

Assignment 6: ANN Model Comparisons

40 Points

Deliverable: Notebook (.ipynb file) with all required code and answers to questions/discussions provided in Markdown cells.

Overview: This assignment expands upon Assignment 5 by asking you to compare models with different configurations or training routines. I would recommend re-running your code from Assignment 5 then continuing with the new analyses. You will work with the Covertype dataset, which is available on the UCI Machine Learning Repository: <https://archive.ics.uci.edu/ml/datasets/Covertype>. It has also been provided as part of the class data. This dataset consists of 581,012 records. The goal is to differentiate 7 forest types using landscape characteristics. The forest types are differentiated in the Cover_Type column using numeric codes as follows:

1 = Spruce/Fir, 2 = Lodgepole Pine, 3 = Ponderosa Pine, 4 = Cottonwood/Willow, 5 = Aspen, 6 = Douglas-fir, 7 = Krummholz

You will use the following predictor variables. Note that there are other predictor variables available which will not be used here.

Elevation, Aspect, Slope, Horizontal_Distance_To_Hydrology, Vertical_Distance_To_Hydrology, Horizontal_Distance_To_Roadways, Hillshade_9am, Hillshade_Noon, Hillshade_3pm

Models

Compare the results obtained in Assignment 5 with results obtained using the following modifications. Including Assignment 5, you should train a total of five models. Make sure to log all training and validation losses by epoch. Train all models using the same training and validation data for 50 epochs.

T1: Replace the ReLU layers with leaky ReLU. Re-train the model. (5 Points)

T2: Re-train the model from Assignment 5 using the AdamW optimizer as opposed to Stochastic Gradient Descent. Use a learning rate of 0.001. (5 Points)

T3: Retrain the model from Assignment 5 using cross entropy loss with class weights. Weights should be scaled relative to the inverse of the abundance of each class in the training dataset. (5 Points)

T4: Implement a one cycle learning rate policy where the learning rate oscillates from a base learning rate of 0.0001 to a high learning rate of 0.1 then back to 0.0001 during the entire training process. (5 Points)

Comparisons

T5: Create a plot of training losses across all 50 epochs for all 5 experiments. (5 Points)

T6: Create a plot of validation losses across all 50 epochs for all 5 experiments. (5 Points)

T7: Using TorchMetrics and the test data, calculate overall accuracy and class-aggregated macro-averaged precision, recall, and F1-score. (5 Points)

T8: Discuss your results. Did any of the four tested alterations improve the model's performance? How do the loss curves and assessment metrics calculated from the withheld test data compare for each model? (5 Points)