

TOPICS IN STATISTICAL SCIENCES II – EXAM EXERCISE 3

SØREN HØJSGAARD AND CHRISTIAN HIRSCH

This exercise is about neural networks. During the oral exam you will have 20 min to present the exercise. You decide what topics to cover and how to present them, however, we will ask questions to any part of the covered curricula, exercise and presentation.

1. BACKPROPAGATION

In the setting of a binary prediction problem, consider a multi-layer perceptron consisting of one input layer of size 50 and one hidden layer of size 100. Assume ReLU as activation function for the hidden layer and a sigmoid function for the output. Use cross-entropy as loss function.

- (1) Write down in mathematical terms the function defined by this network. What is the optimization problem when fitting the model to data and how is the initial value chosen?
- (2) How would stochastic gradient descent be applied to this problem? What happens with given training data during the forward and backward phases of backpropagation?
- (3) Using the chain rule, compute the gradient of the weights in this network. You may assume the biases to be 0.
- (4) How do you choose the learning rate?
- (5) What can be advantages of removing the hidden layer? Think of model interpretability.

2. TITANIC

The Titanic training dataset from the R library `titanic` describes the survival status of 891 passengers. For each of the passengers up to 10 additional features are given. However, be aware that there are missing values. Use a multi-layer perceptron to attack this prediction problem.

- (1) What features do you base your network on? What are preprocessing steps?
- (2) What architecture do you choose? What layers and activation functions do you use?
- (3) How can you regularize your network? What optimization method do you use?
- (4) What are training and validation errors do you obtain?
- (5) What does your network predict for test data?