PREDICTIVE ANALYTICS

Future of Law Lab
University of Toronto Faculty of Law

PRESENTED BY

Jasmine Yu
Kareem Elmankabady
Ziqi Liu
Amy Kwong

FUTURE
OF LAW

Predictive Analytics Introduction Post

Back in 1870, Oliver Wendall Holmes, in his seminal paper, "The Path of the Law," characterized a lawyer's role as a professional predictor of legal outcomes. In the intervening century and a half, our understanding of technology has drastically improved, such that lawyers today have infinitely more tools in their toolbox to not only predict legal outcomes, but also replace much of the work surrounding practice management, marketing, and legal research. But despite being better equipped, many lawyers do not know how to take advantage of these innovations, and many commentators have pointed out issues with these new technologies. Such a context demands further inquiry into what is possible, what is normatively desirable, what is inevitable, and what this all means for the current practitioner.

In this first post, we will set the groundwork for the future topics by setting some basic definitions.

What do we mean by predictive analytics?
Broadly speaking, "predictive analytics" refers to the use of statistics and modeling techniques to make predictions about future outcomes and performance.[1] Using current and historical data patterns, predictive analytics tools can categorize new information or predict future outcomes.

Predictive analytics tools, in large part, draw on a series of underlying techniques including artificial intelligence (AI), and machine learning to analyze the data and ultimately make predictions.

What is Artificial Intelligence?
Artificial intelligence (AI) is a field of computer science that inputs robust datasets to computer programs, enabling them to exhibit rational thinking. It allows machines to comprehend, act, learn, and problem-solve with human-like levels of intelligence. AI can be divided into narrow and general AI: *narrow AI* is AI that is trained to perform specific tasks, while *general AI* would allow machines to solve problems, learn, and anticipate future results – essentially possessing a self-aware consciousness similar to humans.[2]

Artificial intelligence is a broad term that encompasses many other techniques, which are outlined below.

Subsets of Artificial Intelligence:
Machine Learning
Machine learning (ML) is a branch of AI that focuses on the use of data and algorithms to imitate the human learning process in machines, gradually improving their accuracy through repetition. It gives computers the ability to learn without explicitly being programmed – they learn to program themselves through experience.[3]

ML can be divided into supervised and unsupervised ML. Supervised ML involves human interaction within the system, like tagging a document or categorizing information. The computer then constructs its understanding based on these human inputs. Unsupervised ML is

---

[1] https://www.investopedia.com/terms/p/predictive-analytics.asp

[2] https://www.ibm.com/cloud/learn/what-is-artificial-intelligence

[3] https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

where the machine forms an understanding without any human input.. It's good at identifying patterns and trends since it is able to sift through a large quantity of data faster than humans.

Neural Networks
A neural network can be described as a means of conducting machine learning. It is a system that attempts to replicate the workings of the human brain through a set of algorithms. They are primarily used to interpret sensory data through a kind of machine perception, labeling or clustering raw input.[4]

Deep Learning
Deep learning is a system that consists of multiple neural networks to take advantage of large data sets and computing power to recognize and learn patterns. Examples include image and speech recognition.

In the next post, we will share some of the companies using AI and ML to create products that can be used across several areas of private practice. In the third post, we'll provide some of the ways the judiciary is using these technologies.

4

Private Sector Legal Tech Tools

In this post, we'll go over some of the legal tech tools that companies have developed using AI and ML. Specifically, in this post, we'll outline tools used in private practice across different use cases, and in the next post, show how some Canadian jurisdictions are leveraging AI in the judiciary.

Legal Outcome Prediction: Blue J
Blue J is a software platform that combines the power of machine learning and AI with expert understanding of legal factors. It is mostly focused on tax and labour & employment law matters in Canada and the US and can be used by private practitioners as well as certain government agencies.

More specifically, the software can be used by the following practitioners:
- **Tax Lawyers** (to support a legal position, find specific cases)
- **Tax Accountants** (to advise clients on best course of action, assess accuracy of positions, solidify position for audits)
- **Labour & Employment Lawyers** (to fortify labour and employment positions, gain advantage in settlement discussions, find specific decisions)
- **HR Professionals** (to make well-informed, legally compliant decisions and standardize the HR process)
- **In-House Counsel** (to determine fair termination packages, predict how courts would rule in challenging employment law issues, determine when to refer out issues)
- **Public Officials** (to focus on issues warranting further investigation and ensure consistency and reliability in decision-making)

Blue J's process for providing a prediction, or answer, to a specific legal issue can be broken down as follows:
1. **Input your scenario** (answer a series of discrete questions pertinent to the tax/L&E issue (as determined by lawyers))
2. **Run the Analysis** (Blue J uses its AI to predict the outcome)
3. **Fortify your position** (test the impact of different factors, review similar cases, relevant legislation
4. **Document your position** (download a report of Blue J's prediction, including the factors it considers relevant, to serve as an independent assessment)

Blue J, like all legal tech tools, still has some limitations. For example, the company still relies on human lawyers to provide initial research, guidance, and testing. It also relies on humans to read cases and tag them for the relevant factors.

Contract Review: KIRA
KIRA is a machine learning software used for contract review. The software is trained by lawyers to find important information in contracts and to search for discrepancies between documents. Using this software, lawyers can quickly review hundreds of contracts, compare them, and search for terms within them.

The first step in the process is to import a document set into KIRA. KIRA then reads, identifies, and extracts information, flagging any clauses, risks, and information that would likely be significant to a lawyer. The system comes with built-in knowledge of what kind of information might be significant, but users can also customize the search and tell the software what to look

for. Users can also enter a question (e.g., "Does this lease require payment of percentage rent?") and KIRA will deduct the answer from the documents.

KIRA also has a "form deviation" feature, which allows the user to quickly compare a given contract to a standard form agreement. The software will redline any points of deviation from the form.

After reviewing the documents, KIRA summarizes key findings and allows the user to export a report in a preferred format. Users can quickly visualize all ongoing projects quickly using the system dashboard.

How Can KIRA Help Practitioners?

KIRA enables lawyers to review contracts more quickly and more comprehensively. This assists with projects such as conducting due diligence, lease abstraction, checking for compliance, and finding deal points. Instead of spending hours reading the small print of contracts, lawyers can use KIRA to review and analyze the small print in minutes. This frees up time for other important tasks that can't be done by a machine, such as closing deals, client relations, and building new business.

KIRA also enables lawyers to review a larger volume of materials and leave no contract unreviewed. Without automated review, it would be impossible for lawyers to thoroughly review the hundreds of contracts that might be involved in complex commercial deals. Consequently, it was common practice for lawyers to just select a portion of contracts that they knew were the most critical ones for the present deal, and then take a risk on the others. Using KIRA, lawyers can review every single contract and ensure that they are not missing anything.

KIRA is already being used successfully to manage large deals, such as the acquisition of Intelex Technologies by a larger company in 2019. Lawyers handling the acquisition used KIRA to quickly compare and analyze almost 30 years' worth of contracts, amounting to over 110,000 pages.

KIRA also improves equity in contract formation when there is a significant gulf in sophistication between the contracting parties. This is evidenced by a case study of the music industry. When emerging artists are offered a contract, they often lack the experience to recognize whether the contract is fair or consistent with industry standards. Using KIRA, they (or their legal counsel) can quickly review the contract and compare it against industry standards. The time saved by using KIRA translates into cost savings for the client, making legal counsel more accessible to potentially under-resourced individuals.

Potential Limitations
KIRA may be vulnerable to problems related to the reflexivity of law and algorithmic systems (as discussed in another post). The system will produce results that are consistent with the rules derived from the content that was fed into it; consequently, any mistakes in the contracts entered may propagate, and any unnecessary terms may not be automatically flagged as such if they are common to the contracts entered. Using AI for contract review may therefore exacerbate the trend towards increasing complexity in contracts.

Practice Management: Clio

Clio is a suite of web-based tools that help law firms in practice management and client collaboration. It is a private company based out of BC, Canada.

The software is able to automate the intake process, help with scheduling consultations, and accepting payments online, and generally automates redundant tasks, communications, and workflows.

- Features include:
    - Shareable, online intake forms
    - E-signatures
    - Document automation to turn client intake forms into an organized system in the cloud as well as document management once legal documents are created
    - Clio Scheduler, an appointment booking software that is integrated with their payment system to allow clients to pay in advance of their meeting.
    - Lead and revenue tracking insights to allow firms to gauge which referrals generate the most revenue and to learn how clients find the firm.
    - Strong security measures and 24/5 customer support
    - Comprehensive case management functionality and advanced document management

Clio provides these solutions to lawyers working in a plethora of practice areas encompassing both litigation and transactional practices. The product can also scale to help firms of all sizes.

Chatbots: LawDroid
Law firms need answers to many questions about a potential client ASAP, such as contact info, the facts of their case, and whether you are able to offer the legal services they need. Quick answers to these questions will reduce time spent on a dead-end client. Chatbots can essentially automate these first encounters. They use pre-programmed responses or AI algorithms to directly engage with potential clients without the need for a lawyer. Chatbots have become a common customer service and data-gathering tool in other fields, and while not as intelligent as humans, can provide quick responses to legal queries and gather information on a client's needs.

LawDroid is a program that allows users to build a custom chatbot without any programming knowledge. These chatbots on law firm landing pages can help with intake, screening, pricing, and basic legal services such as the likelihood of being held liable for an injury. It's currently offered as part of Clio Grow and has a suite of AIs that can help convert website visitors into leads, intake new clients, and manage cases.

Legal Tech in the Judiciary

In our last post, we shared some of the legal tech tools becoming more prevalent in private practice--including tools capable of predictive analytics using AI and ML. In this post, we'll turn our attention to the current uses of legal tech in the judiciary. Unfortunately, there are not yet many in use in Canada right now, but the technology is gaining popularity in the US, UK, and Europe.

Tools in Canada seem to be limited to tribunals for online dispute resolution (ODR), which we will discuss below. But outside of Canada, predictive analytics tools are most commonly used for criminal justice; tools include:[5]
- Bail and sentencing algorithms that predict future recidivism
- Predictive algorithms that predict who is likely to commit a crime or where a crime will be committed
- "Scoring" algorithms that predict likelihood of being a victim of a crime

Legal Tech in the Judicial System in Canada:

Ontario's Condominium Authority Tribunal (CAT)[6]
Ontario's first online tribunal launched in 2017. It uses an online dispute resolution system to help claimants resolve certain types of condominium-related disputes as quickly as possible.

Jurisdiction
Its jurisdiction is currently limited to some condominium disputes, although the list of disputes that can be brought to CRT is constantly expanding. For example, it can currently handle problems relating to: Condominium Records, Pets and Animals, Vehicles, Parking and Storage, Noise, Odours, Vibration, Light, Smoke and Vapour, and Compliance with Settlement Agreements. The tool has the capacity to order penalties of up to $5,000, but will not order any legal costs unless there are exceptional reasons.

CAT Process[7]
1. File an application - users choose their issues from a list and provide additional details
2. Respondent joins the case
3. **Stage 1**: Negotiation - users work together to try and resolve the case between themselves
4. **Stage 2**: Mediation: CAT Mediator joins the case and tries to help the users
   Note: Agreements are legally binding and can be enforced by courts
5. **Stage 3**: Tribunal Decision: CAT Member joins the case and makes a decision that users must follow
   - Requires leave of CAT Mediator to progress to Stage 3
   - Can bring evidence, including witness evidence, to a hearing
   - Most of the decision will in writing through CAT, but telephone conference call, videoconference, or other live proceedings are possible as needed

---

[5]https://www.lco-cdo.org/wp-content/uploads/2020/10/Criminal-AI-Paper-Executive-Summary-Final-Oct-28-2020.pdf

[6] https://www.condoauthorityontario.ca/

[7]https://www.cci-sw.on.ca/sites/default/uploads/files/CCI-Seminar-Jan31-2018-mysteries-of-the-condominium-authority-tribunal.pdf

- CAT Member makes a decision based on facts and law based on evidence and arguments of the two parties

BC's Civil Resolution Tribunal (CRT)[8]
BC's Civil Resolution Tribunal was Canada's first online tribunal--launched in 2016.
In its first seven months, it tackled almost 14,000 claims. Roughly 85% of cases resolved were settled, while only 12 went to decision at the tribunal.

Jurisdiction
The CRT can be used to resolve strata disputes (condo disputes), small claims disputes up to $5,000, motor vehicle accidents, and injuries. However, the event triggering the dispute must have occurred within BC.

Process
1. Solution Explorer diagnoses dispute and sends you correct application form; you fill it in and send to Respondent
2. Respondent submits a response
3. Negotiation between parties
4. If dispute is not resolved through negotiation, a case manager tries to facilitate an agreement
   - Agreements can be turned into orders and enforced, just as court orders
5. If agreement cannot be reached, a CRT member will make a decision
   - Case manager will help you prepare evidence and submissions
   - Decision is binding and enforceable, just as court order

Sources
https://www2.gov.bc.ca/gov/content/housing-tenancy/strata-housing/resolving-disputes/the-civil-resolution-tribunal
https://civilresolutionbc.ca/
https://www.americanbar.org/news/abanews/publications/youraba/2018/march-2018/british-columbia-odr-system-handles-14-000-cases-in-first-7-mont/

Benefits and Drawbacks with Online Dispute Resolution Technology

Benefits
The online dispute resolution forum encourages collaborative, problem-solving approach to dispute resolution--as opposed to adversarial courtrooms. It also improves access to justice by facilitating and promoting self-help. To that end, it provides legal information and as an online tool, it is available online 24/7. Also, these tools have been proven to be far cheaper than trials - ranges from $0 to $200 per dispute. In Ontario, for example, the cost structure also further incentivizes collaboration by pricing CAT in stages, with each step getting progressively more costly. Similarly, it is also far faster than trial--online dispute resolution timelines range from 1-2 weeks to 3 months. They are also more accessible to non-English/French speakers -- CRT, for example, is available in multiple languages

Drawbacks
We should note that these tools are highly limited in terms of their predictive capabilities. Predictive aspect seems mostly limited to the "Solution Explorer" in CRT, which diagnoses the

dispute. Although definitely a step in the right direction in terms of speeding up the dispute resolution process and providing access to justice, it does not nearly eliminate the need for human intervention and does not make the law completely accessible to someone without a legal background.

Ultimately, if anything goes wrong, human intervention is required through caseworkers or case managers. Nevertheless, it can be complicated to fix errors (typos, math problems, etc.); although caseworkers can fix minor errors, they cannot change final decisions, which instead must be appealed. Similarly, these tools still requires human intervention in later stages if dispute cannot be resolved through negotiation. At this point, costs and time commitment starts to increase for users. Again, these tools also cannot enforce orders -- courts are needed for that function. Similarly, appeals cannot be heard through the platform and must still be sent to courts for judicial review

Implications for lawyers

The following list represents the key takeaways for the lawyers in terms of these tools' impact on their role:
- The tools represents a shift away from courtrooms and litigation (requiring lawyers) towards self-help and negotiation (does not require lawyers)
  - Lawyers are generally prohibited from participating in BC
  - Lawyers are discouraged from participating in ON since costs are rarely ordered
- Although CAT and and CRT are fairly limited now, they may be able to handle more complex issues in the future
  - Lawyers, especially those whose areas of practice are focused on small claims, will need to have additional value propositions in order to maintain clients

Issues with Predictive Analytics

In our first post of this series, we introduced some of the key concepts to understanding predictive analytics and the legal tech tools that are taking over the market. In our second and third posts, we showed you some of these tools and how they can be used by lawyers in private practice for online dispute resolution.

In our last post of the series, we'll outline some of the normative issues we have identified with the proliferation of predictive analytics and other legal tech tools.

Input Data Monopolies

LexisNexis and Westlaw have, seemingly, duopolistic control of the North American market for legal information. Not only are these essential, all-encompassing tools necessary to operate as a competent professional, but they also hold the source data for many legal tech tools.

Democratic theory and traditional copyright theory would suggest that LexisNexis and Westlaw don't own the law. Statutes, regulations, jurisprudence are produced by a public body, enter immediately into the public domain, regardless of public availability. However, these companies provide value in terms of research efficiency--electronic searches on a single online database is far faster and simpler than scouring through a library of different cases and statutes.

It would be difficult to build a minimally comparable database from scratch--this presents a prohibitive barrier to entry. The existing players can easily underprice rival entrants. Similarly, providing these services to law firms and in-house legal departments that have been using WestLaw or LexisNexis for years would be difficult business because it is hard to convince the legal industry to change.

This duopoly imposes broader social costs – the average citizen has no public access to the laws that are supposed to be influencing their behaviour. This indirectly increases the costs of civil litigation and leads to potential unfairness. If litigants have a wide income gap, wealthier party has a superior research advantage. Also, courts must pay LexisNexis and Westlaw to access the fruits of their own labour, so in a sense, taxpayers end up paying for this

There is one free alternative worth noting, CanLII. CanLII makes it faster and easier for legal professionals and the public to access high-quality legal commentary on Canadian court decisions. It is free to use for the public.

Sources:
http://moglen.law.columbia.edu/twiki/bin/view/LawNetSoc/ElliottPaper1
http://openscience.ens.fr/ABOUT_OPEN_ACCESS/BLOGS/2012_01_28_Tim_Worstall.pdf
https://advocatingvalue.com/the-legal-research-duopoly-its-evolution-and-the-new-threats/

Reflexivity of Law

As a reflexive construct of society, law cannot be neutral. It simultaneously reflects society and exerts significant influence on society, and therefore tends to privilege the existing social order and reproduce it in future. Those tendencies may be exacerbated by the use of machine learning in law, as algorithmic systems are also reflexive in the sense that they produce results that are consistent with the context in which they are formed.

Both law and machine learning take their designers' understanding of and goals for society and try to reproduce them in future. This may involve propagating errors or reproducing existing biases that ought to be eradicated. It is therefore important to carefully consider how and where to apply machine learning tools in legal practice.

Some scholars have argued that legal artificial intelligence (AI) cannot bring about the improvements in law's functioning that they predict. Legal AI systems will struggle and may fail to adequately deal with things that were not sufficiently represented in the dataset on which they were trained. This is a significant problem, as the most interesting and important cases in law generally arise where there is no clear or existing legal answer to a question. Although systems can be trained until their error rate is deemed acceptable, they cannot possibly be trained to respond to all (or even most) eventualities.

Source:
Simon Deakin, Is Law Computable? Critical Perspectives on Law and Artificial Intelligence (2020)
https://books-scholarsportal-info.myaccess.library.utoronto.ca/uri/ebooks/ebooks6/bloomsbury6/2021-07-07/2/9781509937097

Rule of Law Issues

Leading scholar R.F. Weber argues that the increased use of artificial intelligence in law may undermine the rule of law, depriving the law of some of its legitimacy and normative power. According to Weber, there are two core pillars of the rule of law: *predictability* and *universality*.

Law is predictable insofar as it is applied consistently with a public-minded rationale rather than in arbitrary fashion. Law is universal insofar as it applies equally to everyone, including policymakers. Weber argues that both the predictability and universality of law may be threatened by the increased use of AI in law. If AI were to replace judges and policymakers as a source of law, it could not be depended upon to produce results that are predictably consistent with the reasonable expectations of the public; rather, it would simply create an automatic logical extension of the existing law, which would be disconnected from any identifiable public purpose and therefore arbitrary. Moreover, AI may undermine universality by treating individuals as data points and attributing too much importance to the differences between legal subjects who are meant to be treated identically under the law.

Source:
Robert F. Weber, "Will the Legal Singularity Hollow Out Law's Normative Core?" Michigan Tech Law Review 97 (2020). Available at

Black Box Issue and Input Data Bias

The "black box" issue in computing refers to the fact that with most AI-based tools, we do not know how they do what they do. This is largely because the only thing we see are the inputs and the outputs; no information regarding the processes and workings in between is shown--nor would it likely be understood by most. AI relies on machine-learning algorithms that internalize data in ways that are not easily audited or understood by humans

There are a few reasons that give rise to the black box issue:
- **Unexplained algorithms:** When an AI-based tool reaches a conclusion, there is little visibility into understanding how the model reached that conclusion. This could occur in image recognition models and neural network models. However, not all machine learning algorithms encounter this issue, such as decision trees.
- **Invisible training data:** If the training data used for the machine learning algorithm is not reflective of real-world data, it will be impossible to infer anything meaningful from the results.
- **Data selection:** Access to the full data set of a machine learning algorithm is not enough to understand the model, since you might not know which parts of the data set were used to train the model. Full transparency requires knowledge of how the data from the available training set were selected.
- **Bias in training data sets:** Bias can refer to three aspects: the weights established in the model, the variance of the datasets, or from informational bias.
- **Model versioning:** Machine learning models constantly evolve with new data and new settings.

As AI-based legal tech tools continue to grow--especially in the judiciary--the impact of its decisions can become more serious. For example, if AI facial recognition tools are used by the police to identify a criminal, it can lead to disastrous consequences if the black box problem is ignored.

Sources:
https://towardsdatascience.com/black-box-theres-no-way-to-determine-how-the-algorithm-came-to-your-decision-19c9ee185a8
https://jolt.law.harvard.edu/assets/articlePDFs/v31/The-Artificial-Intelligence-Black-Box-and-the-Failure-of-Intent-and-Causation-Yavar-Bathaee.pdf
https://www.thinkautomation.com/bots-and-ai/the-ai-black-box-problem/#:~:text=The%20AI%20black%20box%2C%20then,For%20example%2C%20photos%20of%20birds.

A Potential Solution to the Black Box Issue: Anchors

"Interpretable machine learning" is a relatively new concept being developed because humans need to understand "black box" models. The idea is, the better a human can predict what the algorithm will conclude, the better that human can be said to understand the model.

Anchors are a type of "explainer" that explains how the algorithm came to that decision. There are two types of explainers: local and global. Local explainers relate to the prediction in question whereas global explainers relate to the entire data set. Anchors explain both the decision in question and the surrounding observations, combining local and global. An anchor is a rule that "anchors" the prediction locally, such that changes to the rest of the feature values do not matter - as long as the anchor holds, the prediction remains the same. Essentially, it shows the analyst which feature(s) influence the model output. They are useful for explaining black box models with high probability.

Sources
https://medium.com/swlh/ultimate-guide-to-model-explainability-anchors-2deab8239f57
https://homes.cs.washington.edu/~marcotcr/aaai18.pdf

FUTURE OF LAW

Future of Law Lab

UNIVERSITY OF TORONTO
FACULTY OF LAW

University of Toronto Faculty of Law

Web: futureoflaw.utoronto.ca
Twitter: @Futureoflawlab