Applied Multivariate Analysis

Prof. Amit Mitra

Prof. Shramishtha Mitra

Department of Mathematics and Statistics

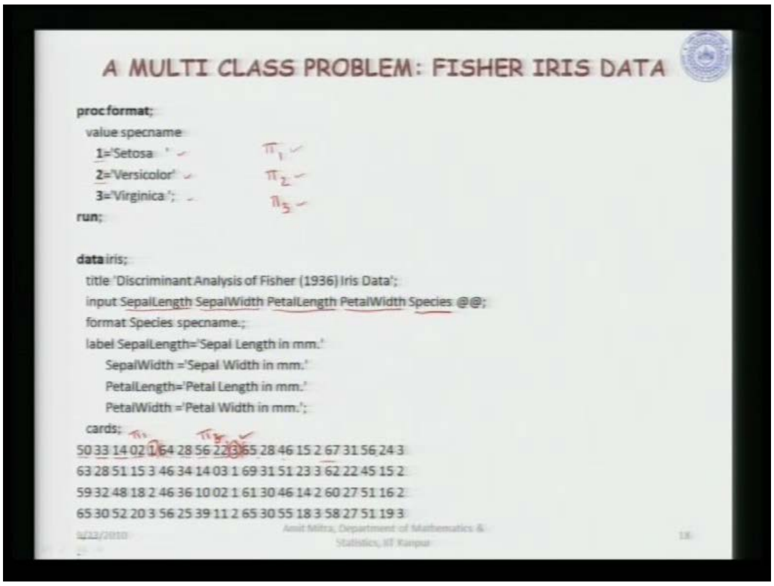
Indian Institute of Technology, Kanpur

Lecture No. # 36

Discriminant Analysis and Classification

In the last lecture, we had started looking at some real life data problems, and was trying to see how classification models can be built on the basis of the theory, that we had learned during the theory lectures. So, we had considered the several problems, couple of problems, when we had two class problems, basically we had looked at that.

(Refer Slide Time: 00:38)



So, today we will look at multi class problems, and look at implementation of the theoretical classification models, that we have learnt for classification in those real life problems. We will also look at an alternate approach of classification which is going to be based on not statistical classification techniques that we have learnt. It is going to be based on artificial neural networks, which provides an alternate way for classification model building..

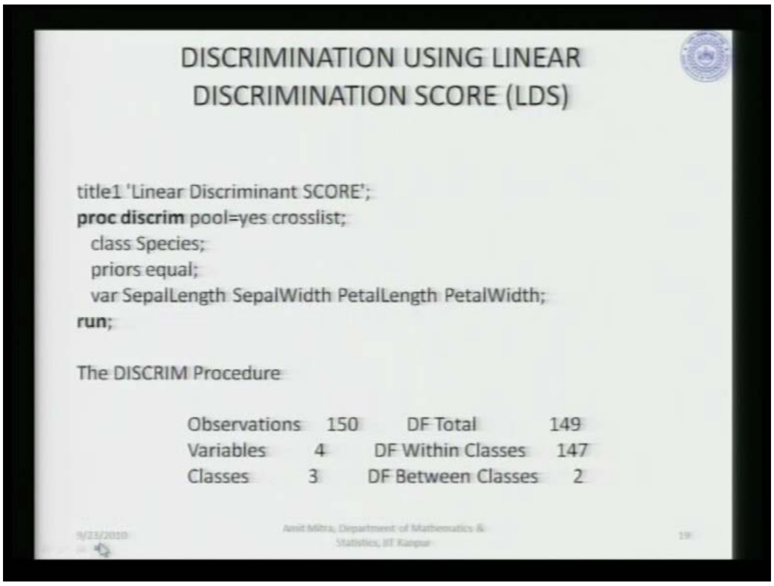
Let us now look at this multi class problems. So, we first have this data which is a very standard data, which is called the fisher iris data. So, it is corresponding to three species: Setosa, Versicolor and Virginica. So, these are the three classes; so, we have three populations pi 1 pi 2 and pi 3; so, these are corresponding to the three different species of this particular data. Now, this is, what is the data? So, it is the input variable the feature vector is comprising of the following variables. We have got sepal length, sepal width, petal length, and petal width corresponding to each of these species; that we observe, and the fifth entry here is the corresponding species identification.

Now, we have made this numerical values corresponding to these species. So, any observation, any feature vector with a particular observation sepal length, sepal width, petal length, and petal width is having a tag of class membership which is one, if it is belonging to this Setosa population, that is the first pi 1 population. It is given two, if it is from the second population, Versicolor; and three for Virginica which is pi 3 in this notation.

So, the data is given in this following form; it is implemented in sass. So, this is the format of the data, that this is the first variable, first dimension of the input space; this is the second dimension, the third, and the fourth dimension so, these are four dimensional data. And the fifth entry, in that order is going to give us the class identification. So, this is the free classified data, the training sample data that is what we are looking at.

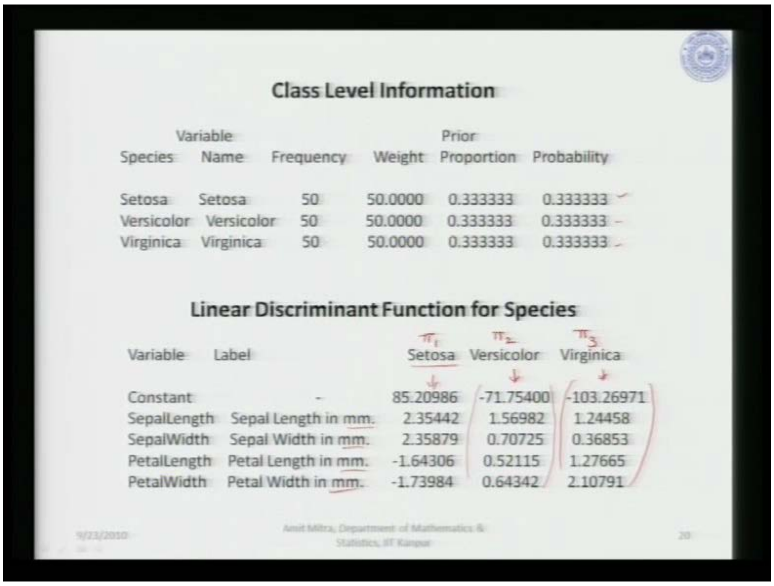
So, similarly, for this is corresponding to the first record, for the second record this is the value of the first variable, second, third, and fourth. So, this is the feature vector and corresponding to this particular feature vector I am sorry corresponding to this feature vector this is the class identification three. So, this feature vector is coming from pi 1 population; this feature vector similarly is coming from pi 2 population, pi 3 population, because it is three here and similarly for the others.

(Refer Slide Time: 03:31)



Now, the problem is, that based on this learning sample data. We are going to build our classification model. Now, as we have seen various different approaches of building such classification model. The first one is when we are trying to look at the fisher, linear discriminant score and classification that is going to be based on linear discriminant score. So, we use proc discrim once again of sass with the class membership as the species, and we take equal priors in the three for the three populations and the feature vector has got these 4 variables right

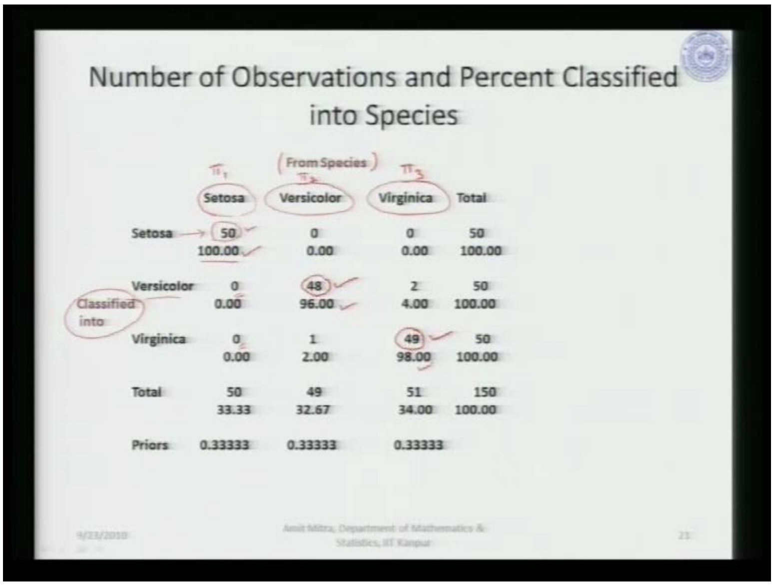
(Refer Slide Time: 04:08)



Now, this discrimination procedure now these are the elementary records what we have. Thus and then, the class level classification what we are looking at this prior probabilities to be equal and hence we are choosing them to be these prior probabilities. These are the frequencies of data in the learning set. So, we take the widths to be equal. Now this is the linear Discriminant function that is building the fisher linear Discriminant function usually under normality assumption.

So, what we have are these are the parameters corresponding to this species. These are the parameters. So, these are the parameters corresponding to the respective variables that we have on this side. Here the labels of course; it is sepal length in millimeter. All these are in millimeter units. These are the parameters linearly combining parameters this with a constant term. So, for the species this pi 2 this was our pi 1 population this was our pi 2 populations and this was our pi 3 populations. So, these are the coefficient vectors corresponding to the respective populations.

(Refer Slide Time: 05:11)

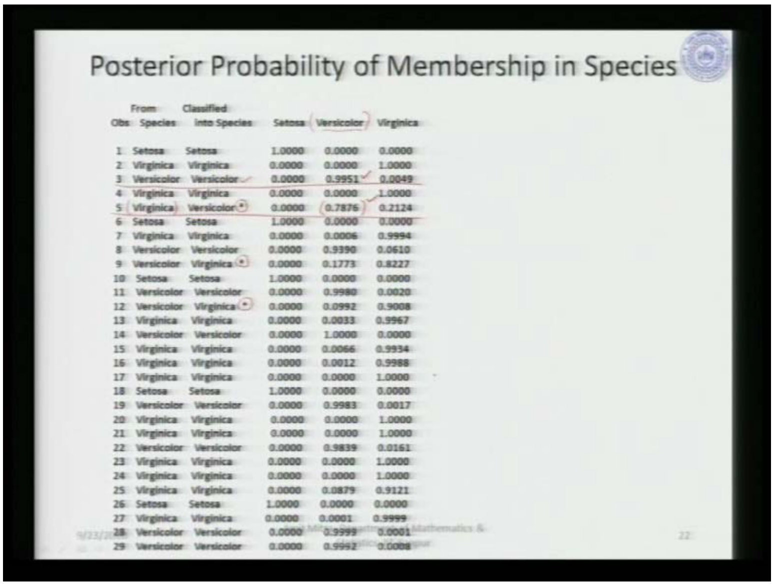


Now, based on these we are going to base our classification records. Now, this is got 2 dimensions that on this side we have got the true memberships. So, an observation which is from this population. It can be from this population and it can be from this population.

Now, let us see where they are getting classified into. So, this is on the basis of the fisher linear Discriminant function. We are classifying 50 observations coming from this species setosa into the class setosa. So, this is a 100 percent correct classification that is what we have, and there are no misclassifications corresponding to observations which are coming from this setosa species here. So, there are no misclassifications 0 and 0 corresponding to this species here which is our pi 2 population. So, these are the 3 populations pi 1, pi 2 and pi 3.

Corresponding to a second population, 48 members coming from this population pi 2 are getting classified into the same class. So, this is the correct number of classifications that we have observed, and we have got one to be misclassified here into this particular class, and corresponding to the third population we have 49 observations correctly classified. So, we see the correct classification percentages is 100 percent, 96 percent and we have got here a 49 percent I am sorry 98 percent. So, this is what is giving us the number of observations and percent classified into the respective species.

(Refer Slide Time: 06:52)

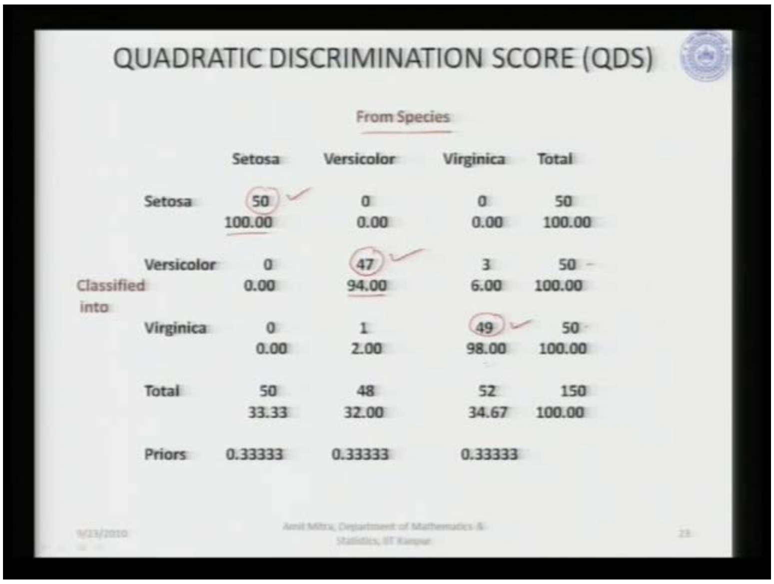


Let me go to the posterior probability of member ship. So, this is this has got the same interpretation as what we had we discussed in the last lecture. So, these are the posterior probabilities now a particular observation vector is going to be classified as to belong to a particular population, if it is posterior probability is the highest for example, if we look at this row of the record then we see that this posterior probability of belonging to this species here which is the pi 2 population is the largest and hence this observation on the basis of posterior probabilities is classified into this class here, and similarly the interpretation.

So, wherever there is a star mark here which indicates that there has been a misclassification corresponding to those records here. So, for example, we had one observations coming from this class Virginica, but the posterior probability we observe that is highest for this versicolor species out here, and hence this observation form this population Virginica is classified wrongly into versicolor species right. So, this is how all these are interpreted?

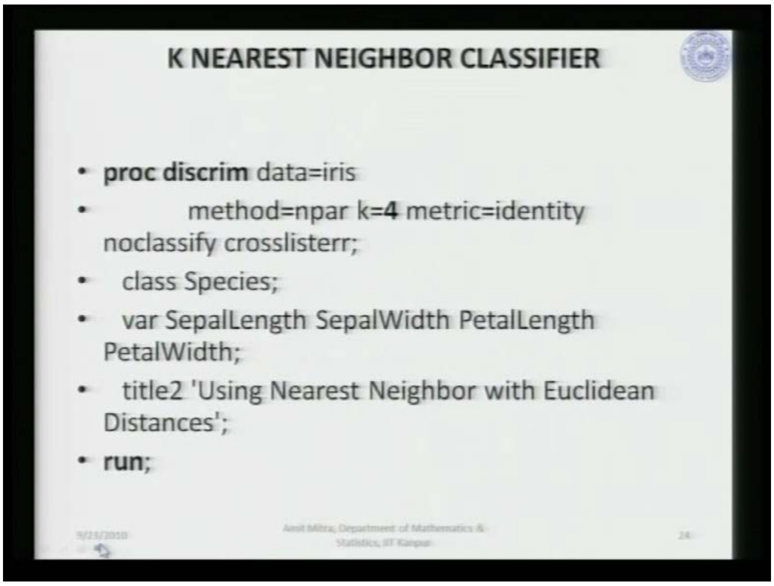
Let us move on and look at what we get for the quadratic Discriminant score? This is once again similar type of table that we had earlier. This is the number of observation and percent classified corresponding to a quadratic Discriminant score. So, this once again is from which species they are coming these are these observations that is what we have.

(Refer Slide Time: 08:22)



So, we have got these observations 50 all of them correctly classified 47 of them correctly classified and we have got 49 out of a total of 50 being correctly classified forty seven out of a total of 50 correctly classified and so on. Once again we observe a very high percentage of correct classifications corresponding to this setup. I think it is marginally it is almost the same. I think comparing with what was that corresponding to linear Discriminant function it is almost what we had for them only.

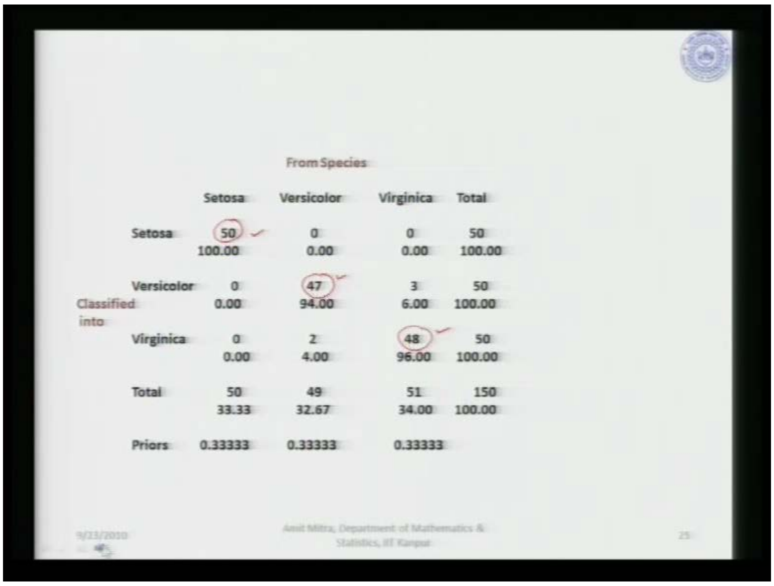
(Refer Slide Time: 09:06)



Now, for the same data, if we adopt another approach of k nearest neighbor classifier So, in the k nearest neighbor classifier what we are trying to do is to look at the input feature vectors and then, look at a neighborhood. A k neighborhood of that particular point in the input space and then, within that k neighborhood we are trying to find out what are the class memberships of the other input feature vectors? that are following within that neighborhood and then, looking at the average of the class memberships of those points which are following within the k nearest neighbor category and that is a non parametric method. So, it has got its own advantages from the point of view of having it as non parametric methods.

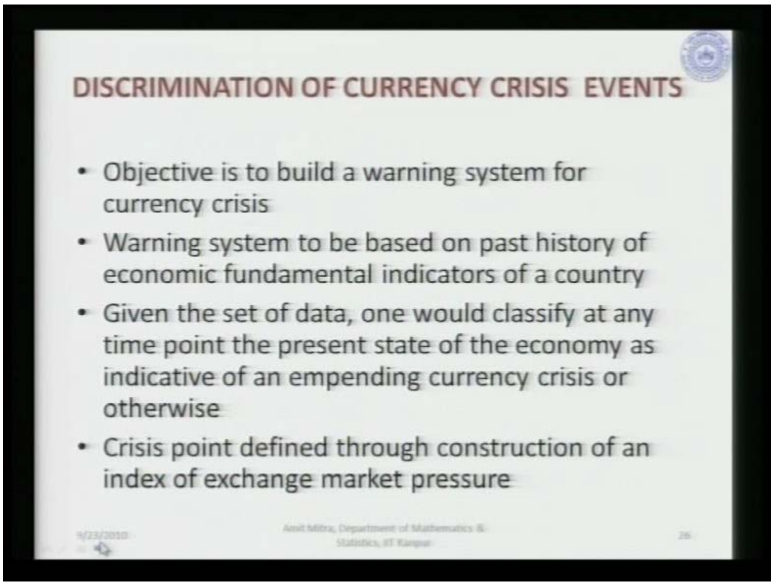
So, we use a non parametric method option from the proc discrim of the sass procedure. So, all other things remaining the same we have got once again class to be the corresponding species of these 3 categories. So, the variables or rather the elements of the feature vector once again are 4 dimensional. We are looking at nearest neighbor classifiers. So, we have got this classification confusion matrix.

(Refer Slide Time: 10:14)



So, these are the correct classification numbers along the diagonals, and the off diagonals denote the misclassified observations. Now once again the performance of the nearest neighbor classifier marginally not as good as may be the quadratic Discriminant function which had 49 here. So, on the correct classification diagonal, but it is performance once again is satisfactory.

(Refer Slide Time: 10:44)



Now, we look at another real life problem which is a problem from the area of currency management. Not exactly, currency management it is foreign exchange management and trying to look at what sort of a classification model? One can build for prediction of currency crisis events. So, it is the discrimination of currency crisis event points. Now, what is the objective of this particular study? Because it is a bit different this example. So, it is better to look at what is the objective? And what the data is all about?

The objective is to build a warning system for currency crisis. So, what one is trying to do is the following? That given the presence state of a particular economy looking at economic fundamentals data at that particular point of time and it is past history. One is trying to build a model. So, as to classify a present time point as indicative of an impending impending currency crisis event.

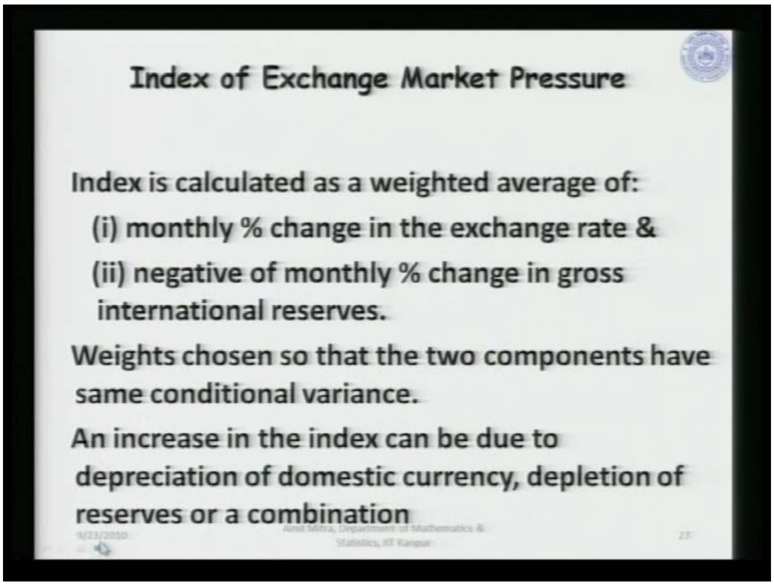
Now, currency crisis event currency from the point of view of foreign exchange reserves. So, it has to be once again defining what we mean by a currency crisis? So, these are rare events; these are extreme events. So, we are basically trying to model some extreme events by saying that at a particular point of time. We are going to classify that state as indicative of a possible crisis or not. So, it is a two class problem that we are looking at.

So, either a present state would be indicative of a future possibility of a future crisis. So, that is may be state one and the other one is not indicative of a future impending crisis. So, that would be the other state which we say that it is the 0 corresponding to that particular variable. So, we have two population problems each time point is going to be classified as coming or rather is indicative of such crisis points or otherwise. So, that is what we are trying to do here?

Now, the warning system is to be based on the past history as I said of economic fundamental indicatives of a particular country. Now this is an event vector which is of interest to any economy as such and hence we are looking at economic fundamental indicators to stand the feature vector space of this classification model right.

So, given the set of data. One would classify at any time point the present state of the economy as I said as indicative of an impending currency crisis or otherwise. So, that is a two class problem that one is talking about. So, either it is belonging to pi 1 which is indicative of crisis or it is belonging to pi 2 which is not indicative of a crisis.

(Refer Slide Time: 13:27)



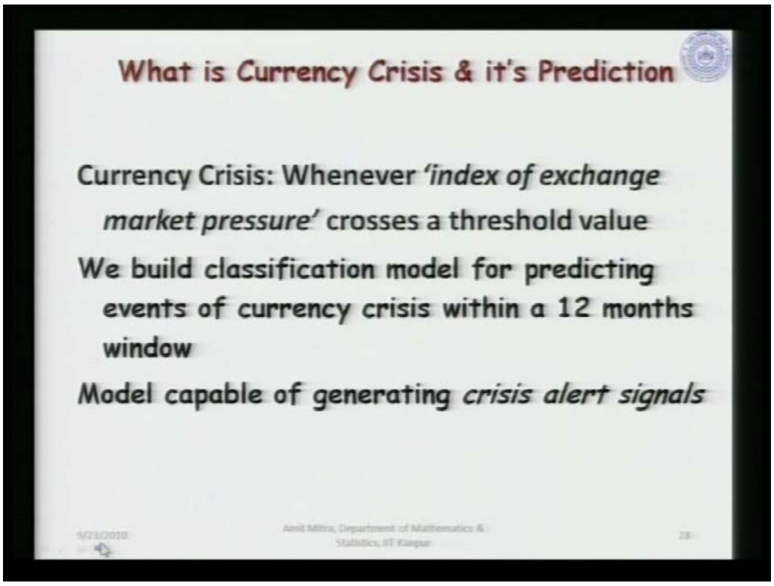
Now, crisis point defined through construction of an index of exchange market pressure. Now, how do we construct such an index for exchange market pressure? It is a very common thing to construct such index for exchange market pressure. Now, a particular country would be in a sort of currency crisis, if two things if two different things can happen either both of them happen or one of them happen in that particular severity. So, this index actually is calculated on the basis of those two aspects of a currency management for an exchange currency management.

So, it is basically one. That is exchange rates which are the say for country dollar, euro dollar, say rupee dollar, rupee euro type of rates. So, it is exchange rate. If it depreciates very heavily then there is a pressure on the currency of that particular country. So, that is going to add to the pressure on the exchange market and hence eventually is going to lead to a situation of a crisis in the currency currency domestic currency with respect to the foreign currency. They can be another aspect aspect of this currency crisis which is in terms of gross international foreign exchange reserves that a country holds. So, if that goes down significantly, then once again there is a pressure on the currency foreign exchange currency and hence it is going towards a situation of a crisis in acute situation of depletion of gross international reserves one is going to have some sort of a currency crisis.

So, an index thus when we are looking at an exchange market index for exchange market pressure is going to be calculated as a weighted average of monthly percentage change in the exchange rate and negative of the monthly percentage change in gross international reserves right and hence this is basically a weighted average of these two quantities if this is too high, monthly percentage change or monthly percentage increase in the exchange rate is too high, and if this goes down significantly, then it is going to have definitely a situation of a currency crisis.

Now, the weights for such a weighted average are chosen. So, that the two components have same conditional variances. So, these two components which are the consistence of the exchange market index actually index for the exchange market pressure. These coefficients here the weights are going to be chosen such that the two components have the same conditional variances and increase in the index thus can be due to either depreciation of the domestic currency that is exchange rate becoming too high say from 48 to 58. So, it is a huge depreciation and such of the domestic currency or land or actually depletion of the reserves or a combination of both right.

(Refer Slide Time: 16:09)



Thus we are going to look at the behavior or the movement of the index for the exchange market pressure and if it exceeds certain threshold, then we are going to call that particular point a point of currency crisis.

So, on the light of the index for exchange market pressure that we have constructed what is thus a currency crisis? And what is it is prediction? So, currency crisis in such a situation is whenever the index of an exchange market pressure as we have defined earlier crosses a threshold value threshold value appropriate to be defined.

Now, what we are trying to do in such a situation? Is to build a classification model for predicting events of currency crisis within next twelve months window. So, as we said that at a particular time point we are trying to predict the state of the economy as 1 which is indicative of a crisis within some time window which is taken in the present example as the twelve months window.

So, given the state of the economy at a particular point of time we are looking at what classification can be done as to looking at that point to indicate some crisis events to happen within the twelve months. Now, usefulness of such a model is important, because usefulness is huge, because if 1 is able to detect such points of currency crisis, then one can take one not one from the point of view the economy or the policy makers of the economy they will actually be looking at controlling the situation. So, that such a crisis event does not actually occur that is why a win prediction win no length of twelve months is taken. So, that enough time actually remains remains in order to correct the situation correct the path of the economy. So, that such a crisis event can actually be avoided and in such a situation we are looking at this model to give us an early warning prediction system right.

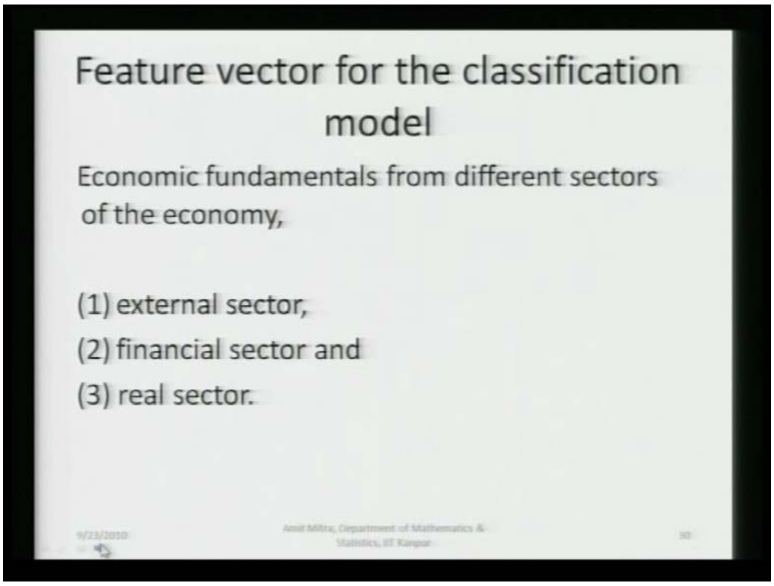
Now, the model thus is capable of generating alert signals. So, at each time point 1 can have this alert signals being generated based on the output of such a classification model.

(Refer Slide Time: 18:21)

Let us see, how we are going to have it now for illustration? We take a particular country which had witnessed huge currency crisis event points in the late nineties which is Indonesia there was a currency crisis in Indonesia Malaysia and other countries there in the south east Asian currency crisis period. So, this is what is the index of exchange market pressure that is constructed from that form that weighted average of monthly percentage change increase in the fore rates and also negative of the monthly percentage change in the foreign exchange reserves. So, these are the points.

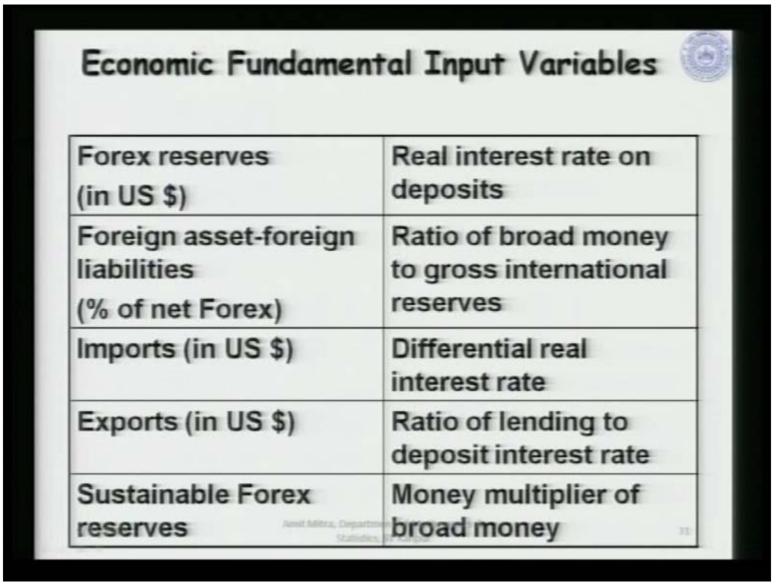
So, accordingly if we choose a certain threshold based on quintiles of this data, one will be able to indicate that well this is perhaps a point of crisis this is perhaps a point of crisis this definitely is a point of crisis these points are points of crisis and So on right. So, based on the threshold one can identify if we choose, say this line to be the threshold corresponding to which we will say, that if the index exceeds that particular point, is going to be called a crisis points. So, after identification of all these crisis points as in here we look at modeling those crisis points.

(Refer Slide Time: 19:37)



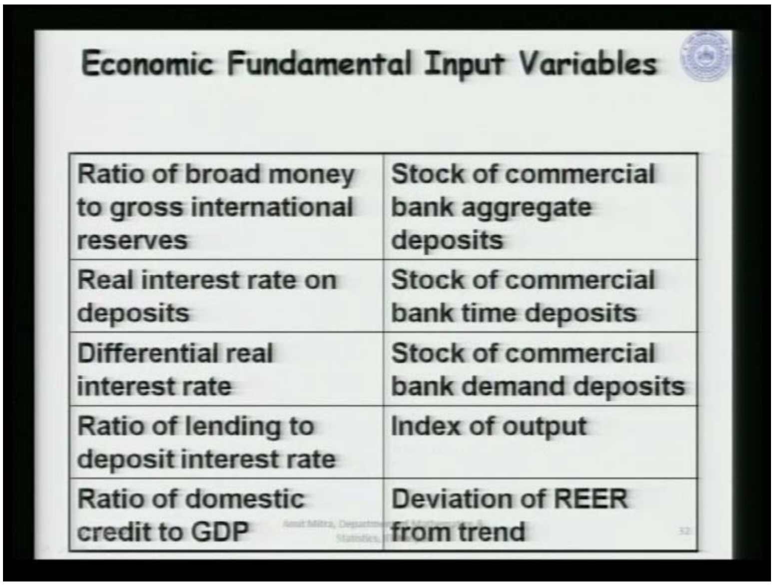
Now, as I said that the input to this particular classification model the feature vector is going to be economic fundamental variables, because a countries health as such countries financial health is judged on the basis of economic fundamental variables and hence we choose here the economic fundamental from different sectors of the economy to represent different sectors of the economy. We choose indicators from external sector. We choose indicators from financial sector. We choose indicators from real sector of the economy.

(Refer Slide Time: 20:10)



Specifically, we use these variables as to comprise of the input features space. So, these are the variable foreign exchange reserves in US dollar, foreign assets minus foreign liabilities imports, exports sustainable for reserves real interest rates on deposits ratio of broad money into gross international reserves differential real interest rates ratio of lending to deposit interest rates money multiplier. So, these span actually more rather go across various sectors of the economy as such.

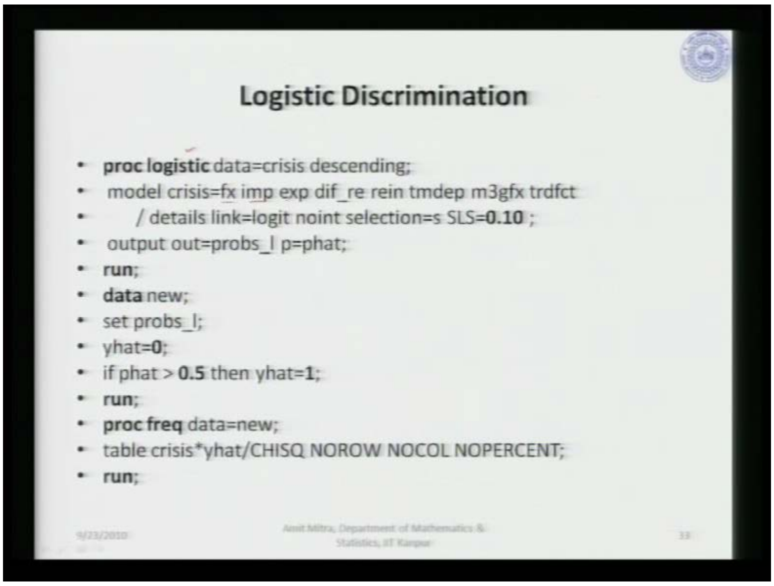
(Refer Slide Time: 20:40)



Similarly, these are the other variables that are chosen huge number in huge number of variables have been chosen. So as to, So as not to miss out any important variables which might affect the future course of the country?

So, we shall broad money to gross international reserves real interest rates to deposits differential interest rates ratio of lending to deposit rates ratio of domestic credit to GDP indicating the output sector. So, stock of commercial bank aggregate deposits time deposits demand deposits index of output once again to indicate what sort of output is there in the economy deviation of the real effective exchange rate? Or the rear from the trend? and So on. right

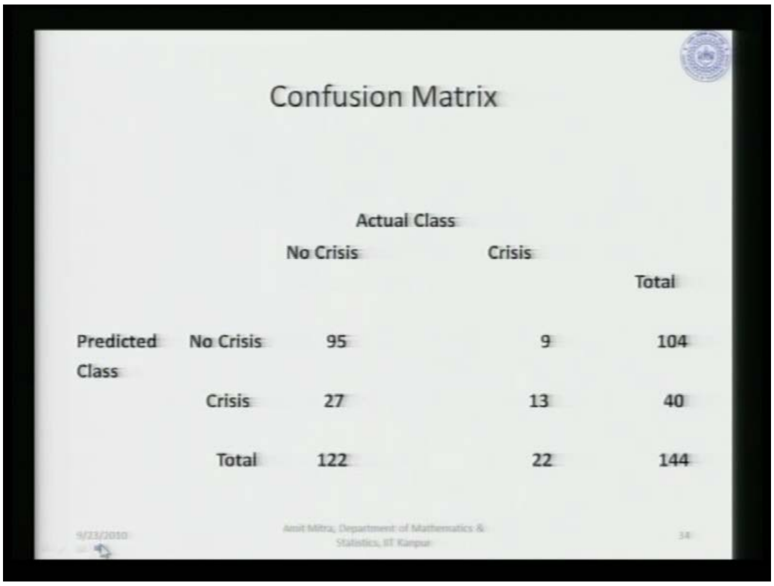
(Refer Slide Time: 21:22)



So, having chosen these as the input variables, we first look at what sort of classification model. We can have for a logistic discrimination based model. So, we use a proc logistic from the sass and then, look at these two spans the input space. So, foreign exchange and all those variables are plugged in here.

Now, we are looking at a stepwise selection procedure for selection of the variables with appropriate criterions and then, the logistic discrimination model is finally, going to be classified on the basis of or rather would be used for classification based on predicted probabilities, now if the predicted probabilities based on a particular feature vector exceeds 0.05 I am sorry 0.5 that is half then, we classified that to belong to population with class membership as one that is indicative of a crisis and if it is less than that 0.5we classify it otherwise that is not indicative of a crisis.

(Refer Slide Time: 22:27)

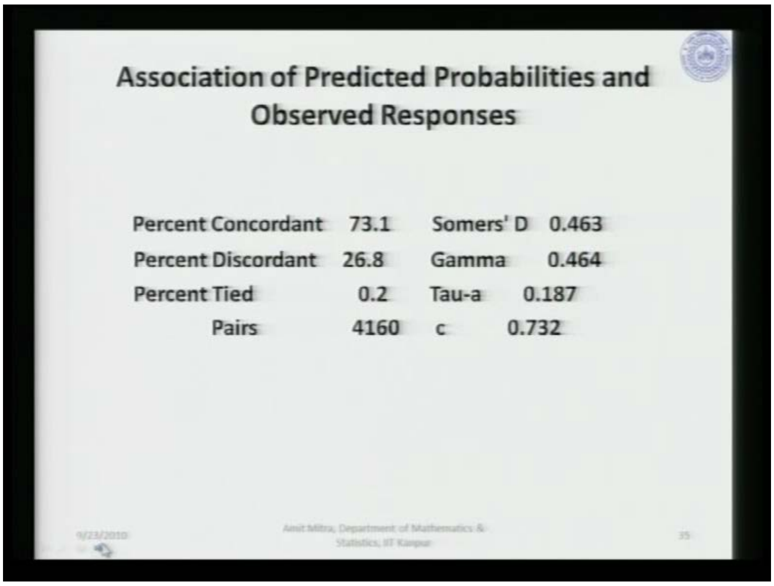


The output of such a logistic discrimination function is a given in the following confusion matrix. So, these are the actual classes now each of the data points are now on the in the learning sample they are classified either as a no crisis point or as a crisis point.

Now, after the prediction model or the classification model is built. So, it can either go to a no crisis point being being classified as a no crisis point or it can go to a point classified as a crisis point. So, what we observe is that 95 no crisis points have been correctly classified. Thirteen crisis points have been correctly classified and there are nine cases of crisis points which are wrongly classified as no crisis points. So, this is a wrong classification that we have done here and if we look at no crisis points being classified as crisis points being classified as crisis points it is huge 27.

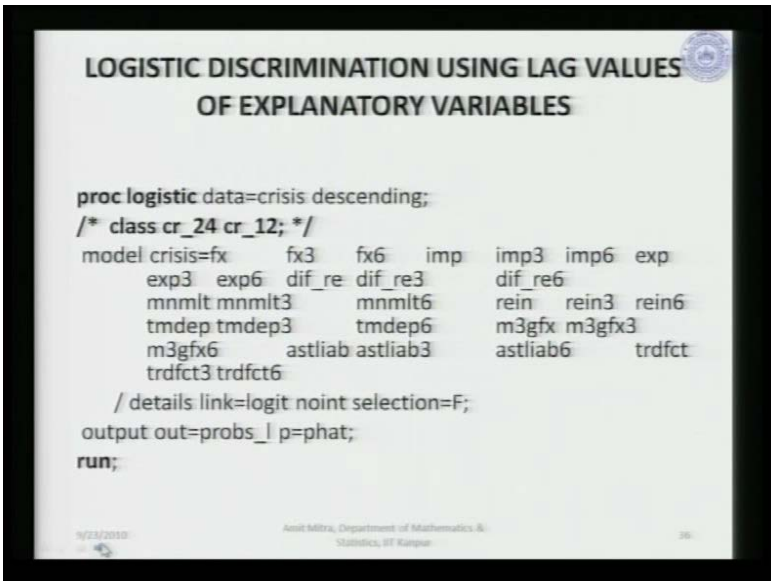
So, this is a is cross this is, because this is misclassification. So, these 2 numbers here indicate wrong classification or misclassification of the present data using a logistic discrimination function.

(Refer Slide Time: 23:47)



So, the results of logistic discrimination function is not that encouraging actually the situation changes little bit now, these are the association of predicted probabilities and the observed corresponding observed responses. So, we have percent concordant just a 73 percent actually percent discordant quite high at 26.8 percent right. So, these are measure of association which are not that high as such to call such a classification model to give us a good classification results.

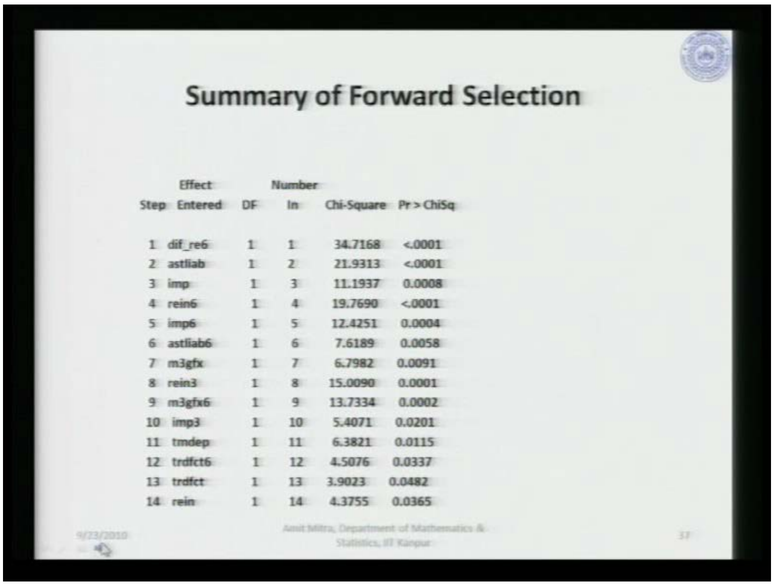
(Refer Slide Time: 24; 14)



Now, form the this simple logistic discrimination we move on to still a logistic discrimination using lag values of the explanatory variable in the previous logistic model. We did not actually consider any lag values. We are just considered the values of the explanatory variables that is the values of the economic fundamental indicators at that particular point of time say t and then, trying to see whether one can correctly classify any crisis events, that is there within next twelve months time period.

Now, here what we are doing? What the difference? We are doing is we are including some lag values of these economic fundamental indicators which are acting as explanatory additional explanatory variables to this particular system. So, we add lags at 3 6 and so on.

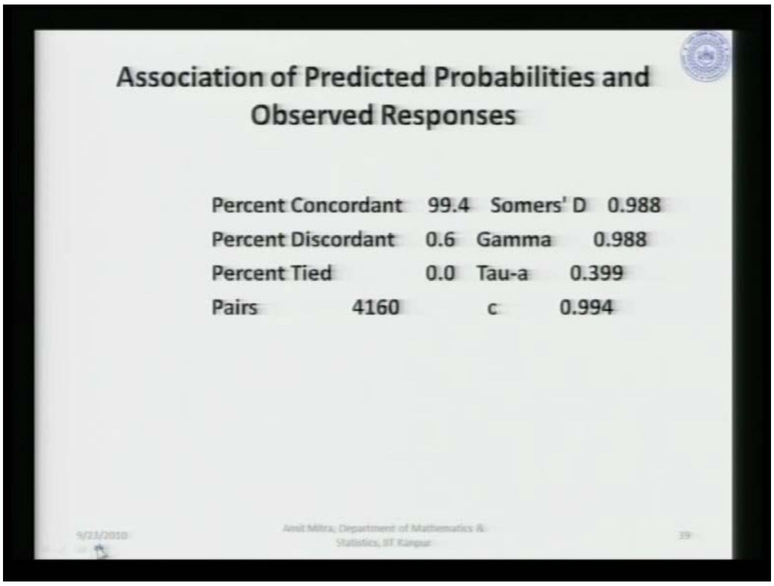
(Refer Slide Time: 25:07)



Step	Effect Entered	DF	Number In	Chi-Square	Pr > ChiSq
1	dif_re6	1	1	34.7168	<.0001
2	astliab	1	2	21.9313	<.0001
3	imp	1	3	11.1937	0.0008
4	rein6	1	4	19.7690	<.0001
5	imp6	1	5	12.4251	0.0004
6	astliab6	1	6	7.6189	0.0058
7	m3gfx	1	7	6.7982	0.0091
8	rein3	1	8	15.0090	0.0001
9	m3gfx6	1	9	13.7334	0.0002
10	imp3	1	10	5.4071	0.0201
11	tmdep	1	11	6.3821	0.0115
12	trdfct6	1	12	4.5076	0.0337
13	trdfct	1	13	3.9023	0.0482
14	rein	1	14	4.3755	0.0365

So, what we observe is that now this is the summary of the forward selection procedure for the logistic regression model and these are the variables that are basically kept after the forward selection procedure, because we have huge number of variable including the lags of various economic fundamental indicators. So, we use a selection method here.

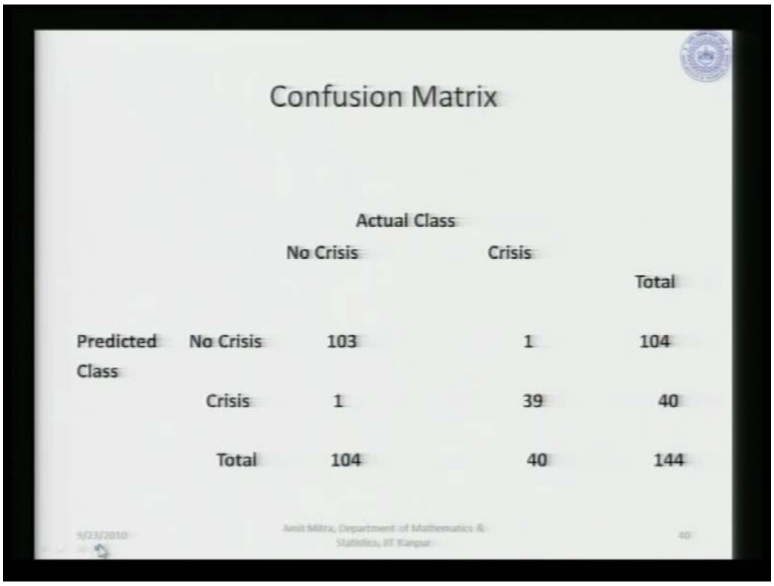
(Refer Slide Time: 25:29)



Now, the results that we get here this are the association table association of the predicted probabilities, and the observed responses. We see there is a huge change marked change in this percent concordant data. We have a 99 percent concordant actually of the predicted probabilities with the observed responses.

Accordingly, once again the measures of association also now shows marked improvement and hence, they are pretty high indicating that the classification is now good. Actually, now the thing actually changes if you look at the confusion matrix, its huge difference than to what we had earlier.

(Refer Slide Time: 25:56)



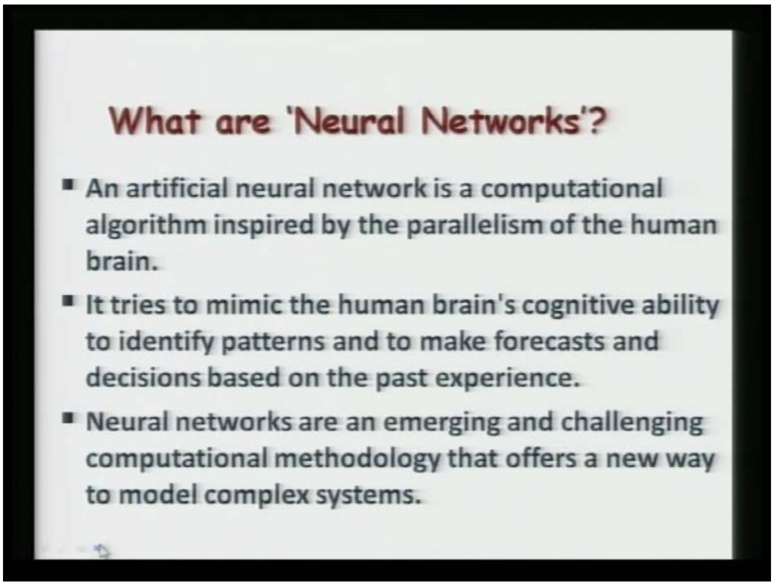
Now, 103 of no crisis points have been correctly classified into no crisis points and 39 crisis points have been correctly classified as crisis points and there adjust 2 points of misclassification. So, it is a huge improvement as to what we had earlier. When we had looked at without including any lag values. So, one basically have to clear on with the data what feature vector one is having and have a proper understanding of what that data basically means have proper proper domain knowledge about the data, and that can vastly improve the performance of the classification model even a simple logistic discrimination model can performance of it can be vastly improved. If we look at including such lags, which for any economists or any analysis is natural, because very many economic fundamental variables can have the effect at a future time point not necessarily the time or rather the value of that particular variable at that time point.

One can actually look at the values of the those economic fundamental variables at a lag three months, six months, they are going to still have say important role to play while building such classification models right. So, this is what we see for this logistic regression model.

Now, as I said that we are in this lecture also going to see how the classical problem of classification to a statistician can be solved using an alternate approach of artificial intelligence? Wherein we try to build up artificial neural network neural network models which can be used once again for the same purpose as what a fisher linear discriminant function or a logistic discrimination function or a k means k nearest neighbor classifier was basically doing.

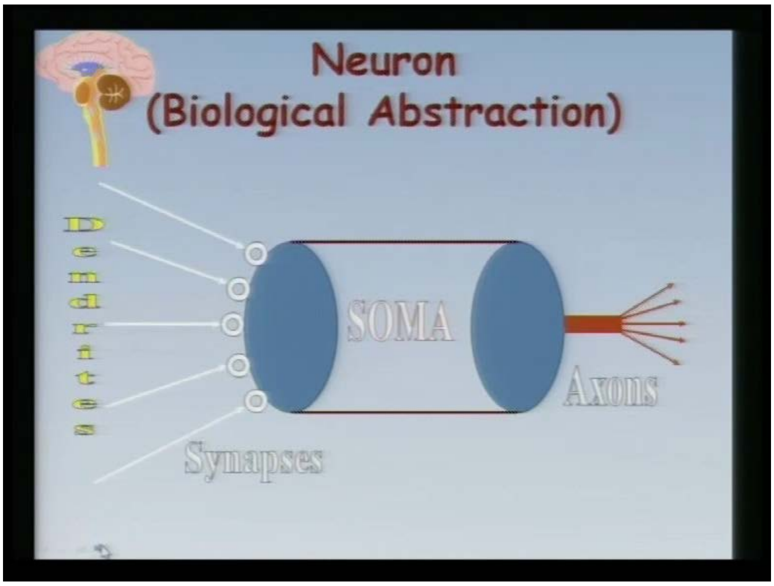
So, one can have the same problem translated to ANN framework and then, try to see if 1 gets the better result. So, a need might arise, that if there are complex dynamic systems such that simple statistical models are unable to resolve this classification problem one can under such a situation have understanding of alternate procedure that can be used in order to have the classification improperly right.

(Refer Slide Time: 28:31)



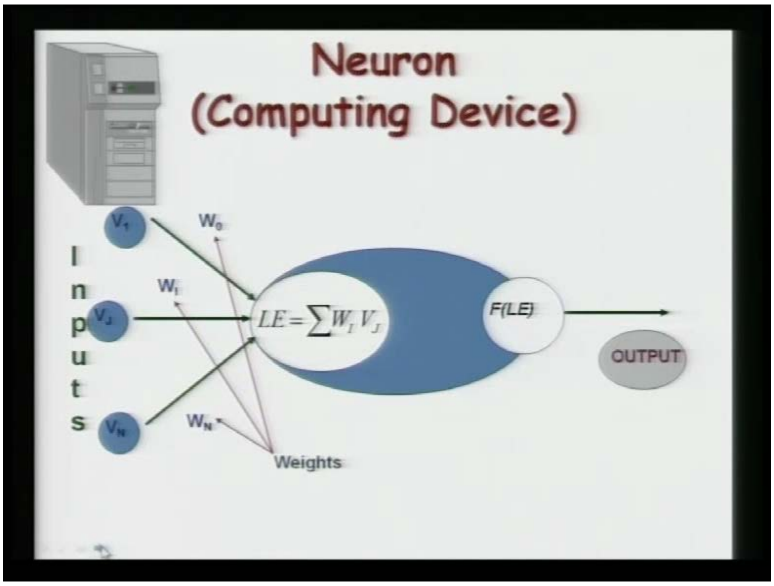
So, here we are trying to build up ANN model for classification now how we are going to do that before that what are neural networks it is a standard thing. So, for completion we just write it here and artificial neural network is a computational algorithm which is inspired by the parallelism of human brain. So, it is basically the parallel processing of the human brain that the artificial neural network system tries to mimic. So, it tries to mimic the human brains cognitive ability to identify patterns and to make focus and decisions based on its past experience the historical data. So, the neural networks of course, it is not only merging it is a well establishing mimic and it is a challenging computational methodology that offers a new way to model complex systems as such

(Refer Slide Time: 29:13)



Now, this is what is a neuron the building block of any neural network system? The artificial neural network system is what is what are called neuron? So, they are the basic processing units of any neural network processing model. So, it has got its name from the biological abstraction actually, when neurons are the basic processing units in the human brain cell also with a cell structure called as soma and axons actually sending out signals with dendrites providing the inputs and a particular time point and these denoting the sanative gap.

(Refer Slide Time: 29:54)



So, remember this particular structure of this figure and this exactly the same structure is going to be replicated, when we talk about neurons or the processing units in a computing device.

Now, wherein a computing device, when we talk about artificial neural network works what we are trying to do is to have sets of inputs. So, these are n inputs actually coming to one processing unit. So, that processing unit suppose, we say that it is the processing unit it is the basic computational unit of an ANN model and the name is still a neuron.

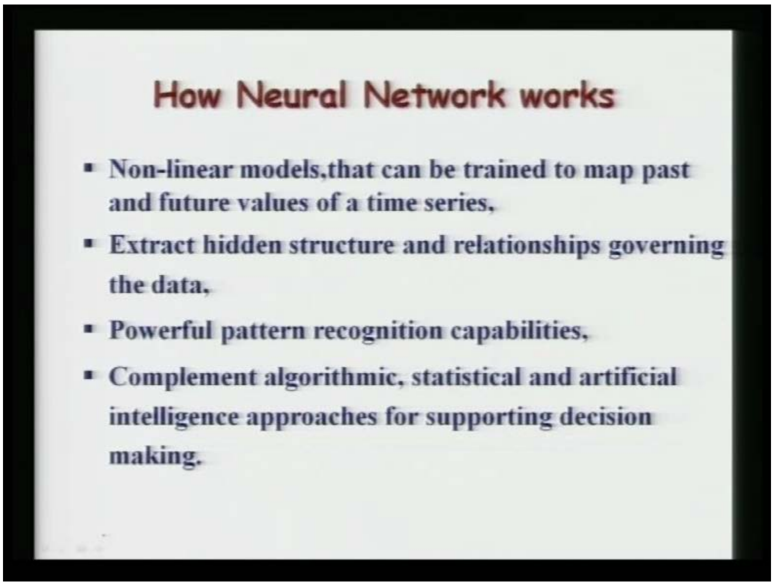
So, we are looking at this two process this information information in terms of these input sets of variables. Now, these inputs come to the processing unit which are of course, multiplied by corresponding weights. Here, which are which signify, this is going to be actually w one and this is going to be w two. So, these are the ways of importance that are attached to each of these inputs. So, it seems for the biological processing of impulses wherein all the inputs may not be equally important. So, one has different weight ages associated with each of these inputs and then, the work of this processing unit is to process this entire information. The values of the input variables along with the weights that are attached to each of those inputs.

So, it combines the input N inputs in terms of looking at a weighted average of these inputs and then, applies a transfer function to process this net input value what we receive at this particular neuron in terms of having some function of the net input value which is arriving at this particular input or neuron or processing unit of this artificial neural network structures, and then, that output that after it processes the input at this neuron sends put an output which is the output of this particular neuron when we talk about an artificial neural network architecture, then we are may be huge number of such processing units arranged in layers in order to lead us to a complete architecture of a neural network wherein, we can have layers of this processing units which are usually termed as hidden layers.

So, we will have those layers of processes in. So, in between the input vector and the output of the neural system, what we will be having? Is series and say a number of such processing units on neurons arranged in layers and then, these parallel processing of the information goes on until we get an output from this particular system.

Now, there are various types of structures of artificial neural network say feed forward networks feedback type of networks and So on. So, it is not ANN class.

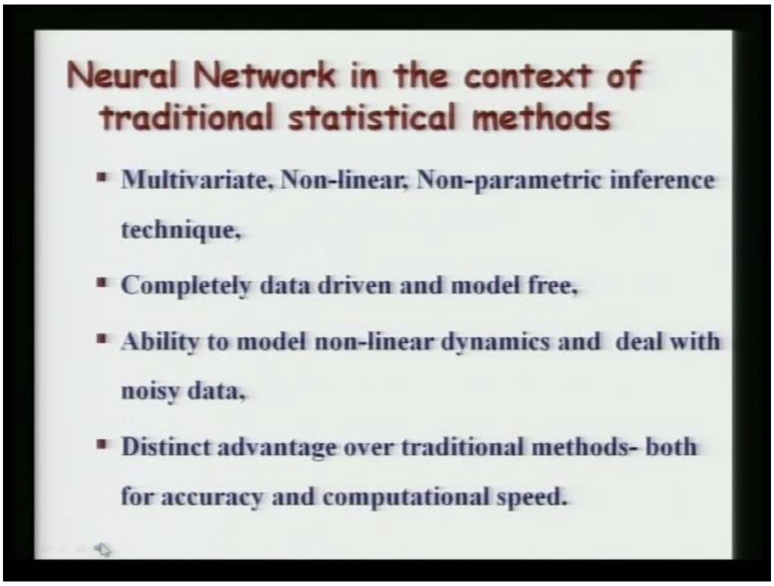
(Refer Slide Time: 32:42)



So, we will not get into the detail of that. Now, how neural network works? It is basically a non-linear model that can be trained to map the past and the future values of the time series type of data or any other type of data.

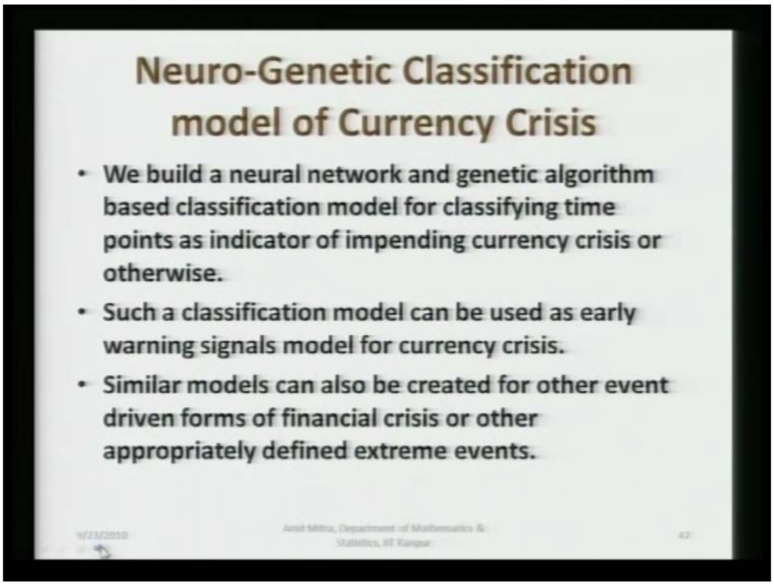
Now, try to extract by doing. So, we try to extract hidden structure and relationships governing the data. So, it is a it is definitely a powerful pattern recognition capabilities of of ANN model there are it is a very powerful technique in that sense. It compliment algorithmic statically and artificial intelligence approaches for supporting decision making.

(Refer Slide Time: 33:16)



Now, what are neural network models as such in the context of traditional statistical methods? A neural network model in the context of traditional statistical methods can be termed as a multi variety a non parametric and non parametric inference technique, that is totally data driven and model free it has the ability to model non-linear dynamics and to deal with noisy data and it has got distinct advantages as the literature of neural network modeling suggests, that it has got distinct advantage over the traditional methods both in terms of accuracy and computational speed.

(Refer Slide Time: 33:52)



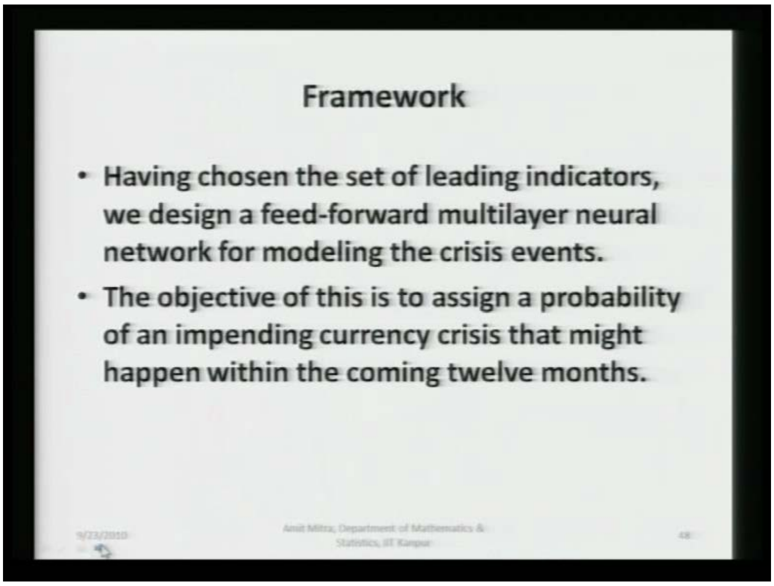
Now, we use the neural network models. We also use a genetic algorithm approach also in order to optimize the structure of neural network models. So, what we are trying to do in the given problem of that currency crisis prediction? Is to build neural genetic classification model of currency crisis.

So, we are still building classification model under a different paradine though it is basically on an artificial neural network framework. So, we are trying to build ANN model for classification and while building the neural network model. We are using a genetic algorithm based search g based search procedure systematic search procedure in order to optimize the architecture of the neural network.

So, what we are doing? We are building a neural network model and genetic algorithm based classification model for classifying time points as indicators of impending currency crisis or otherwise. So, the problem remains the same under a different paradine though now such a classification model can be used once again as an early warning signals model system actually for currency crisis.

Now, a similar models now this is just one of example similar models can be definitely created for other type of extreme events any other form of currency crisis or appropriately defined extreme events extreme events in the sense that one be can have multi class also. So, one can have three types of events say for a loan classification problem one can look at a particular loan application to fall into one of three classes which is say low risk medium risk high risk type of classifications.

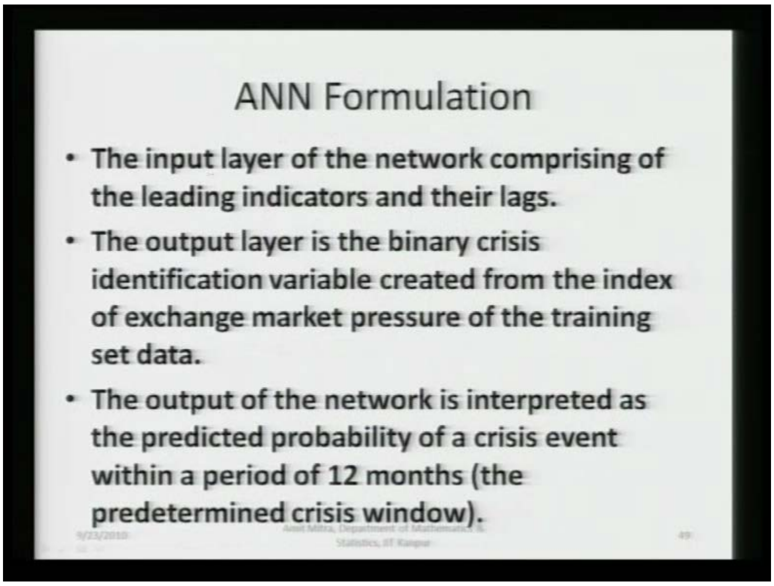
(Refer Slide Time: 35:36)



Once again that type of classification cases can be tackled under such a paradine of ANN model. Now, what is a framework? We already know that what is the what basically are the inputs to the particular classification model system? It is the economic fundamental indicators along with its various lags are the input to the system.

So, having chosen the set of leading indicators. We will have to design a network a neural network. We take here a feed forward multilayer neural network for modeling the crisis events now the objective of this is to assign a probability of an impending currency crisis, that might happen within the coming next coming twelve months. So, the output of the system is going to be once again predicted probabilities. The predicted probabilities based on the values of the economic fundamental variables along with it is lags and it is going to give us an output of the ANN system as the predicted probabilities.

(Refer Slide Time: 36:26)

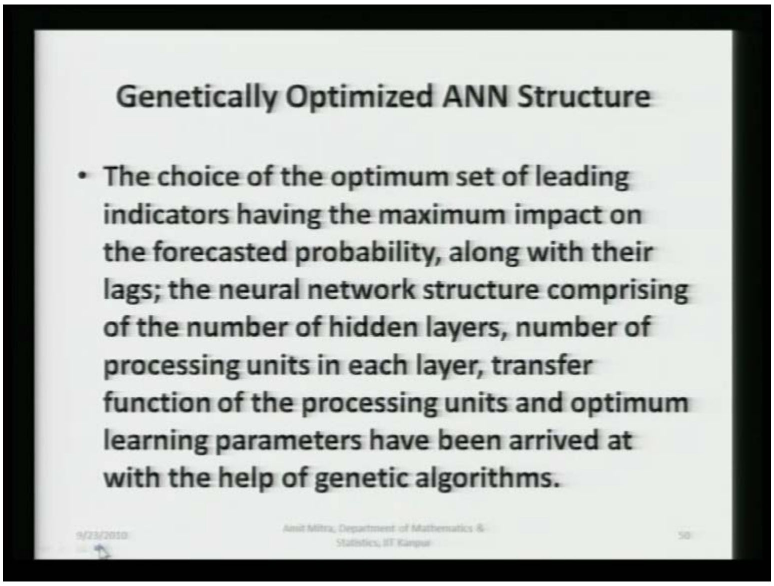


Now, what is the ANN formulation of this problem? This is thus, having the input layer comprising of the leading indicators leading indicators basically at term that is used ,because we are using those economic fundamental variables as leading indicators leading in the sense of those variables providing some insight about a possible future events and as such they are called leading indicators along with a lags.

The output layer is a binary crisis identification variable created from the index of exchange market pressure which exceeding certain threshold exchange market pressure of the training set data. So, based on that particular data records we are basically going to look at what is the binary variable taking the values of the binary variables? say is taking the value 1 or 0 1. If a point is a crisis point and is 0 if it is otherwise and hence, we are creating that binary variable from the index of exchange market pressure by seeing whether that exceeds certain threshold or not if it exceeds the value of the binary variable is one. So, it is going to population, if it is otherwise, it goes to y equal to 0 and that is towards the second population say pi 2.

Now, the output of the neural network is interpreted as the predicted probability of a crisis event within twelve month. We have already discussed this PDF determined crisis window.

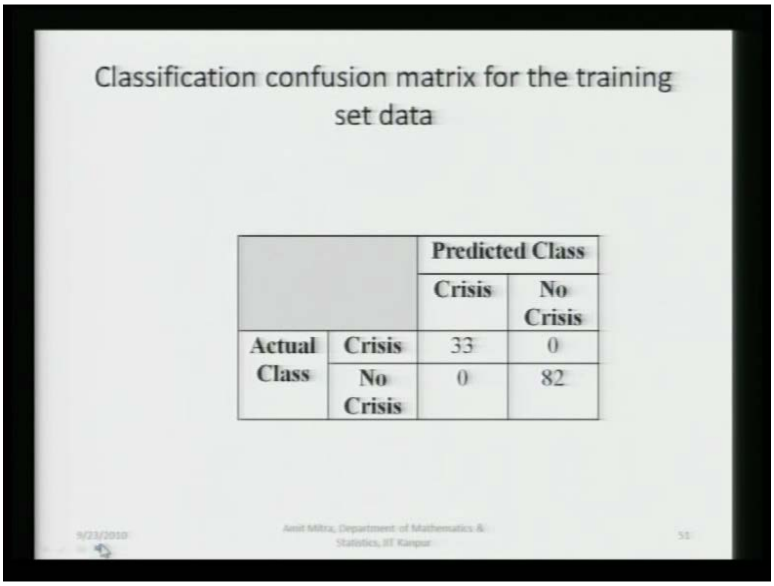
(Refer Slide Time: 37:50)



So, having seen the ANN formulation of this problem classification problem. Here, what we are now going to do? Is that we are going to build that ANN model and as I said that we are going to use genetic algorithms in order to optimize the structure of the neural network.

So, the choice of the optimum set of leading indicators that is the variables that are going to be finally chosen in order to have this classification model in place. So, those indicators having the maximum impact on the forecasted probability along with their lags. So, this is one aspect the neural network structure comprising of the number of hidden layers. The layers of processors and the number of processing units the neurons in each of these hidden layers then transfer function of the processing units which can of course, be different for different processing units. Neurons and optimum learning rate parameters typically when we if we are looking at say a back propagation algorithm for training the neural network, some learning rate parameters learning parameters like learning rate or the momentum rate are associated with such learning algorithm. So, optimum learning parameters have been arrived at with the help of genetic algorithms stochastic search procedures right.

(Refer Slide Time: 39:01)



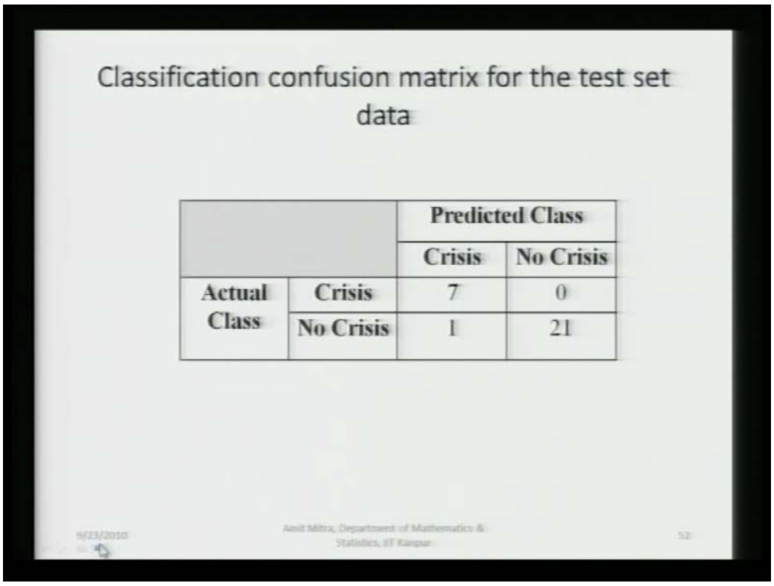
Now, once the model is built, then we look at classification confusion matrix. So, these are the classification results. So, once we have built this particular model here using such a genetically optimized ANN structure with the input and the output variable as formulated the ANN population we when we look at building? The data we actually divide the data into two parts. It is standard when we are looking at a neural network learning system we divide the entire data set into two parts one is termed as a training set data and the other which is called the test set data.

Now, this split of the total data into two parts. Training and test set can be extended. One have a third set of data which is the second test set data called a validation set data. So, what one tries to do? Is to build the model on the basis of a training set data and then, see how the model is performing on the data? Which it has not seen before. So, that is what looks one looks at the performance of the model? What one has built on the test set data? right and then, if one wishes, one can have a third data set which is called a validation set which is once again the a type of test set second test set.

So, the performance of these models has to be had to be judged on the basis of its performance that it is giving on the learning set that is the training set data and the data which it has not seen before that is the test set data. So, this is what we see? We have divided the data into two parts learning set data, learning set or the training set data and the test set data. So, we made a split of if I remember correctly it is around 80 percent ,20 percent split

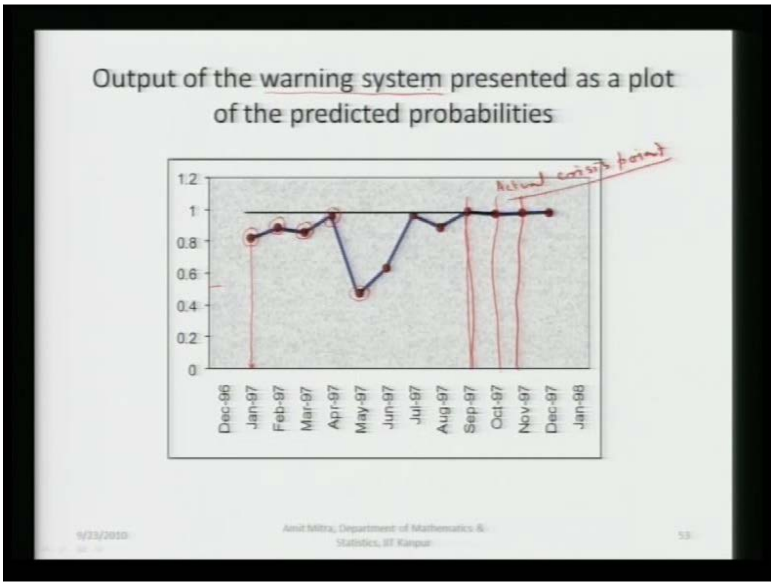
So, 80 percent of the total records that we had was kept in the training set 20 percent of the data was kept in the test set. So, we build the model the classification ANN classification model on the 80 percent of the data and then, use that model and then, work forward to the test set data and see how many in the test set are getting correctly classified with the model? That has been built on 80 percent training set data right. So, what we the results we see here? It is on the training set data. It is a perfect classification on the training set data with no misclassifications zero misclassifications corresponding to such a problem.

(Refer Slide Time: 41:36)



Now, we move on to the test set data. The test is the data which the model has not seen before that is at the time of building the model this is kept aside and kept as a test set data. So, this is a result corresponding to that tests set data we see that the correct classification numbers are 7 plus 21, 28 and only one observation here is misclassified. So, these are the correct classifications 21 no crisis points in the tests set data have been correctly classified into no crisis points and seven crisis points have been correctly classified into crisis points, and only one crisis point was misclassified as coming from a no crisis point. So, the result of course, corresponding to this ANN formulation of this classification problem is showing that it is capable of actually doing a decent or not decent. Actually, it is doing a good job for this classification. Two class classification problem of this crisis identification.

(Refer Slide Time: 42:35)



Now, we say that at the at the beginning that such a classification model when we are looking at this formulation, can work as a warning system for future crisis points and this is what is the such a warning system can look like. So, this is the output we say that of the warning set presented as a plot on the predicted probabilities.

Now, what it is doing? it is that at this particular time point which is this time point out. Here, at this time point the classification model is taking the input of the economic fundamental variables which we are treating as leading indicators leading indicators of possible crisis points and then, with the values of that indicator along with its past lags. What we are now generating? Is a probability that there will be a crisis event in the next twelve months window and one keeps on monitoring the values of those predicted probabilities as in here at every time point on a continuous bases and then, looking at the output of such a system. One is in a position that whether the state of the economy at that point of time is going to lead us to some sort of a crisis situation.

So, looking at this actually one looks at. Now, they were actually severe crisis points round these these times here the lines are not exactly straight, but these are the time points say September 97 October 97 November 97. So, these are the time points of actual crisis.

So, in the cross validation part we see that if these were actually the crisis, actual crisis points, then if one considers the output of the ANN warning system to give us some warning system, then all these points which are having probabilities of crisis prediction within twelve months very high. It is over eight for all these points only one point here was below.

5 below 0.5 all other points previous to the actual points of crisis really had actually given very high probability. Actually, very prominent signals very clear signals of a possibility of a crisis event happening and as such we can thus say that such a system is a definitely an acceptable warning system, when we are looking at prediction of such crisis events. So, it is basically it is just a classification model and then, the output of the classification model is taken in terms of having the outputs as some sort of giving warning signals to such extreme events.

So, we will end the this particular section on Discriminant analysis with this lecture, and we will look at the two other remaining concepts in this class in this course which we are going to look at factor analysis, and which is the concept of the canonical correlations

Thank you