Problem 1. Unsupervised Learning

(a)

(i) At the first iteration, C1, C2 and C3 has the following points closed to them:

C1: X1, X5, X6, X9, X10 C2: None C3: X2, X3, X4, X7, X8

$$C1_{next} = \frac{X1 + X5 + X6 + X9 + X10}{5} = (0.94, 4)$$

$$C2_{next} = C2 = (0.5, 0.5)$$

$$C3_{next} = \frac{X2 + X3 + X4 + X7 + X8}{5} = (2.68, 2.3)$$

(ii) Labeling each point with it's nearest neighbors. (In the table)

	Nearest Ci After 1 K-Means Iteration (i.e., C1, C2 or C3)	3 Nearest Neighbors, xi, xj, xk	3-NN Predicted Class	Cluster Majority-Vote Predicted Class
X1	C1	X6, X10, X5	F	T
X2	C3	X4, X7, X8	T	T
X3	C3	X9, X8, X4	T	T
X4	C3	X2, X7, X8	T	T
X5	C1	X10, X6, X1	T	T
X6	C1	X1, X10, X5	T	T
X7	C3	X2, X4, X8	T	T
X8	C3	X9, X3, X4	T	T
X9	C1	X8, X3, X6	F	T
X10	C1	X1, X6, X5	F	T

(b) Labeling each point xi with it's 3 nearest neighbors. (In the table)

(c) Labelling each point xi with majority class label of its 3 nearest neighbors. (In the table)

(d) Labelling each point xi with the class of its nearest cluster center (In the table)

C1 : T C2 : None C3 : T

(e)

3-NN accuracy: $\frac{3}{10} = 30\%$

Nearest Cluster Center accuracy: $\frac{6}{10} = 60\%$

The Nearest Cluster Center assigned the most correct labels.
The fraction of the number of correctly labeled points to the total number of points is 60%.