Úvod do neurónových sietí

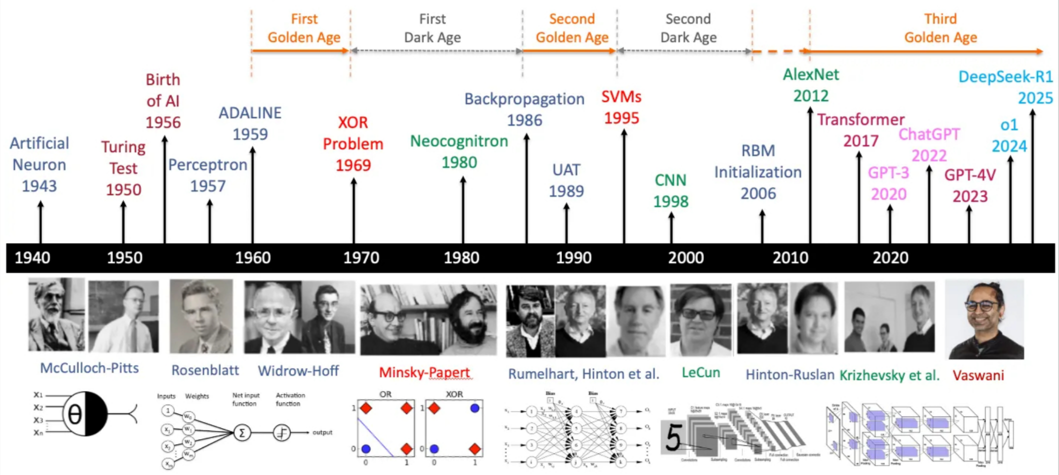
Jozef Jakubík
jozef.jakubik.jefo@gmail.com

Ústav merania SAV, v. v. i.
503a

13. 5. 2026

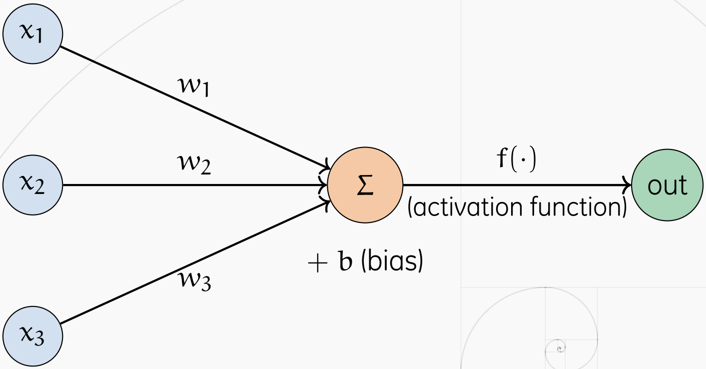
- History
- Basics
- Training
- Standard Neural Network Architectures
- Conclusion

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)



Po, L. (2024). A brief history of AI with deep learning: From artificial neuron (1943) to DeepSeek-R1 (2025). https://medium.com/@lmpo/a-brief-history-of-ai-with-deep-learning-26f7948bc87b. Medium. Updated Dec. 23, 2025

Perceptron

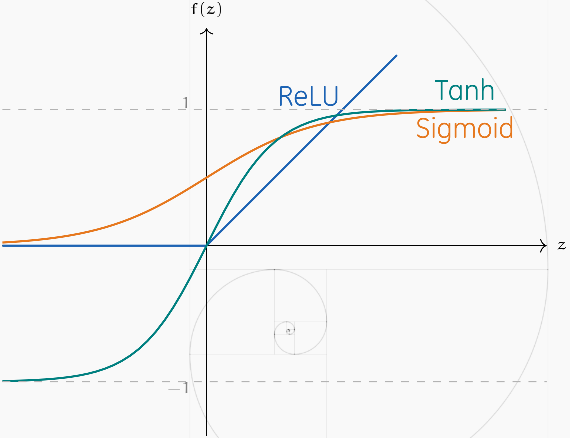


$$\text{out} = f\left(\sum_i w_i x_i + b\right)$$

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

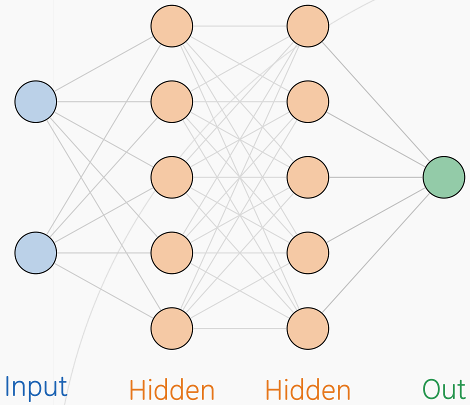
Activation Functions

- $\text{ReLU}(z) = \max(0, z)$
- $\text{Sigmoid}(z) = \frac{1}{1+e^{-z}}$
- $\text{Tanh}(z) = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}}$



Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Multilayer perceptron

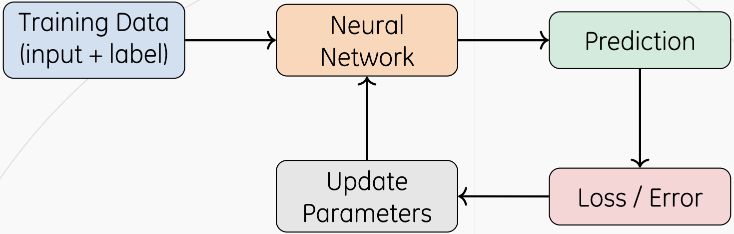


$$\mathbf{a}^{(l)} = f\left(W^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}\right)$$

- **Universal Approximation Theorem** – a single layer *can* approximate any function, but says nothing about how many neurons it needs or how easy it is to train
- **Exponential expressivity** – deep networks represent functions that shallow networks need exponentially more neurons to match
- **Hierarchical features** – each layer builds on the previous: edges → shapes → parts → concepts
- **Parameter efficiency** – depth allows reuse of lower-level features across many higher-level concepts

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

The Idea of Training



- Start with **random weights**
- Pass data through the network (**forward pass**)
- Measure how wrong the prediction is
- Adjust parameters to be **less wrong** (**backward pass**)
- Repeat over many examples

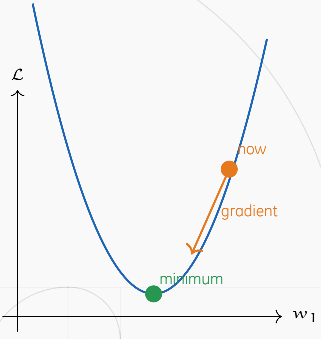
Loss Functions

Mean Squared Error (regression):

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2$$

Cross-Entropy (classification):

$$\mathcal{L}_{\text{CE}} = -\sum_{c} y_c \log \hat{p}_c$$



Goal

Minimise $\mathcal{L}$ by adjusting all weights and biases.

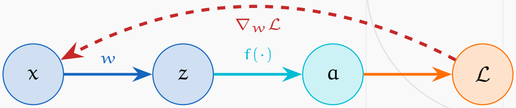
The Chain rule

$$\frac{\partial \mathcal{L}}{\partial w^{(l)}} = \frac{\partial \mathcal{L}}{\partial a^{(l)}} \cdot \frac{\partial a^{(l)}}{\partial z^{(l)}} \cdot \frac{\partial z^{(l)}}{\partial w^{(l)}}$$

$$\frac{\partial \mathcal{L}}{\partial a^{(l)}} = \text{change on output}$$

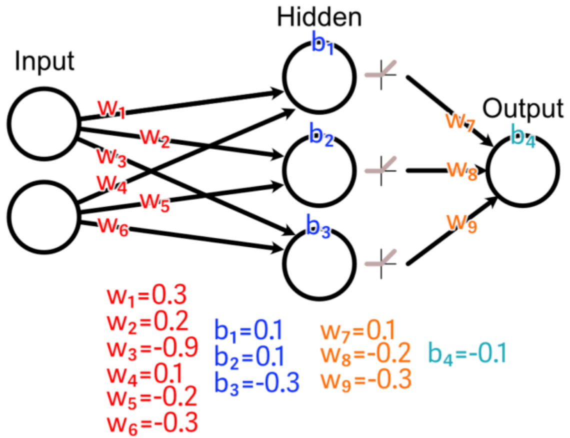
$$\frac{\partial a^{(l)}}{\partial z^{(l)}} = \text{derivation of activation function}$$

$$\frac{\partial z^{(l)}}{\partial w^{(l)}} = \text{input to weight}$$



Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Backpropagation



Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

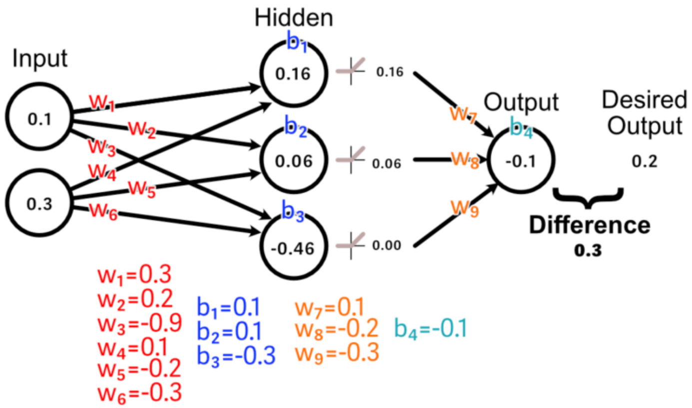
Backpropagation

Input	Desired Output
0.1, 0.3	0.2
0.1, -0.1	0.0
-0.8, 0.2	-0.3
⋮	⋮

Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

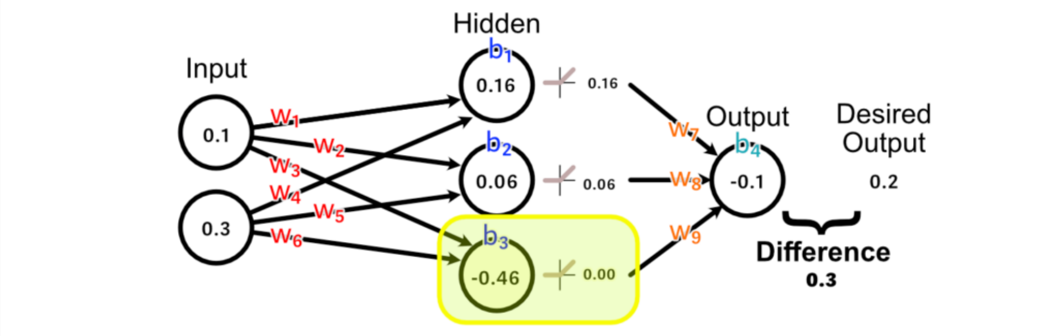
Backpropagation



Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

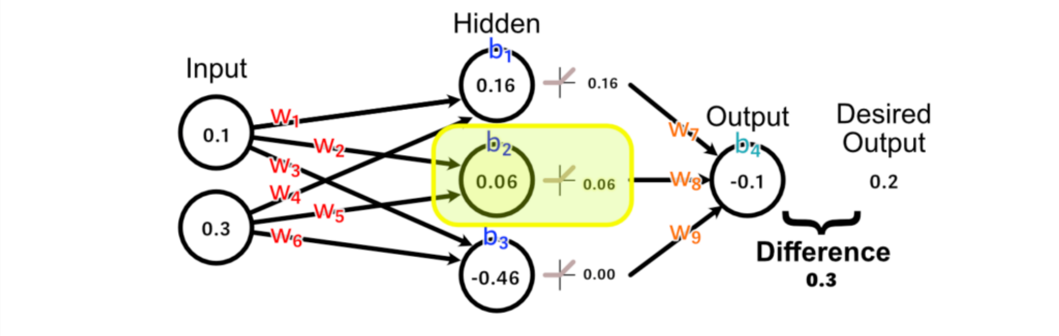
Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Backpropagation



Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

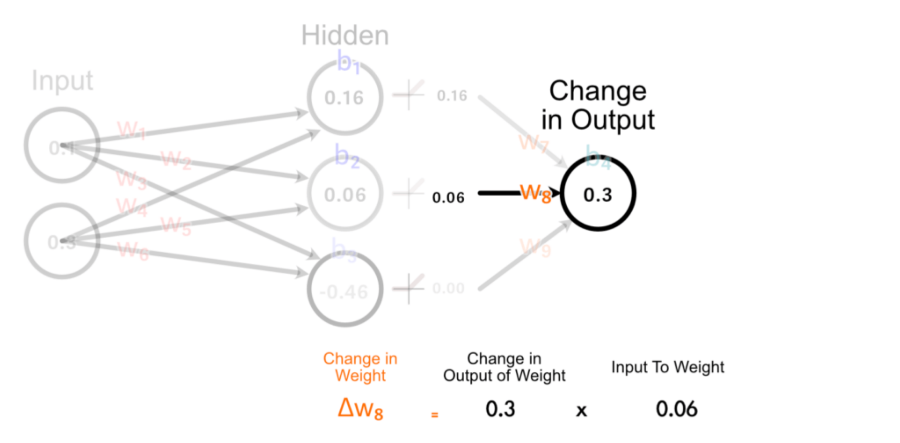
Backpropagation



Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

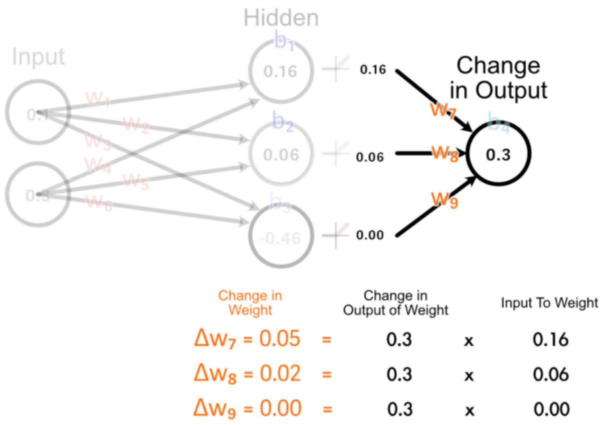
Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Backpropagation



Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

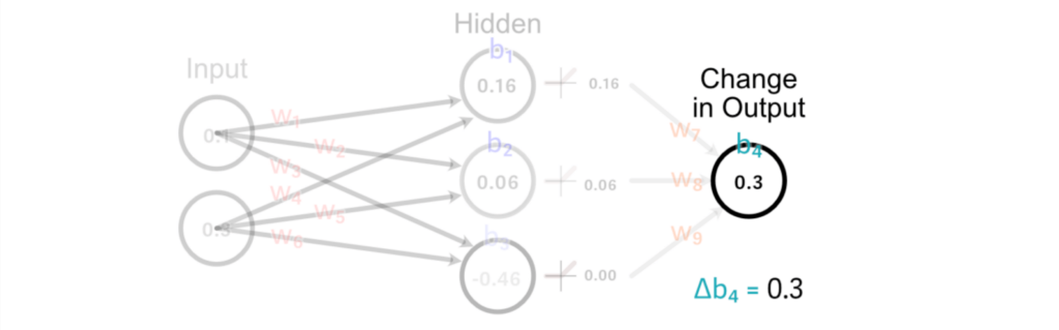
Backpropagation



Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

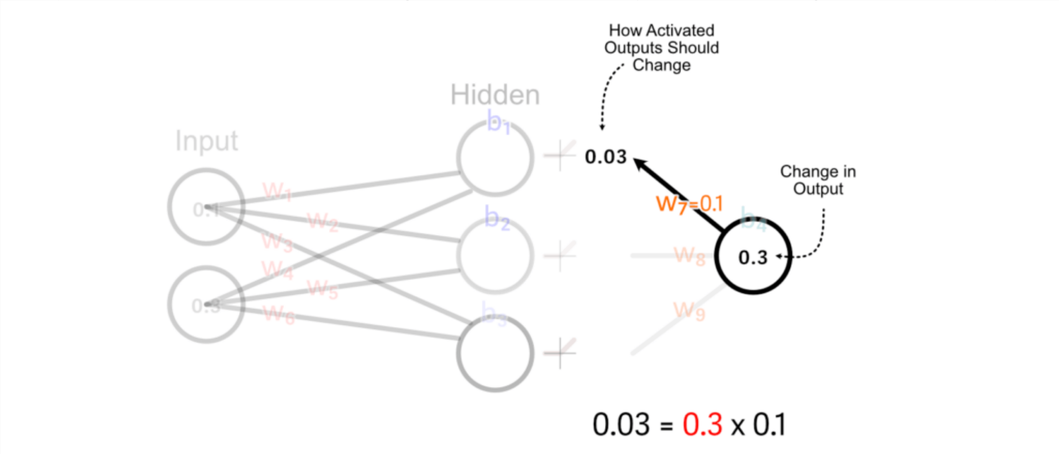
Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Backpropagation



Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

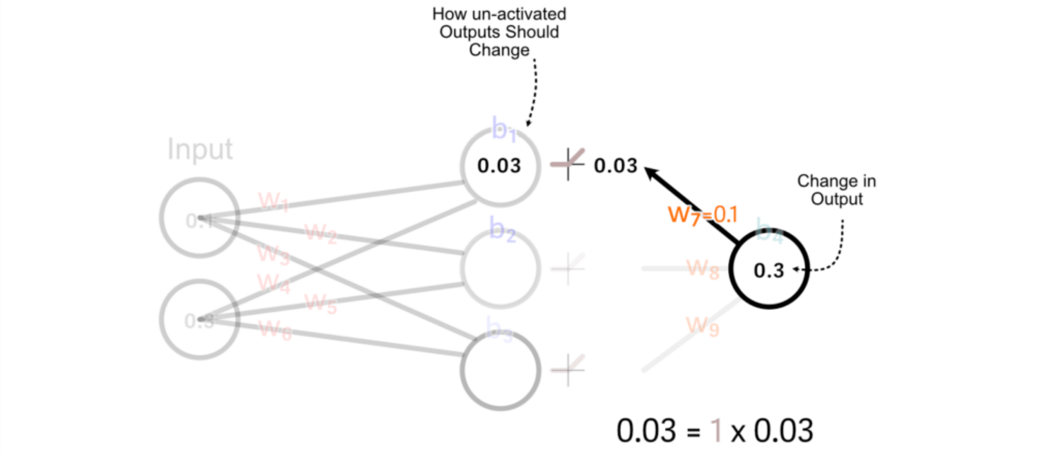
Backpropagation



Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

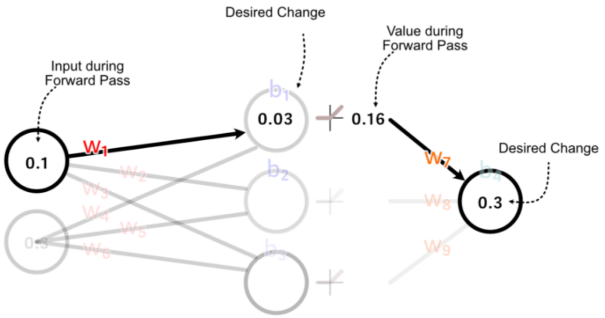
Backpropagation



Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Backpropagation

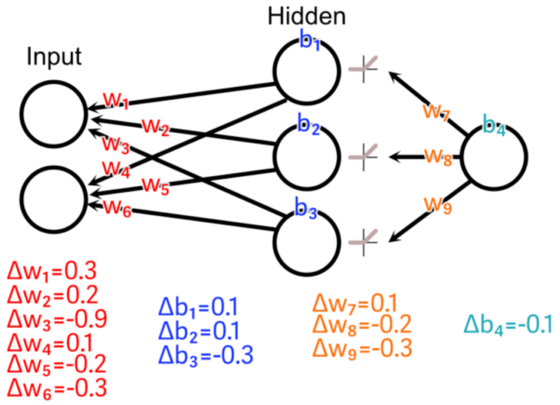


Change in Weight		Change in Output of Weight		Input To Weight
$\Delta w_7 = 0.050$	=	0.30	x	0.16
$\Delta w_1 = 0.003$	=	0.03	x	0.10

Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Backpropagation



Warfield, D. (2025). Neural networks – intuitively and exhaustively explained. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

Learning Rate η

- One training example shouldn't completely change the model
- Scale down all weight updates by a small factor η

$$w \leftarrow w + \eta \cdot \Delta w$$

- Typical values: 10^{-4} to 10^{-2}
- **Too large** $\Rightarrow$ overshooting, divergence
- **Too small** $\Rightarrow$ slow training
- Usually tuned as a *hyperparameter*

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Optimisation Algorithms

Stochastic Gradient Descent (SGD)

$$w \leftarrow w - \eta \cdot \Delta w$$

Adam optimiser

Combines **Momentum** and **RMSprop**:

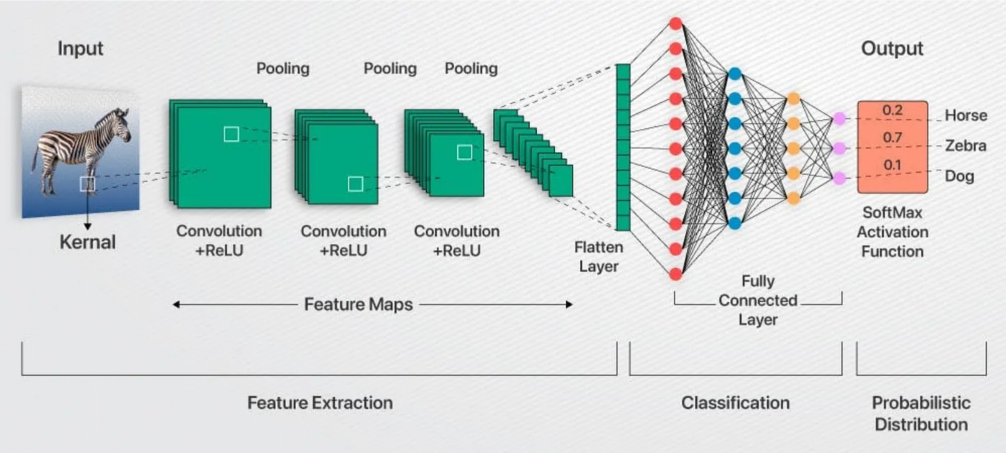
$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\nabla_w$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\nabla_w^2$$

$$w \leftarrow w - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Convolutional neural networks (CNN)



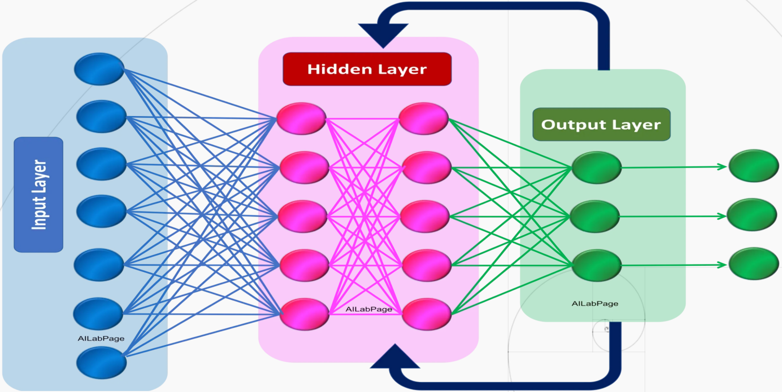
Nadargi, S. (2025). Convolutional neural networks (CNNs). https://medium.com/@sanket.nadargi1/convolutional-neural-networks-cnns-1c55997a6c5a

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

CNN applications

- **Image classification** – assigning a label to an image (e.g. cat vs. dog); backbone of models like ResNet and EfficientNet
- **Object detection** – locating and classifying multiple objects in a single image; used in YOLO and Faster R-CNN
- **Semantic segmentation** – classifying every pixel in an image; core component of self-driving car perception
- **Face recognition** – identifying individuals from images or video; used in security and biometric systems
- **Medical imaging** – detecting tumors, analyzing MRI/CT scans, identifying cancer cells in histopathology
- **Autonomous systems** – lane detection, traffic sign recognition, and pedestrian detection in self-driving vehicles

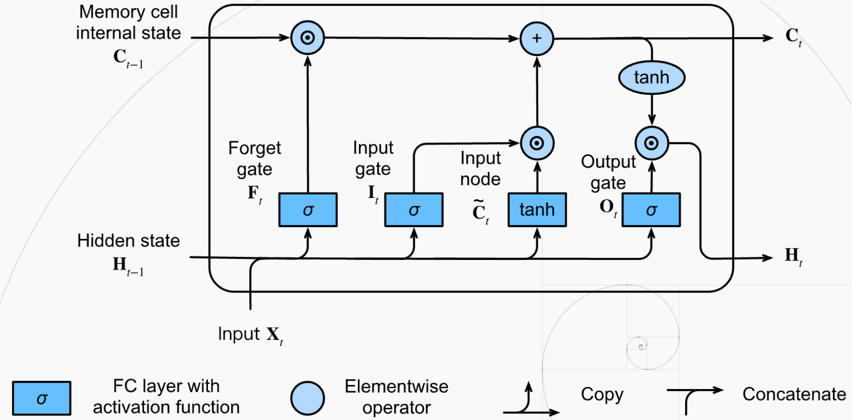
Rrecurrent neural networks (RNN)



Boufeloussen, O. (2020). Simple explanation of recurrent neural network (RNN). `https://medium.com/swlh/simple-explanation-of-recurrent-neural-network-rnn-1285749cc363`. The Startup, Medium

Jozef Jakubík - `jozef.jakubik.jefo@gmail.com`
Ústav merania SAV, v. v. i. (503a)

Long short-term memory (LSTM)



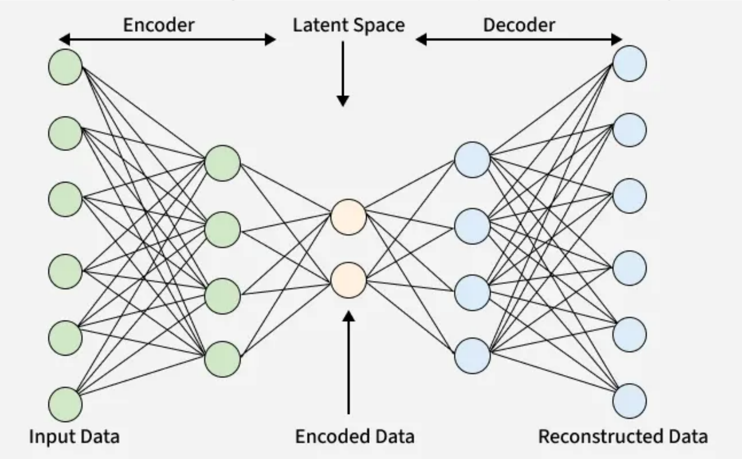
Zhang, A., Lipton, Z. C., Li, M., and Smola, A. J. (2023). *Long Short-Term Memory (LSTM), chapter 10.1.* Cambridge University Press, 1.0.3 edition

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

RNN & LSTM applications

- **Natural language processing** – text generation, machine translation, and sentiment analysis
- **Speech recognition** – converting audio sequences to text; used in virtual assistants
- **Time series forecasting** – stock prices, weather prediction, energy consumption
- **Music generation** – composing melodies by learning patterns in sequential audio data
- **Video analysis** – action recognition and video captioning using frame sequences
- **Handwriting recognition** – decoding sequences of pen strokes into text
- **Anomaly detection** – identifying unusual patterns in sensor or network traffic data

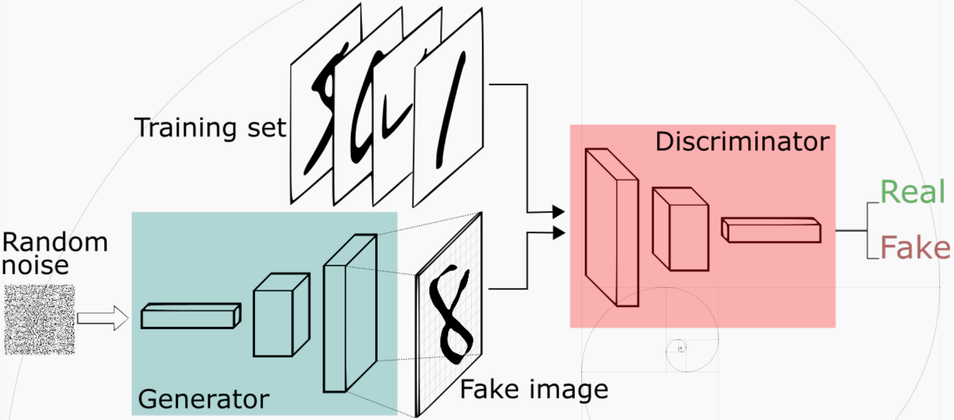
Autoencoders



GeeksforGeeks (2025a). Types of autoencoders. https://www.geeksforgeeks.org/numpy/types-of-autoencoders/

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

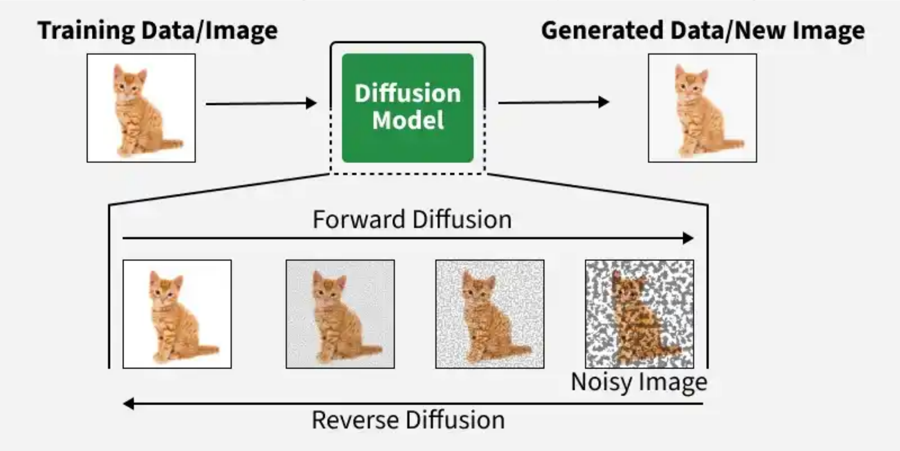
Generative adversarial network (GAN)



Silva, T. S. (2017). A short introduction to generative adversarial networks. https://sthalles.github.io/intro-to-gans/

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Diffusion model



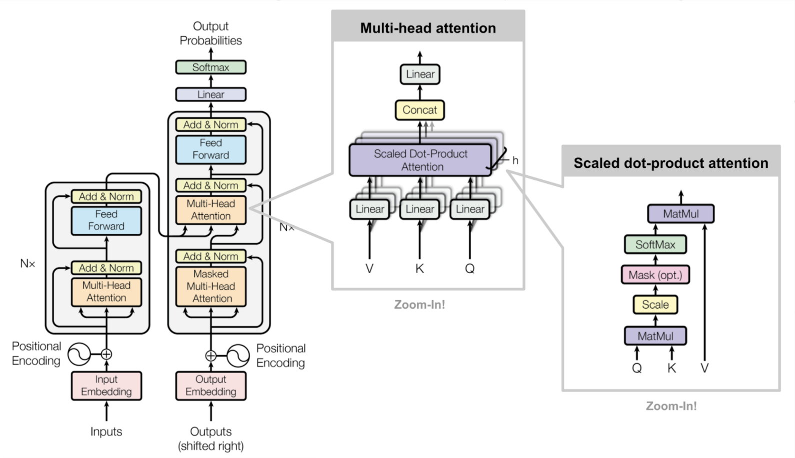
GeeksforGeeks (2025b). What are diffusion models? `https://www.geeksforgeeks.org/artificial-intelligence/what-are-diffusion-models/`

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Autoencoders & GAN & Diffusion model applications

- **Autoencoders**
  - Dimensionality reduction and feature extraction
  - Anomaly detection in fraud and network intrusion detection
  - Image denoising and restoration
  - Data compression for efficient storage and transmission
- **GANs**
  - Photorealistic image synthesis and style transfer
  - Data augmentation for training with limited labeled data
  - Super-resolution – enhancing low-quality images
  - Deepfake generation and face aging simulation
- **Diffusion Models**
  - Text-to-image generation (Stable Diffusion, DALL-E)
  - Video and audio synthesis
  - Medical image generation and enhancement
  - Molecular and drug design in computational biology

Transformer and Attention



Amato, B., Durocher, A., Hurtado, G., Jouandin, A., and Marois, V. (2019). The attention mechanism and the transformer model. https://deepfrench.gitlab.io/deep-learning-project/. Georgia Tech – CS 7643 Deep Learning Class Project

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Transformer & Attention applications

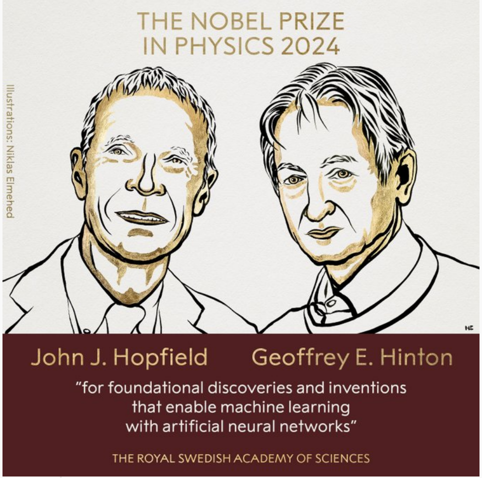
- **Natural language processing** – machine translation, text summarization, and question answering; foundation of models such as BERT and GPT
- **Large language models** – GPT-4, Claude, and LLaMA use transformer architectures scaled to billions of parameters
- **Code generation** – GitHub Copilot and similar tools generate and complete source code from natural language prompts
- **Computer vision** – Vision Transformers (ViT) apply self-attention to image patches for classification and detection
- **Speech recognition** – Whisper and similar models use transformers to transcribe audio to text
- **Protein structure prediction** — AlphaFold uses attention mechanisms to model relationships between amino acids
- **Multimodal models** – CLIP and DALL-E combine vision and language understanding via cross-attention

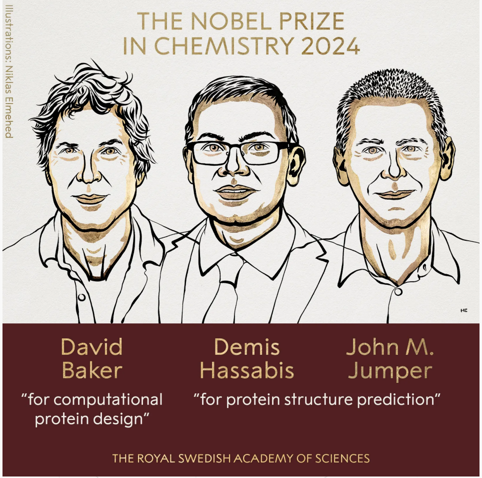
Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)

Neural Networks vs. Classical ML Methods

Criterion	Log. Regression	SVM / RF	Small NN	Deep NN
Interpretability	✓ High	~ Medium	× Low	× Very low
Data volume	Small	Small–Medium	Medium	Large
Training time	Fast	Medium	Medium	Long (GPU)
Unstructured data	×	×	~	✓ Excellent
Performance on img./text	Poor	Poor	Good	Excellent
Feature Engineering	High	High	Medium	Low
Hyperparameters	Few	Moderate	More	Many

Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)





Jozef Jakubík - jozef.jakubik.jefo@gmail.com
Ústav merania SAV, v. v. i. (503a)