The impartiality of some French judges undermined by machine learning

Supra Legem (Michael Benesty) · Dec 19, 2016 · 7 min read

The French administrative judge must be independent and impartial (Art. L721–1, R721–1 et seq. of the Code of Administrative Justice on the disqualification or the famous Article 6 of the European Convention on Human Rights). Is it seriously conceivable to expect from a man, however virtuous, a perfect impartiality? Probably not, and this is what "predictive" algorithms show us.

Technical details related to this article (= Deep Learning stuff) have been presented in march 2016 at the Paris Machine Learning meetup, slides are available here.

1 / A new context and new possibilities

As part of its open data policy, the French State [1] has made available to all over 350,000 judicial decisions corresponding to the content of the Legifrance website [2]. Due to the quantity of texts, the existing research tools are sometimes inadequate.

At the same time, the technological community (universities, research center, small and large companies) has dramatically improved predictive algorithms dedicated to the understanding of human language. The state of the art allows us to extract reliable and accurate information from documents as a human would.

This new combination, where legal data and powerful algorithms are available, offers the possibility of massively analyzing jurisprudence in order to extract a synthetic and unpublished vision.

Supralegem.fr applies those algorithms to extract, among others, the quality of the plaintiff [3] and the defendant, the nature of the legal provision [4] or the theme of the decision [5].

To do this, we use algorithms to read many texts and then we ask them millions of questions, all texts combined. When the answer is false, the algorithm update itself to avoid that error in the future (or at least make it happen less often). The procedure is repeated a hundred times. A specially equipped server, whose gross power is equivalent to approximately fifty conventional computers, takes between 5 and 6 days to run everything. Finally, the results can be queried by lawyers in a few seconds.