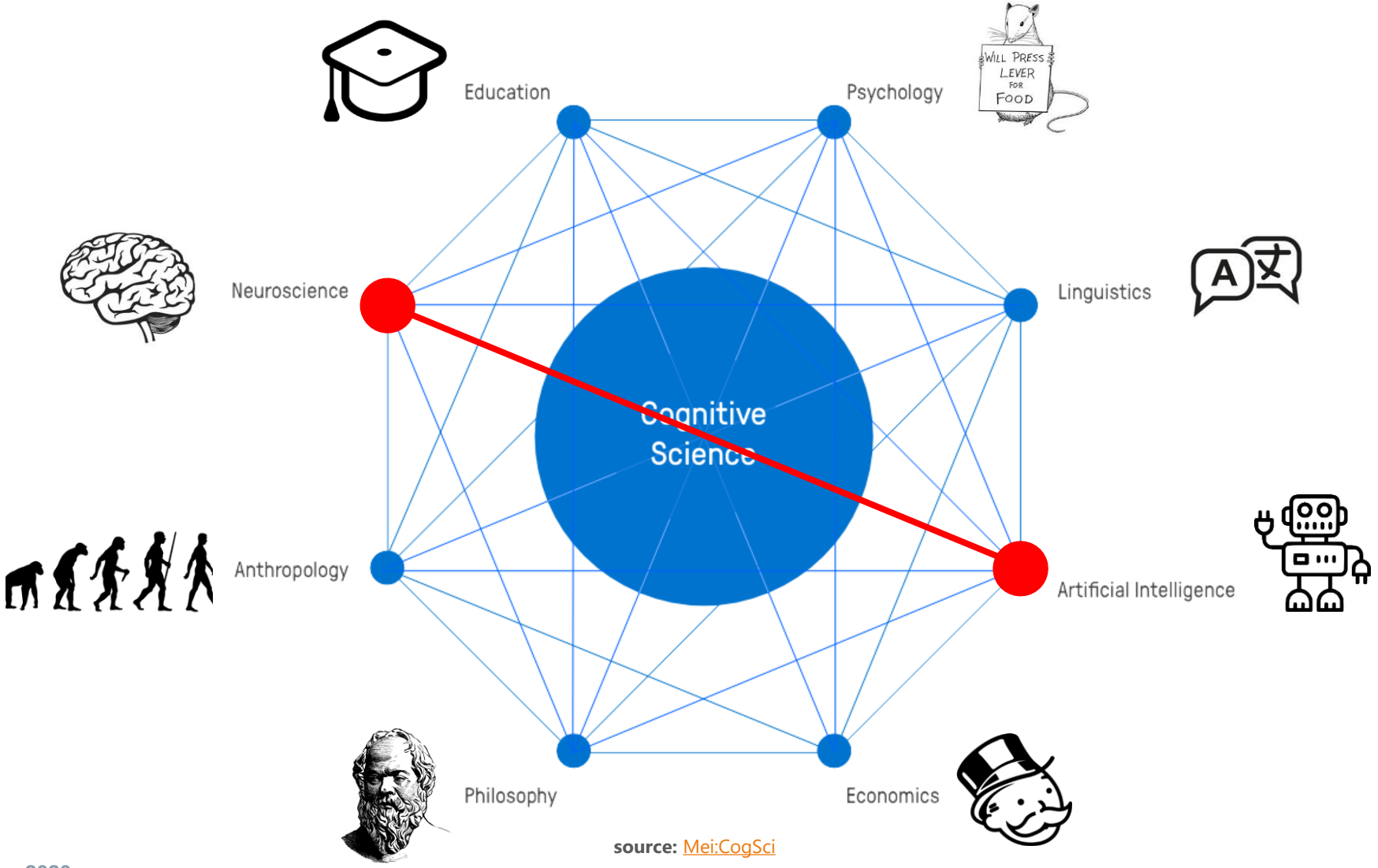
Professional Experience

- Data Analyst
- Test Automation Engineer
- DevOps Engineer
- ML Consultant

source: deepdreamgenerator

Unrestricted © Siemens 2020

Cognitive Science



source: Mei:CogSci

Unrestricted © Siemens 2020

Brain-Computer Interfaces

SIEMENS
Ingenuity for life

source: NBC Learn

Unrestricted © Siemens 2020

Advisors

SIEMENS
Ingenuity for life

Prof. Hugo Ferreira

Research Areas:

- Brain Connectivity
- Neuropsychiatric Disorders
- Neuromarketing
- Brain Stimulation
- Brain-Computer Interfaces

Prof. Arlindo Oliveira

Research Areas:

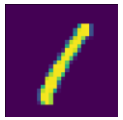
- Algorithms and Complexity
- Machine Learning
- Bioinformatics
- Digital Circuit Design

Unrestricted © Siemens 2020

Administrivia

SIEMENS
Ingenuity for life

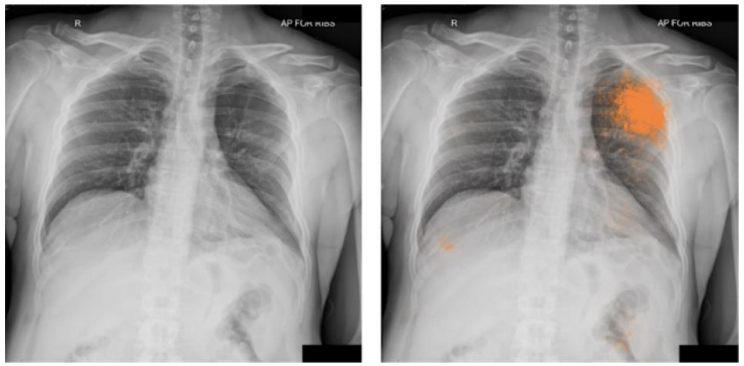
Agenda



Overview of AI and ML

ML in Medicine and Healthcare

Opportunities, Challenges and Future Directions

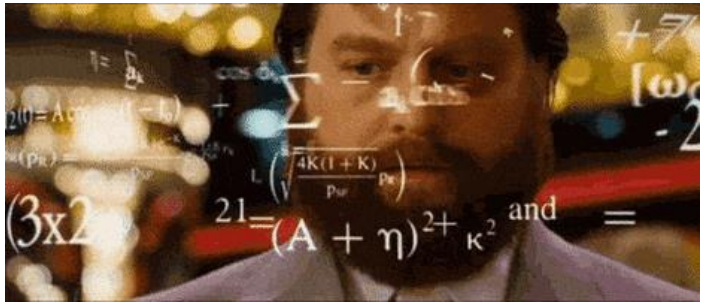


source: Google AI blog

Prerequisites

No knowledge of programming languages

No knowledge of algebra, statistics or calculus



For more information, please visit https://github.com/JGalego/WBME-ML-Workshop

SIEMENS
Ingenuity for life

Overview of AI and ML

Unrestricted © Siemens 2020

Artificial Intelligence

Historical Perspective

SIEMENS
Ingenuity for life

1943 — McCulloch and Pitts publish *'A Logical Calculus of the Ideas Immanent in Nervous Activity'* laying the foundations for artificial neural networks

1950 — Turing publishes *'Computing Machinery and Intelligence'*

1956 — **Dartmouth Conference** — The Birth of AI

1980-7 — The rise of expert systems

1997 — IBM Deep Blue beats world chess champion Garry Kasparov

2011 — IBM Watson wins 1st prize on *Jeopardy!*

2012 — **Deep Learning Revolution** — AlexNet wins the ImageNet challenge by a large margin

2016 — AlphaGo beats the #1 ranked Go player in the world

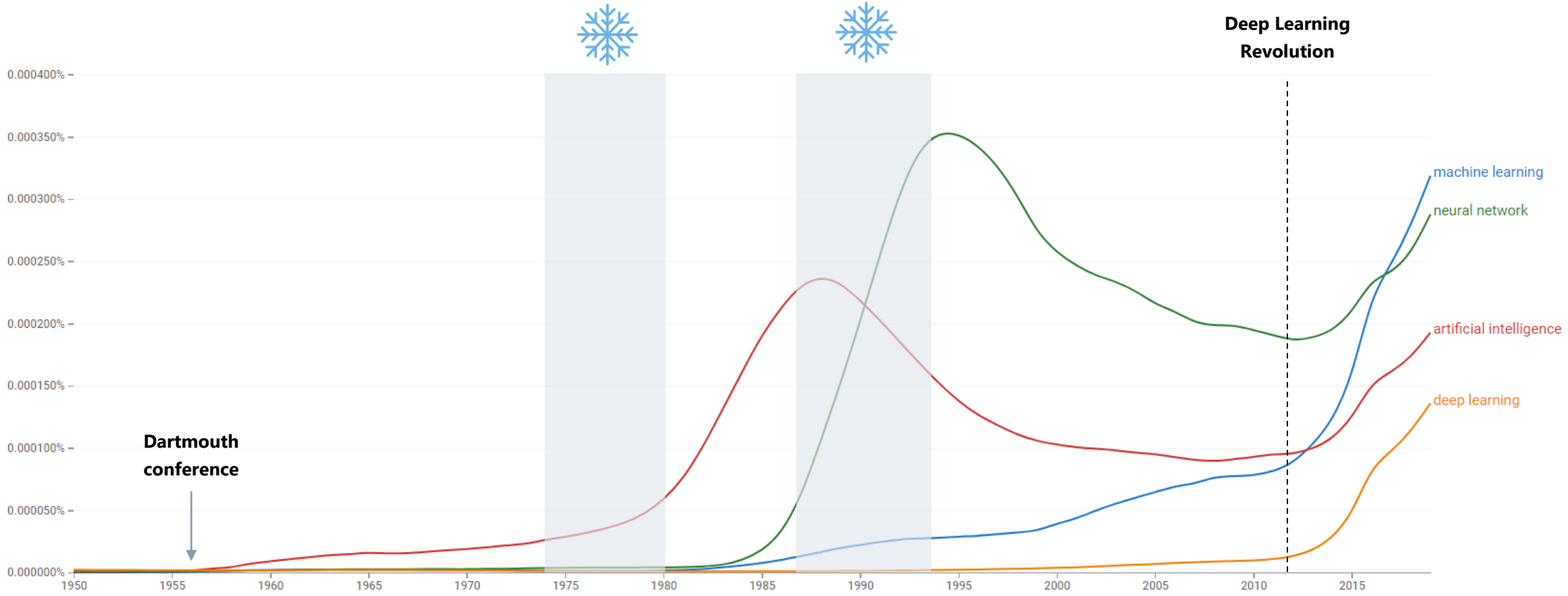
"I like the dreams of the future better than the history of the past."

- Thomas Jefferson

Unrestricted © Siemens 2020

Artificial Intelligence

Historical Perspective



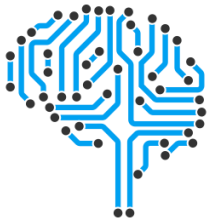
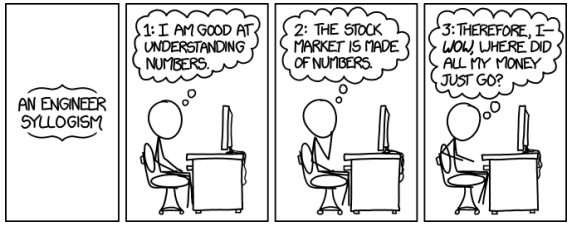
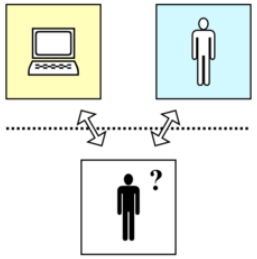
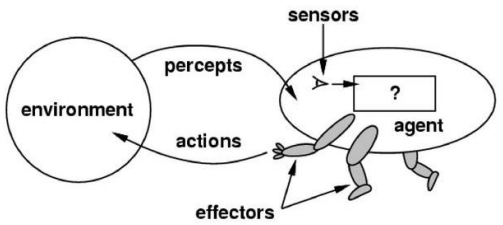
source: Google Books Ngram Viewer

Unrestricted © Siemens 2020

What is AI?

Matrix Definition

SIEMENS
Ingenuity for life

	Humanly	Rationally
Thinking	 ***The cognitive modelling approach*** "The exciting new effort to make computers think… *machines with minds*, in the full and literal sense."	 ***The "laws of thought" approach*** "The study of the computations that make it possible to perceive, reason, and act."
Acting	 ***The Turing test approach*** "The art of creating machines that perform functions that require intelligence when performed by people."	 ***The rational agent approach*** "AI… is concerned with intelligent behavior in artifacts"

source: Russel & Norvig (2020)

Unrestricted © Siemens 2020

What is AI?

Basic Definition

SIEMENS
Ingenuity for life

source: Google

"AI is the science of making machines smart."

Demis Hassabis, co-founder of Google DeepMind

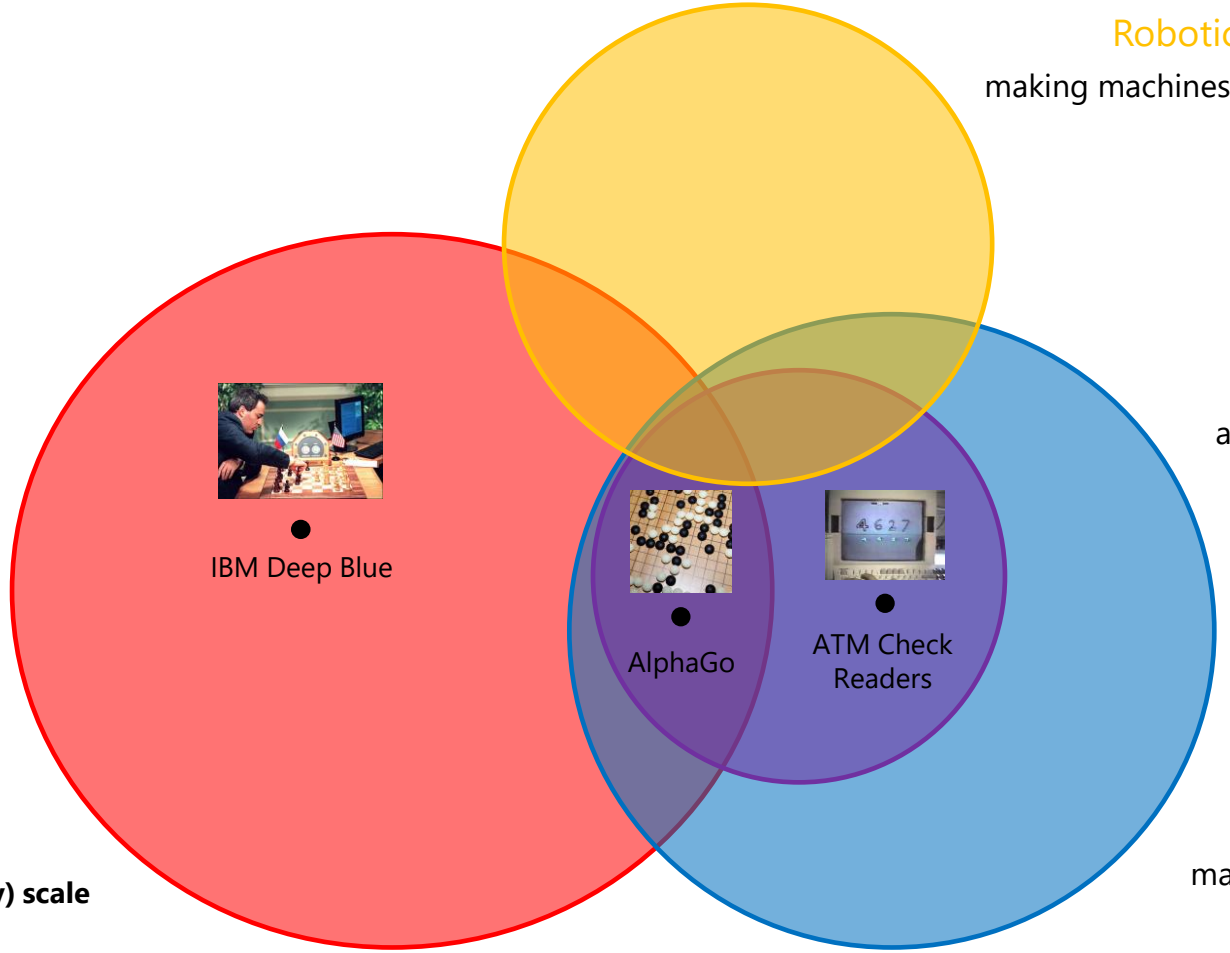
source: Google Zeitgeist

Unrestricted © Siemens 2020

AI and ML

Interdisciplinary Exchange

SIEMENS
Ingenuity for life



(*) fields not drawn to (any) scale

source: *adapted from* Corrado (2018)

Unrestricted © Siemens 2020

What is ML?

SIEMENS
Ingenuity for life

ML and Statistics

Tibshirani's Glossary

Machine learning	Statistics
network, graphs	model
weights	parameters
learning	fitting
generalization	test set performance
supervised learning	regression/classification
unsupervised learning	density estimation, clustering
large grant = $1,000,000	large grant= $50,000
nice place to have a meeting: Snowbird, Utah, French Alps	nice place to have a meeting: Las Vegas in August

source: Tibshirani – Statistics 315A

ML and Linear Algebra

source: xkcd

Unrestricted © Siemens 2020

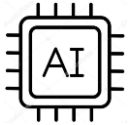
What is ML?

SIEMENS
Ingenuity for life

Docker image

DevOps

Artificial Intelligence



Machine Learning

Backpropagation

Blockchain

yoga

X is just glorified Y

glorified
Collins COBUILD

Word Frequency

(glɔːrɪfaɪd)
ADJECTIVE [ADJECTIVE noun]

You use glorified to indicate that something is less important or impressive than its name suggests.

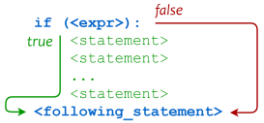
Sometimes they tell me I'm just a glorified waitress.
Some places were no more than glorified pubs.

COBUILD Advanced English Dictionary. Copyright © HarperCollins Publishers

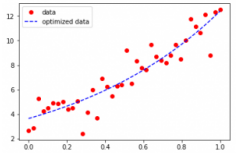
tarball

SysAdmin

IF-THEN-ELSE rules



Statistics
Linear Algebra
Curve Fitting



Chain rule

$$\frac{d}{dx}[f(g(x))] = f'(g(x))g'(x)$$

linked lists

stretching

Unrestricted © Siemens 2020

Intermezzo

Logic Programming

[glorified.pl]

```
% This program is just glorified Prolog programming
%
% "You use glorified to indicate that something is less important
% or impressive than its name suggests." (Collins dictionary)
%
% @author João Galego <joao.galego@siemens.com>
%

% Facts (according to the Internet)
glorified('Docker image', tarball).
glorified('DevOps', 'SysAdmin').
glorified('Artificial Intelligence', 'IF-THEN-ELSE rules').
glorified('Machine Learning', 'Statistics').
glorified('Machine Learning', 'Linear Algebra').
glorified('Machine Learning', 'curve fitting').
glorified('Backpropagation', 'chain rule').
glorified('Blockchain', 'linked lists').
glorified('Physics', 'Mathematics').
glorified('Chemistry', 'Physics').
glorified('Mathematics', counting).
glorified(yoga, stretching).

% Let's make the 'glorified' relation transitive
glorified(X, Y) :- glorified(X,Z),
                   glorified(Z,Y).
```

```
Welcome to SWI-Prolog (threaded, 64 bits, version 7.6.4)
SWI-Prolog comes with ABSOLUTELY NO WARRANTY. This is free software.
Please run ?- license. for legal details.

For online help and background, visit http://www.swi-prolog.org
For built-in help, use ?- help(Topic). or ?- apropos(Word).

?- [glorified].
true.

?- glorified(X, 'Statistics').
X = 'Machine Learning' .

?- glorified('Physics', Y).
Y = 'Mathematics' ;
Y = counting .
```

Extra Credit: *give Prolog a chance* - read *The Power of Prolog's* chapter on AI

Unrestricted © Siemens 2020

What is ML?

Some Definitions

SIEMENS
Ingenuity for life

Mitchell (1997):

"a computer program is said to learn if its performance at a task T*, as measured by a performance* P*, improves with experience* E*"*

Bishop (2006):

"automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions"

Goodfellow, Bengio & Courville (2016):

"the ability to acquire their [AI systems] own knowledge, by extracting patterns from raw data"

Mohri, Rostamizadeh & Talwalkar (2018):

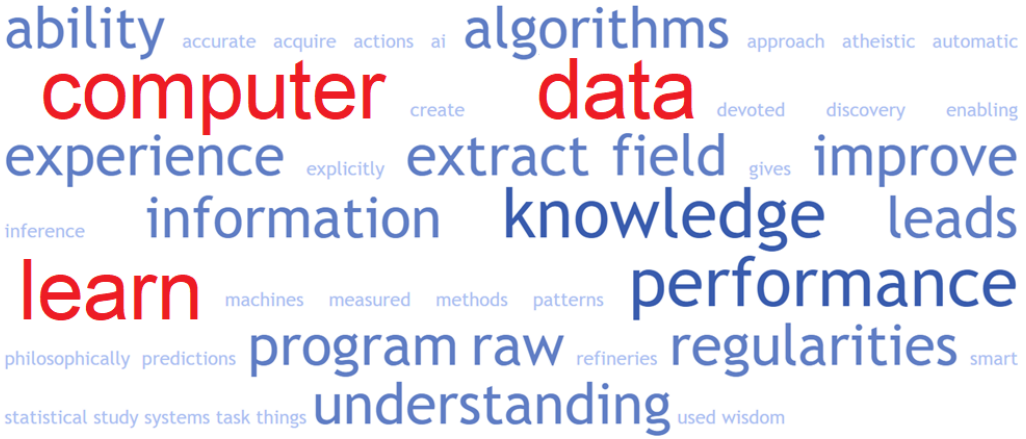
"computational methods using experience to improve performance or to make accurate predictions"

Gerrish (2018):

"field devoted to enabling machines to do smart things by learning from data"

Efron & Hastie (2019):

"a philosophically atheistic approach to statistical inference"



source: TagCrowd

Unrestricted © Siemens 2020

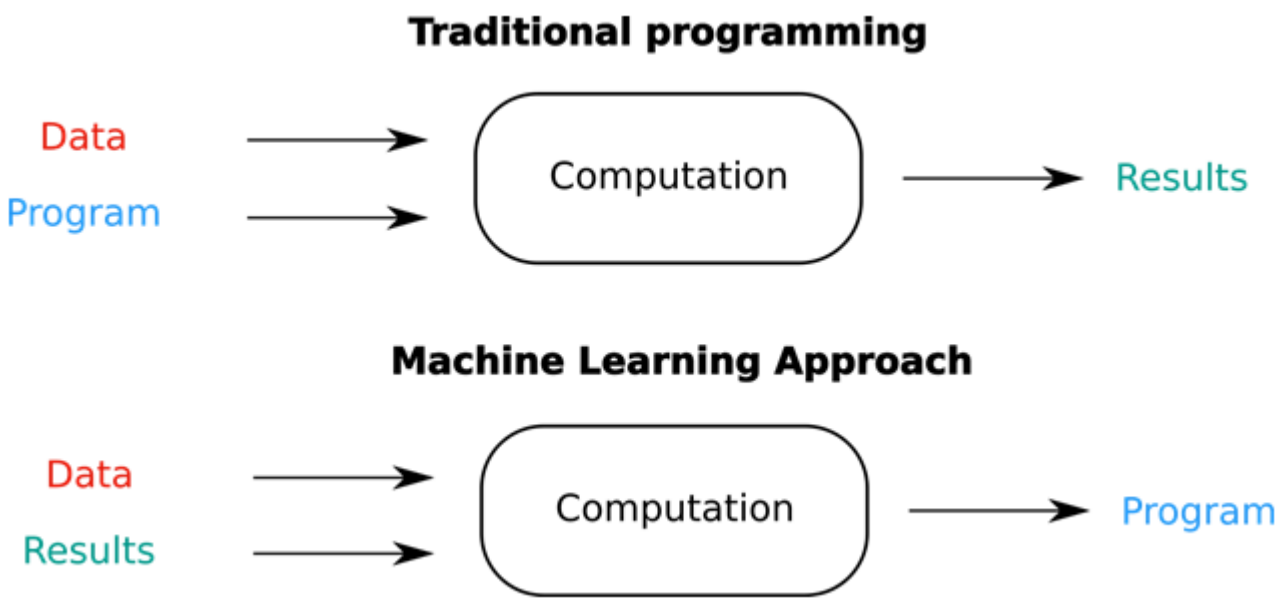
What is ML?

ML as a paradigm shift

SIEMENS
Ingenuity for life

"Field of study that gives computers the ability to learn without being explicitly programmed."

source: *adapted from* Samuel (1959)



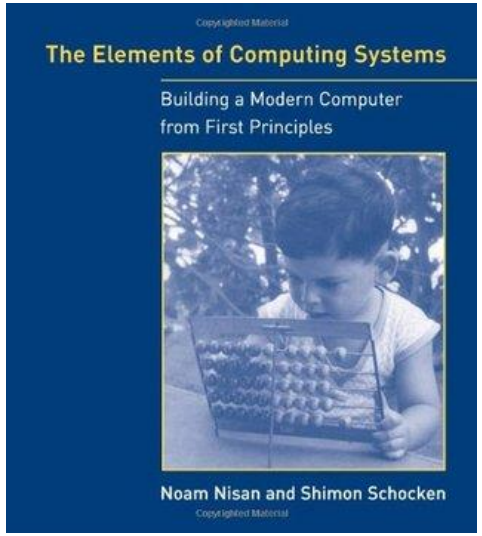
source: Zero to Singularity

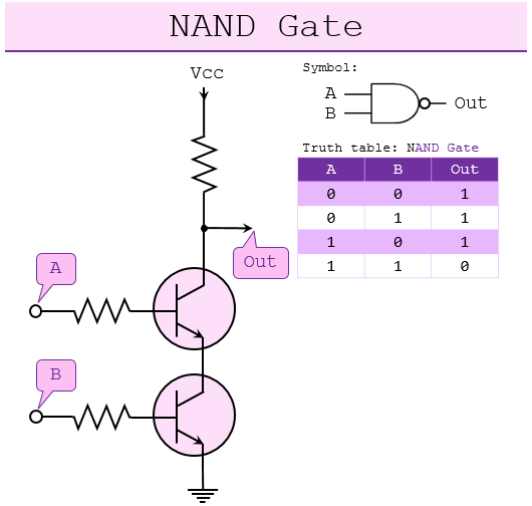
source: Google Arts & Culture

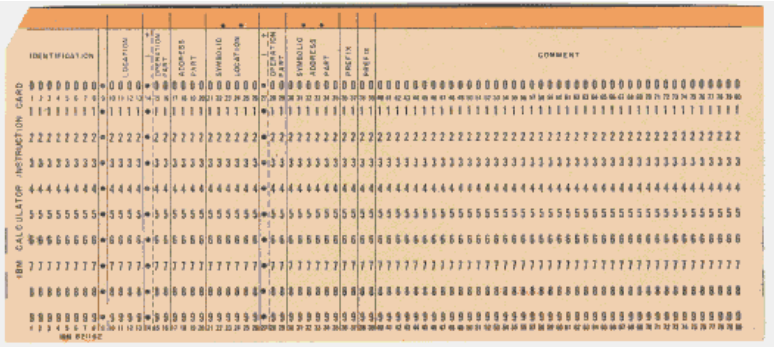
Unrestricted © Siemens 2020

Intermezzo

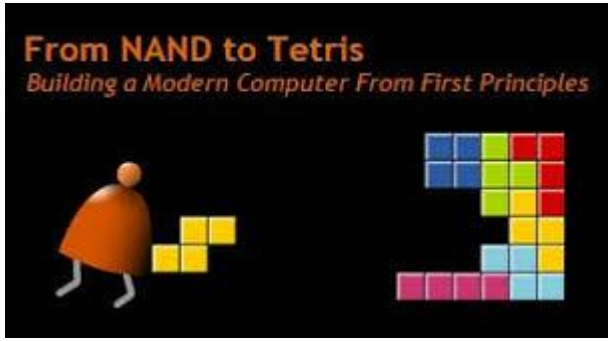
On Computers



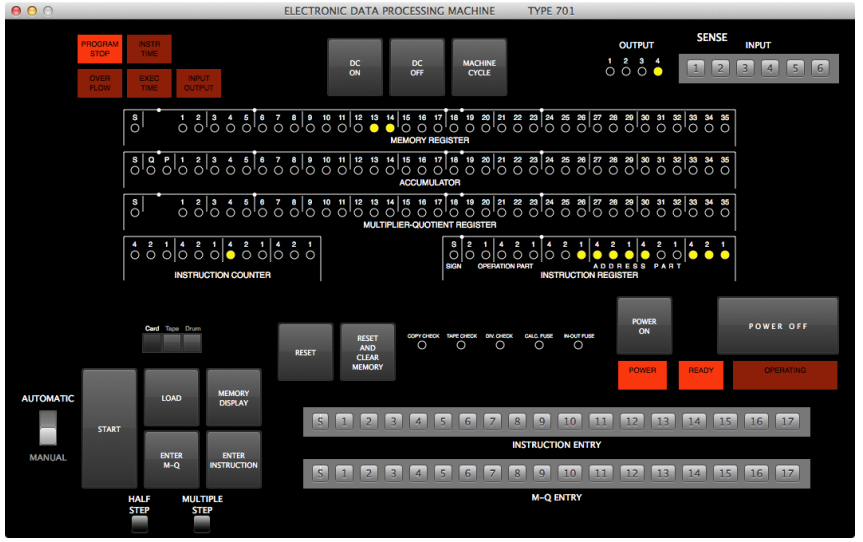




source: Douglas Jones



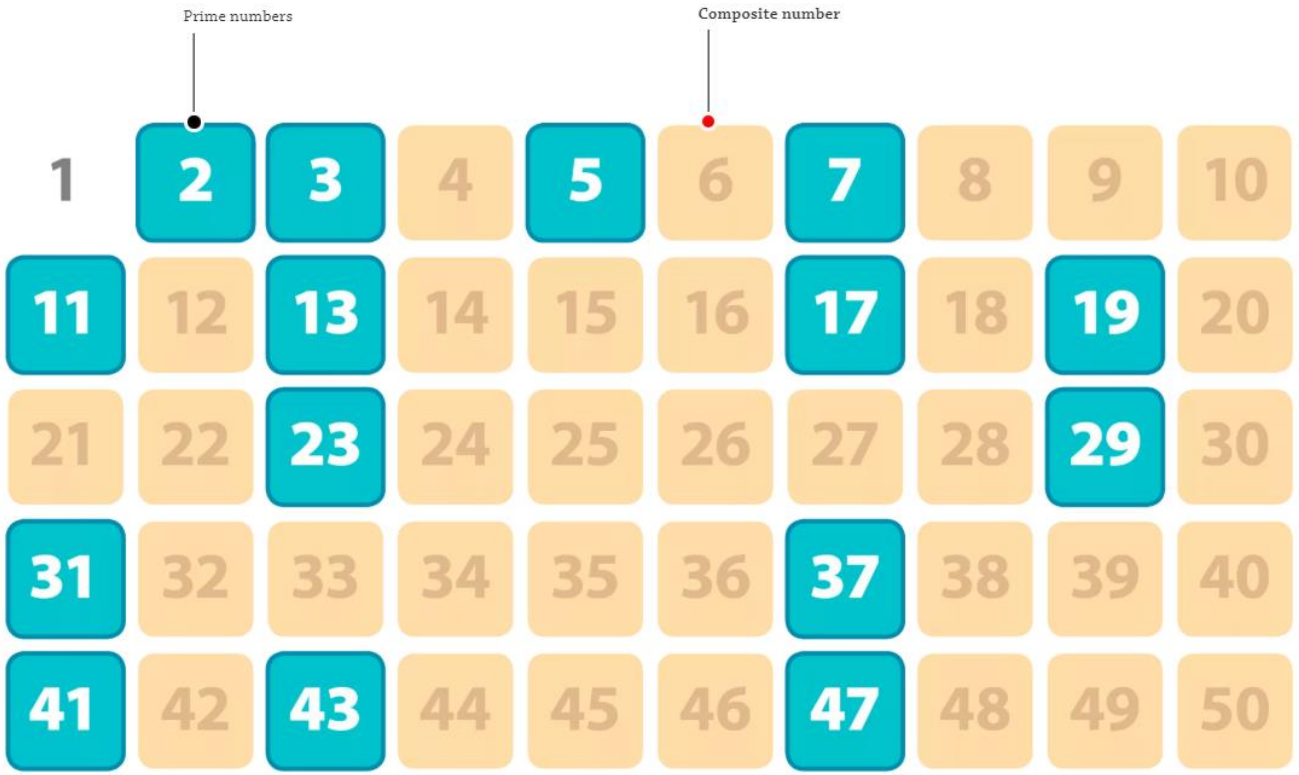
source: nand2tetris



source: Marc Rochkind's IBM 701 Emulator

Extra Credit: check out the NAND2Tetris course and/or the IBM 701 emulator

Prime or Composite?




```
#include <iostream>
#include <math.h>

using namespace std;

int main()
{
    int i, num;

    cout << "Enter a number:";
    cin >> num ; //Taking input

    for(i = 2 ; i < num -1 ; i++)
    {
        if(num % i == 0)
        {
            cout << "COMPOSITE";
            exit(0);
        }
    }

    cout << "PRIME";
}
```

Unrestricted © Siemens 2020

Unrestricted © Siemens 2020

Cat or Dog?

SIEMENS
Ingenuity for life

sample

dog

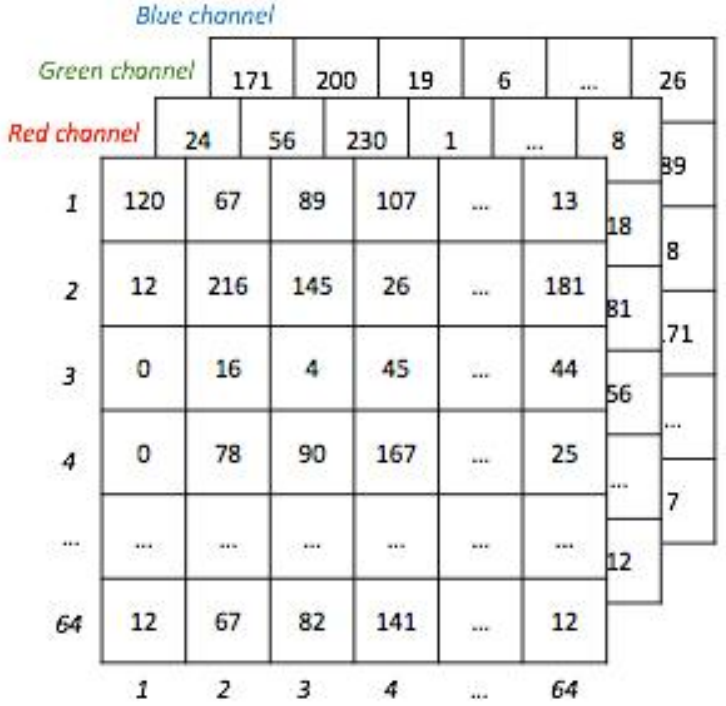


Image array: [64 x 64 x 3]

Unrestricted © Siemens 2020

Cat or Dog?

SIEMENS
Ingenuity for life

dataset

{cat, dog}

Shuffle Data + Cross Validation: Holdout, Leave-p-Out, Leave-One-Out, K-Fold

Unrestricted © Siemens 2020

Cat or Dog?

Classifier from Raw Pixels

In[ ]:= catdogTrain = { → "cat", → "cat", → "cat", → "dog", → "dog", → "dog", → "dog"};

In[ ]:= catdogTest = { → "cat", → "cat", → "cat", → "dog", → "dog", → "dog", → "dog"};

```
In[ ]:= classifier = Classify[catdogTrain, FeatureExtractor → "PixelVector"]
```

Out[ ]= ClassifierFunction[Input type: Image
Classes: cat, dog]

Data not in notebook; Store now »

```
In[ ]:= ClassifierMeasurements[classifier, catdogTest, "Accuracy"]
```

Out[ ]= 0.416667

~50%

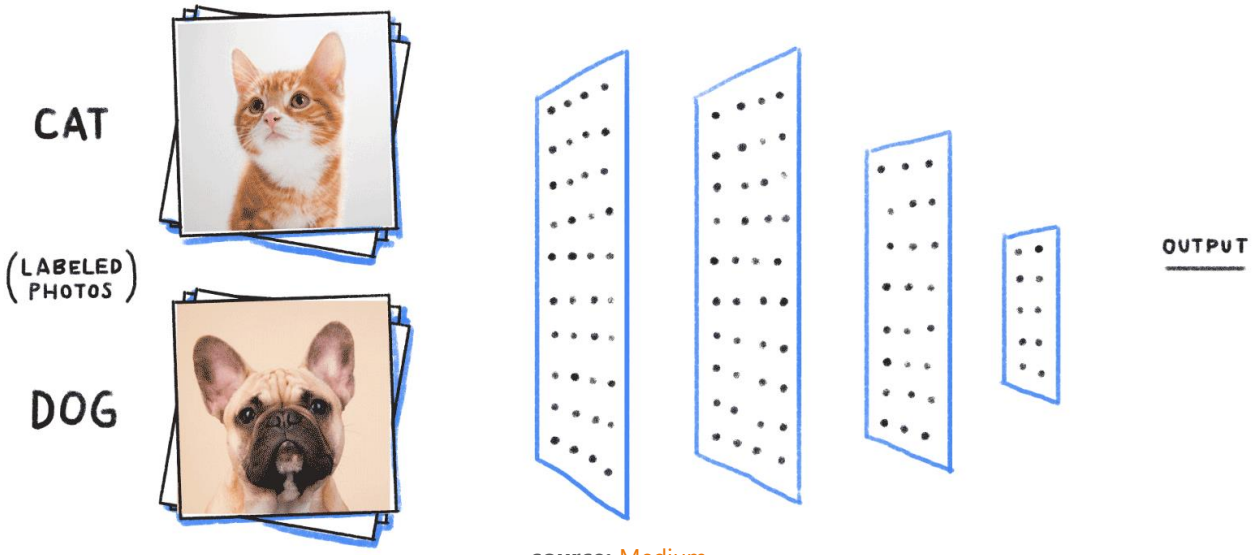
Unrestricted © Siemens 2020

source: *adapted from* Wolfram

Cat or Dog?

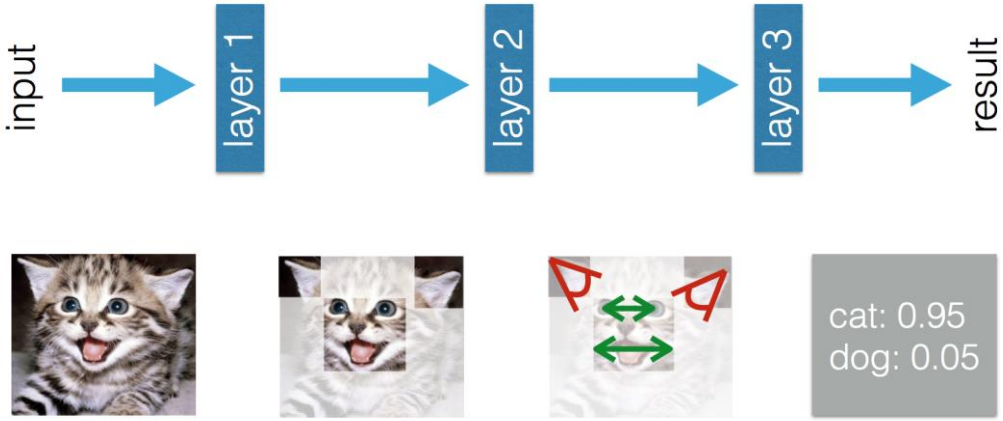
SIEMENS
Ingenuity for life

neural network



source: Medium

features



source: StackExchange

Unrestricted © Siemens 2020

Cat or Dog?

SIEMENS
Ingenuity for life

feature extraction

Train a feature extractor on a dataset of images:

In[1]:= **fe = FeatureExtraction[**

Out[1]= FeatureExtractorFunction[Input type: Image
Output type: NumericalVector (length: 5)]

Use the feature extractor on the training set:

In[2]:= **fe[**

Out[2]= {{16.6901, 0.6639, 2.31437, −8.67619, 22.3044},
{8.66663, 12.1656, 11.059, −15.3779, −16.1057}, {23.1908, −8.59534, −12.0011, 14.9865, −8.70454},
{−14.2674, −21.1947, 18.7504, 4.11292, −0.570131},
{−19.8211, −4.7849, −22.4557, −11.3487, −1.03039}, {−14.459, 21.7455, 2.33303, 16.3033, 4.10636}}

source: *adapted from* Wolfram

Unrestricted © Siemens 2020

Cat or Dog?

Classifier with ResNet-50

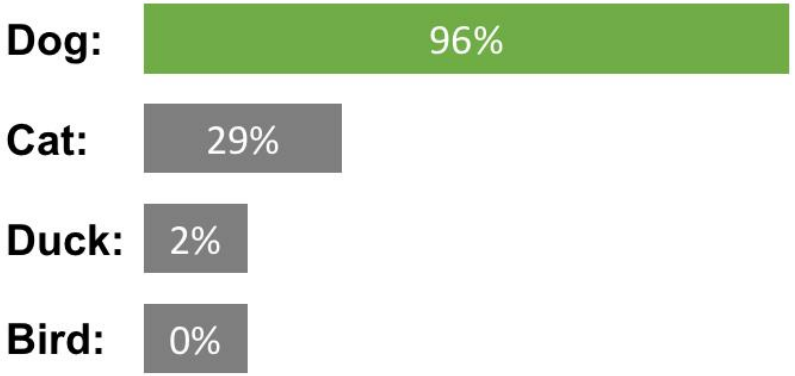
```
In[ ]:= net = NetModel["ResNet-50 Trained on ImageNet Competition Data"]

Out[ ]= NetChain[ Input port: image
                  Output port: class
                  Number of layers: 25 ]

In[ ]:= netFeature = NetDrop[net, -2]

Out[ ]= NetChain[ Input port: image
                  Output port: vector (size: 2048)
                  Number of layers: 23 ]

In[ ]:= classifier = Classify[catdogTrain, FeatureExtractor → netFeature]

Out[ ]= ClassifierFunction[ Input type: Image
                            Classes: cat, dog ]
        Data not in notebook; Store now »

In[ ]:= ClassifierMeasurements[classifier, catdogTest]["Accuracy"]

Out[ ]= 0.916667
```

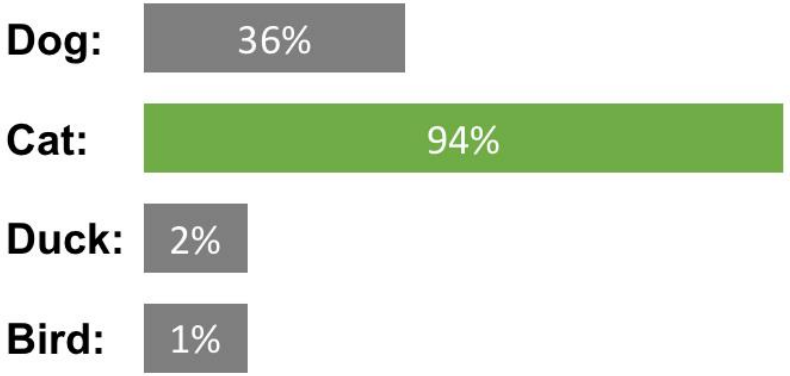
~92%

source: *adapted from* Wolfram

Unrestricted © Siemens 2020

Cat or Dog?





Why probabilities?

Unrestricted © Siemens 2020

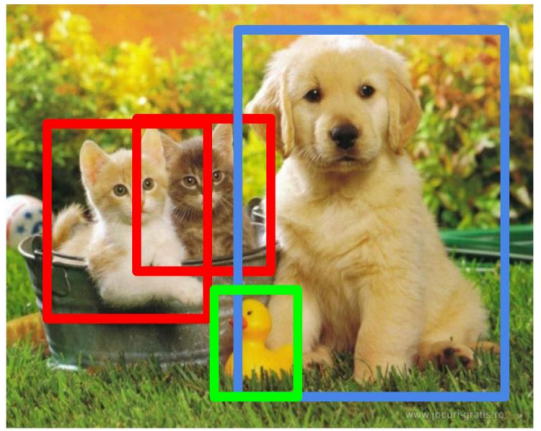
source: reddit

Cat or Croissant?

Unrestricted © Siemens 2020

Beyond *Cat or Dog?*

SIEMENS
Ingenuity for life

Classification	Classification + Localization	Object Detection	Instance Segmentation
			
CAT	CAT	CAT, DOG, DUCK	CAT, DOG, DUCK
Single object		Multiple objects	

source: StackExchange

Unrestricted © Siemens 2020

ML Use-Cases

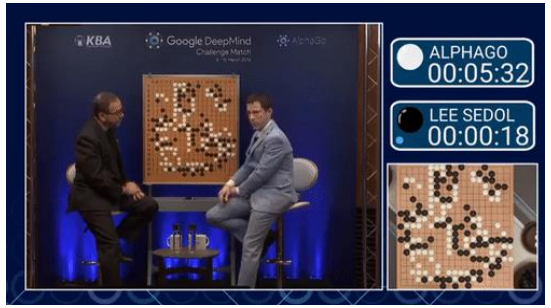
SIEMENS
Ingenuity for life

IBM Watson

Deep Fake

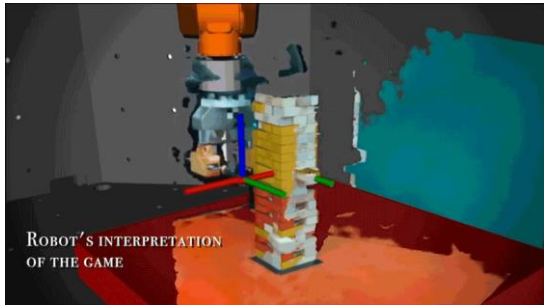
Deep RL and Atari 2600

Autonomous Driving

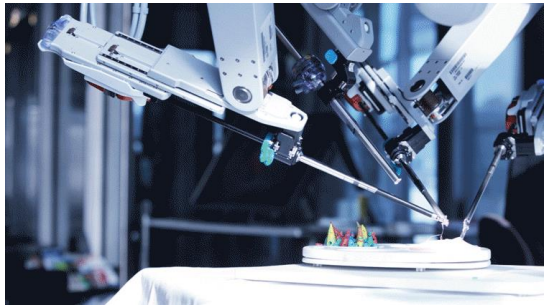


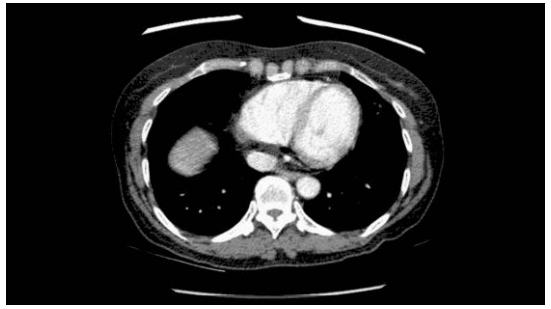
AlphaGo

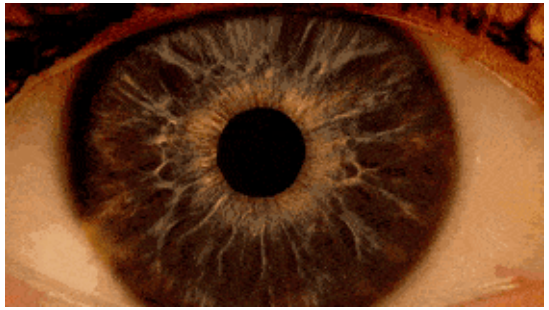
Image Annotation

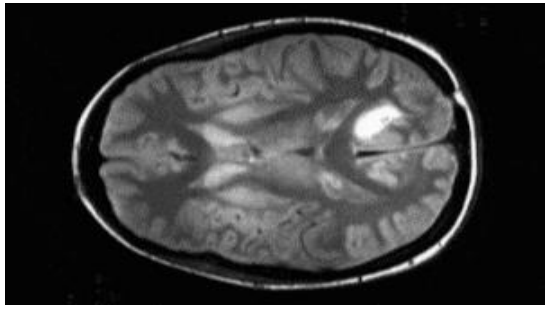


MIT Jenga


Robot Surgery


Chest CT Interpretation


Diabetic Retinopathy Diagnosis

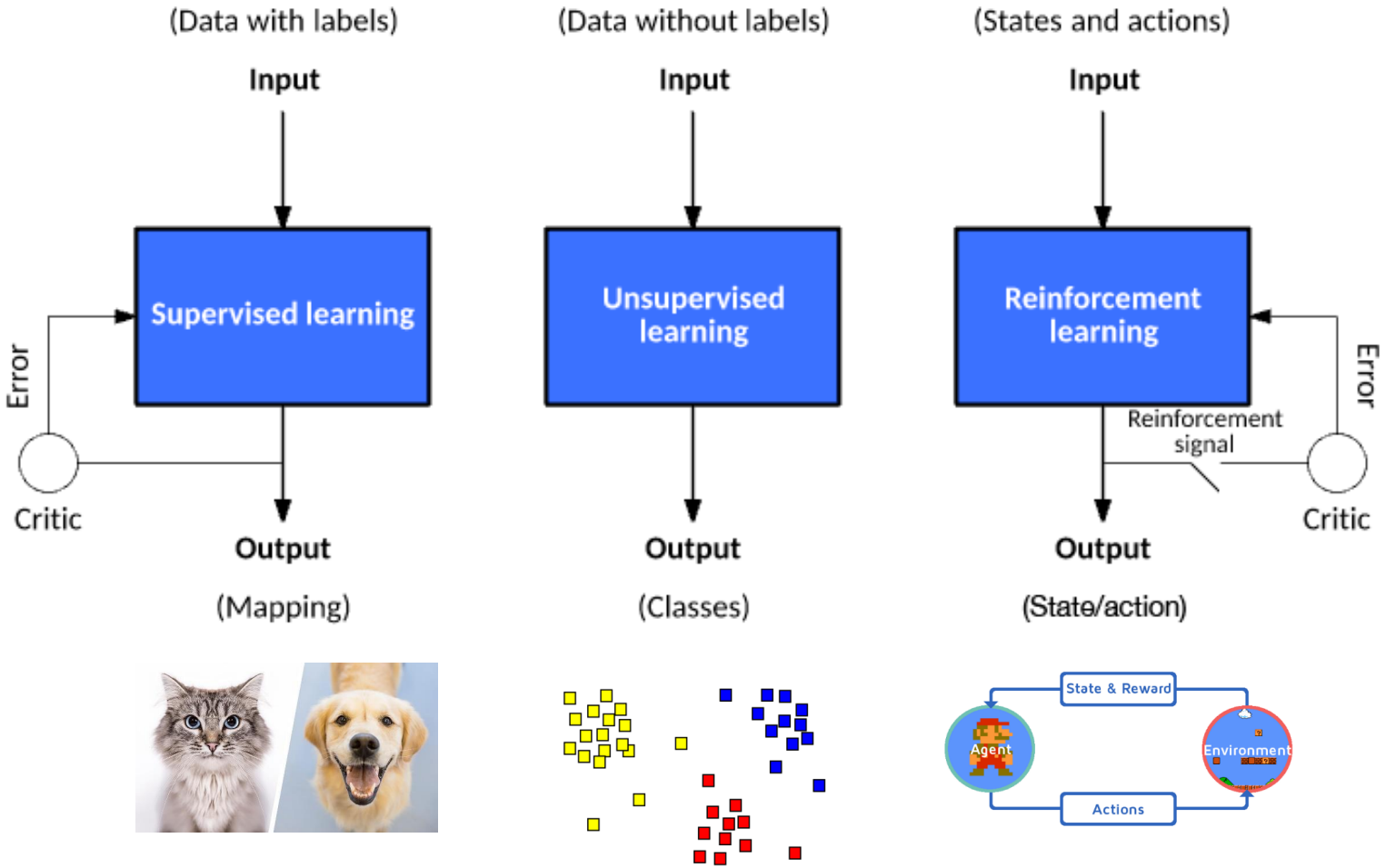

Brain Tumor Detection

Brain-Computer Interfaces

Unrestricted © Siemens 2020

Learning Scenarios

SIEMENS
Ingenuity for life



Keep in mind that there are other learning scenarios that we will **not** be covering in this workshop *e.g.*

- Semi-supervised learning
- Transductive inference
- Online learning
- Active learning

Unrestricted © Siemens 2020

Learning Tasks

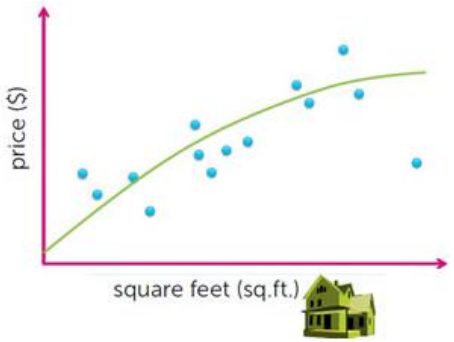
SIEMENS
Ingenuity for life

Classification

assign a category to each item

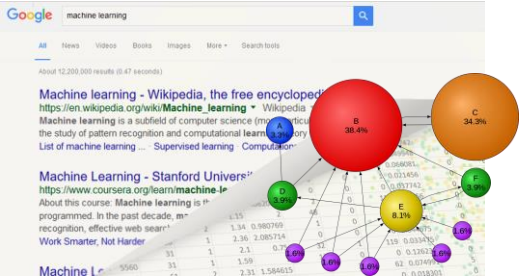
Regression

predict a real value for each item



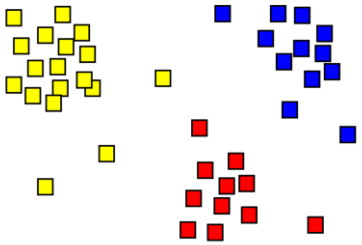
Ranking

learn to order items
according to some criterion



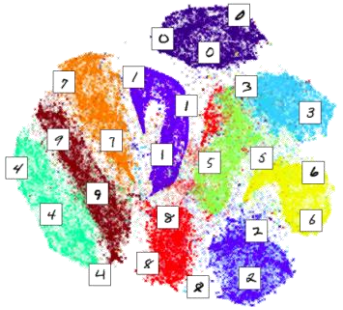
Clustering

partition a set of items into
homogeneous clusters (subsets)



Dimensionality Reduction

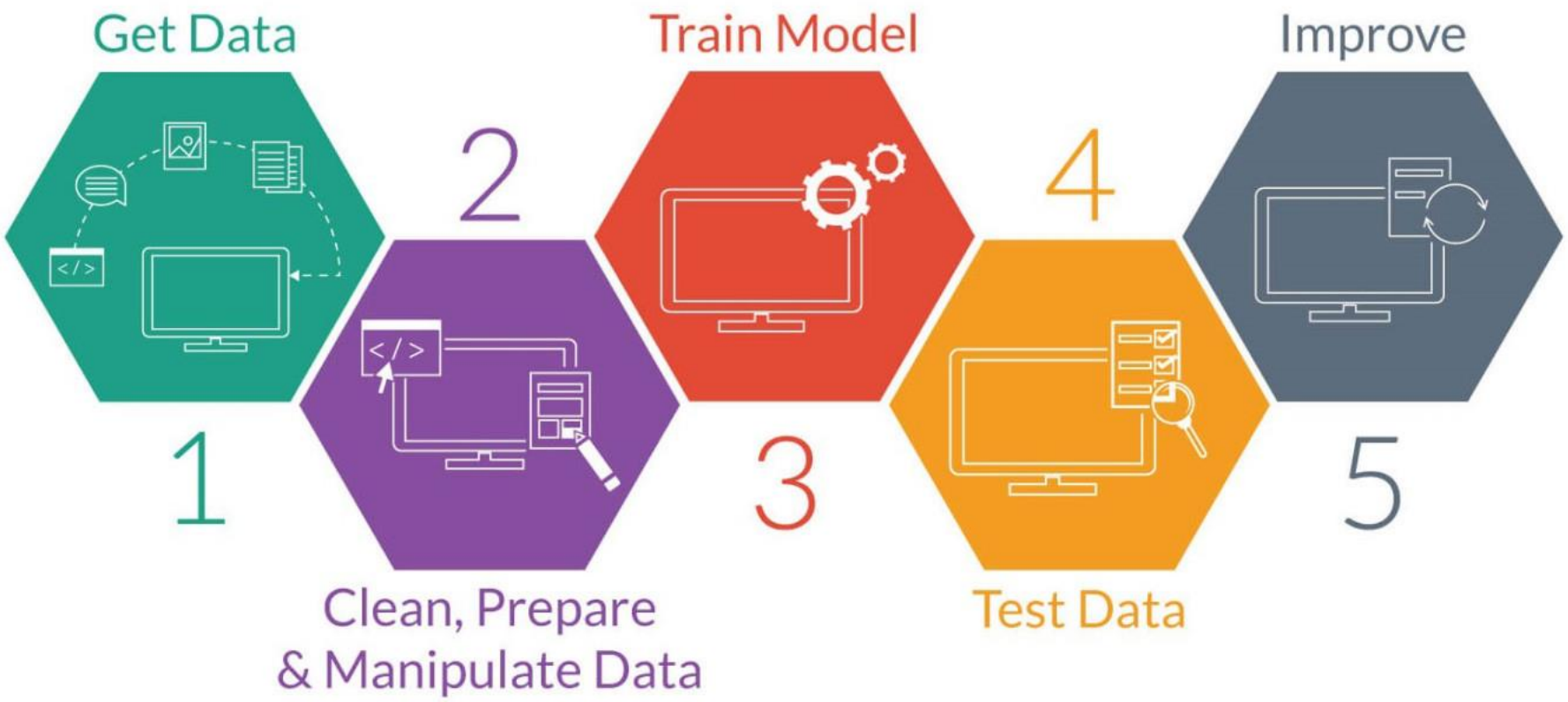
transform an initial repr. of items into a
lower-dim. repr. while preserving some
properties of the initial repr.



source: Mohri, Rostamizadeh & Talwalkar (2018)

Unrestricted © Siemens 2020

ML Pipeline



source: *adapted from* freeCodeCamp

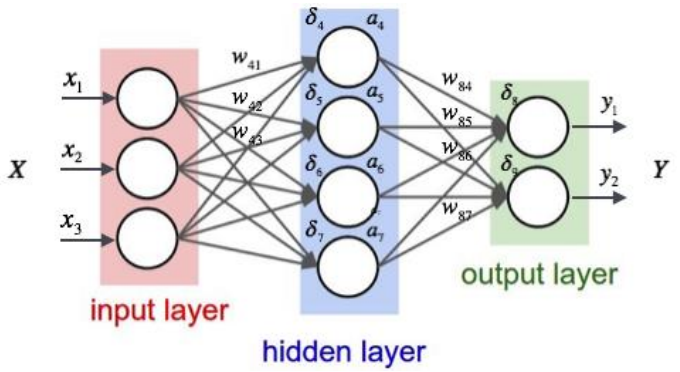
Unrestricted © Siemens 2020

What is an ML Algorithm?

SIEMENS
Ingenuity for life

1 Representation

The formal language in which the learner expresses its models



2 Evaluation

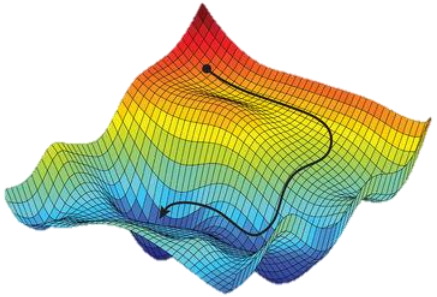
A scoring function to distinguish good models from bad ones

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

makes result quadratic
average over all results
true y
estimate of y

3 Optimization

The algorithm that searches for the highest scoring model

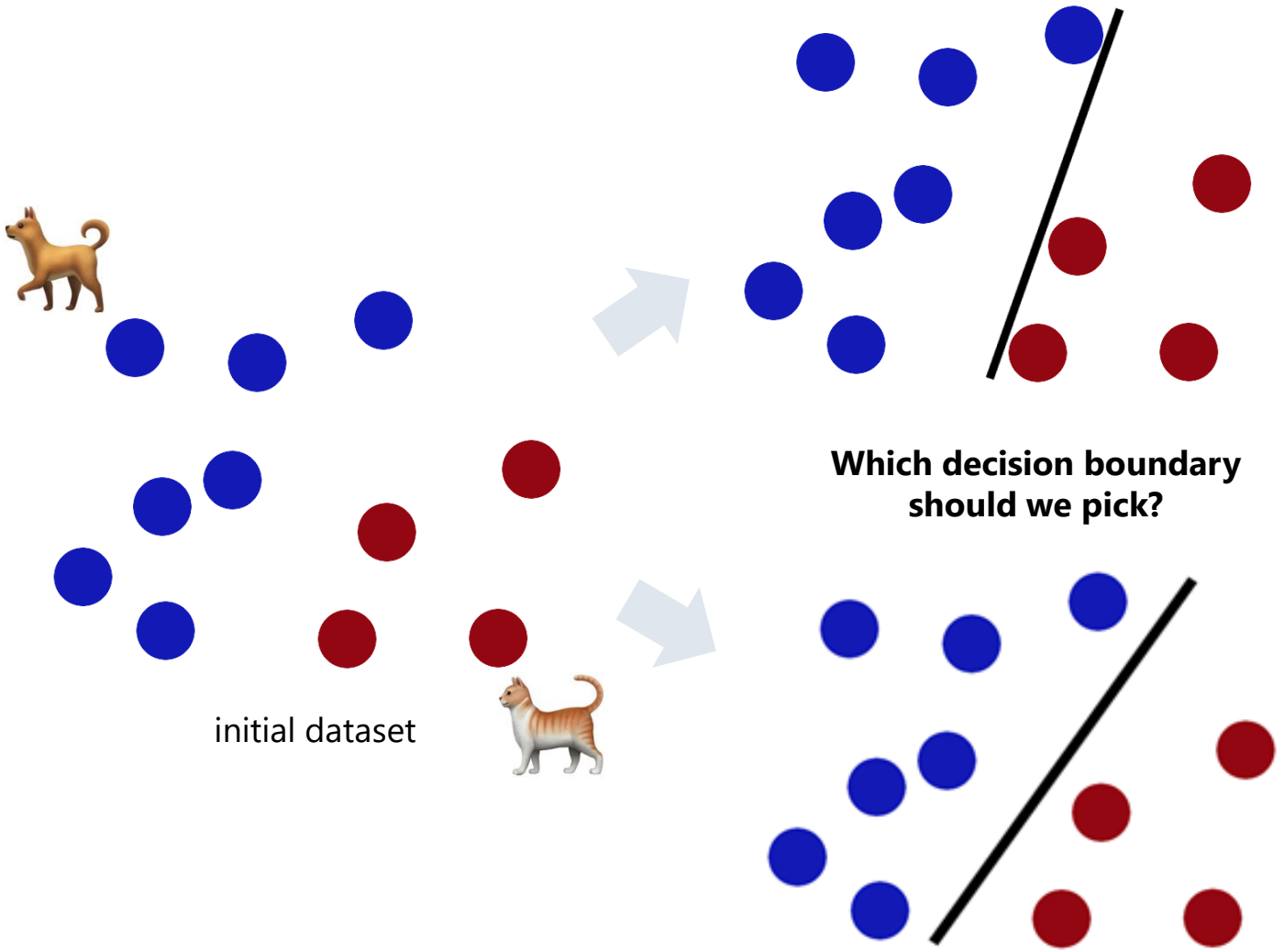


source: Domingos (2012)

Unrestricted © Siemens 2020

ML Algorithms

Support Vector Machines



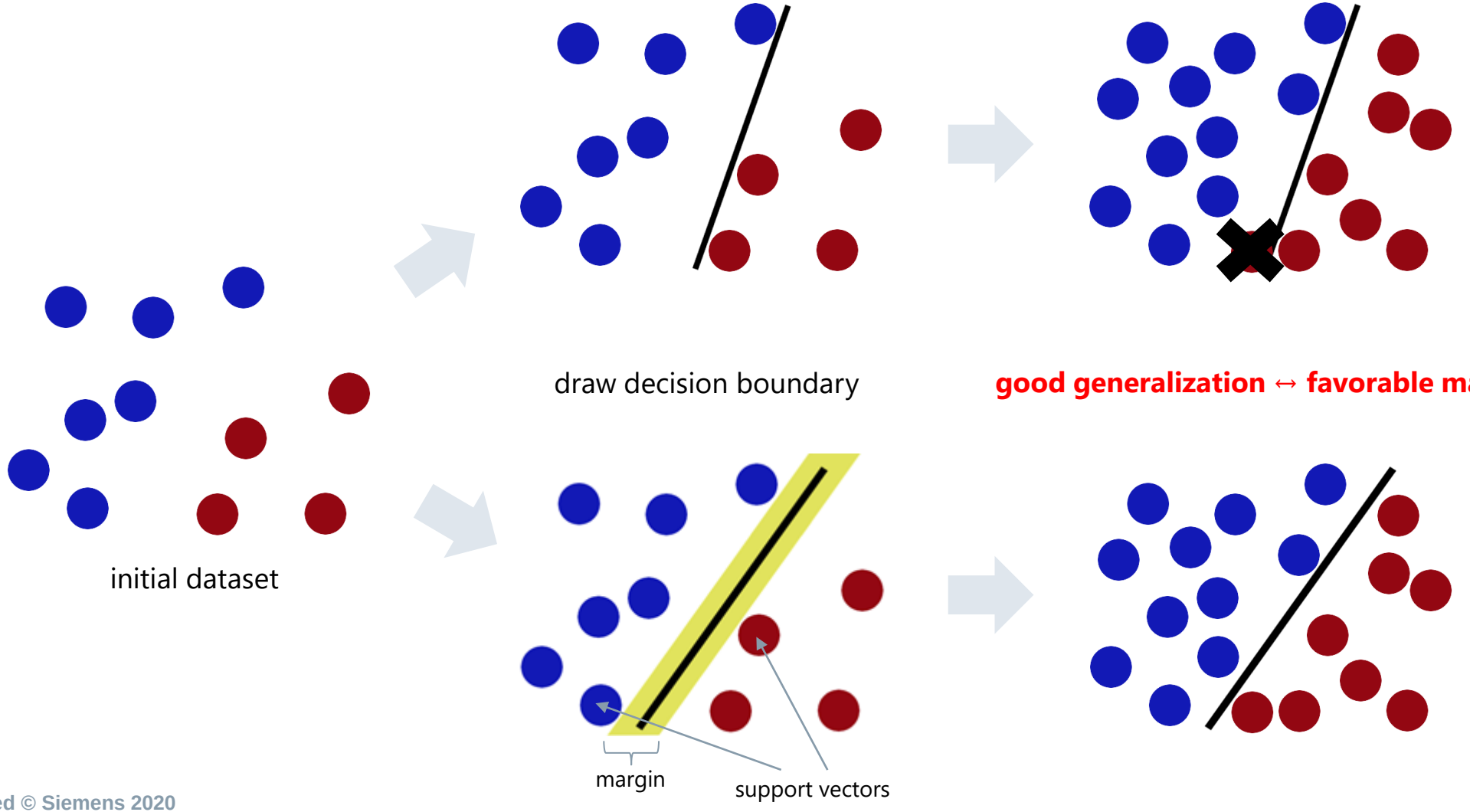
source: reddit

Unrestricted © Siemens 2020

ML Algorithms

Support Vector Machines

SIEMENS
Ingenuity for life



good generalization ↔ favorable margins

source: reddit

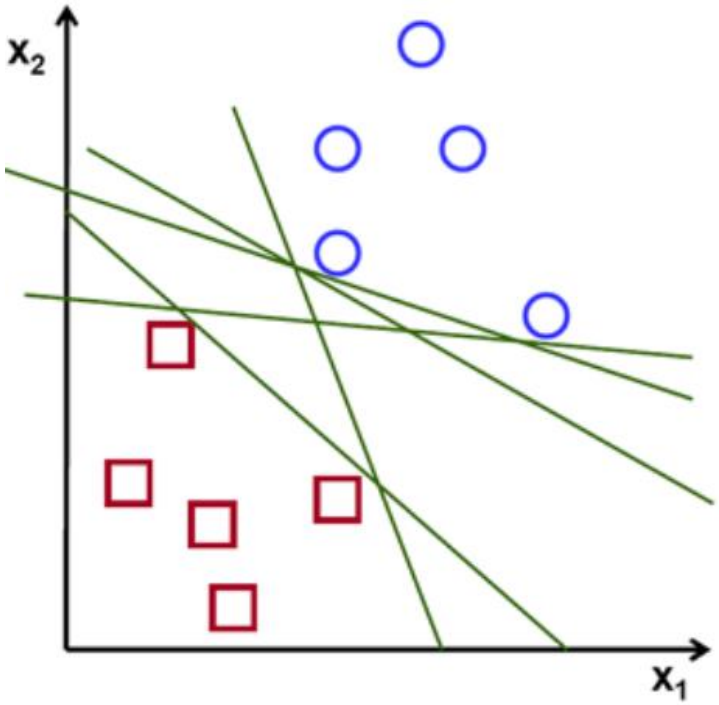
Unrestricted © Siemens 2020

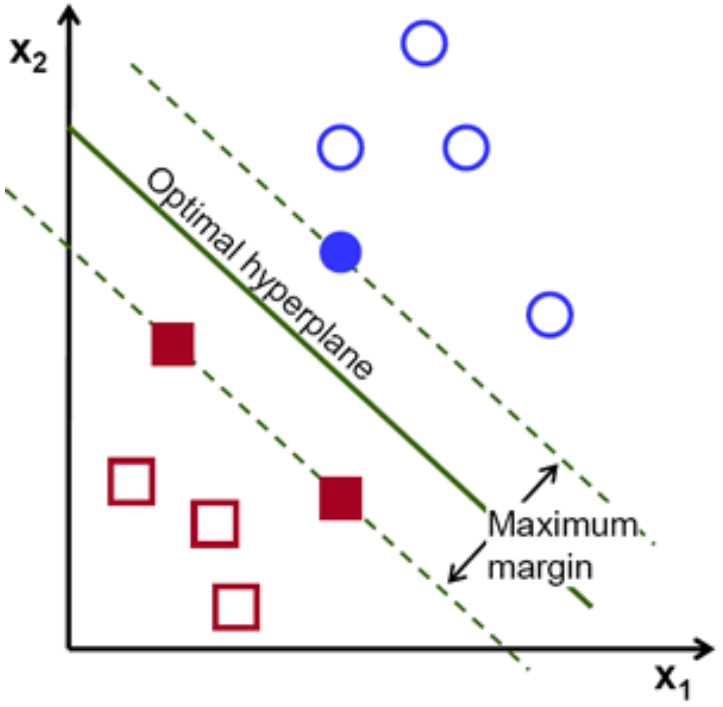
ML Algorithms

Support Vector Machines

SIEMENS
Ingenuity for life

Hypothesis Space





source: reddit

$$H = \{\mathbf{x} \mapsto \operatorname{sign}(\mathbf{w} \cdot \mathbf{x} + b) \colon \mathbf{w} \in \mathbb{R}^N, b \in \mathbb{R}\}$$

source: Mohri, Rostamizadeh & Talwalkar (2018)

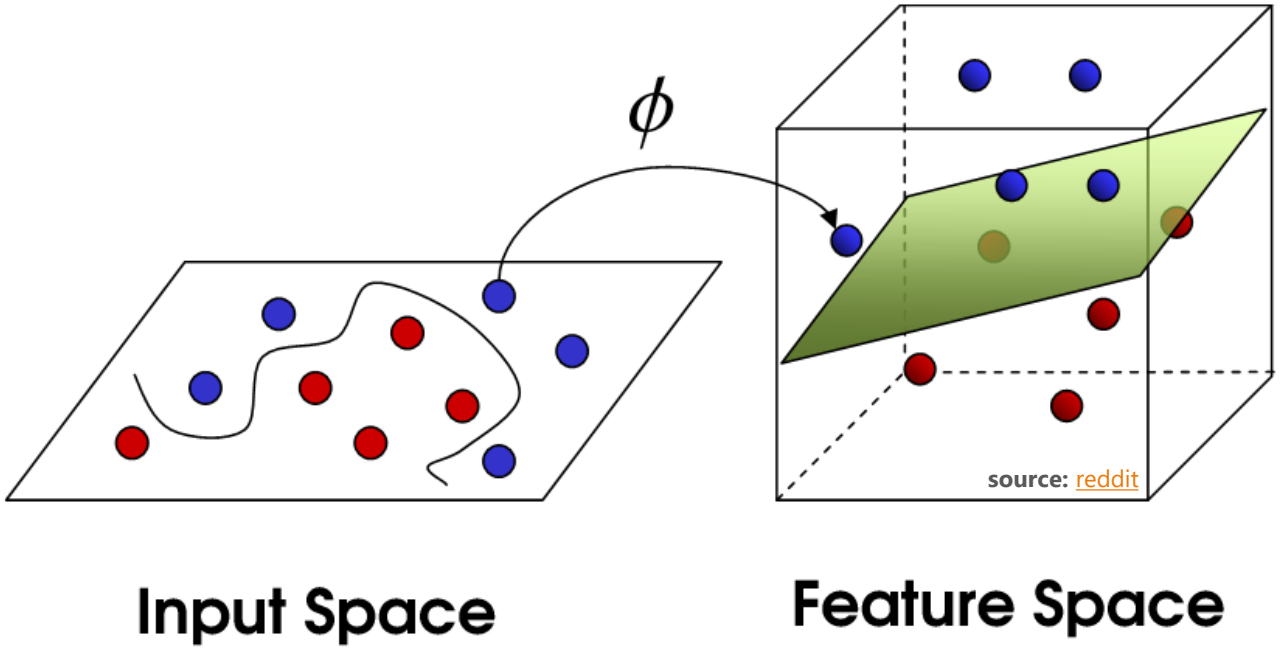
Unrestricted © Siemens 2020

ML Algorithms

Support Vector Machines

SIEMENS
Ingenuity for life

Feature Transformation



Kernel trick: ignore ϕ by focusing on pairwise similarity comparisons

Unrestricted © Siemens 2020

Intermezzo

Statistical Performance Measures

SIEMENS
Ingenuity for life

Reality

Predicted	True	False
True		
False		

source: *adapted from* Medium

Statistical Measures

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

$$Precision = \frac{TP}{TP + FP}$$

$$Sensitivity\ or\ Recall = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{FP + TN}$$

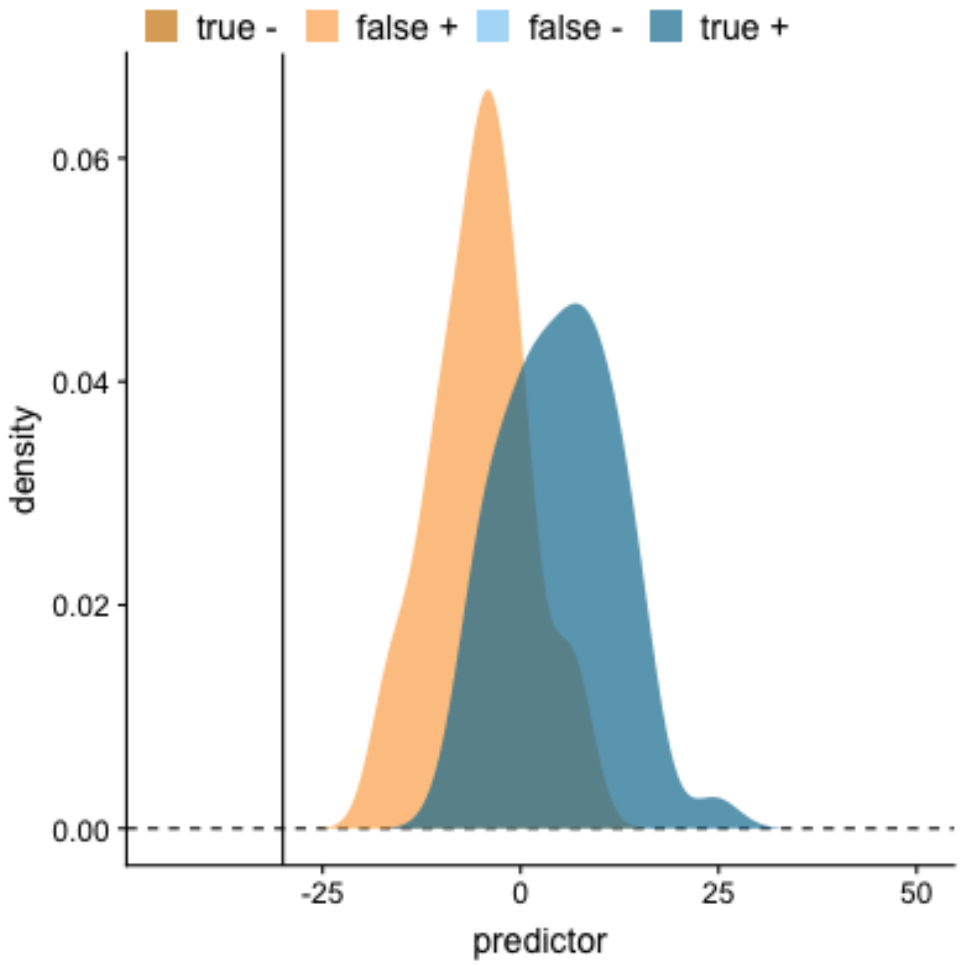
$$F_1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

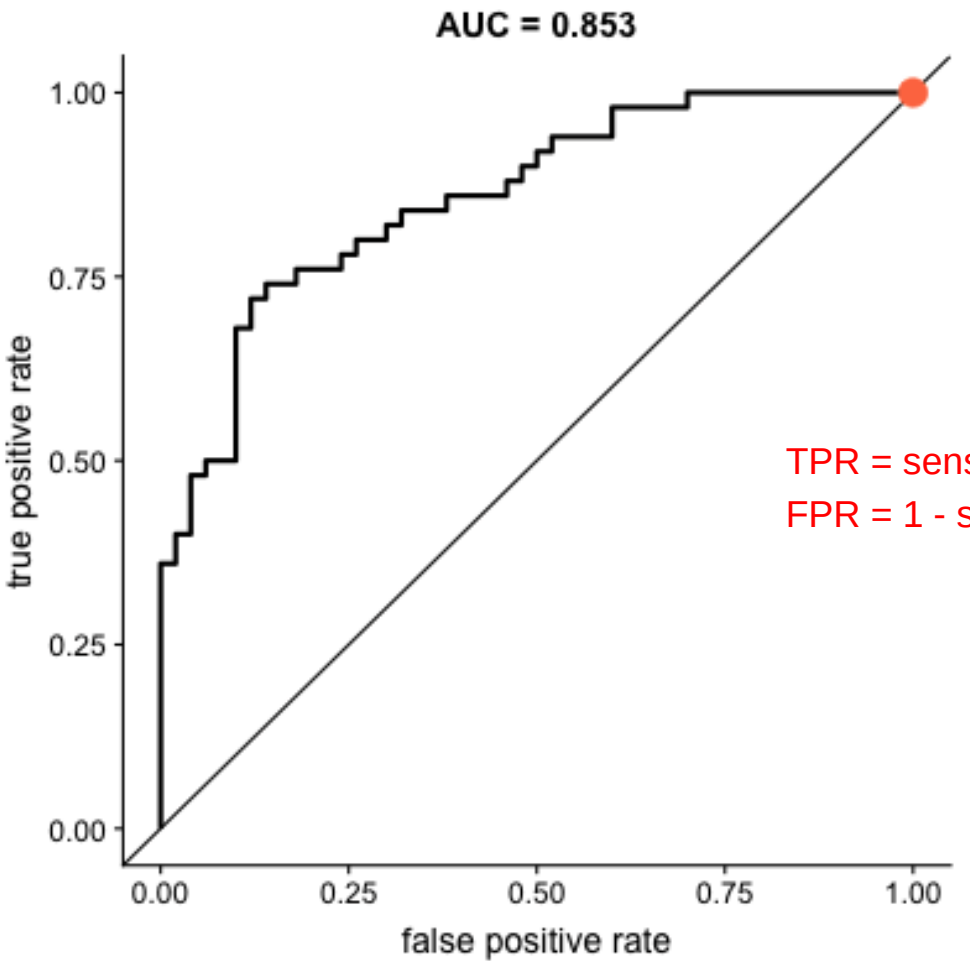
Mnemonic: SnNout and **SpPin**

Unrestricted © Siemens 2020

Intermezzo

ROC and AUC





TPR = sensitivity

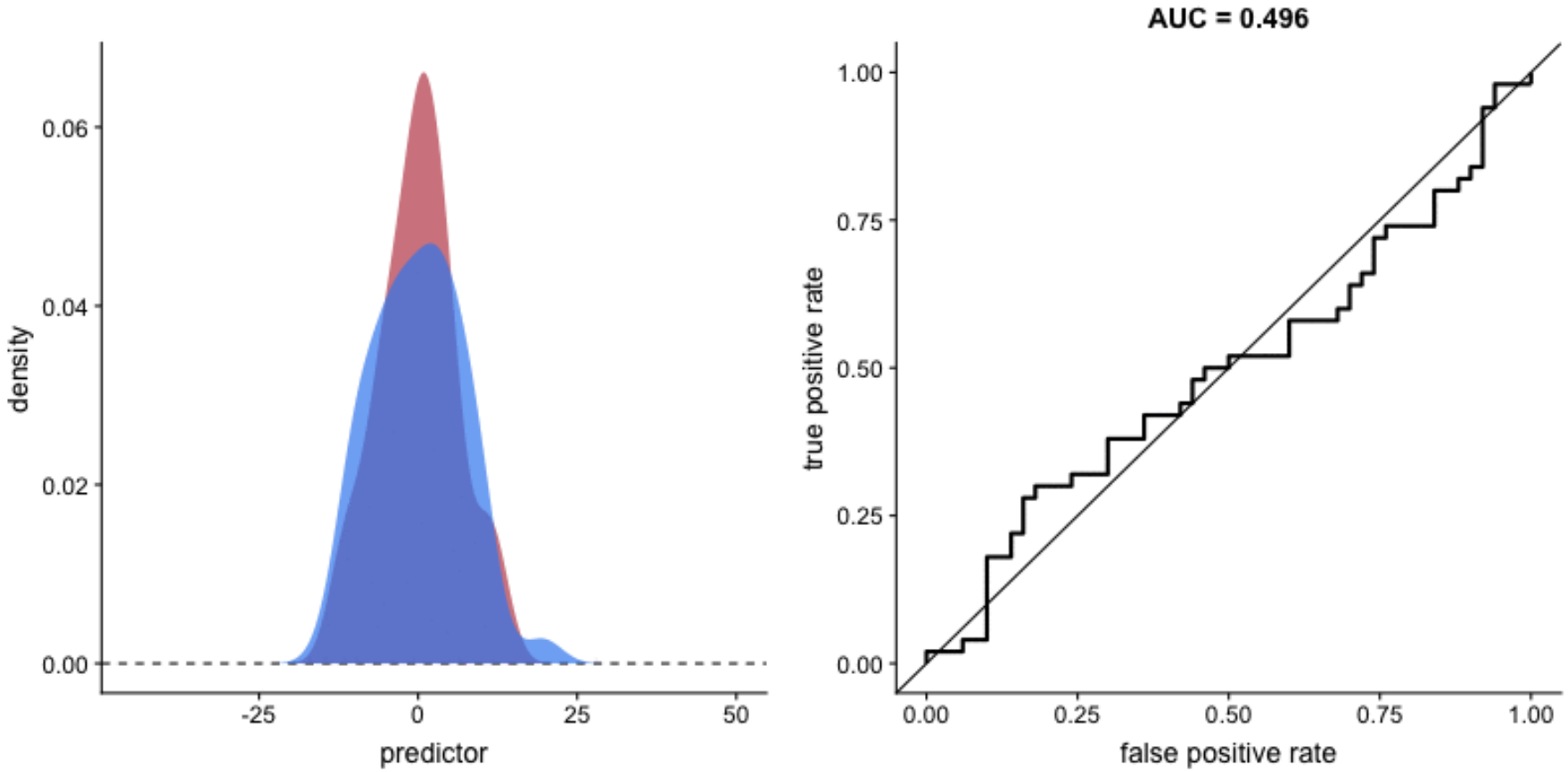
FPR = 1 - specificity

source: paulvanderlaken

Unrestricted © Siemens 2020

Intermezzo

ROC and AUC



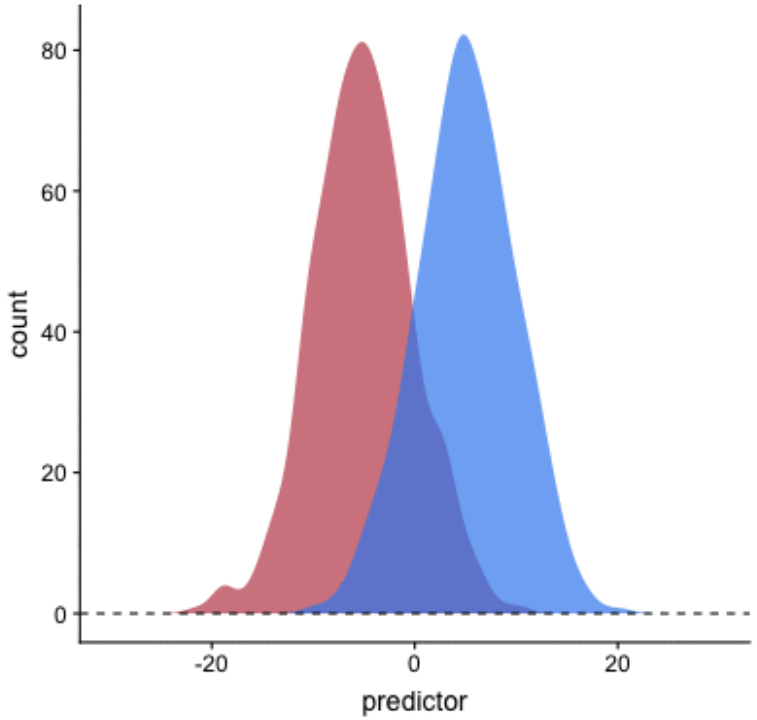
source: paulvanderlaken

Unrestricted © Siemens 2020

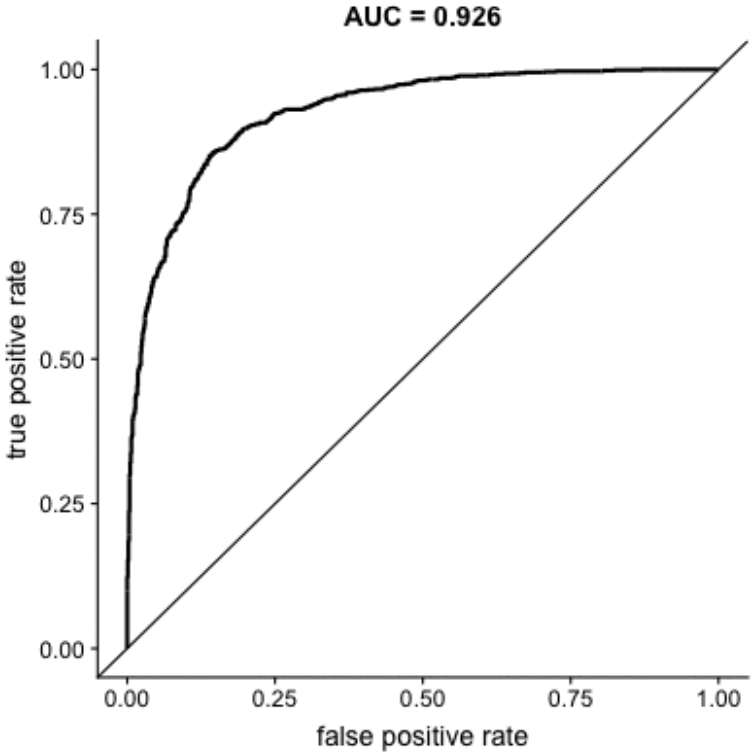
Intermezzo

Class Imbalance

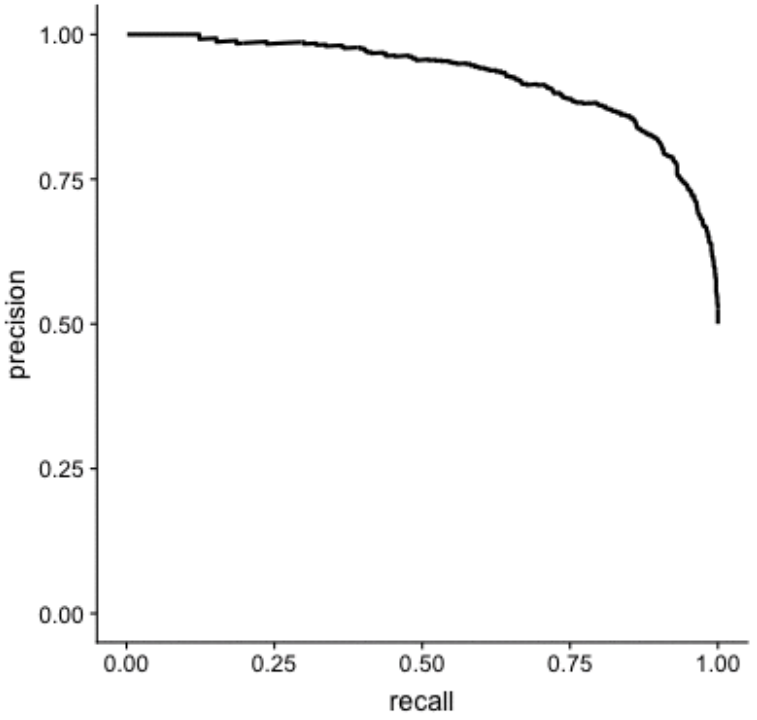
SIEMENS
Ingenuity for life



ROC Curve



Precision / Recall Curve



source: paulvanderlaken

Unrestricted © Siemens 2020

ML Algorithms

Support Vector Machines: Diabetes Prediction

SIEMENS
Ingenuity for life

Variables

- family history
- age
- race and ethnicity
- weight
- height
- waist circumference
- body mass index (BMI)
- hypertension
- sex (II)
- physical activity (II)

(scheme I)

diagnosed/undiagnosed diabetes **VS.** pre-diabetes or no diabetes

(scheme II)

undiagnosed diabetes or pre-diabetes **VS.** no diabetes

Table 1 Description of the National Health and Nutrition Examination Survey data set used for the study

Diagnostic category	Definition	N	Classification Scheme I	Classification Scheme II
Diagnosed diabetes	Answered "yes" to question "Have you ever been told by a doctor or health professionals that you had diabetes?"	1,266	Cases	Excluded from analysis
Undiagnosed diabetes	Answered "no" to question "Have you ever been told by a doctor or health professionals that you had diabetes?" AND Fasting plasma glucose level ≥ 126 mg/dl	195	Cases	Cases
Pre-diabetes	Fasting plasma glucose level 100-125 mg/dl	1,576	Non-cases	Cases
No diabetes	Fasting plasma glucose level <100 mg/dl	3,277	Non-cases	Non-cases

source: Wu *et al.* (2010)

Unrestricted © Siemens 2020

ML Algorithms

Support Vector Machines: Diabetes Prediction

SIEMENS
Ingenuity for life

Table 3 The performance of support vector machine models for the Classification I and Classification II

Model	Data set	Sensitivity	Specificity	PPV	NPV	AUC
Classification Scheme I*	Test	0.7715	0.7503	0.4926	0.9127	0.8347
	Training	0.7938	0.7169	0.4550	0.9211	0.8383
	10-fold cross- validation	0.7765	0.7027	0.4388	0.9130	0.8242
Classification Scheme II*	Test	0.7359	0.6254	0.5061	0.8195	0.7318
	Training	0.7092	0.6590	0.6729	0.8087	0.7393
	10-fold cross- validation	0.7059	0.6589	0.5293	0.8054	0.7357

PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve.

source: Wu *et al.* (2010)

Unrestricted © Siemens 2020

ML Algorithms

Decision Trees

TRAINING DATASET

COLOR	DIAMETER	LABEL
RED	3	APPLE
YELLOW	3	LEMON
PURPLE	1	GRAPES
RED	3	APPLE
YELLOW	3	LEMON
PURPLE	1	GRAPES

CONDITIONS

COLOR== PURPLE?

DIAMETER=3

COLOR== YELLOW?

COLOR== RED?

DIAMETER=1

source: simplilearn

Unrestricted © Siemens 2020

ML Algorithms

Decision Trees

SIEMENS
Ingenuity for life

TRAINING DATASET

COLOR	DIAMETER	LABEL
RED	3	APPLE
YELLOW	3	LEMON
PURPLE	1	GRAPES
RED	3	APPLE
YELLOW	3	LEMON
PURPLE	1	GRAPES

CONDITIONS

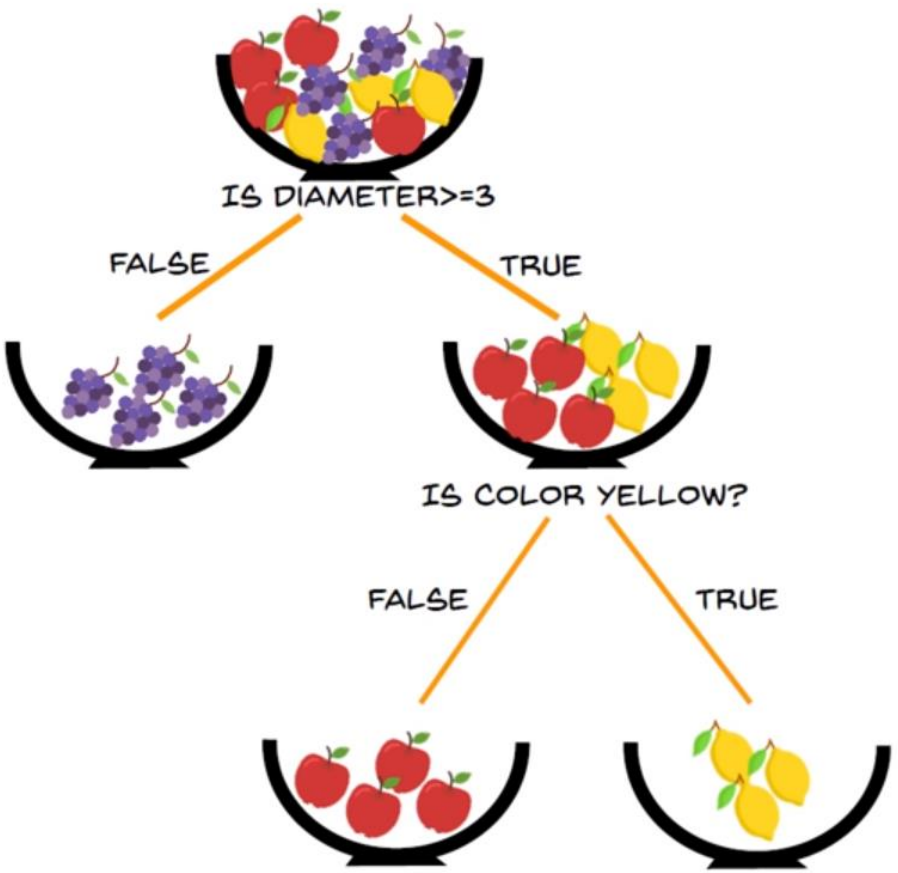
COLOR== PURPLE?

DIAMETER=3

COLOR== YELLOW?

COLOR== RED?

DIAMETER=1



source: simplilearn

Which questions should we ask and when?

Impurity Criterion

Gini Index

$$I_G = 1 - \sum_{j=1}^{c} p_j^2$$

p_j: proportion of the samples that belongs to class c for a particular node

Entropy

$$I_H = -\sum_{j=1}^{c} p_j log_2(p_j)$$

p_j: proportion of the samples that belongs to class c for a particular node.

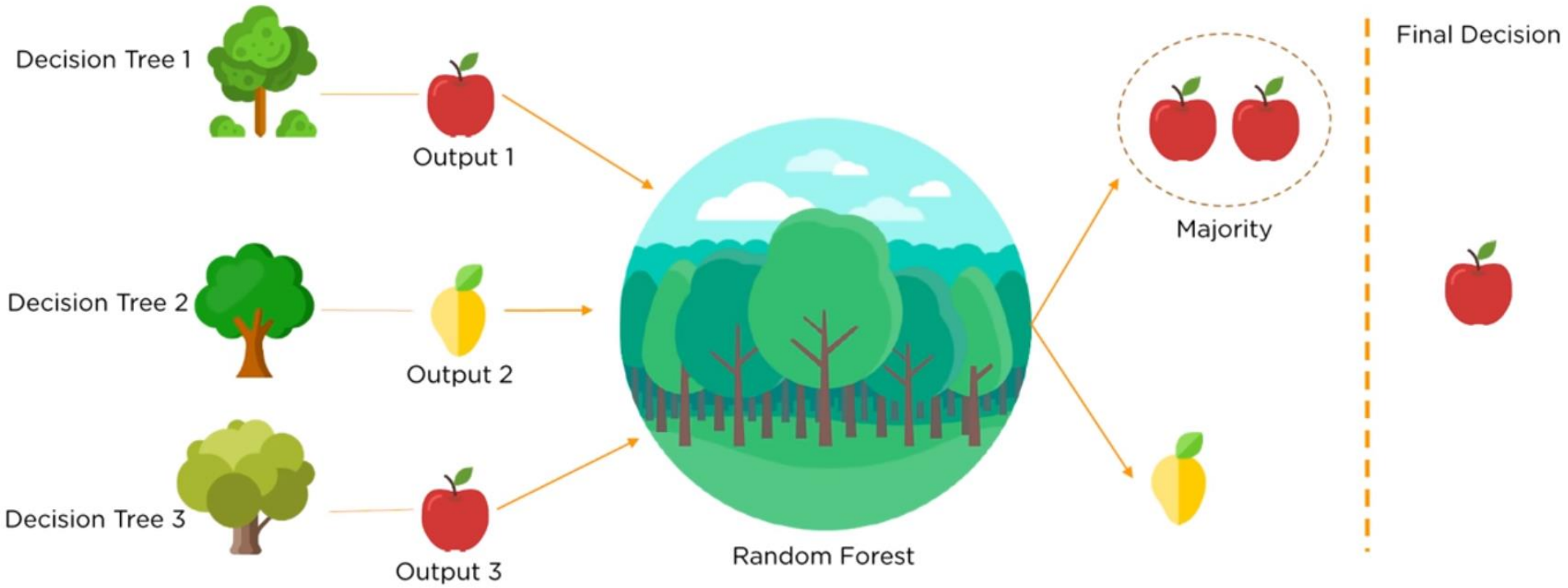
*This is the the definition of entropy for all non-empty classes ($p \neq 0$). The entropy is 0 if all samples at a node belong to the same class.

The **Gini index** quantifies the chance of being incorrect if one randomly assigns a label to an exemple in the same set

Unrestricted © Siemens 2020

ML Algorithms

Random Forest

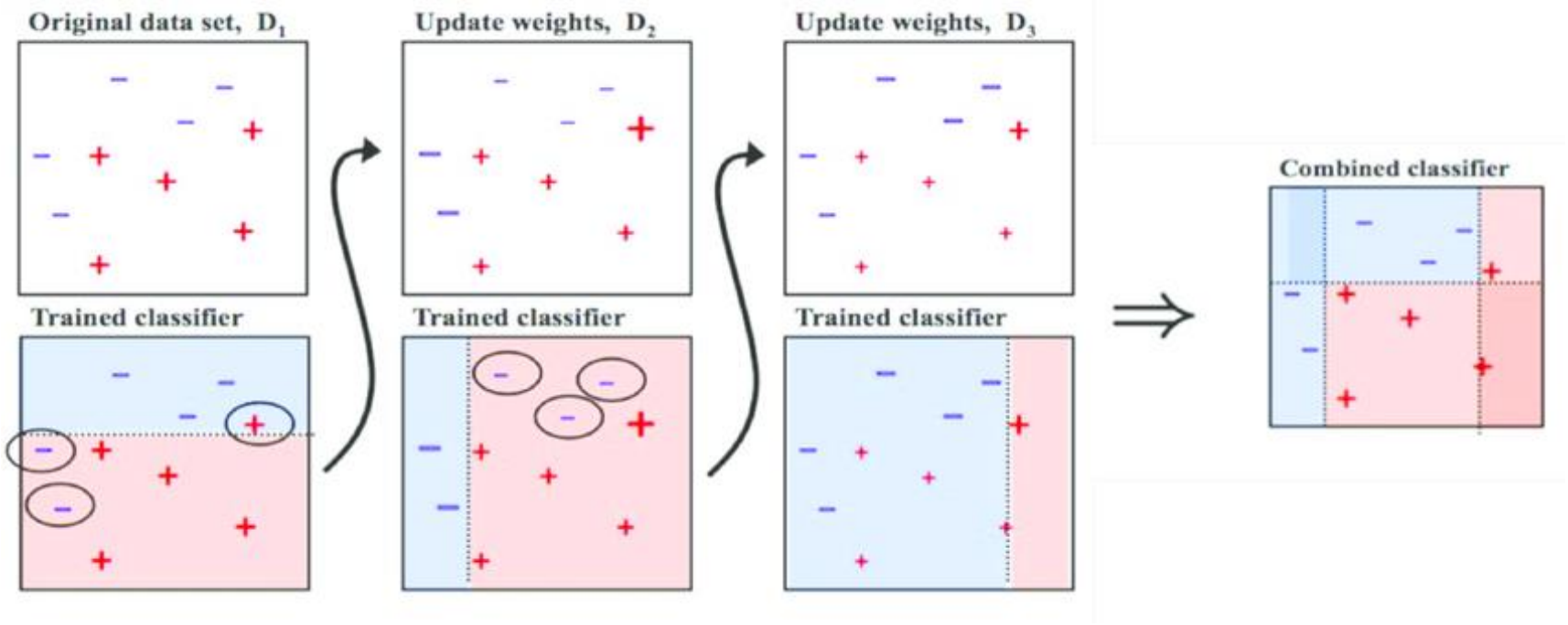


source: simplilearn

Unrestricted © Siemens 2020

Intermezzo

Boosting



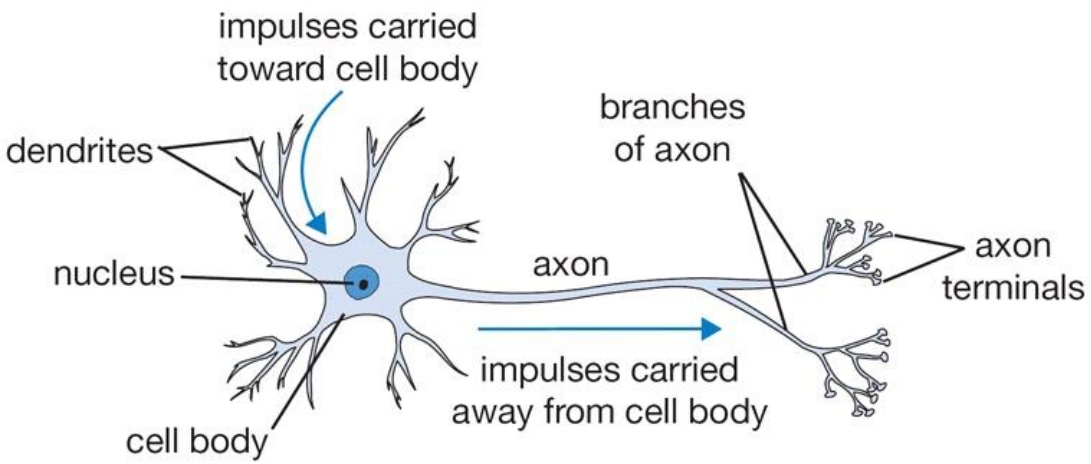
source: Medium

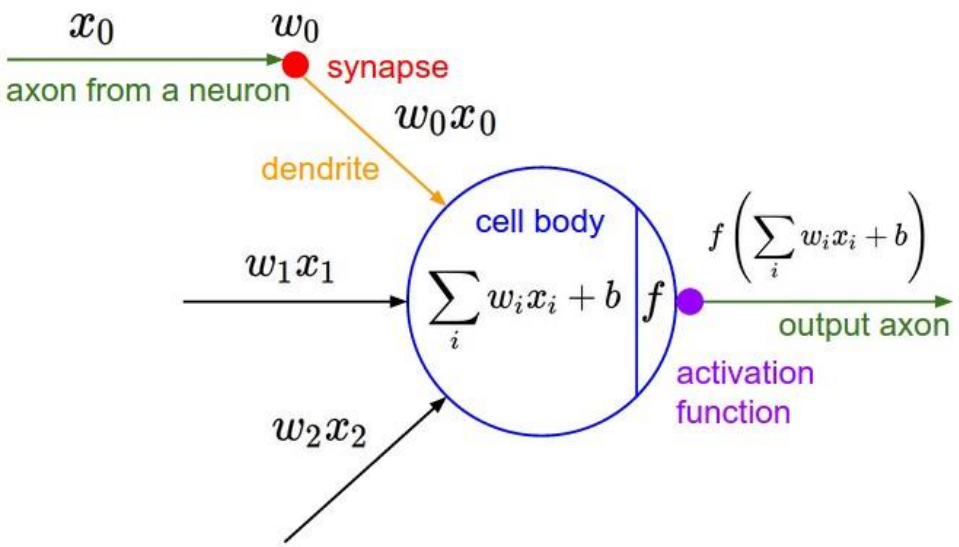
TL;DR Boosting turns weak/simple learners into strong learners

Unrestricted © Siemens 2020

Neural Networks

Biological vs. Artificial Neurons



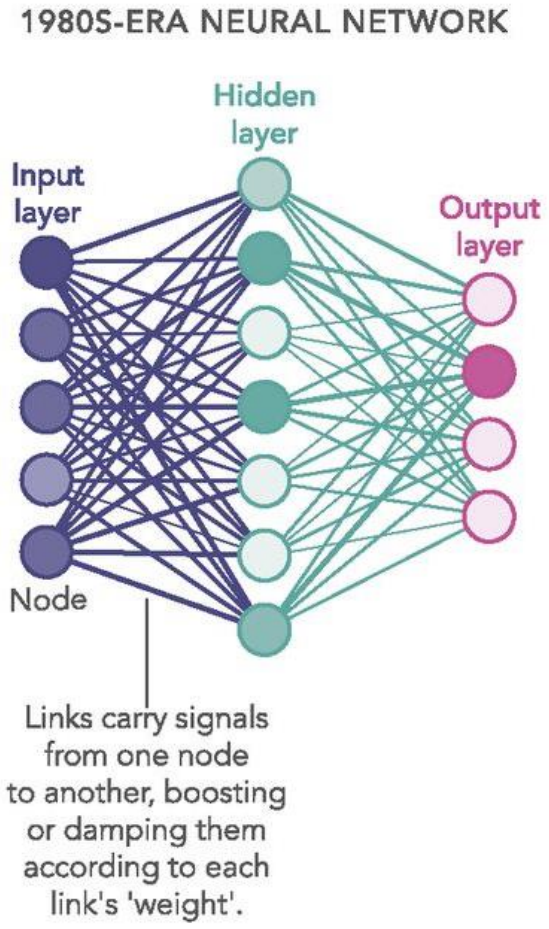


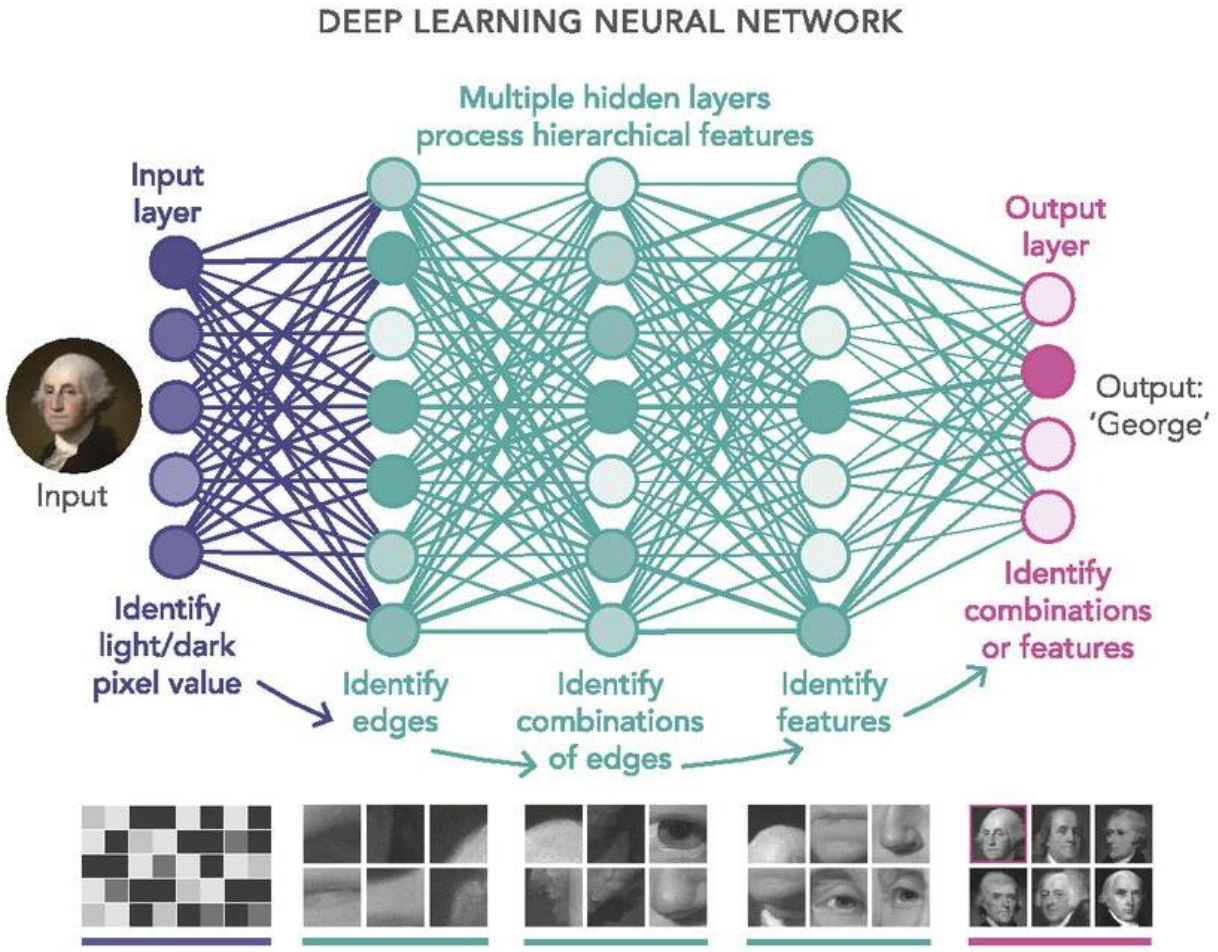
source: CS231n

Unrestricted © Siemens 2020

Neural Networks

From Shallow to Deep Learning



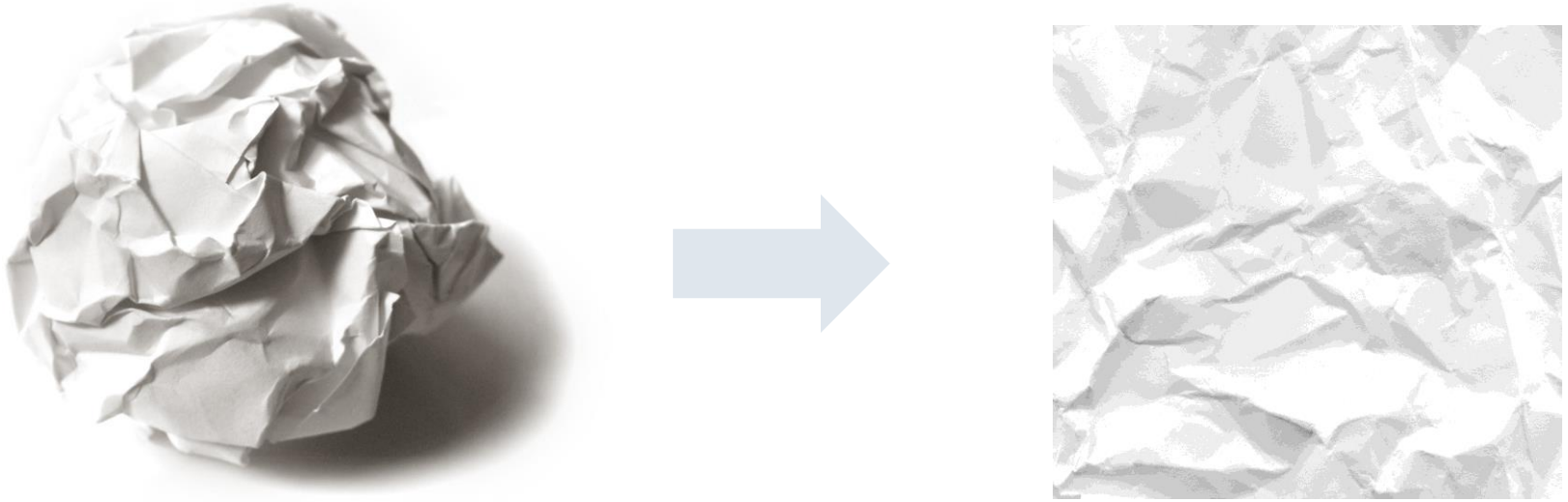


source: PNAS

Unrestricted © Siemens 2020

Neural Networks

Geometric Intuition



"Uncrumpling paper balls is what machine learning is about:
finding neat representations for complex, highly folded data manifolds"

source: Chollet (2015)

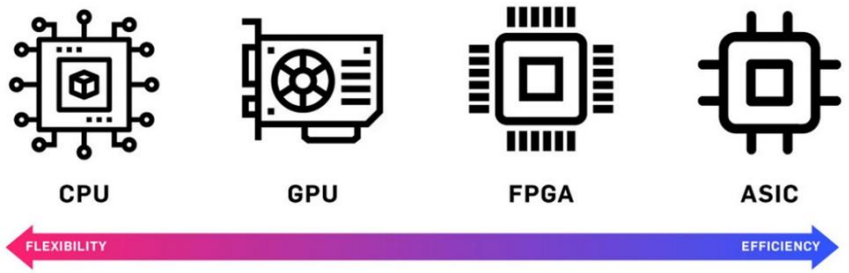
Unrestricted © Siemens 2020

Neural Networks

Deep Learning: Why Now?

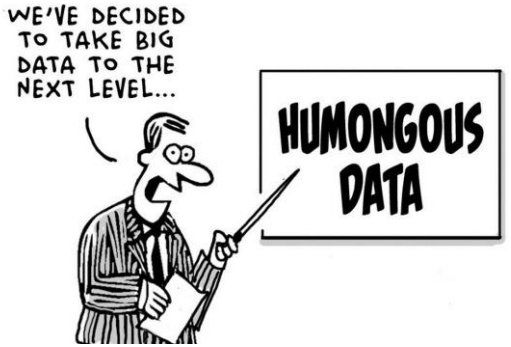
SIEMENS
Ingenuity for life

Hardware

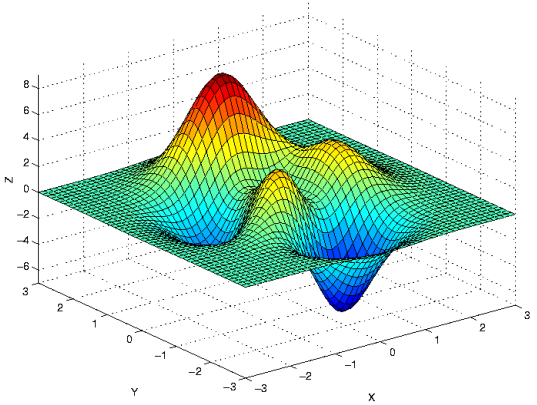


Deep Learning Revolution

Data



Algorithms



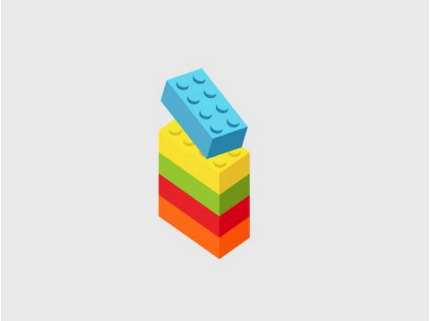
source: Sejnowski (2018)

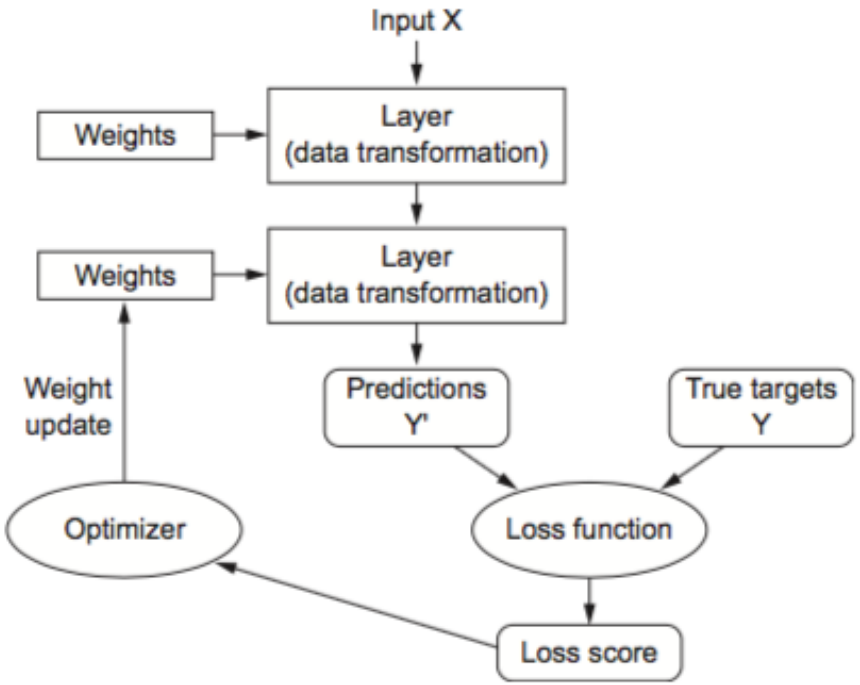
Unrestricted © Siemens 2020

Neural Networks

Ingredients

1. **Input data** – what we will feed into the network
2. **Layers** – which are combined into a *network* or *model*
3. **Loss function** – the *feedback* signal used for learning
4. **Optimizer** – determines how learning should proceed



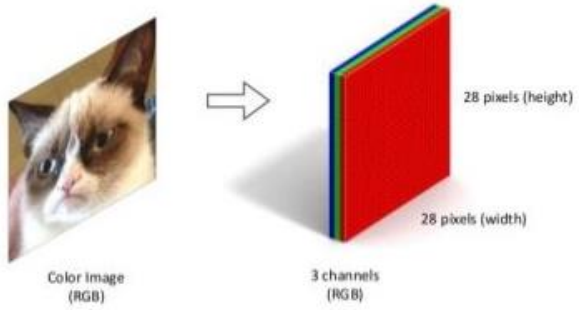


source: Chollet (2015)

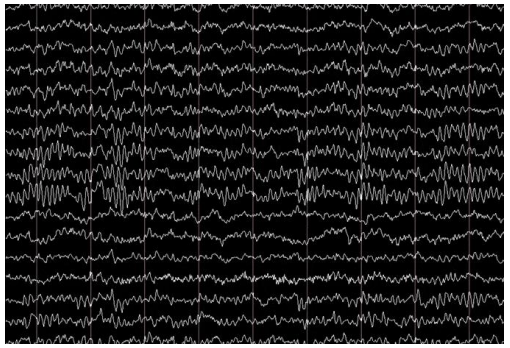
Unrestricted © Siemens 2020

Neural Networks

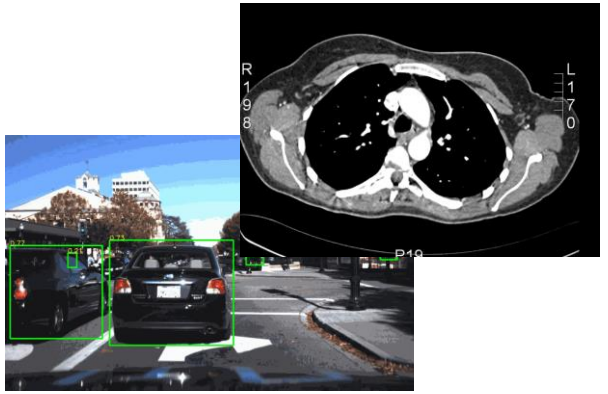
Data Types and Tensors



images

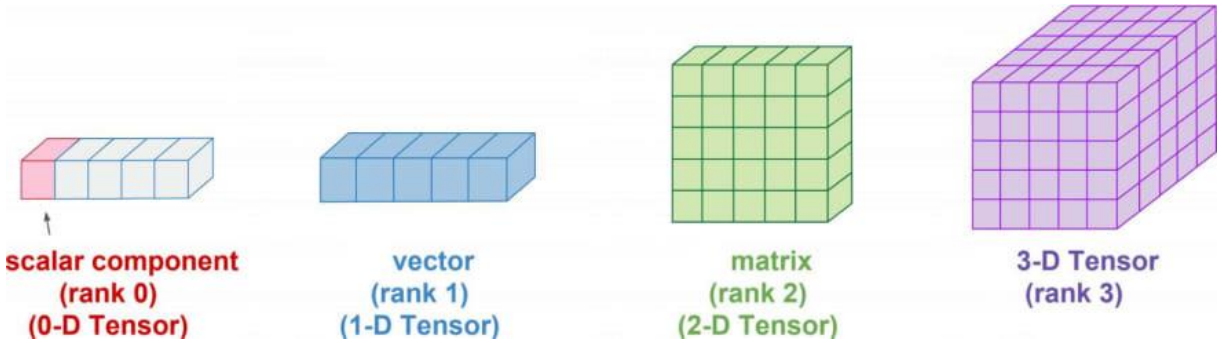


timeseries or **sequence data**



video

tensor:
multidimensional array



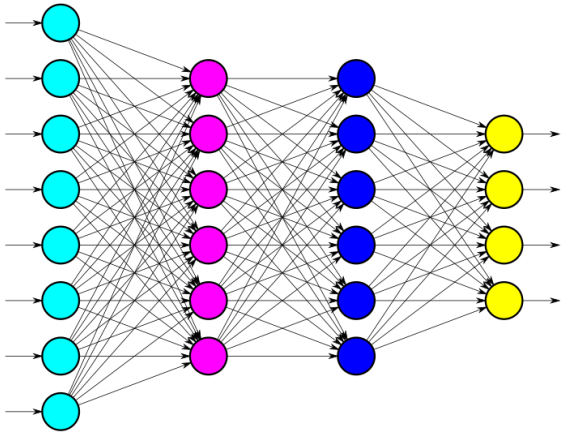
Unrestricted © Siemens 2020

Neural Networks

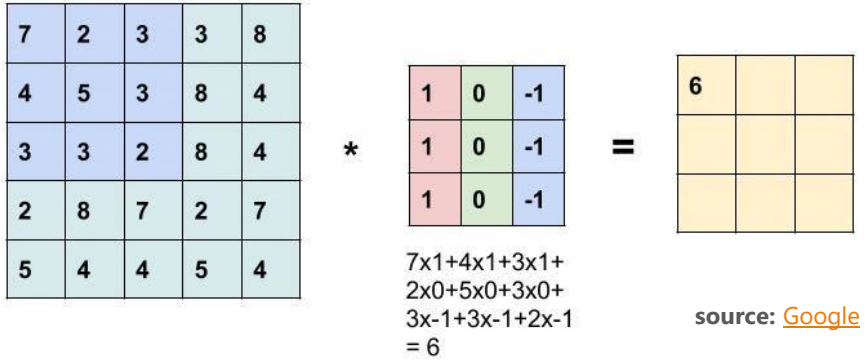
Common Layers

SIEMENS
Ingenuity for life

Dense

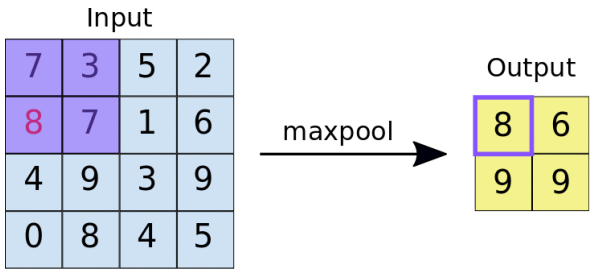


Convolution



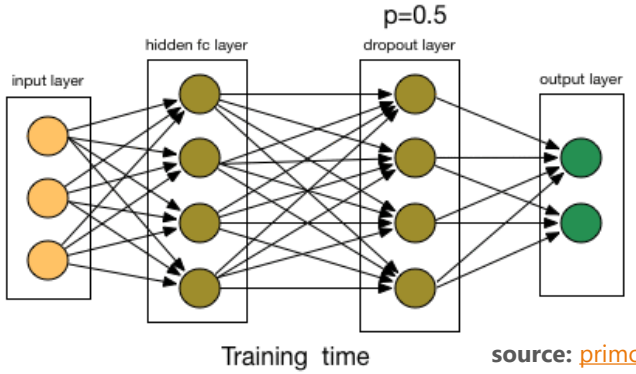
source: Google

Pooling



source: Google

Dropout

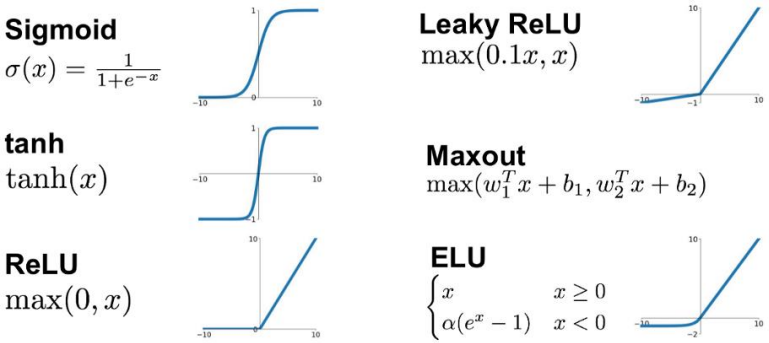


source: primo.ai

Unrestricted © Siemens 2020

Neural Networks

Activation and Loss Functions



source: mc.ai

Problem type	Last-layer Activation	Loss function
Binary classification	Sigmoid	Binary_crossentropy
Multi-class, single-label classification	Softmax	Categorical_crossentropy
Multi-class, multi-label classification	Sigmoid	Binary_crossentropy
Regression to arbitrary values	(None)	MSE
Regression to values between 0 and 1	Sigmoid	MSE or Binary_crossentropy

source: Chollet (2017)

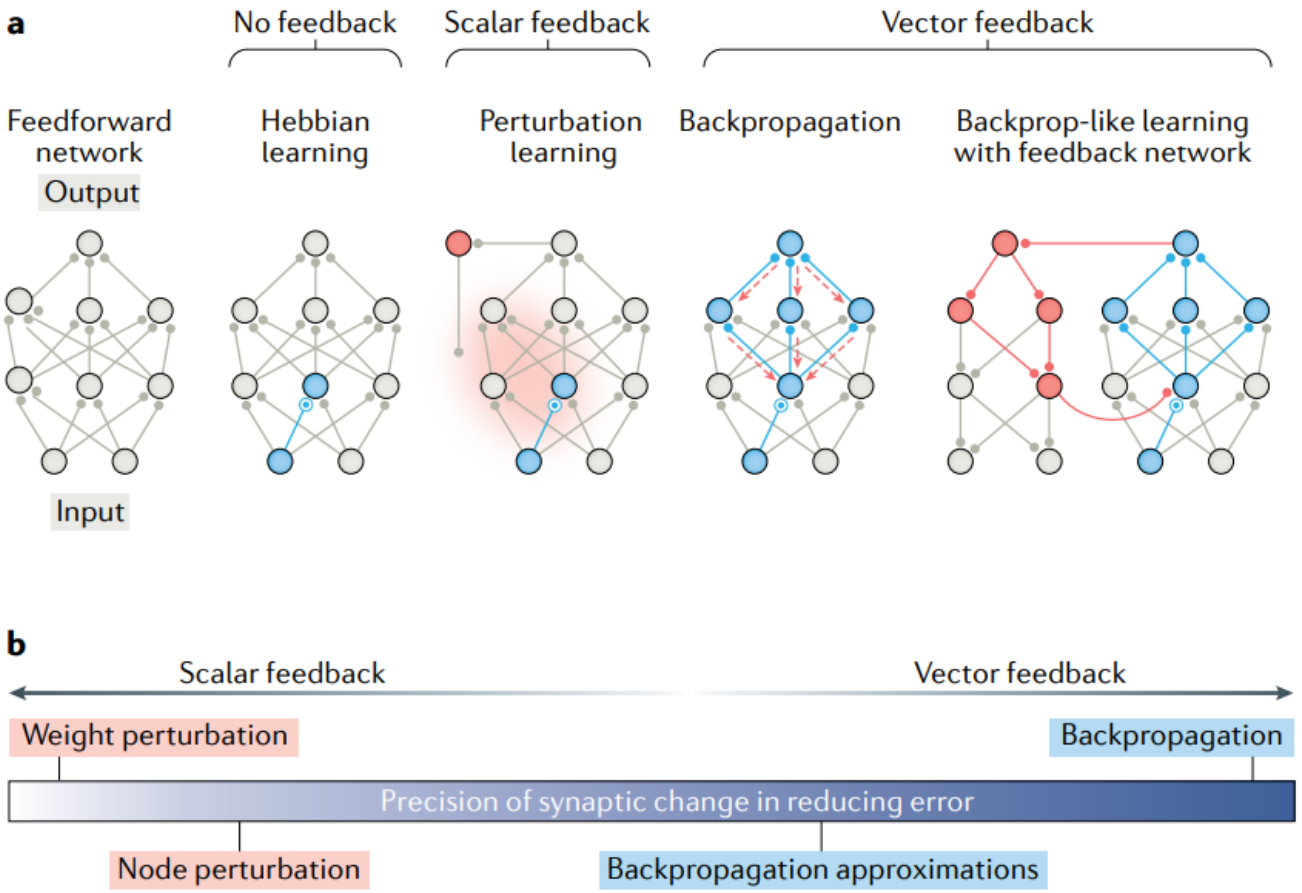
$$CCE = -\sum_{i=1}^{n} y_i \log \hat{y}_i$$

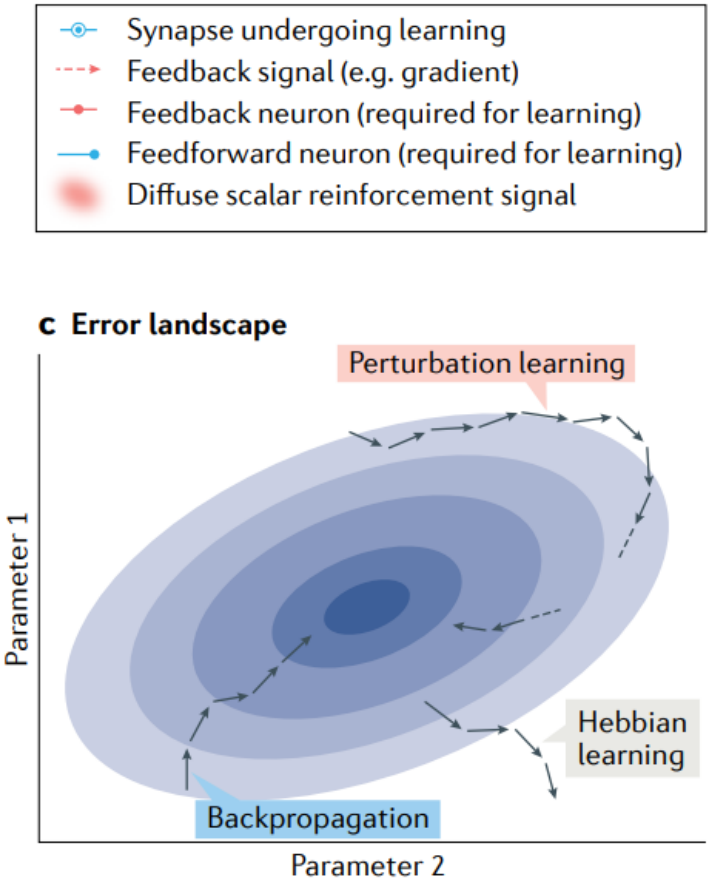
$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Unrestricted © Siemens 2020

Neural Networks

Learning Algorithms





source: Lillicrap *et al.* (2020)

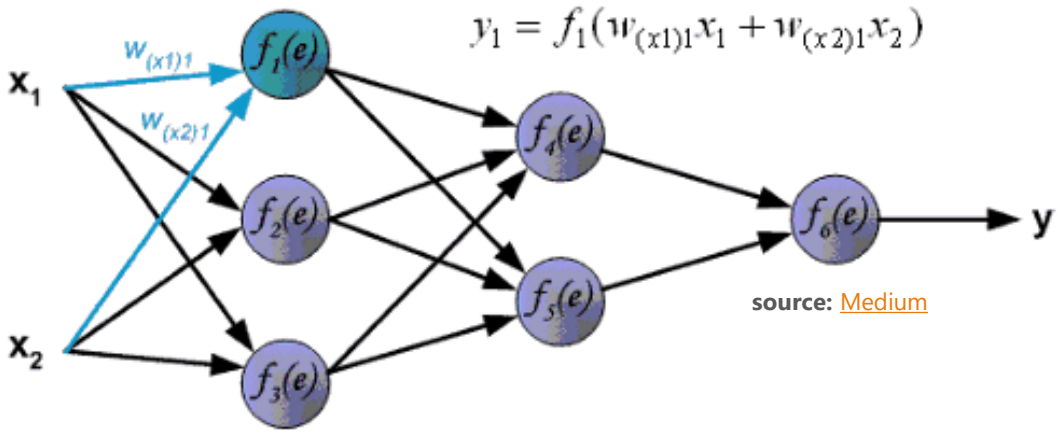
Unrestricted © Siemens 2020

Neural Networks

Backpropagation: Introduction

SIEMENS
Ingenuity for life

FP



source: Medium

Input: dataset D, learning rate η, neural network
Output: gradient of the cost function w.r.t. each weight

Seppo Linnainmaa

ALGORITMIN KUMULATIIVINEN PYÖRISTYSVIRHE

YKSITTÄISTEN PYÖRISTYSVIRHEIDEN TAYLOR-KEHITELMÄNÄ

Pro gradu-tutkielma · ohjaaja professori M.Tienari

source: Jürgen Schmidhuber

Yes you should understand backprop

Andrej Karpathy Follow
Dec 19, 2016 · 7 min read

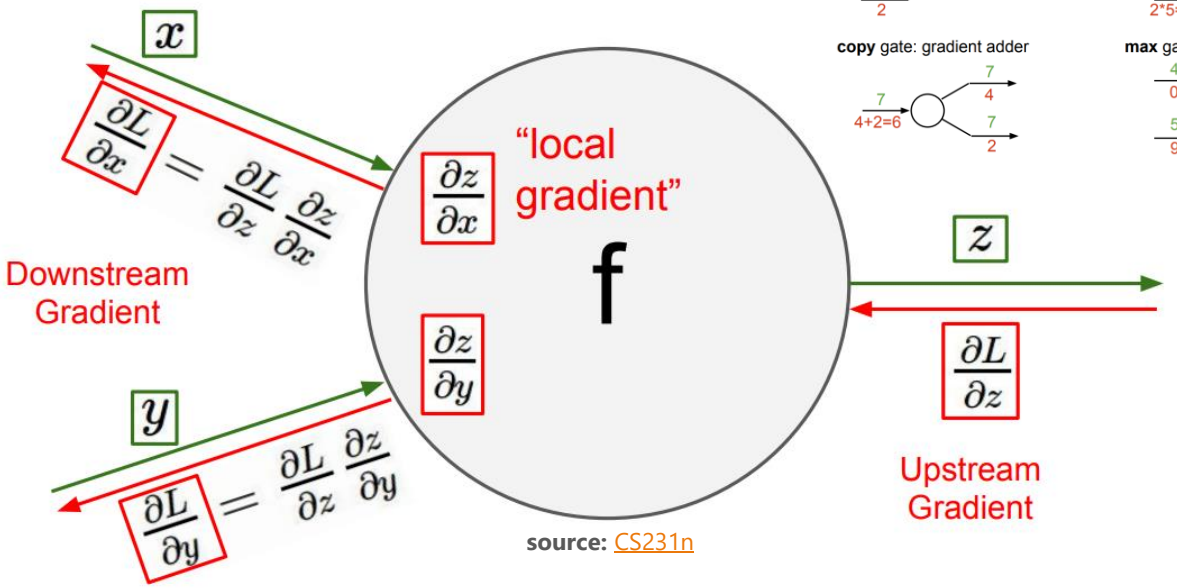
"The problem with Backpropagation is that it is a **leaky abstraction**."

source: Medium

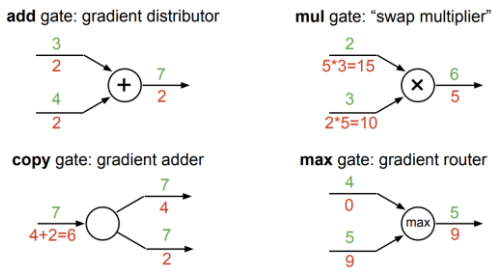
Homework: watch **3Blue1Brown**'s video on backpropagation

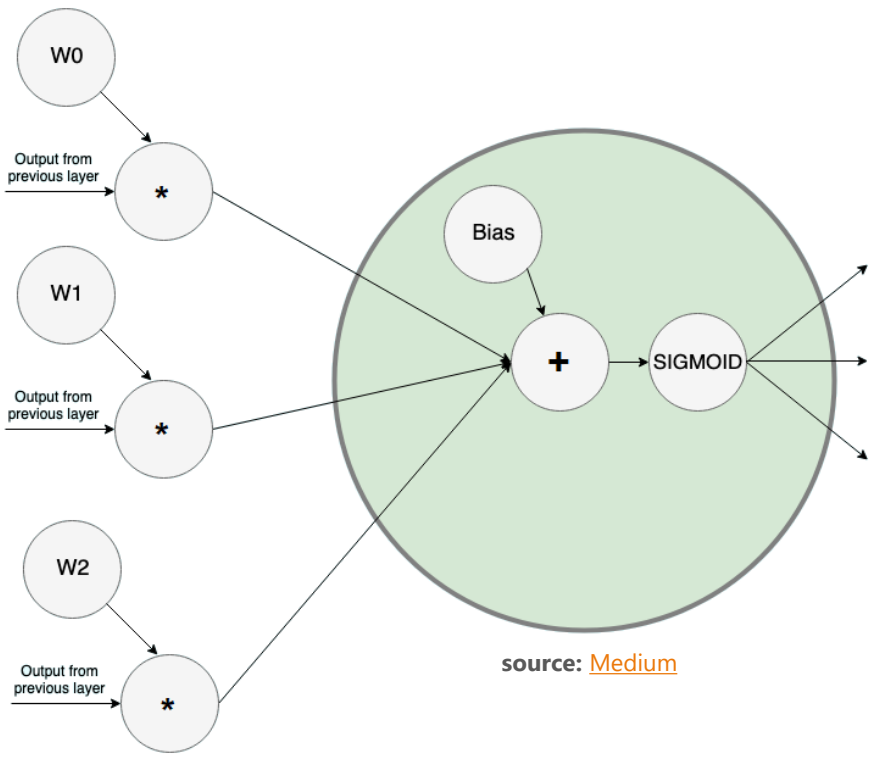
Intermezzo

Backpropagation, Computational Graphs and the Chain Rule



source: CS231n





source: Medium

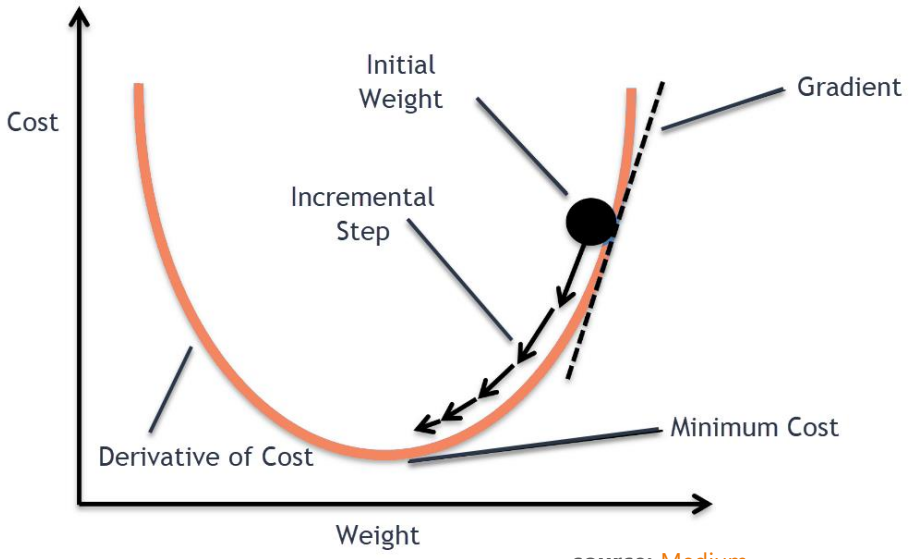
*"(...) back-propagation is often misunderstood as being specific to multi-layer neural networks, but in principle it can compute derivatives of **any** function"*

source: Goodfellow, Bengio & Courville (2016)

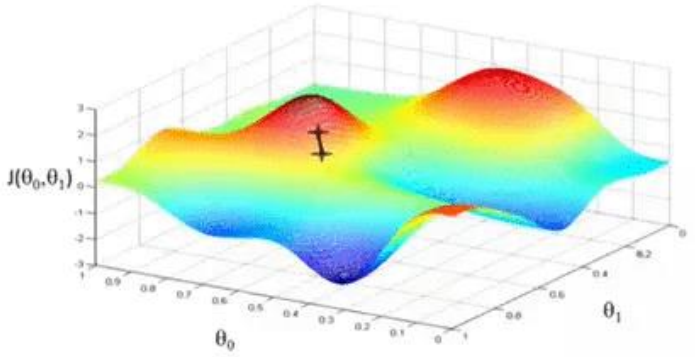
Unrestricted © Siemens 2020

Neural Networks

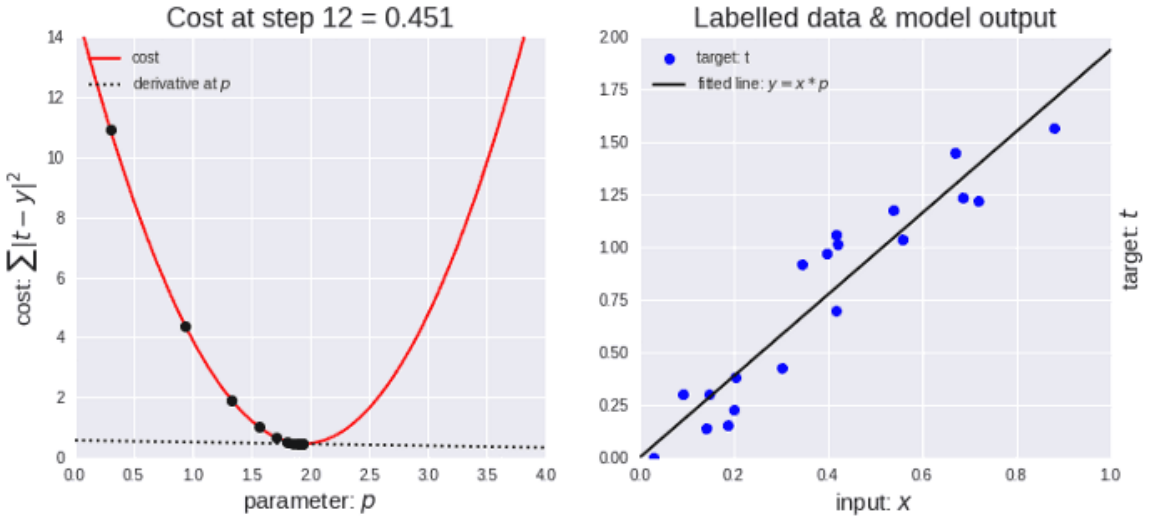
Gradient Descent: Introduction



source: Medium



source: Andrew Ng

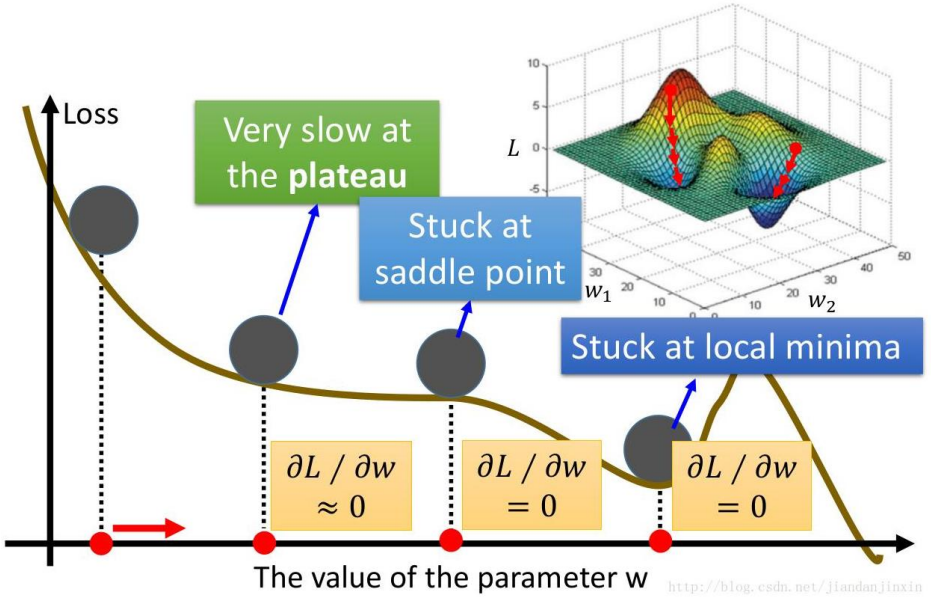


source: Medium

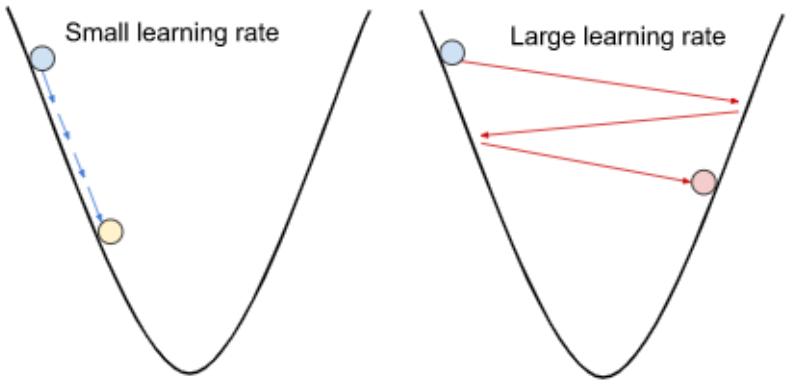
Unrestricted © Siemens 2020

Neural Networks

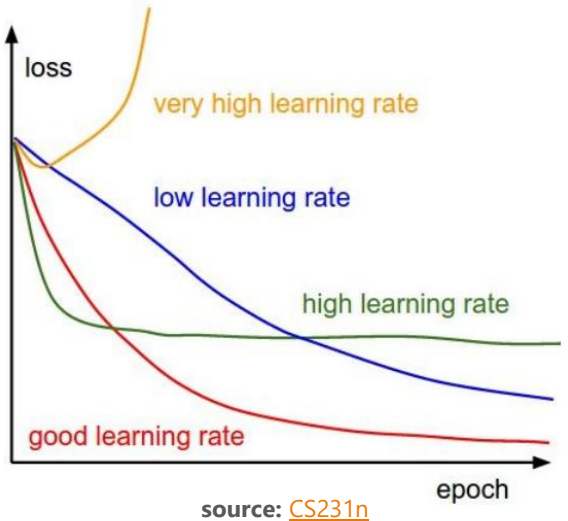
Gradient Descent: Common Problems



source: *adapted from* ProgrammerSought



source: Imperial College



source: CS231n

Unrestricted © Siemens 2020

Neural Networks

Gradient Descent: Variants and Extensions

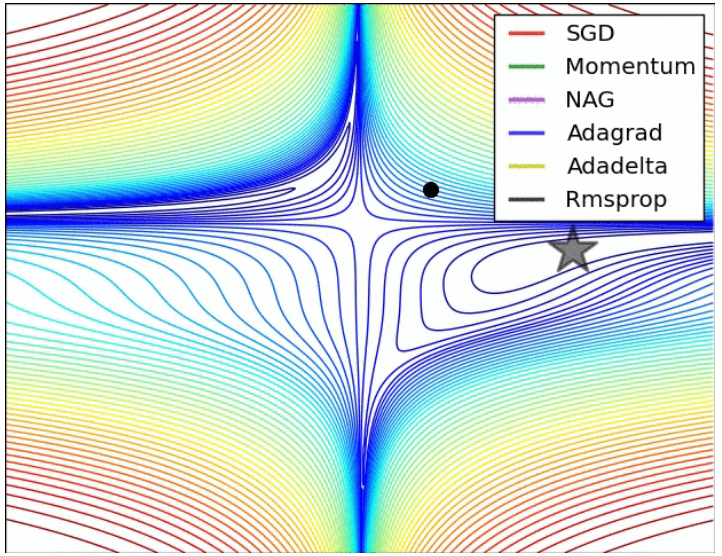
SIEMENS
Ingenuity for life

Adam Optimization Algorithm

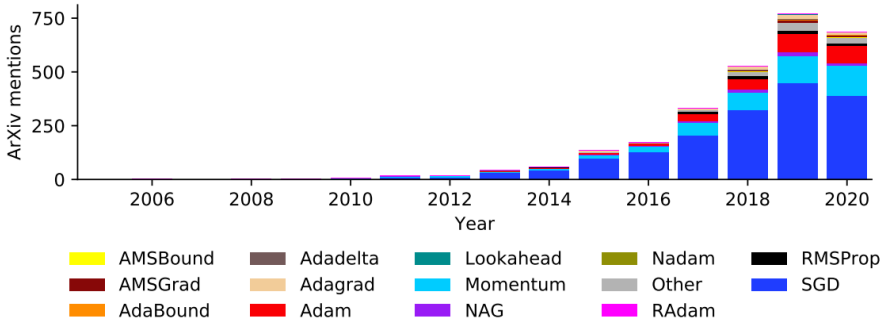
Require: Objective function $f(\theta)$, initial parameters θ_0, stepsize hyperparameter α, exponential decay rates β_1, β_2 for moment estimates, tolerance parameter $\lambda > 0$ for numerical stability, and decision rule for declaring convergence of θ_t in scheme.

procedure ADAM(f, θ_0 ; α, β_1, β_2)
  m_0, v_0, $t \leftarrow [0, 0, 0]$ # Initialize moment estimates
  # and timestep to zero
  # Begin optimization procedure
  while θ_t has not converged **do**
    $t \leftarrow t + 1$ # Update timestep
    $g_t \leftarrow \nabla_\theta f_t(\theta_{t-1})$ # Compute gradient of objective
    $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ # Update first moment estimate
    $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot (g_t \odot g_t)$ # Update second moment estimate
    $\widehat{m}_t \leftarrow m_t/(1 - \beta_1^t)$ # Create unbiased estimate $\widehat{m}_t$
    $\widehat{v}_t \leftarrow v_t/(1 - \beta_2^t)$ # Create unbiased estimate $\widehat{v}_t$
    $\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \widehat{m}_t/(\sqrt{\widehat{v}_t} + \lambda)$ # Update objective parameters
  return θ_t # Return final parameters

source: *adapted from* Nick Winovich



source: Sebastian Ruder

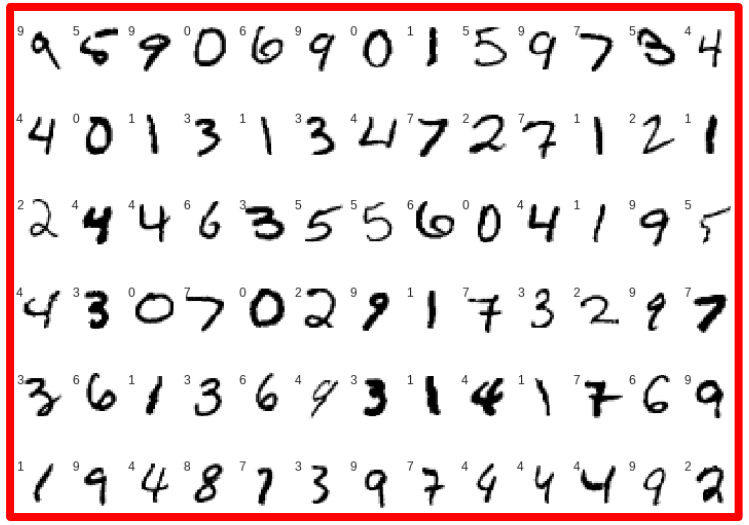


source: Schmidt, Schneider & Hennig (2020)

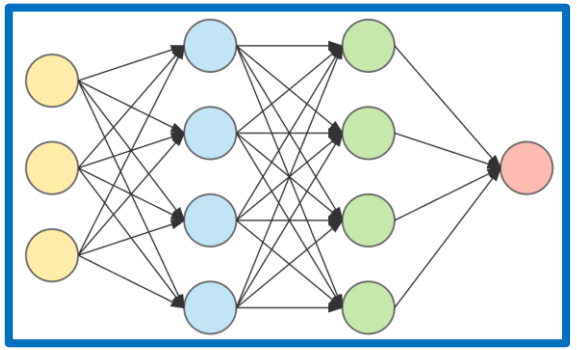
Unrestricted © Siemens 2020

Neural Networks

MNIST Digit Classification



Data + Labels



Neural Network

0
1
2
3
4
5
6
7
8
9

source: *adapted from* javatpoint

Neural Networks

MNIST Digit Classification

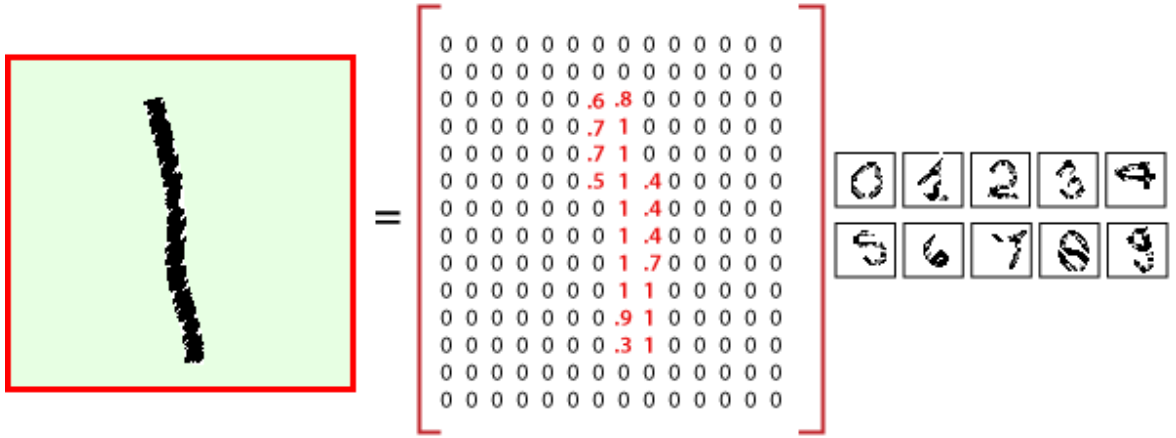
Initial Setup

```python
import numpy as np
from tensorflow import keras
from tensorflow.keras import layers
```

Data Preparation

```python
# Model / data parameters
num_classes = 10
input_shape = (28, 28, 1)

# the data, split between train and test sets
(x_train, y_train), (x_test, y_test) = keras.datasets.mnist.load_data()

# Scale images to the [0, 1] range
x_train = x_train.astype("float32") / 255
x_test = x_test.astype("float32") / 255
# Make sure images have shape (28, 28, 1)
x_train = np.expand_dims(x_train, -1)
x_test = np.expand_dims(x_test, -1)
print("x_train shape:", x_train.shape)
print(x_train.shape[0], "train samples")
print(x_test.shape[0], "test samples")

# convert class vectors to binary class matrices
y_train = keras.utils.to_categorical(y_train, num_classes)
y_test = keras.utils.to_categorical(y_test, num_classes)
```

source: Keras



Unrestricted © Siemens 2020

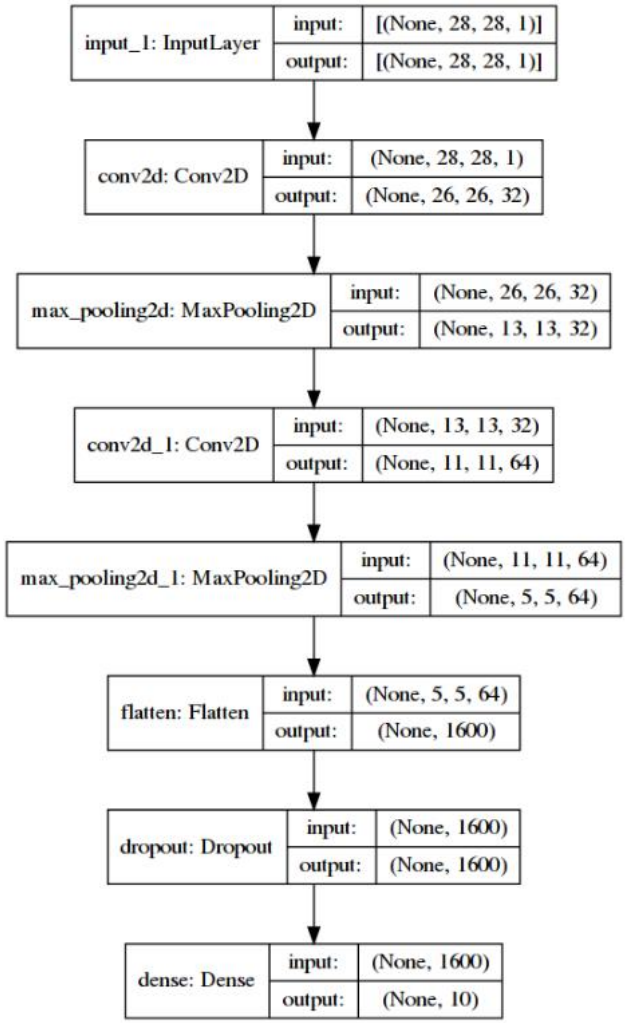
Neural Networks

MNIST Digit Classification

Model Building

```
model = keras.Sequential(
    [
        keras.Input(shape=input_shape),
        layers.Conv2D(32, kernel_size=(3, 3), activation="relu"),
        layers.MaxPooling2D(pool_size=(2, 2)),
        layers.Conv2D(64, kernel_size=(3, 3), activation="relu"),
        layers.MaxPooling2D(pool_size=(2, 2)),
        layers.Flatten(),
        layers.Dropout(0.5),
        layers.Dense(num_classes, activation="softmax"),
    ]
)

model.summary()
```

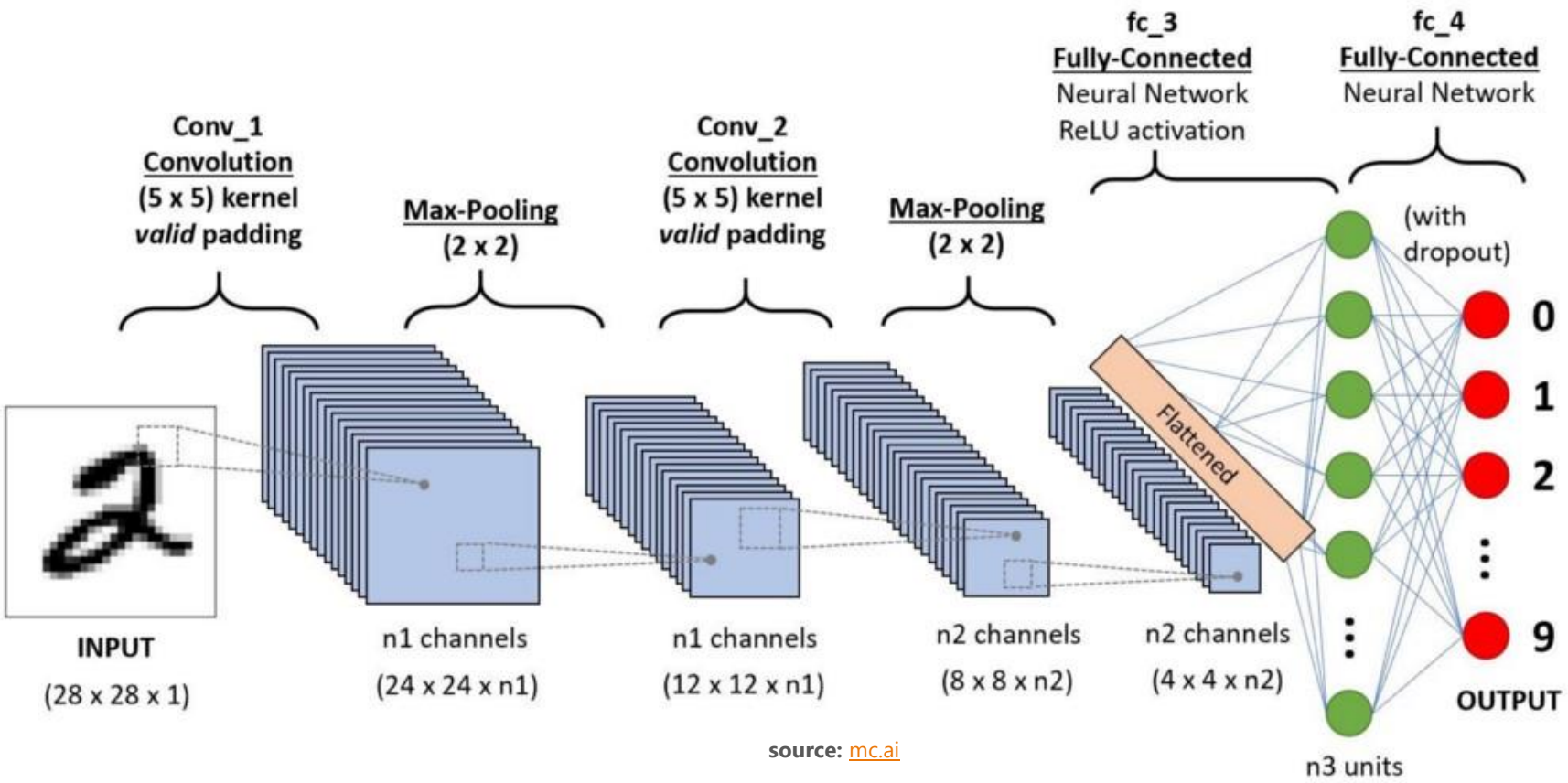
source: Keras



Unrestricted © Siemens 2020

Neural Networks

MNIST Digit Classification



source: mc.ai

Unrestricted © Siemens 2020

Neural Networks

Ingredients

Training

```
batch_size = 128
epochs = 15

model.compile(loss="categorical_crossentropy", optimizer="adam", metrics=["accuracy"])

model.fit(x_train, y_train, batch_size=batch_size, epochs=epochs, validation_split=0.1)
```

```
Epoch 1/15
422/422 [==============================] - 13s 30ms/step - loss: 0.3697 - accuracy: 0.8900 - val_loss: 0.0869 - val_accuracy: 0.9768
Epoch 2/15
422/422 [==============================] - 13s 32ms/step - loss: 0.1178 - accuracy: 0.9637 - val_loss: 0.0612 - val_accuracy: 0.9835
Epoch 3/15
422/422 [==============================] - 15s 35ms/step - loss: 0.0865 - accuracy: 0.9730 - val_loss: 0.0501 - val_accuracy: 0.9857
Epoch 4/15
422/422 [==============================] - 15s 36ms/step - loss: 0.0723 - accuracy: 0.9779 - val_loss: 0.0456 - val_accuracy: 0.9877
Epoch 5/15
422/422 [==============================] - 16s 39ms/step - loss: 0.0653 - accuracy: 0.9793 - val_loss: 0.0394 - val_accuracy: 0.9892
Epoch 6/15
422/422 [==============================] - 17s 41ms/step - loss: 0.0562 - accuracy: 0.9823 - val_loss: 0.0409 - val_accuracy: 0.9875
Epoch 7/15
422/422 [==============================] - 17s 41ms/step - loss: 0.0527 - accuracy: 0.9841 - val_loss: 0.0397 - val_accuracy: 0.9895
Epoch 8/15
422/422 [==============================] - 17s 40ms/step - loss: 0.0485 - accuracy: 0.9848 - val_loss: 0.0367 - val_accuracy: 0.9897
Epoch 9/15
422/422 [==============================] - 17s 39ms/step - loss: 0.0464 - accuracy: 0.9851 - val_loss: 0.0358 - val_accuracy: 0.9888
Epoch 10/15
422/422 [==============================] - 17s 40ms/step - loss: 0.0416 - accuracy: 0.9868 - val_loss: 0.0322 - val_accuracy: 0.9910
Epoch 11/15
422/422 [==============================] - 17s 41ms/step - loss: 0.0417 - accuracy: 0.9864 - val_loss: 0.0306 - val_accuracy: 0.9925
Epoch 12/15
197/422 [=============>...............] - ETA: 9s - loss: 0.0418 - accuracy: 0.9859
```

Testing

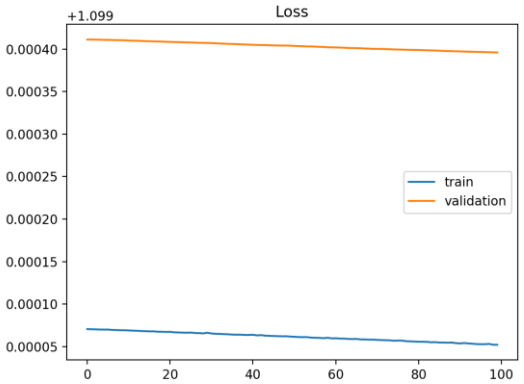
```
score = model.evaluate(x_test, y_test, verbose=0)
print("Test loss:", score[0])
print("Test accuracy:", score[1])
```

source: Keras

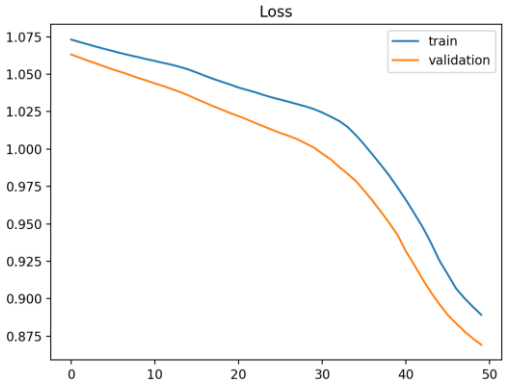
Unrestricted © Siemens 2020

Neural Networks

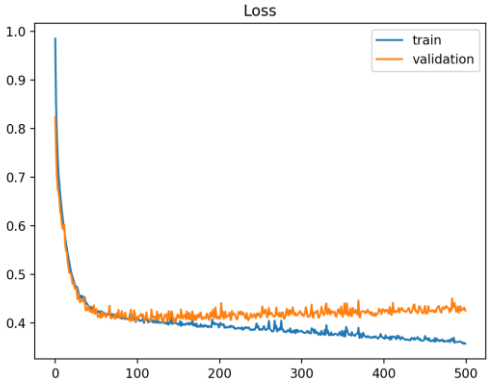
Diagnosing Problems with Learning Curves



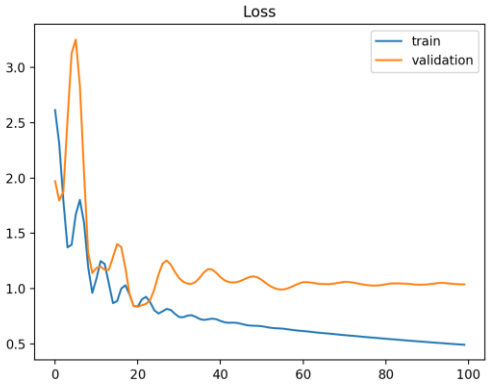
Underfitting
(insufficient capacity)



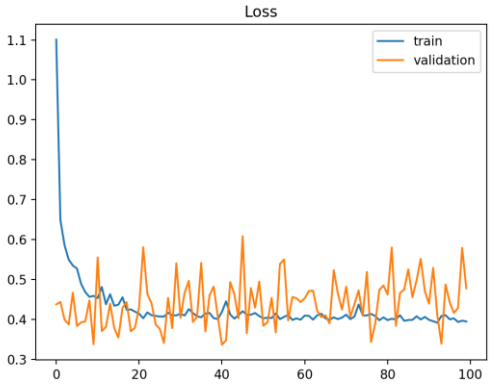
Underfitting
(requires further training)



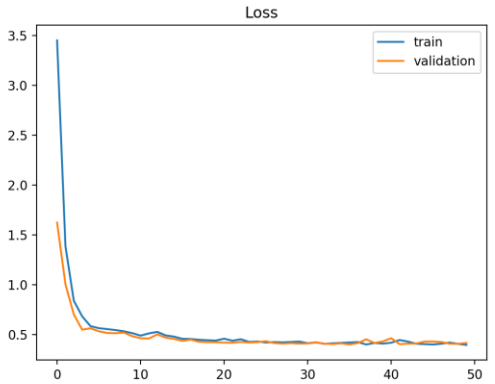
Overfitting



Unrepresentative Dataset
(train)



Unrepresentative Dataset
(validation)



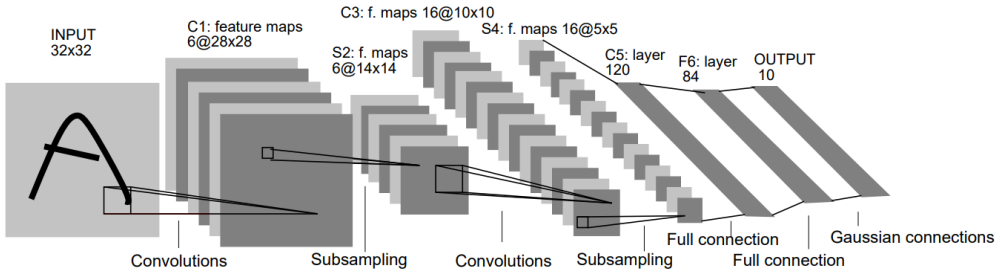
Good Fit

source: machinelearningmastery

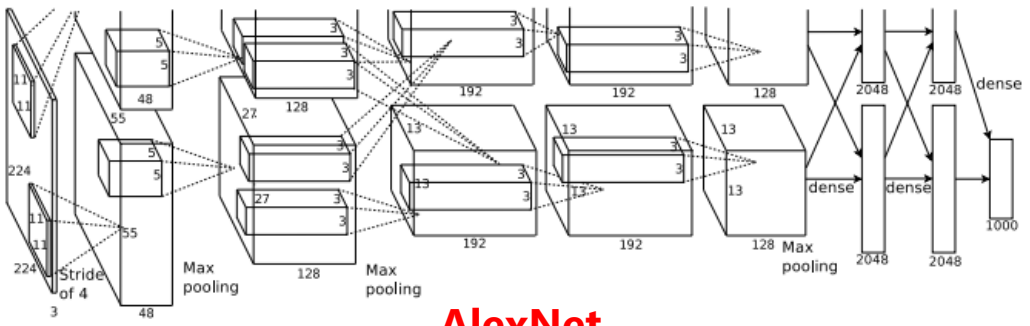
Unrestricted © Siemens 2020

Neural Networks

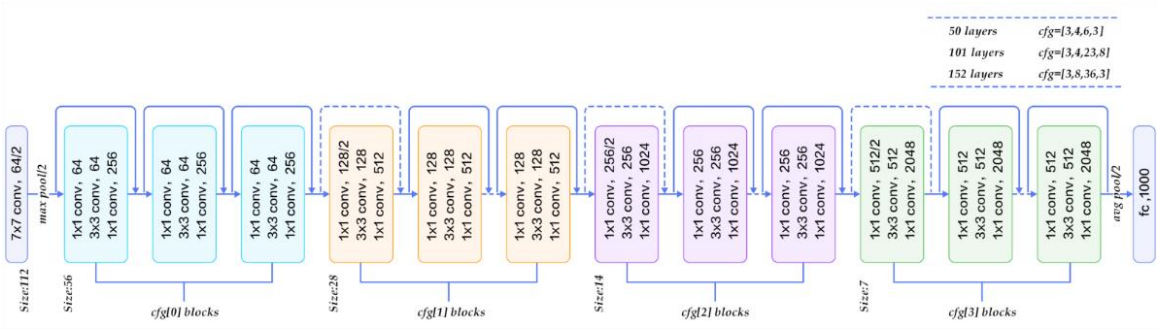
Architectures



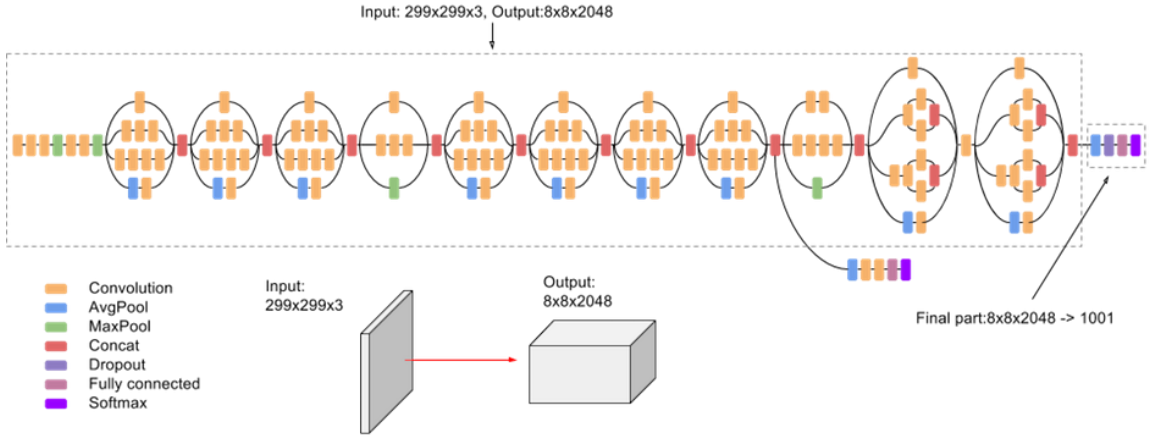
LeNet



AlexNet



ResNet

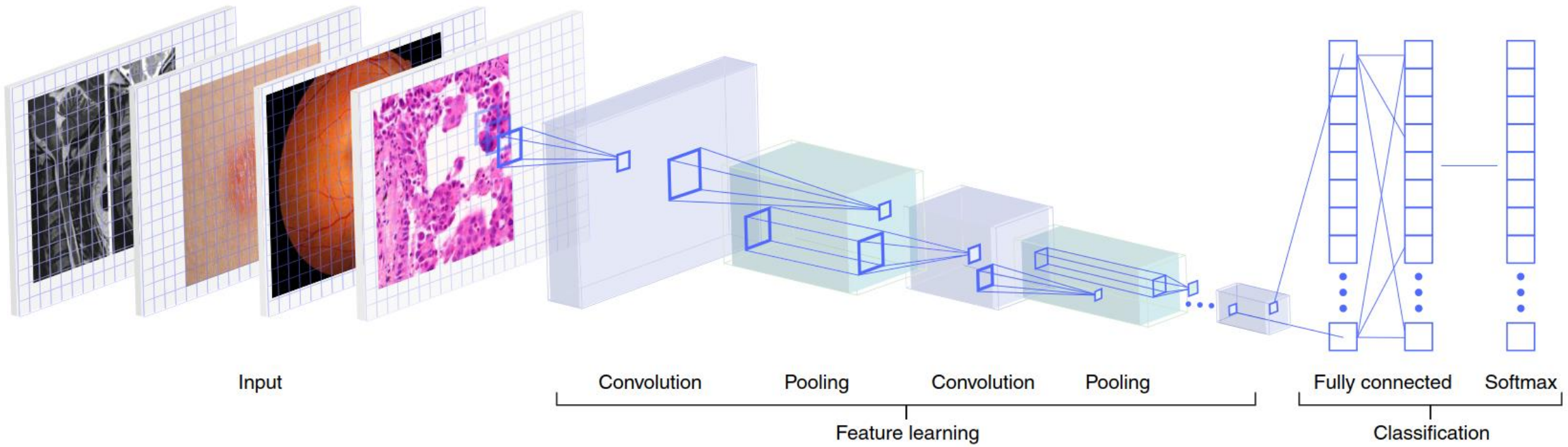


GoogleNet/Inception

Unrestricted © Siemens 2020

Neural Networks

Architectures



source: Esteva *et al.* (2019)

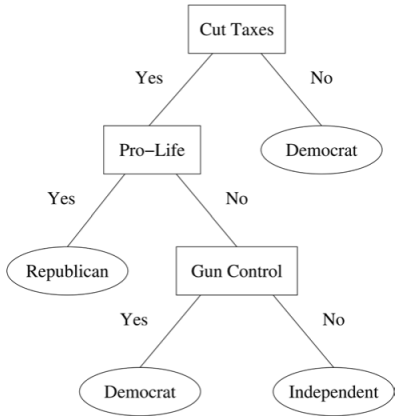
Unrestricted © Siemens 2020

“The Five Tribes”

SIEMENS
Ingenuity for life

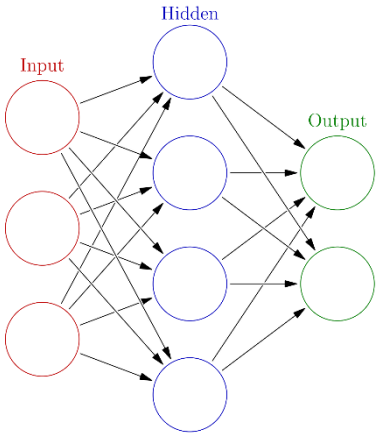
Symbolists

All intelligence can be reduced to symbol manipulation



Connectionists

Learning is what the brain does, and what we need to do is reverse engineer it



Bayesians

All learned knowledge is uncertain, and learning itself is a form of uncertain inference

Prior Probability

Likelihood of the evidence ‘E’ if the Hypothesis ‘H’ is true

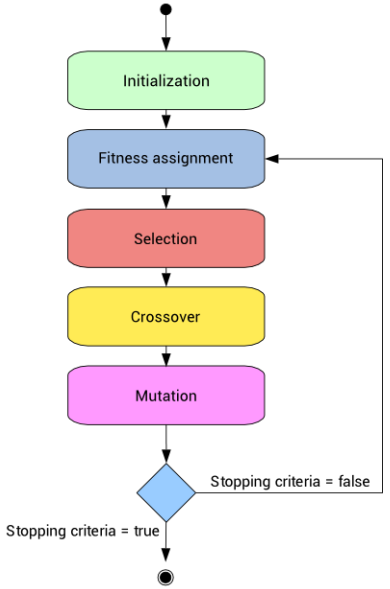
$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)}$$

Posterior Probability of ‘H’ given the evidence

Priori probability that the evidence itself is true

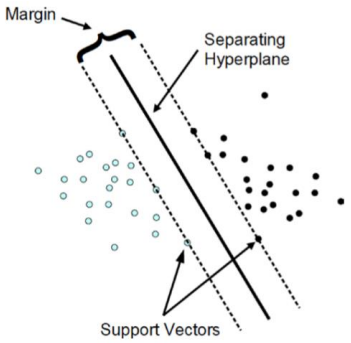
Evolutionaries

The mother of all learning is natural selection



Analogizers

The key to learning is recognizing similarities between situations



source: Domingos (2015)

Unrestricted © Siemens 2020

Caveats
TANSTAAFL

No Free Lunch Theorem (informal statement)

If you make absolutely no assumption about the data, then there is no reason to prefer one model over another

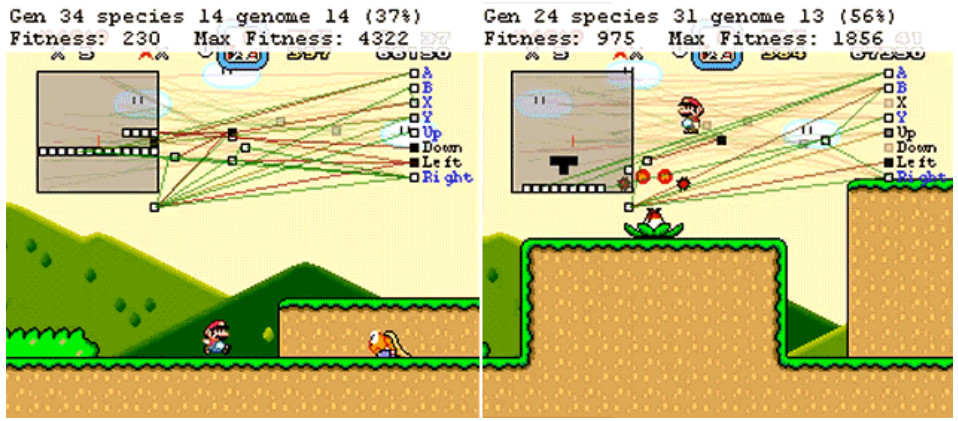
source: Wolpert (1996)

Unrestricted © Siemens 2020

Caveats

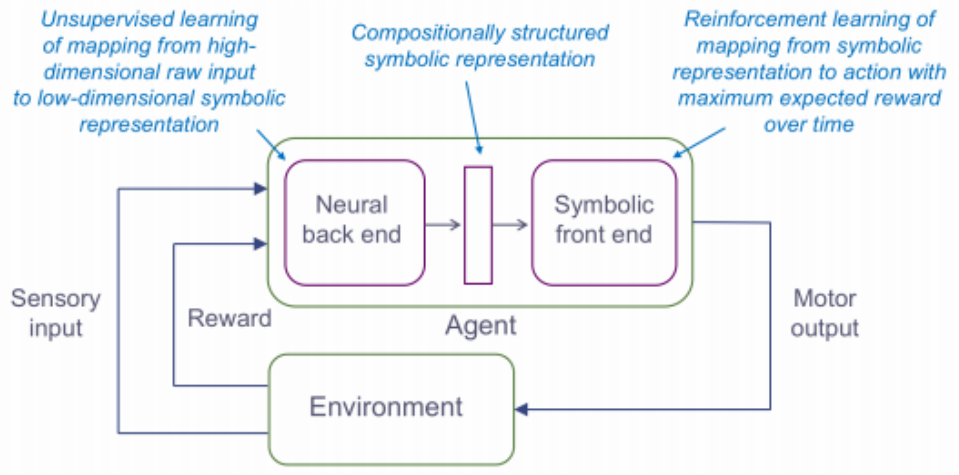
Tribal boundaries are blurry!

NEAT and MarI/O



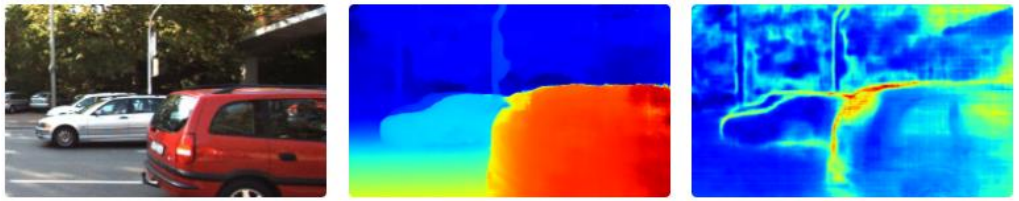
source: SethBling

Deep Symbolic RL



source: Garnelo, Arulkumaran & Shanahan (2016)

Bayesian Deep Learning



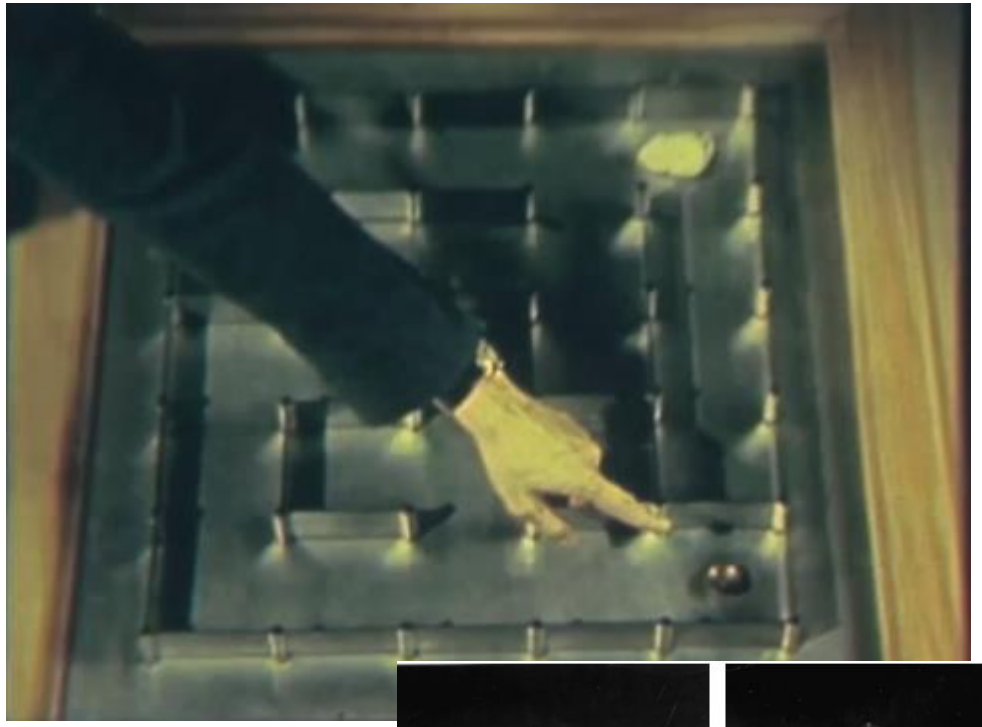
An example of why it is really important to understand uncertainty for depth estimation. The first image is an example input into a Bayesian neural network which estimates depth, as shown by the second image. The third image shows the estimated uncertainty. You can see the model predicts the wrong depth on difficult surfaces, such as the red car's reflective and transparent windows. Thankfully, the Bayesian deep learning model is also aware it is wrong and exhibits increased uncertainty.

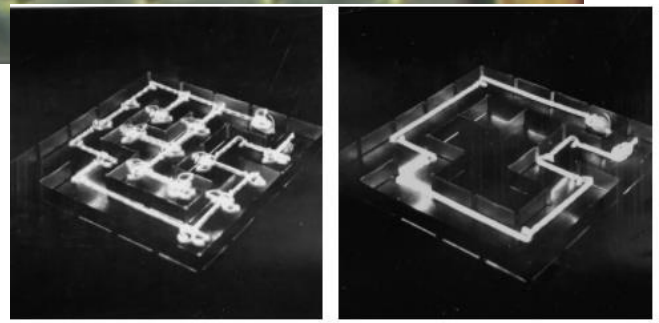
source: Alex Kendall

Unrestricted © Siemens 2020

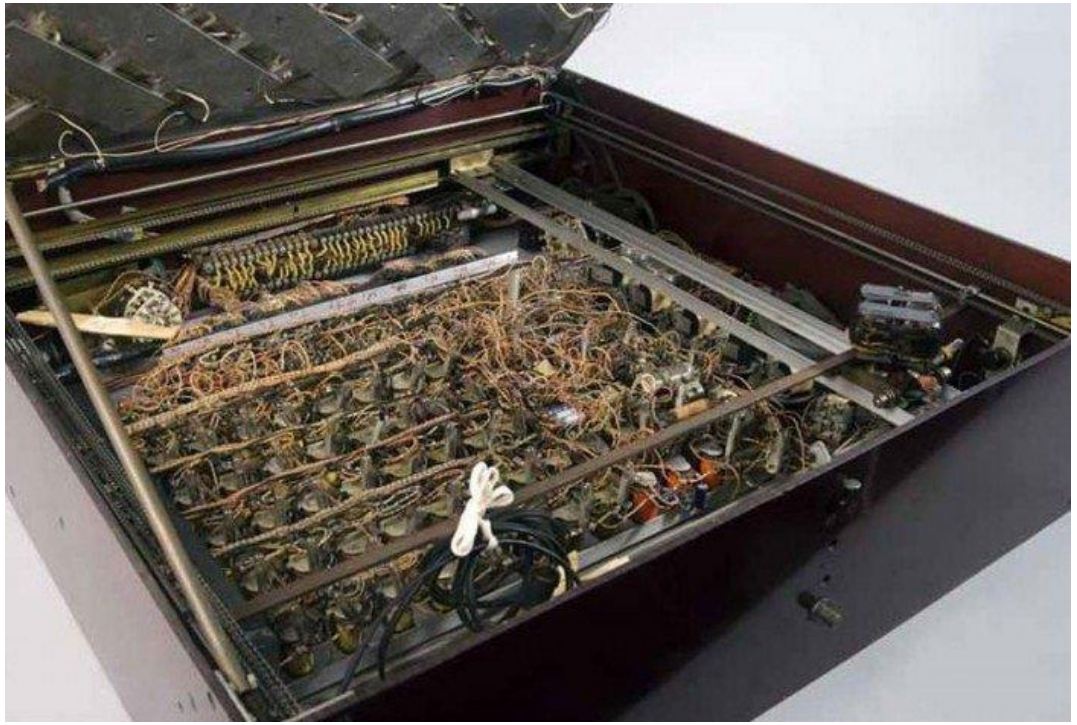
Intermezzo

Shannon's Mighty Mouse





before after

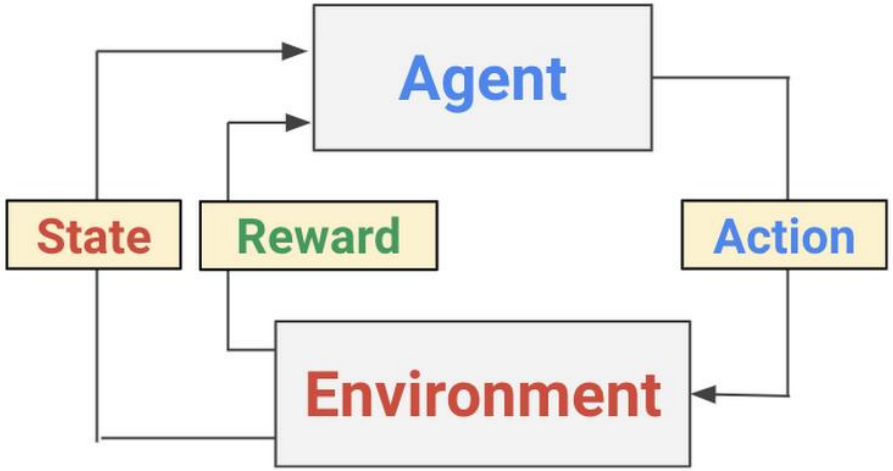


source: MIT museum

Unrestricted © Siemens 2020

Intermezzo

Reinforcement Learning and Atari 2600



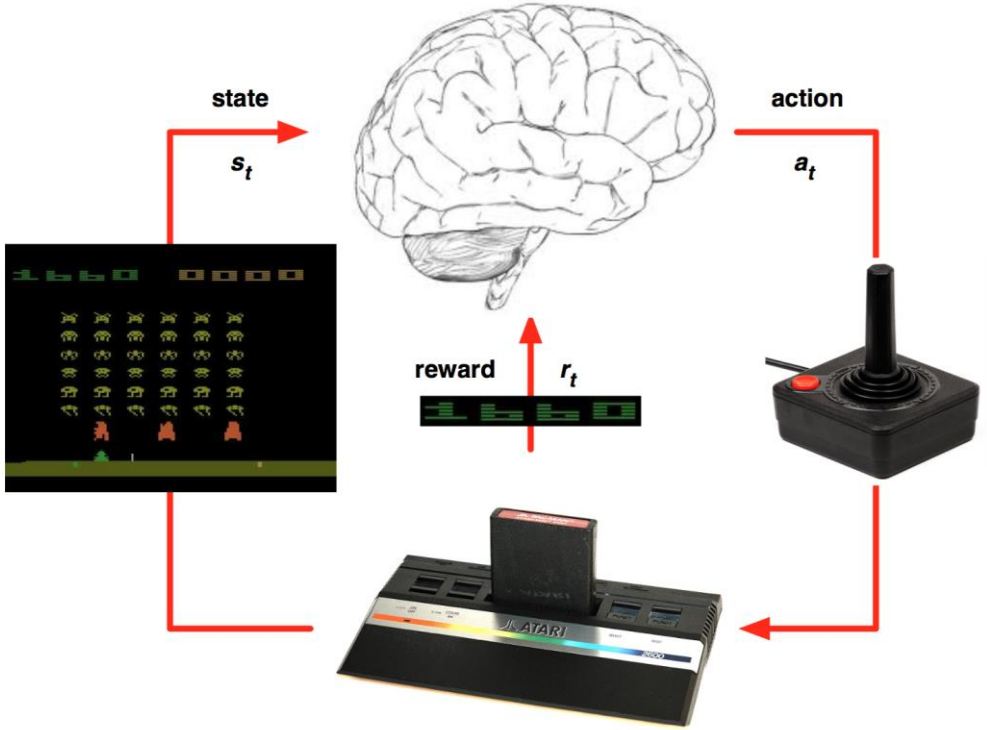
Goal: Maximize the *expected return*

$$v_\pi(s) = \mathbb{E}_\pi \left[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots \mid S_t = s \right]$$

Expected

Reward discounted

Given that state

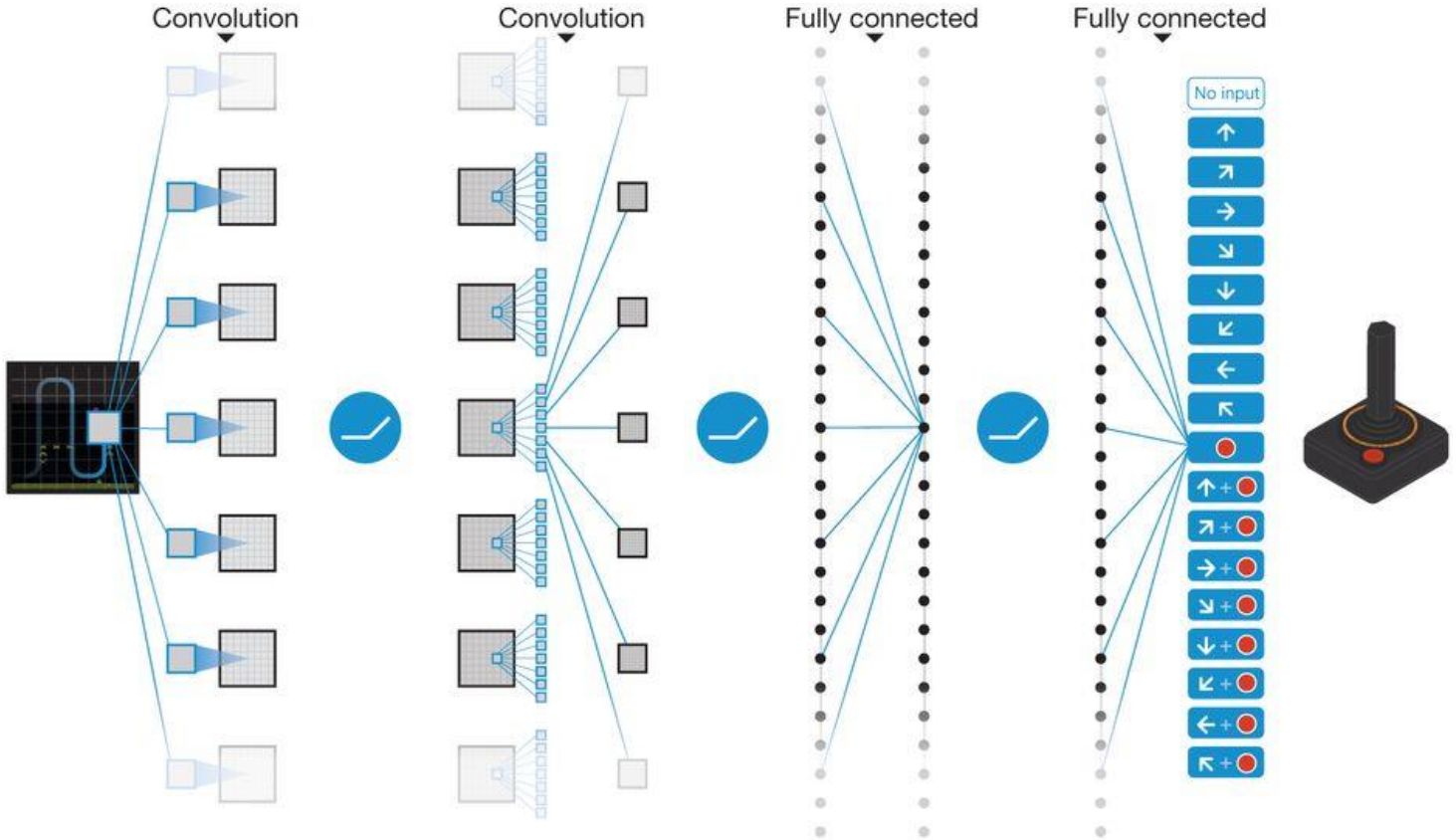


source: ResearchGate

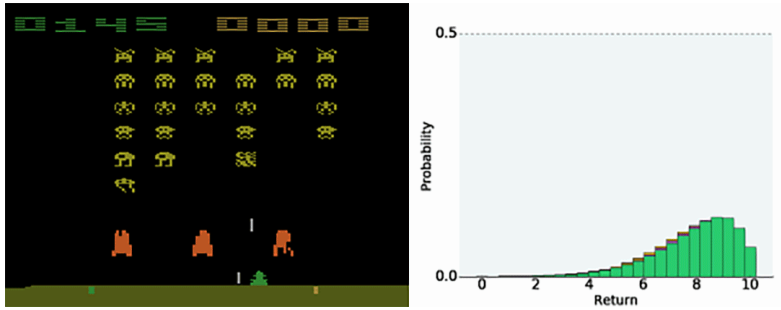
Unrestricted © Siemens 2020

Intermezzo

Deep Reinforcement Learning



source: Mnih *et al.* (2015)



source: DeepMind

$$\text{New } Q(s,a) = Q(s,a) + \alpha\,[R(s,a) + \gamma\, \max Q'(s',a') - Q(s,a)]$$

- New Q Value for that state and the action
- Learning Rate
- Reward for taking that action at that state
- Current Q Values
- Maximum expected future reward given the new state (s') and all possible actions at that new state.
- Discount Rate

Unrestricted © Siemens 2020

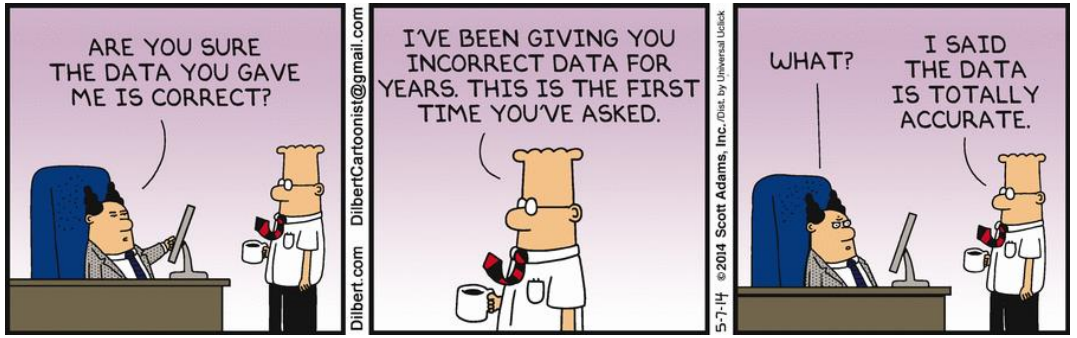
Main Challenges

SIEMENS
Ingenuity for life

"Bad" Algorithms

Learning is about **generalization**, not memorization

"Bad" Data



Garbage in, garbage out!

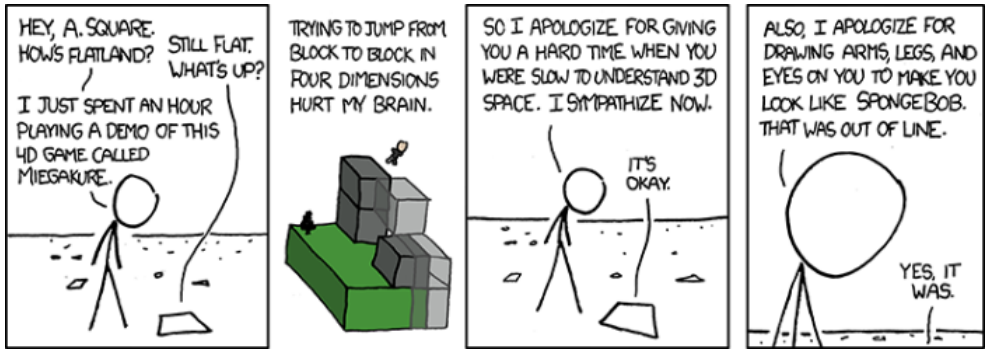
- Insufficient
- Non-representative
- Poor quality
- Irrelevant features

Homework: repeat these mantras *ad nauseam*

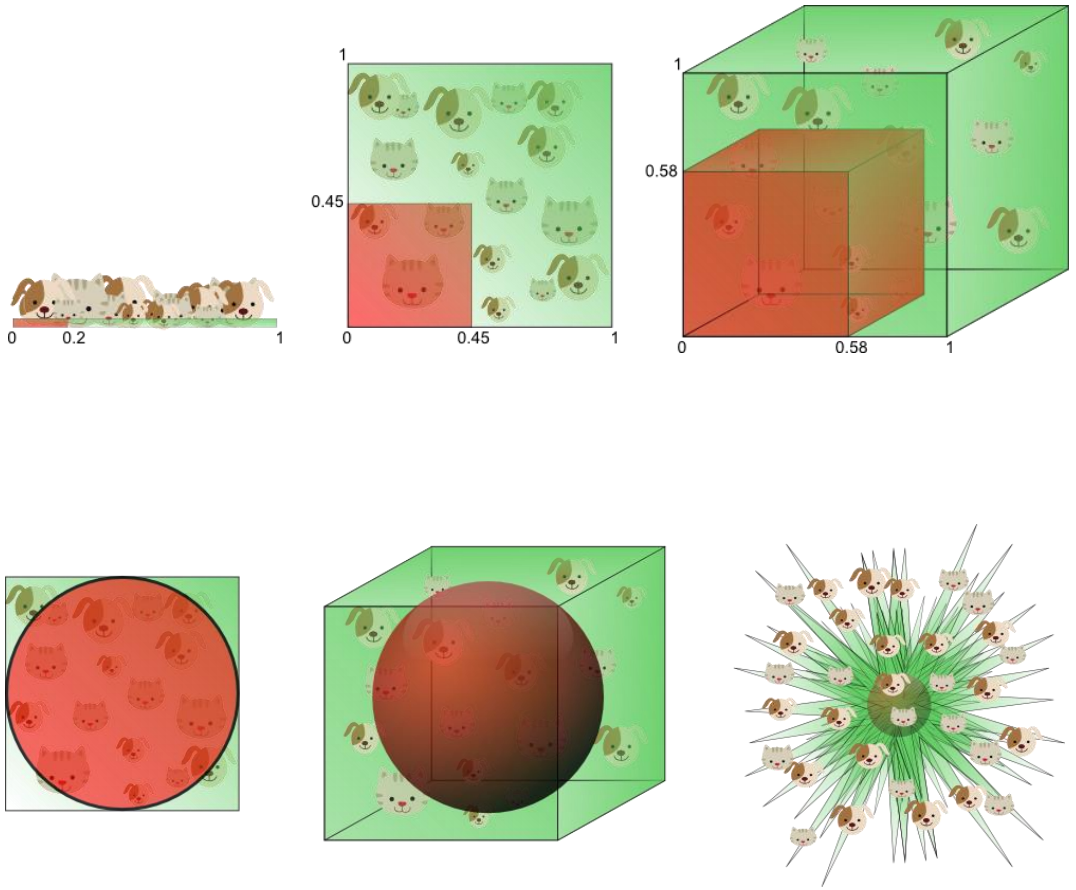
Main Challenges

The Curse of Dimensionality

TL;DR Intuition breaks down in higher dimensions



source: xkcd



source: visiondummy

Homework: read Brian Hayes' An Adventure in the Nth Dimension

Hello kaggle

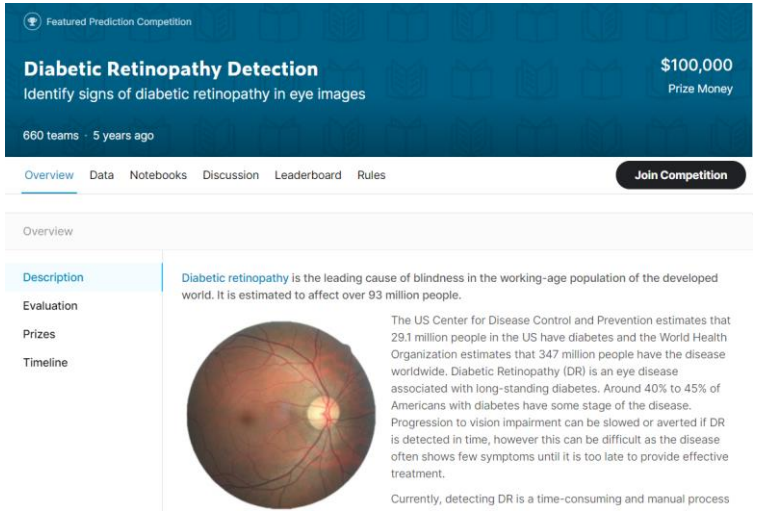
Data Science as a Competition

SIEMENS
Ingenuity for life

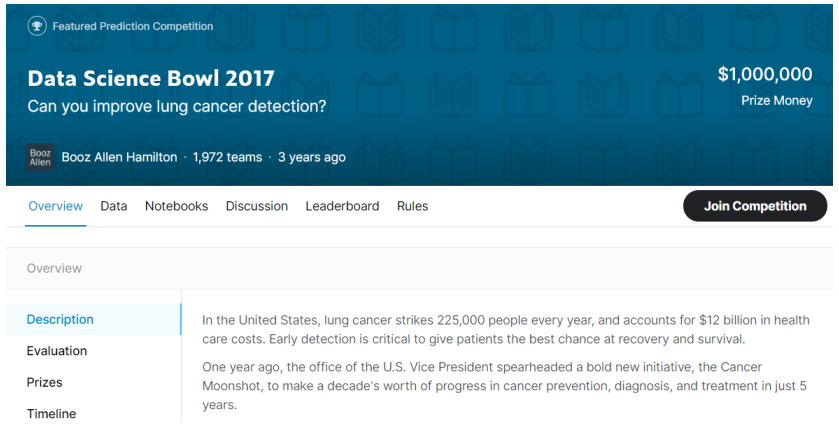
Goal: predict which passangers survived the Titanic shipwreck

Method	Accuracy	AUC	Brier Score
Everyone has a 39% chance of surviving	0.639	0.500	0.232
All females survive, all males do not	0.786	0.578	0.214
Simple decision tree	**0.806**	0.819	**0.139**
Logistic regression	0.789	0.824	0.146
Random forest	0.799	**0.850**	0.148
Support Vector Machine (SVM)	0.782	0.825	0.153
Neural network	0.794	0.828	0.146
K-nearest neighbor	0.774	0.812	0.180

source: *adapted from* Spiegelhalter (2019)



source: Kaggle



source: Kaggle

Unrestricted © Siemens 2020

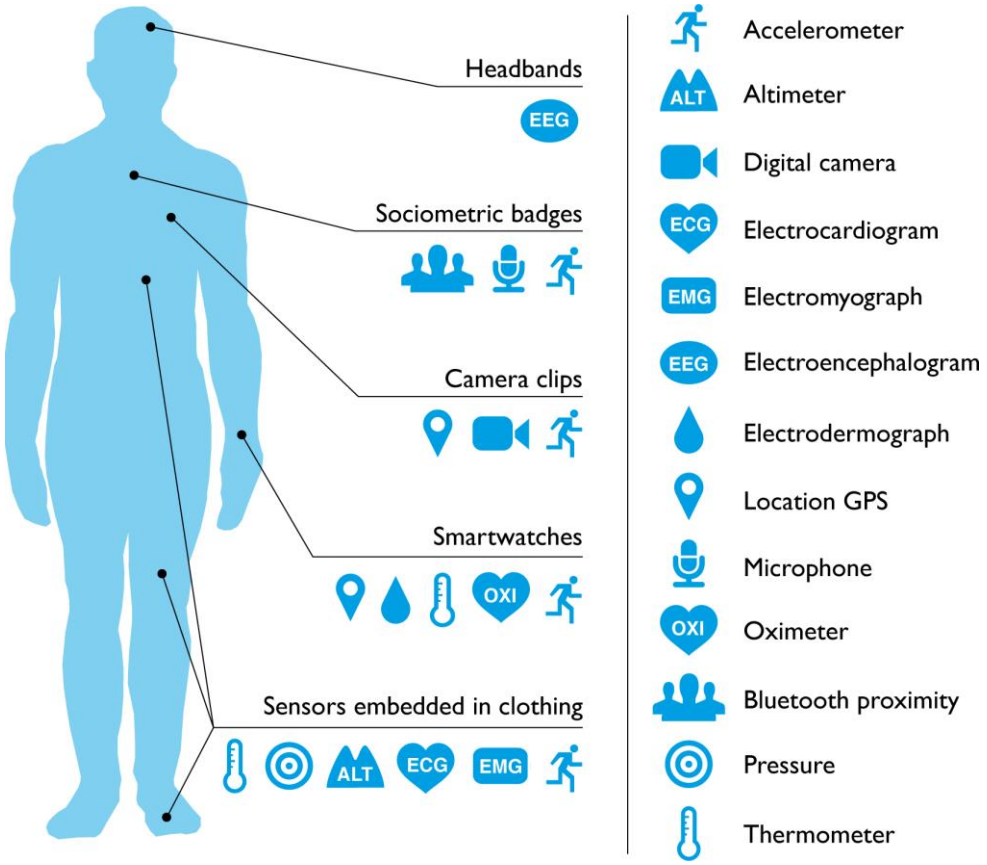
SIEMENS
Ingenuity for life

ML in Medicine and Healthcare

Unrestricted © Siemens 2020

Patient Data

Consumer Health Wearables



source: Piwek *et al.* (2016)

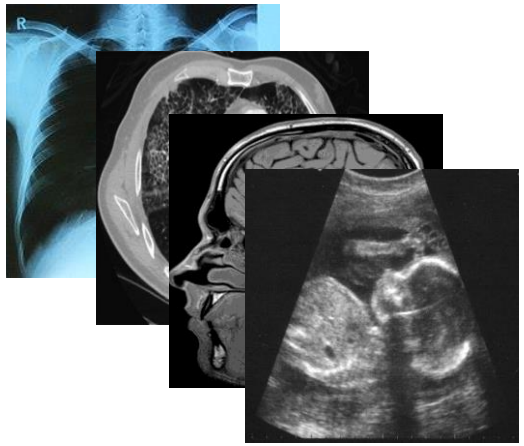
source: Technosoft

Unrestricted © Siemens 2020

Patient Data

Electronic Health Records

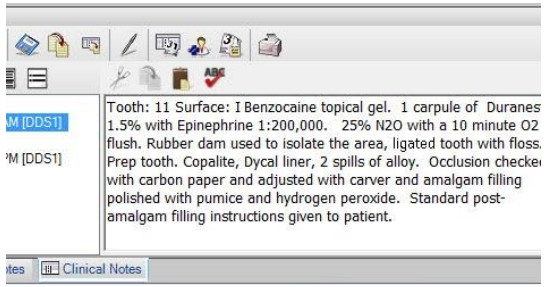
SIEMENS
Ingenuity for life



medical images

Sources: X-ray, CT scan, MRI, Ultrasound

Formats: Analyze, Nifti, Minc, DICOM



clinical notes

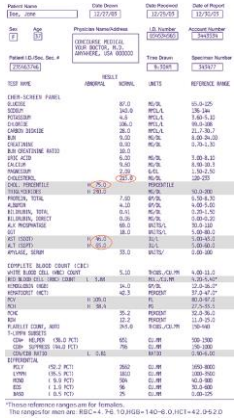
source: Phillips

administrative and billing data
patient demographics
progress notes
medication

vital signs

Sources: EEG, ECG, blood oxygen level

Formats: EDF, EDF+, SCP, HL7



lab results

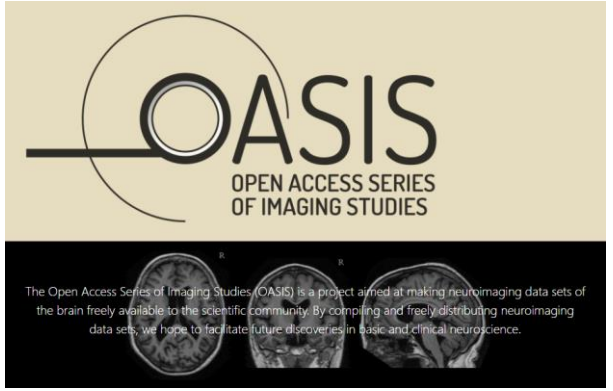
Unrestricted © Siemens 2020

Patient Data

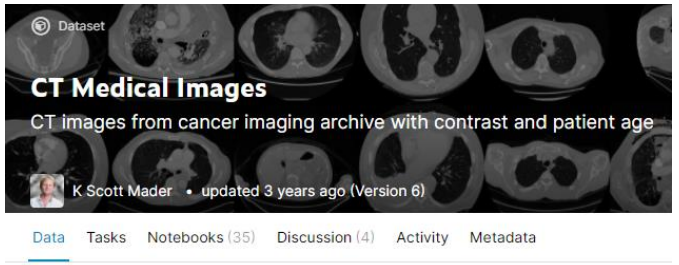
Clinical Datasets

MIMIC is an openly available dataset developed by the MIT Lab for Computational Physiology, comprising deidentified health data associated with ~60,000 intensive care unit admissions. It includes demographics, vital signs, laboratory tests, medications, and more.

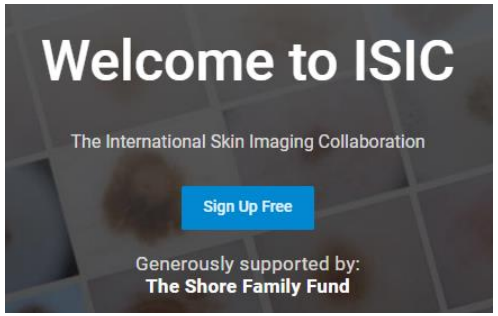
source: MIMIC



source: OASIS



source: Kaggle



source: ISIC

source: UC Irvine

About

This registry exists to help people discover and share datasets that are available via AWS resources. Learn more about sharing data on AWS.

source: AWS

18 Free Life Sciences, Healthcare and Medical Datasets for Machine Learning

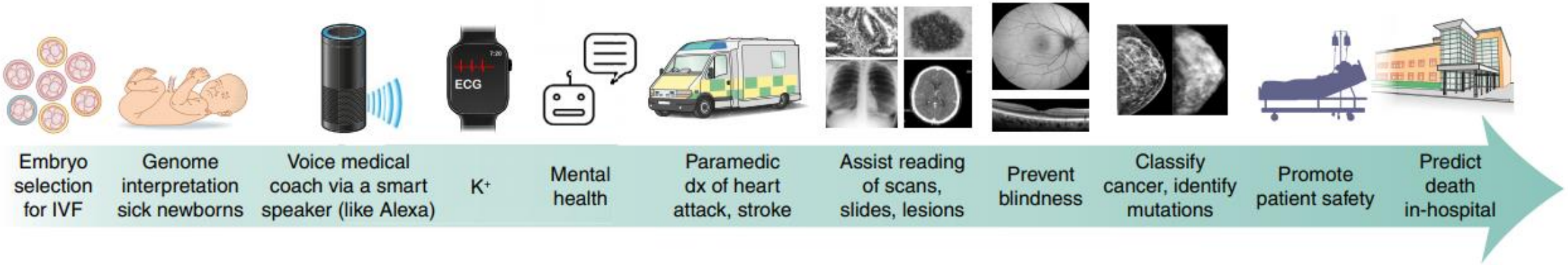
Article by Rei Morikawa | March 12, 2019

source: Lionbridge

Unrestricted © Siemens 2020

AI/ML in Healthcare

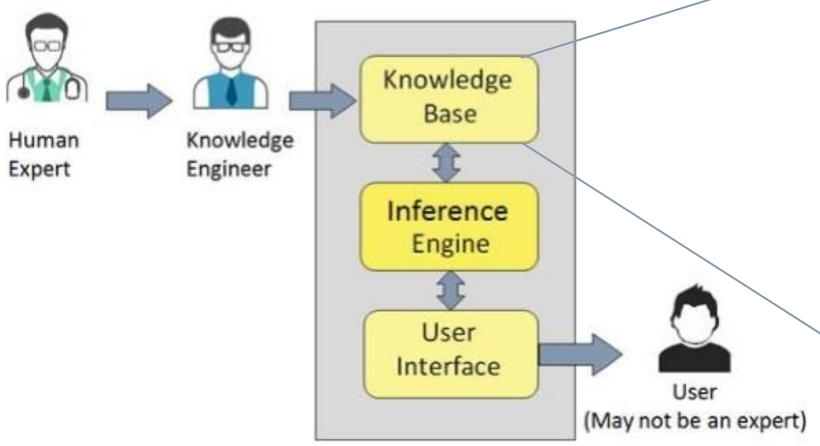
From birth to death



source: Topol (2019)

Unrestricted © Siemens 2020

Expert Systems
Introduction



source: tutorialspoint

```
Rule 75

IF:  1) The site culture is blood, and
     2) The strain of the organism is gram negative, and
     3) The patient is a compromised host

THEN: There is suggestive evidence (.6) that the identity of the organism is pseudomonas

(...)

Rule 107

IF:  1) The strain of the organism is gram negative, and
     2) The morphology of the organism is rod, and
     3) The organism is aerobic

THEN: There is suggestive evidence (.8) that the identity of the organism is Enterobacteriaceae
```

Unrestricted © Siemens 2020

Expert Systems

MYCIN (1970s)

```
(defun emycin (contexts)
  "An Expert System Shell.  Accumulate data for instances of each
  context, and solve for goals.  Then report the findings."
  (clear-db)
  (get-context-data contexts))

(defun get-context-data (contexts)
  "For each context, create an instance and try to find out
  required data.  Then go on to other contexts, depth first,
  and finally ask if there are other instances of this context."
  (unless (null contexts)
    (let* ((context (first contexts))
           (inst (new-instance context)))
      (put-db 'current-rule 'initial)
      (mapc #'find-out (context-initial-data context))
      (put-db 'current-rule 'goal)
      (mapc #'find-out (context-goals context))
      (report-findings context inst)
      (get-context-data (rest contexts))
      (when (y-or-n-p "Is there another ~a?"
                      (context-name context))
        (get-context-data contexts)))))
```

```
(defmacro defrule (number &body body)
  "Define a rule with conditions, a certainty factor, and
  conclusions.  Example: (defrule R001 if ... then .9 ...)"
  (assert (eq (first body) 'if))
  (let* ((then-part (member 'then body))
         (premises (ldiff (rest body) then-part))
         (conclusions (rest2 then-part))
         (cf (second then-part)))
    ;; Do some error checking:
    (check-conditions number premises 'premise)
    (check-conditions number conclusions 'conclusion)
    (when (not (cf-p cf))
      (warn "Rule ~a: Illegal certainty factor: ~a" number cf))
    ;; Now build the rule:
    `(put-rule
      (make-rule :number ',number :cf ,cf :premises ',premises
                 :conclusions ',conclusions))))
```

Inference Engine

+

```
;;; Parameters for patient:
(defparm name patient t "Patient's name: " t read-line)
(defparm sex patient (member male female) "Sex:" t)
(defparm age patient number "Age:" t)
(defparm burn patient (member no mild serious)
  "Is ~a a burn patient?  If so, mild or serious?" t)
(defparm compromised-host patient yes/no
  "Is ~a a compromised host?")

;;; Parameters for culture:
(defparm site culture (member blood)
  "From what site was the specimen for ~a taken?" t)
(defparm days-old culture number
  "How many days ago was this culture (~a) obtained?" t)

;;; Parameters for organism:
(defparm identity organism
  (member pseudomonas klebsiella enterobacteriaceae
          staphylococcus bacteroides streptococcus)
  "Enter the identity (genus) of ~a:" t)
(defparm gram organism (member acid-fast pos neg)
  "The gram stain of ~a:" t)
(defparm morphology organism (member rod coccus)
  "Is ~a a rod or coccus (etc.):")
(defparm aerobicity organism (member aerobic anaerobic))
(defparm growth-conformation organism
  (member chains pairs clumps))
```

Parameters

```
(defrule 73
  if (site culture is blood)
     (gram organism is neg)
     (morphology organism is rod)
     (aerobicity organism is anaerobic)
  then .9
     (identity organism is bacteroides))

(defrule 75
  if (gram organism is neg)
     (morphology organism is rod)
     (compromised-host patient is yes)
  then .6
     (identity organism is pseudomonas))

(defrule 107
  if (gram organism is neg)
     (morphology organism is rod)
     (aerobicity organism is aerobic)
  then .8
     (identity organism is enterobacteriaceae))
```

Knowledge Base

```
------ PATIENT-1 ------
Patient's name:  John Doe
Sex: male
Age: 30
------ CULTURE-1 ------
From what site was the specimen for CULTURE-1 taken? ?
A SITE must be of type (MEMBER BLOOD)
From what site was the specimen for CULTURE-1 taken? blood
How many days ago was this culture (CULTURE-1) obtained? 3
------ ORGANISM-1 ------
Enter the identity (genus) of ORGANISM-1: ?
A IDENTITY must be of type (MEMBER PSEUDOMONAS KLEBSIELLA
ENTEROBACTERIACEAE STAPHYLOCOCCUS BACTEROIDES STREPTOCOCCUS)
Enter the identity (genus) of ORGANISM-1: unknown
The gram stain of ORGANISM-1: ?
A GRAM must be of type (MEMBER ACID-FAST POS NEG)
The gram stain of ORGANISM-1: neg
Is ORGANISM-1 a rod or coccus (etc.): rod
What is the AEROBICITY of ORGANISM-1? why
[Why is the value of AEROBICITY being asked for?]
It is known that:
    1) THE GRAM OF THE ORGANISM IS NEG
    1) THE MORPHOLOGY OF THE ORGANISM IS ROD
Therefore,
Rule 107:
  If
    1) THE AEROBICITY OF THE ORGANISM IS AEROBIC
  Then there is suggestive evidence (0.8) that
    1) THE IDENTITY OF THE ORGANISM IS ENTEROBACTERIACEAE
What is the AEROBICITY of ORGANISM-1? aerobic
Is John Doe a compromised host? yes
Is John Doe a burn patient?  If so, mild or serious? mild
Findings for ORGANISM-1:
 IDENTITY: ENTEROBACTERIACEAE (0.800)   PSEUDOMONAS (0.600)
Is there another ORGANISM? (y/n) n
Is there another CULTURE? (y/n) n
Is there another PATIENT? (y/n) n
```

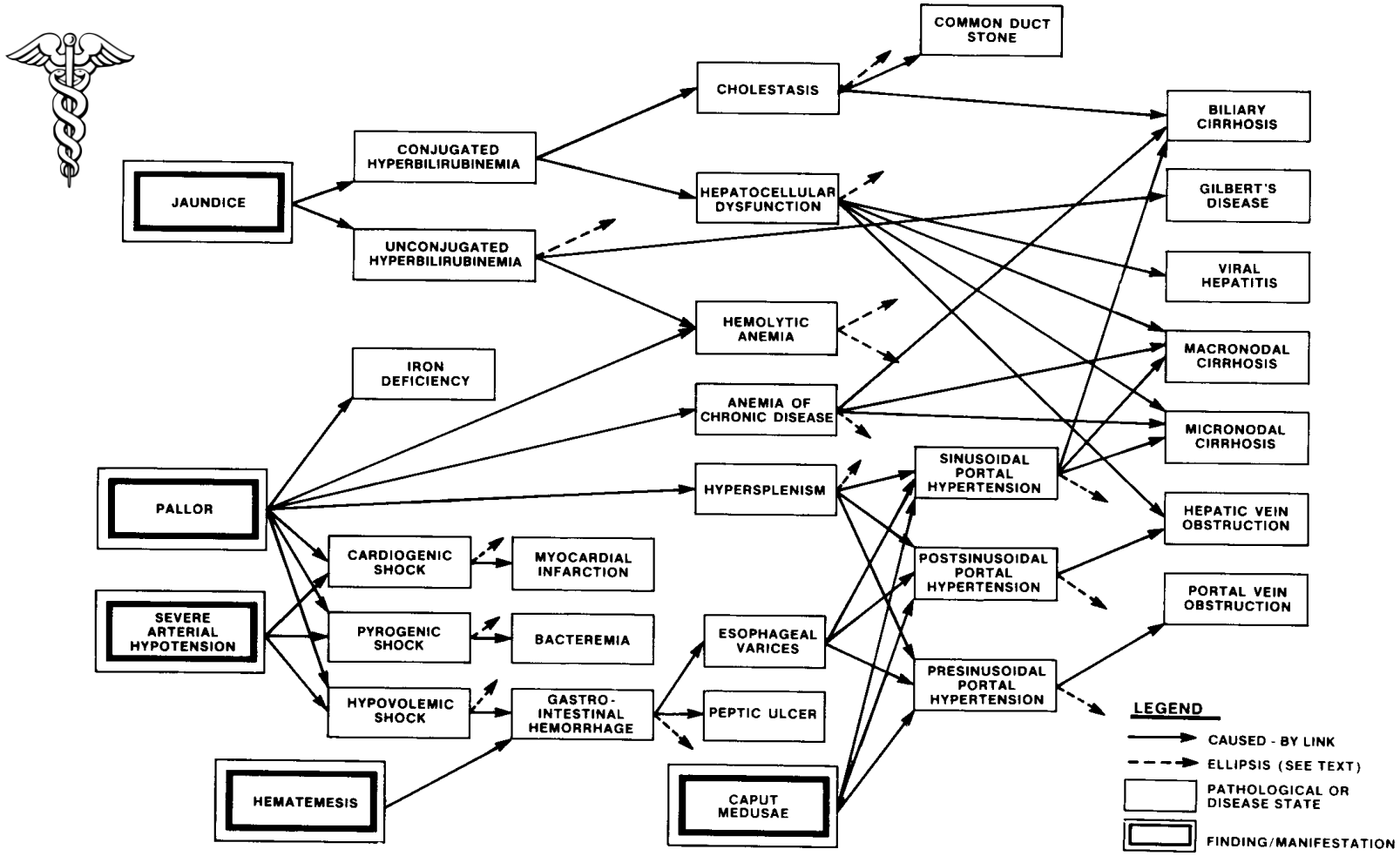
User Interaction

source: Norvig PAIP

Unrestricted © Siemens 2020

Expert Systems

CADUCEUS (1980s)



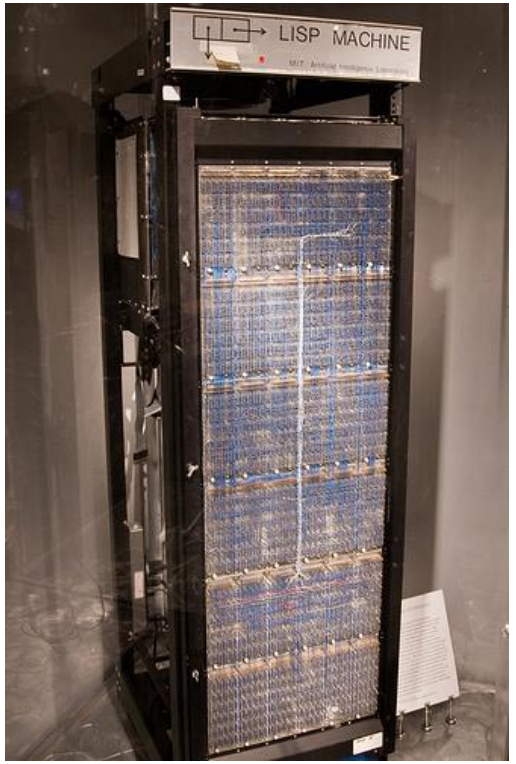
source: Szolovits (1982)

Unrestricted © Siemens 2020

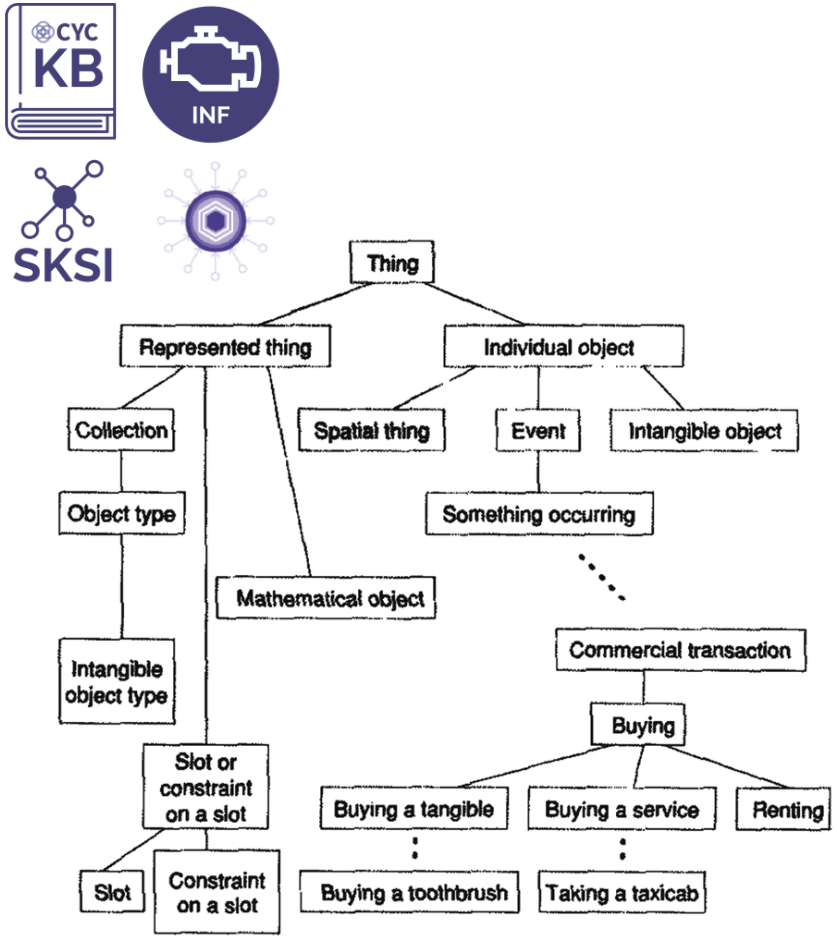
Expert Systems
Drawbacks

SIEMENS
Ingenuity for life

- Knowledge Acquisition
- Performance Issues
- Knowledge Base Size
- Rule Consistency
- Maintenance Process



source: victor.se

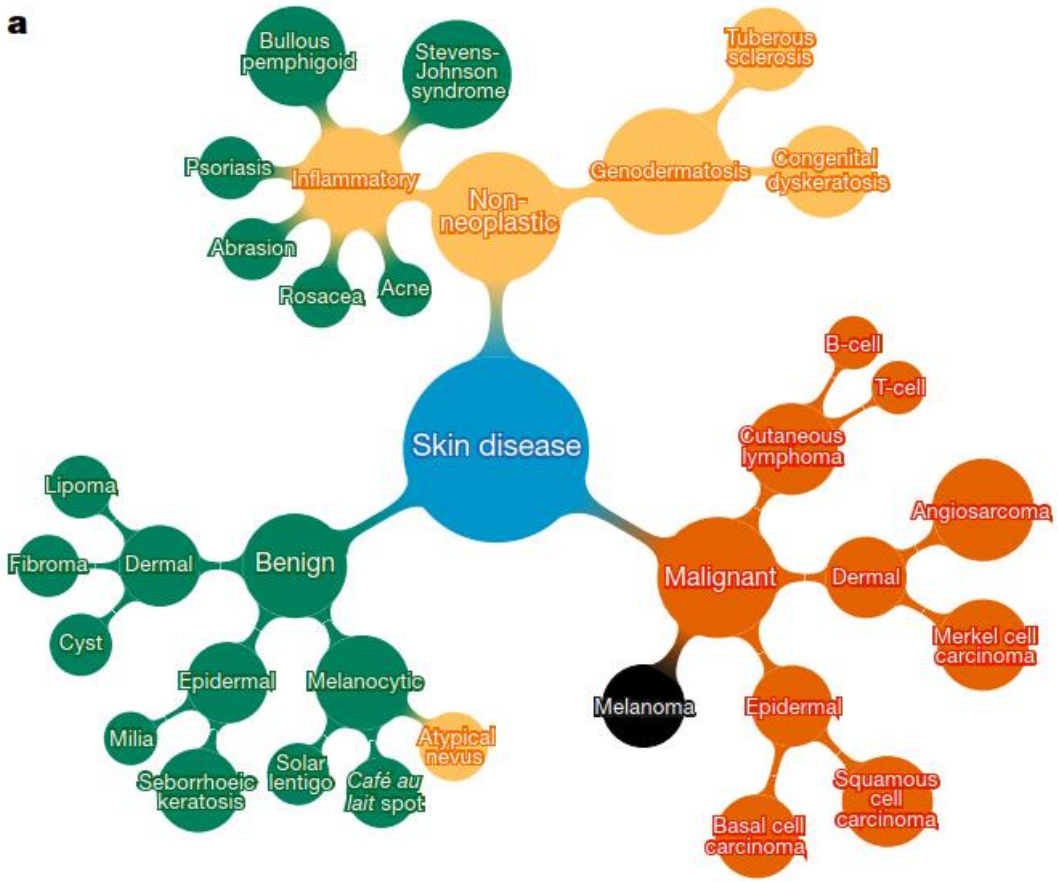


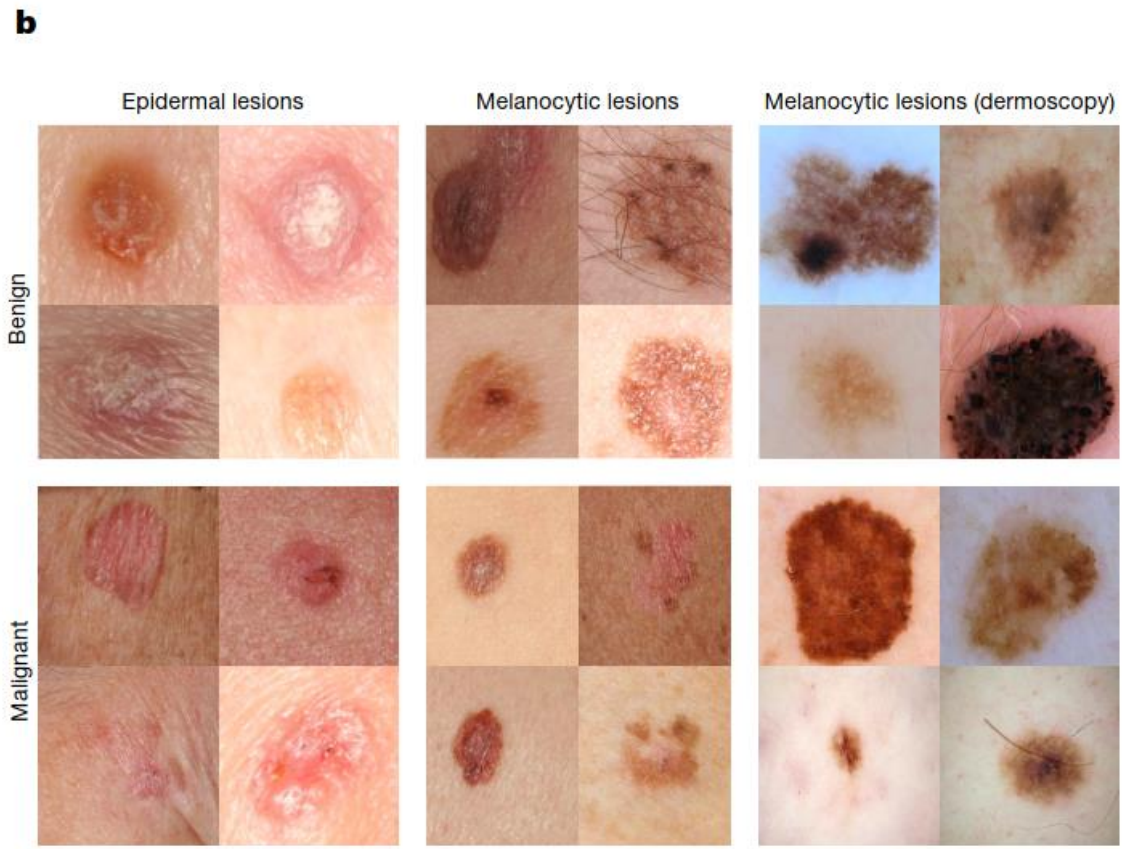
source: Cycorp

Unrestricted © Siemens 2020

AI/ML in Dermatology

Skin Lesion Classification



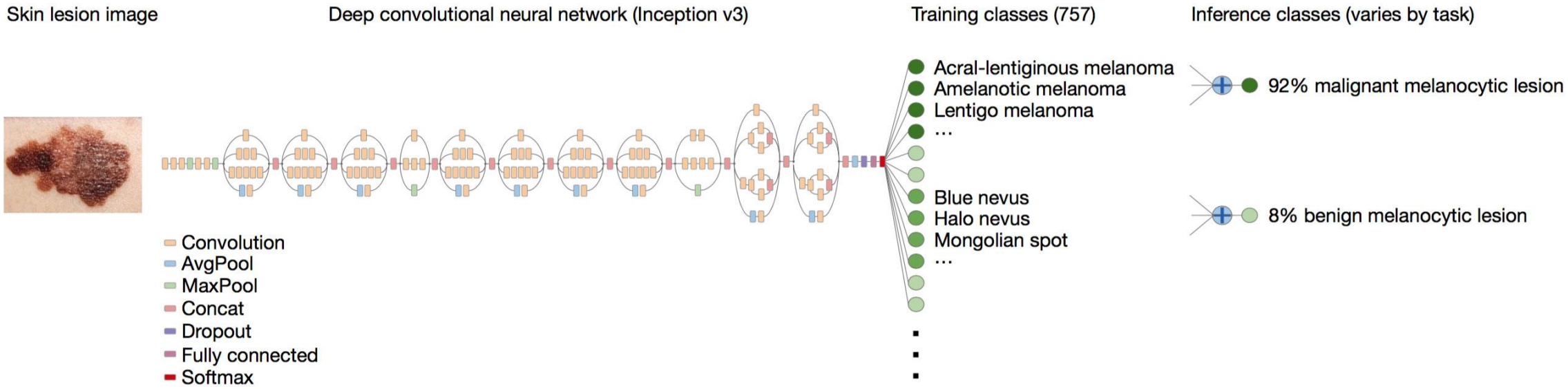


source: Esteva *et al.* (2017)

Unrestricted © Siemens 2020

AI/ML in Dermatology

Skin Lesion Classification



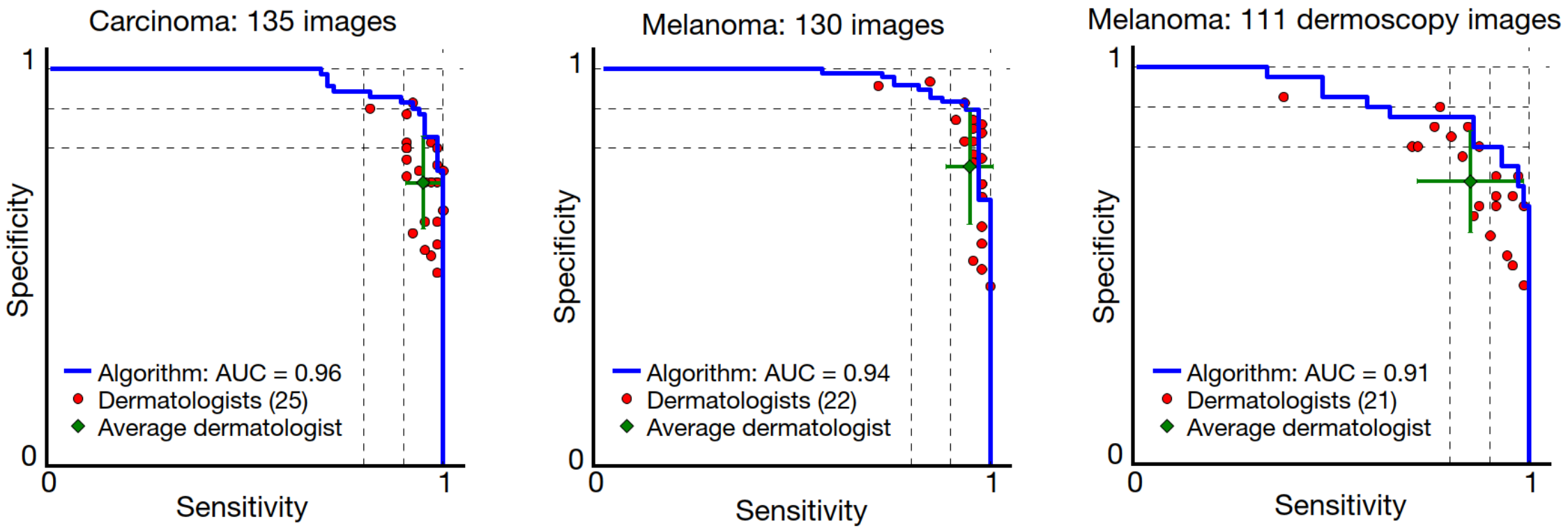
source: Esteva *et al.* (2017)

Unrestricted © Siemens 2020

AI/ML in Dermatology

Skin Lesion Classification

SIEMENS
Ingenuity for life



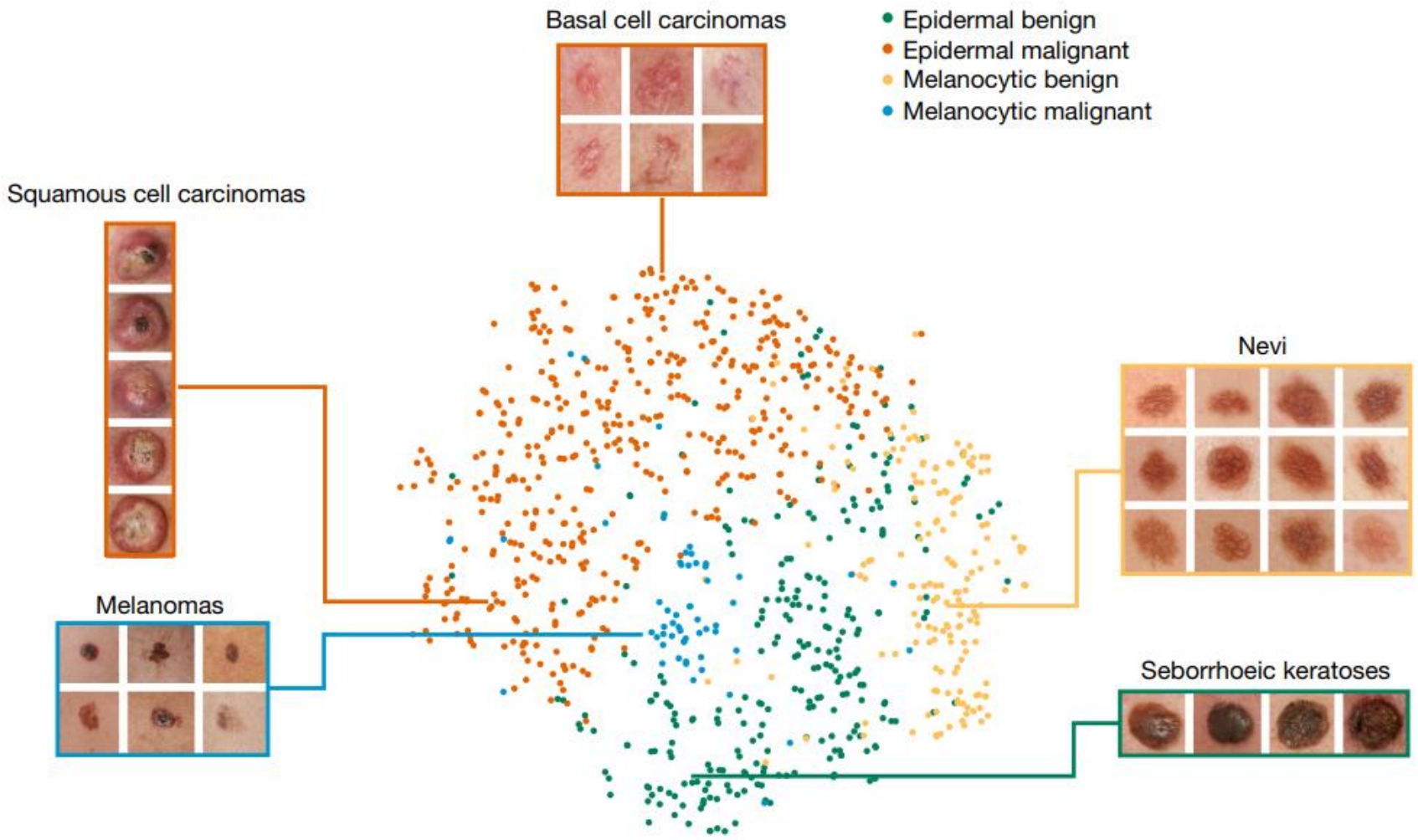
source: Esteva *et al.* (2017)

Extra Credit: use the InceptionV3 application from Keras to reproduce these results

AI/ML in Dermatology

Skin Lesion Classification

SIEMENS
Ingenuity for life



source: Esteva *et al.* (2017)

t-SNE

similar objects
closer to each other

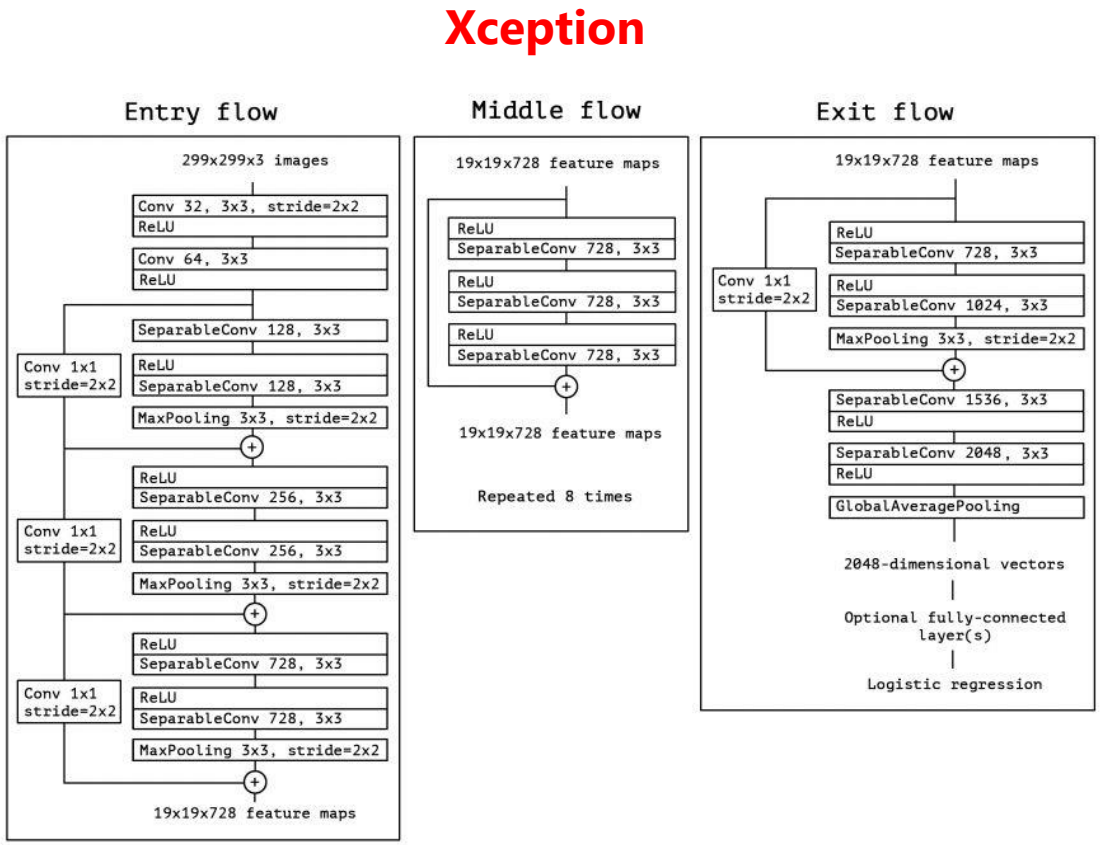
dissimilar objects
distant w/ high probability

Unrestricted © Siemens 2020

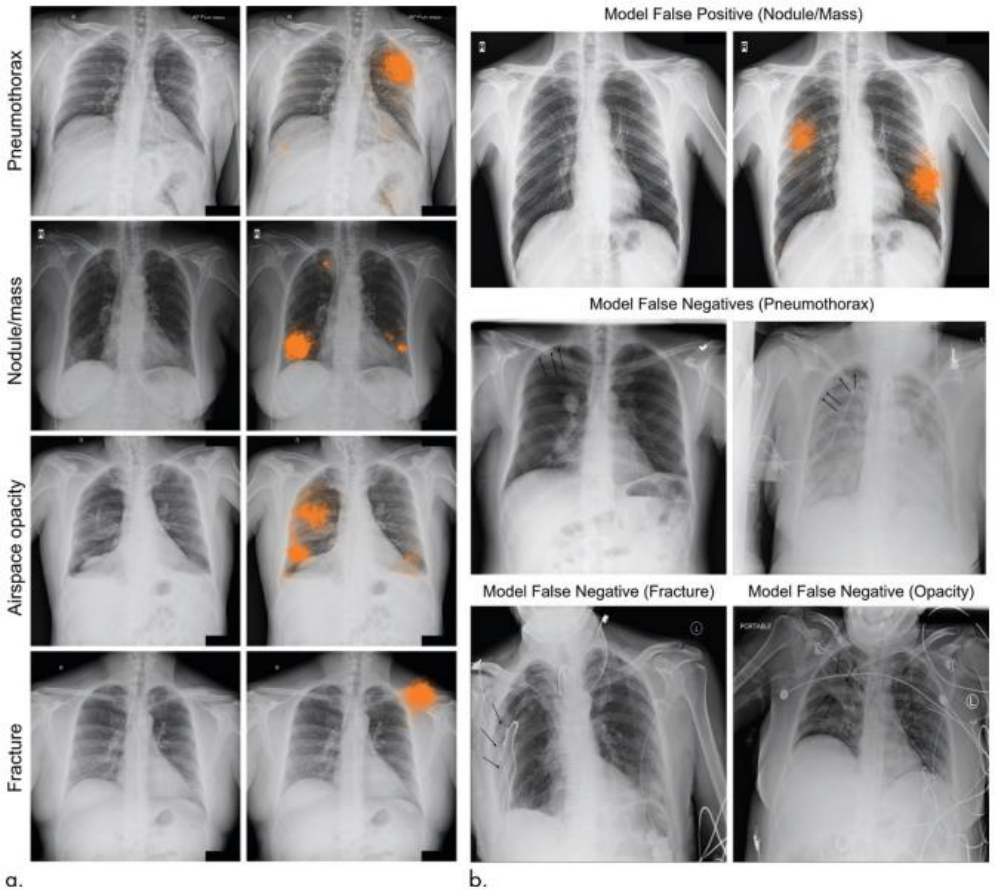
AI/ML in Pneumology

Chest X-Ray Interpretation

SIEMENS
Ingenuity for life



source: Towards Data Science



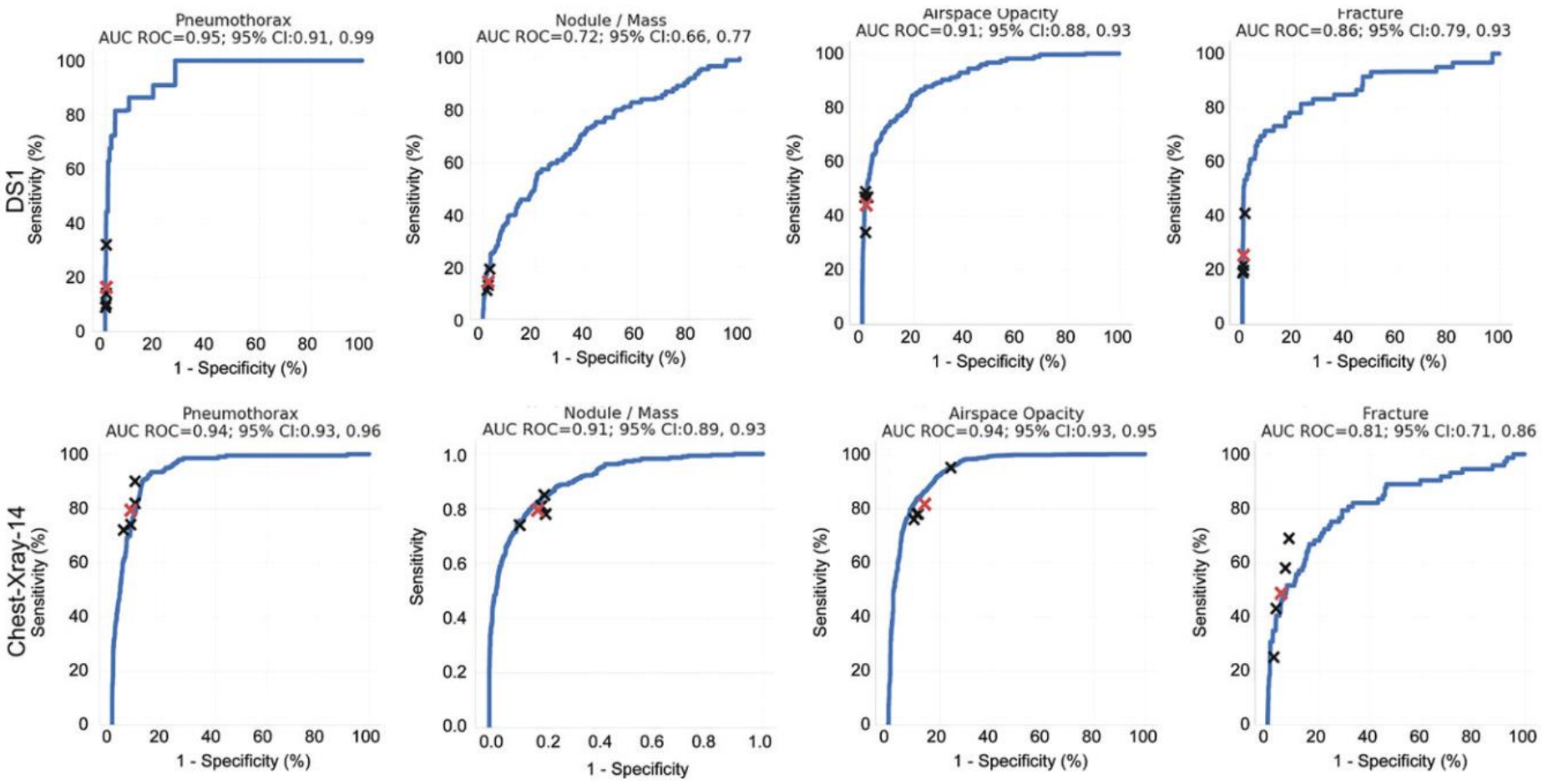
source: Majkowska *et al.* (2019)

Unrestricted © Siemens 2020

AI/ML in Pneumology

Chest X-Ray Interpretation

SIEMENS
Ingenuity for life



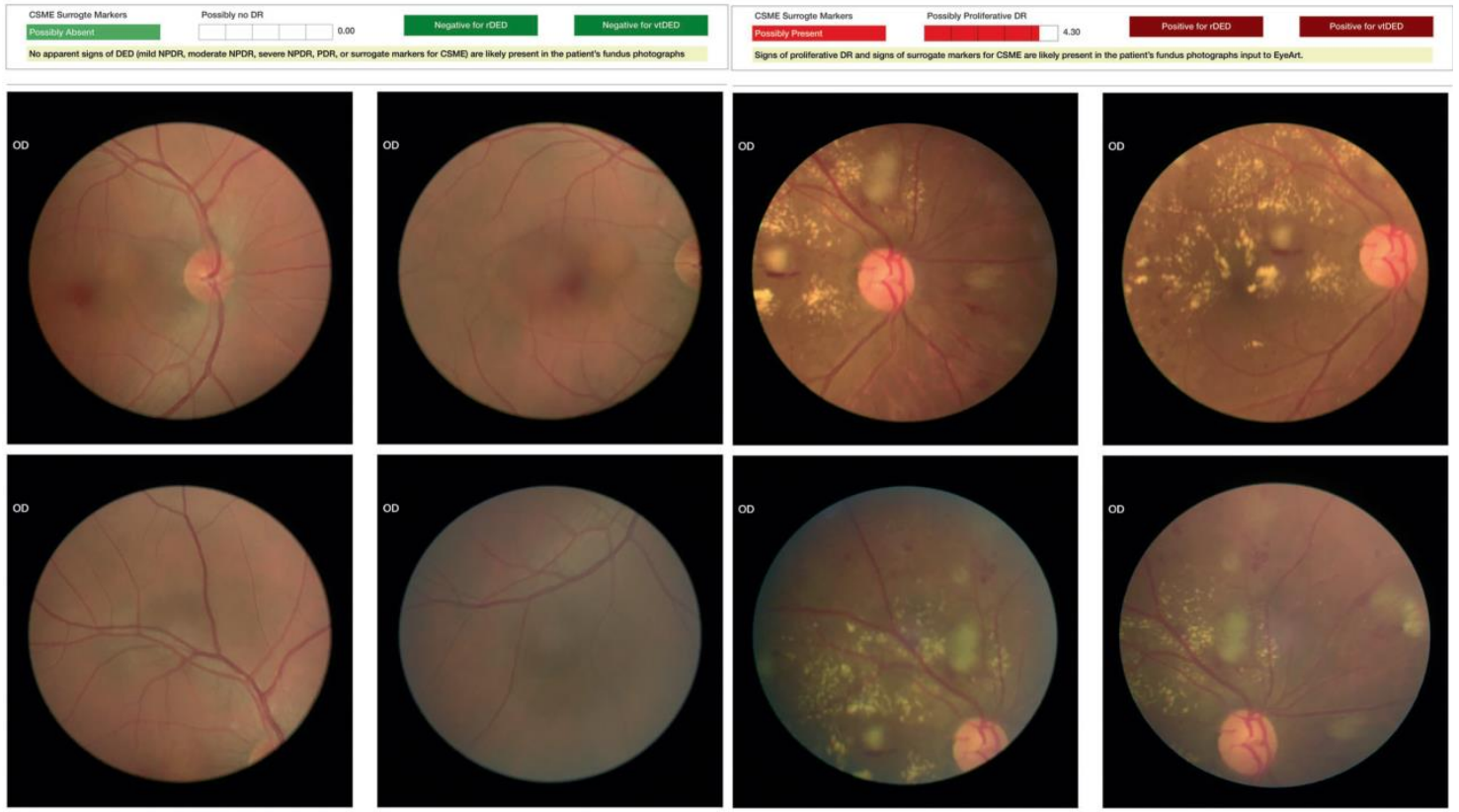
source: Majkowska *et al.* (2019)

Unrestricted © Siemens 2020

AI/ML in Ophthalmology

Diabetic Retinopathy Detection

SIEMENS
Ingenuity for life



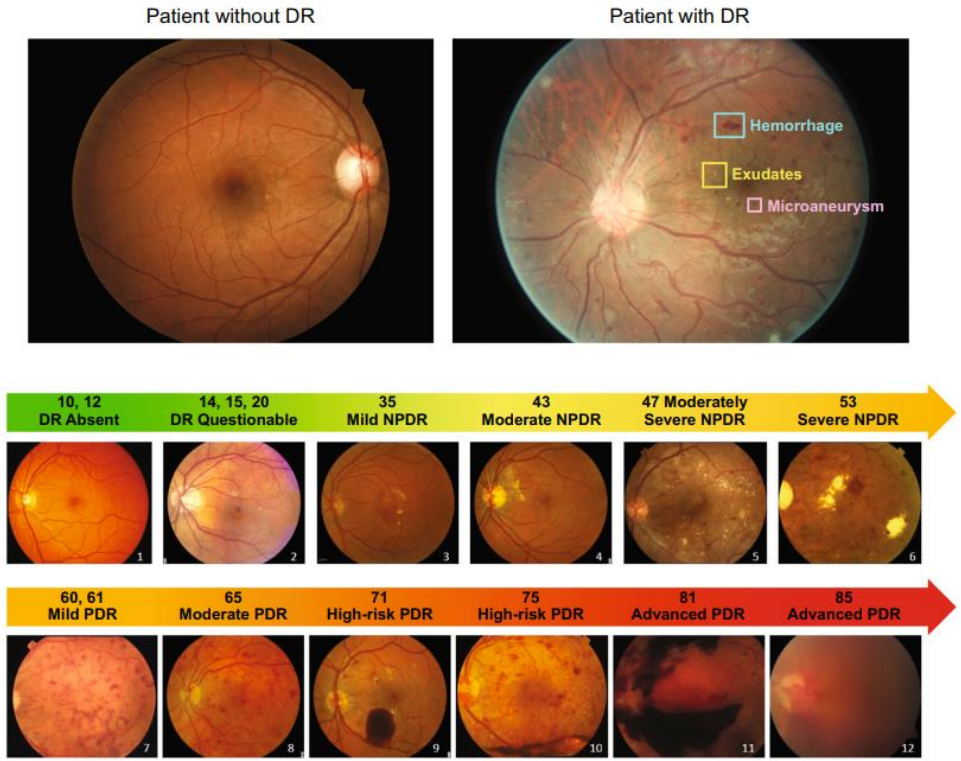
source: Raman *et al.* (2019)

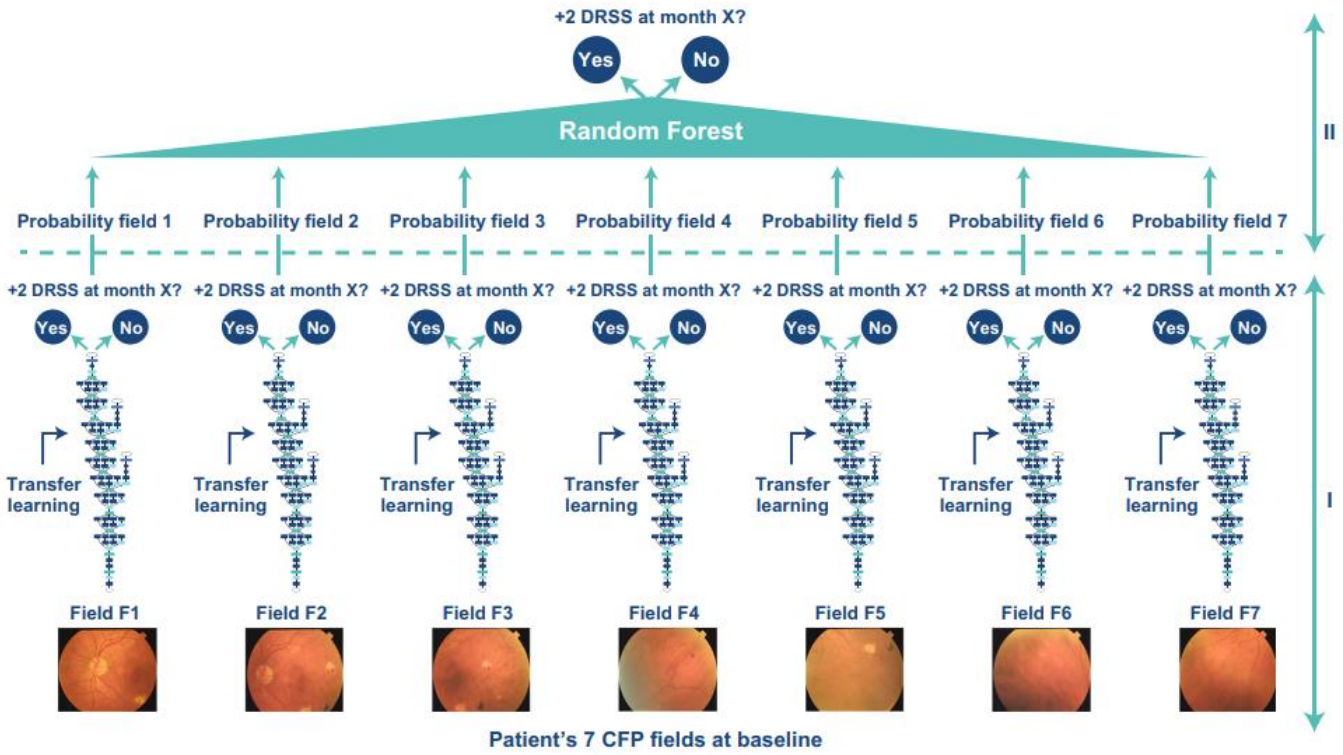
Unrestricted © Siemens 2020

AI/ML in Ophthalmology

Diabetic Retinopathy Progression

SIEMENS
Ingenuity for life





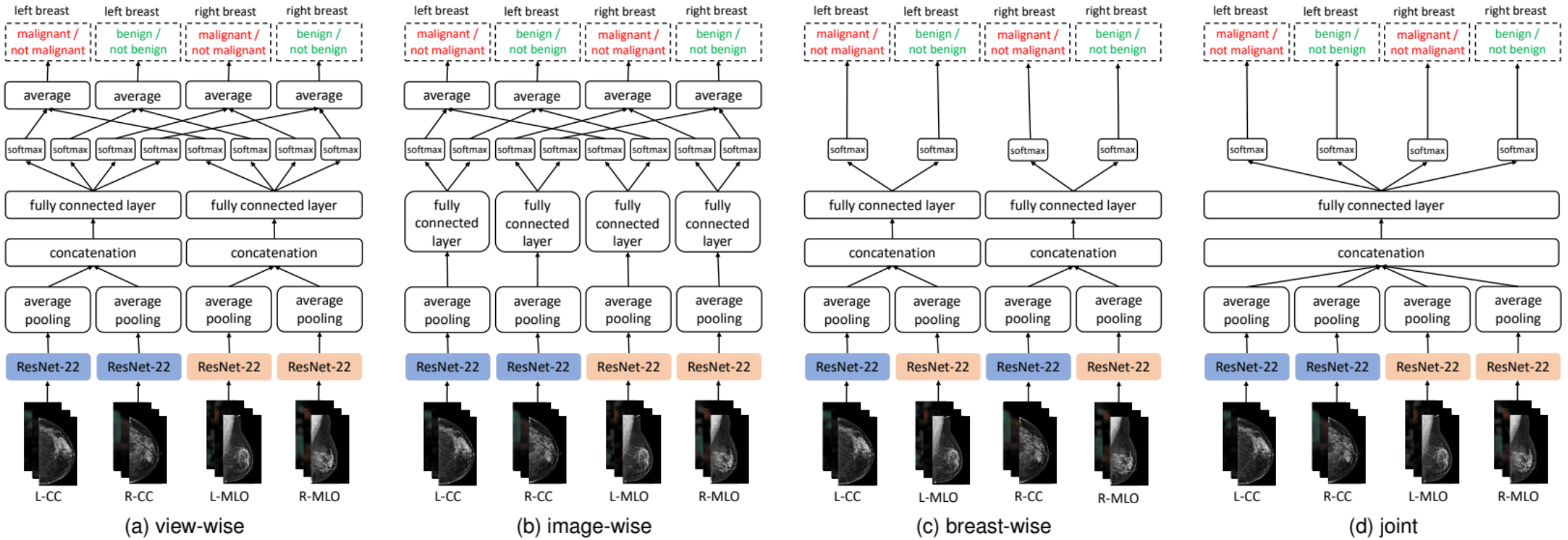
source: Arcadu *et al.* (2019)

Unrestricted © Siemens 2020

AI/ML in Oncology

Breast Cancer Screening

SIEMENS
Ingenuity for life



source: Wu *et al.* (2019)

Unrestricted © Siemens 2020

AI/ML in Oncology

Breast Cancer Screening

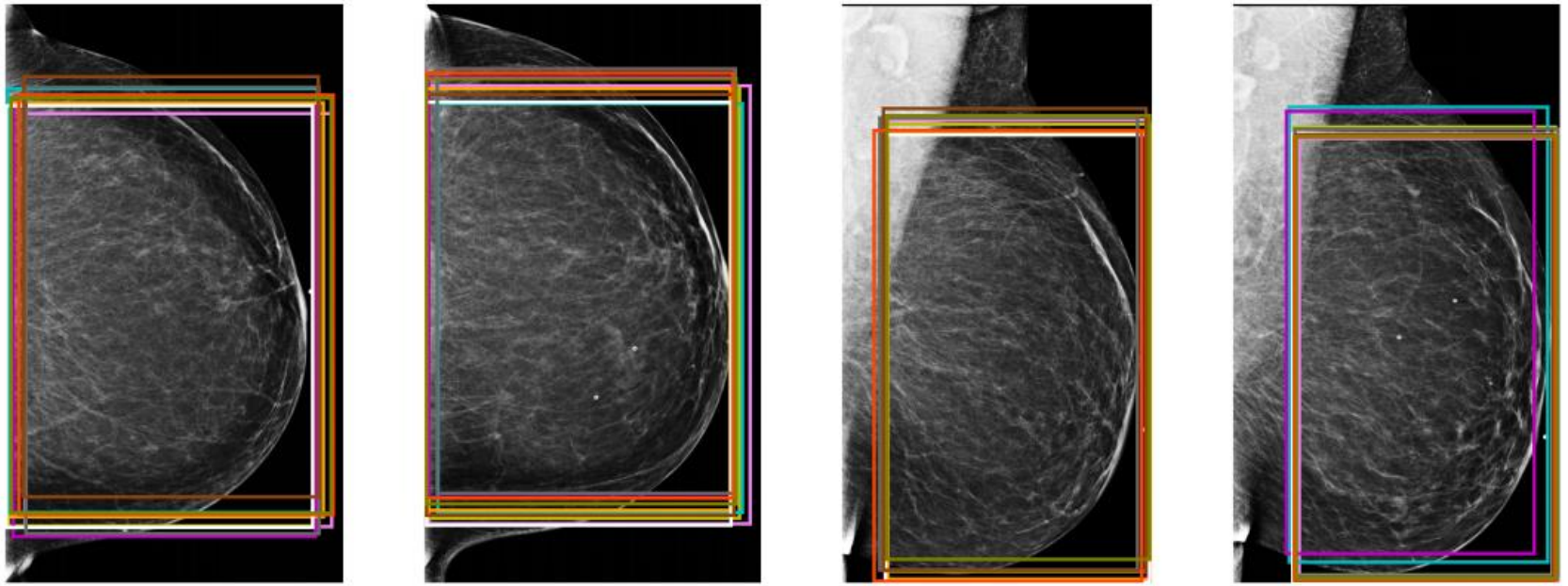
SIEMENS
Ingenuity for life

Data Augmentation

{rotate, flip, crop, color noise, &c.}

↓

natural and *consistent* variations



source: Wu *et al.* (2019)

Unrestricted © Siemens 2020

AI/ML in Oncology

Breast Cancer Screening

SIEMENS
Ingenuity for life

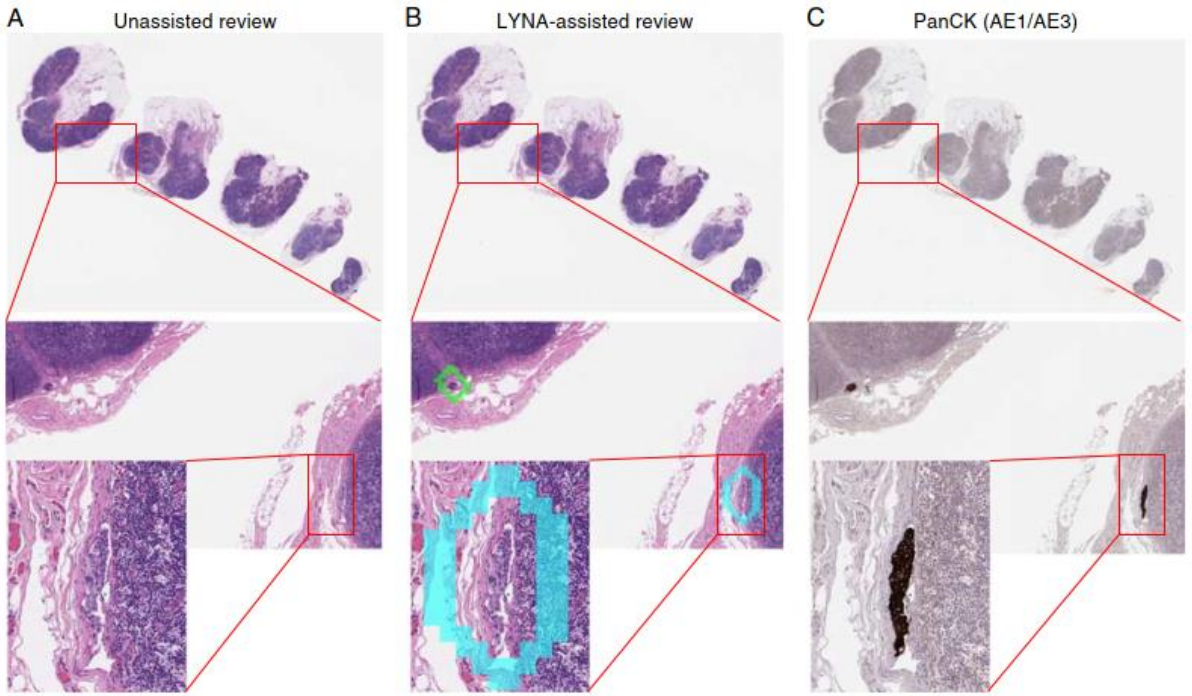
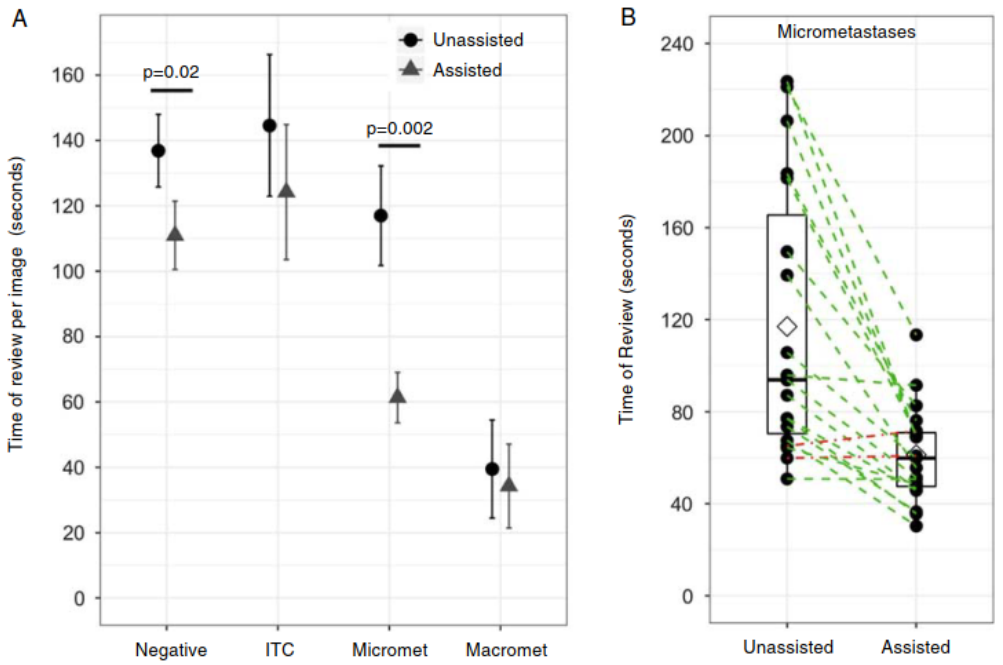


FIGURE 2. Digital image assistance by presenting algorithm predictions as a direct overlay; without assistance (A) versus with assistance (B). Suspicious regions of interest are highlighted in cyan for high confidence and green for moderate confidence (based on the algorithm predictions). In this image, both the high and moderate confidence regions of interests are confirmed as metastatic tumor as indicated by cytokeratin IHC (C). Another image example is shown in Supplemental Figure 2 (Supplemental Digital Content 1, http://links.lww.com/PAS/A677).

Model + Expert Collaboration



source: Steiner *et al.* (2018)

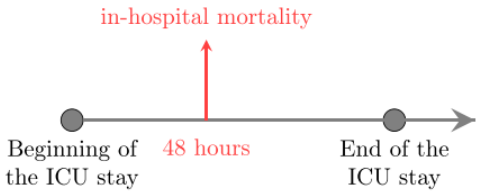
Unrestricted © Siemens 2020

AI/ML in Healthcare

Clinical Event Prediction

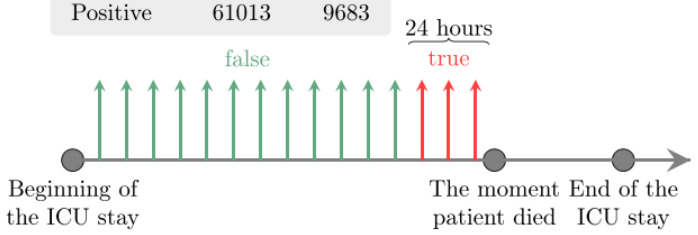
SIEMENS
Ingenuity for life

	Train	Test
Negative	15480	2862
Positive	2423	374



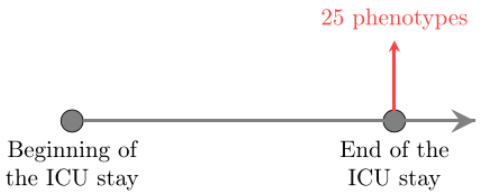
(a)

	Train	Test
Negative	2847401	513525
Positive	61013	9683



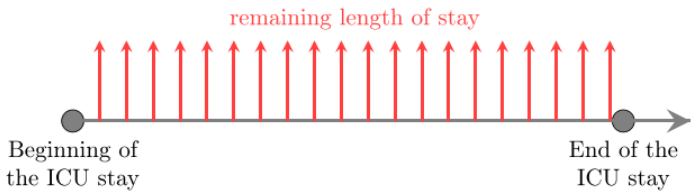
(b)

Train	Test
35621	6281



(c)

Train	Test
2925434	525912



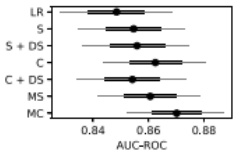
(d)

source: Harutyunyan *et al.* (2018)

LR – logistic regression
S – standard LSTM
C – channel-wise LSTM
DS – deep supervision
MS – multitask standard LSTM
MC – multitask channel-wise LSTM

In-hospital Mortality

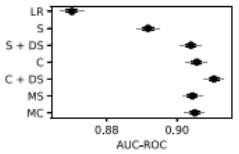
Model	AUC-ROC	AUC-PR
SAPS	0.720 (0.720, 0.720)	0.301 (0.301, 0.302)
APS-III	0.750 (0.750, 0.750)	0.357 (0.356, 0.357)
OASIS	0.760 (0.760, 0.761)	0.311 (0.311, 0.312)
SAPS-II	0.777 (0.776, 0.777)	0.376 (0.376, 0.377)
LR	0.848 (0.828, 0.868)	0.474 (0.419, 0.529)
S	0.855 (0.835, 0.873)	0.485 (0.431, 0.537)
S + DS	0.856 (0.836, 0.875)	0.493 (0.438, 0.549)
C	0.862 (0.844, 0.881)	0.515 (0.464, 0.568)
C + DS	0.854 (0.834, 0.873)	0.502 (0.447, 0.554)
MS	0.861 (0.842, 0.878)	0.493 (0.439, 0.548)
MC	0.870 (0.852, 0.887)	0.533 (0.480, 0.584)



	LR	S	S + DS	C	C + DS	MS	MC
LR	-	15.8	12.9	1.9	18.1	4.1	0.0
S	84.2	-	38.5	3.4	53.8	13.7	0.2
S + DS	87.1	61.5	-	10.7	63.4	19.2	0.6
C	98.1	96.6	89.3	-	95.5	62.3	4.7
C + DS	81.9	46.2	36.6	4.5	-	12.9	0.1
MS	95.9	86.3	80.8	37.7	87.1	-	0.9
MC	100.0	99.8	99.4	95.2	99.9	99.1	-

Decompensation

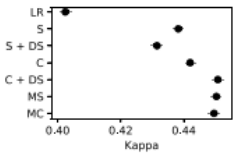
Model	AUC-ROC	AUC-PR
LR	0.870 (0.867, 0.873)	0.214 (0.205, 0.223)
S	0.892 (0.889, 0.895)	0.324 (0.314, 0.333)
S + DS	0.904 (0.901, 0.907)	0.325 (0.314, 0.335)
C	0.906 (0.903, 0.909)	0.333 (0.323, 0.344)
C + DS	0.911 (0.908, 0.913)	0.344 (0.334, 0.354)
MS	0.904 (0.902, 0.907)	0.321 (0.312, 0.331)
MC	0.905 (0.902, 0.908)	0.317 (0.307, 0.328)



	LR	S	S + DS	C	C + DS	MS	MC
LR	-	0.0	0.0	0.0	0.0	0.0	0.0
S	100.0	-	0.0	0.0	0.0	0.0	0.0
S + DS	100.0	100.0	-	6.2	0.0	35.1	13.2
C	100.0	100.0	93.8	-	0.0	91.4	73.1
C + DS	100.0	100.0	100.0	100.0	-	100.0	100.0
MS	100.0	100.0	64.9	8.6	0.0	-	12.5
MC	100.0	100.0	86.8	26.9	0.0	87.5	-

Length of stay

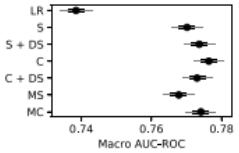
Model	Kappa	MAD
LR	0.402 (0.401, 0.404)	162.3 (161.8, 162.8)
S	0.438 (0.436, 0.440)	123.1 (122.6, 123.5)
S + DS	0.431 (0.430, 0.433)	110.9 (110.5, 111.4)
C	0.442 (0.440, 0.444)	136.6 (136.1, 137.1)
C + DS	0.451 (0.449, 0.453)	143.1 (142.6, 143.6)
MS	0.450 (0.449, 0.452)	112.0 (111.5, 112.5)
MC	0.450 (0.448, 0.451)	122.8 (122.3, 123.3)



	LR	S	S + DS	C	C + DS	MS	MC
LR	-	0.0	0.0	0.0	0.0	0.0	0.0
S	100.0	-	100.0	0.0	0.0	0.0	0.0
S + DS	100.0	0.0	-	0.0	0.0	0.0	0.0
C	100.0	100.0	100.0	-	0.0	0.0	0.0
C + DS	100.0	100.0	100.0	100.0	-	79.7	99.5
MS	100.0	100.0	100.0	100.0	20.3	-	91.0
MC	100.0	100.0	100.0	100.0	0.5	9.0	-

Phenotyping

Model	Macro AUC-ROC	Micro AUC-ROC
LR	0.739 (0.734, 0.743)	0.799 (0.796, 0.803)
S	0.770 (0.766, 0.775)	0.821 (0.818, 0.825)
S + DS	0.774 (0.769, 0.778)	0.823 (0.820, 0.827)
C	0.776 (0.772, 0.781)	0.825 (0.822, 0.828)
C + DS	0.773 (0.769, 0.777)	0.822 (0.819, 0.826)
MS	0.768 (0.763, 0.772)	0.818 (0.815, 0.822)
MC	0.774 (0.770, 0.778)	0.823 (0.819, 0.826)

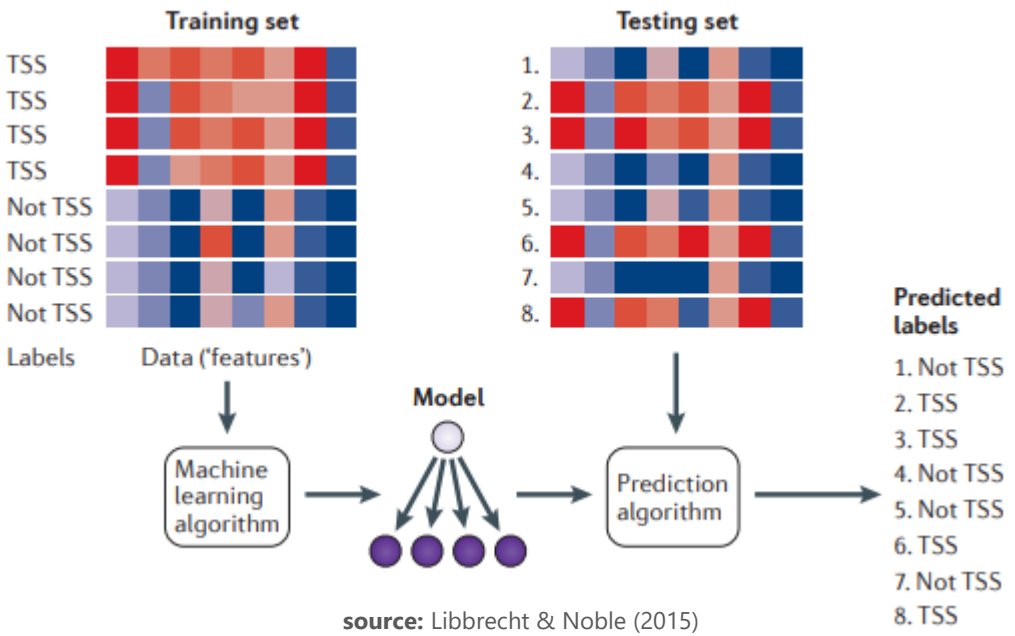


	LR	S	S + DS	C	C + DS	MS	MC
LR	-	0.0	0.0	0.0	0.0	0.0	0.0
S	100.0	-	0.1	0.0	1.7	97.6	0.2
S + DS	100.0	99.9	-	1.3	68.7	100.0	36.5
C	100.0	100.0	98.7	-	99.6	100.0	96.9
C + DS	100.0	98.3	31.3	0.4	-	100.0	20.1
MS	100.0	2.4	0.0	0.0	0.0	-	0.0
MC	100.0	99.8	63.5	3.1	79.9	100.0	-

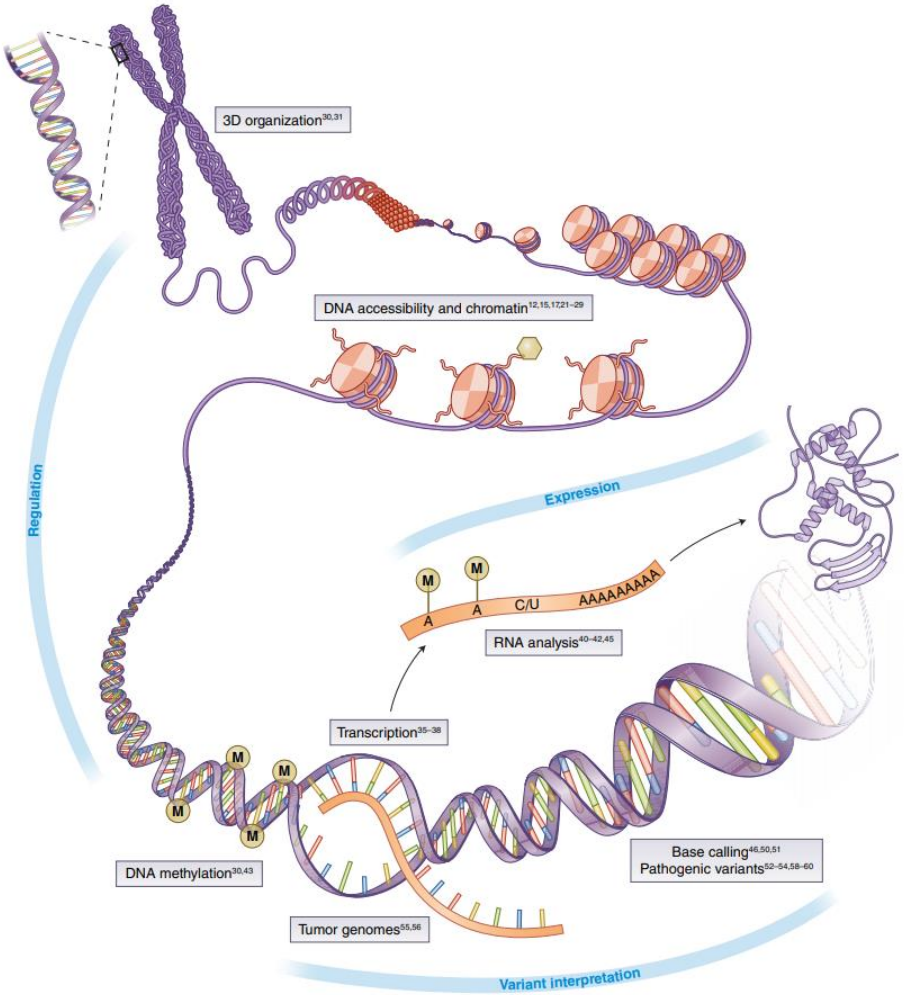
Unrestricted © Siemens 2020

AI/ML in Genetics and Genomics

SIEMENS
Ingenuity for life



source: Libbrecht & Noble (2015)

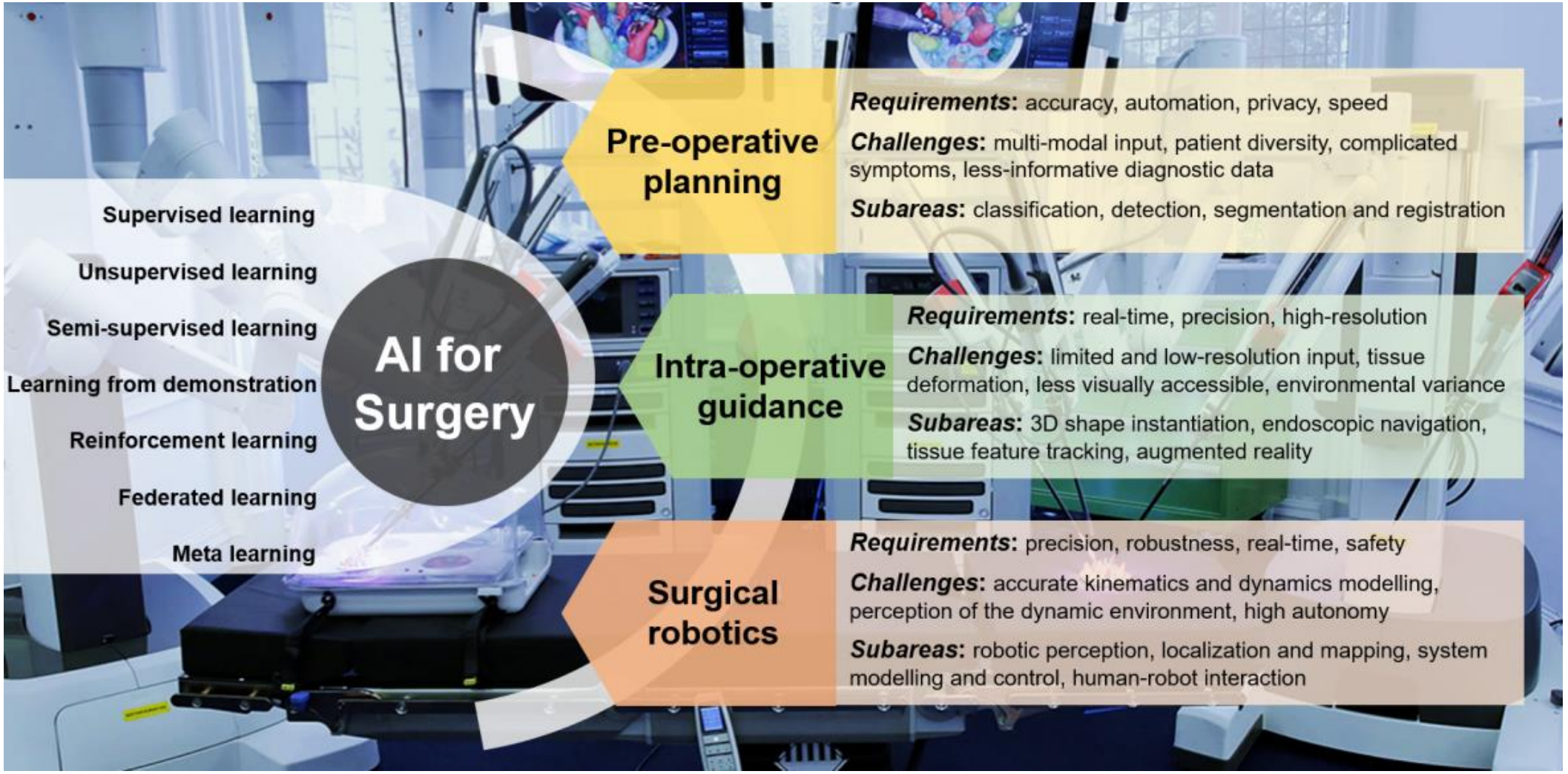


source: Zou *et al.* (2019)

Unrestricted © Siemens 2020

AI/ML in Surgery

SIEMENS
Ingenuity for life



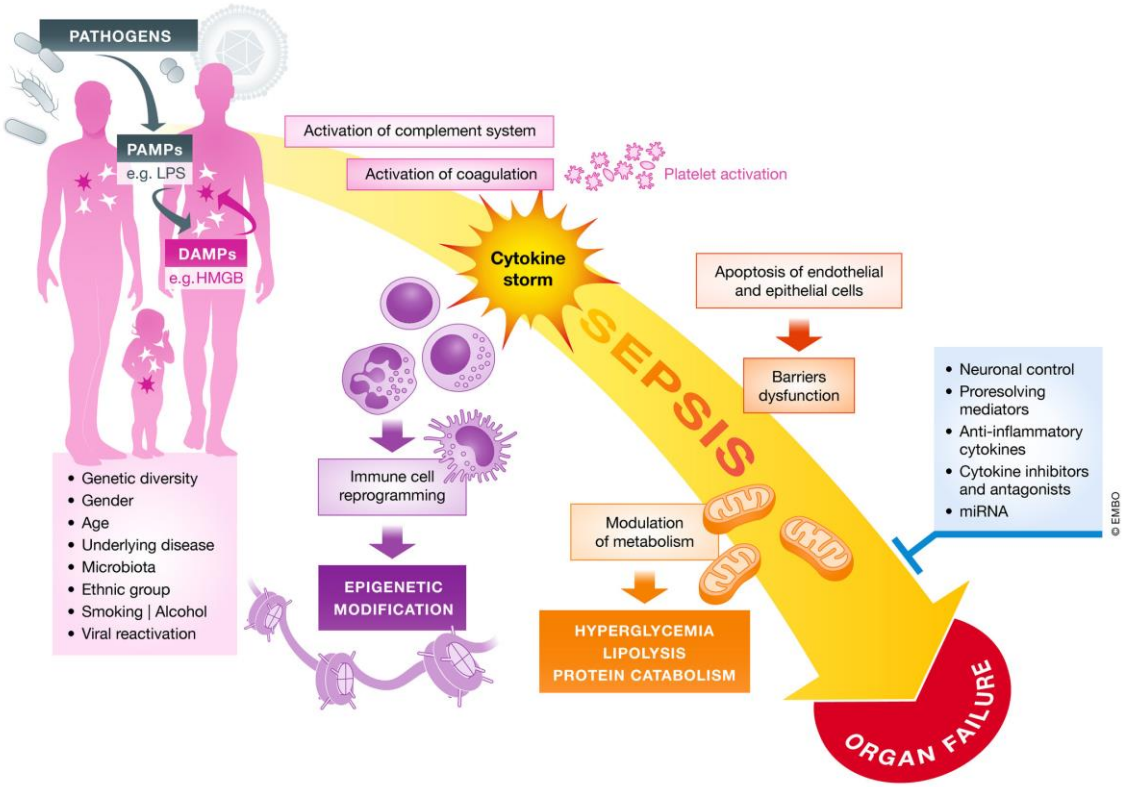
source: Zhou *et al.* (2020)

Unrestricted © Siemens 2020

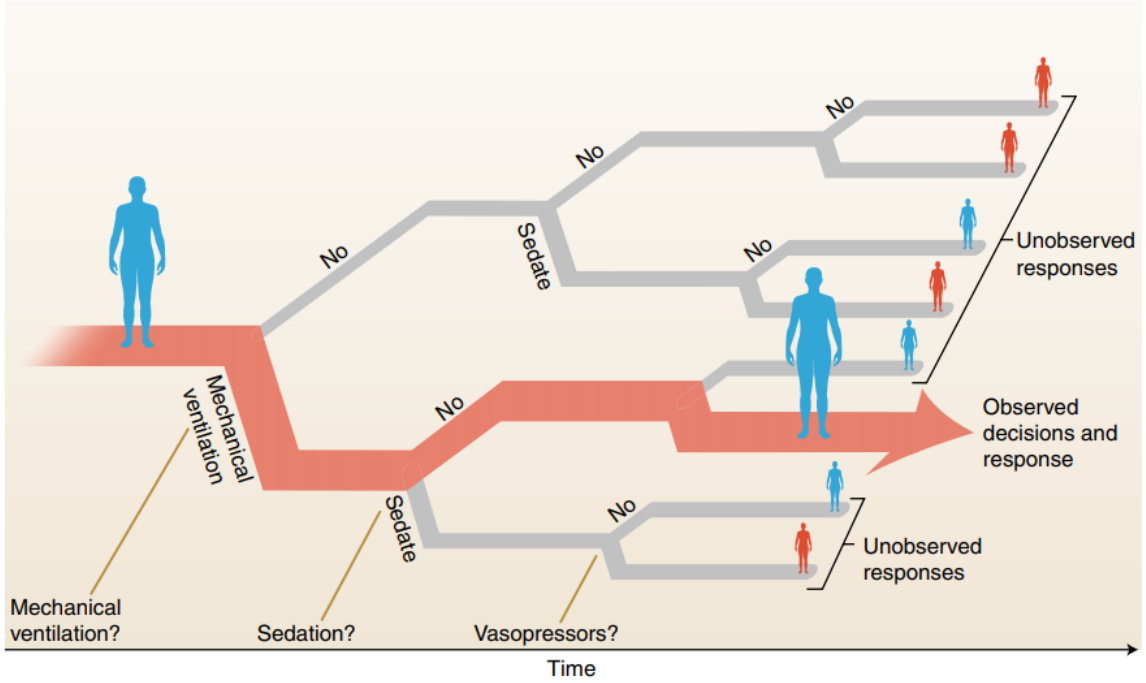
AI/ML in the ICU

Reinforcement Learning and Sepsis Treatment

SIEMENS
Ingenuity for life



source: EMBO



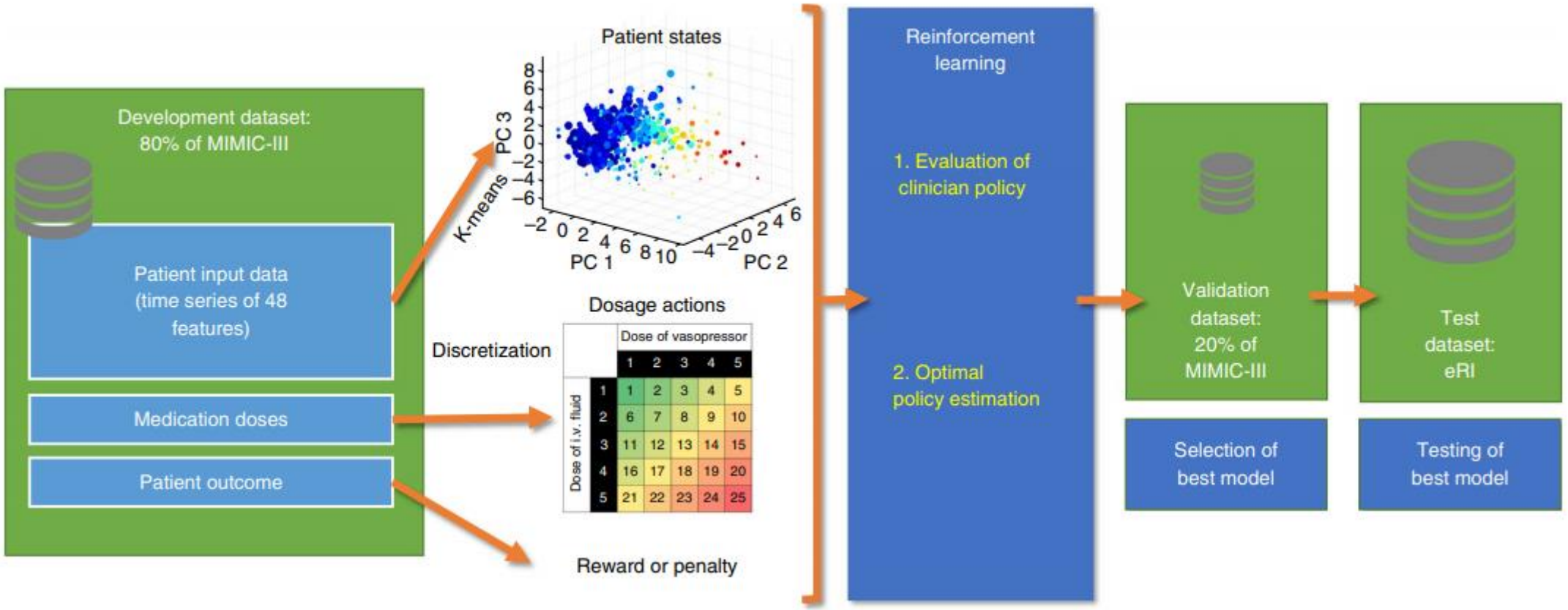
source: Gottesman *et al.* (2019)

Unrestricted © Siemens 2020

AI/ML in the ICU

Reinforcement Learning and Sepsis Treatment

SIEMENS
Ingenuity for life



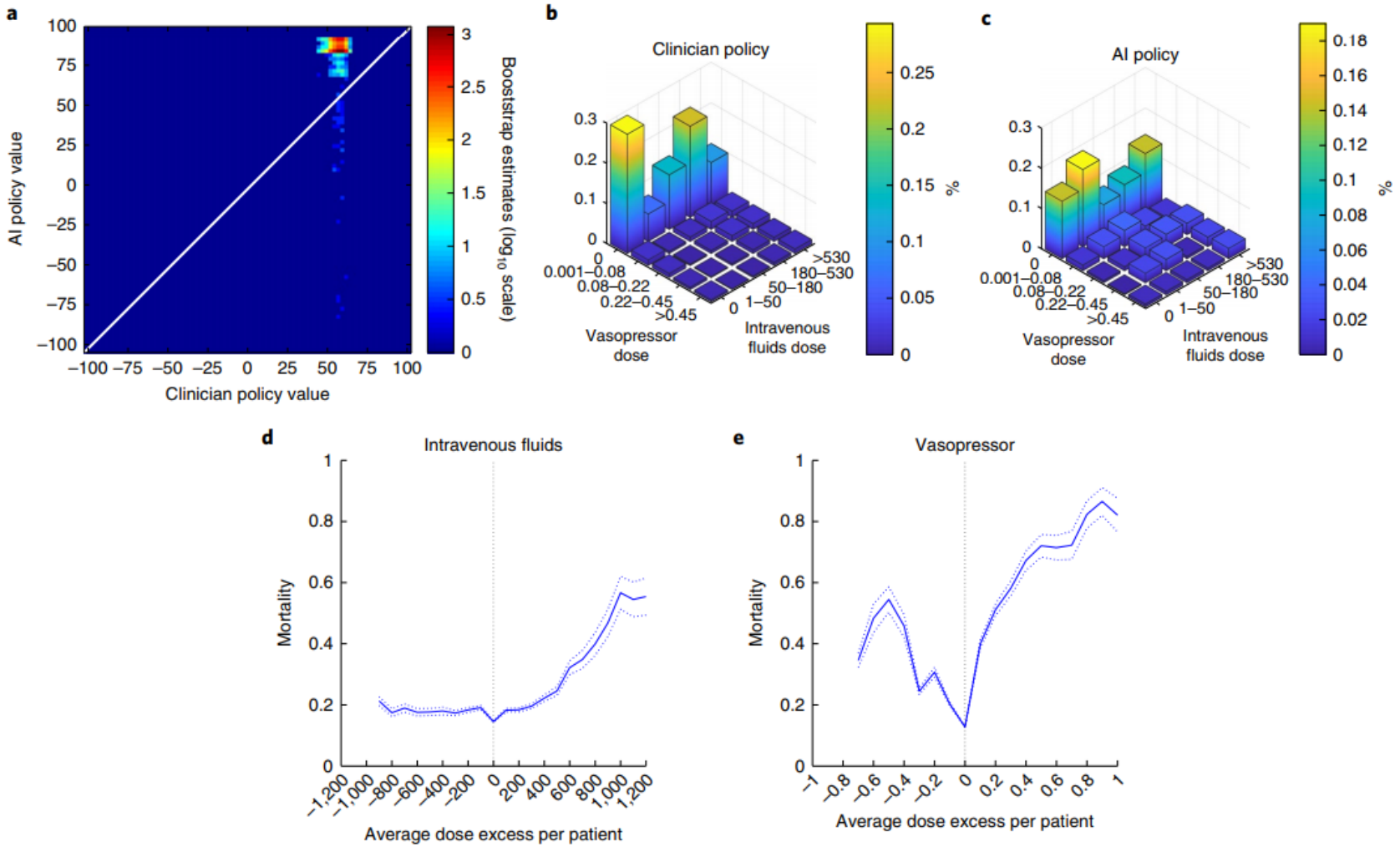
source: Komorowski *et al.* (2018)

Unrestricted © Siemens 2020

AI/ML in the ICU

Reinforcement Learning and Sepsis Treatment

SIEMENS
Ingenuity for life



source: Komorowski *et al.* (2018)

Unrestricted © Siemens 2020

NLP in Healthcare

Amazon Comprehend Medical

SIEMENS
Ingenuity for life

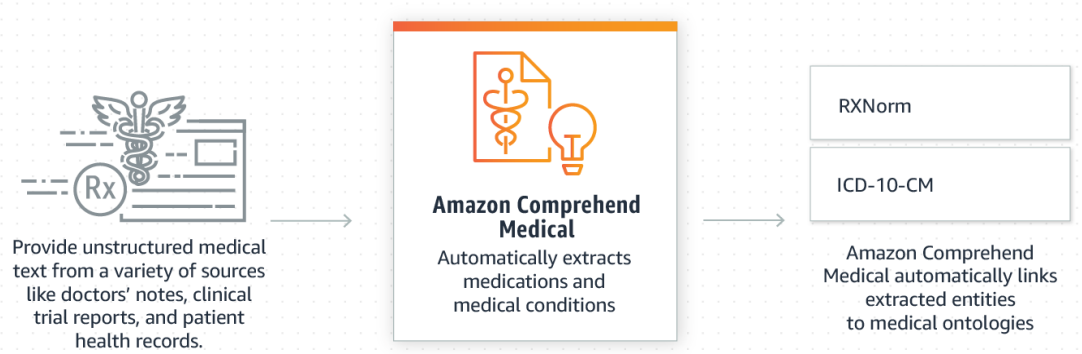
Pt is 40yo mother, software engineer

HPI: Sleeping trouble on present dosage of Clonidine. Severe Rash on face and leg, slightly itchy

Meds: Vyvanse 50 mgs po at breakfast daily, Clonidine 0.2 mgs – 1 and 1/2 tabs po ghs

HEENT: Boggy inferior turbinates, No oropharyngeal lesion. Lungs: clear. Heart: Regular rhythm. Skin: Papular mild erythematous eruption to hairline

Follow-up as scheduled



40yo 0.99+ score	Age	Protected health information	-
software engineer 0.98 score	Profession	Protected health information	-
Sleeping trouble 0.81 score	Dx name	Medical condition	Symptom 0.68 score
Clonidine 0.98 score	Generic name	Medication	-
Rash 0.99+ score	Dx name	Medical condition	Symptom 0.75 score
face 0.98 score	System organ site	Anatomy	-
leg 0.99+ score	System organ site	Anatomy	-

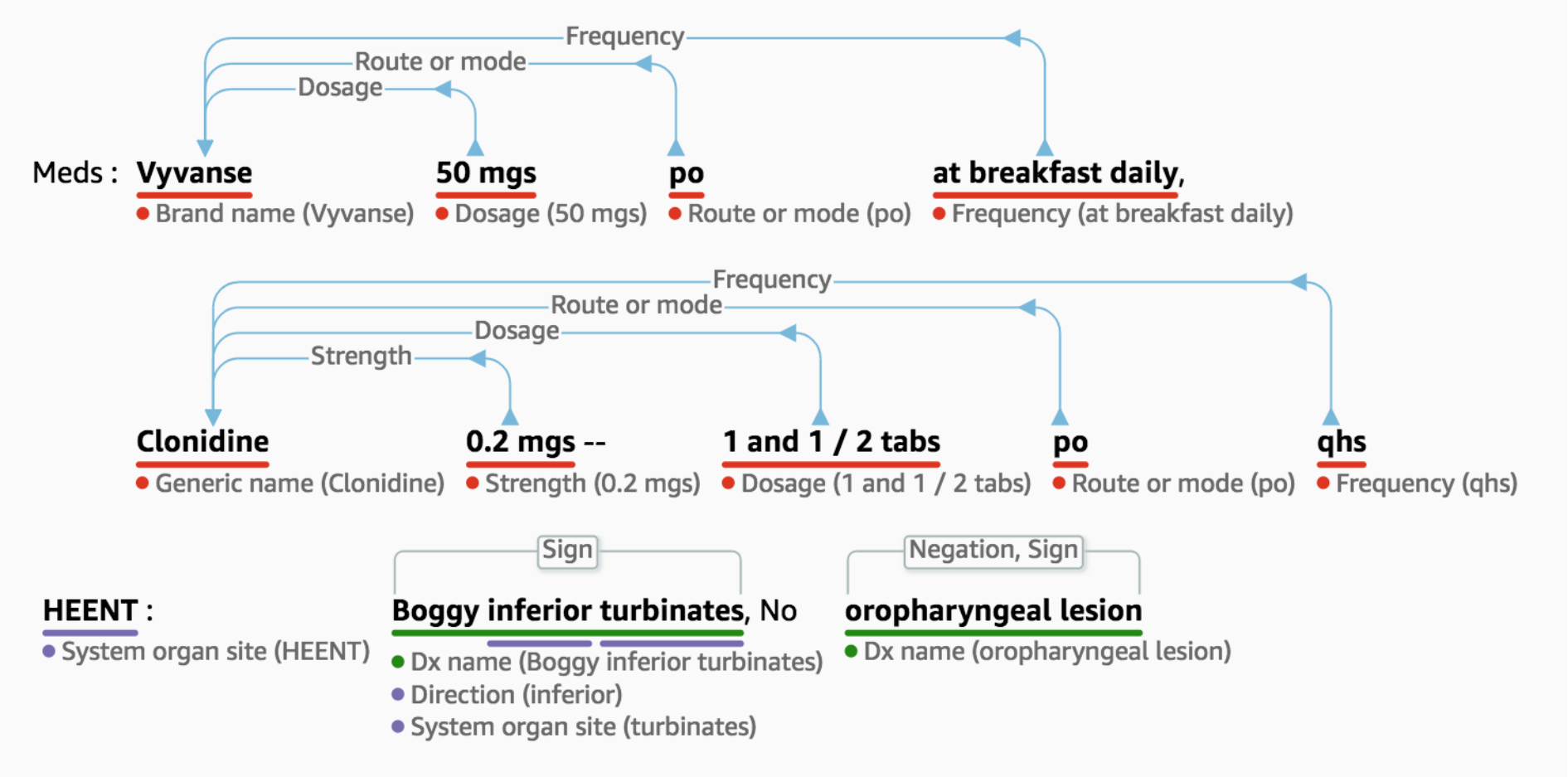
source: Amazon

Unrestricted © Siemens 2020

NLP in Healthcare

Amazon Comprehend Medical

SIEMENS
Ingenuity for life



source: Amazon

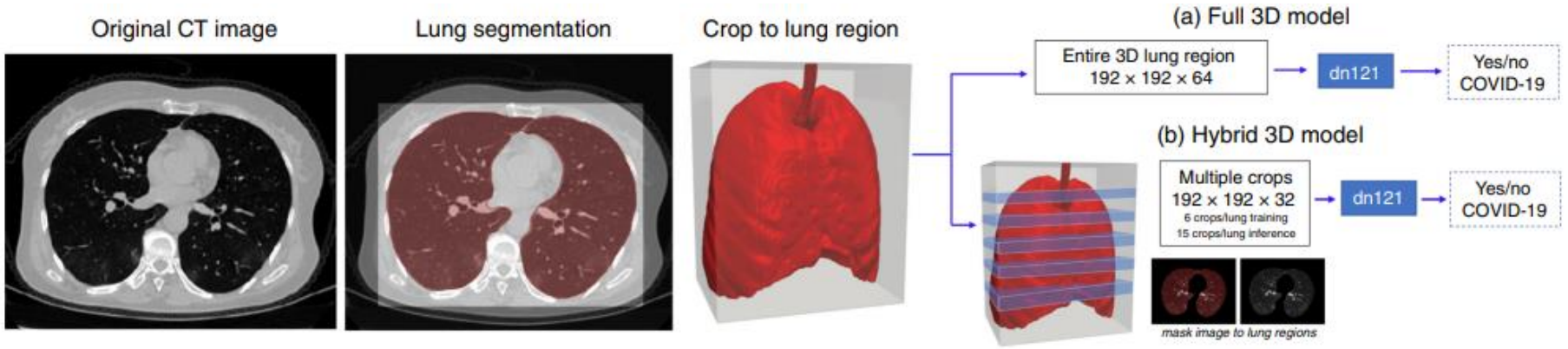
Unrestricted © Siemens 2020

AI/ML in Epidemiology

COVID-19 Diagnosis

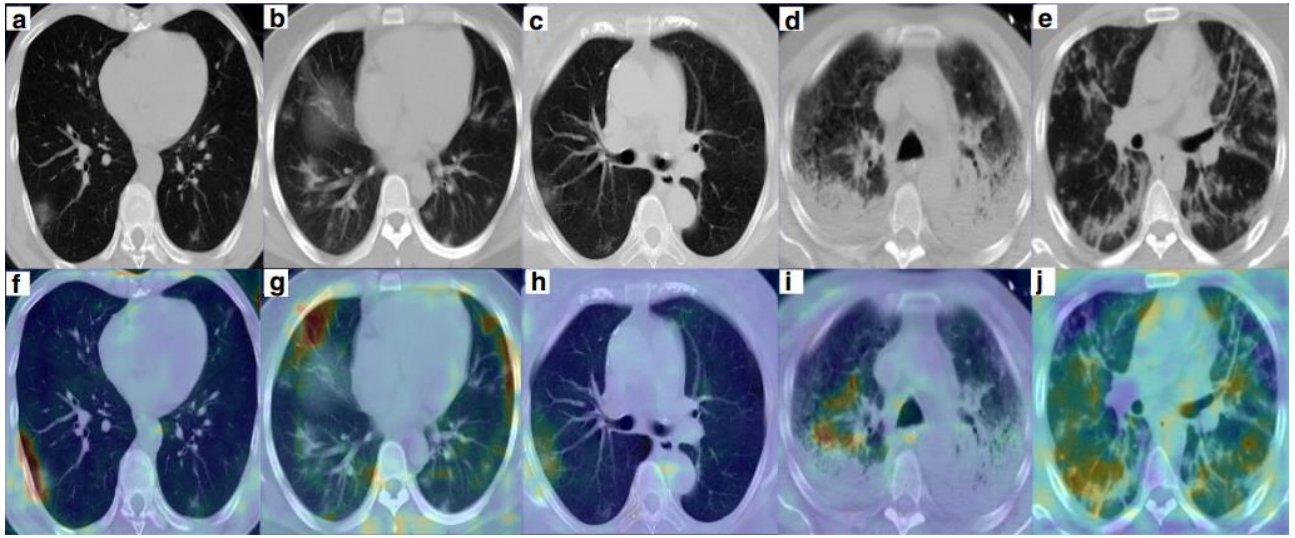
SIEMENS
Ingenuity for life

3D Classification Model



Grad-CAM

"visual explanations"



source: Harmon *et al.* (2020)

Unrestricted © Siemens 2020

AI/ML in Epidemiology

COVID-19 Prognosis

SIEMENS
Ingenuity for life

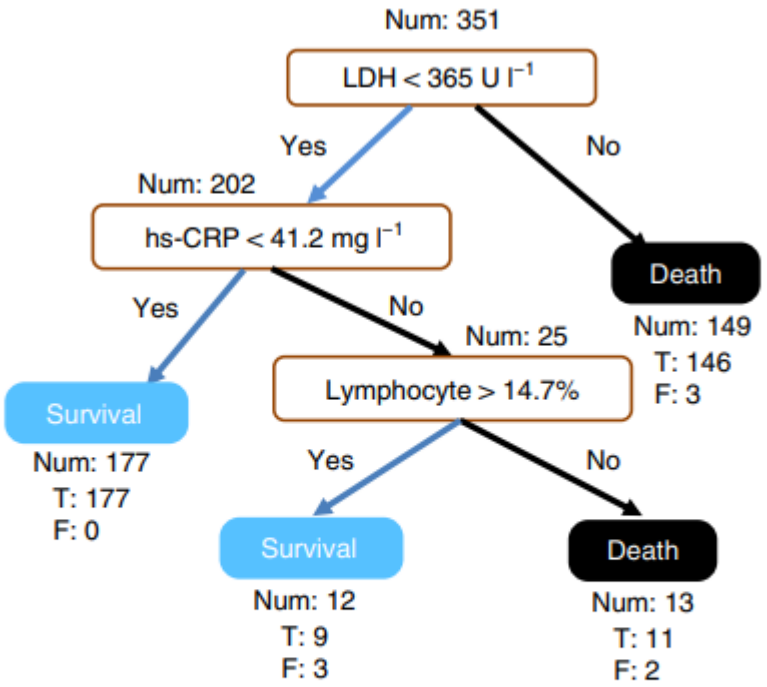


Fig. 2 | A decision rule using three key features and their thresholds in absolute value. Num, the number of patients in a class; T, the number of correctly classified; F, the number of misclassified patients.

Table 3 | Performances of the Multi-tree XGBoost classification in discriminating between mortality outcomes using 100-round fivefold cross-validation using Supplementary algorithm 1

Features	AUC score for training sets (%)	AUC score for validation sets (%)
LDH	94.27 ± 0.82	92.29 ± 2.62
LDH, lymphocyte	96.74 ± 0.45	94.40 ± 2.31
LDH, lymphocyte, hs-CRP	97.84 ± 0.37	95.06 ± 2.21

Data presented as mean ± s.d.

Table 4 | Performance of the proposed interpretable model on the external test dataset

	Precision	Recall	F1 score	Support
Survival	1.00	0.97	0.98	97
Death	0.81	1.00	0.90	13
Accuracy			0.97	110
Macro avg	0.91	0.98	0.94	110
Weighted avg	0.98	0.97	0.97	110

source: Yan *et al.* (2020)

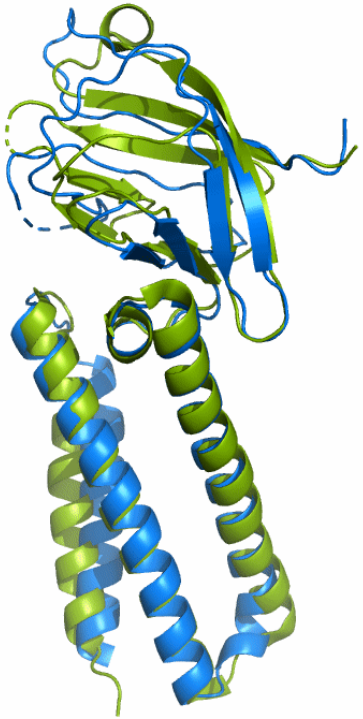
Unrestricted © Siemens 2020

AI/ML in Epidemiology

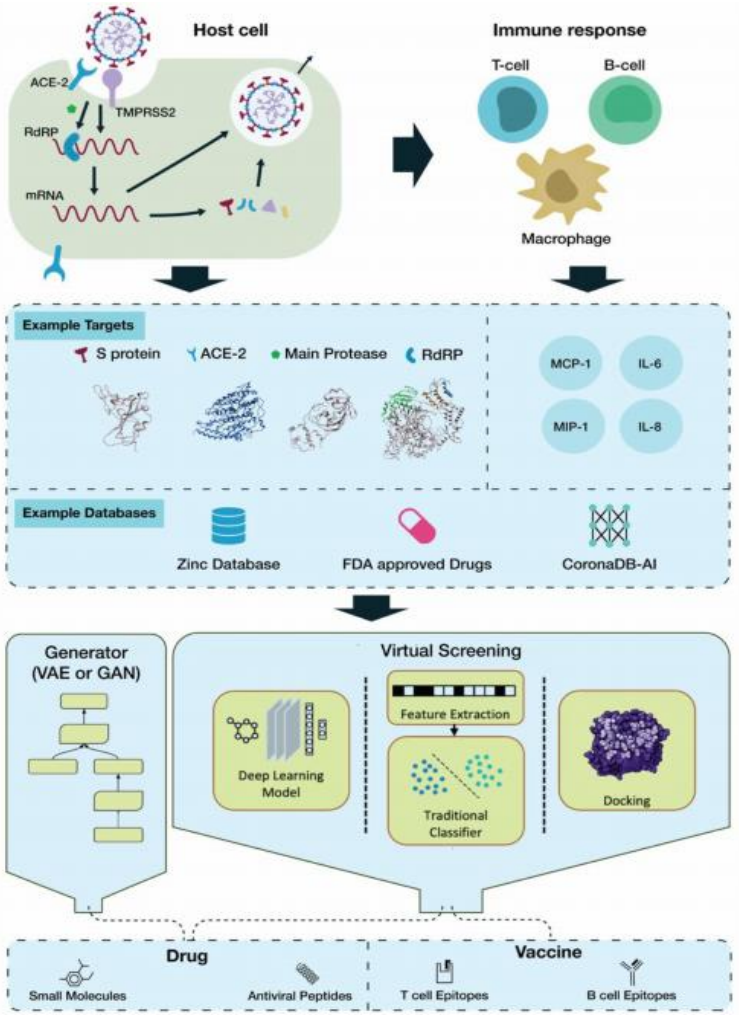
COVID-19 Treatment

SIEMENS
Ingenuity for life

Protein structures with AlphaFold



source: DeepMind



source: Arshadi *et al.* (2020)

Unrestricted © Siemens 2020

Opportunities, Challenges and Future Directions

Unrestricted © Siemens 2020

AI/ML in Healthcare

The Good

SIEMENS
Ingenuity for life

source: Nature

Feature

Apple's future healthcare market moves will rely heavily on AI analysis

By Malcolm Owen | 1 year ago

source: Apple Insider

Amazon Makes $1 Billion Splash in Health Care, Buying PillPack

By Robert Langreth and Zachary Tracer
June 28, 2018, 8:39 AM EDT *Updated on June 28, 2018, 11:06 AM EDT*

- Shares of drug retailers and pharmacy-benefit managers swoon
- Deal likely to hasten wider shakeout in the drug supply chain

source: Bloomberg

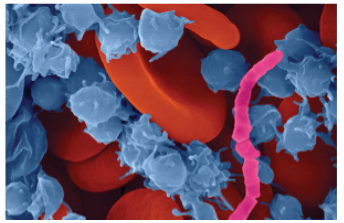
19 Oct 2018 | 15:00 GMT

Hospitals Roll Out AI Systems to Keep Patients From Dying of Sepsis

Septic shock kills 50 percent of people who are affected—Sepsis Watch could save their lives

By Eliza Strickland

In hospitals, doctors and nurses keep vigilant watch over patients' vital signs and blood tests to catch the first symptoms of sepsis. In this life-threatening condition, the body responds to an infection with widespread inflammation that can lead to organ failure. Cases can

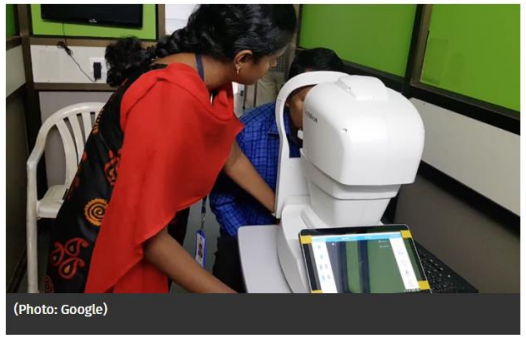


source: IEEE Spectrum

Google, Verily using AI to screen for diabetic retinopathy in India

The machine learning algorithm can also help with screening for diabetic macular edema, a boon for patients in a country where physicians are in short supply.

By Mike Miliard | February 26, 2019 | 03:17 PM



(Photo: Google)

source: Healthcare IT News

Google to Store and Analyze Millions of Health Records

The tech company's deal with Ascension is part of a push to use artificial intelligence to aid health services.

source: New York Times

Microsoft Healthcare is a new effort to push doctors to the cloud

Microsoft wants to be a big part of the cloud and AI healthcare race

By Tom Warren | @tomwarren | Jun 27, 2018, 6:50am EDT

SHARE

source: The Verge

Unrestricted © Siemens 2020

AI/ML in Healthcare

The Bad

SIEMENS
Ingenuity for life

Unmet Expectations

source: Strickland (2019)

Data Security and Privacy

source: Medical Device Network

Bias and Fairness

source: PCMag

Adversarial Attacks

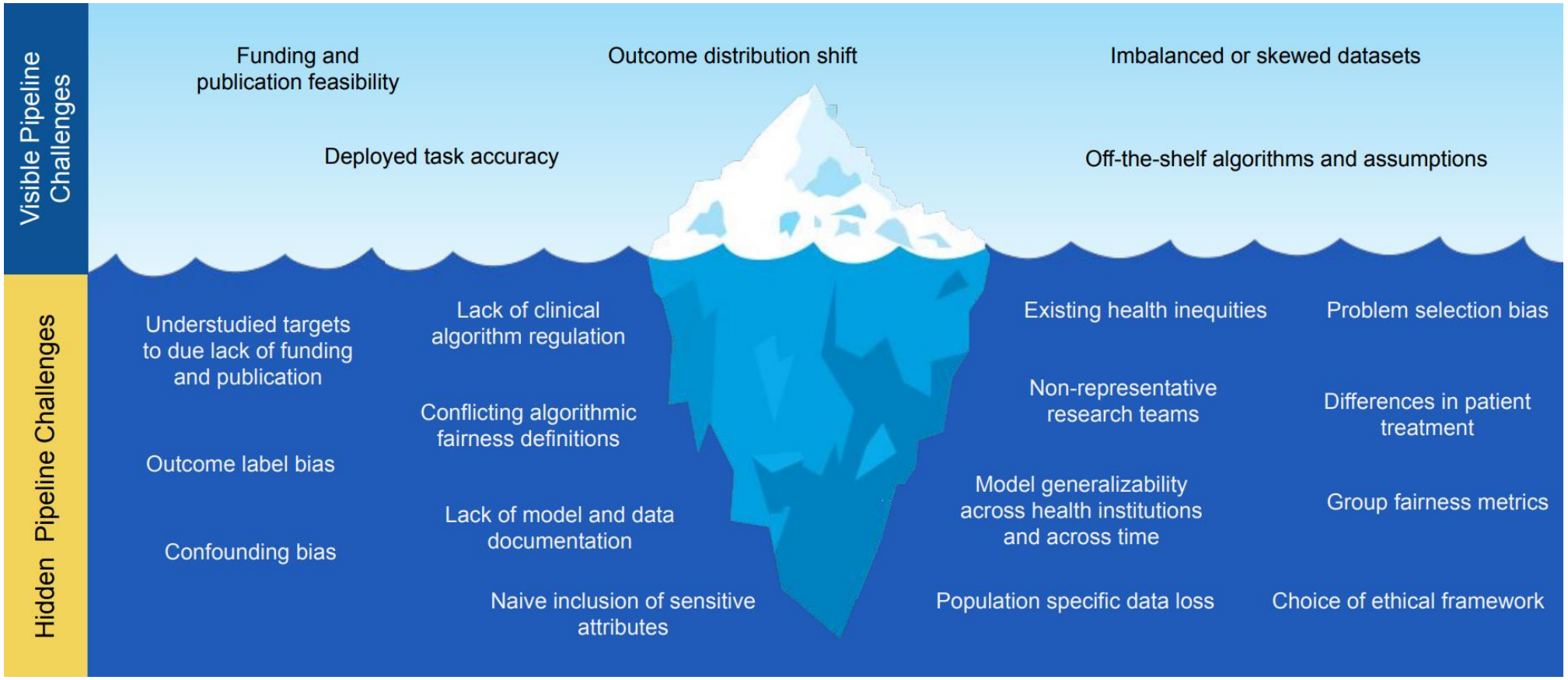
source: New York Times

Unrestricted © Siemens 2020

Opportunities

ML and Ethics in Healthcare

SIEMENS
Ingenuity for life

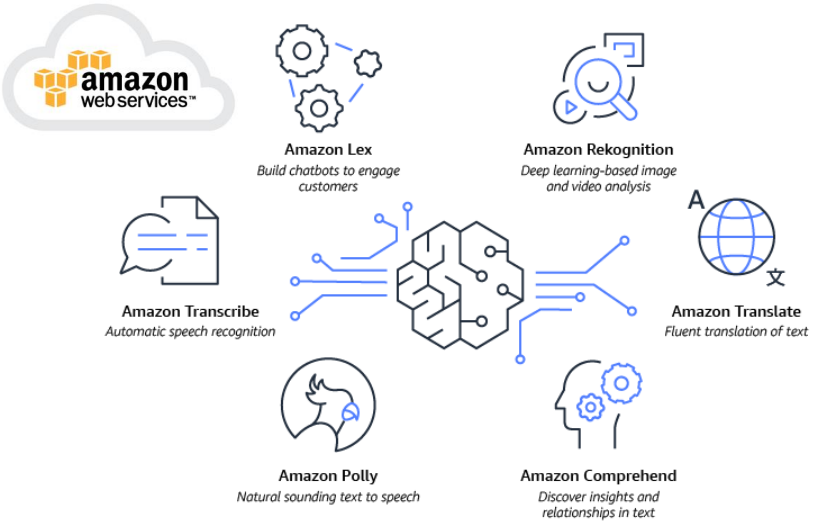


source: Chen *et al.* (2020)

Unrestricted © Siemens 2020

Opportunities

ML and the Cloud ☁️

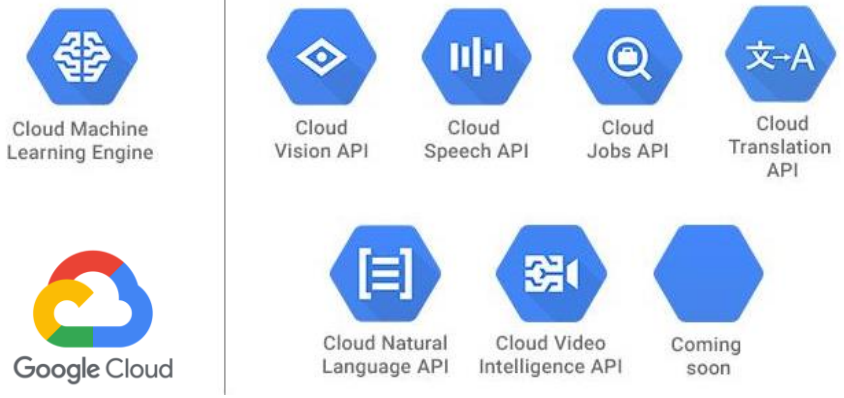


source: AWS

Azure Machine Learning

source: Azure

source: Azure

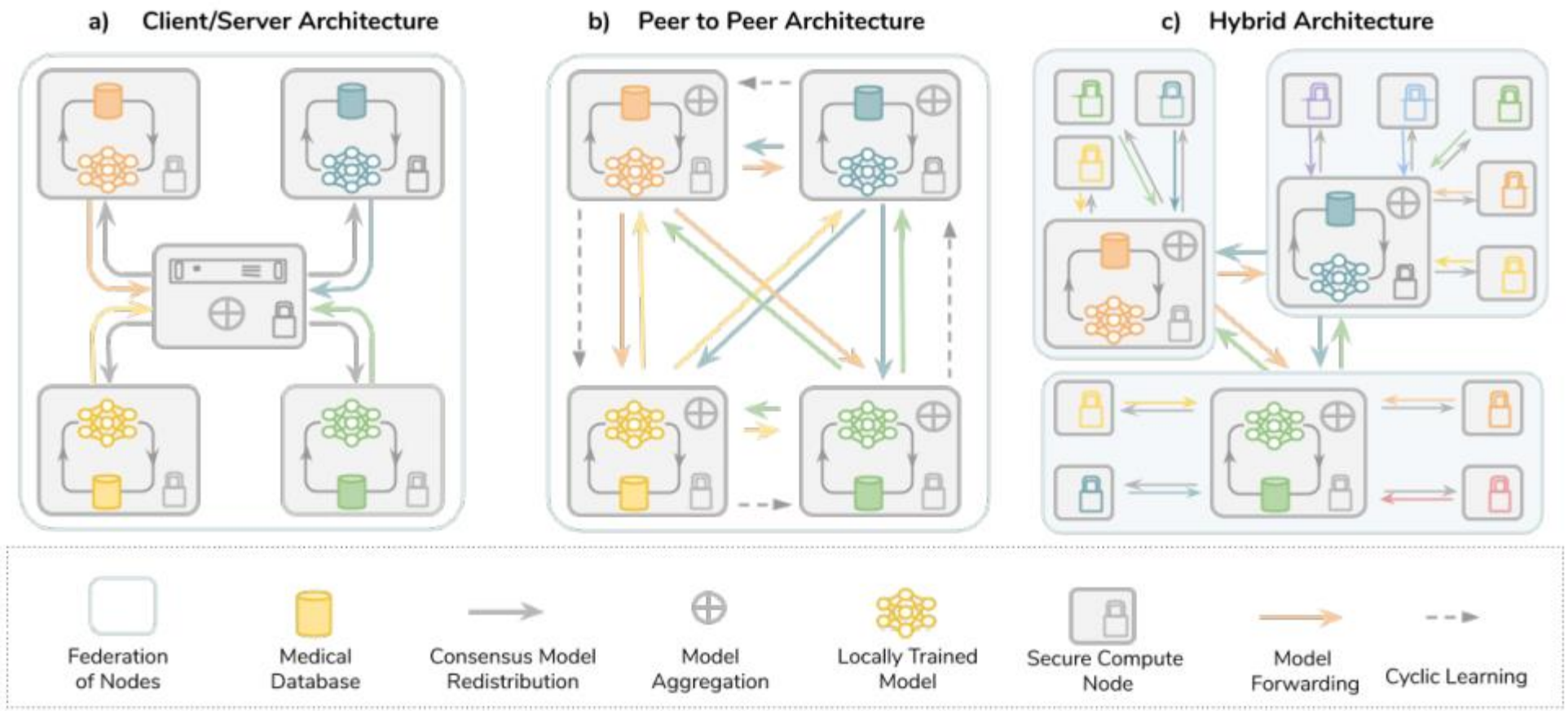


source: GCP

Unrestricted © Siemens 2020

Opportunities

Federated Learning in Healthcare



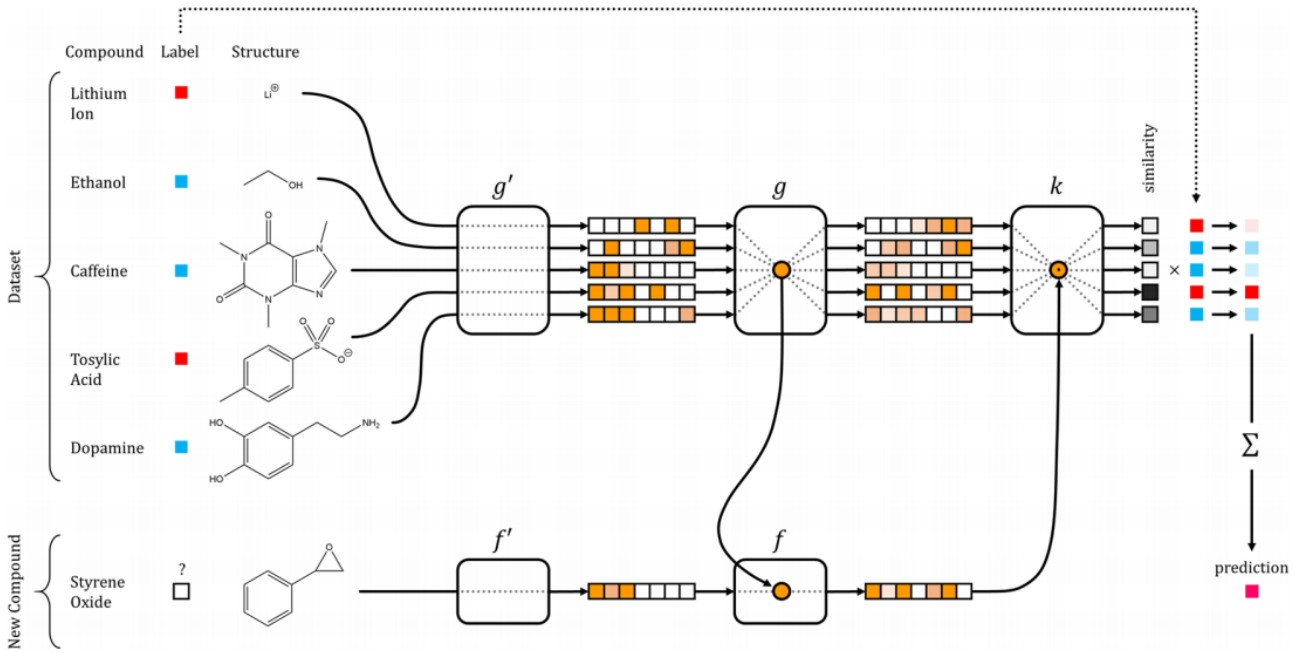
source: Rieke *et al.* (2020)

Unrestricted © Siemens 2020

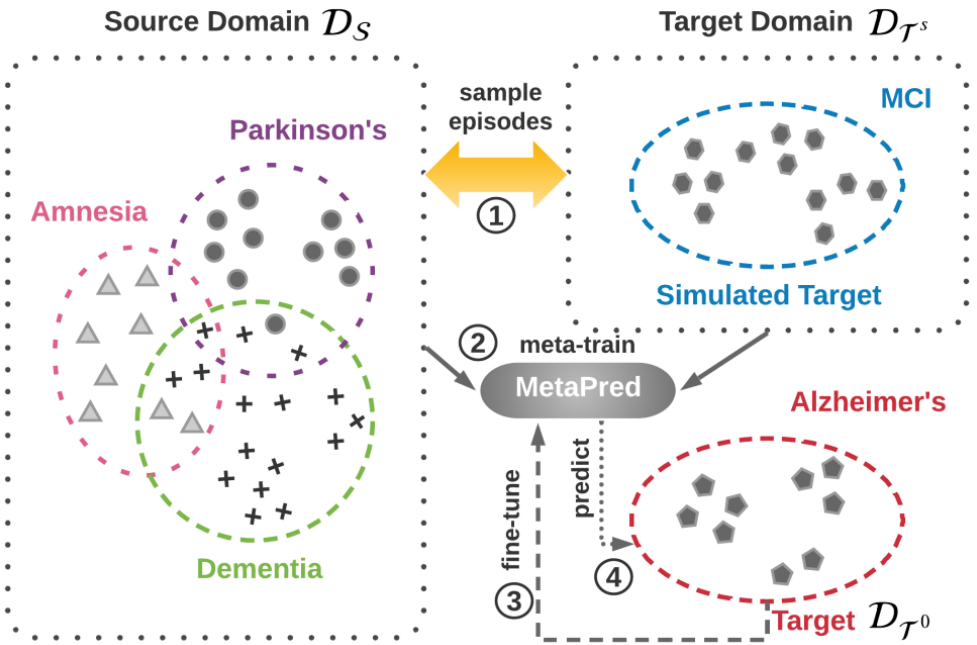
Opportunities

Meta-Learning in Healthcare

SIEMENS
Ingenuity for life



source: Altae-Tran *et al.* (2017)



source: Zhang *et al.* (2019)

Unrestricted © Siemens 2020

Future Directions

Greener Models

SIEMENS
Ingenuity for life

AI in the 2020s Must Get Greener—and Here's How

The push for energy efficient "Green AI" requires new strategies

By **Ameet Talwalkar**

source: Spectrum IEEE

Shrinking deep learning's carbon footprint

Through innovation in software and hardware, researchers move to reduce the financial and environmental costs of modern artificial intelligence.

Kim Martineau | MIT Quest for Intelligence
August 7, 2020

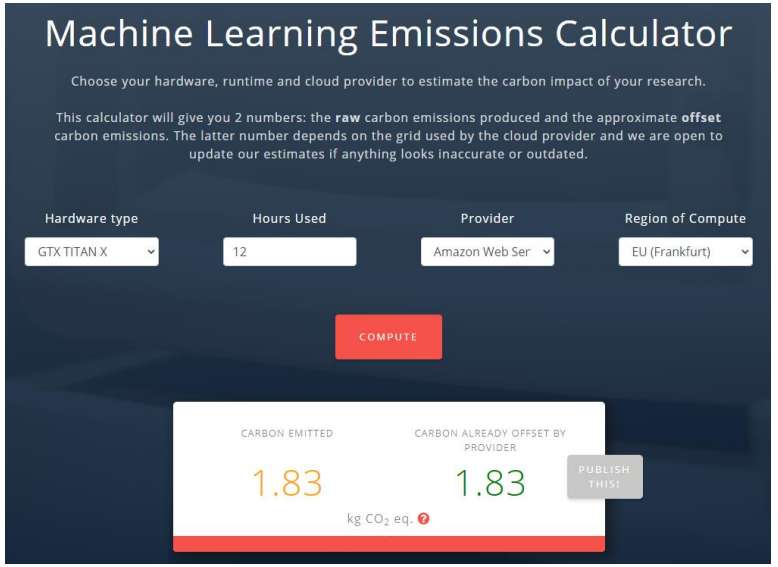
source: MIT News

Jun 17, 2020, 11:54am EDT

Deep Learning's Carbon Emissions Problem

Rob Toews Contributor
AI
I write about the big picture of artificial intelligence.

source: Forbes



ML CO2 Impact

Consumption	CO_2e (lbs)
Air travel, 1 passenger, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000

Training one model (GPU)	
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468
Transformer (big)	192
w/ neural architecture search	626,155

Table 1: Estimated CO_2 emissions from training common NLP models, compared to familiar consumption.[1]

source: Strubell, Ganesh & McCallum (2019)

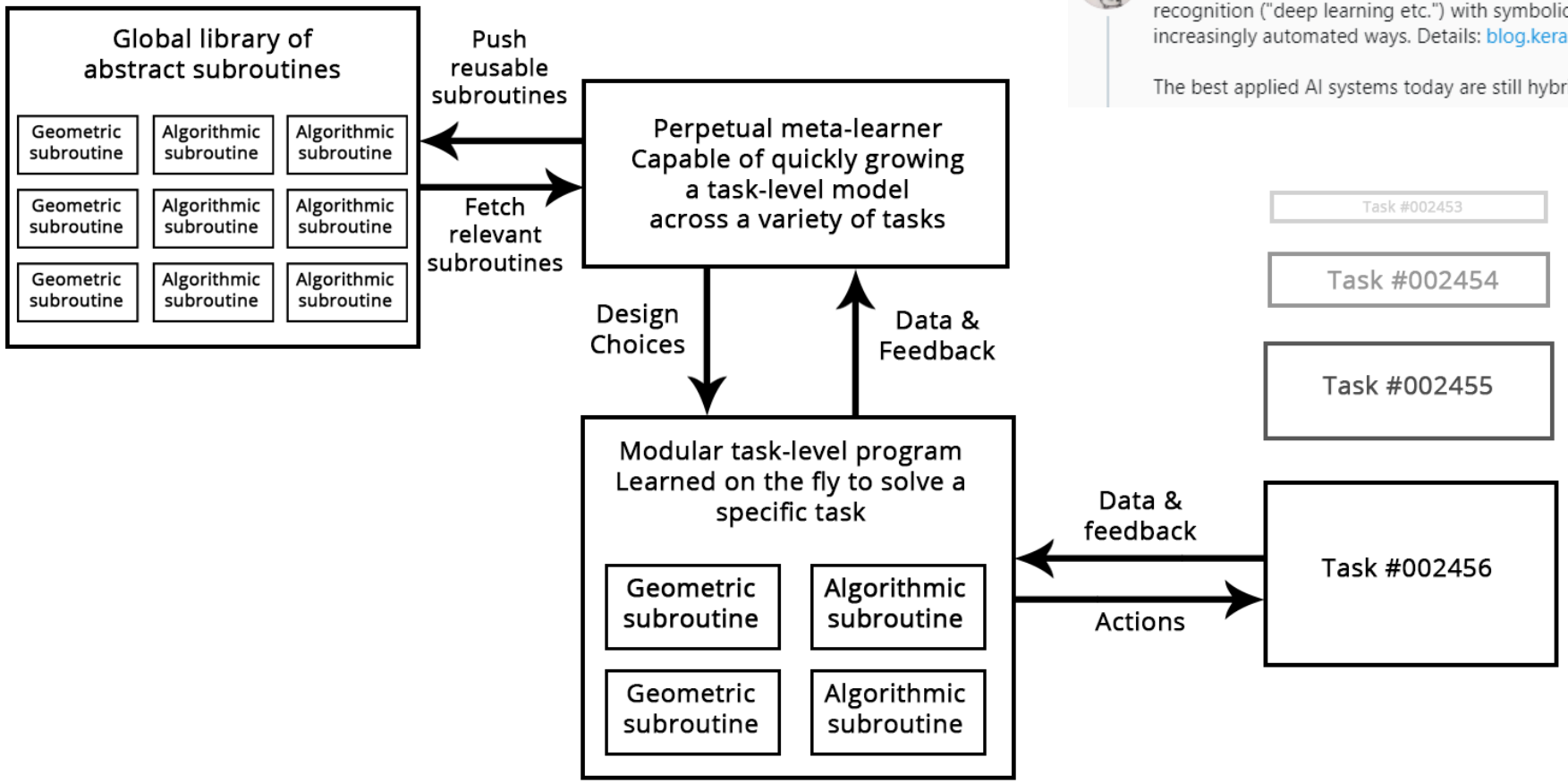
Homework: How many GPU time would it take to train GPT-3?

Unrestricted © Siemens 2020

Future Directions

Pattern Recognition + Symbolic AI

SIEMENS
Ingenuity for life



source: Chollet (2017)

source: Twitter

Unrestricted © Siemens 2020

Future Directions

MLOps

Homework: read this article

Hidden Technical Debt in Machine Learning Systems

D. Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips
{dsculley,gholt,dgg,edavydov,toddphillips}@google.com
Google, Inc.

Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-François Crespo, Dan Dennison
{ebner,vchaudhary,mwyoung,jfcrespo,dennison}@google.com
Google, Inc.

Abstract

Machine learning offers a fantastically powerful toolkit for building useful complex prediction systems quickly. This paper argues it is dangerous to think of these quick wins as coming for free. Using the software engineering framework of *technical debt*, we find it is common to incur massive ongoing maintenance costs in real-world ML systems. We explore several ML-specific risk factors to account for in system design. These include boundary erosion, entanglement, hidden feedback loops, undeclared consumers, data dependencies, configuration issues, changes in the external world, and a variety of system-level anti-patterns.

1 Introduction

As the machine learning (ML) community continues to accumulate years of experience with live systems, a wide-spread and uncomfortable trend has emerged: developing and deploying ML systems is relatively fast and cheap, but maintaining them over time is difficult and expensive.

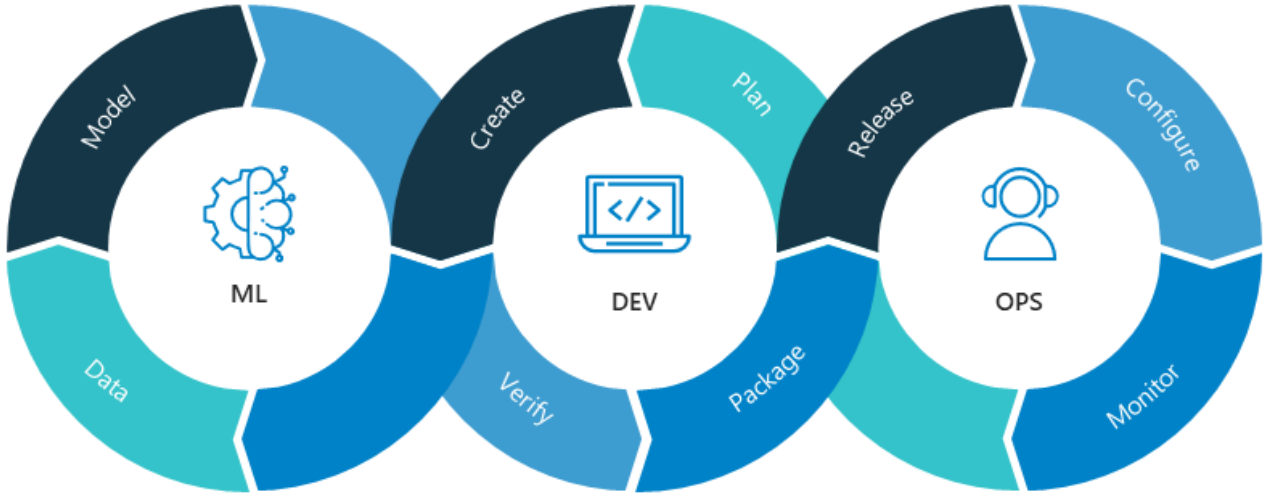
This dichotomy can be understood through the lens of *technical debt*, a metaphor introduced by Ward Cunningham in 1992 to help reason about the long term costs incurred by moving quickly in software engineering. As with fiscal debt, there are often sound strategic reasons to take on technical debt. Not all debt is bad, but all debt needs to be serviced. Technical debt may be paid down by refactoring code, improving unit tests, deleting dead code, reducing dependencies, tightening APIs, and improving documentation [8]. The goal is *not* to add new functionality, but to enable future improvements, reduce errors, and improve maintainability. Deferring such payments results in compounding costs. Hidden debt is dangerous because it compounds silently.

In this paper, we argue that ML systems have a special capacity for incurring technical debt, because they have all of the maintenance problems of traditional code plus an additional set of ML-specific issues. This debt may be difficult to detect because it exists at the *system* level rather than the code level. Traditional abstractions and boundaries may be subtly corrupted or invalidated by the fact that data influences ML system behavior. Typical methods for paying down code level technical debt are not sufficient to address ML-specific technical debt at the system level.

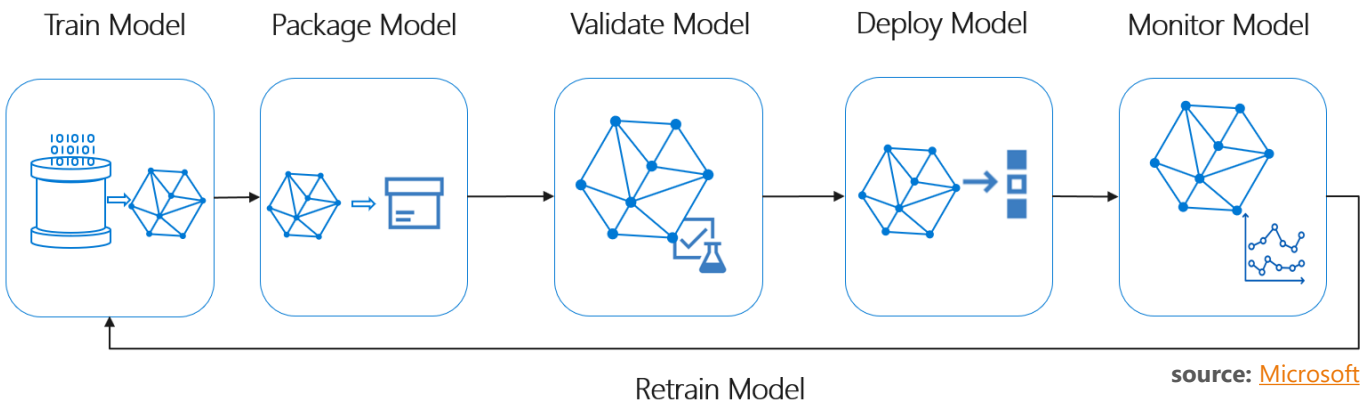
This paper does not offer novel ML algorithms, but instead seeks to increase the community's awareness of the difficult tradeoffs that must be considered in practice over the long term. We focus on system-level interactions and interfaces as an area where ML technical debt may rapidly accumulate. At a system-level, an ML model may silently erode abstraction boundaries. The tempting re-use or chaining of input signals may unintentionally couple otherwise disjoint systems. ML packages may be treated as black boxes, resulting in large masses of "glue code" or calibration layers that can lock in assumptions. Changes in the external world may influence system behavior in unintended ways. Even monitoring ML system behavior may prove difficult without careful design.

1

source: Sculley *et al.* (2015)



source: Neal Analytics



source: Microsoft

Unrestricted © Siemens 2020

Future Directions

Quantum Machine Learning

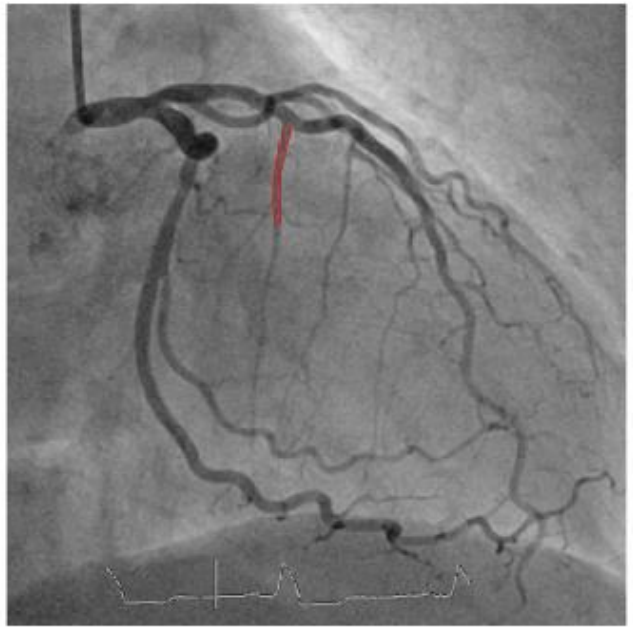
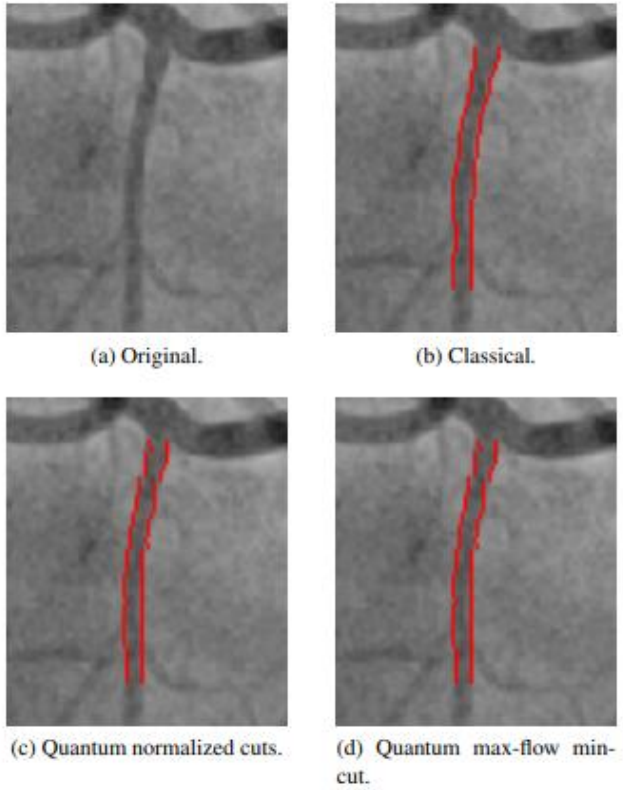


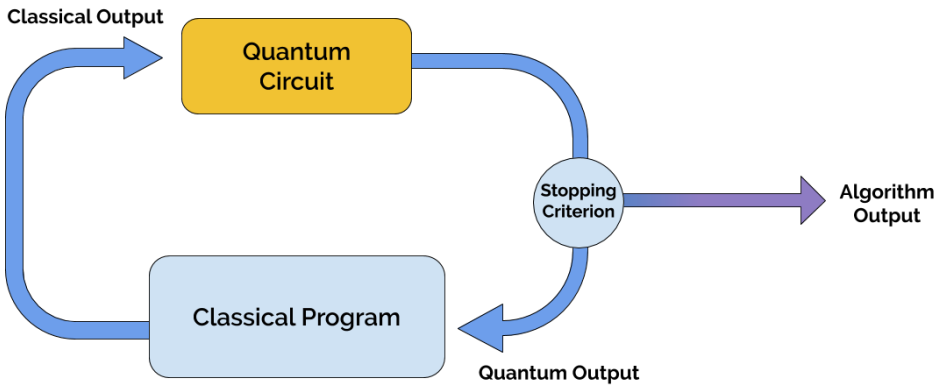
Figure 10: Quantum segmentation result, with the boundaries of the segmentation marked in red on the entire image for the normalized cuts approach.



(a) Original. (b) Classical. (c) Quantum normalized cuts. (d) Quantum max-flow min-cut.

Figure 11: Comparison of the classical and quantum segmentation results of the artery, with the boundaries of the segmentation marked in red. The result is given for one of the 3 runs.

source: Tse *et al.* (2018)



source: Entropica Labs

Unrestricted © Siemens 2020

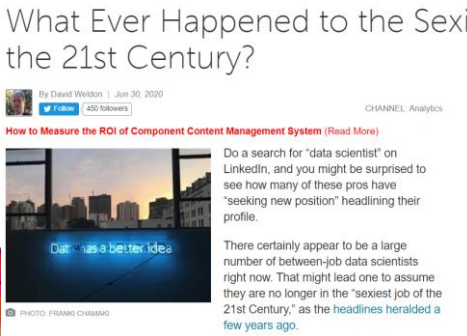
Caution

Hype, Hubris and Hogwash

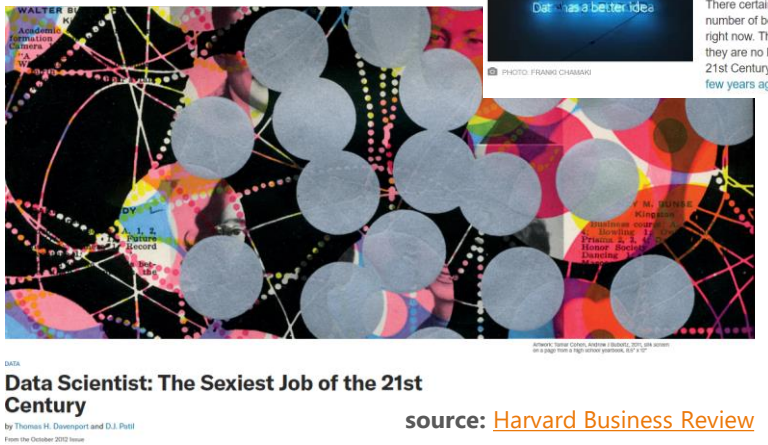
SIEMENS
Ingenuity for life

Hype Cycle for Emerging Technologies (2020)

source: Gartner



source: cmswire



source: Harvard Business Review

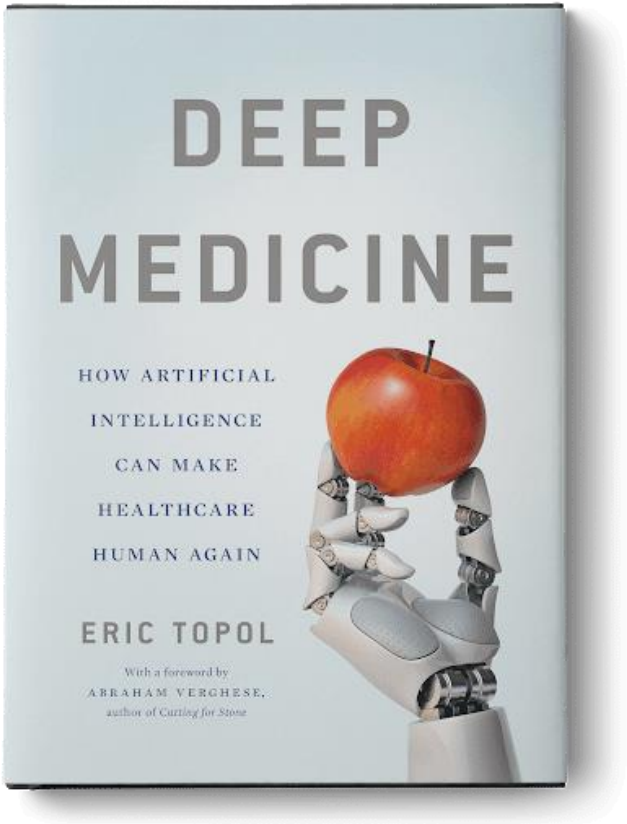
source: Independent

Unrestricted © Siemens 2020

Final Thoughts

SIEMENS
Ingenuity for life

"The greatest opportunity offered by AI is not reducing errors or workloads, or even curing cancer: ***it is the opportunity to restore the precious and time-honored connection and trust—the human touch—between patients and doctors****. Not only would we have more time to come together, enabling far deeper communication and compassion, but also we would be able to revamp how we select and train doctors."*



Unrestricted © Siemens 2020

Final Thoughts

SIEMENS
Ingenuity for life

"If we don't know what we want we're less likely to get it."

source: Tegmark (2017)

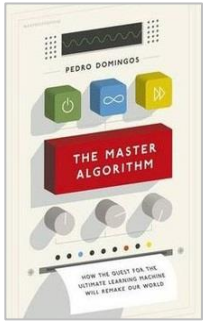
Unrestricted © Siemens 2020

References

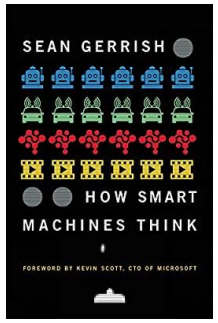
Books

SIEMENS
Ingenuity for life

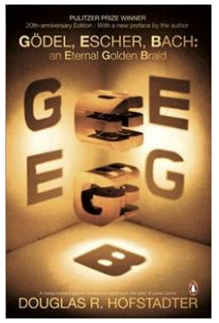
General



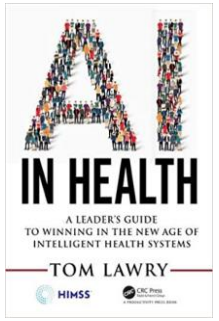
Domingos, P. (2015)
The Master Algorithm
1st edition
Basic Books



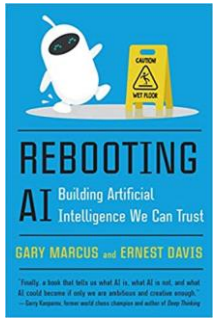
Gerrish, S. (2018)
How Smart Machines Think
1st edition
MIT Press



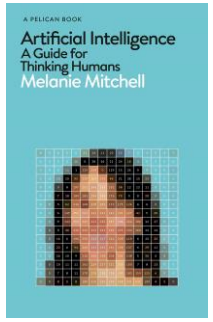
Hofstadter, D. (1999)
Gödel, Escher, Bach:
An Eternal Godel Braid
20th Anniversary Edition
Penguin Books



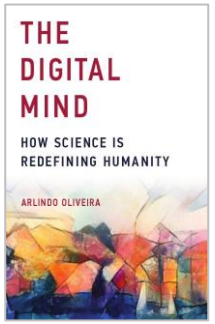
Lawry, T. (2020)
AI in Health
1st edition
Routledge



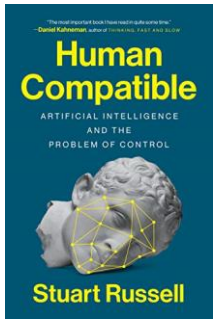
Marcus, G. & Davis, E. (2019)
Rebooting AI
1st edition
Pantheon



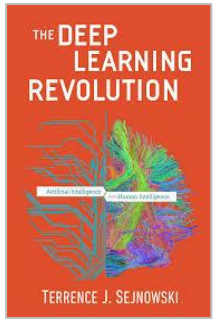
Mitchell, M. (2019)
Artificial Intelligence
1st edition
Pelican Books



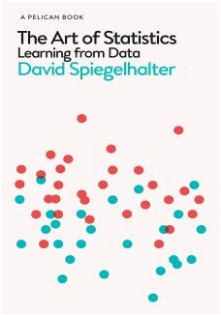
Oliveira, A. (2017)
The Digital Mind
1st edition
MIT Press



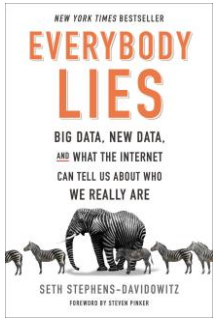
Russell, S. (2019)
Human Compatible
1st edition
Viking



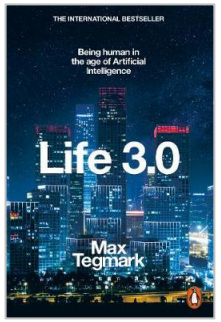
Sejnowski, T. (2018)
The Deep Learning Revolution
1st edition
MIT Press



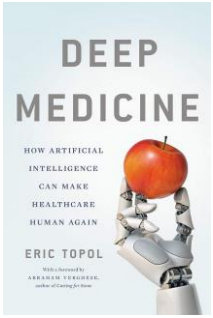
Spiegelhalter, D. (2019)
The Art of Statistics
1st edition
Pelican Books



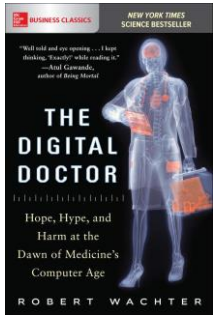
Stephens-Davidowitz, S. (2017)
Everybody Lies
1st edition
Dey Street Books



Tegmark, M. (2017)
Life 3.0
1st edition
Vintage



Topol, E. (2019)
Deep Medicine
1st edition
Basic Books



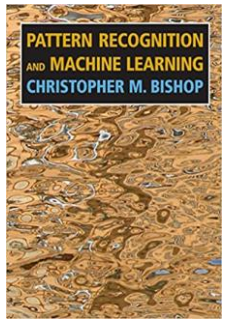
Wachter, R. (2015)
The Digital Doctor
1st edition
McGraw-Hill Education

Unrestricted © Siemens 2020

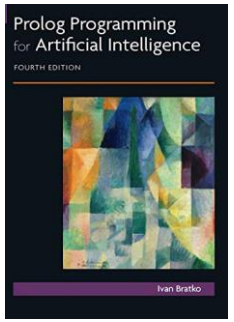
Books

SIEMENS
Ingenuity for life

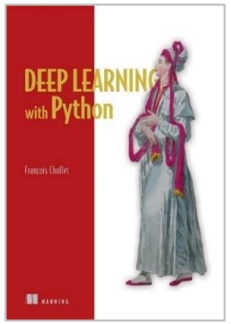
Advanced



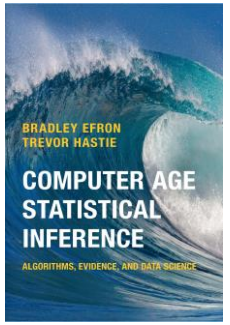
Bishop, C. (2006)
Pattern Recognition and Machine Learning
1st edition
Springer



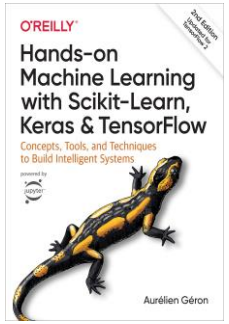
Bratko, I. (2011)
Prolog Programming for Artificial Intelligence
4th edition
Addison-Wesley



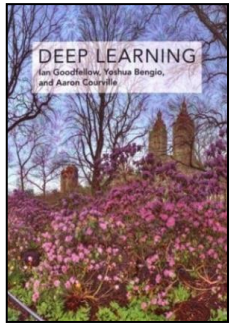
Chollet, F. (2017)
Deep Learning with Python
1st edition
Manning



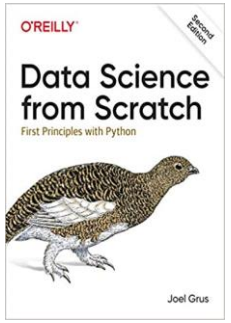
Efron, B. & Hastie, T. (2016)
Computer Age Statistical Inference
1st edition
Cambridge University Press



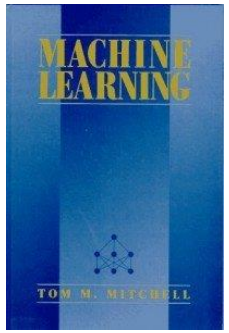
Géron, A. (2017)
Hands-On Machine Learning with Scikit-Learn and Tensorflow
2nd edition
O'Reilly



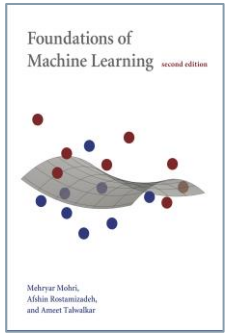
Goodfellow, I., Bengio, Y. & Courville, A. (2016)
Deep Learning
1st edition
MIT Press



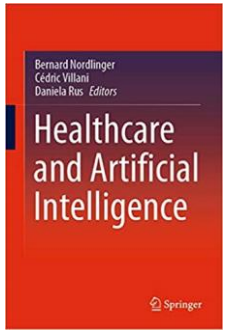
Grus, J. (2019)
Data Science from Scratch: First Principles with Python
2nd edition
O'Reilly



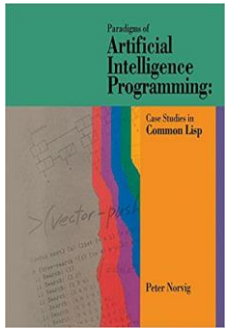
Mitchell, T. (1997)
Machine Learning
1st edition
MacGraw-Hill



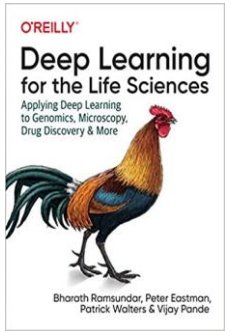
Mohri, M., Rostamizadeh, A. & Talwalkar, A. (2018)
Foundations of Machine Learning
2nd edition
MIT Press



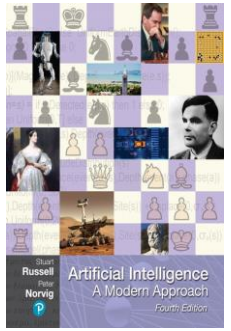
Nordlinger, B., Villani, C. & Rus, D. (2020)
Healthcare and Artificial Intelligence
1st edition
Springer



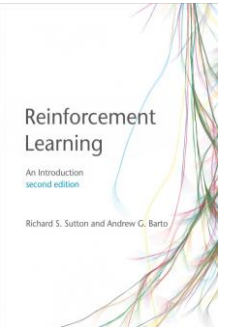
Norvig, P. (1992)
Paradigms of Artificial Intelligence
1st edition
Morgan Kaufmann



Ramsundar, B. *et al.* (2020)
Deep Learning for the Life Sciences
1st edition
O'Reilly



Russell, S. & Norvig, P. (2020)
Artificial Intelligence: A Modern Approach
4th edition
Prentice Hall



Sutton, R. & Barto, A. (2018)
Reinforcement Learning: An Introduction
2nd edition
MIT Press

Unrestricted © Siemens 2020

Articles

SIEMENS
Ingenuity for life

Altae-Tran, H. et al. (2017). Low data drug discovery with one-shot learning. *ACS Cent. Sci.*, 3: 283-293

Arcadu, F. *et al.* (2019). Deep learning algorithm predicts diabetic retinopathy progression in individual patients. *Npj Digit. Med.*, 2: 92-101

Arshadi, A. *et al.* (2020). Artificial intelligence for COVID-19 drug discovery and vaccine development. *Front. Artif. Intell.*, 3: 65

Biamonte, J. *et al.* (2018). Quantum Machine Learning. *Nature*, 549: 195-202

Chen, I. *et al.* (2020). Ethical machine learning in healthcare. *Retrieved from* arXiv:2009.10576

Ching, T. *et al.* (2018). Opportunities and obstacles for deep learning in biology and medicine. *J. R. Soc. Interface*, 15: 20170387

Domingos, P. (2012). A few useful things to know about machine learning. *Comm. ACM*, 55(10): 78-87

Esteva, A. *et al.* (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542: 115-118

Esteva, A. *et al.* (2019). A guide to deep learning in healthcare. *Nat. Med.*, 25: 24-29

Garnelo, M., Arulkumaran, K. & Shanahan, M. (2016). Towards deep symbolic reinforcement learning. *Retrieved from* arXiv:1609.05518

Ghassemi, M. *et al.* (2018). A review of challenges and opportunities in machine learning for health. *Retrieved from* arXiv:1806.00388

Gottesman, O. *et al.* (2019). Guidelines for reinforcement learning in healthcare. *Nat. Med.*, 25: 14-18

Harmon, S. *et al.* (2020). Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat. Commun*. 11: 4080

Harutyunyan, H. *et al.* (2019). Multitask learning and benchmarking with clinical time series data. *Sci. Data*, 6: 96

Hassabis, D. *et al.* (2017). Neuroscience-inspired artificial intelligence. *Neuron*, 95(2): 245-258

Jiang, F. *et al.* (2017). Artificial intelligence in healthcare: past, present and future. *Stroke Vasc. Neurol.*, 2: e000101

Kalis, B., Collier, M. & Fu, R. (2018). 10 promising AI applications in health care. *Retrieved from* Harvard Business Review

Komorowski, M. *et al.* (2018). The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nat. Med.*, 24: 1716-1720

Libbrecht, M. & Noble, S. (2015). Machine learning applications in genetics and genomics. *Nat. Rev. Genet.*, 16: 321-332

Lillicrap, T. *et al.* (2020). Backpropagation and the brain. *Nat. Rev. Neurosci.*, 21(6): 335-346

Litjens, G. *et al.* (2017). A survey on deep learning in medical image analysis. *Retrieved from* arXiv:1702.05747

Lotte, F. *et al.* (2018). A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *J. Neural Eng.*, 15: 031005

Majkowska, A. *et al.* (2020). Chest radiograph interpretation with deep learning models: assessment with radiologist-adjudicated reference standards and population-adjusted evaluation. *Radiology*, 294(2): 421-431

Mnih, V. *et al.* (2015). Human-level control through deep reinforcement learning. *Nature*, 518: 529-533

Nam, C. *et al.* (2018). *Brain-Computer Interfaces Handbook: The Technological and Theoretical Advances*. Boca Raton, FL: CRC Press

Panch, T., Szolovits, P. & Atun, R. (2019). Artificial intelligence, machine learning and health systems. *J. Glob. Health*, 8: 020303

Piwek, L. *et al.* (2016). The rise of consumer health wearables: promises and barriers. *PLOS Medicine*, 13(2): e1001953

Unrestricted © Siemens 2020

Articles

SIEMENS
Ingenuity for life

Rajkomar, A., Dean, J. & Kohane, I. (2019). Machine learning in medicine. *N. Engl. J. Med.*, 380: 1347-1358

Raman, R. *et al.* (2019). Fundus photograph-based deep learning algorithms in detecting diabetic retinopathy. *Eye*, 33: 97-109

Ravì, D. *et al.* (2017). Deep learning for health informatics. IEEE *Journ. Biom. Health Inform.*, 21(1): 4-21

Ribeiro, M., Singh, S. & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Retrieved from* arXiv:1602.04938

Rieke, N. *et al.* (2020). The future of digital health with federated learning. *Retrieved from* arXiv:2003:08119

Samuel, A. (1959). Some studies in machine learning using the game of checkers. *IBM Journal*, 3(3): 535-554

Schmidt, R., Schneider, F. & Hennig, P. (2020). Descending through a crowded valley: benchmarking deep learning optimizers. *Retrieved from* arXiv:2005.01547

Sculley, D. *et al.* (2015). Hidden technical debt in machine learning systems. *Adv. Neur. Inform. Proc. Syst.*, 28: 2503-2511

Sheller, M. *et al.* (2020). Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports*, 10(1): 1-12

Simeone, O. (2017). A brief introduction to machine learning for engineers. *Retrieved from* arXiv:1709.02840

Steiner, D. *et al.* (2018). Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. *Am. J. Surg. Pathol.*, 42(12): 1636-1646

Strickland, E. (2019). IBM Watson, heal thyself: how IBM Watson overpromised and underdelivered on AI healthcare. *Retrieved from* IEEE spectrum

Strubell, E., Ganesh, A. & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *Retrieved from* arXiv:1906.02243

Szolovits, P. (1982). Artificial Intelligence in Medicine. *Retrieved from* Peter Szolovits' personal website (CSAIL)

Topol, E. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25: 44-56

Tse, L. *et al.* (2018). Graph cut segmentation methods revisited with a quantum algorithm. *Retrieved from* arXiv:1812.03050

Wiens, J. & Shenoy, E. (2017). Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clin. Infect. Dis.*, 66: 149-153

Wolpert, D. (1996). The lack of a priori distinctions between learning algorithms. *Neur. Comp.*, 8: 1341-1390

Wu, N. *et al.* (2019). Deep neural networks improve radiologists' performance in breast cancer screening. *Retrieved from* arXiv:1903.08297

Yan, L. *et al.* (2020). An interpretable mortality prediction model for COVID-19 patients. *Nat. Mach. Intell.*, 2: 283-288

Yu, K. *et al.* (2018). Artificial intelligence in healthcare. *Nat. Biomed. Eng.*, 2: 719-731

Yu, W. *et al.* (2010). Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. *BMC Med. Inform. Decis.*, 10: 16-23

Zhang *et al.* (2019). MetaPred: meta-learning for clinical risk predictions with limited patient electronic health records. *Retrieved from* arXiv:1905.03218

Zhao, A. et al. (2019). Data augmentation using learned transforms for one-shot medical image segmentation. *Retrieved from* arXiv:1902.09383

Zhou, X. *et al.* (2020). Artificial intelligence in surgery. *Retrieved from* arXiv:2001.00627

Zou, J. *et al.* (2019). A primer on deep learning in genomics. *Nat. Gen.*, 51: 12-18

Unrestricted © Siemens 2020

Links

SIEMENS
Ingenuity for life

Courses

- (Berkeley) CS 189/289A: Introduction to Machine Learning
- (deeplearning.ai) Andrew Ng's AI for Medicine Course
- (Harvard) CS 181: Machine Learning
- (MIT) Brains, Minds & Machines Summer Course
- (MIT) Deep Learning and Artificial Intelligence lectures
- (MIT) 6.S897: Machine Learning for Healthcare
- (MIT) 6.874: Deep Learning in the Life Sciences
- (Princeton) COS597C: Machine Learning for Health Care
- (Stanford) BIODS220: Artificial Intelligence in Healthcare
- (Stanford) CS231n: CNNs for Visual Recognition
- (UCL) David Silver's course on Reinforcement Learning
- (UToronto) APS360 Artificial Intelligence Fundamentals

Repositories

- (GitHub) Awesome MLOps
- (GitHub) Awesome Machine Learning
- (GitHub) Awesome Production Machine Learning

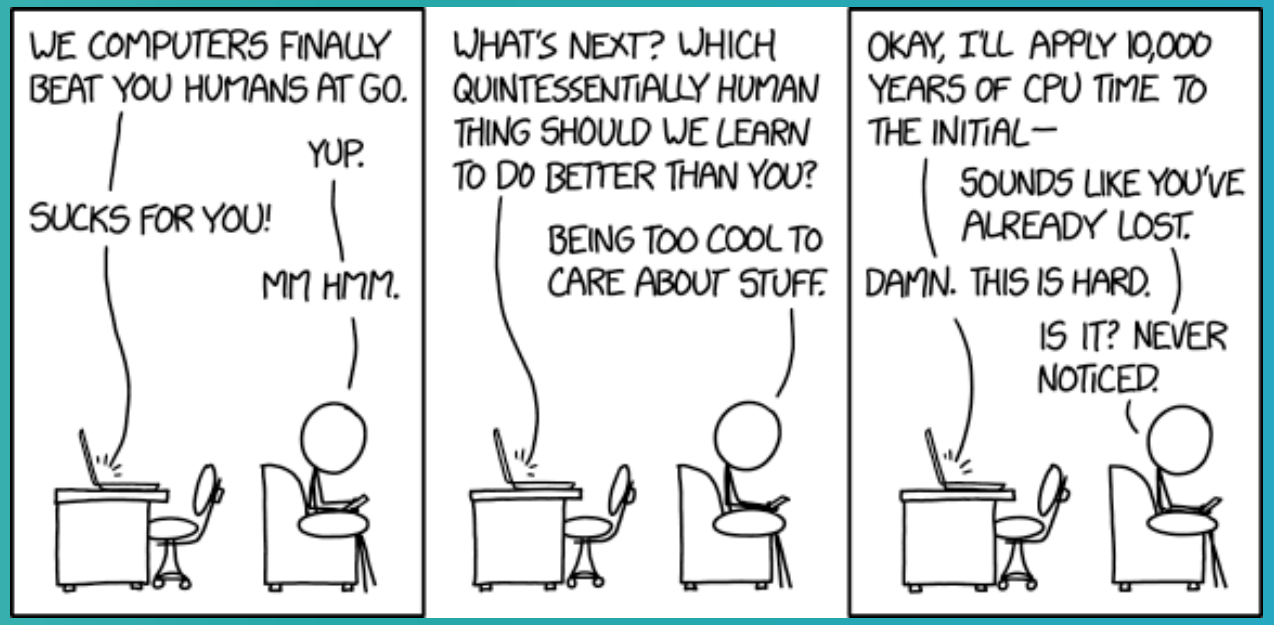
Groups

- Siemens AI Lab
- Siemens Machine Intelligence Research Group
- Volkswagen Group ML Research
- MIT Clinical Machine Learning Group
- MIT CSAIL Clinical Decision Making Group
- Center for Artificial Intelligence in Medicine & Imaging

Miscellaneous

- (OpenAI) AI and Compute
- (Siemens) ML/DL 2019 Whitepaper
- Learn X in Y minutes

Unrestricted © Siemens 2020

SIEMENS
Ingenuity for life

Q&A