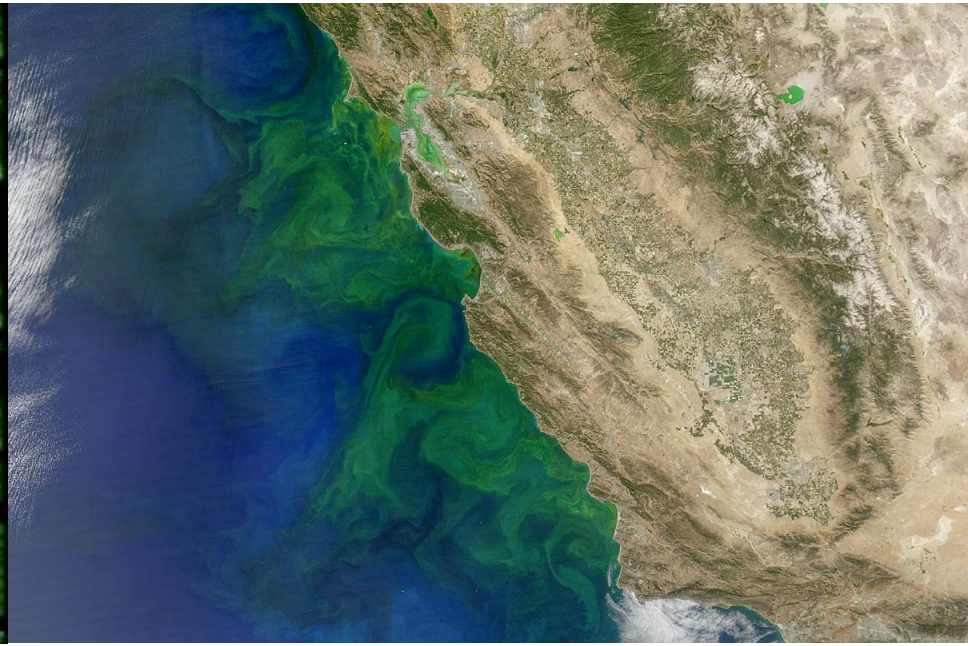
Machine Learning and Ocean Optics



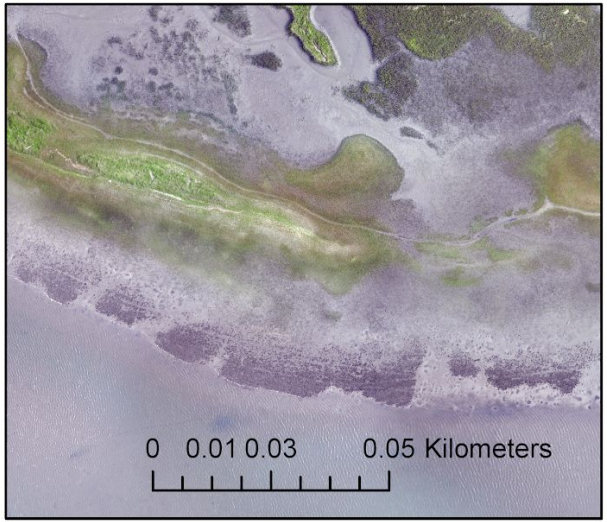
Patrick Gray

Ocean Optics Class 2023

Machine Learning

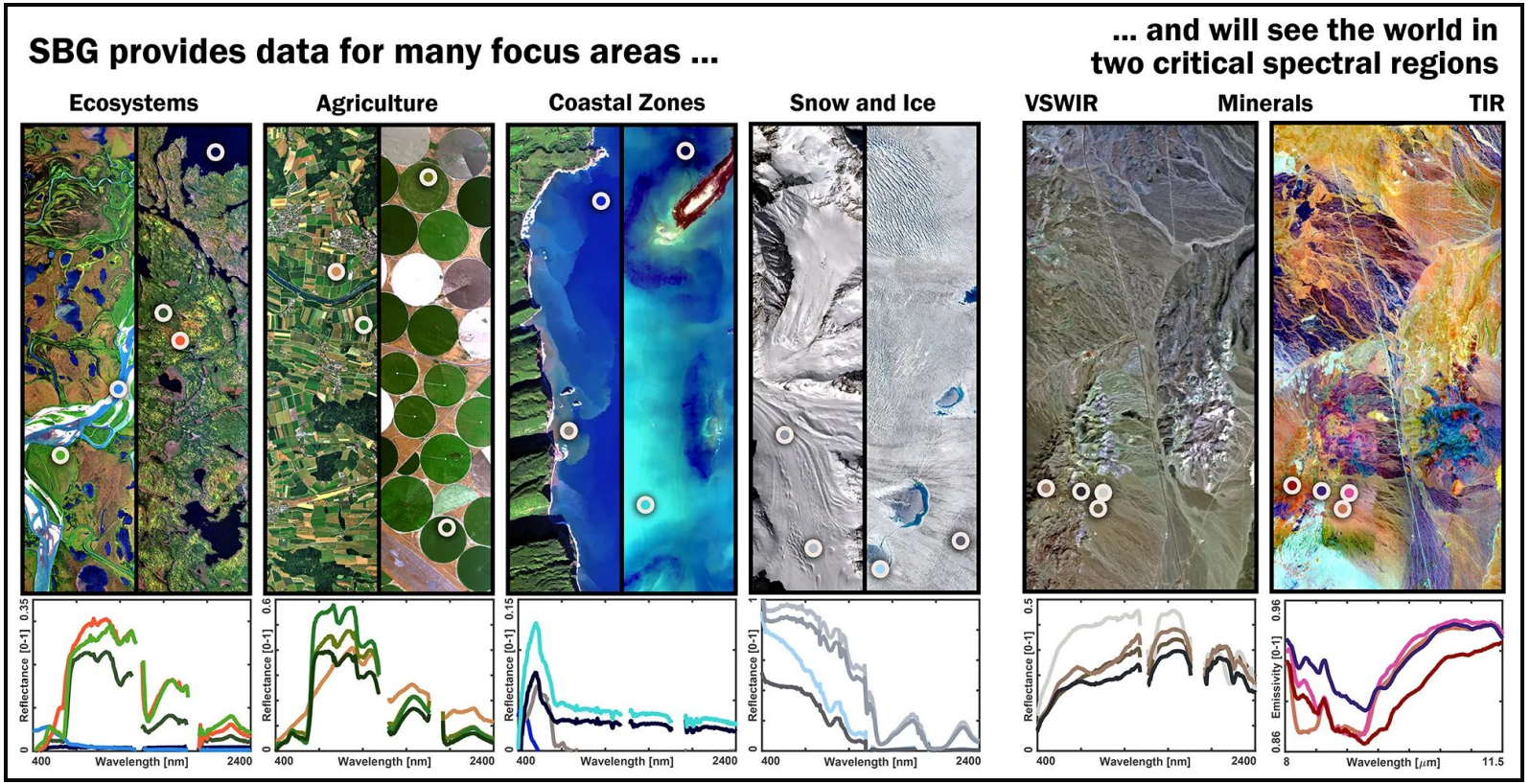
ML algorithms perform a specific task without using explicit instructions, relying on patterns and inference instead.

Why Machine Learning for Remote Sensing



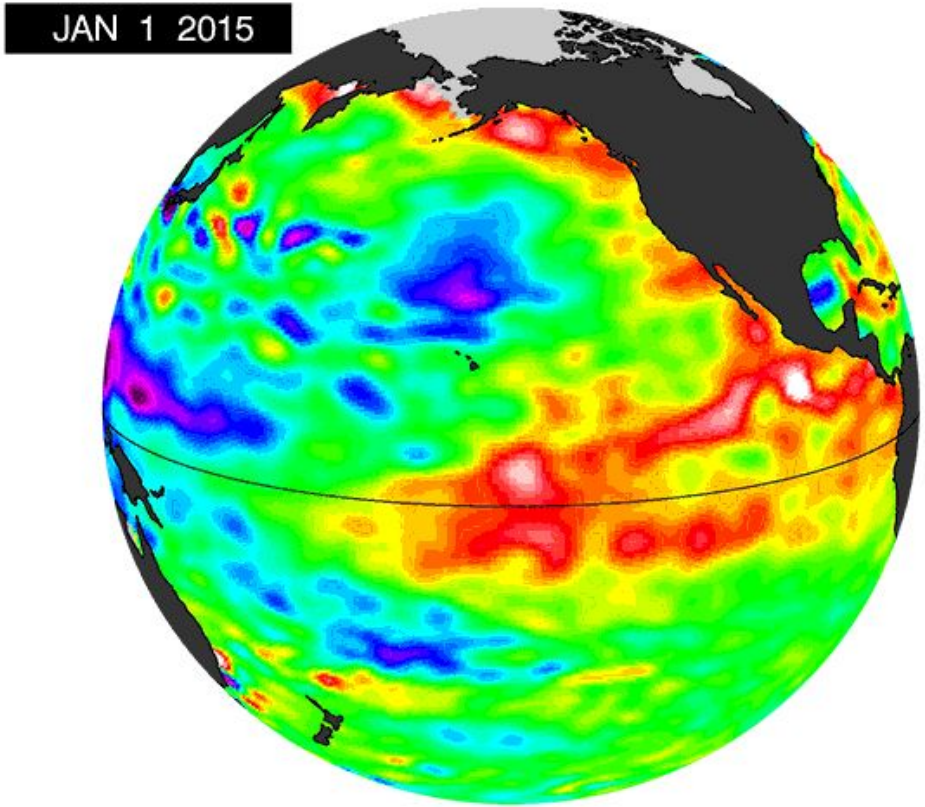
Ridge et al 2020

Why Machine Learning for Remote Sensing



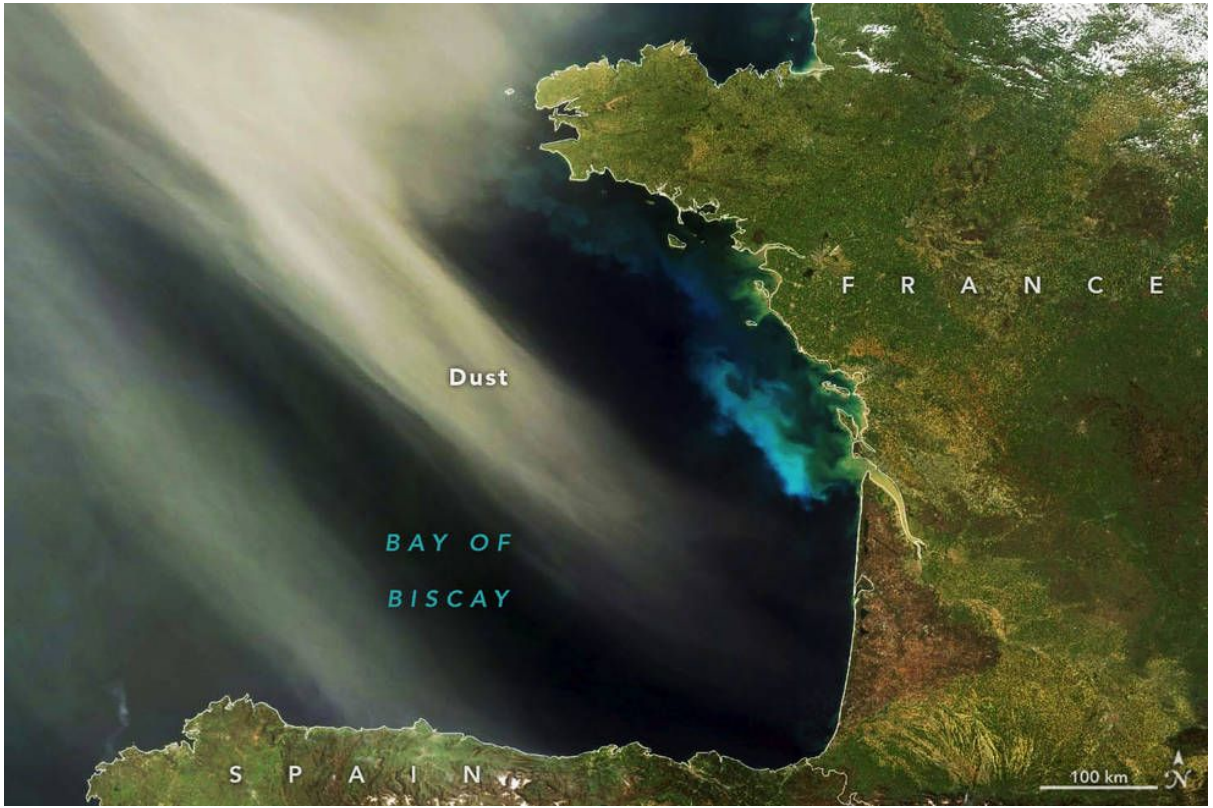
Schneider et al 2019

Why Machine Learning for Remote Sensing



NASA Goddard

Why Machine Learning for Remote Sensing



Wanmei Liang

What does machine learning do?

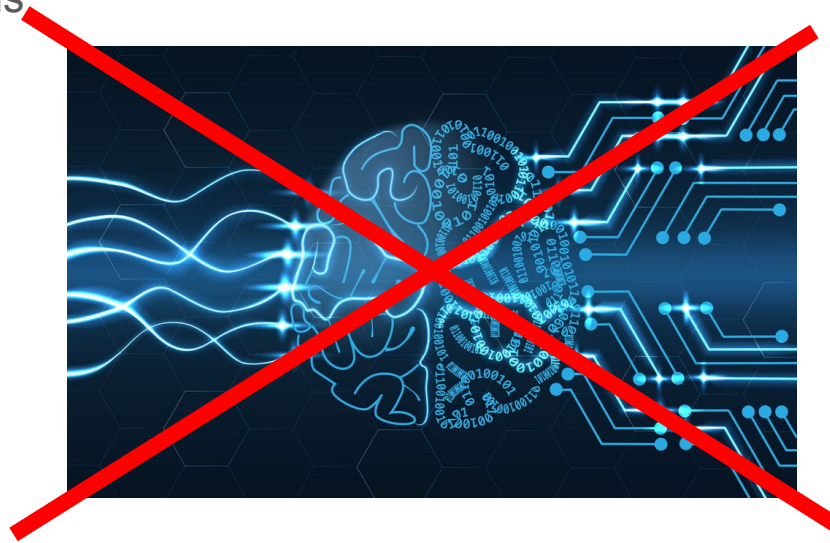
At its most basic ML systems classify input data, cluster data, recognize patterns, and regress data into models

What does machine learning do?

At its most basic ML systems classify input data, cluster data, recognize patterns, and regress data into models

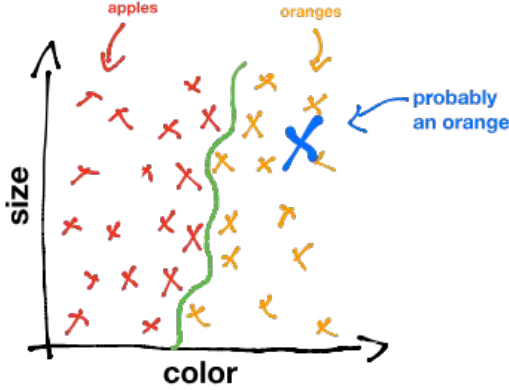
What does machine learning do?

At its most basic ML systems classify input data, cluster data, recognize patterns, and regress data into models



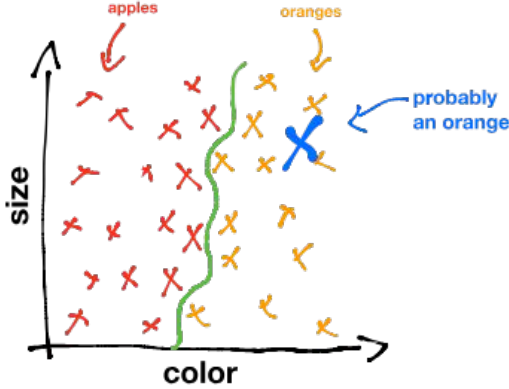
What does machine learning do?

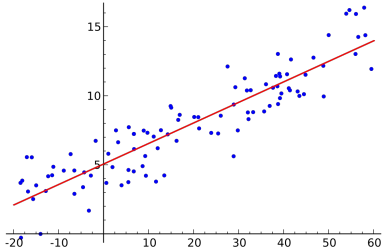
At its most basic ML systems classify input data, cluster data, recognize patterns, and regress data into models



What does machine learning do?

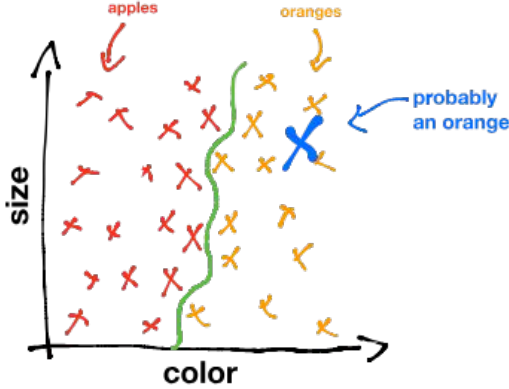
At its most basic ML systems classify input data, cluster data, recognize patterns, and regress data into models

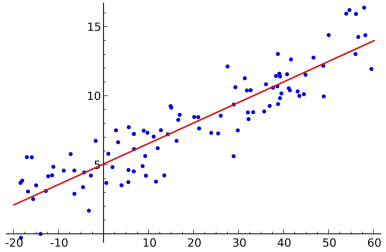


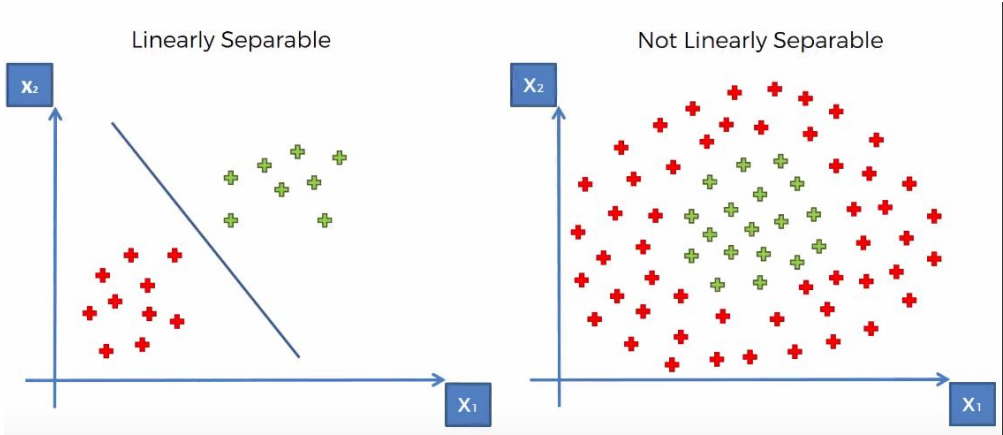


What does machine learning do?

At its most basic ML systems classify input data, cluster data, recognize patterns, and regress data into models





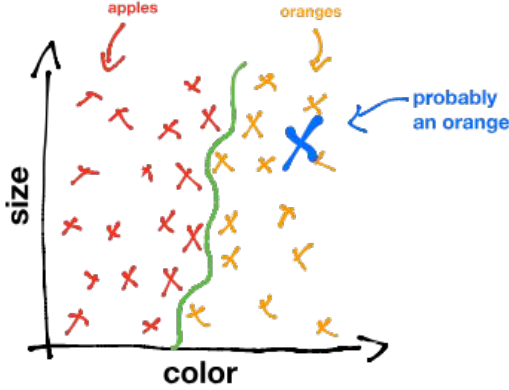


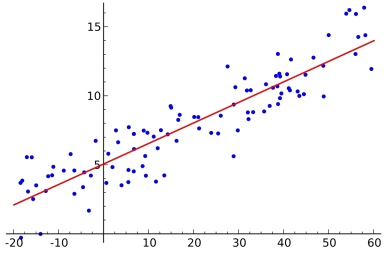
What does machine learning do?

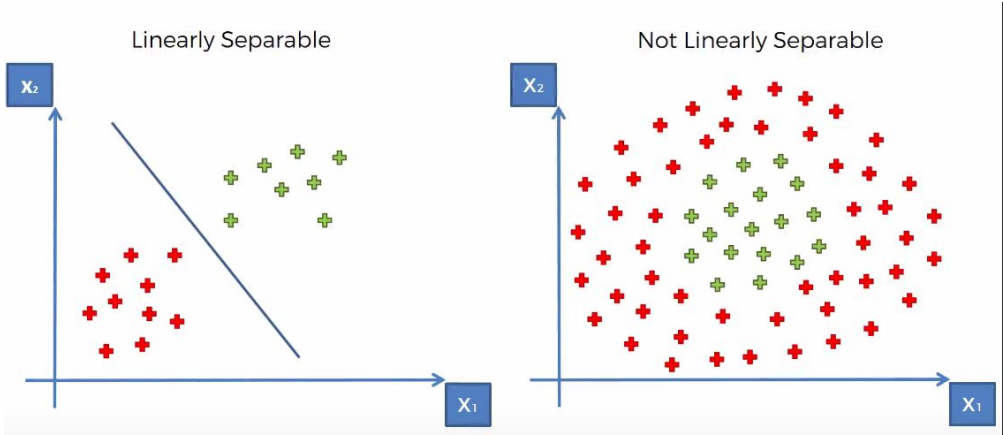
At its most basic ML systems classify input data, cluster data, recognize patterns, and regress data into models

This could mean:

- face recognition
- predicting home prices
- modeling animal habitat
- **estimating chl-a**
- predict what you will buy
- **filling cloud gaps**
- predict netflix preferences
- **predicting PCC**







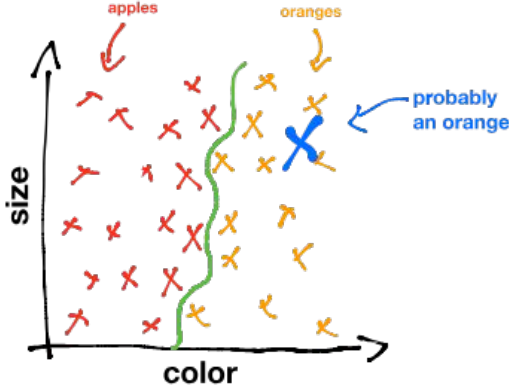
What does machine learning do?

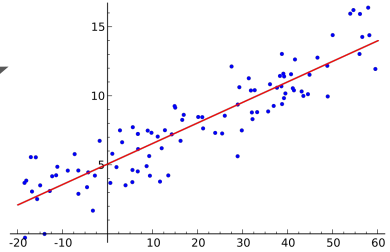
At its most basic ML systems classify input data, cluster data, recognize patterns, and regress data into models

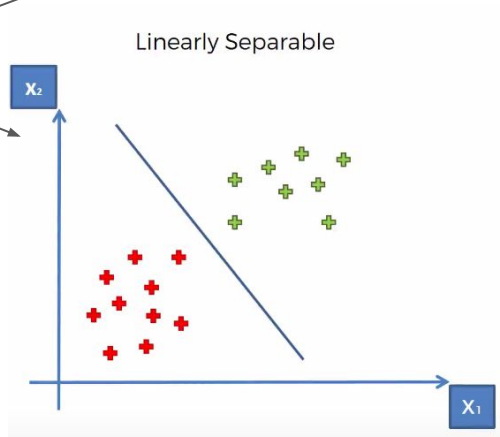
This could mean:

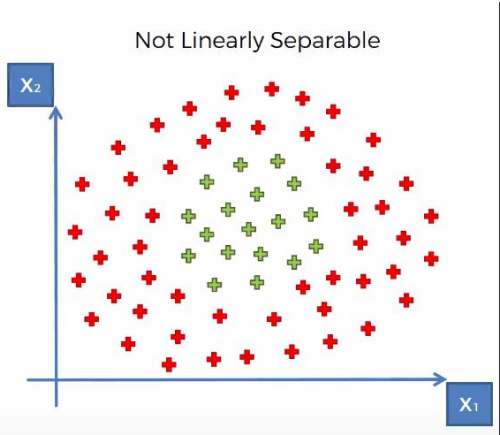
- face recognition
- predicting home prices
- modeling animal habitat
- **estimating chl-a**
- predict what you will buy
- **filling cloud gaps**
- predict netflix preferences
- **predicting PCC**

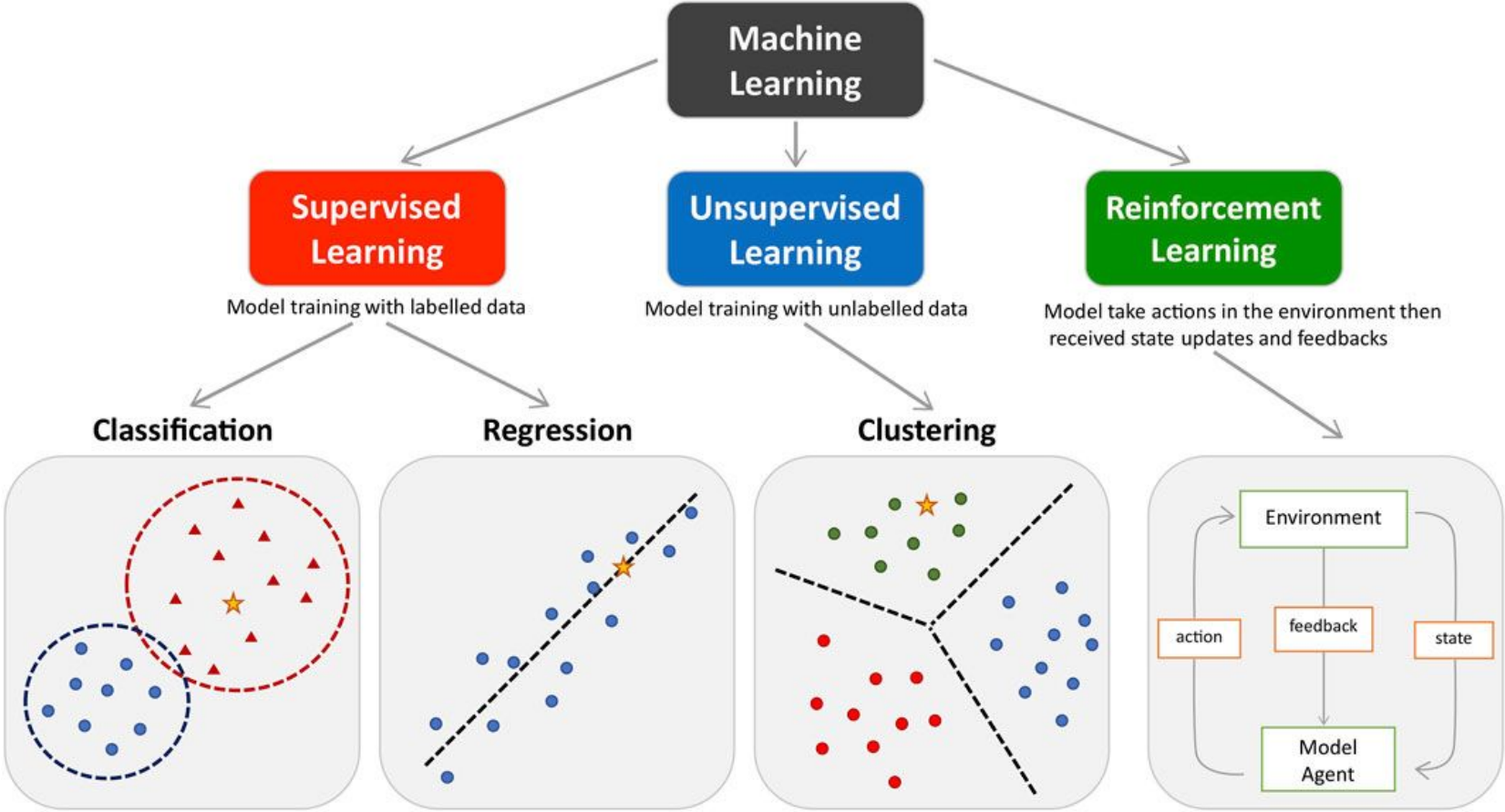
Just high-dimensional

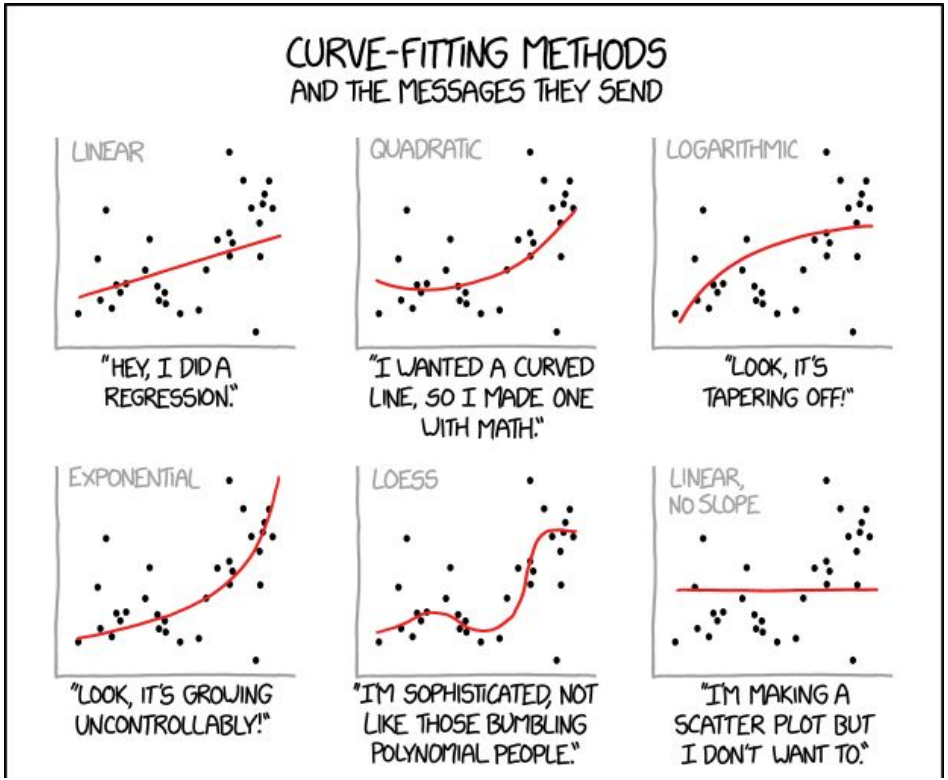


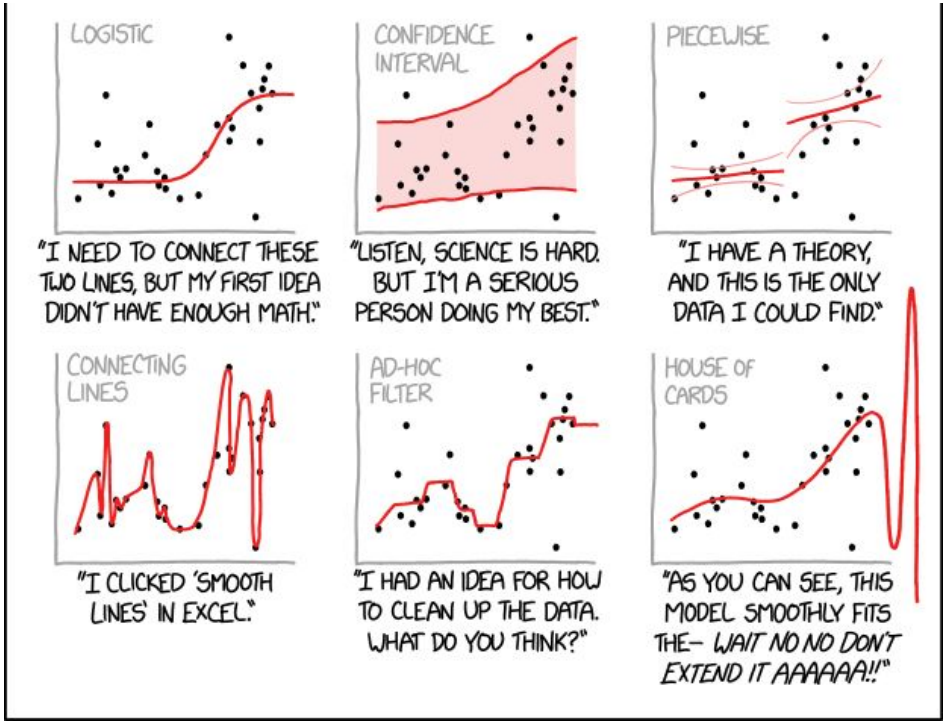






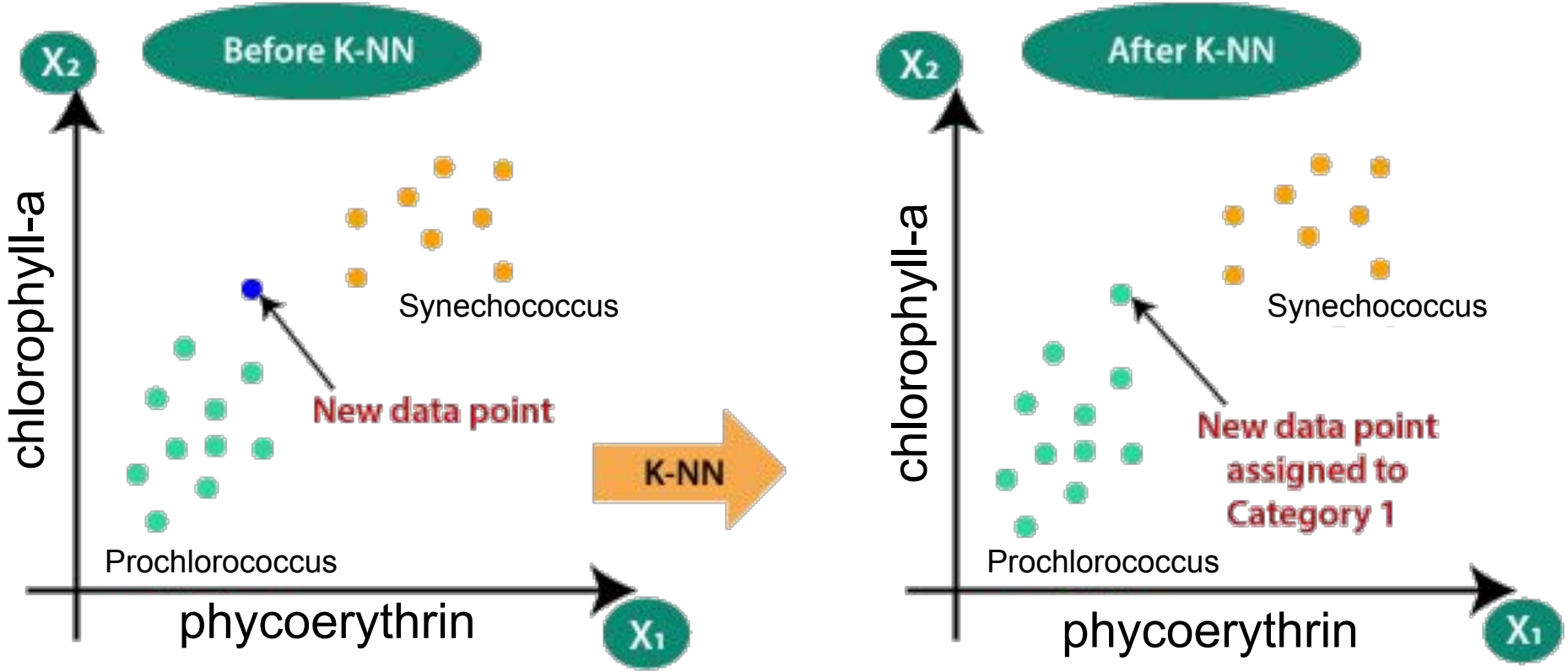

Machine Learning
Supervised Learning
Unsupervised Learning
Reinforcement Learning
Model training with labelled data
Model training with unlabelled data
Model take actions in the environment then received state updates and feedbacks
Classification
Regression
Clustering
Environment
action
feedback
state
Model Agent


CURVE-FITTING METHODS
AND THE MESSAGES THEY SEND
LINEAR
"HEY, I DID A REGRESSION."
QUADRATIC
"I WANTED A CURVED LINE, SO I MADE ONE WITH MATH."
LOGARITHMIC
"LOOK, IT'S TAPERING OFF!"
EXPONENTIAL
"LOOK, IT'S GROWING UNCONTROLLABLY!"
LOESS
"I'M SOPHISTICATED, NOT LIKE THOSE BUMBLING POLYNOMIAL PEOPLE."
LINEAR, NO SLOPE
"I'M MAKING A SCATTER PLOT BUT I DON'T WANT TO."

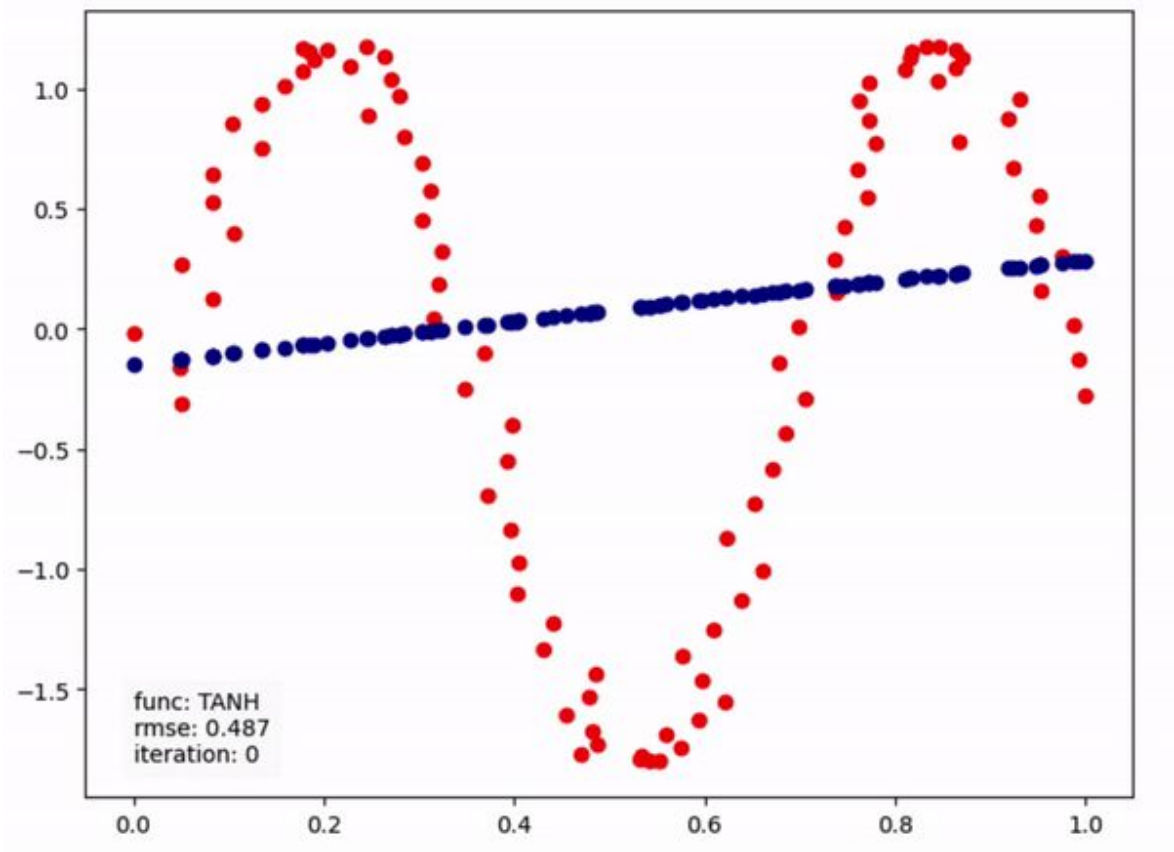

LOGISTIC
"I NEED TO CONNECT THESE TWO LINES, BUT MY FIRST IDEA DIDN'T HAVE ENOUGH MATH."
CONFIDENCE INTERVAL
"LISTEN, SCIENCE IS HARD. BUT I'M A SERIOUS PERSON DOING MY BEST."
PIECEWISE
"I HAVE A THEORY, AND THIS IS THE ONLY DATA I COULD FIND."
CONNECTING LINES
"I CLICKED 'SMOOTH LINES' IN EXCEL."
AD-HOC FILTER
"I HAD AN IDEA FOR HOW TO CLEAN UP THE DATA. WHAT DO YOU THINK?"
HOUSE OF CARDS
"AS YOU CAN SEE, THIS MODEL SMOOTHLY FITS THE— WAIT NO NO DON'T EXTEND IT AAAAAA!!"

XKCD

Classify phytoplankton types

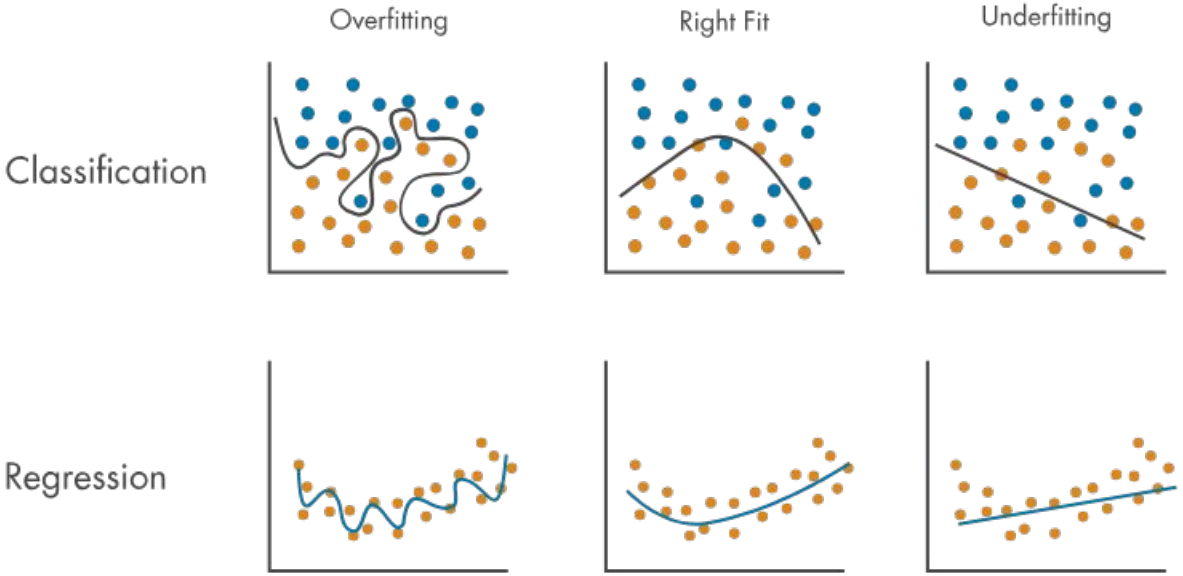


Regress chlorophyll-a



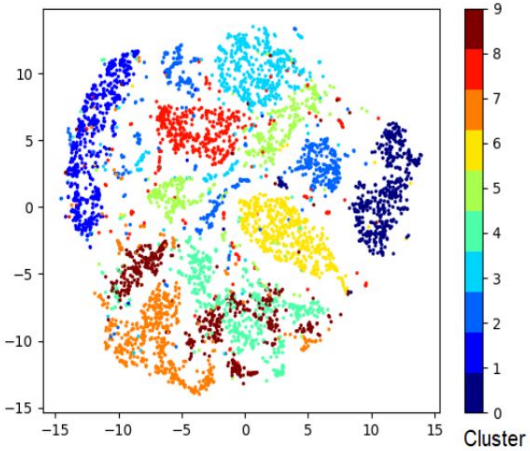
“All models are wrong, some are useful”

- some ML models can be wrongly in wildly unpredictable ways
- must aggressively test and convince yourself it is robust

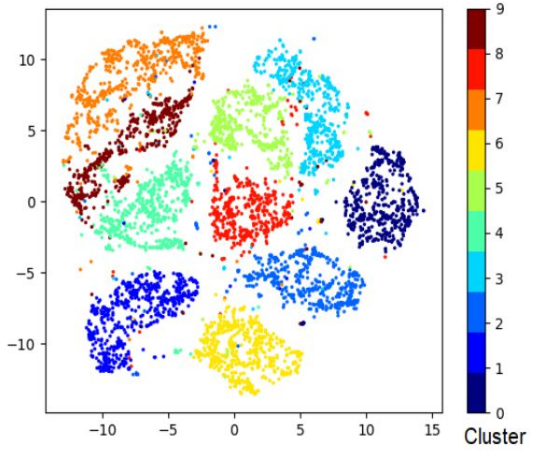


Clustering

Finding spatial/spectral/temporal patterns, outliers, etc

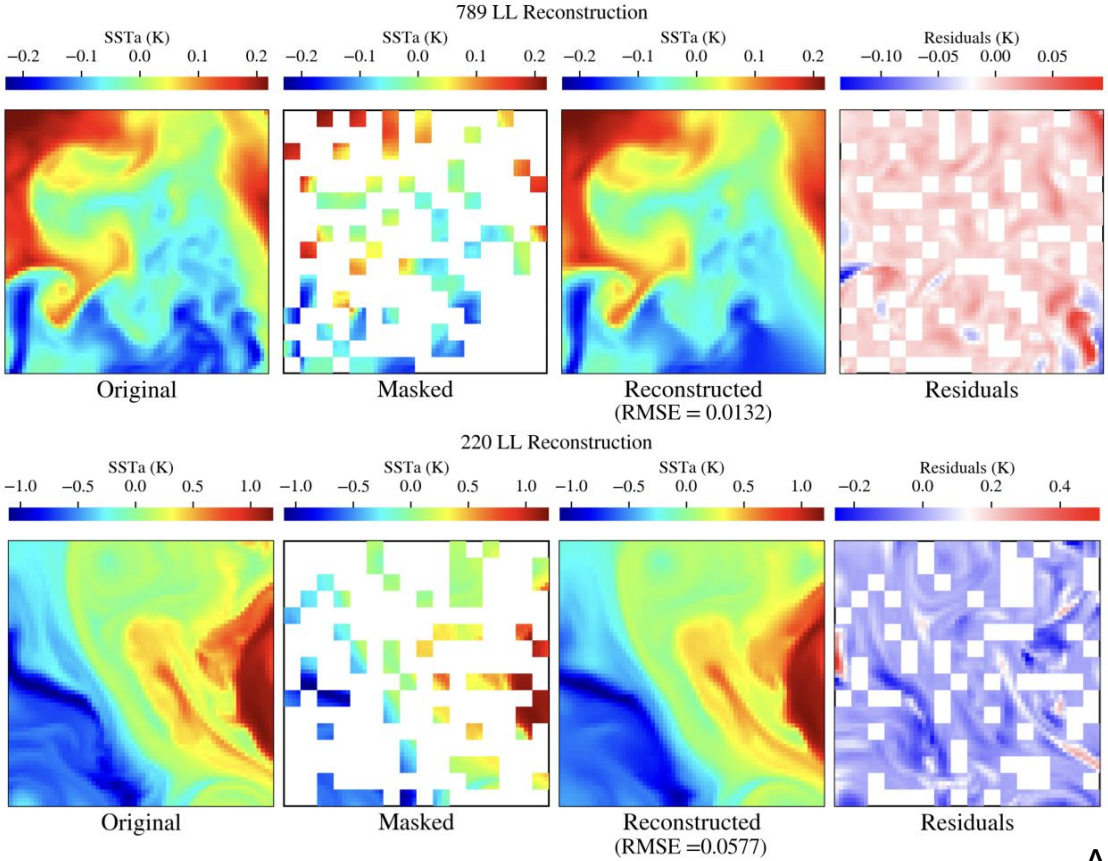


(a) k-Means



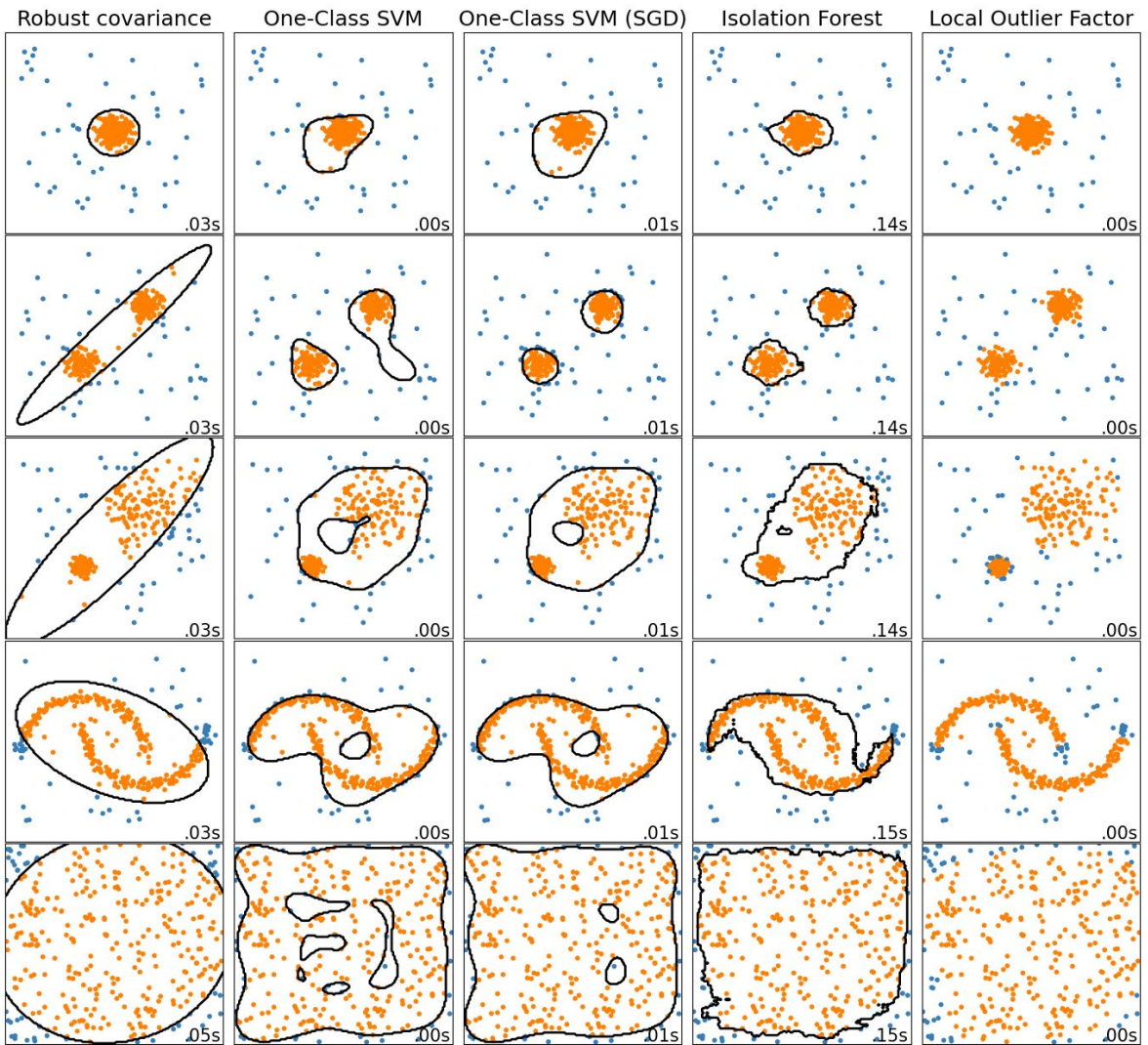
(b) Autoencoder + k-Means

Interpolation



Agabin & Prochaska 2023

Outlier Detection



So what is Deep Learning?

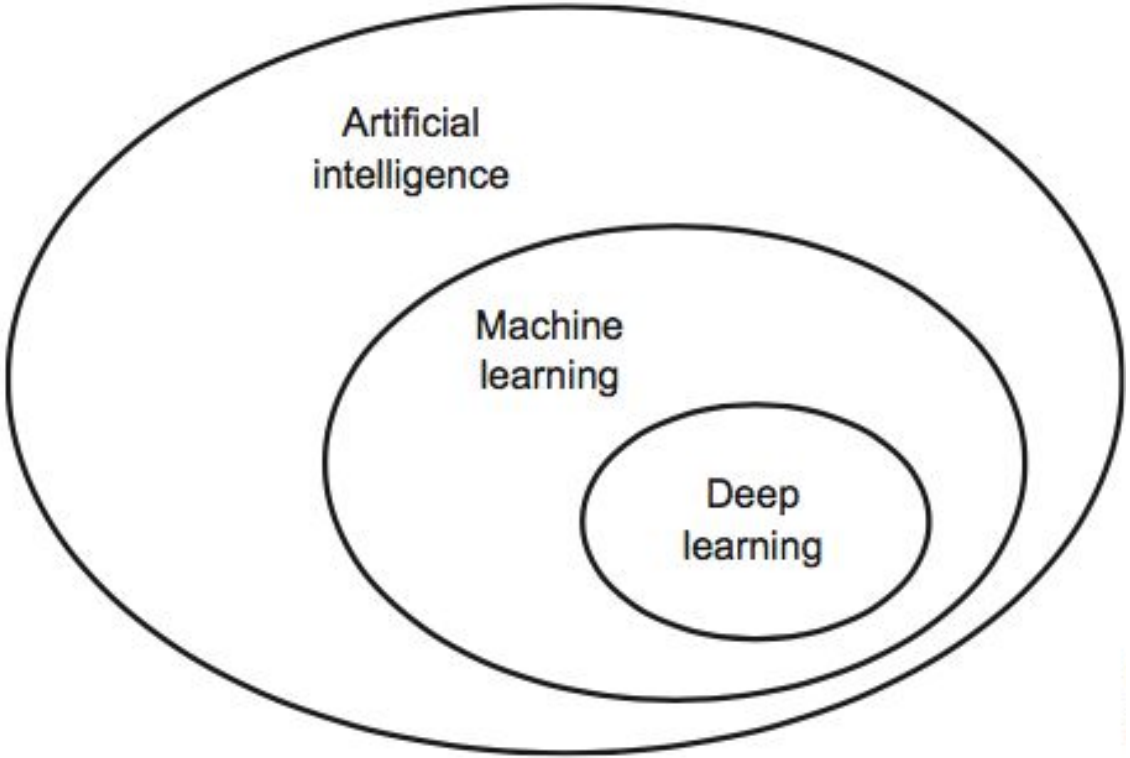


Figure Credits: DL with Python by Francois Chollet

So what is Deep Learning?

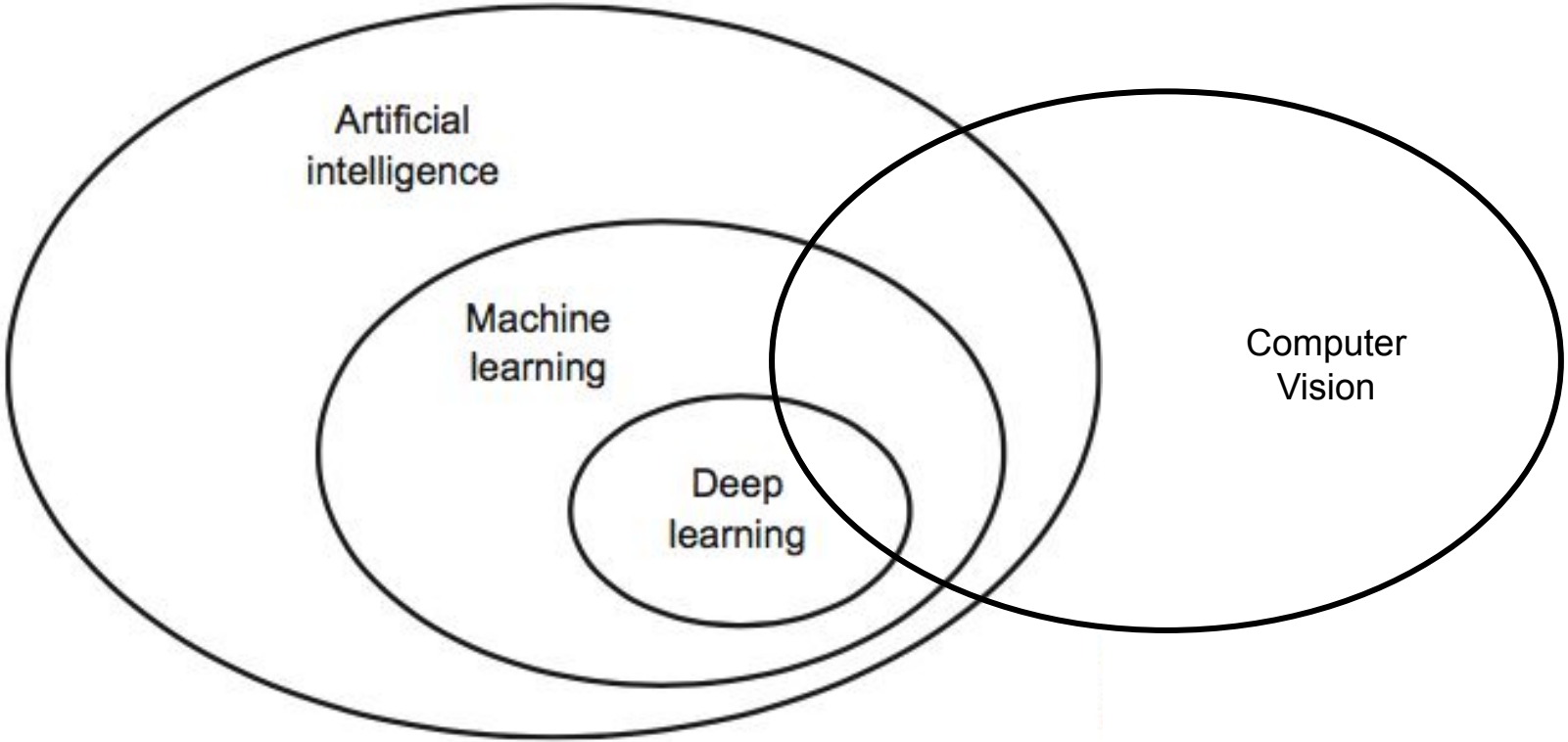
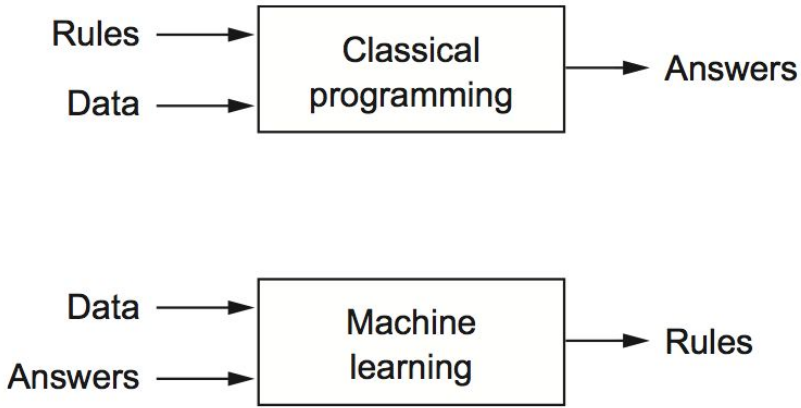


Figure Credits: DL with Python by Francois Chollet

So what is Deep Learning?



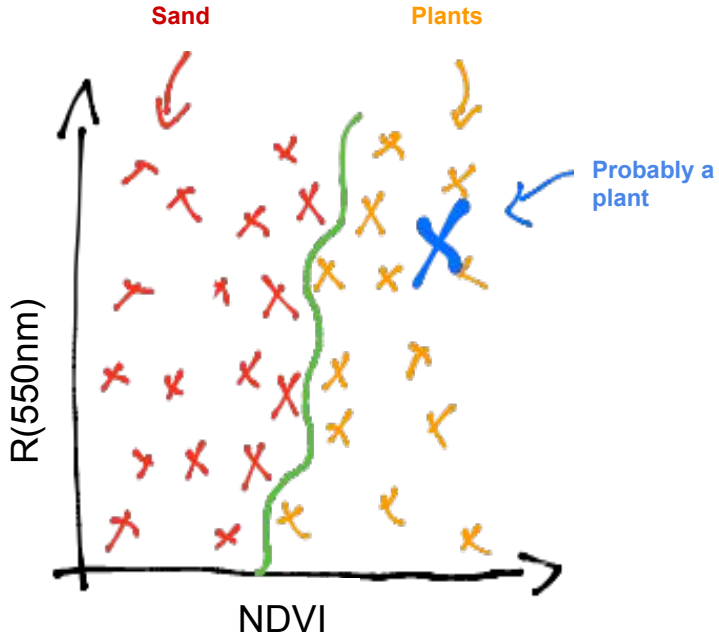
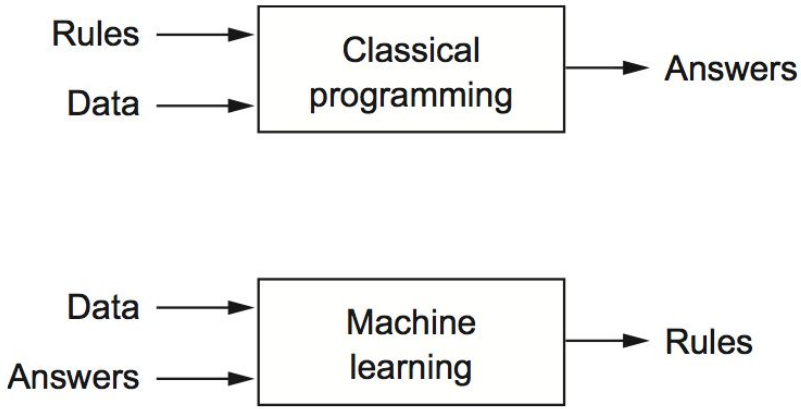


Figure Credits: DL with Python by Francois Chollet

So what is Deep Learning?

Rule: NDVI > 0.5 is a plant



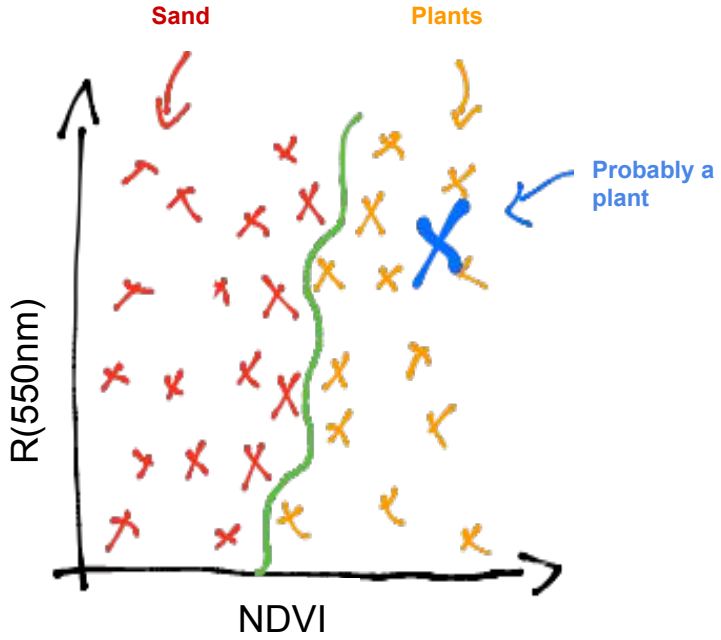
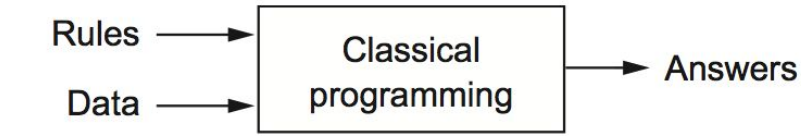
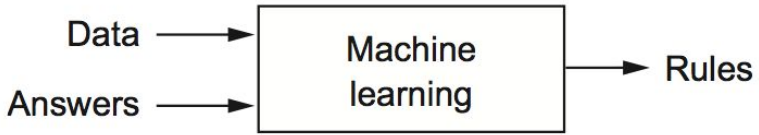


Figure Credits: DL with Python by Francois Chollet

So what is Deep Learning?

Rule: NDVI > 0.5 is a plant





Instead: give it NDVI and pixels you know are plants

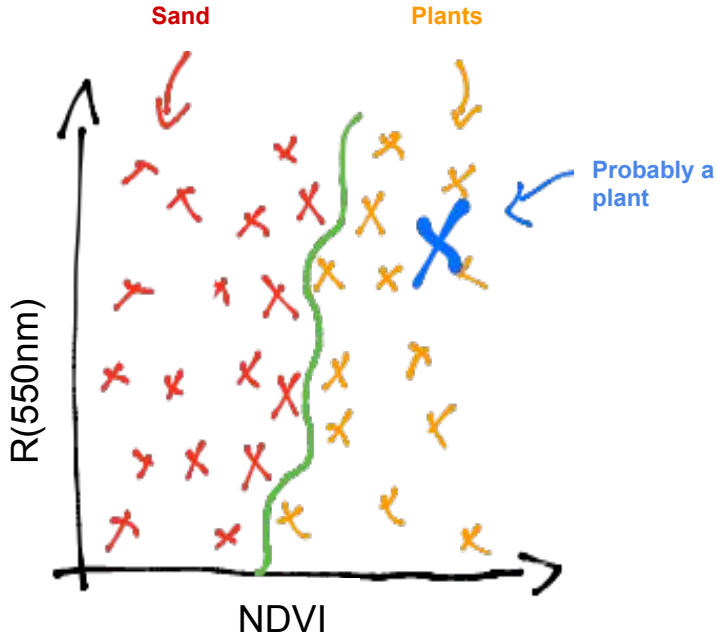
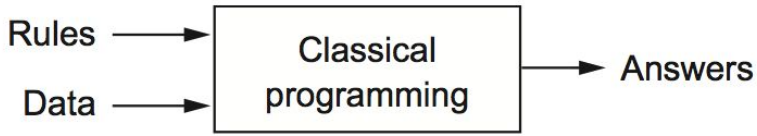
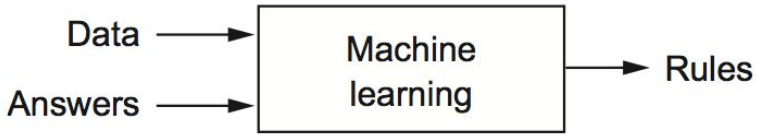


Figure Credits: DL with Python by Francois Chollet

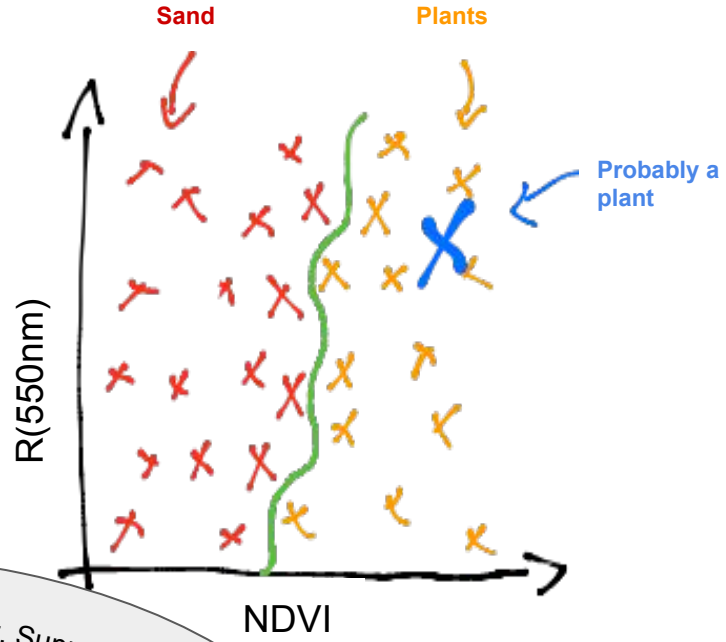
So what is Deep Learning?

Rule: NDVI > 0.5 is a plant





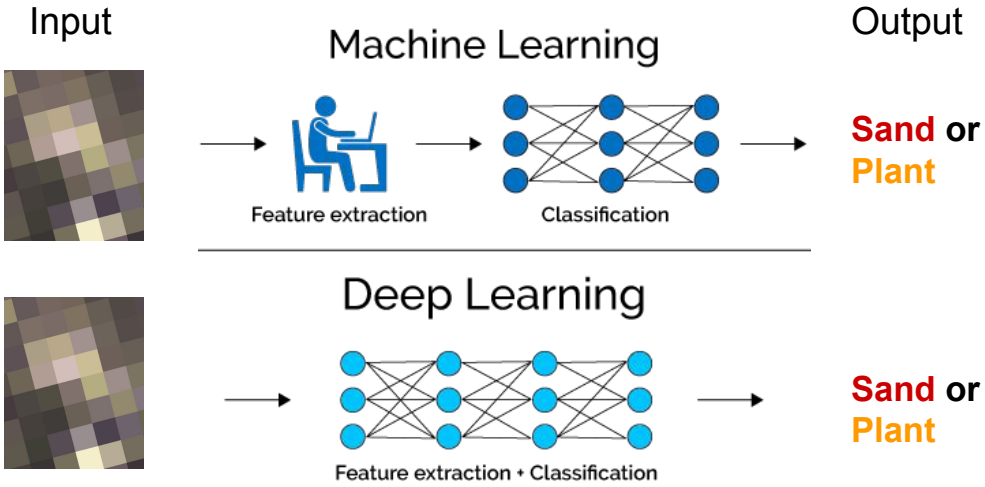
Instead: give it NDVI and labeled pixels you know are plants



Random Forest, Support Vector Machine, Naive Bayes, Max Likelihood, K-Nearest Neighbor, AdaBoost, Spectral Angle Mapper, etc

Figure Credits: DL with Python by Francois Chollet

So what is Deep Learning?



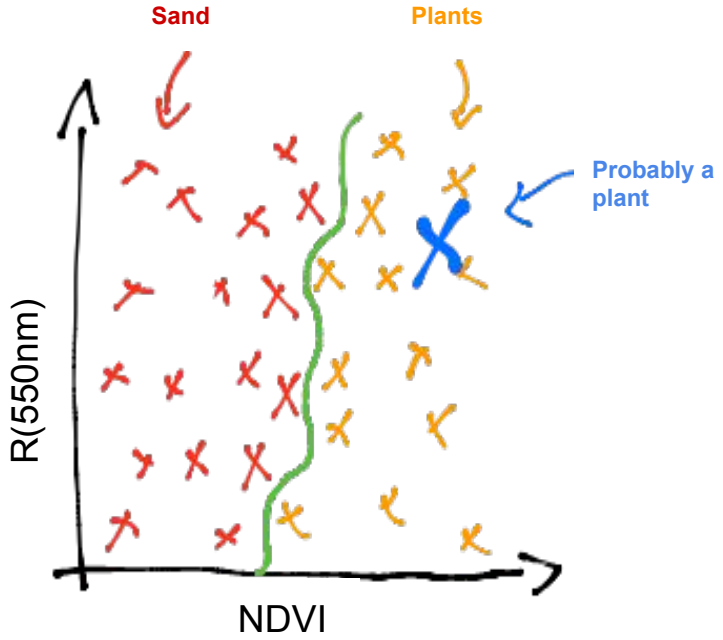
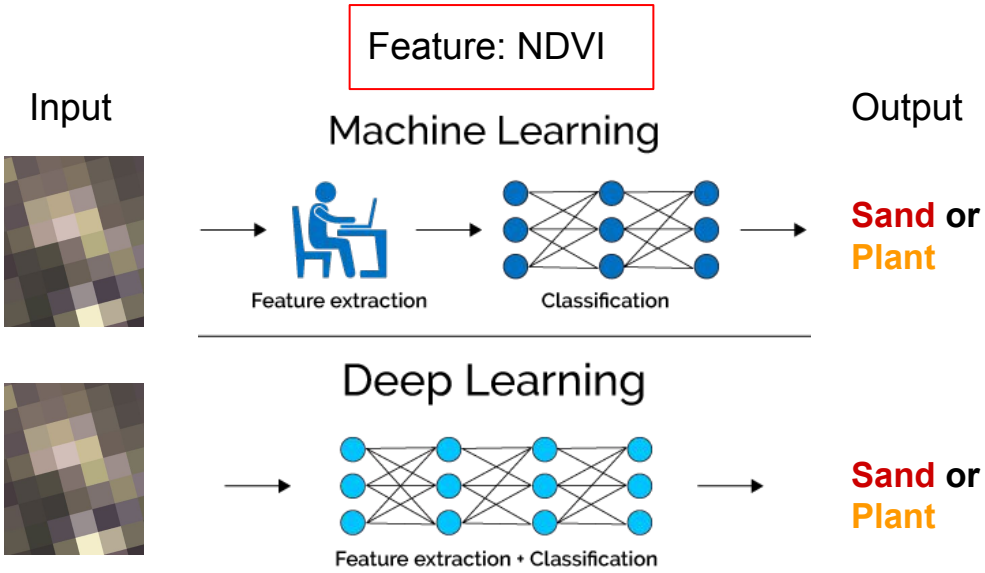


Figure Credits: Modified from DL with Python by Francois Chollet

So what is Deep Learning?



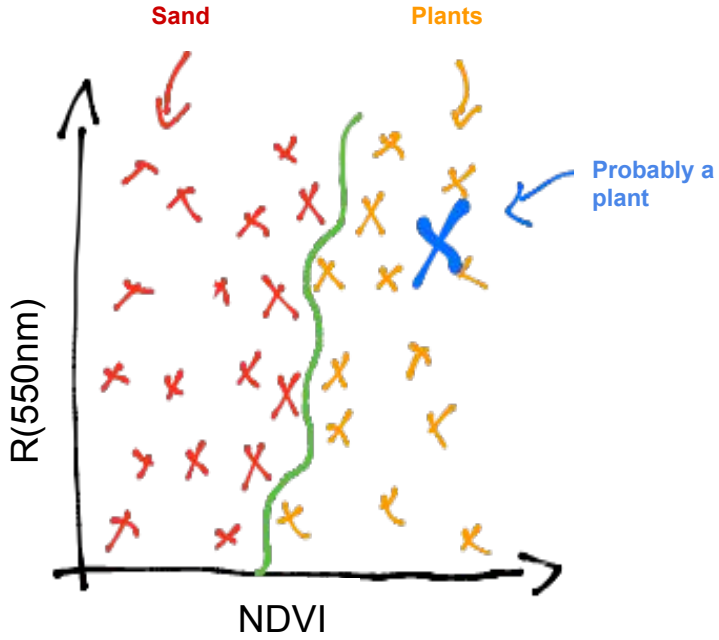


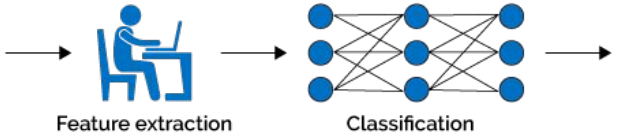
Figure Credits: Modified from DL with Python by Francois Chollet

So what is Deep Learning?

Input

Feature: NDVI

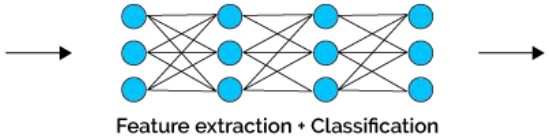
Machine Learning



Output

Sand or **Plant**

Deep Learning



Sand or **Plant**

Instead: give it the raw spectra and let it discover important combinations itself

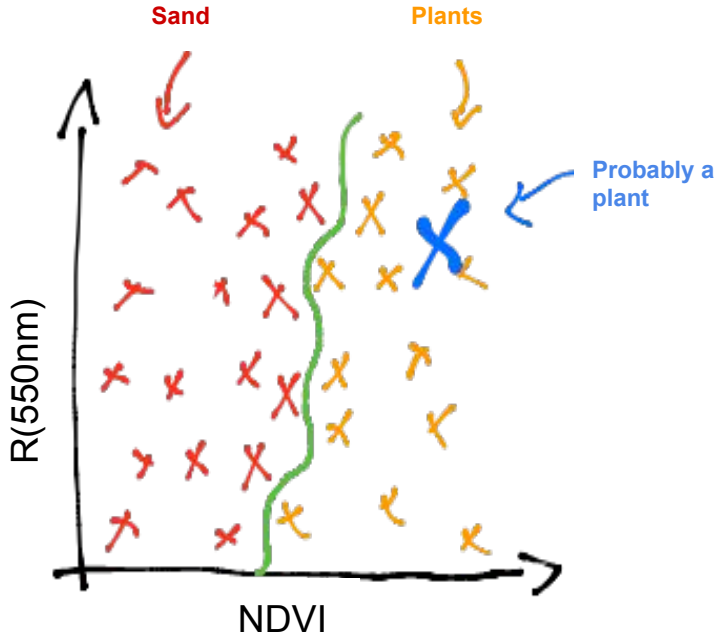


Figure Credits: Modified from DL with Python by Francois Chollet

So what is Deep Learning?

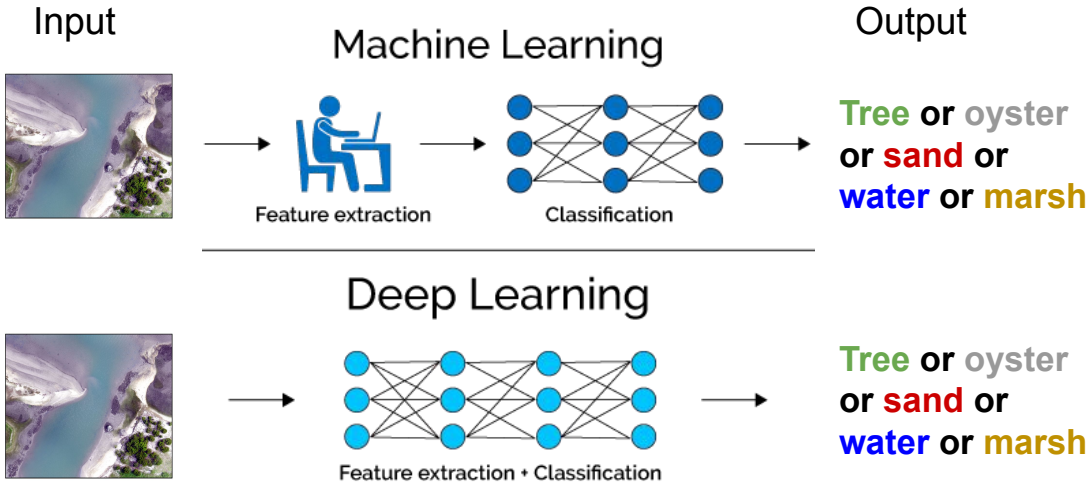


Figure Credits: Modified from DL with Python by Francois Chollet

So what is Deep Learning?

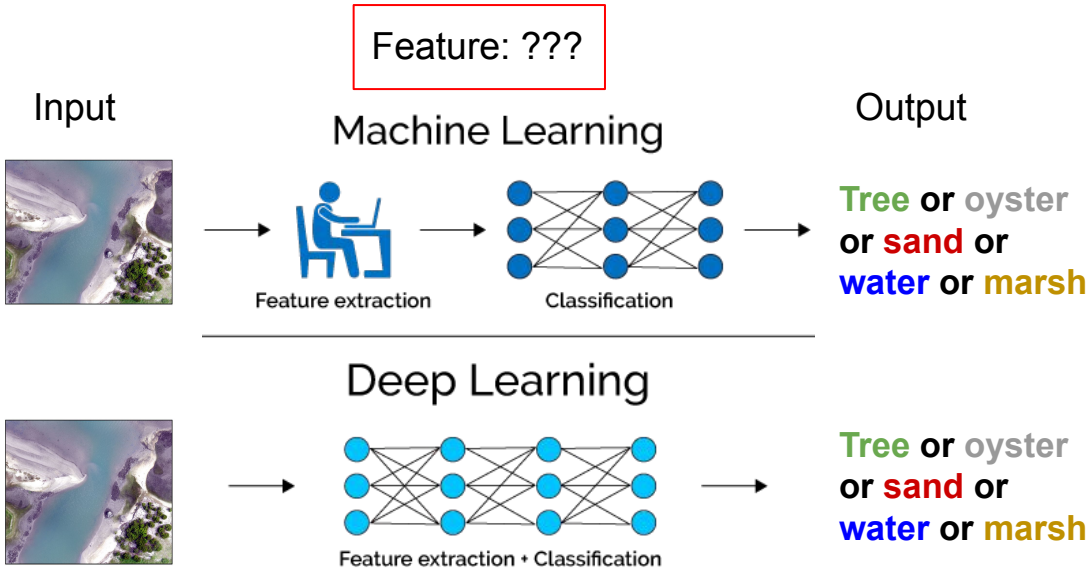
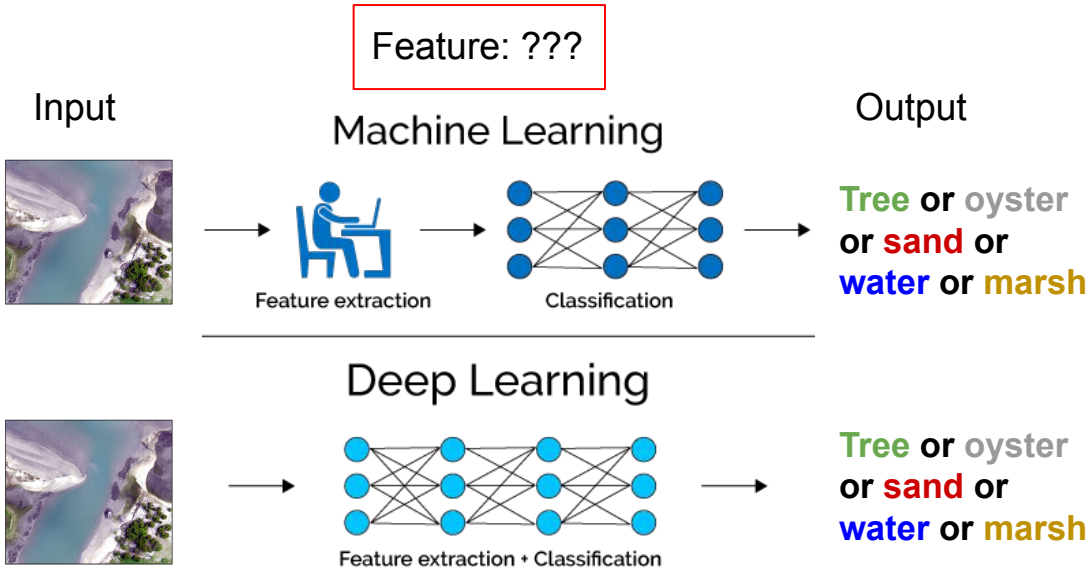


Figure Credits: Modified from DL with Python by Francois Chollet

So what is Deep Learning?



Instead: give it the multispectral imagery and let it discover important features itself

Figure Credits: Modified from DL with Python by Francois Chollet

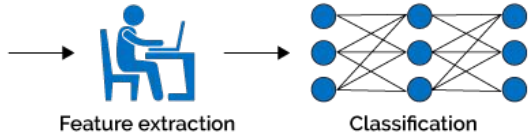
So what is Deep Learning?

Feature: ???

Input

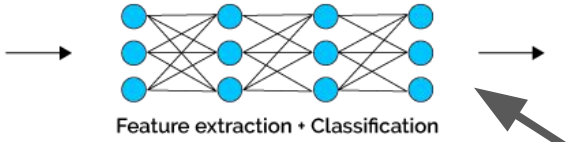
Output

Machine Learning



Tree or oyster or sand or water or marsh

Deep Learning



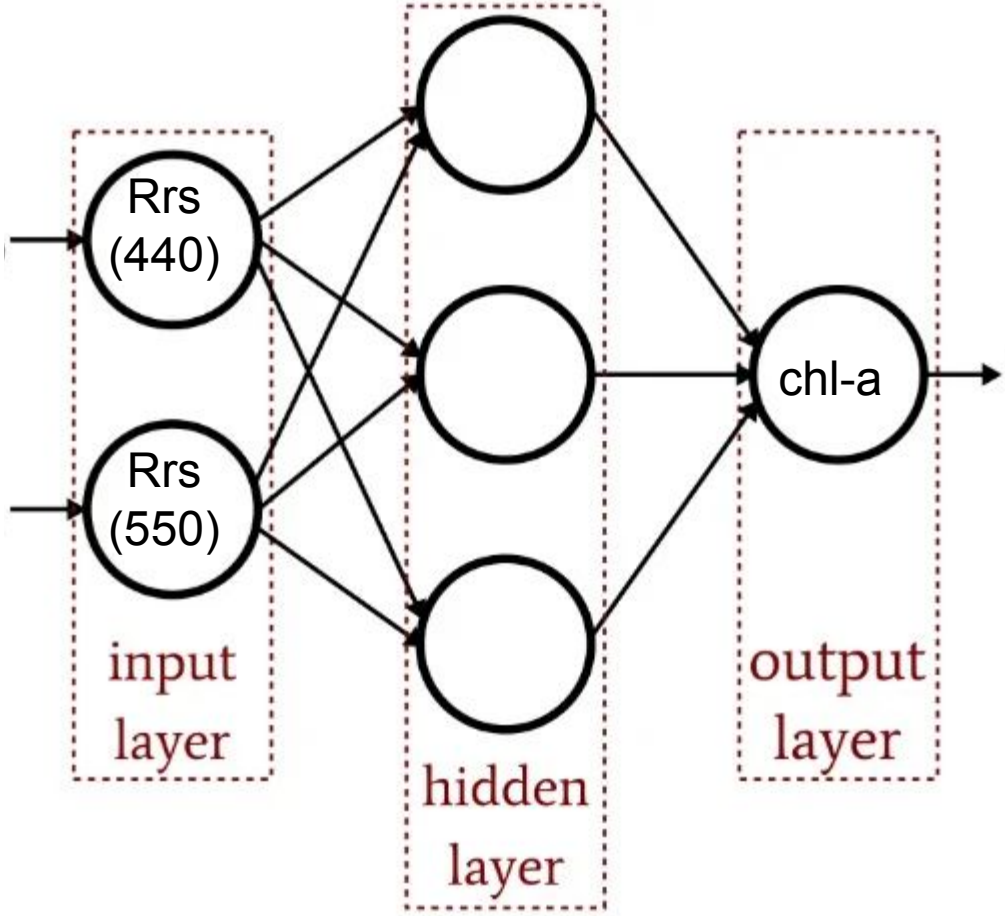
Tree or oyster or sand or wat

Instead: give it the multispectral imagery
let it discover important features itself

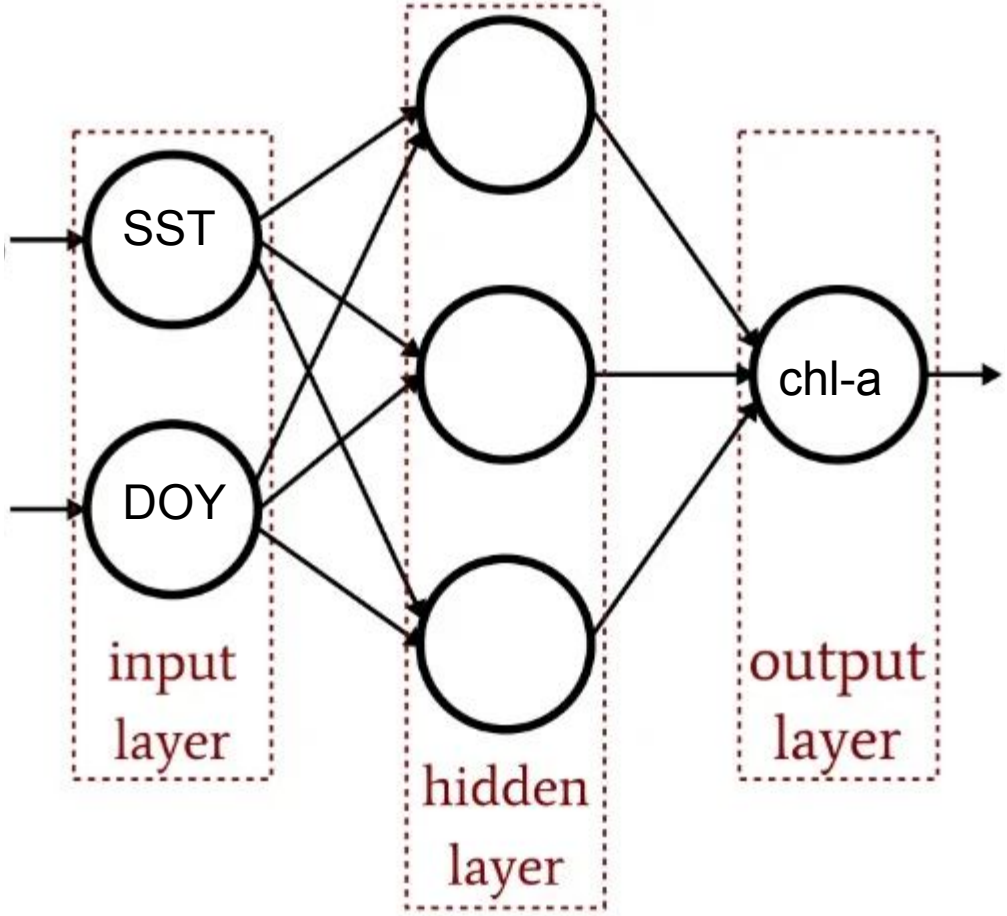
Neural networks

Images: Modified from DL with Python by Francois Chollet

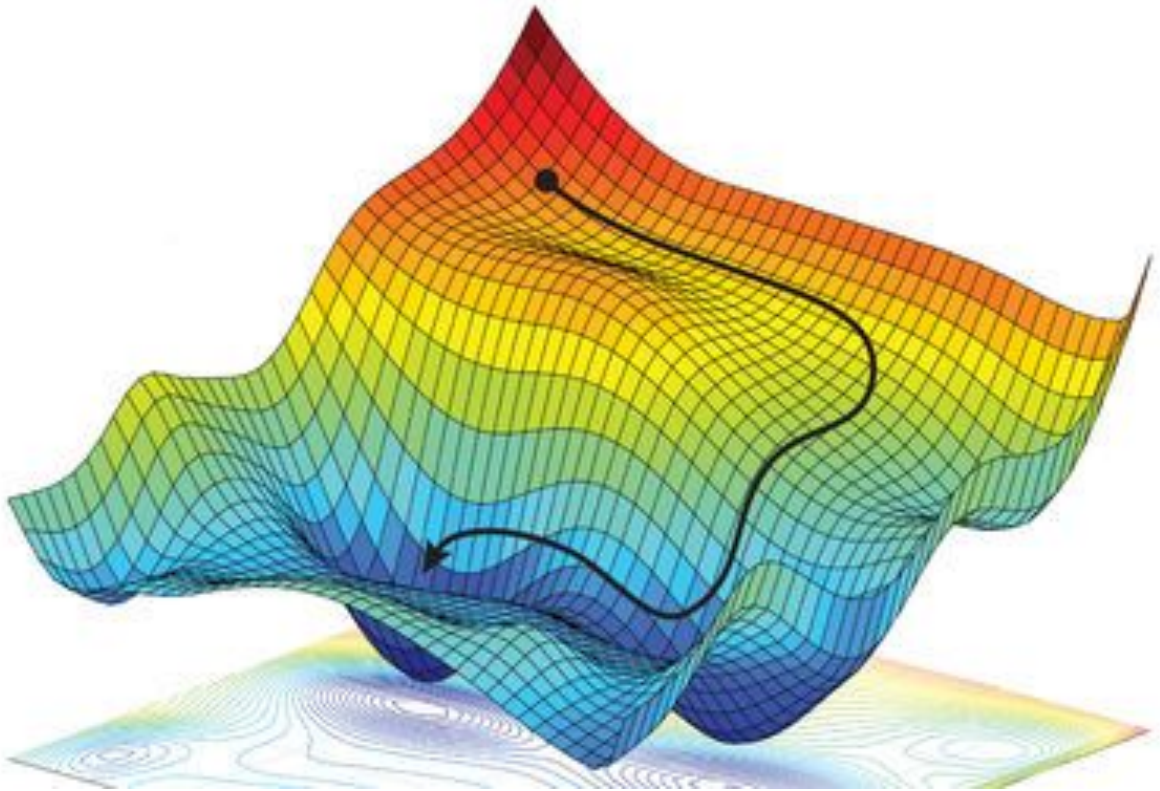
The Simplest Neural Network



The Simplest Neural Network

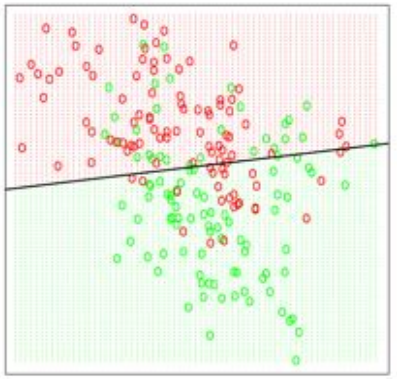


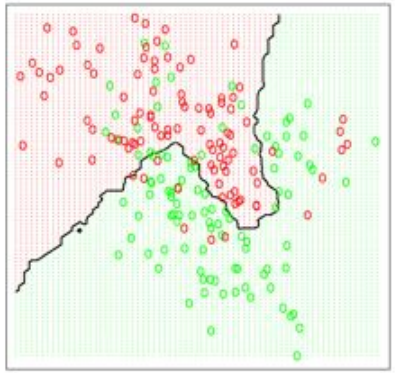
So what is Deep Learning?



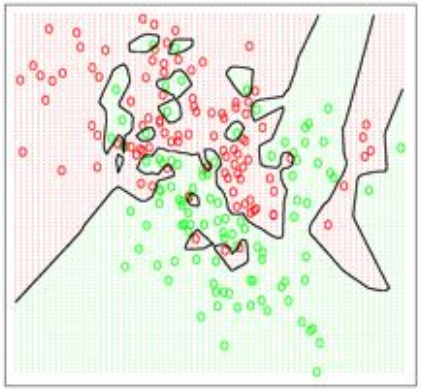
So what is Deep Learning?

Learning means finding a combination of model parameters that minimizes a loss function for a given set of training data samples and their corresponding targets.

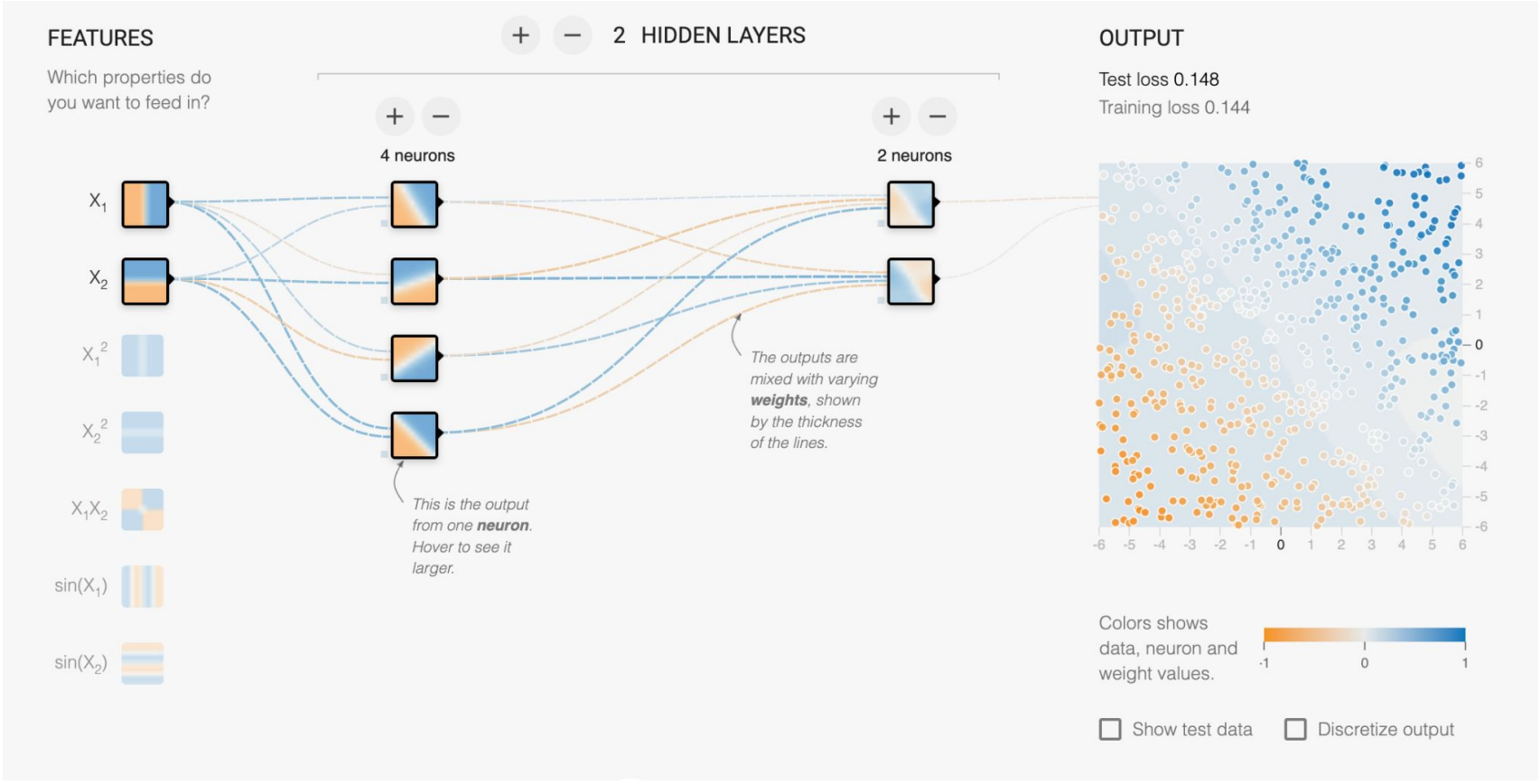




Overfitting?



Play with it: http://playground.tensorflow.org/



So what is Deep Learning?

Convolutional neural networks...

- learn to recognize high-level structure in images by building hierarchical representations
- extract features via spatial convolutions with filters
- learn filters via iterative minimization of a risk function
- **have shown capabilities beyond human performance for image analysis**

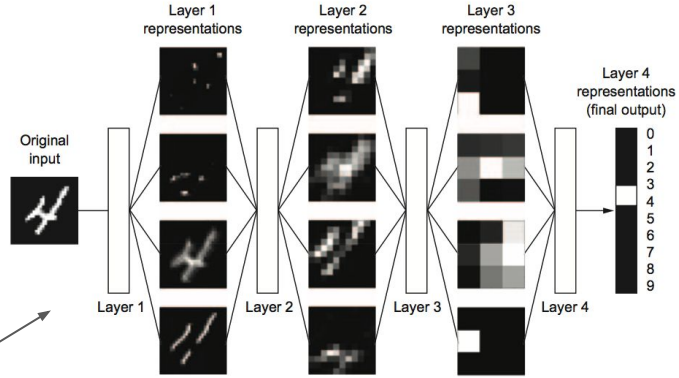


Figure Credits: DL with Python by Francois Chollet

So what is Deep Learning?

Convolutional neural networks...

- learn to recognize high-level structure in images by building hierarchical representations
- extract features via spatial convolutions with filters
- learn filters via iterative minimization of a risk function
- **have shown capabilities beyond human performance for image analysis**

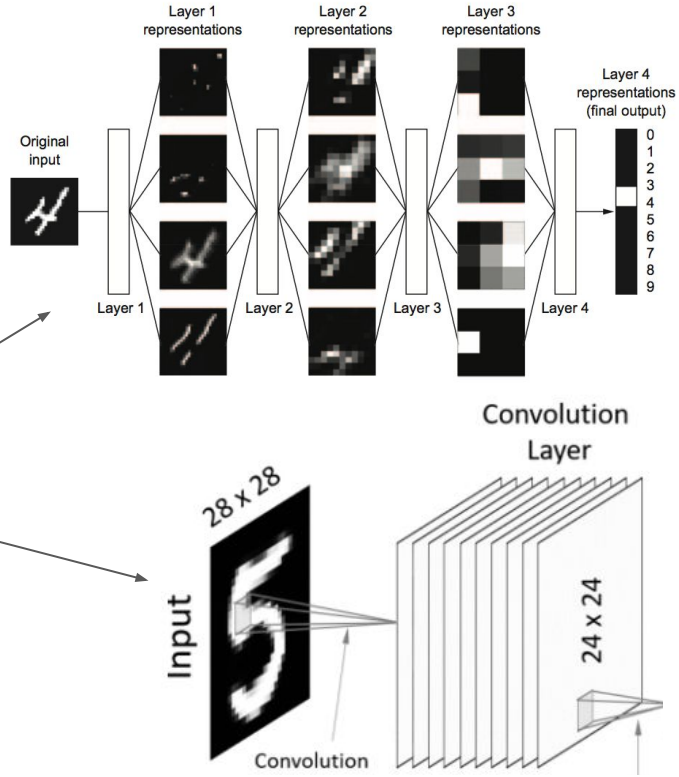
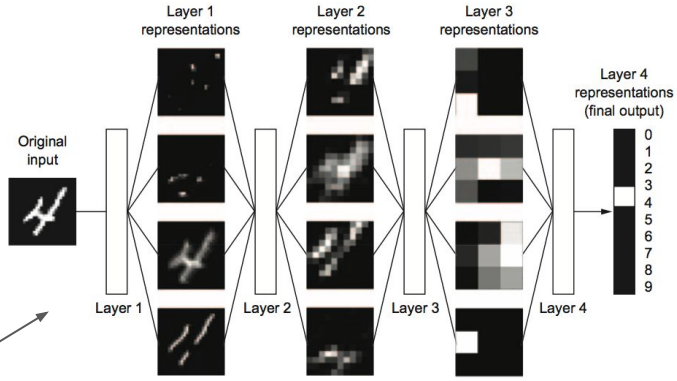


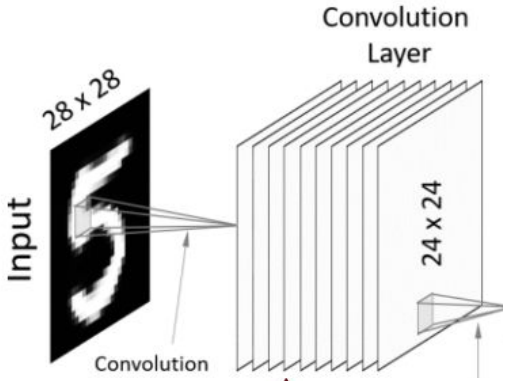
Figure Credits: DL with Python by Francois Chollet

So what is Deep Learning?

Convolutional neural networks...

- learn to recognize high-level structure in images by building hierarchical representations
- extract features via spatial convolutions with filters
- learn filters via iterative minimization of a risk function
- **have shown capabilities beyond human performance for image analysis**





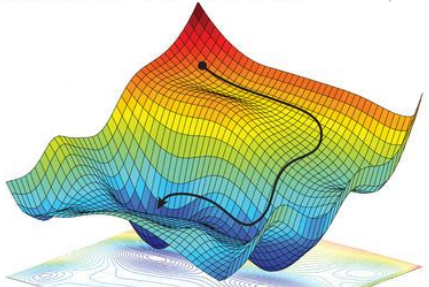


Figure Credits: DL with Python by Francois Chollet

So what is Deep Learning?

Recurrent neural networks...

- learn to extract patterns in sequences
- carries information from each time step to every future timestep
- learn weights via iterative minimization of a risk function
- **have shown capabilities beyond human performance for image analysis**

Figure Credits: DL with Python by Francois Chollet

So what is Deep Learning?

Recurrent neural networks...

- learn to extract patterns in sequences
- carries information from each time step to every future timestep
- learn weights via iterative minimization of a risk function
- **have shown capabilities beyond human performance for image analysis**

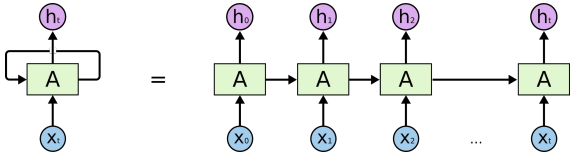
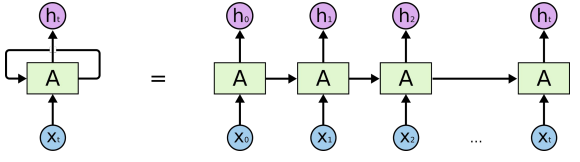


Figure Credits: DL with Python by Francois Chollet

So what is Deep Learning?

Recurrent neural networks...

- learn to extract patterns in sequences
- carries information from each time step to every future timestep
- learn weights via iterative minimization of a risk function
- **have shown capabilities beyond human performance for image analysis**



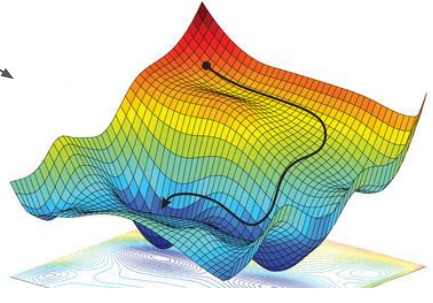


Figure Credits: DL with Python by Francois Chollet

Recap: Why Machine Learning in Remote Sensing?

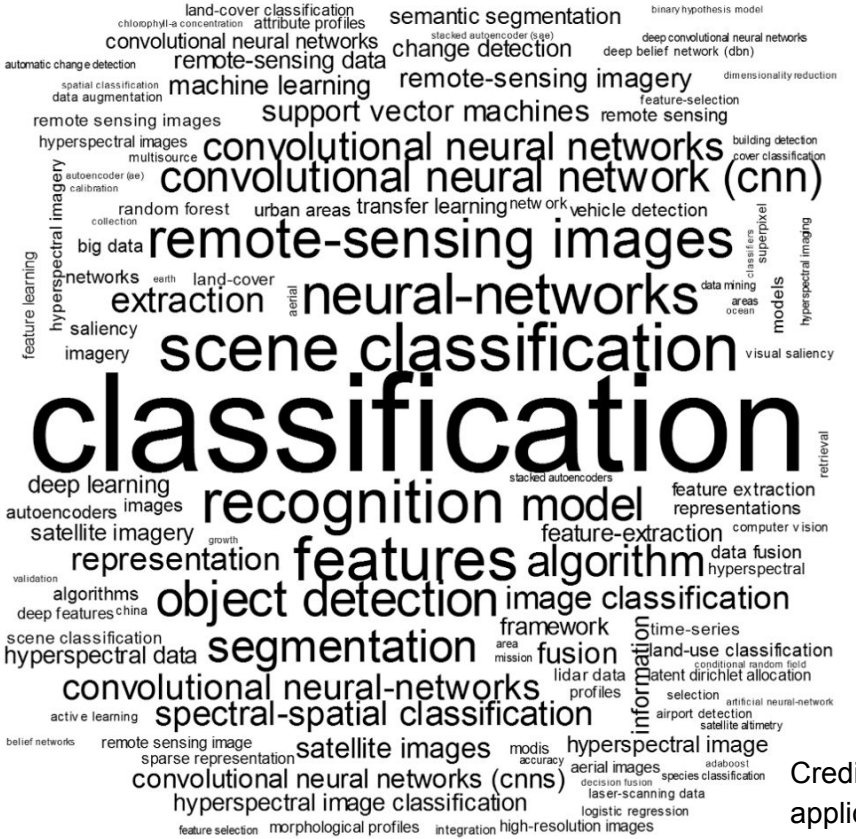
It is useful for its ability to represent complex, nonlinear relationships across spectral, temporal, and spatial dimensions.

Recap: Why Machine Learning in Remote Sensing?

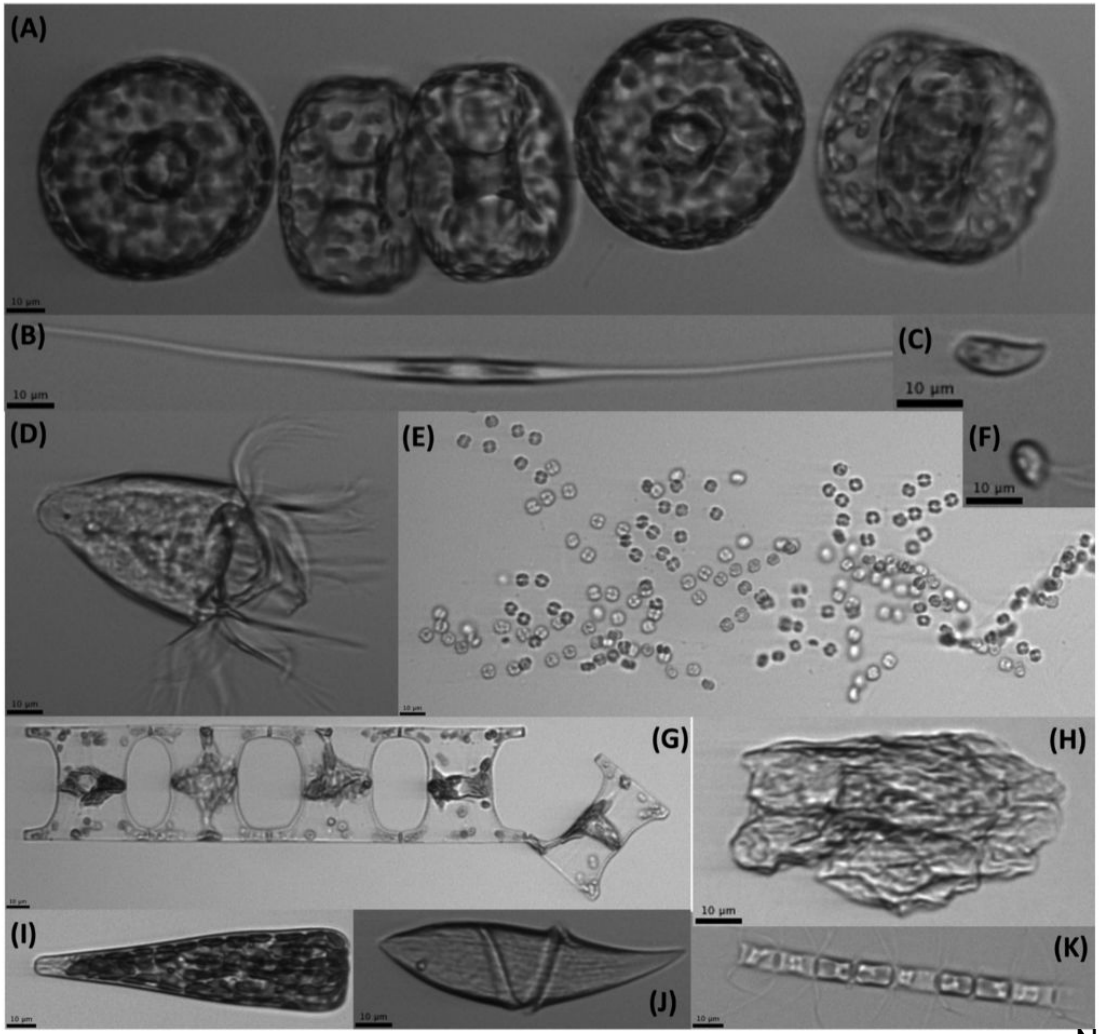
It is useful for its ability to represent complex, nonlinear relationships across spectral, temporal, and spatial dimensions.

- Spectral relationships between bands (and relations between those relations)
- Spatial relationships such as seascape context
- Temporal relationships such as phenology or bloom cycles
- Combining sensing modalities (e.g. Optical + SAR)

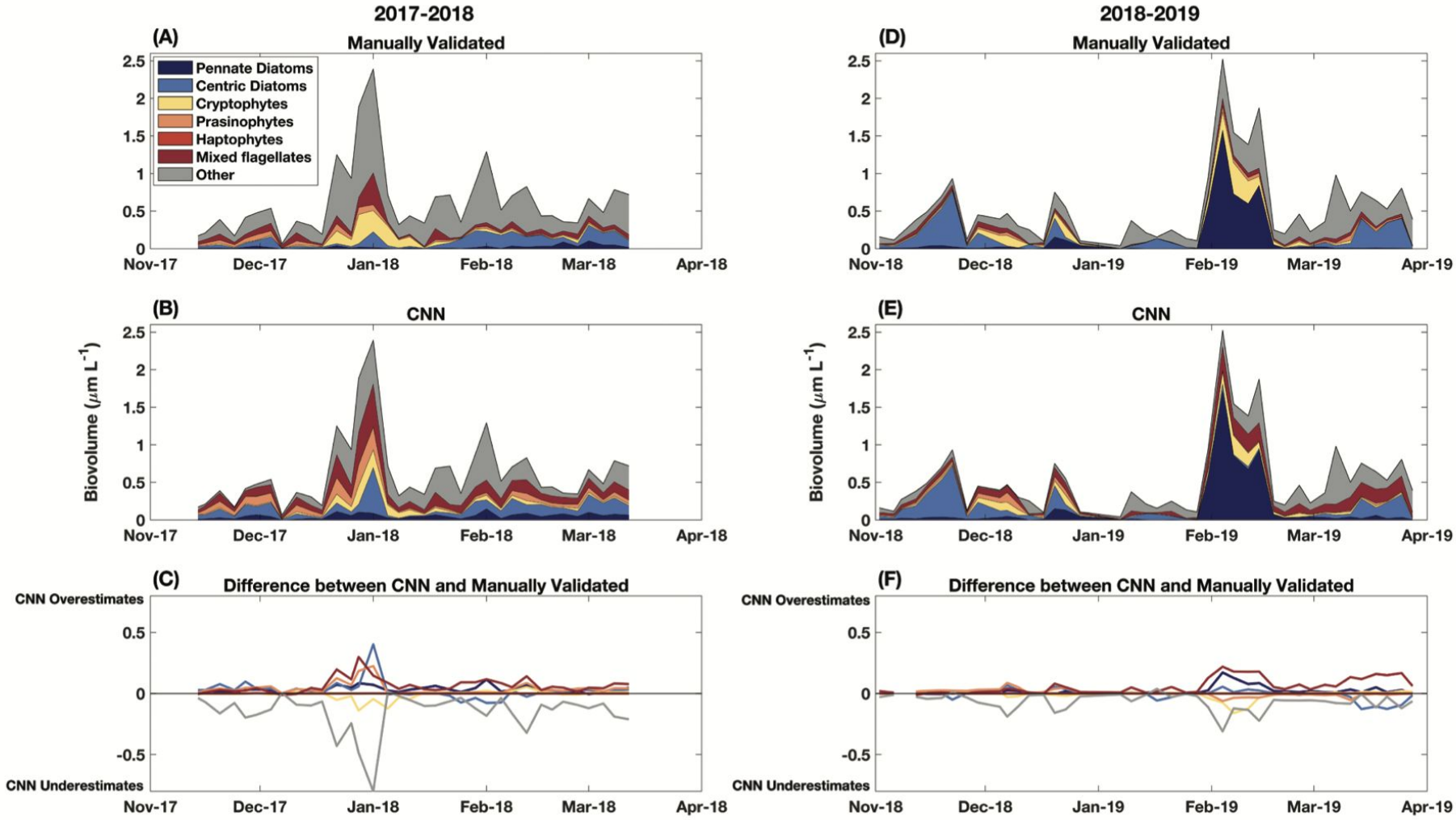
Recap: Why Machine Learning in Remote Sensing?



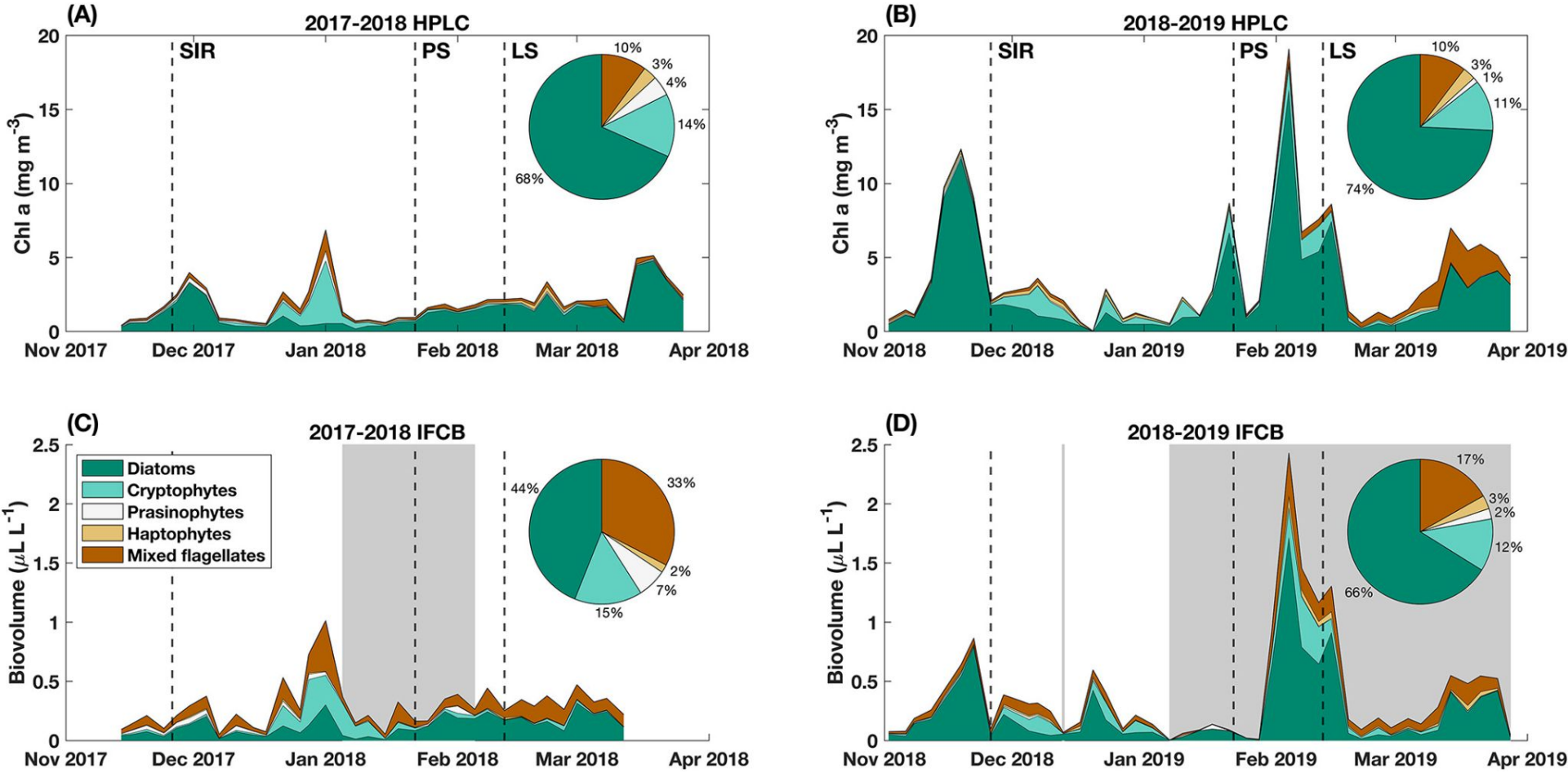
Credit: Ma et al. Deep learning in remote sensing applications: A meta-analysis and review



Nardelli, Gray, Schofield 2022

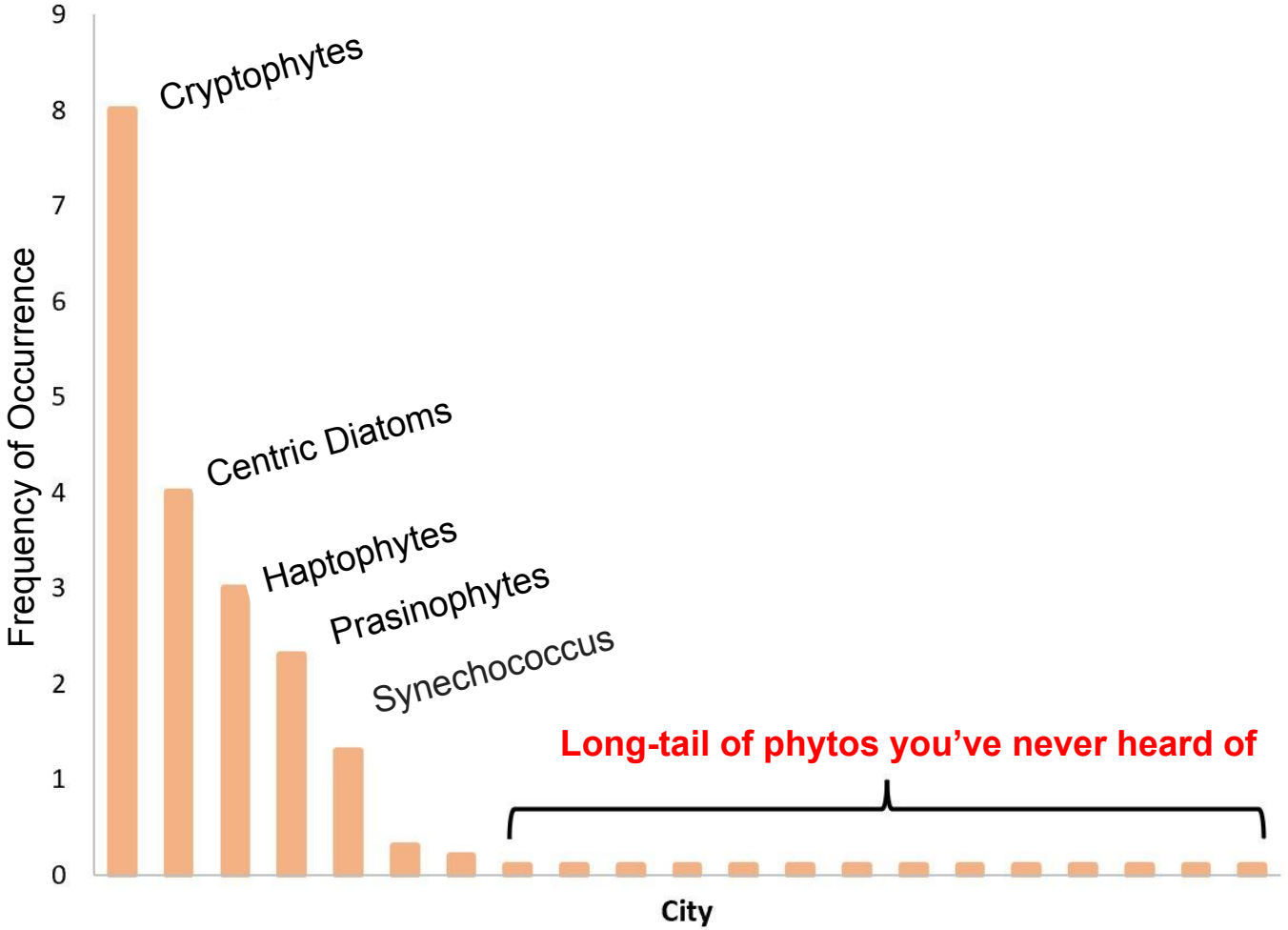


Nardelli, Gray, Schofield 2022

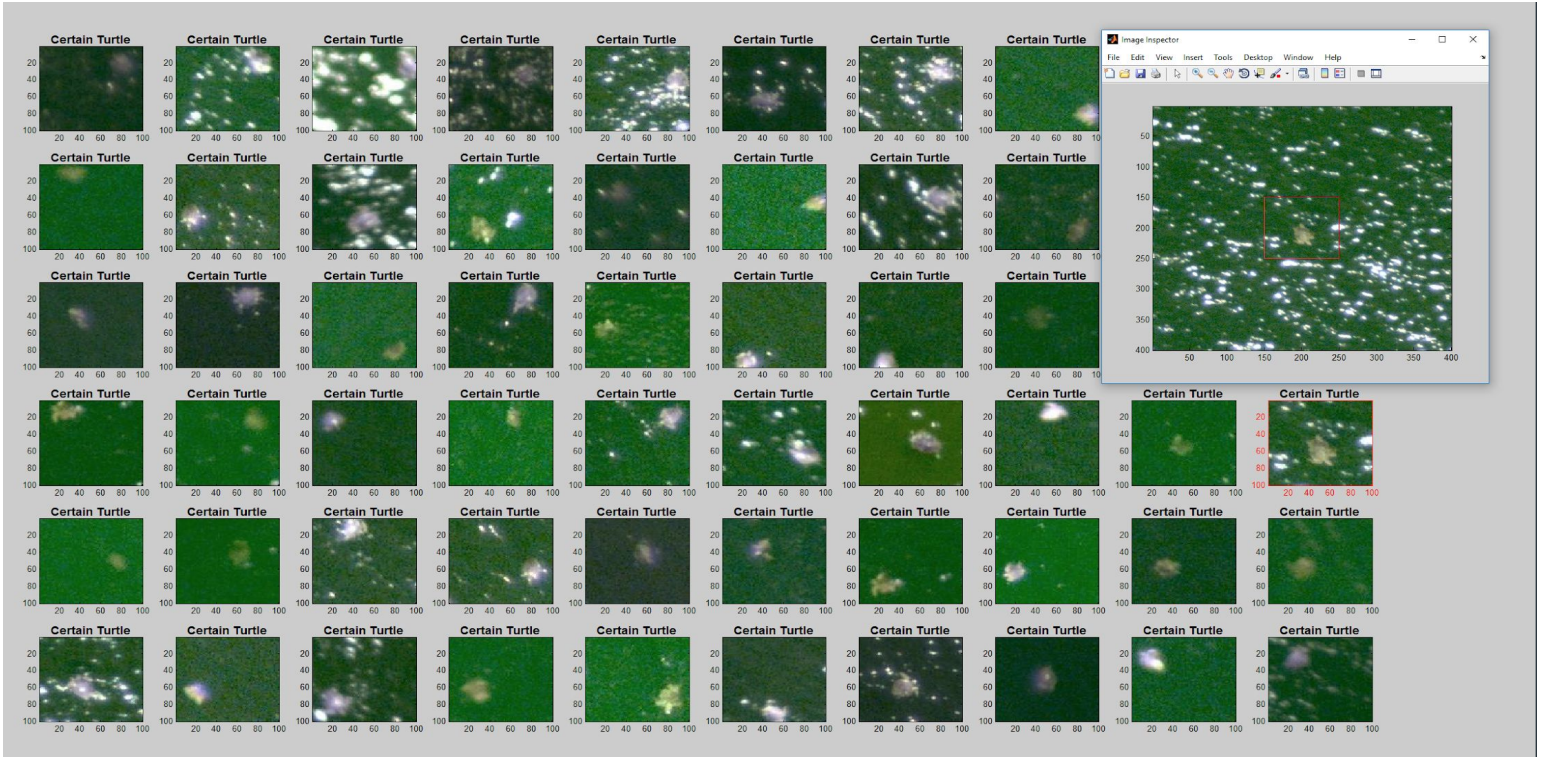


Nardelli et al 2023

Long Tailed Distribution of Phytoplankton Species



Object Detection: Patch Based

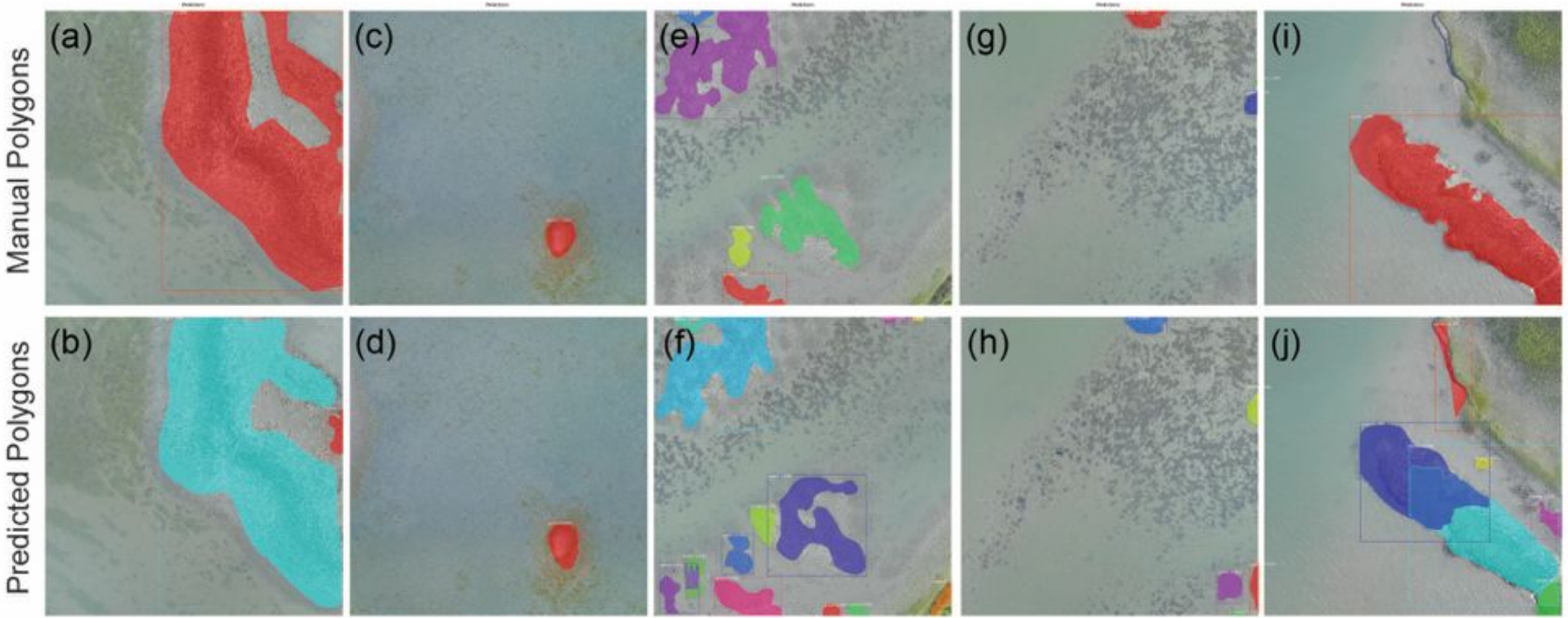


Gray et al. 2018

Object Detection: Bounding Box

Green et al. 2023

Object Detection: Semantic Segmentation



Ridge et al. 2019

Recap: Why Machine Learning in Remote Sensing?

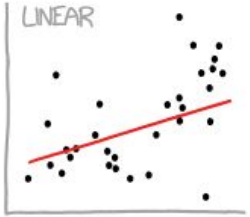
It is useful for its ability to represent **complex**, nonlinear relationships across spectral, temporal, and spatial dimensions.

But simplicity is beautiful if it is feasible.

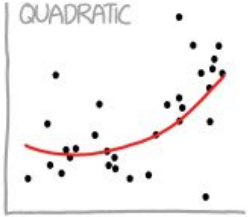
Specifically for Remote Sensing Inversions

- you can’t extrapolate beyond train/test data
- the simpler the better and comparing to a simple baseline is critical
- assessment needs to be spatially and temporally independent
- if we improve 1-2% over a polynomial regression of a band ratio (to predict chla) with a complex NN is this beneficial?
- sometimes what we care about in the ocean in the context of climate change is anomalies and things different from expectations
- think about encodings - you might be interested in the encoding of a light attenuation profile into a single Kd - or you might be interested in the encoding of a spatial pattern into some more semantically meaningful vector.

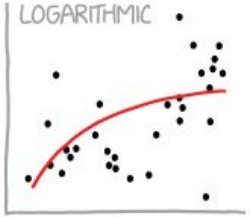
CURVE-FITTING METHODS
AND THE MESSAGES THEY SEND



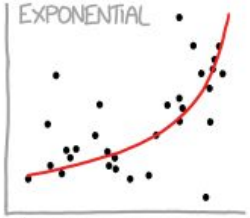
"HEY, I DID A REGRESSION."



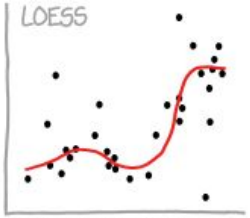
"I WANTED A CURVED LINE, SO I MADE ONE WITH MATH."



"LOOK, IT'S TAPERING OFF!"



"LOOK, IT'S GROWING UNCONTROLLABLY!"

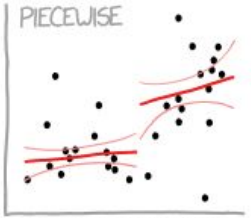


"I'M SOPHISTICATED, NOT LIKE THOSE BUMBLING POLYNOMIAL PEOPLE."

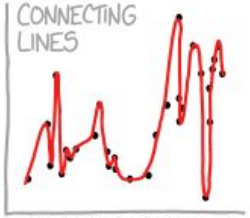
"I'M MAKING A SCATTER PLOT BUT I DON'T WANT TO."

"I NEED TO CONNECT THESE TWO LINES, BUT MY FIRST IDEA DIDN'T HAVE ENOUGH MATH."

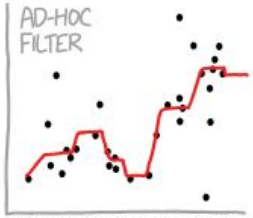
"LISTEN, SCIENCE IS HARD. BUT I'M A SERIOUS PERSON DOING MY BEST."



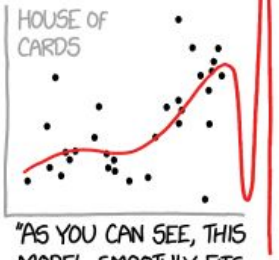
"I HAVE A THEORY, AND THIS IS THE ONLY DATA I COULD FIND."



"I CLICKED 'SMOOTH LINES' IN EXCEL."



"I HAD AN IDEA FOR HOW TO CLEAN UP THE DATA. WHAT DO YOU THINK?"

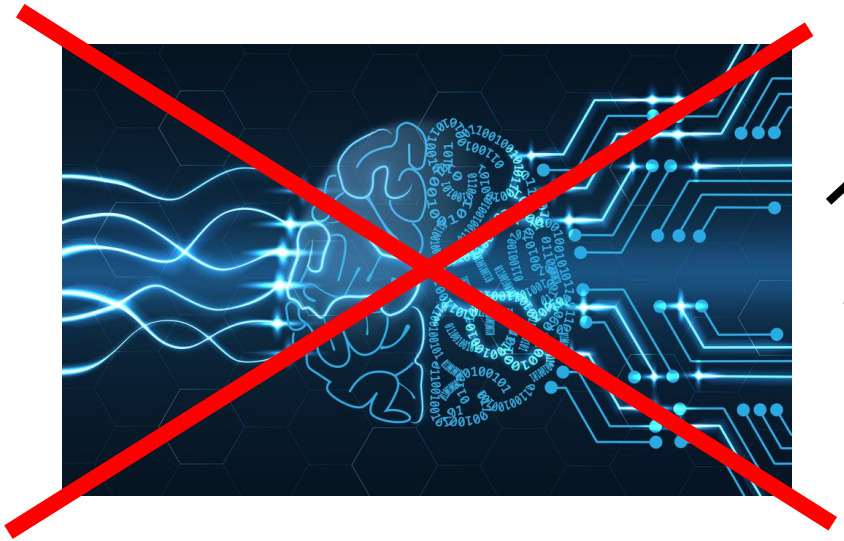


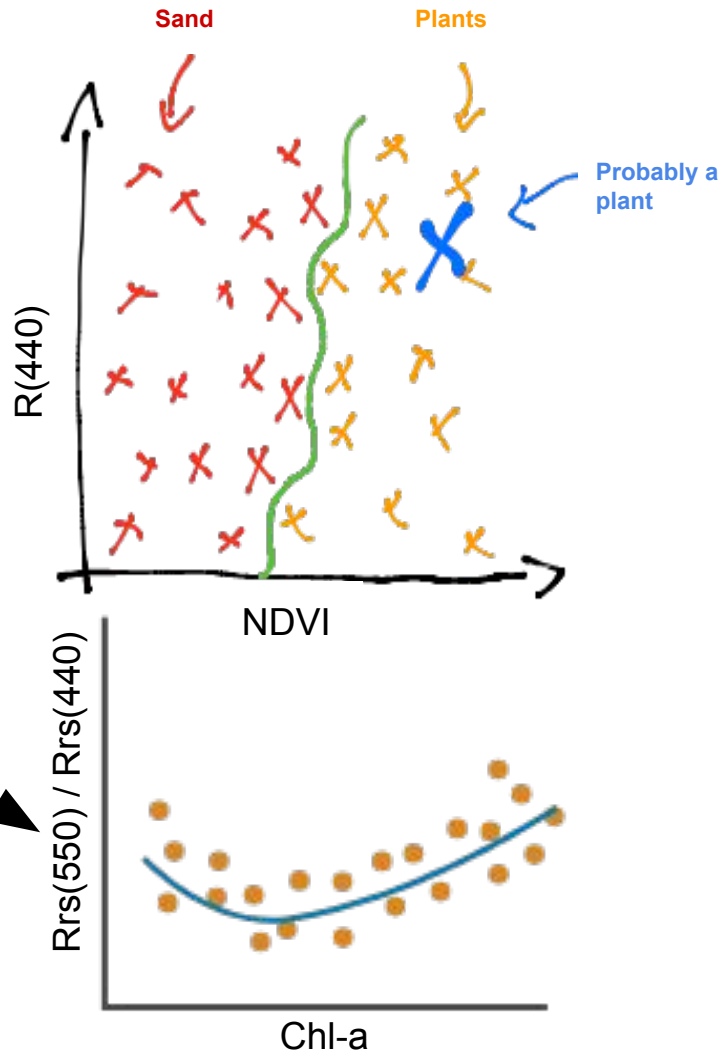
"AS YOU CAN SEE, THIS MODEL SMOOTHLY FITS THE– *WAIT NO NO DON'T EXTEND IT AAAAAA!!*"

Python and Machine Learning Resources

- Building a strong but basic python foundation (e-textbook):
  - https://www.pythonlikeyoumeanit.com/
- Become python proficient (textbook):
  - https://www.amazon.com/dp/0134034287/
- Open Geospatial Tutorials using Python for Geospatial and Remote Sensing Analysis
  - https://github.com/patrickcgray/open-geo-tutorial/
- ML Crash Course (online class from Google):
  - https://developers.google.com/machine-learning/crash-course/
- Python and Deep Learning with Keras (textbook):
  - https://www.amazon.com/Deep-Learning-Python-Francois-Chollet/dp/1617294438
- Neural Networks Video Series (highly recommend)
  - https://www.youtube.com/watch?v=aircAruvnKk
- Papers:
  - Ma et al. Deep learning in remote sensing applications: A meta-analysis and review
  - Zhang et al. Deep learning for remote sensing data: A technical tutorial on the state of the art
  - Lecun et al. Deep Learning

Questions?





Quick Exploration

https://github.com/patrickcgray/oceanoptics2021/blob/main/oci_ml_oo23.ipynb

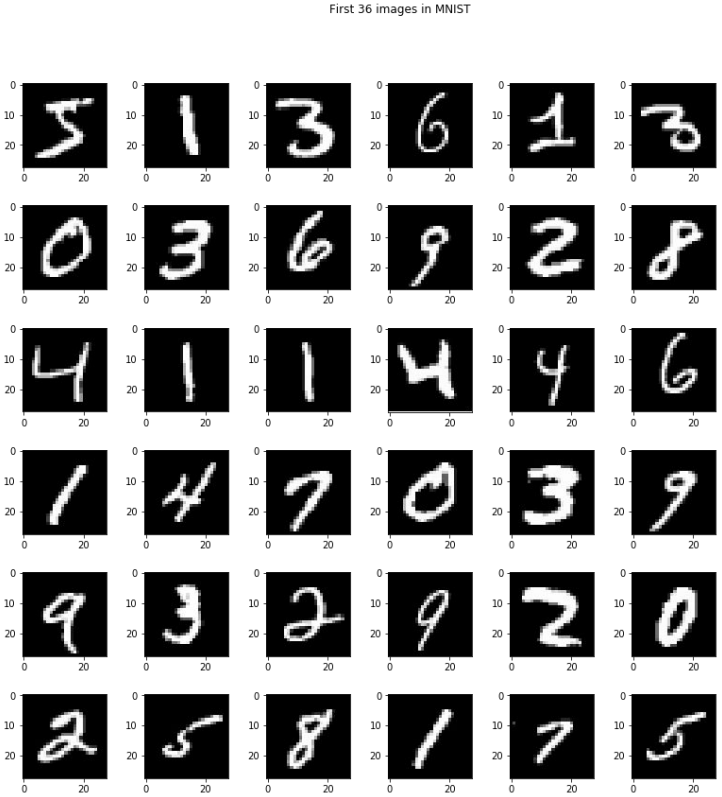
Deep Learning Technical Overview

Surpassing Human Performance

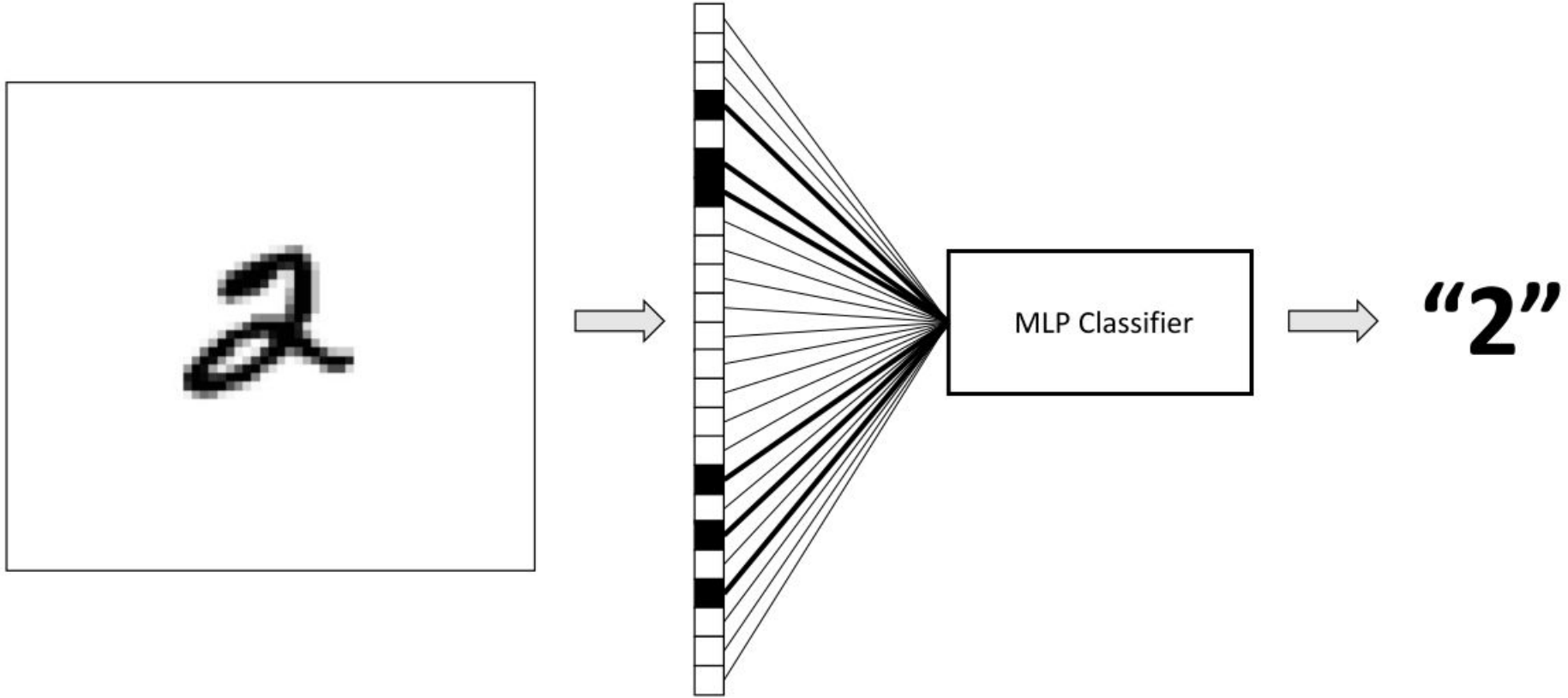
DL is now beating humans at many tasks:

- Chess
- Driving
- Language Translation
- Voice Recognition
- Predicting stock prices
- Object Detection
- **Go**

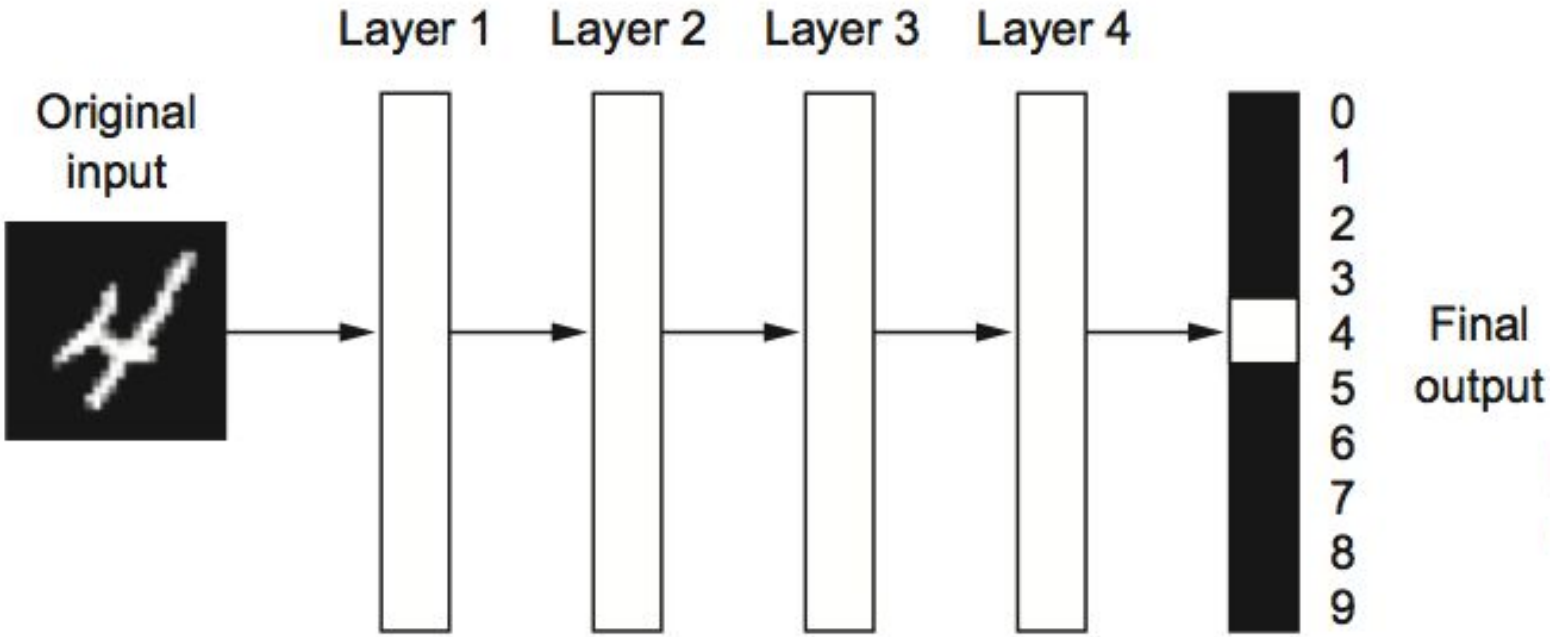
So what is Deep Learning?



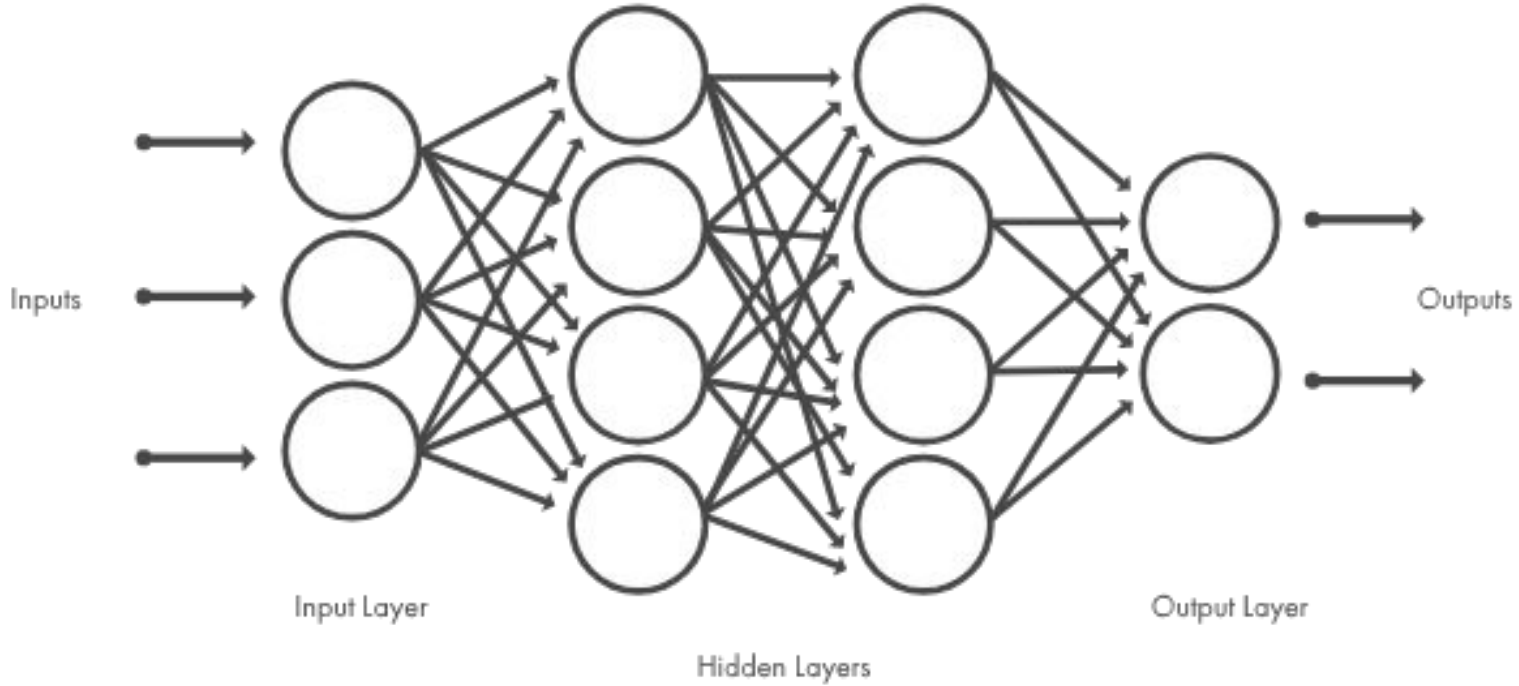
So what is Deep Learning?



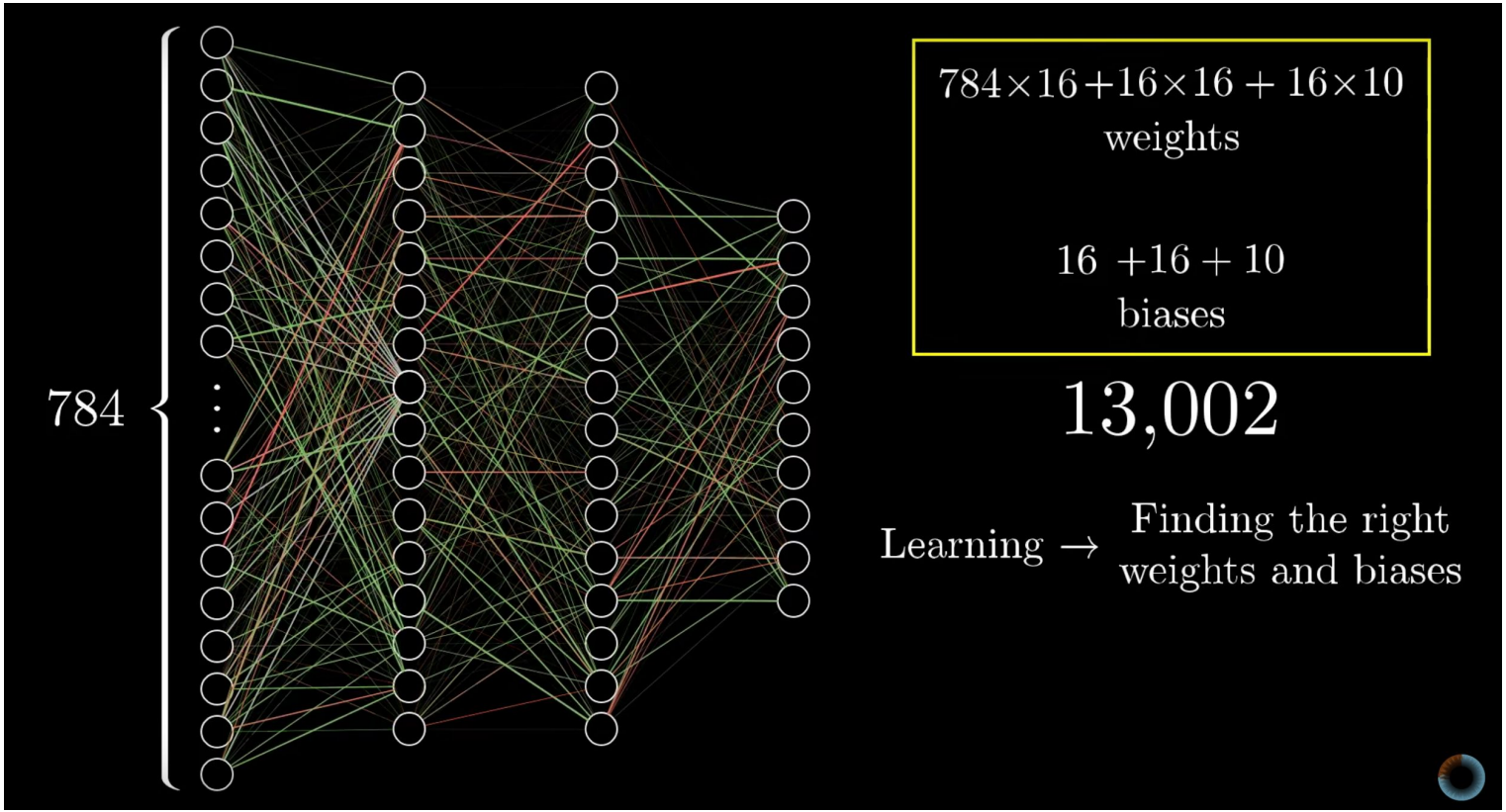
So what is Deep Learning?



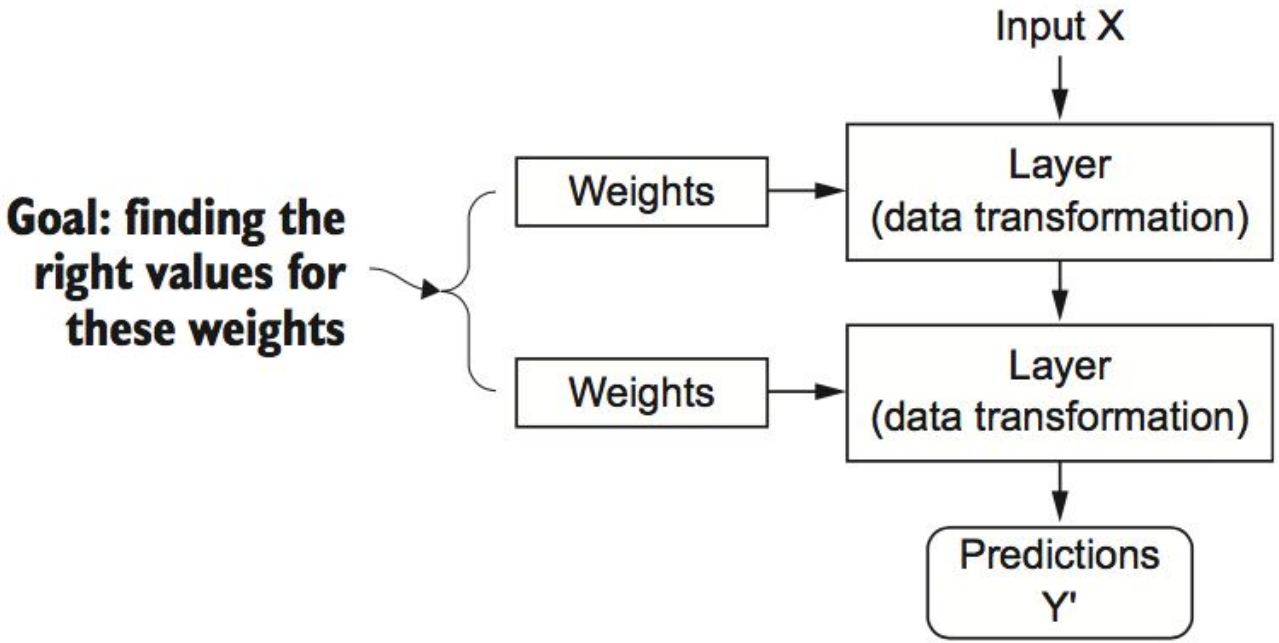
So what is Deep Learning?



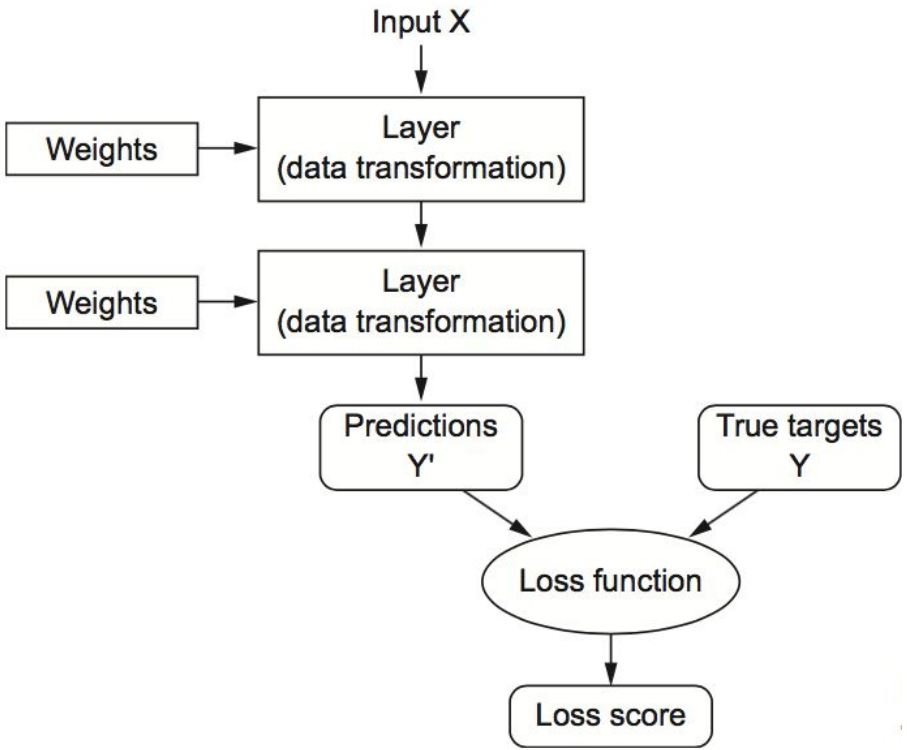
So what is Deep Learning?



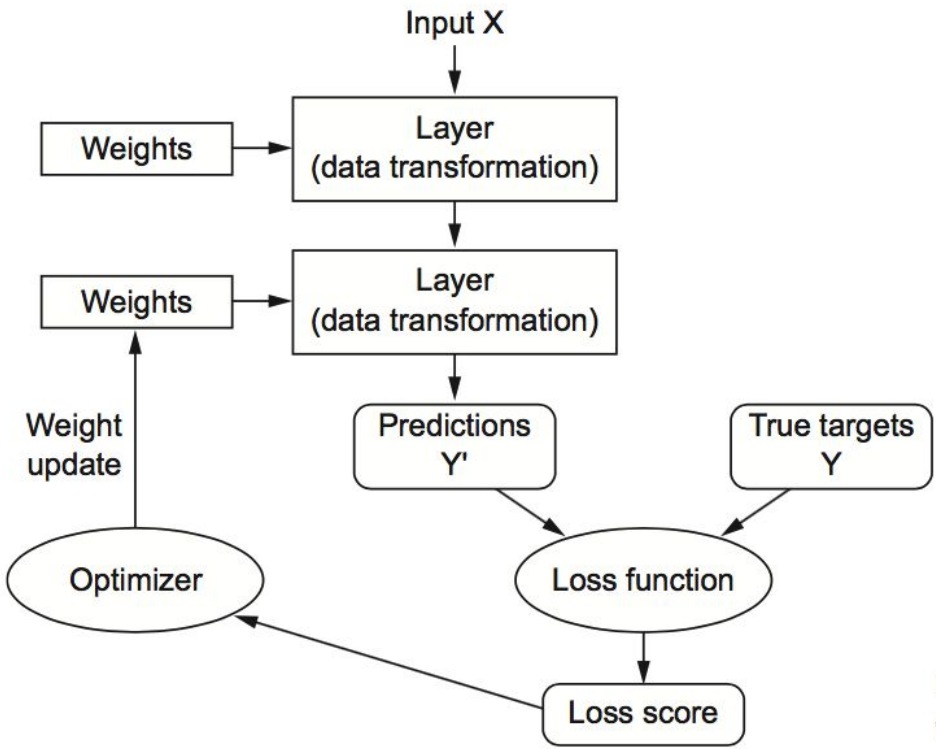
So what is Deep Learning?



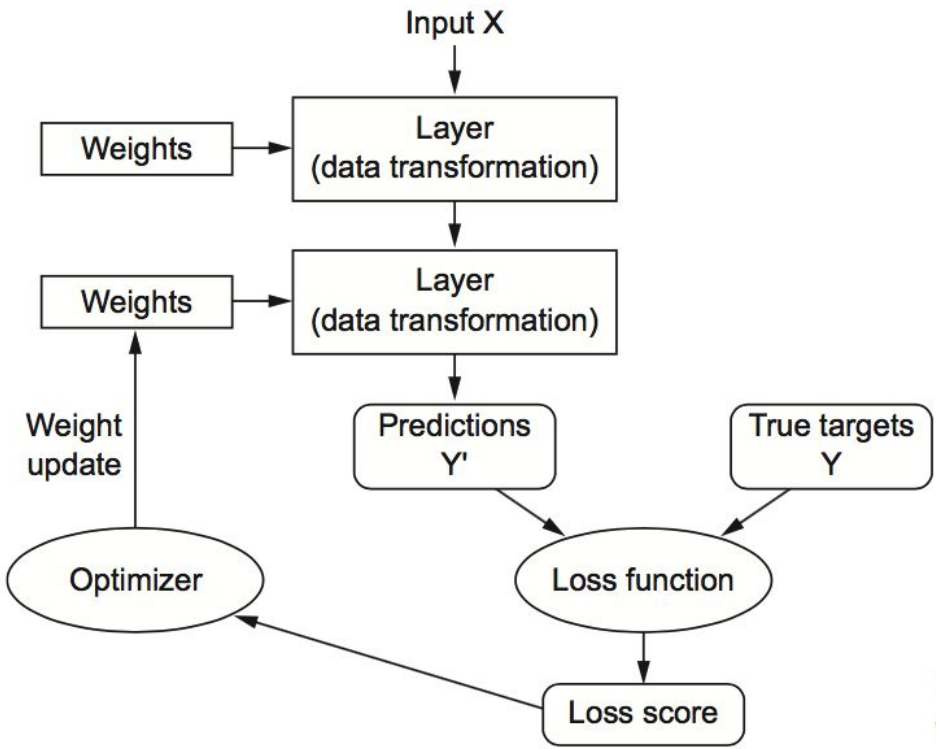
So what is Deep Learning?



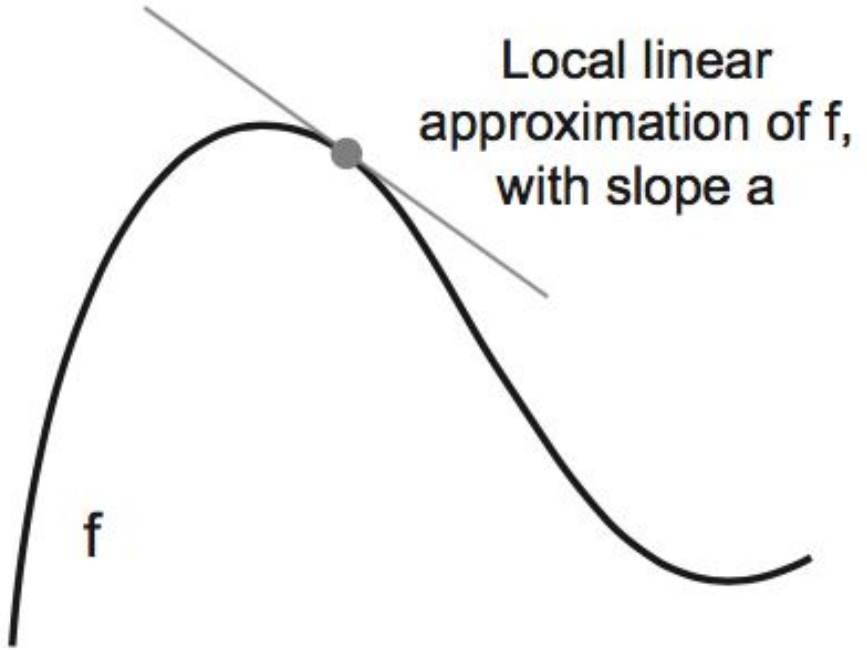
So what is Deep Learning?



So what is Deep Learning? - Questions???



So what is Deep Learning?



So what is Deep Learning?

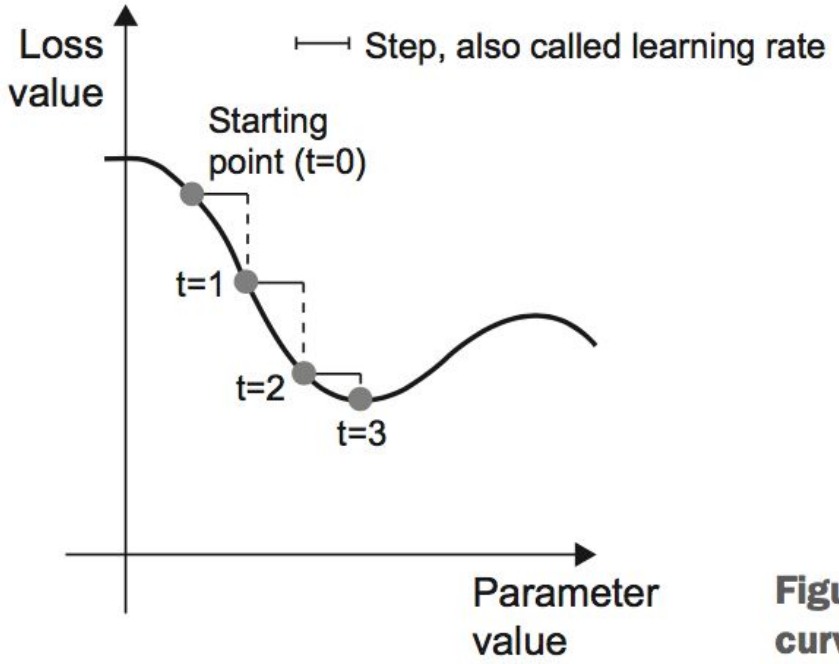


Figure 2.11 SGD down a 1D loss curve (one learnable parameter)

So what is Deep Learning?

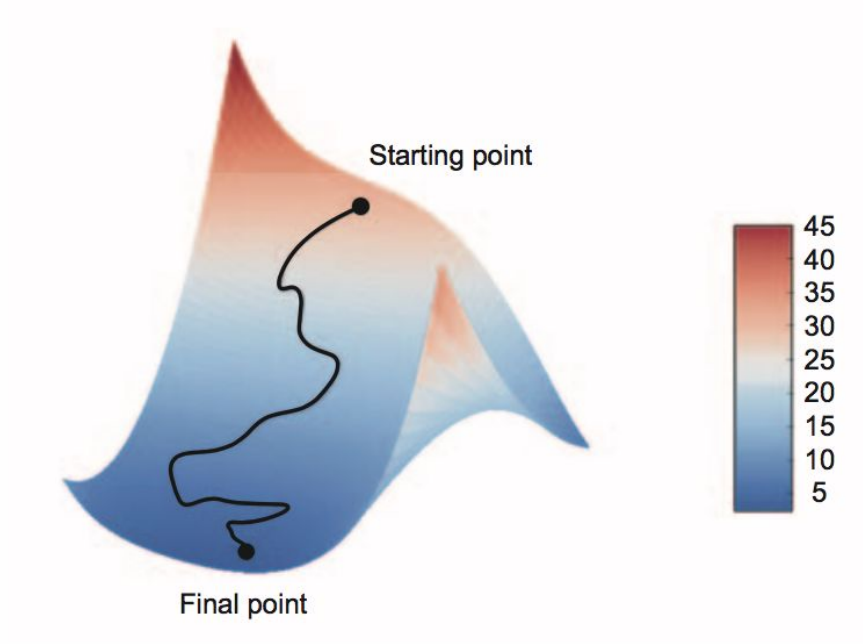


Figure 2.12 Gradient descent down a 2D loss surface (two learnable parameters)

So what is Deep Learning?

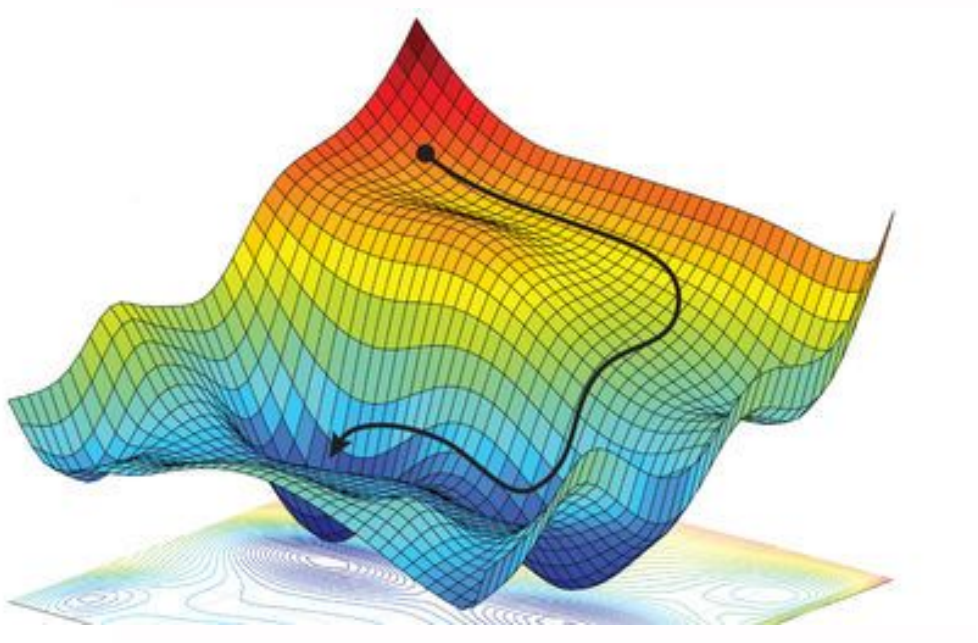
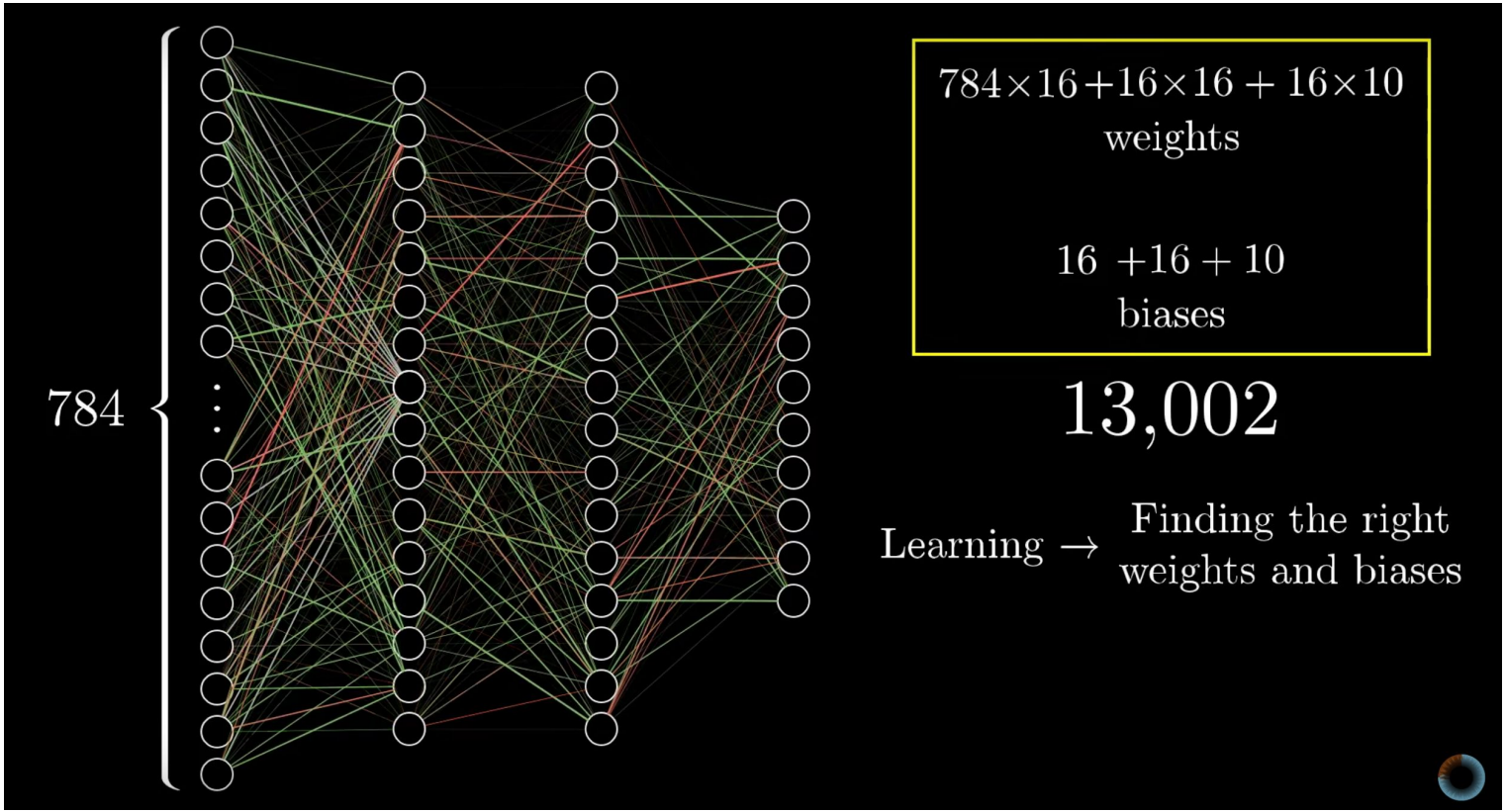


Figure 2.12 Gradient descent down a 2D loss surface (two learnable parameters)

So what is Deep Learning?

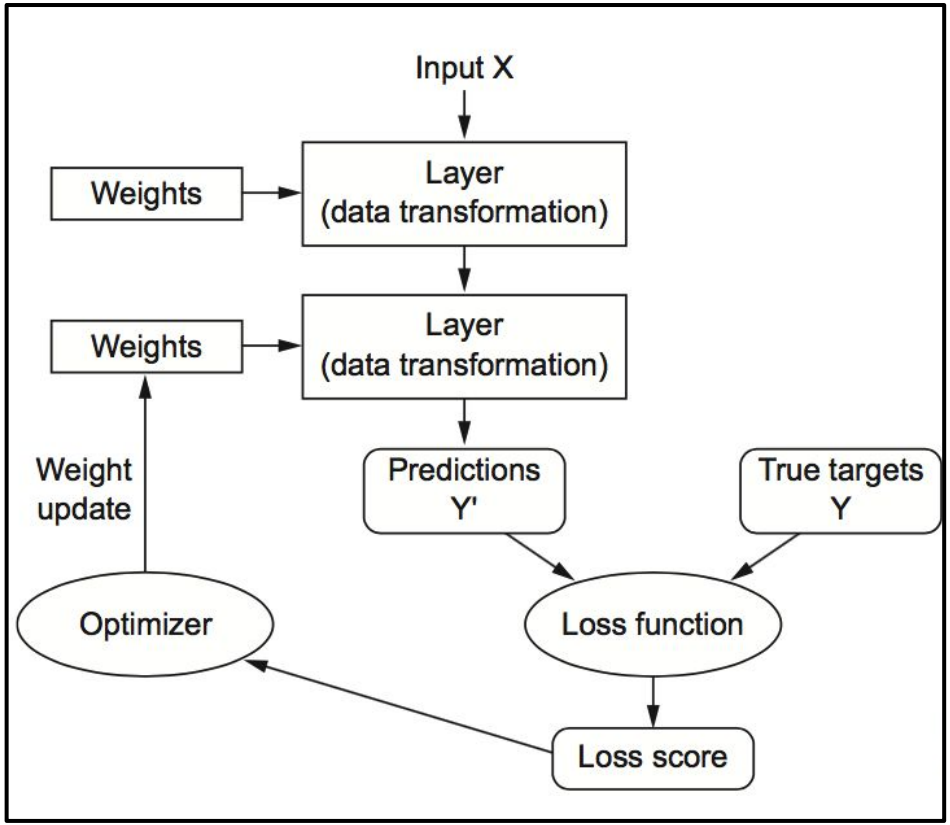


So what is Deep Learning?

Learning means finding a combination of model parameters that minimizes a loss function for a given set of training data samples and their corresponding targets.

Loss is the quantity you'll attempt to minimize during training → so it should represent a measure of success for the task you're trying to solve.

The optimizer specifies the exact way in which the gradient of the loss will be used to update parameters in your model.

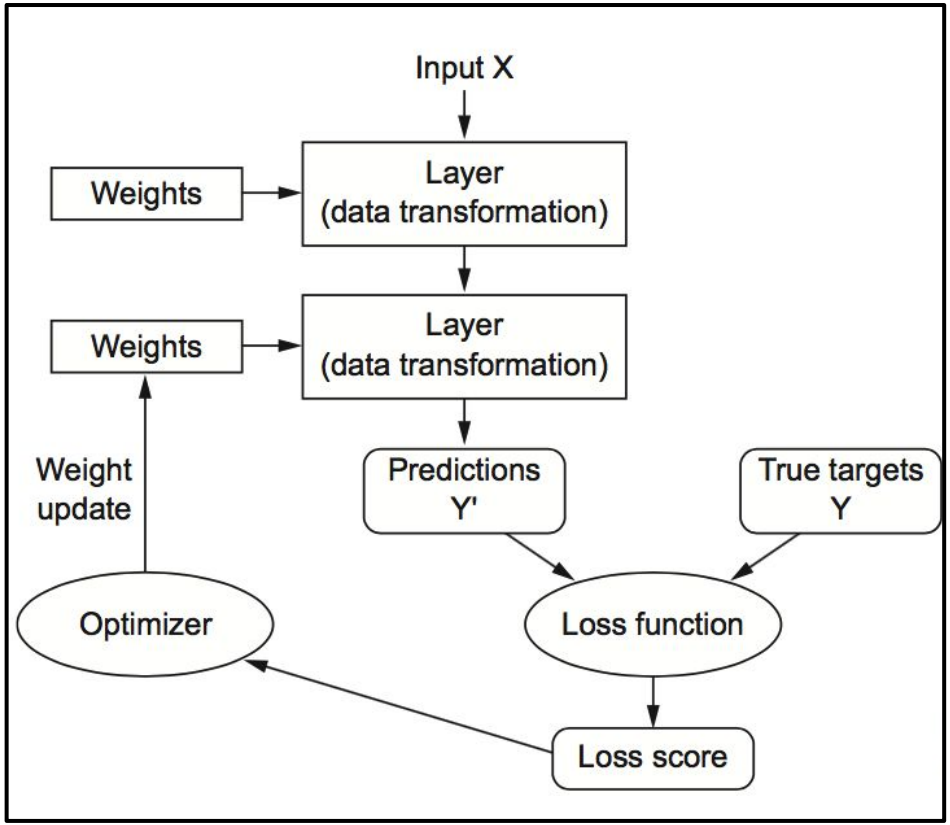


So what is Deep Learning?

Learning means finding a combination of model parameters that minimizes a loss function for a given set of training data samples and their corresponding targets.

Navigate to:

http://playground.tensorflow.org/

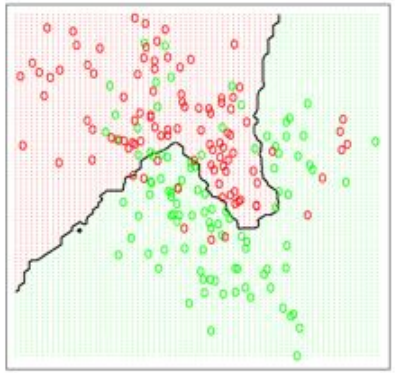


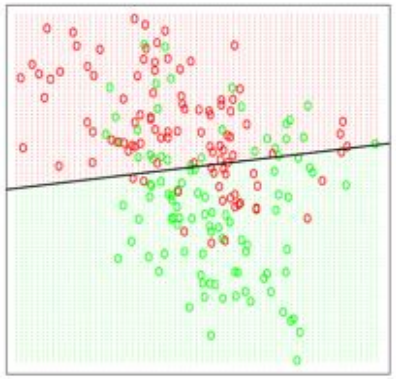
So what is Deep Learning?

Learning means finding a combination of model parameters that minimizes a loss function for a given set of training data samples and their corresponding targets.

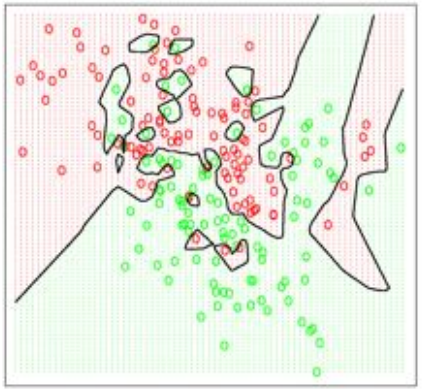
Navigate to:

http://playground.tensorflow.org/





Overfitting?



Let's build a super simple Neural Network!

Navigate to:

https://github.com/patrickcgray/deep_learning_ecology

So what is Deep Learning?

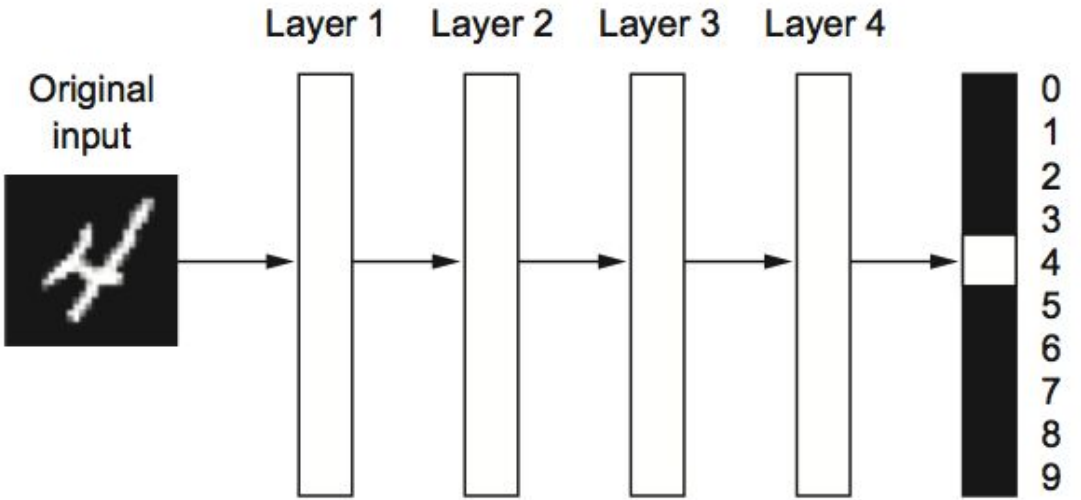
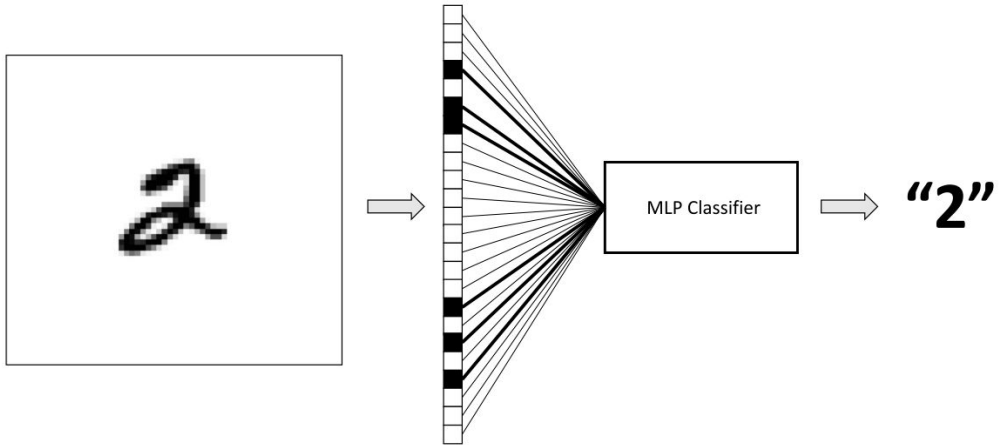


Figure 1.5 A deep neural network for digit classification

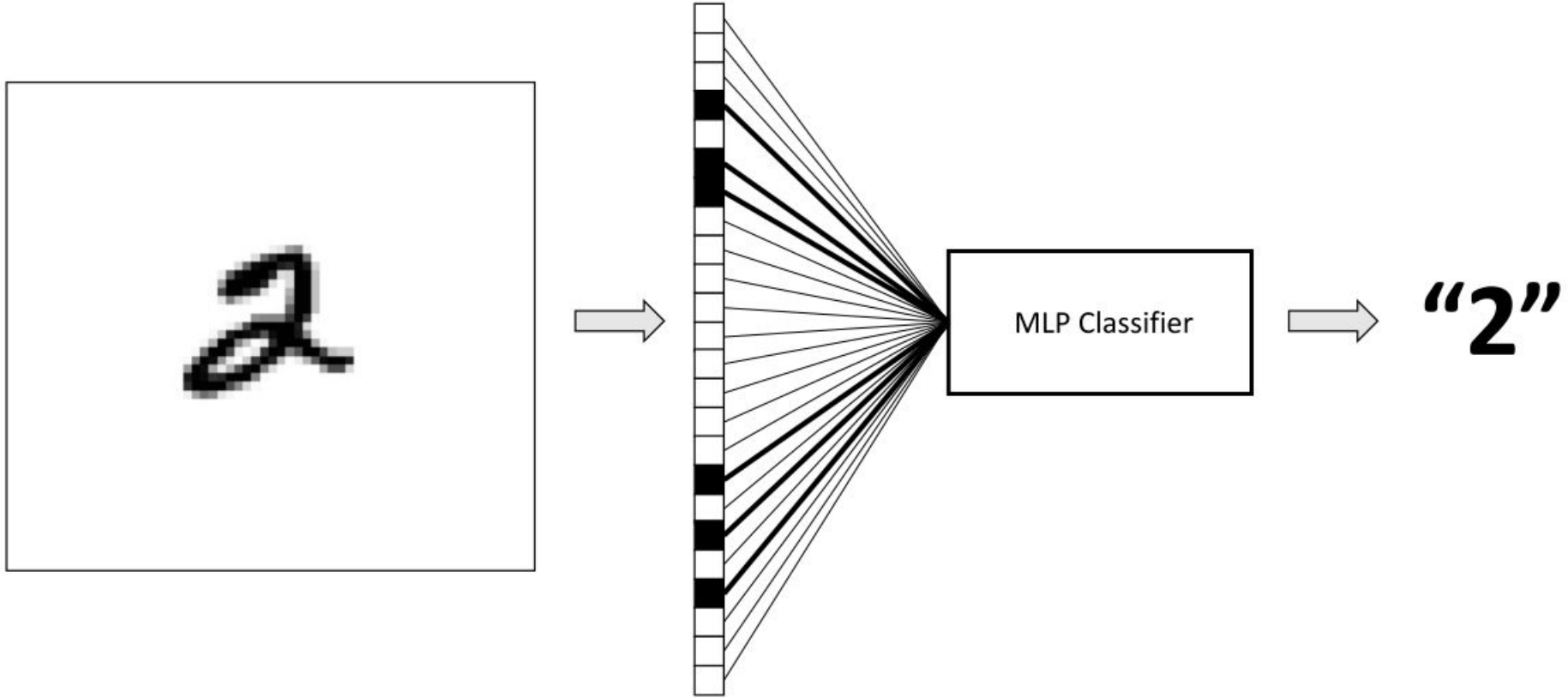
Train your own Multi-layer Perceptron!

Navigate to:

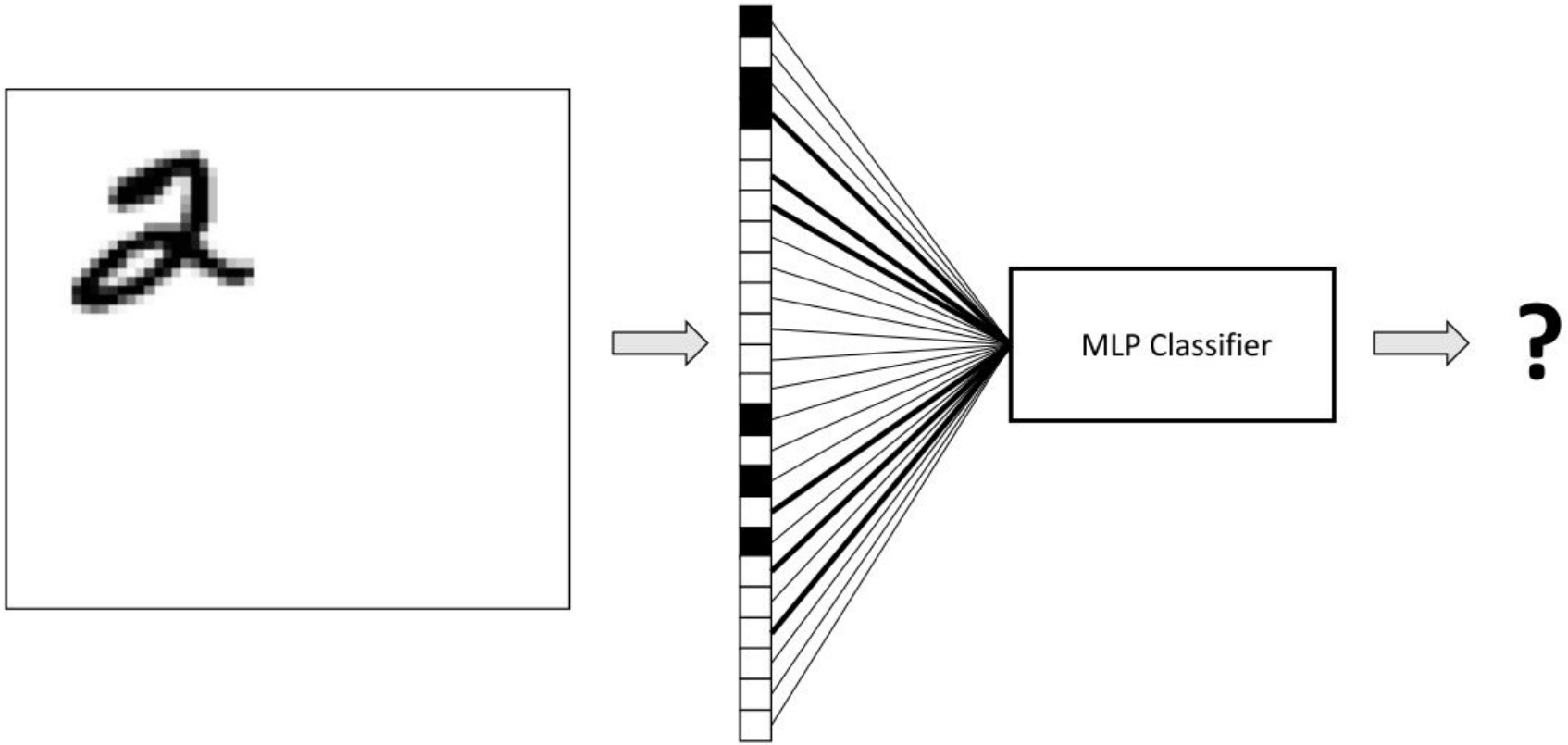
https://github.com/patrickcgray/deep_learning_ecology



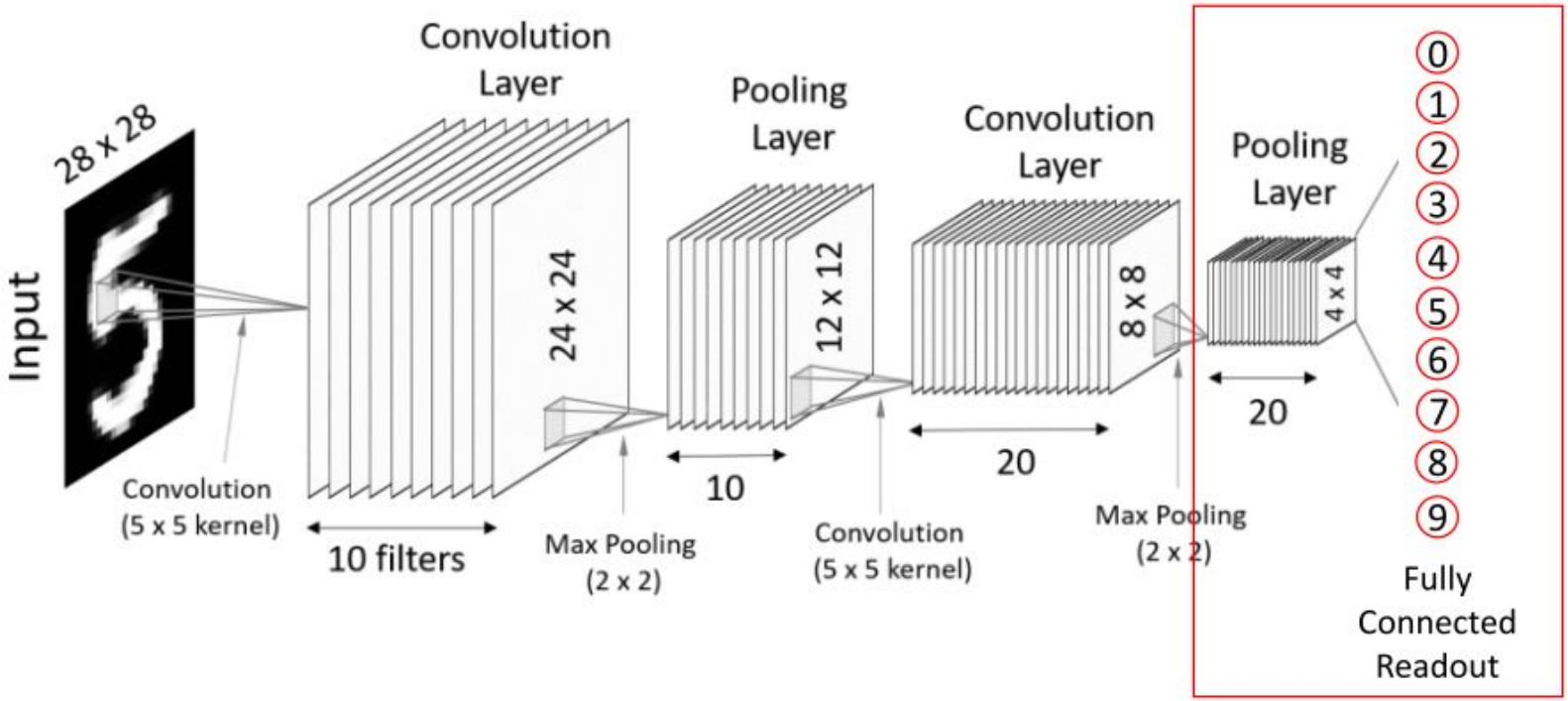
So what is Deep Learning?



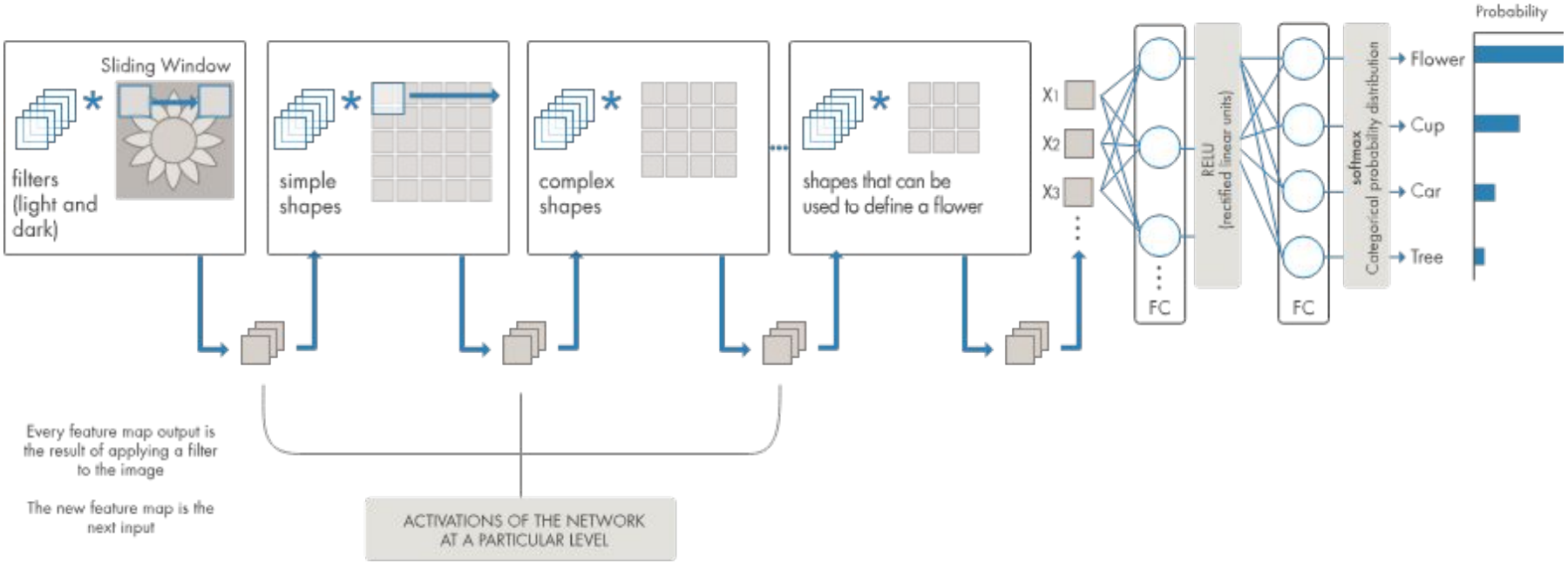
So what is Deep Learning?



So what is Deep Learning?



So what is Deep Learning?



So what is Deep Learning?

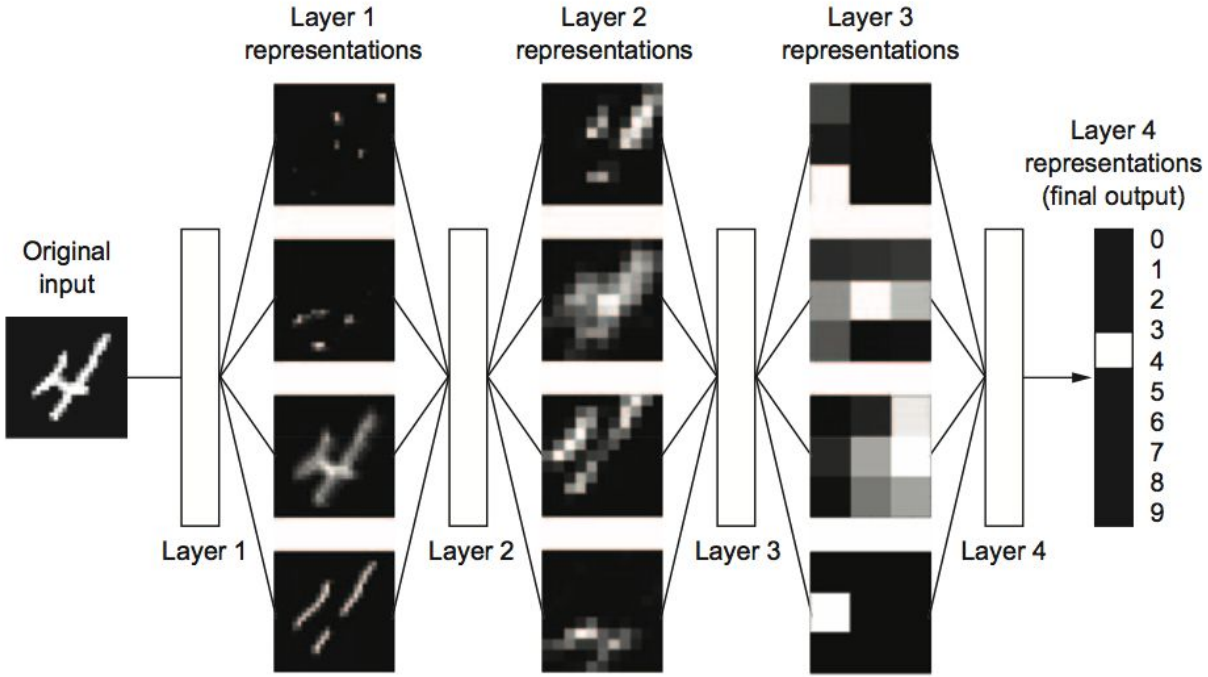
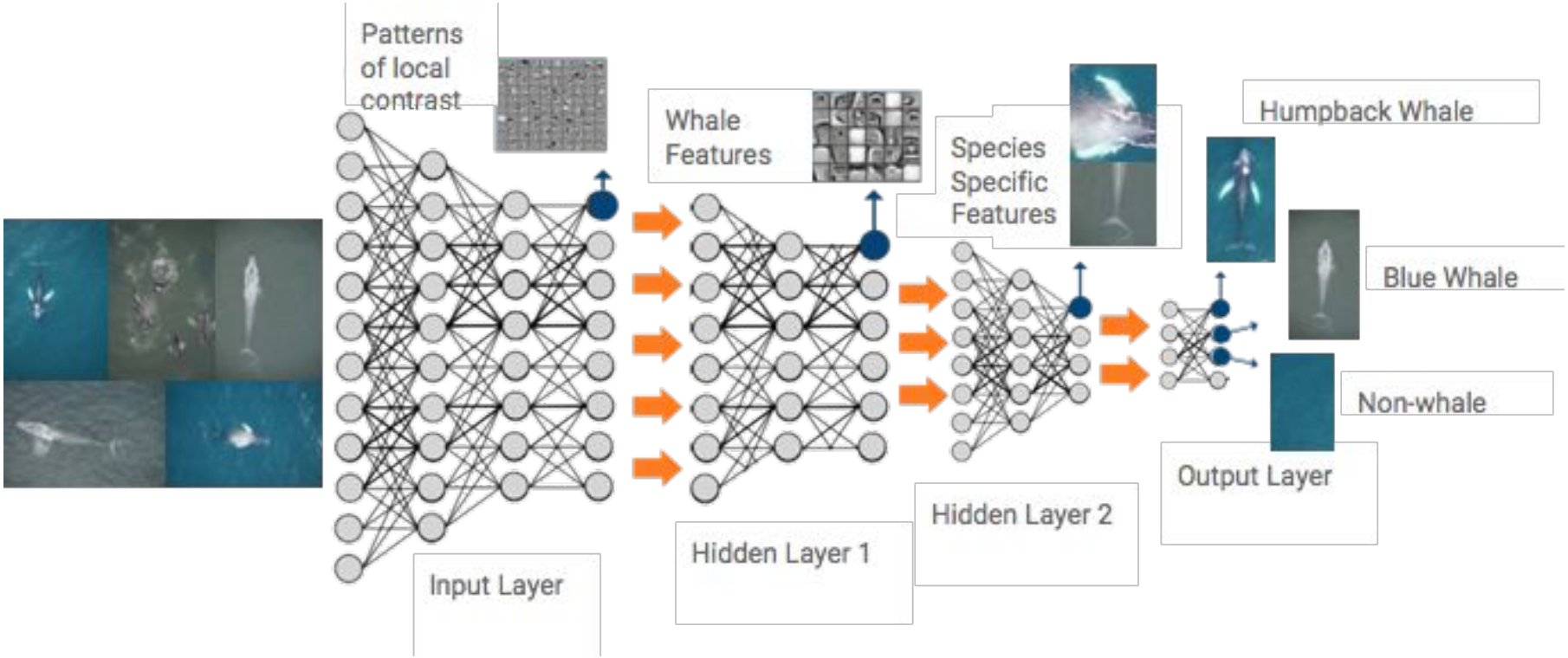


Figure 1.6 Deep representations learned by a digit-classification model

So what is Deep Learning?



Train your own Convolutional Neural Network!

Navigate to:

https://github.com/patrickcgray/deep_learning_ecology

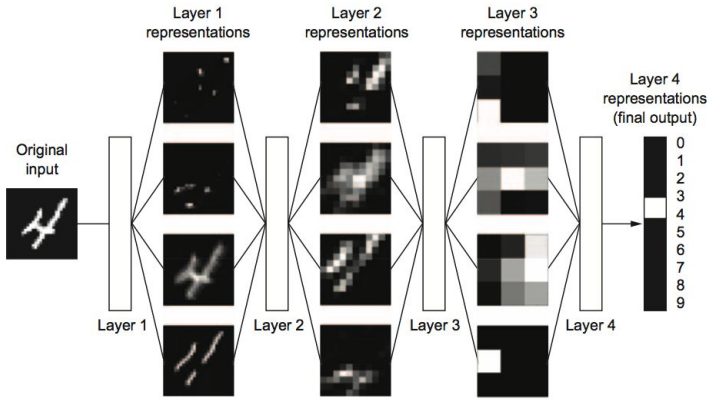
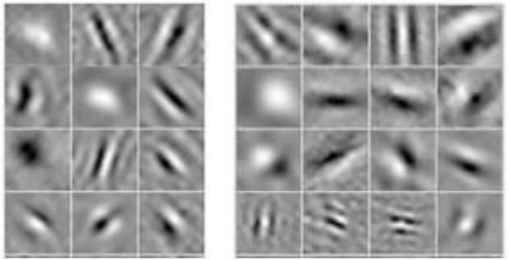


Figure 1.6 Deep representations learned by a digit-classification model

So what is Deep Learning?

Layer 1 Filters,
Convolutional Neural Network

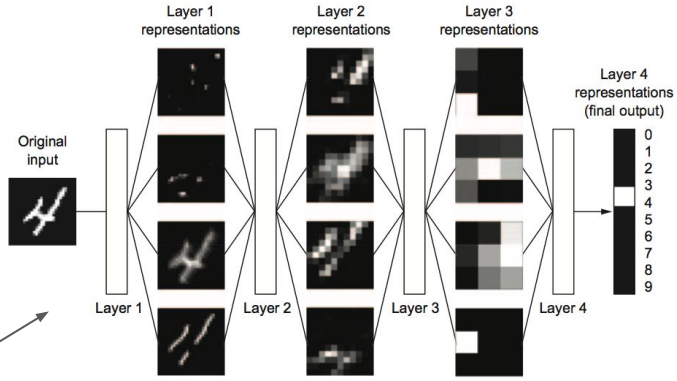
Neuron Receptive Fields,
Macaque Visual Cortex

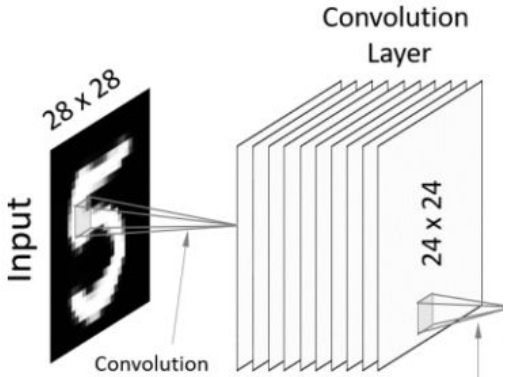


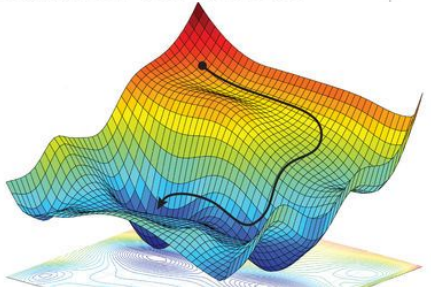
So what is Deep Learning?

Convolutional neural networks...

- learn to recognize high-level structure in images by building hierarchical representations
- extract features via spatial convolutions with filters
- learn filters via iterative minimization of a risk function
- **have shown capabilities beyond human performance for image analysis**







Train your own CNN for Real small Images!

Navigate to:

https://github.com/patrickcgray/deep_learning_ecology