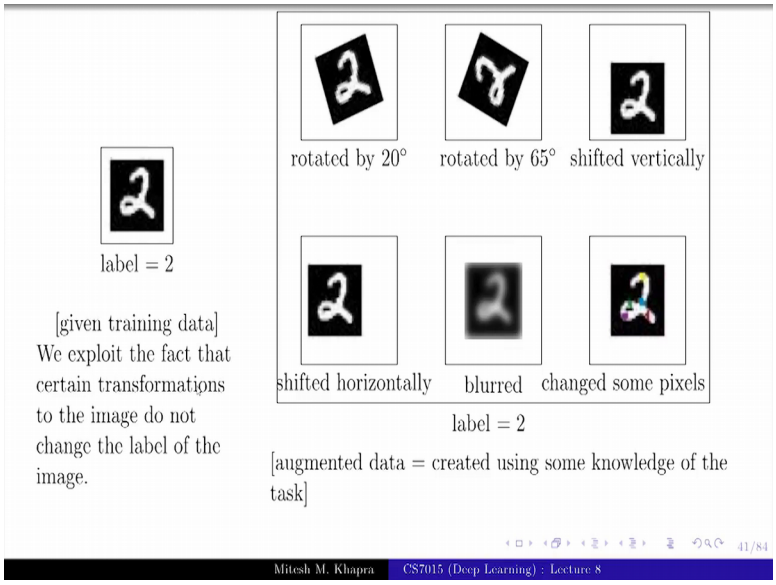
**Deep Learning**
**Prof. Mitesh M. Khapra**
**Department of Computer Science and Engineering**
**Indian Institute of Technology, Madras**

**Module - 8.5**
**Lecture – 08**
**Dataset augmentation**

So, how with that heavy math I will just interline with this, something very simple, which is, something known as dataset augmentation.

(Refer Slide Time: 00:21)




So, what is dataset augmentation mean? So, you always given some training data. So, in the case of nmnist, you had this training data, where you are given these digits, images of digits and you wanted to train some classified. So, in dataset augmentation, what we do is. So, now, we have, what is happening here right. Conceptual is that there some seeing, some training data and try to build a classifier and what you doing actually is minimizing the empherical train error.

That mean it will ensure that whatever you have seen in training is going to look. It is going to be perfectly classified. Whatever we have seen in training that is going to look very good. It is going to be the training error on, that is the error of those training examples, it is going to be very easy.

Now, my question is this; if a training time you are seeing all this 2's which are roughly vertically drawn right and a test time, you see at 2, which is written like this, which is slightly tilted what would happen? It will not be able to do a good job on that; that means, your model is not think of terms that you have used in this lecture, not generalizing ok.

Can you think of a simple trick based on your domain knowledge of how people right digits to kind of overcome. These part overcome, this part you get the question right. I am telling you that it is possible that someone writes to in a very tilted manner, can you prepare for eventualize, eventuality the title of this module was dataset augmentation. So, what would happen is, are given some training data.

You can always generate for training data from that, see here is another training insist that I have created, I have just rotate it to 2 by some random angle. I took this image, I just rotate it and this is a simple operation that all pixels are moving by a certain angle, I could have rotate it more. I could have shifted it vertically; that means, in all my image the 2 was actually exactly at the centre. I just shifted at a bit vertically.

So, I am. So, think that you are reading, one of those KYC forms or bank forms, most people would write at the center of the block provided, but some people could write to the extreme right or extreme left right. So, you are preparing for that, they saying that ok. All my data, the digits are well written at the centre, but let me just shift them bit; so, that I can also deal with people who write it at the corner ok.

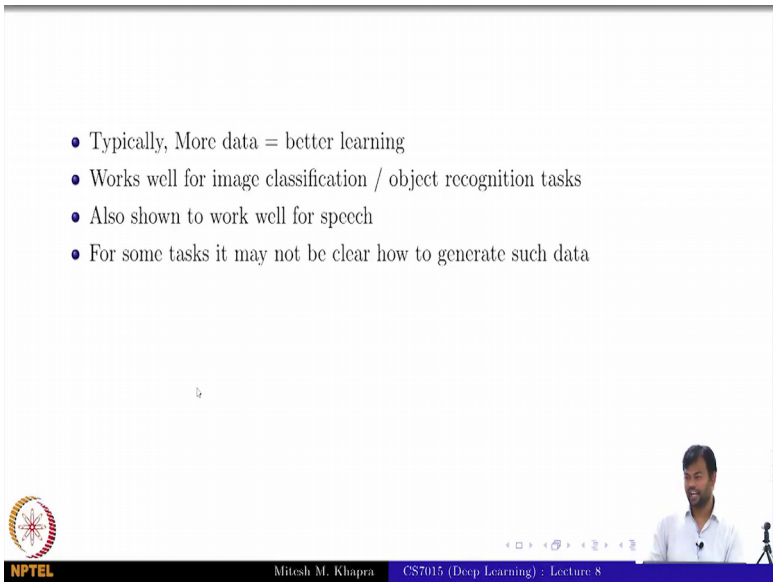
Left align or right align instead of center align. I could have even shift it horizontally. Most people would write at the center, but some people would write at the top or at the bottom. I could blur the image, but someone has taken a photo and send it to me and the photo is not very clear or I could just change, some pixels randomly right. I could add noise. All of this is dataset augmentation, with the hope that I am capturing with these variations. I am capturing enough variations in the data.

So, that I have a better chance of doing something better on the test data. Is that fine? This is all still training data, mind you, I am still going to compute the empherical train error. It is just that. Now, I have blown up my data, but much more than what I had initially do? You all see by doing this, you could have done better on the mnist assignment, you could have done better again, I am not asking you to do this.

So, now I will do this then I will have supervised data, because I know that by this small variations, the label is not going to change and what am I using, there I am using my domain knowledge right. I cannot do this always right. I hope you appreciated that suppose, that changes the domain a bit and I am given images of defects of motor parts right, where I have taken a image and there is a black spot somewhere, which indicates, defect I cannot go about doing the same thing, there I cannot change some other pixels. It will just means that the defects is at a different location right, but in many cases you can do that.

So, if you are given picture, because of dogs and cats, because the entire world case about classifying cats and dogs then you could do some rotations, you could blur them a bit, you could occlude certain questions of the picture and so on and still generate training data right. And what you are trying to do is, trying to take care of cases that you would end up dealing at test (Refer Time: 04:27) right. Is that clear ok, and please be aware that we are exploiting, some domain knowledge here.

(Refer Slide Time: 04:32)




Typically, more data is better learning, works well for image classification in object recognition. These are the task, where this is already been tried out and they have shown to work very well in these tasks also shown to work well for speech, where the people have some speech training data. They try to augment it for some task; it may not be very easy to generate such data right.

So, you could think of various NLP applications, whereas given you a data document right, because always do what Joe does in that Friends episode. Do you remember what I am talking you, see what I am talking about. See you wants to write a recommendation letter for Monika and Chandler ok, and he has a letter written, any replaces, every word by it's best synonyms from the (Refer Time: 5:20) right, that says a way of generating noisy data and in that case, it was actually noisy right. So, you could think of doing here, but as happened in that case, it will not result in very good transformation.

Next, for example, I remember something right. They are very warm hearted people got translated as they have some warm cardiograph or something like that which do not make sense. So, it is not very easy, in almost all application should do it, but in some applications, typically in vision application, this is easy to do and you would gain a lot by doing this right.