

# The Future of Judging: Algorithms in the Courtroom

Dr. Inbar Levy

Melbourne Law School

## Abstract

*These keynote lecture notes explore the potential and limitations of using algorithms and machine learning tools in the courtroom. It identifies different areas in which algorithms would be suitable for legal decision-making and areas in which human judgments should be preserved.*

## Keywords

AI • Algorithmes • Courts • Legal Decision-Making

The topic of this lecture is the bright but modest potential of algorithms in the courtroom. In these notes, I explore the potential of algorithms and machine learning tools to improve decision-making and, particularly, legal decision-making. The goal is to consider the best roles for algorithms while also considering in which circumstances elements of human judging should be maintained. In other words, I investigate different areas in which algorithms might be more or less suitable for use in the courtroom. My assumption is that there are essential human skills in judging and that algorithms could help systemize the judicial function and reduce the risk of human error and individual bias. However, algorithms are still not suitable for every legal task, for the reasons I am going to mention. The lecture, therefore, evaluates the risks and benefits of using algorithms in adjudication by pointing out specific elements of legal skill and expertise and identifying those that are better suited for an algorithm and those that are better suited for a human. As I hope you will see, for the time being, there are significant limitations to using artificial intelligence to make legal decisions. However, AI and algorithms can be useful as tools to support human legal decision-making.


I would like to begin with the premise that human cognition is limited. This leads to the shift we see from human intelligence to artificial intelligence. Psychological studies show us that human intelligence is fallible. Humans are not perfect. We know that humans suffer from biases, memory problems, and other cognitive limitations. The issue we face is whether AI, algorithms, machine learning tools, and digital tools could provide a solution for our limited human intelligence. Before we move on to discuss this issue, I would like to mention a figure who represents the legal world perhaps more than any other figure, and that is Themis, the goddess of justice in Greek mythology. Themis represents the supreme law, and she is the divine voice that presented humanity with the laws of justice and morality. When we look at the sculptures and paintings of Themis, she is presented as a woman who is blindfolded. The goddess of justice is blind, and the meaning behind this depiction is that she is blind to any considerations that are irrelevant to the legal case at hand. The goddess of justice is not supposed to see anything aside from the application of the law to the facts of the case. She is unbiased, neutral, and blind to any extra legal considerations. This leads to the question of whether we see such neutrality in the legal world. Are judges really objective, rational, and blind in the same way as Themis? Admittedly, this question is a bit unfair because the standard posed by Themis is a godlike standard. None of us is always blind to all irrelevant considerations. We are all human, and indeed judges are human as well. There are benefits to being human, but that means that we also have human limitations. The suggestion to replace human judges with algorithms and artificial intelligence could be perceived as an attempt to create a present form of Themis, a modern-day goddess of justice. But are algorithms indeed capable of providing the desired solution?

The goal of the legal system is to achieve justice. The administration of justice is the very basis of every legal system. Justice is considered to be “the first virtue of social institutions” [John Rawls]. Tim Scanlon refers to justice as “what we owe to each other,” demonstrating the ancient idea that justice is “in relation to another person” [Aristotle]. One of the most important elements in the definition of justice is the principle of neutrality. Ensuring an impartial and nonarbitrary legal decision-making process is imperative to the administration of justice. For this reason, our discussion is crucial in the legal context. It relates

<sup>†</sup> Corresponding author: Inbar Levy

E-mail: inbar.levy@unimelb.edu.au

Open Access. © 2021 Inbar Levy, published by Sciendo

 This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License.

to the most basic goal of the legal world: to achieve justice and to provide a fair, neutral, and unbiased mechanism for the resolution of disputes. Therefore, it is necessary to empirically test how legal decisions are being made in practice and to then examine whether indeed algorithms could improve those legal decisions.

I would now like to discuss the human cognitive limitations in more detail before considering the role algorithms could play in mitigating these human limitations. Herbert Simon coined the term “bounded rationality,” demonstrating that our cognitive abilities are limited. Humans have limited calculative abilities, limited attention, and limited willpower. They have imperfect memory—to say the least; humans are very bad at memorizing or recalling events accurately. In order to deal with the fact that our cognitive abilities are limited, we use mental shortcuts and rules of thumb. These shortcuts, heuristics, might lead to mistakes in decision-making. Everyone who has read the book *Thinking Fast and Slow* by Daniel Kahneman, who won a noble prize, would be familiar with the work he did with Amos Tversky about heuristics and biases. Kahneman describes two modes of thinking: system 1 and system 2. System 1 is the intuitive system. It is quick, intuitive, emotional, and this is the system that uses mental shortcuts, and is also more likely to lead to mistakes in decision-making. System 2 is the logical, slow, deliberate system that is less likely to lead to mistakes, but system 2 also requires more mental resources. Since our cognitive resources are limited, we cannot always use system 2. An example for an operation of system 1 is driving on autopilot. For instance, if you ever drove home and were preoccupied with something else other than the road on the way there and yet found your way home nonetheless, your driving was in part an operation of system 1. An example for the operation of system 2 is filling in a tax form, which requires more effort. It is worth noting that our cognition is not neatly divided into two modes of thinking; we all use both systems, normally in conjunction. We use system 1 because it normally works, despite the fact it sometimes leads to mistakes. When it comes to judicial decisions, it seems that legal theory would encourage the use of system 2, as judges are required to apply the law to the facts of the case in a logical manner. It makes sense to discourage the use of the intuitive system that may lead to biased decisions. However, as with any other human, judges too use both systems. Not only that, but judicial intuition may be an important part of judging. Interestingly, Kahneman and Tversky found that there is consistency in the operation of system 1. Meaning, they showed that it is possible to predict the mistakes of system 1, showing a range of biases and heuristics that are the result of system 1. The ability to predict the operation of system 1 makes it possible to try to mitigate its mistakes, for example, via judicial training, or artificial intelligence.

An example of a cognitive tendency relevant to our discussion is studies conducted about the purchase of chocolate and jam. These studies test whether people are more likely to purchase chocolate or jam when they have more choice or less choice, for example, if they are presented with three or four flavours of chocolate or 12 flavours of chocolate. Presumably, under the long-list-of-options condition, one would be more likely to find the flavour of chocolate she or he or they enjoys—mint chocolate, for example—and therefore more likely to purchase the chocolate. However, these studies show that the more choices we have, the less likely we are to purchase the item. The reason is cognitive overload, which is our tendency to avoid making a decision when there is too much choice or too much information, which leads to too much mental effort. When presented with too many options, most people prefer to postpone the decision or to not make a decision at all. This is not only relevant for consumers buying chocolate or jam. There are empirical studies conducted with real judges, showing that they, too, are affected by this kind of cognitive overload. There are psychological studies demonstrating that when judges have to take into account a long list of factors, for example, if the legislation requires judges to consider a list of, say, 10 or 15 factors to reach a decision on the case at hand, they are likely to ignore many of these considerations, because there are just too many. Therefore, the phenomenon of cognitive overload is very much relevant to the legal world. Another legal example of cognitive overload is the “terms and conditions” of online platforms or phone apps. It is quite rare for anyone to read the terms and conditions of an app before or after purchasing and downloading it. The vast majority of people do not read these documents. There is just too much information, and companies behind online platforms count on the fact that users are not going to read the agreement and contest it, and so we all end up signing our privacy and information away. There are several aspects to the cognitive overload here. There is a quantity problem of too much information and also a complexity problem when the information is too dense, too complicated, or if the language is too professional, people struggle to understand it. There is also an issue of having a lot of information, not only for that particular decision, but generally in life. Even if we were able to read the terms of use of one app, it would be too much to do it for the 20 or 30 apps that many of us have on our phones. In practice, we all have limited time and limited attention, and we have to give our attention to other, more pressing tasks. The result is that we sign consumer contracts without giving our real consent. This issue is very much relevant for algorithms because, as opposed to human beings, algorithms can process vast amounts of information. This is one example for which algorithms could be extremely useful, as they can process large sets of data and assist human decision-making. Algorithms could support human judgment when a human judge has to take into account large amounts of data, by summarizing the data or presenting it in a simple way that is easier for a human being to process.

I now turn the discussion toward different legal skills and whether these would be better performed by a human judge or by an algorithm, diving into particular jurisprudential issues.

The first skill I explore is legal interpretation. One legal skill that is arguably beyond the realm of artificial intelligence is that of interpretation. Language skills are integral to legal expertise because the law is expressed in language. The meaning and interpretation of language is frequently the focus of legal debates and legal decisions. Legal interpretation is a part of the judicial function. Different words could mean different things in different contexts, and the judge must use her discretion in any given set of circumstances in order to give the legal norm its practical meaning. Here, algorithms face a challenge when interpreting laws or evidence because AI cannot understand language in the same way as humans understand it.

We already have algorithms that write poetry and even novels, and yet they cannot understand language similarly to humans. In the legal context, there are attempts to use natural language processing (NLP), which is a form of supervised learning, to analyze legal decisions. Under a system of supervised learning, humans must label great amounts of data to enable the machine to “understand” the language. Nikolaos Aletras and colleagues conducted a study that used natural language processing tools and managed to predict legal decisions only based on language. This is a novel study that succeeded in predicting the outcome of cases tried by the European Court of Human Rights based only on textual content, achieving a high accuracy rate of 79 percent prediction on average. The study was set to predict whether a particular article of the European Convention of Human Rights has been breached given textual evidence extracted from a case. The researchers used textual features to train support vector machine (SVM) classifiers in this type of supervised learning. But even though this study was successful in predicting legal outcomes, it is limited to a specific context in which the extracted text from the legal documents bore a sufficient number of similarities. In other words, in this case, the judgments and the applications were written in a particular format that enabled the algorithm to learn how to predict the result of the case. Usually, court judgments are not written in any particular format. Judges have the freedom to write a judgment in a style of their choosing. For this reason, the results of this study cannot be replicated in the majority of legal cases to date without significant development. The applications and judgments in this study had a distinctive structure which made them particularly suitable for a text-based analysis. It is a success story of supervised learning, but it is limited to that particular context. We are still not at the point of using algorithms to predict all types of legal decisions and, in fact, most legal decisions would not enjoy this kind of prediction rate using an algorithm because of the challenge of understanding language. However, there are scholars who are more optimistic about the ability of AI to develop language skills using a semi-supervised learning and even through artificial neural networks.

One solution that could promote the use of NLP is to ask or to mandate judges to write judgments in a particular format. Currently, it is difficult to train an algorithm to read human judgments because each judgment is written in a different structure and a different style. Limiting the way judges write their decisions could help, but it also could be detrimental to the development of the law, as legal reasoning is the main legal skill used to explain the decision-making process and to allow for legal developments and change through precedents.

The next legal skill I explore is applying legal standards. Here as well, the elusive nature of legal expertise and legal rules is demonstrated. Even though the idea of law suggests that there is a need for clear rules that produce just and predictable results in order to govern society, in reality, many rules are uncertain and in some areas no rules have yet been developed. As written by Robert Sharpe: “Judges are often confronted with the task of deciding cases for which the law seems to provide no clear answer.” And that happens quite often. This uncertainty is a result of the need to maintain flexibility. There are circumstances that cannot be predicted in advance. In these cases, we use legal standards instead of legal rules, since rules can be under- or overinclusive and lead to unjust results. An example of a standard is “one has to drive reasonably,” while an example of a rule is “it is illegal to speed above 80 km” on a particular road. Terms such as “reasonable,” “proportionate,” and “just” allow for more flexibility and judicial discretion in future legal cases. While there is a legal need to maintain a level of uncertainty, from the point of view of algorithms, standards are difficult to model, because they do not provide a clear answer (intentionally). Standards are often open to future interpretation, while algorithms search for an answer based on existing data. In this case, the nature of law and legal rules does not correspond with the nature of algorithms, at least for the time being.

This relates to another legal skill: legal reasoning. Law is a social institution, and the social nature of the law is inherently linked to the process of legal reasoning. Legal reasoning is necessary to explain the outcome of a legal decision, and the only way to contest a decision is by referring to the reasons behind that decision. But legal reasoning also is a skill that is a necessary condition for the development of legal systems, especially in the common law world in which legal development is based on precedents. This underlines the need for the community to be involved in the resolution of legal ambiguity. We often require legal reasoning when the legal question involves a controversial social issue. Social change issues cannot simply be decided based on existing data, meaning based on judgments fed into a model, because the process must allow for legal development and legal change. Social development is not a task that can be led by an algorithm, as it requires a social decision. Not only

that, algorithms face explainability issues as well as transparency and black-box issues, which go against the very idea of legal reasoning. Indeed, we cannot be sure that what the judge writes in her legal reasoning is truly the reasons that led to her decision, but at least there is a process and a written document with which to work, while with certain types of algorithms we cannot track the decision-making process at all.

The next issue is life experience and empathy. Life experience and the ability to empathize are qualities often looked for in judicial appointments. This leads to the question of whether an algorithm could perhaps be even more empathic than a human judge. That could be possible, as humans can suffer from empathy or compassion fatigue. Perhaps we could teach an algorithm to be empathetic without getting tired from trying large quantities of difficult cases. It is interesting to note that there is some tension between the ideal of a wise and experienced judge and the notion that justice is blind. We began with Themis, who is portrayed as blind, and yet now we are saying that judicial personal life experience is important. These ideas contradict each other. And indeed, formal views of the role of the judge suggest that the judge should have the ability to exclude all her own prior knowledge and experience in deciding the case before her and only decide based on the evidence of the case. However, if we were to consider life experience as important, then algorithms would yet again face a challenge as they do not absorb the world in the same way as does a human judge.

Lastly, there are issues of bias in relation to algorithms. As stated at the beginning of the article, humans suffer from biases and use mental shortcuts and intuition in their decision-making. However, algorithms could suffer from the same problem and even amplify it, since they are based on human data and human modelling and programming. A fundamental objective of the law is to provide an impartial decision-making process. Achieving this goal is not easy when we talk about human judgment, nor is it in the case of artificial intelligence and machine learning. If the data used by the algorithm contains bias, then the algorithm will systemize this bias instead of reducing it. It is very difficult to “remove” biases from the data because we cannot always identify them.

To conclude, machine learning tools are often compared to human decision-making and, while algorithms are not perfect, neither are humans. We are comparing an imperfect system to another imperfect system. Therefore, the choice between human intelligence and artificial intelligence should depend on the specific task. There will be tasks that humans are better at and tasks that are better performed by an algorithm (or several different algorithms). We can envision decisions that are a combination of both humans and algorithms. Currently, there are challenges in the application of algorithms and machine learning tools in the courtroom, particularly from the point of view of jurisprudence and procedural justice. While algorithms could help reduce the risk of certain human error, there are specific elements of legal skill and legal expertise that are more compatible with human decision-making. Therefore, for now, algorithms should be used as a decision support tool in the legal decision-making process, rather than as a replacement for human judges.

## References

- Higgins, Andrew, et al. “The Bright but Modest Potential of Algorithms in the Courtroom.” In *Principles, Procedure, and Justice Essays in Honour of Adrian Zuckerman*, Festschrift ed., edited by Andrew Higgins and Rabeea Assy, 113. Oxford University Press, 2020.
- Kahneman, Daniel, et al., eds. *Judgment under uncertainty: Heuristics and biases* (Cambridge University Press, 1982).
- Miller, David, “Justice”, *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2017/entries/justice/>>.
- Aletras, Nikolaos et al. “Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective.” *PeerJ Computer Science* 2 (2016): e93.
- Levy, Inbar. “Simplifying Legal Decisions: Factor Overload in Civil Procedure Rules.” *Melb. UL Rev.* 41 (2017): 727.
- Sharpe, RJ, *Good Judgment: Making Judicial Decisions* (University of Toronto Press, 2018).
- Simon, Herbert A. “Theories of bounded rationality.” *Decision and organization* 1.1 (1972): 161.

## Questions and Answers

Q1: How would you design an algorithm that can deliver just outcomes?

A1: This is a challenging task because every legal case is different and has different circumstances. There is a common law principle of treating like cases alike, but the difficulty is in determining when two cases are similar enough to each other in a way that would justify the same decision. There is currently no agreed way to clearly quantify when two or more cases are “alike.” Since each case requires striking

a particular balance, it is difficult to know when a result was just. When it comes to legal decisions, we cannot really know whether the decision was right or wrong. The notion of right or wrong in the legal context is illusive. It is not like forecasting the weather, where there is a test mechanism to check whether the prediction was right or wrong—it either rained or it did not. In the legal context, we never truly know whether the decision was “just.” And, we must bear in mind that the law changes and our social and moral views change as well. Acts that were criminalized in the past (such as homosexual acts) are no longer considered a crime (refer to the example of Alan Turing in the 1950s). For this reason, we cannot use an algorithm to decide what is a fair outcome. And there would be difficult questions that we would not be able to model, questions for which there is no clear yes or no answer. I therefore believe that judges will not be replaced by algorithms in the near future.

Q2: *Could we design an algorithm to write dissenting opinions?*

A2: Dissenting opinions are incredibly important and have an influence on the future law. If we were to have a “superalgorithm” that is able to know what is right or wrong, then perhaps we would not need dissenting opinions. In that case, we also would not need the mechanism of appeal. However, this too is unlikely to happen in the near future for the reasons I mentioned earlier, namely the incapability of algorithms to determine what is “fair” and “right.” If algorithms were able to produce reasons, which is an issue in itself, as there are explainability issues of AI and ML, then we could envision an algorithm that presents reasons for a dissenting opinion. Human judges already do that, since a part of legal reasoning is often to rebut arguments that go against the decision. It is also possible to imagine a system in which we have a few algorithms just as we have few judges. In that kind of system, there will be several different algorithms that are designed differently, and they would each make a decision similarly to humans who are different from each other. We also could think about a system in which there is a combination of human judges and algorithms, for example, six human judges and one algorithm. That could be very interesting.

Q3: *How could we prevent systemic bias when using algorithms?*

A3: There is a good reason to be concerned about algorithms amplifying existing biases. There are many examples for that in the real world, including, for instance, facial-recognition systems that discriminate minorities (the technology has more difficulty recognizing women, Asians, and African-Americans), partly because the database is biased. An algorithm is only as good as its data. The databases we use are created by humans and the data is from the real world. We already mentioned that humans suffer from biases and use heuristics, and so many times the data has bias as well. It is very difficult to “clean” a database because we have to identify the biases, and that could be an impossible task. On the other hand, we have to remember that human decision-making is not perfect either. Humans can make decisions with their own individual biases and so we cannot always ensure that human decisions are better. For this reason, I suggest identifying specific tasks better suited for a human-algorithm team, instead of choosing one over the other.

Q4: *Could algorithms develop and change with society?*

A4: This question is very much related to the issue of life experience. What is life experience? It is not only documents that we read; it is also the things that we see, feel, and experience in the real world. Maybe one day we will have robots walking around among us and they will absorb society in the same way as human judges. But I don't think that, in the legal context, we are close to that. For this reason, we should not have algorithms that are able to decide on crucial social issues and social change. That should be left to members of society (deciding questions of same-sex marriage, abortions, and more). These issues cannot simply be resolved based on existing data. They require something else that includes shifts in values that are absorbed by members of society.