Backpropagation

1. Why do we need to rescale the input data before training an MLP?

2. How does the error propagate from the output layer to the hidden layer?

3. Why de we compute the gradient of the error?

4. Derive the update rule of the output neurons according to backpropagation.

5. Derive the update rule of the hidden neurons according to backpropagation.