Machine Learning **Universität Hamburg**
Spring 2024 Sören Laue

Exercise Sheet 5

Exercise 1

There is a simple way of smoothing the max-function. For instance, if we have two smooth functions $f_1(x)$ and $f_2(x)$ the pointwise maximum $f(x) = \max(f_1(x), f_2(x))$ might not be smooth. Think for instance of $|x| = \max(-x, x)$ which is not smooth at $x = 0$. The following function

$$\text{smoothmax}(f_1(x), f_2(x)) = \frac{1}{M} \log(\exp(M \cdot f_1(x)) + \exp(M \cdot f_2(x)))$$

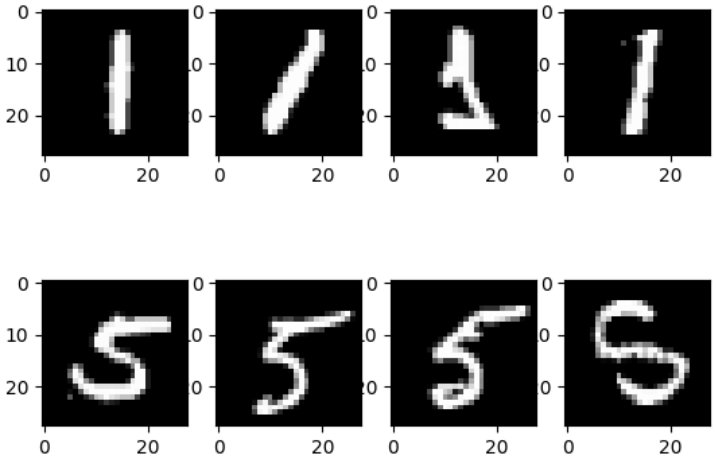
for a parameter $M > 0$ which controls how much to soften the max. Show that the following inequality

$$\max(a,b) \leq \text{smoothmax}(a,b) \leq \max(a,b) + \frac{1}{M} \log(2)$$

holds.
How does the smoothmax relate logistic regression to SVMs? What conclusions can you draw from it for comparing logistic regression to SVMs?

Exercise 2

Your task is to find a linear classifier to classify images of handwritten digits. The data set contains two digits, 1 and 5. The following figures show four samples of each number.



You will find the training and the test data set `dataset_numbers_train.npy` and `dataset_numbers_test.npy`. If $y_i = 0$ then the i-th image represents a one, otherwise $y_i = 1$ represents a five. What is your best test error?

Please turn in your solutions by Friday, May 10th. (Thursday is Ascension day)