

## Q1-2

- If we need to seek a balance between Precision and Recall, there is an uneven class distribution (large number of Actual Negative).
- In PLA, maybe a step function is a better choice, but in neural network, we need smooth signal to train and converge easily, or  $E_{in}$  is NP-hard to solve.
- Bias: performance of consensus, and is also known as underfitting. Variance: expected deviation to consensus, and is also known as overfitting.
- Pruning is required in decision trees to avoid overfitting. However, random forests would give good performance with full depth. It uses bootstrap aggregation along with random selection of features for a split.
- One hot encoding is a process by which categorical variables are converted into a form that could be provided to ML algorithms to do a better job in prediction.
- 1. Adding more data
  2. Early stop if your validation error starts
  3. Parameter Regularization. (L1, L2)
  4. Dropout