INTRODUCTION TO MACHINE LEARNING

Problem set 3
Burkhardt Funk – Winter 2020/2021
due: December 11th 2020

Task 1

Given the feature transform $z = \Phi(x) = (1, x_1^2, x_2^2)$ and the hypothesis $\tilde{h} = sign(\tilde{\mathbf{w}}^{\mathsf{T}}\Phi(\mathbf{x}))$: By using R, visualize the boundary a hyperplane in Z creates in X for the following examples: $\tilde{\mathbf{w}} = (1, -1, -1)$, $\tilde{\mathbf{w}} = (-1,1,1)$, $\tilde{\mathbf{w}} = (1, -1, -2)$, $\tilde{\mathbf{w}} = (1,1, -1)$.

Task 2

Design a Multilayer Perceptron that implements the Boolean functions

a) $\boldsymbol{A \wedge \neg B}$ and

b) $\boldsymbol{A\ X\bar{O}R\ B}$

Task 3

Assume that N training examples are given and that the hypothesis set is $h(x) = \sum_{n=1}^{N} w_n \exp(-\gamma\|x - x_n\|^2)$. The functions $\exp(-\gamma\|x - x_n\|^2)$ are called radial basis functions. Let Φ denote the following matrix.

$$\begin{bmatrix} \exp(-\gamma\|x_1 - x_1\|^2) & \cdots & \exp(-\gamma\|x_1 - x_N\|^2) \\ \exp(-\gamma\|x_2 - x_1\|^2) & \dots & \exp(-\gamma\|x_2 - x_N\|^2 \\ \vdots & & \vdots \\ \exp(-\gamma\|x_N - x_1\|^2) & & \exp(-\gamma\|x_N - x_N\|^2 \end{bmatrix}$$

(i) Generate some training data $\{(x_i, y_i)\}$ (e.g. $y \sim x^2 + \epsilon$) and use $\boldsymbol{w} = \boldsymbol{\Phi^{-1} y}$ to determine w. Understand the behavior when γ is changed. Argue about the expected generalization performance. (ii) Use $w_{ridge} = (\Phi^T\Phi + \lambda I)^{-1}\Phi^T y$ instead, where I is the unit matrix of dimension N and λ is a constant.