

# Machine Learning I: Foundations

## Exercise Sheet 0

Prof. Marius Kloft

TA: Billy Joe Franks

27.04.2021

Deadline: 07.05.2020

### 1) (MANDATORY) 5 Points

- Please organize yourself in groups of up to 3 people.
- You must hand in L<sup>A</sup>T<sub>E</sub>X compiled pdfs, unless the mandatory parts only include programming exercises. Then you may hand in only the .ipynb files. If the mandatory tasks are mixed, please zip your pdf and .ipynb files for upload on OLAT.
- Put the full names of all participants onto the solutions.
- Do not copy solutions from another group.
- You may work with people outside of your group, however in any case your group should write its own solutions.

- 2) a) The scalar projection,  $\Pi_{\mathbf{w}}(\mathbf{x})$ , of a vector  $\mathbf{x}$  onto another vector  $\mathbf{w}$  is defined as the coordinate of the orthonormal projection of  $\mathbf{x}$  onto a line parallel to  $\mathbf{w}$ , or  $\text{span}\{\mathbf{w}\}$ . Assume  $\mathbf{b} := \mathbf{x} - a_1 \frac{\mathbf{w}}{\|\mathbf{w}\|}$  and  $\mathbf{b}^T \mathbf{w} = 0$  then  $\Pi_{\mathbf{w}}(\mathbf{x}) := a_1$ . Show that the scalar projection of vector  $\mathbf{x}$  onto  $\mathbf{w}$  is equal to:

$$\Pi_{\mathbf{w}}(\mathbf{x}) := \frac{1}{\|\mathbf{w}\|} \mathbf{w}^T \mathbf{x}.$$

- b) Let  $f(\mathbf{x}) := \mathbf{w}^T \mathbf{x} + b$ , with  $\mathbf{w}, \mathbf{x} \in \mathbb{R}^d$ ,  $b \in \mathbb{R}$ . Let  $H$  be a hyperplane defined as  $H := \{\mathbf{x} \in \mathbb{R}^d | f(\mathbf{x}) = 0\}$  and let  $\tilde{\mathbf{x}} \in H$ . The signed distance of  $\mathbf{x} \in \mathbb{R}^d$  to  $H$  is  $d(\mathbf{x}, H) := \Pi_{\mathbf{w}}(\mathbf{x} - \tilde{\mathbf{x}})$ . Show that the signed distance can be equivalently defined as:

$$d(\mathbf{x}, H) := \frac{\mathbf{w}^T \mathbf{x} + b}{\|\mathbf{w}\|}.$$

- 3) In this question we will have you do some exploration of the applications of machine learning. One of the goals of this lecture is for you to be able to identify learning problems on your own. To this end, describe three learning problems that have not been mentioned in the lecture. For each problem address the following:
- What is the data? (What are the inputs; what are the labels?)
  - What is the goal? Are humans good at solving this task? Why, or why not?
- 4) (MANDATORY) 5 points
- Make sure to set python notebook up on your computer, you can use the following link to do so <https://jupyter.readthedocs.io/en/latest/install.html>. Then solve programming task 0. The mandatory part is only 3.4  $k$ -nearest neighbours, however doing the rest of the programming tasks should help prepare you for 3.4.