lavery
Lawyers

Neural Network and Liability: When the information lies in Hidden Layers

October 8, 2019

Author

Eric Lavallée

Partner, Lawyer Partner, and Trademark Agent

Many of the most advanced machine-learning techniques rely on artificial neural networks, which allow systems to "learn" tasks by considering examples, without being programmed specifically to perform those tasks.

Neural networks are nothing new, however the emergence of deep learning[1], and of computers able to rapidly manipulate large amounts of data, have led to the development of a myriad of solutions incorporating machine learning for various aspects of life. From image recognition to financial data processing, machine learning is becoming ubiquitous.

From a mathematical perspective, modern neural networks almost always incorporate what are known as “hidden layers”, which process information between the input and output of a neural network system. Hidden layers’ nodes are not specifically assigned any task or weight by a human programmer, and typically there is no direct way of knowing how information is processed within them.

In plain language, most of the current machine-learning techniques rely on methods which function in such a way that part of what is happening is not known by the human operators. For this reason, the systems that incorporate such methods will give rise to new legal challenges for lawyers.

Scholars have been studying this issue for more than a decade now[2], but have failed to provide definitive answers.

Such questions are at the forefront of current legal debates. In a much-publicized case before the U.S. Supreme Court on gerrymandering[3], machine learning was mentioned as a source of concern by the dissenting opinion. This is not surprising given that the lower courts were presented with evidence on *Markov chain Monte Carlo algorithms*[4], which share this characteristic of not providing the human operator with a detailed explanation of how each data entered affects the results.

In some jurisdictions, for example the United States, a technology user may be able to ward off

requests for disclosure of the technology's algorithms and the details on the machine-learning process by arguing that they are protected as trade secrets of the vendor of that technology[5]. Even then, it might still be necessary to disclose at least some information, such as the results of the machine-learning process for various situations to demonstrate its reliability and adequacy.

Even such a defence may not be available in other jurisdictions. For example, in France, the Constitutional Council recently held that a public administration may rely on algorithmic processes in making decisions *only* if it is able to disclose, in detail and in an intelligible format, the way in which this algorithmic process makes its decisions[6]. From a computer-science standpoint, it is difficult to reconcile such requirements with the notion of hidden layers.

More importantly, there might be cases in which a person may wish to disclose how they made a decision based on a machine-learning technology, in order to show that they acted properly. For instance, some professionals, such as in the field of health care, could be required to explain how they made a decision assisted by machine learning in order to avoid professional liability.

A recent decision of the Court of Queen's Bench of Alberta[7] concerning the professional liability of physicians shows how such evidence can be complex. In that case, one of the factors involved in assessing the physicians' liability was the fetal weight, and the different formulas that could have been used in determining it.

The court made the following statement : "[…] the requisite expertise would concern the development of the algorithms used in the machine-based calculations of the composite birth weight as reflecting empirical research respecting actual birth weights and the variables or factors used to calculate composite birth weights. No individual or combination of individuals with such expertise testified. I draw no conclusions respecting the February ultrasound report calculations turning on different formulas and different weight estimates based on different formulas."

For developers and users of machine-learning technologies, it is therefore important at least to document the information used to train the algorithm, how the system was set up, and the reasoning followed in choosing the various technological methods used for the machine learning.

Computer scientists who have developed applications for use in specific fields may wish to work closely with experts in those fields to ensure that the data used to train the algorithm is adequate and the resulting algorithm is reliable.

In some cases, it may even be necessary to develop additional technologies to track the information traveling through the neural network and probe those hidden layers[8].

Things to remember

The risks associated with the use of a system incorporating automatic learning must be assessed from the design stage. It is recommended to consult a lawyer at that time to properly guide the project.
Where possible, technological choices should be directed towards robust approaches with results that are as stable as possible.
It is important to document these technological choices and the information used when developing automatic learning algorithms.
Contracts between technology developers and users must clearly allocate risks between the parties.

---

1. See, in particular: Rina Dechter (1986). *Learning while searching in constraint-satisfaction problems*. University of California, Computer Science Department, Cognitive Systems Laboratory, 1986.; LeCun, Yann; Bengio, Yoshua; Hinton, Geoffrey (2015). "Deep learning". Nature. 521 (7553): 436–444.
2. For example: Matthias, Andreas. "The responsibility gap: Ascribing responsibility for the actions of learning automata." *Ethics and information technology* 6.3 (2004): 175-183; Singh, Jatinder, et al. "Responsibility & machine learning: Part of a process." Available at SSRN 2860048 (2016); Molnar, Petra, and Lex Gill. "Bots at the Gate: A

Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System." (2018).
3. *Rucho v. Common Cause*, No. 18-422, 588 U.S. ___ (2019).
4. 279 F.Supp.3d 587 (2018).
5. *Houston Fed. of teachers v. Houston Independent*, 251 F.Supp.3d 1168 (2017); *Brennan Ctr. for Justice at New York Univ. Sch. of law v. New York City Police Dept*. 2017 NY Slip Op 32716(U) (NY Supreme Court).
6. Decision no. 2018-765 DC dated June 12, 2018 (*Loi relative à la protection des données personnelle*s).
7. *DD v. Wong Estate*, 2019 ABQB 171.
8. For example: Graves, Alex, Greg Wayne, and Ivo Danihelka. Neural Turing Machines. arXiv:1410.5401, [cs.NE], 2014.