**CSE – 6005 – Machine Learning**

**Lab Experiment – 12 - Implement K-modes Clustering to find natural pattern in data.**

> set.seed(1)

> x <- rbind(matrix(rbinom(250, 2, 0.25), ncol = 5),

+ matrix(rbinom(250, 2, 0.75), ncol = 5))

> colnames(x) <- c("a", "b", "c", "d", "e")

>

> ## run algorithm on x:

> (cl <- kmodes(x, 2))

K-modes clustering with 2 clusters of sizes 53, 47

Cluster modes:

a b c d e

1 0 0 0 0 0

2 2 2 2 2 2

Clustering vector:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66

1 1 1 1 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88

2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2

89 90 91 92 93 94 95 96 97 98 99 100

2 1 2 2 2 2 2 2 2 2 2 2

Within cluster simple-matching distance by cluster:

[1] 116 85

Available components:

[1] "cluster" "size" "modes" "withindiff" "iterations" "weighted"

>

> ## and visualize with some jitter:

> plot(jitter(x), col = cl$cluster)

> points(cl$modes, col = 1:5, pch = 8)

