**For Multiple Choice Questions, circle the best answer.**

Q1. (2 points) Consider a dataset for binary classification that is clearly not linearly separable. The features in the dataset are not independent when conditioned on a class. Which of the following classifiers will be your first choice to try out in this case?

- **a. Decision Trees**
- b. Logistic Regression
- c. Naive Bayes Classifier
- d. MLP with linear activation function

Q2. (2 points) Which of the following statements is true about the Bayes Decision Rule?

- **a. It provides the lowest expected misclassification error for a given choice of features used for classification.**
- b. Since it is the optimal decision, It will result in zero error if class priors and class-conditional likelihoods are known
- c. It may not always provide the lowest expected error when training data is limited.
- d. For any given validation dataset, the misclassification error that it yields will always be lower than (or equal to) that achieved by any other decision logic.

Q3.1 (1 point) Decision trees with more depth usually have higher bias than decision trees with lower depth (T/F)?

- a. True
- **b. False**

Q3.2 (1 point) Decision trees with more depths usually have higher variance than decision trees with less depths (T/F)?

- **a. True**
- b. False

Q4. (2 points) Which of the following statements is TRUE about logistic regression model

- a. It is designed for regression tasks.
- **b. It uses the sigmoid function to model the probability of a binary outcome.**
- c. It relies on the mean square error for optimizing the model.
- d. It cannot be used for image classification.

Q5. (2 points) For a binary classification problem (with equal number of points in the positive and negative classes), make a plot of the expected ROC curve of a random classifier (a classifier that predicts 0 or 1 randomly). What is the value of AUC-ROC for this curve?


AUC-ROC = 0.5

**Quiz 3, Fall 23 (10 points)**
**Section B**
**NAME and EID: _____**

**For Multiple Choice Questions, circle the best answer.**

Q1 (2 points) Which of the following statements is FALSE about Logistic Regression?

a. Logistic Regression models the conditional distribution p(y|x).
b. Logistic regression is a linear classifier but retains probabilistic semantics.
c. Parameters in Logistic Regression are learned by iterative optimization.
**d. The decision boundary obtained through Logistic regression will be parallel to the vector of learnt parameters (the beta vector).**

Q2.1 Decision Trees (1 point) Decision trees with depth of 1 will always give a linear decision boundary (T/F)?

**a. True**
b. False

Q2.2 (1 point) Does the split criterion in a decision tree aim to partition the data in a manner that creates an equal distribution of classes in each child node (T/F)?

a. True
**b. False**

Q3. (2 points) What is the naive assumption made in the Naive Bayes Classifier?

**a. For any given class, the features are independent of one another.**
b. All the features are independent of one another.
c. All classes are independent of one another.
d. The most probable feature for a class is the most important feature for classification.

Q4. (2 points) Which of the following is true for the shape of the decision boundary learned using logistic regression for a two-class problem:

a. Depends on the threshold.
b. **Is always linear.**
c. Behaves like a sigmoid function.
d. Depends on the training data.

Q5. (2 points) Qualitatively sketch below how you would expect the precision-recall curve to look like for a binary classifier that is not perfect but does have some predictive power. Compare it with the expected P-R curve obtained using a random classifier. Each class is modeled as a mixture of two Gaussians, and the Prior for the positive class is 0.6.