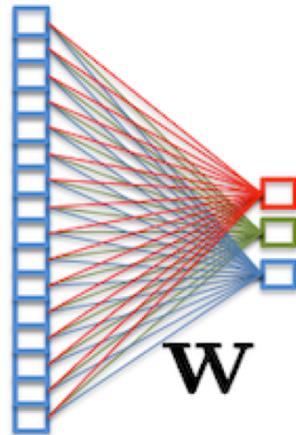


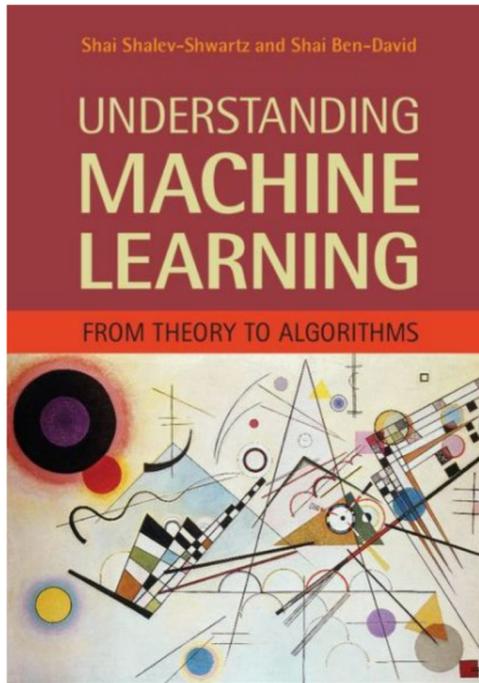
5IF - Deep Learning and Differentiable Programming

1.3 Some extremely short basics of machine learning



To go deeper ...

These next 15 (!!) slides will never be able to replace a full lecture in the theory of machine learning. The interested reader is referred to:



Shai Shalev-Shwartz and Shai Ben-David

Understanding Machine Learning,
from Theory to Algorithms

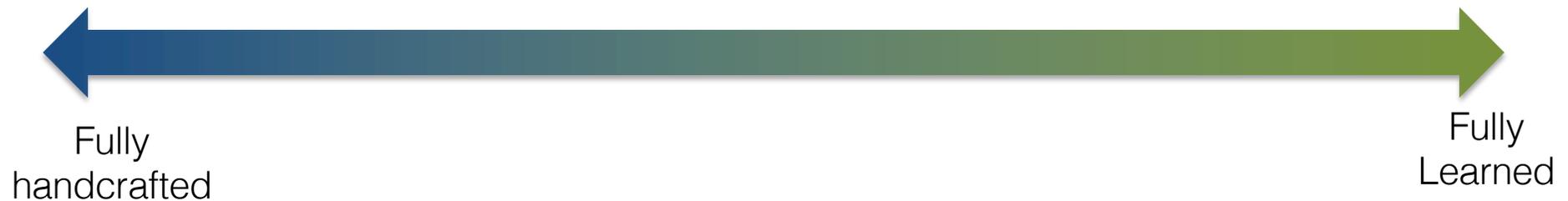
Cambridge University Press, 2014

We would like to learn to predict a value y from observed input x

$$y = h(x, \theta)$$

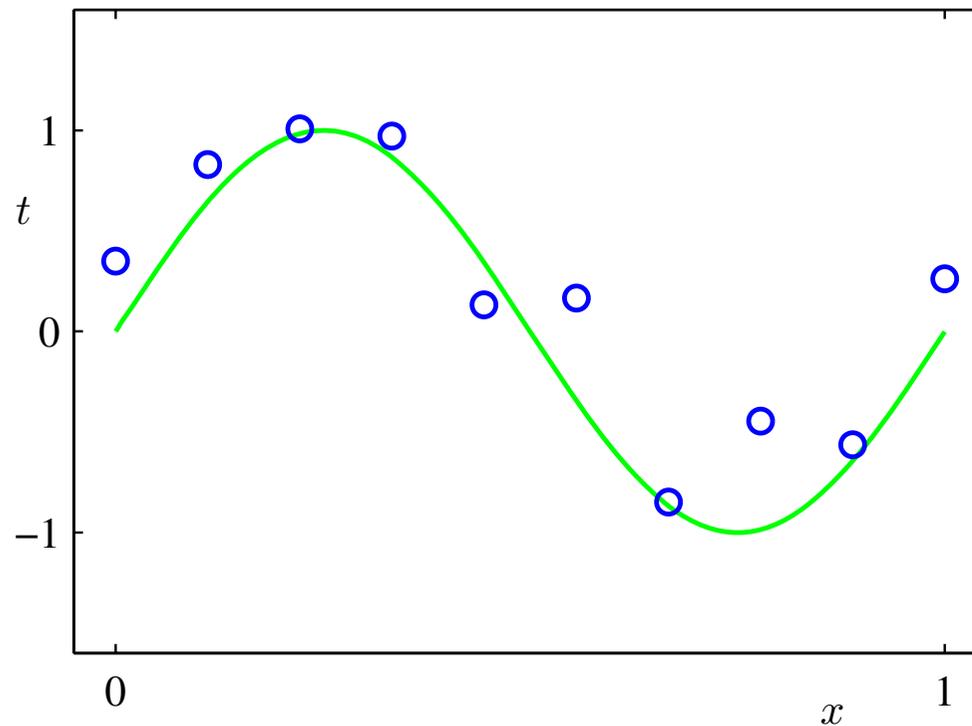
Handcrafted from domain
knowledge

Learned from data or
interactions



Fitting and Generalisation

- Data are generated with function $t = \sin(2\pi x)$
- Objective: assuming the function unknown, predict t from x



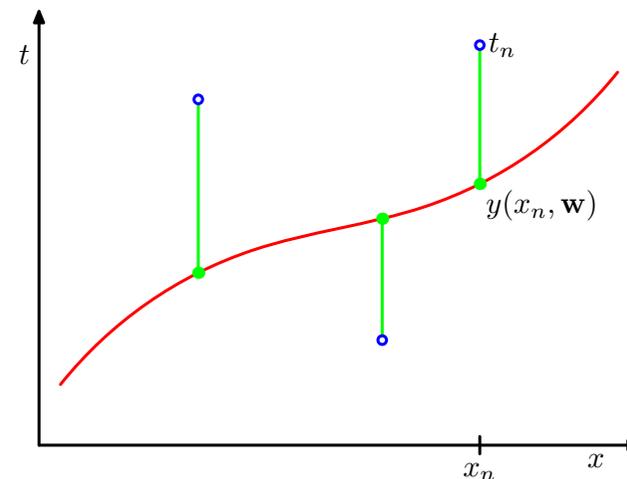
Fitting and Generalisation

Example: « Fitting » of a polynomial of order M

$$y(x, \mathbf{w}) = w_0 + w_1x + w_2x^2 + \dots + w_Mx^M = \sum_{j=0}^M w_jx^j$$

« Least squares » (of errors) criterion

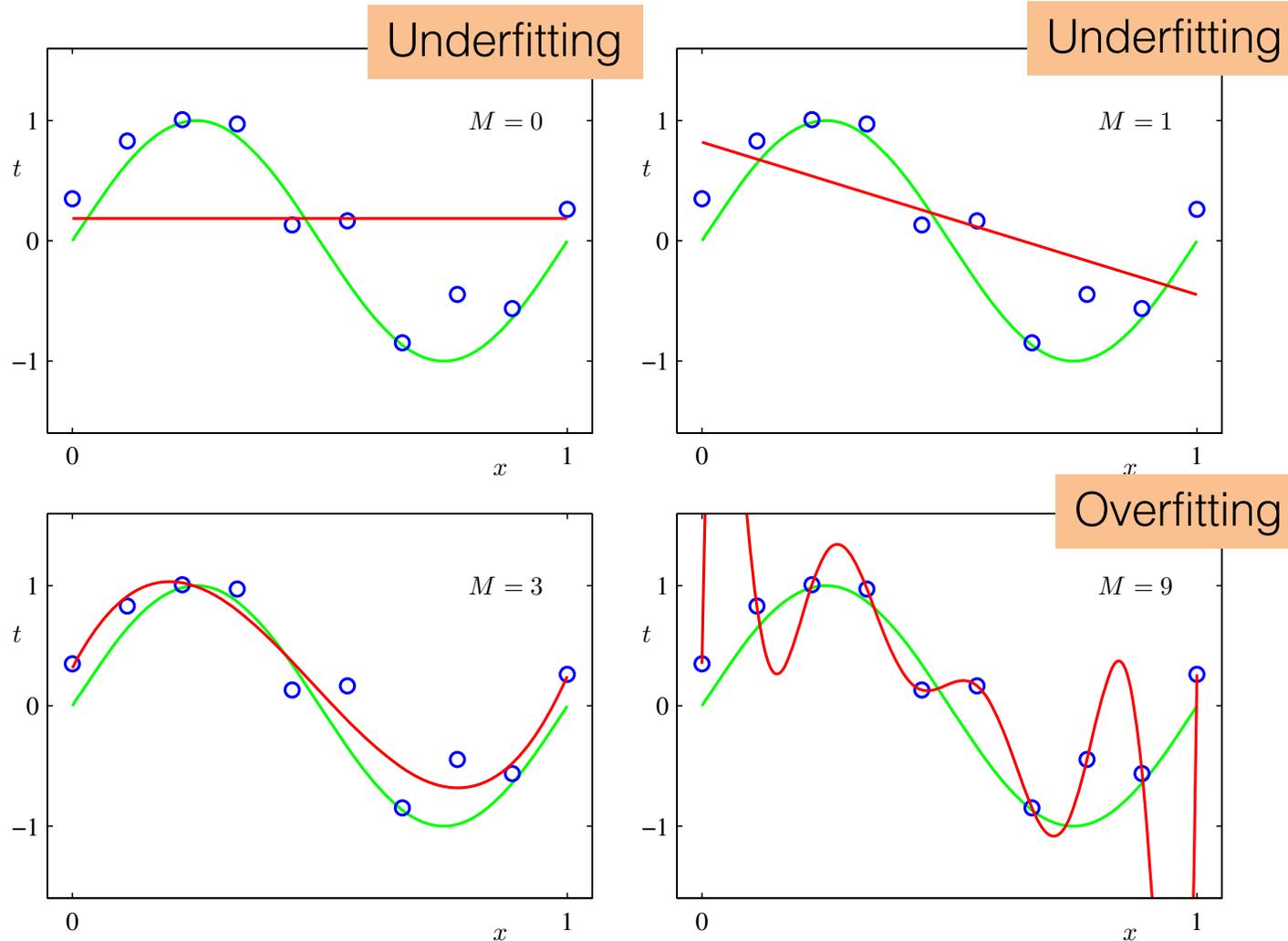
$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2$$



Linear derivative -> direct solution

Model selection

Which order M for the polynomial?



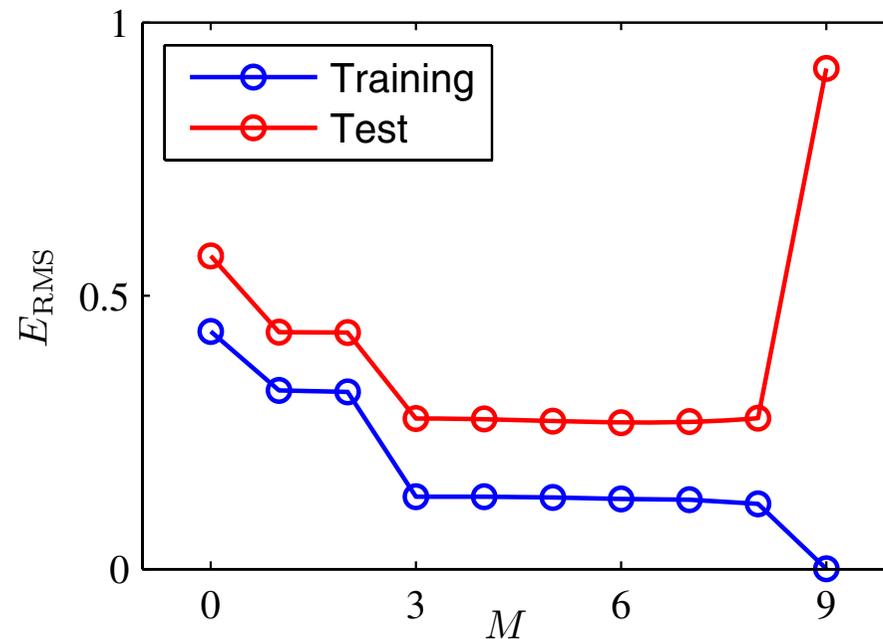
Model selection

Separation into (at least) two sets

- Training set
- Validation set (hold out set)

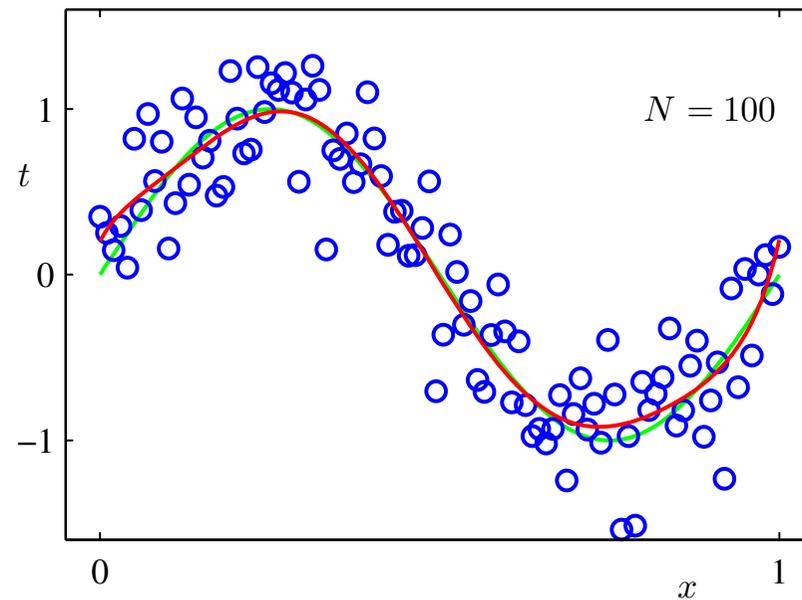
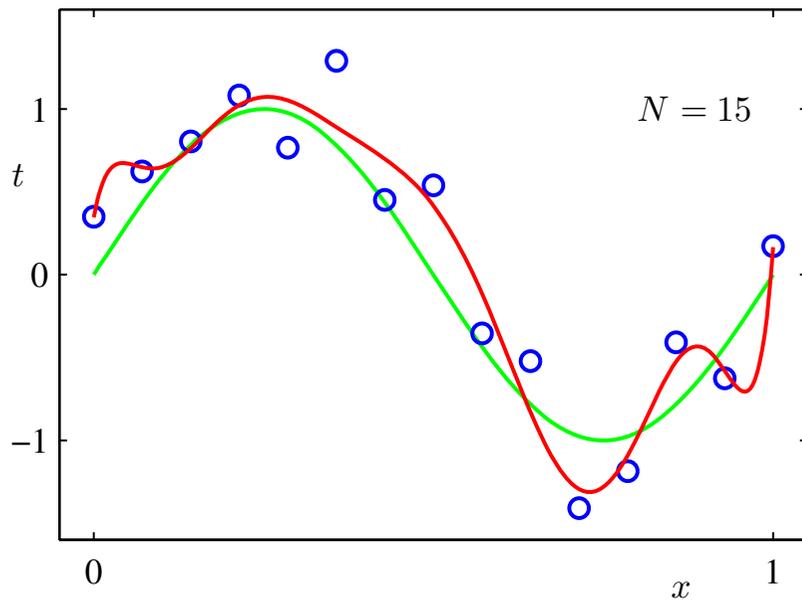
Root Mean Square Error (RMS)

$$E_{\text{RMS}} = \sqrt{2E(\mathbf{w}^*)/N}$$



Big Data!

Overfitting increases if we increase the size of the training set.



$M=9$

The 3 problems of Machine Learning

1. Expressivity

- What is the complexity of the functions my model can represent?

2. Trainability

- How easy is training of my model (i.e. solving the optimization problem)?

3. Generalization

- How does my model behave on unseen data?
- In presence of a shift in distributions?

(D'après Eric Jang & Jascha Sohl-Dickstein)

Learning formulations

Supervised learning — Labels y^* are available during training:

$$\hat{\theta} = \min_{\theta} \mathcal{L}(h(x, \theta), y^*)$$

Unsupervised learning — no labels, discovery of regularities in the data. Different objectives are possible.

Self-supervised learning — prediction of masked parts of the data itself, for instance the future:

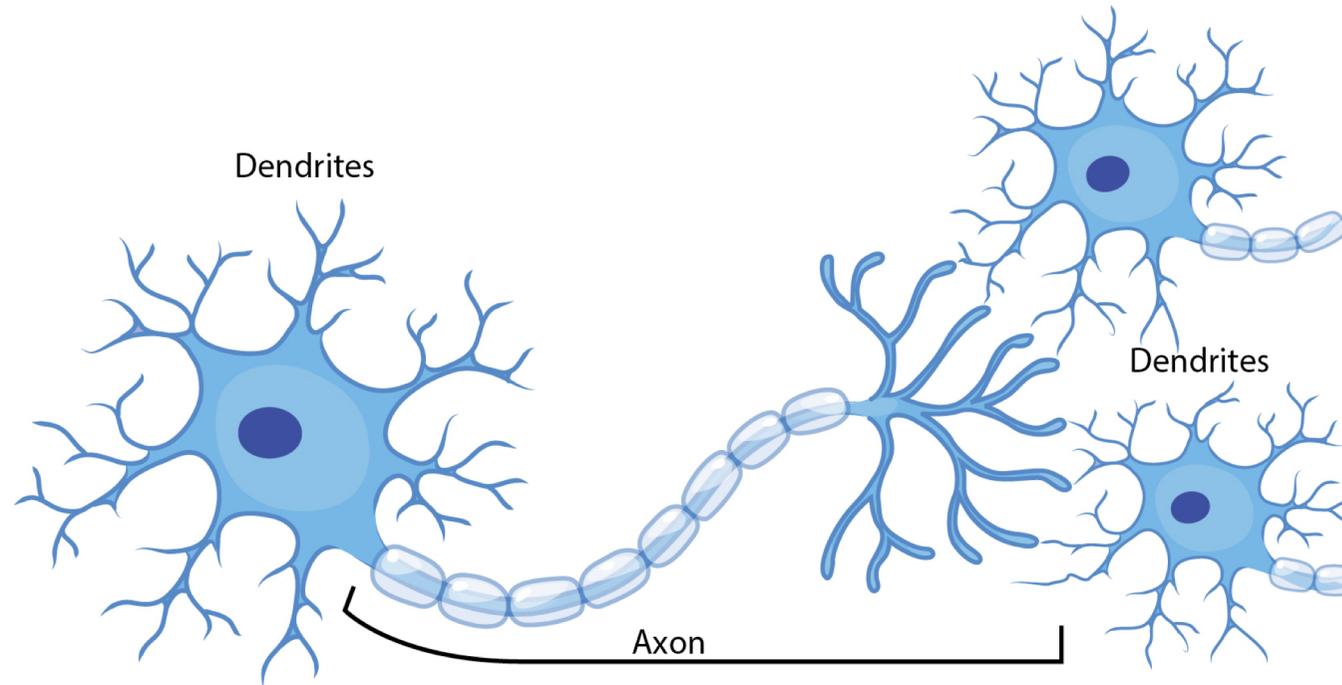
$$\hat{\theta} = \min_{\theta} \mathcal{L}(h(x_{t-\Delta:t-1}, \theta), x_t)$$

⇒ Pretraining step, usually followed by task oriented training.

Reinforcement learning — learning from interactions, maximizing the cumulated reward R over a horizon:

$$\hat{\theta} = J(\pi_{\theta}) = \mathbb{E}_{\tau \sim \pi_{\theta}} [R(\tau)]$$

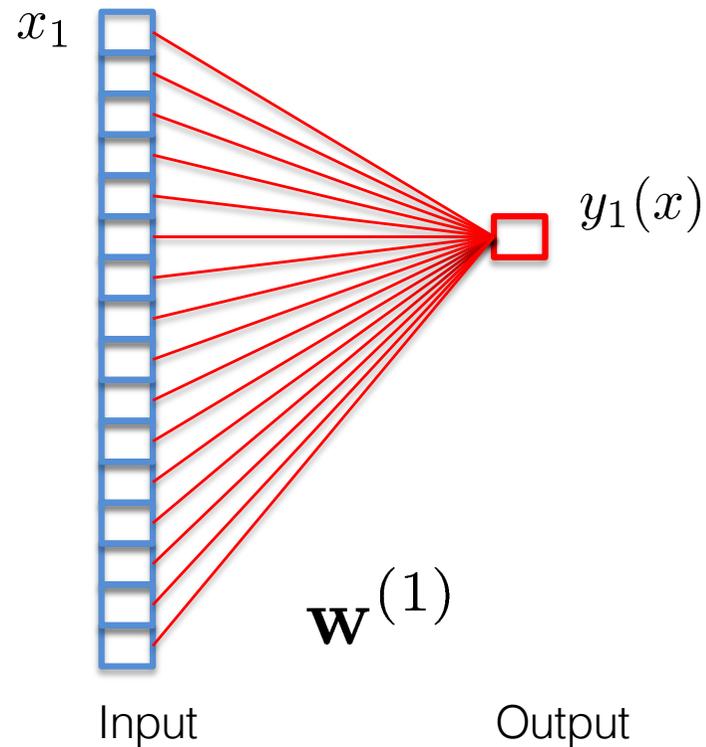
Biological neurons



Devin K. Phillips

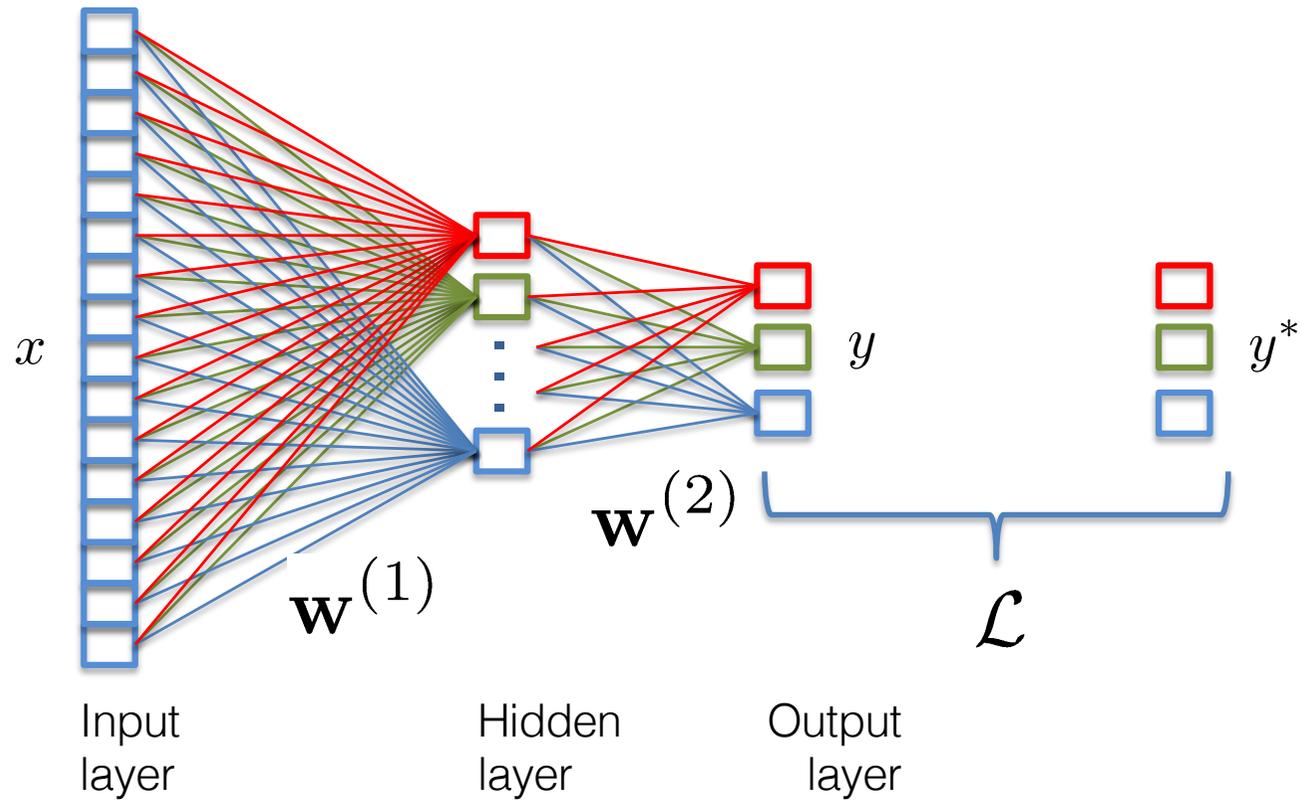
Neural networks

« Perceptron »

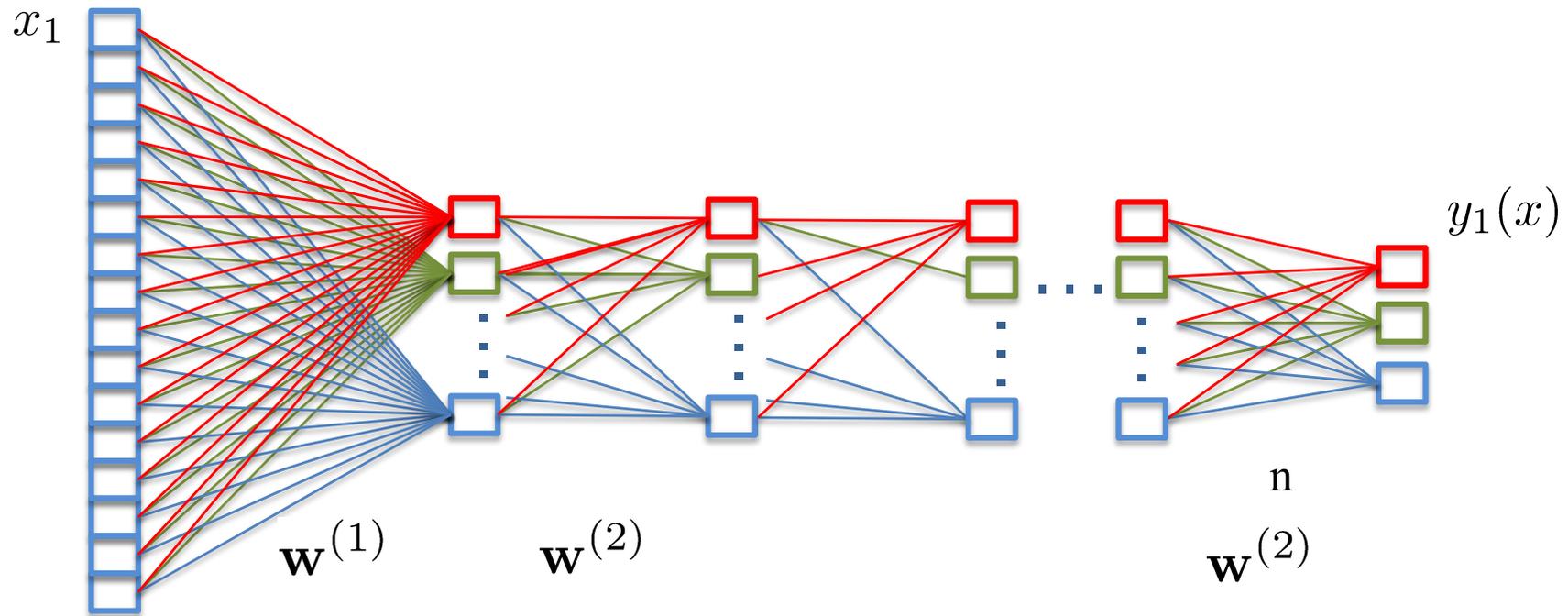


$$y(\mathbf{x}, \mathbf{w}) = \sum_{i=0}^D \mathbf{w}_i \mathbf{x}_i$$

Deep neural networks



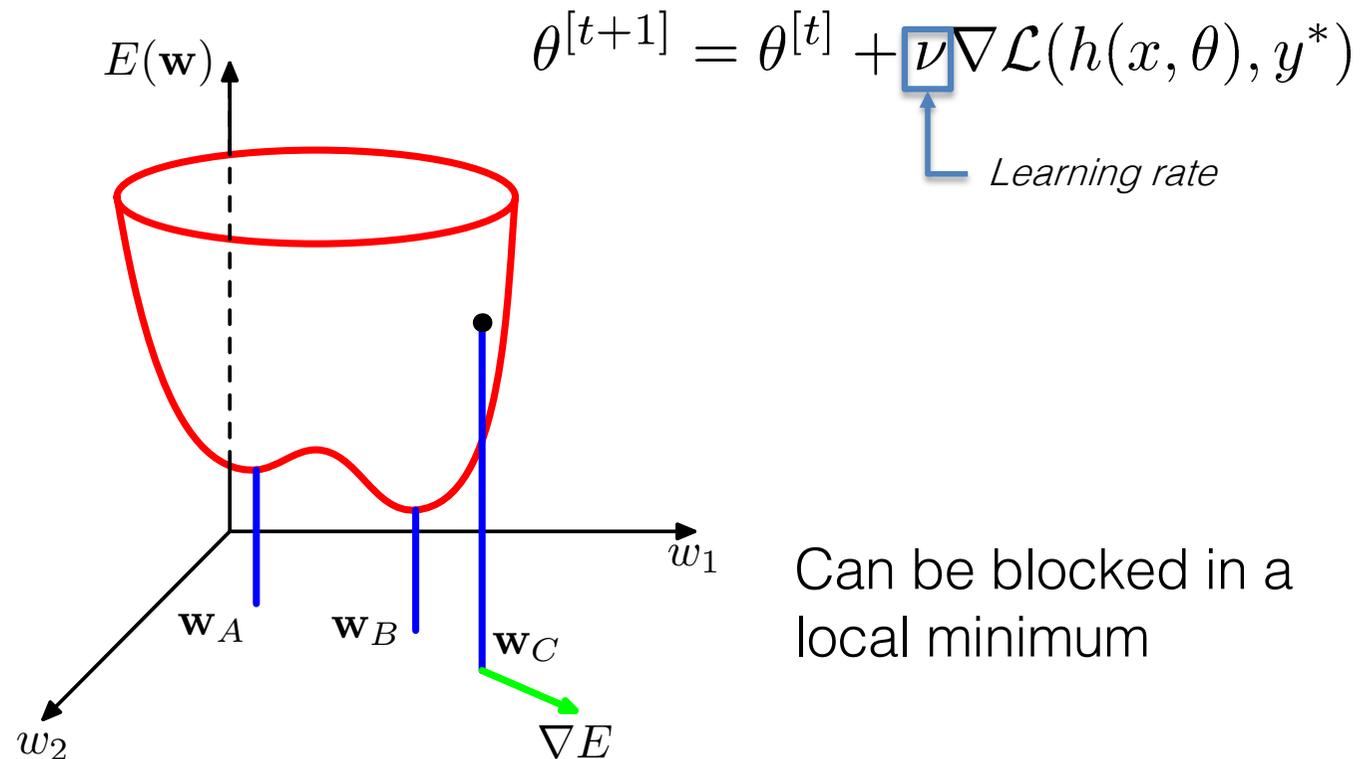
Deep neural networks



Gradient descent

Minimize the error on known data

"Empirical Risk Minimization"



Demo session: Tensorflow playground

Tensorflow Playground

The screenshot displays the TensorFlow Playground interface with the following components:

- DATA:** A section for selecting a dataset and adjusting parameters. It includes a "REGENERATE" button, a "Ratio of training to test data: 50%" slider, a "Noise: 0" slider, and a "Batch size: 10" slider.
- FEATURES:** A section for selecting input features. The selected features are X_1 and X_2 , each represented by a blue and orange bar. Other available features include X_1^2 , X_2^2 , X_1X_2 , $\sin(X_1)$, and $\sin(X_2)$.
- HIDDEN LAYERS:** A control for the number of hidden layers, currently set to 0. It includes "+" and "-" buttons.
- OUTPUT:** A scatter plot showing the model's output. The plot is divided into two regions: orange on the left and blue on the right. The axes range from -6 to 6. A legend below the plot indicates that colors represent data, neuron, and weight values. The legend also includes checkboxes for "Show test data" and "Discretize output", both of which are checked.
- Losses:** The output section displays "Test loss 0.516" and "Training loss 0.496".

<https://playground.tensorflow.org>