Deep Learning

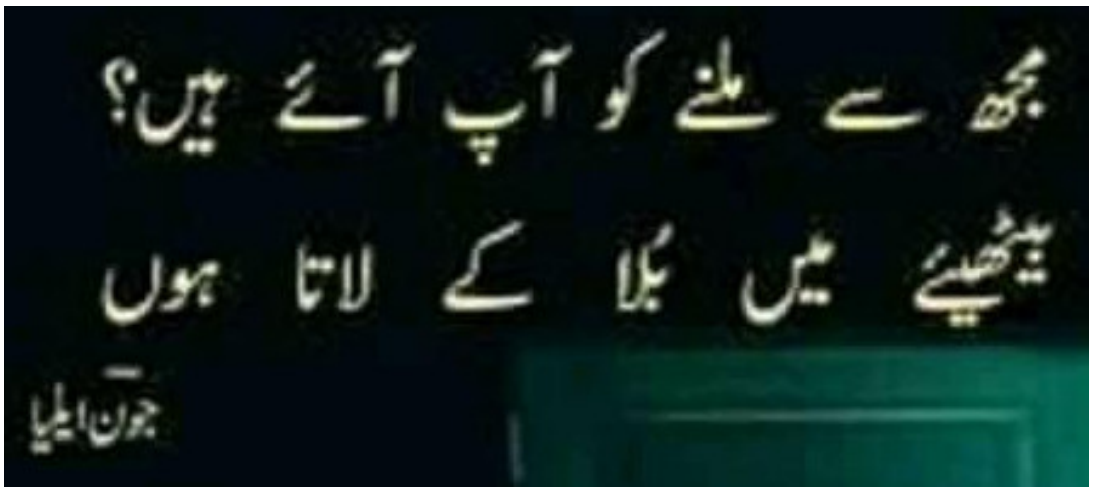
Syed Irtaza Muzaffar

Introduction to Deep Learning

Intelligence

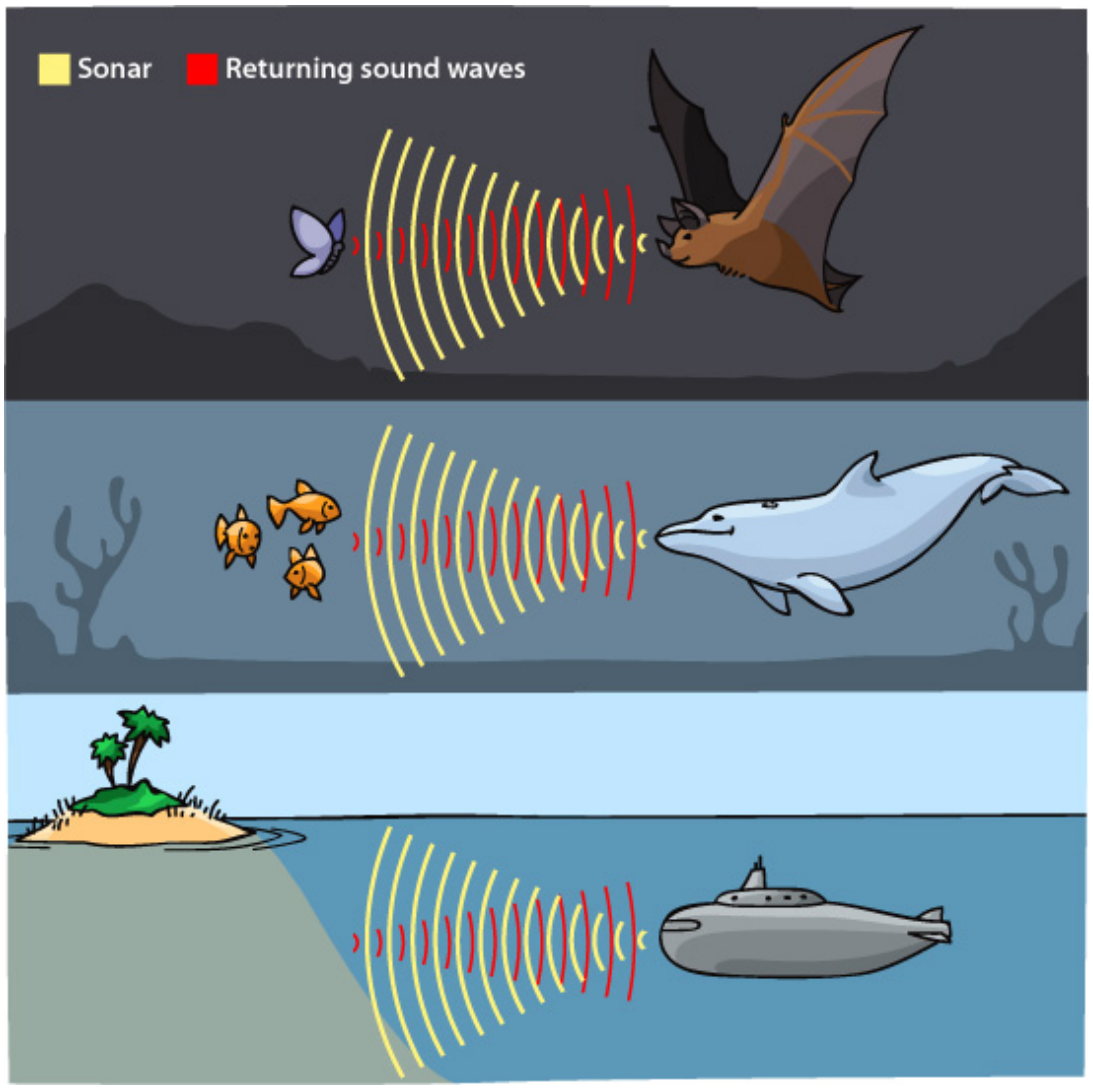
Moravec's paradox: The most difficult things to teach a computer are the ones that a two-year old has already learned – talking, listening, seeing, smelling, walking, grasping, memory and recall, thinking.

Adult humans have uniquely human attributes – metaphor, poetry, satire, sarcasm.



Birds can fly through very small holes at full-speed *while fighting each other*.

In complete darkness, bats can locate, identify and catch their *flying* prey by sending, receiving and analysing sound waves.



The ability of biological brains to sense, perceive, analyse and recognise patterns can only be described as stunning.

- They also have the ability to learn from new examples with or without being taught.
- Mankind's understanding of biological brains and how they operate exactly is embarrassingly limited.
- We are clueless regarding the most fundamental questions.
  - What is intelligence? Are you intelligent if you can't make a mistake?
  - Where in our brains does intelligence lie?
  - What is our brain?
  - Are our brains *just* computational devices or do they *do something more*?
  - What is consciousness?

The Brain

The average human brain has *about* **86 billion neurons** (or nerve cells) and **many more neuroglia** (or glial cells) which serve to support and protect the neurons [and *perhaps* even assist in their functionality]. Each neuron *may* be connected to *up to* **10,000 other neurons**, passing signals to each other via as many as **1,000 trillion synaptic connections**, equivalent *by some estimates* to a **computer with a 1 trillion bit per second processor**. Estimates of the human brain's memory capacity *vary wildly* from **1 to 1,000 terabytes** (for comparison, the 19 million volumes in the US Library of Congress represents about 10 terabytes of data).
Source: `https://human-memory.net/brain-neurons-synapses/`

- Claims about the brain are vague.
- We know *something* about the brain, but we do not know *most* of the crucial functioning.

So what is this course about?

- Modelling what we do not understand seems foolish.
- However, there do exist numerous *practical* techniques that give machines the *illusion of being intelligent*.
- This is the domain of artificial intelligence, statistical pattern recognition, machine learning and deep learning.
- Instead of attempting to mimic the complex workings of a biological brain, this course
  - aims at explaining mathematically well-founded techniques for analysing patterns and learning from them, and is therefore
  - a *mathematically involved* introduction into the field of pattern recognition and machine learning.
- It will prepare you for further study/research in machine learning, computer vision, natural language processing and others areas attempting to solve AI type problems.

Prerequisites

- The course is designed to be self-contained.
- *Required mathematical details will be covered in the lectures*.
- However, this is a *math-heavy course*. Students are encouraged[1] to brush up on their knowledge of
  - Probability (Bernoulli, Binomial, Gaussian, Discrete, Continuous)
  - Calculus (Differentiation, Partial derivatives, Chain rule)
  - Linear Algebra (Vectors, Matrices, Dot-product, Orthogonality, Eigenvectors, SVD)
- This is also a *code-heavy* course. Be ready to become good at coding.

[1] ordered

Prerequisites

The only way to benefit from this course is to be prepared to *spend lots of hours reading the slides, text books, tutorials and attempting exercises* preferably alone or with a class-fellow.

Learn to be mostly alone for the next five months.
Learn to be (reasonably) selfish.
Your social life will be adversely affected. It should be! Time to grow up.
You will need to work harder than ever before.
And even that might not be enough!

Assignments

- There will be 5 programming assignments.
- You are encouraged to help each other but not to cheat.
- *Do not cheat*.
- No assignment will be dropped. Everything counts.

Project

- Solve an interesting problem via Deep Learning.
- Recorded video presentation.
- Place code on Github.
- Create web-page containing everything about the project.
- Report prepared in Latex.
- Templates will be provided to help you out.

An attention based method for offline handwritten Urdu text recognition

Tayaba Anjum, Nazar Khan

17th International Conference on Frontiers in Handwriting Recognition (ICFHR 2020)

Abstract

Compared to derivatives from Latin script, recognition of derivatives from Arabic hand-written script is a complex task due to the presence of two-dimensional structure, context-dependent shape of characters, high number of ligatures, overlap of characters, and placement of diacritics. While significant attempts exist for Latin and Arabic scripts, very few attempts have been made for offline, handwritten, Urdu script. In this paper, we introduce a large, annotated dataset of handwritten Urdu sentences. We also present a methodology for the recognition of offline handwritten Urdu text lines. A deep learning based encoder/decoder framework with attention mechanism is used to handle two-dimensional text structure. While existing approaches report only character level accuracy, the proposed model improves on BLSTM-based state-of-the-art by a factor of 2 in terms of character level accuracy and by a factor of 37 in terms of word level accuracy. Incorporation of attention before a recurrent decoding framework helps the model in looking at appropriate locations before classifying the next character and therefore results in a higher word level accuracy.

Presentation Video

Files

Paper Presentation Code Dataset

Bibtex

```
@inproceedings{anjum2020urdu ohtr,
    author = {Anjum, Tayaba and Khan, Nazar},
    title = {An attention based method for offline handwritten Urdu text recognition},
    booktitle = {International Conference on Frontiers in Handwriting Recognition (ICFHR)},
    month = {September},
    year = {2020},
}
```

Acknowledgements

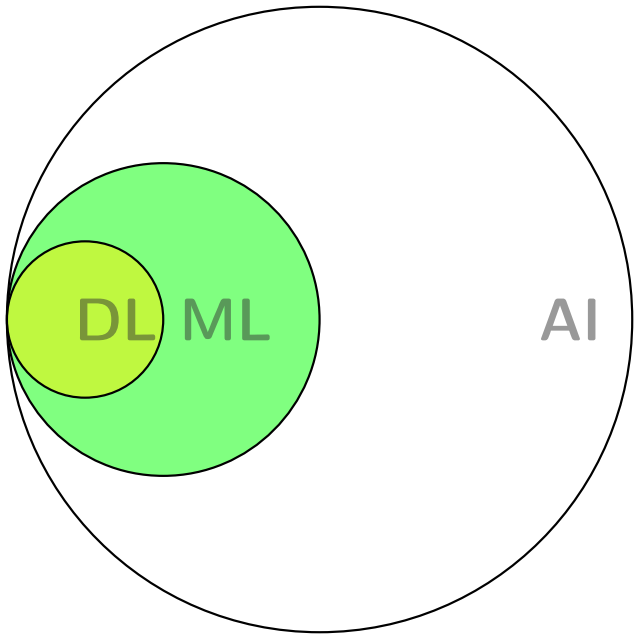
This work was supported by the Higher Education Commission (Pakistan) under Grant 8329/Punjab/NRPU/R&D/HEC/2017. We thank all the undergraduate students at University of Management and Technology (UMT) for agreeing to become scribes for our dataset. We also thank Bilal Rasheed and Faizan Saleem for the tedious process of ground-truth annotations and corrections.

Programming Environment

We will be using

- Python as our programming environment.
- PyTorch as our deep learning framework.
- Google Colaboratory as our GPU enabled machine on the cloud.
- Jupyter notebooks as our interactive tutorials.

Deep Learning vs. Machine Learning vs. AI



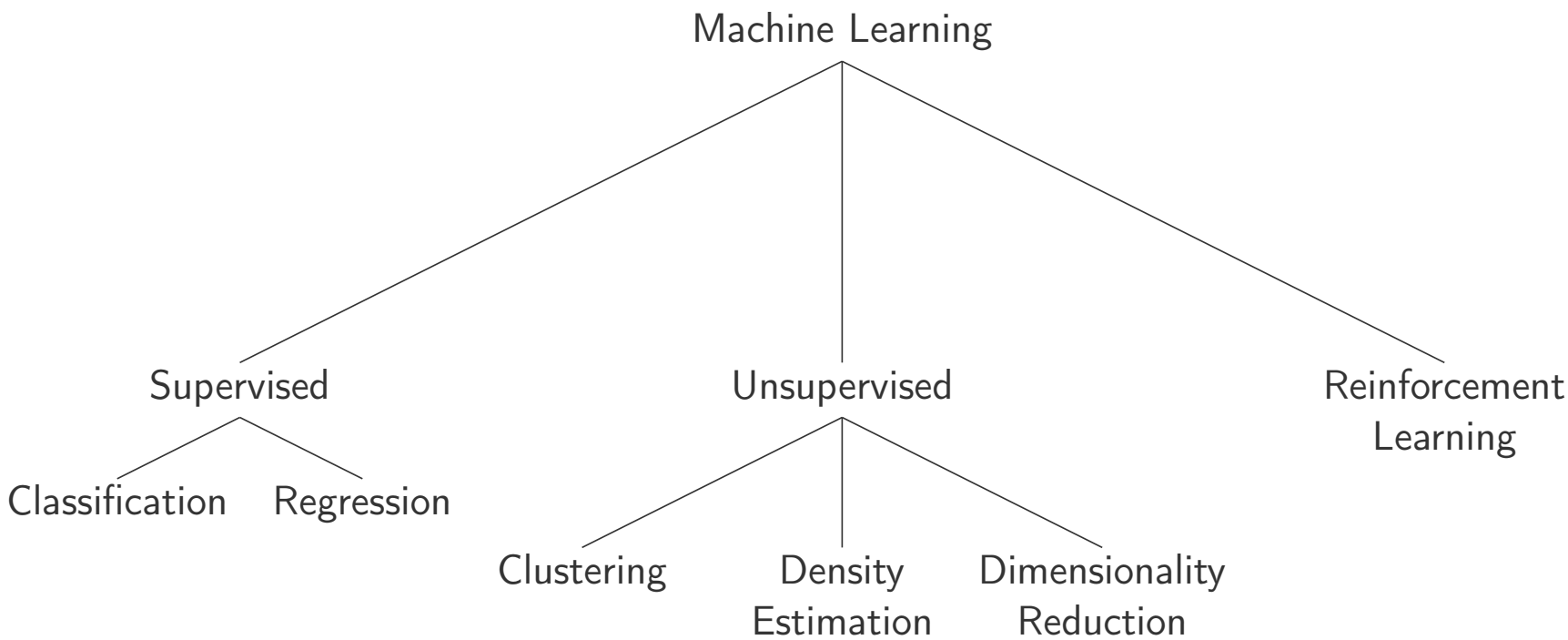
ML and AI problems are increasingly being solved using DL.
Laymen have started considering DL to be the same as AI.

AI: software that solves problems by itself.
ML: algorithms and models that *learn* from processed data.
DL: *neural networks* that *learn better* from *less processed* data.

Introduction

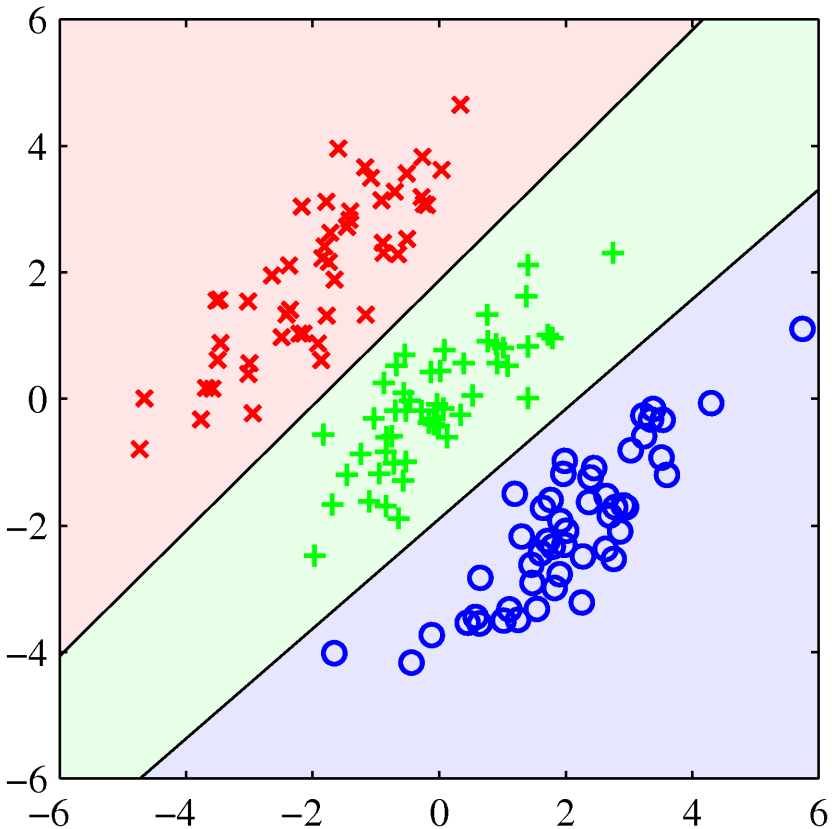
- Machine Learning is concerned with automatic discovery of regularities in data.
- Regularity implies order.
- Learning implies exploiting order to make predictions.

Machine Learning

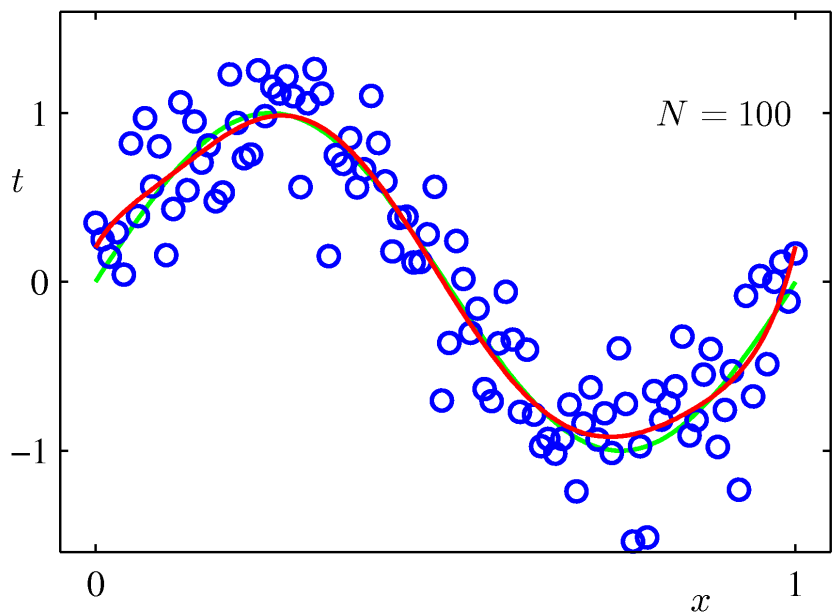


Supervised Learning

- **Classification**: Assign **x** to *discrete* categories.
  - Examples: Digit recognition, face recognition, *etc.*.
- **Regression**: Find *continuous* values for **x**.
  - Examples: Price prediction, profit prediction.



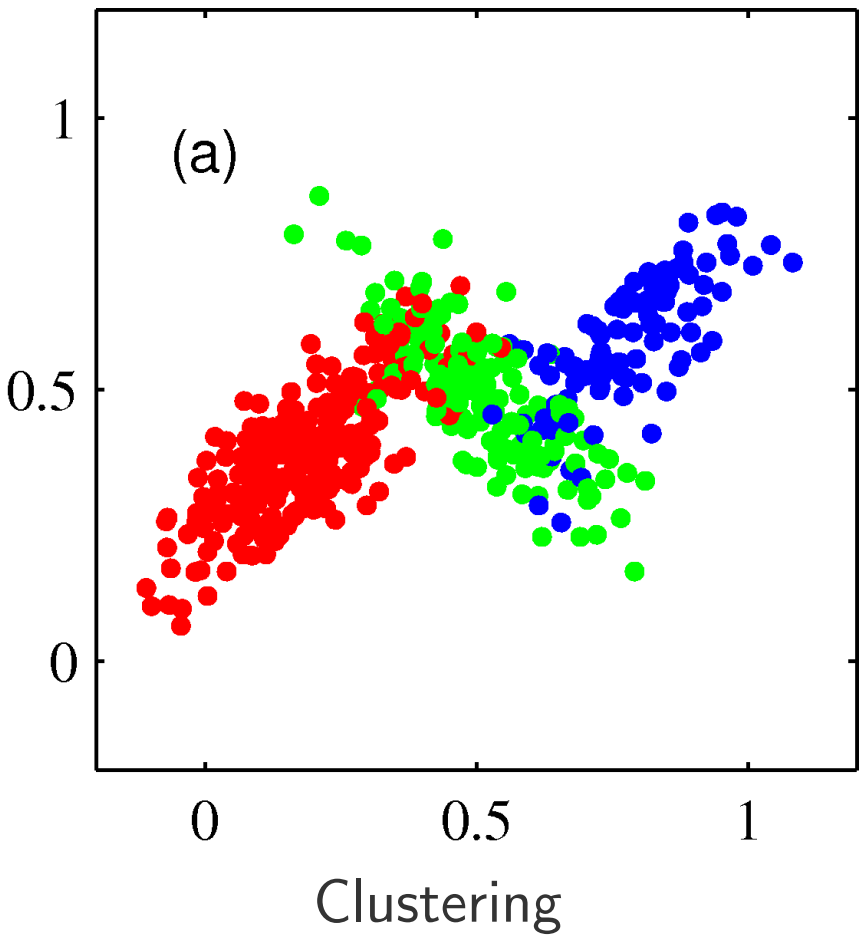
Classification

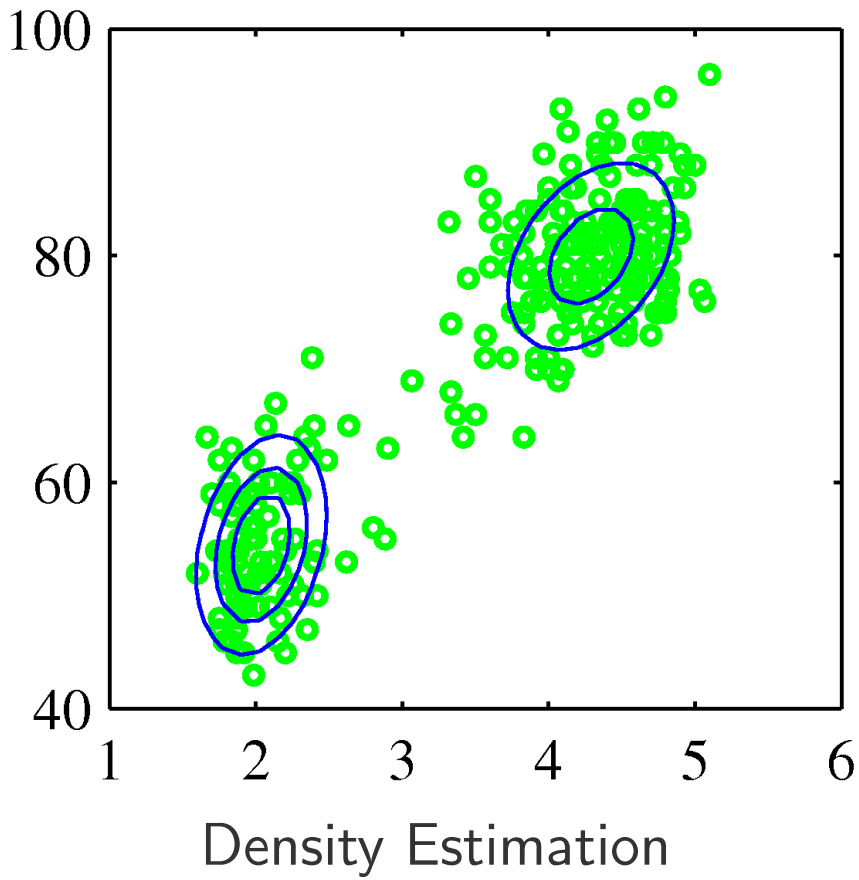


Regression

Unsupervised Learning

- **Clustering**: Discover groups of similar examples.
- **Density Estimation**: Determine probability distribution of data.
- **Dimensionality Reduction**: Map data to a lower dimensional space.



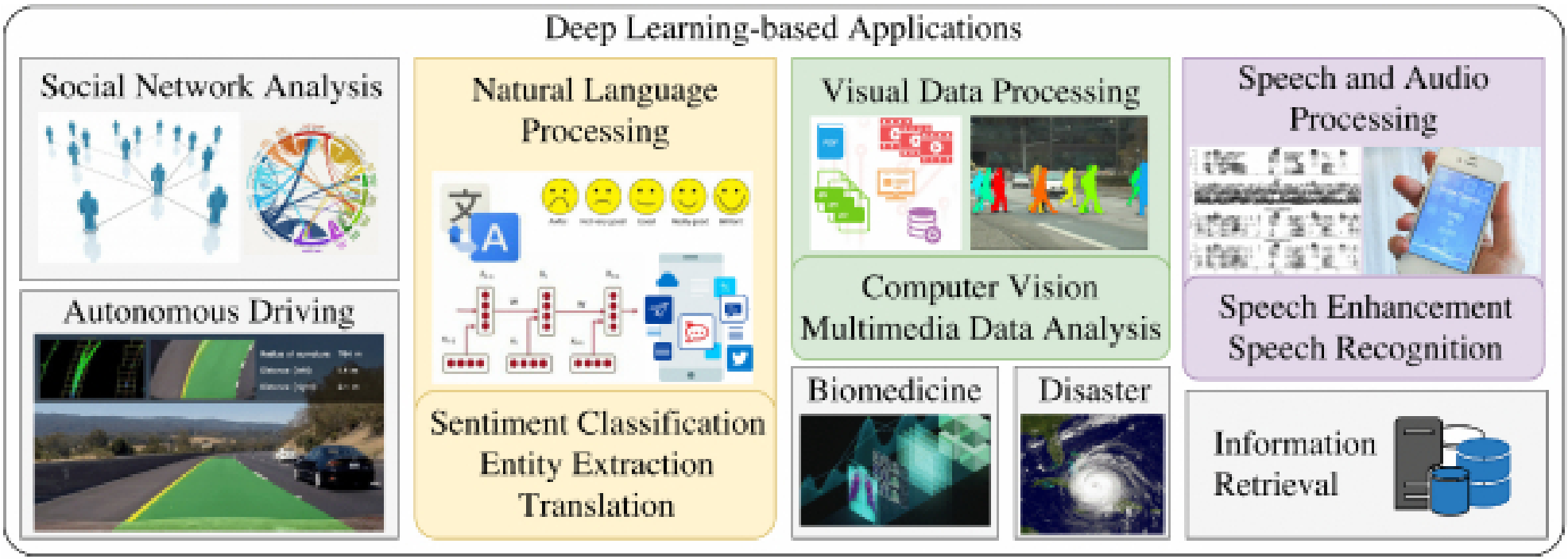


Reinforcement Learning

- Find actions that maximise a reward within an environment.

Figure: Based on the current state of the game (environment), each action of the player changes the state and yields a reward – points or death. The player learns to reinforce taking actions that lead to positive reward and not taking actions that lead to negative reward. Source: https://www.freecodecamp.org/news/a-brief-introduction-to-reinforcement-learning-7799af5840db/

Applications of Deep Learning



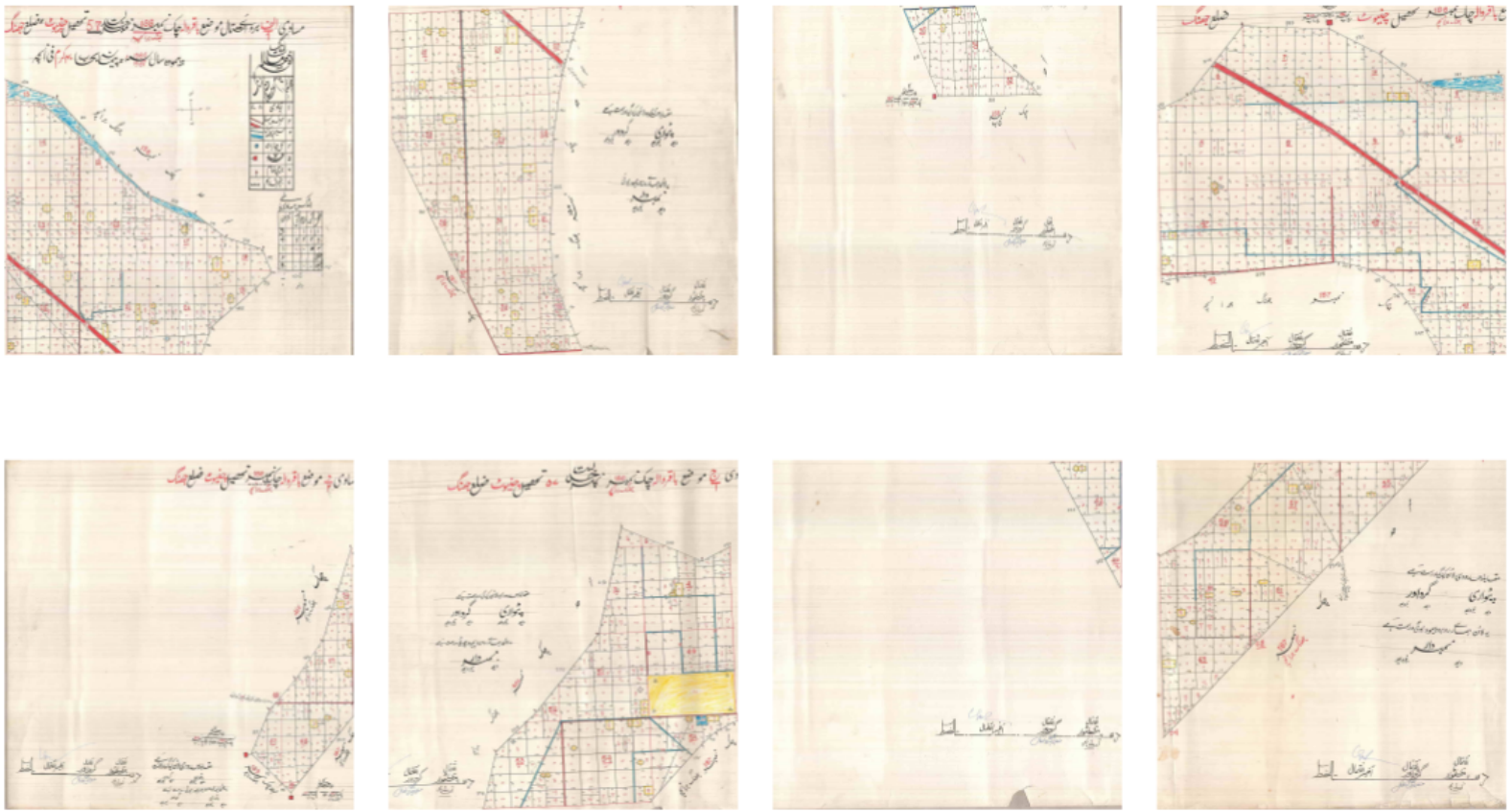
https://doi.org/10.1145/3234150

DL Applications

Recognition of handwritten Urdu text

DL Applications

Learning to solve jigsaw puzzles

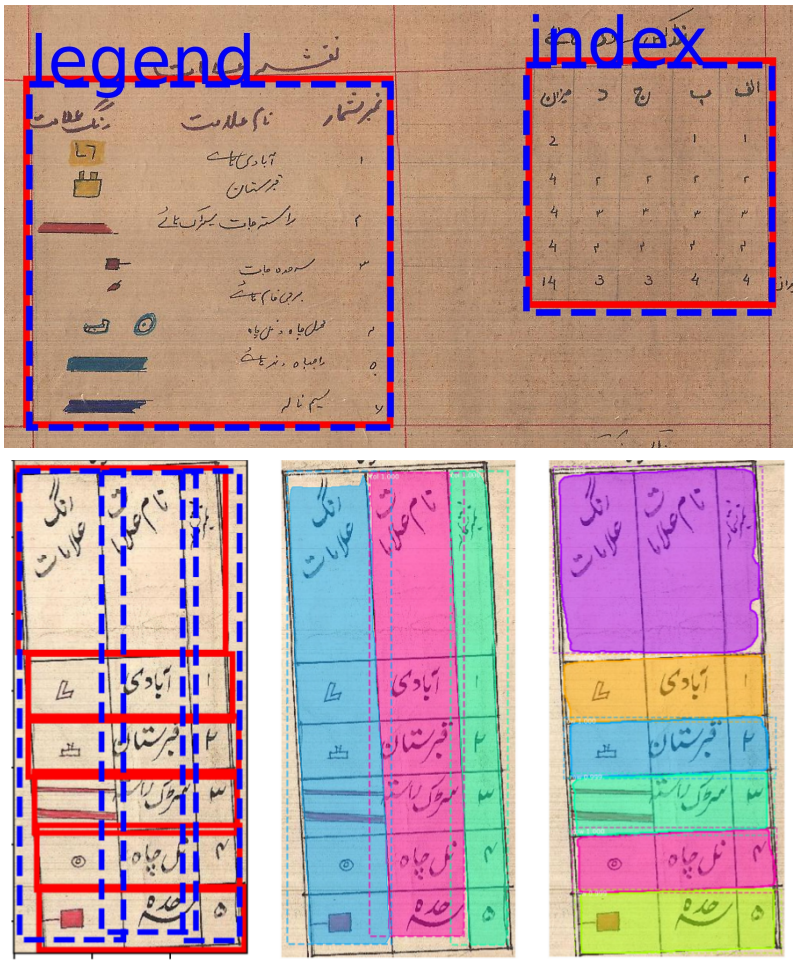


DL Applications

Learning to solve jigsaw puzzles

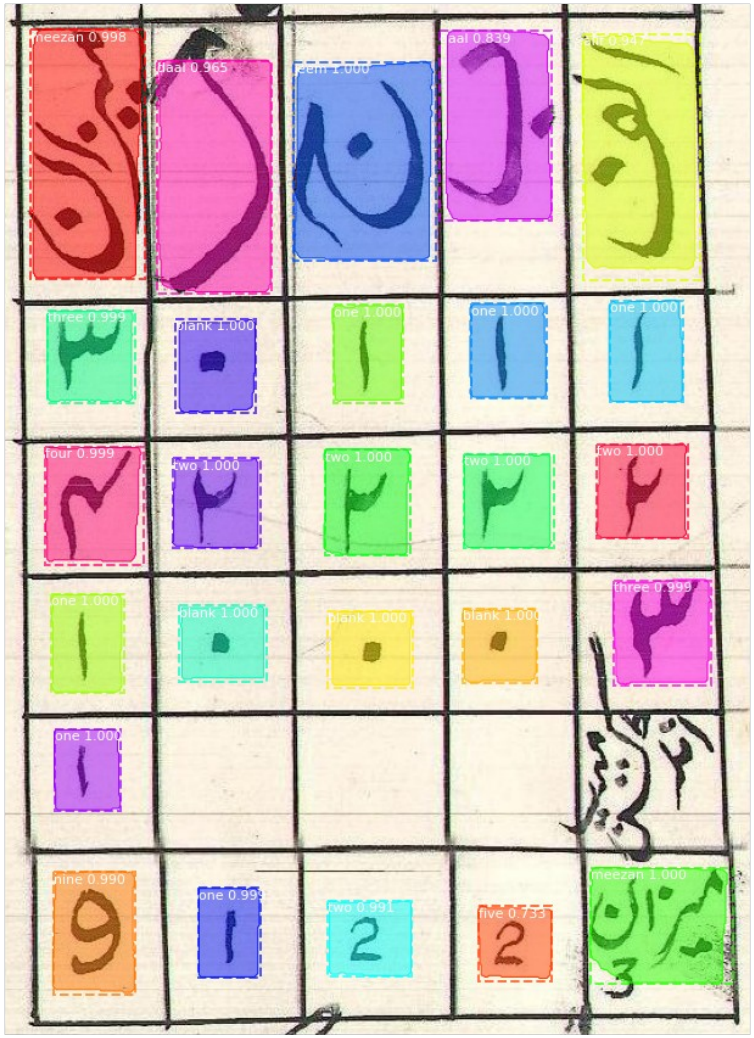
DL Applications

Table detection and understanding



DL Applications

Table detection and understanding



Neuroscience:
A short introduction to ourselves

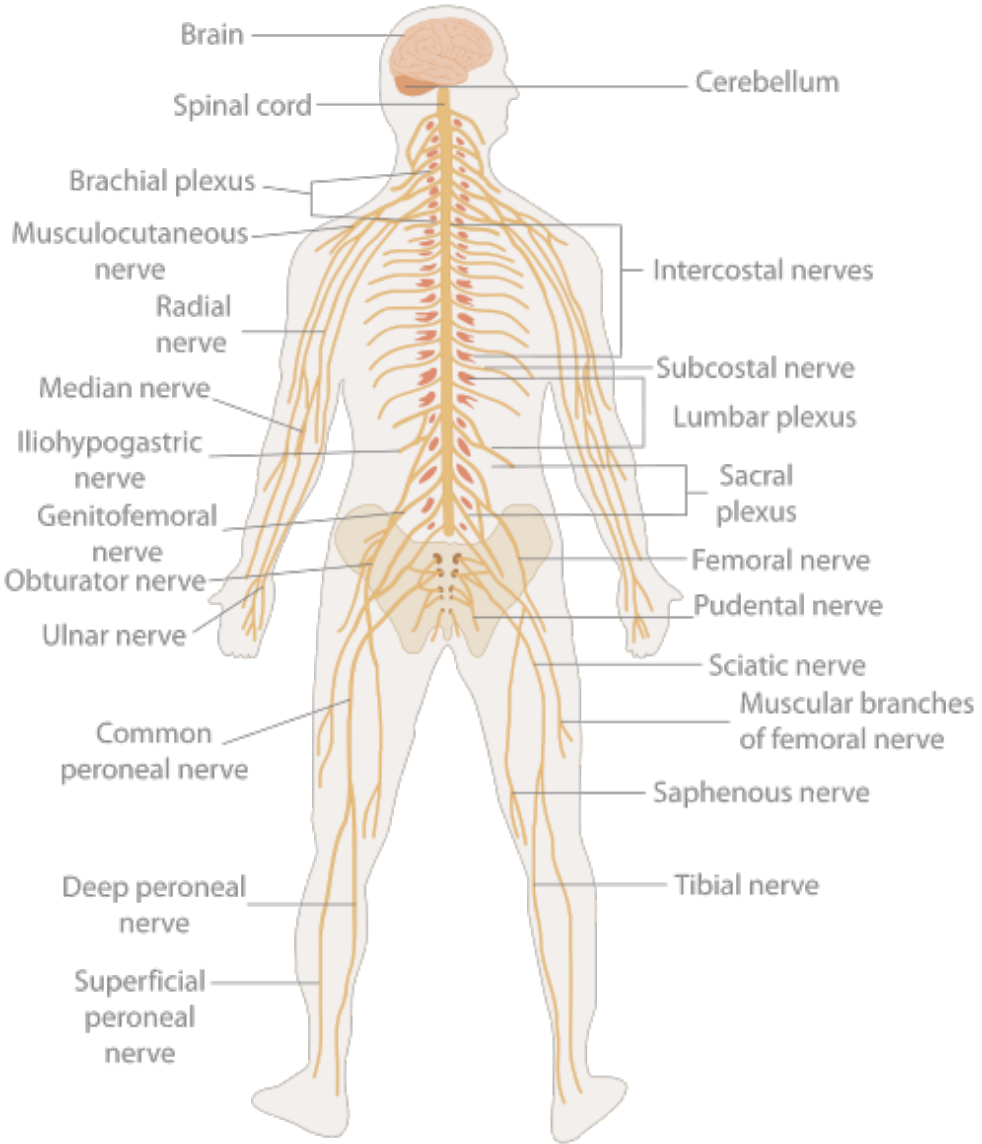
It's all in the brain!

- Your whole existence, your experience of the universe, your happiness, your pain, your memories, your hopes – everything is essentially electrical signals between neurons.
- Sight, smell, sound – it's all in your brain.

Models of Human Cognition

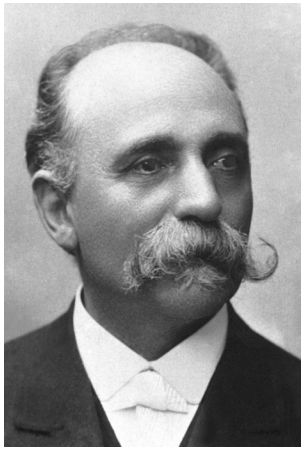
1. Associationism – from Aristotle till 19th century.
   - Humans learn by associating concepts.
   - Aristotle's 4 laws of association
     1.1 contiguity
     1.2 frequency
     1.3 similarity
     1.4 contrast
   - But where and what exactly is the physical mechanism behind associations?
2. Connectionism – last two centuries.
   - Animal systems work through extremely sophisticated networks of inter-connected neurons.

The Nervous System

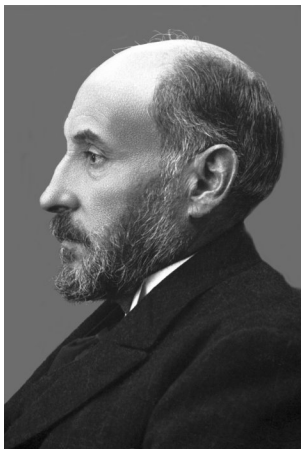


- Coordinates your actions and sensory information by transmitting signals to and from different parts of your body.
- Your brain and spinal cord constitute your *central nervous system*.

The Neuron Doctrine



Camillo Golgi



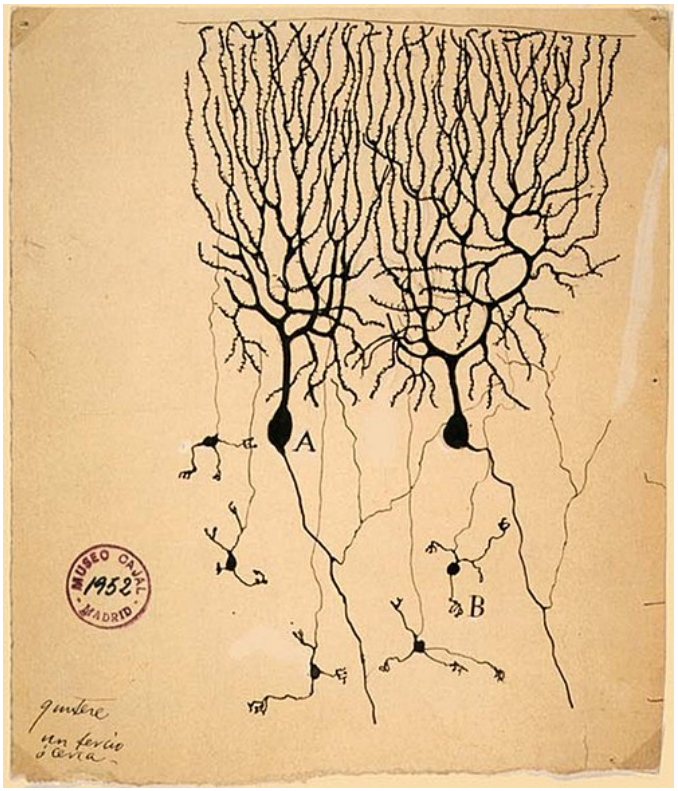
Santiago Ramón y Cajal

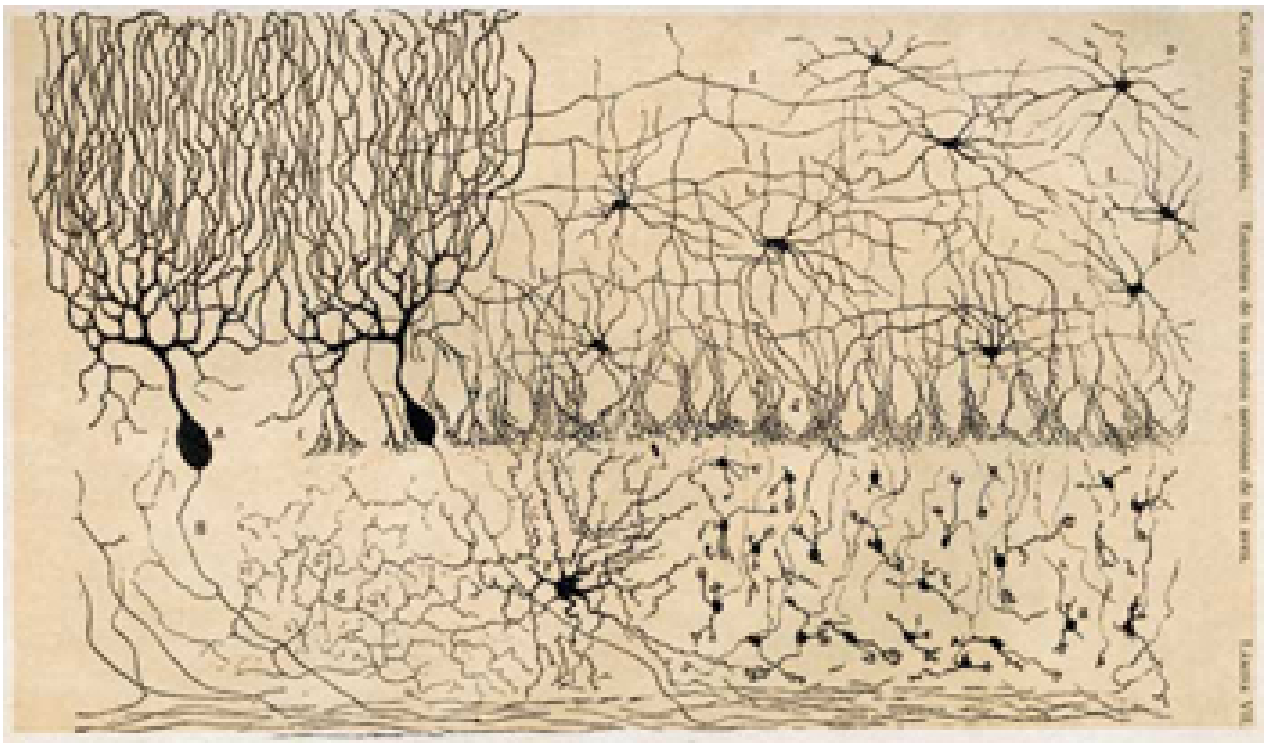
- 1906 Nobel Prize in Physiology or Medicine for studying how our bodies are controlled by the interaction between the brain and the nervous system.
- 1870s – Golgi discovered nerve cells could be coloured using silver nitrate.
- Cajal showed that *each nerve cell is an independent entity* and *nerve impulses travel from one cell to another*.

Source: https://www.nobelprize.org

The nervous system is made up of discrete individual cells, called *neurons* and they form a *communication network*.

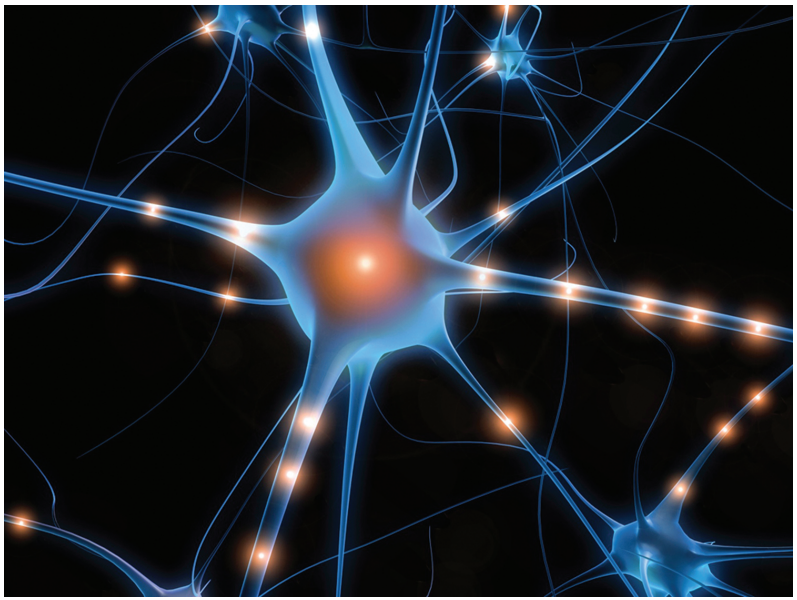
The Neuron Doctrine





Cajal's drawings showing branching and interconnectedness between nerve cells. Left: pigeon cerebellum from 1899. Right: Chick cerebellum from 1905.

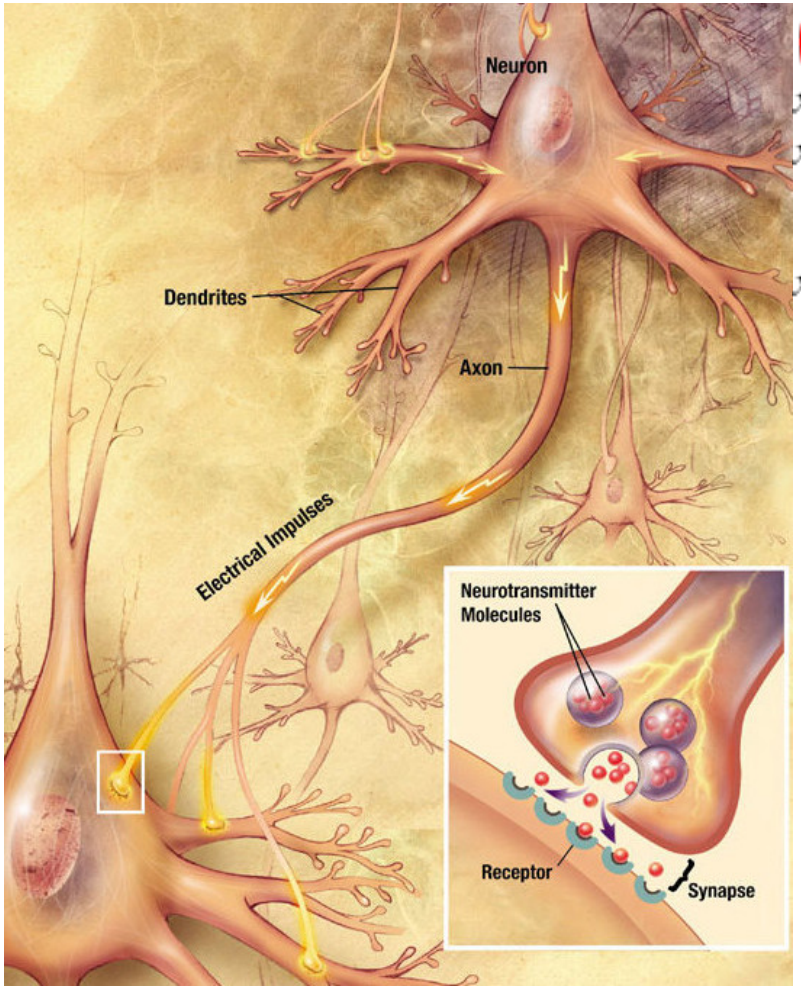
The Human Brain

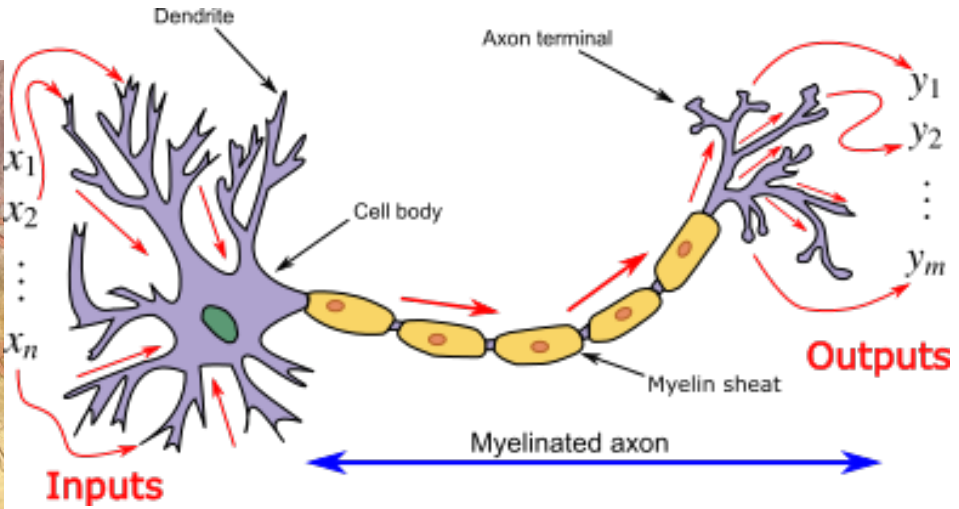


- Around 86 billion neurons[6].
- Each neuron has around $7,000$ synaptic connections to other neurons on average.
- Around 1000 trillion connections in a 3 year old.
- Around 100 to 500 trillion in adults.

[6] Herculano-Houzel, 'The human brain in numbers: a linearly scaled-up primate brain'.

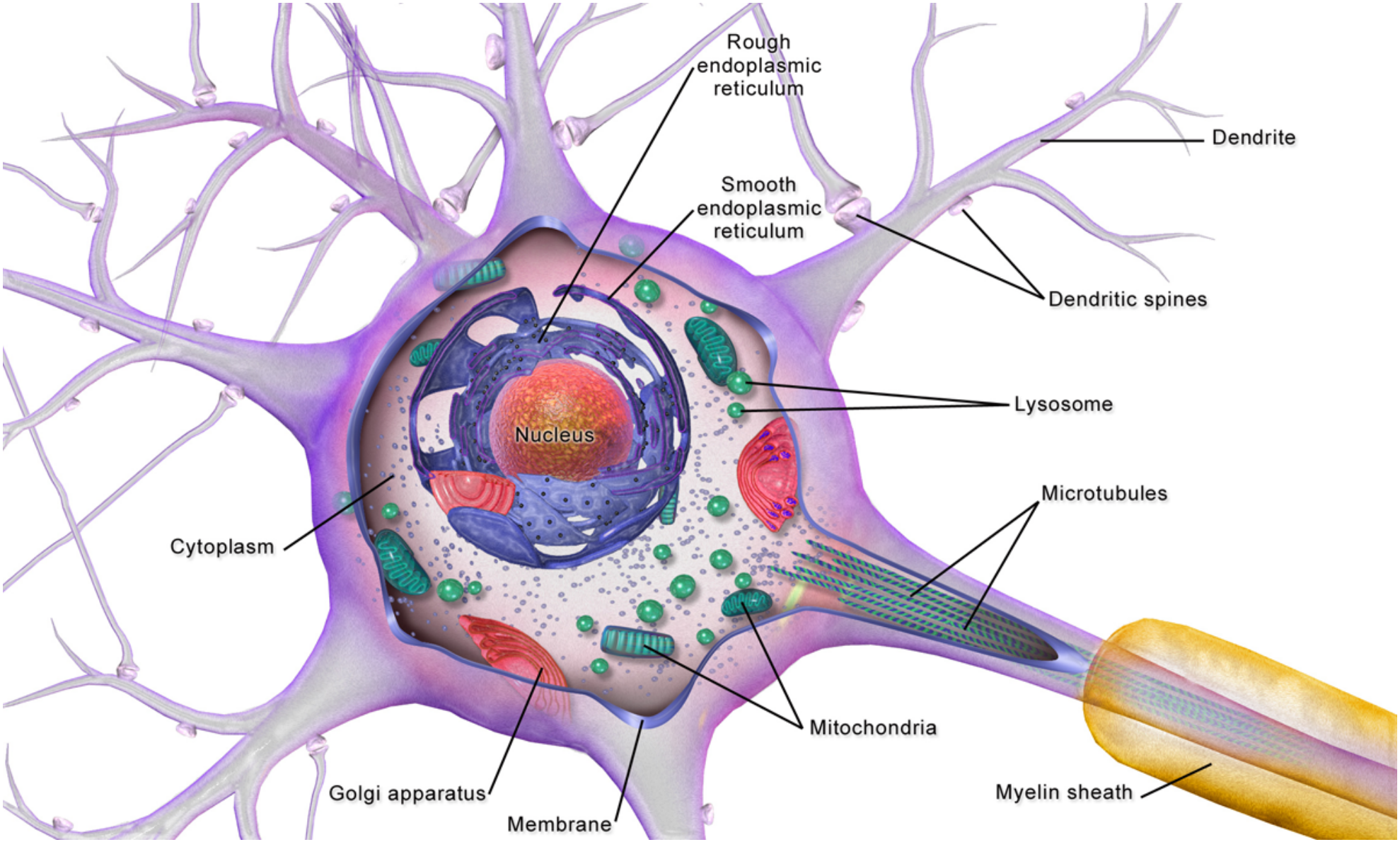
The Neuron
A Simplified View



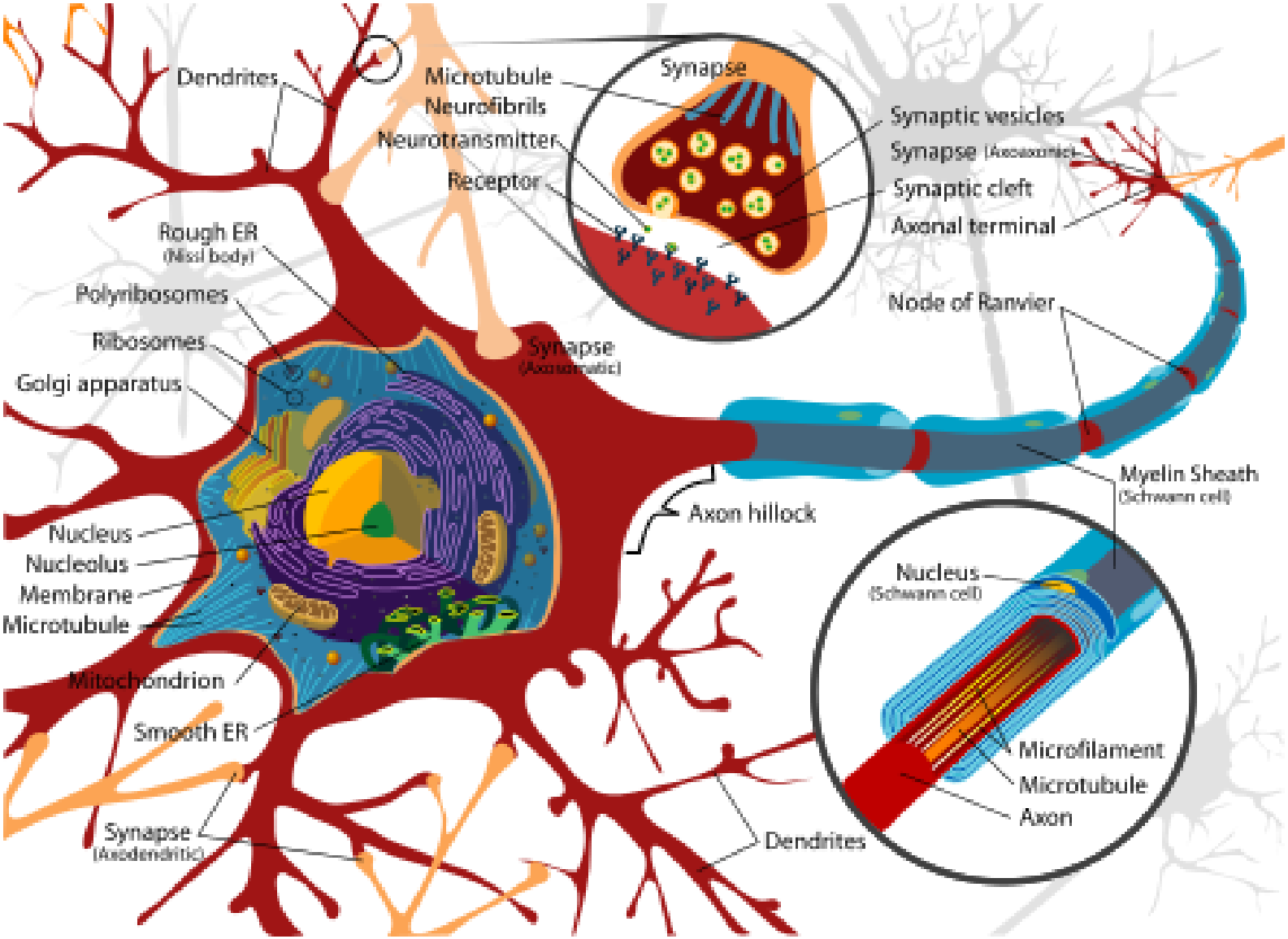


- An electrically excitable cell.
- Cell body called *soma* receives input via filaments called *dendrites*.
- Outputs to other cells via *axon*.
- Axon terminals and dendrites connect via *synapses* where neurotransmitters from one neuron transfer into the next neuron.

Inside the Neuron

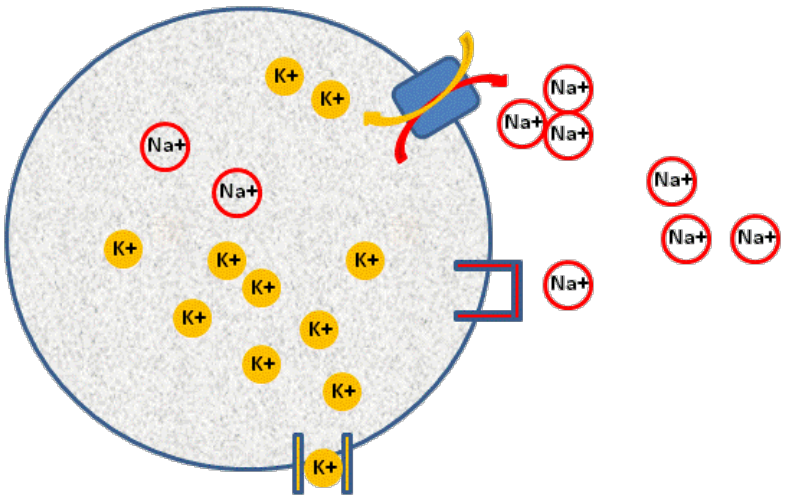


Inside the Neuron

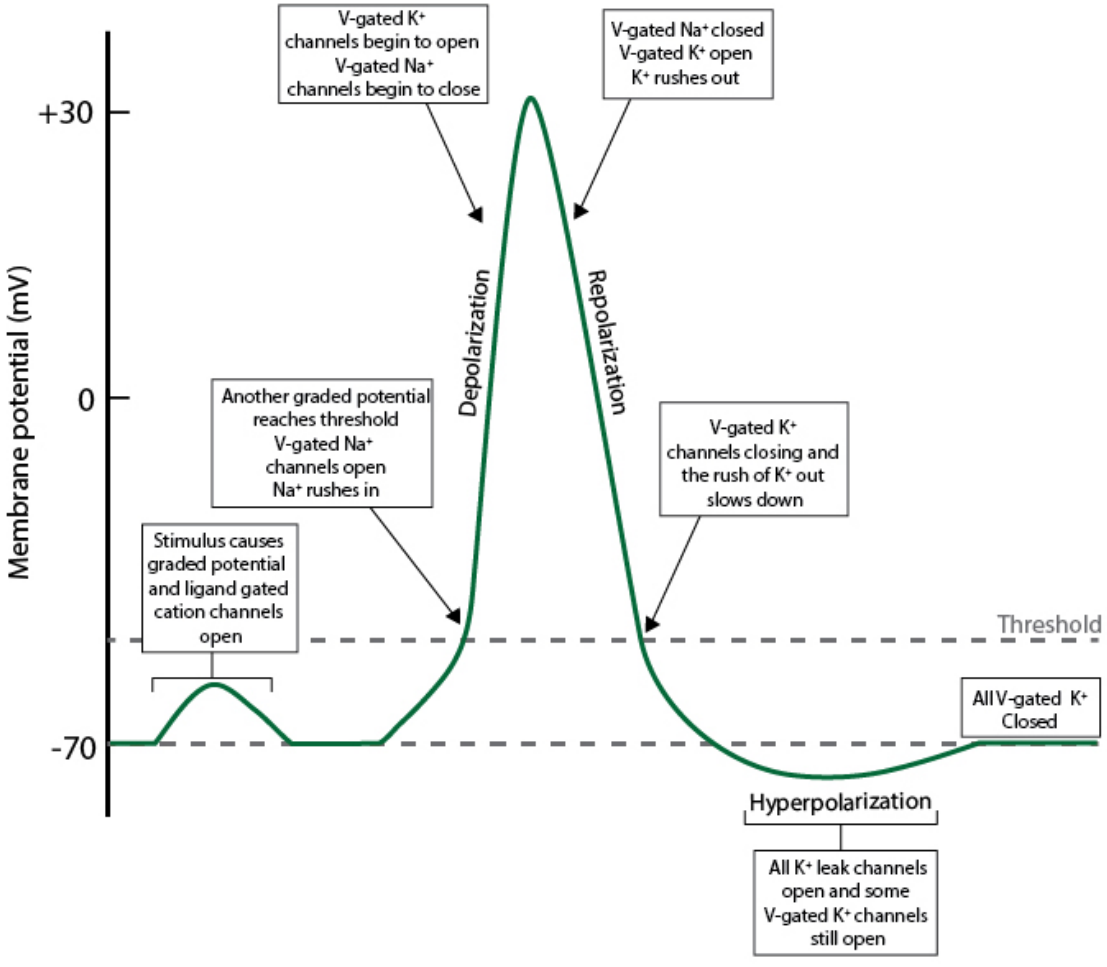


Neuron Spiking

- Neurons maintain voltage gradients across their cell membranes through sodium, potassium, chloride, and calcium ions.
- Sharp voltage changes cause the neuron to emit an electrochemical pulse or spike called an *action potential*.
- This pulse travels rapidly along the axon and activates synaptic connections.
- Synaptic signals may be excitatory or inhibitory, increasing or reducing the net voltage that reaches the soma (cell body).



How does a neuron fire?



- Action potential: a rapid change in voltage across the cell membrane.
- Due to influx of sodium ions, followed by a rapid return via efflux of potassium ions.
- This is the basis of transmitting signals in nerve cells, causing all movement and perception.

A convenient abstraction: a neuron fires when net charge across its membrane exceeds some threshold.