

5 UNSUPERVISED LEARNING

PJD's Opening Remarks

As you have seen, neural networks are made of simple parts, which are mathematical models of biological neurons. When arranged in many layers containing thousands of neurons each, these mathematical neurons can do some astonishing feats such as recognizing objects in photos, interpreting speech, and translating languages.

To completely understand how they work and are trained, you need to learn a lot of math. It can be daunting. Even if you skip the math, there is some comfort in knowing that these machines have a solid mathematical basis and are not hacked together by inexperienced programmers.

The neural networks we've looked at so far have a weakness – they must be trained on massive data sets to learn the function we want from them. This is a weakness because the training data are expensive to obtain and can be untrustworthy or biased.

I saw a news video recently of people in an Asian country hired to produce training data in colonoscopy. They were shown interior images of colons and were asked to draw red circles around any red spots they saw on the colon wall. Hundreds of people had been hired to do this day in and day out. The result was a large database of colonic images with circled red spots. These databases were then sold to medical instrument companies that used them to train neural networks that would diagnose precancerous colon polyps. What's wrong with this picture? A medical device company trying to produce a diagnostic system that outperforms trained professional physicians, using massive data from untrained amateurs?!?

There are numerous other news stories about organizations that used data available to them to build face recognizers, only to find out that their neural network failed miserably to recognize faces of color because their training data were predominantly white faces. This is one of the reasons there is such a strong interest in avoiding "training set weaknesses". Is it possible to train machines without having to rely on an external data set? Could the machine learn from the data themselves?

Indeed, it is sometimes possible for machines to learn without any training data at all. A recent spectacular example is AlphaZero, a machine by the Google subsidiary DeepMind, which by playing hundreds of millions of GO games against itself became a grandmaster level player in 13 days and beat the world champion Lee Sodol in a tournament. Using the same method, AlphaGo learned grandmaster

chess in 4 hours. Compare that with the 50 years it took for chess experts to develop rule-based systems that beat grandmasters.

Today's speaker is Professor Chris Darken. He is an associate professor in the CS department and has been part of the MOVES group for many years. He has been applying neural networks to virtual reality simulation and training. He will speak to you about machines that learn without external supervision.