
Ritika Ranka
PhD Scholar, Gujarat National Law University (GNLU), Gandhinagar, Gujarat

ABSTRACT

The principle that ‘everything that is not forbidden is allowed’ is linked to the principle of legality, which means that there can be no crime or punishment without criminal law. With the disruption of artificial intelligence, algorithmic decision-making in criminal law has become significant. Through a systematic literature review, the authors investigate different tools used in automated decision-making in varied jurisdictions and their prediction techniques to understand the public perception of crime analysis, ethical grounds and limitations. The paper would contribute to assessing the concerns raised regarding the potential for bias in these systems exist. Furthermore, the question about an individual's accountability (mens rea) concerning criminal conduct (actus reus) vis-à-vis an algorithmic agent’s accountability in predicting crime and delivering a verdict would determine the preparedness for adopting artificial intelligence in the criminal justice system. The literature review will utilise RepOrting standards for Systematic Evidence Syntheses (ROSES) standards of publication involving articles and research papers from three databases: SCOPUS and Web of Science. The paper intends to highlight the gaps for further improvement of artificial intelligence algorithms in the decision-making under the criminal justice system for future researchers.

Key Words: Automated Decision-Making, Criminal Law, Systematic Literature Review, Crime Prediction, Accountability of AI.

1 Introduction

The technological innovations and the advent of artificial intelligence tools and techniques could diminish the problem of delayed justice in the judicial system. The effects of delay in justice majorly arise from securing the accused's presence, non-availability of counsel, staying by the courts, etc. The courts have been following standard protocols for delivering justice, however, with the focus on artificial intelligence, there is a possibility of algorithmic tools to predict judicial decisions. Therefore, this paper aims to study the systematic literature review to identify the approach and the accuracy/ performance of different automated decision-making tools.

2 SYSTEMATIC REVIEW – THE ROSES REVIEW PROTOCOL

ROSES provide “the research synthesis community with detailed reporting standards tailored to the field. After carefully considering systematic mapping, narrative and qualitative synthesis methods when producing ROSES, these developments as being equally important for other fields where reporting standards in systematic reviews are needed.”¹ For doing a systematic literature review, it is necessary that the below-mentioned steps are followed.

Formulation of Research Questions

For this, defining the research scope is the most important. The framework in Population, Intervention, Comparison, Outcome and Context (PICOC) helps determine the research scope. Formulation of research questions happens only when the scope of research is defined.

<table>
<thead>
<tr>
<th>Population</th>
<th>Tools and Techniques using AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>Automated Decision Making</td>
</tr>
<tr>
<td>Comparison</td>
<td>Efficient and Accurate</td>
</tr>
<tr>
<td>Outcome and Context</td>
<td>Crime-based Judicial Decisions</td>
</tr>
</tbody>
</table>

After defining the research scope, the author has framed the following research questions:

**RQ1:** What prediction has been made by AI tools in decision-making?

**RQ2:** What are the different tools and techniques for automated decision-making in criminal laws?

**RQ3:** How accurate were the AI tools in decision-making?

**Searching Strategy and Search String**

The search string and relevant databases identify the relevant area of research. Generally, the following syntax is used:

**TITLE-ABS-KEY**

The TITLE-ABS-KEY search performs a text search using the TITLE+ABSTRACT+KEYWORDS fields as one, combining all three fields into one. As a result, the result will return the paper if it includes (for instance) "climate refugee displacement" in its title and "global warming" in its keywords. The three fields are independently searched by TITLE() or ABS() or KEY(). The identical sample paper wouldn't be returned as a result.

<table>
<thead>
<tr>
<th>Database</th>
<th>Search String</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scopus</td>
<td>(&quot;predict*&quot; OR &quot;prediction*&quot; OR &quot;predicting*&quot; OR &quot;forecast&quot;) AND (&quot;court decision*&quot; OR &quot;legal decision*&quot; OR &quot;judicial case&quot;) AND (&quot;machine learning*&quot; OR &quot;artificial intelligence*&quot; OR &quot;AI*&quot; OR &quot;supervise* machine learning*&quot;))</td>
</tr>
<tr>
<td>Web of Science</td>
<td>(&quot;predict*&quot; OR &quot;prediction*&quot; OR &quot;predicting&quot;) AND (&quot;court decision*&quot; OR &quot;judicial decision*&quot; OR &quot;legal decision&quot;) AND (&quot;machine learning*&quot; OR &quot;AI*&quot;))</td>
</tr>
</tbody>
</table>

The author has put the filter of open-access journals for convenience.

**The inclusion and exclusion criteria**

Grey literature, extended abstracts, presentations, keynotes, review articles, and publications written in languages other than English were mainly left out. Even if automated court decisions are not included in the title, keywords, or abstracts of certain publications or articles, they may nevertheless be covered in such pieces. These papers, however, were excluded from the evaluation because they were outside of its purview.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timeline</td>
<td>2000-2023</td>
<td>Before 2000</td>
</tr>
<tr>
<td>Language</td>
<td>English</td>
<td>Non-English</td>
</tr>
<tr>
<td>Methods</td>
<td>Automation Techniques Used</td>
<td>Other than Automation Technique</td>
</tr>
</tbody>
</table>

**Research Databases**

The flow chart showed the general screening procedures and the order of choosing pertinent books. A total of 128 records were discovered in the initial phase (96 from Scopus and 32 from Web of Science). The number of literary works was decreased to 117 pieces after the exclusion of works such as grey literature, extended abstracts, presentations, keynotes, book chapters, non-English language papers, and inaccessible publications. After that, only 57 publications met the requirements for additional abstract reading. There were only 36 articles left to read after finishing the article abstract. 08 duplicate papers and articles were manually deleted during the reading of the main body. In the end, 24 publications were left that met every requirement for inclusion in this SLR effort.
Figure 1: Flow Chart Diagram of the Screening Process for Articles

Selected Research Articles

The author has selected 24 articles for systematic literature review; from the code of S1 to S24. The research articles are S1: Predicting the Outcome of Construction Litigation Using Boosted Decision Trees\(^2\) (2005), S2: Predicting the Outcome of Construction Litigation Using Particle Swarm Optimisation\(^3\) (2005), S3: Prediction of Construction Litigation Outcome - A CBR Approach\(^4\) (2006), S4: Predicting the Outcome of Construction Litigation Using an Integrated AI Model\(^5\) (2010), S5: Litigation Outcome Prediction of Differing Site Condition Disputes Through Machine Learning Models\(^6\) (2012), S6: Study of Termination of Parental Rights: An Analysis


decision-making is automated\(^{24}\) (2023).

**Publication Years**

The studies that were picked were released between 2000 and 2023. However, the first study on this subject to be published was in 2005. The quantity of research that was released within the given time period is shown in Figure 2. According to the findings, numerous studies were found to have been published in the previous ten years. Consequently, the automated decision-making method can serve as one of the strategies for enhancing the legal system by foreseeing outcomes.

**Figure 2: Publication Years of the Articles**

![Publication Years of the Articles](image)

**Synthesis and Analysis**

The research questions and their analysis are discussed in this section. The first research question was: (RQ1) What prediction has been made by AI tools in decision-making? In the world of the legal system, making a decision involves many different tasks that must be taken into account. The legal system is challenging for the general public to comprehend because it involves interactions with lawyers, the hiring of attorneys, decisions about how to proceed with cases, the effects of those decisions, and the meanings of words in case files\(^ {25}\). This study looked into the use of machine learning to anticipate judicial judgements in court. The prediction can take many different forms, such as foreseeing the outcome of a court case or the charges that call for multilabel text classification. Charges, penalty terms, and fines are comprehensive and complex subclauses frequently appearing in several subtasks in judicial judgements.\(^ {26}\) However, the majority of studies only tested a binary task, which only recognises two alternative outcomes. Several nations that use the civil law system, including Germany, France, and China, considered the prediction of pertinent articles to be a crucial subtask that directs and supports the prediction in addition to predicting the outcome of judicial judgment\(^ {27}\).

Seven research publications that discussed projecting the outcome of construction lawsuits were located in this SLR. Arditi and Phulket\(^ {28}\) highlighted that, specifically concerning major contracts, construction litigation is common in many construction projects. Miscommunication, inadequate specifications and plans, strict contracts, modifications to the site’s conditions, non-payment, catch-up profits, a small workforce, inadequate tools and


\(^{27}\) Ibid.

equipment, ineffective supervision, notice requirements, constructive changes that the owner does not acknowledge, delays, and accelerated measures are some of the factors that provoke claims and lead to disagreements. Therefore, Arditi and Phulkert\(^\text{29}\) presented a technology to forecast the outcome of litigation to reduce construction problems brought on by issues that are difficult to resolve amicably\(^\text{30}\).

Due to the high expense of the litigation process and the complexity of the problems involved, legal action needs a higher settlement fee. Additionally, the dispute between the client and the contractor could harm each party's reputation. Additionally, legal action is time-consuming for complex building issues and, depending on the jurisdiction, may take two to six years before trial\(^\text{31}\). As a result, the researchers suggest a number of machine learning techniques to guarantee the precision of outcome prediction in court cases. The techniques effectively cut down on the frequency of disagreements that result in increased expenditures due to the litigation process.\(^\text{32}\)

The results of the present investigation showed that nine research papers predicted the resolution of criminal cases. However, there are a few categories into which crime-related instances can be split. Aletras presented the first systematic study that used textual analysis to forecast case outcomes in the European Court of Human Rights.\(^\text{33}\) Based on text taken from prior cases, the authors divided the predicted outputs into "violation" and "non-violation." By increasing the number of publications and using new factors with the same dataset, further research were carried out.\(^\text{34}\) This idea can be helpful to solicitors and judges as a supplementary tool to locate cases and extract text that directs judgement.

Luo\(^\text{35}\) said that the skill of textual fact analysis is essential for legal assistant systems where non-legal experts can locate analogous cases or potential punishments by describing a case with their own words and comprehend the legal foundation of their search cases. Furthermore, Luo\(^\text{36}\) provided a standard way to forecast charges and retrieve pertinent articles in a unified framework using an attention-based neural network. The results showed that by offering related articles, charge prediction results could be improved and charges for situations with various expression styles could be efficiently anticipated.

Zhong \textit{et. al.}\(^\text{37}\) by asserting that prior studies only developed approaches for specific subtasks set and found it challenging to scale to other subtasks despite being developed to simultaneously predict law articles and charges, the author proposed a different approach in modelling the judgement prediction framework that makes use of multiple subtasks. It also concentrated on cases involving murder by conducting such analysis. To find the specifics of case-specific legal variables, extraction of legal judgement can be used, but it is time-consuming and difficult job. As a result, crucial elements that will influence the prediction for cases involving murder are assessed by creating a dataset to identify the elements as descriptors for prediction results. The outcome prediction is seen as a binary classification for the accused person's "acquittal" and "conviction."

The results of the current study are further explored in situations that don't involve civil law and concentrate primarily on family law disputes. The cases that are highlighted include dowry, parental rights, divorce, and disengagement. Ben-David\(^\text{38}\) conducted a significant study on the balance between the best interests of the child, the parent's right, and the privacy of the family unit in court decisions in favour or against terminating parental

\(^{29}\) \textit{Ibid.}


\(^{36}\) \textit{Ibid.}


The study of algorithms and statistical models used in artificial intelligence, or "machine learning," allows the system to automatically learn from test data and enhance its performance. Information extraction and analysis from already existing legal texts are the main study facets of utilizing machine learning in jurisprudence. In the past, all work was performed manually by solicitors and judges. However, machine learning has helped society as a whole become more intelligent by deciphering written documents and extracting the information contained therein.

The second research question was: (RQ2) What are the different tools and techniques for automated decision-making in criminal laws? The focus of Legal professionals is on Artificial Intelligence and its application in the court-based decisions. Using historical datasets to inform judicial decisions is not a novel concept and is currently being used extensively throughout the world's legal systems. The study of algorithms and statistical models used in artificial intelligence, or "machine learning," allows the system to automatically learn from test data and enhance its performance. Information extraction and analysis from already-existing legal texts are the main study facets of utilizing machine learning in jurisprudence. In the past, all work was performed manually by solicitors and judges. However, machine learning has helped society as a whole become more intelligent by deciphering written documents and extracting the information contained therein.

The third research question was: (RQ3) How accurate were the AI tools in decision-making? Before comprehending the method utilised, the performance of the prediction model should be evaluated. Through k-fold cross-validation, accuracy, sensitivity, specificity, recall, precision, and F-measure, any machine learning model's efficacy can be assessed. Based on the findings from the 24 publications that were analysed, the majority of researchers evaluated the performance of their models using accuracy, precision, recall, and F-measure. The SLR's most surprising discovery was that 16 of the 24 review papers it had chosen had achieved over 80% accuracy.
precision, or prediction rate during the evaluation process. Only four publications attained accuracy or precision in the 50% to 70% range. In contrast, four studies didn’t go into much detail about how well their prediction models worked.

CONCLUSION
This paper has given an examination into applying machine learning techniques to anticipate court judgements. Numerous cases and the findings of the research can be used to demonstrate the value of anticipating judicial decisions. The legal system can be improved with this strategy by becoming more organised and dependable. Future scholarly work in the field of the study may be able to fill any gaps using the methodologies and features derived from the findings. This systematic review study is anticipated to add to the body of knowledge by offering an overview of the models now in use for forecasting judicial judgements, an analysis of how well the model performs, and a discussion of various sorts of legal cases that have used this strategy.