

HTML FORMAT

In this activity, the task is to learn how to design and train a model that does well on the unseen (validation) dataset. The weights and biases of a neural network model are its parameters. The parameters such as the number of layers of neurons, numbers of neurons in each layer, number of epochs, batch size, activation functions, choice of optimizer, choice of loss function, etc. are the hyperparameters of a model. When training a model for a new dataset a crucial question is - what combinations of hyperparameters yield the maximum accuracy on the validation set? Remember, when playing with activation functions, the activation of the last layer should not change - it should always be sigmoid for binary classification and ReLU or linear for regression. The task in this activity is to try as many hyperparameters as possible to obtain the highest possible accuracy on the validation set. For a classification dataset of your choice, the first step is to create a notebook where you can train the model using the training set and evaluate on the validation set. Then, the objective is to find the optimal (best) hyper-parameters that maximize the accuracy (or minimize MAE) on the validation set.

- **Reference Video:**
- [What are hyperparameters and why are they important](#)
- [\(Links to an external site.\)](#)



•

Below are the hyperparameters to optimize:

1. The number of layers in the neural network (try 1, 2, 4, 8, 16, etc.).
2. The number of neurons in each layer (try 2, 4, 8, 16, 32, 64, 128, 256, 512, etc.).
3. Various batch sizes (8, 16, 32, 64, etc.).
4. Various number of epochs (2, 4, 8, ..., 5000, etc.).
5. Various optimizers (rmsprop, sgd, nadam, adam, gd, etc.).
6. Various activation functions for the intermediate layers (relu, sigmoid, elu, etc.).