

Neural Networks KU WS18

708.076 18W

Exercise Sheet 2

Neural Networks [5P]

For this exercise, form groups of two students (see also last paragraph).

Use TensorFlow 1.5 in Python 3 to train a neural network for classification of a variant of the isolet dataset. The data set contains features extracted from the utterances of four human speakers saying the characters “a” to “z” several times. Hence, there are 26 classes, one for each character. The test set contains the same features extracted from the utterances of a fifth speaker. Note that the data samples are sorted by speaker identity. From the original data set, we only use the first 300 features to make the task more challenging.

Download the dataset from the homepage. The file `isolet_crop_train.pkl` contains the training set and `isolet_crop_test.pkl` contains the test set. A small script that loads the dataset is provided.

The class labels are in the range $1 \dots 26$. To train the network, these labels have to be converted to a one-out-of- K coding (K is the number of classes, i.e., 26). Use a softmax activation function for the output layer. Optimize cross entropy error of the network, however, all performances should (also) be reported in terms of the misclassification rate (percentage of misclassified examples).

Task details [5P]:

- a) Make yourself familiar with the dataset. Normalize the data if necessary.
- b) Use stochastic gradient descent with mini-batches of size 40. Use the test set only for final evaluation. For any validation use a validation set consisting of samples from the training data. Use early stopping to avoid overfitting.

Meta-parameter-search: Test a number of learning rates and architectures with varying numbers of hidden units and hidden layers. Compare and report the validation errors in terms of the misclassification rate. Now fix the determined architecture. Document your results. Provide a table where training and validation errors (in terms of the misclassification rate)

of various hyper-parameter settings are compared. Discuss your experiences and motivate your final choice. Provide plots where the evolution of the training/test error during training is shown for a number of parameter settings.

- c) Once the architecture has been fixed, you can perform the final training with the full training set. Make use of early stopping in the following manner: Retrain the network for the number of epochs n_{ep}^* that achieved previously the best validation error for this parameter set (an epoch is a run through all the training set). In this retraining use as initial parameters those that were also used as initial parameters in the initial training for this hyper-parameter set.

Submit the code as a single .py file and a report as PDF in the teach center. Do **not** provide one zip with code and PDF but submit them separately. The code should be executable under the assumption that the files provided at the homepage are available (unzipped) at the working directory. Both students of a group have to submit their python code and report (it can be identical to the one of the group partner). In the report, indicate your group partner on top of the first page (write "Group partner: ⟨First name⟩, ⟨Last Name⟩, ⟨Matrikelnummer⟩").