When Machine Learning Implies Intelligence

By Patrick van der Smagt

In the early 1990s, we thought neural networks would be commonplace soon and used everywhere. It took 30 years longer, but that point has now been reached. Neural networks can learn any function from complex and high-dimensional data sets, image or pattern recognition is solved, and probability theory serves as a sound mathematical basis for it all.

Neural network-based solutions to supervised learning have provided us with many applications related to image recognition, human–machine interaction, and so on, where large numbers of human-annotated data are available. Where these scale, they scale fast. This gives sound business models for IT companies with large user bases.

However, only very few types of applications can be tackled by human annotation. Europe is centered around manufacturing, to a large part covered by small and medium enterprises (70% of the gross domestic product in Europe versus 40% in the United States). For applications to scale, we need more than supervised learning.

We now have to look into fortifying research on more complex machine learning problems: unsupervised learning on small data sets, incorporating previous knowledge. If we ever want to get close to biological intelligence, these are key problems that must be tackled.

Digital Object Identifier 10.1109/MRA.2020.2986567
Date of current version: 10 June 2020

Beyond Frequentist

Why are neural networks so data hungry? An important reason is a mathematical one. At the core of a neural network loss function lies the maximum likelihood estimation principle, which assumes a specific probability density in the distribution of the data. For regression, the standard candidate is the normal distribution: following the central limit theorem, this assumption is reasonable as long as we have enough data. The Bayesian neural network solves this, where uncertainty is represented in the weights of the network. This approach allows for learning with very few data, but these networks are hard to train and do not scale like their frequentist counterparts. Practical applications are not near.

Incremental Learning

At least as serious a problem is incremental learning. Biological systems do not learn everything from scratch but combine the previously learned with new data. Technical approaches should be able to do the same and also incorporate handcrafted physical models into learning systems. How do we put model knowledge in a neural network? Possible solutions exist—e.g., in latent-variable time-series models—but are not commonplace, nor do they scale. This is a problem that must be solved before we can proceed to the next state of system intelligence.

Learning Without Oracle

Learning in biology is never supervised, needing large numbers of precise examples. Instead, we learn from combining observation, previous experience, and experimentation. How does this work?

A neural network solution to supervised learning has been available since the early 1980s, when the autoencoder was proposed. Recently, these models have proven to be useful as efficient latent-variable models that combine probabilistic approaches with neural networks. Time-series extensions have been proposed in the last few years, but scaling to real applications is not near.

A neural network solution to supervised learning has been available since the early 1980s, when the autoencoder was proposed.

It may be as in the 1990s: we have some first working approaches, but these have to scale and extend to wide applicability. Even if investment in machine learning may wane, if we continue focusing research on unsupervised Bayesian learning and latent-variable models, there's a good chance that, next time around, we'll have the resurrection covered.