Indian Institute of Technology Madras
Presents

NPTEL
National Programme on Technology Enhanced Learning

Pattern Recognition

Module 06

Lecture 08

Examples of Uses or Application of Pattern Recognition; And When to do clustering

Prof. C. A. Murthy
Machine Intelligent Unit,
India Statistical Institute, Kolkata

I am just going to give just one or two examples on some of the uses and applications of pattern recognition this thing I should have told in the very first lecture which I did not include which it did not I observe that I did not say this thing so I am telling it now this is a example that I really like this example it is crossing the road crossing the road so suppose you and your friend you have gone out and there is a road in front of your IIT okay you would like to cross the road to the other side.

Now my question to you is it necessarily true that you will cross the road at the same time no though you are go you when they are at the same time probably on the other side one of you has some work other one is accompanying okay and so reliably you are trying to cross the road and one person may cross first another person may cross later now my question to you is for the same situation and exactly everything is same your output is different if you look at it from the point of view of you need to make a judgment.

You need to make a judgment or you need to make a decision everything is same but the judgments are different the decisions are different do you think one of you is right and the other one is wrong, no both of you are equally right this is an example where exactly for the same situation in every respect everything is same but you have two equally good answers you have two equally good answers.

Now let me tell you something regarding this when you are making the decision note that your information that is given to you one car is coming from on the other side you have looked at it maybe two or three cars do you know the speed of the car exactly no note that a person who never drove a car or who never drove a vehicle can cross the road a person who has driven a vehicle probably he has he or she has some idea about the speed of the vehicle but even if a person has not driven any vehicle whatsoever he or she can cross the road an illiterate person can cross.

The road a child of age 7 or 8 years can cross the road now you want to write a program for crossing the road how do you write it the input to this is not precise is the input precise no the input is not precise and in fact is your logic precise no even then you are able to cross the road this is a nice example to say that many times human beings have imp resized imperfect and imp resized imperfect ambiguous information and from there we make judgments and we go ahead one of the aims of pattern recognition is to model this to model this imprecision ambiguity and two and two make a system so that the system works.

Note that the decision I mean the output need not be unique if you need unique output that is fine but many times we do not need unique output this is one example that I wanted to tell you and there is another example means the chalk piece right many times in our trailer in my childhood at least what I try to do was I try to balance this in my hand I hope many of you have done it when you are a child and probably failed okay this problem of balancing any object like this there is a name for it this is inverted pendulum problem.

A pendulum you know there is a thread and then there is a small blob attached and towards the end of the thread now that is pendulum and then you invert it that means the thread portion instead of thread you put a stick it comes like this and the blob is up the stick is like this the blob is up that is inverted pendulum this you need to balance it you would like to write a program for balance singlet what is the meaning of writing a program if it is falling in this direction the speed at which it is falling and the angle at which it is falling and somehow you need to change your hand you need to probably bend the hand are it at the same speed or maybe slightly more speed you need to do all these things.

Writing a program for this you may ask me a question why balancing a stick is a problem why does one need to do it what is the use of this solving this problem the use I will tell you one or two uses from this probably you will be able to get many other uses suppose you would like to make an unmanned aircraft so that this aircraft would fly into enemy territory to find out what the enemies are doing enemy territory are the nearby country territory whatever it is you would like to send an unmanned aircraft.

Why unmanned aircraft if you send a human being inside the aircraft since it is done illegally you are not you are not taken any permission from the other place to fly into it they may as well shoot this aircraft so that the person may die and that will become a Scand also you would like to have an unmanned aircraft here so that even then someone shoots this one no problem only machine is getting wasted some money is going into the drain but human beings lives are not lost.

So you would like to have an unmanned aircraft to fly into some unknown territory so unmanned aircraft means how does it navigate again you need to have some sensors and then the sensory is going to get the input and some decisions are to be made note that many of these decisions which must be taken they are to be taken very fast because the object is moving the aircraft is moving it is not standing still aircraft is moving suddenly it sends something there okay.

Then immediately it needs to change the direction there is something it has seen immediately it needs to change a direction that is very similar to this is falling you need to change it again something is happening you need to change it have you understood what I wanted to say if you solve this problem then the similar logic you can apply to solve the problem of having an unmanned aircraft I mean how do you actually collect the collecting the information is one part analysis.

And interpretation and you need to have actually an online algorithm online algorithm for doing this thing. So that problem is very much similar to it this is about aircraft similarly you may want to have unmanned trains any unmanned trains are there in the world train is moving and then there is a signal it needs to stop and there is a bend there also the information how much bend is there.

You see you may not have the exact value similarly when you are taking a photograph from such a distance to the earth from low-flying aircraft you may not have the exact information there is ambiguity present, but you need to write your rules based on those ambiguities if you have something precise it is fine but you cannot guarantee that everything that you are able to sense is precise or I mean you cannot guarantee that and you need to work on these imprecisely.

So you need to in order to make all these systems you need to have the corresponding logic which deals with this ambiguities the it is very easy to say that you need to have logic which deals with this ambiguities that is very easy to say, but I do not think till now we really have the complete logical partially at some places yes we have it and note that the moment there is an ambiguity and the moment we made decisions on the base of this ambiguity sometimes our decisions can go wrong.

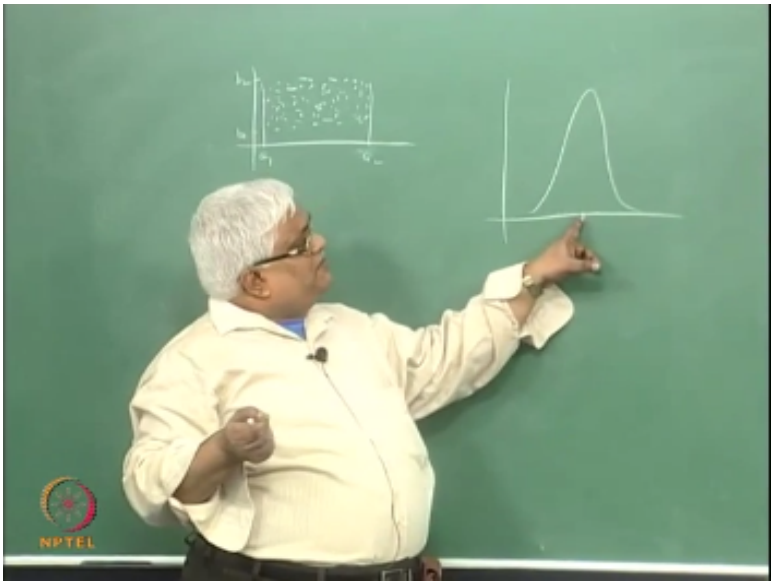
It is not necessarily true at every time the decisions are right, then there is a very basic question, the very basic question is if you want to ensure that no decision would be wrong then can you do it and if you cannot do it that means you are able to you are saying that some decisions can go wrong then how much percentage and once you decide the percentage do you have the corresponding logic automatically.

See these are all very big questions for which till now there are no answers till now there are no good answers and I mentioned in one of the first lectures, first few lectures one or two lectures I do not remember compared to the aim of what we want to achieve in pattern recognition what we have achieved is only Epsilon compared to what we wanted to achieve what we have at you till now is only epsilon and even that epsilon will be teaching epsilon. Here in that epsilon will be teaching epsilon here I think that is.

(Refer Slide Time: 13:46)

It given a data set how do we know that we need to do clustering, this question can be put in many ways.

(Refer Slide Time: 14:02)



Suppose your dataset is say like this then actually there are no clusters here, in fact suppose say this is say a minimum say the maximum is say b here and then such is a1 this is a2 and say here it is say b1 and then the highest is say b2 one can say it is uniformly distributed data x it is a1, a2 and why it is b1, b2 so in when you have uniform distribution you cannot have any clustering. So the first question is if you are given a data set before you want to apply any clustering algorithm what you need to first find out is whether the data has clusters.

One way of putting it is that check whether the data follows uniform distribution if the data follows uniform distribution then there are no clusters. How do you know that the data follows uniform distribution or not for this you need to refer to dupes and Janes book and algorithms for clustering data where they discussed all these tests whether the data follows uniform distribution or not how do you judge it those tests are given there.

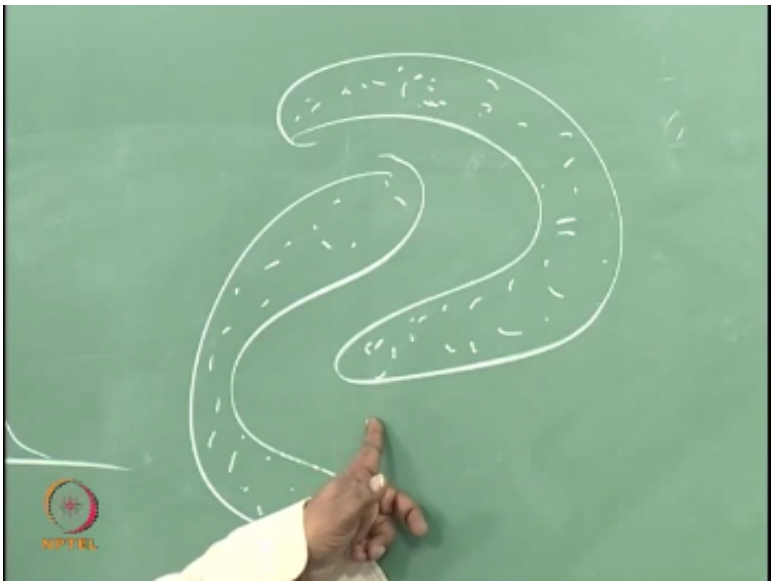
Now so the first one is that the data should have at least one cluster, what is the meaning of the data having one cluster suppose the data is say unimodal data mode is only one say something like this single dimensional data this point occurs maximum number of times and then slowly and slowly the number of occurrences decrease then you have a single mode here that means you have only one cluster.

If you have only one cluster you do not need to do any clustering you look at the terminology here I said that there are no clusters because somehow the points are occurring equally likely and hear some points are occurring more times and then some points are occurring less number of

times I am using the word you have only one cluster I am differentiating this situation from this. And the next thing is that you should have more than one cluster okay, you should have more than one cluster.

So basically if the whole data follows uniform distribution you cannot have you need not have to apply any clustering algorithm that is one and secondly if your data has data has exactly one mode or the data is unimodal then you need not have to apply any clustering method only when the data is not following uniform distribution a case like this, okay.

(Refer Slide Time: 18:25)



The data is distributed in this way if you look at each one there is uniform here and on this one there is uniform, but if you put these two things together then it is not following uniform distribution in this whole space, because there are many blanks here okay. So if the data is not following uniform distribution or if the data is not if it has at least two modes then you need to do clustering over questions. You do not know what to ask how I will just say one thing for many of these things the corresponding tests are given oops and drains book algorithms for clustering data all the statistical tests you will find there.

End of
Moudule 06 – Lecture 08

Online Video Editing / Post Production
M. Karthikeyan

M. V. Ramachandran
P. Baskar

Camera
G. Ramesh
K. Athaullah
K. R. Mahendrababu
K. Vidhya
S. Pradeepa
D. Sabapathi
Soju Francis
S. Subash
Selvam
Sridharan

Studio Assistants
Linuselvan
Krishnakumar
A. Saravanan

Additional Post – Production
Kannan Krishnamurty & Team

Animations
Dvijavanthi

NPTEL Web & Faculty Assistance Team
Allen Jacob Dinesh
Ashok Kumar
Banu. P
Deepa Venkatraman
Dinesh Babu. K.M
Karthick. B
Karthikeyan. A
Lavanya. K
Manikandan. A
Manikandasivam. G
Nandakumar. L
Prasanna Kumar. G
Pradeep Valan. G
Rekha. C
Salomi. J
Santosh Kumar Singh. P
Saravanakumar. P
Saravanakumar. R
Satishkumar. G

Senthilmurugan. K
Shobana. S
Sivakumar. S
Soundhar Raja Pandian. R
Suman Dominic. J
Udayakumar. C
Vijaya. K.R
Vijayalakshmi
Vinolin Antony Joans

Administrative Assistant
K.S. Janakiraman

Principal Project Officer
Usha Nagarajan

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