Hello everyone, I’d like to give a speech about the next BIG high technology that may greatly change our life.

I think it is AI with unsupervised learning that can become the leader of the next generation.

As we all know that Artificial Intelligence has become a widely used technology.

People apply AI to many fields such as computer vision, natural language processing, Information Retrieval and so on.

For example, nowadays we often spend much time on short video apps, even someone are addicted to them.

So how can they be so attractive?

I believe that it is recommendation that plays the key role.

The system behind may analysis your browsing history and give you the next video that you may interested in, thus making you hard to leave your smart phone.

For interest, I got into a lab and did some research about it in the past one year.

I know that it is Deep-Learning based methods that have the best recommendation affects. Actually not only the rs field, almost every other field is trying to use AI to replace their traditional methods.

Well, it seems that AI is a big high technology.

But, what I have to say is that, AI still has many limitations preventing it for more useful things.

One problem is the gap between scientific research and industrial production.

Many papers just improve and refresh the result on certain existing datasets, which are meaningless for extending and applying.

So people now pay more attention to unsupervised learning, which has high scalability and are highly consistent with the initial goal of Machine Learning.

One of the most famous researcher on Deep Learning, 杨立昆， said that Supposing machine learning is a cake, reinforcement learning is a cherry on the cake, and supervised learning is a layer of sugar coating, then unsupervised learning is the main body of the cake.

In conclusion, I believe that unsupervised learning can lead AI to be the biggest technology of the next generation.

Let me introduce the general process of training an AI model.

To start with, we have datasets and our designed model.

What we are supposed to do is feeding the model for several epochs with many batches of data until the training and validation loss decrease to a certain level.

Lastly, apply the model to test the final result.

To be specific, in the most common scene of computer vision, image classification, we have a folder of images.

Each of them is associated with a label.

For example, dog images have label zero and cat images have label one.

We use the image as input, and the model can be treated as a black box.

It transfer the image to different types of data and eventually the image become a probability between zero and one.

If it is closer to zero, it means that the model thinks the input is a dog image.

Simultaneously, if it is closer to one, it means that the model thinks the input is a cat image.

The gap between the true label and the predicted probability can be used for updating the parameters of the model, so the model may perform better in the next epoch.

So after understanding the basic framework of getting a AI model, which part do you think that contain the human specific information?

Yes, the label of images. It is what we tell the model to imitate.

Then the problem comes.

It is wished that the more data we have, the better performance we will get.

However, while the internet generates plenty of data every day, supervised methods which use explicit label as supervised signal require enormous labor resources.

So how can we take advantages of those unlabeled data?

That’s why we explore unsupervised learning.

It doesn’t need label information, but can still perform good results.

The important thing is that it can be extended to large-scale datasets even using the real world’s datasets.

Actually, some famous Internet company, like face-book, have already applied this method to their production.

As for the specific detail of unsupervised learning, there are many kind of method to generate supervised signal, and this process is called pretext task.

If we crop and flip, or change the color of one image, it becomes a new sample. However, it’s semantic doesn’t change.

That’s to say, we now have two images that belong to the same semantic category.

It is a kind of supervised signal, and this method is called instance discrimination.

By using pretext task and design suitable structure of models, the gap between supervised learning and unsupervised learning is getting closer and closer.

More researcher begin to devote to this field.

If we say AI is the future, unsupervised learning is the future of AI.