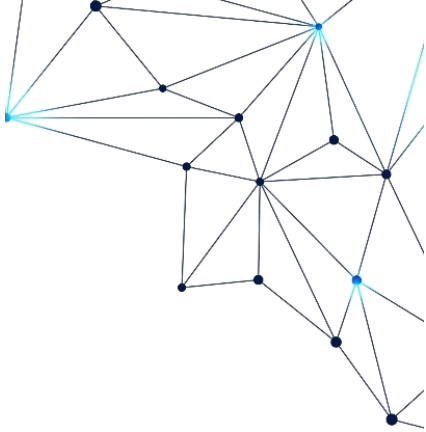
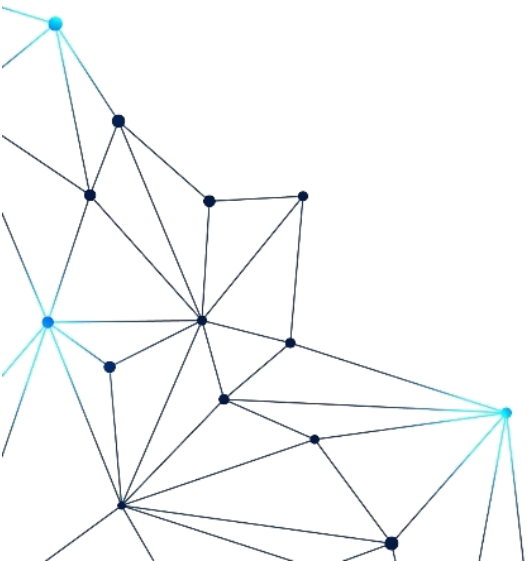


Introduction to AI

Alexandre Guilbault, P.Eng, MBA

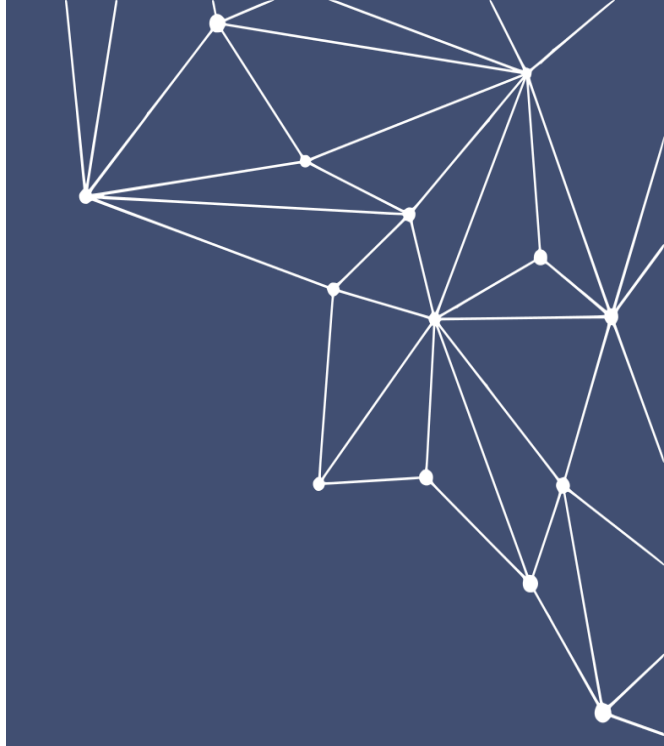


March 12th, 2019

RectifAI

Objectives

- Understand where AI is coming from
- Understand it’s different applications
- Get a sense of it’s mechanic

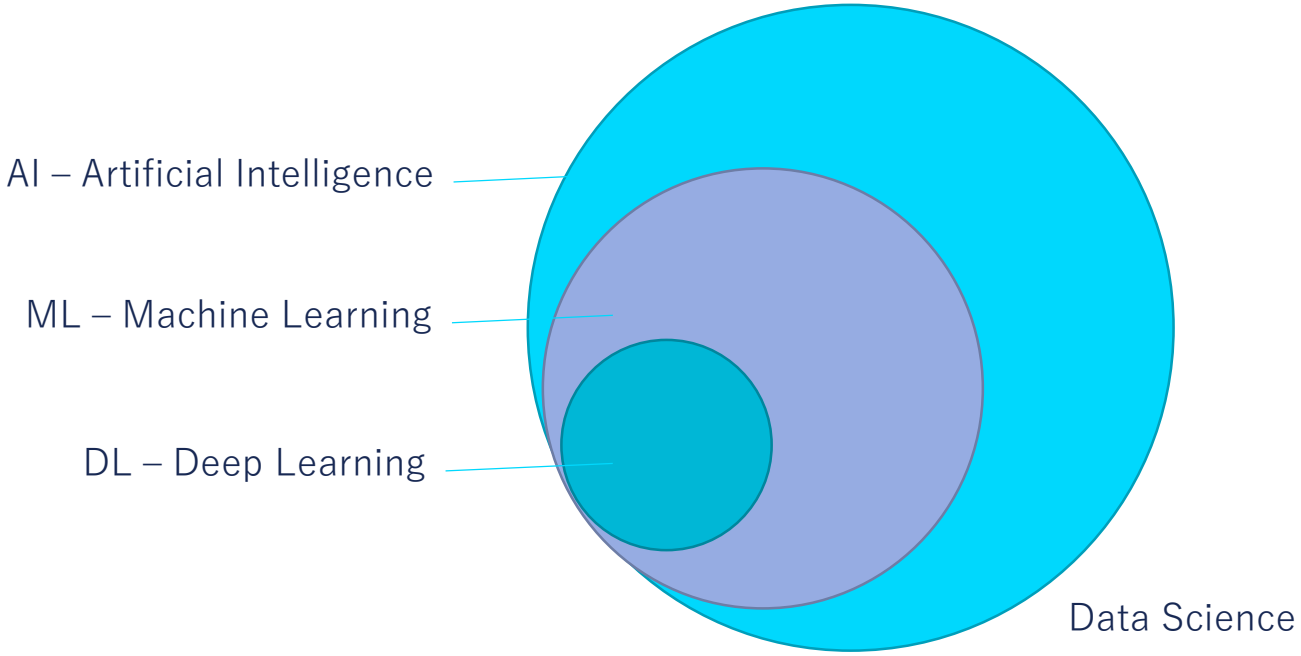


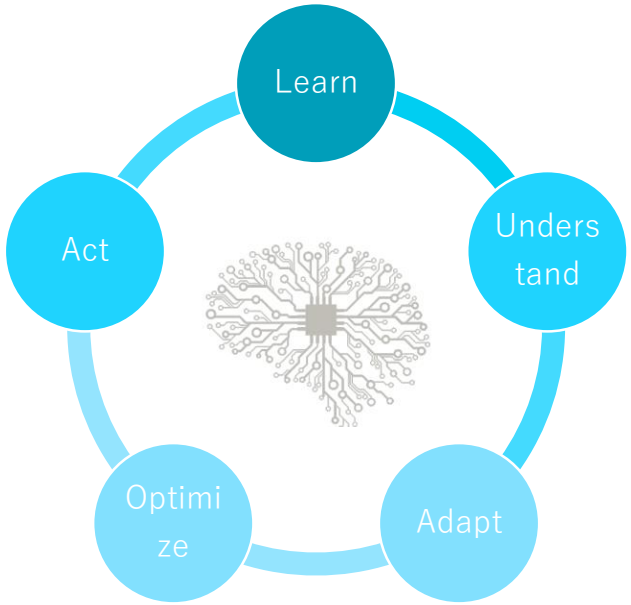
Agenda

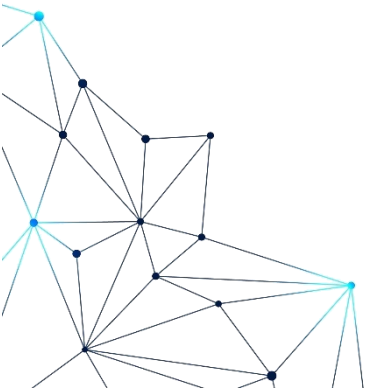
1. General introduction
2. Setting up an AI project
3. Some interesting use cases
4. Classic vs. modern approaches
5. How it works
6. Q&A

General Introduction to AI

What is Machine Learning





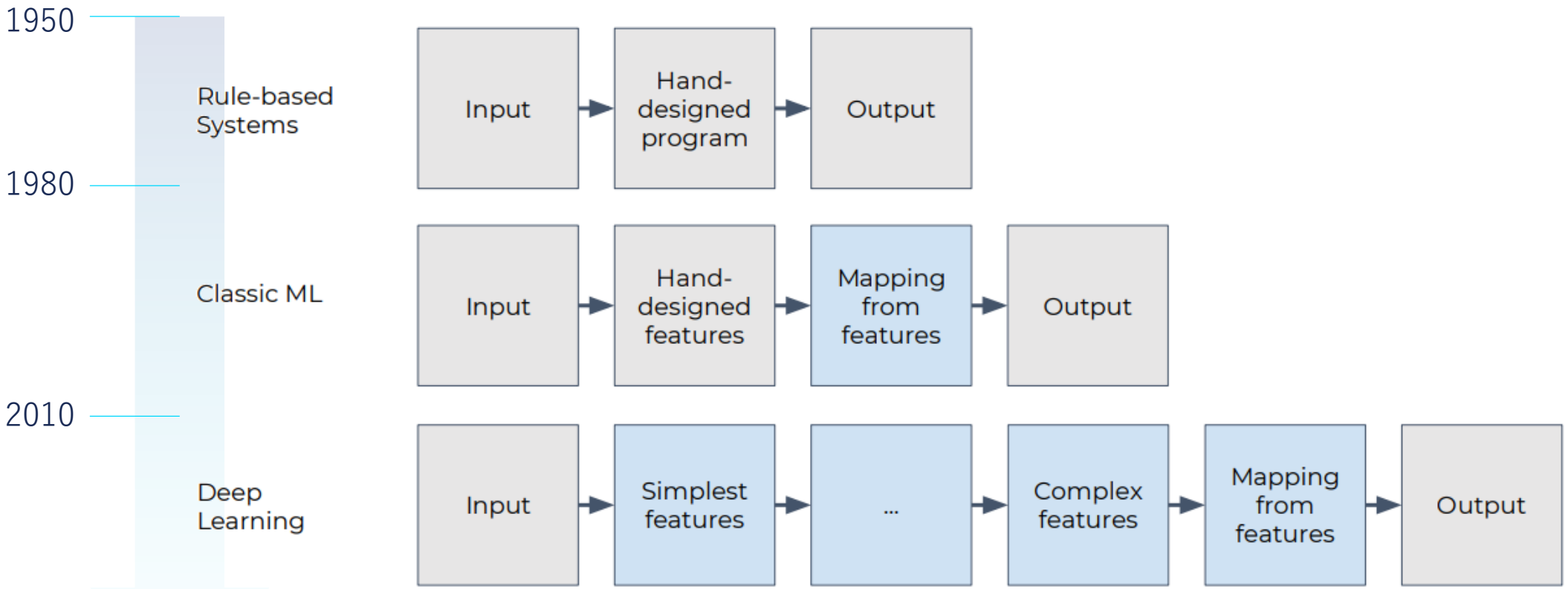


Machine learning gives computers the ability to learn without being explicitly programmed

Arthur Samuel, 1959

RectifAI

Machine Learning evolution



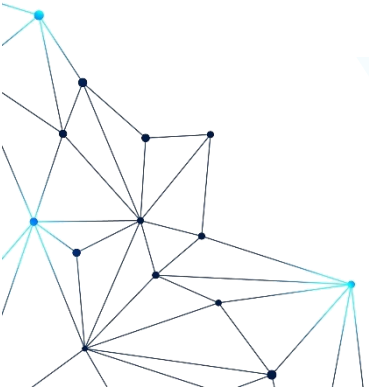


Image adapted from: Splunk Webinar: Getting Started with Machine Learning

RectifAI

Machine Learning timeline

1950 — First use of the term “computer science” and start of Artificial Intelligence research

1966 — **Winter of AI**

Failure of machine translation, cutbacks in academic research, decrease in research, difficulty to train a neural network…
Basic research was still funded in Canada

1990

2006 — First Deep Neural Network that has been trained - Geoff Hinton University of Toronto

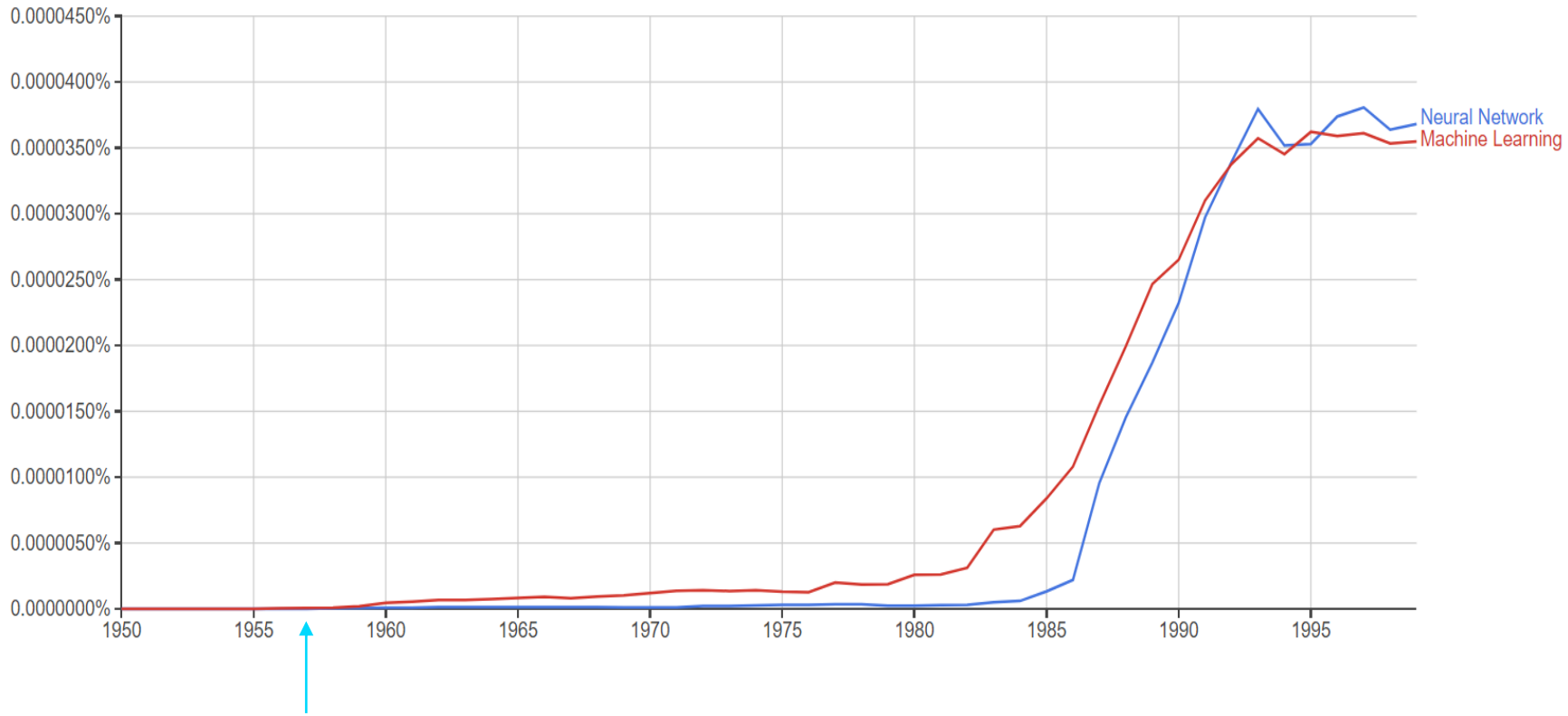
2011 — First use of rectifiers - Yoshua Bengio Universite de Montreal

2015 — Breakthrough in human voice recognition
Breakthrough in computer vision
Computer reaches human performances in image recognition

RectifAI

Machine Learning timeline

Google Books Ngram view of the usage of “Machine Learning” and “Neural Network”

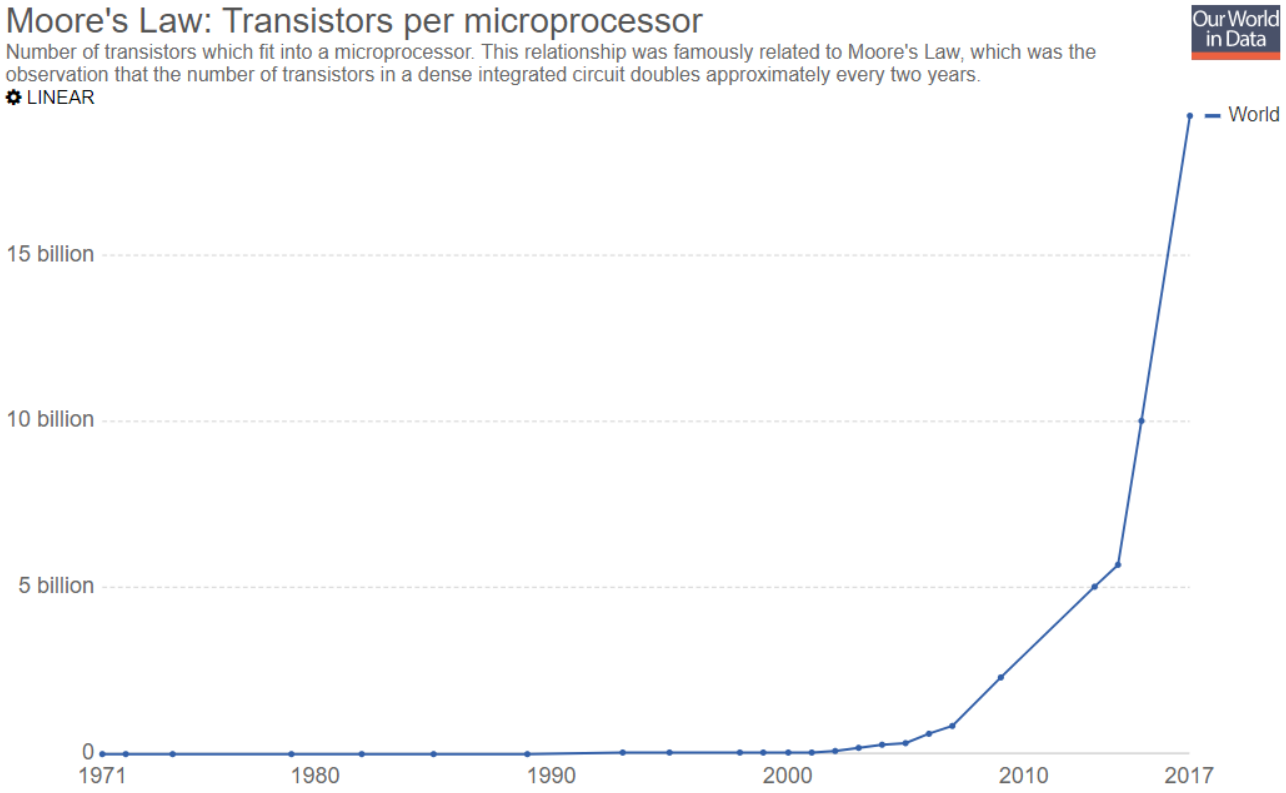


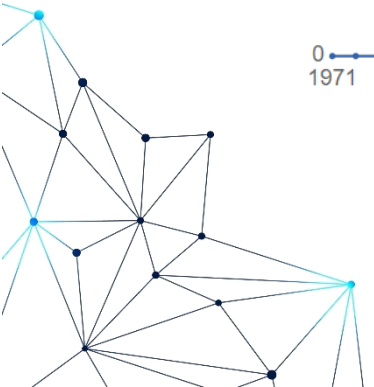
First use of the term “computer science”
and start of Artificial Intelligence research

So…
Why now?

Why now?

> Way more processing power





\+ Rise of the GPU computing

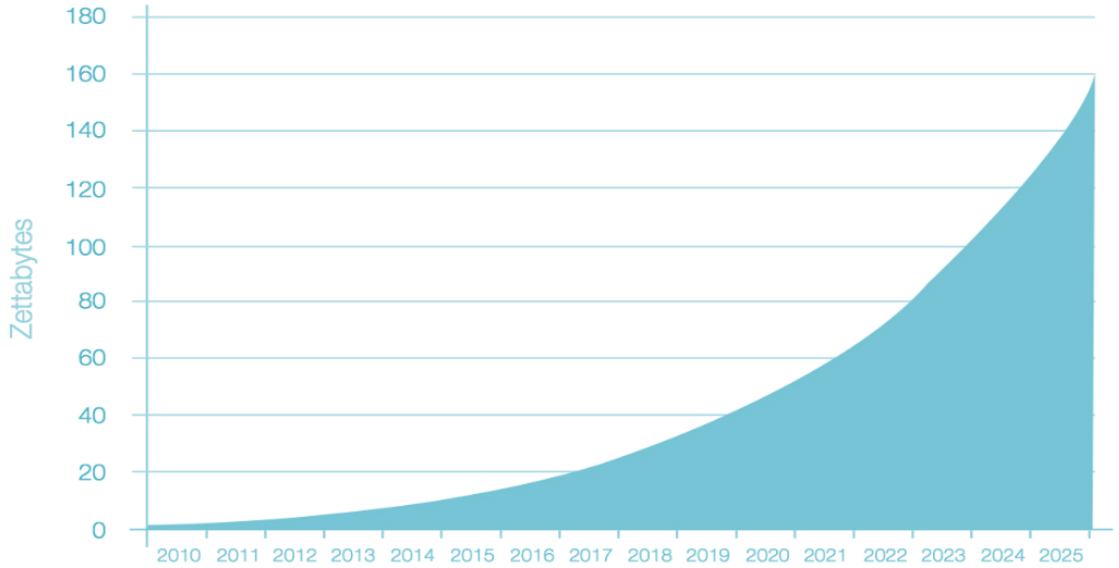
NVIDIA

\+ CUDA toolkit

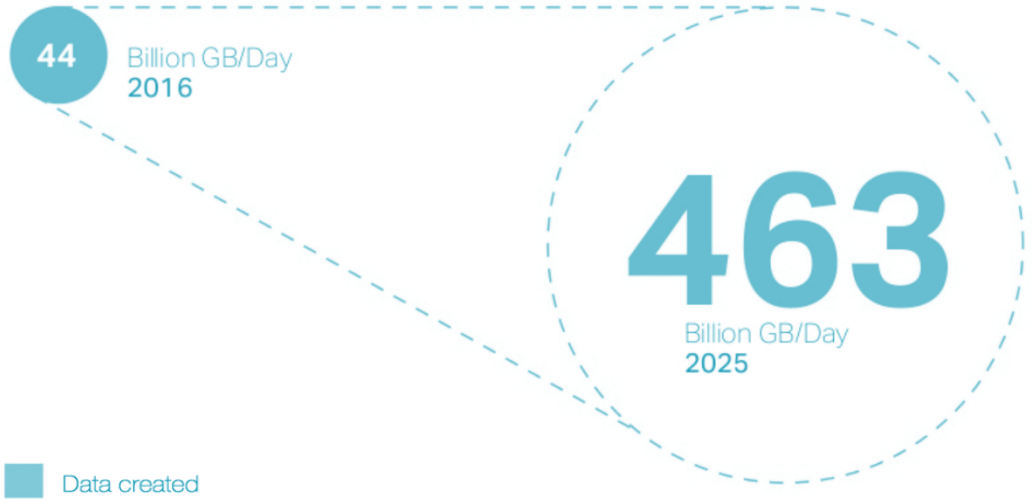
Karl Rupp, 40 Years of Microprocessor Trend Data

Why now?

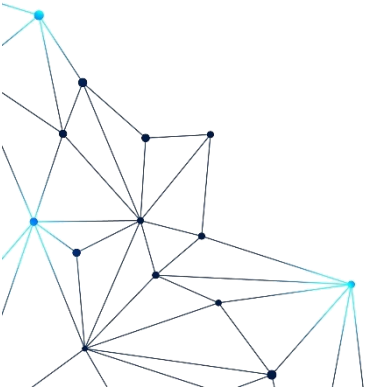
> Way more data



David Reinsel, John Gantz and John Rydning, analysts at IDC Corp. on March 2017



https://www.slideshare.net/Micro-Focus/growth-of-internet-data-2017



90% of the data on the internet has been created since 2016

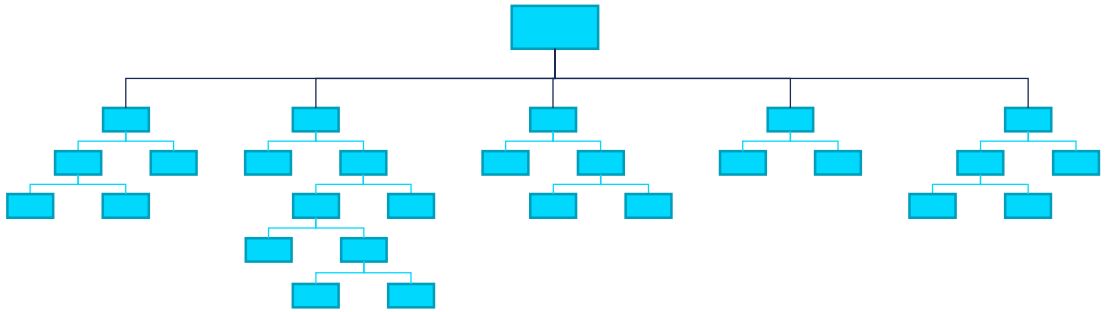
according to an IBM Marketing Cloud study

RectifAI

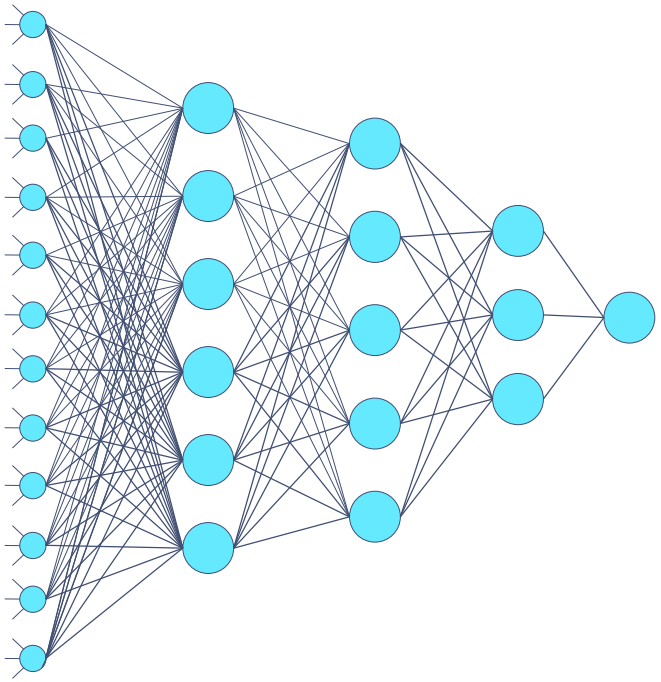
Why now?

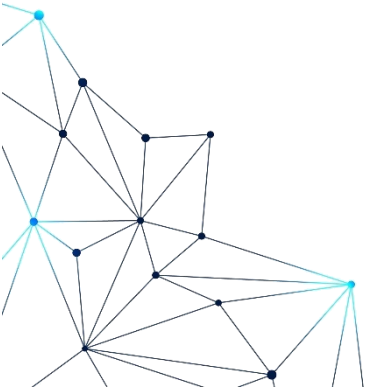
> New architectures

Tree Based Model – ex. Gradient Boosting



Neural Networks

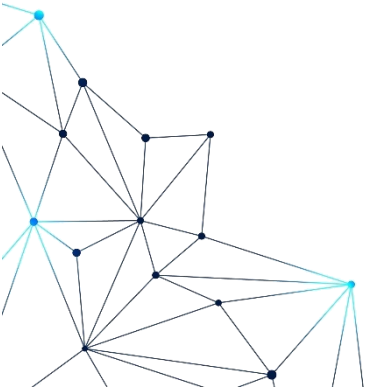




Why now?

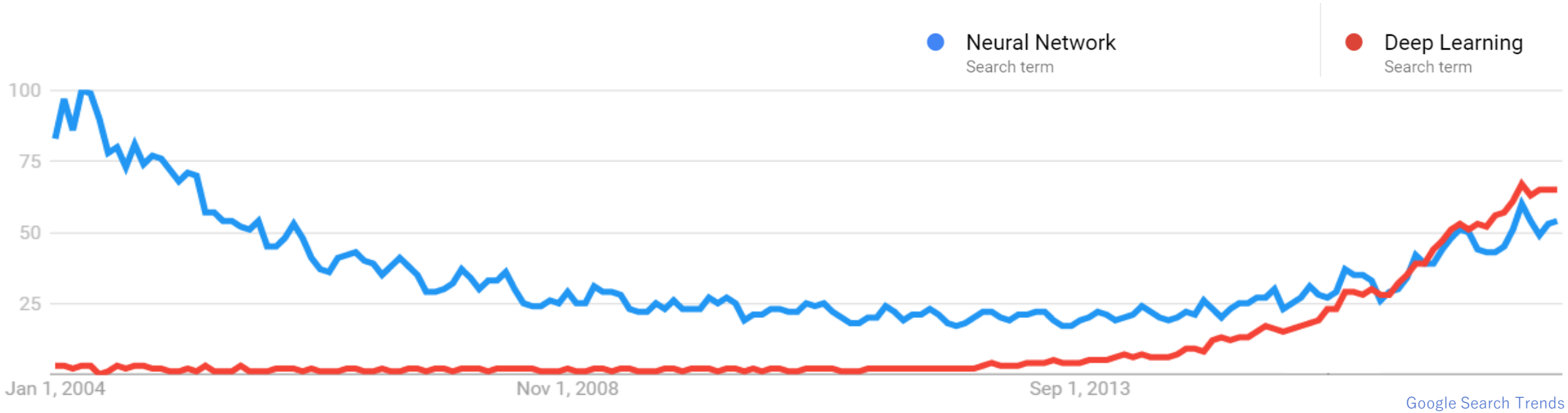
> Easier to program

- Important **Open Source** community (Github)
- Opened **research** paper (arXiv)
- Research papers released with **source code**
- **High Level** programming (Python)
- Simpler and simpler **wrapping** (Scikit-learn, TensorFlow, Theano, PyTorch...)
- Graphical interfaces (Watson studio, Tableau, Alteryx, RapidMiner, Splunk...)



Why now?

> Branding



Google Search Trends

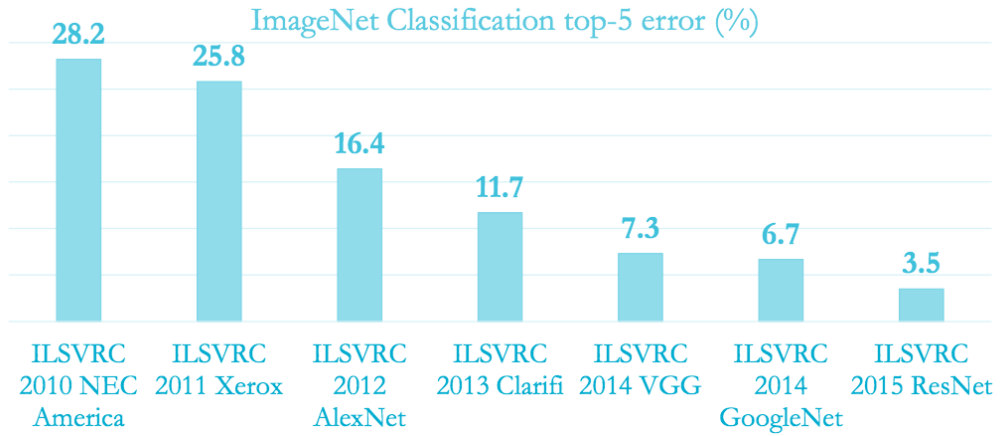
Where are we today?

> Machines can now **replicate or even exceed human capabilities** in certain domains

> AI can **speed up processing**, **increase precision**, **automate dangerous tasks**, **produce content**…

Today : “Pretty much anything that a normal person can do in < 1 sec, we can now automate with AI.”

Andrew Ng

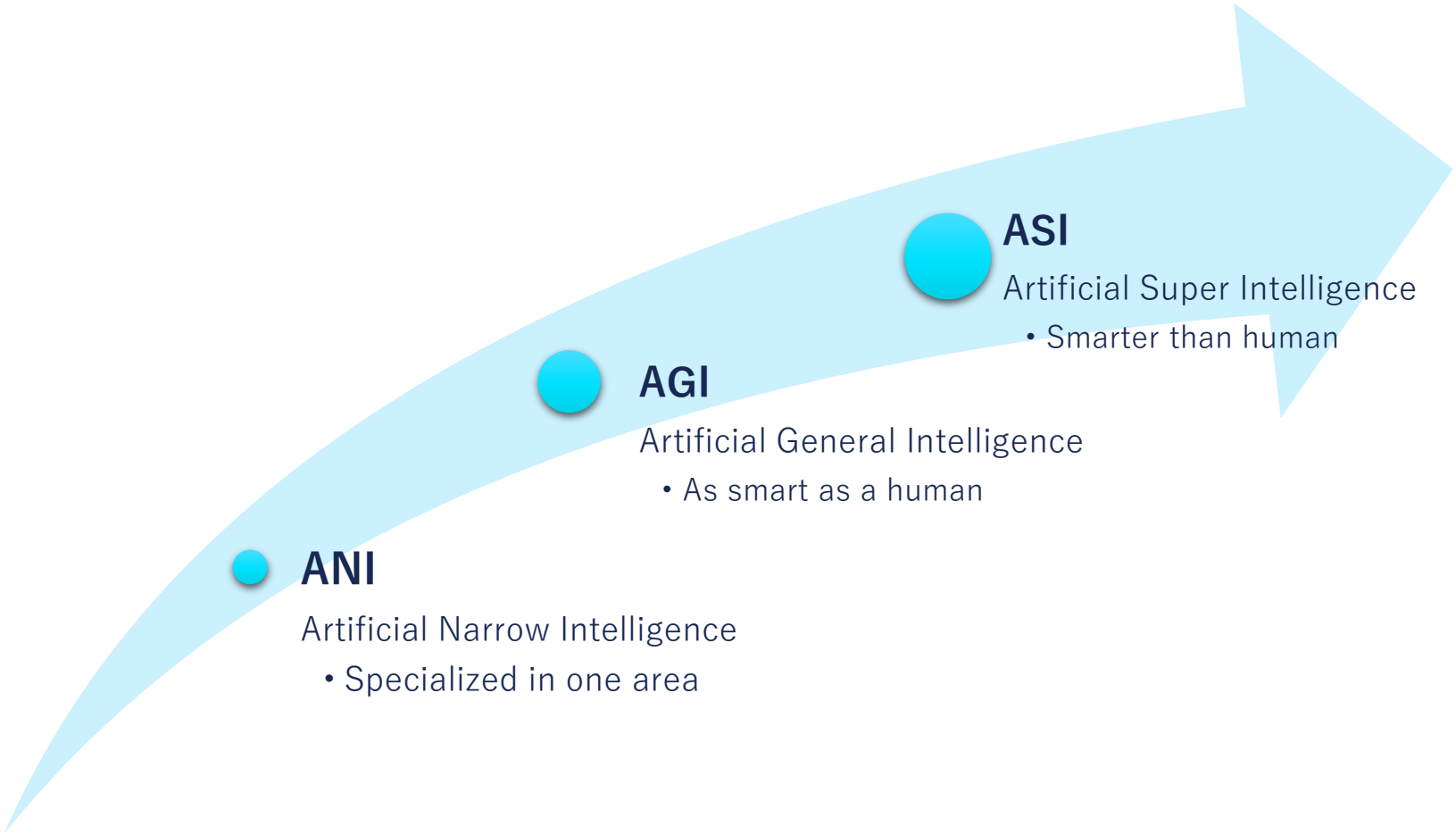


An application of AI to the right processes can create tremendous value for a company and can be a source of significant competitive advantages

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In Proceedings of the IEEE international conference on computer vision (pp. 1026-1034).

RectifAI

Future of AI



ANI

Artificial Narrow Intelligence

- Specialized in one area

AGI

Artificial General Intelligence

- As smart as a human

ASI

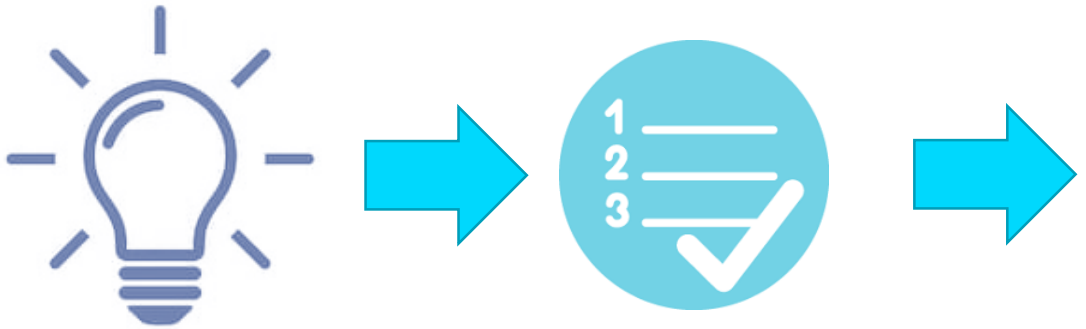
Artificial Super Intelligence

- Smarter than human

Setup an AI project

What's needed

Well Defined objective



Data

< 100,000 examples for classic ML
> 100,000 for Neural Network

Processing power

CPU for simple trainings
GPU for complex trainings

ML Model

Classic ML
Deep Learning

AI is well suited for business processes that…

… can **tolerate** a certain **margin of error**

… require more **predictions and estimates** than direct cause-effects

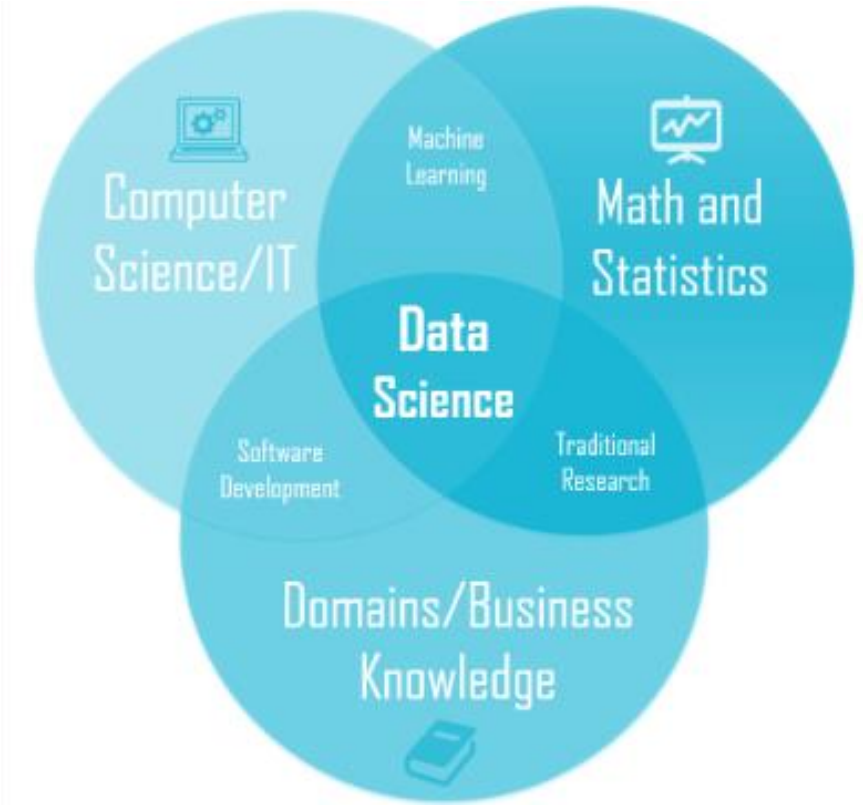
A Data Science team

Domain expert

Data analyst /
Data Engineer

Data Scientist /
Machine Learning Engineer

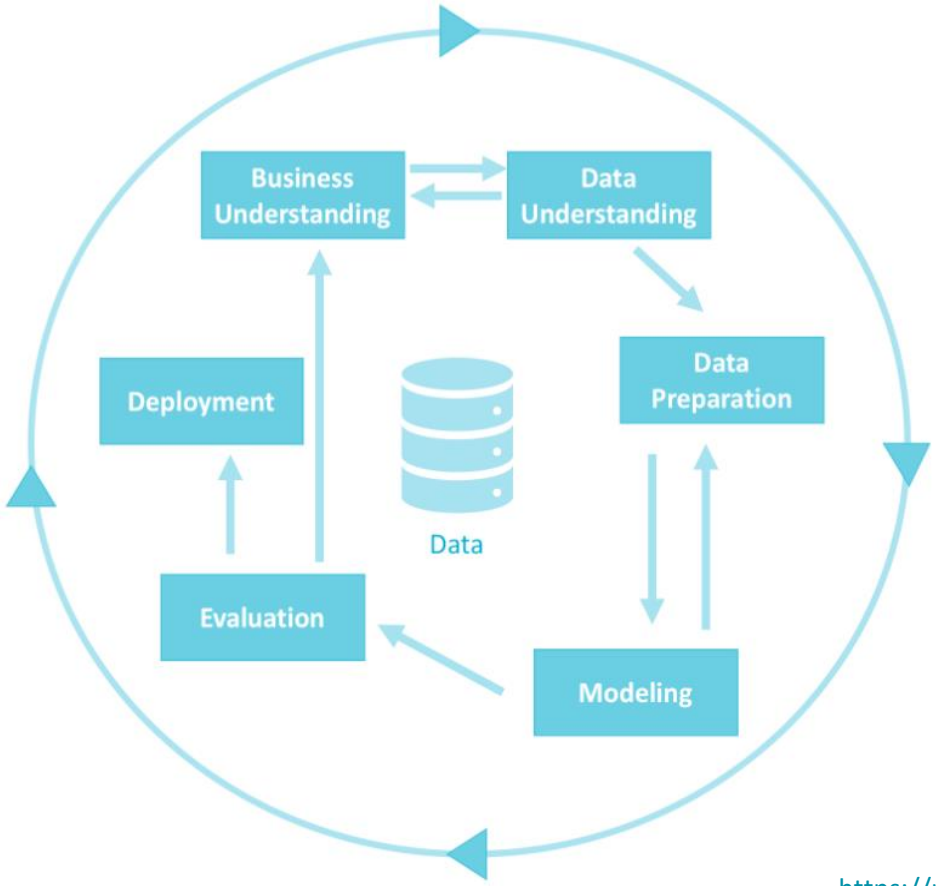
Research scientist /
Research Engineer



https://medium.com/craftdata-labs/on-building-effective-data-science-teams-4813a4b82939

Typical steps of an AI project

Collecting Data → Cleaning Data → Modeling → Evaluating → Deploying



https://medium.com/craftdata-labs/on-building-effective-data-science-teams-4813a4b82939

RectifAI

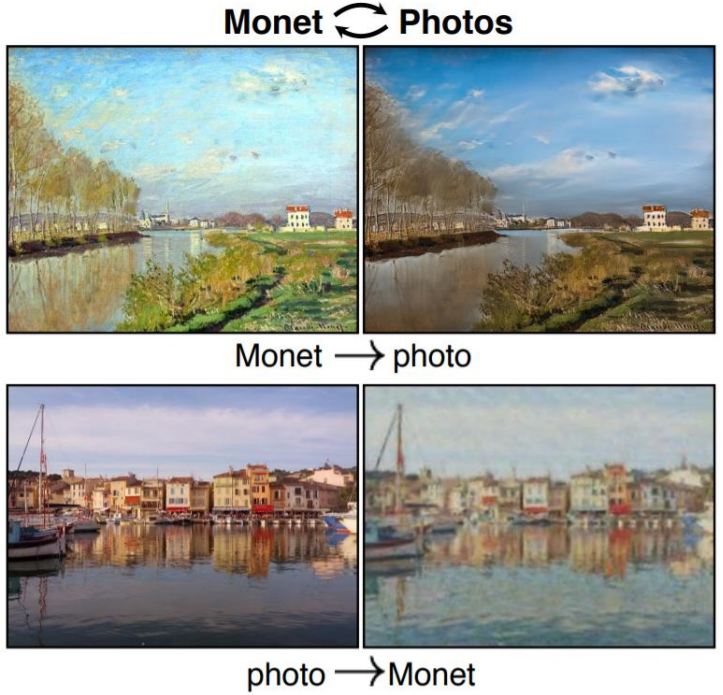
Some interesting use case

Generating art

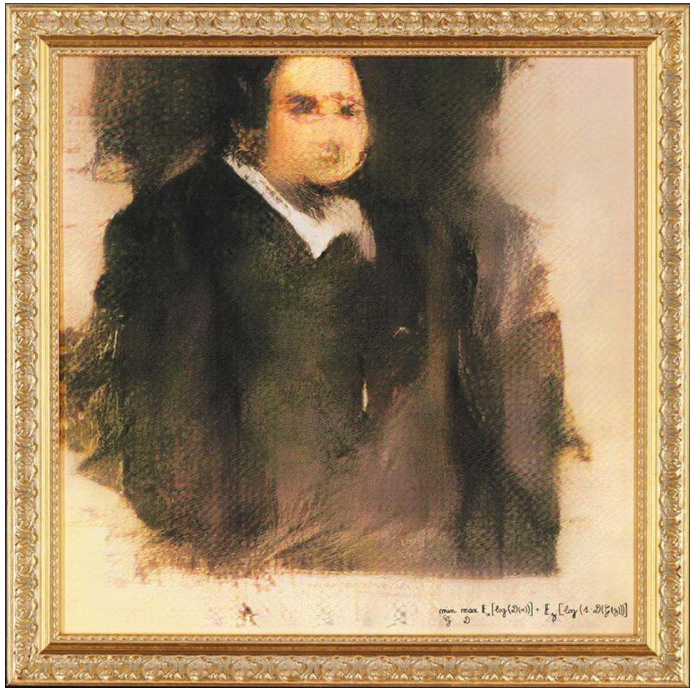
CNN/GANs

https://deepdreamgenerator.com/#gallery - 2015

[NVIDIA 2017]



https://arxiv.org/pdf/1703.10593.pdf



AI artwork sells for $432,500 – [Forbes 2018]

Mozart :

https://soundcloud.com/psylent-v/samplernn_torch-mozart-mu-law-peak-normalization

Electronic :

https://www.youtube.com/watch?v=3hvhgDiklIk

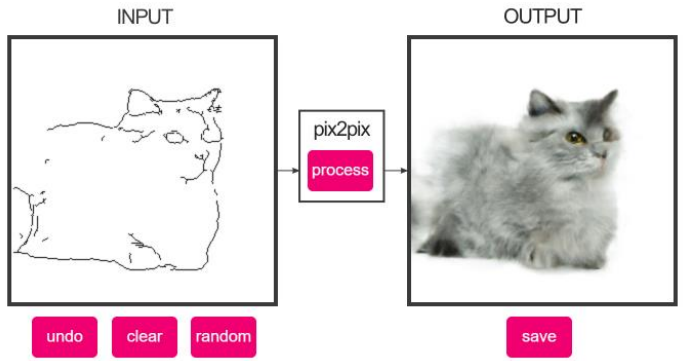
Barack Obama :

https://www.youtube.com/watch?v=YfU_sWHT8mo

RectifAI

Learning to finish pictures

Drawing to picture



https://affinelayer.com/pixsrv/

Coloring picture

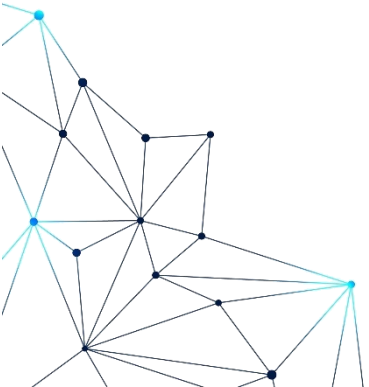
http://stanford.edu/~lucy3/CS231N.pdf

Denoising picture

https://arxiv.org/pdf/1704.05838.pdf

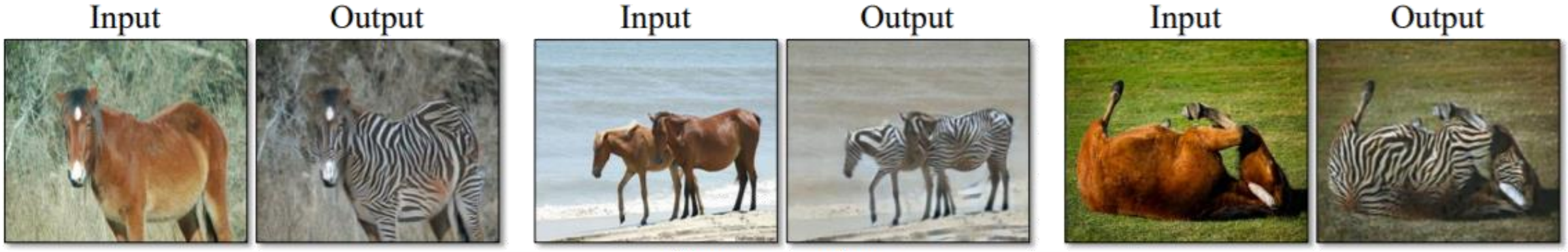
Generating maps

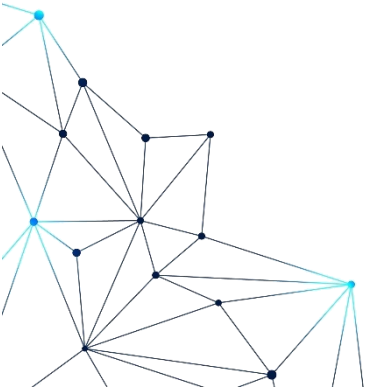
http://stanford.edu/~lucy3/CS231N.pdf



RectifAI

Transform images

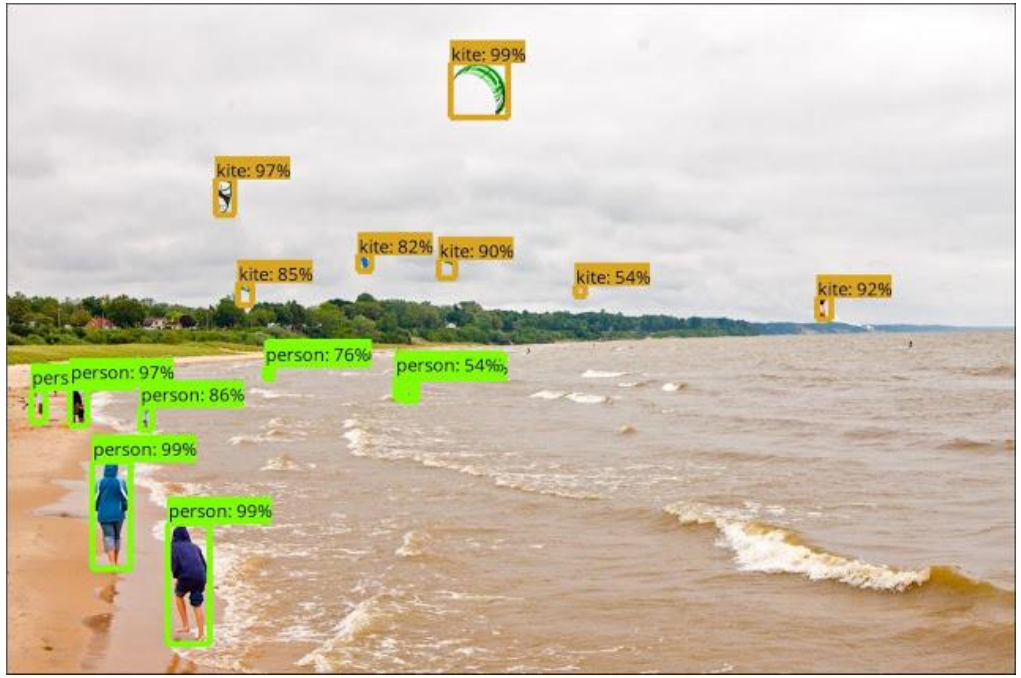




https://arxiv.org/pdf/1703.10593.pdf

Labelling

Image to labels



https://research.googleblog.com/2017/06/supercharge-your-computer-vision-models.html

Labels to image

This bird is red and brown in color, with a stubby beak

The bird is short and stubby with yellow on its body

A bird with a medium orange bill white body gray wings and webbed feet

https://arxiv.org/pdf/1612.03242.pdf

Natural Language Processing (NLP)

RNN/GANs

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla at dui nec lacus ornare tristique. Nunc ut ornare urna. Etiam tristique, leo at porttitor tincidunt, sapien ex porta lorem, vitae rutrum turpis enim a purus. Etiam dignissim facilisis faucibus. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Suspendisse vehicula iaculis felis. Donec

mollis posuere sapien, a finibus nisi auctor nec. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam sit amet neque a leo dignissim ultrices non sit amet nulla. Fusce pulvinar feugiat erat ac eleifend. Aliquam fringilla, eros bibendum suscipit

aliquam, lacus dolor euismod mi, ac cursus urna nulla ac dolor. Donec vel dolor eget enim luctus molestie. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam vitae nulla a nisl fermentum euismod. Duis finibus ac metus a rhoncus.
Maecenas a velit elit. Ut molestie, nulla et auctor facilisis, nibh tellus ullamcorper quam, in ullamcorper lacus nulla et libero. In augue tellus, auctor sed justo convallis, sodales feugiat

- **Translation**
  ex. Google Translate
- **Summary**
- **Sentiment Analysis**
  ex. Twitter API
- **Chatbots**
  ex. Siri
- **Question answering**
  ex. Maluuba
- **Categorization**

? / !

Rectif AI

Robotics/Self-driving cars

https://www.youtube.com/watch?v=gn4nRCC9TwQ :

Photo from : uber.com

What else?

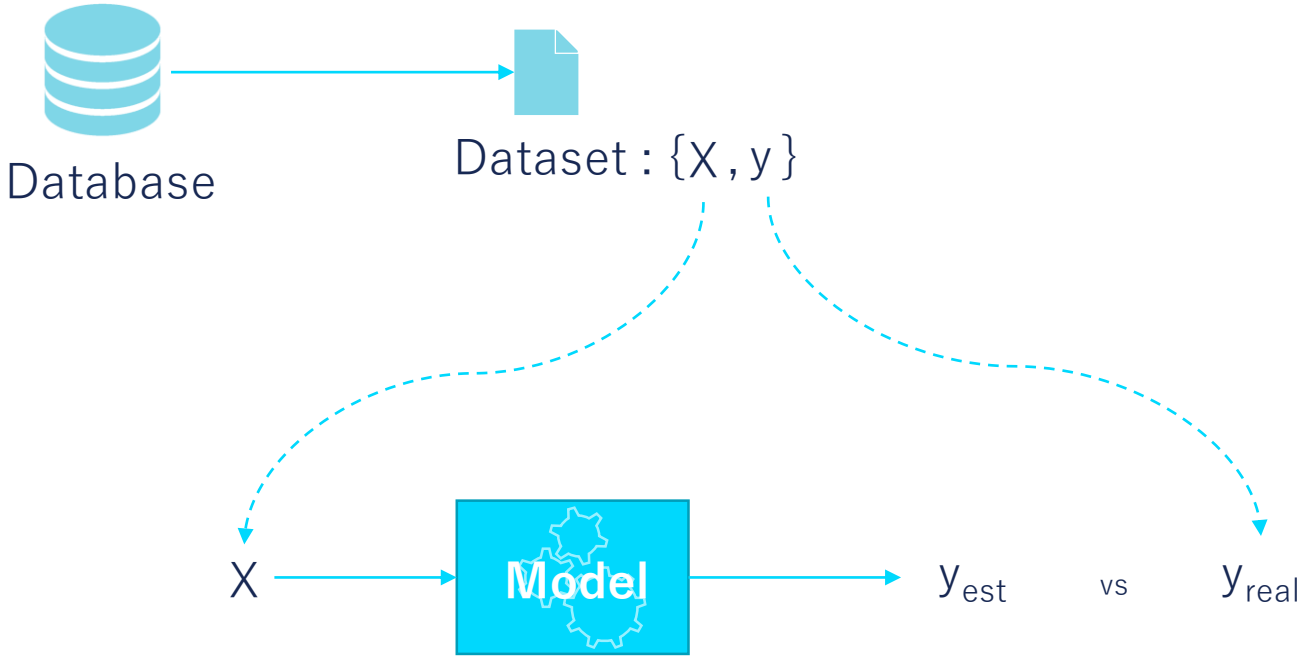
- Recognize objects in images
- Transcribe speech
- Translate between languages
- Speak
- Pick out the bit of a paragraph that answers your question
- Recognize emotions in images of faces
- Recognise emotions in speech
- Drive / Fly a drone
- Predict parking difficulty by area
- Discover new uses for existing drugs
- Spot cancer in tissue slides
- Predict hypoglycemic events in diabetics
- Identify diabetic retinopathy from retinal photos
- Analyze the genetic code of DNA
- Detect a range of conditions from images
- Detect crop disease
- Spray pesticide with pinpoint accuracy
- Predict crop yields/ Sort vegetables
- Spot burglars in your home
- Write its own encryption language
- Unscramble pixelated images
- Detect malware
- Verify your identity
- Anticipate fraudulent payment attacks before they happen
- Trade stocks
- Handle insurance claims
- Predict the outcomes of cases
- Do legal case research
- Do due diligence on M&A deals
- Flag errors in legal documents
- Beat humans at Jeopardy
- Play Go better than humans
- Beat the best human players at Texas Hold 'Em poker
- Be your personal trainer
- Write software unit tests
- Identify potentially threatening weather
- Write poems / Write music / Paint
- Design logos
- Come up with its own recipes
- Write sports articles
- Write film scripts
- Recommend songs you'll like
- Lip-read better than humans
- Optimize energy usage

Business Insider 2017

RectifAI

Classic vs. Modern Approaches

General Idea



Some terminology

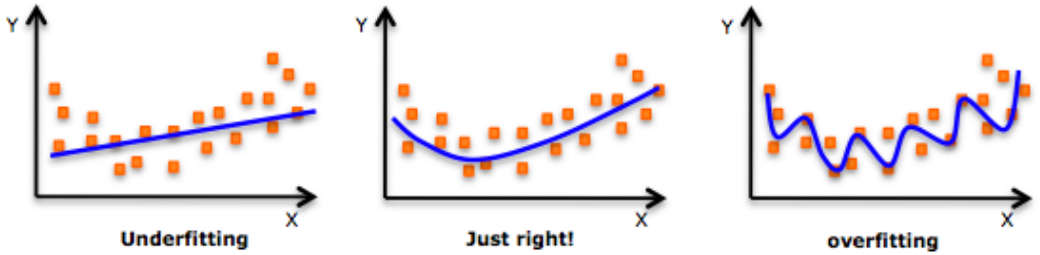
Classification vs. Regression

Classification

vs

MILA Deep Learning Winter School 2018

Regression



https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/

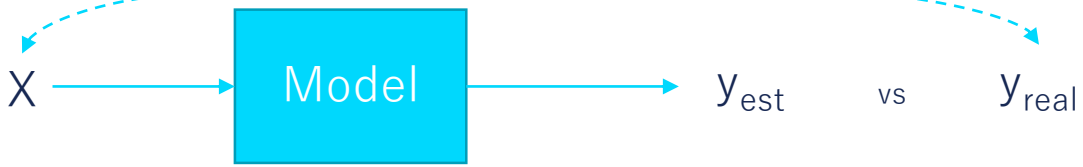
Some terminology

Supervised vs. Unsupervised

Supervised

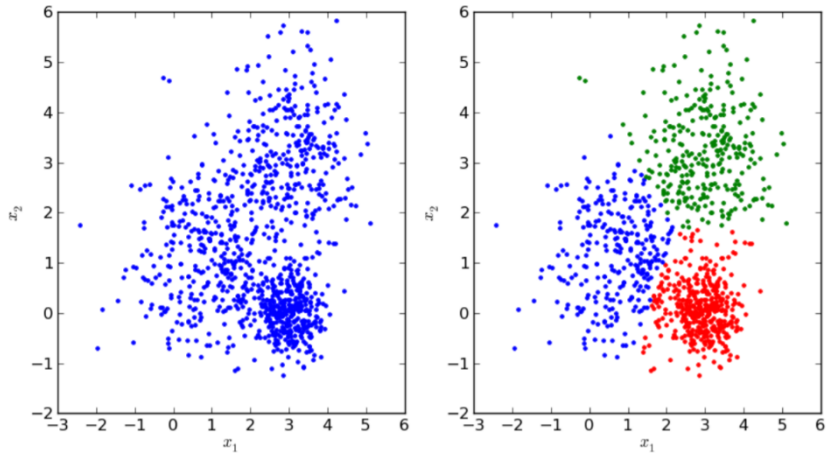
✓ Labels

Dataset : {X,y}



Unsupervised - Clustering

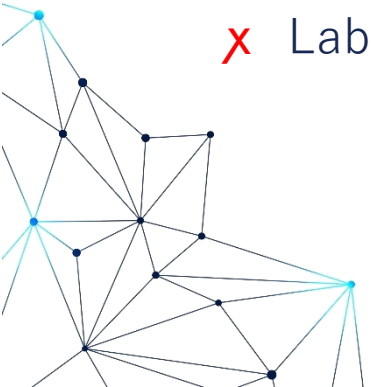
✗ Labels



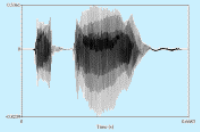
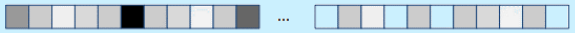
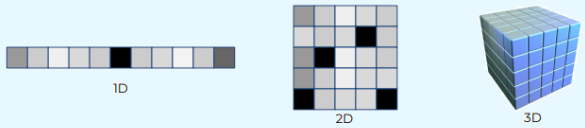
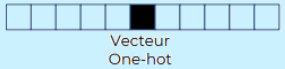
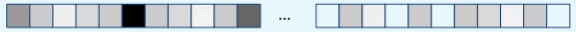
MILA Deep Learning Winter School 2018

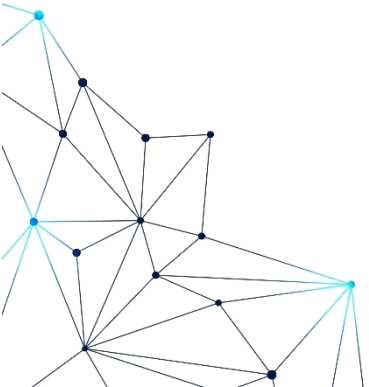
Unsupervised - GANs

[NVIDIA 2017]



Convert everything to numbers

Data type	Natural representation	Input representation	Description	Example
Sound/timed sequence			Float sequence	[12.4 , 14.2 , 15.5 ⋯]
Image			Integer 1D, 2D or 3D	[[255, 134, 12], [254, 134, 13], ⋯]
Category	Macro Micro Pico …		Binary value for each category	[[0 0 0 0 1 0] , [0 1 0 0 0 0], ⋯]
Dates	2018/01/01 9:00		Year, Month, Day Of Week, Day Of month, AM/PM⋯ etc.	[12 , 14, 1, 0 ⋯]

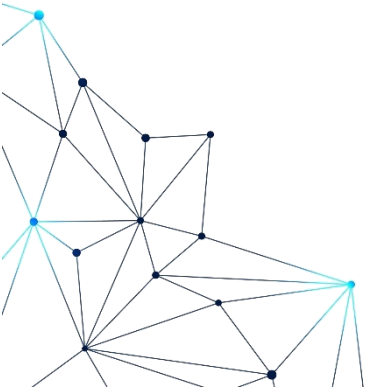


Classic vs. Modern Approaches

Interpretability vs. Accuracy

	Interpretable	Accurate
Modern Algorithms	✗	✓
Classic Algorithms	✓	✗

Interpretable or accurate: **choose one**

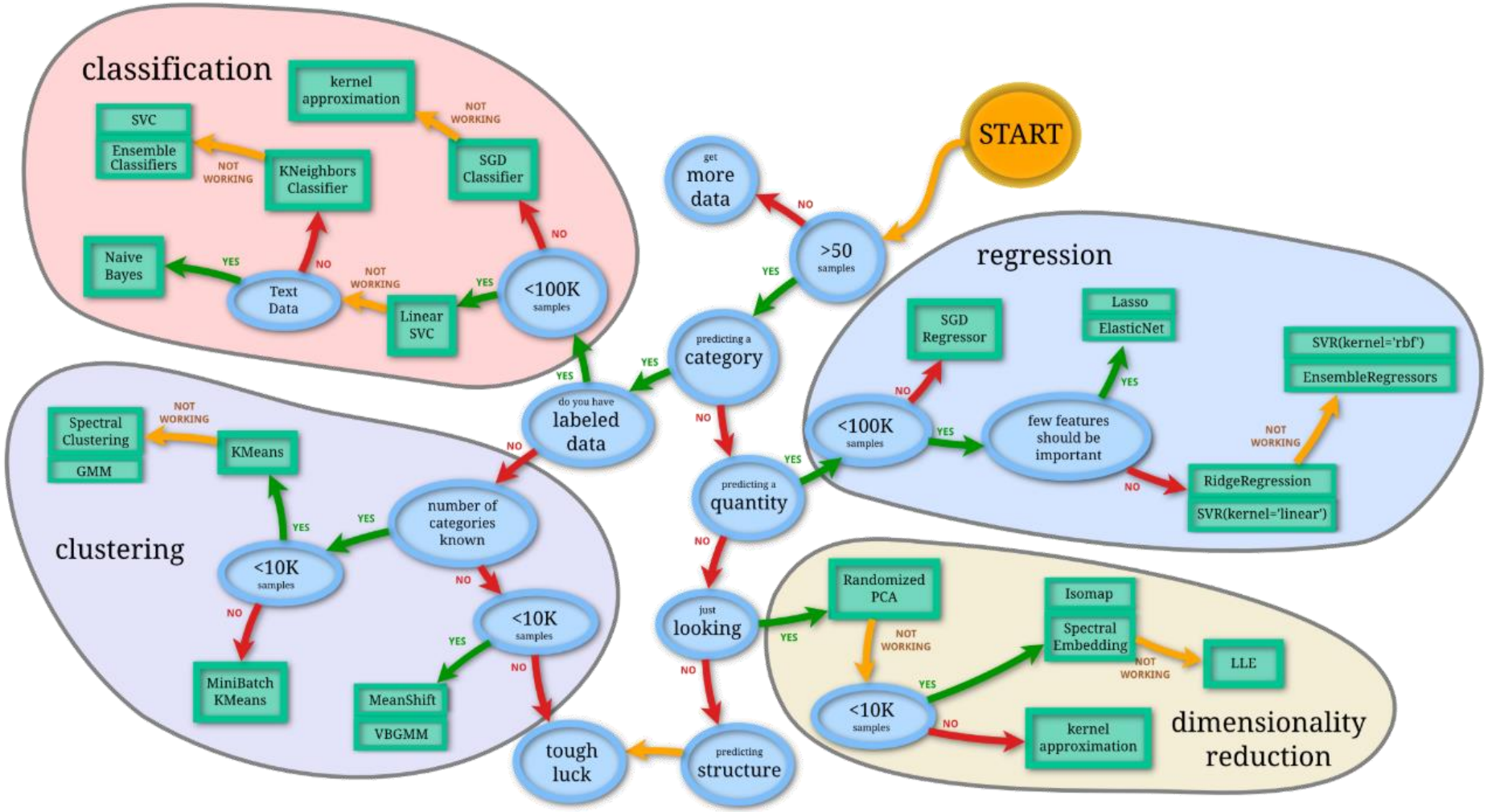


> Not 100% true since NIPS 2017 conference → SHAP Values

https://github.com/slundberg/shap/raw/master/docs/presentations/NIPS%202017%20Talk.pptx

Classic Machine Learning models

Well... most of them... and if you don't consider time series as ML...

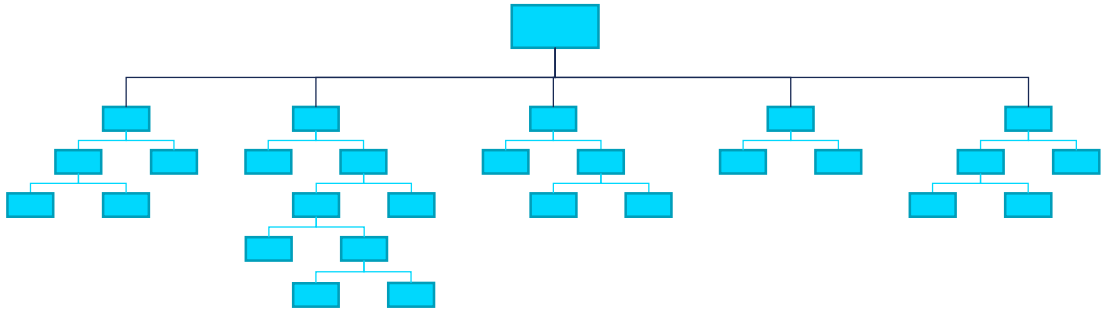


https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html

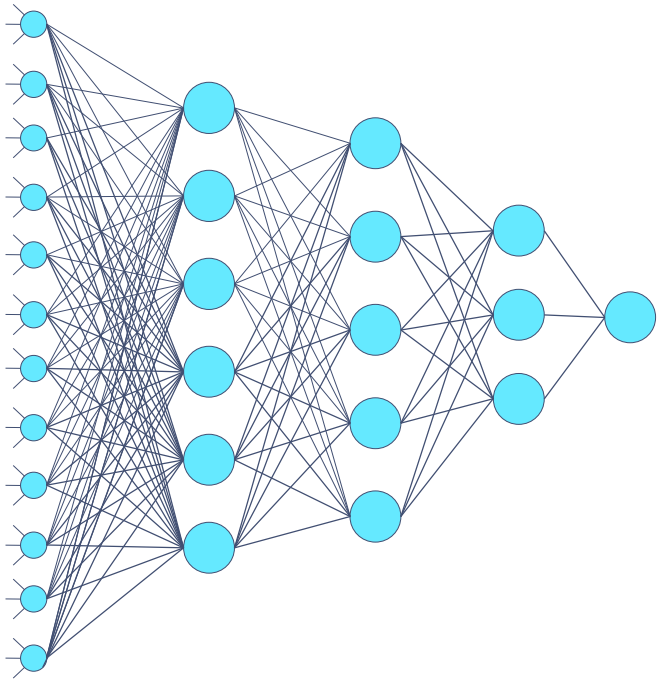
RectifAI

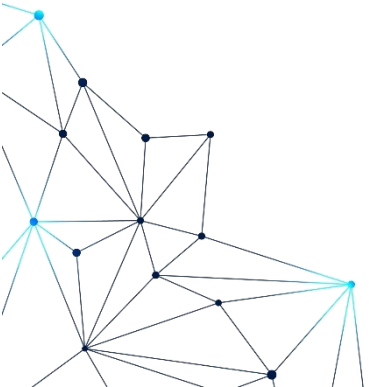
“Modern” models

Tree based model – ex. Gradient Boosting



Neural Networks



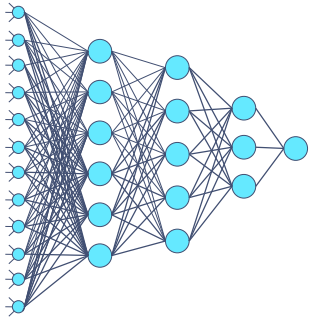


Neural Networks

RNN, CNN, GAN, GNN

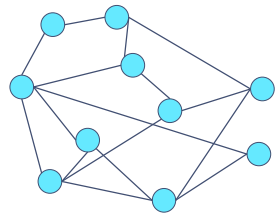
Neural Network (NN)

"Pure data"



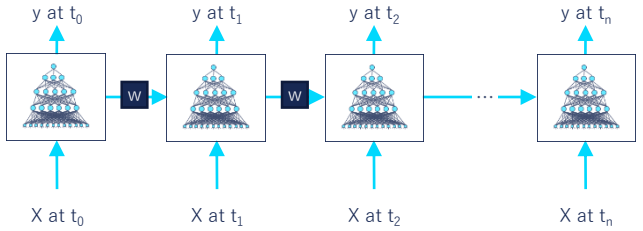
Graphical Neural Networks (GNN)

Connections



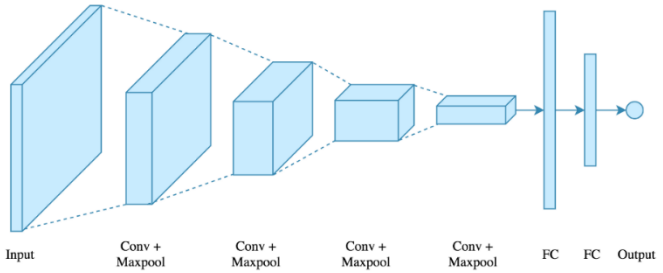
Recurrent Neural Network (RNN)

Time/ordered data



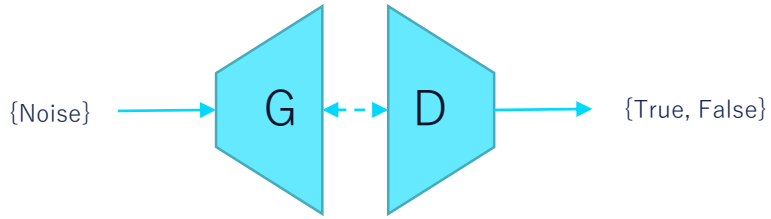
Convolution Neural Network (CNN)

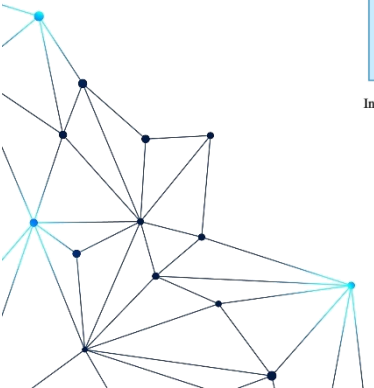
Image/Audio/Linked



Generative Adversarial Network (GAN)

Generation



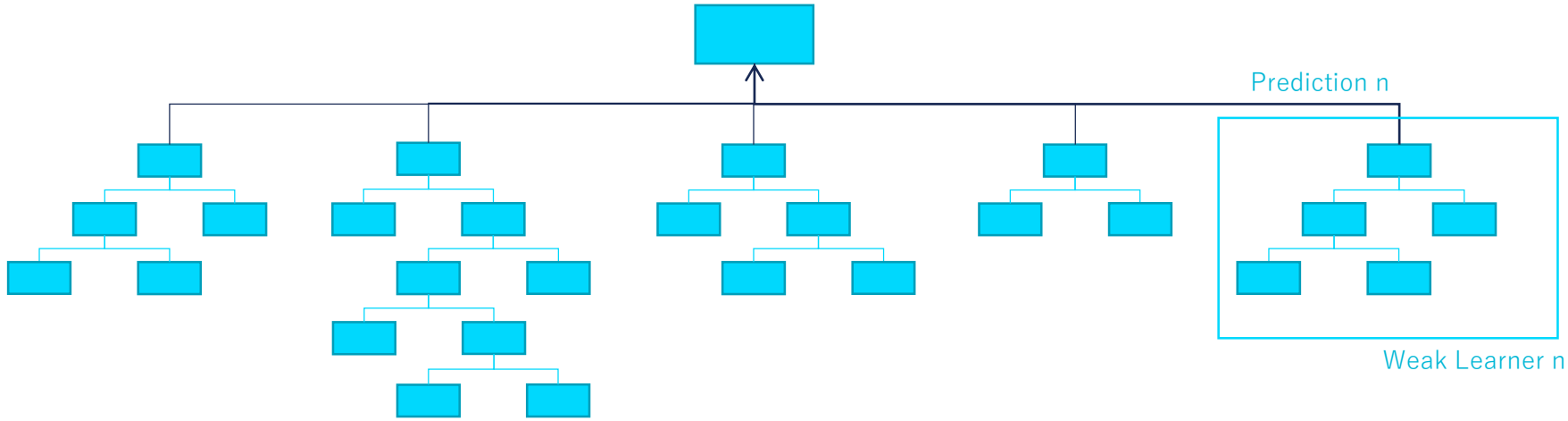


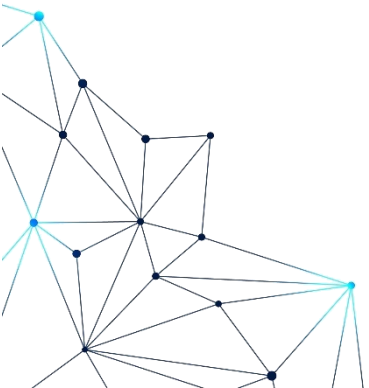
Rectif AI

How it works
Intuition behind modern models

Gradient Boosting

Ensemble of weak prediction models





XGBoost

Open source

LightGBM

Microsoft

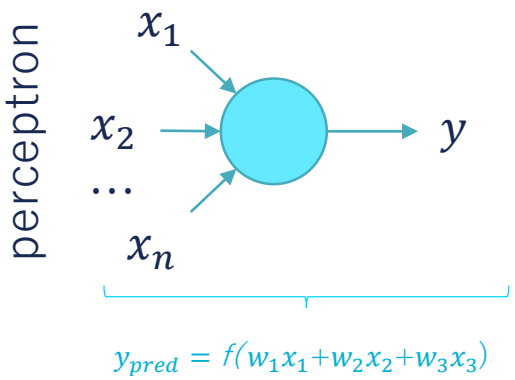
Catboost

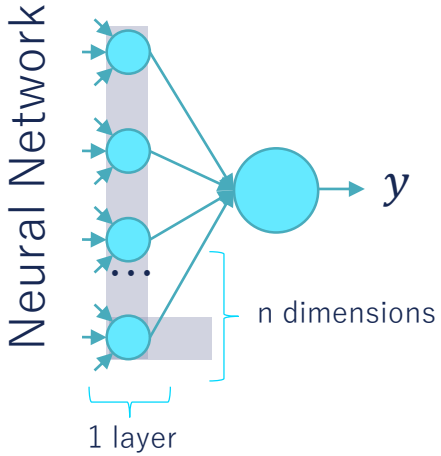
Yandex

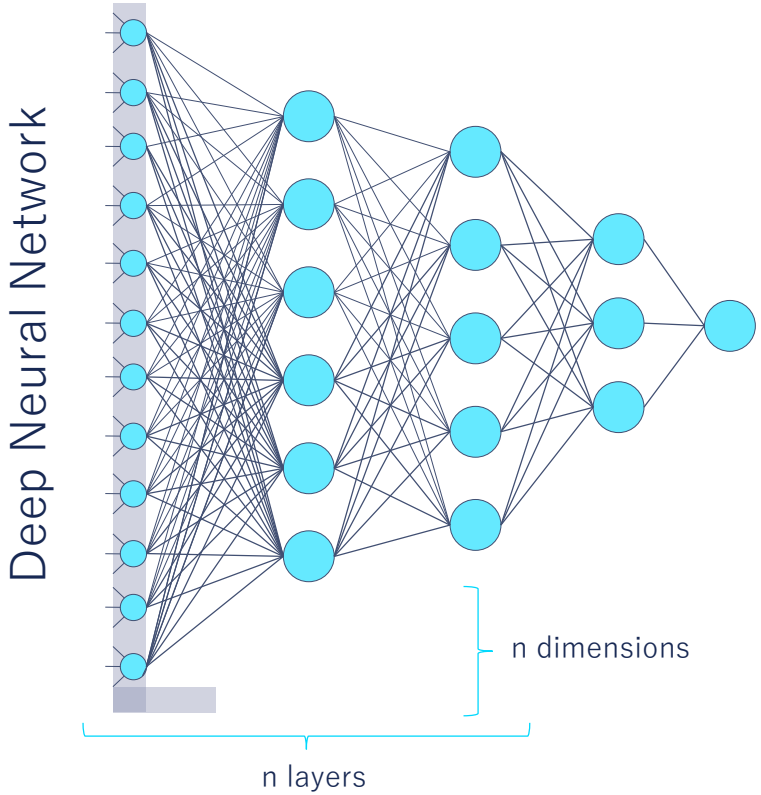
Neural Network

Multi-layer perceptron

Input :
$(x_1, x_2 ; y)$

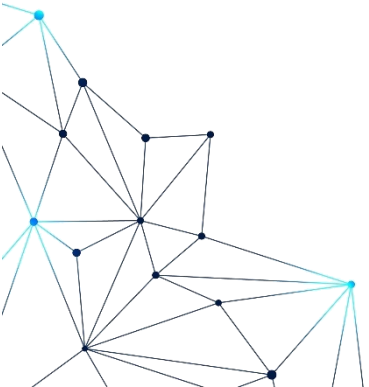


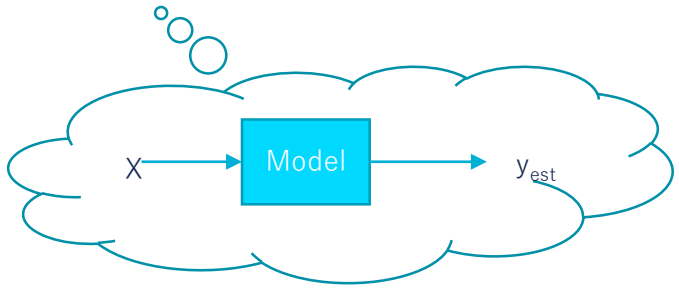




Error :
$y - y_{pred}$

Goal:
$\min(y - y_{pred})$

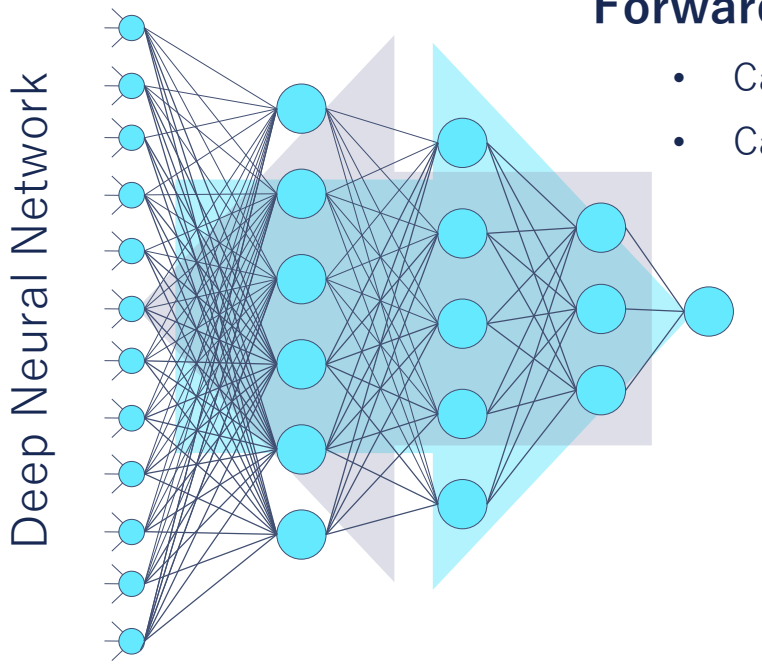




RectifAI

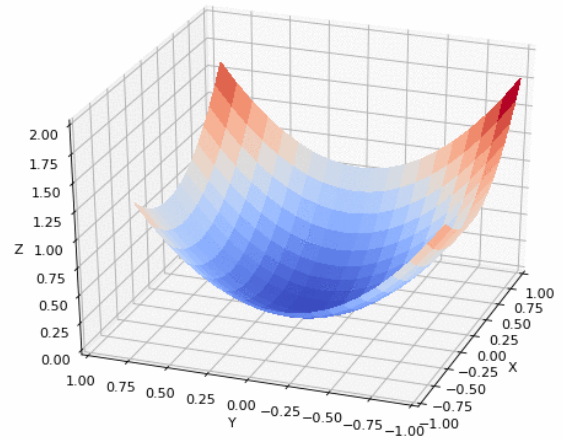
Training process

Prediction → Calculating loss → Optimizing weights to minimize loss



Forward Propagation

- Calculate an estimated y : y_{pred}
- Calculate the error : y - y_{pred}



https://blog.paperspace.com/intro-to-optimization-in-deep-learning-gradient-descent/

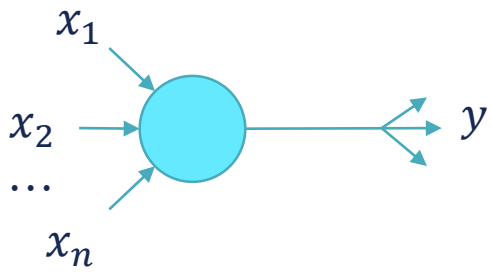
Backward Propagation

- Adjust the weights to minimise the error: y - y_{pred}

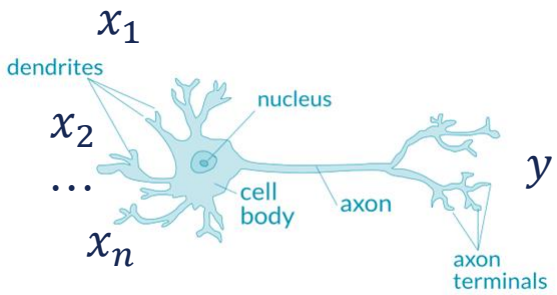
RectifAI

Human brain analogy

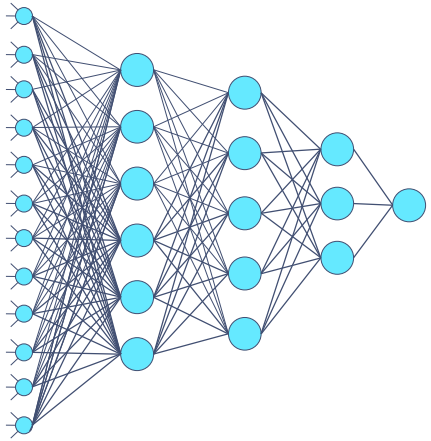
Perceptron



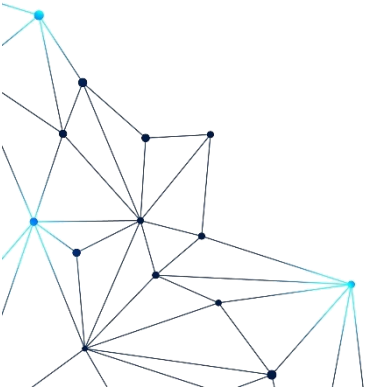
Neurone



Deep Neural Network



Brain

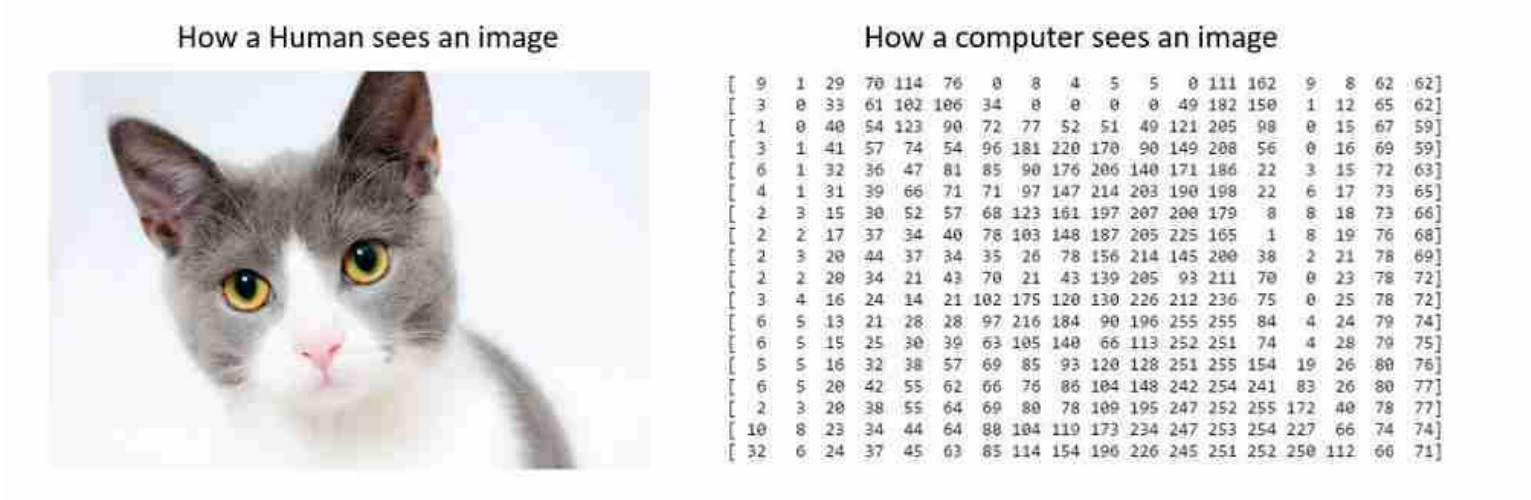


https://appliedgo.net/perceptron/

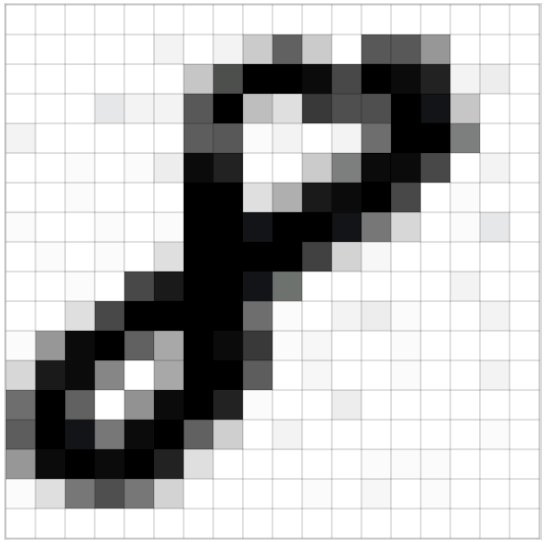
RectifAI

How these neurons could be mapped to other problems…

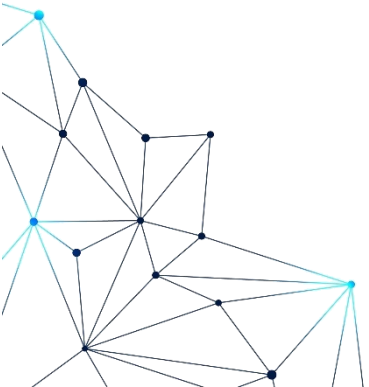
Computer vision – How computers see an image



https://medium.com/x8-the-ai-community/cnn-9c5e63703c3f



https://medium.com/@ageitgey/machine-learning-is-fun-part-3-deep-learning-and-convolutional-neural-networks-f40359318721



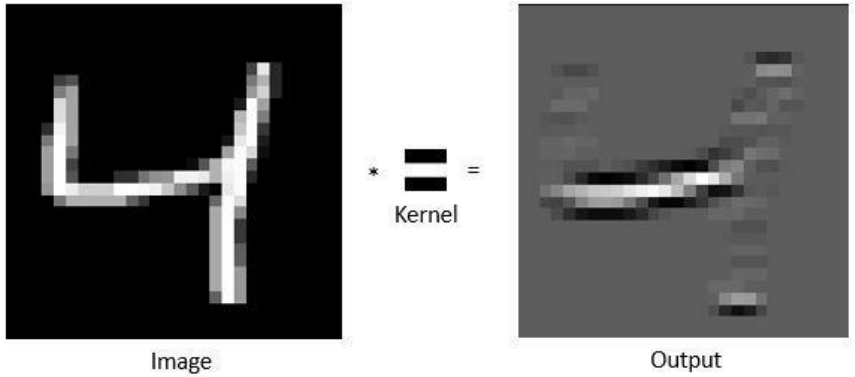
RectifAI

How these neurons could be mapped to other problems…

Computer vision – Applying a digital filter

1x1	1x0	1x1	0	0
0x0	1x1	1x0	1	0
0x1	0x0	1x1	1	1
0	0	1	1	0
0	1	1	0	0

4		



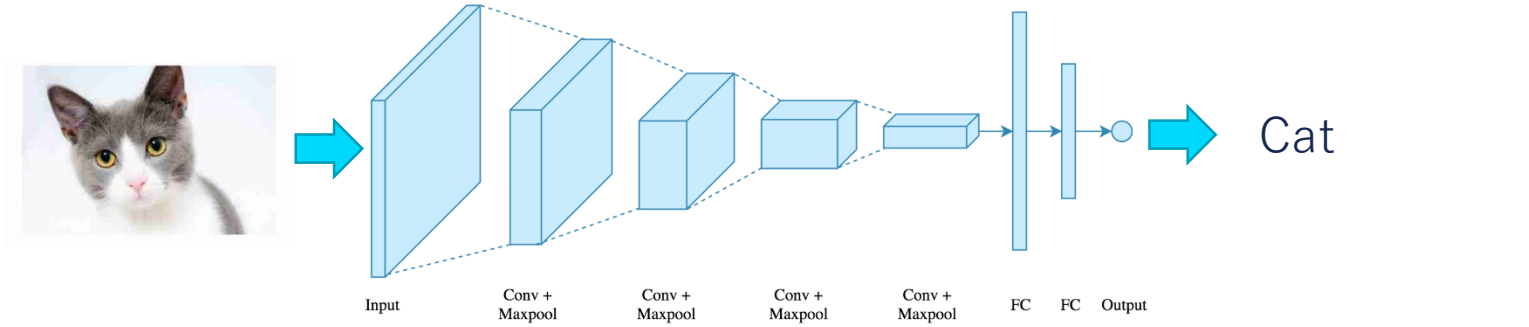
https://medium.com/x8-the-ai-community/cnn-9c5e63703c3f

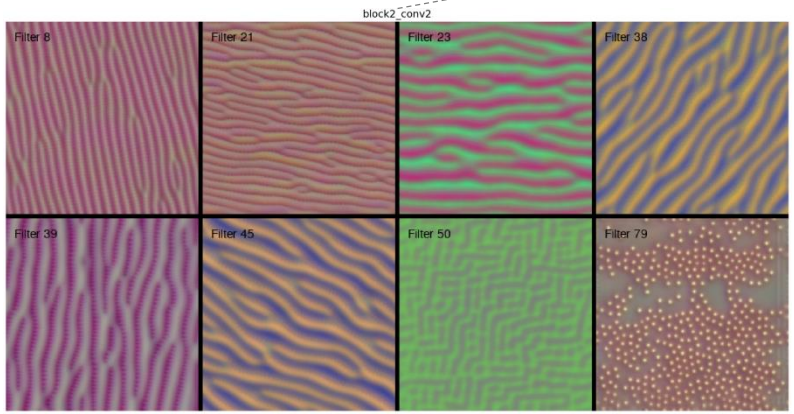
There isn't any Machine Learning on that page!

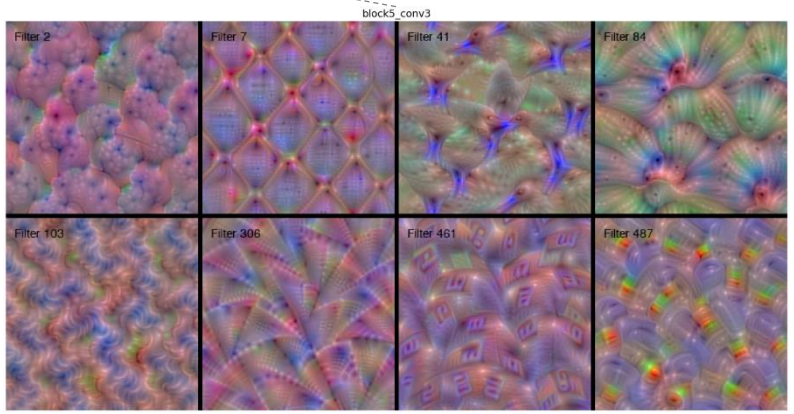
Demo : http://setosa.io/ev/image-kernels/

How these neurons could be mapped to other problems…

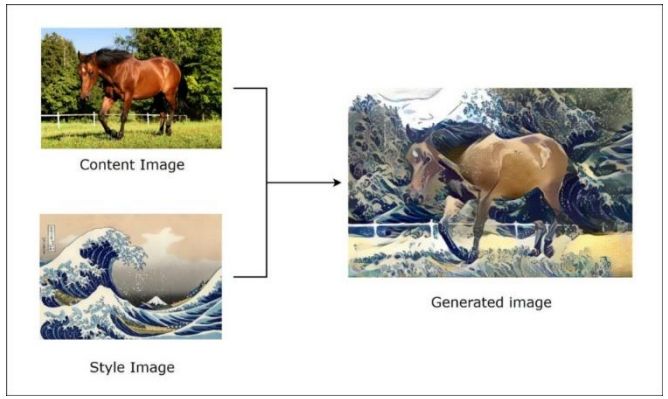
Computer vision – Letting the NN calculate the kernels



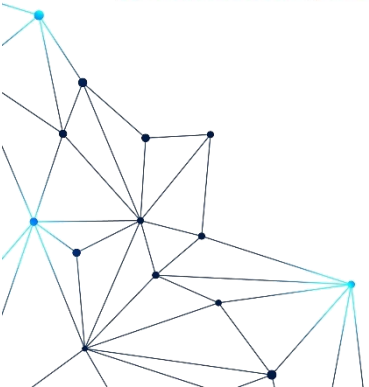




https://deepdreamgenerator.com/#gallery - 2015



https://towardsdatascience.com/neural-style-transfer-tutorial-part-1-f5cd3315fa7f



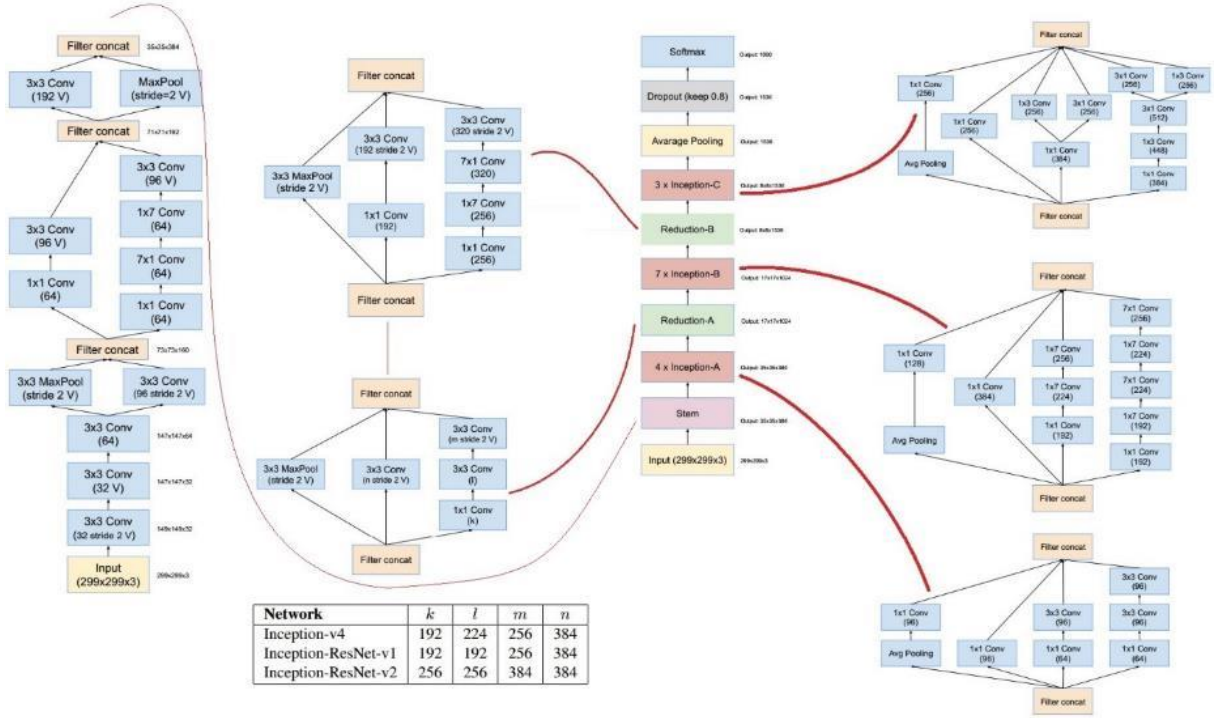
Images from: https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2

RectifAI

How these neurons could be mapped to other problems…

Computer vision – More complex models

Newer models looks much more complex…

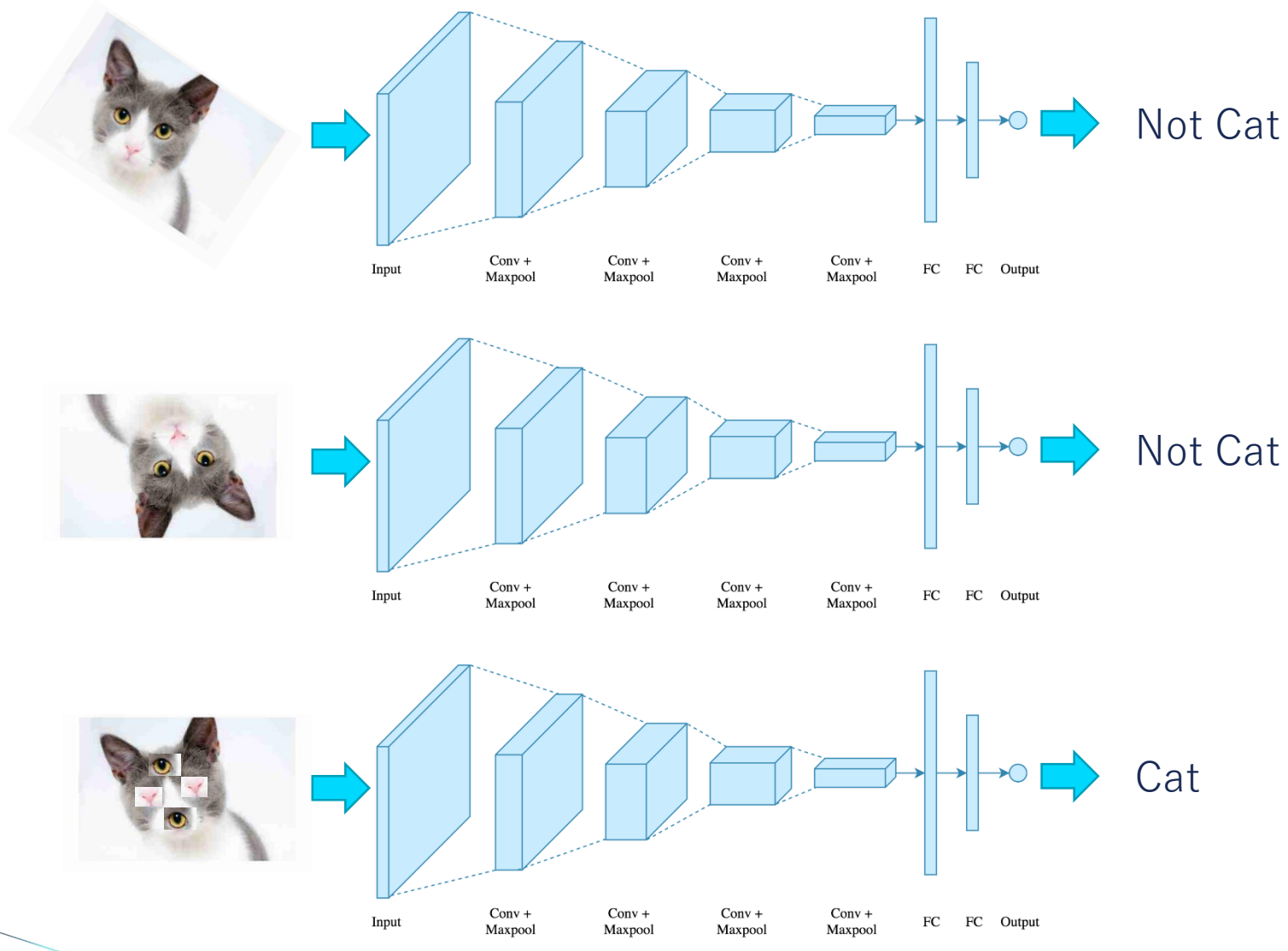


Network	k	l	m	n
Inception-v4	192	224	256	384
Inception-ResNet-v1	192	192	256	384
Inception-ResNet-v2	256	256	384	384

… but are mostly a combination of CNN, MaxPool and ResNETs and are available through transfer learning

How these neurons could be mapped to other problems...

Computer vision – Future evolution

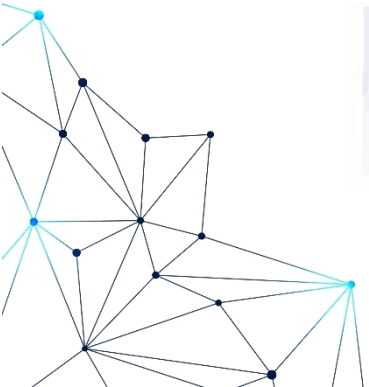


Before

- ✓ Data augmentation

Now

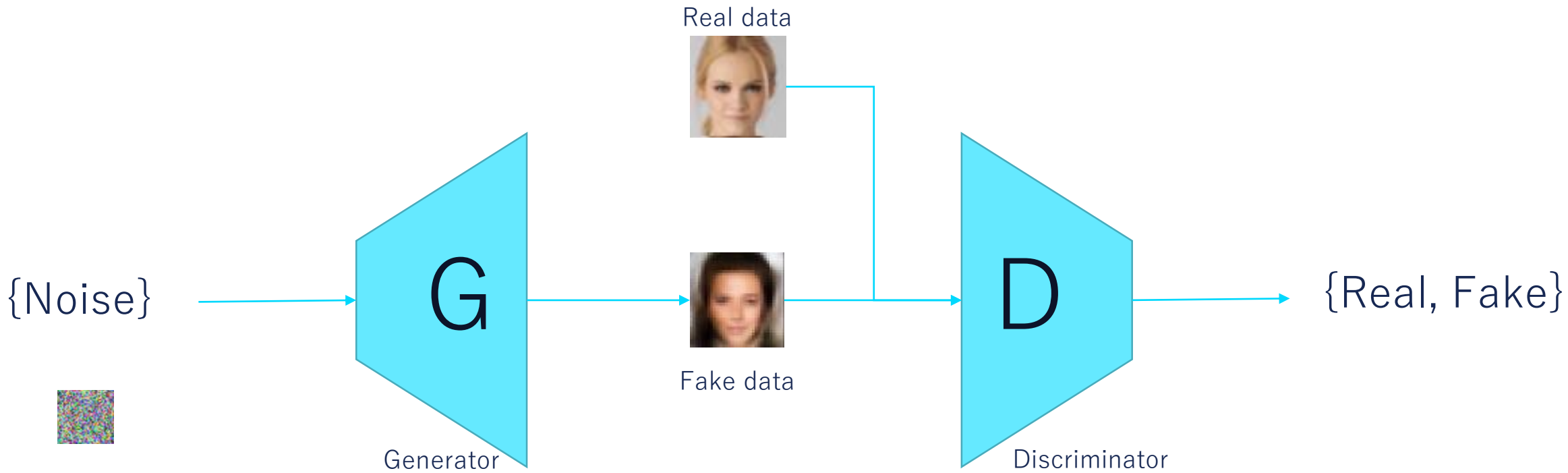
- ✓ Capsule Networks
- ✓ Attention

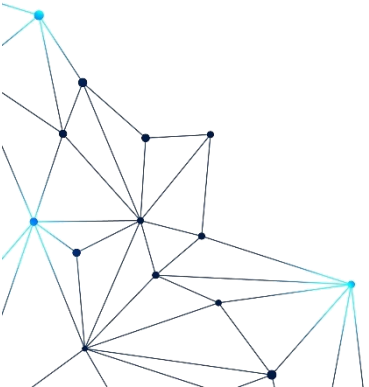


RectifAI

How these neurons could be mapped to other problems...

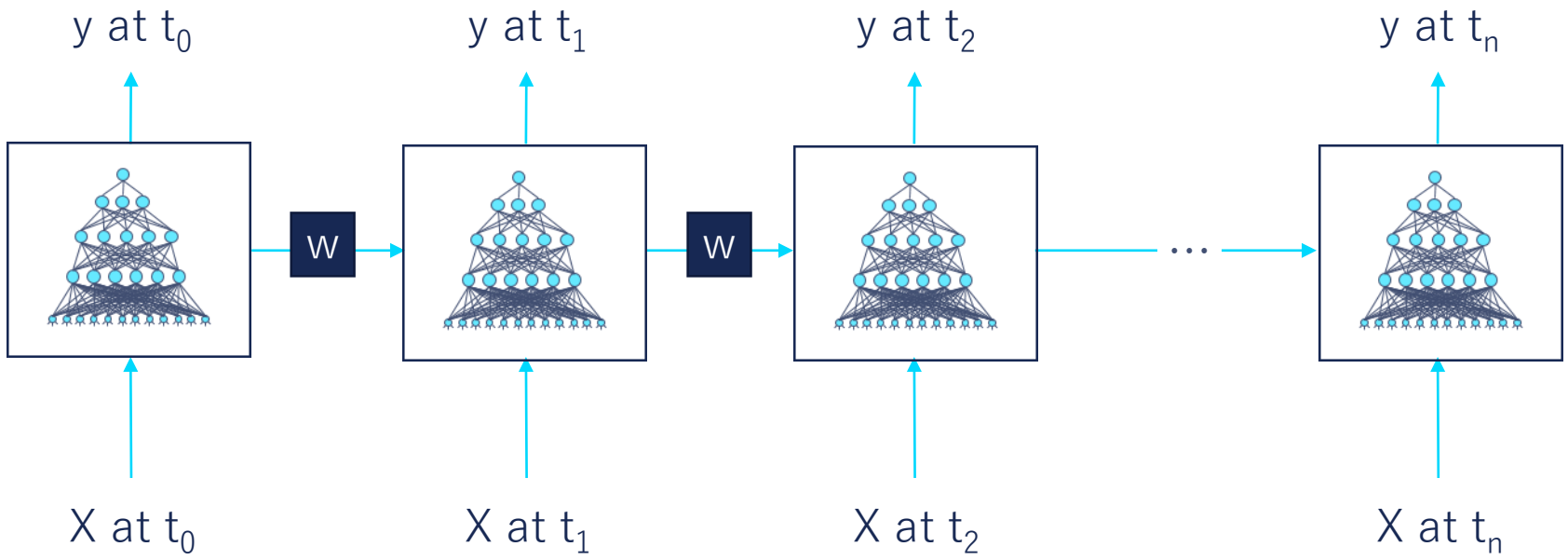
Generative Adversarial Network (GAN)

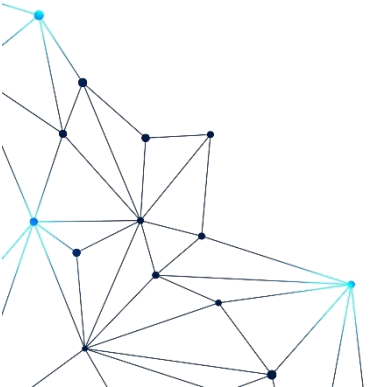


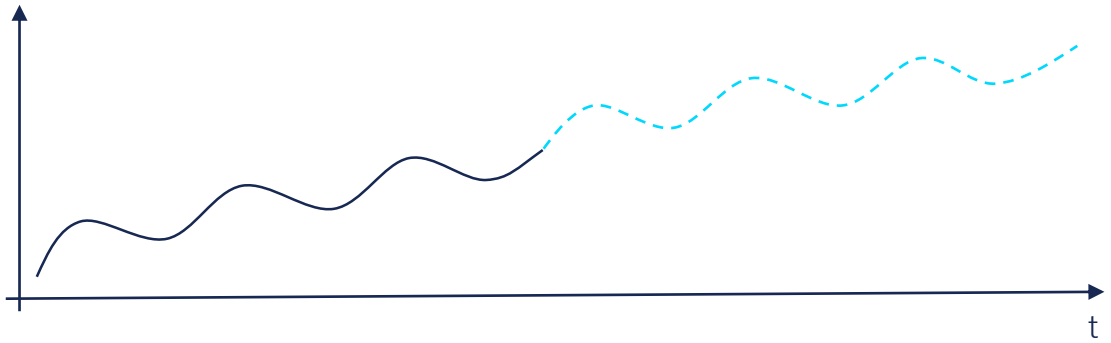


How these neurons could be mapped to other problems...

Recurrent Neural Network (RNN)



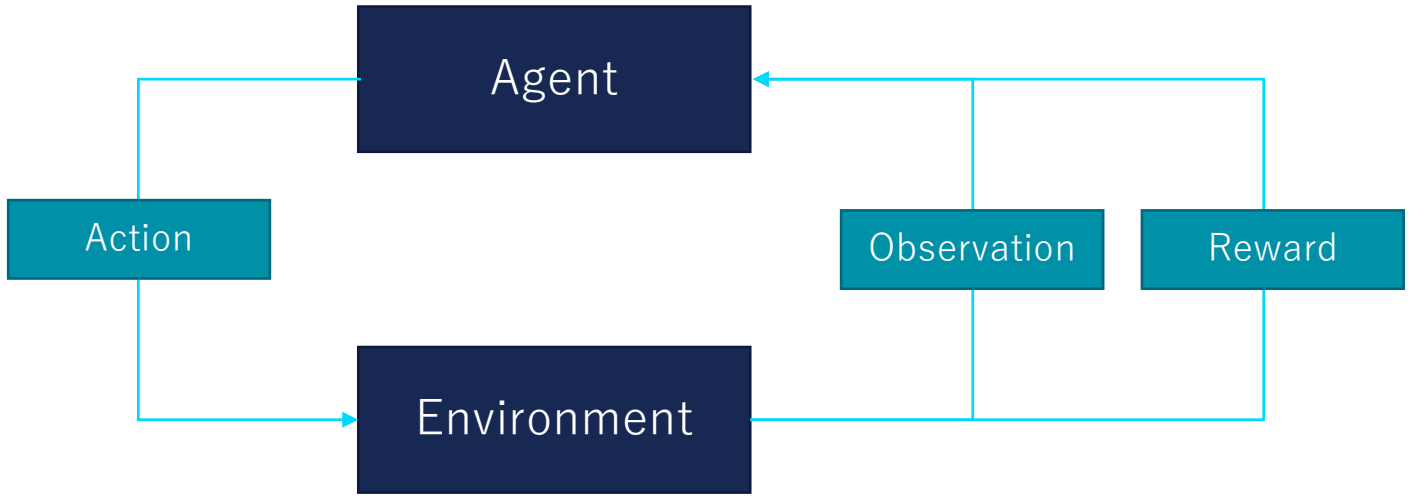


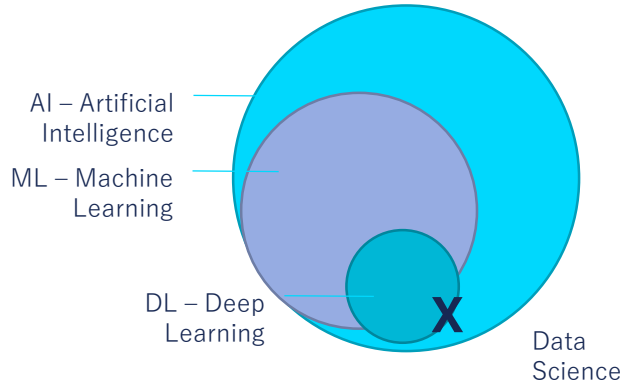


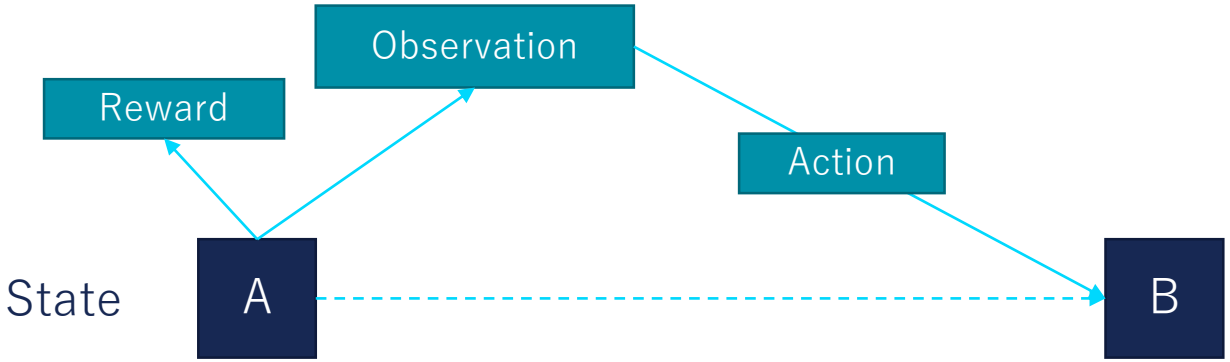
RectifAI

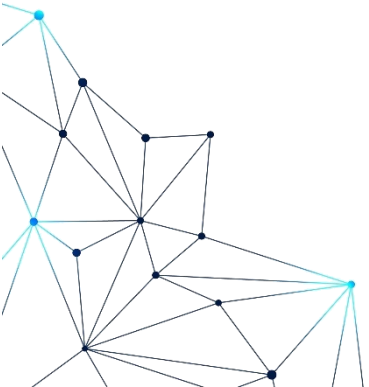
How these neurons could be mapped to other problems...

Reinforcement Learning (RL)









RectifAI

Winning keys to consider

1. Well defined objectives

- ✓ What are we looking to achieve exactly?
- ✓ What is the exact output we need to get?
- ✓ Is it a simple enough task?
  - ➢ If a human can't do it, neither can an AI

2. Data

- ✓ Is the right data available?
- ✓ How can I get it/process it?
- ✓ Is the data reliable?
  - ➢ Garbage in – Garbage out

Require leaders to create **a culture of continuous improvement and learning** It will be critical to implement programs that allow for **constant reskilling** of the workforce

[Harvard Business Review 2017]

3. Research level of the project

- ✓ Is science enough advanced on the subject?
- ✓ How much time, resources do we have for the project?
- ✓ Is there any model out there that can do a similar task?
  - ➢ Transfer learning

Rectif AI


Thank you!

Choosing the right data

> Which represent the conditions of use of your applications

> In sufficient quantity

- < 100,000 examples for Classic Machine Learning
- > 1,000,000 examples for Deep Learning
  - Data augmentation is possible if not enough examples

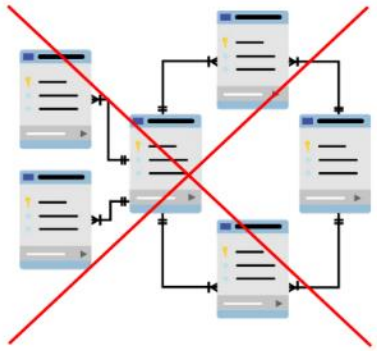
> Structured or not

> Labeled subset on which to train

> Available before starting the project

Getting the data

> A database is not a dataset. It needs processing to extract the right format



Text

Coordinates

Table

Images

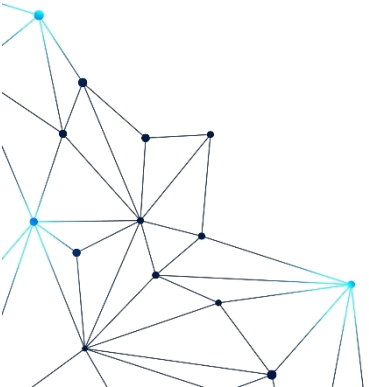
...

Number(s)

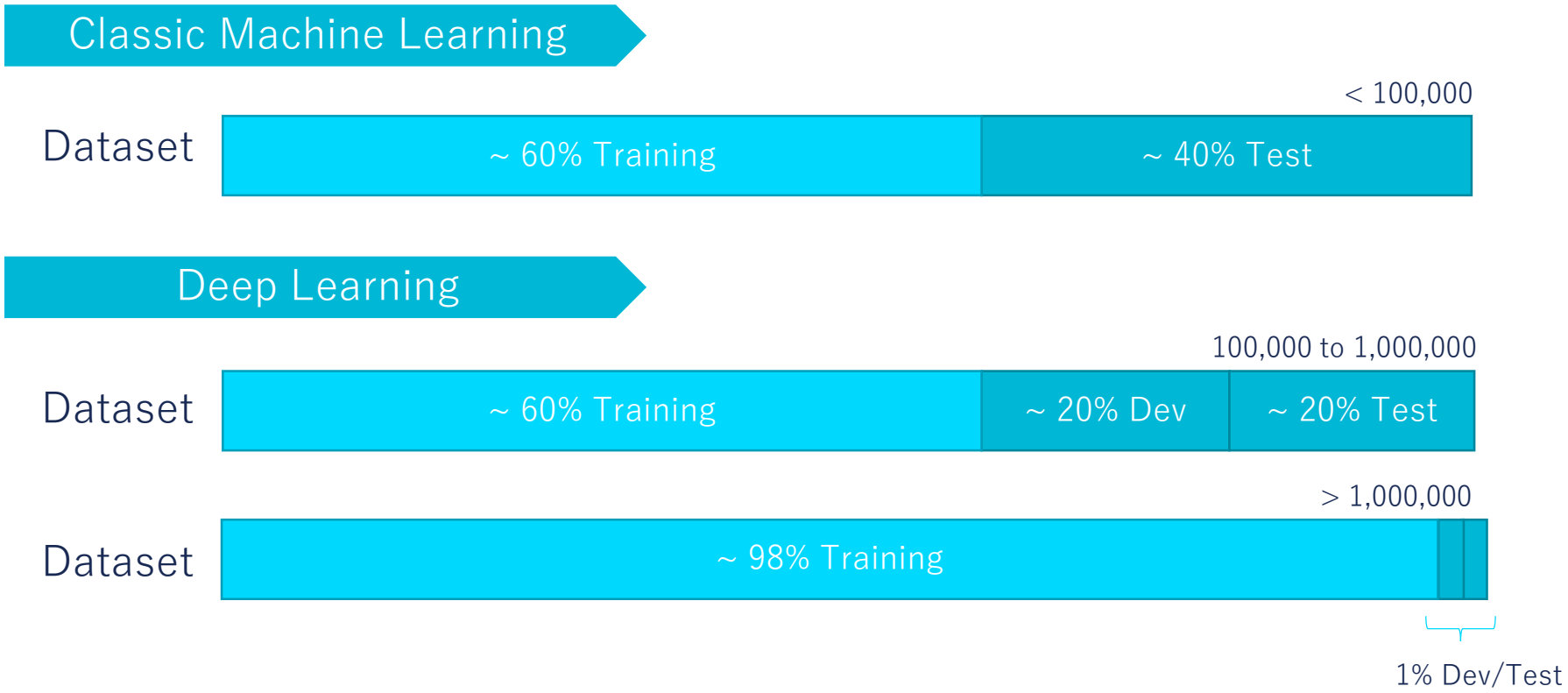
- Single
- Vector
- Matrix

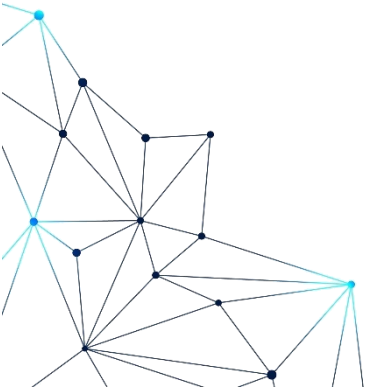
Tensor

Longest, but most important part of the project…

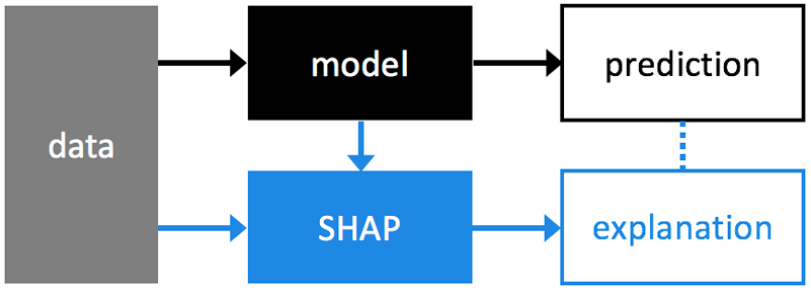


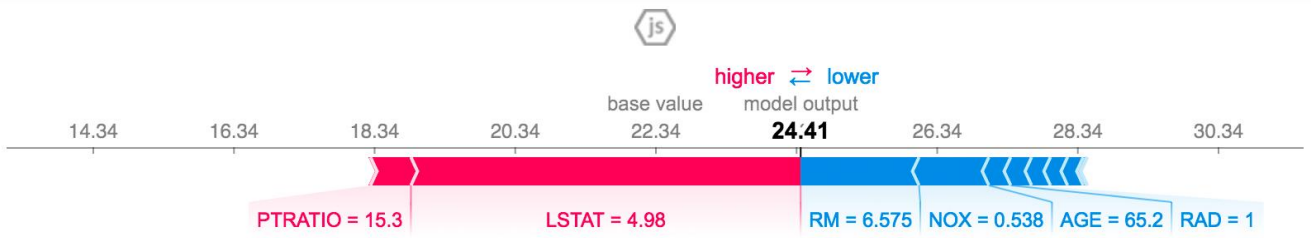
Dataset split

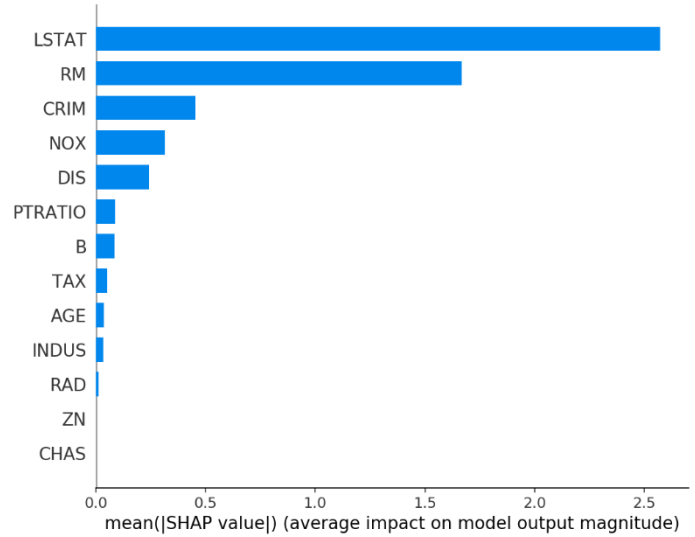


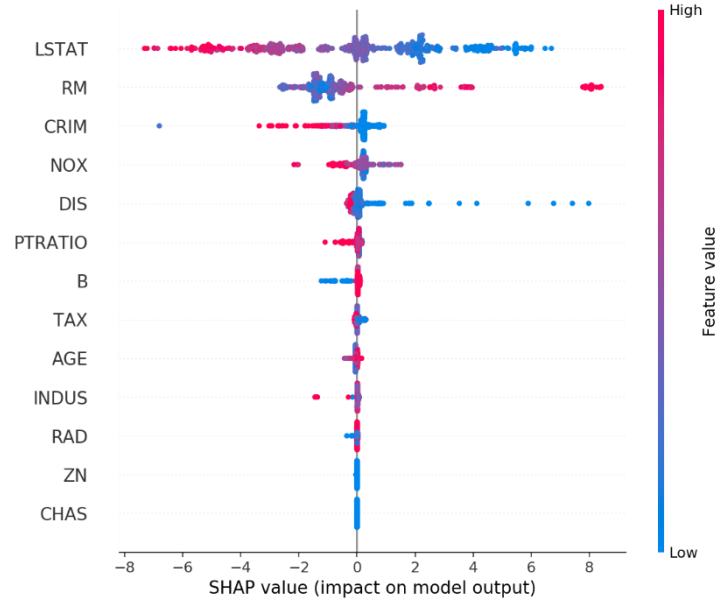


SHAP values









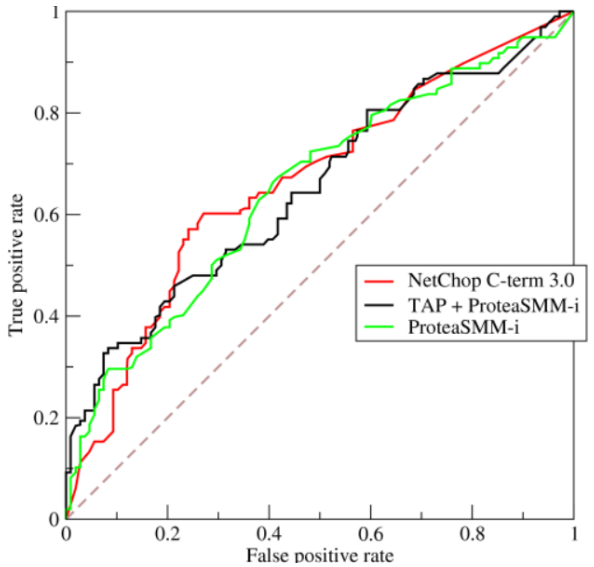
https://github.com/slundberg/shap

RectifAI

Evaluating the model

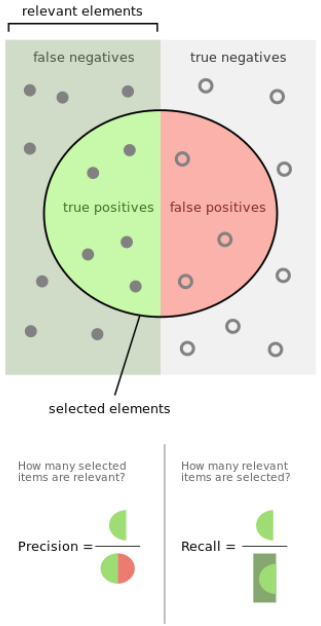
Classification task

ROC AUC score



https://en.wikipedia.org/wiki/Receiver_operating_characteristic

Precision and recall



https://en.wikipedia.org/wiki/Precision_and_recall

Regression task

Root mean squared error (RMSE)

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}$$

Mean Absolute Error (MAE)

$$\mathrm{MAE} = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j|$$

RMSE penalize errors more than MAE

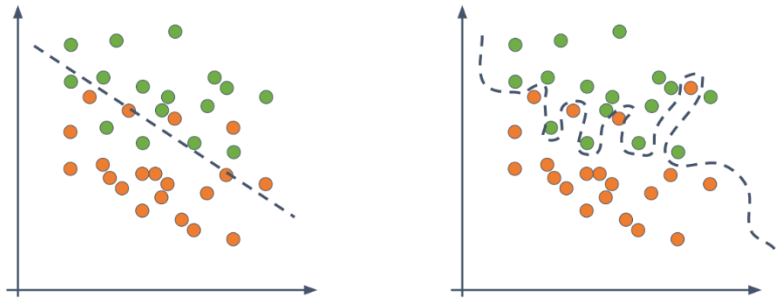
\+ accuracy score, average precision score, f1 score…

RectifAI

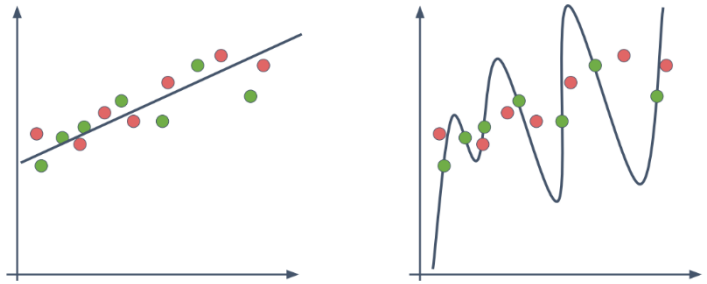
Overfitting

Overfitting

Overfitting : Classification

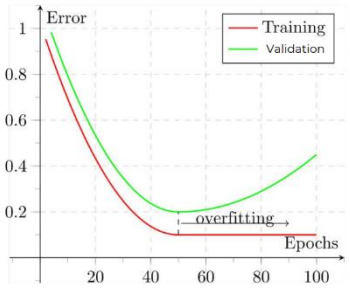


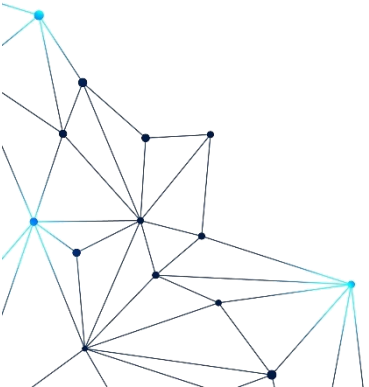
Overfitting : Regression



Regularisation

Ex. Early stopping





RectifAI