The tech-legal aspects of machine learning: Considerations for moving forward

Jatinder Singh
Computer Laboratory
University of Cambridge, UK
Jatinder.singh@cl.cam.ac.uk

Abstract

Earlier this year, an interdisciplinary Symposium was convened to explore the emerging tech-legal landscape regarding machine learning. Participants were invited to suggest topics they felt warranted discussion. This paper summarises these suggestions, to indicate some potential directions for moving forward in this new and important area.

1 Background

Machine learning is a topic of the moment. At a Campus event in April 2016, Google's Rob Craft suggested that *"we are currently at year zero of the machine learning revolution"*. Indeed, there is much hype and excitement regarding the potential of machine learning. However, as such technology looks to become increasingly prevalent, concerns have been mounting towards its use. It follows that the legal and policy concerns surrounding these technologies are increasing salience and prominence.

In September 2016, the Microsoft Cloud Computing Research Centre (MCCRC) hosted a Symposium in Cambridge entitled *"Machine Learning: Technology, law and policy"*. The aim was to provide a forum for some initial, exploratory discussion regarding the emerging tech-legal landscape surrounding machine learning. The Symposium was highly interdisciplinary, participants coming from a broad range of backgrounds, academic and industrial, in areas including law, computer science, government/public-policy, engineering, mathematics, sociology, philosophy, psychology and business.

As part of the Symposium, participants were invited to suggest discussion topics for a 'breakout session'. A number of interesting propositions were raised, far more than there was time to explore. This paper summarises and highlights (without judgment!) the aspects that an interested interdisciplinary community felt warranted consideration. As a new and emerging area, the aim is to share this collection of initial thoughts, in order to spur future directions and discussion.

The suggestions broadly fell into four categories, elaborated below:[1]

1. The effects and impacts of current and upcoming law and regulation;
2. Alternative governance approaches and practical ways forward;
3. The broader social aspects (context); and
4. Fairness and transparency: meaning and mechanisms.

[1] Note that this categorisation is to give a general flavor of the concerns. Naturally, there are no hard boundaries – some topics span a range of categories.

29th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain.

2 Current regimes

It is of little surprise that there was much interest in the impact of existing/upcoming legal and regulatory regimes on machine learning.

As a predominately European group, there was a considerable focus on data protection, particularly in light of the recent EU General Data Protection Regulation (GDPR) that comes into effect in 2018. One suggestion invited predictions as to what might be the GDPR's first machine learning "test case". The idea was that such a thought experiment would ground the more abstract legal discussions, by *practically* identifying aspects of interest, concern and ambiguity regarding the new regulation, and by focusing on its possible effects, including the provisions most likely to be tested (or not), and the sectors, firms and/or applications of machine learning that will likely come under scrutiny, e.g. by Data Protection Authorities or interest groups.

A question was raised as to whether (current and imminent) data protection regimes – particularly Article 22 of the GDPR that concerns "automated individual decision-making, including profiling" [1] – would work to stifle innovation in the EU. Related is the argument that regulation also creates innovation opportunities, e.g. where the regulatory regime incentivises the development of new tools and techniques (technical or otherwise) to facilitate compliance.

There was a general interest in issues of liability in a ML-driven world – as it was aptly put: "*where should the liability in machine learning systems lie?"* Indeed, this was a key focus of the Symposium. Reed et al. [2], who presented at the event, suggests that the existing legal/regulatory settlements would not be adequate for widely adopted machine learning, noting that *"it is far too early to devise a liability regime for machine learning generally because in the current state of development of the technology the law would rapidly fail to accord with technological change"*.

3 Governance models

Another area of interest was in alternative governance models and practical approaches for moving forward.

A number of suggestions were around the idea that governance mechanisms, rather than focusing on the individual and/or their consent (as is often the focus of data protection regimes), should perhaps shift to a model that regulates (controls/limits) *specific uses* of data. Relevant here, as one participant put it, is to consider *"when should machine learning algorithms be used, and when should they not?"* Related questions included whether there needs to be categories defining the acceptable uses of machine learning, and what those categories might be, e.g. what of life-threatening situations, or applications where fairness is crucial – should machine learning, or certain learning approaches, be encouraged or prevented in such situations? There was a proposal that the algorithms used by the public sector should be made public, though it is arguable that similar standards should also apply to firms.

It was thought that in order to make better policy decisions in this emerging area, more data is required. The idea of a registry was floated, to record who is using machine learning and what for, the possible groups affected, and any interactions with other systems (a factor particularly important as the world becomes increasingly instrumented [3]). Though challenged as impractical, it was explained that the vision for the registry, if one could exist, was to be indicative as opposed to an absolute 'source of truth', to provide policy makers with at least some visibility over the current state of play. Clearly representativeness would be an issue in such a scenario.

The role of standards and certification in regulating ML systems was also raised as a discussion topic. The focus was on the practicalities: if we want to encourage machine learning best practices, how do we come up with such standards? Who should lead such efforts? What would the standards look like? How do they maintain relevance given the rapid pace of innovation? Related are issues of incentivisation, concerning how best to encourage adoption and adherence, and whether this comes through a regulatory 'stick', market forces/consumer demand, voluntary codes-of-conduct, or some other means. There was a general consensus that ethical considerations should be embedded in ways forward, though the mechanism for how was felt to be an open question.

4 Social & society

Machine learning does not occur in a 'vacuum', but rather the vision (and concerns) regarding machine learning relate to its real-world, practical effects. It follows that a number of social considerations were raised. Note that though relevant and interesting, the Symposium did not delve into broader issues such as the 'future of work', super-intelligence, existential risk, and so forth, so as to focus more on the immediate tech-legal specifics.

Several participants challenged the validity of the focus on the 'individual', arguing that data is social; certain data may be shared, collaborative, and/or protected, all of which varies according to the circumstances. Context matters! Questions were raised as to whether current governance approaches – not only political and legal, but also the technical mechanisms for management, transparency and control – are sufficient and capable to handle the social nature of data. If not, and alternatives are needed, one needs to consider what forms might these take.

An interesting discussion point concerned whether current discourse is *too* focused on the technology. Given that machine learning technologies are created by people, used by people, and the outputs of which affect people, the proposition was *"might we blame the machines too much when, for example, policy or societal [or economic] factors are also at play?"*

5 Fairness and transparency

Fairness and transparency in machine learning is currently a hot topic of research. Indeed, forums dedicated to such issues, such as FATML (fatml.org), are rapidly expanding.

A number of topics were suggested concerning fairness, prompting questions such as what does fairness actually mean? How do notions of fairness vary by context? In what situations, and to what extent is 'unfairness' tolerable? How do we measure and evaluate fairness, and what are the pros/cons of different approaches? Tackling these sorts of foundational questions is crucial for progressing the area. There were also some specific technically-oriented questions, including fairness in unsupervised and reinforcement learning, whether particular learning algorithms (of a more transparent and/or controllable nature) should be mandated in particular scenarios (and *vice-versa*), and how the technology can usefully and effectively present to users its operations, functionality and indicate ('levels') of fairness.

In response to some discussion during the Symposium, several submissions considered whether machines could be made to be more objective, fair and transparent in their decision-making than humans. It was also questioned whether such a vision was misplaced (the participant noting that these 'tech-utopian' beliefs were *"especially [held by] those in computer science"*). As another articulated, given that humans (at the moment, at least) are involved in building machine learning systems (which entail a number steps, including data

selection, management and "feature engineering", algorithm selection, implementation, training, testing, run-time management, etc. [3]), some degree of bias will likely be 'built-in' (sometimes invisibly) to the systems that aim to be fair or fairer.

6 Concluding remarks

The tech-legal issues regarding machine learning are of real, and increasing concern. This paper presents some initial questions and thoughts on the topic, as raised by an interested, interdisciplinary community. Of course, as a new area, this is far from a definitive list, but even at these early stages, various themes appear to be emerging. The hope is that illuminating these considerations will provide some starting points for future work.

Undoubtedly moving forward requires an interdisciplinary approach. Indeed, as we progress, new areas for collaboration may well arise – in the words of one participant *"perhaps a new group of assurance and risk management disciplines [will develop] as machine learning becomes more prevalent."* What the Symposium made clear is that given the potential impacts of machine learning, there must be input from all domains into the discussion, including the economic, social and ethical aspects. There is a lot to do!

Acknowledgements

Thanks are given to the participants of the MCCRC Symposium 2016 for their valuable contributions, and to Microsoft for their financial support.

References

[1] Dimitra Kamarinou, Christopher Millard and Jatinder Singh, 'Machine Learning with Personal Data'. Queen Mary School of Law Legal Studies Research Paper No. 247/2016. Available at SSRN: https://ssrn.com/abstract=2865811, Oct 2016

[2] Chris Reed, Elizabeth Kennedy, and Sara Nogueira Silva, 'Responsibility, Autonomy and Accountability: Legal Liability for Machine Learning', Queen Mary School of Law Legal Studies Research Paper No. 243/2016. Available at SSRN: https://ssrn.com/abstract=2853462, Oct 2016

[3] Jatinder Singh, Ian Walden, Jon Crowcroft, Jean Bacon, 'Responsibility & Machine Learning: Part of a process', Available at SSRN: https://ssrn.com/abstract=2860048, Oct 2016