Binary Classification:
 • Output Activation: Sigmoid
 • 1 neuron in output
 • Loss: binary cross entropy But during prediction/inference time, the features came in the following order: gender, age, height # 3= 100 4 classes Multiclass classification :Output Layer: Softmax, 4 neuronLoss: categorical cross entropy Regression:
 Output: No activation
 Loss mean sqr error A: We must store the function and their params.
Min-Max Scaling: min, max
Standardization: mean, sd
OneHotEncoder Epoch: Keep training and stopping when it overfits.
Learning Rate:
If the training is diversing, reduce the LR.
If learning is too slow, try increasing. Use LR Scheduling 1. Load previously trained model or Initialize weights randomly for all layers - he or glorat 2. for epoch in range(epochs): a. Call callbacks b. for X_batch in batches[can be parallel]: i. Call callbacks ii. compute the outputs of each layer. Activation, Batch Normalization, Left (input) to right (output) iv. Compute the gradients for all layers. Use backprop to reduce the computations. v. Call callbacks vi. Update the weight with -eta * gradients vii. Checkpointing [can be parallel]

 $\frac{3}{\sqrt{2}} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}$