

## Fiche TP2 : Training an MLP (Keras)

- 1- Load and preprocess (data normalization) the MNIST dataset.
- 2- Split the train set into Train and validations sets (20% for validation).
- 3- By using the keras library, design an MLP architecture, which is composed from one input layer, two hidden layers (128, 64 neurons) and one fully connected layer. Use the RELU activation function for hidden layers and the softmax for the fully connected layer. Use the cross-entropy loss function.
- 4- Train the designed architecture based on SGD, Mini batch SGD (64 the size batch), batch SGD in 50 epochs and set the learning rate to 0.01. Plot the learning curves for each strategy (loss and accuracy).
- 5- Compare between the training time and the learning curves of mini batch SGD, mini batch SGD with decay ( $1e-6$ ) and SGD with decay ( $1e-6$ ) and momentum.
- 6- Compare between SGD ( $lr=0.01$ ), Adam ( $lr=0.001$ ), and RmsProp ( $lr=0.001$ ) optimizers.
- 7- Export the best model on the validation set, import it, and make predictions on some samples from the test set.
- 8- Presents all obtained results (graphs, accuracy, running time) in a report and discuss them.