

We are going to save some time!

Course Information

- [BILD 62 Syllabus \(WI25\)](#) ✓
- [BILD 62 Incoming Survey](#)
- [BILD 62 MAGIC MATERIALS LINK](#) ✓
- [BILD 62 Mid-Quarter Feedback Survey](#)

ML Class Activity

- [brain_tumor_model.h5](#)

/ BILD62_WI25 / ML Class Activity /

- brain_tumor_classification_activity.ipynb
- brain_tumor_utils.py
- README.md
- training_history.csv
- training_history.pkl

/ BILD62_WI25 / ML Class Activity /

Name	Modified	File Size
brain_tumor_classification_act...	7 minutes ago	646.8 KB
brain_tumor_model.h5	10 seconds ago	273.3 MB
brain_tumor_utils.py	16 minutes ago	5.9 KB
README.md	16 minutes ago	1 B
training_history.csv	16 minutes ago	1.6 KB
training_history.pkl	16 minutes ago	862 B

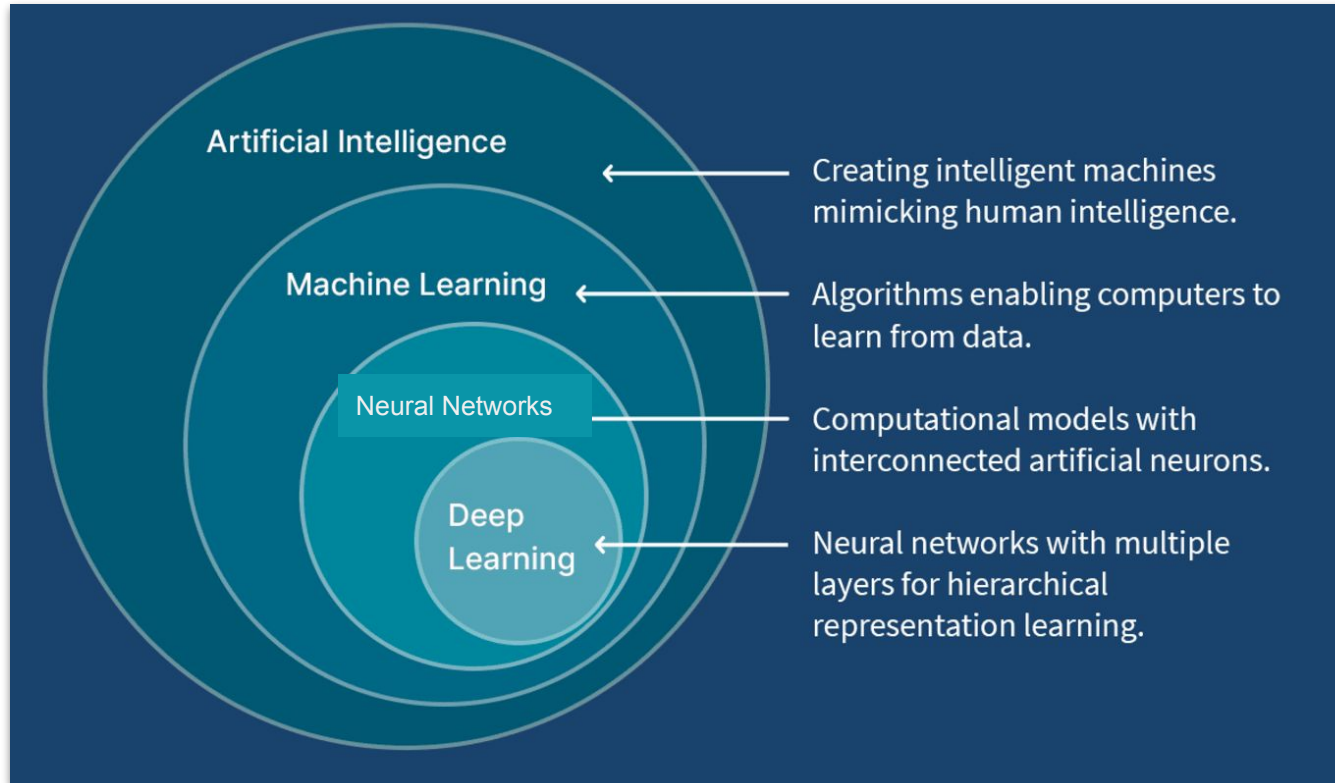
Introduction to Machine Learning

BILD 62 - Introduction to Python for Biologists

Artificial Intelligence in real-world!



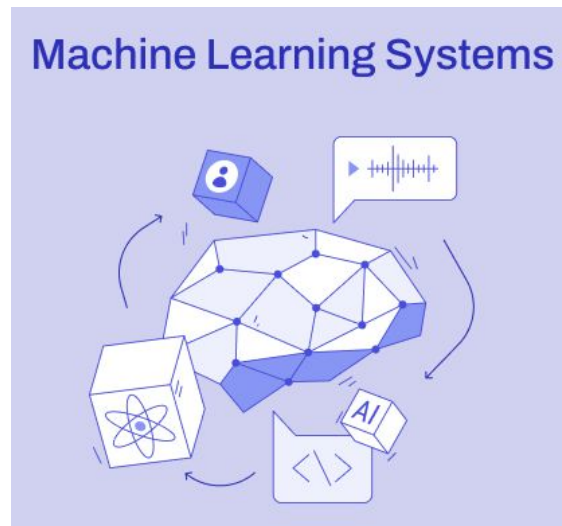
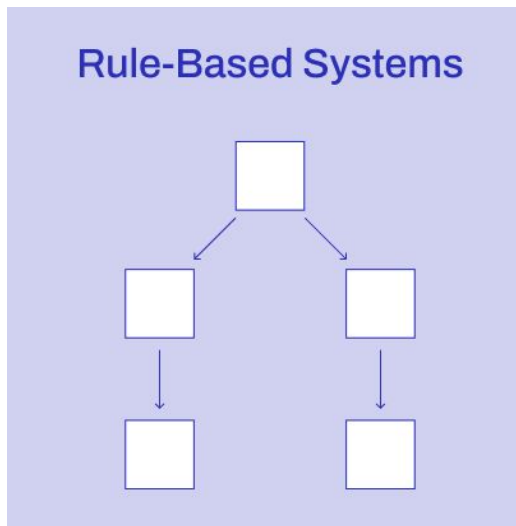
Artificial Intelligence vs Machine Learning



What is Machine Learning?

Definition: Machine Learning is a way for computers to **learn patterns from data** without being explicitly programmed.

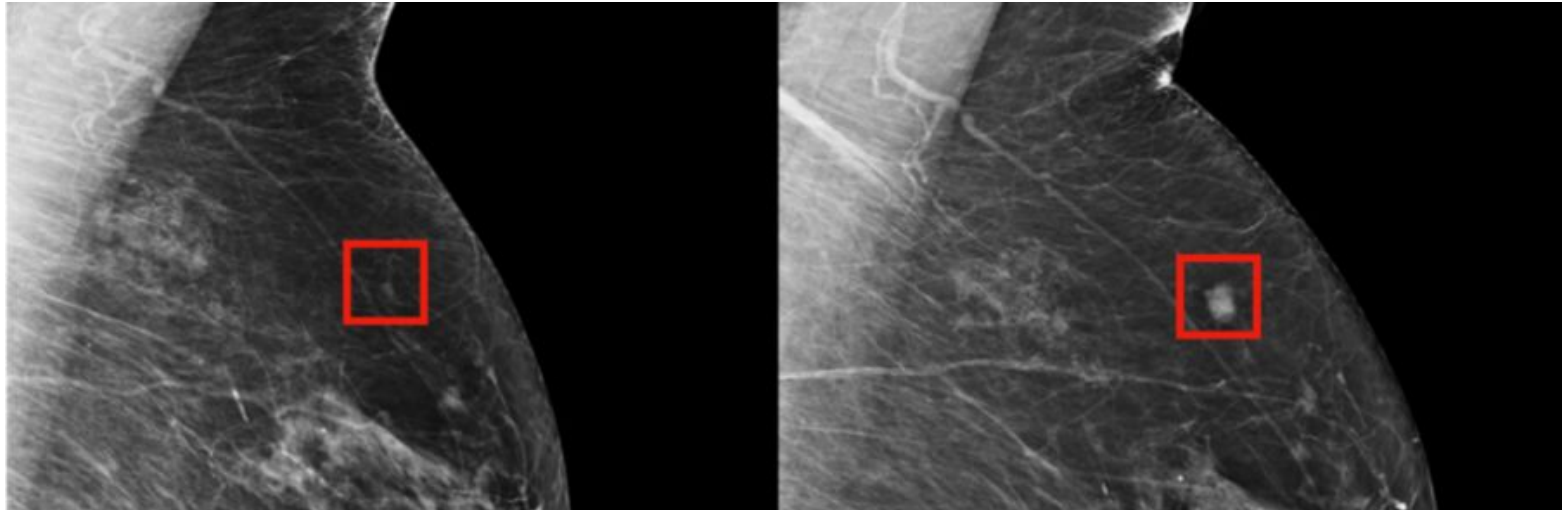
Simple Analogy: Like a child learning to recognize animals—seeing many examples helps them learn.



Why we need machines to learn?

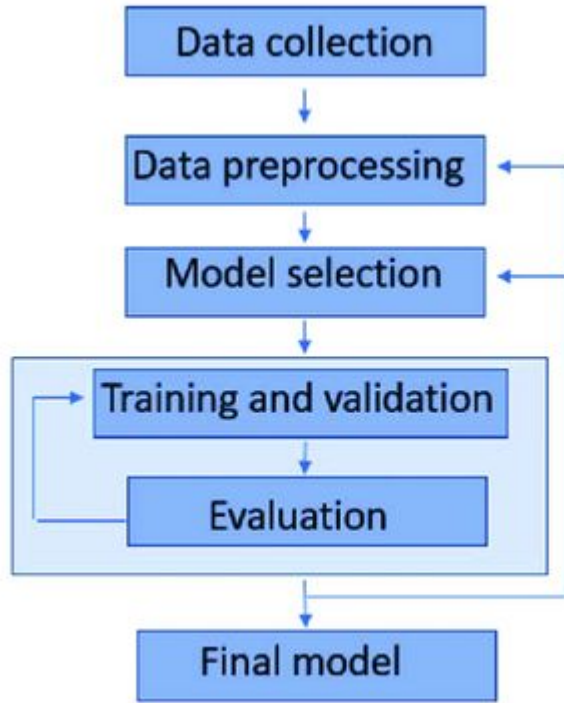
- 1) Automation - Saves Time & Effort ⌚
- 2) Some Problems Are Too Complex for Humans 🧠
- 3) Machines Can Handle Big Data 📊

Detecting risk of cancer development within 5 years



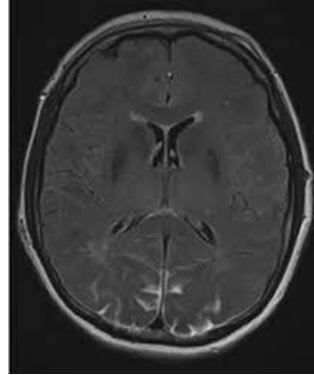
Yala et al., 2019, Radiology

How can machine learn? The Workflow

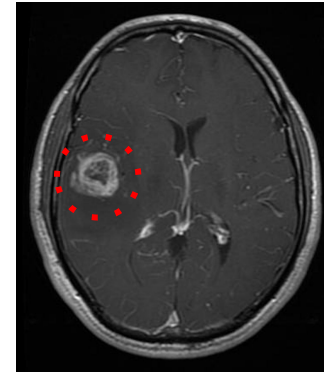


Task:

No-tumor



Tumor

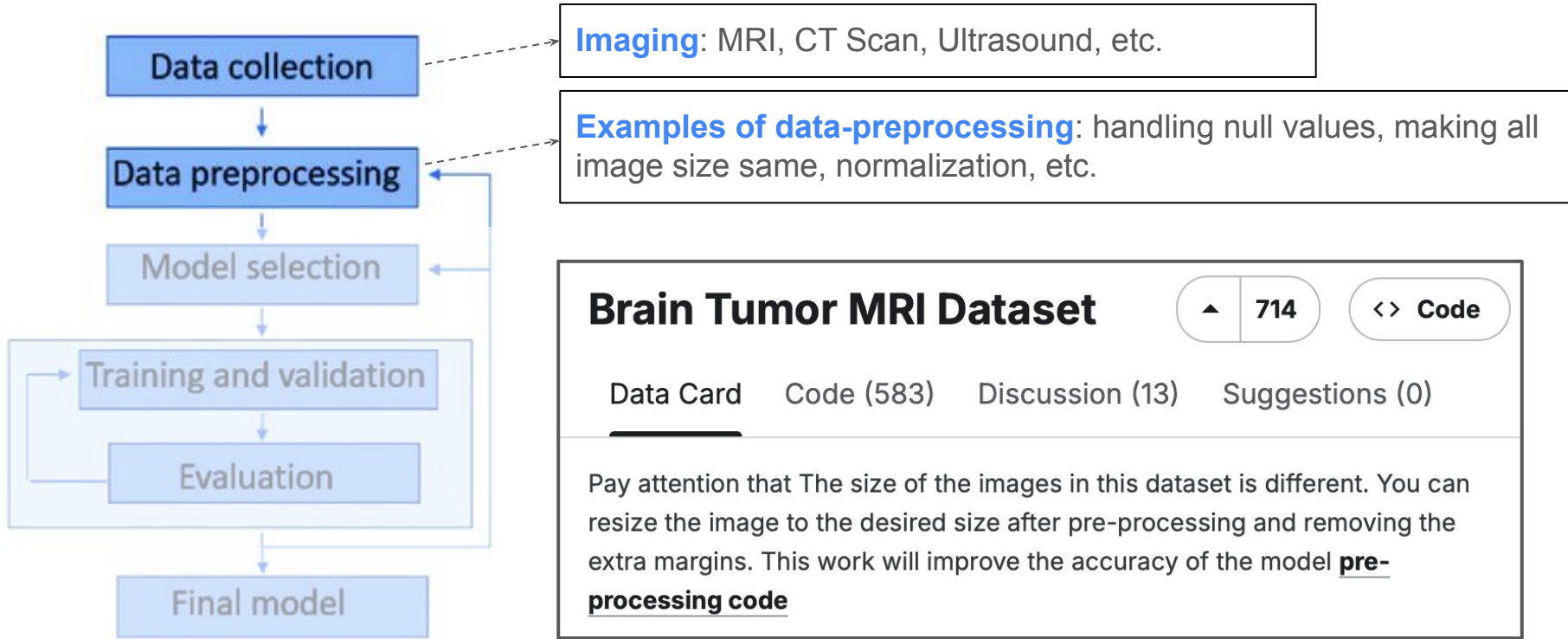


Distinguish between tumor and no-tumor given a patient brain MRI images.

Dataset: [Brain Tumor MRI Dataset](#)

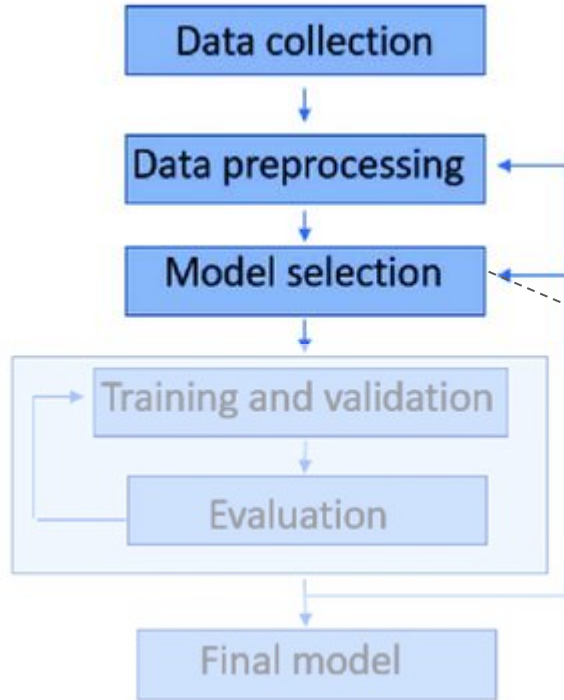
The Workflow

Task: Distinguish between tumor and no-tumor given a patient brain MRI images.



The Workflow

Task: Distinguish between tumor and no-tumor given a patient brain MRI images.



Classic models (e.g. Regression, Decision Trees, Support Vector Machines (SVM), etc.)

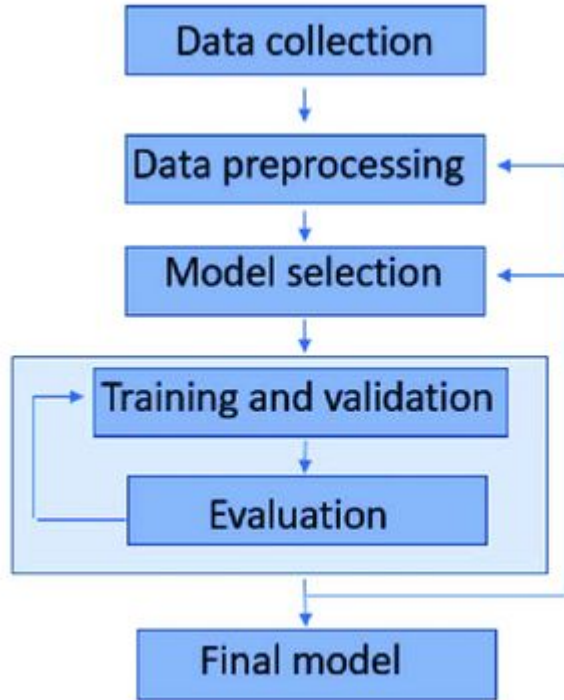
Deep learning models (e.g. Deep Neural Network, Convolution Neural Network, Recurrent Neural Network, Transformer, Graph Neural Network, Autoencoder, Generative Adversarial Network, etc.)

How to choose a mode?

No simple guide! Depends on your data, task, resources you have, etc. Sometime you try out different models and see which one works the best.

The Workflow

Task: Distinguish between tumor and no-tumor given a patient brain MRI images.



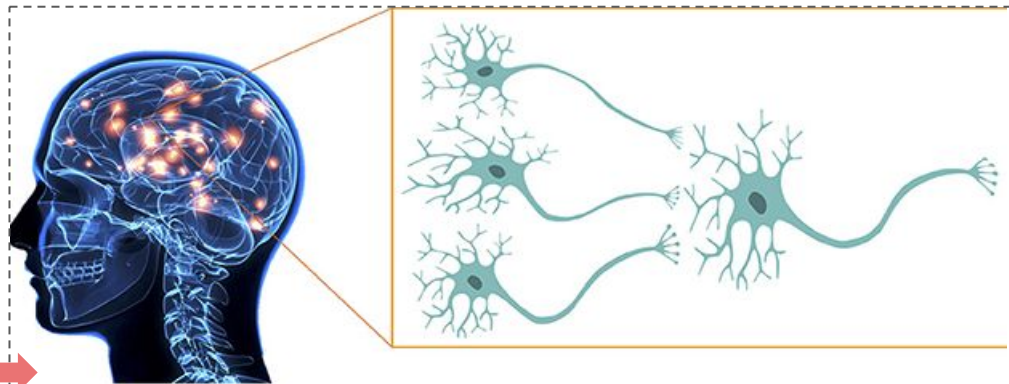
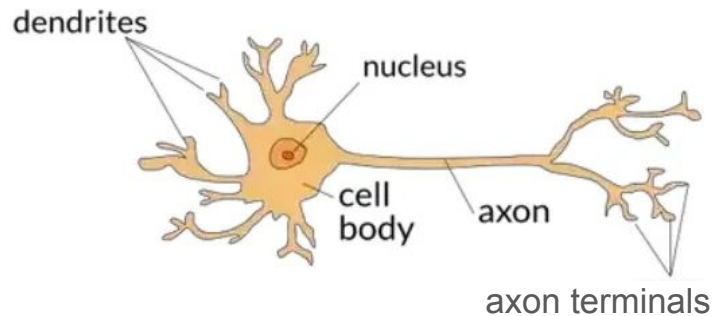
Training Data: Pass the data to the model and optimize the model parameters.

Validation Data: A dataset used during training to assess model performance on unseen data.

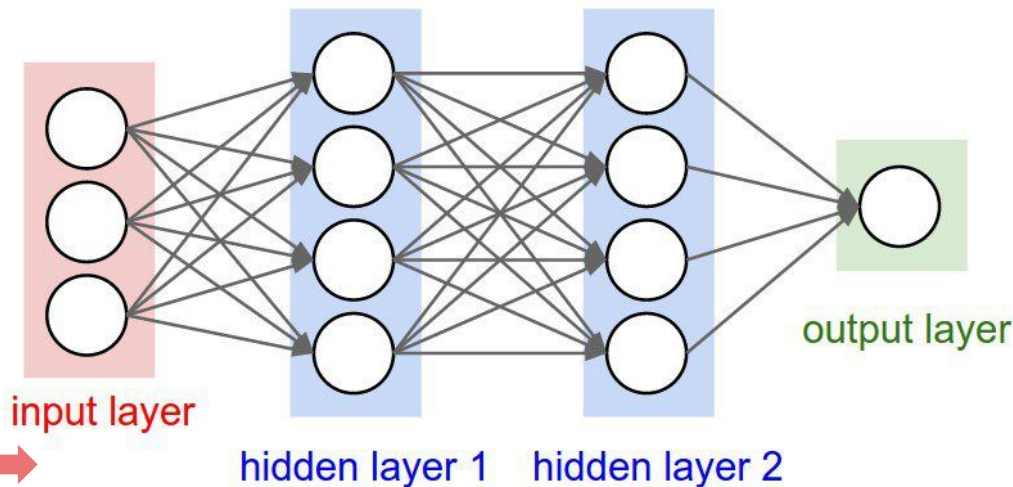
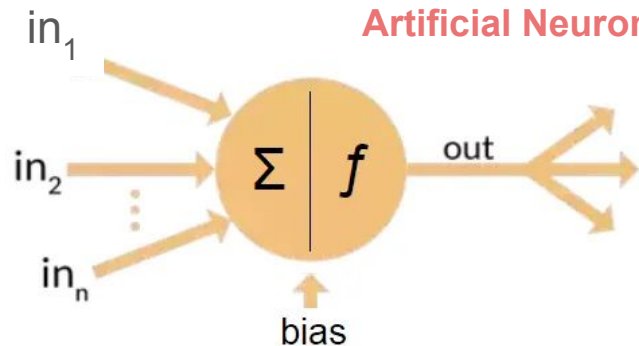
Testing Data: A completely separate dataset used after training to measure the final performance of the model

A Neural Network

Biological Neuron



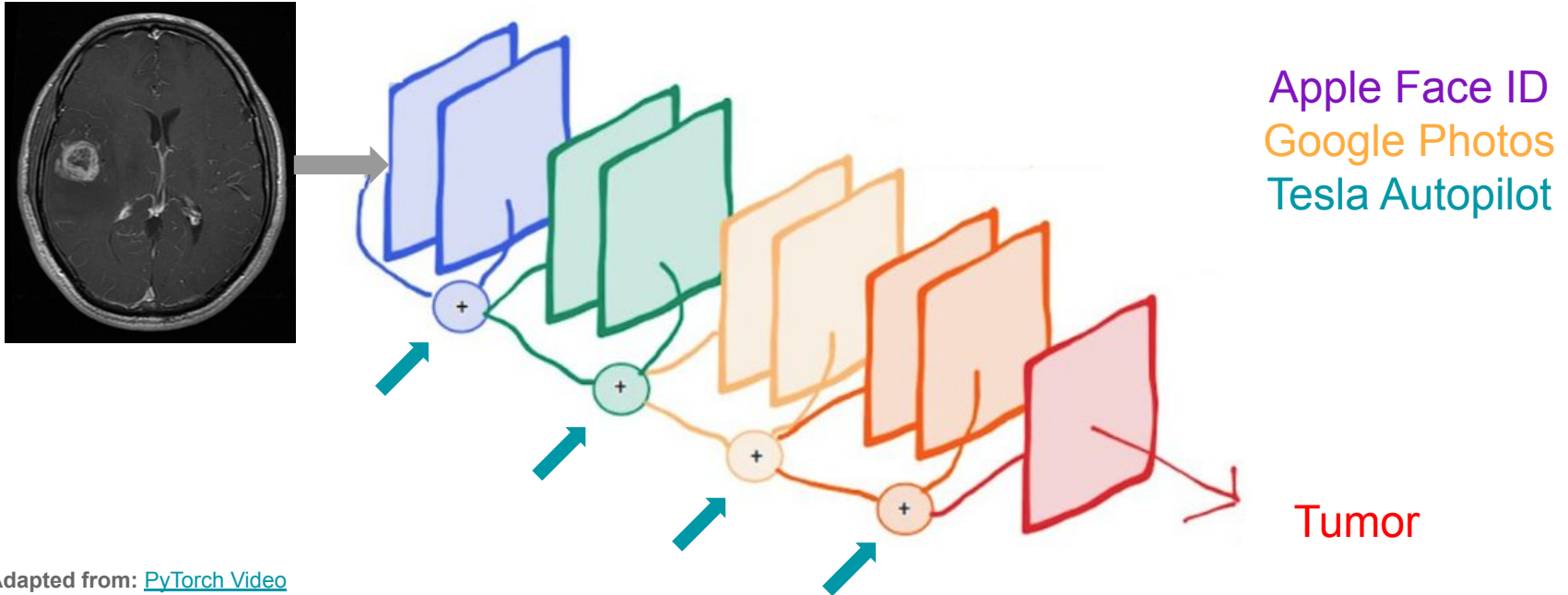
Artificial Neuron



Residual Neural Network (ResNet)

is a type of deep learning model used for image recognition. It solves the problem of **training very deep neural networks by using skip connections (or residual connections)**, which help the network learn better without forgetting important information. This makes it more accurate and efficient for tasks like identifying objects in pictures

ResNet50 simply means it has 50 deep layers



Let's make our first Machine Learning model!

Notes:

1. **DON'T RUSH!** I will slowly walk you through each step.
2. The goal of this activity isn't to master all the machine learning terminology. Instead, it's about seeing how we can use an existing model with just a basic understanding of Python. **Focus on understanding the workflow.**
3. There are **three Class Activities** we will do:
 - a. #1 - visualize the amount of training and testing data we have.
 - b. #2 - visualize the patient MRI images
 - c. #3 - Test the model for new images and check it's accuracy
4. **Model is already trained** for you. It take 12 hours to train the model. So, that part is commented to save time.

Open “brain_tumor_classification_activity.ipynb” on Datahub.

How AI is improving the world?



Automation



Improved Customer Experience



Reduces Errors



Smarter Decision Making



Research and Data Analysis



Improved Business Efficiency



Solves Complex Problems



Manages Repetitive Tasks



Strengthens the Economy

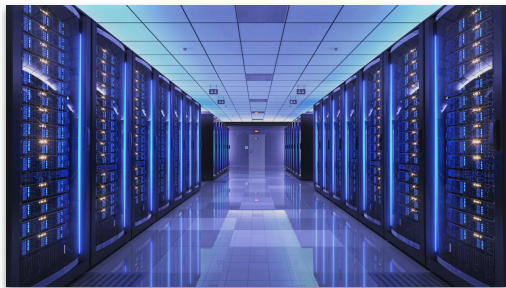


Enhances Lifestyle



Disaster Management

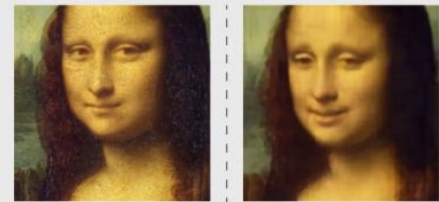
Current Drawbacks of Machine Learning



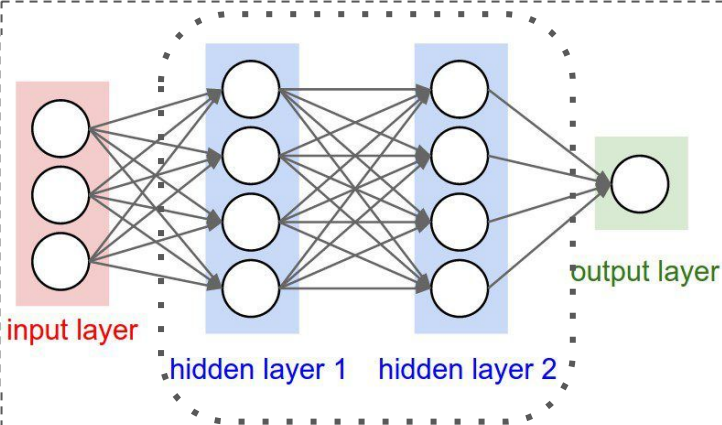
Lot's of Data and Cost



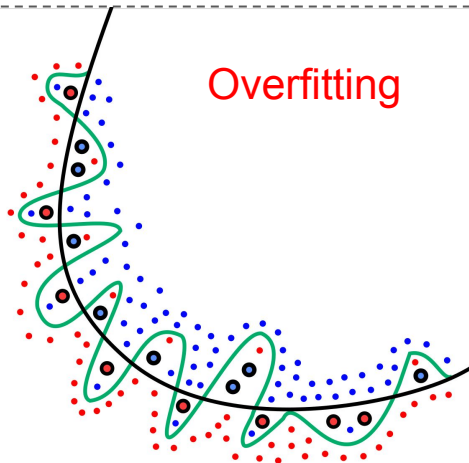
Ethical Bias



Deep Fakes



Interpretability ("black box")



Overfitting

Environmental Cost

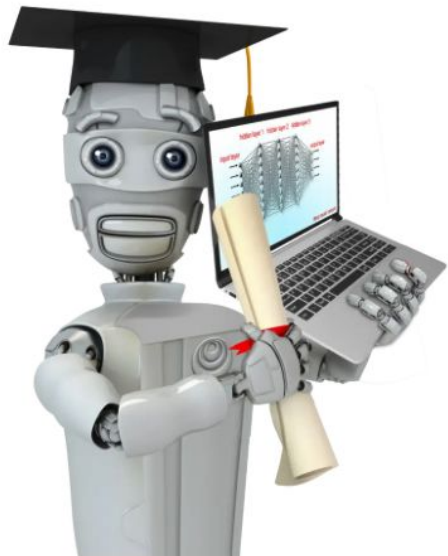
Training a model can generate emission upto 300 round-trip flights between NY and SF ([source](#))

Job Displacement

AI could potentially replace around 800 million jobs worldwide by 2030 ([source](#))

Resources

Machine Learning Specialization



About the original course

2012

Year launched



Rated 4.9 out of 5 by 170K learners

4.8 Million

Learners enrolled

About the instructor



A pioneer in the AI industry, Andrew Ng co-founded Google Brain and Coursera, led AI at Baidu, and has reached and impacted millions of learners with his machine learning courses.