innovate achieve lead

BITS Pilani

Machine Learning (IS ZC464) Session 13:

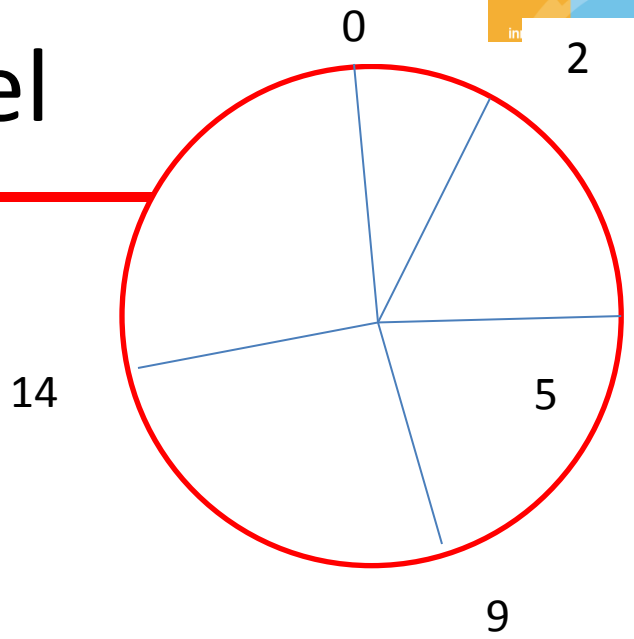
Genetic Algorithm – Parent selection and applications of GA

Roulette wheel

```
sum = 0.0
for all members of population
        sum += fitness of this individual
end for
sum of probabilities = 0.0
for all members of population
        probability = sum of probabilities + (fitness / sum)
        sum of probabilities += probability
end for
loop until new population is full
        do this twice
                number = Random between 0 and 1
                for all members of population
                        if number > probability but less than next probability
                                then this member gets selected
                end for
        end
        create offspring
end loop
```

Source of code: Google

Roulette Wheel



chromosome	fitness	cumulative	probability	Cumulative prob.
1	2	2	2/20 = 0.1	0.1
2	3	5	3/20 =0.15	0.25
3	4	9	4/20 = 0.2	0.45
4	5	14	5/20 = 0.25	0.7
5	6	20	6/20 = 0.3	1.0

Generate a random number in [0,1] and select a fit parent according to its probability

How many fit chromosomes should be selected

- If the number of chromosomes with fitness value greater than a specified threshold in generation t is p and if the crossover probability is p_c, then $p_c \times p$ chromosomes are selected as parents
- These participate pair wise in crossover to produce pair of offsprings of the new generation.

Mutation:Exploitation

- If the mutation probability is p_m then $p_m \times p$ chromosomes undergo the mutation process.

Crossover operation

- One point crossover
- Two point crossover
- Crossover mask
- Uniform crossover – combines bits sampled from the two parents

Feature Selection

- Given a set of features, select few most informative and discriminative features
- **Class assignment:**
  - Represent as a search problem
  - Represent as a chromosome
  - Define fitness function

23	29	10	2
5	30	12	25
52	16	9	2
15	11	12	13

(a)

c1	c2	c3	c4
c5	c6	c7	c8
c9	c10	c11	c12
c13	c14	c15	c16

(b)

c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	c14	c15	c16

(c)

23	29	10	2	5	30	12	25	52	16	9	2	15	11	12	13

(d)

1	1	0	0	1	1	1	0	1	0	0	1	1	1	0	1

(e)

0	1	0	1	1	1	1	1	0	0	1	1	0	0	0	

(f)

Genetic Encoding of Face Image (a) 4×4 window of 16 features as the initial pool of features (b) coding of features (c) codes as a vector (d) Feature vector (e)and (f) Two different Chromosomes with 1's representing inclusion of the corresponding features

Summary of GA

Characteristic	GA
Population	Chromosomes
Search Heuristic	Survival of Fittest
Exploration Process	Crossover
Exploitation Process	Mutation
Modality	Single
Convergence	Slowest
Overall merit in finding optimal solution	Weak: May get trapped to Local Optima

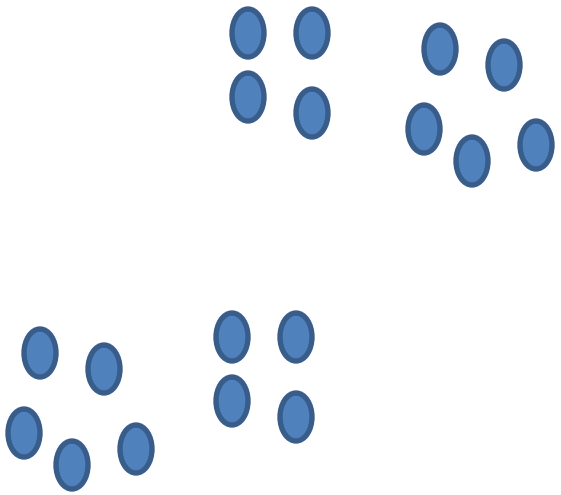
Comparison of the Evolutionary Algorithms

Characteristic	GA	PSO	FA
Population	Chromosomes	Particles	Fireflies
Search Heuristic	Survival of Fittest	Swarm Behavior	Attractiveness
Exploration Process	Crossover	Global Best	Attractiveness
Exploitation Process	Mutation	Local Best	Randomization
Modality	Single	Single	Multiple
Convergence	Slowest	Slow	Fast
Overall merit in finding optimal solution	Weak: May get trapped to Local Optima	Moderate: Only single optimal solution can be obtained	Strong: Reaches Global, Multimodal Optimal Solutions and is very fast

Application of Genetic Algorithm in Unsupervised Clustering

- Given training data without their categories (classes), the objective of unsupervised clustering is to group the data based on a similarity criteria.
- Training and testing data samples (e.g. images, speech, text etc.) are represented as feature vectors (with selected features).
- The data sample feature vectors are represented as points in a hyperdimensional space.

2D data



Given 2D points without any class labels.

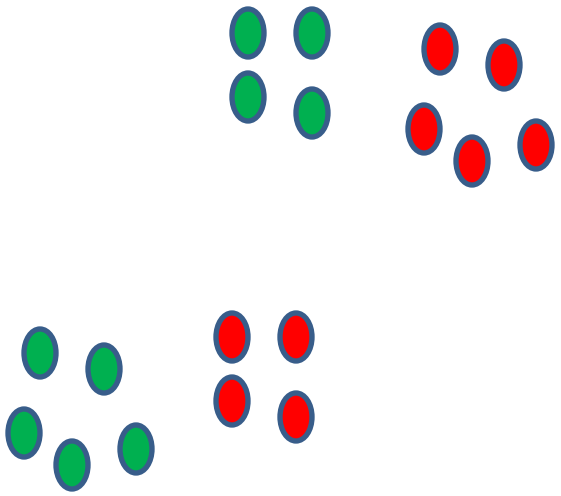
Objective: to group them in clusters

Specification required: Number of clusters

Criteria of similarity: closeness of features

2D data

If number of clusters = 2,
Then the following clustering is one way, but not same as humans do



Given 2D points without any class labels.

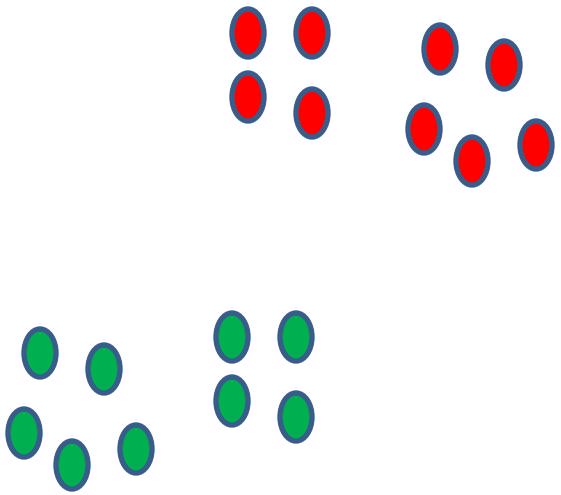
Objective: to group them in clusters

Specification required: Number of clusters

Criteria of similarity: closeness of features

2D data

If number of clusters = 2,
Human intelligence based clustering is as follows



We can call this the best clustering

Given 2D points without any class labels.

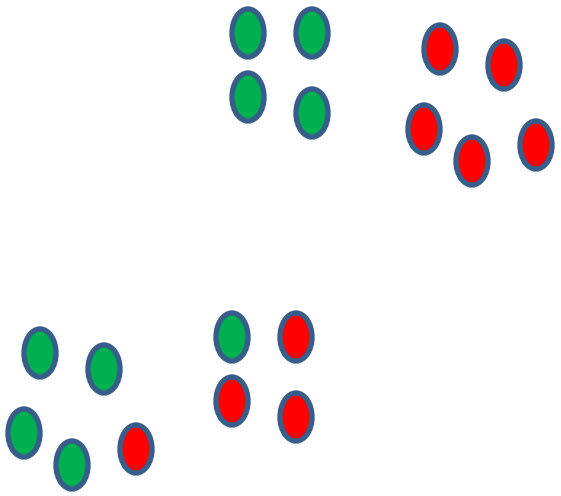
Objective: to group them in clusters

Specification required: Number of clusters

Criteria of similarity: closeness of features

Number of clusters = 2

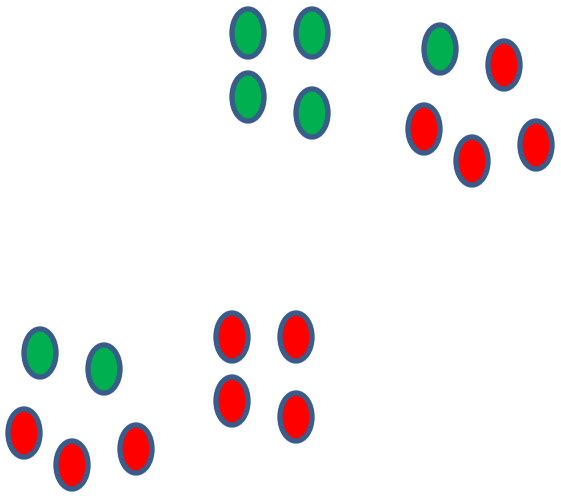
If number of clusters = 2,
Human intelligence based clustering is as follows



Is this good clustering?

Number of clusters = 2

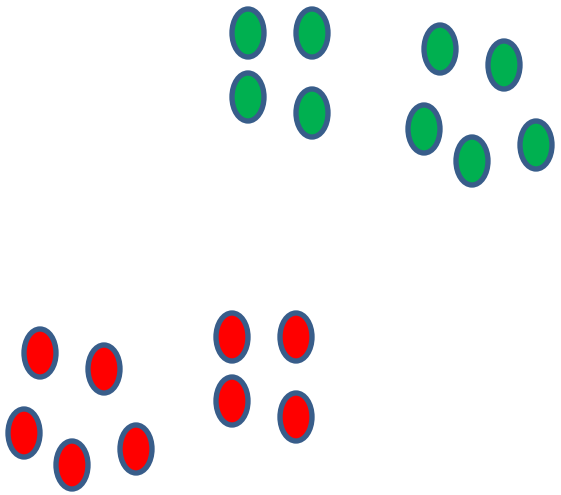
If number of clusters = 2,
Human intelligence based clustering is as follows



Is this good clustering?

Number of clusters = 2

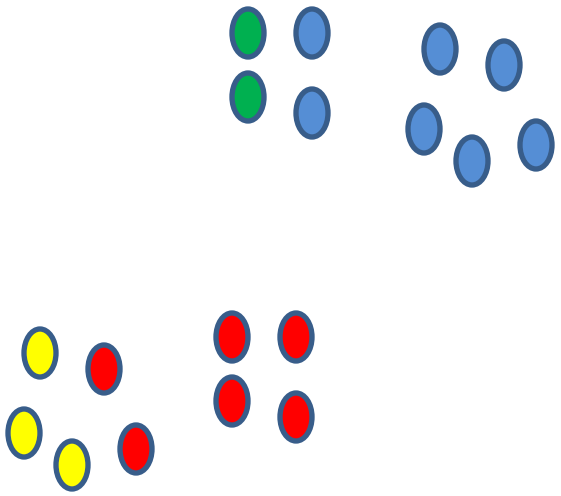
If number of clusters = 2,
Human intelligence based clustering is as follows



Is this good clustering?

Number of clusters = 4

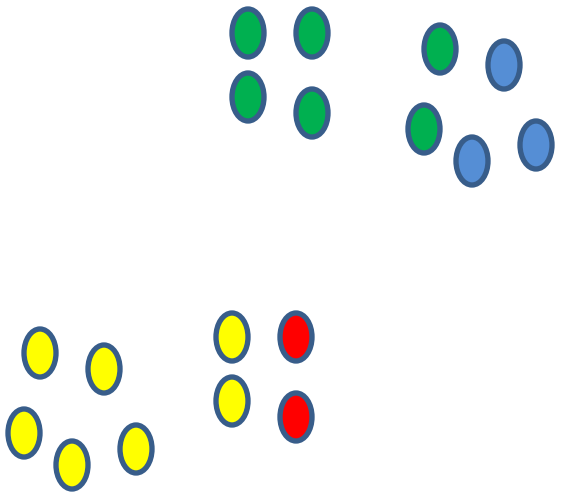
If number of clusters = 4,
Human intelligence based clustering is as follows



Is this good clustering?

Number of clusters = 4

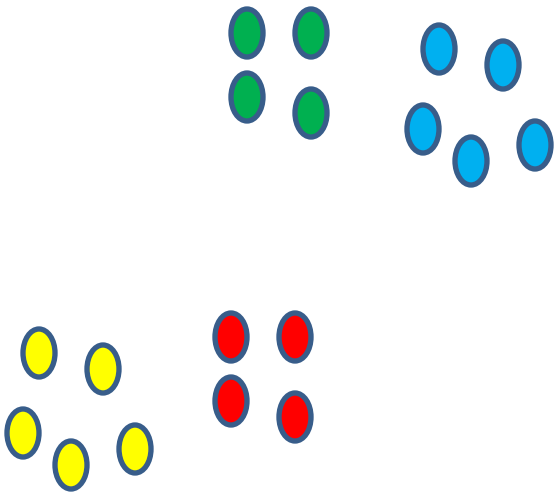
If number of clusters = 4,
Human intelligence based clustering is as follows



Is this good clustering?

Number of clusters = 4

If number of clusters = 4,
Human intelligence based clustering is as follows

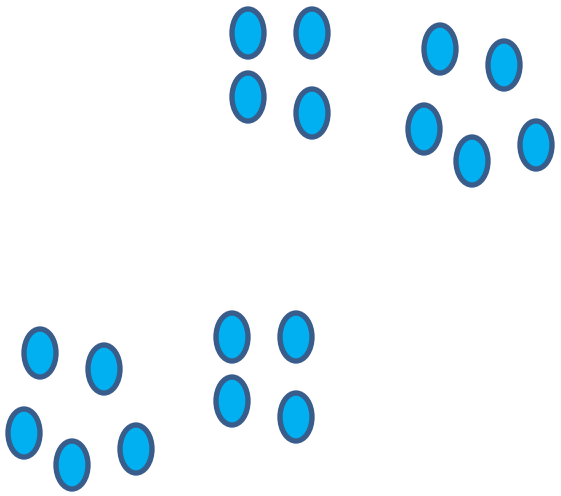


Is this good clustering?

What is optimized?

Distances of data points to the respective cluster centers

Clustering as an optimization problem



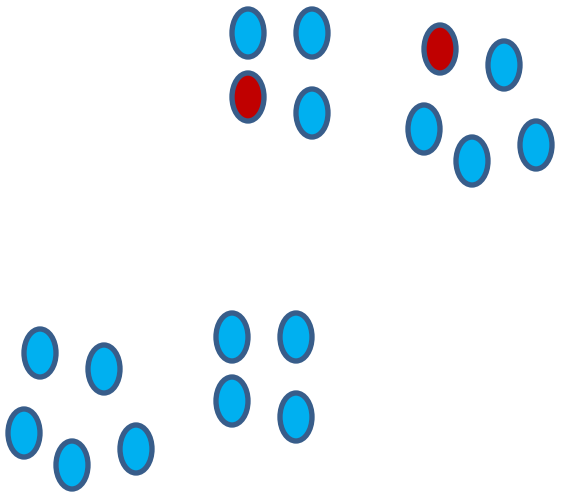
How do you think machine learns to group the data on its own?

Take some points randomly to mark them as cluster centers

If number of clusters = n

Then take ‘n’ random points as centers

Clustering as an optimization problem



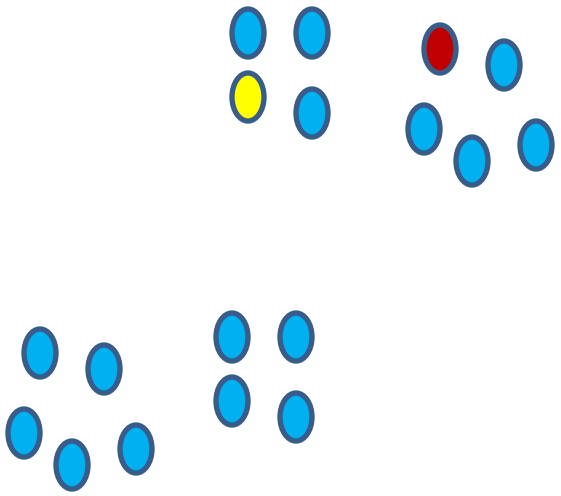
How do you think machine learns to group the data on its own?

Take some points randomly to mark them as cluster centers

If number of clusters = n

Then take ‘n’ random points as centers shown in red

Clustering as an optimization problem

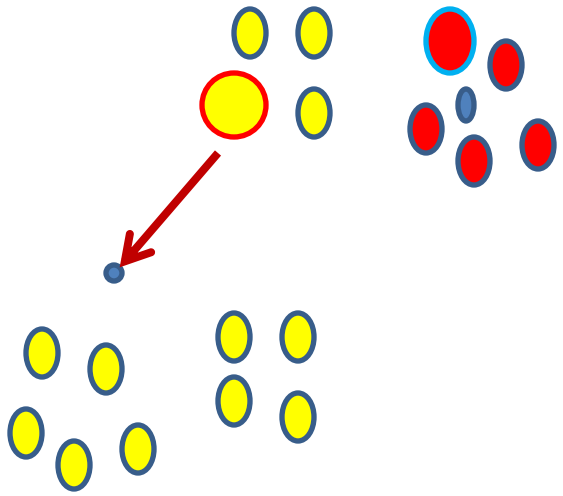


How do you think machine learns to group the data on its own?

Give different colors to these centers – say yellow and red

Iterate over all points to color them according to its closeness to one of the two points

Clustering as an optimization problem



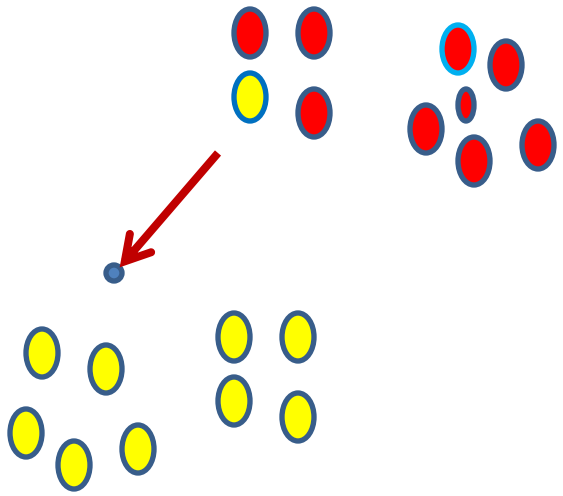
How do you think machine learns to group the data on its own?

Give different colors to these centers – say yellow and red

Iterate over all points to color them according to its closeness to one of the two points

Compute the mean vector of the two clusters

Clustering as an optimization problem

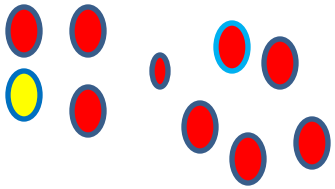


How do you think machine learns to group the data on its own?

Iterate over all points to color them according to its closeness to one of the two points

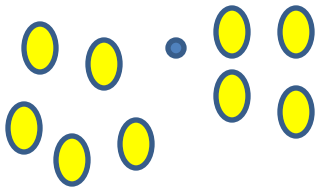
Compute the mean vector of the two clusters

Clustering as an optimization problem



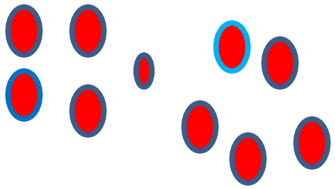
How do you think machine learns to group the data on its own?

The centers keep moving



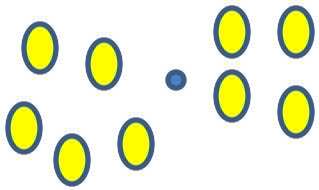
Compute the mean vector of the two clusters

Clustering as an optimization problem



How do you think machine learns to group the data on its own?

The centers keep moving



Compute the mean vector of the two clusters

View the sum of the distances of points in a cluster being reduced every time

Fitness function for clustering

$$F = \sum_j \sum_i d_{ij}^2$$

Where d_{ij} is the distance of i^{th} sample of j^{th} cluster from the corresponding cluster mean.

How to view the problem GA solvable?

- Represent the centers of the clusters as strings of real numbers representing coordinates of a center.
- Generate initial population of ‘n’ centers by randomly generating the values to form the chromosome.
- Compute fitness of each chromosome
- Select the pool of fit parent chromosomes and use Roulette wheel method for selecting best parents