Contents

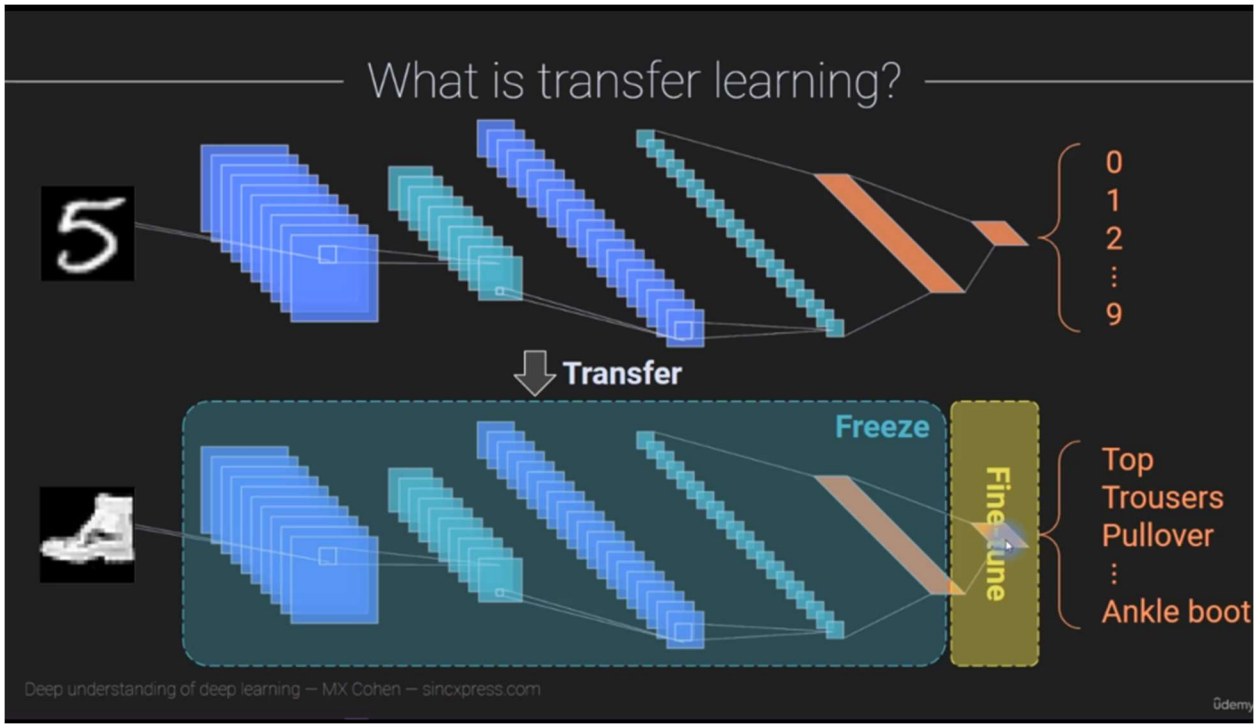
Transfer Learning: What, why and when?

In transfer learning you are transforming a trained model to a different data set or a different problem. you cannot simply just copy and paste everything in the model and expect to work immediately and that's because the nature of the problem is different.

For example, in MNIST digit dataset we have 10 categories from 0 through 9 so these are the digits here but in FMNIST dataset we have 10 categories which are qualitatively different.

When we use the same model for another problem it doesn't mean to follow the same architecture and the same set of parameters, like the same learning rate and so on, literally exactly the same model. So you can take all of the weights matrices, all the bias vectors, all of the filter kernels that this models have learned, all the learned parameters. There is a couple of ways to do this to fine tune the models so that it's going to work for another particular problem, which is slightly different from the problem that you train the model on.

One common way to implement transfer learning is to freeze part of the model and then you would do some retraining with your data so you continue training the data or the model based on your data, which is different from the data that the model was originally trained on. But notice, you are not training the entire model.so a lot of the parameters in the model are fixed, they are frozen and you are only training the output layer.



So this is called fine tuning the model and freezing all of these previous layers. This was one way for transfer learning.

Another way is to fine-tune the entire model. So in this case we are not freezing any of the weights, any of the layers or anything. We are allowing the entire model to be retrained. But you can imagine that the model doesn't need to trained from scratch. Because we have already have a model that has learned lots

of low level image statics, features like edges and curves and things like this. Sometimes this procedure is also called trained weight initialization because instead of initializing this model to have weights that are just purely random numbers, the weights here are initialized to be the weights here from model that is already trained on a slightly different problem. But you should know that transfer learning doesn't work in all problems cases.

Why transfer learning is useful?

Transfer learning allows you to Leverage HUGE datasets and existing (source)models to solve new problems. This is particularly useful when you are working with a fairly small data set you want to be able to use a fairly large model. So basically you just download a model that someone else trained with tons od data and you don't actually need a huge amount of data.

Transfer learning allows you to fine-tune existing trained models to your unique data. So transfer learning can be useful when you have a relatively small amount of data but bother people have used large amounts of data to train a model.

Transfer learning is also useful if you want to deploy your models quickly and if you have limited computational and data resources.

When is transfer learning appropriate?

You would use transfer learning when your problem is similar to a problem that someone else has already solved and this is one of the points.

Transfer learning is also appropriate when the model that you are transferring from the source model was trained on a lot more data that what you have. And also transfer learning is appropriate when the source model is deep. That's because shallow models tend not to transfer as well unless target data and the source data, the data that the model were trained on are very similar. Furthermore, shallow models are actually easier to train and it takes less data to train relatively shallow, model. therefor if you only need a shallow model you might not even need transfer learning in the first palce.so transfer learning is mainly advantageous for deep learning models.

[Reference: COURSE: A deep understanding of deep learning ,TEACHER: Mike X Cohen, cxpress.com

COURSE URL: udemy.com/course/dudl/?couponCode=202108]