By now you should know what the following terms are:

* **Feature:** The input(s) to our model
* **Examples:** An input/output pair used for training
* **Labels:** The output of the model
* **Layer:** A collection of nodes connected together within a neural network.
* **Model:** The representation of your neural network
* **Dense and Fully Connected (FC):** Each node in one layer is connected to each node in the previous layer.
* **Weights and biases:** The internal variables of model
* **Loss:** The discrepancy between the desired output and the actual output
* **MSE:** Mean squared error, a type of loss function that counts a small number of large discrepancies as worse than a large number of small ones.
* **Gradient Descent:** An algorithm that changes the internal variables a bit at a time to gradually reduce the loss function.
* **Optimizer:** A specific implementation of the gradient descent algorithm. (There are many algorithms for this. In this course we will only use the “Adam” Optimizer, which stands for *ADAptive with Momentum*. It is considered the best-practice optimizer.)
* **Learning rate:** The “step size” for loss improvement during gradient descent.
* **Batch:** The set of examples used during training of the neural network
* **Epoch:** A full pass over the entire training dataset
* **Forward pass:** The computation of output values from input
* **Backward pass (backpropagation):** The calculation of internal variable adjustments according to the optimizer algorithm, starting from the output layer and working back through each layer to the input.