Is there a place for Machine Learning in Law?

Stephan Ralescu
CETANA, LLC
ralescu@gmail.com

Anca Ralescu
Senior Member IEEE
EECS Department, machine learning 0030
University of Cincinnati
Cincinnati, OH 45221, USA
Anca.Ralescu@uc.edu

Abstract

Research in artificial intelligence and law goes back approximately 40 years. It remains largely based on formal logic, including non-monotonic logic, case-based reasoning, and logic programming. However, some researchers in and practitioners of law have argued in favor of quantitative approaches (e.g. probability) to account for uncertainties in legal arguments. Other researchers have pointed some of the shortcomings of the current artificial intelligence and law research, e.g. inability to take context into account. At the same time, machine learning has made huge inroads in many different fields and applications, and therefore, the question is whether machine learning has anything to offer to the theory, and, equally important, the practice of law. As a position paper, this is a preliminary study towards the exploration of a synergistic integration of current artificial intelligence approaches in law, with machine learning approaches. It puts forward the idea that formal, logic-based approaches, currently very popular the Artificial Intelligence & Law research, could benefit from an extension with a machine learning component, and discusses some ways in which machine learning could be integrated into these approaches.

Introduction

When it comes to machine learning and law, there are two, quite unrelated, directions of study. On one hand, researchers are interested in legal issues raised by the research in machine learnig. For example, the symposium *Machine Learning and the Law*, held in conjunction with NIPS-2016[1], had as goal to *"explore the key themes of privacy, liability, transparency and fairness specifically as they relate to the legal treatment and regulation of algorithms and data*. On the other hand, the second direction is that of actually use of machine learning *in* law research and practice.

Stimulated on one hand, by progresses, as well as by shortcomings of artificial intelligence approaches in law (as perceived by various researchers), and on the other hand, by the tremendous recent machine learning succeses in many different directions (not including law), this paper suggests that there is scope for using machine learning in law, and moreover, that a formal treatment may be used towards this end.

The paper is inspired by work on logic-based formalization of legal reasoning, (Prakken and Sartor 1996), (Prakken and Sartor 1997), (Prakken and Sartor 1998), (Prakken and Sartor 2002), (Sartor 2002), as well as by ideas from (Tillers 2011) (Tillers 1993), (Franklin 2012) making the case for continuous mathematics tools (probability, mathematical evidence, fuzzy sets and logic), and the promise[2] that machine learning holds for law. An example of a predictive system can be found in (Campbell et al. 2016), for the restricted area of patent law.

As in many areas of research, ranging from science, engineering, medicine, and social sciences including the legal field, artificial intelligence has brought about possibilities, which excited some, intrigued others. Pioneering work done by Edwina Rissland and her students and collaborators, (Rissland and Skalak 1991), (Skalak and Rissland 1992), and Hafner and Berman (Hafner 1978), (Berman and Hafner 1993) has gone a long way towards understanding the promise and challenges that face formalization, with goal of developing a computer system, of legal reasoning.

Artificial intelligence and law

Formal logic is the approach of choice for artificial intelligence and Law, as evidenced by a wealth of articles, including those already mentioned, and others, (Bench-Capon 1997), (Prakken and Sartor 1996), (Prakken and Sartor 1997), (Prakken and Sartor 1998),published in a series of artificial intelligence journals, including the specialized journal of *Artificial Intelligence and Law*[3]. A critical review of the logic-based approach can be found in (Prakken and Sartor 2002).

Moreover, analyzing the current results, Franklin (Franklin 2012) lists challenges, not yet met by current artificial intelligence approaches in law, for formalization of legal reasoning. These challenges include:

[1] Annual meeting of the Neural Information Processing Society

[2] According to Google's Rob Craft *We are currently at year zero of the machine learning revolution.* (Singh 2016)

[3] https://link.springer.com/journal/10506

1. *"The open-textured or fuzzy nature of language (and of legal concepts)"*
2. *"Degrees of similarity and analogy"*
3. *"The representation of context"*
4. *"The symbol-grounding problem"*
5. *"The representation of causation, conditionals and counterfactuals"*
6. *"The balancing of reasons"*
7. *"Probabilistic (or default or non-monotonic) reasoning (including problems of priors, the weight of evidence and reference classes)"*
8. *"Issues of the discrete versus the continuous"*
9. *"Understanding"*

Franklin discusses the use of fuzzy set based approaches, to capture the nature of some concepts, or the similarity of a case to precedents. As an example, he considers, the concept 'vehicle', in the ordinance "No vehicles are allowed in the park", which obviously would refer first and foremost to cars, less to motorcycles/bicycles, and even less to roller skates. A fuzzy set of vehicles, defined by a membership function $\mu_{vehicle} : \mathcal{U} \to [0, 1]$, where $\mathcal{U}$ denotes a universe of discourse of 'things', would assign different degrees to *cars*, *motorcycles*, *bicycles*, *roller skates*, for example, respectively

$$\mu_{vehicle}(v) = \begin{cases} 1 & \text{if } v \text{ is a car} \\ 0.8 & \text{if } v \text{ is a motorcycle} \\ 0.5 & \text{if } v \text{ is a bicycle} \\ 0.1 & \text{if } v \text{ is "roller skates"} \end{cases}$$

The ordinance has as goal prevention of accidents in the park, and by consequence, the definition of the fuzzy set is meant to reflect the common sense knowledge that cars can cause serious accidents, motorcycles less, bicycles even less, and so on. The actual assignment of membership degrees is seen in (Franklin 2012) as one of the difficulties of adopting a fuzzy set based approach. However, the researchers in fuzzy systems know that while this issue is not trivial, fuzzy set based approaches have a rich collection of choices to address it, including, *learning* the membership function. Furthermore, it should be noted that (1) in many applications, the relative magnitudes of the membership degrees matter more than their actual magnitude, and (2) where the absolute magnitudes matter, they could and should be subject to a (machine) learning approach.

The issue of similarity is of utmost importance in legal reasoning and to illustrate the difficulties in similarity evaluation (Franklin 2012) refers to a celebrated case, *Popov v Hayashi*, centered on the issue of *possesion*.[4] The two precedents considered for the case, both involved hunting: in the first, hunting a fox with hounds did not confer rights of possession; in the second, the whale harpooned by one individual, and found by another on the beach was found, based on customs of whalers, to belong to the man who harpooned it not to the one who found it. The decision in *Popov v Hayashi* was that Popov and Hayashi had equal interests in the ball; to reach such a decision, issues such as *context*, *continuity* play an important role, and an intelligent (artificial intelligence) legal system must be able to deal with such issues. According to Franklin (Franklin 2012), none of the current artificial intelligence in law approaches, based on similarity with the two precedents, could have actually reached this decision. Achieving it, would require quantitative approaches including probability, fuzzy sets, and evidential reasoning, which may go a long way to complement logic based approaches towards an artificial intelligence based law systems.

Machine learning in rules with legal values

In (Sartor 2002) several (legal) theory constructors are given in terms of rules and (legal) values promoted by them, in order to formalize the legal argument. First *factors*, i.e. abstract features of a case which may influence the outcome of the case, are considered. Following (Berman and Hafner 1993), *values* underlying a case are introduced. For example, πLiv stands for the fact " π was pursuing his livelihood", (π denotes the *plaintiff*), or $\delta Nposs$ stands for "δ (δ denotes defendant) was not in possession". A legal value V is an objective pursued by the legal argument. Examples of values include *Less Litigation(LLit), More productivity(MProd), More security of possession(MSec)*. A case may be formalized as a collection of rules such as

$$\begin{array}{l} \lceil \pi Liv \Longrightarrow \Pi \rceil \text{ promotes } Mprod \\ \lceil \pi land \Longrightarrow \Pi \rceil \text{ promotes } Msec \\ \lceil \pi Nposs \Longrightarrow \Pi \rceil \text{ promotes } LLit \\ \lceil \delta Liv \Longrightarrow \Delta \rceil \text{ promotes } Mprod \end{array} \tag{1}$$

where $\pi Liv \Longrightarrow \Pi$ means *"π was pursuing his livelihood is a reason why π should have a legal remedy against δ"*.

To formalize, following (Sartor 2002), let $\{V_i, i = 1, \ldots n\}$ be a collection of legal values, where a *minimal approach to ordering* is adopted, such that the theory may specify

$$V_i < V_j; \;\; i \neq j$$

More over, it is assumed that

$$V_i < V_i \cup \bigcup_{j \neq i} V_j \tag{2}$$

[4]http://www.miblaw.com/lawschool/popov-v-hayashi-2002-wl-31833731-cal-super-ct-2002/: Popov v. Hayashi 2002 WL 31833731 (Cal. Super. Ct. 2002) **Case Name:** Popov v. Hayashi **Plaintiff:** Popov **Defendant:** Hayashi **Citation:** 2002 WL 31833731 (Cal. Super. Ct. 2002) **Issue:** Whether the defendant is liable for conversion when he picked up the home run ball that was dropped by the plaintiff. **Key Facts:** Barry Bonds 73rd Homerun. The plaintiff caught the ball in the upper portion of his glove but was tackled and thrown to the ground by the crowd. The ball fell out and the defendant picked it up and put it in his pocket. The plaintiff sued for conversion. **Holding:** The plaintiff and defendant had equitable claims and could not prove their case either way. **Reasoning:** Although the plaintiff proved intent to possess the ball, he could not establish that he would have fully possessed the ball had he not been tackled by the crowd. If he could have established this, his pre-possessory interest would have constituted a qualified right to possession which can support a cause of action for conversion. **Judgment:** The ball was sold for $450,000 and the proceeds were divided equally.

Replacing $\cup$ in (2) by the maximum $\vee$, and using $\wedge$ for minimum, it follows that

$$\begin{array}{l} V_i < \max\left(V_i, \bigvee_{j\neq i} V_j\right) \\ V_i > \min\left(V_i, \bigwedge_{j\neq i} V_j\right) \end{array} \tag{3}$$

Equality of values must also be specified as part of the theory. Then, enlarging upon (Sartor 2002), given the rules

$$\begin{array}{l} \lceil \alpha_1 \Longrightarrow \gamma \rceil \text{ promotes } V_1 \\ \lceil \alpha_2 \Longrightarrow \gamma \rceil \text{ promotes } V_2 \\ \dots \\ \lceil \alpha_n \Longrightarrow \gamma \rceil \text{ promotes } V_n \end{array} \tag{4}$$

one can construct the following:

$$\lceil \alpha_1 \& \alpha_2 \& \dots \& \alpha_n \Longrightarrow \gamma \rceil$$

promotes

$$[\min(V_1, V_2, \dots, V_n), \max(V_1, V_2, \dots, V_n)]$$

Thus, considered together, rules (4) promote *at least the smallest* value, *at most the largest* value, and possibly *values in between*, i.e., those which lie in $[\min(V_1, V_2, \dots, V_n), \max(V_1, V_2, \dots, V_n)]$. All promoted values can be expressed as convex combinations of $V_1, V_2, \dots, V_n$, that is,

$$\lceil \alpha_1 \& \dots \& \alpha_n \Longrightarrow \gamma \rceil \text{ promotes } w_1 V_1 + \dots + w_n V_n \tag{5}$$

where $w_i \geq 0, i = 1, \dots, n$ and $w_1 + \dots + w_n = 1$. For different values of $w_i, i = 1, \cdots, n$ (5) can generate any subset of the set of values $\{V_i, i = 1, \cdots, n\}$. Interpreted as a probability, $w_i = Prob($ to promote $V_i)$ can be obtained through a *machine learning algorithm* based on history of (similar) cases.

The mechanism outlined above has the effect of producing a *continuum* of legal values (even though to begin with, these form a discrete set), which in turn may lead to a *continuum* of possible decisions.

Inference and machine learning - legal theory and practice

This section touches upon the issue of inference in legal reasoning. It takes its cue from (Tillers 1993) and references therein, according to which *"the governing assumption of this body of law has been that all or practically all facts are uncertain and that proof of facts is always or almost always a matter of probabilities"*. The *necessity* of mathematical models of uncertainty (currently missing) in legal reasoning is furthermore discussed in (Tillers 2011) and (Franklin 2012) among others.

Since complex arguments about inferences from evidence rest on almost innumerable subjective judgments, (Tillers 2011) proposes several purposes for mathematical and formal analysis of inconclusive arguments about uncertain factual questions in legal proceedings, as follows:

1. "To *predict* how judges and jurors will resolve factual issues in litigation.
2. To devise methods that can *replace* existing methods of argument and deliberation in legal settings about factual issues.
3. To devise methods that *mimic* conventional methods of argument about factual issues in legal settings.
4. To devise methods that *support* or facilitate existing, or ordinary, argument and deliberation about factual issues in legal settings by legal actors (such as judges, lawyers and jurors) who are generally illiterate in mathematical and formal analysis and argument."
5. To devise methods that *capture some* but not all 'ingredients of argument' in legal settings about factual questions questions.

It can be claimed that achieving these purposes *predict - replace - mimic - support* falls into the machine learning realm, requiring machine learning algorithms of possibly different levels of sophistication.

Machine learning in the practice of law – the low hanging fruit

From the point of view of a typical approach to machine learning, data (usually, a lot) is needed to construct a machine learning algorithm - a classifier, or a clustering algorithm. Usually, such data is thought of as history on which to base future predictions. The need to take into account *history* is discussed in the conclusion section of (Sartor 2002), which suggests a *history-subtheory*. That would add a 'sense of history' to a case, predicting a judge's handling of a case based on that judge's history of opinions and their context. All of these could be attacked by machine learning methods. Issues on the representation of an argument, of an opinion, measures of similarity must be considered. Law, like other social sciences, seldom uses a *quantitative* language, rather, it is *text-based*. This means that solving the issues mentioned above is not trivial.

Using machine learning to analyze judges' personalities and ruling tendencies helps tailor pleadings to their personalities. Machine learning helps analyze attorney personalities and use those to decide who writes what in a law firm, and evaluate a firm's previous work and identify strength, weaknesses and faults.

This added dimension to legal theory and practice straddles several disciplines, including psychometry, representation of uncertainty (e.g., fuzzy logic to represent meanings of utterances, and similarity measures), probabilistic (point, interval valued or imprecise probabilities), all to be used in machine learning to build *predictive algorithms of behavior*.

As a recent example, with far reaching consequences, of behavior prediction, comes from the 2016 USA presidential elections: Cambridge Analytica[5] used machine learning to *specifically target* independents and other voters disenchanted with the status quo, with messages that appealed to *their* personalities. A similar system that does the same - for judges, courts - could help build a litigation strategy, tailor language, and develop legal reasoning to the personality of the particular court.

[5] https://cambridgeanalytica.org/

Conclusion

We have discussed some preliminary ideas on the challenges/issues that law research faces, which could be approached from an machine learning point of view. This paper only hinted at these issues and possible solutions using machine learning. Much is to be done, including a very thorough understanding of quantitiative ideas in legal theory put forward by researchers in the legal profession.

Acknowledgments

The authors are grateful for the reviewers comments, some very enthusiastic, some very negative, on the first draft of this paper. All served to stimulate the authors' thinking on how to further approach the use of machine learning in law, and are likely to inform their future study of this field.

References

Bench-Capon, T. 1997. Argument in artificial intelligence and law. *Artificial Intelligence and Law* 5(4):249–261.

Berman, D. H., and Hafner, C. D. 1993. Representing teleological structure in case-based legal reasoning: the missing link. In *Proceedings of the 4th international conference on Artificial intelligence and law*, 50–59. ACM.

Campbell, W.; Li, L.; Dagli, C.; Greenfield, K.; Wolf, E.; and Campbell, J. 2016. Predicting and analyzing factors in patent litigation. *NIPS2016, ML and the Law Workshop*.

Franklin, J. 2012. Discussion paper: How much of commonsense and legal reasoning is formalizable: A review of conceptual obstacles. *Law, Prob. & Risk* 11:225.

Hafner, C. D. 1978. An information retrieval system based on a computer model of legal knowledge.

Prakken, H., and Sartor, G. 1996. A dialectical model of assessing conflicting arguments in legal reasoning. In *Logical Models of Legal Argumentation*. Springer. 175–211.

Prakken, H., and Sartor, G. 1997. Argument-based extended logic programming with defeasible priorities. *Journal of applied non-classical logics* 7(1-2):25–75.

Prakken, H., and Sartor, G. 1998. Modelling reasoning with precedents in a formal dialogue game. In *Judicial Applications of Artificial Intelligence*. Springer. 127–183.

Prakken, H., and Sartor, G. 2002. The role of logic in computational models of legal argument: a critical survey. In *Computational logic: Logic programming and beyond*. Springer. 342–381.

Rissland, E. L., and Skalak, D. B. 1991. Cabaret: rule interpretation in a hybrid architecture. *International journal of man-machine studies* 34(6):839–887.

Sartor, G. 2002. Teleological arguments and theory-based dialectics. *Artificial Intelligence and Law* 10(1-3):95–112.

Singh, J. 2016. The tech-legal aspects of machine learning: Considerations for moving forward. *NIPS2016, ML and the Law Workshop*.

Skalak, D. B., and Rissland, E. L. 1992. Arguments and cases: An inevitable intertwining. *Artificial intelligence and Law* 1(1):3–44.

Tillers, P. 1993. Intellectual history, probability, and the law of evidence.

Tillers, P. 2011. Trial by mathematics - reconsidered. *Law, probability and risk* 10(3):167–173.