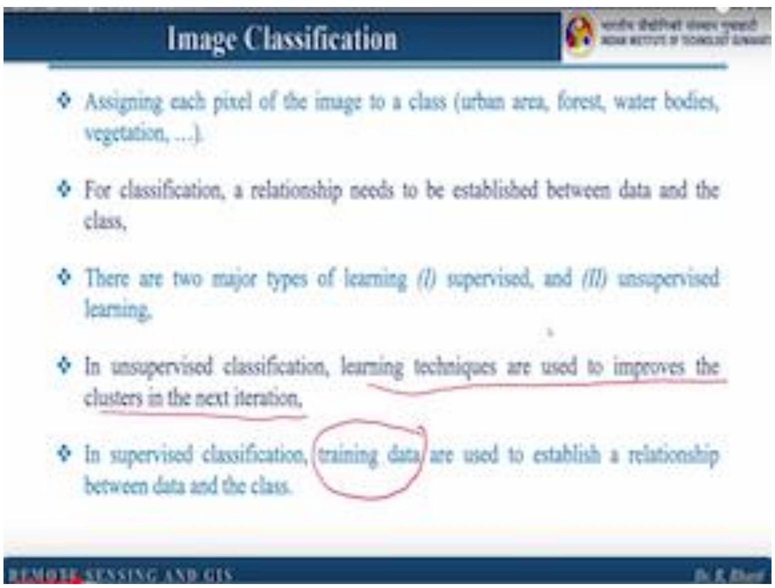
Remote Sensing and GIS
Rishikesh Bharti
Department of Civil Engineering
Indian Institute of Technology - Guwahati

Lecture – 12
Image Classification - 2

This is the second lecture on image classification. So today, we will continue the topics which I left in the previous lecture. So let us just briefly go through what exactly we mean by image classification and what are the different types of image classifiers available to classify a remotely sensed image.

(Refer Slide Time: 00:54)



So in image classification, what exactly we do, we assign each pixel of the image to a class right.. And class means, our land use or land cover. So suppose you have an image and you have pixels right., and these pixels are nothing but the digital numbers. So only for visual purpose, we see colours represented by these numbers right.. So here these classes are urban area, forest, water bodies or vegetation.

So exactly, we want to locate which pixel belong to urban area or which pixel belong to forest, which pixel belong to water body, and which pixel belongs to vegetation. So here we need to establish the relationship between the digital number and the classes, and that relationship will be used to classify or to assign each pixel of this image to a class. Right so this is what in the second point I have highlighted.

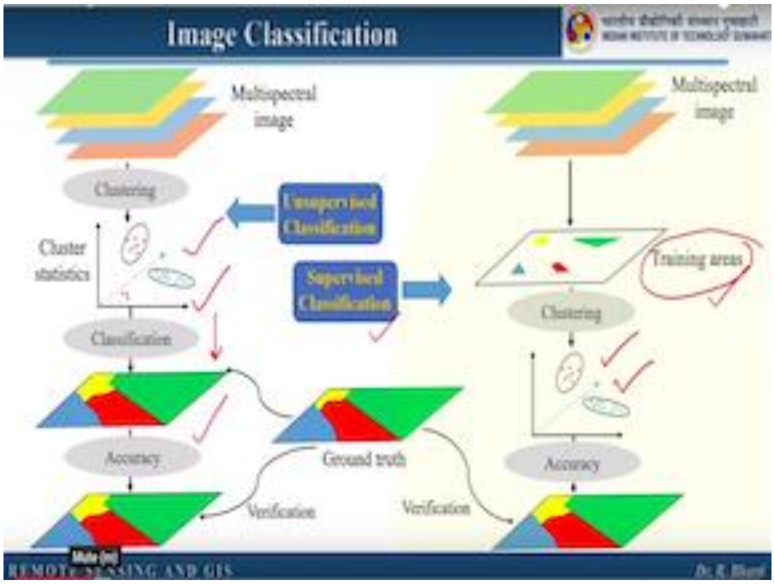
So you can see for classification, a relationship needs to be established between data and the class right. And there are two major types of learning, first one is supervised and second one is unsupervised. I hope you remember my previous lecture where, we have discussed this supervised and unsupervised classification techniques.

And I have given you some examples, in case of supervised classification, I am teaching you how to solve a numerical right., and then slowly you will learn that and you will be mastered in that. So you will reduce the time required in solving that numerical with your learning right. In unsupervised classification technique, we use some logic through which image DN numbers will be used to make clusters right. So there are two major types of learning technique, first one is supervised, and second one is unsupervised.

I hope you remember the examples of my previous lecture, like cooking, driving, solving a mathematical problem. So these are problems which can be learnt either in the supervised manner or in unsupervised manner right. So if you use supervise, that means you will have an instructor. And in case of image classification an analyst is required to provide the dataset. dataset means here we are referring training data right.

So in unsupervised classification learning techniques are used to improve the clusters in next iteration right. And in supervised classification, basically we provide training data and that are used to establish the relationship between data and the class. And subsequently we will classify the whole input images into several classes right.

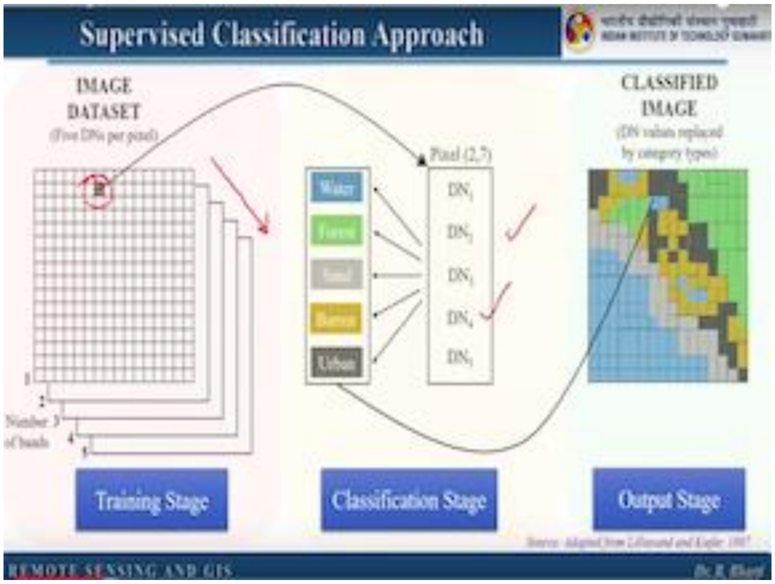
(Refer Slide Time: 04:35)



This already, I have shown you in the previous lecture where, I have highlighted the unsupervised and supervised classification technique. In unsupervised classification technique, basically we start with clustering algorithm and we cluster or we find the clusters based on the statistical similarities. And then subsequently we use either ISODATA or K-Means Algorithms and then we classify this image in two different classes right.

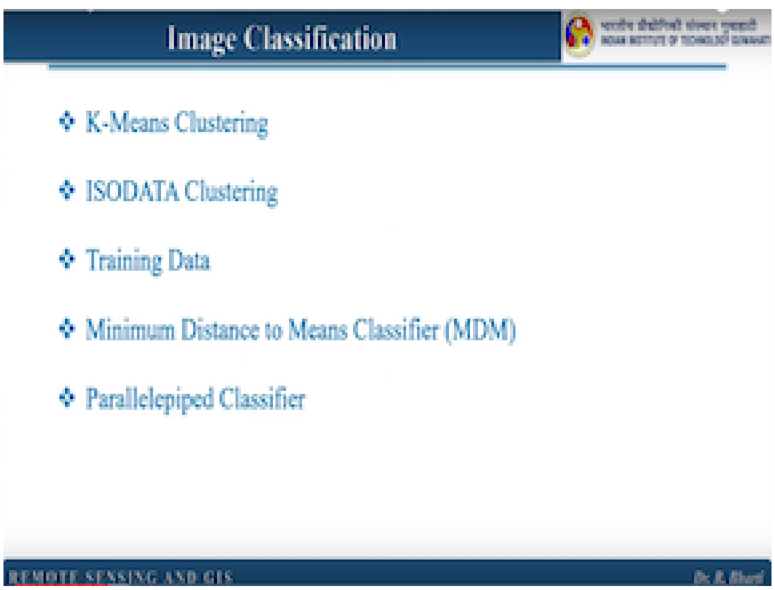
In supervised classification technique, what we do, we provide training dataset for the classification and based on this training data, we form clusters right. And those clusters will be used for this classification. So here in unsupervised classification, basically clusters are formed based on this statistical similarity. In supervised classification, these clusters are formed based on the provided data and that we are referring as training data and in both classification technique we get classified image.

(Refer Slide Time: 05:57)



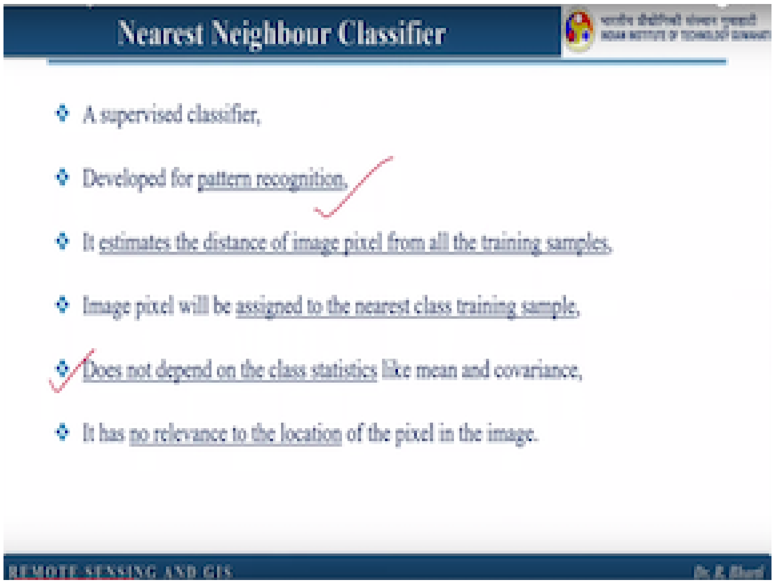
Here in this one, this slide is used to highlight the image dataset where we have 4 or 5 bands and then for a particular pixel we will have 5 values right. And the same time we will have training data from the field and these DN numbers and that information which we have collected in the field, we will develop a relationship and that will be used to produce a classified image right. So this is what we do in unsupervised and supervised classification technique.

(Refer Slide Time: 06:42)



In my last lecture I have covered K-Means clustering then second one was ISODATA clustering. Then in supervised classification, I started with how do we select or how do we provide training data in supervised classification. So that we have already discussed, then in supervised classification, I have started with minimum distance to mean classifier. Then subsequently we discussed Parallelepiped classifier, now in this lecture I will start with Nearest Neighbour Classifier.

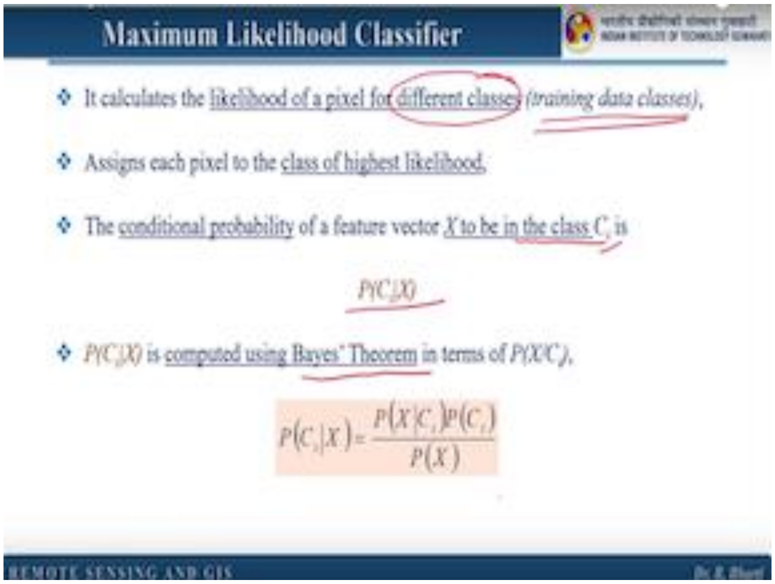
(Refer Slide Time: 07:21)



So let us see what exactly Nearest Neighbour classifier does right. So this nearest neighbour classifier is a supervised classifier and we know if we are working with a supervised classifier then we need to provide the training data. So I am not going to repeat or we are not going to discuss here what exactly these training datasets are. I assume that you have understood in the previous lecture.

So now here this nearest neighbour classifier, this was developed for pattern recognition right. where it estimates the distance of image Pixel from all the training samples. The fourth point is image pixel will be assigned to the nearest class training sample and then it is very important to note here that it does not depend on the class statistics like mean and co-variance and it has no relevance to the location of the pixel in the image.

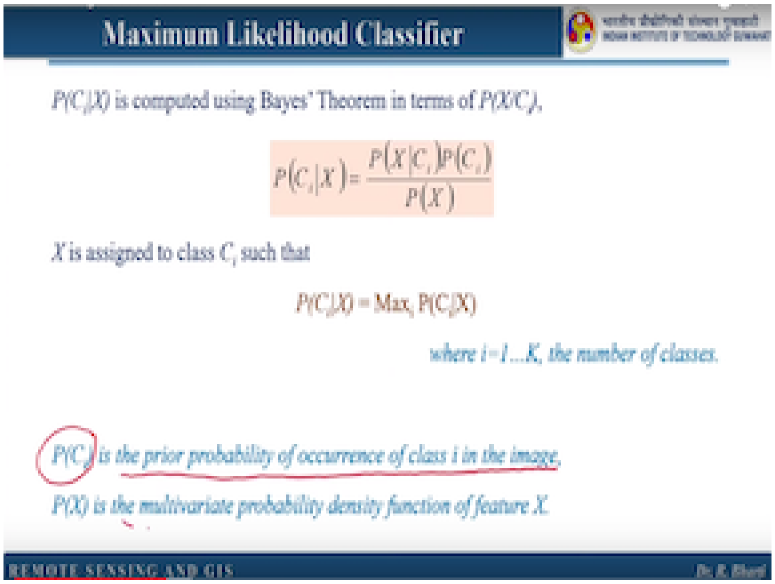
(Refer Slide Time: 08:38)



Then the next classifier is maximum likelihood classifier. Here we calculate the likelihood of a pixel for different classes and here different classes means training data classes. So if you have given 4 or 5 classes, then you have to have probability of assigning this pixel into all those 4 to 5 classes. In this method, we assign each pixel to the class of highest likelihood right. So here we have assumed that we have provided 4 or 5 training data right.

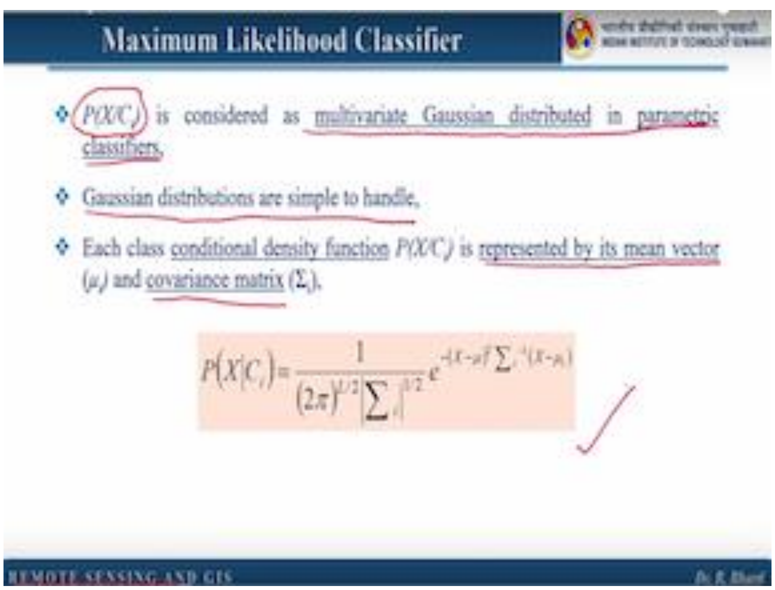
And the highest probability class will be used to assign or identify the class of unknown pixel, unknown pixel means I am referring the image pixel, which we want to classify it right. And the conditional probability of a feature vector X to be in the class of Ci is this one right. So this is the normal expression in general we use this. So here we use this Bayes Theorem right. For identifying the probability of X to be in the class of Ci right. And this is the mathematical expression for Bayes theorem.

(Refer Slide Time: 10:08)



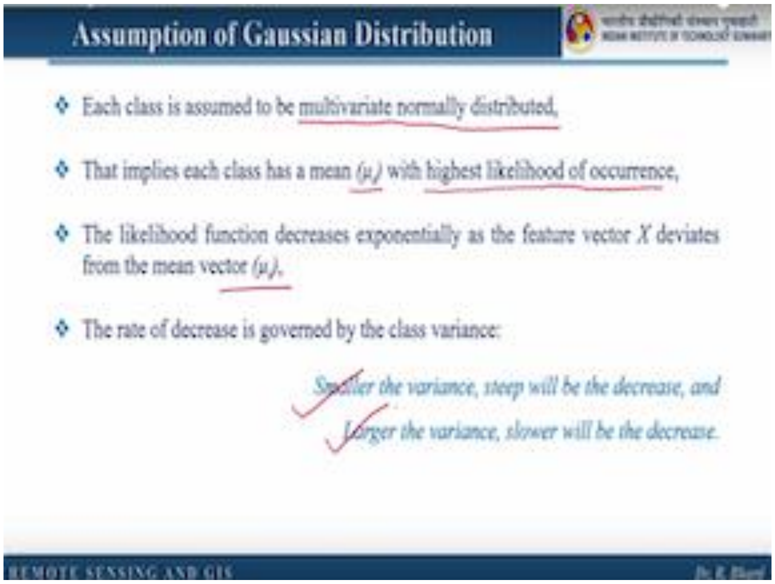
Now the probability of X pixel to be classified in Ci class can be calculated using this formula right. Where X is assigned to class Ci such that, that maximum probability we will get for Ci class, where I can be 1 to K, the number of classes. So here if you have 4 or 5 classes that this will go to 4 or 5 right. Here, probability of Ci class is the probability of occurrence of the class I in the image. And here this is the multivariate probability density function of feature X right.

(Refer Slide Time: 10:57)



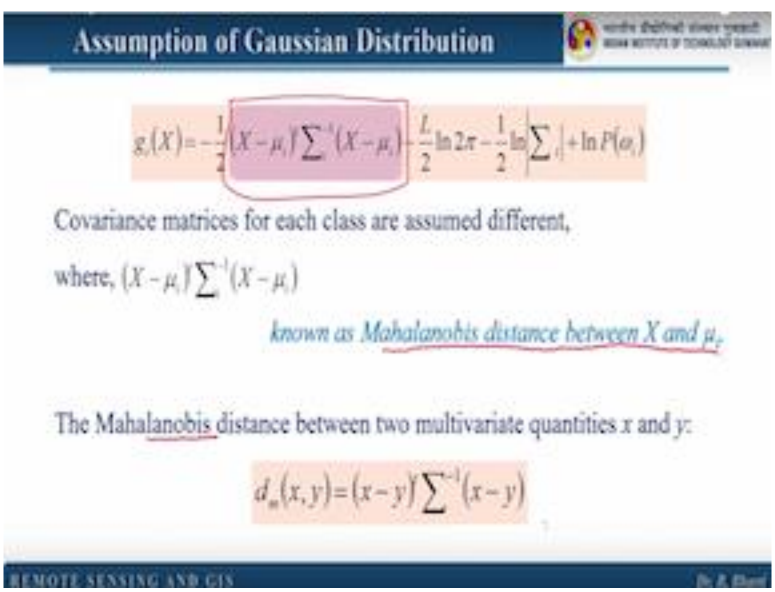
Probability of X pixel to be classified in Ci class is considered as Multivariate Gaussian distributed in parametric classifier right. And this Gaussian distributions are very simple to handle. But when you see the real examples or real data, it is not always we get Gaussian distributed sample. So in that case it will be a little bit difficult to resolve that class with this maximum likelihood classifier. So each class conditional density function is represented by mean vector and covariance Matrix right and this mathematical expression can be used for that.

(Refer Slide Time: 11:50)



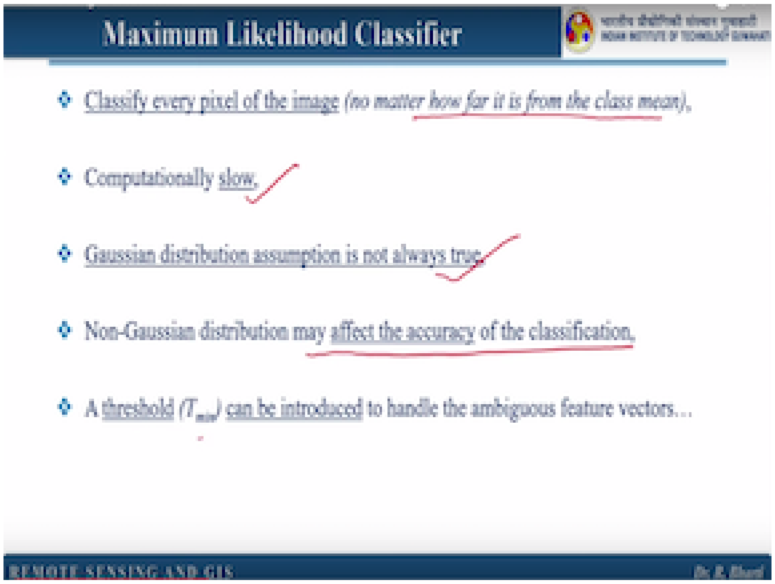
Now, what are the assumptions we take into account in Gaussian distribution. So each class is assumed to be multivariate normally distributed right. That implies each class has a mean right with highest likelihood of occurrence. The likelihood function decreases exponentially as the feature vector X deviates from the mean vectors and which is this one right. The rate of decrease is governed by the class variance where smaller the variance steep will be the decrease and larger the variance slower will be the decrease right.

(Refer Slide Time: 12:42)



So the covariance matrix for each class are assumed different rate, where you see this particular part of this equation is basically the Mahalanobis distance between X and Mu i. So here this is the mean vector right. So if you want to calculate the distance using this Mahalanobis, then you can use this particular function right.

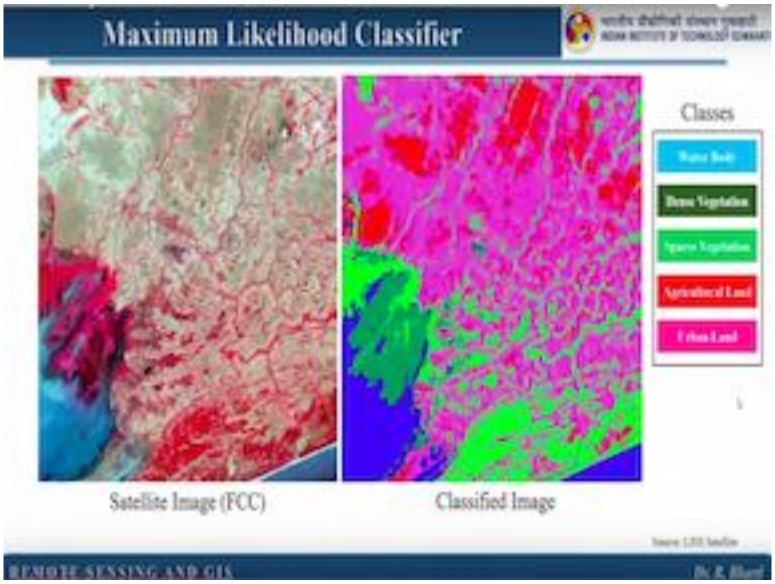
(Refer Slide Time: 13:18)



So in maximum likelihood classifier, each pixels are classified no matter how far they are from the mean class right and this is computationally slow compared to other techniques and Gaussian distribution assumption is not always true. That I also indicated in the previous slide because in nature it is very difficult to find this Gaussian distributed sample right. Non-Gaussian distribution may affect the accuracy of the classification.

A threshold, T minimum can be introduced to handle the ambiguous feature vector. So here basically a threshold can be introduced to resolve the ambiguity of the feature vector right.

(Refer Slide Time: 14:10)

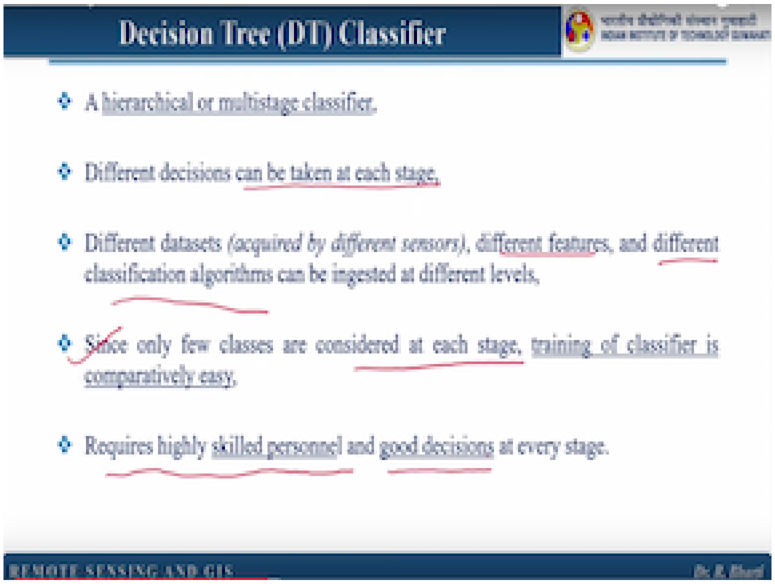


This is the example and you can see the result of maximum likelihood classifier in this particular image where we have given 1,2,3,4,5 5 classes right. So you can increase this or you can decrease this based on your application and requirement. So this is the result you get if you are using maximum likelihood classifier and there are certain limitation with each of these classifiers that you need to understand. So as I told you that here, no matter how far the pixels are away from the mean vector, they will be classified.

So if you introduce a threshold, then there will be several pixels here which will be unclassified. So that will be better when we do not want to resolve or you do not want to classify the whole area into the available classes. So here, suppose if you want to identify or classify these image into only two classes right. Water and the forest right, so you can take these samples from these two and you can run this maximum likelihood classifier with threshold.

So what will happen, you will find these two classes available in your classified image and rest of the pixels will be unclassified. So you need to understand all these methods very carefully and accordingly you need to modify this as per your application.

(Refer Slide Time: 15:59)



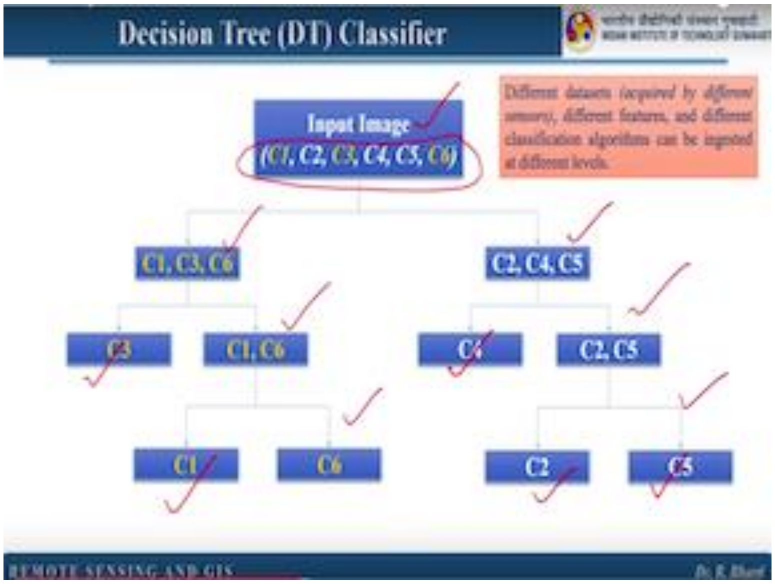
The next classification technique is decision tree classifier. This is a hierarchical or multistage classifier. So multistage means here you can have multi stages or multi levels where you can change your decision and the input data. Because in the previous cases, in the previous methods we have seen, we have classified the image using only one method where the data is fixed. method is fixed right.

But here in decision tree classification you have all the flexibilities. So here the sources of error will be more. So you need to be very careful while using this decision tree classifier. So decision can be taken at each stage right. And different datasets can be ingested in each of these stages and different features, different classification algorithm can also be changed or ingested time to time right. Since only few classes are considered at each stage, training of this classifier is comparatively easy.

So here this is a mechanism or this is how we are going to apply the classification technique or we want to identify a particular class type using this decision tree classification where you have all the flexibility to use unsupervised, supervised classification at different levels right. But as I told you, that errors will be more here because you will have several stages and each stage require some decision right.

So if you are not very careful, then you will have more error in your classified image. So that is why you need to have a skilled personnel and a very good decision maker at every stage right.

(Refer Slide Time: 18:12)

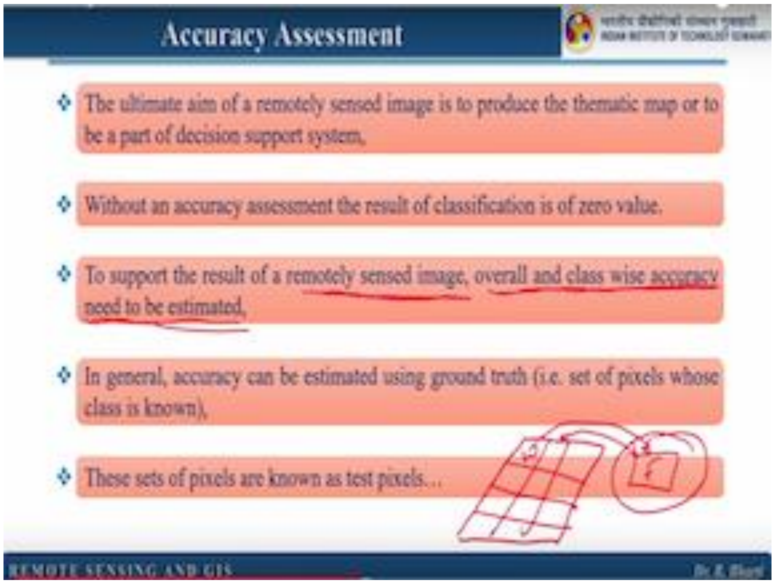


So let us take one example where we have a input image, and I want to classify this image in to C1,C2,C3,C4,C5 and C6 classes right. Now what I can do is I can segregate based on this particular image right, I can identify only C1,C3,C6 in one class right, and C2,C4,C5 in another class. So here you will have only two levels of classification. So C1,C3,C6 and another set is C2,C4,C5 right. Now, here what you can do, you can provide more input data right. So that you will be able to segregate all this classes separately.

Now you can see further, you have segregated C3 from C1 and C6 here also, C1 and C6 are segregated to further subdivide this C1,C6 class in separate group right. So now we have C1,C6,C3 in separate output. Now similarly you can ingest here some dataset are different logic or different classification technique. And subsequently you will segregate all the classes here right. So ultimately we have all these classes separate right.

So in this is not tree classification, what exactly we are doing, here we are ingesting, different satellite images or airborne images to segregate all these classes of an area right and here at different levels you can change the input data, you can change your decision or you can change your classification algorithm. So this is the beauty of this decision tree classification. Here you have all the freedoms, but you will have more errors right.

(Refer Slide Time: 20:36)



So errors can be introduced at any stage and here you can use this equation to calculate the probability of error. So error committed at any stage propagate downwards. So here for example, for any 6 stage decision tree classification, if the probability of error is 8% the overall classification error will be approximately 40%. So you can see how serious this 8% could be right. So you have to be very careful while using this. Decision tree classifier.

Now let us understand why exactly we use this classifiers because we want to identify some land use land cover classes using this satellite images right. But ultimately these products will be used in some decision support system or it can be an input for your model right. So you need to be very careful about the accuracy of all these outputs. So the ultimate aim of a remotely sensed image is to produce the thematic map or to be a part of decision support system without an accuracy assessment of the result of classification is of 0 value.

Because if I do something like whatever I have shown you the classified image, but I did not mention what was the classification accuracy. So in that case what happens, that classification result was of no use because nobody will accept without an accuracy right. So here you need to support the results of remotely sensed image, overall and class wise accuracy need to be estimated.

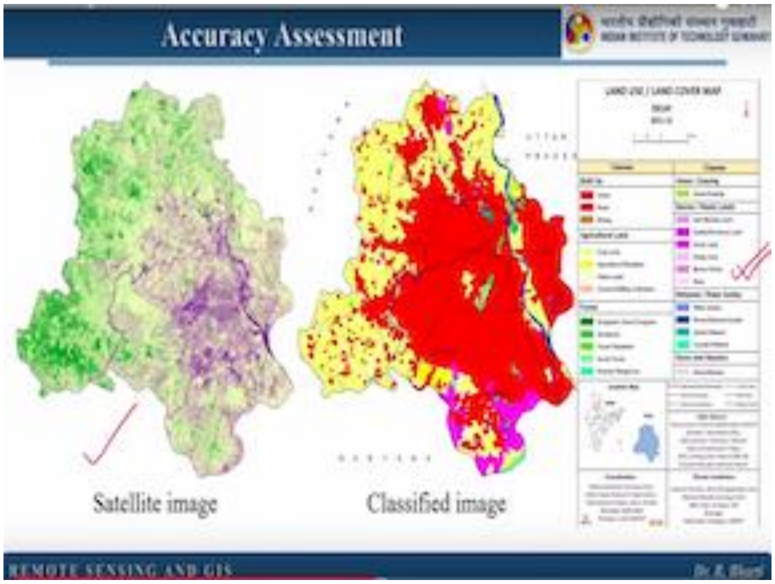
So for example, if you have classified your input images into 4 or 5 classes like vegetation, urban area, water or shrubs, grassland, dense forest, so some classes you have taken and you have classified the whole image into all these classes. But what if I say the classification accuracy of my image is 20%, 30% or maybe 80% right. Does it mean that all the classes, whatever we have considered in the classification, all the classes have this similar accuracy?. No, so it cannot be.

So you need to identify or you need to estimate the class-wise accuracy as well as the overall accuracy of your classified images. So that's why this third point I have highlighted to support the result of remotely sensed images overall and class wise accuracy need to be estimated right. In general accuracy can be estimated using ground truth because ultimately we want to compare with the original scenario.

So suppose you have an image right, and this is the classified, let us assume and you have found that this pixel, it belongs to water. But when you visited the area and you found that the latitude longitude of this particular pixel is basically forest, this is not water in the field. So what we found, we found there is a mismatch of the classes. So I have reported this pixel as water, but in field it is forest. So that means there is a mismatch.

So I need to calculate the accuracy with respect to my ground information right. So the set of pixels, whose classes are known, are known as test pixel right. So this is similar as your other ground truths.

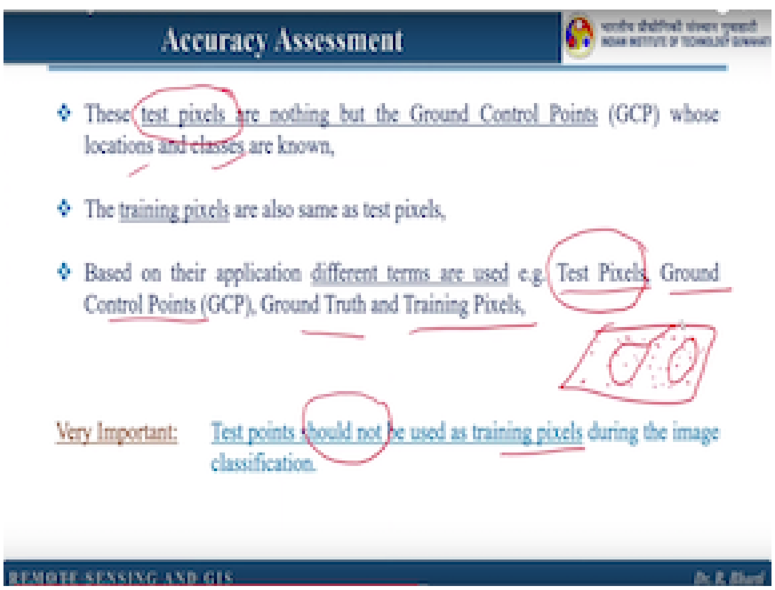
(Refer Slide Time: 25:01)



Let us understand this in detail. So as you assuming this is the satellite image of an area and this is the classified image of this particular area right and here all these colours indicate some classes and that classes can be seen here right. But exactly whether they are matching or mismatching with the real world that you need to establish. So what do you have to do? You have to visit field and then you have to collect the location values like latitude and longitude using a GPS.

And then that you will plot here right, and these plotted information will be used to compare the classified image class and the field class right. So that is known as ground truth.

(Refer Slide Time: 26:00)



So here we will understand it, so these test pixels are nothing but the ground control point, which we have plotted on the classified image collected from the field. So here, what are the things we know? We know the location and the class right, so you had a classified image and here suppose let us say it has three classes. This is water, forest and urban area, when you collected the ground control point, which is known as test pixel in accuracy assessment.

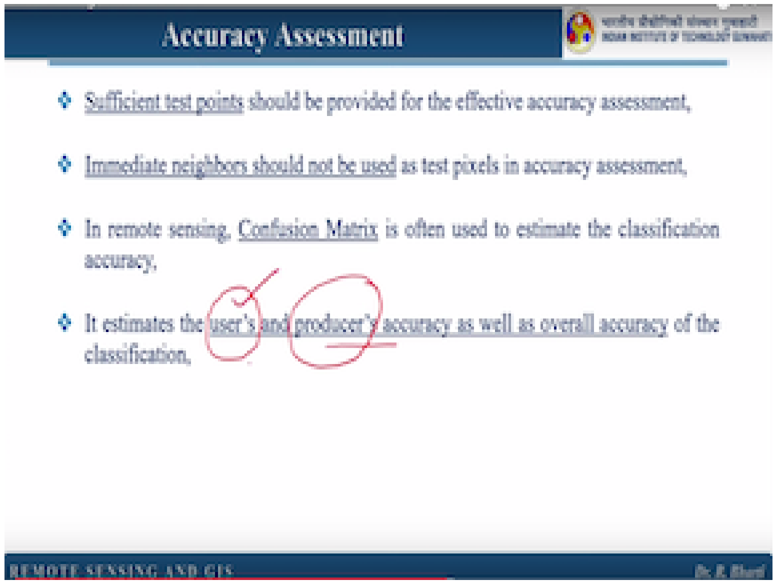
So here the test pixel, I will use this triangle right, and here location and classes are known. So if I plot the collected information from the field right, and let us assume this pixel in the classified image it belongs to water but in field it was maybe let us say urban area, the next boundary. Similarly if you consider this pixel here in the classified image it is forest, but in the field it is water. So like that you need to find out the corresponding class of the ground control points in the classified image and in the field.

The training pixels are also same as test pixel. I hope you remember this training pixel which we have used in classification there also we have provided the location and the known classes. Here we are calling it test pixel. So both are same only thing is based on the application, different terms are used like test this pixel, ground control point, ground truth and training pixel. So here in this case in accuracy assessment, we use test pixel right.

It is very important to note test pixel should not be used as training pixel during the image classification. Otherwise what will happen, your accuracy assessment will be 100%. So accuracy of your produced result will be 100% so that cannot be true right. So here, suppose you have visited this particular area and you have collected maybe 1000 ground control point or maybe 100 ground control point.

So out of 100 you have to use some in your classification, some in your accuracy assessment right. So that should be the strategy you have to formulate beforehand.

(Refer Slide Time: 29:04)



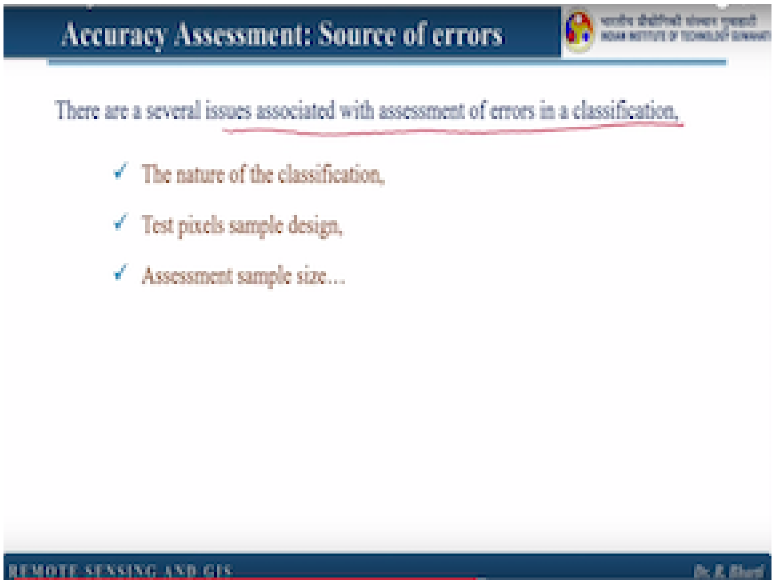
It is very important to know that sufficient test points should be provided for the effective accuracy assessment. Otherwise, if you provide less number of test points, then your accuracy will be more. And that cannot be true right, so you need to follow this. The next rule is immediate neighbour should not be used as test pixel in accuracy assessment.

That was immediate neighbour means, suppose this was your image and you have provided this ground control point in your classification right and this another pixel or another ground control point you have used in error estimation. So if they are just next to each other, then you will end up in again introducing errors. So it is better not to use the immediate neighbours as test pixel.

In remote sensing, Confusion Matrix is often used to estimate the classification accuracy. It estimates the user and producers accuracy as well as overall accuracy of the classification. So here you have these two terms like users accuracy and producers accuracy. So in classified image, if I am the analyst and if I am reporting a accuracy, that will be the accuracy of producers right. But if you are the user, then you will evaluate that accuracy in different term right.

So then your accuracy will be something different from the produced accuracy. So the users and producers accuracies are different from each other and depending upon which side you are, you will follow the accuracy. ow what are the different types of error? We introduce or which can be introduced in classification and in error estimation.

(Refer Slide Time: 31:26)

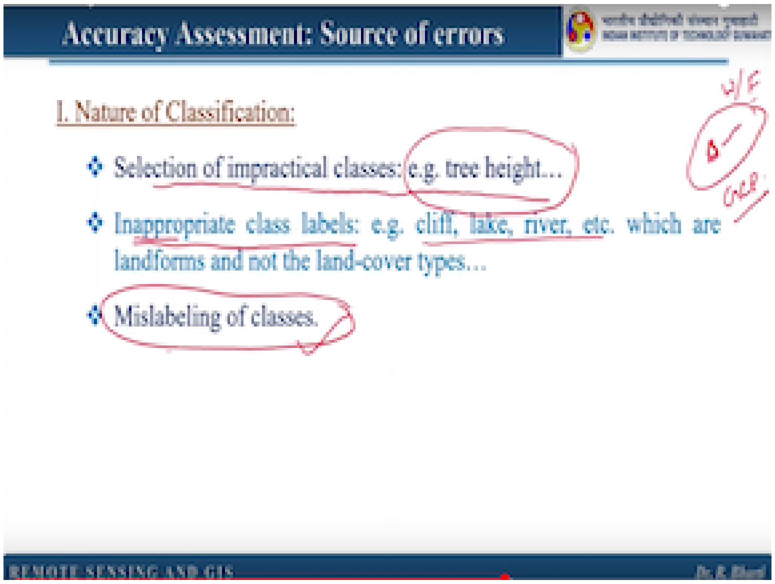


So here let us see in accuracy assessment, there are several issues associated with assessment of errors in classification. So first one is the nature of the classification, second one is test pixel design right, and the third one is assessment sample size. So here the first one indicates the nature of the classification. So which algorithm you are using, depending on that you will have accuracy right.

The next is test pixel sample design. So how do you design your test pixel? Suppose if you use immediate neighbours that will introduce more error right. If you distribute this test pixel in your image, then your accuracy will be more right. Then third one is assessment, sample size. How did you plan your sampling right? So how did you decide the number of pixels or number of GCPs required in classification required in assessment.

So that is also going to matter here, so let us see how these source of error are introduced in classification as well as in the assessment. So there are several issues associated with assessment of errors in classification. First one is the nature of the classification, second one is test pixel sample design, then third one is assessment sample size. So let us see one by one how these are different from each other and what exactly they mean right.

(Refer Slide Time: 31:26)

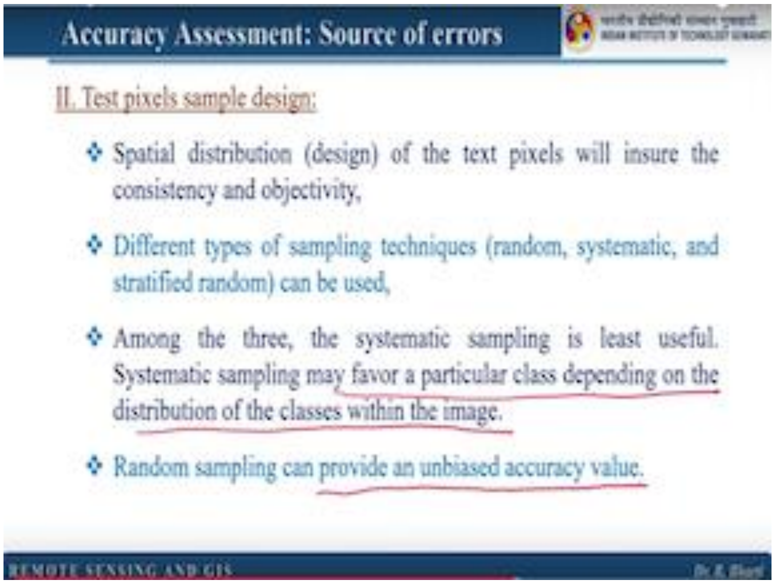


So let us see nature of classification. So here selection of impractical classes. For example, tree height. So if you have used ground control point, whose location is known. So latitude, longitude is known, but the information the class information is basically tree height to let us say a 100 meter tall tree or 200 meter tall tree. So this is not going to help you in the classification. So you will introduce more error right.

In appropriate class labels like cliff, lake, river, etcetera, which are landform and not the land cover type. So this is very important to label the class appropriately because you will be using the same nomenclature in the whole classification and subsequently you will be able to compare with the earlier results. In that case, in an appropriate class labelling, will introduce error in your classification.

Then mislabelling of the classes, so suppose you have collected this particular ground control point right. And in the field you have noticed that this particular pixel or the location, latitude, longitude belongs to water right. But by mistake you have named it or you have written forest, then what will happen? This error will get propagated in your classification. So this is the meaning of mislabelling of classes right.

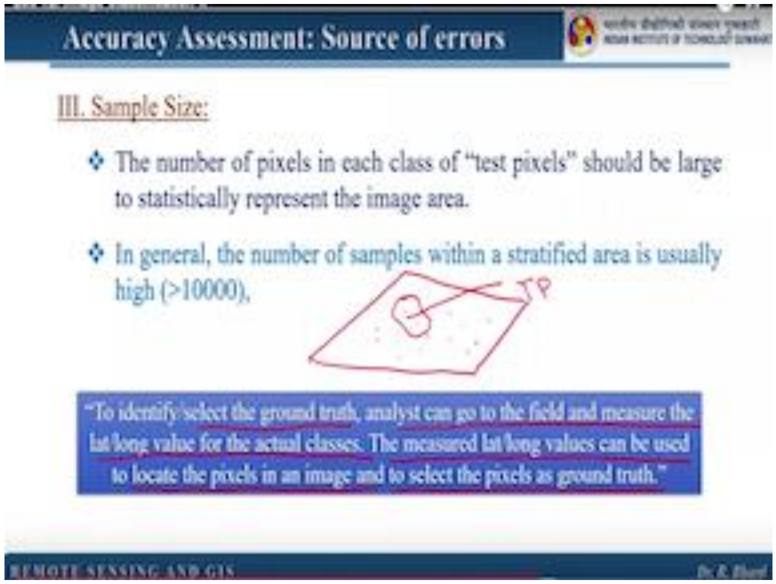
(Refer Slide Time: 35:10)



Now the next is test pixel sample design. So the spatial distribution of the test pixel will ensure the consistency and objectivity right. Different types of sampling technique like random systematic stratified random can be used here. Among the three, the systematics, a systematic sampling is least useful and systematic sampling may favour a particular class depending on the distribution of the classes within the image.

And random sampling can provide an unbiased accuracy value. So it is very important, how do you collect, how do you plan to collect the ground control point and that subsequently you will use as test pixel and the training pixel, right. So you have to be very, very careful. How do you plan to here the sampling technique, we have a random systematic stratified random, so depending on your application and the accessibility of that area, you need to decide which sampling technique you are going to use?

(Refer Slide Time: 36:34)

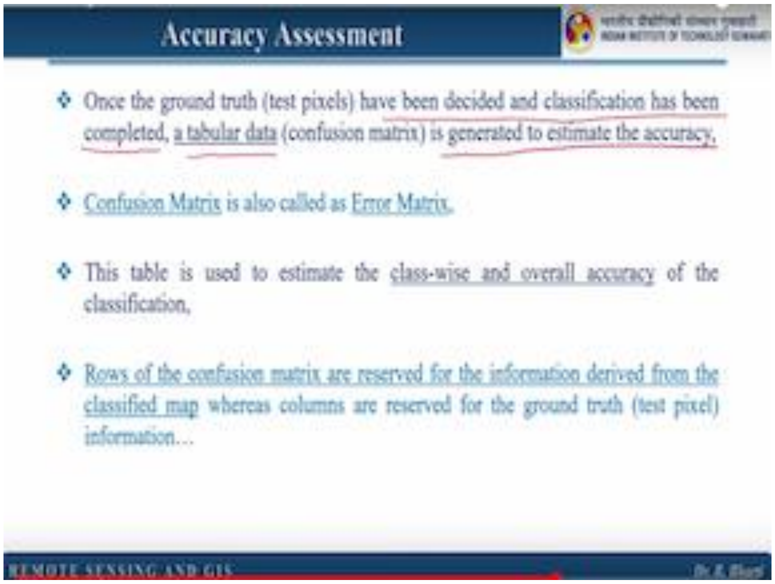


Then the next point is sample size, the number of pixel in each class of test pixels should be large to statistically represent the image area right. In general, the number of samples within a stratified area is usually high, so to identify, select the ground truth analyst can go to the field and measure the latitude, longitude value for the actual classes in different. Remember the measured latitude, longitude values can be used to locate the pixel in an image and to select the pixel as ground-truth.

Because in image we have marked earlier for the training Pixel right. But here now we are marking these pixel to evaluate the accuracy of my classification. So it is very important to understand all these information. Does not matter whether you are talking about test pixel or training pixel, both have been collected in the field using a GPS and that has been plotted on the image and the selected pixel will be used as the rise training pixel or the Test Pixel right.

So everything we are doing on the image right, now assuming you have collected the field information.

(Refer Slide Time: 38:08)



So let us start the accuracy assessment, so once the ground truth have been decided and classification has been completed, a tabular data which is known as confusion matrix is generated to estimate the accuracy right. So confusion matrix is also known as error matric. So sometimes we will find people have written error matrix. Sometimes people have used confusion matrix. This table is used to estimate the class wise and overall accuracy of the classification.

The rules of the confusing metrics are reserved for the information derived from the classified map, whereas columns are reserved for the ground truth right. So here it is important to understand this is the classified image right. Now you have plotted the ground control point which you are using as test pixel, you have plotted here and using this information you have formulated or you have generated this confusion matrix.

(Refer Slide Time: 39:30)

Confusion Matrix

		Ground Truth (Test Pixels)						
		Class-I	Class-II	Class-III	Class-IV	Class-V	Total no. of classified pixels	User's Accuracy
Map (Classified Map)	Class-I	7	1	4	1	1	14	
	Class-II	0	8	0	0	0	8	
	Class-III	1	0	25	1	1	28	
	Class-IV	0	0	1	12	1	14	
	Class-V	0	0	1	1	8	10	
	Total No. of ground truth pixels	8	9	31	15	11	74 Total no. of pixels used for accuracy assessment	
	Producer's Accuracy							

REMOTE SENSING AND GIS — Dr. R. Bharti

In confusion matrix rules are fixed for map information, map information means classified map right. And columns are fixed for ground truth and which is basically test pixels right. So if you see this table has been prepared, now let us assume, that 74 pixels were used. So you have collected the 74 ground control points, which you have used as test pixel. You have plotted on the image and then you have calculated or you have estimated what was the real class name in the field and what is the classified class name right.

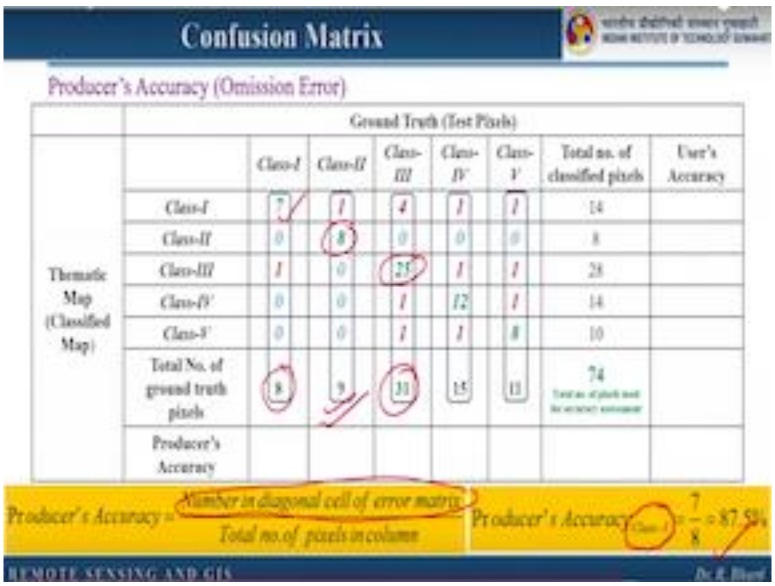
So here you will find that 7 pixels, which were class 1 in the field and there also class 1 in the classified image right. So if you see the diagonal elements so they are the correctly classified image pixels right. So 7 pixels, which were class 1 and in the classified images they are class 1. In class 2 if you see 8 pixels are correctly classified as class 2 in the image. Here class 3 25, class 4 12 and class 5 8.

So basically what exactly it means out of 74 these are the pixels which have been correctly classified by your classifier rest are wrong right. So if you see individual class wise, so in the classified image right 7 pixels are correctly classified as class 1, in class 2 8 pixels, in class 3 25, in class 4 12, in class 5 8 right. And if you see the total number of pixels in ground truth for class 1 was 8 right. And out of 8 only 7 are correctly classified and one is coming into another class.

So that is the problem with the classification. Now how will we evaluate? So we will calculate the producers accuracy. Now here you need to remember this, that this row is fixed for classified map information and these columns are fixed for ground truth information right. So if you see the last column of this confusion matrix that is fixed for users accuracy, and the last row of this table is basically for producers accuracy.

So now if you have any intention to change this matrix, then what you have to do, you have to reorient this maintaining this structure right. So let us calculate producers accuracy.

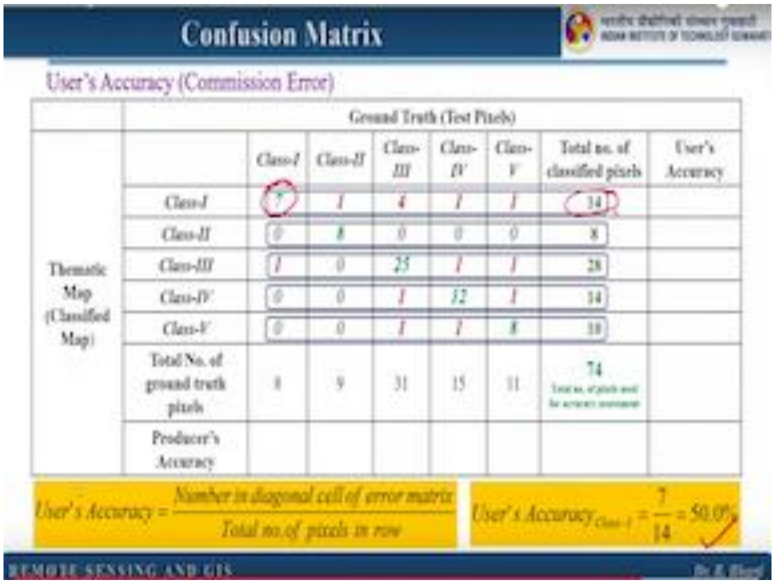
(Refer Slide Time: 42:52)



Now here, if you see the number of diagonals sale in error matrix, they are the pixels correctly classified by your algorithm. So here what is happening out of 8 only 7 are correctly classified. So if you calculate that, you will get the producer accuracy of class 1 right. If you calculate class 2 producers accuracy, then you have to take the diagonal element. So that will be 8/9 because that is the total pixel.

In class 3 it will be 25/31, in class 4 that will be 12/15, in class 5 it will be 8/11 right. So that will give you class-wise, producers accuracy.

(Refer Slide Time: 43:52)



Now let us see what we mean by users accuracy. Now here, if you see total number of classified pixels in your image in class 1 is 14 now out of 14 only 7 are correct as our ground information. So what we will do, we will take the ratio, and here we will calculate 7/14 right. That will be the user accuracy of class 1, so if you remember the last slide, let us note that value right.

So in the producer accuracy, class 1 producer accuracy was 87.5% in users accuracy, class 1 accuracy value is 50% now you can compare these two. So for a producer, it will be always more, not always but in general it is more right. Because I am producing a result, so I will say my results are more accurate. But for a user it may be less, so here basically we want to highlight this committed error, in producer accuracies only we highlight the omitted error.

To understand this commission error, let us understand only 8 pixels, which we have collected in the field for class 1 right. But in classified image, we have total 14 pixels right. And these pixels are basically when we compare with our ground control point, 1 belongs to class 2, 4 belongs to this 3, 1 to this 4 and 1 to this 5. So basically here you have committed more errors right, in classification.

So that is why the user accuracy will be reduced compared to producer accuracy. So this is why you have 50% for class I. But in case of producer accuracy, it was more right.

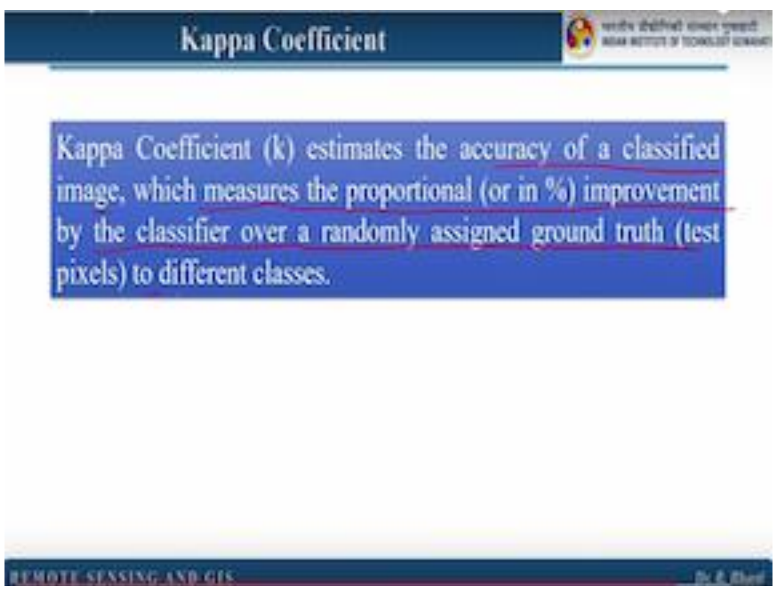
(Refer Slide Time: 46:16)

		Ground Truth (Test Pixels)						
		Class-I	Class-II	Class-III	Class-IV	Class-V	Total no. of classified pixels	User's Accuracy
Thematic Map (Classified Map)	Class-I	7	1	4	1	1	14	
	Class-II	0	8	0	0	0	8	
	Class-III	1	0	25	1	1	28	
	Class-IV	0	0	1	12	1	14	
	Class-V	0	0	1	1	8	10	
	Total No. of ground truth pixels	8	9	31	15	11	74	
	Producer's Accuracy							

$$\text{Overall Accuracy} = \frac{\text{Total no. of correctly classified pixels}}{\text{Total no. of pixels used for accuracy assessment}}$$

Now how do we evaluate the overall accuracy of this classification?. In overall accuracy, what we will do, we will calculate the total number of correctly classified pixels by adding these right., and that will be divided by the total number of pixels used in this accuracy assessment. So if you see this will be 60/74 right. So now these users, producers, overall accuracy, omission, error, commission error, everything I have explained.

(Refer Slide Time: 47:10)



Now let us understand there is a new way or there is a different way of calculating the accuracy of our classification that is known as Kappa Coefficient. So Kappa Coefficient estimates the accuracy of a classified image, which measures the proportional improvement by the classifier over a randomly assigned ground truth, which is basically test pixel to different classes right.

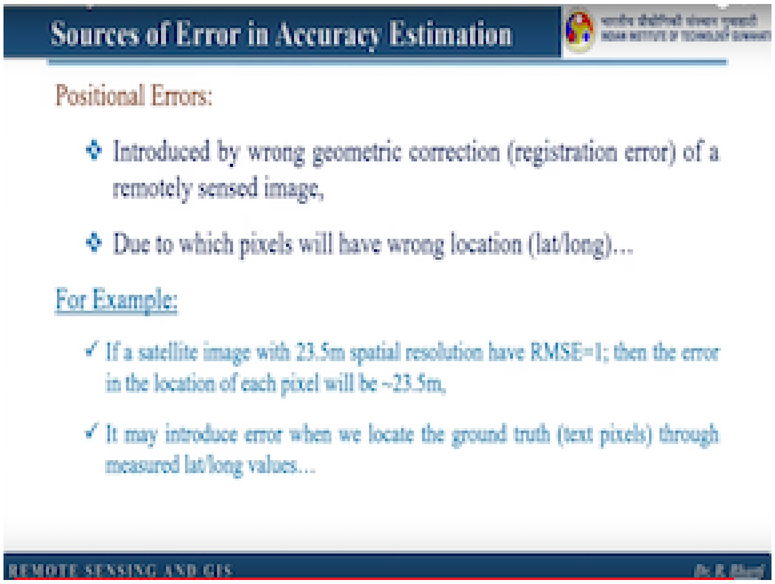
(Refer Slide Time: 47:34)

Now let us see, how do we calculate Kappa Coefficient? So the same table we have to generate does not matter whether you are going for users accuracy, producers accuracy, omission error, commission error, or the overall accuracy and Kappa coefficient. So here only thing is the calculation will be different.

So for Kappa coefficient we use this expression right, and where N is the total number of pixels used for accuracy assessment. So that is 74 right, and A is the total number of correctly classified pixel. That is this diagonal pixel right and that was 16 right. And B is basically the multiplied or the product of this total number of ground truth for that class and total number of classified pixel for that class.

So here if you see 8 into 14+9 into 8+31 into 28 then 15 into 14+11 into 10 that will be your B right. Now once you have that, you can easily calculate this Kappa coefficient. So this is the way we calculate all these accuracies for classified image does not matter whether you have used unsupervised classification or supervised classification right. You have to produce or you have to report the accuracy of your classification. Let us understand what are the different sources of error in accuracy estimation. So basically this can be from your pre- processing right.

(Refer Slide Time: 49:37)



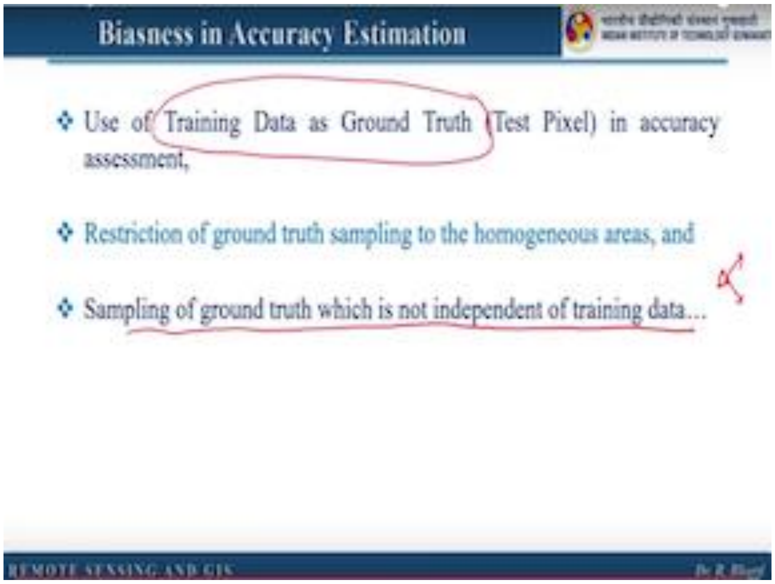
So let us understand this positional error. So how do we get this positional error in our classified images?. So this is basically introduced by a wrong geometric correction because if you remember, you have used some ground control points to relocate the pixels or to correct the geometric errors of your image right. So in case if you have wrongly given the information of your ground control point, then what will happen?.

The location of this Pixel, which was supposed to be here, but because of the wrong information, it has been shifted here to the next pixel right. So that may introduce error in your classification. So due to which pixel will have, wrong location right, latitude and longitude. For example, if a satellite image with 23.5 meters spatial resolution have RMSE 1, then the error in the location of each pixel will be approximately 23.5 meter.

So if you have like this pixel which was supposed to be here right, if this is the image, but because of the error it has been shifted. So this X value, which was supposed to be here, it got shifted here right. And which was supposed to be the position of Y, so if there is a change in the position, basically the size of 1 pixel is 23.5 meters. So these, this is the positional error which are associated with your classification.

So you have to be very careful while doing each and every step starting from correction to the analysis and you have to be very careful while introducing any ground control point in the classification, in the accuracy assessment or in the correction methods. Now there are few biasness in accuracy assessment, like use of training data as ground truth. in accuracy assessment.

(Refer Slide Time: 51:55)



In accuracy assessment, so if you have used the same data in your training data and the ground truth, then what will happen, you will have more accuracy because the same information is used for classification the same information has been used for accuracy so it will be 100% right Restriction of ground truth sampling to the homogeneous area, sampling of ground truth which is not independent of training data.

So here basically this is very important sampling of ground truth, which is not independent of training data. So this should be different so that is all for today so we will see more classification technique in my future lectures where we will discuss thermal remote sensing, hyperspectral remote sensing, microwave remote sensing. And I will introduce you to many different or many advanced classifiers that you can use in your application right. Thanks.