Student ID:

Name:

**Notably, please illustrate steps clearly to get the full marks.**

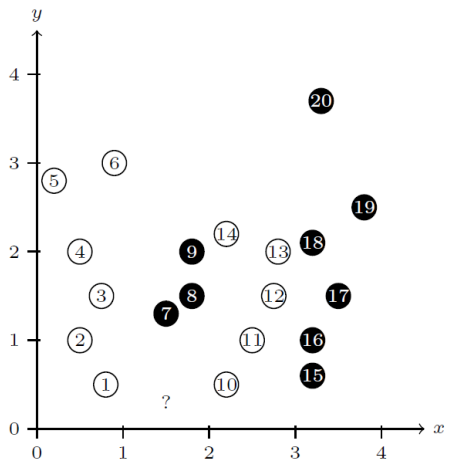
**Question 1**: Consider the following data set:

| Restaurant | Type | Price | Neighborhood | Restriction | OK |
|---|---|---|---|---|---|
| $R_1$ | Fast Food | $ | Oakland | Vegetarian | 0 |
| $R_2$ | Ethnic | $$ | Squirrel Hill | Gluten Free | 0 |
| $R_3$ | Casual Dining | $$ | Squirrel Hill | None | 0 |
| $R_4$ | Casual Dining | $$$ | Shadyside | Vegetarian | 0 |
| $R_5$ | Casual Dining | $ | Oakland | Vegetarian | 1 |
| $R_6$ | Fast Food | $$ | Squirrel Hill | None | 1 |
| $R_7$ | Ethnic | $ | Squirrel Hill | None | 1 |
| $R_8$ | Casual Dining | $ | Shadyside | Gluten Free | 0 |
| $R_9$ | Fast Food | $$$ | Oakland | None | 0 |
| $R_{10}$ | Ethnic | $$ | Shadyside | Vegetarian | 1 |
| $R_{11}$ | Casual Dining | $$ | Shadyside | Gluten Free | 1 |

Suppose we decide to construct a decision tree using multiple splits and the Gini index impurity measure. Which attributes would be the best to use as the root node assuming that we consider each of the input features to be unordered? (12 points)

**Question 2:** For the given datasets with the k-Nearest Neighbors algorithm. As shown in the following figure, we show a set of training points classified as being either black or white. (10 points)

1 Predict the point marked by "?" with k=1,2,3 and Euclidean distance (2 points)
2 Are there any points in the training set that would be misclassified using k=1 using Euclidean distance? If so, identify them. (2 points)
3 If we modify the distance metric to 'the distance on the x-axis', please answer the question 1 and 2 again. (4 points)
4 What happens when k=5 using your distance metric? (2 points)



**Question 3:** Consider the training data set. There are three attributes A, B, and C. The class label is in column Y.
(1) Predict the class label for a test sample (A=1, B=0, C=0) using the naïve Bayes classifier. The answer can be +, -, or cannot decide. (10 points)
(2)We modify the data record 9 from (A: 0, B: 2, C=0) to (A: 0, B: 0, C=1). Then predict the class label for a test sample (A=1, B=0, C=0) using the naïve Bayes classifier. (10 points)

| Record | A | B | C | Y |
|---|---|---|---|---|
| 1 | 1 | 0 | 1 | + |
| 2 | 0 | 2 | 0 | - |
| 3 | 1 | 1 | 0 | - |
| 4 | 0 | 1 | 1 | + |
| 5 | 0 | 0 | 0 | + |
| 6 | 0 | 2 | 1 | - |
| 7 | 1 | 1 | 0 | + |
| 8 | 1 | 2 | 1 | + |
| 9 | 0 | 2 | 0 | - |
| 10 | 1 | 1 | 1 | + |

**Question 4:** Check the following binary classifiers whether they are able to correctly separate the training data (circles vs. triangles) given in following Figure and illustrate the reasons. (8 points)
Logistic regression
SVM with linear kernel
Decision tree
3-nearest-neighbor classifier (with Euclidean distance).
Perceptron