

Intro

Exercise 1 : Machine Learning (general) File: ex-en-ml-introduction004

(a) Define the terms “supervised learning”, “unsupervised learning”, and “reinforcement learning”.

Answer

- Supervised: Learn a function from a set of input-output pairs
- Unsupervised: Identify structures in unlabeled data.
- Reinforcement: Maximizing rewards

(b) Sketch for each learning paradigm a typical problem together with a description of its technical realization.

Answer

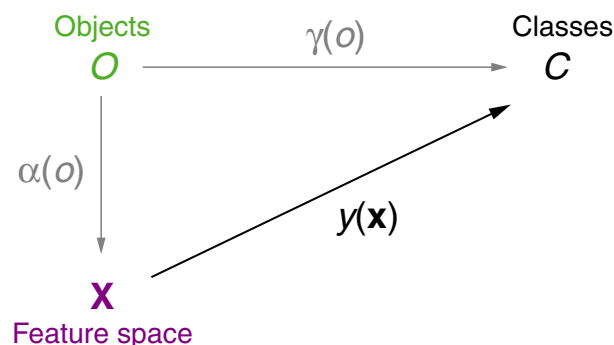
- Supervised learning. Text classification, Spam filtering. - given a database of emails, each annotated as “spam” or “not spam”, learn a function that classifies new emails
- Unsupervised learning. Sort fruits in a basket, clear up a writing-desk, compression of data.
- Reinforcement learning. Replace a light bulb, touch a hot hot plate, drink alcohol. (Simpsons episode: Homer wants to tap a telephone line)

Exercise 2 : Specification of Learning Tasks File: ex-en-ml-introduction001

Answer

Explanation: <https://youtu.be/QAnxg2IzmNU>

The following picture from the lecture slides describes the relationship between Real World and Model World, when it comes to the specification of learning tasks.



Assume you are building a machine learning system that predicts whether a given mushroom is poisonous or edible. For the following list, decide which symbol from the picture most closely matches the given list item:

- (a) A pile of Mushrooms.
- (b) A table with the columns “size”, “weight”, and “color”, as well as one row for each possible mushroom, and the respective measurements in the cells.

- (c) A human mushroom expert who can tell whether any mushroom you show them is poisonous or edible.
- (d) A device that measures size, weight and color of a mushroom.
- (e) The set {Poisonous, Edible}
- (f) The machine learning system that you are trying to build.

Answer

(a)  $O$ , (b)  $\mathbf{X}$ , (c)  $\gamma(o)$ , (d)  $\alpha(o)$ , (e)  $C$ , (f)  $y(\mathbf{x})$ .

Exercise 3 : Linear Algebra File: ex-en-ml-prerequisites001

Answer

Explanation: <https://youtu.be/PMadbrOEBF0>

The three  $(2, 3)$ -matrices  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  are given as

$$\mathbf{A} = \begin{bmatrix} 3 & 5 & 2 \\ 1 & 4 & 0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 2 & 1 & 2 \\ 5 & 8 & -1 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & -1 & 2 \\ 5 & -2 & 0 \end{bmatrix}.$$

Evaluate the following expressions:

(a)  $3\mathbf{A} + 2\mathbf{B} - 5\mathbf{C}$

Answer

$$\begin{aligned} 3\mathbf{A} + 2\mathbf{B} - 5\mathbf{C} &= 3 \begin{bmatrix} 3 & 5 & 2 \\ 1 & 4 & 0 \end{bmatrix} + 2 \begin{bmatrix} 2 & 1 & 2 \\ 5 & 8 & -1 \end{bmatrix} - 5 \begin{bmatrix} 1 & -1 & 2 \\ 5 & -2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 9 & 15 & 6 \\ 3 & 12 & 0 \end{bmatrix} + \begin{bmatrix} 4 & 2 & 4 \\ 10 & 16 & -2 \end{bmatrix} + \begin{bmatrix} -5 & 5 & -10 \\ -25 & 10 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 8 & 22 & 0 \\ -12 & 38 & -2 \end{bmatrix} \end{aligned}$$

(b)  $3(\mathbf{B} - \mathbf{A})^T - \mathbf{C}^T$

Answer

$$\begin{aligned} 3(\mathbf{B} - \mathbf{A})^T - \mathbf{C}^T &= 3 \left( \begin{bmatrix} 2 & 1 & 2 \\ 5 & 8 & -1 \end{bmatrix} - \begin{bmatrix} 3 & 5 & 2 \\ 1 & 4 & 0 \end{bmatrix} \right)^T - \begin{bmatrix} 1 & -1 & 2 \\ 5 & -2 & 0 \end{bmatrix}^T \\ &= 3 \begin{bmatrix} -1 & -4 & 0 \\ 4 & 4 & -1 \end{bmatrix}^T - \begin{bmatrix} 1 & -1 & 2 \\ 5 & -2 & 0 \end{bmatrix}^T \\ &= 3 \begin{bmatrix} -1 & 4 \\ -4 & 4 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 5 \\ -1 & -2 \\ 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -3 & 12 \\ -12 & 12 \\ 0 & -3 \end{bmatrix} + \begin{bmatrix} -1 & -5 \\ 1 & 2 \\ -2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -4 & 7 \\ -11 & 14 \\ -2 & -3 \end{bmatrix} \end{aligned}$$

(c)  $\mathbf{A} \cdot \mathbf{C}^T$

Answer

$$\begin{aligned}\mathbf{A} \cdot \mathbf{C}^T &= \begin{bmatrix} 3 & 5 & 2 \\ 1 & 4 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 5 & -2 & 0 \end{bmatrix}^T \\ &= \begin{bmatrix} 3 & 5 & 2 \\ 1 & 4 & 0 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ -1 & -2 \\ 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 5 \\ -3 & -3 \end{bmatrix}\end{aligned}$$

Exercise 4 : Calculus File: ex-en-ml-prerequisites002

Answer

Explanation: <https://youtu.be/FbPFxOL5m4s>

Calculate the first-order partial derivatives  $\frac{\partial z}{\partial x}$  and  $\frac{\partial z}{\partial y}$  of the following functions:

(a)  $z = (2x - 3y^2)^5$

Answer

$$z = u^5 \text{ with } u = 2x - 3y^2$$

$$\begin{aligned}\frac{\partial z}{\partial x} &= \frac{\partial z}{\partial u} \cdot \frac{\partial u}{\partial x} = 5u^4 \cdot 2 = 10u^4 = 10(2x - 3y^2)^4 \\ \frac{\partial z}{\partial y} &= \frac{\partial z}{\partial u} \cdot \frac{\partial u}{\partial y} = 5u^4 \cdot (-6y) = -30yu^4 = -30y(2x - 3y^2)^4\end{aligned}$$

(b)  $z = x^2 \cdot e^{-xy}$

Answer

$$\begin{aligned}z &= uv \text{ with } u = x^2, v = e^{-xy} \\ \frac{\partial u}{\partial x} &= 2x \\ v &= e^t \text{ with } t = -xy \\ \frac{\partial v}{\partial x} &= \frac{\partial v}{\partial t} \cdot \frac{\partial t}{\partial x} = e^t \cdot (-y) = -y \cdot e^{-xy} \\ \frac{\partial z}{\partial x} &= u_x v + v_x u = 2x \cdot e^{-xy} - y \cdot e^{-xy} \cdot x^2 = (2x - x^2 y) \cdot e^{-xy} \\ \frac{\partial z}{\partial y} &= \frac{\partial z}{\partial t} \cdot \frac{\partial t}{\partial y} = x^2 \cdot e^t \cdot (-x) = -x^3 \cdot e^{-xy}\end{aligned}$$

Exercise 5 : Gradient Descent File: ex-en-ml-regression001

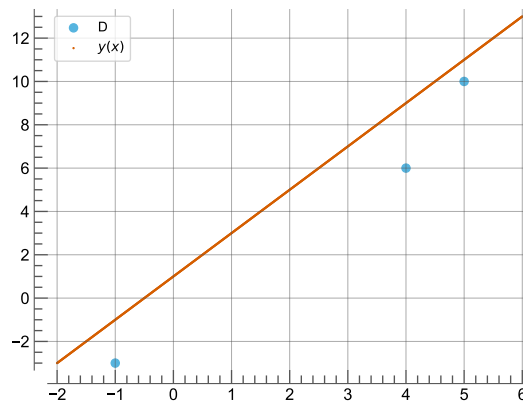
Answer

Explanation: <https://youtu.be/xuyqzgxYrfE>

In this exercise you will be calculating one iteration of the LMS algorithm. The task is to fit a straight line given as  $y(x) = w_0 + w_1 x$  through a given set of points  $D$  by adjusting the parameters  $w_0$  and  $w_1$ . The pairs  $(\mathbf{x}, c) \in D$  are given as  $(4, 6), (-1, -3), (5, 10)$ .

- (a) Plot the line parametrized with  $w_0 := 1, w_1 := 2$  as well as all samples from  $D$  into the same coordinate system.

Answer



- (b) For the first iteration,  $(4, 6)$  is ‘randomly’ selected as the ground-truth pair  $(\mathbf{x}, c)$ . Compute the loss w.r.t.  $\mathbf{w}$ , which is defined as

$$l_2(c, y(\mathbf{x})) = \frac{1}{2} \cdot (c - y(\mathbf{x}))^2.$$

Answer

$$\begin{aligned} l_2(c, y(\mathbf{x})) &= \frac{1}{2} \cdot (c - y(\mathbf{x}))^2 \\ &= \frac{1}{2} \cdot (c - (w_0 + w_1 \cdot x))^2 \\ &= \frac{1}{2} \cdot (6 - (w_0 + w_1 \cdot 4))^2 \\ &= \frac{1}{2} \cdot (6 - w_0 - 4w_1)^2 \end{aligned}$$

- (c) Derive the gradient

$$\begin{pmatrix} \frac{\partial l_2}{\partial w_0} \\ \frac{\partial l_2}{\partial w_1} \end{pmatrix}.$$

Verify that this is indeed  $-\delta \cdot \mathbf{x}$ .

Answer

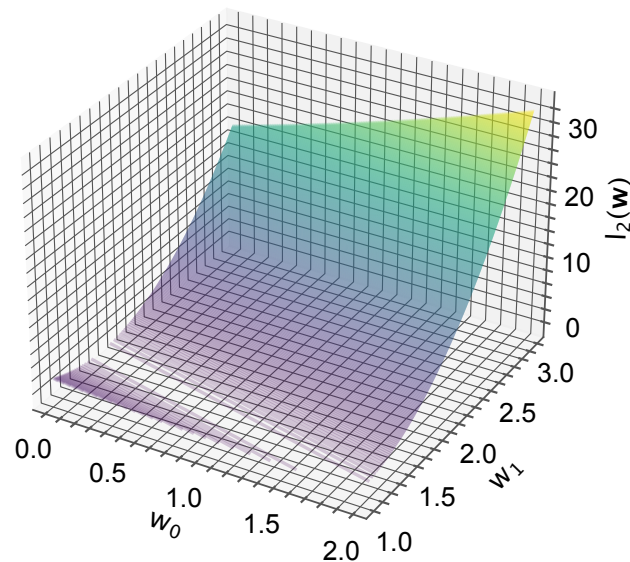
$$\begin{aligned}\frac{\partial l_2}{\partial w_0} &= \frac{1}{2} \cdot 2 \cdot (6 - w_0 - 4w_1) \cdot (-1) \\ &= -6 + w_0 + 4w_1\end{aligned}$$

$$\begin{aligned}\frac{\partial l_2}{\partial w_1} &= \frac{1}{2} \cdot 2 \cdot (6 - w_0 - 4w_1) \cdot (-4) \\ &= -24 + 4w_0 + 16w_1\end{aligned}$$

$$\delta = c - y(\mathbf{x}) = 6 - (w_0 + w_1 \cdot 4) = 6 - w_0 - 4w_1$$

$$\begin{aligned}-\delta \cdot \mathbf{x} &= \begin{pmatrix} -(6 - w_0 - 4w_1) \cdot x_0 \\ -(6 - w_0 - 4w_1) \cdot x_1 \end{pmatrix} \\ &= \begin{pmatrix} -(6 - w_0 - 4w_1) \cdot 1 \\ -(6 - w_0 - 4w_1) \cdot 4 \end{pmatrix} \\ &= \begin{pmatrix} -6 + w_0 + 4w_1 \\ -24 + 4w_0 + 16w_1 \end{pmatrix}\end{aligned}$$

(d) The following plot shows the loss landscape defined by  $l_2$  with the current choice of  $(\mathbf{x}, c) = (4, 6)$ .



(d1) The model is initialized with  $w_0 := 1, w_1 := 2$ . Mark the current location of the model in the loss landscape, i.e.,  $\begin{pmatrix} w_0 \\ w_1 \end{pmatrix}$ .

(d2) Draw the line of gradient descent, which is defined as

$$\underbrace{\left\{ \begin{pmatrix} w_0 \\ w_1 \end{pmatrix} - \eta \cdot \begin{pmatrix} \frac{\partial l_2}{\partial w_0} \\ \frac{\partial l_2}{\partial w_1} \end{pmatrix} \mid 0 \leq \eta \leq 0.03 \right\}}_{\Delta \mathbf{w} \text{ for increasing } \eta},$$

by connecting the start and end point from the set.

Answer

