

Fiche TP2_suite : Optimizing hyperparameters (Keras)

- 1- Load and preprocess the Cifar-10 dataset by normalizing the data.
- 2- Use the same architecture as in the previous TP (TP2).
- 3- Split the train set into training and validation sets, with a 20% validation set.
- 4- Train the designed architecture using Mini-batch Stochastic Gradient Descent (SGD) with a batch size of 128, a learning rate of 0.01, for 50 epochs. Plot the learning curves on the train and validation data.
- 5- Add L2 norm regularization to the second fully connected layer, and plot the learning curves on the train and validation data. Compare the results obtained between the architecture with and without L2 norm.
- 6- Add a dropout layer with rates of 0.2, 0.3, and 0.5 to the second fully connected layer, and compare the results obtained.
- 7- Use the early stopping technique, and specify the epoch at which the training stopped.
- 8- Add a batch normalization layer after the first hidden layer, and train using mini-batch SGD. Present the results obtained.
- 9- Use random search to find the best hyperparameters for `learning_rate`, `dropout_rate`, and `batch_size`, and specify the parameters that ensure the best results.
- 10- Present all obtained results, including graphs, accuracy, and running time, in a report and discuss them.