"Learning" simply means finding a combination of model parameters that minimises a loss function for a given set of training data samples and their corresponding targets

LOOKING BACK ON OUR MNIST EXAMPLE