

## Homework 3)

### 2.3.

a)

**Full set of features:**

Value of K	Purity
2	0.4930
3	0.5399
4	0.4783
5	0.5165

**Filtered set of features:**

Value of K	Purity
2	0.3422
3	0.3663
4	0.3667
5	0.8910

b)

**Full set of features:**

	Cluster 1	Cluster 2	Cluster 3
Percentage Case	31.3%	0.0%	11.1%
Percentage Control	16.9%	7.8%	58.9%
Percentage Unknown	51.8%	92.2%	30.0%

**Filtered set of features:**

	Cluster 1	Cluster 2	Cluster 3
Percentage Case	33.4%	0.0%	88.9%
Percentage Control	34.1%	0.0%	0.0%
Percentage Unknown	32.5%	100.0%	11.1%

## 2.4

a)

**Full set of features:**

Value of K	Purity
2	0.4783
3	0.4783
4	0.4783
5	0.4797

**Filtered set of features:**

Value of K	Purity
2	0.4329
3	0.7779
4	0.8544
5	0.8620

b)

**Full set of features:**

	Cluster 1	Cluster 2	Cluster 3
Percentage Case	7.6%	28.0%	0.0%
Percentage Control	24.3%	25.7%	38.1%
Percentage Unknown	68.1%	46.3%	61.9%

**Filtered set of features:**

	Cluster 1	Cluster 2	Cluster 3
Percentage Case	28.2%	0.2%	0.0%
Percentage Control	0.0%	99.5%	0.0%
Percentage Unknown	71.8%	0.3%	100.0%

3)

b)

### Full set of features

Value of K	Purity
2	0.4783
3	0.4783
4	0.4783
5	0.4783

### Filtered set of features

Value of K	Purity
2	0.3374
3	0.3398
4	0.3467
5	0.3549

c)

### Full set of features

	Cluster 1	Cluster 2	Cluster 3
Percentage Case	29.5%	25.3%	25.2%
Percentage Control	25.4%	25.5%	27.7%
Percentage Unknown	45.1%	49.2%	47.1%

### Filtered set of features

	Cluster 1	Cluster 2	Cluster 3
Percentage Case	32.9%	33.9%	33.7%
Percentage Control	32.6%	32.2%	33.5%
Percentage Unknown	34.5%	33.9%	32.8%