IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 6, December 2026, pp. 5559~5567
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i6.pp5559-5567

Data-driven framework for sentencing consistency in Indonesian traffic courts

Mamok Andri Senubekti[1,4], Pujiyono[2], Foni Agus Setiawan[3]
[1]Doctoral Program of Information System, School of Postgraduate Studies, Diponegoro University, Semarang, Indonesia
[2]Department of Law, Faculty of Law, Diponegoro University, Semarang, Indonesia
[3]Research Center for Data and Information Sciences, National Research and Innovation Agency, Bandung, Indonesia
[4]Department of Informatics Management, Digitech University, Bandung, Indonesia

Article Info

Article history:

Received Jul 18, 2025
Revised Aug 18, 2026
Accepted Sep 20, 2026

Keywords:

Explainable artificial intelligence
Judicial decision support system
Sentencing disparity
Shapley additive explanations
XGBoost

ABSTRACT

Sentencing disparity poses a critical threat to judicial trust, particularly in high-volume jurisdictions like Indonesia's traffic courts. While judicial decision support systems (JDSS) exist, they often lack transparency, creating a "black box" problem that hinders practical adoption. This study addresses this gap by developing a novel, data-driven framework that integrates high-performance predictive modeling with explainable artificial intelligence (XAI) to audit sentencing consistency. By analyzing approximately 500,000 court records from the Bale Bandung District Court, an optimized extreme gradient boosting (XGBoost) regressor was trained and achieved strong predictive performance (coefficient of determination (R^2) ≈ 0.68; root mean squared error (RMSE) = IDR 16,741). Unlike prior works that solely predict outcomes, Shapley additive explanations (SHAP) were utilized to mathematically quantify the drivers of sentencing variance. The results reveal a significant discovery: the presiding judge's identity is frequently a more dominant predictor of fine amounts than the violated legal article, providing empirical evidence of systemic disparity. Based on these findings, the proposed framework introduces a "consistency audit" mechanism designed not to replace judicial discretion, but to serve as a reflective, data-informed instrument for judges. This research contributes a responsible artificial intelligence architecture for civil law jurisdictions, demonstrating how XAI can transform raw judicial data into actionable insights for a fairer and more transparent judiciary.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Mamok Andri Senubekti
Doctoral Program of Information System, School of Postgraduate Studies, Diponegoro University
Semarang, Indonesia
Email: mrymodion@students.undip.ac.id

1. INTRODUCTION

The fundamental pillar of a judicial system grounded in the rule of law is public trust, a quality heavily dependent on the consistency and predictability of sentencing [1], [2]. In an ideal legal framework, analogous cases should yield comparable penalties to ensure equitable justice [3]. However, realizing this ideal presents a monumental challenge, particularly in high-volume jurisdictions where judges contend with immense caseload pressures [4], [5]. Such administrative burdens often precipitate cognitive fatigue, leading to unintentional sentencing disparity—a phenomenon where identical offenses receive significantly divergent penalties, thereby eroding the integrity of the justice system [3], [6].

This challenge is acute in Indonesia, particularly within the context of *perkara tilang* or traffic violations. As the highest-volume category of cases in district courts, traffic violations create an

administrative bottleneck that severely hampers judicial efficiency. The Bale Bandung District Court alone, for instance, processed nearly 500,000 traffic records over seven years. In such a high-velocity environment, the risk of erratic sentencing increases, potentially compromising public confidence in the judiciary [1], [7]. While the sheer scale of data generated by these proceedings offers a unique vantage point for analyzing judicial behavior, it remains largely underutilized for monitoring consistency compared to jurisdictions in developed nations [4], [8].

To mitigate efficiency challenges, the field of legal informatics has advanced various judicial decision support systems (JDSS) and legal judgment prediction (LJP) models [9], [10]. Although recent strides in machine learning (ML) have empowered these systems to predict outcomes with remarkable accuracy, their practical integration into courtrooms is stalled by the "black box" problem [11], [12]. Algorithmic models that render predictions without transparent justifications are fundamentally incompatible with judicial ethics, which mandate reasoned and explicable deliberations [13], [14]. As Chesterman [12] argues, opacity in artificial intelligence (AI) systems can obscure latent errors and biases, rendering them perilous for high-stakes decision-making.

Despite the proliferation of LJP literature, a critical research gap persists regarding the integration of explainable artificial intelligence (XAI) specifically for auditing sentencing disparity in civil law jurisdictions like Indonesia [15], [16]. The majority of existing studies, such as [9], [17], predominantly focus on maximizing prediction accuracy (e.g., guilty vs. not guilty) or forecasting judicial disagreement in common law contexts [18]. Scant attention has been paid to quantifying the subtle variance in sentencing amounts (regression) attributable to judicial discretion using interpretability tools like Shapley additive explanations (SHAP) [19], [20]. Furthermore, frameworks that visually decouple the influence of a judge's identity from objective legal facts remain rare [21], [22]. The absence of a holistic framework combining high-fidelity predictive modeling with interpretability tools limits the capacity of court administrators to diagnose and rectify systemic inconsistencies [15], [23].

Bridging this gap, this study proposes a novel, data-driven framework designed to enhance sentencing consistency. Diverging from traditional automation tools, this framework operates as a reflective audit instrument. The extreme gradient boosting (XGBoost) algorithm [24], in tandem with SHAP [20], is employed to analyze approximately 500,000 court records to not only predict fines but also mathematically quantify the drivers of sentencing variance. This research delivers two primary outcomes: i) it provides empirical evidence that the presiding judge's identity frequently exerts a more dominant influence on fines than the violated legal article, and ii) it introduces an XAI-driven consistency audit mechanism that empowers judges to benchmark their decisions against historical consensus, thereby fostering a more transparent and accountable judiciary [1], [23].

Building on these gaps, this study makes three distinct contributions that differentiate it from existing legal AI systems. First, whereas most prior work frames legal AI as an outcome-prediction problem, this study reframes the predictive model as an auditing instrument: the objective is not to recommend sentences but to quantify the extent to which sentencing outcomes deviate from the legal facts of a case. Second, it introduces a judge-level disparity metric derived from SHAP attributions, which converts the abstract notion of sentencing inconsistency into an interpretable monetary quantity attributable to individual adjudicators—a measurement that conventional feature-importance analyses cannot provide. Third, it demonstrates this framework at scale on approximately 500,000 authentic records from an Indonesian civil law jurisdiction, a judicial context that remains markedly underrepresented in the legal AI literature relative to common law systems. Together, these contributions position the framework as a supervisory and reflective tool for judicial oversight bodies rather than as an automated decision-maker.

2. METHOD

This study employs a quantitative approach. It follows the knowledge discovery in databases (KDD) process. The proposed framework, depicted in Figure 1, integrates advanced machine learning regressors with post-hoc explainability techniques to audit sentencing consistency. The pipeline in this figure proceeds from data curation (phase 1) to comparative modeling using XGBoost and benchmarks (phase 2), and culminates in an XAI-driven disparity audit using SHAP (phase 3).

2.1. Data curation and feature engineering

The dataset was obtained from the case tracking information system (*sistem informasi penelusuran perkara* (SIPP)) of the Bale Bandung District Court, covering seven years (2016–2023). The raw dataset comprised 500,488 traffic violation records. To ensure data integrity, a rigorous cleaning pipeline was executed to handle missing values, inconsistent currency formats, and typographic errors in legal text.

Feature engineering was critical for capturing judicial nuances. The target variable is the fine amount (*denda*), modeled as a continuous regression task. The predictor variables include: i) legal article (*pasal*): the specific traffic regulation violated (e.g., article 281, 287); ii) vehicle type: categorical data representing the vehicle involved (e.g., motorcycle, truck). This was encoded using one-hot encoding; iii) offender age: continuous variable to detect potential bias against specific age groups; iv) presiding judge: a high-cardinality categorical variable representing the specific judge adjudicating the case. To preserve the individual identity of the judge for disparity analysis, label encoding was applied. This allows the model to learn specific sentencing patterns associated with each judge ID. This large-scale, real-world dataset provides a robust empirical foundation to analyze sentencing patterns and develop the proposed data-driven framework.

2.2. Predictive modeling framework

Three regression algorithms of increasing flexibility were trained on identical data partitions, so that any gain in accuracy could be attributed to model capacity rather than to differences in preprocessing:

- Linear regression (baseline): included to establish how much of the variation in fine amounts can be explained by an additive, strictly proportional relationship among the legal predictors.
- Random forest regressor: an ensemble of decorrelated regression trees whose averaging behavior restrains variance while still accommodating conditional dependencies among predictors.
- XGBoost: the principal model of this study. XGBoost was adopted because gradient-boosted trees remain the strongest performers on heterogeneous tabular data of this scale [24]. Its controlling parameters (learning_rate, max_depth, n_estimators) were tuned against out-of-sample squared error.

2.3. Explainable artificial intelligence architecture

Addressing the "black box" limitation in legal AI, SHAP were integrated into the modeling pipeline. Rooted in the Shapley value of cooperative game theory, SHAP apportions the gap between a single prediction and the dataset baseline among the predictors that produced it [20]. Such attribution answers the explainability obligation increasingly imposed on automated decision-making in legal settings, where an unexplained output cannot satisfy fundamental ethical requirements [25].

Conventional impurity-based importance scores summarize a model only in aggregate. SHAP was preferred here because it operates at two levels that this study requires:

- Global interpretability: ranking predictors by their mean absolute attribution answers whether the adjudicator weighs more heavily than the statute.
- Local interpretability: decomposing a single case shows precisely which attributes pushed its fine above or below the expected amount, and by how many rupiah.

Mathematically, the SHAP value ϕ for a feature i is obtained by averaging its marginal contribution over every subset of the remaining predictors, weighted by the number of orderings each subset represents. Because the resulting values are signed and expressed in rupiah, they reveal not merely that the adjudicator matters but in which direction and by what magnitude a given judge departs from the court-wide baseline.

2.4. Performance evaluation

Model accuracy was assessed on a held-out partition comprising one-fifth of the records, using three complementary indicators: the coefficient of determination (R^2) for the share of variance explained, mean absolute error (MAE) for the typical rupiah discrepancy between predicted and imposed fines, and root mean squared error (RMSE), which weights large deviations more heavily and is therefore the stricter criterion for an auditing application. The complete experimental configuration is documented below so that the reported figures can be regenerated exactly. Records lacking a value in any of the five modeling variables were dropped listwise, leaving 500,487 usable cases out of the 500,488 retrieved. The two categorical predictors were treated differently by design: vehicle type was expanded into indicator columns with the first level omitted to keep the design matrix full rank, while the adjudicator was mapped to an integer code so that each judge remained a distinguishable entity throughout the audit. The final feature space therefore comprised the judge code, the legal article, the offender's age, and the one-hot vehicle-type indicators, with the fine amount as the continuous target. Partitioning was randomized once under a fixed seed and reused unchanged by every algorithm, so the three models were fitted and scored on the same cases.

The tuned gradient-boosting regressor comprised 500 boosting rounds at a shrinkage of 0.05, trees limited to depth 7, and 80% subsampling of both rows and columns at each round; the forest benchmark comprised 100 trees. At half a million cases, the sampling error of a single partition is negligible, so repeated cross-validation was not pursued. The pipeline was written in Python, drawing on scikit-learn for the baselines, the XGBoost package for the boosted model, and the SHAP package for post-hoc attribution. This unified experimental setup ensured a consistent basis for comparing predictive performance and interpreting the factors influencing sentencing outcomes.

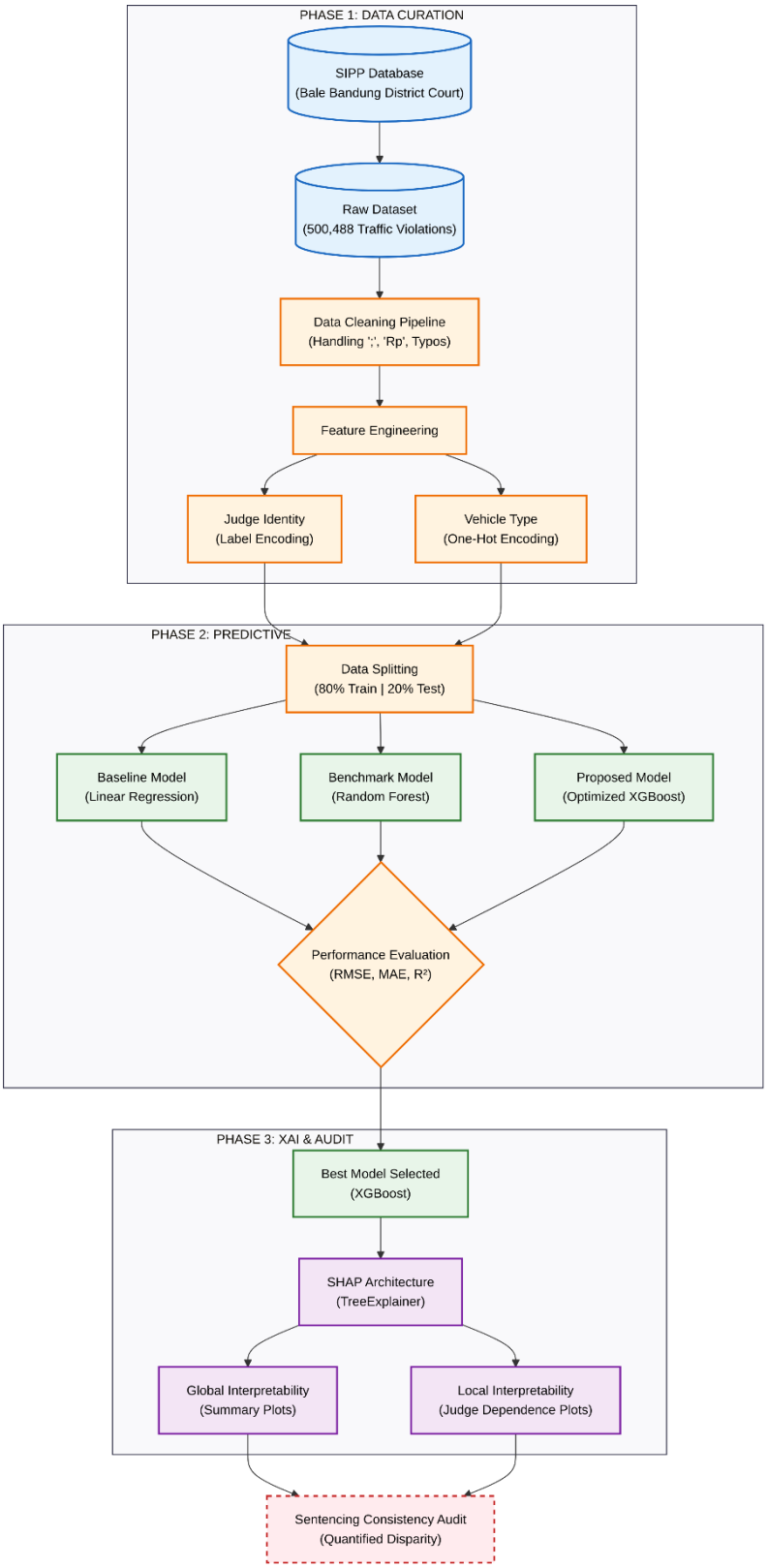

Figure 1. The proposed research workflow using the KDD framework

3. RESULTS AND DISCUSSION

This section presents the empirical findings of the proposed framework. It is structured into three parts. First, a comparative analysis of predictive performance to justify the choice of the XGBoost algorithm.

Second, an XAI-driven investigation into the global drivers of sentencing. Third, a granular audit of judicial disparity using SHAP dependence analysis.

3.1. Predictive performance and model selection

The proposed XGBoost was set against a linear baseline and a random-forest benchmark in order to establish whether the added model complexity is warranted by the data. Table 1 reports the resulting metrics. As presented in Table 1, the linear regression model yielded suboptimal performance ($R^2 \approx 0.32$), indicating that the relationship between legal facts (e.g., articles, vehicle types) and fine amounts is not strictly linear. In contrast, the tree-based ensemble models demonstrated a significant leap in predictive capability. This pattern is consistent with prior reports that tree ensembles outperform classical estimators on judicial outcome data [26]. The proposed XGBoost attained the strongest fit, with an R^2 near 0.68 and the smallest error dispersion at an IDR 16,741 RMSE. This substantial improvement over the baseline confirms that judicial decision-making involves complex, non-linear interactions that simple statistical methods fail to capture. Consequently, the optimized XGBoost model was selected as the core engine for the subsequent interpretability analysis.

Table 1. Model performance

Model	R^2	RMSE (IDR)	MAE (IDR)
Linear regression	0.316704	24410.865084	12059.471522
Random forest	0.672121	16909.698618	7937.291029
XGBoost (optimized)	0.678609	16741.547593	8071.794423

3.2. Unveiling the "black box": global determinants of sentencing

Predictive accuracy is a means rather than an end here; the contribution of this study lies in what the fitted model discloses. To recover the reasoning behind individual fine predictions, SHAP was applied. Figure 2 presents the summary plot, in which predictors are ordered by the mean absolute attribution they receive across the evaluated cases. In this figure, the 'Judge code' exhibits a high impact on the model output, indicating significant influence on sentencing decisions.

The analysis in Figure 2 reveals a critical insight into the judicial process. Surprisingly, the presiding judge's identity (judge code) emerges as one of the most dominant features, often competing with or surpassing the influence of the legal article (*pasal*). In the plot, red dots represent high feature values (e.g., specific judge IDs), while blue dots represent low values. The wide horizontal dispersion of the "judge code" feature indicates that for similar cases, the identity of the judge can swing the predicted fine significantly to the right (higher fine) or left (lower fine). This provides quantitative evidence that sentencing in traffic cases is not solely determined by the rule of law (*pasal*), but is heavily influenced by the discretionary patterns of individual judges.

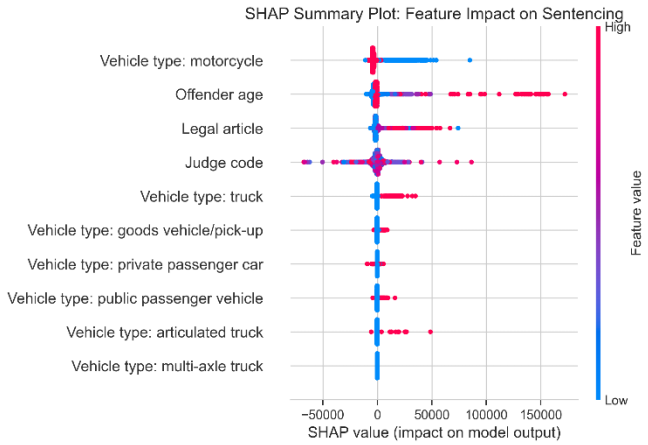

Figure 2. SHAP summary plot revealing the global importance of features

3.3. Auditing judicial disparity: a forensic analysis

A granular forensic audit was performed using SHAP dependence plots to further investigate the extent of the discretionary influence identified in the global analysis. Unlike standard feature importance metrics, which only indicate that a feature matters, the dependence plot isolates how specific values of a

feature (in this case, the unique identity of the judge) drive the model's prediction up or down relative to the baseline. As illustrated in Figure 3, the model reveals significant heterogeneity in sentencing behavior. In this figure, the x-axis represents individual judge codes, while the y-axis represents the marginal impact on the fine (SHAP value in IDR). The visual spread indicates that specific judges (clusters above the zero line) consistently increase the fine amount, while others (below the line) decrease it, quantifying the exact monetary value of judicial discretion. The zero line represents the average prediction; points scattered above this line indicate cases where the judge's identity contributed to a higher fine, while points below indicate a lenient contribution. The plot exposes that certain judges consistently impose fines IDR 20,000–50,000 higher than the court's average for identical offenses, independent of the legal article violated. This quantifiable variance serves as the core metric for the proposed consistency audit framework.

To ensure these AI-driven insights are not artifacts of the model but reflect ground-truth data, the findings were validated using a traditional statistical distribution analysis. The most frequent traffic violation (article 281) was isolated, and the fine distributions were compared across the top 10 most active judges. Figure 4 corroborates the SHAP analysis. This statistical validation in this figure corroborates the SHAP findings in Figure 3, showing observable variations in median fines (box centers) and outliers among different judges for identical traffic violations. Even when controlling for the legal violation (article 281), the median fine amounts vary noticeably among judges. Some judges exhibit a very narrow interquartile range (indicating rigid consistency), while others show wide variance with numerous outliers. This discrepancy validates the "cognitive noise" theory [3], confirming that factors extraneous to the case facts—such as the presiding judge's personal sentencing philosophy—significantly affect the outcome. By surfacing these patterns, the framework empowers the Judicial Commission to move from anecdotal observation to data-driven supervision. Nevertheless, consistent with prior cautions regarding algorithmic support in criminal matters, such tools must remain strictly assistive so that they support, rather than replace, human judgment [27]. The institutional adoption of AI-generated analytical insights also raises emerging legal questions—including the economic and moral rights attached to AI-generated works under Indonesian copyright law—that court administrators should anticipate [28].

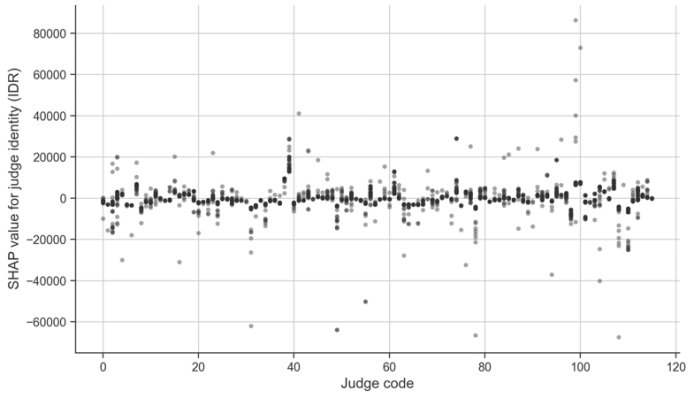

Figure 3. SHAP dependence plot for the 'judge identity' feature

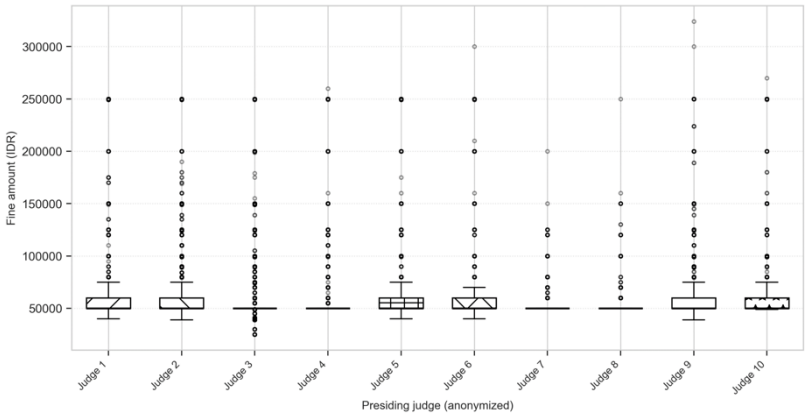

Figure 4. Comparative distribution of fines for article 281 across the top 10 judges

Several limitations should nevertheless be acknowledged. The framework is trained on historical sentencing data from a single jurisdiction and therefore reflects the prevailing practices of that court, including any biases embedded in them; the random hold-out partition also does not model temporal drift arising from evolving regulations. Moreover, the present framework operates exclusively on structured case attributes, whereas recent advances in transformer-based legal language models have substantially expanded the state of the art in legal text understanding and judgment explanation [29], [30]. Positioning consistency auditing within that broader trajectory of legal AI [31], future work should combine the disparity metric proposed here with document-level representations of full verdicts, and extend the evaluation through external validation on independent court datasets together with fairness auditing across demographic subgroups before any institutional deployment.

4. CONCLUSION

This study established a data-driven framework to quantify and audit sentencing disparity in Indonesian traffic courts based on approximately 500,000 records from the Bale Bandung District Court. The optimized XGBoost model substantially outperformed the linear baseline ($R^2 \approx 0.68$; RMSE = IDR 16,741), confirming that sentencing patterns involve complex non-linear interactions, while the SHAP analysis provided empirical evidence that the presiding judge's identity frequently exerts a stronger influence on fine amounts than the violated legal article. By transforming the judicial "black box" into a transparent and auditable process, the proposed framework offers oversight bodies such as the Supreme Court and the Judicial Commission a reflective, data-informed instrument to identify outlier sentencing behavior, standardize practices, and strengthen public trust while preserving judicial discretion. The framework currently relies on structured tabular data; future work should incorporate natural language processing of full verdict documents to capture the reasoning embedded in judicial decisions.

ACKNOWLEDGMENTS

The author would like to express gratitude to the Bale Bandung District Court for the provision of public access to the case tracking information system (SIPP) data, which formed the empirical foundation of this research. During the preparation of this paper, the authors used Claude (Anthropic) for language editing, reference formatting, drafting selected passages, and coding assistance for figure-plotting scripts; all AI-assisted text and code were reviewed and revised by the authors, who take full responsibility for the content. All figures were produced from the authors' own data and models, and no AI tools were used to generate or alter data or results.

FUNDING INFORMATION

The first author's doctoral studies are supported by the *program pascasarjana berbasis riset* (research-based postgraduate program) of the National Research and Innovation Agency (BRIN), Republic of Indonesia, under Decree of the Deputy for Human Resources for Science and Technology No. 16/II.5/HK/2025 dated 11 February 2025. The scheme covers tuition and educational expenses. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Mamok Andri Senubekti	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓		✓	✓
Pujiyono	✓			✓			✓			✓		✓	✓	
Foni Agus Setiawan		✓		✓	✓				✓	✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

This study utilized anonymized data derived from publicly available court records from the Case Tracking Information System (SIPP) of the Bale Bandung District Court. All analyses were performed on aggregated data, and personal identifying information, such as names and specific vehicle registration numbers, was excluded from the final dataset used for modeling. As this research is based on the analysis of public records and does not involve direct interaction with human subjects, obtaining individual informed consent was deemed not applicable.

ETHICAL APPROVAL

This study is based on the analysis of secondary data obtained from publicly available judicial records. The research involves no interaction with or intervention on human subjects. All data were fully anonymized to protect the privacy of the individuals concerned. In accordance with the institutional policies of Diponegoro University and national regulations, this type of research using public, anonymized archival data is exempt from formal ethical review by an institutional review board. Nevertheless, all procedures were conducted in strict adherence to the ethical principles for scientific research.

DATA AVAILABILITY

The data that support the findings of this study are available from the Case Tracking Information System (SIPP) of the Bale Bandung District Court. Restrictions apply to the availability of these data, which were used for this study, as they contain information that could compromise the privacy of research participants. The processed and anonymized data that directly support the model's findings are available from the corresponding author, [MAS], upon reasonable request.

REFERENCES

[1] S. Greenstein, "Preserving the rule of law in the era of artificial intelligence (AI)," *Artificial Intelligence and Law*, vol. 30, no. 3, pp. 291–323, Sep. 2022, doi: 10.1007/s10506-021-09294-4.

[2] J. B. Rajendra and A. S. Thuraisingam, "Artificial intelligence and its impact on the legal fraternity," *UUM Journal of Legal Studies*, vol. 13, pp. 129–161, 2022, doi: 10.32890/uumjls2022.13.2.6.

[3] J. Ryberg, "Sentencing disparity and artificial intelligence," *The Journal of Value Inquiry*, vol. 57, no. 3, pp. 447–462, Sep. 2023, doi: 10.1007/s10790-021-09835-9.

[4] F. F. Vasconcelos, R. M. Sátiro, L. P. L. Fávero, G. T. Bortoloto, and H. L. Corrêa, "Analysis of judiciary expenditure and productivity using machine learning techniques," *Mathematics*, vol. 11, no. 14, Jul. 2023, doi: 10.3390/math11143195.

[5] R. S. de Oliveira, A. S. Reis, and E. G. S. Nascimento, "Predicting the number of days in court cases using artificial intelligence," *Plos One*, vol. 17, no. 5, May 2022, doi: 10.1371/journal.pone.0269008.

[6] F. Borgesano, A. De Maio, P. Laghi, and R. Musmanno, "Artificial intelligence and justice: a systematic literature review and future research perspectives on Justice 5.0," *European Journal of Innovation Management*, vol. 28, no. 11, pp. 349–385, Dec. 2025, doi: 10.1108/EJIM-01-2025-0117.

[7] A. N. Vidaki and V. Papakonstantinou, "Democratic legitimacy of AI in judicial decision-making," *AI and Society*, vol. 40, no. 8, pp. 6025–6035, Dec. 2025, doi: 10.1007/s00146-025-02411-w.

[8] D. Ramachandran and D. R. S. Rana, "AI-driven jurisprudence: navigating legal landscapes in the digital age," *International Journal of Law, Justice and Jurisprudence*, vol. 4, no. 1, pp. 128–139, Jan. 2024, doi: 10.22271/2790-0673.2024.v4.i1b.103.

[9] Y. Feng, C. Li, and V. Ng, "Legal judgment prediction: a survey of the state of the art," in *Thirty-First International Joint Conference on Artificial Intelligence*, Vienna, Austria, 2022, pp. 5461–5469, doi: 10.24963/ijcai.2022/765.

[10] X. Shang, "A computational intelligence model for legal prediction and decision support," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–8, Jun. 2022, doi: 10.1155/2022/5795189.

[11] J. Zeleznikow, "The benefits and dangers of using machine learning to support making legal predictions," *WIREs Data Mining and Knowledge Discovery*, vol. 13, no. 4, Jul. 2023, doi: 10.1002/widm.1505.

[12] S. Chesterman, "Through a glass, darkly: artificial intelligence and the problem of opacity," *The American Journal of Comparative Law*, vol. 69, no. 2, pp. 271–294, Nov. 2021, doi: 10.1093/ajcl/avab012.

[13] W. H. Gravett, "Judicial decision-making in the age of artificial intelligence," *Multidisciplinary Perspectives on Artificial Intelligence and the Law*, vol. 58, pp. 281–297, 2024, doi: 10.1007/978-3-031-41264-6_15.

[14] N. Majeed and A. K. Mahmood, "The story model of judicial decision-making and reasoning with evidence," *Global Legal Studies Review*, vol. 6, no. 2, pp. 7–13, Jun. 2021, doi: 10.31703/glsr.2021(VI-II).02.

[15] K. M. Richmond, S. M. Muddamsetty, T. Gammeltoft-Hansen, H. P. Olsen, and T. B. Moeslund, "Explainable AI and law: an evidential survey," *Digital Society*, vol. 3, no. 1, 2024, doi: 10.1007/s44206-023-00081-z.

[16] A. Tagarelli and A. Simeri, "Unsupervised law article mining based on deep pre-trained language representation models with application to the Italian civil code," *Artificial Intelligence and Law*, vol. 30, pp. 417–473, 2022, doi: 10.1007/s10506-021-09301-8.

[17] N. Sivaranjani and J. Jayabharathy, "Legal logit model for predicting judicial disagreement in Indian courts," *Frontiers in Artificial Intelligence*, vol. 8, Art. no. 1671474, Oct. 2025, doi: 10.3389/frai.2025.1671474.

[18] F. Bex and H. Prakken, "On the relevance of algorithmic decision predictors for judicial decision making," in *Eighteenth International Conference on Artificial Intelligence and Law*, Jun. 2021, pp. 175–179, doi: 10.1145/3462757.3466069.
[19] L. K. Branting *et al.*, "Scalable and explainable legal prediction," *Artificial Intelligence and Law*, vol. 29, no. 2, pp. 213–238, Jun. 2021, doi: 10.1007/s10506-020-09273-1.
[20] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems 30*, Long Beach, California, 2017, pp. 4765–4774.
[21] R. Tinarrage, H. Ennes, L. Resck, L. T. Gomes, J. R. Ponciano, and J. Poco, "Empirical analysis of binding precedent efficiency in Brazilian Supreme Court via case classification," *Artificial Intelligence and Law*, May 2025, doi: 10.1007/s10506-025-09458-6.
[22] J. Lee, "Legal text classification in Korean sexual offense cases: from traditional machine learning to large language models with XAI insights," *Artificial Intelligence and Law*, May 2025, doi: 10.1007/s10506-025-09454-w.
[23] F. Contini, E. A. Ontanu, and M. Velicogna, "AI accountability in judicial proceedings: an actor–network approach," *Laws*, vol. 13, no. 6, Nov. 2024, doi: 10.3390/laws13060071.
[24] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
[25] H. Surden, "Values embedded in legal artificial intelligence," *IEEE Technology and Society Magazine*, vol. 41, no. 1, pp. 66–74, Mar. 2022, doi: 10.1109/MTS.2022.3147542.
[26] T. Ansari, H. S. Dhillon, and M. Singh, "Machine learning model to predict results of law cases," in *2024 Fourth International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*, IEEE, Jan. 2024, pp. 1–8, doi: 10.1109/ICAECT60202.2024.10468789.
[27] K. Karsai, "The use of algorithms to support judicial decision-making in criminal matters with a special focus on trial decisions," *Studia Iuridica Lublinensia*, vol. 33, no. 5, pp. 103–124, Dec. 2024, doi: 10.17951/sil.2024.33.5.103-124.
[28] R. F. Mayana, T. B. Yudhya, T. Santika, A. M. Ramli, and S. Suseno, "Economic & moral right for artificial intelligence generated works: Perspective from Indonesia copyright law," in *2024 5th International Artificial Intelligence and Blockchain Conference*, Tokyo, Japan, Sep. 2024, pp. 39–43, doi: 10.1145/3702359.3702365.
[29] C. M. Greco and A. Tagarelli, "Bringing order into the realm of transformer-based language models for artificial intelligence and law," *Artificial Intelligence and Law*, vol. 32, no. 4, pp. 863–1010, Dec. 2024, doi: 10.1007/s10506-023-09374-7.
[30] I. Benedetto *et al.*, "Boosting court judgment prediction and explanation using legal entities," *Artificial Intelligence and Law*, vol. 33, no. 3, pp. 605–640, Sep. 2025, doi: 10.1007/s10506-024-09397-8.
[31] E. Francesconi, "The winter, the summer and the summer dream of artificial intelligence in law," *Artificial Intelligence and Law*, vol. 30, no. 2, pp. 147–161, Jun. 2022, doi: 10.1007/s10506-022-09309-8.

BIOGRAPHIES OF AUTHORS



Mamok Andri Senubekti holds a Master of Computer Science (M.Kom.) in Information Systems from STMIK LIKMI (2012) and a Bachelor of Computer Science (S.Kom.) in Informatics Engineering from Universitas Komputer Indonesia (2005). He is currently the head of the subdivision for planning, IT, and reporting at the Bale Bandung District Court and serves as head of the D3 Informatics Management Study Program at Digitech University. He is pursuing a Ph.D. in Information Systems at Diponegoro University. His research interests include application development, data science, and UI/UX, which are reflected in his numerous scientific publications. He is also an active lecturer and journal reviewer. He can be contacted at email: mrymodion@students.undip.ac.id.



Pujiyono is a professor at the Faculty of Law, Diponegoro University, specializing in Criminal Law. He has dedicated his entire academic career to Diponegoro University, where he completed his Bachelor of Laws (S.H.) in Criminal Law in 1988, his Master of Humanities (M.Hum.) in Criminal Law in 1997, and his Doctor of Law (Dr.) in 2011. As a distinguished academic, he holds the functional position of professor. His research and contributions are indexed across various academic platforms. He is an active contributor to the field of law with a significant scholarly presence. He can be contacted at email: pujiyono@lecturer.undip.ac.id.



Foni Agus Setiawan holds bachelor's and master's degrees in Computer Science and earned his doctoral degree in the same field from Universitas Indonesia, Jakarta, in 2019. Since 2009, he has been a researcher at the Research Center for Data and Information Sciences, National Research and Innovation Agency (BRIN), Indonesia. In 2020, he obtained a data scientist certification from IIT Kanpur, India, and certifications as a Knowledge Engineering Specialist and Semantic Web Associate from PoolParty Academy, Vienna, Austria. His research interests include artificial intelligence, data science, knowledge engineering, uncertainty reasoning, and software engineering. He is a national journal (SINTA) assessor and Editor-in-Chief of Informatics Research Journal (Jurnal Riset Informatika). He also reviews for IEEE Access, Information and Software Technology (Elsevier), Journal of Engineering, Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM), Jurnal Ilmu Pengetahuan dan Teknologi Komputer (JITK), and Krea-TIF. He can be contacted at email: foni001@brin.go.id.