Neural Network and Applications

Prof. S. Sengupta
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 35
Introduction to Self Organizing Maps

This is the topic for today is, Introduction to Self Organizing Maps in short form they are referred to as SOM. So, this is new chapter that we are going to learn in this lecture. We all ready had an idea about the principal component analyses and we had seen that, how the principal component analysis can be realized using the neural networks using the linear neural network model.

Where, we had seen that it ultimately evolves into the determination of the principal component. So, if there is only one neuron at the output, then that results in the first principal component and if, there are more then it extracts the 2nd, 3rd and so on, of the principal components which is, very much useful for the case of data reduction. Now, this in effect is a sort of an unsupervised learning, because we are not providing there any desire. Or target response everything evolves to the repeated presentation of the patterns, which is basically a self learning or rather to say an unsupervised learning that takes placed.

Now, we had all ready touched upon the topic of competitive learning. When we were discussing about the different learning mechanisms, where I think if you recollect that what we did there, we had taken several neurons. And we had presented and those neurons are connected essentially, they are fully interconnected to the inputs. And there the out of the neurons, which are there at the output they are competing amongst themselves in order to determine the winner.

And those learning mechanisms, we refer to as the winner takes all mechanism in the sense that, whoever is the winner the neuron which emerges as the winner, all the synaptic weights are adjusted in favor of the winning neuron.

So, that if the same pattern is presented once more, then it wins the competition then it has got a great a chance of winning the competition. In the sense that, even if not the same pattern even if you feed a pattern that is, very close to the pattern that causes the winning of the neuron even for that, pattern the chances of they that neuron winning the competition improves.

So, that was the winner takes all mechanism, now the self organizing maps that we are going to talk about also works on the principal of competitive learning. But then why a it is being treated differently from that of the competitive learning as, because it although each uses the competitive learning mechanism. But here, there is a kind of organization, organization means the there is a special organization, that we are talking of in the distribution of the neurons.

So, essentially here we are talking in terms of a lattice of output neurons, the lattice that can be arranged as either one dimensional lattice or two dimensional lattice or even higher dimensional lattice space, the neurons will be organized. Although for all practical applications we normally make use of one dimensional and two dimensional lattices and higher dimensional lattices or not that popular. Because of the complexity that is brought about, but essentially by organizing the neurons in the in a very structured manner, in the structure of a lattice.

If we now present, if those neurons are connected to the inputs in some manner and then, we feed the input patterns then those input patterns will be actually acting as a stimuli to those neurons, which are there are the outputs. So, that when thus stimulus is present, then out of the different neurons that are existing in the lattice, one of them will be the winner.

And the synaptic connections, from the input layer to the output layer will be adjusted in such a way, that the synaptic connection will move the synaptic connection will be move. The weight updating will take this in such way, that the Euclidean distance between the input vector and the weight vector that is minimized. So, a the minimization of the Euclidean distance effectively means, the maximization of the w transpose x output, that we will be seeing otherwise.

So, now what happens that one of the neurons will emerges as a winner and, we are feeding different patterns from the input space. We will be feeding various types of input patterns to the systems, to the system. And then depending upon the input distribution, because the input distribution is actually not a uniform distribution, the inputs will be distributed in some random way throughout the input space, but if we start with a regular lattice structure depending upon our input statistics.

The ultimate organization of the lattice would be slightly different, but the ultimate organization of the lattice that results, would be indicative of the statistics of the input pattern that, we are applying as a stimuli right. So, this is the essential philosophy, so that means, to say the that there will be a weight adjustment and essentially. The synaptic weight adjustment that takes place, will physically bring you can imaginate this way that as if to say, that it will disturb the lattice structure and it will moves the wining neuron physically to the input.

The adjustment of the weight vector that, takes place should be such that it is moved closer, so in otherwise the lattice that ultimately results will be indicative of the statistics. In fact these aspect will be more clear, will be take up the specific examples of the self organizing map, because first of all we will be explaining the different models. In fact, there are two popular models of self organizing maps, which we will be talking about that is, lattice Von-Der Malsburg model and then, the Kohonen model.

And of these Kohonen model is the more general one, since it permits data reduction and that why we will be dealing move with the Kohonen model, or what is known as what is more popularly known as Kohonen self organizing maps. So, very often the literature talk about this self organizing map as Kohonen SOM, or they call it as Kohonen competitive learning. Basically they are one of the same as referring to the self organizing map based on Kohonen model, that also we will be discussing.

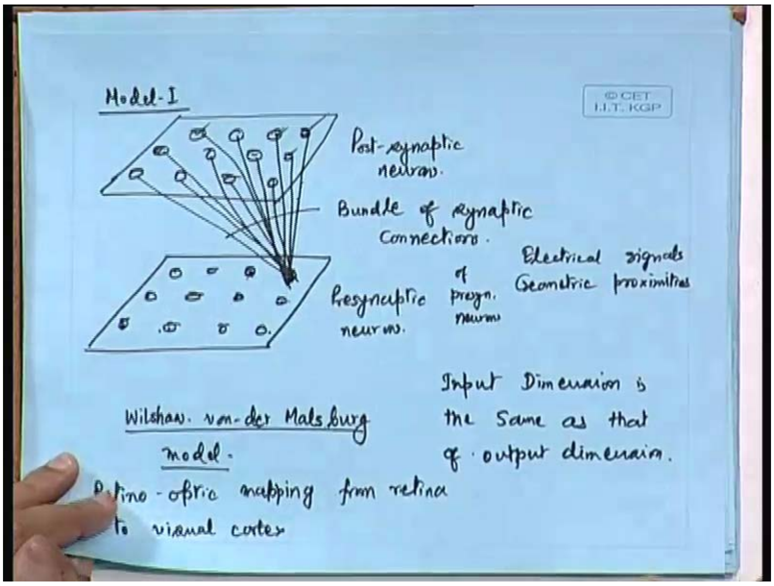
And In fact, another point which should be noted at this stages that that the neurons, which are there at the output they act in a competitive manner in the sense, that they inhibit the responses of each other. But what happens is that, like it is also neurobiologically inspired one fact exist that. The neurons, which are close to the winning neuron they tend to behave they tend to have an excitatory response that means, to say that around the winning neuron and excited to the response is generally created.

Whereas, an inhibitory response is created for the neurons which are there for distant apart that means, to say that this exhibits this sort of networks will exhibit, long range inhibitions and short range excitations. So, these mechanisms also we will be explaining, but before going into that there is one very important aspect, which should be born in mind and that is, to say that what actually motivated the development of this self organizing maps.

In fact, it is neurobiologically motivated why, because you see our brain receives the signals from various sources of inputs, some inputs are visual which are obtained from our eyes. And then, ultimately obtained through the retinas, the responses of the retinas they are feed to the brain. Then the acoustic responses tactile, so acoustic inputs, tactile inputs, visual inputs all these different sensory inputs, that we are processing it is seen that, they are processed in different regions of our cerebral cortex in the brain. And since, they are actually processed in different physical regions, one can really believe that there is a topological ordering of computation that takes place.

And, because of that it inspired, the topologically ordered computation was neurobiologically inspired and so, the field of the self organizing map developed inconsistent with the neurobiological model. So, let us first see the different models that exists, the two popular models of the self organizing map as I had mention a brief while back are the von-der Malsburg model and, the other is the more general Kohonen model.

(Refer Slide Time: 13:00)



So, let us see the structure of these two models now, this von-der Malsburg model actually is called as Willshaw von-der Malsburg model. Now, here the structure is like this that we will be having an array of pre-synaptic neurons, and an array of post-synaptic neurons. So, let us draw two arrays, now this an array of pre-synaptic neurons and this is, array of post-synaptic neurons and here, these are the elements of the array and pre-synaptic neurons also will be having similar arrays.

Now, let us take any one of the neurons belonging to the pre-synaptic layer, and it will be connected to the post-synaptic layer fully interconnected. In fact all the postsynaptic neurons will be fully interconnected to the pre-synaptic neurons. So, that we can draw the connection from, let us say take this neuron it will be connected to all the neurons which are there in the postsynaptic layer, so we can draw the connections.

So, this essentially will form a bundle of synaptic connections, so this will form a bundle of synaptic connections, in fact this model was used to explain the retino-optic mapping, so this very nicely explains the retino-optic mapping from retina to visual cortex. So, retina will form the input layer or pre-synaptic layer and visual cortex will form the postsynaptic and they, are interconnected to each other to a bundle of synaptic connections.

So, this what Willshaw von-der Malsburg model says, now here one point to notice that the dimension of the input and the outputs they are the same, so input sdimensions, so here input dimension is the same as that of output dimension. Now, one point which should be noted here, is that the basic idea of this Malsburg model is for the geometric proximity of the pre-synaptic neurons. So, we have to consider the geometric proximities, in fact the electrical signals they are based on geometric geometrical proximities.

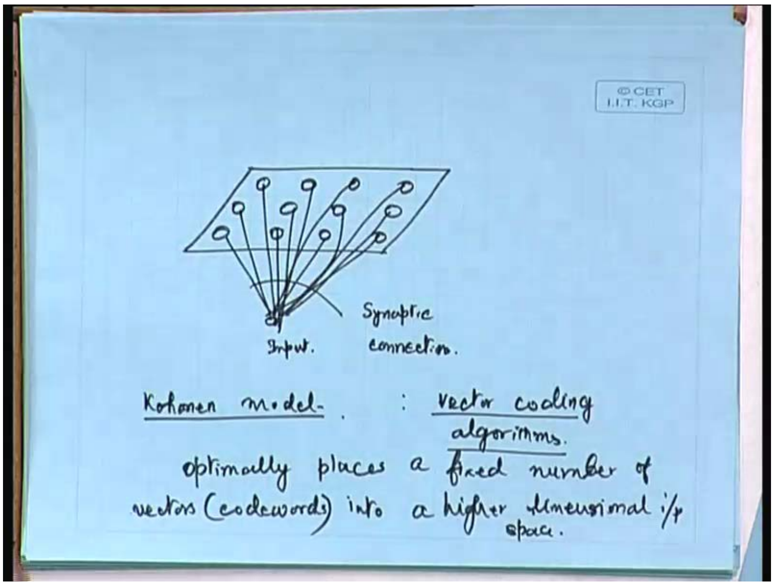
So, the electrical signals of pre-synaptic neurons, they are based on geometric proximities in the sense that, if this neuron, this pre-synaptic neuron is geometrically close to these neurons. Then, they are electrical signals also will be highly correlated in the sense that, here the electrical responses that will emerged from this will be somewhat similar, to this it will not be absolutely uncorrelated they are highly correlated. May be that, when we come to think of the neurons lying here, which are at a distance from these neuron which are not geometrically close, there the electrical signals can be different.

Now, what happens as a result of this is, that you see it is I have shown the connection from one input to all the outputs like this, all the inputs will be connected to all the outputs, but let us now take the case of some post-synaptic neuron, let us say the post-synaptic neuron out here. Now, here you will be finding that all these responses will be strengthened, because this one if this neuron is active.

Then all these connections will be strengthened and again, this is connected to all these, so if this signal is strong even this signal is going to be strong, so that ultimately specially it will enhance, those neurons which are there in the similar special location. In the sense that, this will ultimately fire this neurons will have, a chance to fire this neurons more than firing those neurons, so that ultimately there will be a special correlation of activities.

So, that activities which are exciting here will be ultimately mapped into the similar neuron activities at the post-synaptic layer as well. This can be shown through mathematical analysis also, but we can going for a more generalize model as proposed by Kohonen. So, this you can call has the model number 1, that is Willshaw von-der Malsburg model.

(Refer Slide Time: 19:52)



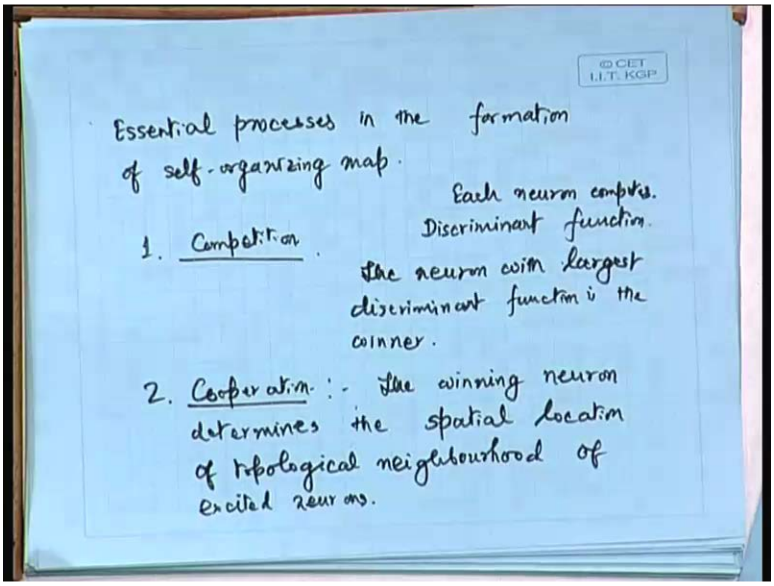
But, model number 2 is by Kohonen, where you can think of an output layer only, only the output layer is organized in a lattice, let us say that it is organized in a two dimensional lattice although some times, we going for a one dimensional lattice also and let us say, that we have got and input lying over here. Now, this in input need not be connected in the form of a lattice, the input could be unorganized also, but the thing is that this inputs will be, you take any input that will be connected to all these outputs.

And then, there will be a bundle of synaptic connections, in this case actually, it is possible to half the data compression, because the number of inputs could be less than that of outputs, in fact what one can do is that, this is what the Kohonen model is, so this is based on the Kohonen model, where it belongs to the vector coding algorithm. This is used for the vector coding algorithms, which optimally places a fixed number of vectors into a higher dimensioning input space fixed number of vectors.

So in fact, those vectors will be used as code words, it is like those who are familiar with the data compression techniques like say entropy coding will be knowing, that if you have a higher dimensional input. You know that, if you have got a higher dimensional input let us like say for example, a code word of an input of length l, let us say could be actually compressed into a code word of length much less than l.

There by doing a data compression, where we will be exploiting the entropy that is present in the data, that is how people do the Huffman coding and all this things. Now, basically a fixed number of code words, those code words can form a vector. So, the this is Kohonen models, in fact can be used in the vector coding algorithms also, where these kind of fixed number of vectors will be placed. And we will be seeing that later on, but there are some essential processes, that one has to fulfill for the self organizing maps.

(Refer Slide Time: 23:23)

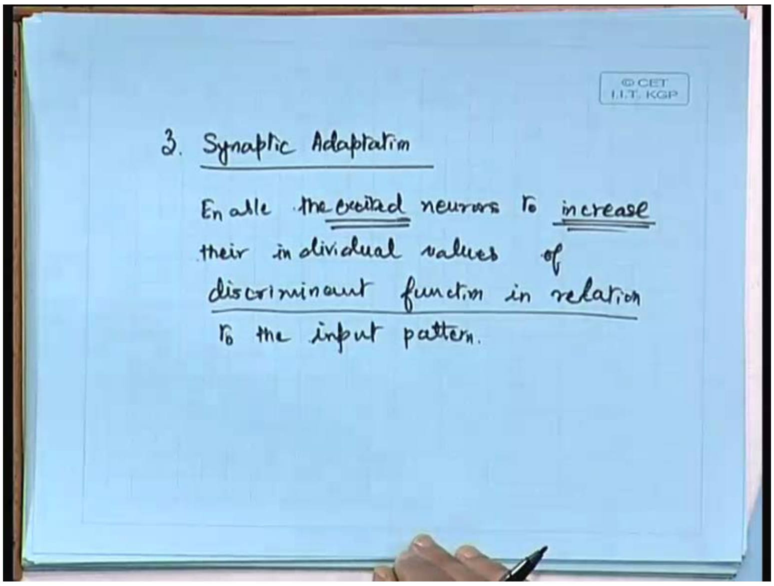


So, we can talk about some of the essential processes in the formation of self organizing map, here the first process is what is called as the competition, now competition basically means that for each input pattern the neurons in the output layer, they will determine the value of a function, that function we will be calling as, the discriminant function. So, each neuron will compute a discriminant function, so each neuron computes the discriminant function. And this function, provides a basis of the competition and the particular neuron with the largest discriminant function, so the neuron with largest discriminant is the winner.

And then, next comes the step of cooperation, now the winning neuron that determines the topological neighborhood of excited neuron, because as I was telling you that the neuron that wins the competition that excides, the neighboring neurons surrounding it. So, that the winning neuron will determine the special location of topological neighborhood, so it determines the topological neighborhood of the exited neurons and excitation is obviously, a cooperation, because what happens is that it not only strengthens the neuron which is the winner.

But is it, but it also strengthens the neurons, which are closer to it whereas, it by the process of competition, the neurons which are far apart they are eliminated by the winner takes all mechanism. So, here this refers to the long range inhibition whereas, the cooperation will essentially enforce the short range excitation. So, these two steps actually go hand in hand it is always a process of competition followed by cooperation.

(Refer Slide Time: 27:11)

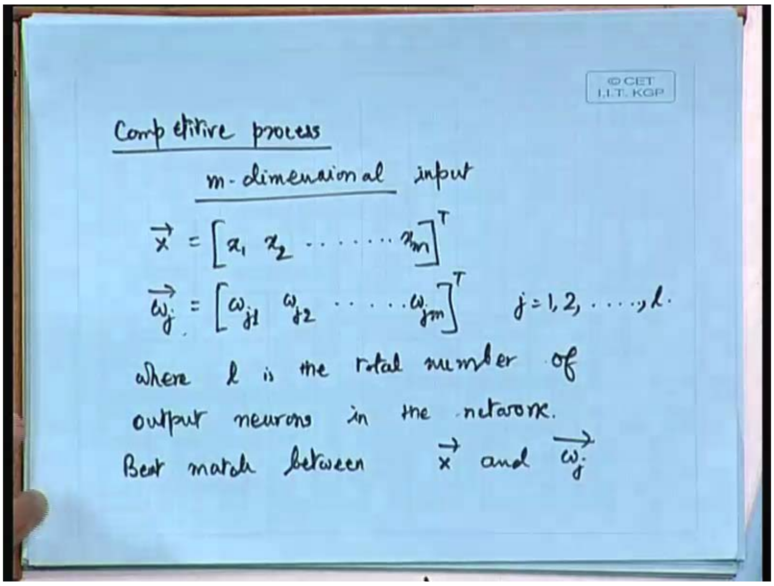


And then, the third step is of course, the step of synaptic ((Refer Time: 27:11)) adaptation, like we find in every network structure every neural network structure that there is a synaptic adaptation. So, what it means is that it enables the excited neurons to increase their discriminant function in response to the stimulus, which has caused the winning of the neuron. So, it enables the excited neurons to increase their enables the excited neurons, to increase their individual values of discriminant function in relation to the input pattern.

So, mind you we are in this case talking about the excited neurons only, than that only the excited neurons will have their discriminant value increased, so otherwise the discriminant function values for the non excited will be kept unchanged. So that means, to say that when the similar pattern is if fate, then the response of the winning neuron in response to a very similar flat pattern that and, which is fade again will increase, the response will increase next time.

So, this is, so these three are the essential steps, the competition the cooperation and the synaptic updation, synaptic adaptation and, we will be first talking about the competition mechanism. Let us go over to the mathematical modeling of the competitive process, so there as before we are assuming, so we are talking about the competitive process.

(Refer Slide Time: 29:51)



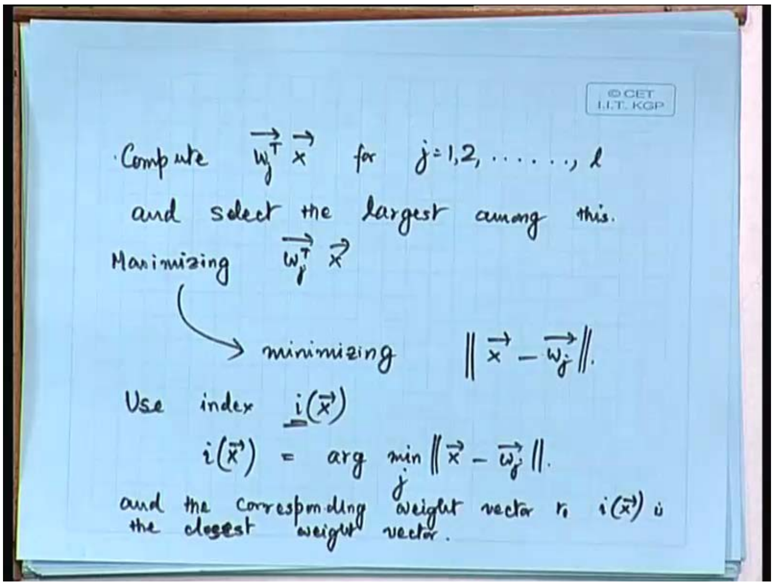
So, here we are considering an m dimensional input, so that the input x vector will be given by, x 1 x 2 with the elements x 1 x 2 up to x n transpose of this. And the weight w j will be given by, w j 1 w j 2 up to w j m, where this transpose, where j is equal to 1, 2 l. If we have got, what is this l, l is the number of output neurons, l is the number of output neurons, because m is the number of input and l is the number of output.

So, what happens is that every output is connected to the input, so that for the neuron one for the output neuron one, we will be having an m dimensional vector it is connected to all the inputs like that, for j is equal to 2. That is, the second output neuron we will be again having m number of such connections. So, totally there will be an array of m into l those many number of weights, so we will be having l such different w j vectors, so l, where l is the total number of neurons, total number of output neurons.

Now, what we have to do is to find a best match between x and w j, so we have to determine the question that exists is that, what is the best match between x and w j that is, what we want to determine. That out of see there, will be now a competition between this l number of output neurons. So, x is now going to compete with all the x, x is now going to find the match with w j and whichever, w j is having the best match that will, that j will emerge as the winner.

So, the winning index will be that j and, the corresponding weight will be the winning weight vector, in fact for some applications we may like to know only the winning neuron index and for some applications we may like to know, the actual winning vector. So, the question is that if we have to determine the best match between x and w j, what we have to do we have to compute the w j transpose x for different j’s.

(Refer Slide Time: 33:25)



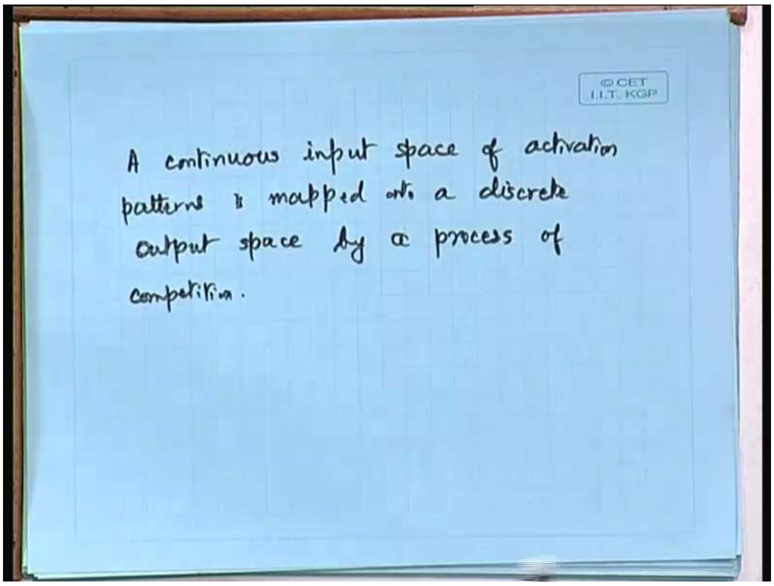
So, what we need to do is to compute, w transpose j x for j is equal to 1, 2, l and then, we have to select the largest, select the largest amongst this. So, the j that gives us the largest value is the winner, but in this case what we are doing essentially is maximizing w j transpose x and as I told you, that these in effect is nothing but, minimizing the Euclidean distance between the x vector and the w j.

So, what we do that, if we use the index i of x, so use the index as i of x vector, so where i is the index and it is index is based on some input function x on some input vector x. And the index i has a function of the input vector, x will be given by the argument of the minimum of this x minus w j, minimum is determined over all j’s. And that j which gives the minimum value of this, which is in effect the argument of this one, argument of minimum j of this that means, to say that that particular index is, the index of the winning neuron.

So, this is how we can find the index and, the corresponding weight vector, and the corresponding weight vector to, weight vector to i x is the closest weight vector. Now, one point to note here is that, when we are feeding the our input is actually in a continuous space, because input we are not discretizing the input x vector is in a continuous m dimensional space, but what we are doing is that we are mapping it into a discrete space of one outputs.

So, it is mapped into a discrete space and what happens is that, it is going to find the best matching to it so that means, to say that we are doing an a process of approximation from the continuous space to the discrete space.

(Refer Slide Time: 37:20)



So, a continuous input space of activation patterns is mapped on to a discrete output space of neurons, so any question on this. So, we have covered the competition part of it. So far, and the cooperation that is to say the excitation of the winning neuron, the excitation caused by the winning neuron into the neighborhood these are the things, which we are going to take up in the next lecture. So, if you have any questions in the mean time you can ask me nothing, than that is all for today.