

CS776: DEEP LEARNING FOR COMPUTER VISION

Assignment

Assignment Due Date: 16 Feb 2025 11:59 PM IST (Strict Deadline)

General Instructions:

1. Late submissions will not be accepted.
2. Any sort of plagiarism will be penalised. If you are referring to any materials online or books, please cite them accordingly otherwise, it will be considered plagiarism.
3. Please ensure that the report is written as per the instructions given in the assignment.
4. You are free to use Google Colab or Kaggle to train the models.
5. Use of the numpy library is allowed.

Question (100 Marks):

1. Download the [Fashion MNIST \(https://www.kaggle.com/datasets/zalando-research/fashionmnist\)](https://www.kaggle.com/datasets/zalando-research/fashionmnist) dataset. Split the dataset into training, validation and testing sets. A common split is 80% of the data to train, 10% to validate, and 10% to test scenarios, but you can adjust this as needed. Normalize the images. This involves scaling the pixel values to a range between 0 and 1.
2. Implement a MLP for classification. (total 40 marks)
 - a. Flatten the images into a single dimensional vector before feeding it to the model. (1 marks)
 - b. Write a pre-processing module for all the images. (3 marks)
 - c. Write the Forward pass from **scratch**. Use of the inbuilt forward pass function will result in 0 marks for this sub-question. (8 marks)
 - d. Write the Backward pass from **scratch**. Use of the inbuilt back propagation function will result in 0 marks for this sub-question (12 marks)
 - e. Write the custom coded module for cross entropy loss (1 marks)
 - f. Experiment with different hyperparameters like number of layers, dropout, activation functions such as RELU, Leaky-RELU, Tanh, and GELU include comparisons in the report and settle with a combination which performs the best for the given problem. (15 Marks)
3. Implement a CNN backbone model using pytorch. (total 40 marks)
 - a. Build a small CNN model consisting of 5 convolution layers. Each convolution layer would be followed by a ReLU activation. In pooling layer, implement different pooling methods such as max pooling, average pooling and global average pooling to compare their effects. (10 Marks)
 - b. Experiment with different kernel size, number of kernel each layer (keep number of filter same in each layer, double it in each layer etc) and settle with a combination which performs the best for the given problem. (10 Marks)
 - c. Try different weight initialization methods (random, Xavier, He) (5 Marks)

- d. After extracting feature from CNN model use MLP for classification (use code from question 2) (15 Marks)
4. Submit a report clearly explaining how you have built both the models, the architecture of the models, learning rate, epochs used for training, evaluation metrics and the instructions for running the models. Compare the performance of the models on the different hyperparameters you tried and justify the observed behavior. (20 Marks)

Deliverables:

1. The solution for the assignment should be submitted along with the readme file as a zip file. Please submit a **single zip file per group**. The readme file should include the exact steps to run your code. Write the filename as your group name.
2. Please mention the names and email ids and the percentage contribution of each group member in the report. Submit a report in pdf format.