k-Nearest Neighbor Homework

Given Training Set:

x	y	Class Label	Regression Label
0.3	0.8	A	0.6
-0.3	1.6	B	-0.3
0.9	0.0	B	0.8
1.0	1.0	A	1.2

New Point:

$$(0.5,\ 0.2)$$

Manhattan Distances to Training Points:

1. **Point 1 (0.3, 0.8):**

$$d_1 = |0.5 - 0.3| + |0.2 - 0.8| = 0.2 + 0.6 = 0.8$$

2. **Point 2 (-0.3, 1.6):**

$$d_2 = |0.5 - (-0.3)| + |0.2 - 1.6| = 0.8 + 1.4 = 2.2$$

3. **Point 3 (0.9, 0.0):**

$$d_3 = |0.5 - 0.9| + |0.2 - 0.0| = 0.4 + 0.2 = 0.6$$

4. **Point 4 (1.0, 1.0):**

$$d_4 = |0.5 - 1.0| + |0.2 - 1.0| = 0.5 + 0.8 = 1.3$$

Sorted Distances (Nearest Neighbors):

1. **Point 3:** $(d_3 = 0.6)$ (Class B, Regression 0.8)
2. **Point 1:** $(d_1 = 0.8)$ (Class A, Regression 0.6)
3. **Point 4:** $(d_4 = 1.3)$ (Class A, Regression 1.2)

1. Output Class for 3-NN with No Distance Weighting:

- **Classes of Nearest Neighbors:** B, A, A
- **Majority Class: A**

2. Output Class for 3-NN with Squared Inverse Distance Weighting:

Calculating Weights:

Weights are calculated using:

$$w_i = \frac{1}{d_i^2}$$

- **Weight for Point 3:**

$$w_3 = \frac{1}{(0.6)^2} = \frac{1}{0.36} \approx 2.7778$$

- **Weight for Point 1:**

$$w_1 = \frac{1}{(0.8)^2} = \frac{1}{0.64} \approx 1.5625$$

- **Weight for Point 4:**

$$w_4 = \frac{1}{(1.3)^2} = \frac{1}{1.69} \approx 0.5917$$

Summing Weights by Class:

- **Class A Total Weight:**

$$w_A = w_1 + w_4 = 1.5625 + 0.5917 \approx 2.1542$$

- **Class B Total Weight:**

$$w_B = w_3 \approx 2.7778$$

Determining the Output Class:

Since $(w_B > w_A)$, the output class is **B**.

3. 3-NN Regression Value:

a) Without Distance Weighting:

- **Regression Values:** $0.8, 0.6, 1.2$
- **Predicted Regression Value:**

$$\hat{y} = \frac{0.8 + 0.6 + 1.2}{3} = \frac{2.6}{3} \approx 0.8667$$

b) With Squared Inverse Distance Weighting:

- **Weighted Sum:**

$$\text{Numerator} = (w_3 \times 0.8) + (w_1 \times 0.6) + (w_4 \times 1.2)$$

$$\text{Numerator} = (2.7778 \times 0.8) + (1.5625 \times 0.6) + (0.5917 \times 1.2) \approx 3.8697$$

- **Sum of Weights:**

$$\text{Denominator} = w_3 + w_1 + w_4 = 2.7778 + 1.5625 + 0.5917 \approx 4.9320$$

- **Predicted Regression Value:**

$$\hat{y} = \frac{\text{Numerator}}{\text{Denominator}} = \frac{3.8697}{4.9320} \approx 0.7846$$