Machine Learning

Nicolas P. Rougier

July 2, 2018

Inria – Institute of Neurodegenerative Diseases – Bordeaux

What to learn?

Considering an input $\mathbf{x}$ (that can be an image, a text, a vector, a sequence, a set of features, a scalar, etc.)

Classification

To find the category $C_{\mathbf{x}}$ of some input $\mathbf{x}$

Regression

To find the value $f(\mathbf{x})$ of some input $\mathbf{x}$

Clustering

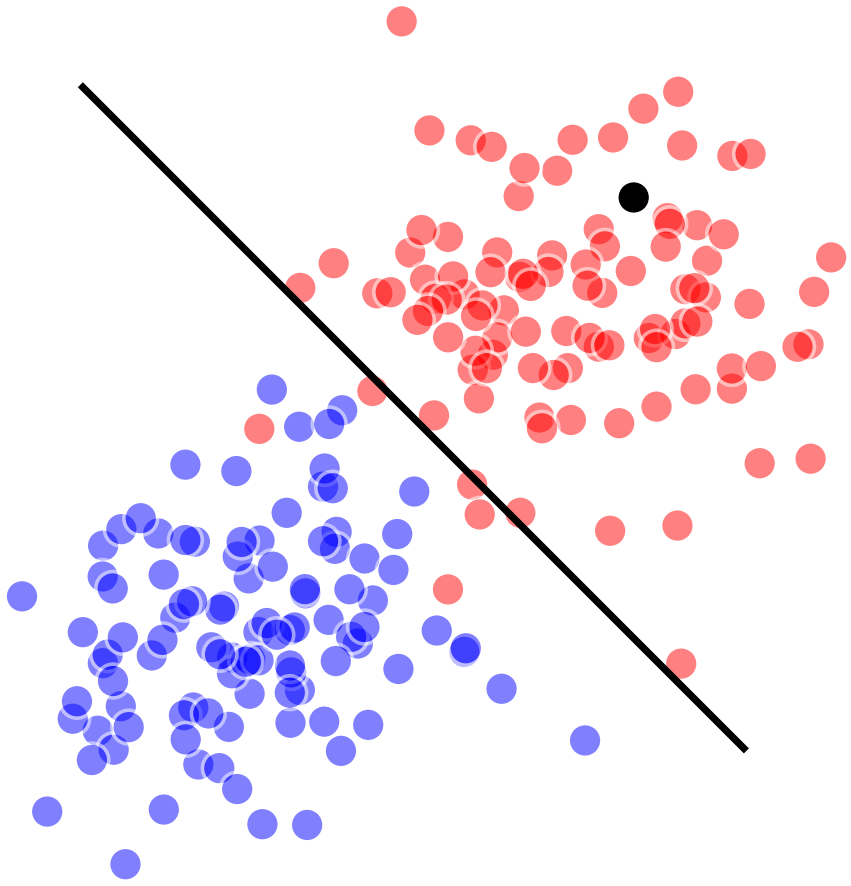
To divide all inputs $\{\mathbf{x}\}$ into distinct groups

Optimization

To find a sequence $\{\mathbf{a}_i\}$ such as to optimize $\sum f(\mathbf{a}_i)$

Classification

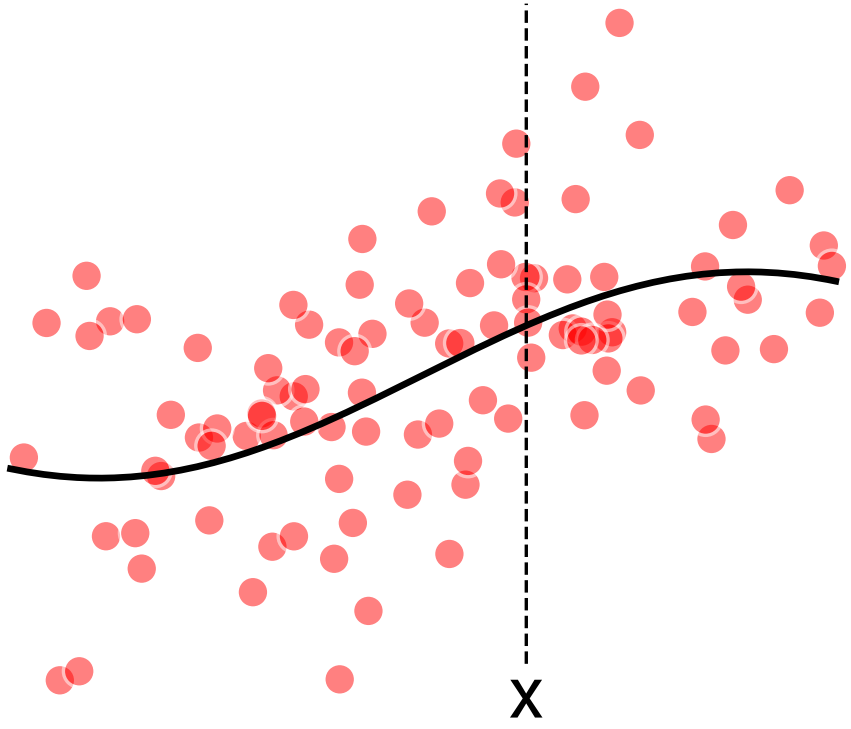
We know that (x_1, y_1) is blue, (x_2, y_2) is blue, (x_3, y_3) is red., etc.



What is the color of (x, y)?

Regression

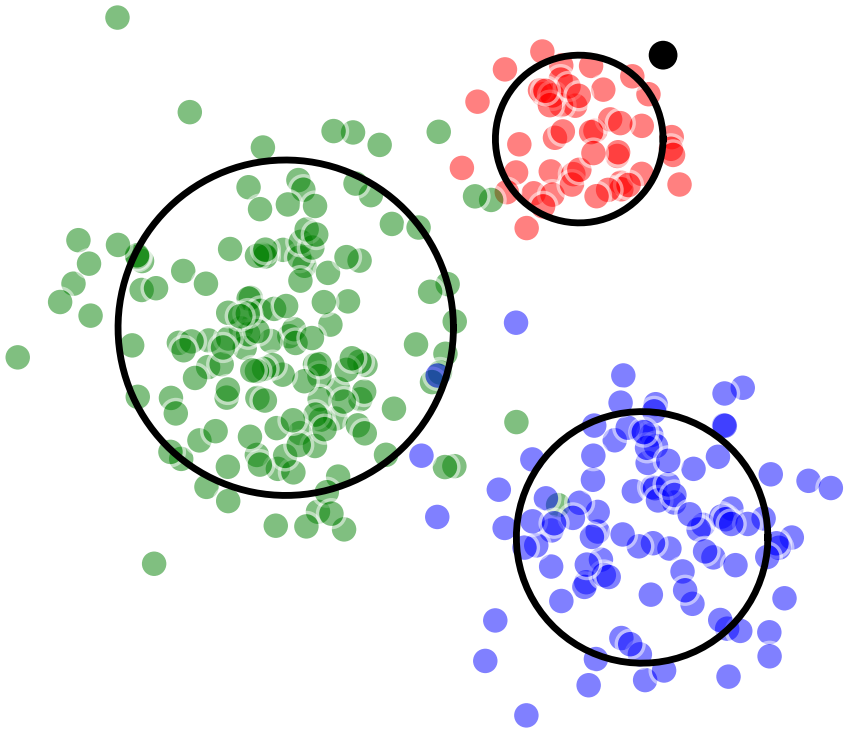
We know that $f(x_1) = y_1$, $f(x_2) = y_2$, $f(x_3) = y_3$, etc.



What is the value of $f(x)$?

Clustering

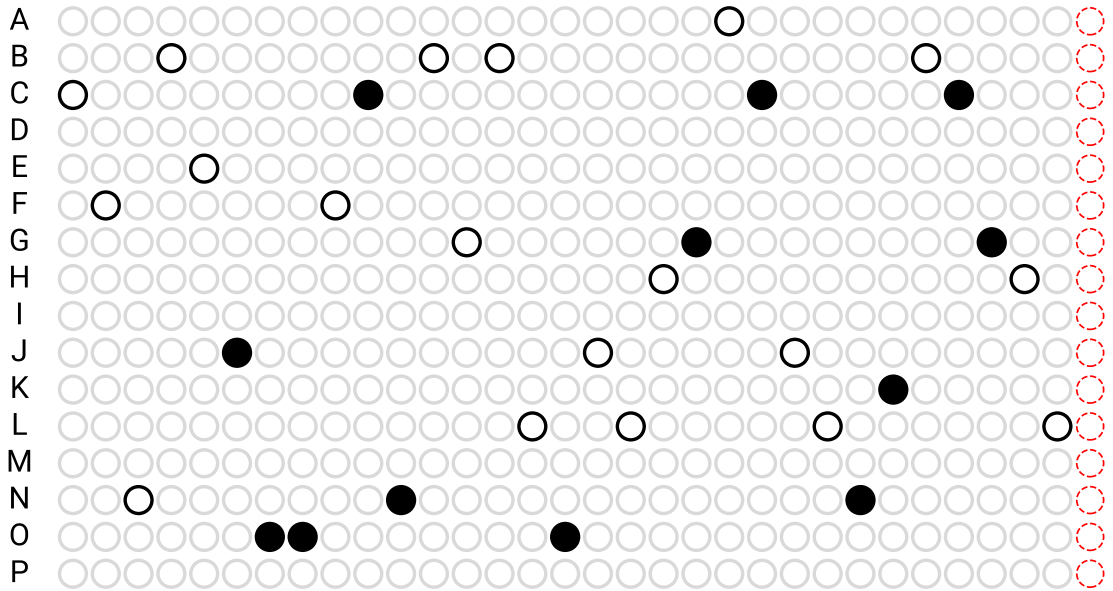
We have (x_1, y_1), (x_2, y_2), (x_3, y_3), etc.



What group (x, y) belongs to?

Optimization

You chose C and got no reward, you chose F and got no reward, you chose N and got no reward, etc.



What do you choose next?

How to learn?

Supervised learning

To learn a function that maps an input to an output based on example input-output pairs

Reinforcement learning

To find a policy that maps states to action such as to maximize long-term reward

Unsupervised learning

To infer a function that describes the structure of some (unlabeled) data

Machine learning ≠ Human learning

To learn anything, at any time, using any method and a few samples.

Explicit & implicit learning, lifelong learning, imitation learning, one-shot learning, declarative and procedural memory, short and long term memory, working memory, forgetting, motivation, emotion, etc.

The machine-learning pipeline

Data (60% of your time)

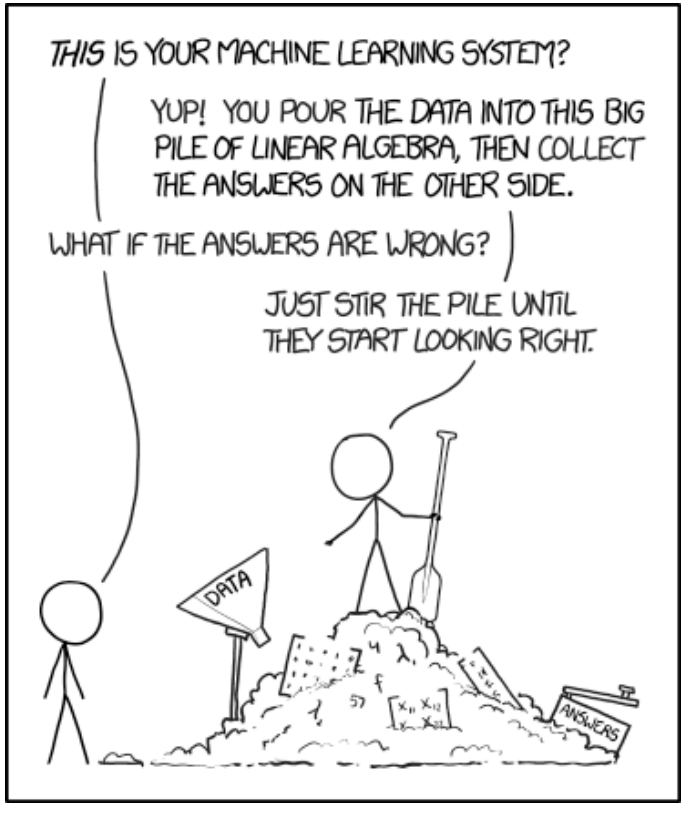
Analysis, acquisition, cleaning, preparation, augmentation, balance

Model (30% of your time)

Hyperparameters, pre-training, training, tweaking, testing

Deployment (10% of your time)

Optimization, dockerization, evaluation



Common pitfalls

Data

Small, unbalanced, biased, leakage, variance, etc

Model

Initialization, local minima, curse of dimensionality, etc

Deployment

Scalability, fragility, ethics, etc

a woman riding a horse on a dirt road

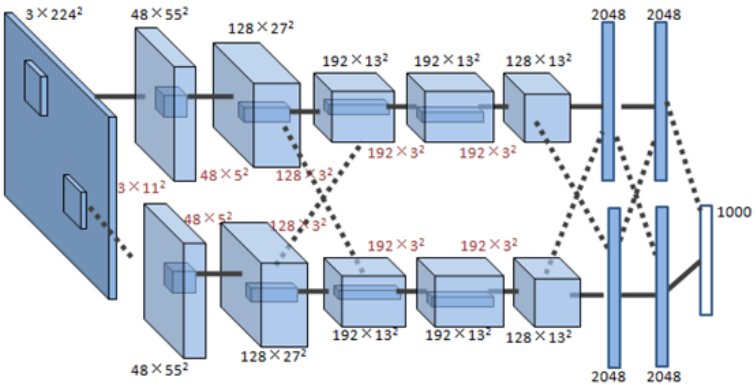
an airplane is parked on the tarmac at an airport

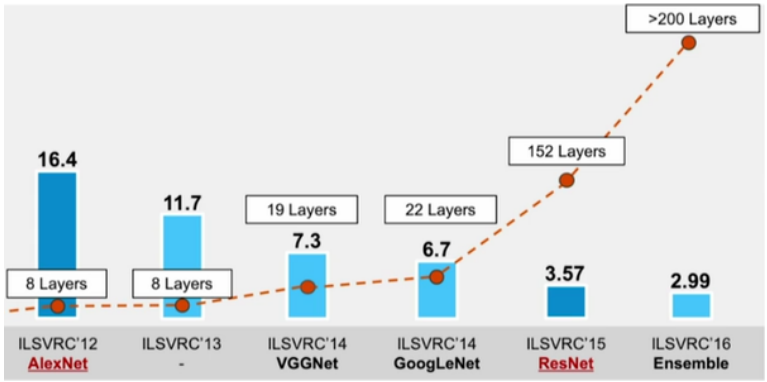
a group of people standing on top of a beach

How big is “big”?

Deep convolutional network (Hinton, 2010)

- 5 convolutional layers
- 650,000 units
- 60,000,000 parameters
- 1,200,000 train images
- 100,000 test images
- 1,000 classes

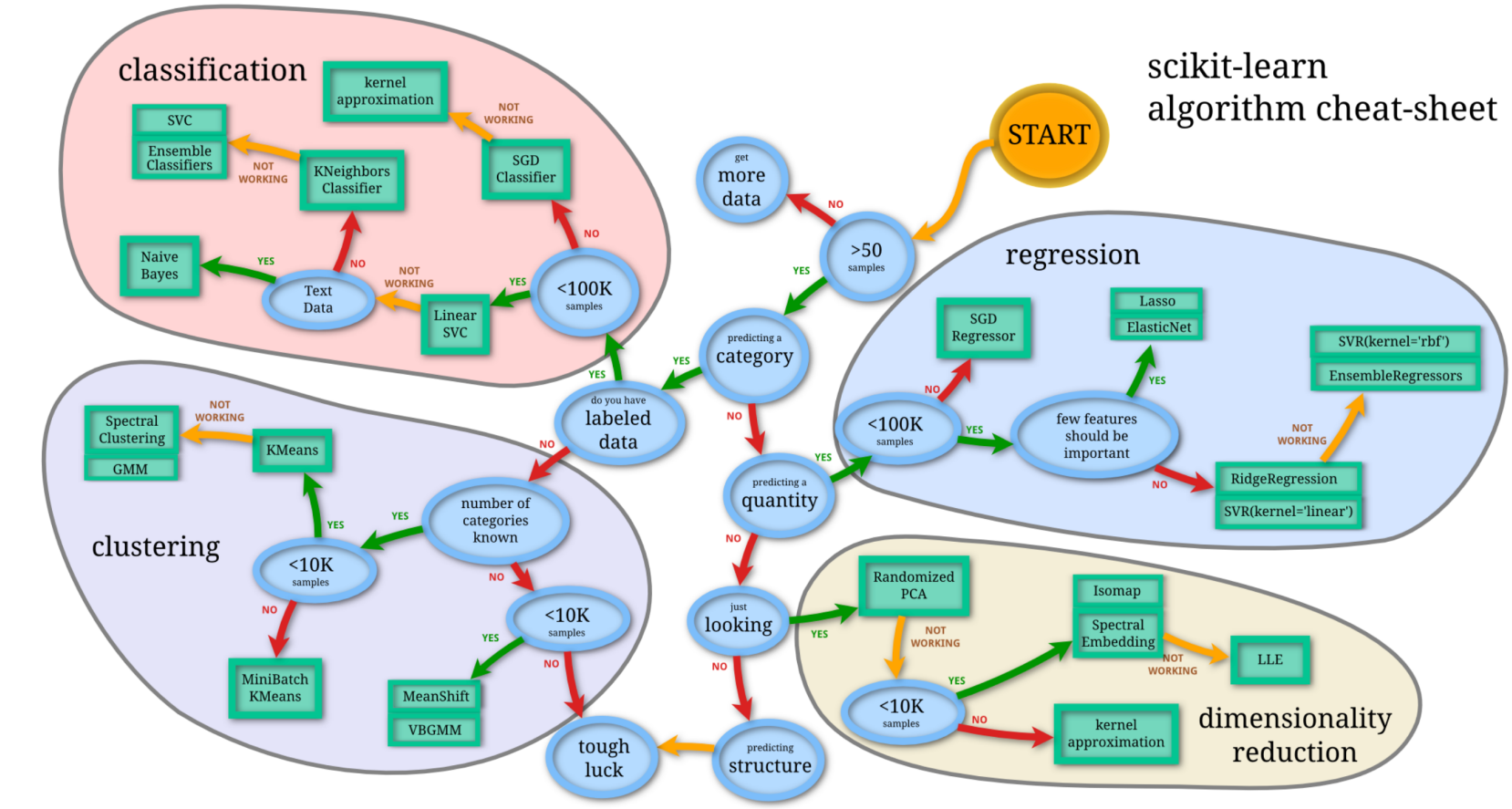




Latest OpenAI DOTA player (June 2018): 180 years worth of games, 256 GPUs and 128,000 CPU cores.

Do I need deep learning?

If all you have is a hammer, everything looks like a nail.



Conclusion

Lot of challenges ahead

- Learn with a few samples
- Lifelong learning
- Transfer learning
- Unbiased learning
- Robust learning
- Unsupervised learning
- Reinforcement learning

Lot of opportunities as well...