

INTRODUCTION TO MACHINE LEARNING

Problem set 1

Burkhardt Funk – Winter 2020/2021

```
Given a linearly separable training set S and learning rate  $\eta \in \mathbb{R}^+$ 
 $w_0 \leftarrow 0$ ;  $b_0 \leftarrow 0$ ;  $k \leftarrow 0$ 
 $R \leftarrow \max_{1 \leq i \leq l} \|x_i\|$ 
repeat
  for  $i = 1$  to  $l$ 
    if  $y_i(\langle w_k \cdot x_i \rangle + b_k) \leq 0$  then
       $w_{k+1} \leftarrow w_k + \eta y_i x_i$ 
       $b_{k+1} \leftarrow b_k + \eta y_i R^2$ 
       $k \leftarrow k + 1$ 
    end if
  end for
Until there are no mistakes within the for loop
Return the list  $(w_k, b_k)$ 
```

This is the perceptron algorithm (PLA) based on Christianini & Shawe-Taylor (2000). In this pseudocode x_i is a n -dimensional vector of real numbers, w_i is a n -dimensional vector of weights (real numbers as well), b_0 is a real number (the axis-intercept) and R is the maximum euclidean norm of all the x vectors. The training set S consists of l observations $(x_1, y_1), \dots, (x_l, y_l)$.

1. Let the points $x_1 = (1, 2)$, $x_2 = (3, 2)$, $x_3 = (2, 1)$ and $x_4 = (3, 3)$, where x_1 and x_3 are labeled -1 and x_2 and x_4 are labeled +1. Calculate appropriate weights manually using the perceptron algorithm manually.
2. What happens if data is not linearly separable?
3. Simulate data for two classes in a 2D feature space, classes should be linearly separable. Generate data in such a way, that there is a gap between the two classes. You can also use the "make_blobs" function from the sklearn.datasets module.
4. Implement the perceptron algorithm and run the perceptron algorithm on your simulated data
5. Study the number of mistakes made by the algorithm in systematical manner (e.g. vary the number of examples, the size of the maximal margin). What do you see?