Indian Institute of Technology Madras
Presents

NPTEL
NATIONAL PROGRAMME ON TECHNOLOGY ENHANCED LEARNING

Pattern Recognition

Module 01

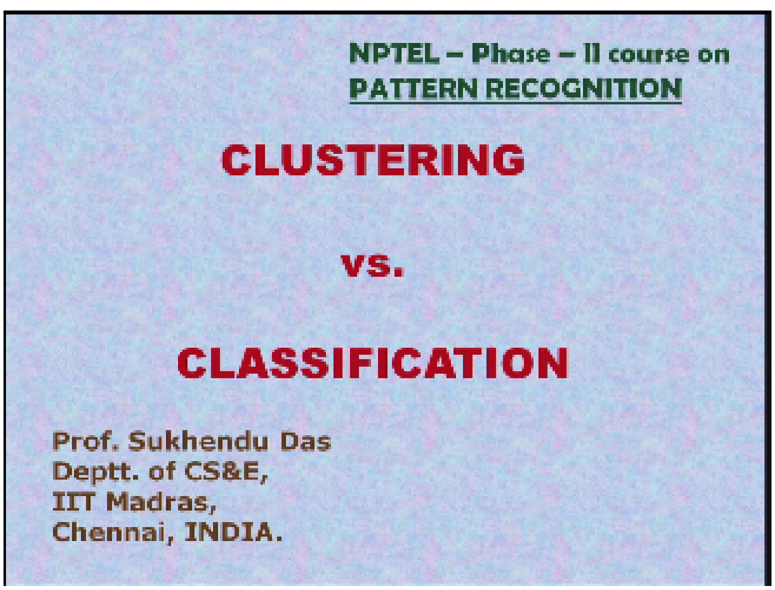
Lecture 04

Cluttering vs. Classification

Prof. Sukhendu Das
Department of CS&E, IIT Madras

Welcome to the 4th lecture of the course pattern recognition under the NPTEL phase 2 program in the 1st lecture you have heard for Prof. C. A. Murthy about importance significant and different applications and the purpose of pattern recognition in the 4th lecture now we will consider an important aspect in the field of pattern recognition in fact 2 aspect and we will try to understand the difference between the concepts of culturing and classification.

(Refer Slide Time: 01:06)



The difference of culturing and classification can be understood best with an example let us start with the example of classification first which is easy to understand and then look at the problem of clustering okay.

If you look into the slide what you now have actually or two images or two pictures of two different flowers if we ignore the presence of the background if we ignore the presence of background in these images then we were given the problem of classification to distinguish and say whit flower is one class and that appears different from the second flower which is of class 2 let us say.

Then you have to identify one property one attribute or one difference between these two images in terms of something which belongs to the flower and that property could be color here let us say because on the left hand side what you see is an image of a flower and on the right hand side you see also an image of flower but there colors are different on the left hand side the flower is of white and the right hand side you have the color which is yellowish of the flower.

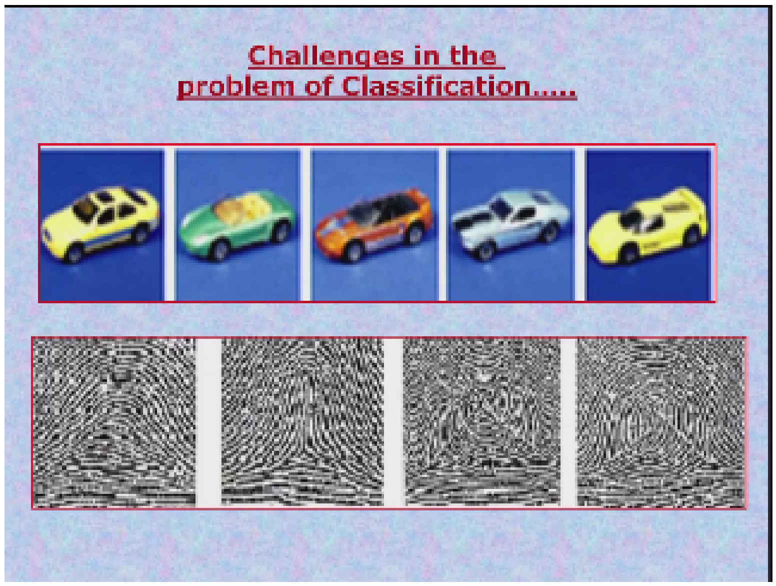
So if I pick up color as an important property which we will actually specify by the name feature later on and define what is a feature for the time being in a very simple sense let us say I take the word characteristics or attribute or property of the object from which we are solving the problem of classification then color is a very important feature here I will go back to the slide and can extract color from these 2 images basically the flowers.

Again I repeat ignoring the background then I can easily distinguish or anybody can easily distinguish between the flower on the left and the flower on the right of course there are few richer features more than color like the shape of the pattern of each flower which we will consider later on now let me bring a third object which is completely different as you can see that this object is different from the other two it is a not a flower of course in a animal but if you had used or if you use color as an important feature here you will not b able distinguish between the left flower on this white as well as the animal which is also having a white color.

These example shows that on certain vocations so if you want to distinguish between the 3 different types of images shown on the screen to you just now color will not satisfy the criteria of classification between a flower and the animal it might work in certain cases between the yellow flower and that the white animal but between the flower as also a white color and the animal also as a same color in this case white then you will not be able to discriminate between the 2.

So this example shows that you will need more than 1 property more than one feature more than one characteristic or attribute form objects from signals which we see which we perceive to distinguish on with respect to the other we will move forward and see more example.

(Refer Slide Time: 05:06)



Of classification look at this problem you have several set of cars the problem of classification now is to distinguish between different models of cars let us say correct if is use color again you can see that the car on the extreme left is the same as the one which is the color on the extreme right sample, so you have to use more than one feature many be the shape of the object the shape of the windows the position of the wheels with respect to the car various attributes of shape in addition to the color to distinguish between different categories of cars which are looking almost similar almost same.

I am not trying to distinguish ort classify or categories between flower as and cars, cars and other animals but between color between cars itself all cars look more or less the same they have 4 wheels they have front screen doors, windows and so and so forth that the shape and the size almost look similar of course the other types of vehicles as well but between cars when the objects look very similar the problem of classification can be more difficult look at the example in the same slide.

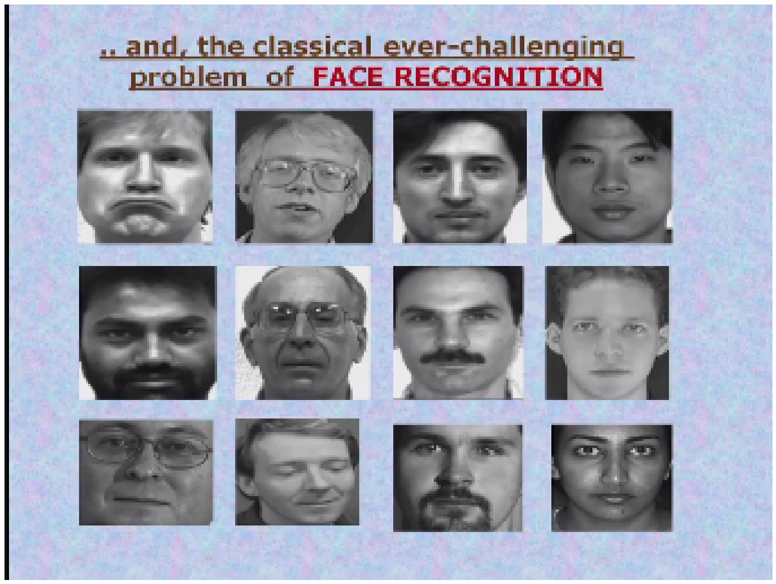
These are 4 different examples are figure prints scan from different individuals and there are exports who can actually distinguish between these 4 different categories of figure prints how I will not go into details but there are features extract from this type of lien diagrams as they appear the figure prints after they have been scanned and digitized and rectified exports who will find certain patterns like rich orientations the initial points okay rigid openings and closing to distinguish which mean one figure print and other which is a very good biometric trade. You can use figure print to uniquely identify an individual it as a huge application that is a huge application in various field of biometry access control and criminal identification in the case of forensics.

(Refer Slide Time: 07:38)

Let us look at this as another example of biometry you are seeing here examples of 12 different face images what is the problem of classification in this case it is called a problem of face recognition involves trying to identify an individual recognize it from the patterns of the face, now all faces look similar all of us have 2 eyes 1 nose lips and ear etc. We start may different then may be added patterns in the face.

So the variability of the patterns within the ways from one individual to the other in terms of eyes, nose gravel pixel and variation may be structured and something else I did not exploit by pattern recognition scientists to design a classifier. Which can identify an individual on that certain conditions it may not work well under all conditions but in certain yes it is possible.

(Refer Slide Time: 08:48)

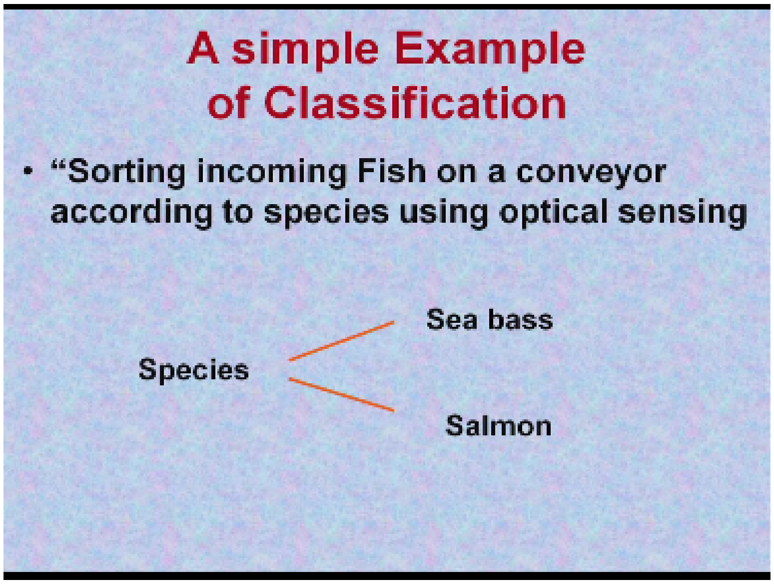


Now if you look back to this problem of classification of trying to distinguish and classify between the faces between finger prints between faces and finger prints hands between cause this problem was really different from the first easy problem which I showed you between two flowers or between the flower and an animal why the samples are very different the images were different the objects where different one with respect to the other in the case of facing condition all the faces in the identical.

The would form properties if you extract properties their properties will be very similar similarly for finger print similarly for cars or similar objects, so the problem of classification becomes difficult if you are trying to distinguish different type of objects within the same category different type of objects within category face recognition finger print recognition cars okay between type of say trying to identify patterns within a signature okay then very similar and the problems becomes very close to a clustering problem.

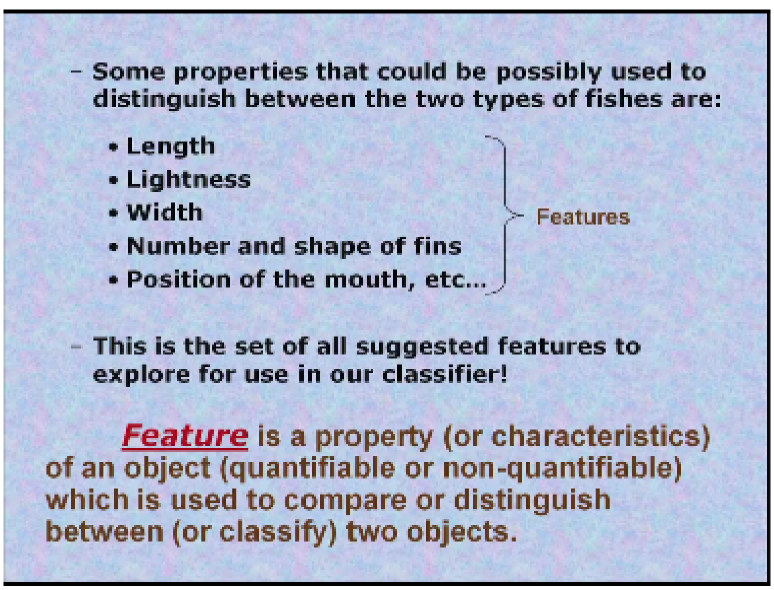
Which will understand little bit let us let it will become more specific with features and classifications if you look into this very simple example.

(Refer Slide Time: 10:19)



Which has been taught of in many books sorting incoming fish on a conveyor according to the species using well optical engineering or optical sensing, so what we do with the sample face there may be two classes of fish which we are trying to classify a categorize these are the two for examples but you can of course take any other a samples of fish which you want it could be very classic verities of the hills and the row which you find in India the what you do with these species.

(Refer Slide Time: 10:52)



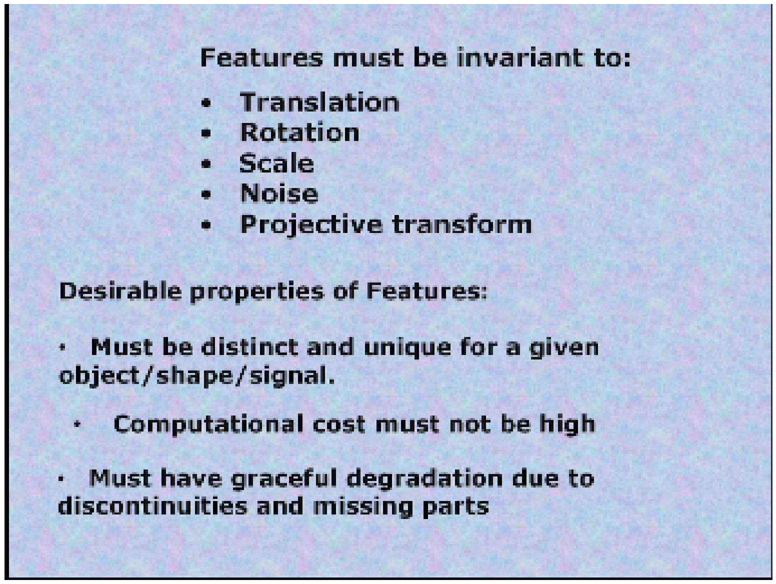
You try to extract certain properties that could be used to distinguish between the two type of fishes the length of the fish the brightness of the lightness of the fish the lightness could be the brightness of the intensity the width of the fish number of shapes of the fin position of the mouth which respect to the body et. These are some examples of properties or characteristics which we can measure of course there are few other piece which you can measure like the weight of a fish something which you probably cannot measure like the smell.

Or the touch so these are important properties of characteristics which will help us to distinguish one sample of a category from the other in this case could here in different categories of fish, okay so these are some examples as you can see in the slide or example of water called on the right hand side examples of features which one has to extract so that the steps in distinguishing one class of object from the other one these are all suggested features which can be used of course you can use a few more if you extract and then use it in your classifier this is a definition of the word feature.

Which it may find possibly in some dictionary it is a property or characteristics of an object sometimes quantifiable sometimes not then which is used that is a main purpose which is used to compare or distinguish between or classify two objects I repeat it is a property of an object which is used to compare or distinguish between two objects or classify two objects.

So hence first in the rest part of the lecture will deal with these features we have seen some examples of features repeat them from the beginning of this lecture color, shape, intensity, size, initiate features, length, width shear type features we will some examples some more examples position of certain important landmarks of the you and account of certain landmarks in images are objects using these features the question comes as comes is how can you do classification and of course the other question is what is clustering.

(Refer Slide Time: 13:14)



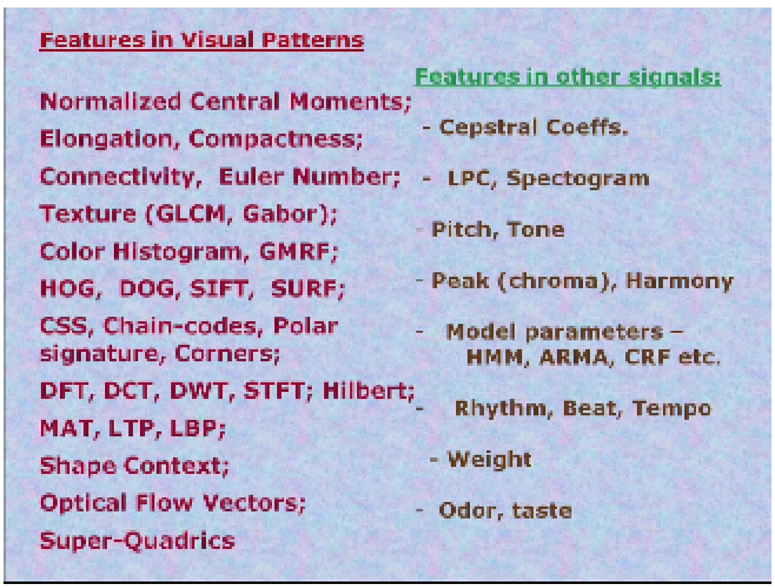
But before we go there we will try to see some very important properties are desirable properties of features, features must be in variant to translation rotation scale noise and other type of projective plan sorts this is much to do with images what does it mean even image if you rotate the object keep the object anywhere in your screen scale the object make it bigger or small or if there are noisy artifacts in image or if you have projected transforms that sequence that if I am going away from you.

And appearing smaller this classic example of scale where it is due to projective transforms the classifier should be able to still economize my face in variant to where am I compared to the senor of the camera which is speaking here features are expected to be invariant to all sorts of an transformations go back to the slide some examples are rotation translation scale noise and projective transformation other desirable properties of features it must be distinct and unique fro a given shape object or signal.

Okay now if you have these two objects okay one is USB mouse the other is basically a pen okay you can see that the features like as for example if you taking a intensity as an example and the short of the ratio of the hide versus with if you take that a feature there will be different for two different objects, so going back to the slide it must be distinct or a unique for a given object the cost of competition must not be very high for a feature because otherwise the classifier will let take lot of time to do the processing and must have graceful degradation due to discontinuity is an missing parts that means.

If the picture is not very clean and there are certain discontinuities in the object which has been images which has been imaged still the features should not show much of variations the feature should not show much of variations it would still pick up similar identical value of the feature from the image.

(Refer Slide Time: 15:36)



These are some examples of features which are using visual patterns it pick up you do pattern classification mostly with the images you can perform classification with non visual signals will also come to that in moment but let us look at visual patterns which is one of the most important example of pattern recognition some of these will be explain later on we just let us name them these are something to do with shade features normalized central moments elongation, compactness, connectivity, Euler number they need to sometimes model the background or a foreground texture of an object, so you use measures like grey level cohesiveness matrix or GLCM you can use filters like Gabber filter, wavelet filters and so on.

You can use color histogram, GMRF or Gaussian Marker Random Field and these abbreviations which you see are very rich features which have been just you know discovered or worked up on over the last decade also. HOG is histogram of oriented gradients I repeat again histogram of oriented gradients which is HOG, DOG, dog is called the Difference of Gaussian I am sorry I should repeat again HOG is histogram of oriented gradients, DOG, dog is difference of Gaussian.

Scale in Variant Feature Transform is shift and SURF is Speeded UP Robust Features, Curvature Scale Space CSS, Chain Codes, Polar signatures, Corners, DFT, Discrete for a Transform, DCT Discrete Cosine Transform, Discrete Wavelet Transform DWT, short term for a transform the STFT and Hilbert Transform these are different type of transform which I used.

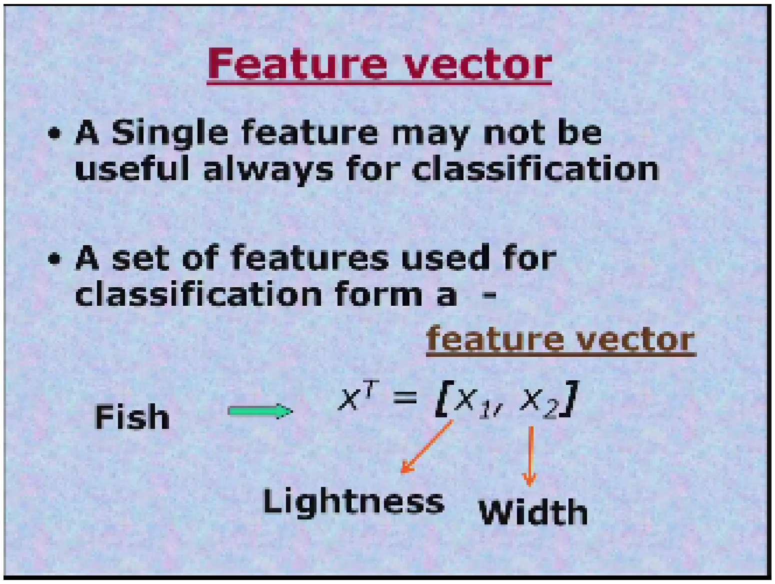
MAT is Media Access Transform, LTP is Local Turnery Pattern, LBP is Local Binary Pattern, Shape Context very rich features for representing shapes. For motion people use optical flow vectors and also super quadrics for molding blobs when they see it in images. These are some examples for visual patterns which you can detect from images; the pattern recognition area is applied for any type of signal.

Visual as where as non visual something which you hear say an audio signal let us go back, features use in other signals these are examples of features which are used in audio, music's, speech at the Cepstral Coefficients, LPC Linear Predictive Coefficients, Spectograms, Pitch, Tone, Peak, chroma, Harmony. Then of course you can have model parameters which can do for visual and non visual like the Hidden Markup Model HMM.

The RMM model, Auto regressive moving average, Condition Random Field CRF, some more similar to GMRF. It also extracted feature like Rhythm, Beat and Tempo from music and of course you can measure Weight, Odor, taste fine you can use them as features also, that is a huge list of features some of them an odor exhaustive, some of them by could not mention because there is a technology in volume and some of them.

Like PCA features we will talk of them later on, all the LD or the CIA features we will talk of them as the course goes on.

(Refer Slide Time: 18:47)



With respect to facture we should know what is feature vector, because the single feature is not useful for the purpose of classification and a set of features are used together to perform a classification they form a feature vector. So what is the difference between feature and a feature vector well you can say a group of features are combined together to form a feature vector, what is the dimension of the feature vector.

The number of components or the number of features which have extracted to form the feature vector. Let us say, you take three different features color, the area of an object and say the extensity of an object I repeat again color, area and any arbitrary moment like extensity or moment driven parameters like perimeter you can take perimeter area and color. Let us say this very three simple features color, area and perimeter are the three simple features which you want, you want to use all the features because you know that if I use all these features together I can do a problem of classification much more accurately then picking up any one or two of these.

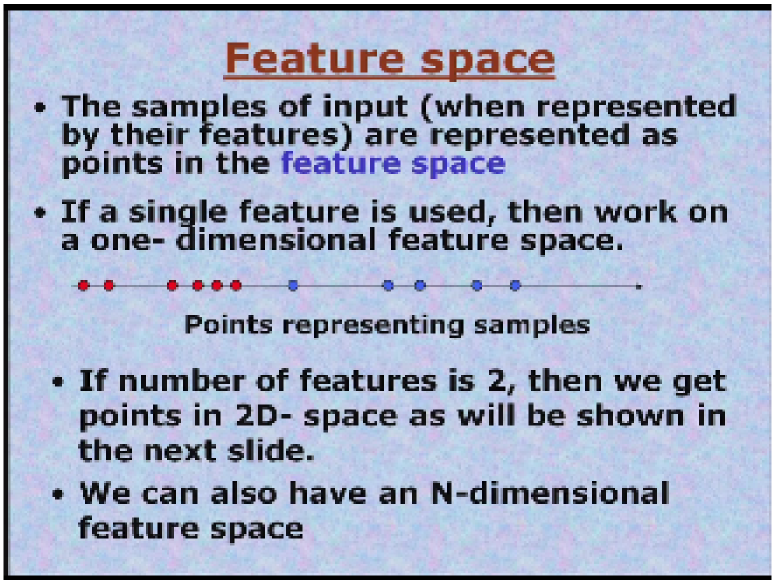
So when you take all of these three features together you form a feature vector whether the first component all the feature vector is color, the second could be area, the third could be perimeter or it could be something else also does not matter. So what do you expect this feature vector extracted from a sample belonging to one class or category of an object should be different from the feature vector extracted from another object.

And the problem of classification would be to distinguish between these two feature vectors belonging to two different class of objects or if possible learn and know from a priory samples as to what this feature vectors would be for sample belonging to class A and how they are different from samples belonging to class B and that is the problem of classification and this will lead as automatically to the idea of clustering as well.

As for example if you go back to the slide, this is an example of a two dimensional feature we talked about this pair of features to be extracted from fish and if the two features which we extract are say brightness and length forming the two components x1 and x2 as they are indicated in your slide then X to your left is a feature vector the transpose indicating that we typically talk of this is a column vector then a row vector.

So the transposes given as a row vector then the feature vector will be of dimension 2, one could be brightness the other could be width or they could be something different. So if you select or identify instant of 10 feature vectors, sorry 10 features from on object if selected of 10 features from an object your feature vector will be of dimension 10 or even larger if you have more number of features.

(Refer Slide Time: 22:09)



Feature vectors lie in a feature space this has to be learned it is something like the vector space, samples of input when represented by their features together are represented as points in feature space. If a single feature is used then you are working in a one dimensional feature space, so now you have to learn and you have to know that a feature vector is now represented as a point you know vector has a magnitude and orientation depending up on the value of its components.

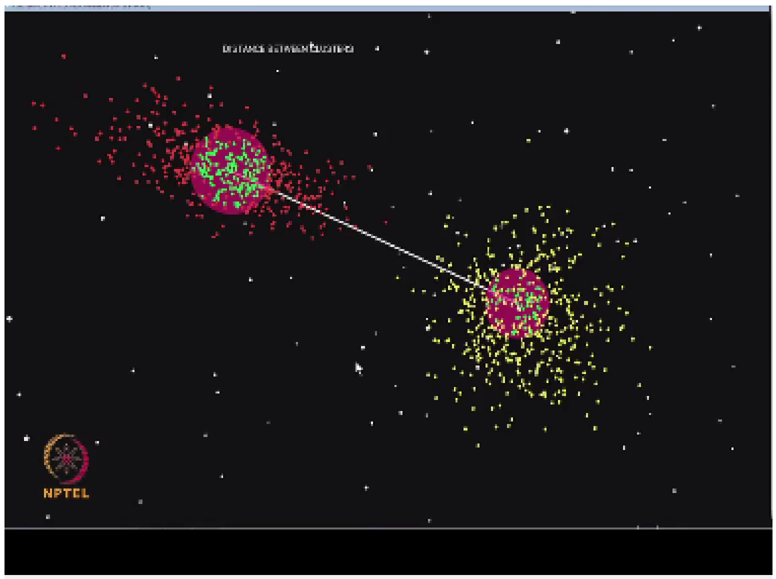
But in a feature space which is like a vector space which could be one deep to the minimum by generally very hard dimensions of course we can visualize up to three dimensional only, but if we have a large dimensional feature vector it is basically indicating a point in that feature space, okay. The feature space dimensionality is dictated by the feature space dimensionality is dictated by the dimension of the feature vector.

Of course you can have variations like one dimensional case where these are examples of points representing sample in one dimensional feature space, what does this color indicate it is possible at the red colors could be indicating points of features extracted from one category of the object and the blue color could be representing another category okay, to your right.

But you have just extracted one feature and hence the feature space is one dimension and the points are marked in 1D of course you can have features in two dimensional also where you can have points in two dimensional space which you will see in the next slide. But in general you can have an N dimensional feature space. Next what we are going to see is an animation which will take you into the feature space, since our visualization is restricted to 3D.

So you will see this animation which will take you in 3d space and you will see a set of feature vectors sometimes they form a cluster sometime they form two disting clusters or a vied looking cluster will see what is this cluster, and what is this classification see the difference between the two in this animation and come back and continuing the discussion between of the difference between clustering and classification.

(Refer Slide Time: 24:34)



So let us look at the demonstration in this screen now this will give you a visualization of a feature space in 3d you may actually fell that you are traversing in space in a shuttle what you have to realize is that each of this dots on the screen represents a point in a feature space is also feature vector. Now as we are passing by you will not notice any particular type of pattern or group or cluster or anything remarkable which would be happening here.

But as we move ahead in this screen you will start to notice a distance at a points which are now appearing as a dense set of points a compact cluster a dense cluster the distance between appear of points which are marked in yellow are much less in general compare to any two points which are in the background mark in white, so you see a cluster being formed here which is marked in yellow this set of points form a dense cluster the distance average distance you can say if you are able to measure it between any pare of point sis much less then if you take any points which are marked in white and there in the background let us say this distance is much larger than any pare of distance is which you see here.

The question comes is how would you estimate that this is the dense cluster, you must have some measure by which say the cluster density can be estimated from this feature space, remember we are absorbing this space in three dimension but in general a vector can be a very large dimension and the same thing may occur in a large n dimensional space. So how to measure the density if you move on this space you will find after sometime 2 fears 2 spherical volumes or fears marked in two different colors around the set of cluster points marked in yellow.

If you pick up any one of this pare of fears you will find what we have done is the points in yellow which are within this fear are actually marked in red you can see here they are marked in red. Although the revel acting remember this is a 3 dimensional visualization, so points within the blue fear are marked in red points yellow points within the red fear are marked in green I have shown two fears just to indicate that if we make a count or do an estimate of the number of points inside this fear it will give you an indication of density the definition of density from the basic definition of density we know that it is mass divided by the volume.

So assuming that the volume of both this fears are identical if the number of points within a particular fear is large then we estimate this the density at this point we much larger than the density at some other place, but in general the density estimated around any region within the cluster or any space within the cluster will be much larger than if I would have put this fear here in the outer space where you do not have a cluster but a few spots set of points it is possible that this cluster could be generated by a set of feature vectors obtain for a certain class or category or type of object.

What are these points they may not belong to a particular category or class of object. So this is how you measure density of course you can replace this fear by unit cube and an estimate the number of points which actually synonymous to represents the mass at that region divided by the volume unit volume this you can visualize this to be units fear or unit cube will actually give you an idea of density.

So cluster density is one common measure use to detect the presents of a cluster I repeat again the density will be having reasonably large value when you measure it somewhere in this region of the cluster if you measure it somewhere else you will have a much less value of cluster density indicating the absence of a cluster. So that is what we saw as a cluster do all clusters appear same well may not let us move a little bit forward in space.

On left hand side appear screen now you see two different clusters one marked by red dots the other which we had earlier where in red dots they are two different clusters, what is this significant of this two different colors which I have used to indicate the two clusters well it is possible that the feature vectors here indicated by this cluster could have been measured from images or signals or certain object categories resembling one type of object and this could have been obtain from another class of object you have two classes class A and B class 1 and2.

And this picture although resembles two different clusters in which I can put this two difference fears find the count of the number of green points within this fear out of the red and the green points out of yellow within the red and estimate density of both this clusters indicating that I have two disting clusters here there is no cluster is that because when you measure density at the back ground somewhere here you will not detect the presents of any cluster because the cluster density value will be much less.

Whatever you estimate will be much less in the background, this figure also illustrates the problem of classification visa vise clustering we have two different clusters all right but the clusters are separated in space there is a distance between them and the distance between the clusters indicates how far or this two different classes are categories or object types for which we want to do pattern classification task.

Remember the difference you have to detect clusters is one job of clustering, after you have detect a clusters let us say you have detected more than one clusters it is possible that the two difference clusters could be belonging to two different object types and then you can actually find the distance between the two clusters by one of the simple method is we can compute average or the centroid of the set of clusters for the red points or red markers for the yellow markers.

You can also find the mean or what is called of the class mean average or centroid of this cluster and look at the distance this distance is called as separablility between clusters I repeat again this distance will give an indication of the separablility between clusters the larger the classes or categories are separated in space the more over will be the distance the less the separation we will see the illustration of that less the separation the distance will start to collapse or the distance will start to come down if you move a head.

Now what you see here is the interesting phenomena of I have removed the color of the dots I will remove the color of the dots if you remove the color of the dots that means you do not know which point is belonging to which class you can still see the two clusters is space very clear because in fact now I have removed the back grounding site you can detect this two clusters using cluster density estimation techniques one of them which you have just discussed is this put a sphere and measure the number of the points inside this sphere.

You will get some idea of the cluster density you can also measure the distance between these two clusters if you are able to detect them but the question comes is if the class labels as they have got which you are given by the color markers remember this cluster was not in red this color was marked this cluster was marked by yellow markers what we only have coloring is the points which are insert this sphere.

At the only indicating which we have we have lost the cooler information then it becomes a different problem of classification sometimes it is called an unsupervised classification this is the case when you have to detect the clusters and assume that each cluster belongs to a separate class or category or type of an object this could resemble a particular type of fruit this could represent a particular category of flowers let us say okay.

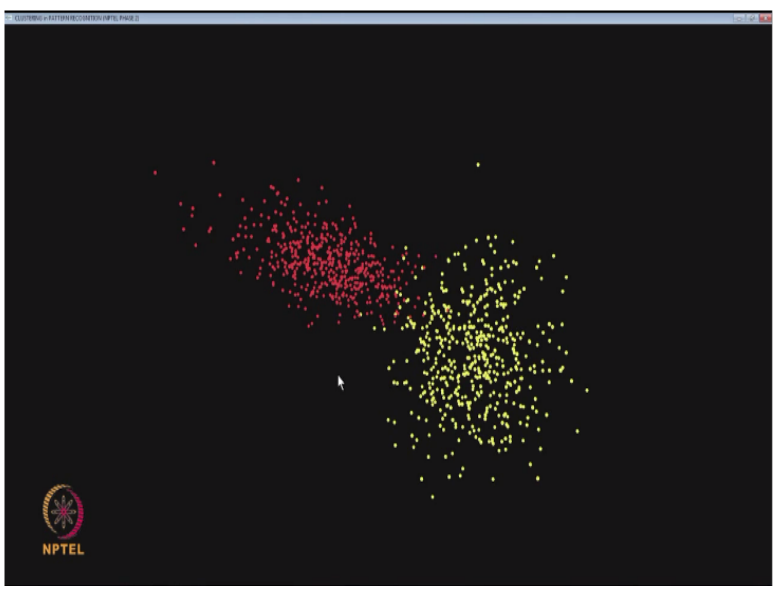
This still good point here which we now have at this point of time until this point of time is this the cluster a still separated in space that means you have some distance based on which you can separately detect this clusters the problem can be more diffciu7lt as you can see now that6 I am bringing one cluster close to the other one.

And you can see the region here very clearly you can see the region very clearly that the point sustaining overlap this is the region were classification becomes every difficult p0robelm if you visualize all the set of the points together can we have the visualization yes, look at it now.

(Refer Slide Time: 35:31)

This could actually represent the cluster by itself the cluster will not have regular appearance the cluster may not have regular appearance as we had for this two distance clusters when they separated out but this could actually represents the cluster and it is this situation where the clusters starts over lap that the problem of both clustering or the problem of classification if we introduce back the colors.

(Refer Slide Time: 36:01)



Let us see now how it looks like look at this problem this problem is not that difficult but we are entering a situation where the classify given the fact or the class labels that these points belong to one category class A the points in yellow belong to another category class B the classifier will work reasonably well perhaps if you have a classify design but it will a like a erase here.

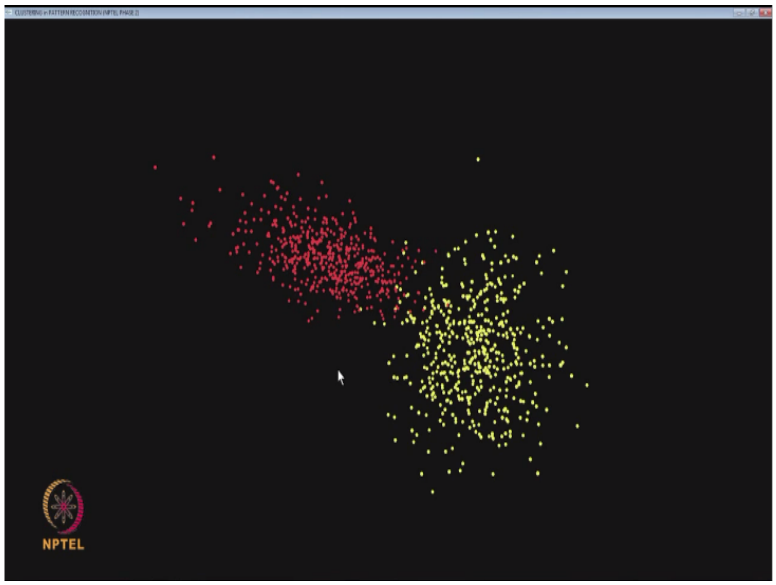
The region where the cluster points starts overlap between classes so you look into the problem of classification clustering here which we are trying to discuss this is can be consider as problem of classification if you remove up the colors.

(Refer Slide Time: 36:51)

It can become a problem for clustering even here you can detect two clusters but I am not sure whether a very naive classifiers or naive clustering mechanisms can detect two clusters permanently it might detect two it might be detect three it might detect four if the number of clusters or not known if numbers of clusters are not known you night think that this is the cluster which has a very regular shape.

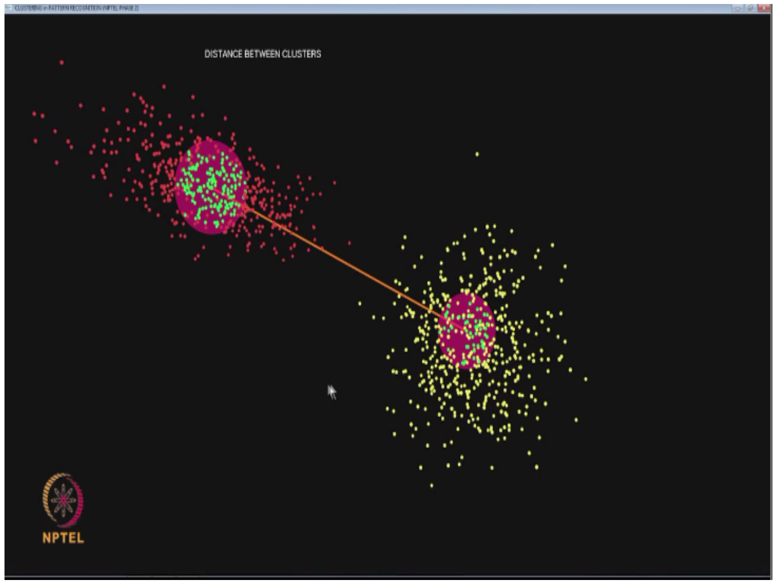
That sir possible we will give some examples later on of certain shapes which are geometrical nature or like but then is not very smooth but of course the other problem which is not clustering is the problem of classification as you see here.

(Refer Slide Time: 37:33)



In this problem of classification I have reintroduce the class levels with the help of color markers and the problem of classification is risk imitate features or feature vectors belonging to one class this or this or with respect to features in the other class the difficulties what in this case for classification that there is an overlap. The distance between the classes which much less, is much less compare to what we had earlier look at this problem.

(Refer Slide Time: 38:03)

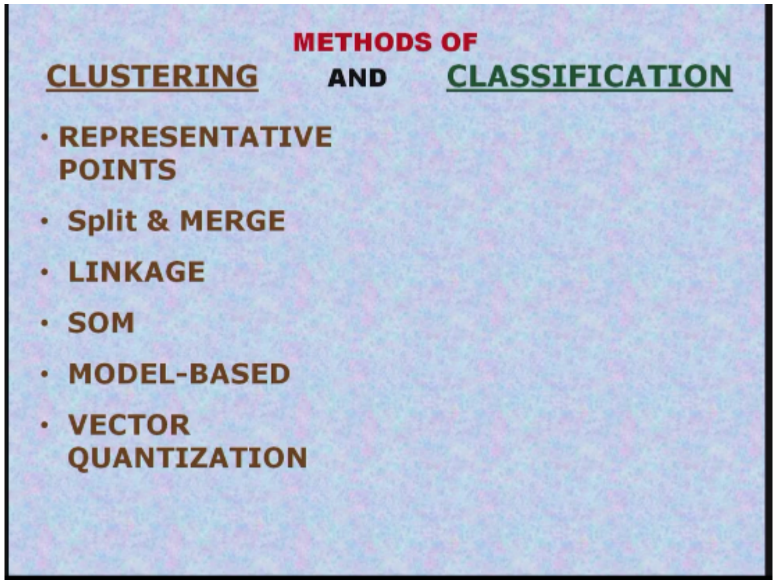


Is this much better situation for both classification and clustering were what this called as a inert class distance between these two classes of the feature points is much, much larger compact of course this cater or this spread or this width of the individual clusters there is hardly there is in fact no overlap they can actually lie draw a line here and say to my left of this line I have all the red points think of an imaginary line here or a point along this line.

To the left I have all the red points with respect to this limn as you can see here which I m drawing on the screen to right I have all the yellow points so classification becomes easy if I bring them close see if I bring them close as you can see that the cluster distance now is very minimal there is lot of overlap there is the problem of classification to produce accurate results.

So we have to see two scenarios now were classification and clustering could be easy these re the situations were classification with class levels or If I remove the color information clustering also could be difficult problem I will leave it to for visualization in these two clusters over lap quite a lot more then you can actually have just one visible cluster of out of two and even if the class information of the color levels on the overlap is given the classification might give certain good results.

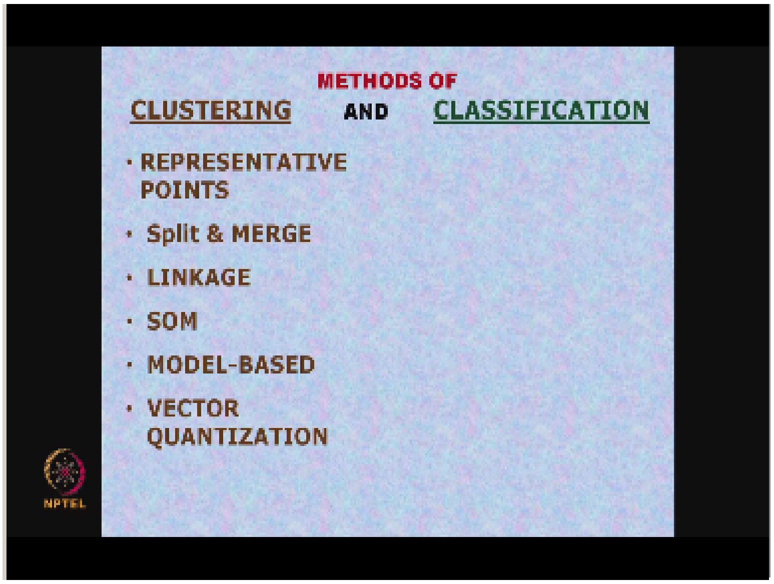
(Refer Slide Time: 39:34)



So will now look at two different methods relative methods few methods of clustering nd classification we have so far understood the difference between the two what it takes or what does clustering problem of clustering solve we saw the problem of classification we have seen it we have seen some animations some static pictures and also a table containing discriminating points between the two. Now we will just list a few methods of clustering and classification a few state of the art algorithms we just name them mind you.

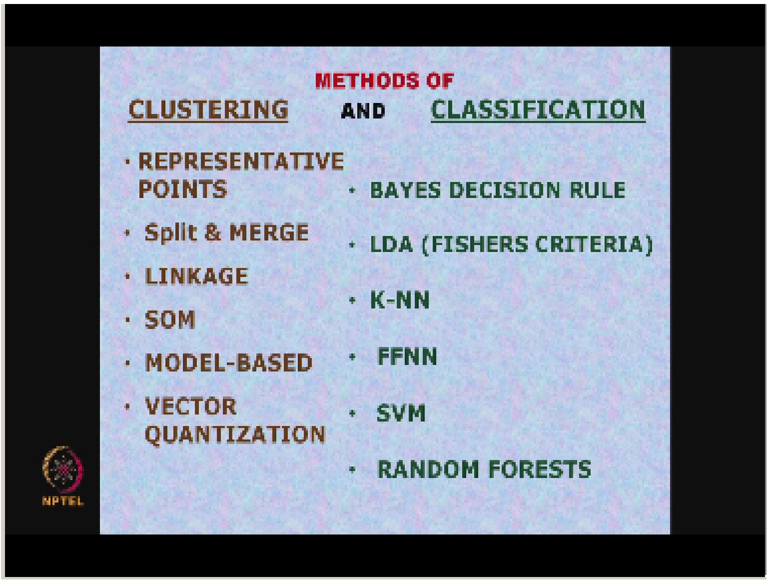
Because now what this you are at a point where you can understand how we can write a algorithm to solve this problem and we are going to do that to the rest of this course discuss many methods of clusteri8ng as well as classification but we will just name them now keep them in mind as well go through the rest of the course where these methods will be discussed in detail we will go back to the slides.

(Refer Slide Time: 40:43)



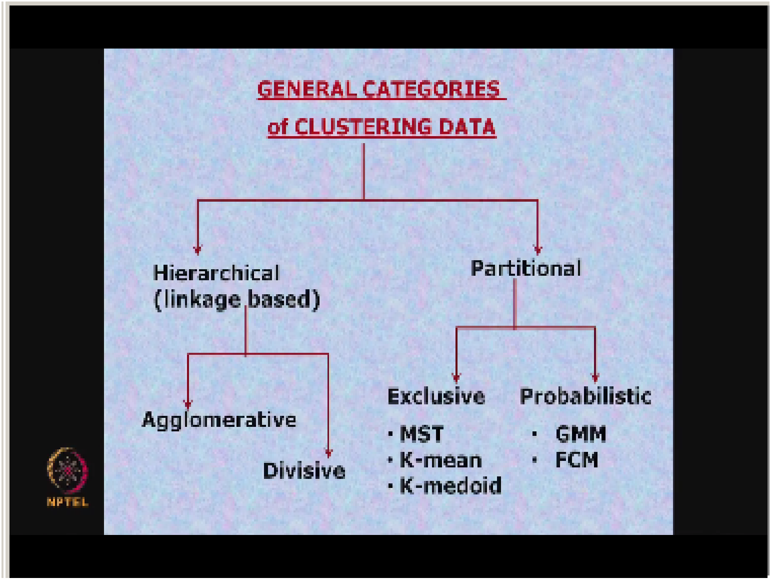
You have a list of methods on your left hand side column which indicate which are names of the methods used for solving the problem of clustering representative points, spilt and merge, linkage there are many methods of linkage we will discuss that later on SOM indicates self organizing map a special class of new networks which can also serve the problem of clustering although actually they do solve a problem of mapping model based methods vector quantization we will discuss most of this to the rest of the lecture series of pattern recognition not today what are the examples of classification.

(Refer Slide Time: 41:30)



Well a classic best example is the best decision rule which will be discussed in detail then we have the linear discriminate analysis or LDA which is called the fishers criteria then we have methods based on k nearest neighbor called the KNN the feed forward network is the most commonly used architecture under the class of artificial networks or ANN I repeat it is called the feed forward network which uses the back propagation learning law support vector machines the most popular and the power one which is rate and random forests will discuss most of these as we go on before we end up the talk.

(Refer Slide Time: 42:11)

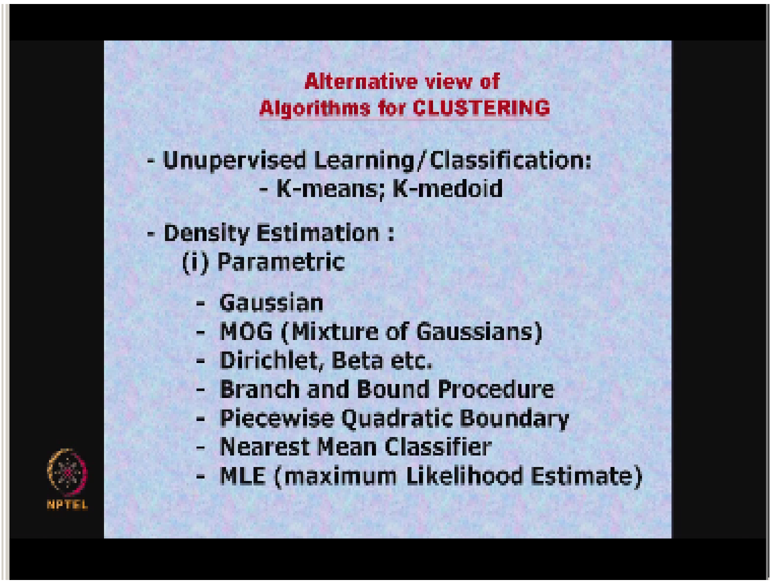


We would like to see the general categories of methods of clustering data usually most books will be categorizing them into two different classes one is called the hierarchical or linkage based the other is called the partitioned methods within the hierarchy kicker you will have a divisive policy and at agglomerative method of clustering again I will repeat you have two methods called agglomerative and the divisive methods of clustering under the hierarchical category of methods partitioned methods.

Some examples again two categories exclusive or probabilistic under exclusive you have the minimum spanning tree which use the representation of graphs the came in or the camel and under probabilistic you have the GMM called the gossans mixture model and the FCM as it is called which is the extension of the came in or see algorithm this categorization of methods of clustering is a very generic one it may be that when you go to certain literature.

And read from books the clustering these categorization of the different methods of the clustering may appear little bit different then what I gave you but off course you still have methods which are based are partitioned which are hierarchical and some of them are probabilistic and some of them are exclusive I have just confirmed myself to discussion of between these four different categories of lectures.

(Refer Slide Time: 43:52)



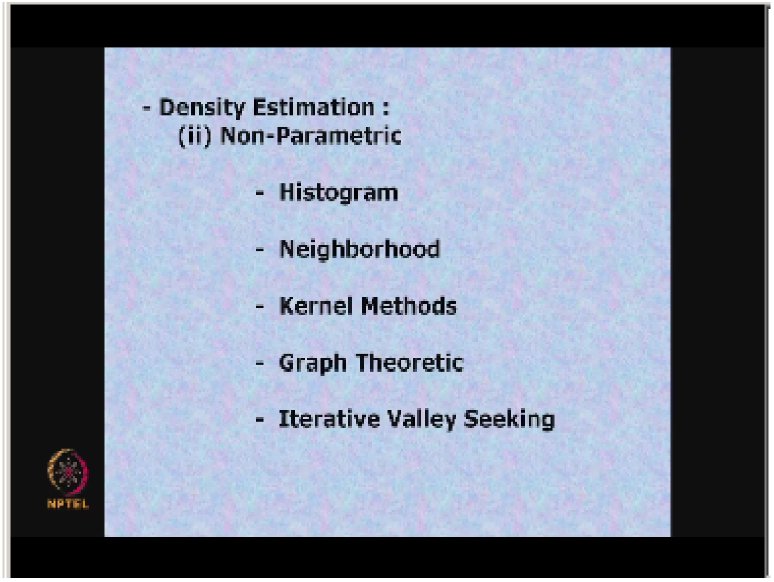
There is alternative view for algorithms of clustering also some would like to say an unsupervised method of learning is also clustering and unsupervised classification is also called clustering examples are K means and K medoid this is where most people will confused do you think unsupervised learning are classification is clustering well may not clustering has no class patterns unsupervised learning has no class levels.

But the purpose is different in the method of clustering you try to group data try to find structure within the data in the case of unsupervised learning or classification you tend to form groups that is also true you may deduct clusters that is also true but actually they not same this is not voice to call a method of clustering as an unsupervised method of learning or classification in a very loose sense you can tell that.

But that cannot be actually true other methods involve density estimation which we talked about when we are using as spheres to indicate how to measure density in a feature space typical examples of density estimation couple of some of them could be parametric type that means it estimate certain parameters from the density Gaussians mixture of Gaussians as it is called GMM is the classic example you can take non Gaussian distributions also.

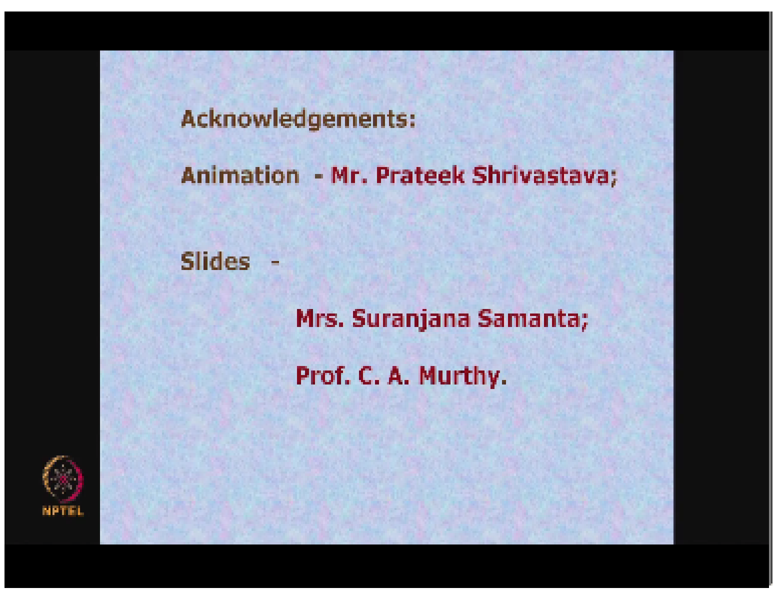
Although they are not popular the dirichlet or beta distributions then there are other methods called the branch and bound piecewise quadratic boundary nearest mean classifier or maximum likelihood estimate these are just names of different methods of clustering based on which are parametric density estimation there are non parametric density estimation also.

(Refer Slide Time: 45:42)



As given in this slide which are also often used methods based on histogram nearest neighborhood or kernel based methods graph theoretic methods interactive valley seeking some of them we will discuss if not all of them in the remaining part of all lectures I would like to conclude the lecture today

(Refer Slide Time: 46:04)



Of clustering versus classification with acknowledgments for my master students Mr.Prateek Shrivastava for helping me in creating this animation and some of the professor C.A.Murthy for helping me to prepare these slides thank you and get back to the remaining set of lectures to understand all different methods of clustering and classification and many other applications thank you.

Online Video Editing/Post Production
K.R.Mahendra Babu
Soju Francis
S.Pradeepa
S.Subash

Camera

Selvam
Robert Joseph
Karthikeyan
Ramkumar
Ramganesh
Sathiaraj

Studio Assistants

Krishnakumar
Linuselvan
Saranraj

Animations

Anushree Santhosh
Pradeep Valan.S.L

NPTEL Web&Faculty Assistance Team
Allen Jacob Dinesh
Bharathi Balaji
Deepa Venkatraman
Dianis Bertin
Gayathri
Gurumoorthi
Jason Prasad
Jayanthi
Kamala Ramakrishnan
Lakshmi Priya

Malarvizhi
Manikandasivam
Mohana Sundari
Muthu Kumaran
Naveen Kumar
Palani
Salomi
Senthil
Sridharan
Suriyakumari

Administrative Assistant
Jnakiraman.K.S

Video Producers
K.R.Ravindranath
Kannan Krishnamurty

IIT Madras Production

Funded By
Department of Higher Education
Ministry of Human Resource Development
Government of India

www.nptel.ac.in