CS 7015 : Deep Learning

Assignment 2

Instructor: Dr. Anurag Mittal (BSB 368)
amittal@cse.iitm.ac.in

Part-A

Tuning the regularization parameter

Please take the best performing model from Assignment-1A. Experiment with various regularization parameter (by changing the *weight_decay* parameter of the optimizer) values and find the best one. Visualize the train vs test **accuracy, loss** for each value. Describe all your findings and inferences in the report.

Part-B

This part is a written assignment. You need to submit hand written copy (detailed step-by-step derivations are expected).

1. Gradient calculation of common Activation functions

Please derive the gradient formulas for the following activation functions. Can you decompose these formulas so that such gradients can be computed more easily via the chain rule. Please show the decomposition of the function and how you would compute the gradients in such a case.

- Sigmoid
- Hyperbolic Tangent
- ReLU
- Leaky ReLU
- PReLU
- ReQU
- Softmax

2. Gradient calculation of common loss functions

Please derive the gradient formulas for the following loss functions. As in part 1, try to decompose these functions for easier computation of the gradient via chain rule:

- Cross entropy loss
- Hinge loss (y_i = correct class of example i)
- L1 loss

- Huber loss
- L2 loss
- Cosine similarity

3. Hand-calculation of gradients

Please perform back-propagation in the following network by hand calculations. Consider sigmoid activation function in both Hidden and Output layers and L2 loss as objective function. (consider bias = -0.2 in both hidden and output layers):

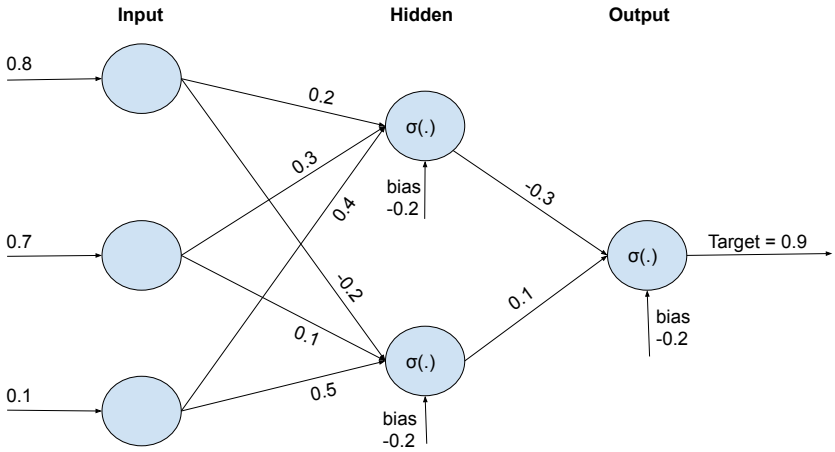

Figure 1: Computation graph with weights on the edges. $\sigma(.)$ represents Sigmoid activation function.

Plagiarism

- You should do the assignment yourself. In case you take help from others, please mention in the pdf submitted.
- No sharing of code/experiments etc. will be allowed under any circumstances and may attract disciplinary action by the institute disciplinary committee.

Submission Guidelines

- **Dead line : 16/09/2019 11:59 PM**
- **PDF & Code Upload:** The details will be shared in Moodle.
- Email submissions will not be accepted. Reduce file size (if required).
- This is not a team assignment.

TAs:

- Gouthaman KV (gouthamankv@gmail.com), BSB 374
- Arulkumar (aruls@cse.iitm.ac.in), BSB 374
- Vikram Singh (vik1409@gmail.com), BSB 374