Data Analisys and Decision Making - II
Prof. Raghu Nandan Sengupta
Department of Industrial & Management Engineering
Indian Institute of Technology, Kanpur

Lecture - 60
General Concepts

Welcome back my dear friends and dear students a very good morning, good afternoon good evening to all of you wherever you are and this is the DADM - II which is data analysis and decision making two course under the NPTEL MOOC series. And as you know this was total course is basically for 12 weeks, which is 30 hours of contact over the lecture series and which is 60 lectures in total because each lectures is for half an hour.

And we have already completed 11 weeks each week with 7 lectures and after each week you have assignments. And if you can see the number here which is the 60 DADM - II lecture number 60 which is 12 week; that means, we are when you wrap up this course with this lecture and; obviously, after this lecture you will do basically taken as this assignment number 12 and with that you will basically have the final examination also which covers the whole course.

If you remember we are discussing about the concept of NN very briefly then we went to the concept of change point detection also briefly and we just mention the type of tests. So, petit test would be utilized considering the concept that we want to basically check the points under the Mann Whitney concept of test which will consider, where we are using two things I will just give a very brief two minute discussion which I actually do before the starting of the class.

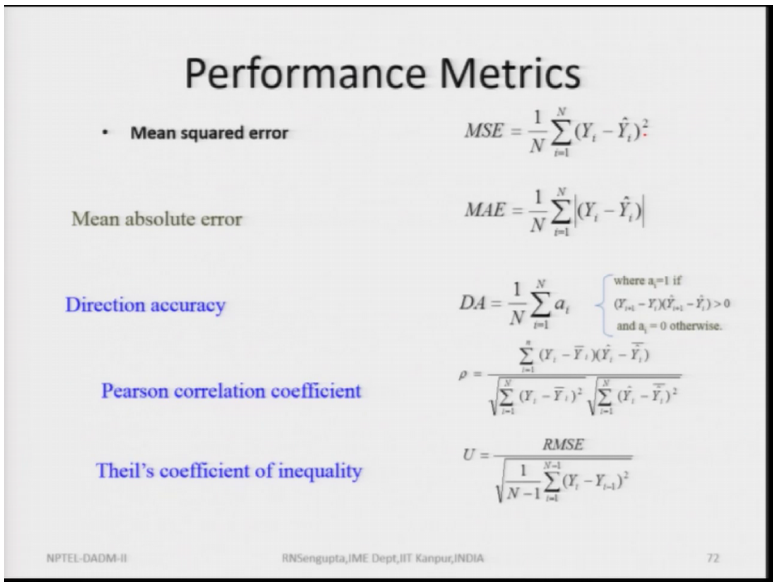
So, you have the NN methods will be utilized because prediction of exchange is a non parametric method, because we do not know which parameters based on which you can predict the exchange it. And as I said they can be different parameters we change and what is the fluctuation also we do not know it can be petroleum price, gold price, silver price, it can be inflation rate, it can be interest rate of that country it can be interest rate of other countries whatever it is.

What is the GDP, GNP population they can be many objective and subjective criteria, say for example, suddenly there is a political approval in one country so; obviously, it will have an effect on the exchange it; consider there is an election. Now we will we basically would try to predict using the artificial neural network, the prices are there we would like to predict, but the actual prediction is not the main concern our going consist of find out the change point depending on where, we can take a decision where they want to enter the exchange of market or come out of that.

Added to that we will try to basically fine tune are NN network based on this either the conjugate gradient concept method; and to top it all basically utilize the NN; ANN along with the conjugate one with genetic algorithm simulated annealing. And we will try to utilize this combination for two sets; one is the change point detection algorithm being utilized and another case the change one an algorithm not being utilized.

Obviously when we when you run any method or tests you have some criteria based on which you want to basically analyze how good or bad your results are like.

(Refer Slide Time: 03:38)



So, some of the results I will just briefly mention them is the mean squared error where we try to. Predict the actual value differences on the actual value and the predicted value and find out the square of that and find out the average. So, this is something to do with trying to minimize that mean squared error is something to do with minimizing the variance; because this is square term in you are trying to find out the average. The mean

absolute error is that we find out the errors in any direction and try to basically give the weightages of plus.

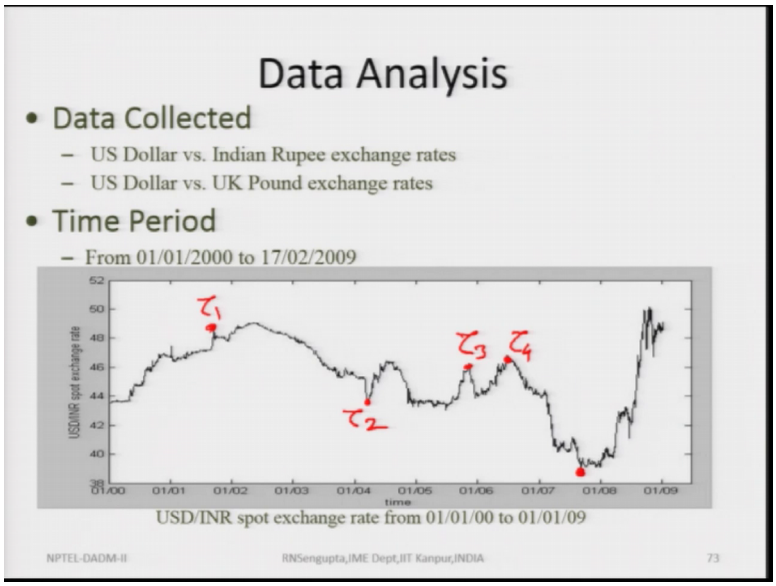
So; that means, any negative or positive direction methods are all are consider positive then again we find out the expected value. The direction accuracy one is basically that depending on the actual value and the predicted value for the time period next i plus 1 whether if it is both are positive you give a point of 1 and if both are negative also you give a point of 1; so, in the direction that the predictions which is happening in the actual value now and the, predicted value later on and the actual value later on.

So, if the difference between the actual value then; that means, t plus 1 and t and difference between the predicted value of t plus one and the actual value of t if they are moving in the same direction; that means, we are trying to predict in the right direction. If it is more positive the first one is positive first one means the difference between Y t plus 1 and Y t and the second difference between Y t plus 1 hat and Y t if both are positive, then it is plus 1 both are negative; that means, I am going in the negative direction my prediction is also right then also it positive.

If they are in opposite direction; that means, actual it is going plus in the in its increasing I am giving a prediction on decreasing or vice versa then; obviously, the value would be negative. We will try to basically find out the direction accuracy on an average the Pearson correlation coefficient and Theil's coefficient of in inequality would be utilized depending, Pearson correlation coefficient we know is basically very simply the correlation coefficient which we use in the concept of covariances.

And the Theil's coefficient in equation would be the root means squared error divided by the prediction value which is happening for two time periods, N minus 1 depending on the value of the degrees of freedom which we already have considered.

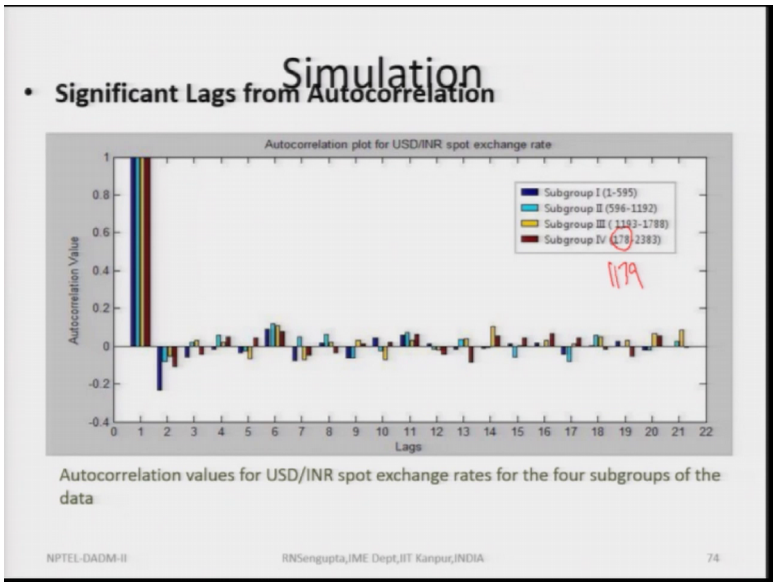
(Refer Slide Time: 06:01)



So, now we basically take the US dollar to Indian rupees and us dollar to UK pound the exchange rates and they are taking from 2000 to 2009 from first January to 17th of February.

And here you see the exchange rate of US to Indian spot rate. So, say for example, you may say well pictorial is thing; this is a change point here or this is a change point here or is the change point here, may be a change point here, maybe a change point here. So, this total time frame is capital T and we will basically try to find out if this is tau 1 this is tau 2 this is tau 3 this is tau 4 and so on and so forth. So, pictorial what you are seeing you will try to basically verify it using the NN added to that the conjugate gradient method plus it is genetic algorithm and the simulated annealing method.

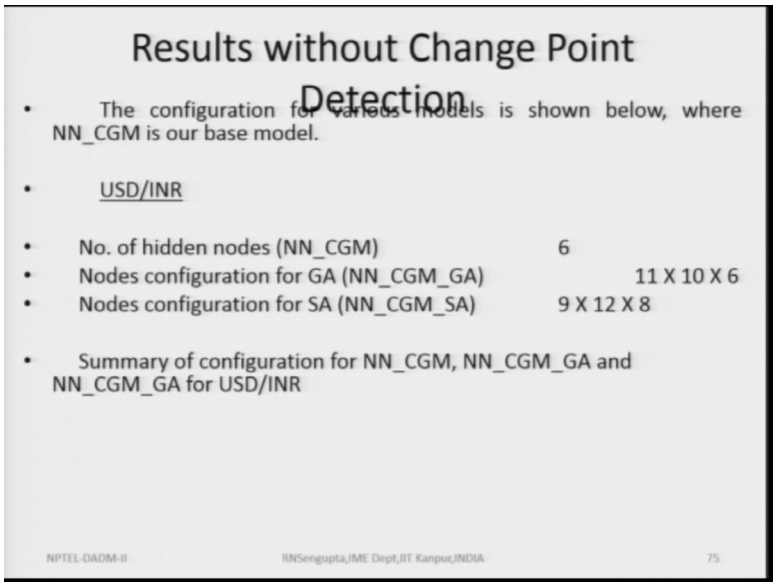
(Refer Slide Time: 06:56)



So, we find out seeing the using the time series concept the lags and the autocorrelation. So, with the sub groups are based on the fact. So, we have basically the data set of 2383 data points. So, we have the sub groups from 1 to 595 data point, 596 to 1192, 1193 to 1788 this should be 1189. So, 1189 to 2383 based on that the sub groups are calculated and you find out the auto correlated values of the US to the Indian spotted for the four subgroups based on which you can divide and then proceed.

Subgroups you remember I mention that if you want to predict standing from one year down the line, they would be time points in between where you want to enter the market or come out of the market and take a decision accordingly. So, those sub points or sub groups where the data change point is happening we have to check that that is why we are divided into subgroups.

(Refer Slide Time: 08:04)



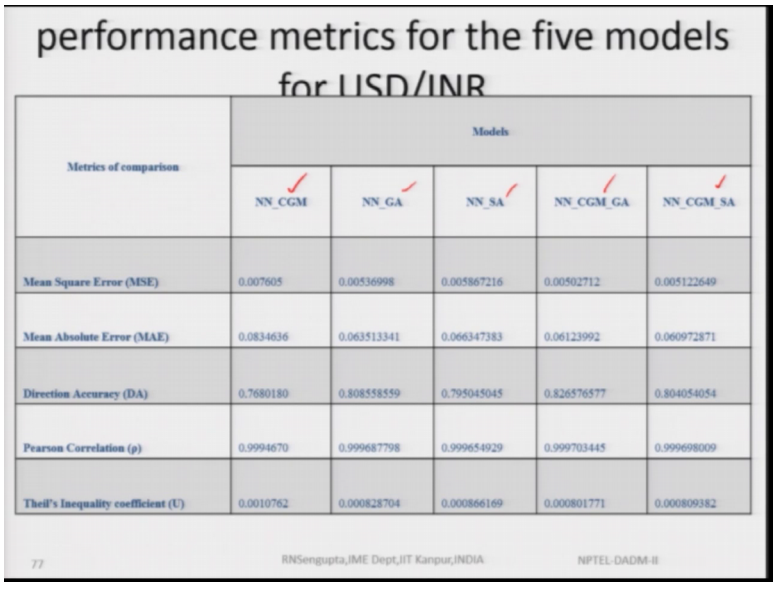
The configuration on the various models is shown below for the US in Indian exchange rate so number of hidden days for the ok. Now remember we had mentioned that the model which will try to use; obviously, we do some learning methods learning means we find out the optimum number of nodes and the hidden layers.

So, the number of hidden nodes which is there for the neural network in the conjugate gradient method is 6, then the nodes configuration for the genetic algorithm being utilize with conjugate gradient method; I am not considering them singly I am basically singly in the sense not NN using the genetic algorithm or NN using simulated annealing. So, the nodes are basically 11 cross 10 into 6 and 6 cross 12 into 8. So, the summary of this results are as follows.

(Refer Slide Time: 08:59)

So, this learning of how the nodes are build up I skip that and come to the results accurately. So, this is the spot exchange rate depending on from 2000 to 2009 which you have already considered.

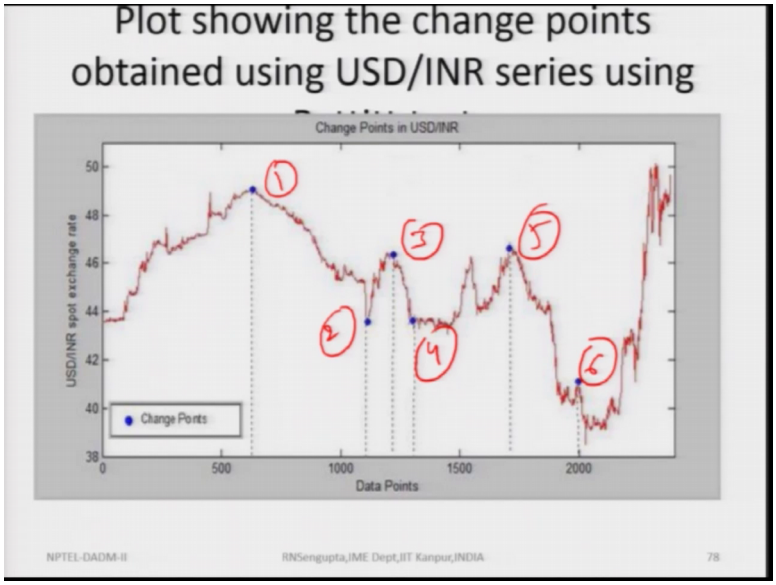
(Refer Slide Time: 09:10)



Now, the metric for comparison for these using the neural network conjugate gradient method, genetic algorithm simulated annealing, genetic algorithm and conjugate gradient method and genetic and simulating annealing using conjugate gradient method. The all the statistics are given statistics means the metric based on which we are going to say how good are bad are predictions are we use the mean squared the mean absolute error the detection accuracy, the Pearson coefficient and the Theil's inequality coefficient.

So, the values come out over this for the mean square it is 10 to the power minus 2 7 this is 8 into 10 to the 7 into 10 to the power minus 3 sorry this is 8 into 10 to the power minus 2 this is directly accuracy is 7 into 10 to the power minus 1. Pearson coefficient is about 1 and Theil's inequality coefficient is about 1 into 10 to the power of minus 3. So, all the values basically are given depending on the performance of the five models.

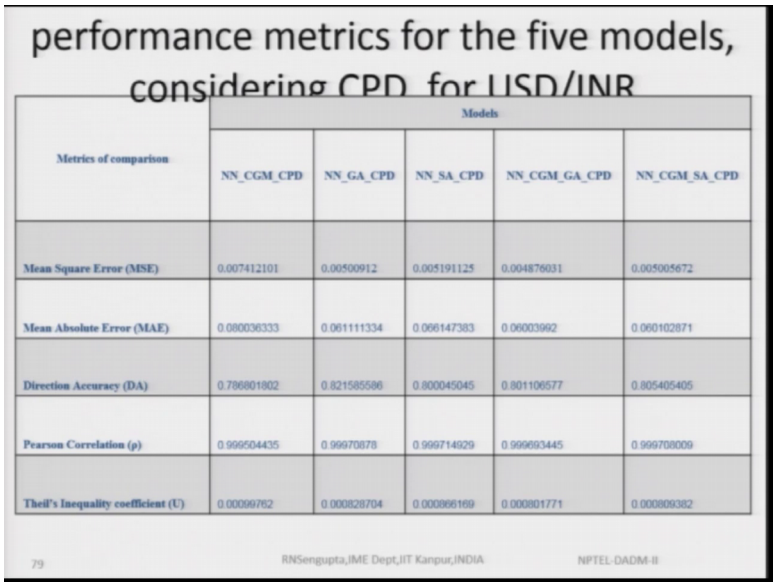
(Refer Slide Time: 10:20)



Now, here if you remember the points which are marked arbitrarily now when I do the change point detection depending on the petit test and when I find out the change points for the US and the Indian rupees are happening like this. So, I have not marked exactly what point so, the change point is happening round about 600 points 600 data point another is happening say for example, 3 points on happening between 1000 to 1500 one is happening say for example, at the range of 1700; and the another one is happening at the range of about 2000.

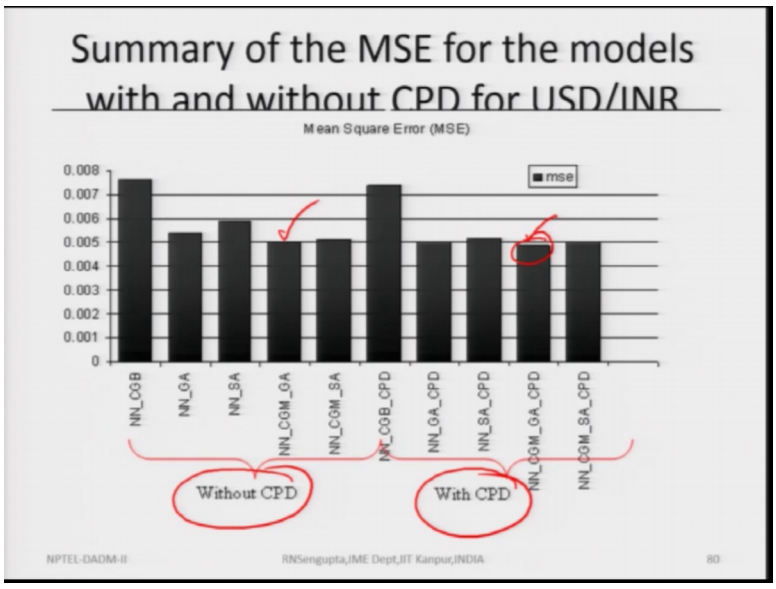
So, in all from 2000 to 2009 if you are take the data point. So, that the number of change points is basically 1 2 3 4 5 6 even though it looks that a lot of data points change at change point is happening, but here we predict the change point is 6 in number.

(Refer Slide Time: 11:23)



Now, performance of the matric based on the change. So, initially the change point was not utilize algorithm now you use the change point and the mean squared the absolute mean square error. The detection accuracy Pearson coefficient values I given I am not going to the details of discussion the values.

(Refer Slide Time: 11:41)



Now I want to basically give the summary of the mean square error. So, I have group them into two groups one is without change point detection algorithm one is with change point detection algorithm, but here we will basically have the without the change point

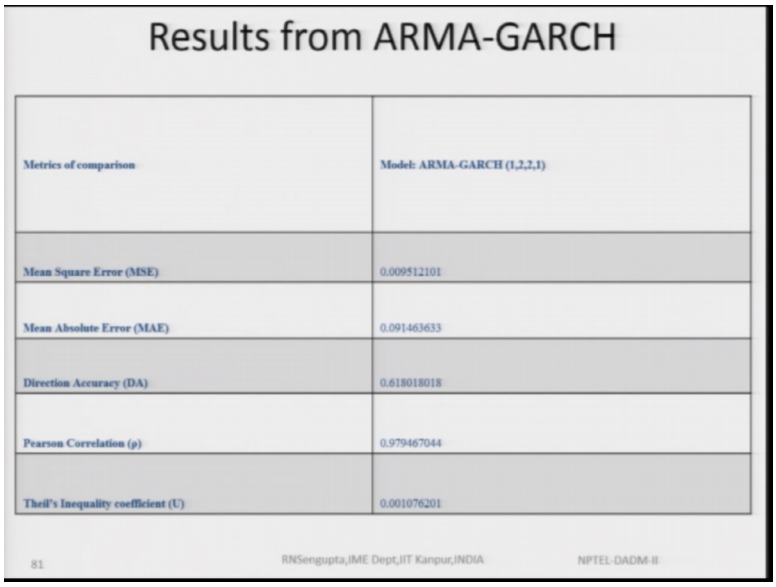
detection algorithm which is happening is conjugate gradient method, genetic algorithms simulated annealing and then utilizing conjugate CGM conjugate gradient method along with SA and conjugate gradient method along with GA.

Similarly, when I use the change point again the same sequence we use neural network with change point detection and conjugate gradient method, then genetic algorithm with change point detection simulated annealing change point detection. And finally, the neural network which considers change point gradient method genetic algorithm and the last one being change point change gradient method and simulated annealing.

So, the errors would basically give me which is the best if you check generally. So, they are very low on the scale of mean square error. So, the best which is coming out to be if I consider? So, the pictorial one is the genetic algorithm being used with change point gradient method without change point and the conjugate gradient method.

Another one the best one is coming out to be the conjugate gradient method along with genetic algorithm and the change point 1. So, the error is the best here error in the least sense.

(Refer Slide Time: 13:12)



So, results for the ARMA GARCH model if you use the means so, we can compare this results along with the results which we get for the different combinations of the of the change point and not change point method being utilize in the ANN along with GA or

and SA ok. Now I have just giving a snapshot of the discussions which we had so I will try to wrap up this DADM - II. So, I will go through.

So, if I consider I will go very briefly. So, this is the set of slides which you have discussed. So, the first one what we discussed was basically the concept of utility theory. So, in utility theory and utility analysis we did the concept of different type of absolutely utility concept, the relative utility concept and the absolute utility concept being utilized. Then we found out or I try to discuss the concept on non satiation and the risk aversion properties.

Then we consider the four different types of utility function quadratic power than the concept of exponential utility function logged logarithmic utility function then we discuss different examples then we went to the very simple concept of how we can find out the safety first principle can be utilized where you want to maximize or minimize the functional form? And then we consider the expected utility; obviously, we can find out and then we consider the concept of rather than, discussing through the slides I am just mentioning what we have covered; then we consider them the safety first principle which I did repeat.

So, they were different then we consider the geometric mean method and companies in the geometric method with the generally the safety for principle. And I did also mention that using the quadratic utility function also means that the returns was normal and vice versa. So, this is the two way equation for that.

So, then I will close and open the slides accordingly. So, I will close this also so it is easy for us to discuss. So, in the next one we discuss the concept of considering the utility functions, the certainty value was important you should do not definitely know that, then we considered how you find out the certainty value and how the expected values can be found out.

Then we went into the problems of certainty value and how it can be compared how certainty value can be utilized to compressing different decisions. Then the mean variance concept along with utility analysis was done that that the safety first principle and the stochastic dominance concept the geometric mean concept were discussed.

Then we discuss the concepts of utility function loss functions then the concept of loss functions being consistent unbiasedness properties, then you have the quadratic loss function, the mod loss function with unequal penalties being there for overestimation underestimation both for the both for the linear part and the non-linear part also.

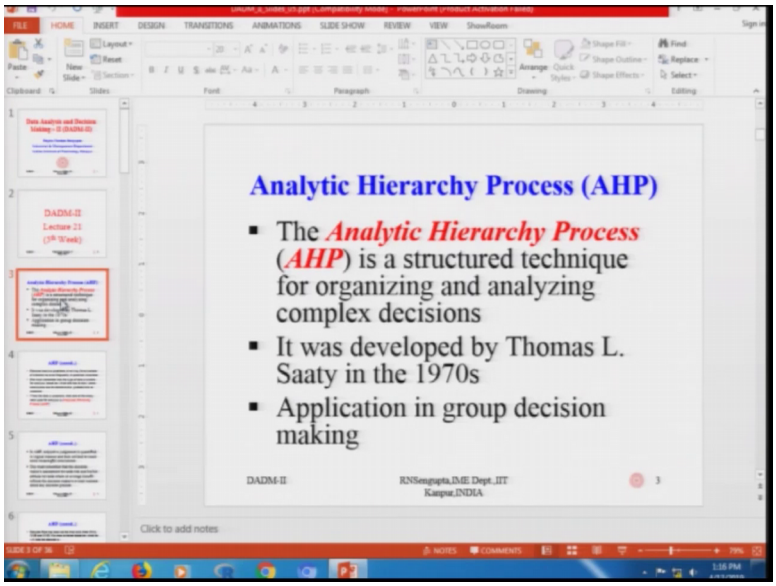
Then we considered the Linux loss function where, overestimation underestimations were can be prioritized and I give you three examples, then we considered very briefly the balance loss function which was considered a part and parcel from multiple linear regression how it can be utilized.

Then we consider the data envelopment analysis in data envelopment analysis, we consider that what is input bundle output bundle what is decision making in unit and how increasing decreasing? Inconstant return to scales will have three different diagrams based on which you can find out the maximization and minimization on the problem which can be converted into linear form. And then we covered the concepts this is I think this is the d a part I will try to.

So, this is the d a part we already considered so, this is no now we repeating it; so, this way gone for quite long means the number of lectures. So, they are basically being saved and given a concept based on the lectures. Then we consider the concept of decision trees, this was need to where we considered the probability of success failure then for the and how you can utilize the concept? The expected value variance to rank the decisions, if both are to be consider how you can take the ratio of expected value to variance and rank them for the highest to the lowest. Or you want to take the ratio of variance to the expected value rank them from the lowest to the highest.

So, we did consider examples accordingly and I did mention that considering that you till the values are negative we will definitely take some of the problems where utility values would be considered 0, in order to compare the models. Even though negative values are possible, then we consider the concept of AHP analytical hierarchy process and how the hierarchies are there, how the properties could be multiplied and how you can utilize the concept of Eigen values, Eigen vectors to find out the ranking consistency ratio consisted ranking? And based on that how you can basically rank the decision process. I give you examples that, if you have trying to buy a house or buy a car they are objective and subjective values also.

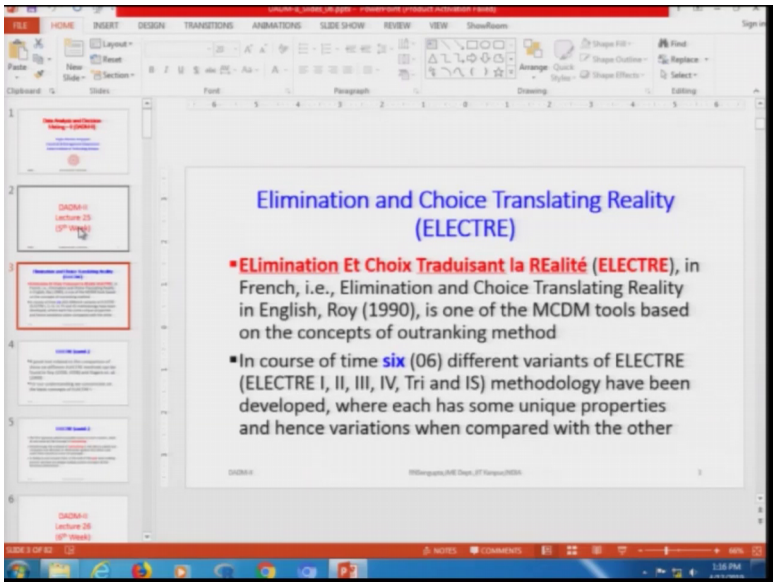
(Refer Slide Time: 19:12)



And in ANN you will should basically consider and the; so sorry analytical hierarchy process you should. Consider the weights to be in such a way that they can be in a scale of 1 to 9 with odd numbers or they can be in the even numbers also.

So, highest value being the liking and disliking means such that if a one is compared to a 2 this is the alternative and if a one gets a point of 9. So, a 2 would basically get a point of 1 minus then; that means, I dislike it and based on that you basically multiply those values in order to get the ranking system. Then if I go to then for the six this numbers which are given are the weak number.

(Refer Slide Time: 19:50)



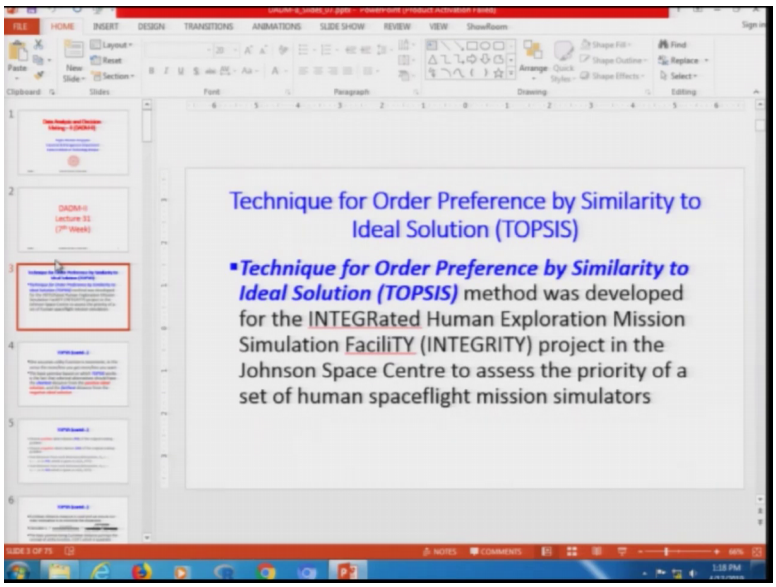
So, then we went to the electre method. And electre method we will only considered one of the methods then we considered the concept of the in electre epsilon electre also, where the liking disliking in the inference sets were there in the initial method or electre method you have the concordance in discordance sets liking and disliking and club them in the sets and basically rank them, and found out the values in the epsilon electre values we basically had the again as I as I mention few seconds back.

Liking disliking in differences such that together basically the constitute the whole universal set and based on that we can do the ranking and you can find out the likings and disliking set and the in difference set based on which we can compare whether alternative a 1 is better than a 2 or a 2 is better than a 3. And you remember that when you are trying to basically combine this concordance discordance in difference set we are trying to take all the criterias into consideration such that collectively they was they would give a overall decision based on all the criterias; when we are comparing alternatives with respect to each other.

And; obviously, here how you do the normalization is important depending on the utility function normalization. If you remember I did mention time and again, if normalization you are doing along the row or along the column remember to check the values is one, because we consistently follow the row one or the column one do that accordingly for all steps point one. Point number 2; I also mention that if you are trying to change the utility

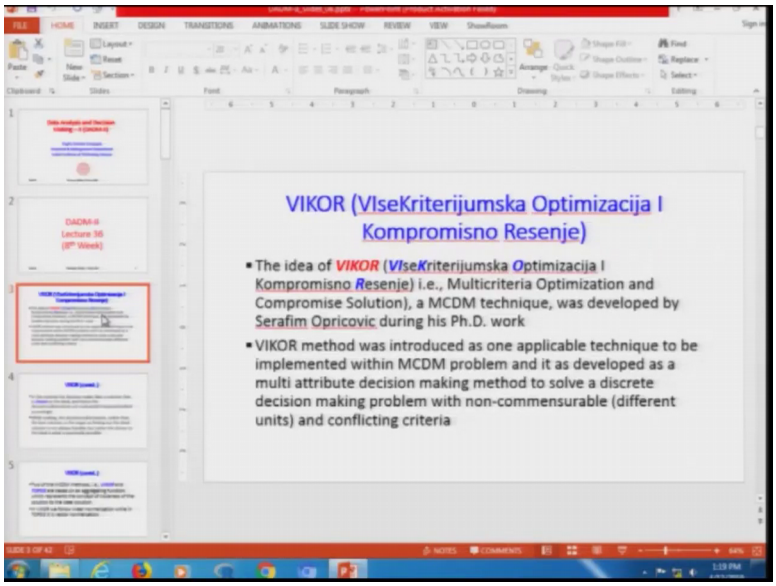
function in between that would not be possible; technically when you are solving the problem because different people may have different utility functions as the age as the income changes as the risk preference changes, but we are not going to consider that in a analysis of the problem. Then late in the 7th week of lectures we considered the concept of TOPSIS method.

(Refer Slide Time: 21:43)



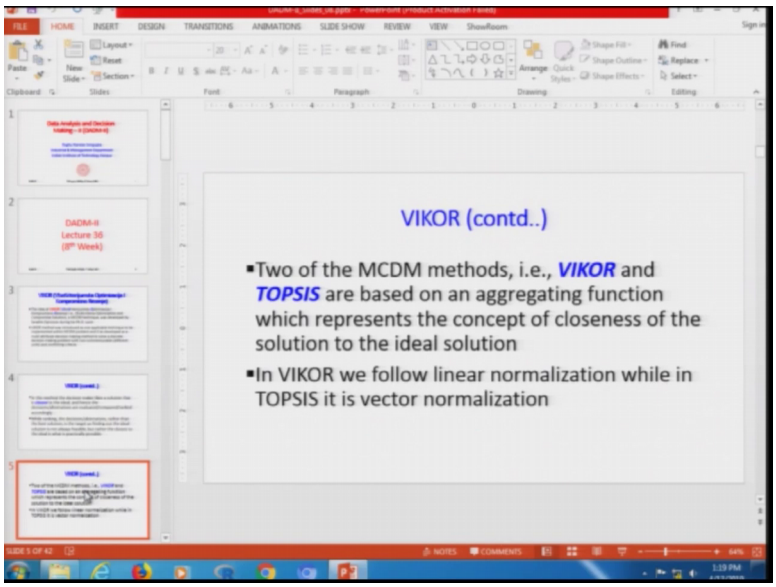
In the TOPSIS method the it was basically the centre of distance of the liking and the disliking. How far it is from the liking point median point is a liking set an and the disliking set, based on that again we differentiated. Again I mention the concept of utility function again I mentioned about, having a weighted metrics which could be normalized based on that we can multiply the priority values or the values which assign for each decision it alternatives and then rank them accordingly. In the 8th lecture 8th week of lecture we consider the VIKOR method.

(Refer Slide Time: 22:21)



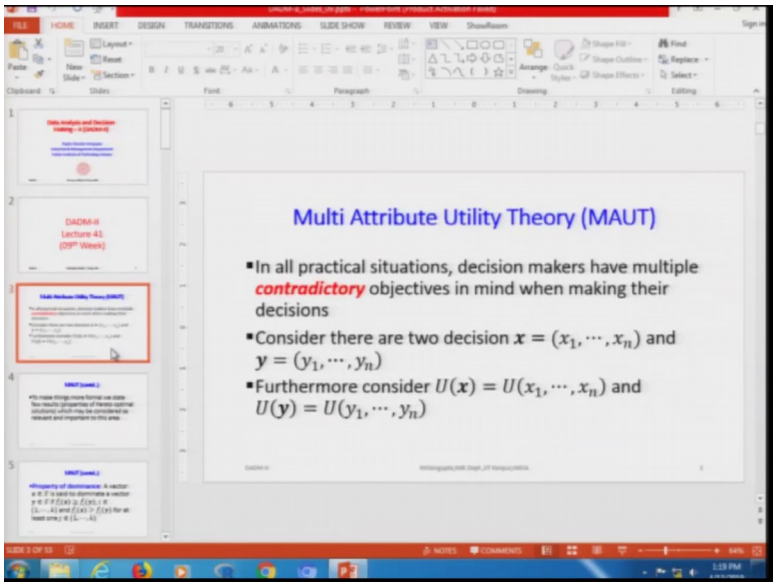
Again the VIKOR method is exactly similar to the electre and the TOPSIS method, but the algorithm based on which you will basically rank them would be different. And I did discuss one of the algorithms in details in one of the method based on that you can build up the models everywhere.

(Refer Slide Time: 22:38)



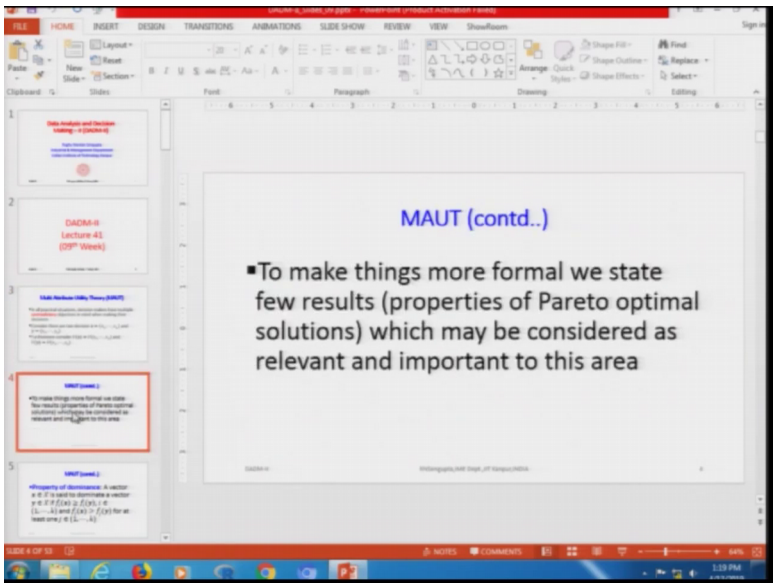
And I also discuss the difference between this multi trade decision making VIKOR and TOPSIS, how the vector nomination could be done and how they can be utilized?

(Refer Slide Time: 22:54)



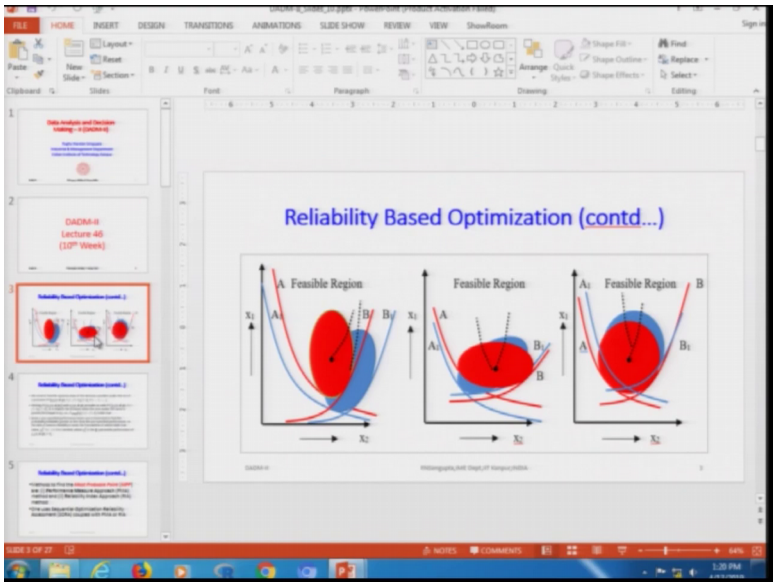
Then in the 9th week lecture we considered MAUT and MAUT; the Multi Attribute Utility Theory or multi attribute.

(Refer Slide Time: 23:01)



The utility is being considered where utilities could be added or depending on the functional form and I did is, here there were no problems is to just a general concept.

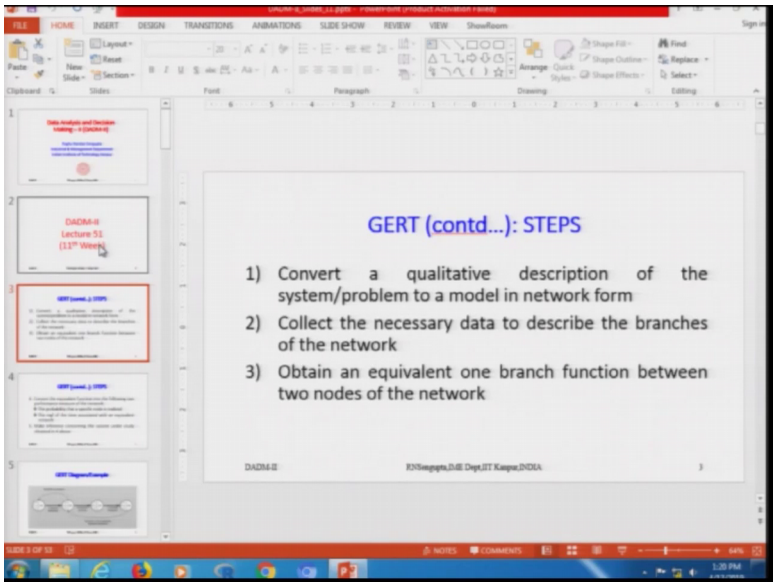
(Refer Slide Time: 23:14)



In the 10th week it was more to do with the concept of, robust optimization reliability and how the concept of utilities going to be utilized or the concept of multi criteria decision making can be utilized from the objective function, it was more of a very conceptual idea. A problems can be solved, but trying to basically develop the problem you have to basically write the code I am not I am did mentioned that I am not going go into the detail of the solution method, it will be considering in the DADM - III.

Then we considered in the 11th lecture we consider the different concept of a GERT method.

(Refer Slide Time: 23:46)



The general evaluation review technique and Q-GERT method and we consider the concept OR network, AND network and how they can be combined to come with the different scenarios. Here looping was allowed where pert and CPM you do not consider lopping.

And the 12th lecture, we considered the artificial immune system the NN method and trying to combine give a example in area of very simple area finance, where bankruptcy prediction was important and they can you can achieve better results using AIS. And in the artificial neural network we considered the prediction and of the change point in foreign currency exchange; considering two different heuristic method with genetic algorithm and simulated annealing.

With this I will like to say thank you all for all your attention and the amount of energy and the time you have spend for doing this course. If you have any queries your most welcome to write it to the tutors and if personally you think you want to further your knowledge in this field you can definitely right to my email id. You can check the net and basically find my email id and write to me; I will; I generally would definitely reply with all the details as may be required. Have a good day and best of luck for your future.

Thank you very much.