Intro to Machine Learning

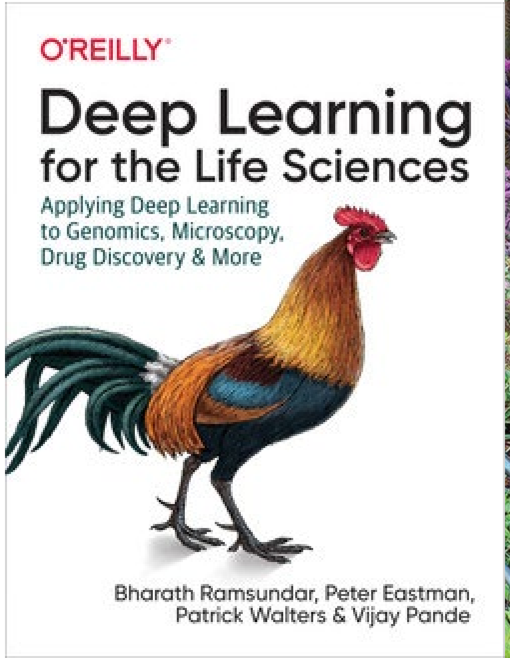
Duke Quantitative Living Systems Center

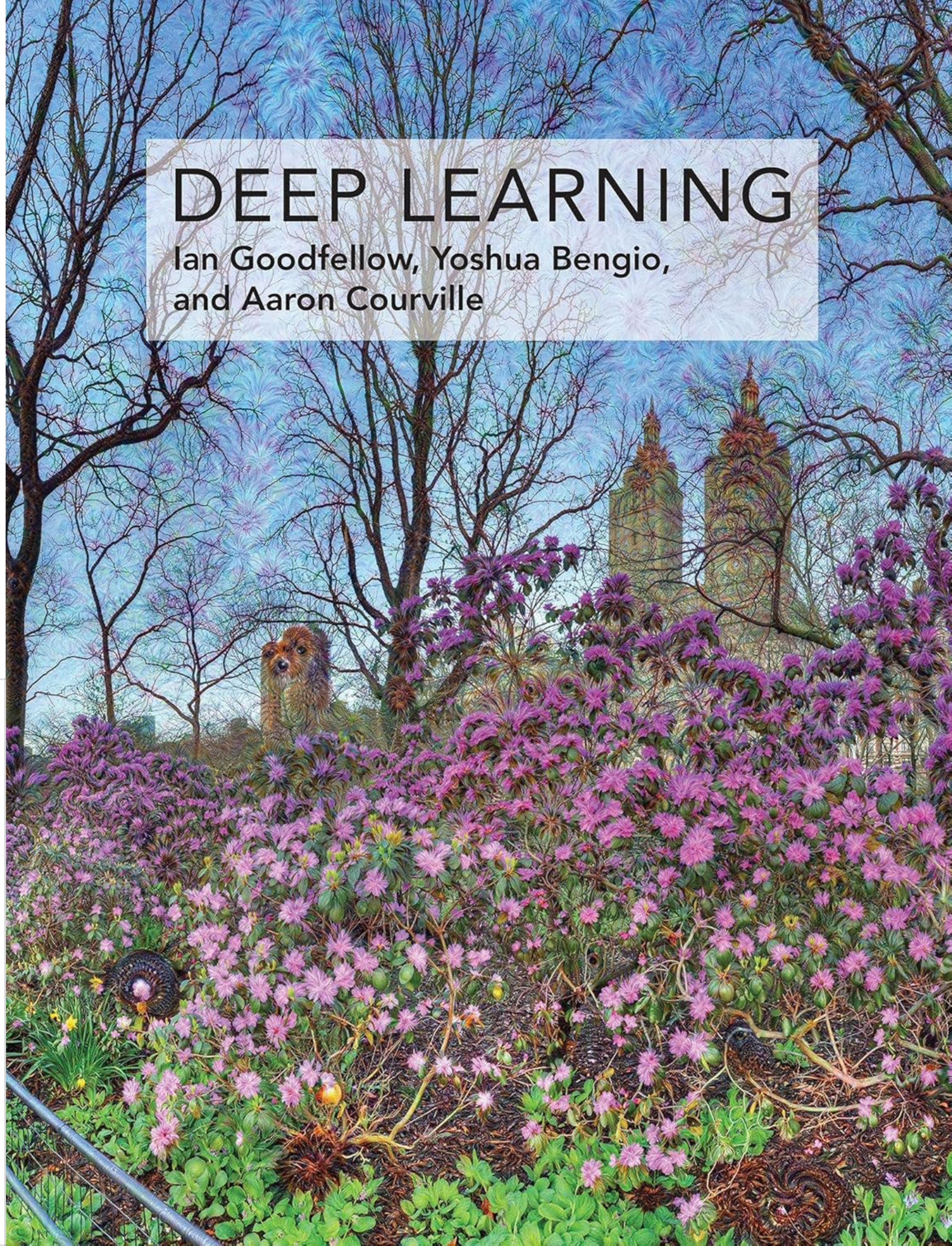
Justin Savage

1/29/25

Resources

- https://www.deeplearningbook.org/
- https://www.oreilly.com/library/view/deep-learning-for/9781492039822/





What types of problems can computers solve?

What types of problems can computers solve?

Easy for Computers

- Math problems
- Working with lots of data
- Games with formal rules

Hard for Computers

- Recognizing spoken words
- Identifying objects in images
- Understanding data

What types of problems can computers solve?

Easy for Computers

- Math problems
- Working with lots of data
- Games with formal rules

Hard for Computers

- Recognizing spoken words
- Identifying objects in images
- Understanding data

Computers excel at tasks where the key representations of the data are well defined

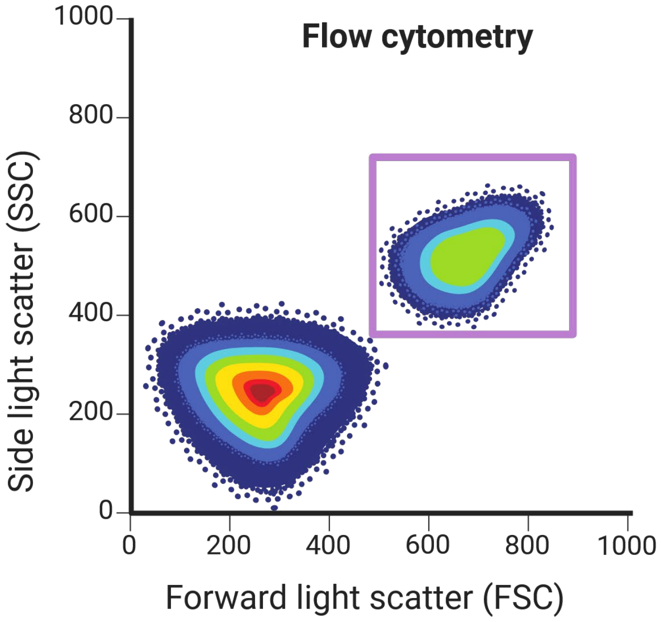
Three flavors of “AI”

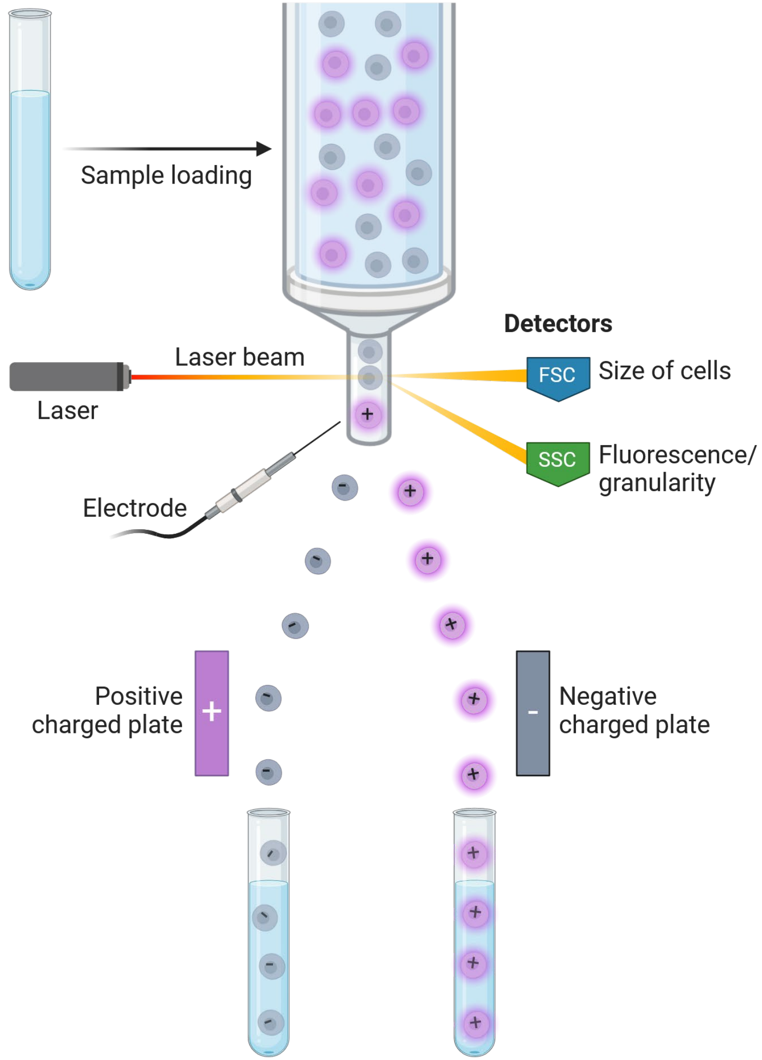
- Rule based systems
- Classic Machine Learning
- Representation Learning

Rule based systems

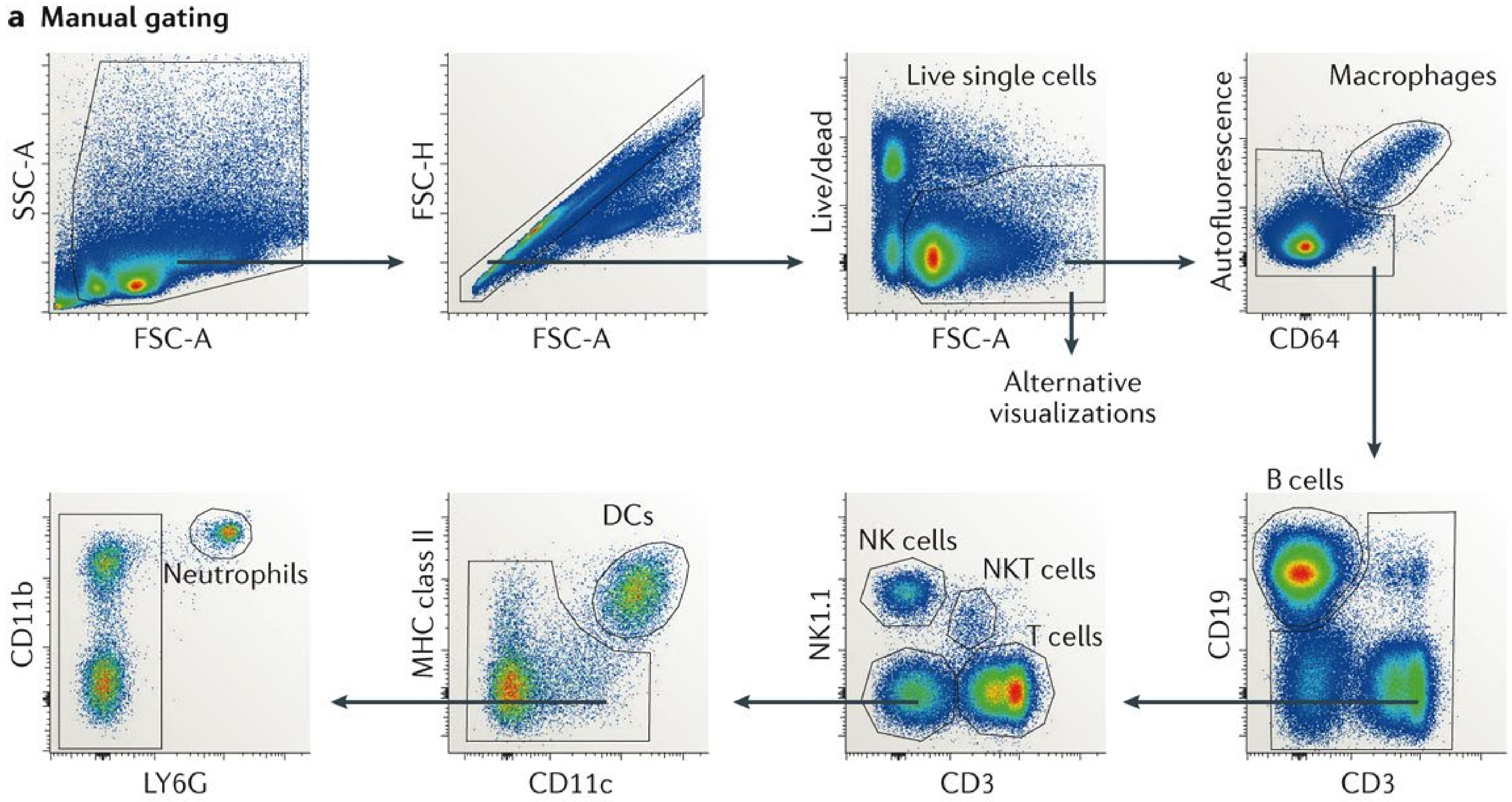
- If condition X then output Y
- Relies on the programmer providing all necessary information for making a decision including:
  - The data in a simple format
  - The exact decision method and boundary values are provided by the programmer
- Example: FACS sorting

Example FACS Sorting





Example FACS Sorting



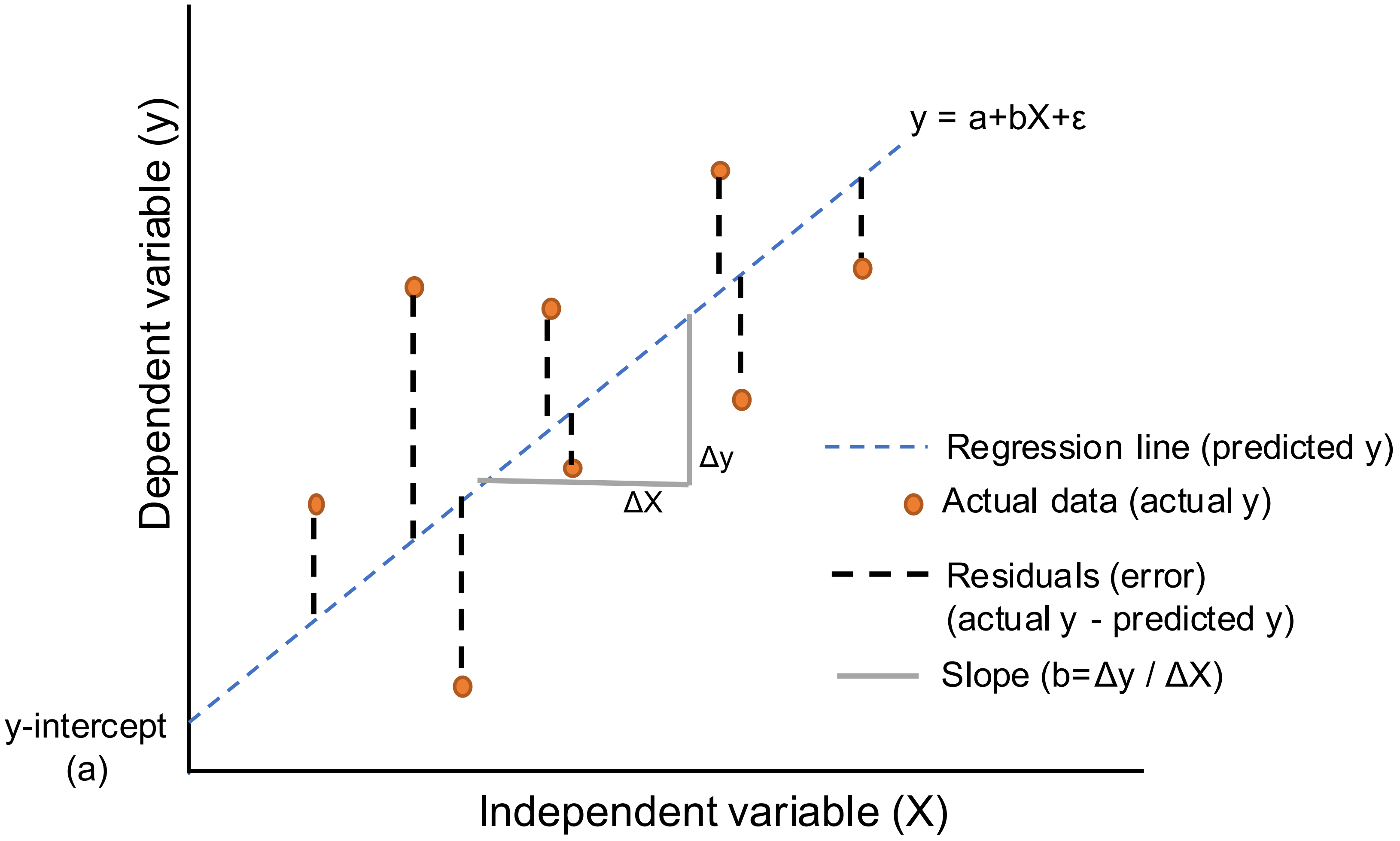
Rule based systems

- “The difficulties faced by systems relying on hard-coded knowledge suggest that AI systems need the ability to acquire their own knowledge, by extracting patterns from raw data. This capability is known as machine learning.”

Classical Machine Learning

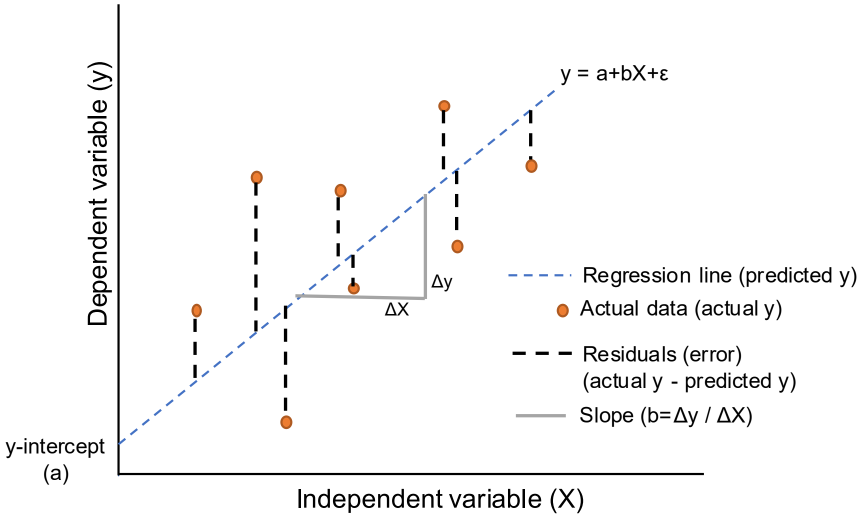
- The program optimizes a small number of parameters (learning)
- Relies on the programmer providing some structure for making a decision including:
  - The data in a simple format
  - A general framework for decision making is provided by the programmer
  - Some feedback mechanism for how far off a prediction is from the truth
- Example: Linear Regression

Linear Regression



Linear Regression

- Program makes a guess for what the parameters (slope and intercept) should be
- Program uses the formula residuals = actual y – predicted y for each point to assess how well it predicted
- Program adjusts parameters to minimize the residuals

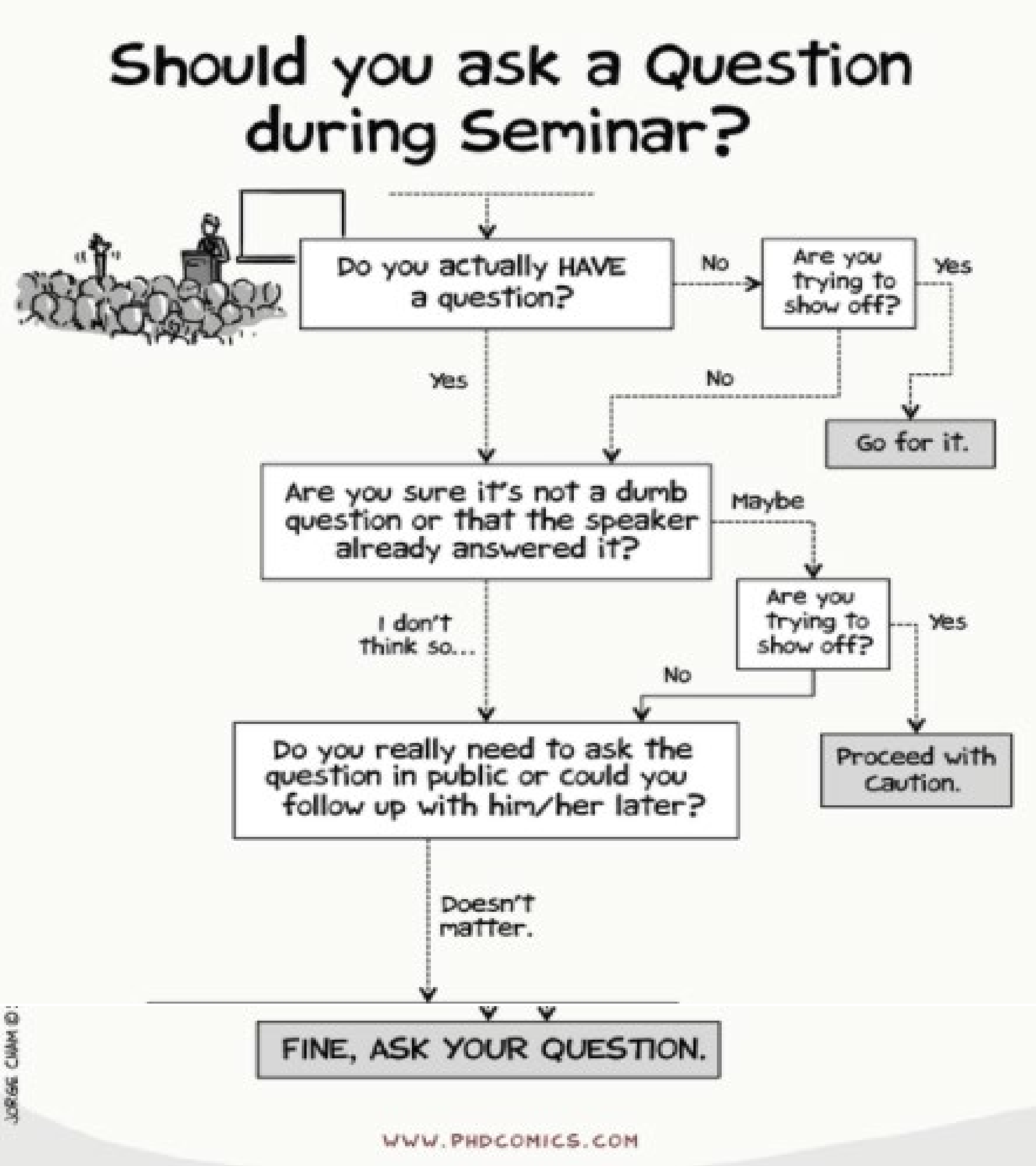


Classical Machine Learning

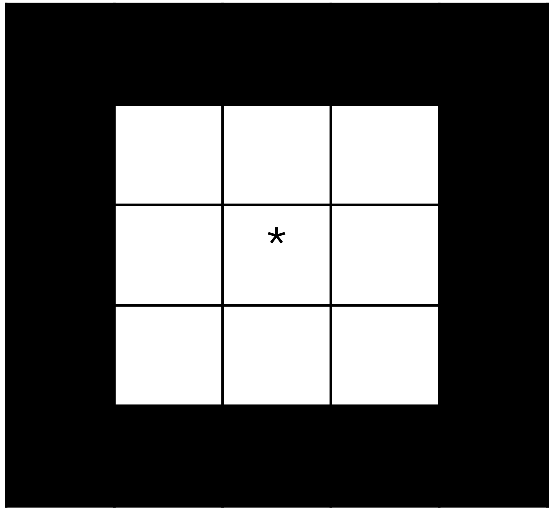
- The program optimizes a small number of parameters (learning)
- Relies on the programmer providing some structure for making a decision including:
  - The data in a simple format
  - A general framework for decision making is provided by the programmer
  - Some feedback mechanism for how far off a prediction is from the truth
- Example: Linear Regression
- Example: Random Forest (ilastik)

How ilastik works

- Random Forest of Decision Trees

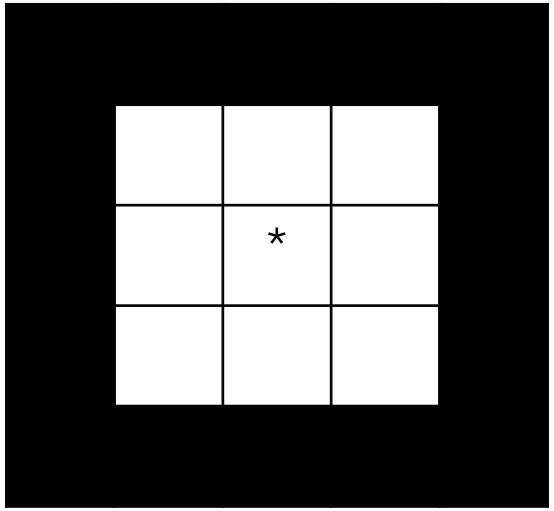


Decision Trees



- Trying to figure out if pixel * is a puncta or background
- We know the Intensity of * (range is 0-255)
- We know the position of * and the intensity of all of the pixels around it
- What would be a good first feature to consider to help decide if * is a puncta or background?

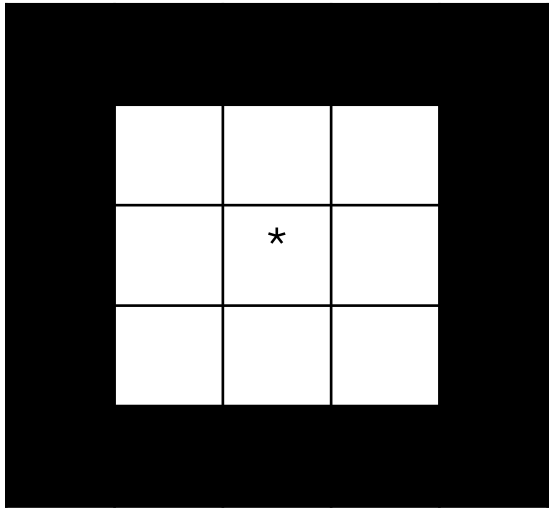
Decision Trees



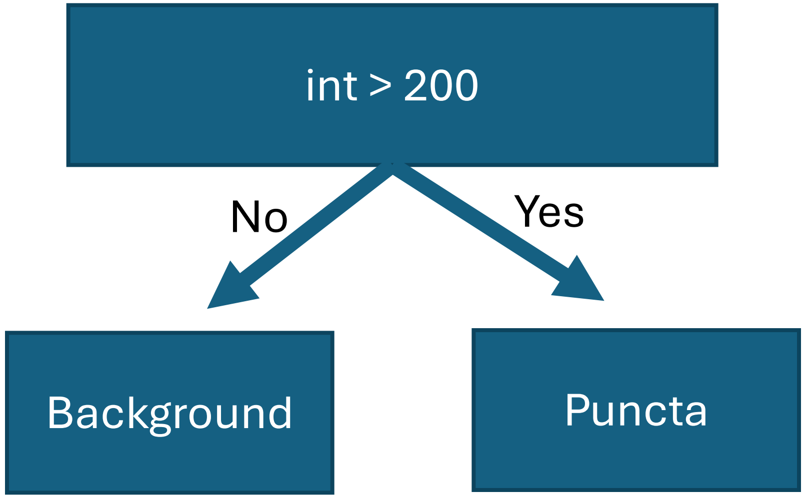
- Trying to figure out if pixel * is a puncta or background
- We know the Intensity of * (range is 0-255)
- We know the position of * and the intensity of all of the pixels around it
- What would be a good first feature to consider to help decide if * is a puncta or background?

Is * bright enough to be a puncta?

Decision Trees

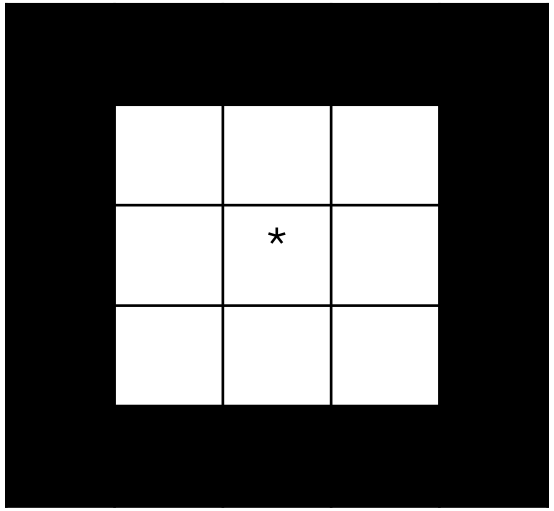


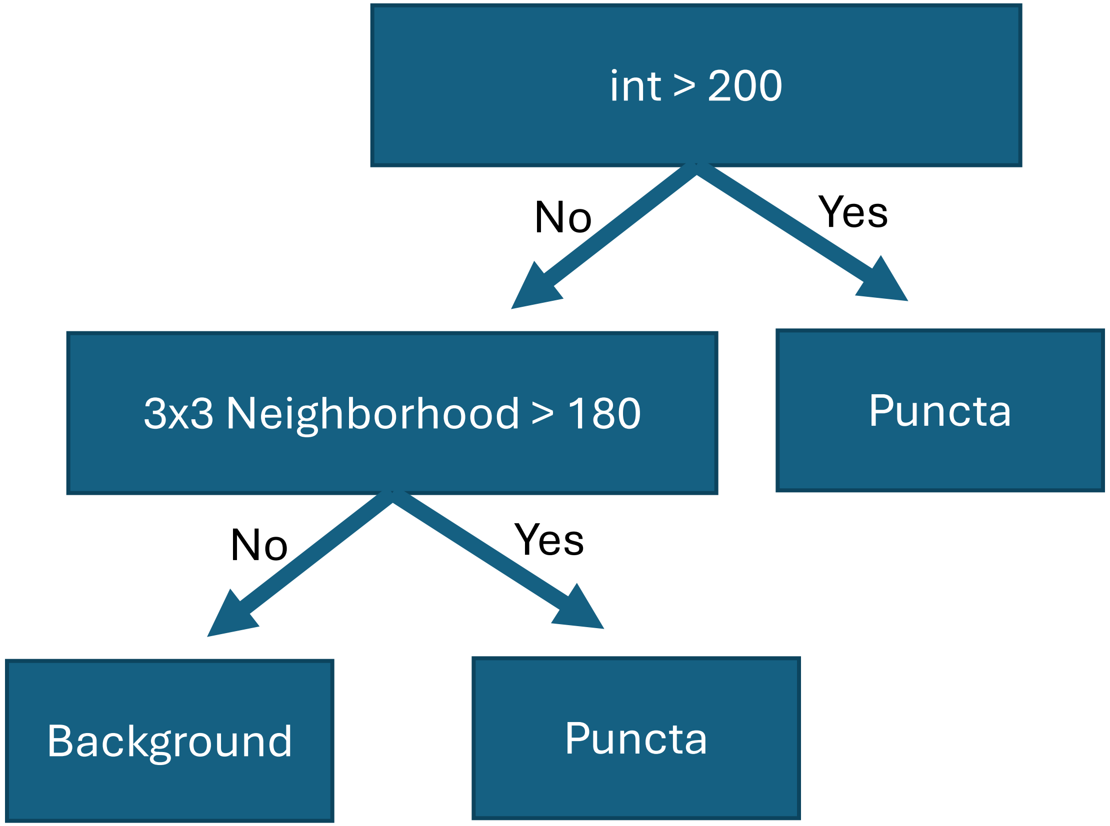
Try an intensity cutoff



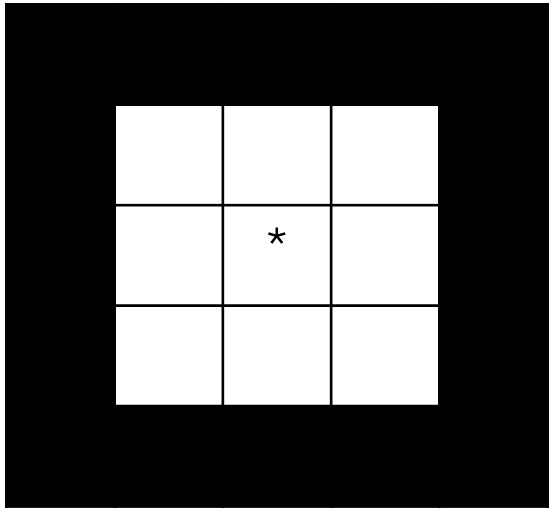
This is the same as traditional thresholding, but we can do better

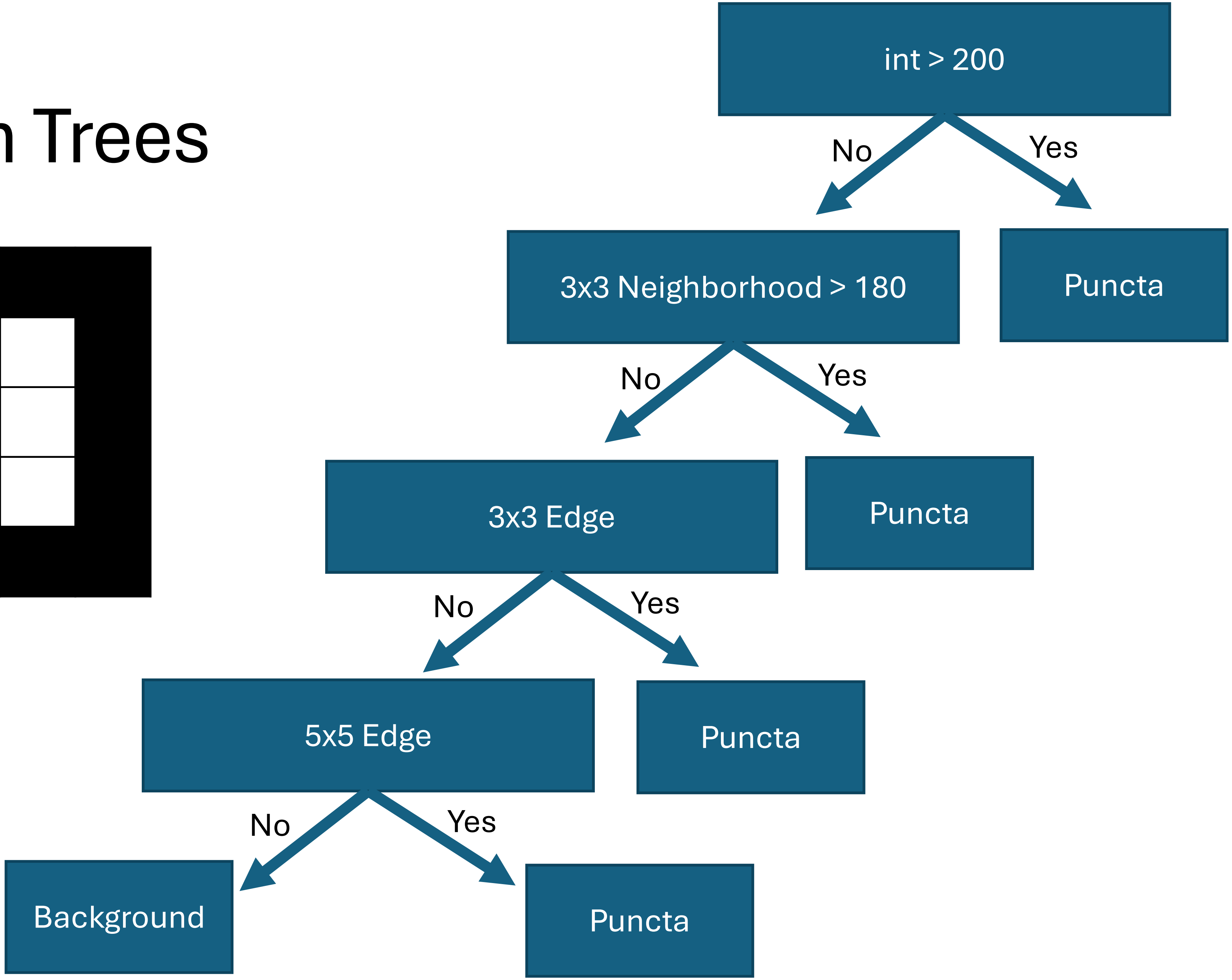
Decision Trees





Decision Trees

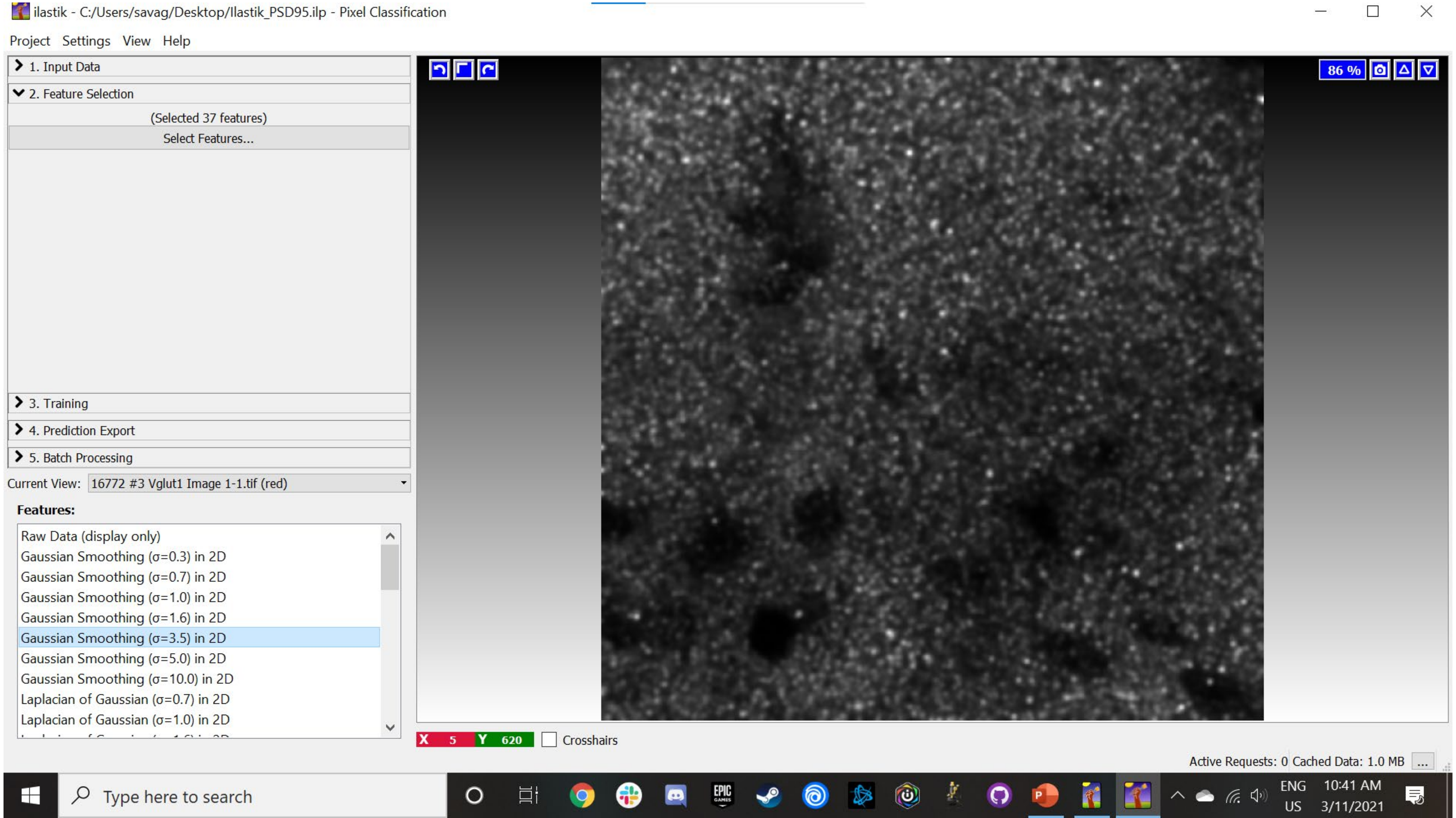




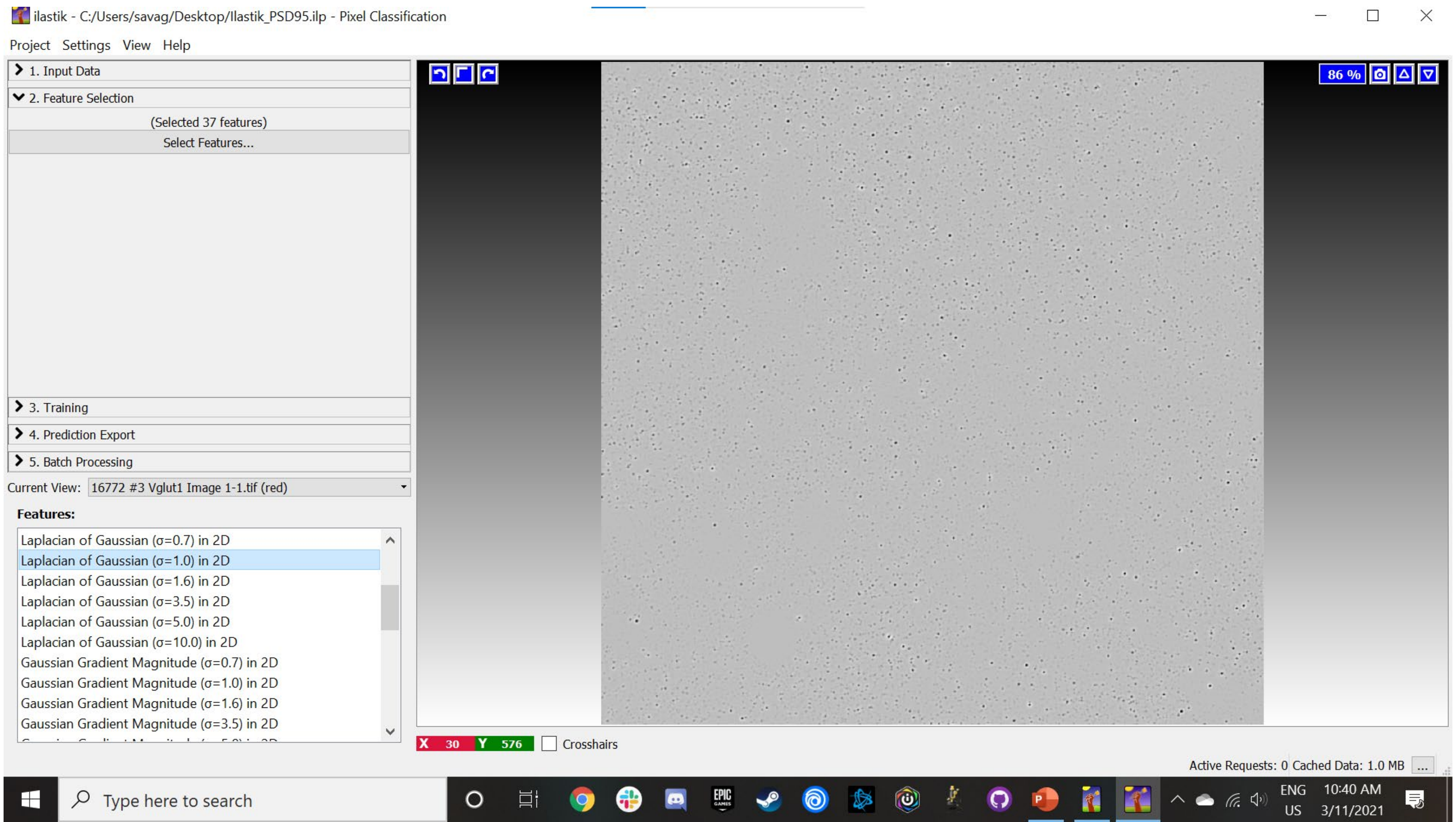
Random Forest

- ilastik uses 100 decision trees that are randomly generated to be slightly different from each other
- Each one ”votes” for whether a pixel is foreground or background

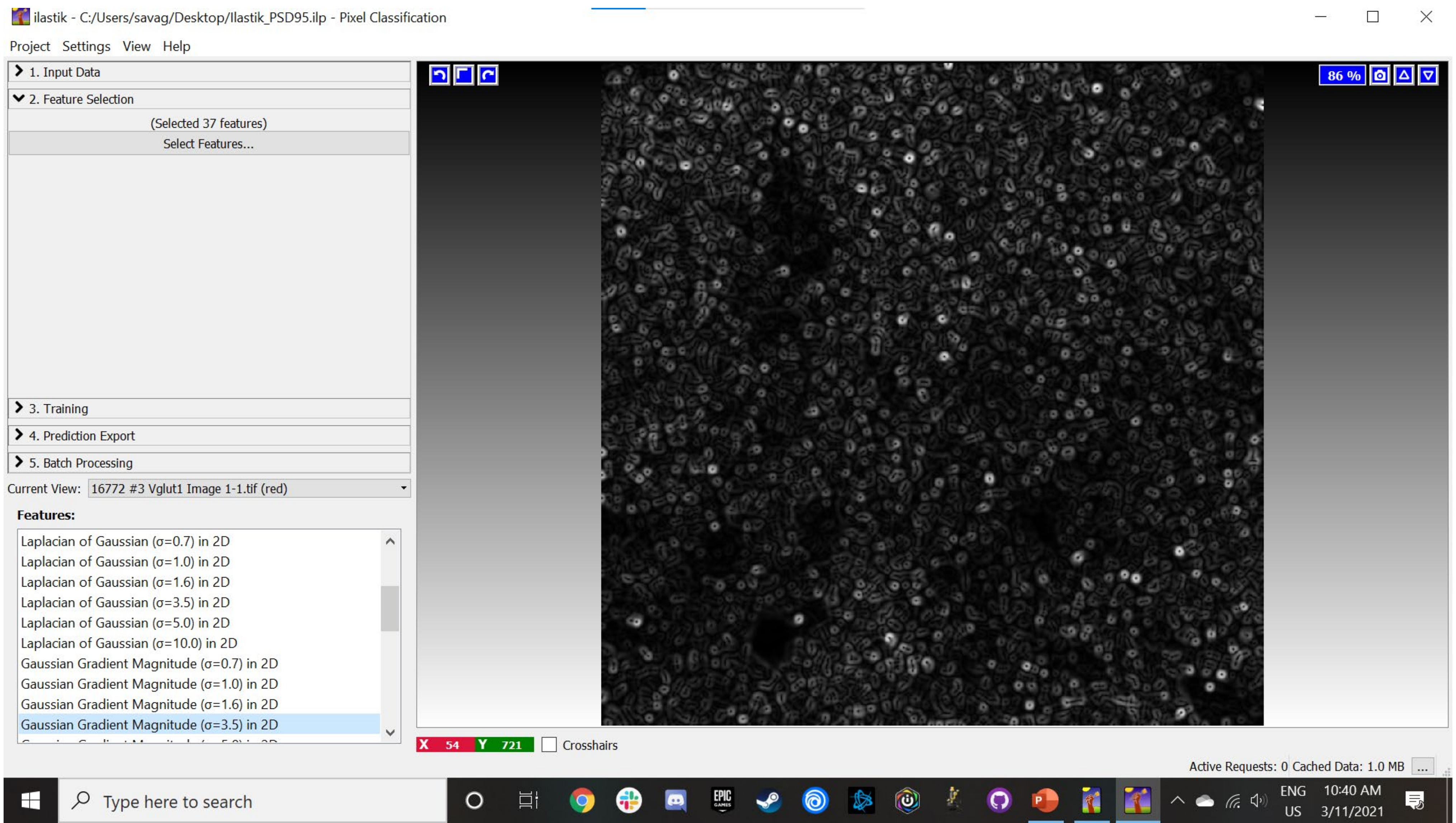
Viewing Features in Feature Selection Window



Viewing Features in Feature Selection Window



Viewing Features in Feature Selection Window



Classical Machine Learning

- The program optimizes a small number of parameters (learning)
- Relies on the programmer providing some structure for making a decision including:
  - The data in a simple format
  - A general framework for decision making is provided by the programmer
  - Some feedback mechanism for how far off a prediction is from the truth
- Example: Linear Regression
- Example: Random Forest (ilastik)

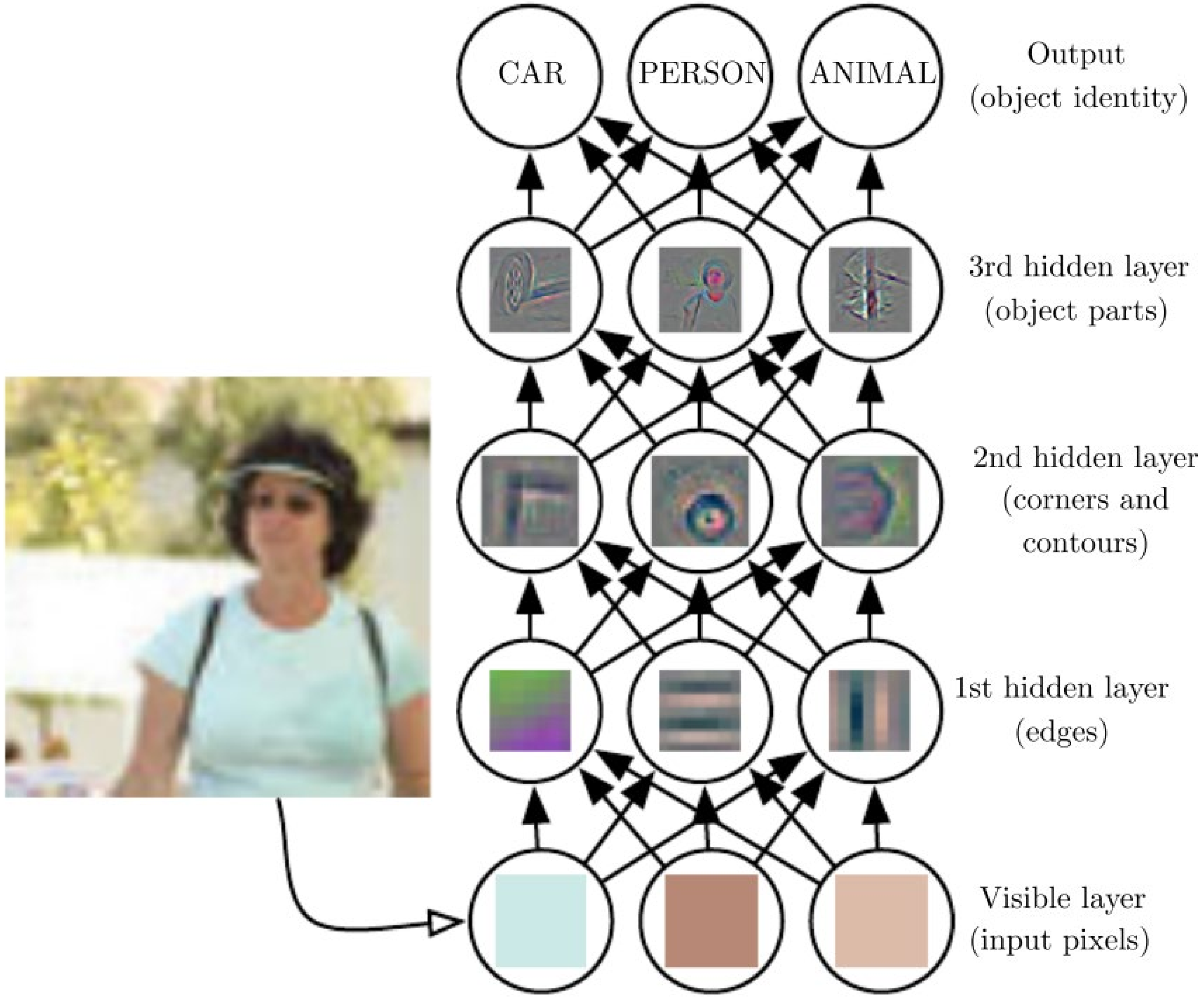
Representation Learning

- For some applications, it is difficult to say which features of a data set are important for a decision
- We may want the algorithm to discover novel combinations of features that could explain the data better than features we have as input
- We may want to eliminate as much of our own bias as possible

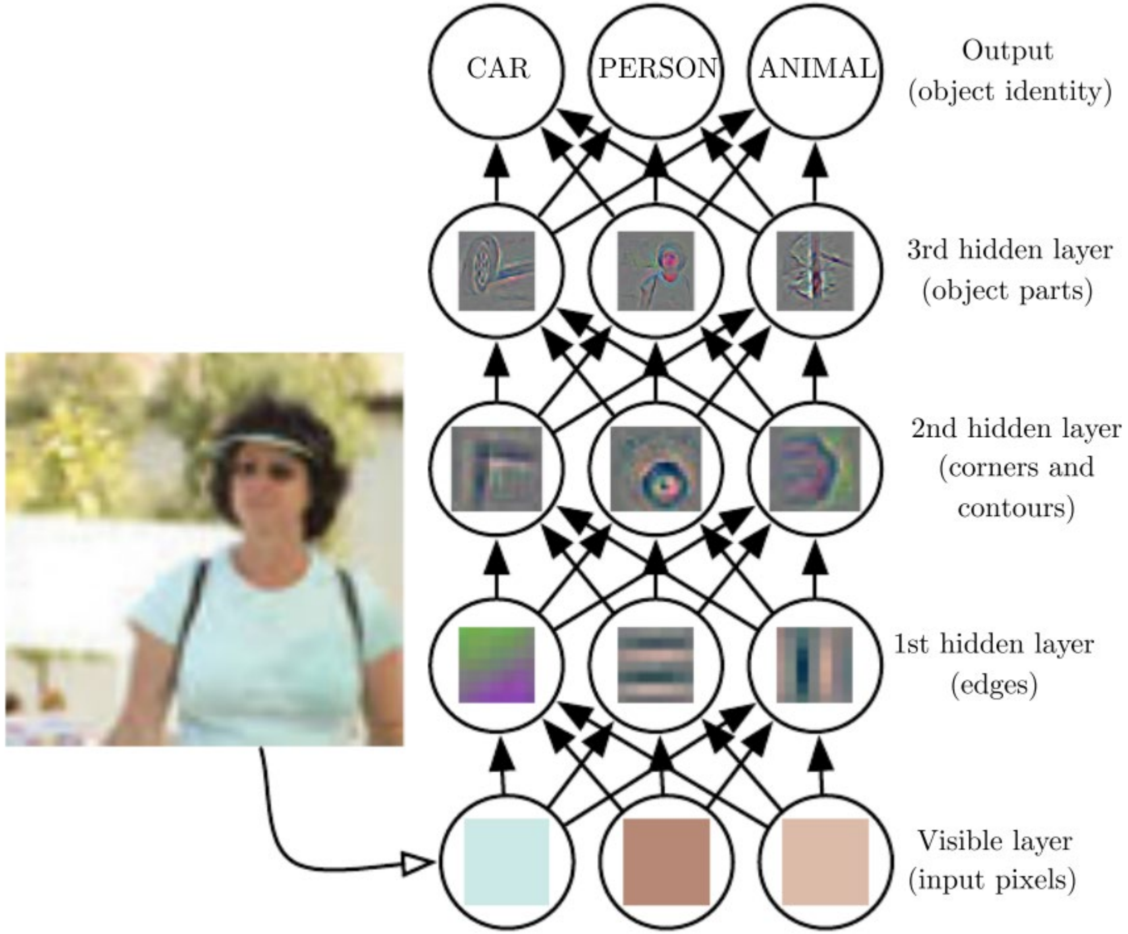
Representation Learning

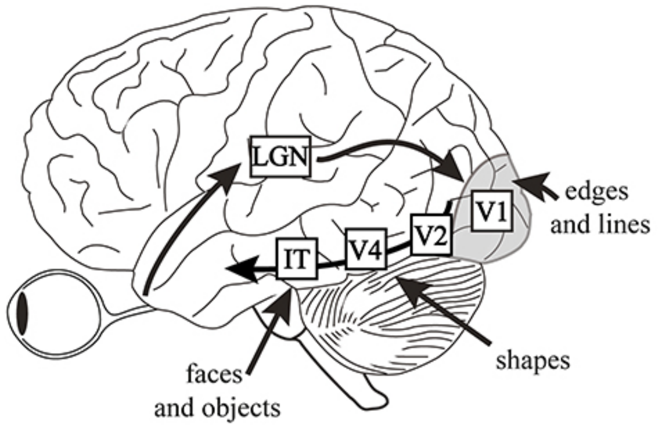
- The program not only learns parameters, but learns which representations of the data are most useful for decision making
- Relies on the programmer providing a minimal framework for making a decision including:
  - The data in a complex, raw format
  - Some structure for how many features to learn and how they should be connected to eachother
  - Some feedback mechanism for how far off a prediction is from the truth
- Example: Artificial Neural Networks

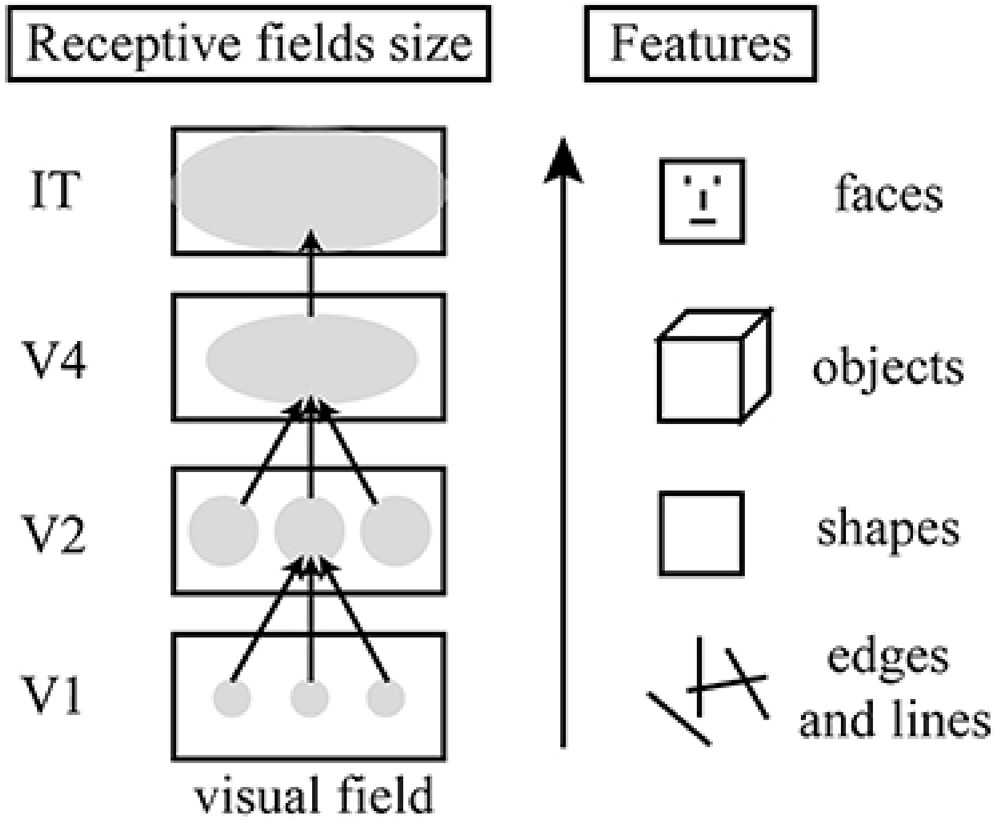
Artificial Neural Networks



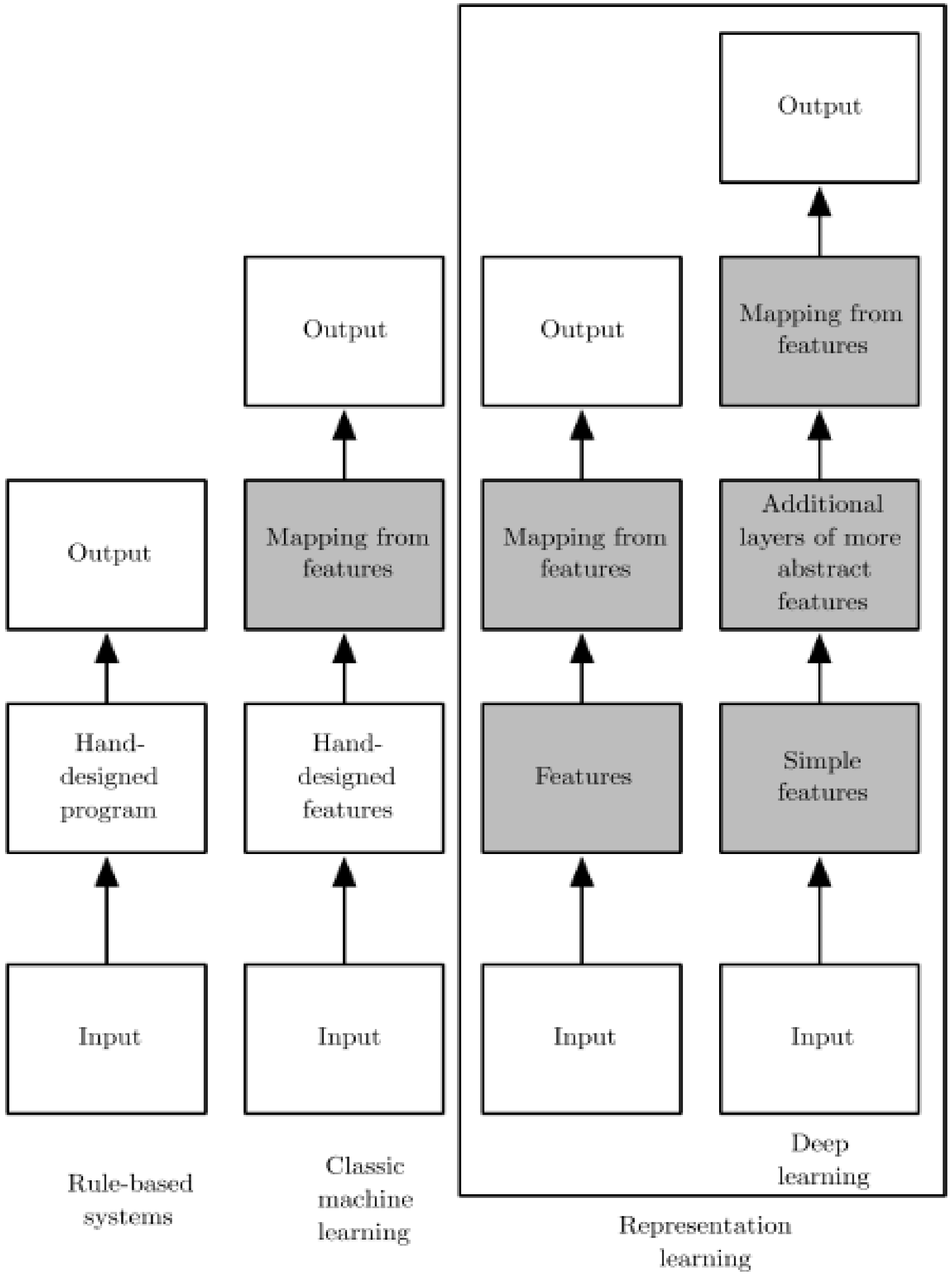
Artificial Neural Networks







Three flavors of "AI"



When should you use each flavor of AI?

- If the relevant values and computations are known by the experimenter

When should you use each flavor of AI?

- If the relevant values and computations are known by the experimenter
  - Use rule-based systems

When should you use each flavor of AI?

- If the relevant values and computations are known by the experimenter
  - Use rule-based systems
- If some relevant features are known, but the best values of the parameters are unknown

When should you use each flavor of AI?

- If the relevant values and computations are known by the experimenter
  - Use rule-based systems
- If some relevant features are known, but the best values of the parameters are unknown
  - Classic machine learning

When should you use each flavor of AI?

- If the relevant values and computations are known by the experimenter
  - Use rule-based systems
- If some relevant features are known, but the best values of the parameters are unknown
  - Classic machine learning
- If there is no clear set of features that best represent the data

When should you use each flavor of AI?

- If the relevant values and computations are known by the experimenter
  - Use rule-based systems
- If some relevant features are known, but the best values of the parameters are unknown
  - Classic machine learning
- If there is no clear set of features that best represent the data
  - Representation learning/ Deep learning

Biological Example: RNAseq data

- RNAseq data from M and F animals collected at different ages
- You want to determine how variable the samples are and if there are effects of sex and/or age
- You've measured thousands of genes and don't know which genes would be changing

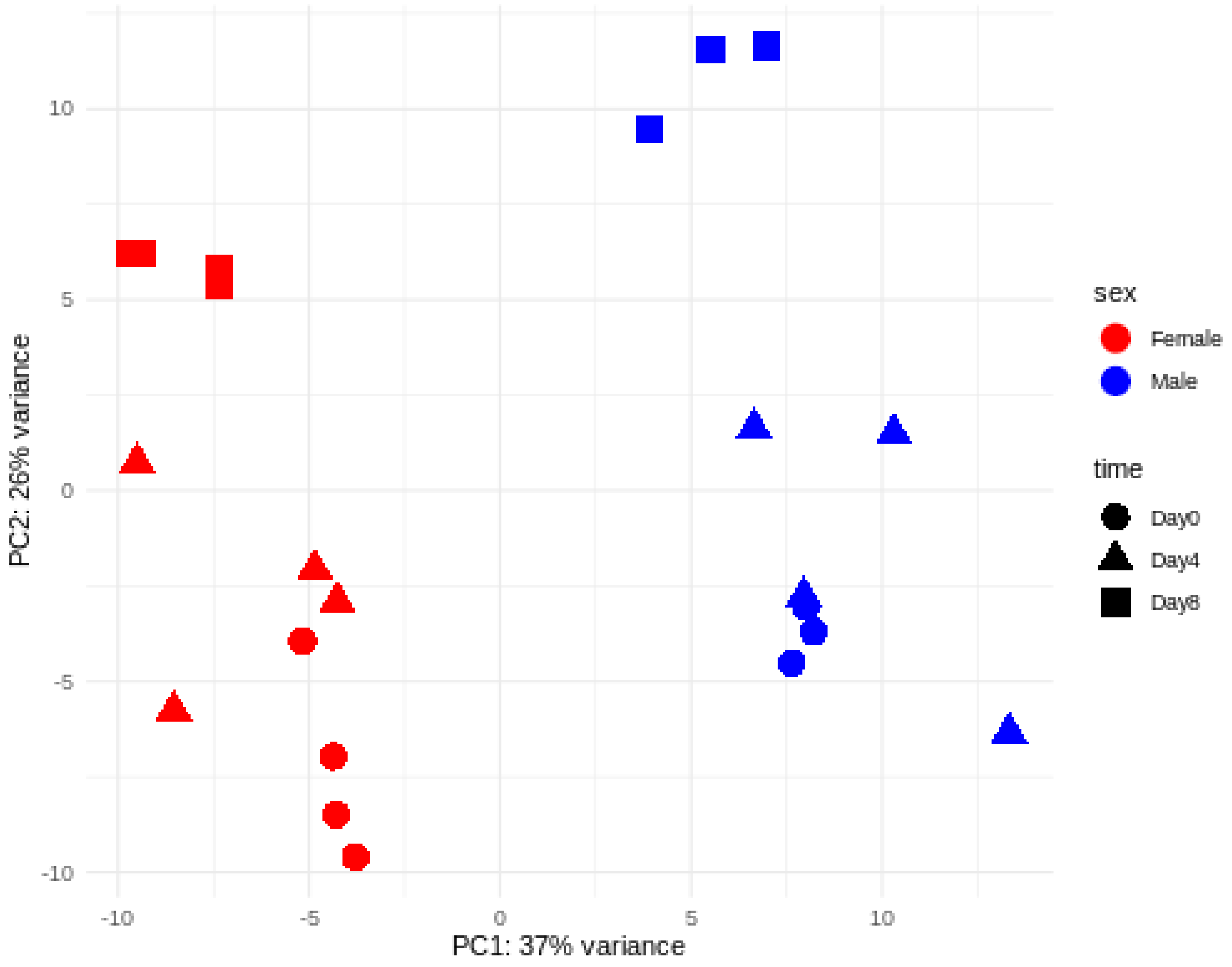
https://carpentries-incubator.github.io/bioc-rnaseq/05-exploratory-qc.html

Biological Example: RNAseq data

- RNAseq data from M and F animals collected at different ages
- You want to determine how variable the samples are and if there are effects of sex and/or age
- You’ve measured thousands of genes and don’t know which genes would be changing
- ^We need a better representation of the data to more easily visualize the differences between samples

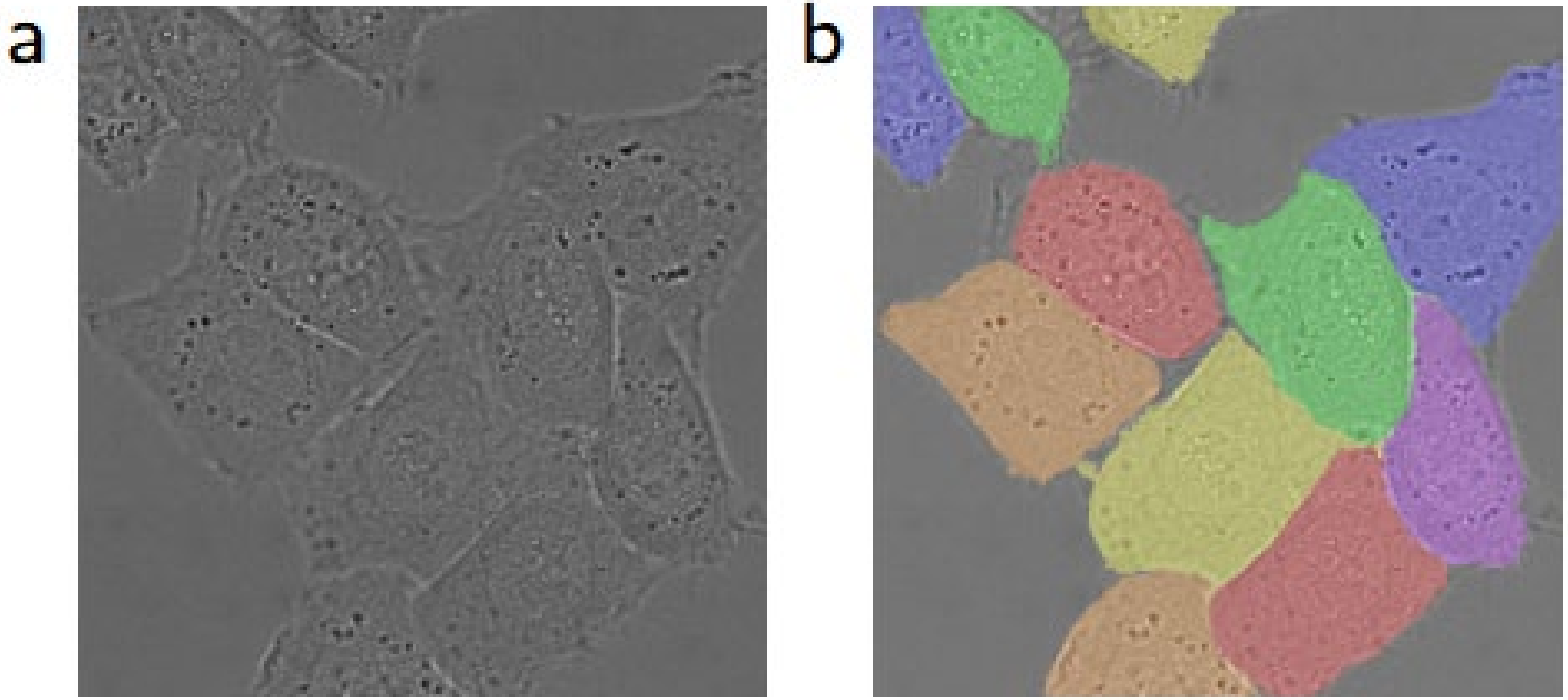
https://carpentries-incubator.github.io/bioc-rnaseq/05-exploratory-qc.html

Biological Example: RNAseq data PCA



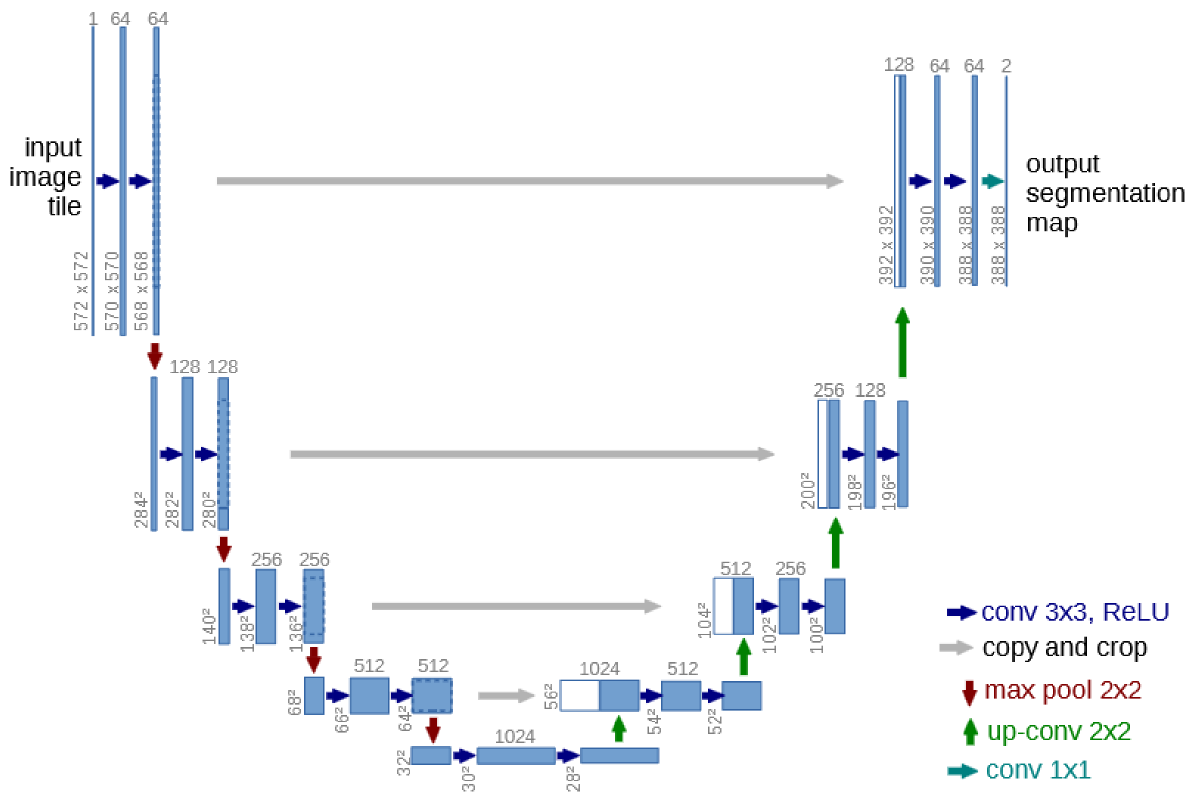
https://carpentries-incubator.github.io/bioc-rnaseq/05-exploratory-qc.html

Biological Example: Image Segmentation



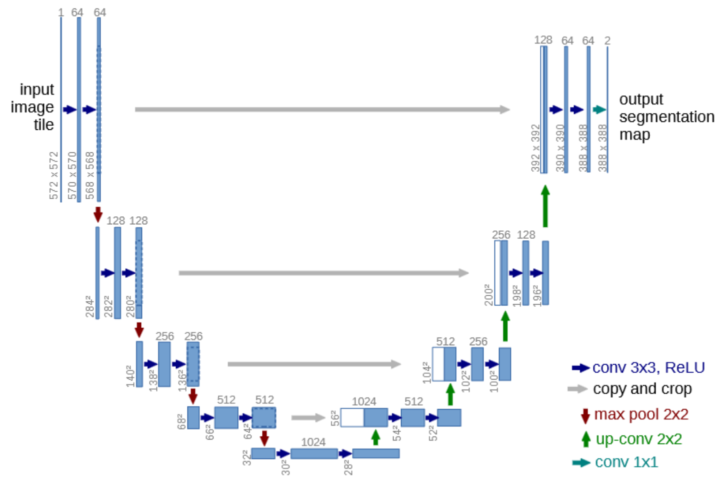
Biological Example: Image Segmentation

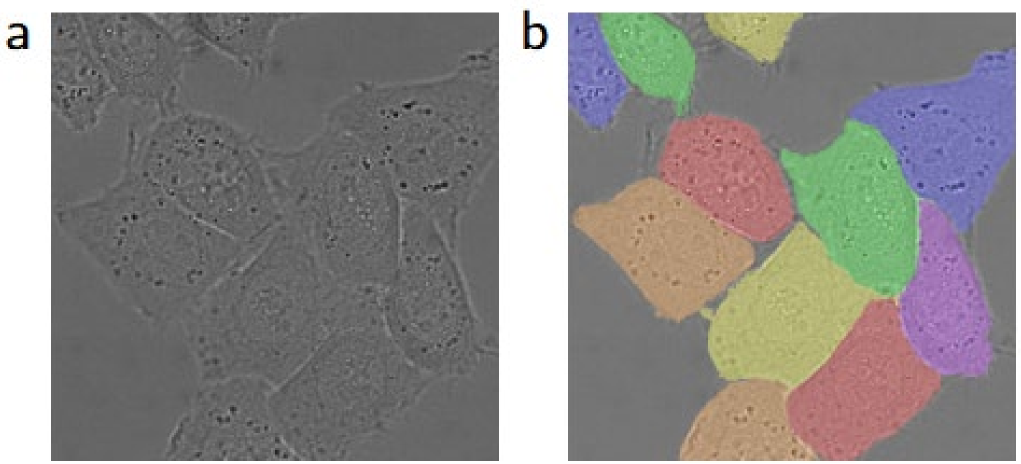
U-net: Convolutional networks for biomedical image segmentation
O Ronneberger, P Fischer, T Brox - … **image** computing and computer …, 2015 - Springer
… use of **convolutional networks** is on classification tasks, where the output to an **image** is a single class label. However, in many visual tasks, especially in **biomedical image processing**, …



Biological Example: Image Segmentation

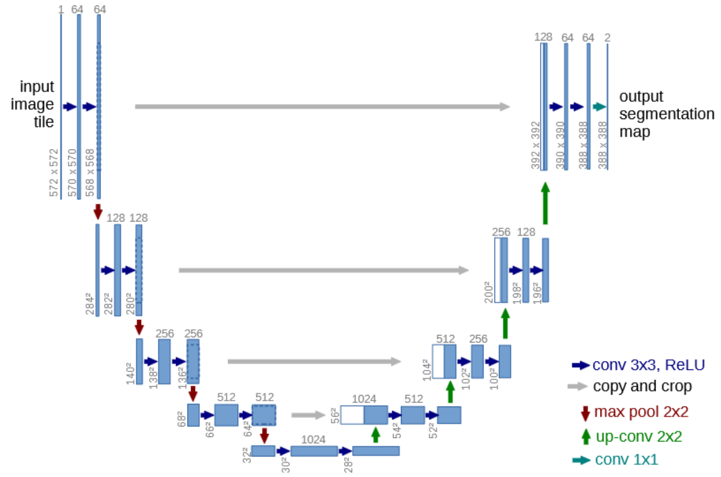
- How does the neural network actually learn?
  - Weights in the network are randomly initialized

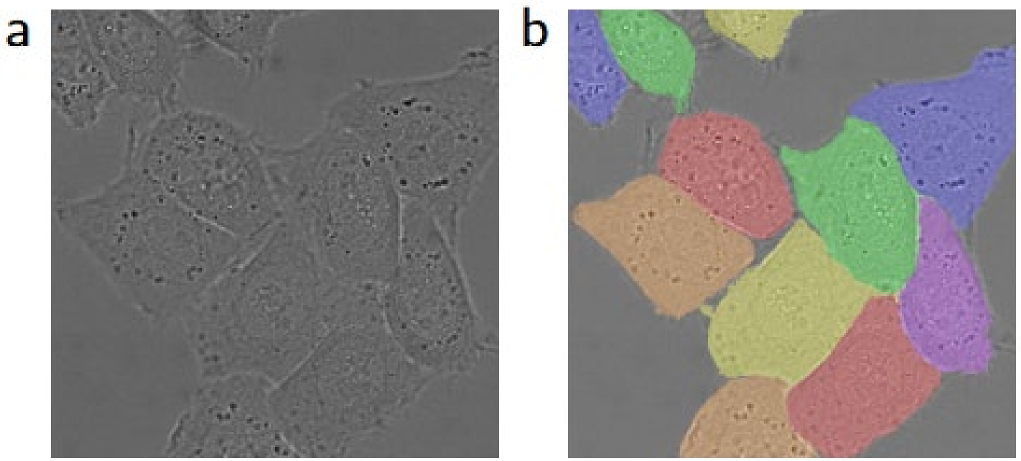




Biological Example: Image Segmentation

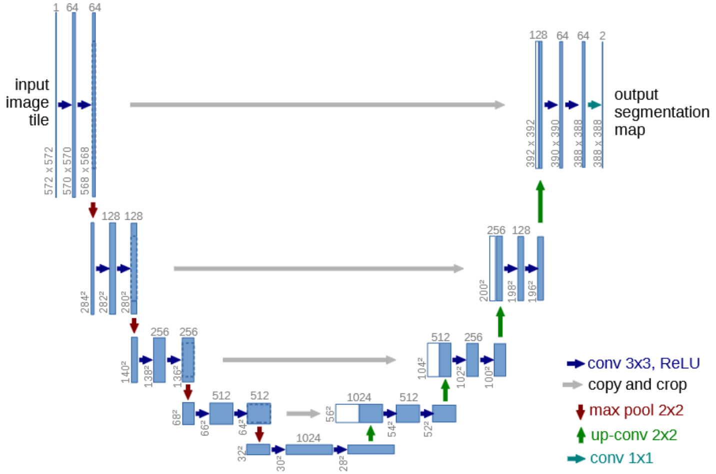
- How does the neural network actually learn?
  - Weights in the network are randomly initialized
  - Loss function compares if each pixel is categorized correctly

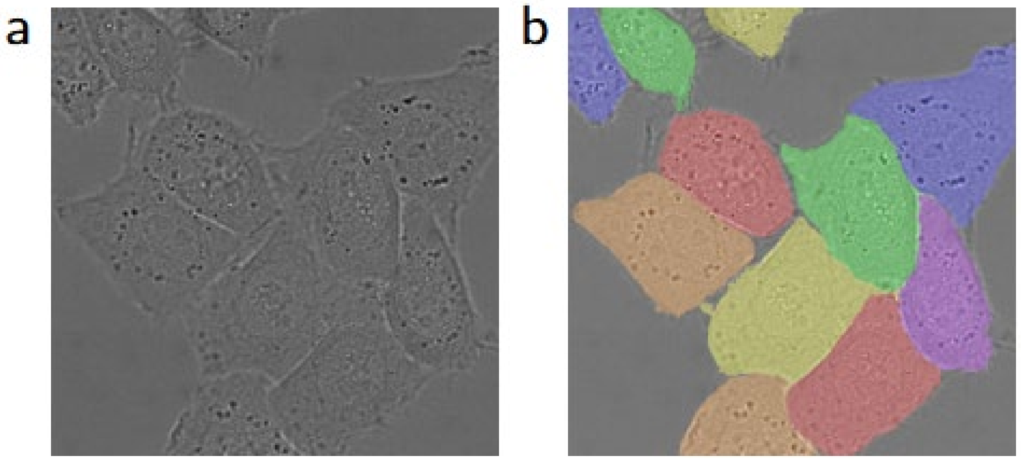




Biological Example: Image Segmentation

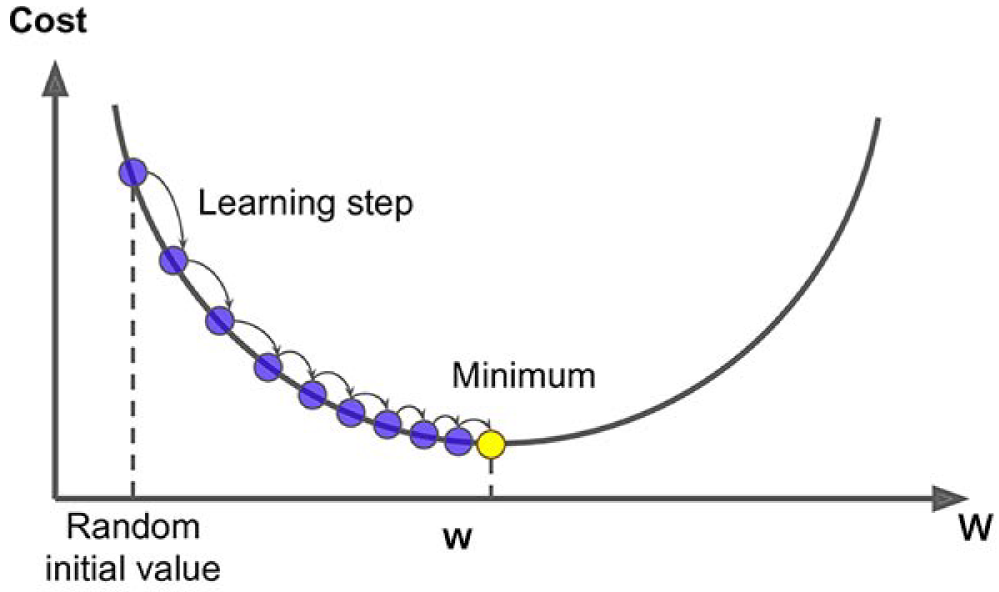
- How does the neural network actually learn?
  - Weights in the network are randomly initialized
  - Loss function compares if each pixel is categorized correctly
  - Gradient descent is used to move the weights in the network towards the correct values





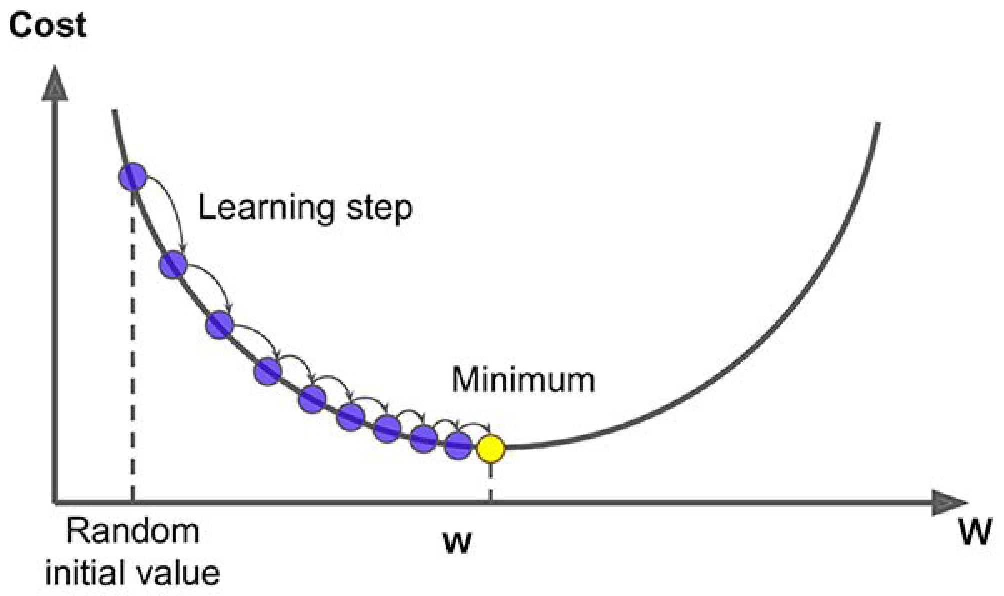
Biological Example: Image Segmentation

- How does the neural network actually learn?
  - Weights in the network are randomly initialized
  - Loss function compares if each pixel is categorized correctly
  - Gradient descent is used to move the weights in the network towards the correct values



Biological Example: Image Segmentation

- How does the neural network actually learn?
  - Weights in the network are randomly initialized
  - Loss function compares if each pixel is categorized correctly
  - Gradient descent is used to move the weights in the network towards the correct values



Biological Example: Image Segmentation