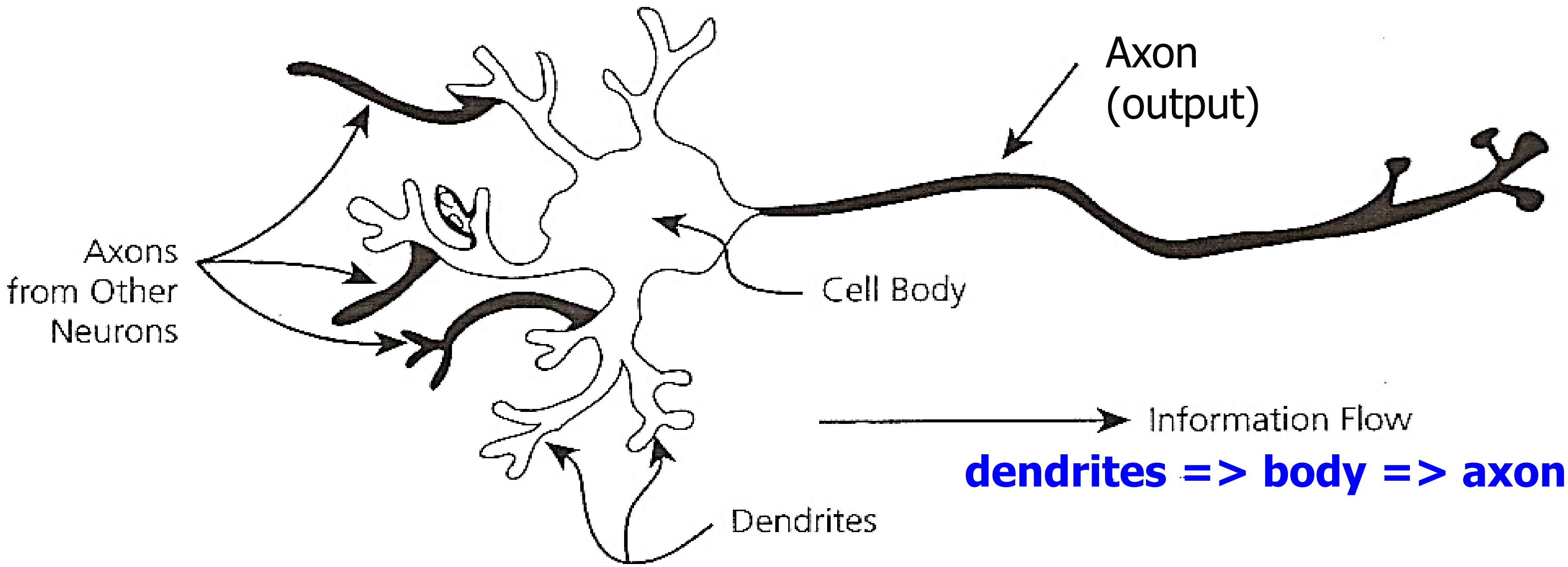
Neural Network Concepts and Paradigms

Biological Neuron

Neurons: nerve cells; consist of dendrites, body and an axon; signals flow through *synapses*.

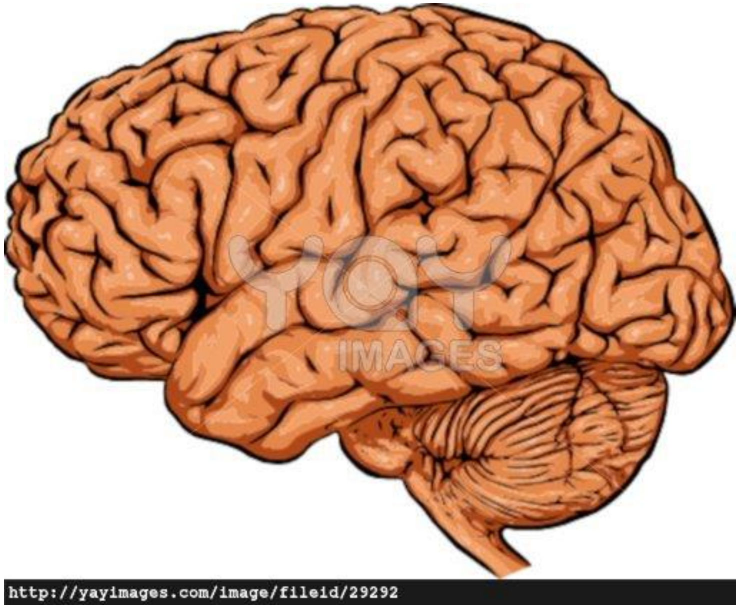
an electrically excitable cell that processes and transmits information through electrical and chemical signals.



dendrites => body => axon

- ➢ Axons connects to dendrites via synapses
- ➢ Synapses
  - ✓ vary in **strength**
  - ✓ may be **excitatory or inhibitory (+ or -)**

Human Brain



On average, the human brain has

- on average 86 billions neurons
- ~1000 main modules /~500 neural network each /~100 000 neurons

Properties of the brain:

- It can learn, reorganize itself - **from experience**
- It adapts to the environment
- It is robust and fault tolerant

Artificial Neural Networks (ANN)

Applies only for conventional (standard) ANN, not for CNN or RNN

❖ An analysis paradigm very roughly modeled after the massively parallel structure of the brain.

❖ Simulates a highly interconnected, parallel computational structure with **numerous relatively simple** individual ***processing elements*** **(*PE*)**, (artificial neuron)

❖ Some differences between biological neurons and artificial neurons (processing elements - PE):

- Signs of weights (+ or -)
- Signals are *ac* in biological neurons, *dc* in *PE*s
- Many types of neurons in a system; usually only a few at most in ANN
- Basic cycle time for PC (<1 ns) faster than brain (10-100ms) {as far as we know!}

What ANNs are and why they are useful

- In the simplest terms, ANNs map input vectors to output vectors

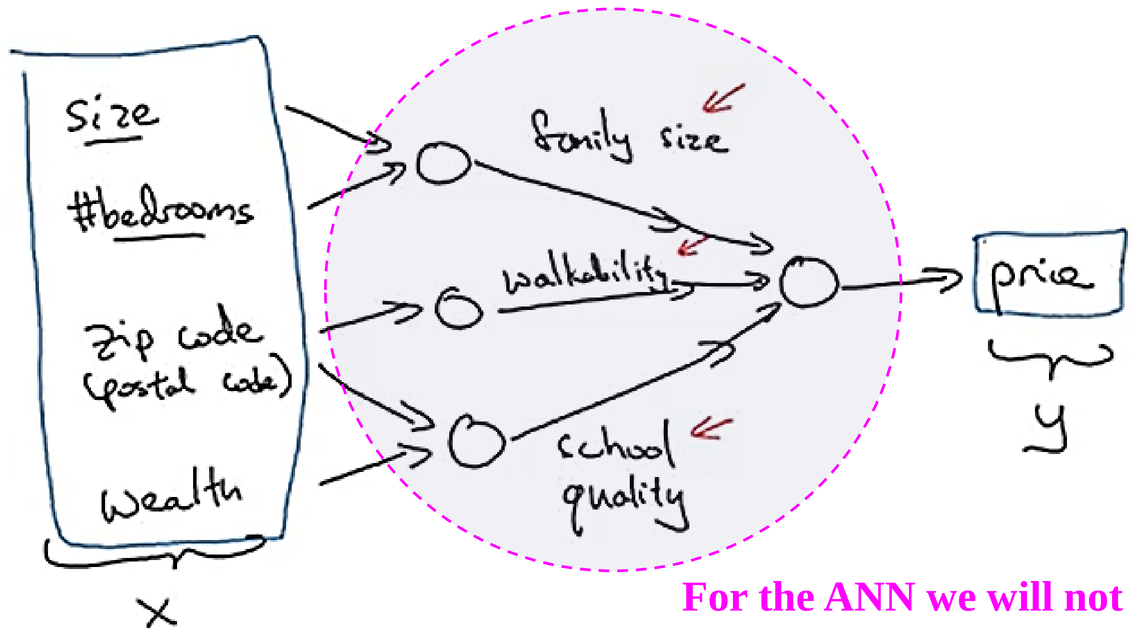
- ANNs consist of **processing elements (*PEs*) or neurons** and **weighted connections**

- Operations performed by ANNs can include:
  * **Fitting** (approximation) and **modeling**
  * **Pattern recognition (Classification)**
  * **Prediction**
  * Pattern matching and completion
  * Noise removal
  * Optimization
  * Simulation
  * Control

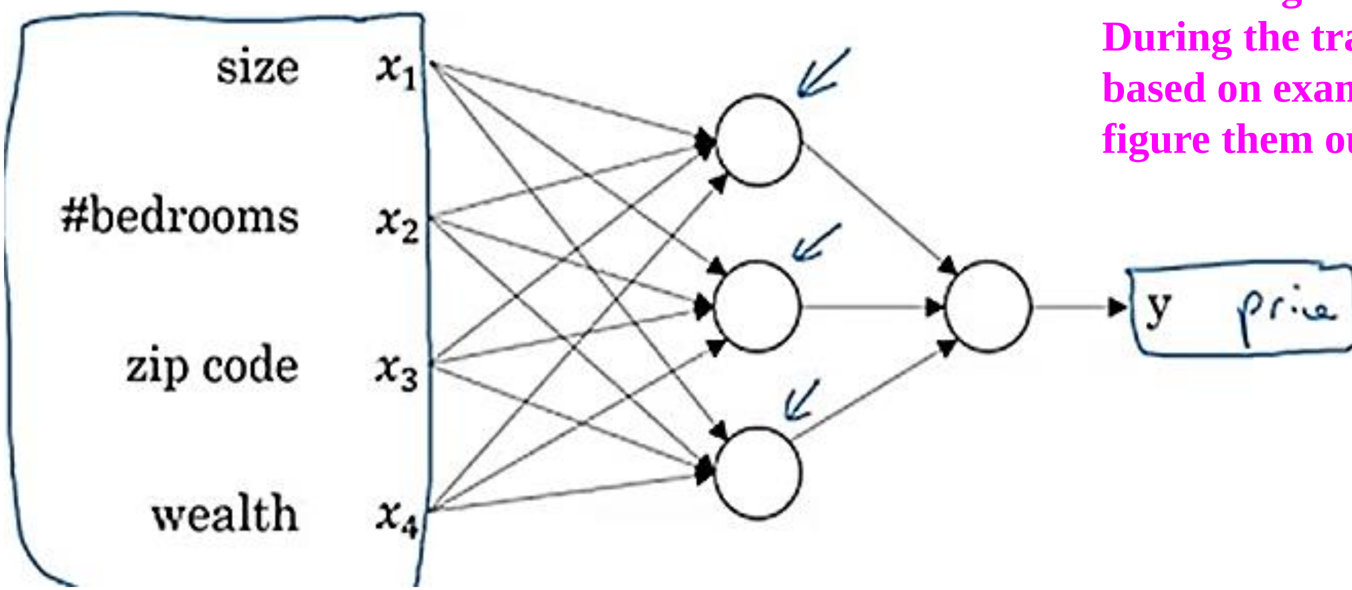
What NNs are and why they are useful

- Connection weights **store the information**
- Weight values usually determined by a **learning** procedure
- Each ***PE*** acts **independently** of all others
- Each ***PE*** relies only on **local** information
- Connection pattern provides **redundancy** and facilitates fault tolerance
- ANNs are able to “learn” arbitrary **non-linear** mappings

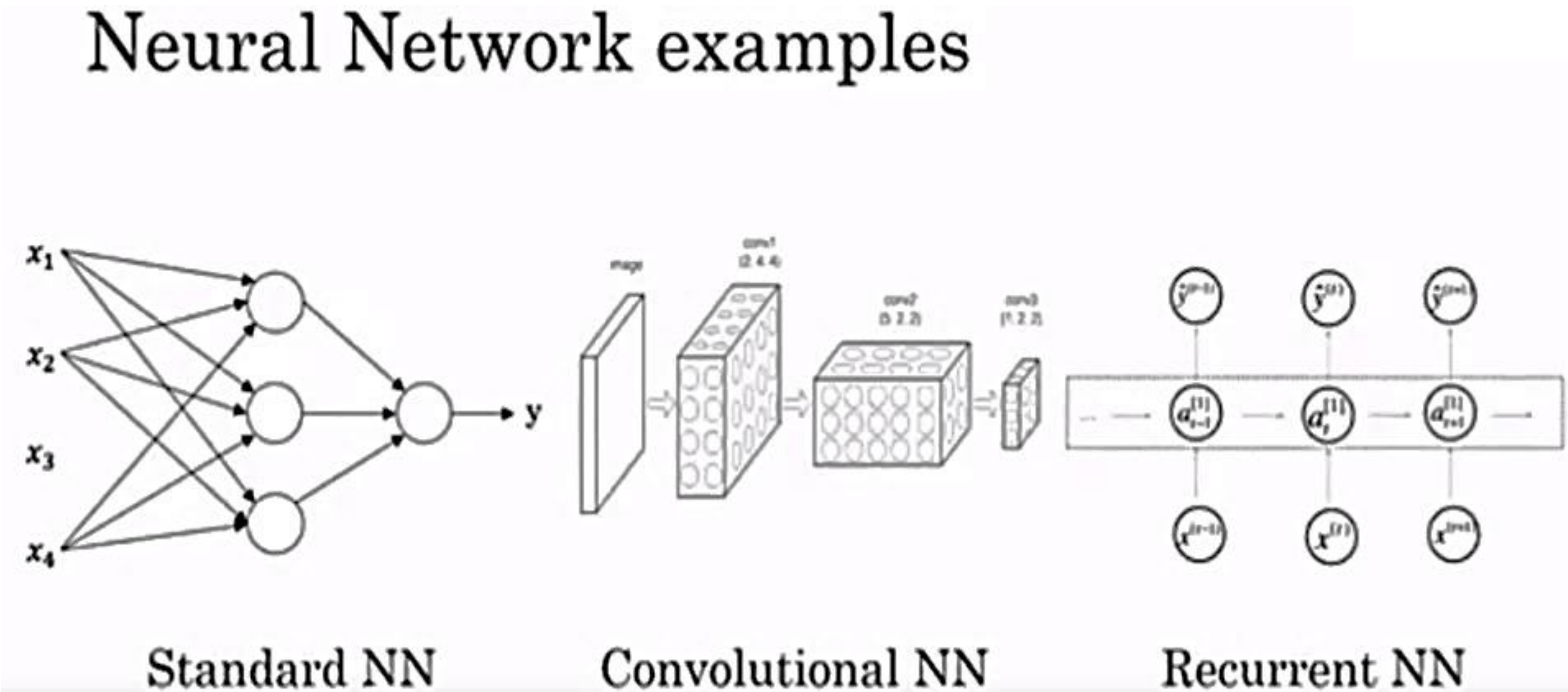
Housing Price Prediction



For the ANN we will not provide those things in the middle. During the training process, based on examples, the ANN will figure them out by itself.



https://www.coursera.org/learn/neural-networks-deep-learning/lecture/eAE2G/what-is-a-neural-network



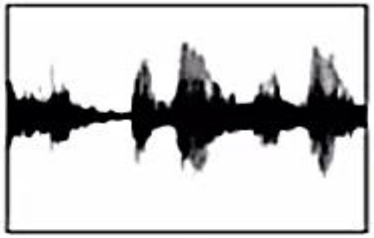
https://www.coursera.org/learn/neural-networks-deep-learning/lecture/2c38r/supervised-learning-with-neural-networks

Structured Data

Size	#bedrooms	…	Price (1000$s)
2104	3		400
1600	3		330
2400	3		369
⋮	⋮		⋮
3000	4		540

User Age	Ad Id	…	Click
41	93242		1
80	93287		0
18	87312		1
⋮	⋮		⋮
27	71244		1

Unstructured Data



Audio

Image

Four scores and seven years ago…

Text

Structured data - each of the features, such as size of the house, the number of bedrooms, or the age of a user, has a very well-defined meaning.

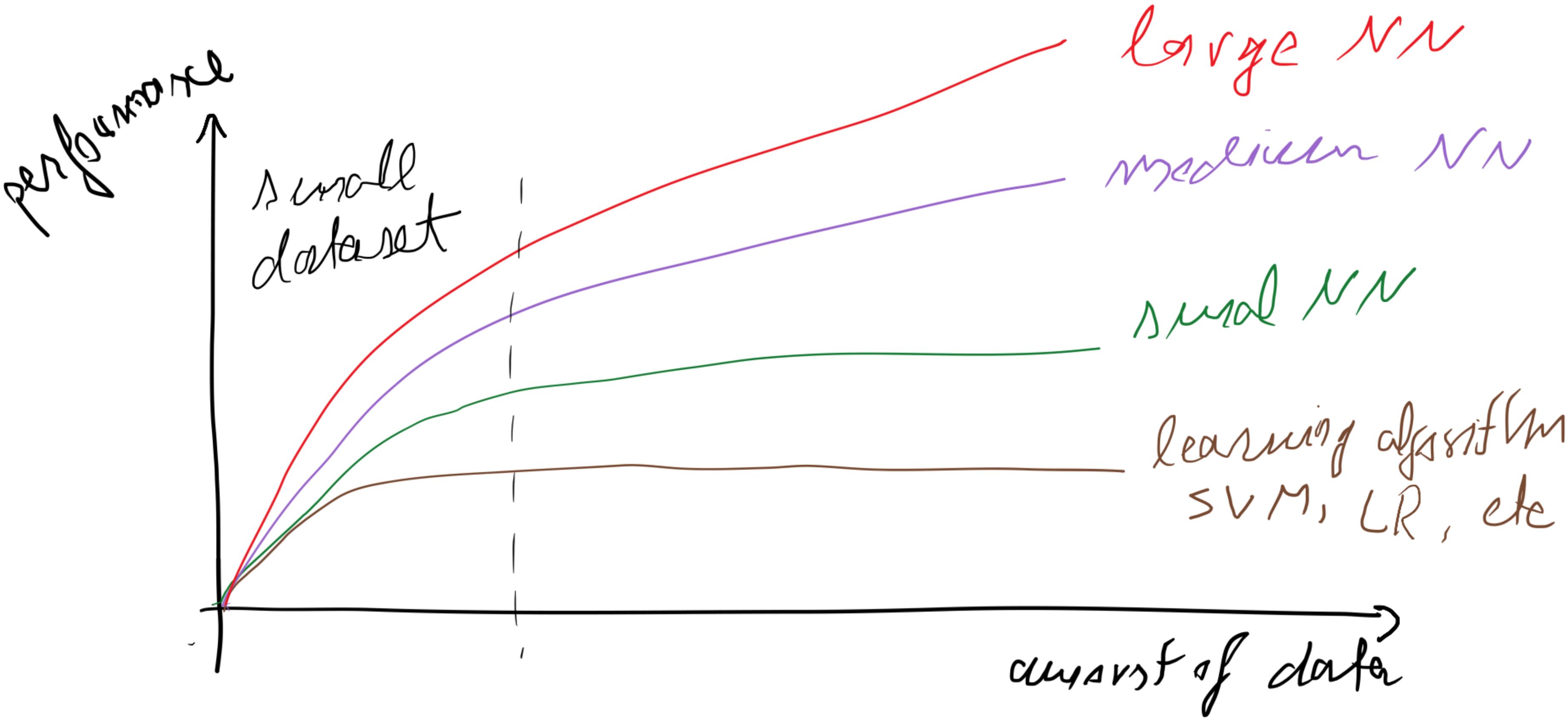
Humans are just really good at interpreting unstructured data.

It has been much harder for computers to make sense of unstructured data compared to structured data.

Thanks to deep learning (neural networks), computers are now much better at interpreting unstructured data.

https://www.coursera.org/learn/neural-networks-deep-learning/lecture/2c38r/supervised-learning-with-neural-networks

Scale drives deep learning progress



- **huge data volumes**
- very sophisticated statistical tools
- **efficient algorithms**
- **gigantic computational power**
- the transformation of our habitats into increasingly AI-friendly places

The gigantic compute power also means carbon footprint, which should imperatively be taken into consideration by the AI community.

Although the training process is a pre-market step, testing different training strategies, neural architectures and datasets solicit the availability of such high-performance machines.

A current focal point of the deep learning research community is the data-oriented neural architecture (Ng, 2021). This means that given a neural model, more effort should be allocated towards finding the optimal data subset or data preprocessing which would increase the accuracy of the final solution.

Instead of adding more complex layers to the neural networks, which would, in turn, require more training data, we will look into optimising the data selection and curation processes.

The novelty of the approach lies in **selecting a few top neural architectures** and evaluating their data requirements, as well as their data specificity, or centricity.