Eddie Zhang HW 6

1. (b) Sigmoid. Firstly, we can immediately eliminate ReLU and Linear because these do not constrain the output from 0 to 1, so it would not be possible to use Cross Entropy Loss or other loss functions designed for classification problems. Between softmax and sigmoid, I would choose sigmoid because this is only a binary classification problem. Softmax is used for multiclass problems, since it expects a vector and outputs a vector.

2. (d) Mean Squared Error (L2 Loss). Cross Entropy and Hinge Loss are not suitable because these are designed for classification problems. While both MAE and MSE could be used for this problem, I am making the assumption that airfare will follow a Gaussian distribution and MSE is most apt for this distribuition. Additionally, I am assuming that our data is clean and that outliers are mostly nonexistent, so MSE will not overly punish these cases.

3. A model is overfitting if the training accuracy is significantly higher than the testing accuracy. Two ways of addressing this is through adding regularization to the loss function and drop-out neurons.