Machine Learning, Lab Class ML:1, ML:2

14.11.2020

Exercise 1:

a.) supervised learning:

b.) - learning a function with a set of input-output pairs
- the data from which the function is learned is already tagged with the correct answer
- "supervised" because the correct answers are already given

Problem:

- The website Classification in Task 6
- At first collect the Training Data via manual annotation of the websites
- determine the features that should be used in the model
- train the model with the Traing Data (Algorithms could be SVM, Neural Networks, Regression, Classification trees)
- for new data the trained model can estimate a class as an output

unsupervised learning:

- identifying structures in unlabled/uncategorized data (e.g. clustering)
- only use input data to learn, no output data
- important subareas: - automated categorization (e.g. via cluster analysis)
  - parameter optimization (e.g. via expectation minimization)
  - feature extraction (e.g. via factor analysis)

Problem:

- detecting different classes of wine based on example data of many different wines with a couple of features
- e.g. via ~~[illegible]~~ Cluster analysis

Reinforced learning:

- learn / adapt / optimize a behaviour strategy in order to maximize the own benefit / reward by interpreting feedback from the environment
- basic reinforcement is modeled as a "Markov decision process":

Problem:

(- computer agents learning to be "friendly" in a conversation ?)
- learning behaviour strategies in hostile environments
- chess-bots, Go-Bots, elevator scheduling, robot control, checkers, backgammon, (games in general ?)

Exercise 2:

a.) a1) Objects O a2) Feature Space X

a3) γ a4) α

a5) Classes C a6) $c \approx y$

b.) 6a) γ 6b) α

Exercise 3:

a.) $y(x) = w_0 + w_1 \cdot x$

$$RSS(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2$$

$$\bar{x} = \frac{5+7+15+28}{4} = \frac{55}{4}$$

$$\bar{y} = \frac{50+79+124+300}{4} = \frac{553}{4}$$

$$\hat{w}_0 = \bar{y} - \hat{w}_1 \cdot \bar{x}$$

$$\hat{w}_0 = \frac{553}{4} - \hat{w}_1 \cdot \frac{55}{4}$$

$$\hat{w}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\hat{w}_1 = \frac{13837}{1307} \approx 10,6$$

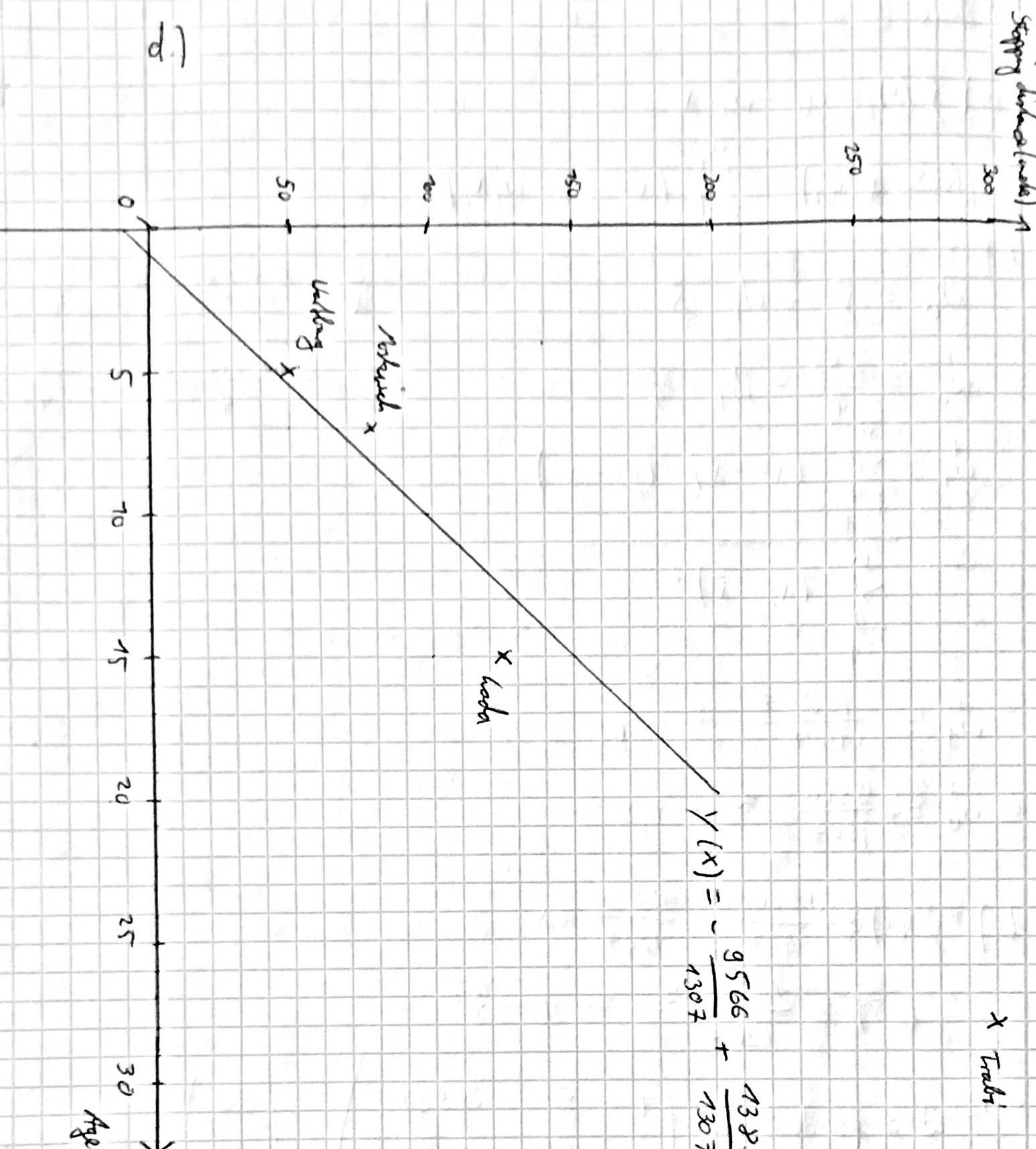
$$\hookrightarrow \hat{w}_0 = -\frac{9566}{1307} \approx -7,3$$

b.) $y(x) = -\frac{9566}{1307} + \frac{13837}{1307} \cdot x$

$$y(15) = \frac{197989}{1307} \approx 151,5$$

c.) $\hat{w} = \text{pseudo inverse} \begin{pmatrix} 1 & 5 & 30530 \\ 1 & 7 & 90000 \\ 1 & 15 & 159899 \\ 1 & 28 & 270564 \end{pmatrix} \cdot \begin{pmatrix} 50 \\ 79 \\ 124 \\ 300 \end{pmatrix}$

d.)



e.) - can't model the exact searched function with linear regression, only a linear estimate

- one should be aware that the extrapolations can be different from the real values (searched function could be much more complex and depending on many more features)

4a) with EXAMPLE : $(x, c(x))$ // concept is attribute of EXAMPLE

$$u(p) = m_1 \cdot \ldots \cdot m_p \cdot 2$$ // m_i is number of values for an attribute

$$= \prod_{i=1}^{p} m_i \cdot 2$$ // concept $(c(x))$ has 2 values

4b) syntactically distinct hypothesis $= \prod_{i=1}^{p} (m_i + 2)$

semantically distinct hypothesis $= 1 + \prod_{i=1}^{p} (m_i + 1)$

// a hypothesis is described ~~as~~ without concept, but ~~$\langle a_1, x, a \ldots, a_p, x \rangle$~~ as a tuple ~~of~~ of the ~~with $x \in \{1, \ldots, m_i\}$~~ the ~~remaing~~ ~~remaining~~ other feature values (vgl. page 9 < sunny, ?, ?, strong, ?, same ?)

// 1 because of contradiction „⊥"

// $m_i + 1$ because of „?"

4c) $\begin{cases} u(1) = m_1 \cdot 2 \\ u(p+1) = u(p) \cdot m_{p+1} \end{cases}$

$\begin{cases} |H_1| = m_1 + 2 \\ \cancel{H_p +} \; |H_{p+1}| = ((|H_p| - 1) \cdot (m_{p+1} + 1)) + 1 \end{cases}$

5a) consistent hypothesis: $h(x) = c(x)$

Inconsistency is caused, when $c(x) = 0$, then the inner if-cause isn't ~~entered and~~ reached and the chance of correcting an $h(x) = 1$ is zero.

• If $c(x) = 0$ ~~them~~ then if $h(x) = 0$ consistent

$h(x) = 1$ inconsistent, no chance of making it consistent (with min-generalisation (h, x)

[illegible] example (positiv)

5b.)

No, the order of the examples is not important.

There are 4 cases when looking at an attribute.

(writen like: (attribute value hypothesis ,attribute value example) -> attribute value hypothesis,

a,b and c are examples for literals)

1) most specific to less specific (⊥,a) -> a

2) same attribute-values: (a,a)-> a

3) different attribute values : (a,b) -> ?

4) most general and a any attribute value (?,c)-> ?

A) Given more than one example (- the first example changes the hypothesis to a copy of itself-),

the first value different to the value of the first example, will cause a change to ? (3))

everything afterwards is of no importance (4)).

-> not dependent on order because if there is a differnent value it will be taken into account at any time

B) If there are no different attribute values in the examples, the initial value will persist. (2))

-> not dependent on order, cause the attribute values are equal

So the outcome is eather A) or B) and there is no Order