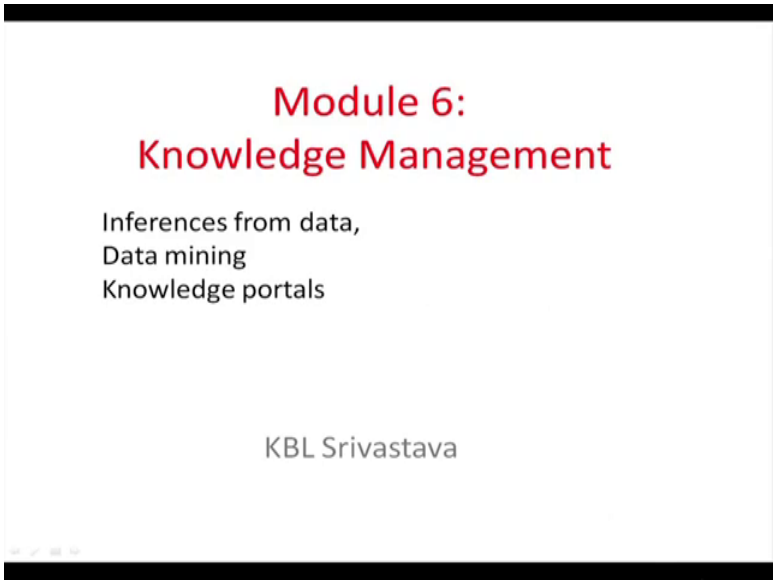
Knowledge Management
Prof. K B L Srivastava
Department of Humanities and Social Science
Indian Institute of Technology - Kharagpur

Lecture – 26
Inferences from Data

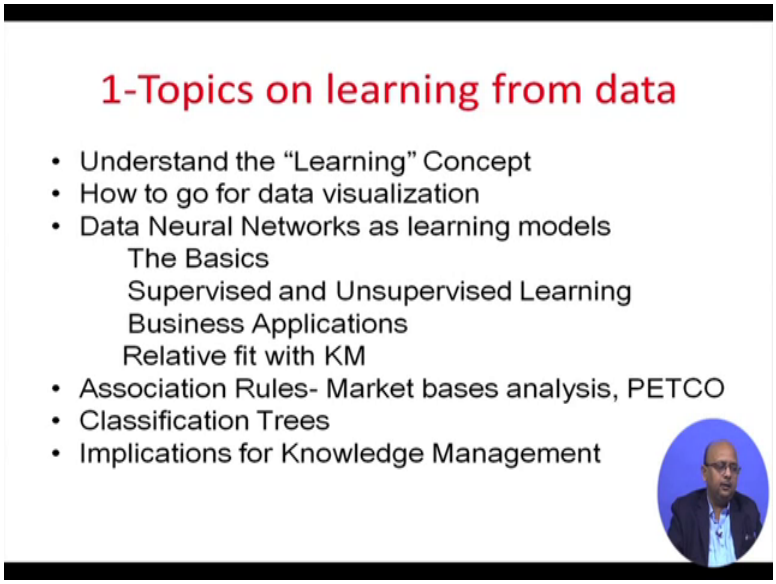
Okay, so now we are going to start the sixth model of knowledge management and this actually talks about that how we are going to draw certain inferences from the data. Now once you have a knowledge management system and the data is there in the warehouse or it is in the repositories now the major problem is that how you are going to make use of that data and draw certain interference centers of data's or recognizing the data and then getting some value out of this data so that you can make use of it right. And then we also discuss some of the techniques.

(Refer Slide Time: 01:07)



And one of the major techniques that are used is data mining and finally we will see that how we can use knowledge portals for dissemination of information and other kind of things. So this model basically talks about three things. And this week we are going to talk about inferences from data that how we are going to learn from the data, we are going to discuss a lot of issues related data mining and then knowledge portals.

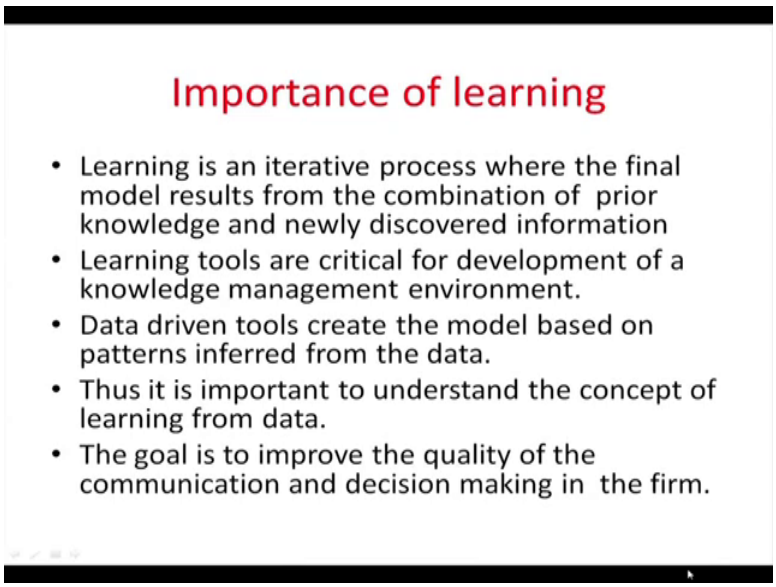
(Refer Slide Time: 01:28)



Moving to the first part that is learning from the data so we will try to understand what is learning then how we go about visualizing the data in different forms it could be in graphic forms and pictorial forms and it could be in a different forms. Then we will also discuss about neural networks and learning models and then we will also talk about some of the rules. Which are used for identifying patterns from the data and then we will see that how we go about classifying data's and developing trees for decision making process and what are the implications for knowledge management right.

So these are topics that were to cover in learning from data. Now when you are going to talk about learning as such so we have to define what is learning okay.

(Refer Slide Time: 02:14)



So in layman terms if you look at learning it is defined as something that changes our behavior okay. So if you are going to learn something it is going to help you to behave in a different way, so learning is a process through which we are going to change and transform our behavior. And that happens either due to practice or rehearsal right, that is the general definition that we can

give for learning. So the basic object of learning is to make changes in the behavior.

Now when we were talking about learning in terms of knowledge management so what is the learning in case of knowledge management that what you have learned from the knowledge that is stored in the data from or some other form in the repositories okay. So you are going to develop certain models are going to frame certain rules for this purpose. And based on these rules and models you are going to derive some inferences from the data okay. And for that purpose you are going to use certain learning tools okay.

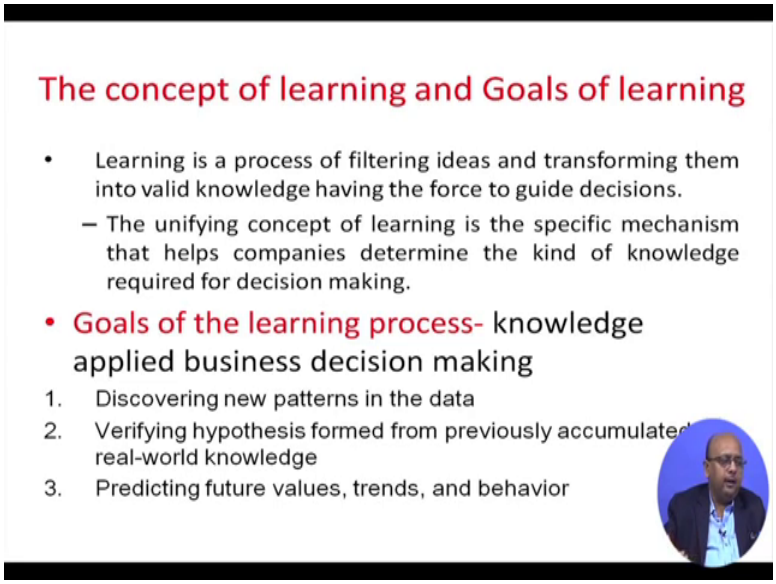
These learning tools are very, very critical because if you are going to use them then you can find your meaning, patterns okay. You can make predictions also you can classify the data's there are a lot of issues that can be resolved with this and that is why we are going to use these learning tools okay and they are very, very important and this can be used for knowledge management basically the information available in explicit form and you are going to make use of this information to derive certain value for business.

So it is very, very important that you do certain models and techniques of learning which is going to be helpful in the process, now when you are going to talk about data driven tools then there are different ways and means through which you can derive inferences it could be in the description form whether you are going to visualize the data or you can use statistical tools of software to analyze the data and identify certain patterns and values or trends out of this data's.

So it is very, very important to understand that how we are going to use this concept of learning from data that is how what you are going to learn out of the data that is available because your understanding is based on the learning from the data which you are going to make use in your job to perform certain things more effectively. And in the process there are two things that is very, very important that is, It is going to help you to communicate and take better decisions, say suppose the data is available in certain form for my unclassified data and based on that you are willing to take second to see.

So what data is communicating to you is very, very imparted and how we are going to learn certain things from the data based on your inferences and then take certain decisions which is going to add value to the business and that is why learning from the data becomes very, very important in the process.

(Refer Slide Time: 05:16)



Now if you look at the learning what happens in the process of learning that you get to know about certain things you have knowledge about certain things and then you are going to transform that knowledge into action okay. So that learning the basic objective is to see that how knowledge is transformed into action it means whatever you have learned. By way of experience either through explicit form or inexplicit form or sharing knowledge that how you are going to use it in order to perform your job okay.

So knowledge leads to action would be correct only if you are able to apply that knowledge to perform your job effectively okay. And that is how you can validate that knowledge that is the kind of knowledge that is available in that database is correct and that you are able to use that to transform certain things to a more practical form and then it is going to help you to take certain decisions related to your work your job or the kind of things right.

So if you look at this concept of learning it is very, very important because learning helps you to transform certain things okay it helps you to visualize and understand certain things in a particular way so that you can decide what is to be done what not is to be done right, and what kind of decision you have to take out of it so it could be related to decision making, it could be related to productivity, it could be related to performance right. So if you look at the major will of learning what is that when you are going to learn something it means that you are going to acquire knowledge.

And this knowledge is going to be acquired either in explicit form are tacit or are through different processes like we talked about the earlier. Information externalization of depending upon the kind of knowledge that you have gained in the process either in explicit form tacit form through different processes like we talked about that earlier like through socialization, Internalization or externalizations okay. Depending upon the kind of knowledge that you have

gained in the process that either in explicit form or implicit or tacit form the idea is that you are going to make use of that knowledge tool work or apply that in business context right.
And that is the major goal of learning so that knowledge that you have gained through the process of learning is going to be useful in business context okay so that you can perform your job or you can take decisions okay . Now what is the objective of learning from the data is that. If you have the data and if you are able to identify certain patterns of the data okay this is how the things are moving here. Say for example if you are having data of past five years of the productivity trends okay these are the productivity of the firm since last five years okay.

Now this is a data now based on this data what kind of decisions you are going to take. Are you going to use these data, analyze the data and then infer something from the data about what are the likely trends which are going to be there next years that is the sixty years so at least it will give you a trend and based on that based on that you will take certain decisions what kind of actions you want to take, so it basically help you to identify trend okay.

It means based on five year data you will be able to predict what is likely to happen in the sixty years. Provided the same kind of situation remain right. Now this is one way to look at it, another way to look at it is that you can form certain hypothesis based on the data that X. Y. Relate X is really to this. Say for example if you go for a particular strategy okay. Then it would result in this much profit. Suppose you are going to spend X. Percentage of your revenue on advertising then this would relate to Y. percentage of profit. Now this X. is to be related the Y.

So you make hypothesis that Spinning on advertisement would be positively related to Sales or Revenue from the market. Now this is the hypothesis which you have developed. Now you are going to collect data on both the variables that is spending on the advertisement and revenue sales right. Now we are going to see that whether this hypothesis is correct or not and this is based on your previous experience that those organizations or even your organizations are spending more on advertising the product have been benefited out of it and got better profits or got better sales from the revenue now you are going to examine this hypothesis okay.

And see whether this hypothesis is correct or not. Now you have data on advertising and you also have data on Sales. Now you can relate both the data and see whether there is a positive correlation between the data or not. So if there is a positive correlation find that yes, spending on advertising has really helped the business to make more profits or more sales in the market then this will help you to take certain decisions okay that how much we should spend on advertising in future right. So that is the second point so you again verify certain hypothesis

based on your experience intuition and knowledge which you have accumulated earlier.

And then you collect data on that and then you can say whether it is correct or not right. Based on this then you can predict the future values, trends and behaviors right. Say for example you have installed a knowledge management system in the organization. Now you want to see that what kind of attitude people have about the KM system what is the use? Okay. Now installation of KM use it means that you have made certain investments. In KM system now you want to see that how this investment in KM system is related to the use of game for system and accordingly the change in the behavior of the employees.

Whether they have become more innovative? Okay with their using the knowledge management system frequently or not. So these kind of behaviors entrants can be identified depending upon the data so what you need to do is you need to collect data about how much investments you have made in the knowledge management system and you can also look into the frequency of use. How many innovative behaviors with have been made by the employees what kind of suggestions are there what kind of creative ideas are given by the implies so you can relate these two things and identify okay.

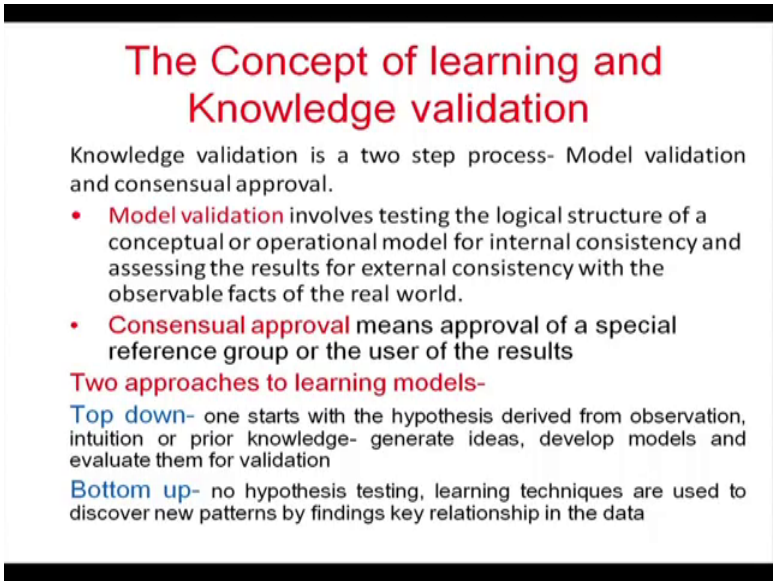
If you are going to invest more in KM system, having a good KM system it is going to help employers to be more creative, and innovative in the process, the idea is that you can predict certain things based on certain things that you have so suppose you have X. Where they will that is independent variable why a variable so whether X. Is going to project away or not that too that is to examine right and then lot of statistical techniques that will be applied to the data and ways on that you can identify with the trends that is there is going to help you to take certain business decisions right.

For example suppose you have gone for a particular strategy okay. Now you have to see that what is the result of this particular strategy once the results are out you are going to see that in what way that particular strategy is related with the results or outcomes of the organization to some effectiveness, productivity say for example you have initiated total quality management program in the organization right. Now you can predict the value of the total quality management program by identifying certain criteria in terms of its outcome okay.

In terms of improvement in quality, efficiency and productivity right. So these things are happening you can say yes the KM program is really helping the organization on these criteria productivity, efficiency and quality. So you are able to predict certain behaviors based on the

data provided the data is available okay. So when you are talking about data and then identifying trends, patterns and values of the data basic idea is that you are going to make use of these inferences to make certain business decisions right.

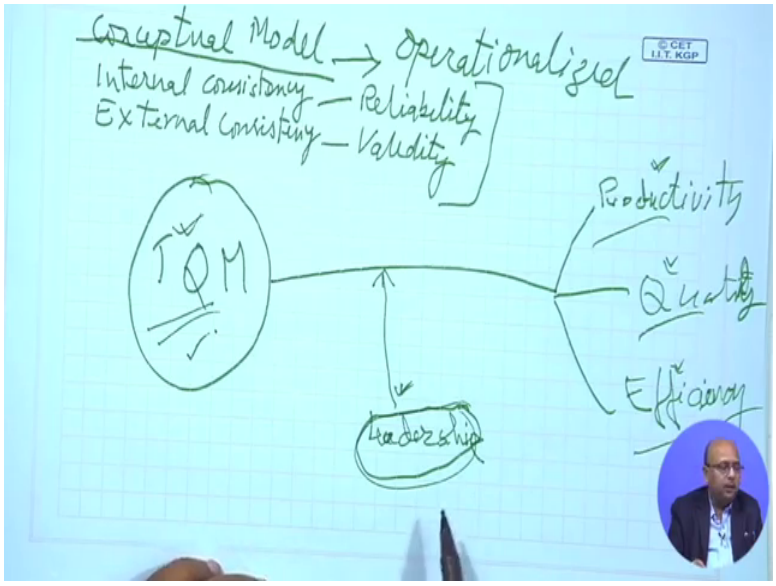
(Refer Slide Time: 13:39)



Now the basic idea is that you are going to validate certain knowledge based on the data right. When I say that okay T Q M is related to improve the quality of the product okay. This is a hypothesis that needs to be validated okay. Right, so we need to validate this knowledge based on the data okay. So there are two ways through which it can be done one is model validation and the other one is based on the consensus. In case of Model validation what you are going to do you are going to develop a structure that okay this is T Q M.

And this T Q M is related to productivity, efficiency, quality right so say for example. I am going to write or develop a model based on this validation so what I am going to write I am going to make one column and I am writing T Q M.

(Refer Slide Time: 14:35)



Now I said this T Q M is related to certain things which could be related to productivity, quality efficiency right. So these are the outcome variables and this is the independent variable. Now this may also be moderated by the kind of support that you get from the management or the leadership okay. Now this leadership is a variable which is going to moderate the relationship between this T Q M and these variables. Now in the process when you are going to validate this model what you are going to do we are going to first implement this total quality management in the organization.

We are going to see that what kind of leadership support is available then we are going to collect data on productivity in those units or departments where this T Q M has been implemented. So how whether there has been significant change in the productivity quality and efficiency or not. So what will do it will try to relate whether this is due to this total quality implementation or some other factors and the extent to which the leadership has been able to support this process right? So this is a conceptual model that we have developed it has a logical structure.

And then this model has to be operationalize in terms of its measurement So we are going to measure it one is known and is a conceptual model and then this model needs to be operationalize in terms of measurement right. So once this is operationalize it means that you are going to measure all these variables and then you are going to use certain structures tickle tool to see whether T.Q.M. is able to predict these things are not and whether there is the interaction of these particular leadership right. I mean this model is to be tested for both that is we are going to check for internal consistency and you are also going to check for external consistency.

Internal consistency is nothing else but you are going to check the reliability. And also the validity okay, so this reliability and validity is to be checked by getting data through observations from the real world in an organization. So when you are going for model validation this is the process that you are going to test a logical structure of a conceptual our operational model that is developed here and then you are going to assess the results for both reliability and validity. Reliability mean how consistent the results you are going to get all the time you are going to implement.

Suppose you have implemented T.Q.M whether it is because of T.Q.M or some other factors because of which it has gone right. So we have to check both reliability and validity. So validity means whether it is really because of T.Q.M or some other factors and reliability whether it is

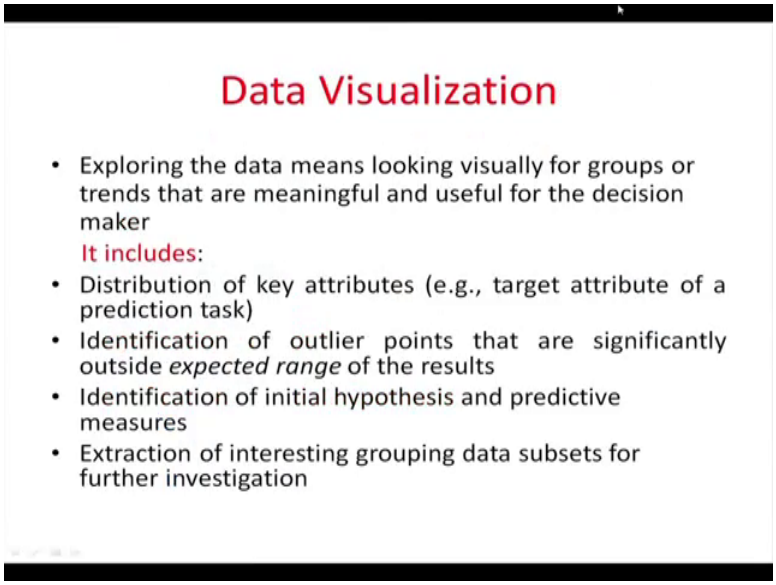
related this or not. So if you use certain statistical technique then find a positive correlation and think that okay this is able to explain certain variances in the level of productivity quality and efficiency then you can say that yes, this model is validated.

Now this model is to be tested operationally by getting data whether this data is available in the Knowledge Management System or whether this data is to be collected from the organization it depends upon that. Now it could be both either you get the data from the system or you collect the data through your observations and then you see whether this model is going to be correct or not. Another way to validate is that you have subject matter experts.

A special reference group and are they agree to this point that yes if you are going to implement T.Q.M then it would result in these kind of things but that is more qualitative approach above subjective, but the first qualitative approach is that when you go for model validation through data then it is a more quantity approach right. Now it could be done in two ways that is top down either you have derived hypothesis from your observation okay.

Or prior knowledge or you generate ideas that if you are going to implement this then it is likely to happen or bottom up do not test in hypothesis so simply go for T.Q.M and then see what happens right. So both kind of things can be done and this is known as top down approach or bottom up approach. So both kinds of things are possible.

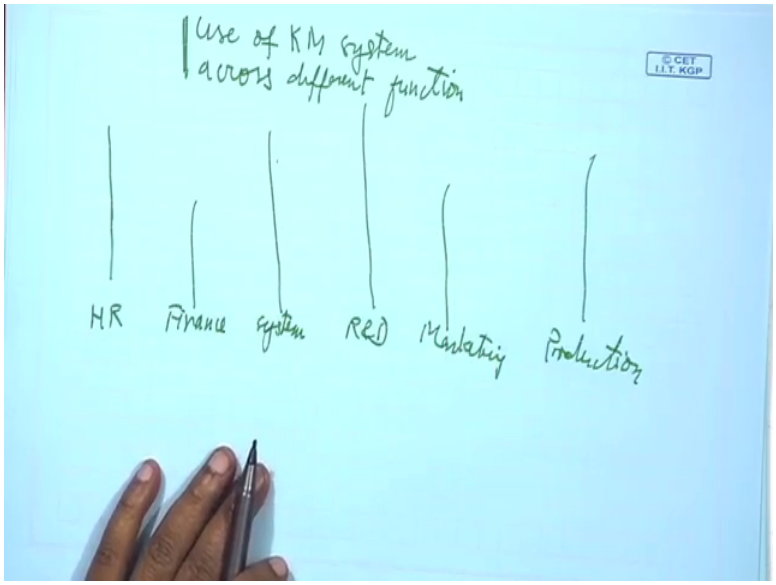
(Refer Slide Time: 19:55)



Another way to look at it is data visualization okay. Now when you are going to explore data it means that you can also look visually what kind of trends are coming up okay. So it would be a graphical presentation for example it could be related to a particular attribute task the way it is changing. So you can see that how a particular distribution K Attributes is being distributed across the people or group right. And you can also identify out layers. Say for example if you

are going to collect data okay about the use of K.M. system.

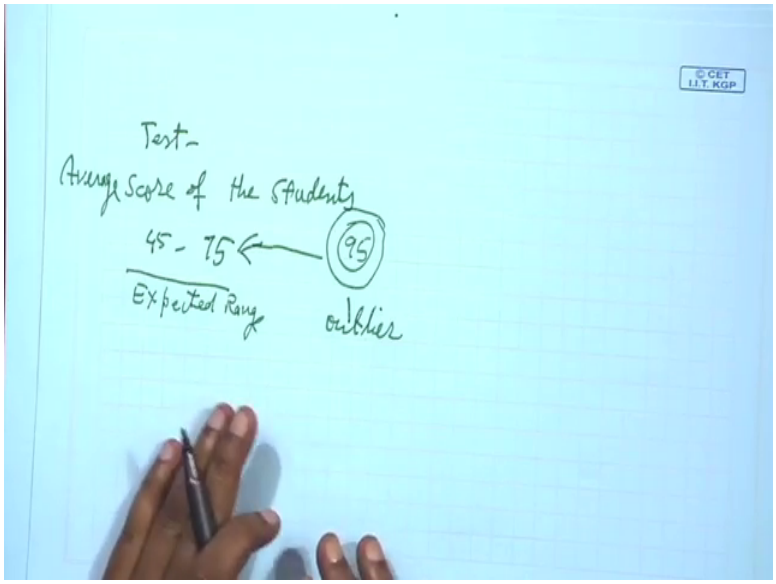
(Refer Slide Time: 20:34)



Across different functional, now you have HR, you have Finance, you have System Department, you have R and D department, and you have marketing okay. You have production so these are the different functional areas which are going to use the K.M System. So this is what one attributes which you are looking at that is use of K.M system across different functional areas now you have to see that how much use is there in terms of quality, frequency, other kind of things and then you simply put a graph like this. So this is basically a visual representation right correct.

And then you can see that okay on an average various departments are going to use the K.M system. So this basically talks about one particular attribute that is use of K.M system across different groups okay. And then you can predict look at this is how they are likely to predict in the future. Another important issue is identifying the out layers okay. May be other departments are using this way but one particular department you find that production department or say R and D department is using it in this way. So this becomes the out layer okay, out layers are those points basically which does not fall within the ambit of the average right.

(Refer Slide Time: 22:24)

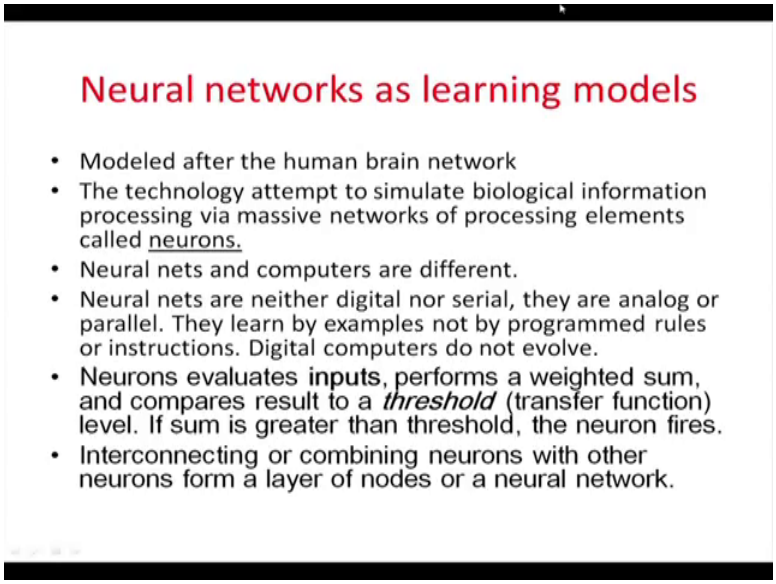


Say for example if in a test the score of average score of the student, average score falls between say forty five to seventy five, right. Now but one student has got ninety five right. So it means this is becomes an out layer because it is going to influence the average of this. So this is known as the out layer because it is out of the expected range this is the expected range and this is called score is an out layer right. So we can also identify the out layers okay.

And then we can develop our hypothesis okay and then we see that what kind of predictive measures we are going to use, predictive measures is what kind of tools and techniques you are going to use in terms of statistics and other things with the help of which you are going to you are going to make certain predictions okay. And then you can also extract interesting group data's for further investigation suppose you have this kind of data this is a data which you can, you have extracted now you can see that how different groups are reacting in the K.M System okay.

You can also get feedback about the K.M system from the different group's right and then see that what is the reason for using or not using K.M system from the different groups. So you can further investigate it and get their feedback from this right. So this is about data visualization.

(Refer Slide Time: 24:12)

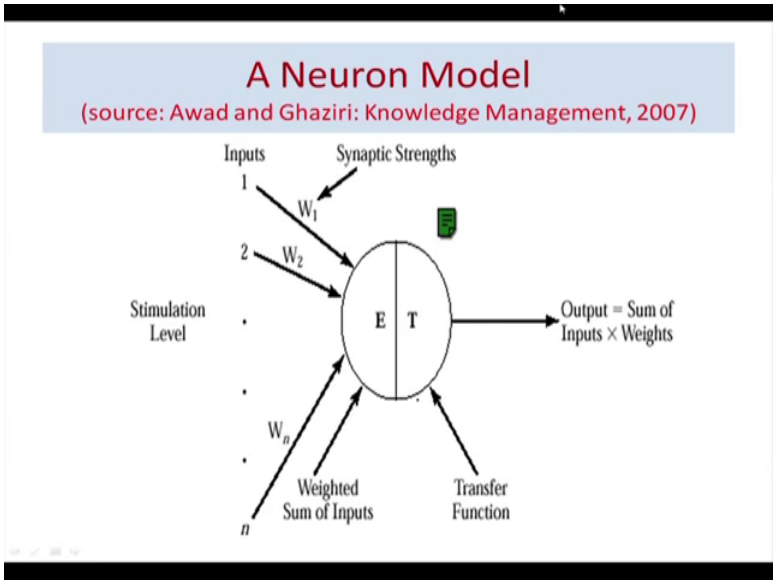


And we are moving to next part where neural networks has being used as learning models right. Neural networks basically are stimulated based on the human brain network but there is a difference because the basic idea is that you are using technology to simulate brain okay. Basically the brain which processes biological information it has a massive networks okay. And these networks are connected through you know this is known as neurons so the idea is taken from there okay but these neural networks are computers and computers are different because neural networks do not evolve.

Evolve over a particular time in the human being but in computers it is more static right. And there is neural nets are not digital or serial but they are no longer parallel because you can learn for example they are not programmed rules for the human brains okay. And digital computers whatever configurations you have they work on the basis of that if they do not evolve. But if you look at these neurons the idea has been taken from the human brain and then you have inputs okay. And then you go for transformation of these inputs okay.

And then you come out with a transfer function, so that is what we call a threshold function okay. So if the sum is greater than the threshold it means that the neuron fires it means that it is stimulated and then it is going to act right then you have interconnections among the various neurons okay and then there are notes through which these neurons are interconnected. Now let us see how they are connected through this.

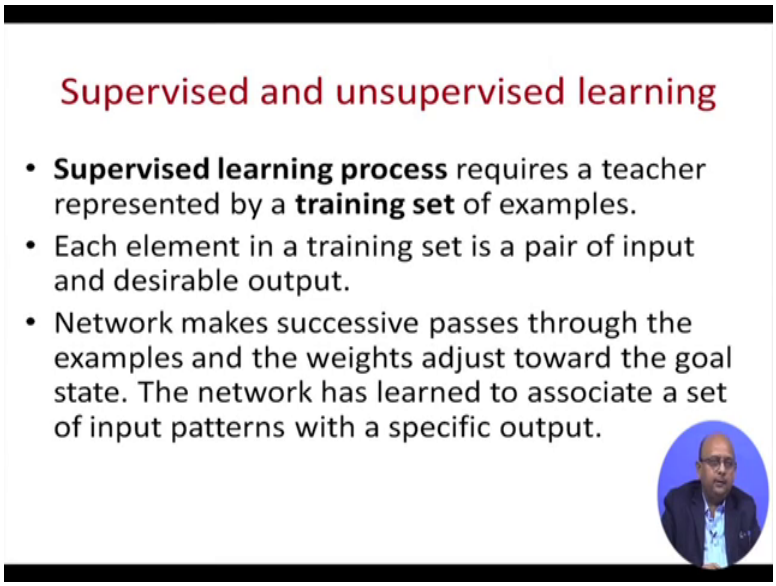
(Refer Slide Time: 25:46)



Look at this, this is the model of a neural network you can see that there are various inputs you can see the various inputs various inputs which are acting as stimulators then these are the sign they the connections are known as synaptic strengths. These are synaptic connections and these are the various weights and this is this is where the transformation takes place okay this is nothing in the weighted sum of inputs that how much stimulation is there and whether it is able to fire or not and firing depends upon the weight of some of the weights of these inputs.

And this transfer function is dependent on what you call the output is nothing else but the sum of inputs that you have multiplied by the weights and then you go for calculating whether it is going to fire or not. Similar idea is to understand that whether you are getting inputs in the form of neurons and whether it is going to help you or not and whether it is able to connect it or not and whether it is able to connect it or not and whether it is going to help you to take certain decisions or not.

(Refer Slide Time: 26:54)



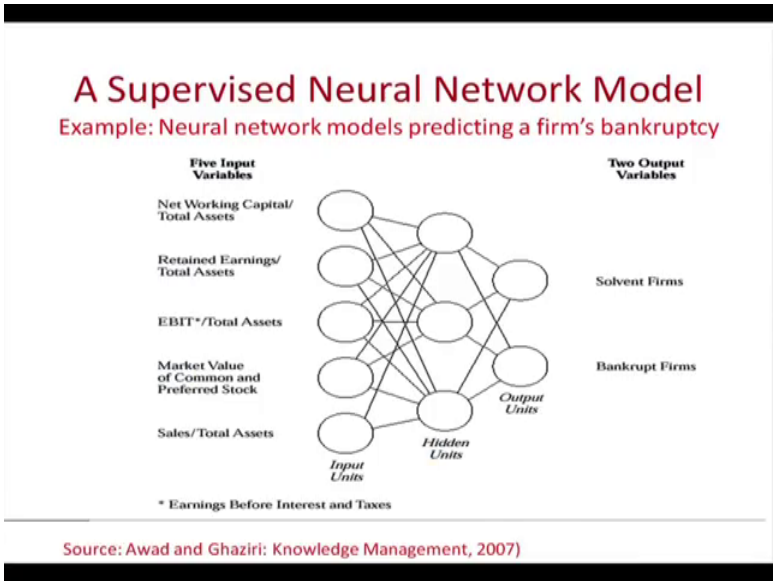
Now related to that there are two kind of learning that happens in neural networks one is more

known as supervised learning and the other one known as unsupervised learning. So in a supervised learning process what happens is a teacher and there is a training set of examples through which you are going to learn the process okay. And each element is a input okay. And then you also expect certain output which is desired after you have gone through those inputs right.

These inputs are processed by you and then you have come out with certain outputs. Now do you have these networks okay which regularly processed through the examples and waits that is adjusted towards the goal? Now the network has learned to associate a set of input patterns with specific outputs. So if you are going to do this, this is going to happen right. It means there is a direct combination between the inputs and the specific output it depends whether the inputs have been able to simulative to fire and come out with certain outputs or not right.

For example suppose you are sleeping and then suddenly you are hearing a sound bang right, what happens it is an input which is strong enough to wake you up right so similarly. So this is a cause of and effect relationship. Similarly you have to see that what kind of inputs are there what is the how the weight is being transferred and what is the output for the people.

(Refer Slide Time: 28:11)



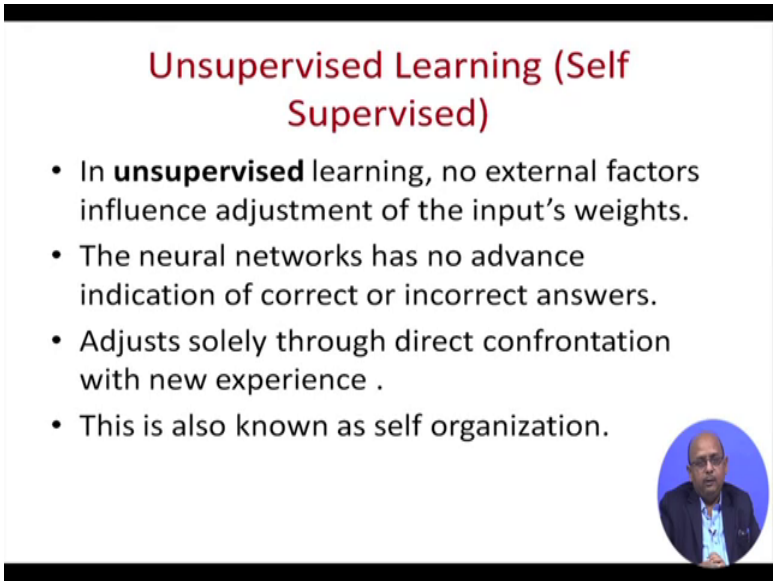
Now one example that I have taken from this source is that how we are going to predict whether the firm is going to bankrupt or not using the neural network model. Supervised Neural Network Model basically now see the number of variables that we have taken here includes what is the networking of the capital or the total assets then what is the Retained Earning then Earning before the Income tax their market value what is the market value and Sales and total assets.

Now if you look these variables okay these are the input units right so they are interconnected

these are these are known as nodes and see all of them are interconnected okay. So now if you look at this and then based on this you are coming out with the output. The output is that whether you are able to save yourself or whether you are going to be bankrupt right. Whether it is going to be solvency or whether it is going to be bankrupt right. So these are the variables which are going to decide what you will do you are going to collect the data on these variables process them.

And based on that you are going to take decisions which could be in the form of either Solvency or in the form of Bankruptcy of the firm right.

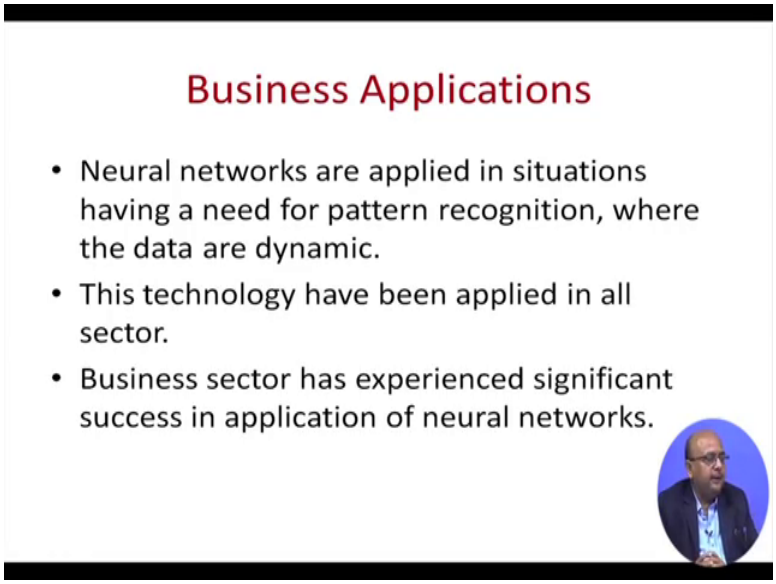
(Refer Slide Time: 29:27)



Now in case of Unsupervised learning what happens is that it is basically self organized because there is no external factors influence adjustment of the input weights and even if you look at neural network it does not have any indication whether the answer is correct or not correct so basically depends upon how you are going to see it because you are going to have direct confrontation with the new experience and it depends upon how you are going to organize the things.

It is see if you look at this it is very, very much organized and you can identify the nodes and the network that is connect to their different levels and then output but in this case it is based on your experience and learning okay. And that is what we know as self surprised.

(Refer Slide Time: 30:10)



Now since we have talked about learning and neural network model for learning the next thing is that we have to see that when they were able to apply it in the context of business or not right so it has been applied in different context where you want to recognize certain patterns based on the data. And it could be applied in any kind of sector okay. And off late you find that a lot of organizations or businesses have been able to use neural networks to take certain. Take it forward and make certain decisions based on that okay. Thank you.