Notes on Loss Functions for Classification

* Loss Function for Single Label Multiple Classes Classification (output is always one class)
  + softmax is the multi-category equivalent of sigmoid
  + softmax, and then the log likelihood (nll) of that - is called cross-entropy loss. In PyTorch, this is available as nn.CrossEntropyLoss
  + This will work only when the output is always a category(class).
    - If input image is something that does not belong to the Classifier’s “n” categories, the output will still be one of those categories which is wrong in real-life scenarios.
* Loss Function for Multiple Label Multi Class Classification (with no classes for certain inputs). Input Image may not belong to any category that is being predicted.
  + Binary Cross Entropy is useful for such cases. It is roughly RMSE loss with log applied.
  + nn.BCEWithLogitsLoss is the Pytorch class instance