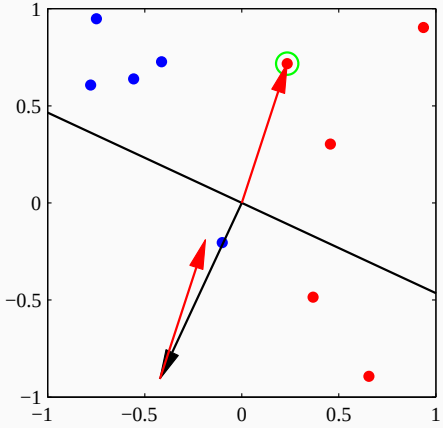
Introduction to neural networks

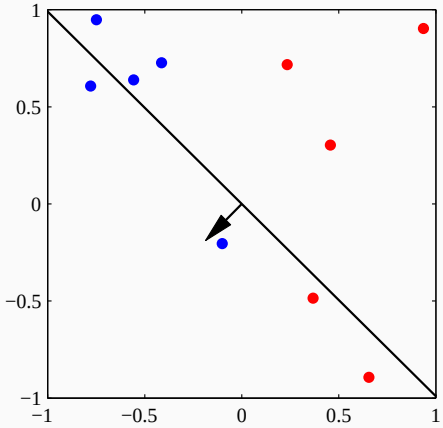
Gianluca Campanella

Perceptrons



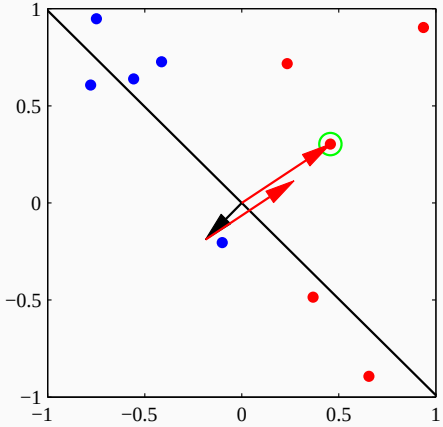
From *Pattern Recognition and Machine Learning*

Perceptrons



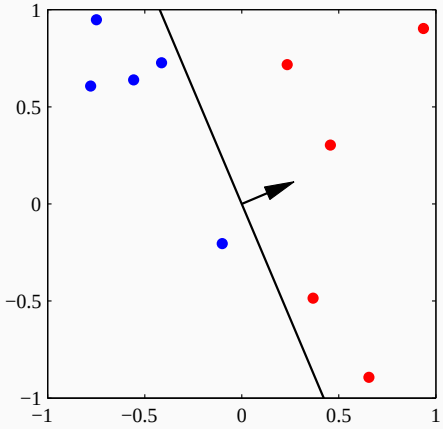
From *Pattern Recognition and Machine Learning*

Perceptrons



From *Pattern Recognition and Machine Learning*

Perceptrons



From *Pattern Recognition and Machine Learning*

Perceptrons

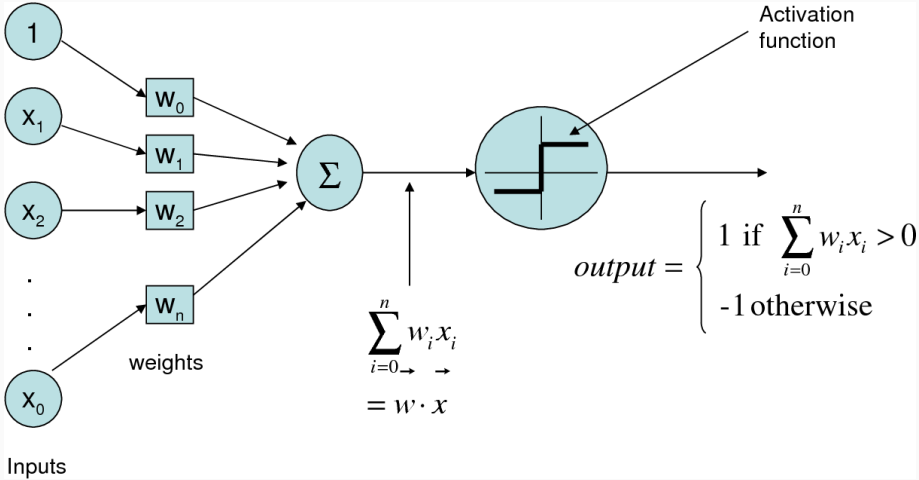
Support vector machines

- Are trained on the entire dataset at once
- Try to find the largest possible margin

Perceptrons

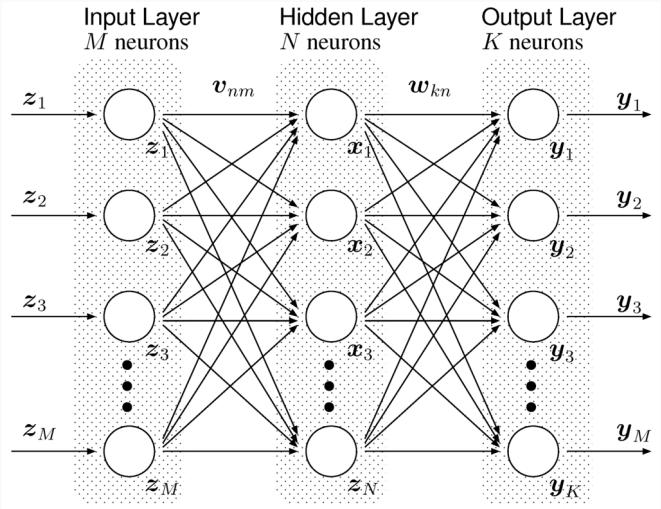
- Can be trained online (as the data arrives)
- Do not necessarily maximise the margin

Perceptrons and neurons



From Verma et al. (2015)

Multi-layer perceptrons



Multi-layer perceptrons

Feed-forward of information

- Receive a new sample X with outcome y
- Compute value for each unit in each layer
- Compute prediction $\hat{y}$ and error $\hat{\varepsilon}$

Multi-layer perceptrons

Back-propagation of error

- Compute ‘blame’…
  - For output units: $y - \hat{y}$
  - For all other layers, as weighted contribution to blame of following layer’s units
- Adjust weights and biases

Multi-layer perceptrons

Questions

- How many hidden layers?
- How many units in each layer?
- Which activation function?
- How do we initialise weights?
- How do we minimise error?

Other common topologies

Convolutional neural networks

- Inspired by the organisation of the visual cortex
- Include convolutional and pooling layers

Recurrent neural networks

- Possess ‘internal memory’
- Can process sequences of inputs

Pros and cons

Pros

- Can handle large datasets
- Effective in high-dimensional spaces ($p > n$)
- Predictions are fast

Cons

- Can require considerable parameter tuning
- Training is somewhat cumbersome
- New data can cause ‘forgetfulness’