

Homework 3 (100 points)

Deep Learning

The homework will be due in 2 weeks from the day of release.

In this homework you will implement a 2-layer neural network using any Deep Learning Framework (e.g., TensorFlow, Pytorch etc.). Upload a .txt file with a link to your file as your submission on Submittity.

You need to perform the following tasks for this homework:

In your project, you will pick a dataset (not the same as in the previous homeworks) and describe the problem you would like to solve. Include a link to the dataset source. Next, you should pick a Deep Learning Framework that you would like to use to implement your 2-layer Neural Network.

Task 1 (25 points): Assuming you are not familiar with the framework, in this part of the homework you will present your research describing the resources you used to learn the framework (must include links to all resources). Clearly explain why you needed a particular resource for implementing a 2-layer Neural Network (NN). (*Consider how you will keep track of all the computations in a NN i.e., what libraries/tools do you need within this framework.*)

For example, some of the known resources for TensorFlow are:

<https://www.tensorflow.org/guide/autodiff>

https://www.tensorflow.org/api_docs/python/tf/GradientTape

Hint: You need to figure out the APIs/packages used to implement forward propagation and backward propagation.

Task 2 (60 points): Once you have figured the resources you need for the project, design, and implement your project. The project must include the following steps (it's not limited to these steps):

1. Exploratory Data Analysis (Can include data cleaning, visualization etc.)
2. Perform a train-dev-test split.
3. Implement forward propagation (clearly describe the activation functions and other hyper-parameters you are using).
4. Compute the final cost function.
5. Implement mini-batch gradient descent to train your model. In this step it is up to you as someone in charge of their project to improvise using optimization algorithms (Adams, RMSProp etc.) and/or regularization.
6. Present the results using the test set.

Task 3 (15 points): In task 2 describe how you selected the hyperparameters. What was the rationale behind the technique you used? Did you use regularization? Why, or why not? Did you use an optimization algorithm? Why or why not?