CS-C1000 – Introduction to Artificial Intelligence

Deep Learning

Arno Solin

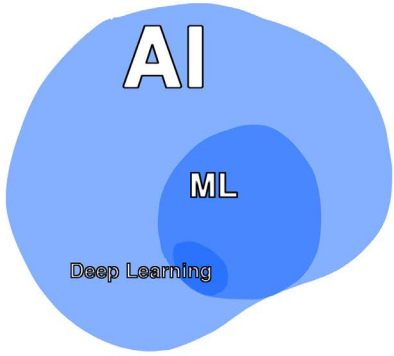
March 19, 2021

@arnosolin arno.solin.fi

Outline

- What is Deep learning?
- Neurons and networks
- Types of deep learning models
- Applicability and examples

AI → ML → Deep Learning

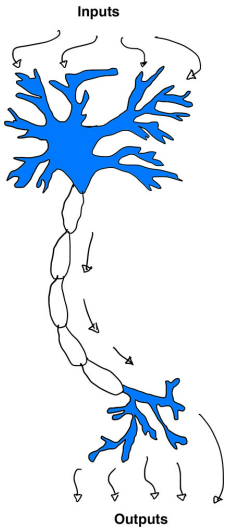


Deep learning

- Deep learning is a class of machine learning methods.
- Uses multiple layers of nonlinear processing units for feature extraction and transformation.
- Learns in supervised (e.g., classification) and/or unsupervised (e.g., pattern analysis) manners.
- Ideally, tries to capture multiple levels of representations that correspond to different levels of abstraction (forms a hierarchy of concepts).

Neural networks

- Artificial neural networks are inspired by actual networks of neurons.
- The human brain has billions of neurons.
- Each neuron can have tens of thousands of connections depending on its type.



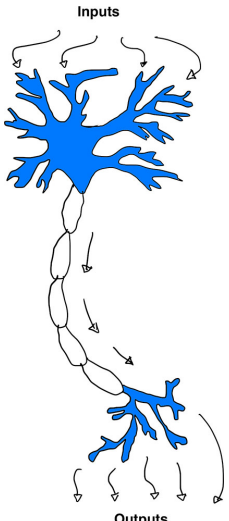
Presented clip

Clip reconstructed from brain activity

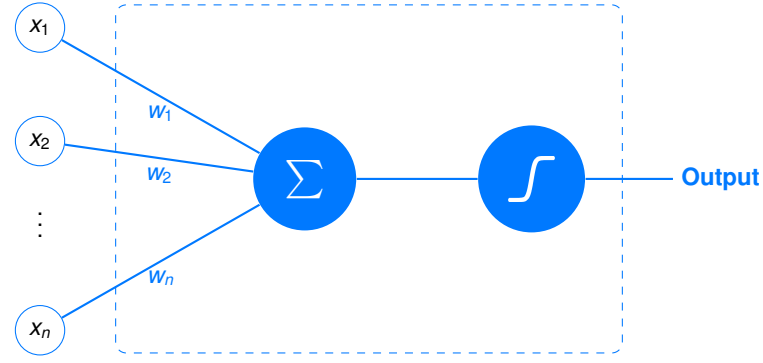
Reconstructing visual experiences from brain activity:
https://www.youtube.com/watch?v=nsjDnYxJ0bo

Neural networks

- Single neurons are simple (nothing ‘intelligent’ about them).
- When there are enough of them, they can power-up the human brain.
- The key is how they are connected and interact.

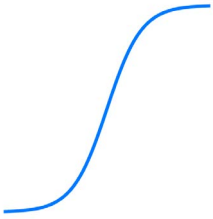


Artificial neuron



Activation functions

Logistic

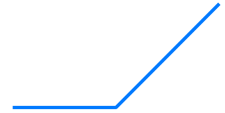


$$g(x) = \frac{1}{1 + e^{-x}}$$

Hyperbolic

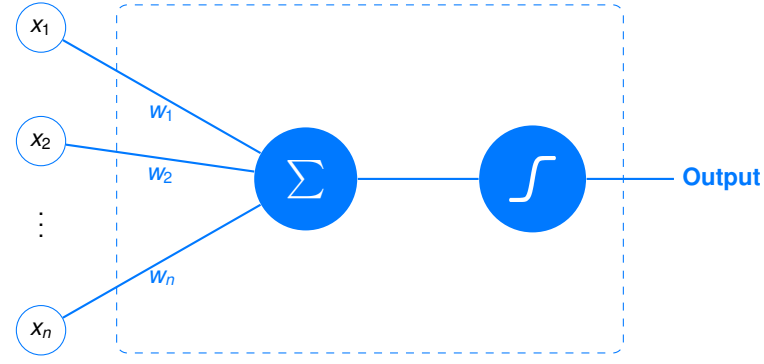
$$g(x) = \tanh(x)$$

Rectifier linear unit (ReLu)



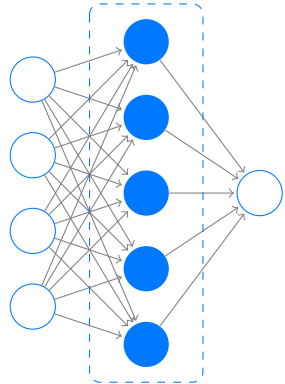
$$g(x) = \max(x, 0)$$

Artificial neuron



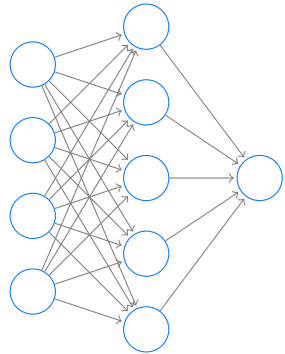
Artificial Neural Networks (ANNs)

- Collection of connected units or nodes called artificial neurons.
- Composing layers consisting many (simple) units.
- Typically you might need many units for capturing interesting phenomena.



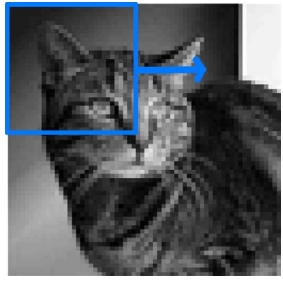
Deep neural networks

- In deep neural networks, there are several layers.
- Each successive layer uses the output from the previous layer as input.
- Makes the model very flexible (see previous lecture).

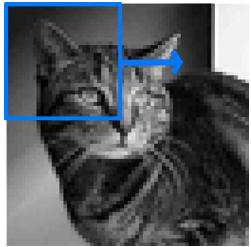


Convolutional Neural Networks (CNNs)

- Most commonly applied to analyzing visual imagery.
- Convolutions sweep the image in order to extract features.
- Good for copturing translation invariant characteristics (no matter where the feature in question appears in the image).

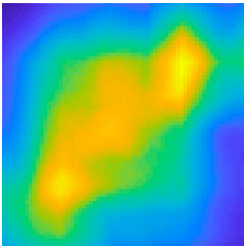


Convolution?



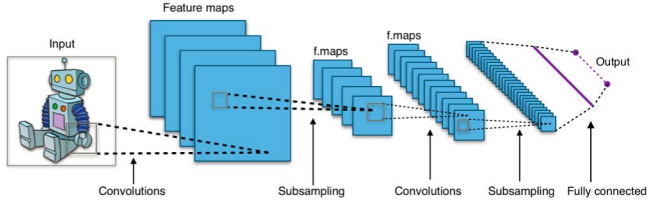
Input image

Convolution kernel
(what to extract)



Output
‘cat-likeness’

How to train these monsters?



- Real-world models are often constructions of many layers.
- Recall that there are weights associated with the units.
- These weights are parameters that need to be ‘learned’ (through optimization).

Image: Typical cnn, Wikimedia Commons.

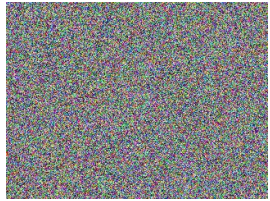
What problems can there be?

Great flexibility comes with some costs:

- Typically the number of weights/parameters is huge.
- Can lead to problems with overfitting (see previous lecture).
- Models are typically big (as in our previous exercise session!) and can require a lot of computational resources (thus GPUs are used).
- Interpretability typically low (black box).

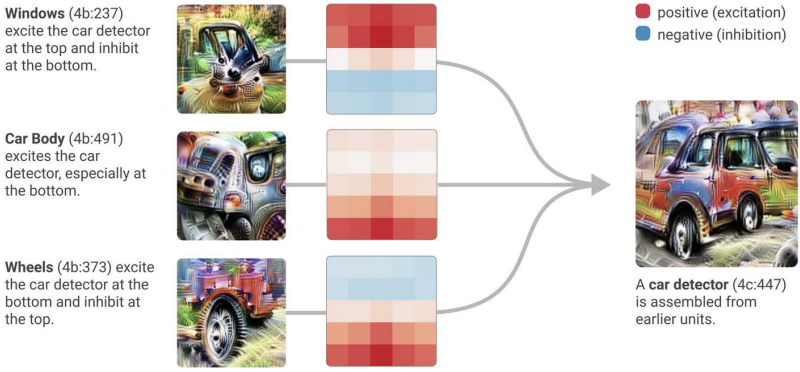
Applicability

- Well-suited especially for applications where the input is complicated (e.g., images or audio).
- But on the other hand, the (sensible) latent space is quite concentrated.
- For example: Of all possible combinations of pixel colors, only a very small (and concentrated) subset actually make sense to us.



A random image.

Interpretability

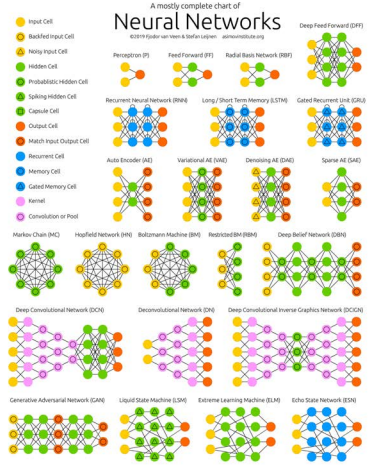


Example of a car detector neuron

By OpenAI (`https://distill.pub/2020/circuits/zoom-in`)

A Zoo of Neural Network Models

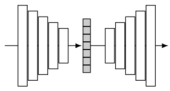
- Deep learning / Neural networks do not form a single model family
- There are numerous different model types that go under this category, and there is no idea going through all of them here
- We only cover the basics of some of the most known ones



See more `https://www.asimovinstitute.org/neural-network-zoo/`

Autoencoders

Input



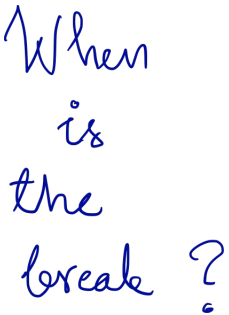
low-dimensional encoding

Output

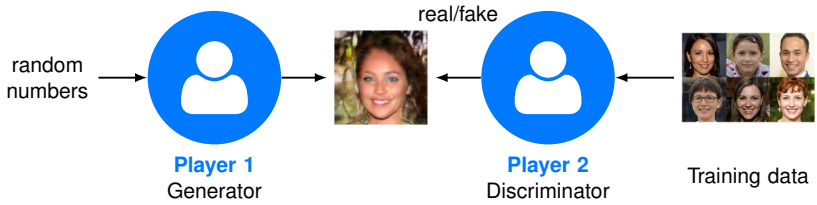
- A type of artificial neural network used to learn efficient data codings in an unsupervised manner.
- Learns a representation (encoding/encoder) for a set of data, typically for dimensionality reduction.
- Also a decoder/reconstructor is learnt, where it tries to reconstruct the origial content.

Recurrent Neural Networks (RNNs)

- The connections between the nodes form a directed graph along a temporal sequence.
- Used for time-evolving phenomena (time-series).
- Applicable to tasks such as unsegmented, connected handwriting recognition or speech recognition.

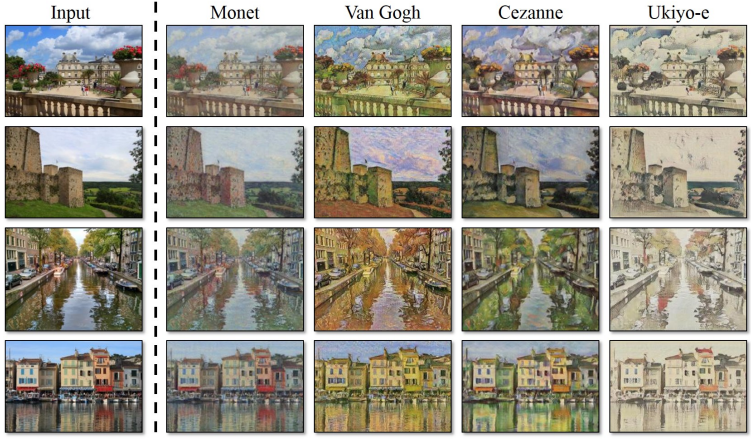


Generative Adversarial Networks (GANs)



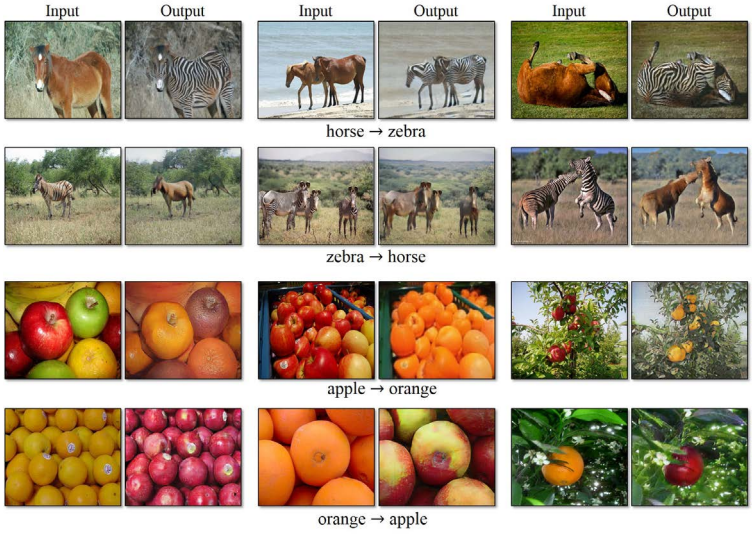
- GANs are like two-player games.
- The generative network generates candidates while the discriminative network evaluates them (real vs. fake).
- Both parts learn and try to get better.

GAN examples



Zhu *et al.* Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV 2017.

GAN examples

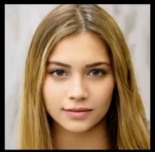


Zhu *et al.* Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV 2017

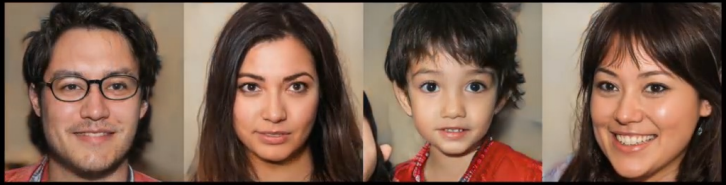
GAN examples

Zhu *et al.* Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV 2017.

Source A: gender, age, hair length, glasses, pose

Source B: everything else



Result of combining A and B

A Style-Based Generator Architecture for Generative Adversarial Networks:
https://www.youtube.com/watch?v=kSLJriaOumA

Transformers

- A Transformer is a deep learning model that utilizes the mechanism of attention
- Mimics cognitive attention: enhances the important parts of the input data and fades out the rest
- Primarily in the field of natural language processing (NLP), in machine translation, text generation, etc.

Transformers Example

TEXT PROMPT

an illustration of a baby daikon radish in a tutu walking a dog

AI-GENERATED IMAGES

- DALL·E a 12-billion parameter version of GPT-3 trained to generate images from text descriptions, using a dataset of text–image pairs.

See `https://openai.com/blog/dall-e/`

Transformers Example

TEXT PROMPT

an armchair in the shape of an avocado [...]

AI-GENERATED IMAGES

- DALL·E a 12-billion parameter version of GPT-3 trained to generate images from text descriptions, using a dataset of text–image pairs.

See `https://openai.com/blog/dall-e/`

Transformers Example

TEXT PROMPT

a store front that has the word ‘openai’ written on it [...]

AI-GENERATED IMAGES

- DALL·E a 12-billion parameter version of GPT-3 trained to generate images from text descriptions, using a dataset of text–image pairs.

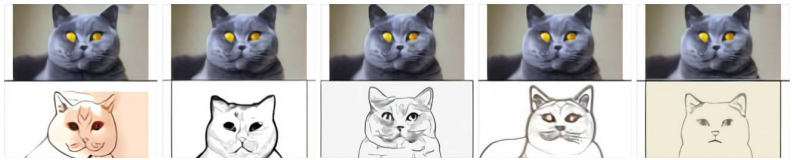
See `https://openai.com/blog/dall-e/`

Transformers Example

TEXT AND IMAGE PROMPT

the exact same cat on the top as a sketch on the bottom

AI-GENERATED IMAGES



- DALL·E a 12-billion parameter version of GPT-3 trained to generate images from text descriptions, using a dataset of text–image pairs.

See `https://openai.com/blog/dall-e/`

Recap

- Deep learning is a class of machine learning methods.
- Multiple layers of nonlinear processing units.
- Each successive layer uses the output from the previous layer as input.
- Capable of great flexibility in feature extraction and transformation.
- Produces state-of-the-art results image, audio, and text modelling.

What next?

- There is a lineup of past visiting lectures in MyCourses. The intention is that you watch at least one of those.
- The second Computer Exercise is next Tuesday.
- The next lecture is next Friday. We will have Janne Pulkkinen from KELA visiting live during the first hour.

AI