

Table 1: Structure of XML Metadata Fields

Field	Description	Role
Identification	Name of the .xml file	Excluded
Date-Mise-Jour	Internal processing date	Excluded
Code-Juridiction	Jurisdiction code (e.g., "CE")	Excluded
Nom-Juridiction	Jurisdiction name	Excluded
Numero-Dossier	Case number	Excluded
Type-Décision	Decision or order	Excluded
Type-Recours	Type of claim	Excluded
Solution	Final judgment outcome	Outcome (target)
Texte-Integral	Full textual content; factual section used as model input (see Appendix ??)	Input (facts)

Table 2: Administrative Court Outcome Categories

Name	Translation	Description
Non-lieu	Discontinuance / No ruling required	Case closed without substantive judgment
Rejet	Dismissal	Claim rejected by the court
Satisfaction partielle	Partial satisfaction	Partial relief granted to the claimant
Satisfaction totale	Full satisfaction	Full relief granted to the claimant
Expertise / Médiation	Expertise / Mediation	Resolution through expert assessment or mediation
Sursis à statuer	Stay of proceedings	Temporary suspension of the judicial decision
Supplément d’instruction	Additional investigation	Further evidence or clarification requested by the court
Désistement	Withdrawal	Voluntary withdrawal of the claim by the applicant
Rejet défaut de doute sérieux	Dismissal for lack of serious doubt	Insufficient legal basis for claim
Désistement d’office	Automatic withdrawal	Case closed automatically by the court
Radiation du registre	Removal from the register	Case removed from the court’s docket
Satisfaction partielle (susp. exécution)	Partial satisfaction with suspension	Partial relief with suspension of enforcement
Renvoi au Tribunal des conflits	Referral to Tribunal of Conflicts	Jurisdictional conflict between courts
Renvoi autres juridictions	Referral to other jurisdictions	Case transferred to another court
Demande d’avis article (12) L.113-1	Advisory opinion request	Court requests higher legal interpretation

Name	Translation	Description
Renvoi au TA	Referral to Administrative Court	Case transferred to Tribunal Administratif
Dessaisissement	Relinquishment of jurisdiction	Court declines competence
Question préjudicielle	Preliminary question	Legal question referred before ruling
Renvoi au CE	Referral to Council of State	Case sent to Conseil d'État
Renvoi à la CAA	Referral to Court of Appeal	Case sent to Cour Administrative d'Appel
Transaction	Settlement	Agreement between parties
QPC - Transmission avec sursis	Constitutional question (suspended)	Question transmise avec suspension
QPC - Refus transmission	Refusal to transmit QPC	Constitutional referral rejected

Table 3: Grouped Outcome Categories

Name	Description
Rejet	Includes dismissal-related outcomes where the court rejects the claim or terminates the case without granting relief.
Satisfaction partielle	Partial acceptance of the applicant's claims, where the court grants only part of the requested relief.
Satisfaction totale	Full acceptance of the applicant's claims, where the court grants the requested relief in its entirety.
Renvoi	Procedural outcomes involving referral, withdrawal, suspension, or other procedural decisions without a substantive ruling on the merits.

Table 4: Summary of Models Used in the Literature on Judicial Decision Prediction

Model Type / Model	References
<i>Statistical Models</i>	
Generalized Gini-PLS (Partial Least Squares)	Tagny-Ngompé et al. (2020)
<i>Machine Learning Models</i>	
Logistic Regression	Bertalan and Ruiz (2024) , Shelar and Moharir (2021)
Support Vector Machine (SVM)	Aletras et al. (2016) , Bertalan and Ruiz (2024) , Shelar and Moharir (2021) , Bharati (2024) , Kumar (2024)
Decision Tree	Kumar (2024)
Random Forest	Katz et al. (2017) , Bertalan and Ruiz (2024) , Kumar (2024) , Jacob de Menezes-Neto and Clementino (2022)
Gradient Boosting	Jacob de Menezes-Neto and Clementino (2022)
Naive Bayes	Kumar (2024)
k-Nearest Neighbors (k-NN)	Kumar (2024)
<i>Deep Learning Models</i>	
Convolutional Neural Network (CNN)	Shelar and Moharir (2021)
Long Short-Term Memory (LSTM)	Shelar and Moharir (2021)
Deep Neural Network (DNN)	Bharati (2024) , Jacob de Menezes-Neto and Clementino (2022)
Feedforward Neural Network	Jacob de Menezes-Neto and Clementino (2022)
Transformer-based Neural Networks	Xia and Luo (2024)
<i>Transformer Models</i>	
BERT	Alcántara Francia et al. (2022) , Shu et al. (2024) , Xia and Luo (2024)
RoBERTa	Shu et al. (2024)
Legal-BERT	Shu et al. (2024)
JuriBERT	Douka et al. (2021)
<i>Large Language Models (LLMs)</i>	
LawLLM	Shu et al. (2024)
GPT	Gray et al. (2024) , Posner and Saran (2025)
Llama	Gray et al. (2024)

Table 5: Reported Prediction Performance of Models in Legal Text Prediction Studies

Model	Performance Range	References
Support Vector Machine (SVM)	Accuracy: 70–85%	Aletras et al. (2016) , Bertalan and Ruiz (2024) , Shelar and Moharir (2021) , Bharati (2024) , Kumar (2024)
Random Forest (RF)	Accuracy: 75–90%	Katz et al. (2017) , Bertalan and Ruiz (2024) , Kumar (2024) , Jacob de Menezes-Neto and Clementino (2022)
Gradient Boosting Classifier (GBC)	Accuracy: 78–92%	Jacob de Menezes-Neto and Clementino (2022)
BERT / JuriBERT	F1-score: 80–94%	Alcántara Francia et al. (2022) , Shu et al. (2024) , Xia and Luo (2024)
GPT (Large Language Model)	Accuracy: 75–90%	Gray et al. (2024) , Posner and Saran (2025)

See Section ?? for the formal definition of the metric (Accuracy / F1-score).

Performance ranges are reported from the original studies and may vary depending on dataset characteristics, preprocessing methods, and evaluation metrics.

Table 6: The confusion table

		Actual values	
		Positive	Negative
Prediction	Positive	TP	FP
	Negative	FN	TN

Table 7: Accuracy and macro F1-score of classical ML models under four resampling strategies

Model	SMOTE		BorderSMOTE		SMOTETomek		ADASYN	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
SVM	0.5995	0.5625	0.5810	0.5466	0.5993	0.5624	0.5907	0.5463
RF	0.7071	0.6174	0.7028	0.6149	0.7038	0.6126	0.7011	0.6130
GBC	0.5986	0.5371	0.5842	0.5228	0.5982	0.5363	0.5932	0.5240

Table 8: Predictive Performance of Transformer-Based and Large Language Models on French Administrative Tribunal Decisions

Model	Accuracy	F1-score (macro)
<i>RF + SMOTE (baseline)</i>	<i>0.7071</i>	<i>0.6174</i>
BERT	0.7243	0.6512
JuriBERT	0.7831	0.7184
GPT (<i>few-shot</i>)	0.6547	0.5721

References

- O. A. Alcántara Francia, M. Nunez-del Prado, and H. Alatrística-Salas. Survey of text mining techniques applied to judicial decisions prediction. *Applied Sciences*, 12(20):10200, 2022.
- N. Aletras, D. Tsarapatsanis, D. Preoțiuc-Pietro, and V. Lampos. Predicting judicial decisions of the european court of human rights: A natural language processing perspective. *PeerJ computer science*, 2:e93, 2016.
- V. G. F. Bertalan and E. E. S. Ruiz. Using attention methods to predict judicial outcomes. *Artificial Intelligence and Law*, 32(1):87–115, 2024.
- R. K. Bharati. Predictive justice in indian courts: Machine learning approaches to case outcome forecasting 1. *Machine Learning*, 2024.
- S. Douka, H. Abdine, M. Vazirgiannis, R. El Hamdani, and D. R. Amariles. Juribert: A masked-language model adaptation for french legal text. In *Proceedings of the Natural Legal Language Processing Workshop 2021*, pages 95–101, 2021.
- M. Gray, J. Savelka, W. Oliver, and K. Ashley. Using llms to discover legal factors. *arXiv preprint arXiv:2410.07504*, 2024.
- E. Jacob de Menezes-Neto and M. B. M. Clementino. Using deep learning to predict outcomes of legal appeals better than human experts: A study with data from brazilian federal courts. *PloS one*, 17(7):e0272287, 2022.
- D. M. Katz, M. J. Bommarito II, and J. Blackman. A general approach for predicting the behavior of the supreme court of the united states. *PloS one*, 12(4):e0174698, 2017.
- M. Kumar. Predictive modelling in legal decision-making: Leveraging machine learning for forecasting legal outcomes. *J. Electrical Systems*, 20(3):2060–2071, 2024.
- E. A. Posner and S. Saran. Judge ai: A case-study of large language models as judges. *Journal of Law & Empirical Analysis*, page 2755323X261433614, 2025.
- A. Shelar and M. Moharir. Predicting outcomes of court judgments-a machine learning approach. In *2021 International Conference on Intelligent Technologies (CONIT)*, pages 1–6. IEEE, 2021.
- D. Shu, H. Zhao, X. Liu, D. Demeter, M. Du, and Y. Zhang. Lawllm: Law large language model for the us legal system. In *Proceedings of the 33rd ACM International Conference on information and knowledge management*, pages 4882–4889, 2024.
- G. Tagny-Ngompé, S. Mussard, G. Zambrano, S. Harispe, and J. Montmain. Identification of judicial outcomes in judgments: a generalized gini-pls approach. *Stats*, 3(4):427–443, 2020.
- Y. Xia and X. Luo. Legal judgment prediction with llm and graph contrastive learning networks. In *Proceedings of the 2024 8th International Conference on Computer Science and Artificial Intelligence*, pages 424–432, 2024.