

CAP4611 Homework 6

Preliminaries

- Visit <https://playground.tensorflow.com>
- Turn on “Show test data” check box.
- Set the batch size to 10.
- Familiarize yourself with the four training sets/ problems for classification
 - P-GAUSS: Gaussian
 - P-XOR: Exclusive or
 - P-CIRCLE: Circle
 - P-SPIRAL: Spiral
- Experiment with choosing the ratio of training to test data (which changes the amount of the training data), and the noise. You need to push “regenerate to create a new dataset”. We will call these training data quality.
 - Q-BEST: 90% training data, 0 noise
 - Q-WORST: 10% training data, 50 noise
 - Q-OK: 70% training data, 20 noise

Four sample models

Consider the following four models. For all of them, the regularization is None.

- Linear regression (LR)
 - 0 hidden layers
 - x_1, x_2 inputs
 - activation linear
- Logistic regression (LO)
 - 0 hidden layer
 - x_1, x_2 inputs
 - activation sigmoid
 - regularization none
- 1980s neural net (NN)
 - 1 hidden layer (5 neurons)
 - x_1, x_2 inputs
 - activation sigmoid
 - regularization none
- Engineered features (EF)
 - 0 hidden layers
 - x_1, x_1x_2 inputs
 - activation linear

Problem 1 - Evaluation (5 pts):

For each of the four sample models (LR, LO, NN, EF), and all six problems (P-GAUSS, P-XOR, P-CIRCLE, P-SPIRAL) and all three quality settings Q-BEST, Q-WORST, Q-OK do the following:

- Try to train the model to learn the task.
- You might want to vary the learning rate. Don't let any of the models train more than 30 seconds.
- Record the best training loss and test loss obtained.
- Record the time it took to finish the learning.
- Judge whether the learning was success and whether the system overfit or not. Use your own criteria of success.

Write it up like this:

EF + P-XOR + Q-BEST

Successful, not overfit, <1 sec L_train = 0.018, L_test = 0.035, rate 0.00001

Write two paragraphs of conclusions. Did any of the models solve all problems? Were some models better for some problems? Did some models train faster than others?

Problem 2 - Network design (5 pts):

Develop three new models, to improve learning, as follows

- MYENG:
 - any input features (try to be creative about this)
 - any number of hidden layers
 - activation linear
- DEEP
 - x_1, x_2 input
 - sigmoid activation
 - any number of hidden layers
- MYBEST
 - x_1, x_2 input
 - any activation
 - any number of hidden layers
 - feel free to use regularization if needed or change the batch size.

Your objective is to find the best solution within the set constraints.

Repeat the evaluation process for the resulting networks.

Write up the conclusions of the work in two paragraphs. Did you manage to create a network which solves all problems in all quality settings? Which settings were the hardest? Did some models train faster than others?