Neural networks

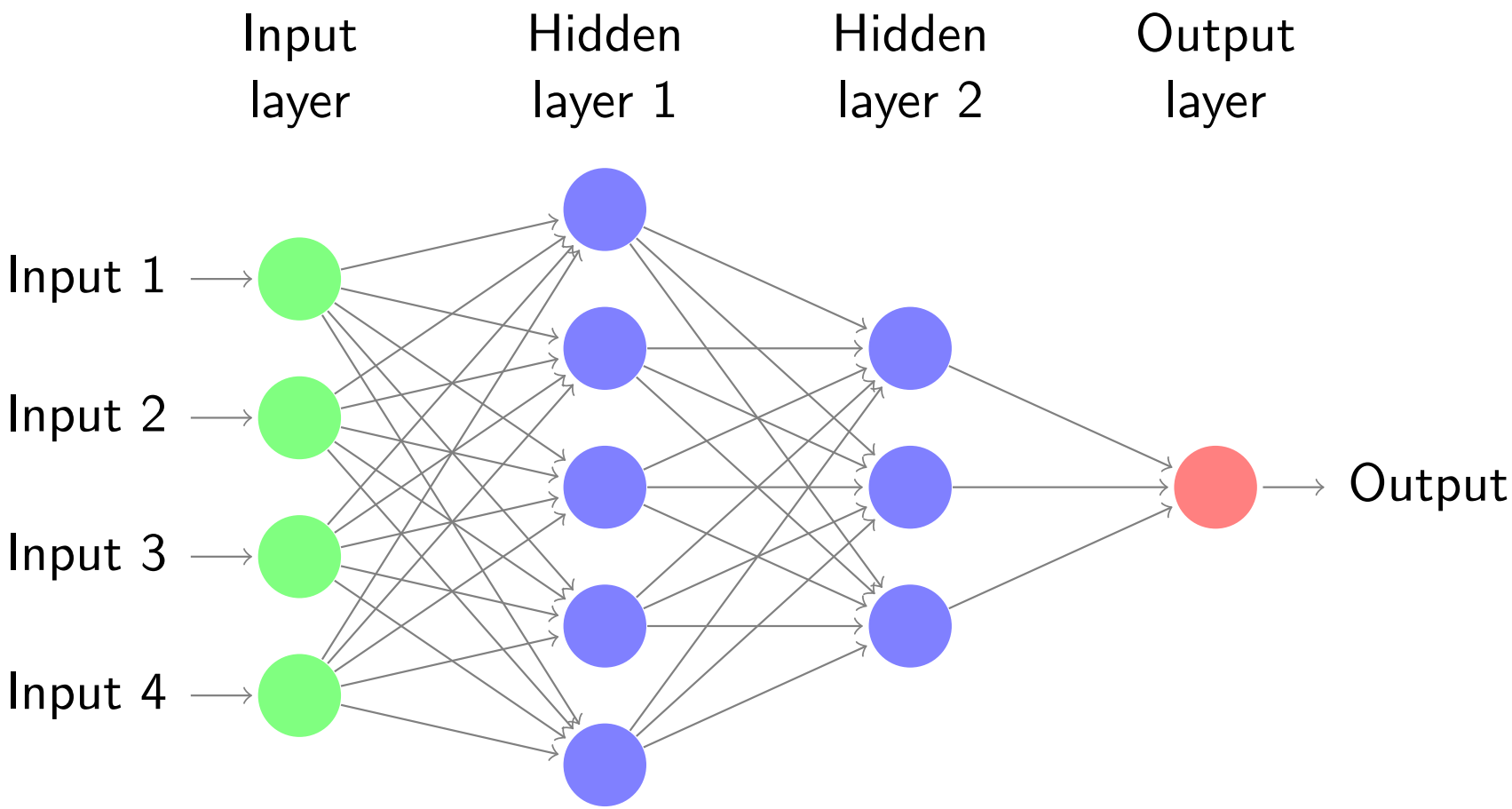
Architectures and training tips

Sebastian Björkqvist

IPRally Technologies

09.01.2019

What is a neural network?



Modified from `http://www.texample.net/tikz/examples/neural-network/`

Why neural networks?

- Can approximate any function [Hornik, 1991]
- May learn to respond to unexpected patterns
- Useful especially when the amount of data is large
- Less need for feature engineering compared to traditional ML methods

Recurrent neural network (RNN)

TODO: Picture here

Processes each element of the input sequence in order, and keeps information about the past elements in a hidden state vector

RNN pros and cons

+ Accepts input of variable size, i.e. sequences (time series, sentences etc)
+ May learn long-term dependencies
– Training may be slow when sequence length is large
– Can be hard to train

Convolutional neural network (CNN)

TODO: Picture here
Extracts features of two-dimensional input (usually an image) using convolutional and pooling layers.

CNN pros and cons

+ Works well with image data
– Pre-existing models can be fine-tuned for specific tasks
– Does not take into account position or orientation of the object

Challenges when training neural networks

- Finding the optimal neural network layout is often time-consuming
- The model may be sensitive to changes in hyperparameters
- A model may take several hours or even days to train.

References

Nielsen, Michael A. *Neural Networks And Deep Learning*. Determination Press, 2015. `http://neuralnetworksanddeeplearning.com/`

Hornik, Kurt. *Approximation Capabilities of Multilayer Feedforward Networks*. Neural Networks, 4(2), 251–257, 1991.