A Systematic Review on Explainable AI in Legal Domain

Mr. Gautam Joshi[1], Sanjana Mali[2], Poonam G. Fegade[3]

Master Of Computer Application, K.K. Wagh Institute of Engineering Education and Research, Nashik, Maharashtra, India.

Abstract: In the past few years, Artificial Intelligence (AI) has become a game-changer in many fields, including law. AI can speed up decision-making and make it more accurate. However, AI often makes decisions in ways that are hard to understand. This lack of clarity is a big problem in law, where decisions can change lives. This review brings together current research on Explainable AI (XAI) in law. It looks at the different AI methods used and how XAI affects legal decision-making. The review also deals with the ethical and practical issues of using XAI, like biases and privacy concerns. By tackling these challenges, the legal world can use XAI effectively while sticking to fairness and responsibility. Our goal is to assist researchers, lawyers, and policymakers in making AI in law more transparent, fair, and trustworthy.
Keywords: Artificial Intelligence, Explainable AI, Legal Profession, Law, Interpretable ML.

I. INTRODUCTION

Artificial intelligence has swiftly ingrained itself into numerous sectors, including the legal field, presenting a myriad of possibilities for advancement. With the spotlight on machine learning, deep learning, and natural language processing, there's a growing recognition of their potential to revolutionize legal research and its multifaceted applications. However, as AI technologies evolve, the challenge lies in unraveling the intricate algorithms that underpin decision-making processes. Often, these algorithms operate as opaque black boxes, rendering it challenging for users, including developers, to discern the rationale behind their outputs. This opacity underscores the necessity for Explainable AI (XAI) methodologies. XAI serves as a conduit for engendering trust in the outcomes produced by machine learning algorithms. By elucidating model outputs and identifying biases, XAI fosters fairness and transparency, particularly in AI-driven decision-making frameworks within the legal domain. It empowers developers to ensure the fidelity of their systems and provides users with a comprehensible rationale for algorithmic decisions. In essence, while machine learning models excel at predicting outcomes based on historical data, XAI bridges the gap by offering insights into the decision-making process.
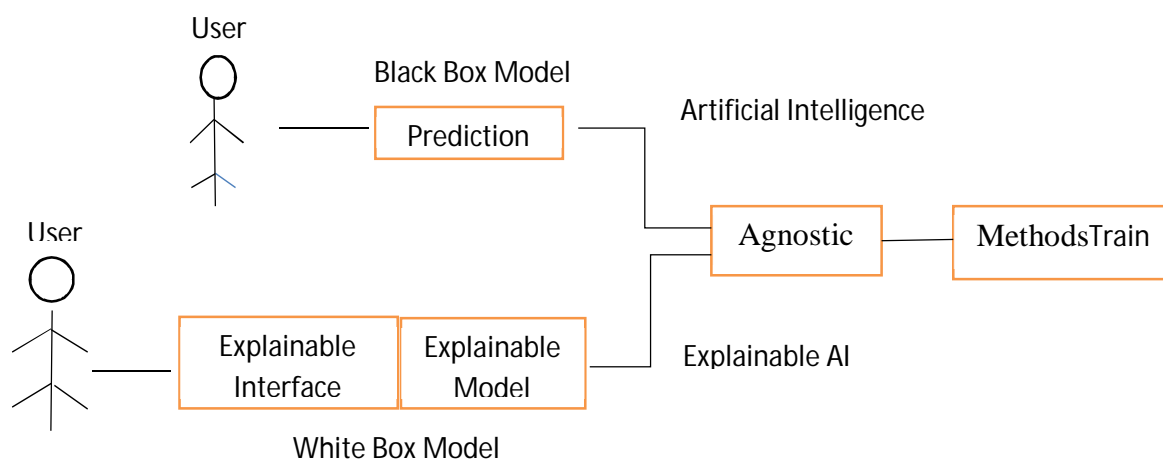

Fig: Artificial Intelligence Vs Explainable AI

The integration of XAI not only instills confidence in users' decisions but also propels firms towards accelerated growth and profitability. When users possess a clear understanding of the factors influencing algorithmic predictions, they can make more informed choices, thereby driving efficiency and efficacy in legal proceedings.

Thus, the advent of XAI heralds a transformative era in the legal profession, characterized by enhanced transparency, accountability, and informed decision-making. The integration of AI in the legal domain has seen rapid growth, with technologies like machine learning, deep learning, and natural language processing gaining prominence. These technologies offer vast potential for revolutionizing legal research and its applications. However, a significant challenge arises from the opacity of the algorithms that drive AI decision-making processes. This lack of transparency makes it difficult for users, including developers, to understand the reasoning behind the outputs produced by these algorithms. Explainable AI (XAI) methodologies address this challenge by elucidating model outputs and identifying biases, thereby fostering trust, fairness, and transparency in AI-driven decision-making frameworks within the legal domain. XAI empowers developers to ensure the fidelity of their systems and provides users with a comprehensible rationale for algorithmic decisions

II. LITERATURE SURVEY

Traditional AI models can be highly effective but often function as "black boxes," making it difficult to understand how they arrive at their decisions. XAI addresses this by providing methods to explain the reasoning behind an AI's output. This can involve techniques like highlighting factors most influential in a decision or offering alternative scenarios with different inputs.In the legal domain, transparency and fairness are paramount. Opaque AI algorithms raise concerns about bias and the rationale used for legal determinations. XAI helps ensure fairness, transparency and accountability.

AI applications in law demonstrate how technology can augment and enhance various aspects of legal practice, from case analysis and prediction to document review and contract management.

Case Prediction: AI is used to analyze past legal cases and predict the outcomes of similar cases based on various factors such as judge rulings, case law, and legal precedents. This helps lawyers and legal professionals assess the likelihood of success for their cases and make informed decisions.

Document Analysis: AI tools are employed to review and analyze large volumes of legal documents, including contracts, briefs, and court rulings. Natural Language Processing (NLP) techniques enable AI systems to extract relevant information, identify key clauses, and flag potential issues or discrepancies, saving time and reducing the risk of human error.

Contract Review: AI is utilized to streamline the contract review process by automatically reviewing and categorizing contracts, identifying important clauses, and highlighting potential risks or inconsistencies. This accelerates the due diligence process in legal transactions such as mergers and acquisitions, enabling lawyers to focus on higher-value tasks.

[1] Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., ... & Herrera, F. (2023). Explainable Artificial Intelligence (XAI):- This research delves into Explainable AI (XAI) and its pivotal role in instilling confidence in AI systems. While strides have been made in revealing AI model operations, there's a call to tailor explanations to users and offer actionable insights. Trustworthy AI transcends mere transparency; XAI stands as a crucial foundation. Future pursuits should explore regulations mandating explanations, crafting tailored explanations for diverse user groups, and fostering collaboration between legal and technical experts. By overcoming these challenges, XAI ensures AI advancement is accompanied by responsible and ethical implementation, shaping a future where AI not only progresses but also operates with transparency and accountability, enhancing trust among users.

[2] Kabir, M. S., & Alam, M. N. (2023). The Role of AI Technology for Legal Research and Decision Making :- AI is transforming law by speeding up research and decision-making for lawyers. It sifts through tons of legal info quickly, finding what's relevant and suggesting strategies. This saves time and helps lawyers see connections they might miss. AI learns from past cases to predict outcomes and assist in decisions. But, there are challenges too. AI's inner workings can be hard to grasp, making transparency an issue. Also, if AI is trained on biased data, it might give biased results. Despite challenges, AI is revolutionizing law practice, making it faster and more efficient.

[6] Chaudhary, G. (2023). Explainable Artificial Intelligence (xAI) :- As machine learning seeps into legal realms like criminal, administrative, and civil matters, worries crop up about the mystery surrounding these algorithms, making it hard for judges to grasp how decisions are made. This paper dives into why Explainable AI (xAI) is crucial in legal contexts for transparency and accountability. It looks at existing research on xAI and talks about ways judges can better understand algorithmic outcomes through feedback. Without transparency, judges might struggle to make fair judgments. But with xAI, decision-making transparency can improve. The paper suggests that the legal system can help develop xAI by customizing it for different areas of law and providing feedback for enhancement. It stresses how xAI is vital for clear, accountable decisions in legal settings and urges cooperation between judges and the public to ensure this technology grows ethically and responsibly.

[3]Richmond, K. M., Muddamsetty, S. M., Gammeltoft-Hansen, T., Olsen, H. P., & Moeslund, T. B. (2024). Explainable AI and law: an evidential survey. Digital Society, 3(1), 1.: - This paper argues that current research on AI in law suffers from a "one-size-fits-all" approach. It highlights the importance of considering the different types of legal reasoning used in various legal sub-domains (administrative, criminal, civil) when designing AI systems for legal decision-making. The core finding is that the suitability of an AI model depends on the dominant form of reasoning and mode of fact-finding employed in a specific legal domain. Legal sub-domains with more transparent reasoning processes (like those relying on deduction) can benefit from complex AI models like deep learning. Conversely, sub-domains with opaque, tacit knowledge (like those using holistic fact-finding) might require simpler models with higher explainability. This research offers a unique taxonomy that links legal reasoning forms to specific AI models, promoting a more nuanced approach to AI in law.

[8] Luo, C. F., Bhambhoria, R., Dahan, S., & Zhu, X. (2022). Evaluating Explanation Correctness in Legal Decision Making. In AI. : - This research examines how trustworthy machine learning models are in legal applications. Explainable AI (XAI) is gaining interest for high-impact decisions, but its effectiveness in law remains unclear. To address this, the study combines computational analysis with user studies involving lawyers to evaluate XAI methods. Faithfulness, a core concept, measures how well explanations reflect the model's true reasoning. Plausibility, assessed through user studies, reflects how reasonable explanations appear to legal experts. The goal is to determine if XAI methods can meet legal decision-making requirements. The study emphasizes XAI's importance in building trust and promoting AI adoption in legal tasks. Researchers assessed the explanation accuracy of two XAI methods (LIME and SHAP) on legal tasks from various legal areas. Similar to prior research, the study found that explanation faithfulness can vary depending on the training data. However, all XAI methods provided significantly better explanations than a random baseline. The study acknowledges that different legal areas can have varying levels of subjectivity, so it proposes new metrics to quantify legal subjectivity and potential bias in explanations. A user study employed a newly developed metric to assess explanation plausibility. The results suggest that SHAP explanations are consistently more plausible to lawyers than LIME explanations, regardless of the legal domain's subjectivity. However, inconsistent faithfulness results across datasets caution against blindly trusting these methods.

[7] Yamada, Y. (2023). Judicial Decision-Making and Explainable AI (XAI)–Insights from the Japanese Judicial System. Studia Iuridica Lublinensia, 32(4 New Challenges to Justice: the Application of Law and ADR in the Perspective of Information Society), 157-173. :- This research delves into the potential of Artificial Intelligence (AI) within court proceedings, focusing on its role as a supportive tool for judges, not a replacement. The rise of AI has ignited discussions about AI judges. However, the author highlights a key difference: traditional legal software functions on pre-programmed logic. In contrast, AI's ability to learn from data raises concerns about transparency and control in its decision-making.

The paper distinguishes AI from current legal software, like case research tools, which operate on a pre-defined foundation. AI, on the other hand, adapts its reasoning based on data it processes, making it less predictable and potentially unsuitable for replacing human judges. The core challenge lies in achieving "explainable AI" (XAI) for legal applications. The author emphasizes the importance of XAI, drawing on insights from the Japanese judicial system and suggesting its wider relevance. The lack of explainability and the potential weakening of the legal system's authority due to opaque AI decision-making are significant hurdles to AI playing a major role in judicial rulings.

III. EXPLAINABLE AI TECHNIQUES

A. Model-Specific Methods

Intrinsic Methods: These methods are inherent to the model architecture and provide explanations without requiring additional post-hoc techniques. Examples include decision trees, rule-based models, and linear models.

Extrinsic Methods: These methods involve modifying the model or its training process to enhance interpretability. Techniques include simplifying complex models, adding constraints, or incorporating domain knowledge during training.

B. Post-hoc Methods

Local Explanation Methods: Provide explanations for individual predictions or instances. Examples include LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations).

Global Explanation Methods: Offer insights into the overall behavior and decision-making process of the model. Techniques include feature importance analysis, model-agnostic surrogate models, and sensitivity analysis.

C. Rule-Based Methods

Rule Extraction: Automatically extract human-readable rules from black-box models to explain their decision logic. Techniques include decision tree induction, rule learning algorithms, and symbolic rule extraction.

Rule Post-processing: Refine or simplify complex rule sets to improve interpretability without sacrificing accuracy. Methods include pruning, rule merging, and rule optimization algorithms.

D. Example-Based Methods

Prototype Selection: Identify representative instances or prototypes from the dataset to explain model behavior. Techniques include k-means clustering, nearest neighbor search, and exemplar-based sampling.

Counterfactual Explanation: Generate alternative input instances that result in different model predictions to highlight the importance of specific features or decision factors.

E. Feature-Based Methods

Feature Importance Analysis: Determine the relative importance of input features in influencing model predictions. Techniques include permutation importance, SHAP values, and feature contribution scores.

Partial Dependence Plots (PDP): Visualize the relationship between individual features and model predictions while marginalizing over other features.

F. Attention-Based Methods

Attention Mechanisms: Highlight relevant parts of the input data that the model focuses on when making predictions. Common in sequence prediction tasks, such as natural language processing and computer vision.
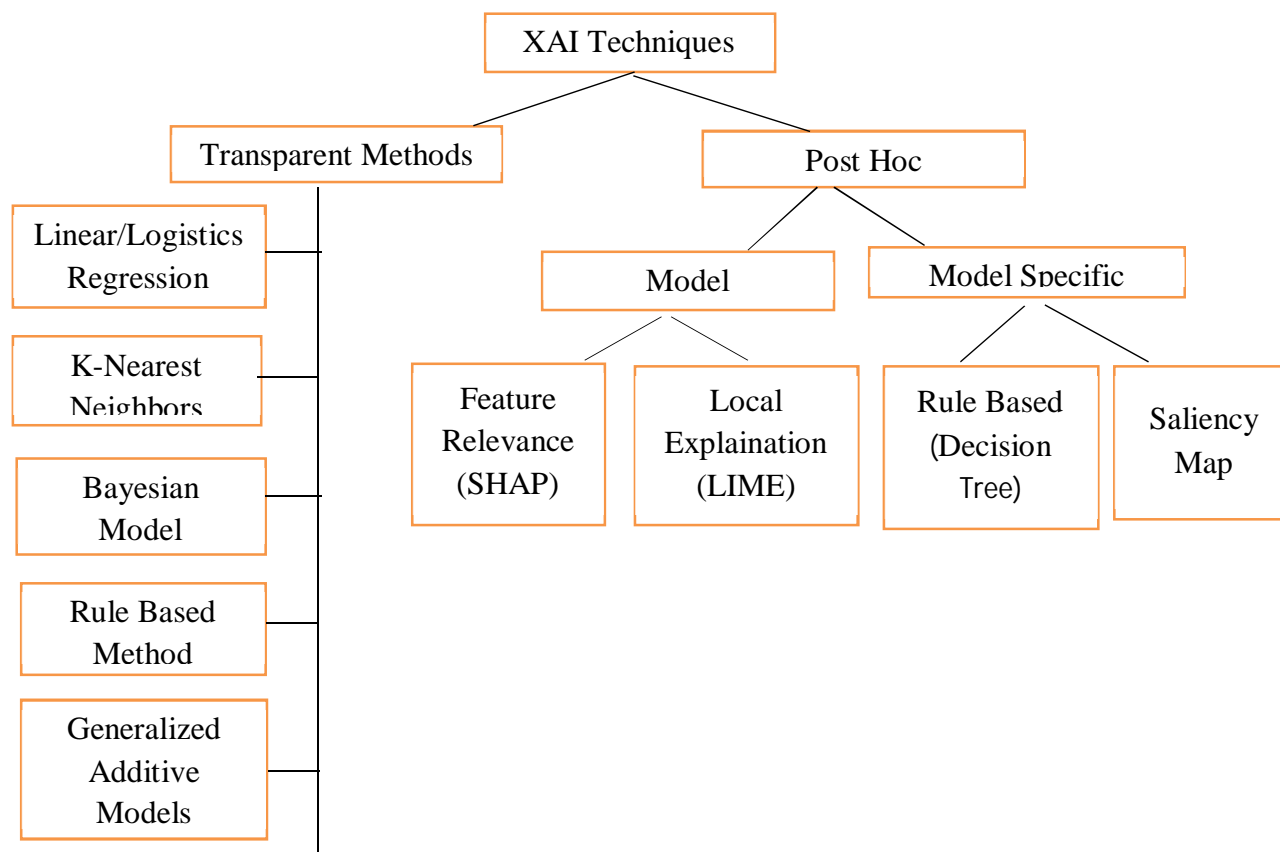

Fig. XAI Techniques

Table 1: XAI methods encountered in literature

Article	Algorithm/Methodology	XAI Methods	Purpose
Górski, L., et al. (2021)	GradCam , SHAP, LIME	Post hoc	Explainability concept for legal texts
Nielsen, A., et al.(2023)	ANGELIC (ADF for kNowledGe Encapsulation of Legal Information for Cases).	Post hoc	Build an AI tool that can explain its legal reasoning in cases involving the European Court of Human Rights.
Steging, C., et al. (2021)	knowledge-driven method for model-agnostic rationale evaluation	Post hoc	This paper is to introduce a novel methodology for evaluating the rationale behind decisions made by black box machine learning systems in the context of AI and law.
Bhambhoria, R.,et al (2021)	Longformer, XGBoost	Post hoc	Exploring how well cutting-edge models perform in forecasting court decisions based on case descriptions. It also delves into the critical question of explainability, can we understand the reasoning behind these models predictions.

IV. CHALLENGES AND CONSIDERATIONS

1) *Complexity and Interpretability:* In the legal profession, explainable AI (XAI) faces challenges due to the complexity of legal problems, leading to potential misinterpretations. Despite efforts to enhance interpretability, the intricate nature of legal issues may still pose difficulties, underscoring the importance of ongoing refinement and scrutiny to ensure accurate and reliable outcomes.

2) *Contextual Understanding:* In the legal profession, explainable AI (XAI) encounters limitations in contextual understanding, as it may struggle to grasp the nuanced meanings and usage of language within specific legal contexts. This challenge underscores the importance of human oversight and expertise in interpreting legal nuances to complement XAI-driven insights effectively.

3) *Risk of Biased Decision:* In the legal profession, explainable AI (XAI) confronts the risk of biased decisions due to its reliance on historical data for learning. If the training data contains biases, XAI algorithms may perpetuate and amplify these biases, highlighting the critical need for transparent and accountable AI systems to mitigate such risks.

A. Case Studies and Applications

1) *Predictive Analysis:* In the legal profession, explainable AI (XAI) facilitates predictive analysis by aiding legal professionals in forecasting outcomes and making informed decisions. Through transparent explanations of its predictions, XAI empowers practitioners to anticipate potential legal scenarios and strategize effectively, enhancing the efficiency and accuracy of legal processes.

2) *Improving Legal Writing:* In the legal profession, explainable AI (XAI) enhances legal writing by analyzing documents and suggesting corrections in legal texts and contracts. By providing transparent explanations for suggested edits, XAI assists legal professionals in improving the clarity, accuracy, and compliance of legal documents, thereby optimizing the quality of written content.

3) *Contract Review and Analysis:* In the legal profession, explainable AI (XAI) streamlines contract review and analysis by scrutinizing documents and offering insights for revisions or corrections. Through transparent explanations of its assessments, XAI empowers legal professionals to make informed decisions and ensure the accuracy and compliance of contractual agreements, enhancing efficiency and reliability

V. DISCUSSION

In the legal world, AI is stepping up to handle tasks like research and decision-making. But as these systems become more common, people are asking for clarity and accountability. They want to know how AI makes its decisions. That's where Explainable Artificial Intelligence (XAI) comes in.

1) *Understanding Legal Texts:* Górski and his team (2021) looked at how AI can explain legal documents. They used fancy algorithms like GradCam, SHAP, and LIME to create a special way of explaining legal stuff. Their work shows how important it is for AI to be clear about how it reaches decisions in legal papers.

2) *Improving Legal Reasoning:* Nielsen and his crew (2023) came up with a cool idea called ANGELIC. They're using it to help with cases in the European Court of Human Rights. Their goal is to build an AI tool that not only makes decisions but also explains why it made them. This helps lawyers and judges trust AI more.

3) *Checking AI's Reasoning:* Steging and his pals (2021) made a smart method to check if AI is making good decisions in legal stuff. They came up with a way to look inside the "black box" of AI and see how it's working. This helps make sure AI systems are fair and understandable.

4) *Predicting Court Decisions:* Bhambhoria and team (2021) checked out how well AI can predict court decisions. They used fancy models like Longformer and XGBoost to do this. But they didn't stop there—they also wanted to see if AI could explain its predictions. This helps make AI predictions easier to trust and understand.

In the end, we've seen how XAI can make the legal system fairer and clearer. But it's not without its challenges. We need to keep working together to make sure AI helps everyone in the legal world and makes justice more accessible for all.

VI. FUTURE DIRECTIONS AND RECOMMENDATIONS

Adapting Explainable Artificial Intelligence (XAI) for legal reasoning is crucial. In legal matters, it's not just about stating facts but also about explaining the legal reasoning behind decisions. One way to enhance this is by merging human expertise with AI processes. This could mean letting lawyers review AI explanations and tweak models as needed.

Bias in legal AI is another concern. We should work on spotting and fixing bias in the data used to train AI models. This involves gathering data in a fairer way and using algorithms less likely to pick up biased patterns.

To build trust, we need common ways to measure how understandable AI systems are in law. These metrics should be easy for legal pros to grasp, not just tech experts. Plus, we should weave explainability into every step of making and using AI in the legal field. This way, users can trust AI decisions and understand how they're reached.

VII. CONCLUSION

Explainable or Trustworthy AI helps to overcome the blackbox processing problem of traditional AI and helps in making better decision and understanding predictive analysis of the outcome by giving an explanation of outcome.

In conclusion, the studies we've explored show how Explainable AI (XAI) can make a big difference in the legal world. By using smart technology like GradCam and SHAP, researchers are helping to explain complicated legal texts in a way that makes sense. Tools like ANGELIC are being developed to help AI systems explain their legal reasoning, which can build trust with lawyers and judges. These findings highlight the importance of XAI in making legal decisions more understandable and fair. By continuing to improve AI tools and work together, we can make sure that technology supports justice and fairness in the legal system. In the future, XAI could be a powerful tool for making law more transparent and accessible to everyone.

REFERENCES

[1] Ali S., Abuhmed T., El-Sappagh S., Muhammad K., Alonso-Moral J. M., Confalonieri R., ... & Herrera F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. Information fusion, 99, 101805.

[2] Kabir M. S., & Alam M. N. (2023). The Role of AI Technology for Legal Research and Decision Making. International Research Journal of Engineering and Technology (IRJET), Volume: 10 Issue: 07, e-ISSN: 2395-0056.

[3] Richmond K.M., Muddamsetty S.M., Gammeltoft-Hansen T. et al. Explainable AI and Law: An Evidential Survey. , https://doi.org/10.1007/s44206-023-00081-z

[4] Górski Ł., & Ramakrishna S. (2021, June). Explainable artificial intelligence, lawyer's perspective. In Proceedings of the eighteenth international conference on artificial intelligence and law (pp. 60-68). https://doi.org/10.1145/3462757.3466145

[5] Nielsen A., Skylaki S., Norkute M., & Stremitzer A. (2023, June). Effects of XAI on Legal Process. In Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law (pp. 442-446). https://doi.org/10.1145/3594536.3595128

[6] Chaudhary G. (2023). Explainable Artificial Intelligence (xAI): Reflections on Judicial System. Kutafin Law Review, 10(4). DOI: 10.17803/2713-0533.2023.4.26.872-889

[7] Yamada Y. (2023). Judicial Decision-Making and Explainable AI (XAI)–Insights from the Japanese Judicial System. Studia Iuridica Lublinensia, 32(4 New Challenges to Justice: the Application of Law and ADR in the Perspective of Information Society), 157-173.

[8] Luo C. F., Bhambhoria R., Dahan S., & Zhu X. (2022). Evaluating Explanation Correctness in Legal Decision Making. In AI., The 35th Canadian Conference on Artificial Intelligence

[9] Steging C., Renooij S., & Verheij B. (2021, June). Discovering the rationale of decisions: towards a method for aligning learning and reasoning. In Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law (pp. 235-239).

[10] Bhambhoria R., Dahan, S., & Zhu X. (2021). Investigating the State-of-the-Art Performance and Explainability of Legal Judgment Prediction. In Canadian Conference on AI.

[11] Collenette J, Atkinson K, Bench-Capon T (2023). Explainable AI tools for legal reasoning about cases: A study on the European Court of Human Rights, Artificial Intelligence, Volume 317, ISSN 0004-3702,
https://doi.org/10.1016/j.artint.2023.103861, (https://www.sciencedirect.com/science/article/pii/S0004370223000073)

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ⓒ (24*7 Support on Whatsapp)