SYDE 556/750

Simulating Neurobiological Systems
Lecture 8: Learning

Andreas Stöckel

February 25 & 27, 2020

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Supervised Learning

$\mathbf{x}_1 =$ $\mathbf{t}_1 = (1, 0, 0)$

$\mathbf{x}_2 =$ $\mathbf{t}_2 = (1, 0, 0)$

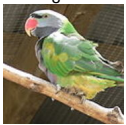
$\mathbf{x}_3 =$ $\mathbf{t}_3 = (1, 0, 0)$

$\mathbf{x}_4 =$ $\mathbf{t}_4 = (0, 0, 1)$

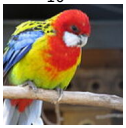
$\mathbf{x}_5 =$ $\mathbf{t}_5 = (1, 0, 0)$

$\mathbf{x}_6 =$ $\mathbf{t}_6 = (0, 0, 1)$

$\mathbf{x}_7 =$ $\mathbf{t}_7 = (1, 0, 0)$

$\mathbf{x}_8 =$  $\mathbf{t}_8 = (0, 0, 1)$

$\mathbf{x}_9 =$ $\mathbf{t}_9 = (0, 1, 0)$

$\mathbf{x}_{10} =$  $\mathbf{t}_{10} = (0, 0, 1)$

$\mathbf{x}_{11} =$ $\mathbf{t}_{11} = (1, 0, 0)$

$\mathbf{x}_{12} =$ $\mathbf{t}_{12} = (1, 0, 0)$

$\mathbf{x}_{13} =$ $\mathbf{t}_{13} = (0, 1, 0)$

$\mathbf{x}_{14} =$ $\mathbf{t}_{14} = (0, 1, 0)$

$\mathbf{x}_{15} =$ $\mathbf{t}_{15} = (0, 1, 0)$

$\mathbf{x}_{16} =$ $\mathbf{t}_{16} = (0, 1, 0)$

$\mathbf{x}_{17} =$ $\mathbf{t}_{17} = (0, 0, 1)$

$\mathbf{x}_{18} =$ $\mathbf{t}_{18} = (0, 1, 0)$

$\mathbf{x}_{20} =$ $\mathbf{t}_{20} = (0, 1, 0)$

$\mathbf{x}_{21} =$ $\mathbf{t}_{21} = (0, 0, 1)$

$\mathbf{x}_{22} =$ $\mathbf{t}_{22} = (0, 0, 1)$

$\mathbf{x}_{23} =$ $\mathbf{t}_{23} = (1, 0, 0)$

$\mathbf{x}_{24} =$ $\mathbf{t}_{24} = (0, 1, 0)$

Supervised Learning – Training

$\mathbf{x}_{10} =$

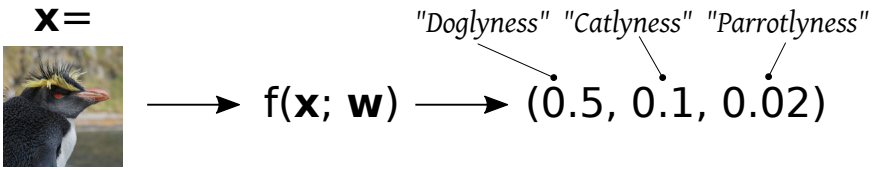
$\rightarrow$ $f(\mathbf{x}; \mathbf{w})$ $\rightarrow$ (0.1, 0.9, 0.25)

"Doglyness" *"Catlyness"* *"Parrotlyness"*

weight adaptation

$\mathbf{t}_{10} = (0.0, 0.0, 1.0)$

Supervised Learning – Inference



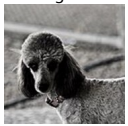
Unsupervised Learning

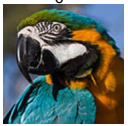
$\mathbf{x}_1 =$

$\mathbf{x}_2 =$

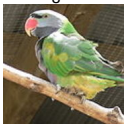
$\mathbf{x}_3 =$

$\mathbf{x}_4 =$

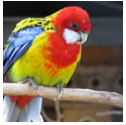
$\mathbf{x}_5 =$ 

$\mathbf{x}_6 =$ 

$\mathbf{x}_7 =$

$\mathbf{x}_8 =$ 

$\mathbf{x}_9 =$

$\mathbf{x}_{10} =$ 

$\mathbf{x}_{11} =$

$\mathbf{x}_{12} =$

$\mathbf{x}_{13} =$

$\mathbf{x}_{14} =$

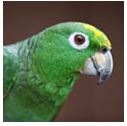
$\mathbf{x}_{15} =$

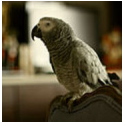
$\mathbf{x}_{16} =$

$\mathbf{x}_{17} =$

$\mathbf{x}_{18} =$

$\mathbf{x}_{20} =$

$\mathbf{x}_{21} =$ 

$\mathbf{x}_{22} =$ 

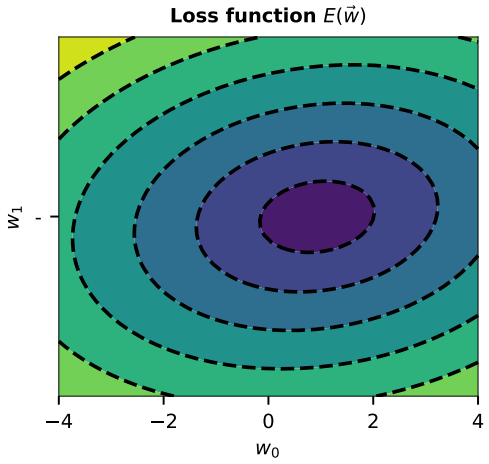
$\mathbf{x}_{23} =$

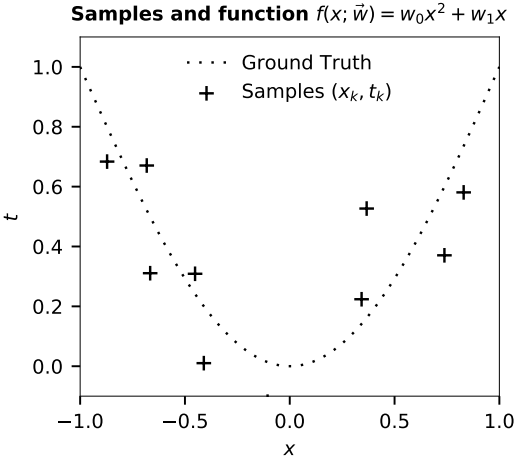
$\mathbf{x}_{24} =$

Unsupervised Learning – Training

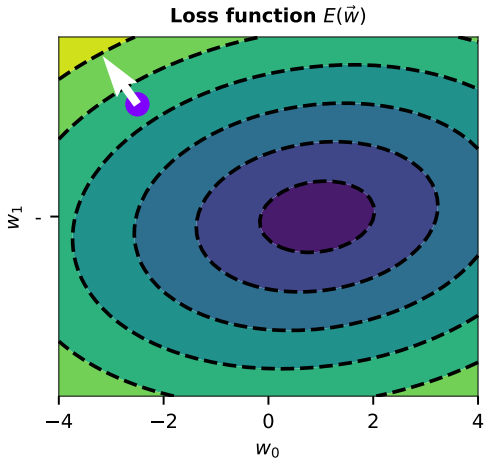
Unsupervised Learning – Inference

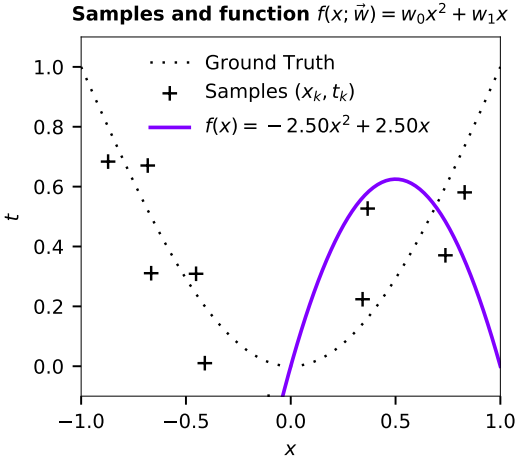
Gradient Descent – Example



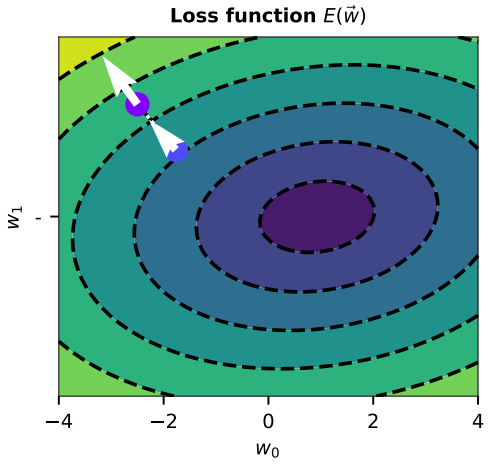


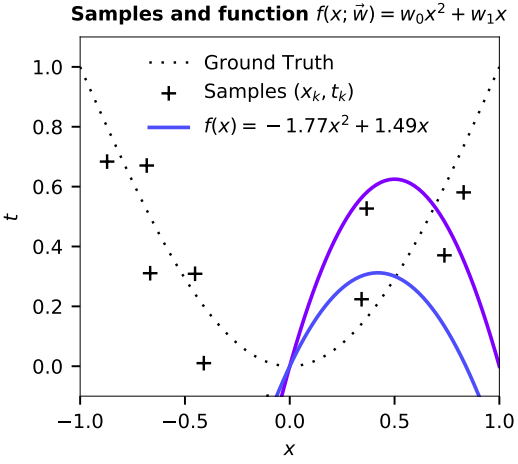
Gradient Descent – Example



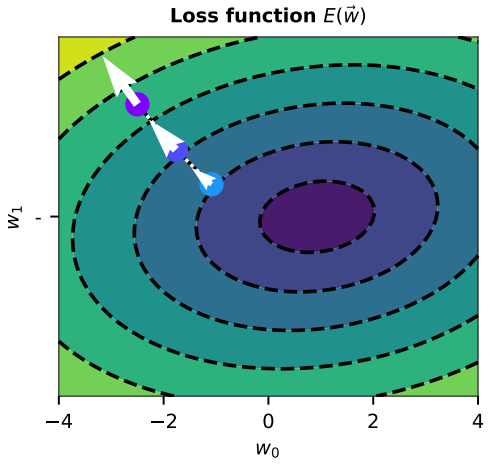


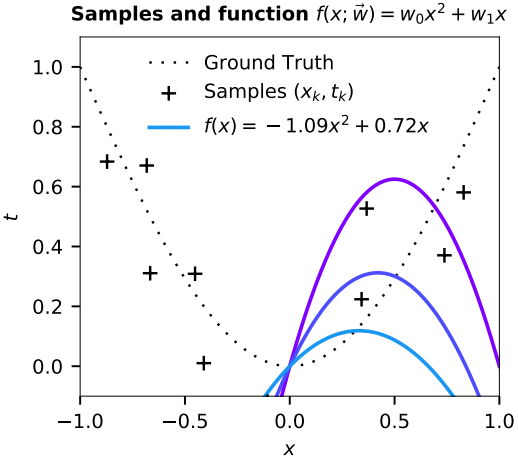
Gradient Descent – Example



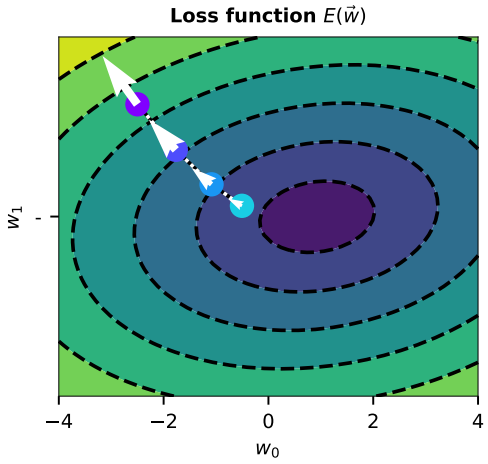


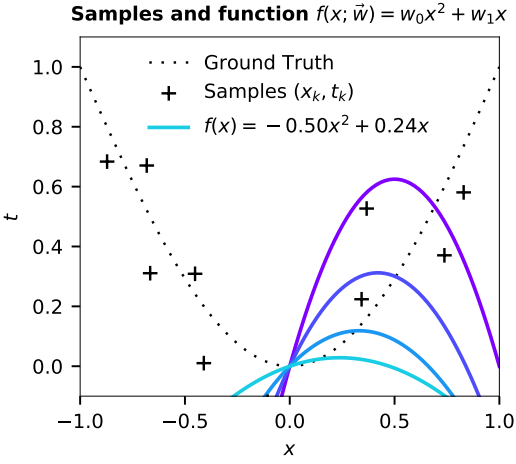
Gradient Descent – Example



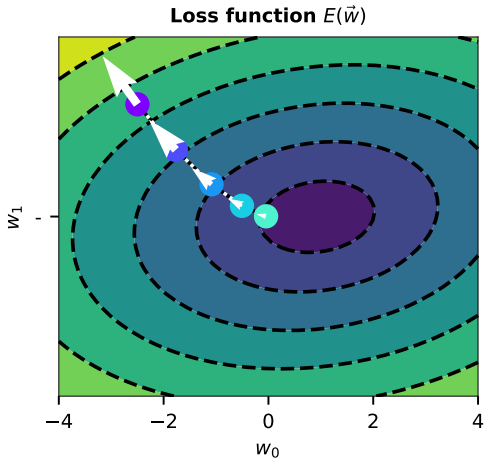


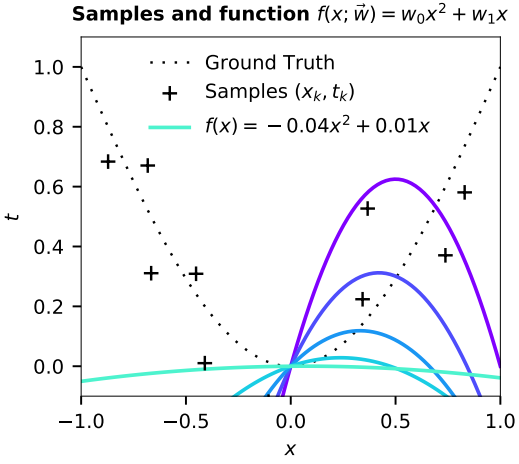
Gradient Descent – Example



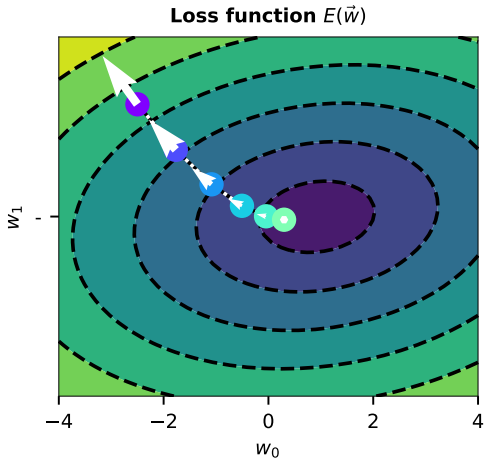


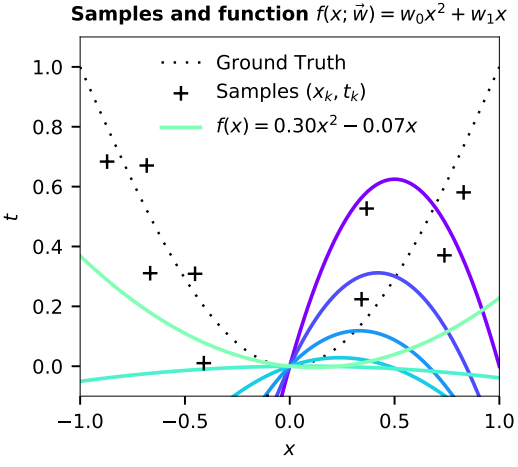
Gradient Descent – Example



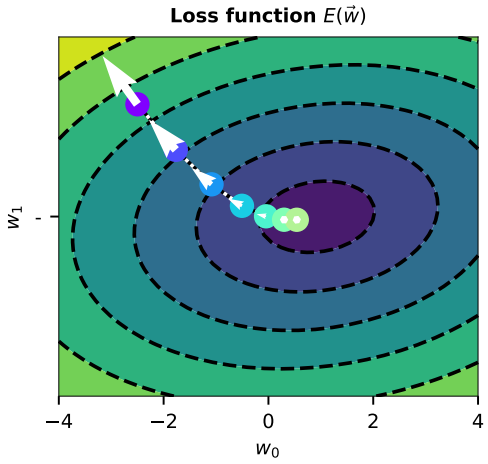


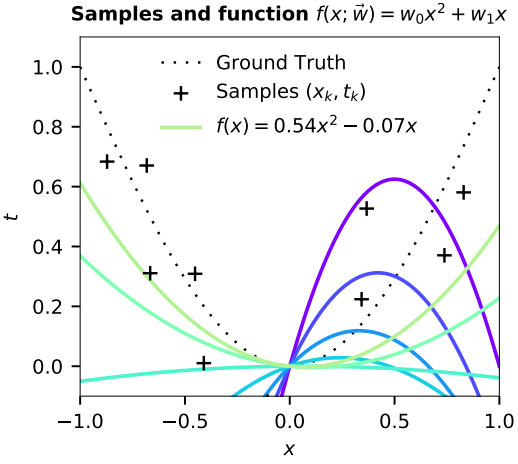
Gradient Descent – Example



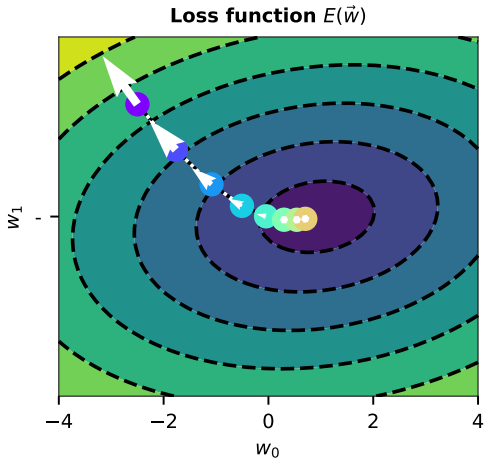


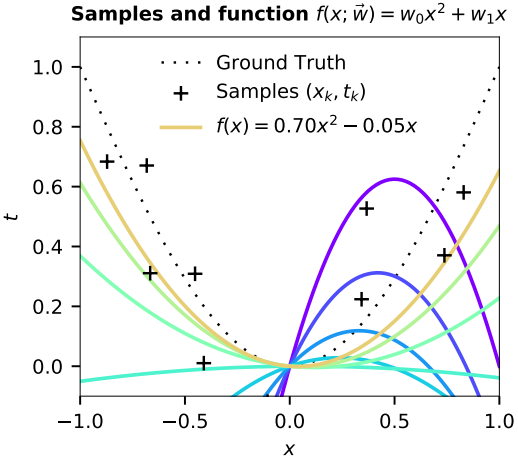
Gradient Descent – Example



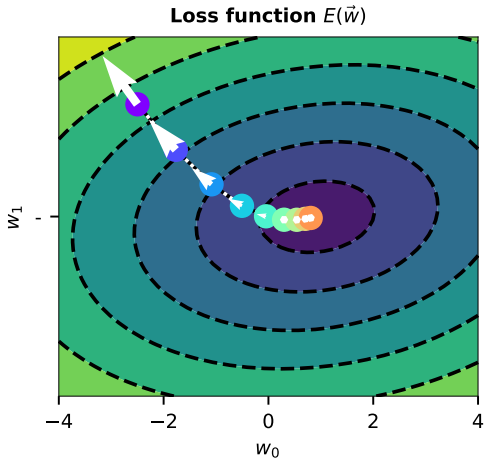


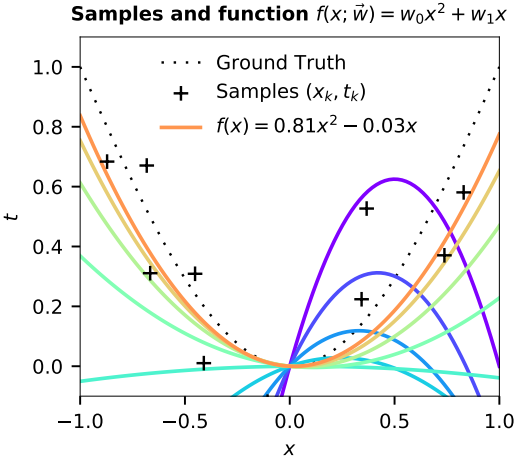
Gradient Descent – Example



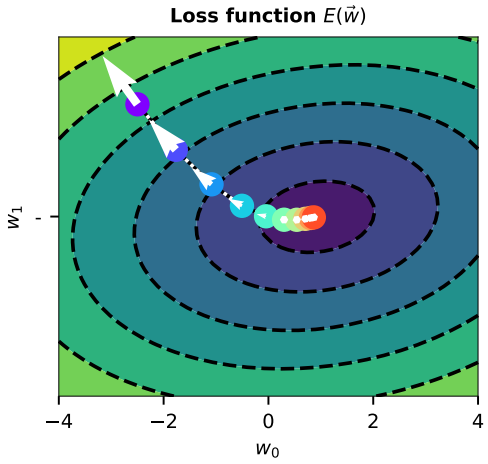


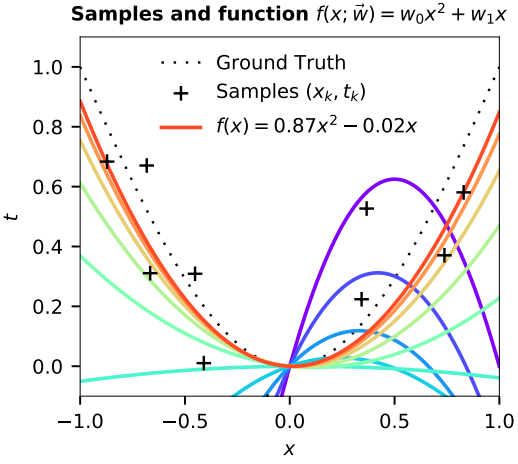
Gradient Descent – Example



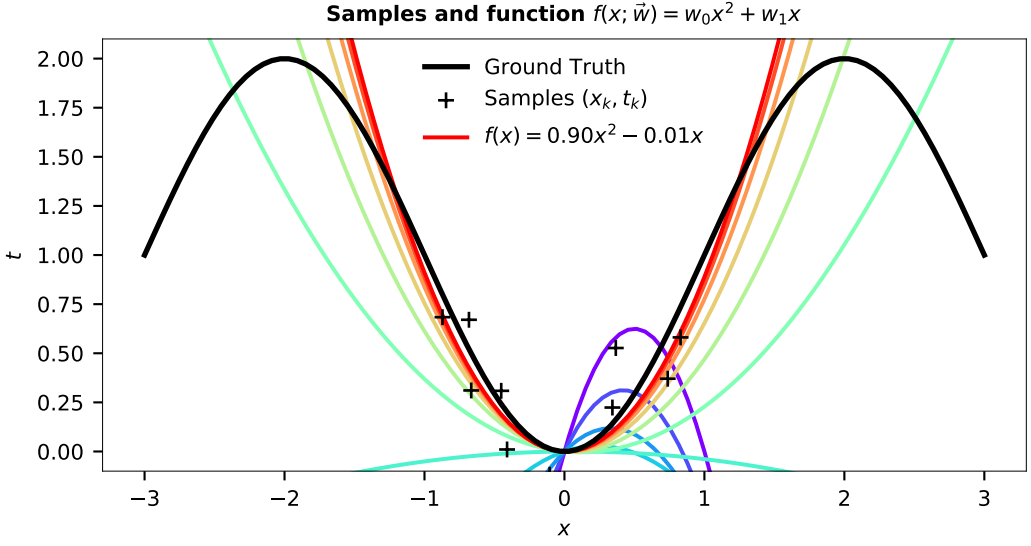


Gradient Descent – Example



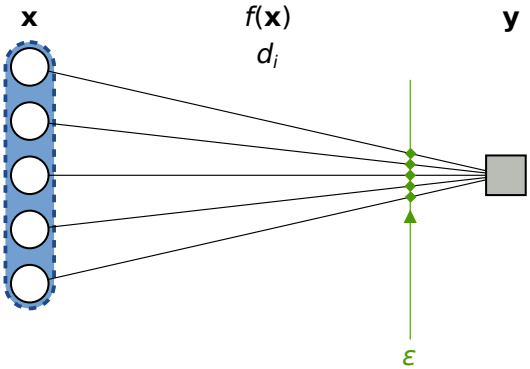


Supervised Learning – Generalisation



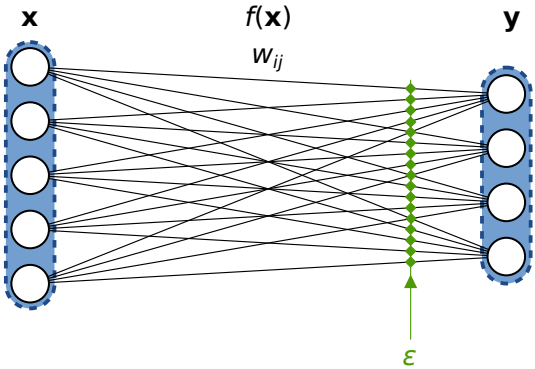
Learning Decoders and Learning Weights

Learning Decoders (Delta Rule)



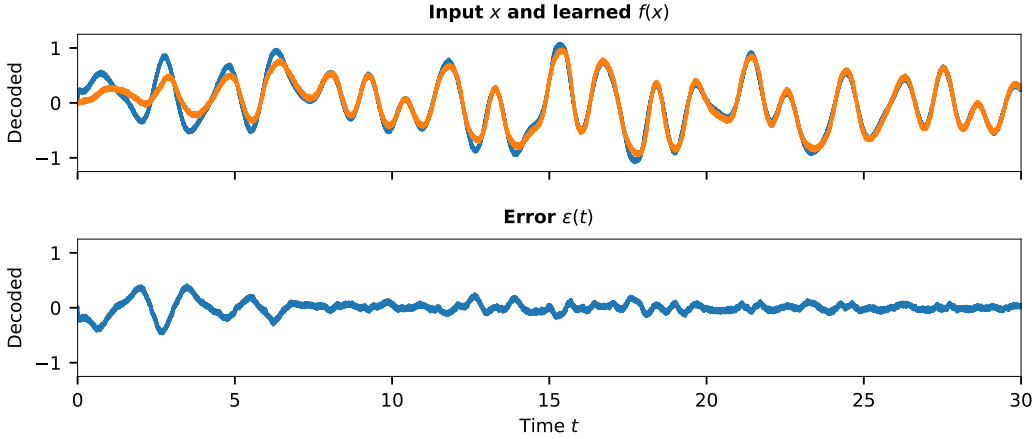
$$\Delta d_i = -\eta a_i(\mathbf{x}) \underbrace{\left(y(t) - y^{\mathrm{d}}(t)\right)}_{\varepsilon(t)}$$

Learning Weights (PES Rule)

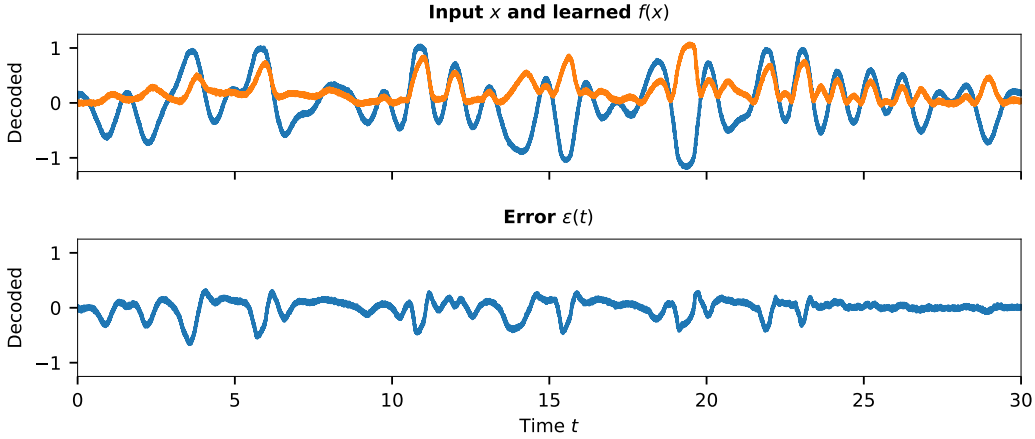


$$\Delta w_{ij} = -\eta a_i(\mathbf{x}) \Big(\alpha_j \langle \mathbf{e}_j, \varepsilon(t) \rangle \Big)$$

Example: Learning a Communication Channel



Example: Learning $f(x) = x^2$



Example: Adaptive Controller

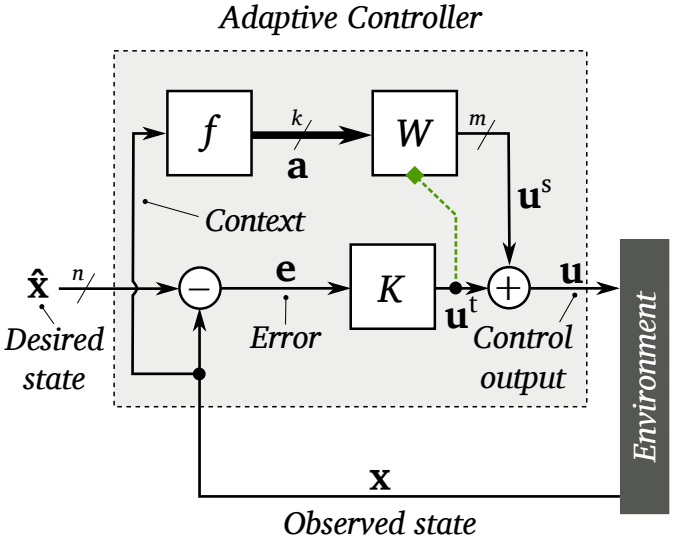


Image sources

Title slide
Page from “Liber ethicorum des Henricus de Alemannia”. Title: “Henricus de Alemannia con i suoi studenti” (Henricus of Germany with his students), second half of 14th century.
From Wikimedia.