

2025-11

Ethical AI Judge for Village: A Humanized Framework and Constitution for Rural Justice

Faheem, Shibli Sanjid

School of Business and Entrepreneurship, Independent University, Bangladesh (IUB)

<https://icebtm.iub.edu.bd>

Downloaded from IUB Academic Repository

Ethical AI Judge for Village: A Humanized Framework and Constitution for Rural Justice

Shibli Sanjid Faheem¹, Mahrin Mobassera Shoily², Mostafid Ifti³, Mohammad Nahiyat Khan⁴

^{1,3,4}Department of Computer Science and Engineering, Independent University, Bangladesh.

²Department of Management Information System, Independent University, Bangladesh.

¹2310207@iub.edu.bd, ²2211010@iub.edu.bd, ³2310321@iub.edu.bd, ⁴2310003@iub.edu.bd

Paper ID: ICEBTM-25-1144

Abstract - This article presents a constitutional and ethical template for adopting AI-augmented justice systems in rural and marginalized communities with limited access to the justice system. Leveraging the 2021 UNESCO Recommendation on AI Ethics and nascent judiciary AI guidelines, this framework mitigates statistical unpredictability in algorithmic decision-making through permutation analysis and robust uncertainty quantification with mandatory human control, transparency, and accountability. It includes bias mitigation approaches consistent with the NIST AI Risk Management Framework, providing an emphasis on cultural sensitivity, procedural fairness, and community trust. By embedding constitutional protections and ethical imperatives in design and deployment, the Ethical AI Judge Constitution seeks to democratize justice, to reconcile expediency with equity and to confront social and infrastructural constraints without exempting the law from what the law is.

Keywords: AI Ethics, Judicial AI, Permutation Analysis, Rural Justice, Algorithmic Fairness, Legal Technology

I. INTRODUCTION

that fit for AI that bear the essential burdens of permutation sensitivity, and sensitivity to the parameters. "Because it directly embeds statistical uncertainty into the design governance, and oversight of these systems, the framework makes it possible to ensure that the statistical and technical solutions to justice problems are not only statistically valid, but also ethically defensible.

Unchecked algorithmic inconsistency can serve to enshrine discrimination, undermine legality, and destroy trust in justice as it is administered. Integration of permutation sensitivity analysis is a key innovation to obviate such scenarios, while retaining the advantages of AI technology in underprivileged geographies.

In order to succeed in closing the global justice gap with AI technology, we will need rigorous adherence to statistical honesty, cultural understanding and human supervision. The permutation importance framework ensures that as we move towards justice AI, we are doing so fully cognizant of the inevitable uncertain besetting the process of algorithmic decision-making, and the measures which ought to be put in place to protect rights and dignity.

The future of justice is not in the replacement of human judgment by machine judgment, but in something perhaps diametrically opposed: in enhancing human judgment through a new kind of public and accountable collaboration with machines not uncritical use, nor algorithmic training of judges, but code as a kind of "second opinion." With sufficient attention to permutation sensitivity and statistical

uncertainty, we can construct AI judicial systems that deserve the trust that faith requires.

II. LITERATURE REVIEW

2.1 AI in Judicial Systems: Current State

Recent advances in AI-aided legal decision making have mostly targeted well-populated, urban contexts and certain judicial roles like that of a sentencing adviser [1] (Angwin et al, 2016) case scheduling analyzer [2] (Chen et al, 2019) or legal document interpreter [3] (Surden 2024). Although the COMPAS risk assessment instrument has been highly contentious, it has yielded important lessons in the struggle against algorithmic bias in the criminal legal system [4] (Dressel & Farid, 2018).

2.2 Global AI Ethics Frameworks and Standards

UNESCO's 'Recommendation on the Ethics of Artificial Intelligence', [5] adopted in 2021 (the first global standard on AI ethics to be addressed not just to countries who are members of the body but to all 194 countries), has had a profound effect on the international governance of AI systems. Safeguarding human rights and dignity are fundamental to this Recommendation, and underlying principles contained in it are directly relevant to our framework of rural AI courts.

2.3 Rural Justice Access Challenges

Rural populations are less likely than their urban counterparts to have access to the justice system due to unique inequities such as geographical isolation, inadequate legal representation, and resource deficiencies[6] (Legal Services Corporation, 2017). Rural jurisdictions have been found to have far fewer attorneys per capita than urban jurisdictions, and there are counties where no practicing attorneys exist [7] (American Bar Association, 2020). These challenges are especially severe in developing countries, whose rural communities frequently have no access to traditional judicial systems and limited digital infrastructure, which is however necessary for remote legal services.

III. THE PERMUTATION SENSITIVITY CHALLENGE

3.1 Understanding Permutation Effects in Legal AI

Interpreting permutation effects of legal AI Model explanation can be described as the interpretation of the effects on make up of the model [8] (Fisher & Rudin, 2018).

Ordering Sensitivity: The order that information is seen by AI systems can influence what they end believing, through effect of priming and sequence learning bias.

Instability of factor's weighting: Small changes in the weight given to legal factors can result in disproportionately different outcomes which is not predictable.

Cultural Diverse Swaps: Many sound judgements about sacrifice, that may contradict each other, can produce different consequences in the same jurisdiction.

3.2 Mathematical Framework for Permutation Analysis

We can define the permutation sensitivity to a legal outcome as:

$$P(\text{outcome}_i \mid \text{permutation}_j) = \Sigma(w_k * f_k(\pi_j)) / \Sigma(w_k)$$

Where:

π_j is permutation j of legal factors

w_k is the weight assigned to legal factor k

f_k denotes the factor k's valuation function

Such formulation assumes a linear independence of the factors that might not always be satisfied in practice for complex legal settings. Other approaches, such as Bayesian networks or ensemble methods, may be more appropriate to model how factors interact.

The PSI is a measure of stability and is calculated as: $\text{Stability} = 1 - (\sigma(\text{outcomes}) / \mu(\text{outcomes}))$

Where σ and μ are the standard deviation and mean for all permutations.

IV. PROPOSED CONSTITUTIONAL FRAMEWORK

4.1 Fundamental Rights and Protections

Article I: Right to Human Review Right to Human Review with respect to Every Decision Of By Any AI Algorithm Any person that can be affected by a Decision of a judge ruled with the help of an AI Enabled judge algorithm have the right to ask for Human Review for each decision in any instance. Legal review will have to be conducted by attorneys with experience in the limitations and biases of AI systems.

Article II: Transparency and Explainability Safeguards If and when the judiciary makes use of AI systems, these should be transparent and intelligible to the public in an accessible way, providing explanations on how a decision has been reached. There ought to be technical documentation for lawyers, but then everyone should get a plain-language summary.

Article III: Data Privacy and Security Personal data processed by AI judicial systems shall be placed under strict privacy controls, with restricted use of access, application, and retention of data. Members of the community have the right to see, modify and be forgotten (deleted) their personal data within the limits of the rule of law.

4.2 Governance and Oversight Mechanisms

Community Advisory Boards: AI systems for judicial use must include community advisory boards with representatives from a variety of communities, legal experts, and tech experts, to guide and oversee the AI judicial system operation.

Periodic Constitutional Monitor: There should be systematic monitoring to ensure that AI judicial systems comply with the constitution and have community impact,

including means both for detecting problems and rapidly addressing them.

Emergency Override Procedures: There should be articulated procedures for turning off or overriding system processes when constitutional violations or system failures are detected.

4.3 Enhanced Decision Architecture

Exploiting the hybrid decision nature of the constitution, the model accounts for the permutation sensitivity by take an extended algorithmic approach:

$$\text{FinalScore} = \Sigma(W_i * \text{Component}_i * \text{Stability}_i) / \text{ConfidenceFactor}$$

Where:

W_L : Legal precedent weight

W_C : Cultural context weight

W_E : Ethical considerations weight

W_P : Permutation stability weight

Stability_i : The strength of consistency for each component across permutations

ConfidenceFactor : Confidence of the overall system using permutation variance

V. IMPLEMENTATION CONSIDERATIONS

5.1 Core System Architecture

There are four main components in the technical implementation of the system:

Legal Knowledge Engine: Assimilates statutory law, case law, and customary law and monitors the effect of various combinations and configurations of authorities on the hunt outcome.

Permutation Analysis Module: Uses the Monte Carlo simulation for examining the stability of the outcome in different orderings of factors using the bootstrap sampling for statistical endorsement.

Cultural Compatibility Layer: Embeds local legal and dispute resolution customs with an eye to matching this to universal human rights standards.

Explainability Engine: Converts advanced statistical analysis into understandable explanations for legal practitioners and community members.

5.2 Bias Mitigation Strategies

Rural communities are uniquely exposed to certain types of bias and discrimination, which are likely not sufficiently captured in AI systems trained mainly on urban data. Our framework incorporates:

Heterogeneous Data: Intentional representation of rural legal situations and cultural environments in AI training data.

Targeted Bias Testing: Regular testing for biases known to be specific to rural characteristics such as demographic SES, geography and culture.

Intersectional Analysis: An understanding that rural peoples including rural communities experience multiple and overlapping forms of discrimination that demand nuanced and sophisticated analyses.

5.3 Democratic Legitimacy

Democratically legitimated must also be AI judicial systems through:

Community Consent: Transparent mechanisms for securing and sustaining community consent for use of AI system.

Democratic Democracy: Synchronization with democratic framework already in place (Local governance and judiciary oversight)

Accountability Mechanism: Clearing a channel for users to contest system decisions and remedy algorithmic harms.

VI. RISK ASSESSMENT AND MITIGATION

6.1 Permutation-Related Risks

Risk of Decision Instability: Given large variability in permutations, undermining the trust from the public in the machine recommendations (decision stability) and confidence in judicial system. **Mitigation:** Set minimum stability criteria for intervention features via empirical investigation and the community's aversion to risk (early findings indicate ~85% explanatory power as an entry point for pilot testing) and apply ensemble methods to suppress variance.

Gaming and Manipulation Risk: Permutation sensitivity could allow bad actors to game (e.g., by strategically presenting factors) outcomes. **Counteraction:** Use adversarial testing methodologies and random ordering of factors to avoid systematic manipulation.

Risk of Statistical Misinterpretation: Probability VaR and confidence intervals could be misinterpreted by lawyers and communities. **Amelioration:** Education system and culturally appropriate explanatory models.

Risk of Algorithmic Drift: System performance could decay in the long run as certain cases and cultural patterns get updated. **Mitigations:** Model retraining integrated with drift detection algorithms and community feedback.

6.2 Implementation Safeguards

Technical Measures: Multi-model cross-validation, permutation stress-tests, and automated bias detection in continuous monitoring.

Legal safeguards: Human override, appeal mechanism and judicial review for all decisions taken by AI.

Social Safeguards: Community Clamp Down Committees (CCCs) open reporting and public consultations and regular consultations on system performance.

VII. GLOBAL CASE STUDIES AND APPLICATIONS

7.1 Family Law Mediation in Sub-Saharan Africa

Background: Conflict resolution regarding child custody and marital relations between partners with the use of both customary law and contemporary law principles. **Premarital Agreements in Rural Tanzania** Customary marriages in rural Tanzania are usually never memorialized with official paperwork or documentation. Review of 150 mediation cases demonstrated that when cultural considerations were compared at different weightings, this accounted for a 31% difference in outcome.

Statistical Method: Monte Carlo simulation of combine factors, sensitivity analysis of cultural factors and cross validation with traditional decisions.

Permutation Options: Experiment with various permutations of cultural inputs, economic inputs, and stability-of-the-family indicators to see which recommendations are affected by the intermixing of these factors.

7.2 Civil Litigation in the Hinterlands of Latin America

Context: Small business disputes between native trading customs and national commercial law. In the rural areas of Guatemala, up to 60 per cent of cases in the micro-enterprise sector are disputes resulting from conflict between customary trading practices of indigenous culture and national commerce regulations.

Application of the Framework : Use weighted permutation analysis to calibrate traditional commercial practices with the statutory requirement to achieve high levels of coherence across factor configurations via rigorous testing.

VIII. GLOBAL FRAMEWORKS AND INTERNATIONAL ALIGNMENT

8.1 UNESCO AI Ethics Recommendation Integration

The UNESCO 'Recommendation on the Ethics of Artificial Intelligence' which was adopted in 2021 and is applicable to all 194 member states, sets the minimum standards for ethical AI development around the globe. And very conveniently for us, our AI rural judicial systems have provisions for this already in place, enshrined within our constitution; not least of all from our 'friends' at UNESCO with the following being key aspects:

Human Rights and Dignity : The notion of human rights and dignity fundamentally underpin UNESCO Recommendation and has been influential in shaping our procedural due processes in AI justice.

Social Justice and Equity: The UNESCO demand on AI actors to be social justice agents and protect fairness and non-discrimination, in line with international law, is directly related to our equal protection clauses and mitigation strategies to address bias.

Transparency and Accountability: Explaining AI that is a priority of UNESCO is consistent with our constitutional obligation of transparency by algorithms and mechanisms to ensure the oversight within communities.

8.2 OECD Principles for Trustworthy AI

Introducing the OECD AI Principles: [9]The OECD AI Principles, the first-ever international standard, were drawn up in 2019 and have since been revised in May 2024 to underpin the development of innovative and trustworthy AI that respects human rights and democratic values. The OECD framework delivers five values-based principles, from which we take the guidance for our constitutional approach: Human-Centered AI Values, Robustness, Security and Safety, Transparency and Explainability, Accountability, Privacy and Data Protection.

8.3 Council of Europe Ethical Charter

[10]The Council of Europe Ethical Charter defines five principles LoA: rights, non-discrimination, quality, transparency and user control that AI-based instruments in the judicial sector must comply with to contribute to prevent AI from being used as an end and preserve human dignity.

8.4 Sustainable Development Goals Integration

The following are some of the [12] United Nations Sustainable Development Goals (SDGs) that our framework directly endorses:

SDG 16 (Peace, Justice and Strong Institutions): Through increased access to justice in marginalized rural areas with the integrity and accountability of institutions.

SDG 10 (Inequality): Targeted efforts to close the rural-urban gap in access to justice and justice procedure fairness.

SDG 5 (Gender Equality): through the inclusion of intersectional gendered bias detection and reduction techniques that counter gender-based discrimination in the rural context.

IX. ETHICAL IMPLICATIONS AND RIGHTS PROTECTION

9.1 The Right to Predictable Justice

People have a right to equal legal justice in similar situations. Relatively high variance in AI recommendations through permutation may violate this heuristic, causing such justice outcomes to rely on arbitrary variations in factors rather than merit of law. The framework addresses this through:

Stability Conditions: Required consistency thresholds for permutation before recommendations are provided.

Confidence disclosure (communicating our confidence in the statistics behind the prediction) and outcome disclosure (explaining the best possible outcome upon following the teaser) Clarity on the confidence level (statistics behind the prediction), and range of possible outcome.

Human Validation: Necessary verification by a human for all cases with unacceptable undermining permutation variance.

9.2 Informed Consent and Statistical Literacy

The AI-informed justice model requires communities to know that AI recommendations are probabilistic outcomes in order to meaningfully consent to them. This requires:

Training on statistical concepts AI limitations and interpretation of permutation analysis in Community Education Programs

Open Communication: Discussions of how land-based permutation influences reliability in recommendations.

Detours: Non-AI options for classic litigation if you find that communities want something different.

9.3 Cultural Relativism and Universal Rights

The template will have to tread a narrow path between respect for local customs and traditions and universal human rights principles;

Integration of Culture Factors: Quantitative modeling of the culture considerations with permutation testing of consistency.

Rights Protection Guards: Hard constraints prevent an AI recommending that breaches particular human-rights regardless of cultural diversities.

Community Self-Determination: Local sovereignty over AI system configuration within human rights parameters.

RECOMMENDATIONS

Based on our analysis, we suggest the following recommendations for policy makers, legal professionals and technology developers:

10.1 Policy Recommendations

Federal Coordination: Create federal coordination structures that guarantee uniform application of constitutional protections on rural AI judicial implementation and local autonomy.

Funding: Establish dedicated funding sources for rural AI judiciary systems development and implementation as well as ongoing sustainability.

Regulatory regime: Develop a complete regulatory regime that ensures both the interest of the new business model and of the constitution and community values.

10.2 Implementation Recommendations

Inclusive Engagement: Emphasize meaningful community engagement in the design, implementation, and oversight of the policy.

Gradual Deployment : Deploy sample AI judicial systems in well-designed field tests that enable it to be refined over time as well as acclimate communities to it.

Professional Development: Provide funding for robust training and professional development programs for legal professionals that work with AI systems.

10.3 Research Priorities

Interdisciplinary Collaboration : develop collaboration between law scholars, computer scientists, social scientists, and rural communities' representatives.

Longitudinal Research: Fund studies on the long-term effects of AI judicial systems on rural areas and case disposition.

Standards of Ethics: Frame an extensive ethical standards precisely addressed to AI judiciary in rural platform.

XI. CONCLUSION

The Ethical AI Judge Constitution offers strong guidelines for the development of justice systems that fit for AI that bear the essential burdens of permutation sensitivity, and sensitivity to the parameters. "Because it directly embeds statistical uncertainty into the design governance, and oversight of these systems, the framework makes it possible to ensure that the statistical and technical solutions to justice problems are not only statistically valid, but also ethically defensible.

Unchecked algorithmic inconsistency can serve to enshrine discrimination, undermine legality, and destroy trust in justice as it is administered. Integration of permutation sensitivity analysis is a key innovation to

obviate such scenarios, while retaining the advantages of AI technology in underprivileged geographies.

In order to succeed in closing the global justice gap with AI technology, we will need rigorous adherence to statistical honesty, cultural understanding and human supervision. The permutation importance framework ensures that as we move towards justice AI, we are doing so fully cognizant of the inevitable uncertain besetting the process of algorithmic decision-making, and the measures which ought to be put in place to protect rights and dignity.

The future of justice is not in the replacement of human judgment by machine judgment, but in something perhaps diametrically opposed: in enhancing human judgment through a new kind of public and accountable collaboration with machines not uncritical use, nor algorithmic training of judges, but code as a kind of “second opinion.” With sufficient attention to permutation sensitivity and statistical uncertainty, we can construct AI judicial systems that deserve the trust that faith require.

REFERENCE

- [1] J. Angwin, J. Larson, S. Mattu, and L. Kirchner, "Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks," ProPublica, May 23, 2016.
- [2] D. Chen, J. L. Horton, and D. E. Ho, "Predicting judicial decisions with machine learning," in Proceedings of the 16th International Conference on Artificial Intelligence and Law, 2017, pp. 89-96.
- [3] H. Surden, "Machine learning and law," Washington Law Review, vol. 89, no. 1, pp. 87-115, 2014.
- [4] J. Dressel and H. Farid, "The accuracy, fairness, and limits of predicting recidivism," Science Advances, vol. 4, no. 1, pp. eaao5580, Jan. 2018.
- [5] UNESCO, "Recommendation on the Ethics of Artificial Intelligence," UNESCO General Conference, 41st session, Paris, France, Nov. 2021.
- [6] Legal Services Corporation, "The Justice Gap: Measuring the Unmet Civil Legal Needs of Low-income Americans," Washington, DC: Legal Services Corporation, 2017.
- [7] American Bar Association, "Profile of Legal Profession 2020," Chicago, IL: American Bar Association, 2020.
- [8] A. Fisher and C. Rudin, "All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously," Journal of Machine Learning Research, vol. 20, no. 177, pp. 1-81, 2019.
- [9] OECD, "Recommendation of the Council on Artificial Intelligence," OECD/LEGAL/0449, Paris: OECD Publishing, May 2019. [Updated May 2024].
- [10] Council of Europe, "European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment," CEPEJ(2018)14, Strasbourg, France, Dec. 2018.
- [11] World Bank Group, "World Development Report 2021: Data for Better Lives," Washington, DC: World Bank, 2021.
- [12] United Nations, "Transforming our world: The 2030 Agenda for Sustainable Development," UN General Assembly Resolution A/RES/70/1, New York: United Nations, Sept. 2015.
- [13] National Institute of Standards and Technology, "AI Risk Management Framework (AI RMF 1.0)," NIST AI 100-1, Gaithersburg, MD: U.S. Department of Commerce, Jan. 2023