

Outcome Prediction in the Practice of Law

MARK K. OSBECK AND MICHAEL GILLILAND

PREVIEW *Business forecasters typically use time-series models to predict future demands, the forecasts informing management decision making and guiding organizational planning. But this type of forecasting is merely a subset of the broader field of predictive analytics, models used by data scientists in all manner of applications, including credit approvals, fraud detection, product-purchase and music-listening recommendations, and even the real-time decisions made by self-driving vehicles.*

The practice of law requires decisions that must be based on predictions of future legal outcomes, and data scientists are now developing forecasting methods to support the process. In this article, Mark Osbeck and Mike Gilliland first examine the traditional tools lawyers employ along with the limitations that prevent these tools from consistently delivering accurate predictions. They then describe how new data-science approaches, including AI, are starting to alter the way law firms operate.

Editor’s Note: This article builds upon Mark Osbeck’s “Lawyer as Soothsayer: Exploring the Important Role of Outcome Prediction in the Practice of Law,” which is forthcoming in the *Penn State Law Review* (Fall 2018) and available at SSRN: <https://ssrn.com/abstract=3138211>.

ROLES OF THE LAWYER

Lawyers are best known for their role as *advocates*, promoting a client’s interests in the courtroom—sometimes even in the media. Yet lawyers also serve an equally important role as *advisors*, considering questions like “Is it worth taking this case to court?” “What outcome is a party likely to accept to resolve a dispute?” “Is a plea bargain advisable?” These questions require the lawyer to make predictions, a vital task and an essential lawyering skill.

Time-series models are not relevant to predicting legal outcomes—there are no appropriate time-series data for lawyers to use. Instead, lawyers have traditionally relied on three principal tools to facilitate outcome prediction: empirical information, legal (element-focused) analysis, and lawyerly experience. Here we examine these tools and their parallels to business forecasting methods.

It might seem that outcome prediction would be easier in law than in business;

the relevant statutes, precedents, facts, etc. should be known to lawyers, and legal prediction isn’t subject to the fads, unforeseen events, or other vagaries of customer demand that confound business forecasting. However, there are intractable challenges for lawyers, suggesting that traditional methods cannot ensure accurate assessment of legal outcomes.

Recent advances in predictive analytics, however, are providing new tools to at least complement traditional outcome-prediction methods. Business forecasters may find parallels between the predictive challenges faced by lawyers and their own forecasting endeavors.

TYPES OF LAWYER PREDICTIONS

Lawyers must predict outcomes in at least three situations:

- Case selection
- Settlement decisions
- Transactional practice

In *case selection*, lawyers have ethical obligations to avoid pursuit of spurious actions, and fiduciary obligations to their clients and their own firms. They must assess the merits of a prospective case, the likelihood of success, and whether an action is worth pursuing.

Regarding *settlements*, the vast majority of criminal and civil cases are resolved through negotiation, not by trial. In a criminal case, the lawyer balances a plea bargain versus the possibility of more adverse results (e.g., longer prison term) by losing at trial. In a civil case, the odds of winning at trial are balanced against the financial ramifications of losing. The driving factor in such negotiations is each party's prediction of litigation outcomes.

SELECTING YOUR ATTORNEY

You've been arrested and charged with a crime. What should you look for in a defense attorney? Surprisingly, a brilliant trial lawyer may not be your best option. Most cases are settled through negotiation, not trial. You may be better off with a skilled negotiator who can advise you on the desirability of accepting a plea bargain.

To provide such counsel, your lawyer will need to balance the prospect of a certain adverse result (e.g., one-year prison term) with a potentially worse result if you lose the case. This requires predicting the likelihood of losing at trial and the length of sentence you might receive.

Transactional lawyers are not litigators but give advice on matters such as business contracts or tax filing. But, they still need to assess potential outcomes of different courses of action, including the risks (and potential costs) if a contemplated transaction leads to litigation (or an audit). So, outcome prediction remains a vital component of client counseling.

PROPERTY TRANSACTION

A client wishes to purchase a tract of land for development, but the title is in dispute and the subject of a pending lawsuit. The client (i.e., the prospective

Key Points

- In their role as advisors, lawyers offer guidance on the course of a case and likely settlements based on their legal-outcome predictions. For example, the driving factor in settlement negotiations is each party's prediction of litigation outcomes.
- The basis for guidance on legal outcomes has rested traditionally with three tools: empirical information, element-focused analysis, and experience, none of which ensures accurate assessments. But recent advances in the science of predictive analytics are providing new tools that complement and enhance legal-outcome prediction.
- In this article, we offer insights into the potential benefits of predictive and text analytics for legal research and outcome prediction.
- Artificial intelligence is also being applied in multiple contexts: to identify those legal arguments that are likely to be most successful, to predict the outcome of trials, and to assess the likelihood of recidivism for criminal sentencing.

buyer) and the seller both believe that the client's pending quiet-title action is likely to succeed, yet it is crucially important to get a sense of just how likely the odds are, so that the client can make a rational investment decision.

The difference between a 5% and 20% likelihood that the quiet-title action would not succeed will make a significant difference in what the client's willing to pay for the property. In this type of situation, the transactional lawyer's job includes predicting the likely outcome of the title action.

The client's decisions may also hinge on outcome predictions regarding other types of proceedings: zoning approval, the prospects for proposed legislation or regulatory action, whether necessary licenses and permits can be obtained. Thus, the transactional lawyer may look to public-policy lawyers for guidance.

With advances in data availability and analytics, empirical information about legal cases is becoming an important basis for predicting outcomes in law.

TRADITIONAL TOOLS FOR OUTCOME PREDICTION IN LAW

Decisions must be made under uncertainty in virtually every endeavor, including practicing law. There are three traditional tools used by lawyers in outcome prediction:

Empirical Information

With advances in data availability and analytics, empirical information about legal cases is becoming an important basis for predicting outcomes in law. Currently available data sources include *jury verdict reports*, enabling lawyers to learn how similar cases resolved.

Another source of information, typically available only to attorneys who work with insurance companies or other entities that are involved in numerous litigation matters, is *confidential settlement data*. Since most cases settle prior to trial, settlement information gives the lawyer a more comprehensive picture of possible outcomes.

The most-costly but least-used source of empirical information is *jury research*, such as the empaneling of a mock jury to test arguments and strategies.

Legal (Element-Focused) Analysis

Here a case is broken down into its constituent elements. Lawyers use their judgment to determine whether each is applicable for predicting the likely outcome.

For example, proving *intentional infliction of emotional distress* usually requires the plaintiff to establish four elements: (1) the conduct was intentional or reckless; (2) the conduct was “outrageous”; (3) the conduct caused the emotional distress; and (4) the resulting emotional distress was severe. An element-focused analysis evaluates each element in light of case precedents, so is heavily dependent on legal research. Once the lawyer assesses the probability of establishing each element individually, the case as a whole can be evaluated.

Lawyerly Experience

Seasoned lawyers may temper an element-focused analysis with their own experience. Factors could involve knowledge of individual judges, the sympathetic (or non-sympathetic) nature of the parties, reputation of opposing counsel, and other anecdotal factors. These lead to a more holistic approach to outcome prediction. Evidence shows that judges and juries don’t decide cases merely by analyzing the individual elements. Rather, they balance that analysis with an intuitive sense of what justice demands (Posner, 2008).

Parallels with Business Forecasting

There are parallels between a lawyer’s approach to outcome prediction and that of the business forecaster.

The latter has empirical information: historical sales, pricing and promotional plans, market/distribution expansion or contraction, etc. The time series of historical sales is nearly always the starting point for statistical forecast modeling, often enhanced by exogenous variables like price changes. And newer methods can add further model enhancements, such as text analytics to capture customer sentiment expressed through online reviews or Twitter.

Element-focused analysis is analogous to techniques such as *attach-rate forecasting*. In this approach, the forecast for an option or accessory combines two elements: a forecast for the number of units of the base item, and one for the rate at which the option/accessory is sold per base item. In effect, a more complex judgment is broken down into simpler judgments and the rationale for each is documented.

Lawyerly experience corresponds well with judgmental overrides made to statistical forecasts, with one important difference. Forecast value added (FVA) analysis (Gilliland, 2013) can tell us the effectiveness of overrides to statistical forecasts (whether they improved accuracy or

not). But we don't have the data to know whether a lawyer's intuition-based "override" of an outcome has an effect on predictive accuracy.

SHORTCOMINGS OF THE TRADITIONAL TOOLS

The accuracy of outcome predictions in law has left much to be desired. Following are some of the fundamental limitations of the traditional approaches.

Empirical Information

Much of the *empirical information* useful for outcome prediction is simply not available. For example, jury verdict reports provide only cursory factual summaries, making it difficult to closely compare cases on the relevant facts. Also, they cover only trial results, so are of no help in assessing important pretrial activities such as a motion to dismiss.

Also, settlement information is critical for civil-trial negotiations yet is usually confidential. It may only be available to lawyers at insurance companies or other organizations that are in frequent litigation.

Element-Focused Analysis

Perhaps the primary weakness in *element-focused analysis* is its dependence upon an overly simplified view of how legal analysis works. The lawyer making the prediction relies on the consistent applicability of legal rules to known facts and assumes these rules will be consistently applied in future cases. (This is akin to a time-series modeler's assumption that the future will follow the patterns of the past.) But there are several inherent problems affecting the reliability of element-focused analysis.

- **Factual Uncertainty** Making accurate factual comparisons between cases may be impossible. At the outset, the lawyer must rely primarily on the client's factual account—a story that may be biased, incomplete, or even contradicted by what's uncovered during discovery or trial. Additionally, witness likability and credibility are important factors in a jury's assessment of the facts, and the judge may not weigh evidence the way

the lawyer initially assumed or may even exclude evidence. Thus, the application of legal rules to facts is more difficult than might appear.

- **Legal Uncertainty** Effective prediction is complicated by uncertainty over legal rules. The exact parameters of rules are often unclear, as is exactly how the rules apply in the given case. In researching precedents, it's often difficult to synthesize cogent legal rules from disparate cases; the rules themselves may be vague or ambiguous. Lastly, legal rules are not entirely static, so interpretation can evolve over time. Determining what "the law" is and how it's applied to a given factual scenario may not be straightforward.

- **Assessing the Legal Significance of Facts** What are the legally significant facts in a judicial ruling? This knowledge is required when determining relevant precedents, yet judicial opinion may not reveal all the factual considerations the judge relied on in reaching a decision.

- **Assessing Significance of Non-legal Considerations** How does a lawyer account for the myriad factors that can have a major impact on the legal outcome? Analysis may be skewed by economic or psychological factors, such as when advocating for a wealthy client who will generate considerable fees. Unwarranted optimism and financial incentives might tempt a lawyer into prolonging litigation, even if early settlement would be best for the client.

Other considerations include personal biases of judges or juries, the likability of the parties, reputations of the attorneys, even the time of day a ruling is made (a study of Israeli judges found that prisoners were more likely to be granted parole early in the day, or after a break such as lunch [Danziger, 2011]). All need to be weighed in an elements-focused outcome prediction, but it is very difficult to do so.

- **Limitations of Available Information** Published judicial opinions are useful in determining whether a cause for action may exist or whether a case is allowed to proceed, but these opinions provide less

insight into the odds of success at trial or the amount of potential recovery.

• **Difficulty in Making Probability Assessments** As described earlier, *intentional infliction of emotional distress* requires four elements to be established. Assuming reasonable probabilities could be determined for each element (a big assumption!), what then is the probability of all four being met?

If the elements are considered *independent*, one could simply multiply the

may be influenced by the lawyer's own beliefs, biases, and memory (where more recent events have stronger influence). A fundamental challenge always exists: identifying which facts are legally significant and how a court will interpret laws and precedents. There is even a danger of adopting what Kahneman (2011) refers to as the "inside view": overly focusing on the specifics of the case and neglecting base-rate data on the underlying empirical probability of a given outcome (the "outside view").

Data science has assumed an increasingly important role in the practice of law, beginning in the area of *e-discovery*. Lawyers can now conduct discovery investigations (the disclosure of all relevant facts and documents to both parties prior to trial) in a more cost-efficient and timely manner.

separate probabilities together. (Thus, if each element had an 80% chance of success, the probability of establishing all four would be $(0.8)^4$ or 41%.) But the nature of judicial decision making tends to be holistic, balancing an intuitive view as to which party should prevail with a strict analysis of the separate elements.

Since the factors influencing a court's resolution of one element may well affect its assessments of the others, there is correlation among the individual elements. So it is not appropriate to assess the overall likelihood of success by simple combination of the probability for each individual element. And currently, there are no data to address this.

Lawyerly Experience

Experience enables a lawyer to broaden the scope of the analysis and take a more holistic approach to outcome prediction. For example, knowledge of a judge's propensities or the likability of a client may result in a prediction that differs from what is suggested by the empirical information and legal analysis alone.

However, even a seasoned lawyer's experience is finite, limited to the relatively small number of clients and cases they have handled. And that experience is a filtered interpretation of past events that

APPLYING DATA SCIENCE AND PREDICTIVE ANALYTICS

Data science has assumed an increasingly important role in the practice of law, beginning in the area of *e-discovery*. Lawyers can now conduct discovery investigations (the disclosure of all relevant facts and documents to both parties prior to trial) in a more cost-efficient and timely manner.

One example is *predictive coding*, a machine-learning process that takes keyword searches (used to find relevant documents) and applies it to much larger data sets to reduce the number of irrelevant documents that need to be reviewed manually (<https://www.exterro.com/basics-of-e-discovery/predictive-coding/>). Legal-research services have used such data-science advances to improve the responsiveness of their searches and demand for these advancements is growing quickly. Although not yet at the forefront of data-science applications, there are significant prospects for using predictive analytics as a complementary tool to improve outcome predictions.

Development of Predictive Tools

Use of predictive analytics can extend the application of empirical information for outcome prediction. For example, the process can compare information about

past cases with a prospective case, under the assumption that similar cases will be similarly decided. (This is akin to *forecasting-by-analogy* for new products.) *Text analytics methods* seek to detect objective patterns in the language of court documents, rather than relying on the lawyer's subjective assessment of similarity. This basic strategy, in rudimentary form, was first described over 50 years ago (Nagel, 1964).

Legal doctrine is a framework or set of rules, often established through precedent. Twenty years ago, Shauer (1998) examined the role of legal doctrine in outcome predictions, contrasting two approaches. In one, lawyers predicted outcomes by evaluating how courts resolved precedents in accordance with traditional legal concepts. The alternative approach denies reliance solely on legal doctrine and suggests there are empirical regularities that are of value in prediction. For example, the party affiliation of the presidents who appointed Supreme Court justices may be the best predictor of their rulings.

Current Applications of Predictive Analytics in Law

Online research services, such as Westlaw and Lexis/Nexis, offer tools that aggregate data from jury verdict reports and related publications, allowing users to search and filter results to find comparable cases. This is much more efficient than manually perusing case summaries for similarities.

Business forecasting professionals think their job is hard—and it is. However, they do benefit from some advantages not available to the lawyer.

More sophisticated tools use artificial intelligence (AI) and machine-learning techniques. Lex Machina offers software to track different variables and case outcomes, and to assist lawyers in drafting motions by identifying which arguments are likely to be most successful. Ravel Law provides tools both for outcome prediction and crafting persuasive arguments. And Ross Intelligence allows lawyers to

pose natural-language questions, providing answers by predicting the most applicable solution to the problem posed. While these tools do not yet ensure accurate outcome predictions, this technology is increasingly used to supplement the lawyer's traditional tools.

Other areas of the legal system are applying AI as well, although not without challenges and controversy. For example, the risk of flight is an important consideration in setting (or denying) bail to the accused individual in a criminal court. Risk of recidivism is an important consideration in sentencing a convicted criminal. AI tools are now in use to assess these risks and provide purportedly "objective" guidance to the judge. However, while AI might eliminate a specific judge's implicit biases, "it also incorporated the aggregate biases of all the decisions it was trained on" (Weber, 2018). A ProPublica exposé (Angwin and colleagues, 2016) found racial bias in risk assessments created by commercial software used in several jurisdictions.

Obstacles to Predictive Analytics in Law

Lack of meaningful data is a principal factor limiting the success of predictive analytics generally. This includes the void of publicly available settlement data, as well as the generic nature of much of the other available data. For example, published opinions may not specify all the facts that weighed upon the court's decision. This makes it difficult to find meaningful factual similarities between cases and to determine what facts were legally relevant.

Additional documents that would be data-rich, such as deposition transcripts, are generally not publicly available. And important extra-legal considerations recognized as affecting trial outcomes, such as the credibility and likability of individual parties, have no reliable source of data.

The other principal factor limiting predictive success generally is discerning signal

from noise in the cacophony of available data (for example, Silver, 2013). There will always be the fundamental challenge of distinguishing relevance and insight from randomness and anomalies. Thus, significant work remains to improve identification of meaningful patterns from the accidental correlations that do not inform predictions.

CONCLUSION

Given the challenge of delivering consistently accurate predictions of legal outcomes, will emerging methods in data science and predictive analytics offer some solutions? They are already having an impact on the practice of law and may eventually improve some outcome prediction. But just as there is no magic algorithm ensuring consistent accuracy in business forecasting, these new data-driven approaches won't entirely solve the problem of outcome prediction.

Business forecasting professionals think their job is hard—and it is. However, they do benefit from some advantages not available to the lawyer: complete and accurate data about the historical product

sales and knowledge of organizational actions (e.g., pricing changes or promotional activities) that will affect future demand. And importantly, they receive continuous feedback on their forecasting performance, as each new period's actual sales can be compared to the forecast. The biggest challenges include the internal politics of the forecasting process and the inherent volatility and randomness in customer behavior.

Legal software should “learn” and improve as new cases come in. Yet accurate outcome prediction still faces intractable challenges, including the incomplete and often ambiguous historical record of facts, interpretations, reasoning, and other potentially relevant information. And lawyers suffer many of the same temptations for bias—for creating “aspirational” rather than objective predictions—that afflict business forecasters.

Perhaps, in light of the difficulties discussed above, business forecasters can find sympathy not just for themselves but for the plight of the lawyer as well!

REFERENCES

- Angwin, J., Larson, J., Mattu, S. & Kirchner, L. (2016). Machine Bias, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Danziger, S., Levav, J. & Avnaim-Pesso, L. (2011). Extraneous Factors in Judicial Decisions, *Proceedings of the National Academy of Sciences*, 108 (17), 6880-6892.
- Gilliland, M. (2013). FVA: A Reality Check on Forecasting Practices, *Foresight*, Issue 29 (Spring 2013), 14-18.
- Kahneman, D. (2011). *Thinking, Fast and Slow*, Farrar, Straus and Giroux.
- Nagel, S. (1964). Applying Correlation Analysis to Case Prediction, *Texas Law Review*, 1006.
- Posner, R. (2008). *How Judges Think*, Harvard University Press.
- Shauer, F. (1968). Prediction and Particularity, *Boston University Law Review*, 773.
- Silver, N. (2013). *The Signal and the Noise: Why So Many Predictions Fail but Some Don't*, New York: Penguin Press.
- Weber, S. (2018). How Artificial Intelligence Is Transforming the Criminal Justice System, <https://www.thoughtworks.com/insights/blog/how-artificial-intelligence-transforming-criminal-justice-system>



Mark K. Osbeck is Clinical Professor of Law at the University of Michigan Law School. He is the author of *Impeccable Research: A Concise Guide to Mastering Legal Research Skills* and coauthor of *Marijuana Law in a Nutshell*. Prior to joining the law school faculty in 2001, Mark litigated sophisticated commercial cases, with extensive trial and appellate experience.

mosbeck@umich.edu



Michael Gilliland is Product Marketing Manager at SAS Institute and *Foresight's* editor for Forecasting Practice. He is the author of *The Business Forecasting Deal* and principal editor of *Business Forecasting: Practical Problems and Solutions*. In 2017 he received the Institute of Business Forecasting's Lifetime Achievement Award.

mike.gilliland@sas.com