Microsoft Cloud Computing Research Centre
3rd Annual Symposium, 9th September 2016

Responsibility, Autonomy and Accountability:
Legal Liability for Machine Learning

Chris Reed
chris.reed@qmul.ac.uk

Consequences of ML decisions

- Incorrect decisions
  - Physical injury or property damage
  - Other losses
    - Reputation, privacy, financial loss

- ‘Correct’ decisions
  - ML discriminates/classifies accurately according to its training
  - But does so on a basis which is not allowed by law, eg on basis of sex

- Impenetrability of ML reasoning
  - Complexity of decision trees and opaque neural networks
  - Ex ante v ex post explanations
  - Impenetrability depends on questioner
Basis for liability

- Liability is imposed on persons, not machines
- Two bases
  - Dangerous/particularly risky activities
    - ‘Strict’ (no-fault) liability
    - Eg defective products liability
      - Tesla Autopilot?
      - Product or cloud service?
  - Negligence
    - Liability imposed because of fault
    - Based on relationship between claimant and defendant

Fundamentals of negligence liability

- Duty of care
  - Foreseeability of harm
  - Fair, just and reasonable to impose duty
- Breach of duty
  - Tested against the reasonable man acting in same circumstances
  - Professionals are tested against other professionals
- Causing loss
  - “but for” causation
  - Problems of multiple causation and causal chains
Negligence and ML

- New challenges to judges
- Duty of care
  - Introduces potential new responsible persons in ML producers
    - Split responsibilities
    - Remote relationship with those who suffer loss
  - May reduce or eliminate duty of care for those currently responsible
    - Product manufacturers
    - Users
  - Breach and causation
    - Standard of performance expected from ML?
    - Identifying how and why the loss occurred
      - Complex expert evidence
    - Multiple interactions as causation

Autonomous vehicle scenarios

- ML as assistant
  - Eg Tesla Autopilot

- Truly autonomous ML vehicle
Can negligence evolve?

- Of course!
  - Continuous evolution
  - Largely driven by new technologies
    - Especially motoring
- But evolution is a slow process
  - Can ML producers, or society, wait 20+ years for the answers?
- And will we like the result?
  - Probably based on false presumptions about knowledge and responsibility

ML and individual autonomy

- Individual autonomy as a societal value
  - Recognised by law eg freedom of contract
  - Embedded in fundamental rights, eg privacy and freedom of speech
- ML replacing individual decisions
  - Should individual's consent to basis of decision be required?
- Unlawful ‘correct’ decisions
  - Based on unlawful reasoning
  - Or based on facts which are impermissible to use
  - Infringe fundamental rights and thus diminish autonomy
A clash of perspectives

- Human rights
  - Takes the perspective of the individual
  - Fair and reasonable treatment in the light of fundamental rights

- ML technology
  - Takes the aggregate perspective
  - Objectively relevant predictive correlations influence decision
    - Even though impermissible from an individual perspective
      - Correlation ≠ causation
    - Eg use of ML to determine criminal sentencing

Can accountability help?

- From A4Cloud project, five attributes
- Transparency is the most important
  - Explains how and why a ML decision was made
    - Causation becomes less of a legal conundrum
    - Protection of fundamental rights can be assessed
- Responsibility can help the law allocate legal responsibility
- Verifiability would require records and audits of the technology
  - Potential legal liability of technology producers provides an incentive!
Should we regulate ML?

- Perhaps in defined and risky areas
  - Autonomous vehicles are an obvious category
  - Self-regulation of ML in medical diagnosis/treatment
  - Use of ML in sentencing – *Wisconsin v Loomis*
- But as ML technology becomes pervasive, this would turn into a project to regulate all human life
- And regulation has a known chilling effect on technology development

Incentives towards accountability

- Revise liability law to provide incentives
  - “Mere” injury or property damage
  - Infringement of fundamental rights
- Strict liability for high risk activities which only risk injury/property
  - Effectively reallocates insurance burden
  - Accountability relevant to premium levels
- Introduce presumptions about unaccountable ML
  - Rebuttable by explaining how and why of ML decision
  - Presumption of negligence
    - Acceptable to use ML which provides no accountability
    - Effective strict liability
  - Presumption of fundamental rights infringement
    - Will cause users to seek accountability
    - Not likely to be an insurable liability
Microsoft Cloud Computing Research Centre
3rd Annual Symposium, 9th September 2016

Responsibility, Autonomy and Accountability: Legal Liability for Machine Learning

Chris Reed
chris.reed@qmul.ac.uk