1. (b) Accuracy is non-differentiable and cannot be used for direct optimization via gradient descent.

2. I will choose precision, recall, and F1 score to evaluate my model because the dataset may be imbalanced.

3. Gradient Descent. Error is defined through MSE (mean squared error)W

4. (a)
   (b)
   (c)
   (d)

5. (a) Which classifier is better depends on what the use case of the model is
   (b) Classifier A would undoubtedly be better in a situation where a low FPR is required. For example a medical situation. In medical situations an accurate prediction could be life or death.
   (c) Classifier B would undoubtedly be better in a situation where a high sensitivity is required.
   (d) The accuracy of predictions affect the ROC and thus, the AUC. (true positives/false positives)

6. When evaluating metrics, it is most efficient to optimize the loss function. That way you can tweak the weights and biases in order to optimize it.

7. (a) for the ROC graphs listed, the first one is the best metric because it has an AUC of 1. This means that the model's predictions are 100% correct. The next best metric is the second with an AUC of 0.8 and the worst metric is the last one with an AUC of 0.5. As the auc decreases that means the amount of false positives that the model predicts is increasing.
   (b) For the first metric with an AUC of 1, it's predictions are 100% correct which means the TPR is 100% and the FPR is 0%. As the curve decreases these rates change. For example, in the last metric, it's a random guess meaning the model randomly diagnoses positive regardless of if it actually is or not.