Business, Management and Economics Engineering
ISSN: 2669-2481 / eISSN: 2669-249X
2024 Volume 22 Issue 01

AN OVERVIEW OF USING MACHINE LEARNING TO FORECAST RISK AND INFORM, SENTENCING DECISION

[1]Ms. Banupriya G, [2]Mr. Prasanna Reddy. V
[1]Research Scholar, Muit Lucknow banupriya.gnanasundaram@gmail.com
[2]III Yr B.B.A LL. B(HONS), Crescent School of Law prasannareddyr6@gmail.com

ABSTRACT

In recent years, there has been significant discourse regarding the representativeness of the judiciary as an institution. Research suggests that judges are somewhat responsive to public opinion, but there remains a misconception regarding how the public influences the courts. Courts are tasked with considering the future, estimating the likelihood and severity of unlawful behaviour, and, within specified boundaries, imposing sentences to mitigate potential harm. Ideally, these forecasts should be highly accurate and based on practical, transparent methods that account for the consequences of prediction failures. However, there is often ambiguity about the best approach to achieving these objectives. Subjective judgment, often referred to as "clinical judgment," relies on intuition and experience. However, the resulting risk assessments can be highly inaccurate, and the reasoning behind them may not be clear. On the other hand, "actuarial" strategies utilize data to establish connections between "risk factors" and various outcomes of interest. Regression statistical methods have traditionally dominated actuarial risk assessments, yielding generally positive results. Nevertheless, with the increasing availability of vast datasets and advancements in data analysis technologies, machine learning is poised to become the primary statistical driver in this field, offering the potential for even greater improvements in the future.

INTRODUCTION

When concerns about public safety are considered in sentencing choices, projections of "future dangerousness" must be established. Forecasts are sometimes effectively mandated. There is already considerable and persuasive literature in statistics and computer science demonstrating that machine learning statistical algorithms forecast at least as accurately, and often more precisely, than older methodologies derived from various types of regression analysis. The experience in a variety of criminal justice situations is consistent.[1] Claims to the contrary are grossly inaccurate. [2] Machine learning-based forecasting approaches present a unique opportunity for decision-makers to obtain more accurate, transparent, and appropriately reviewed predictions of criminal conduct. Although such estimates are simply one sentencing aspect, relying on more traditional forecasting methodologies may disadvantage decision-makers who are devoted to harm prevention and evidence-based sentencing. The machine

[1] Richard Berk, Criminal Justice Forecasts of Risk: A Machine Learning Approach (2012).
[2] 2See Richard Berk & Justin Bleich, Forecasts of Violence to Inform Sentencing Decisions, 30 J. Quantitative Criminology 79–96 (2014).

learning approach is presented in the following pages, with special emphasis on a machine learning procedure known as "random forests.[3] Random forests have already been used successfully in several difficult criminal justice forecasting exercises and will serve as an outstanding instructional tool here. The software is widely available. In the United State('US'), AI is already used in the processing of bail applications, DNA analysis of crimes, gunshot detection, and crime forecasting.[4]

Analysis of US, French, Israeli, United Kingdom ('UK'), and Chilean courts, for example, finds that in various settings, the tone of words used in the first three minutes of a hearing, the frequency of birthdays, the outcomes of sporting events, and even the time of day of a hearing or the defendant's name, affect the outcome of cases. These approaches can detect both conscious and unconscious biases. The research of 18,686 judicial opinions compiled over seventy-seven years by the twelve US circuit courts (also known as courts of appeals or federal appellate courts) revealed that judges exhibit significant prejudice prior to national elections.[5] Similarly, fresh research suggests that sequencing matters in high-stakes decisions: decisions taken in prior cases influence the results of subsequent cases, even if the cases are unrelated. Refugee asylum judges are two percentage points more likely to deny asylum to refugees if their previous decision had granted asylum.[6] AI systems offer huge potential to improve the legal system in India. Human capacity is already recognized as a significant constraint in the system. India has only nineteen judges per million people and twenty-seven million (2.7 crore) pending cases.[7] The justice system has already made significant progress in adopting information technology, releasing enormous amounts of data to court users and encouraging them to use electronic systems. However, legislative, institutional, and resource constraints have limited their overall influence. We claim that combining Machine Learning methods with newly available legal data provides a framework for identifying biases in judicial behaviour and proposing real-time fixes. This can lead to a more streamlined system and a reduction in backlog. Such techniques can detect prejudice and bias even when it is not obvious to court participants, so increasing the judiciary's legitimacy.[8]

FORECASTING MACHINE LEARNING IN THE CRIMINAL JUSTICE SYSTEM

Over the past 15 years, India's courts have made considerable strides in incorporating information technology systems. The e-courts project, which began in 2005, is a big undertaking. The Supreme Court of India carried out the "National Policy and Action Plan for Implementation of Information and Communication Technology (ICT) in the Indian Judiciary." The e-court initiative brought technology into Indian courts in a variety of ways. The system's primary differentiating feature was the employment of technology in courtrooms. Judges were given LCD touchscreen devices, and displays and projectors were linked together

[3] See Leo Breiman, Random Forests, 45 Machine Learning 5–32 (2001).

[4] Rigano (n 3) 7; WJ Epps Jr and JM Warren, 'Now Being Deployed in the Field of Law' 59(1) The Judges' Journal 16-39.

[5] C Berdejo and DL Chen, 'Electoral Cycles Among US Courts of Appeals Judges' (2017) 60(3) The Journal of Law and Economics 479-496.

[6] DL Chen, TJ Moskowitz and K Shue, 'Decision Making under the Gambler's Fallacy: Evidence from Asylum Judges, Loan Officers, and Baseball Umpires' (2016) 131(3) The Quarterly Journal of Economics 1181-1242.

[7] Amirapu (n 8); D Damle and T Anand, 'Problems with the e-Courts Data' (2020) National Institute of Public Finance and Policy Working Paper 314 accessed 16 August 2021.

[8] K Kannabiran, 'Judicial meanderings in patriarchal thickets: Litigating sex discrimination in India' (2009) 44(44) Economic and Political Weekly 88-98; M Galanter, Competing Equalities:

via a local network to communicate information with lawyers. Electronic boards were also installed in courts to display information. To better data management, electronic filing procedures and an online case management system were developed. This involved scanning historical cases, developing digital archives, and setting up direct electronic connection with litigants. These investments resulted in the creation of the National Judicial Data Grid, a database of 27 million cases that allows court users to review pending cases and acquire information from prior sessions.[9]

Typically, the aim is to predict various forms of human misbehaviour, which can range from specific behavioural incidents like a homicide arrest to broader categories such as any criminal arrest. These instances may also encompass undesirable actions that aren't necessarily criminal, like failing to report to a parole office for a drug test. Usually, the emphasis in communication revolves around highlighting negative outcomes, but there's also room for projecting positive results. For instance, when someone is expected to successfully complete probation, it's considered a "success" forecast. These success forecasts can lead to less intrusive or costly interventions in the criminal justice system, or even no intervention at all, which can help mitigate excessive imprisonment. Moreover, forecasts aren't limited to binary outcomes. A significant aspect is that machine learning forecasting algorithms can be fine-tuned explicitly to enhance forecasting accuracy. It's normal for some forecasting inaccuracies to occur, and how these errors are handled depends on the forecasting method used. Forecasting behaviours into different classes offers the advantage of straightforwardly incorporating the relative costs of various types of forecasting errors.

Consider this scenario: an individual under probation review is expected to remain crime-free, but later commits a murder. In this case, there's likely been a missed opportunity to prevent the homicide. Conversely, imagine another scenario where an individual under consideration for probation commits a homicide but subsequently refrains from further criminal activity. Here, resources were probably misallocated, resulting in unintended harm to the perpetrator. The consequences of two types of forecasting errors—false negatives and false positives—differ significantly.

In essence, projections should factor in "asymmetric" costs as necessary, a capability often present in machine learning classification algorithms. Conventional methods frequently assume equal costs for false negatives and false positives, despite stakeholder preferences indicating otherwise. Consequently, forecasts can sometimes be highly misleading.

MACHINE LEARNING: BUILDING AND EVALUATING IN INDIAN COURTS

The legal data released by the Indian judiciary is extensive, disorganized, and intricate.[10] Typically, cases include identifiable tags for important dates (such as filing and order dates), key individuals (petitioners, respondents, judges), and court names. However, crucial details like the type of case, outcome of proceedings, and relevant legal citations are often obscured

---

[9] The e-courts data is public and can be accessed via the district court websites, the e-courts Android/iOS app, or the district court services webpage.

[10] Damle and Anand (n 9).

within the body of orders or judgments. Therefore, cleaning and preprocessing this data is essential for any form of analysis, particularly for supervised algorithms trained on this dataset.

Traditional empirical legal studies have typically dealt with this challenge by relying on small-scale datasets, manually coding legal variables, and limiting the scope of analysis to a narrow range of legal cases relevant to a single issue. A major hurdle in preprocessing this data is the inconsistency in reporting practices across states and districts. The data quality varies greatly, lacking a nationally standardized system for defining variables or reporting on them. For example, some states provide clear delineation of act names and section numbers, with a higher proportion of cases having uploaded orders, while others do not. This discrepancy makes it challenging to compare case types across courts and states.[11]

Moreover, there are no standardized identifiers within the data to track a case through its potential appeals in higher courts. Similarly, tracing a criminal case from its initiation as an FIR to its resolution as a judgment is not straightforward. Issues also arise with identifying information about participants, their characteristics, and the relevant laws or acts pertaining to the case. Incorrect reporting and spelling mistakes further compound these challenges, sometimes resulting in entries from one field appearing in another, necessitating meticulous cleaning and systematic recoding of variables.

To address these issues, various machine learning (ML) tools can be utilized to enhance data quality. A robust pipeline has been developed to scrape, clean, and prepare this data for analysis. Below, we briefly outline some of these methods.

A. Inference About the Identity of Participants
B. Identification of Laws and Acts
C. New Interpretations of Text
D. Identification of Discrimination and Bias
E. Identification of Causal Effects of Legal Rulings

RANDOM FORESTS AS A MACHINE LEARNING IILUSTRATION

Random forests, a machine learning technique, has demonstrated significant effectiveness as a predictive tool in criminal justice contexts. While other machine learning methods can perform similarly well, random forests offer a wide range of output features that greatly aid in interpretations and are relatively straightforward to explain. The process of random forests involves two stages. Firstly, a large number of classification trees are constructed, with a common default being 500 trees. In a typical criminal justice scenario, each tree represents a partitioning of the training data into different case profiles, each associated with a particular outcome class.

For instance, if the outcome to be predicted is a subsequent arrest for intimate partner violence, the random forests algorithm searches for configurations of predictor values linked to such arrests. These configurations, or "portraits," can be visualized as paths down branches of a tree,

[11] Damle and Anand (n 9).

hence the term "classification tree." To facilitate this process, all predictors are treated as binary variables. For example, a predictor like the number of prior arrests might be split into categories like no priors versus one or more priors, with cases assigned to one subset based on which side of the threshold they fall.

Categorical predictors are handled similarly, with the algorithm determining the optimal way to collapse categories into two groups that best associate with the outcome. Each classification tree is grown from a random sample of the training data to account for chance features, and predictor selection within each tree is done sequentially from a small random sample of predictors.[12]

For new observations with unknown outcomes, the known predictor values can be used to estimate the outcomes. In summary, random forests in criminal justice forecasting are a flexible regression procedure for outcome variables with multiple categories. They can account for the asymmetric costs of forecasting errors and automatically generate forecasts from test data. Unlike conventional model-based forecasts, which prioritize explanation, random forests prioritize forecasting accuracy. Therefore, at sentencing, random forests can often provide significantly better forecasts than conventional modelling approaches.

A NOVEL RELATIONSHIP BETWEEN HUMANS AND MACHINES

We believe that the application of machine learning (ML) can significantly enhance the organization and analysis of vast amounts of unstructured data released by the Indian judiciary over the past 15 years. ML algorithms offer the capability to identify participants and analyse their proceedings within the court system. By converting extensive textual data into numerical representations, ML technologies can provide valuable insights into the procedures and outcomes of the justice system. For instance, text analysis can help uncover biases and discrimination, thereby improving the competencies of judges and lawyers, and streamlining review processes.

However, the adoption of these tools in courts must address certain limitations and constraints. Concerns related to data privacy, protection of personally identifiable information, security, and legal data control must be carefully addressed. Additionally, algorithms require preprocessing, training on large and high-frequency datasets, and iterative refinement to ensure their suitability for real-world scenarios. Robust pilot projects, evaluated through randomized control trials, are necessary to gain insights into data privacy, costs, and outcomes. It's crucial that these technologies complement human decision-making rather than replacing it.

One significant challenge is the interpretability of algorithms, often referred to as the 'black-box' problem.[13] Sophisticated algorithms, such as word embedding algorithms, may learn biases present in the data, which can inadvertently influence decision-making processes.

[12] See Leo Breiman, Random Forests, 45 Machine Learning 5–32 (2001).
[13] F Pasquale, The Black Box Society: The Secret Algorithms that Control Money and Information (HUP 2015).

Collaborative and deliberative approaches to technology design, deployment, and evaluation are essential to address these challenges effectively.

AI and ML can enhance human decision-making by providing judges with accurate predictions based on their previous decisions, thereby promoting consistency and reducing the influence of extraneous factors. Furthermore, AI can facilitate decision-making by creating customized communities of experts trained based on data from other experts, potentially spanning different geographic and subject matter contexts. However, it's important to mitigate the risk of 'groupthink' and ensure that judges maintain their individuality in decision-making processes.

Predictive systems for detecting judicial errors can help judges make more suitable decisions by identifying areas where additional assistance may be needed. However, it's essential to ensure that these systems allow for open dialogue between judges and AI, enabling judges to provide confidential information that may not be captured by the algorithms. The interpretability of algorithms is crucial for fostering trust and ensuring justice, as judges may need explanations for algorithmic suggestions.

The progressive integration of AI and ML technologies in courts must be preceded by extensive research and testing to assess their costs and benefits accurately. Randomized controlled experiments can provide valuable insights into the causal effects of algorithm adoption, including cost, efficiency, user satisfaction, and outcomes. Ultimately, responsible and ethical use of AI and ML holds enormous potential for Indian courts, provided that these technologies are implemented thoughtfully and with careful consideration of their implications.

CONCLUSION

Our primary focus has been on using forecasts to assist in making sentencing decisions. The safety of the public, law enforcement personnel, and offenders themselves can all be at risk if a sentence fails to adequately reduce the likelihood of harm. However, overlooking nonviolent defendants who are capable of rehabilitation can also be costly and directly disadvantage such individuals. In both scenarios, the accuracy of forecasting is of utmost importance.

Machine learning offers superior forecasting accuracy compared to traditional methods and unstructured clinical judgment. Actuarial risk assessment tools, including advanced approaches like random forests, are fully compatible with a punishment system based on principles of justice and limited retribution. Predictions of risk derived from machine learning can help judges decide whether to sentence at the upper or lower end of recommended ranges, or even outside those ranges.

In a system based on just deserts, balanced sentencing requires considering the defendant's blameworthiness, proportionality, and likely future behaviour.[14] The integration of accurate forecasting directly aligns with the principles of limited retributivism. This hybrid system employs principles of uniformity and proportionality to establish a sentencing range, with other

[14] See Richard S. Frase, Punishment Purposes, 58 Stan. L. Rev. 67–84 (2005); see also Richard P. Kern & Mark H. Bergstrom, A View from the Field: Practitioners' Response to Actuarial Sentencing: An ''Unsettled Proposition,''25 Fed. Sent'g Rep. 185–89 (2013).

principles providing additional refinement, including deterrence, incapacitation, rehabilitation, and parsimony.

Alternative risk assessment approaches, while sometimes appealing in their simplicity, often share or exacerbate the limitations of machine learning. At worst, they may sacrifice statistical accuracy, responsiveness to stakeholder preferences, and transparency.[15] Disregarding the potential value of modern forecasting methods is a policy decision that could lead to a deliberate disregard for procedures that could save lives and reduce the costs of unnecessary incarceration for individuals and the criminal justice system as a whole. When applied to sentencing decisions, whether for individuals deemed dangerous or those suitable for diversion programs, machine learning offers a promising approach.

[15] See Frase, supra note 27, at 68