

NLP-DRIVEN DIGITAL ADVISORY BOARD USING AI ANALYZER

KAMESH M (2127210501070)

PRANAV P B (2127210501107)

PRANAV V P (2127210501109)

CS18811 – PROJECT WORK

FIRST REVIEW

Batch No	: 26
Name of the Internal Guide	: Mr. V. Krishnamoorthy
Date of Review	: 05/03/2025
Domain	: Machine Learning

ABSTRACT

This project develops an AI-driven legal advisory system to help deliver faster and unbiased legal suggestions. Using NLP and deep learning models like BERT and GPT, it analyzes legal case files to suggest outcomes. Tools like SHAP and LIME ensure transparency by explaining the factors behind predictions. A user-friendly interface provides quick, data-driven suggestions, reducing case backlogs and detention time.

PROBLEM STATEMENT

How might we build an AI-powered tool to efficiently analyze petitions, categorize them into relevant departments, flag urgent and important cases, and send reminders to officials, while also identifying repetitive grievances and tracking progress until resolution? The tool should also include a feature to communicate the status of the grievance to the petitioner, ensuring transparency and accountability.

INTRODUCTION

The legal sector increasingly integrates AI for efficient and accurate suggestions. Digital Advisory Board uses AI and NLP to analyze past cases and predict outcomes, addressing the challenges of manual legal research, including complexity and bias. NLP techniques like tokenization and named entity recognition extract critical features, which AI models such as BERT and GPT process. The proposed AI Analyzer employs machine learning and deep learning methods, along with explainable AI tools like SHAP, to ensure transparency. This system reduces workloads, minimizes biases, and supports legal professionals across civil, criminal, and corporate domains by providing data-driven insights.

ISSUES & CHALLENGES

Data Quality and Availability

- Inconsistent or incomplete legal documents and petitions may affect the accuracy of AI analysis.
- Legal data may be scattered across different systems or formats, making it difficult to access or standardize.

Complexity of Legal Language

- Understanding the nuanced language, terminologies, and jurisdiction-specific rules in legal texts can challenge NLP models.
- Variations in legal phrasing and case-specific complexities require highly refined AI models.

ISSUES & CHALLENGES

Bias and Fairness in AI Suggestions

- AI models trained on biased data can reinforce existing disparities in legal suggestions or grievance categorization.
- Ensuring the fairness and ethical operation of AI models remains critical.

Interpretability and Trust

- Legal professionals require explainable AI (XAI) to understand and trust predictions.
- Black-box models without transparency hinder adoption in legal contexts.

ISSUES & CHALLENGES

Data Privacy and Security

- Handling sensitive legal information and petitioner data poses privacy concerns.
- Systems must comply with data protection regulations, such as GDPR or local legal standards.

Scalability and Performance

- The tool must handle a large and ever-growing volume of petitions without performance degradation.
- Ensuring high availability for continuous monitoring and tracking of cases is necessary.

ISSUES & CHALLENGES

Identifying Urgency and Prioritization

- Accurately flagging urgent or high-priority cases requires robust sentiment and context analysis.
- Errors in prioritization can delay critical case resolutions.

Categorization Challenges

- Mapping petitions to the appropriate departments is complex due to overlapping responsibilities between departments.
- Misclassification can lead to inefficient handling or delays.

LITERATURE REVIEW

S.NO	AUTHORS	RESEARCH ARTICLE TITLE	REVIEW OF THE PAPER
1.	Jose Felix Munoz-Sor, Rafael del Hoyo Alonso , Rosa Montanes, Francisco Lacueva	A neural network to identify requests, decisions, and arguments in court rulings on custody	<ul style="list-style-type: none">Challenges in Legal Text Processing – The study highlights key difficulties, such as judicial language complexity and appellate rulings’ intricate structure, which necessitated a two-stage classification (binary long-text and sentence-level classification).Valuable Dataset & Contributions – The dataset, openly available, provides a significant resource for future legal AI research. The study’s insights on NLP in legal decision-making and argument mining can extend to other legal domains.
2.	Vasiliy A. Laptev , Daria R. Feyzrakhmanova	Application of Artificial Intelligence in Justice: Current Trends and Future Prospects	<ul style="list-style-type: none">AI is revolutionizing justice by automating case management, analyzing evidence, and assisting judges in decision-making. It speeds up legal processes, reduces costs, and minimizes human errors.While AI cannot replace judges, it enhances efficiency, ensures transparency, and improves access to justice. Future advancements will further integrate AI into courts, making legal systems more effective.

LITERATURE REVIEW

S.NO	AUTHORS	RESEARCH ARTICLE TITLE	REVIEW OF THE PAPER
3.	Sabine Wehnert, Shipra Dureja, Libin Kutty, Viju Sudhi, Ernesto William De Luca	Applying BERT Embeddings to Predict Legal Textual Entailment	<ul style="list-style-type: none">• This study explores predicting legal textual entailment using BERT-based models. Three methods were tested: a Graph Neural Network with Sentence-BERT, a fine-tuned LEGAL-BERT model, and a KERMIT+BERT model for syntactic analysis. LEGAL-BERT performed best.• Future improvements include better sentence embeddings, visualization tools, and using external knowledge for more accurate legal text classification.
4.	Benjamin M. Chen , Zhiyu Li , David Cai , Elliott Ash	Detecting the influence of the Chinese guiding cases: a text reuse approach	<ul style="list-style-type: none">• The Chinese Supreme People's Court introduced Guiding Cases in 2011 to help judges decide similar cases. However, courts rarely cite them directly. Instead, judges reuse text from these cases without crediting them.• This study shows that Guiding Cases influence decisions more than expected, even though China's legal system officially denies the use of judicial precedent

LITERATURE REVIEW

S.NO	AUTHORS	RESEARCH ARTICLE TITLE	REVIEW OF THE PAPER
5.	Masha Medvedeva, Martijn Wieling , Michel Vols	Rethinking the field of automatic prediction of court decisions	<ul style="list-style-type: none">• This paper reviews past research on using machine learning to predict court decisions. It explains three key tasks: forecasting outcomes, categorizing judgments, and identifying decisions from case texts.• The study highlights the importance of understanding legal data and encourages collaboration between technical experts and legal scholars to improve the accuracy and usefulness of such research.
6.	Ha-Thanh Nguyen , Vu Tran, Le-Minh Nguyen, Ken Satoh	Transformer-Based Approaches for Legal Text Processing	<ul style="list-style-type: none">• This paper explains how Transformer-based AI models were used to process legal texts in the COLIEE 2021 competition. The team tackled five tasks, including retrieving legal cases, finding relevant paragraphs, and answering legal questions.• They improved accuracy by using pre-trained models and generating extra data. Their methods, including NFSP and NMSP, achieved top results and can help future legal AI research.

LITERATURE REVIEW

S.NO	AUTHORS	RESEARCH ARTICLE TITLE	REVIEW OF THE PAPER
7.	Andrea Tagarelli , Andrea Simeri	Unsupervised law article mining based on deep pre-trained language representation models with application to the Italian civil code	<ul style="list-style-type: none">• Deep-learning model designed to find relevant Italian civil law articles based on user queries. It fine-tunes an Italian BERT model to handle legal texts and tackles challenges like many categories and limited training data.• The system helps lawyers and citizens find legal information efficiently and can be adapted for other legal codes.
8.	Masha Medvedeva, Michel Vols , Martijn Wieling	Using machine learning to predict decisions of the European Court of Human Rights	<ul style="list-style-type: none">• This study uses machine learning to predict decisions of the European Court of Human Rights based on court case texts. It achieved 75% accuracy in identifying human rights violations. However, accuracy dropped when predicting future cases.• The research highlights how language analysis can reveal important patterns in legal decisions, though more advanced methods are needed for improvement.

PROPOSED WORK

Data Collection and Preprocessing:

- Gather legal documents from legal databases. Clean and structure the data.
- Remove irrelevant text, tokenize, and apply lemmatization. Perform stopwords removal. Format data for analysis.

Feature Extraction Using NLP:

- Use NLP techniques like Named Entity Recognition (NER). Apply Term Frequency-Inverse Document Frequency (TF-IDF). Utilize word embeddings like BERT and Word2Vec.
- Extract legal features such as terms, entities, and case facts.

Development of the Legal Suggestion Model:

- Develop a deep learning model with transformers (BERT, RoBERTa, GPT). Train the model to classify case outcomes.
- Explore machine learning classifiers (Random Forest, SVM, LSTM). Optimize the model for better predictions.

PROPOSED WORK

Explainability and Interpretability Using AI Analyzers:

- Use SHAP and LIME for model explainability. Explain which factors influenced the predictions.
- Ensure transparency for legal professionals.

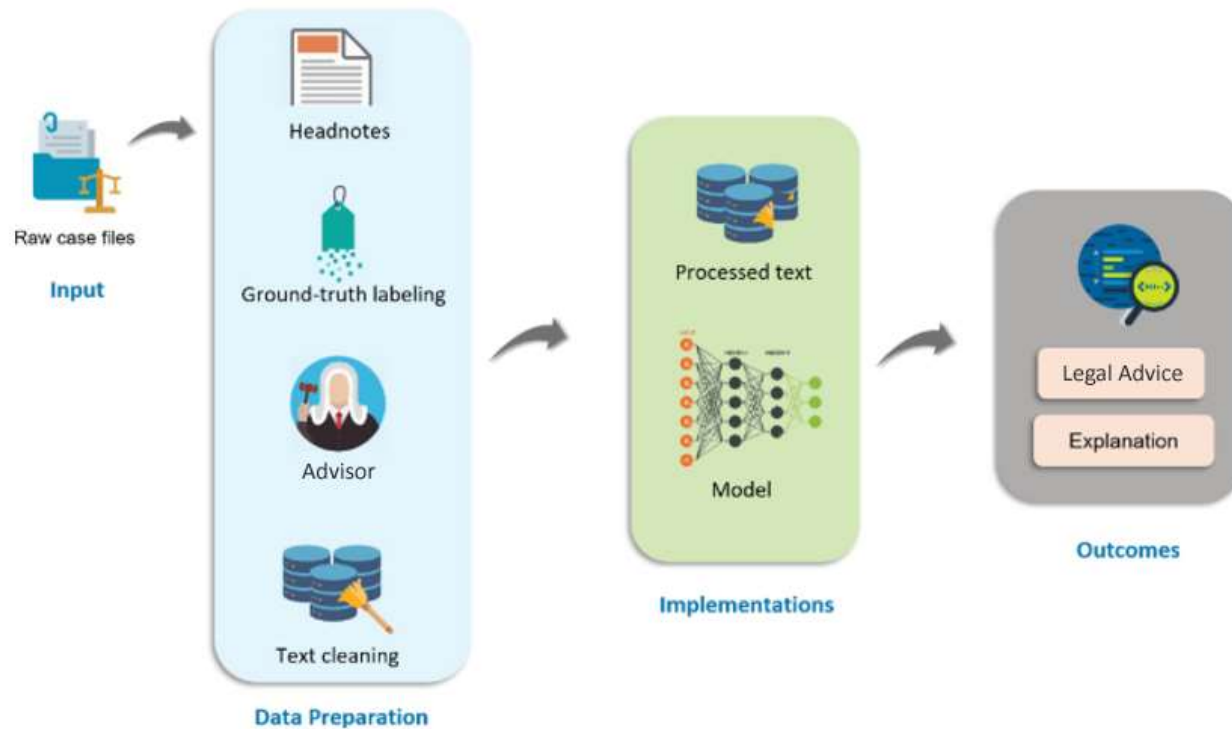
User Interface Design and Development:

- Develop a user-friendly web interface or API. Allow legal professionals to input case details.
- Provide legal suggestions and detailed reports. Ensure scalability and easy integration into legal workflows.

Legal Argument Extraction:

- Use NLP to extract key legal arguments. Focus on reasoning, claims, defenses, and counterarguments.
- Analyze the strength of legal arguments in predicting outcomes.

ARCHITECTURE DIAGRAM



SYSTEM REQUIREMENTS

SOFTWARE REQUIREMENTS:

- Operating System : Windows XP.
- Platform : PYTHON TECHNOLOGY
- Tool : Python 3.6
- Front End : Python anaconda script
- Back End : Spyder

HARDWARE REQUIREMENTS:

- Processor : Intel Core i5 or equivalent.
- RAM : 4 GB or more.
- Storage : 256 GB SSD or more.
- Graphics : Integrated or NVIDIA MX450 or better.
- Internet : Stable connection

MODULE & ITS DESCRIPTION

1. Data Collection and Preprocessing

- This module gathers legal case records from various sources such as court rulings, law databases, and legal repositories.
- Preprocessing involves text cleaning, tokenization, stopwords removal, lemmatization, and formatting to prepare the data for further analysis.

2. Legal Text Feature Extraction

- Utilizes Natural Language Processing (NLP) techniques such as Named Entity Recognition (NER), Term Frequency-Inverse Document Frequency (TF-IDF), and word embeddings (e.g., Word2Vec, BERT) to extract relevant legal terms, case facts, and key arguments.

MODULE & ITS DESCRIPTION

3. AI-Based Legal Suggestion Model

- Implements transformer-based deep learning models (e.g., BERT, RoBERTa, or GPT) along with machine learning classifiers (e.g., Random Forest, SVM, or LSTMs) to analyze legal case texts and predict outcomes.

4. AI Analyzer for Interpretability

- Enhances the explainability of predictions using techniques like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations).
- This module provides insights into which legal factors contributed to the final decision.

MODULE & ITS DESCRIPTION

5. Model Training and Evaluation

- The system is trained on labeled legal datasets using performance metrics such as accuracy, precision, recall, F1-score, and ROC curves to assess its effectiveness.
- Cross-validation techniques ensure model generalizability and reliability across different legal contexts.

6. User Interface and Deployment

- A web-based or API-driven interface allows legal professionals to input case details and receive legal suggestions.
- Provides visualizations and reports summarizing case analysis and prediction results.

IMPLEMENTATION SCREENSHOT

Registration page:

The screenshot displays a web application's registration page. At the top, a red navigation bar contains the following links: DHARMA, POLICE LOGIN, PUBLIC REGISTER, PUBLIC LOGIN, LAWYER LOGIN, ADVISOR LOGIN, and ADMIN LOGIN. The main content area features a large, semi-transparent registration form titled "Register Here" in a maroon header. The form includes the following input fields:

- Username
- Email Id
- First Name
- Last Name
- Phone Number
- Password
- Country
- State
- City
- Address

At the bottom of the form is a yellow button labeled "REGISTER HERE". The background of the page is a photograph of a large, multi-story building with a central courtyard and trees.

IMPLEMENTATION SCREENSHOT

Login Page:



The screenshot displays the login interface for the Dharma system. At the top, a navigation bar contains the following links: DHARMA, POLICE LOGIN, PUBLIC REGISTER, PUBLIC LOGIN, LAWYER LOGIN, ADVISOR LOGIN, and ADMIN LOGIN. The main visual is a large background image of a building with a prominent dome. Centered over this image is a login form. The form has a dark red header with the text "Login Here". Below this header are two input fields: "Username" and "Password". At the bottom of the form is a yellow button labeled "Login".

IMPLEMENTATION SCREENSHOT

Police Dashboard:

The screenshot displays the DHARMA Police Dashboard. On the left is a red sidebar menu with the following options: Dashboard, New Case, Existing Case, More Evidence, Search By Date & Today's Case, and Choose Lawyer. The main content area has an orange header with the DHARMA logo and a search bar labeled "Search By Case No/Case Type" with a yellow "Submit" button. Below the search bar is a white box titled "All Case Details" containing a table with two rows of case information.

ID No	Case No	Case Type	Case Happen Date	Option
1	CASE-YEHQWV0	Civil	Dec. 22, 2024	
2	1002	Civil	Oct. 10, 2024	

IMPLEMENTATION SCREENSHOT

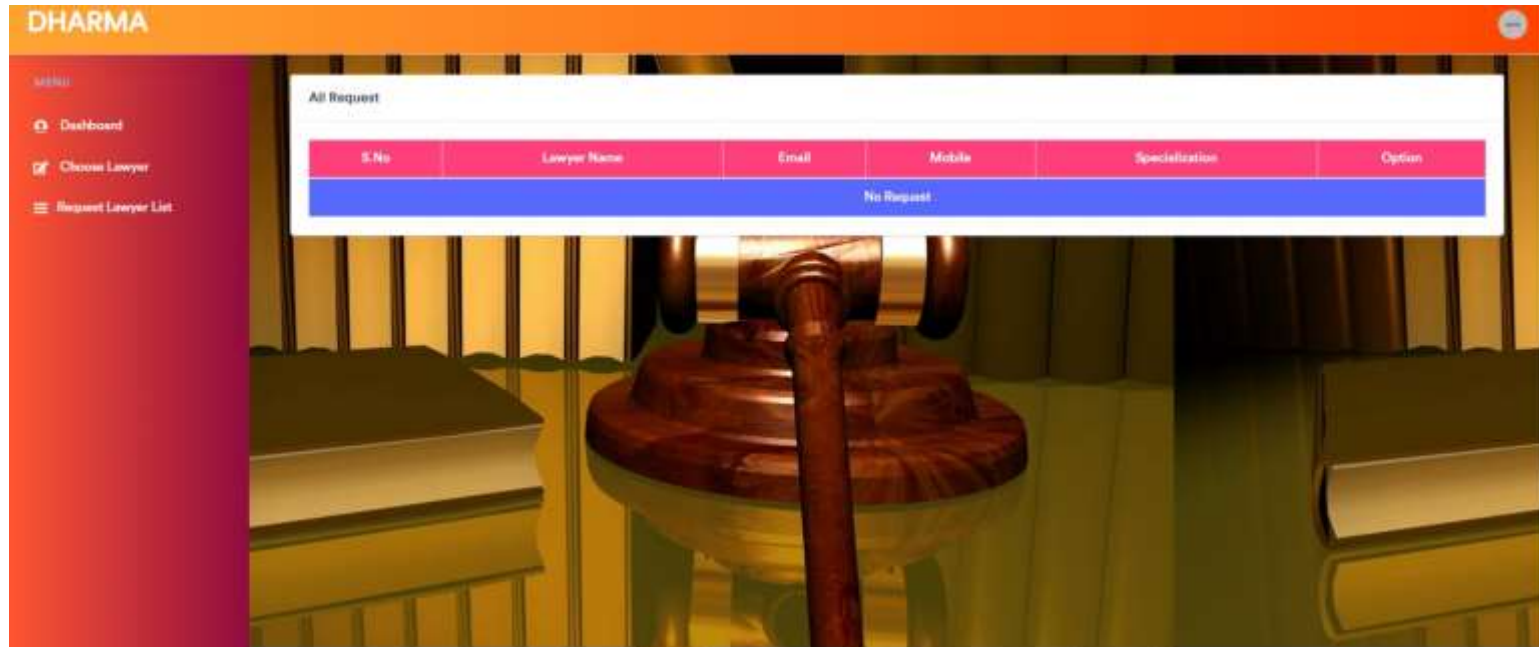
Advisor Dashboard:

The screenshot displays the DHARMA Advisor Dashboard. The interface features a sidebar on the left with navigation links: Dashboard, Solved Cases, Unsolved Cases, Today's Case & Search By Date, and All Case List. The main content area includes a search bar labeled 'Search By Case No/Case Type' with a 'Submit' button. Below the search bar, a section titled 'All Solved Case Details' contains a table with the following data:

S.No	Case No	Case Type	Case Happen Date	Case Status	Option
1	1003	Civil	Oct. 10, 2024	Solved View Advice	
2	CASE-VEHQG9V0	Civil	Dec. 22, 2024	Solved View Advice	

IMPLEMENTATION SCREENSHOT

Public Dashboard:

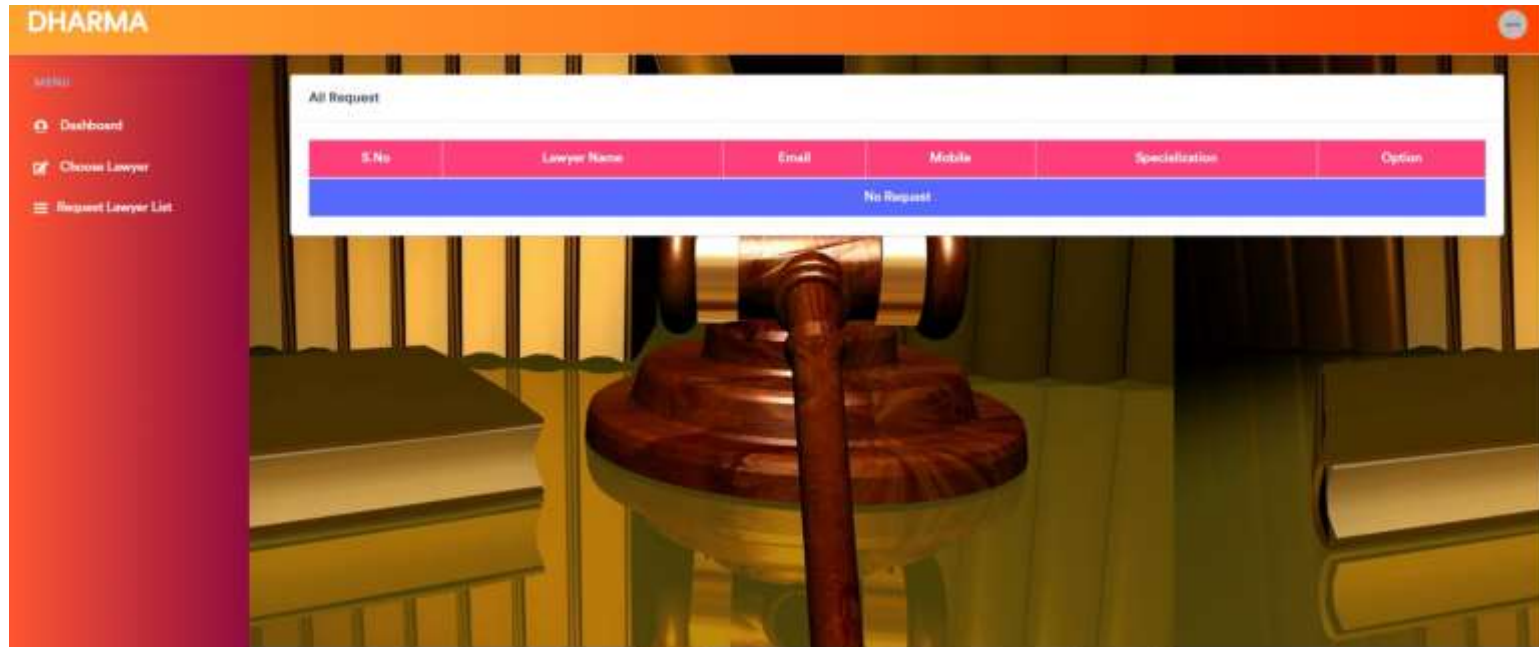


The screenshot displays the DHARMA Public Dashboard. The interface features a red sidebar on the left with the following menu items: Dashboard, Choose Lawyer, and Request Lawyer List. The main content area has a background image of a wooden gavel on a stack of books. A white modal window titled "All Request" is centered on the screen, containing a table with the following structure:

S.No	Lawyer Name	Email	Mobile	Specialization	Option
No Request					

IMPLEMENTATION SCREENSHOT

Lawyer Dashboard:



The screenshot displays the DHARMA Lawyer Dashboard. On the left is a red sidebar with navigation links: Dashboard, Choose Lawyer, and Request Lawyer List. The main content area features a white modal window titled "All Request" containing a table. The table has columns for S.No, Lawyer Name, Email, Mobile, Specialization, and Option. A blue bar at the bottom of the table indicates "No Request". The background of the dashboard is a 3D illustration of a wooden gavel on a stack of books.

S.No	Lawyer Name	Email	Mobile	Specialization	Option
No Request					

REFERENCES

- Jose Felix Munoz-Soro, Rafael del Hoyo Alonso, Rosa Montañes, Francisco Lacueva, “A neural network to identify requests, decisions, and arguments in court rulings on custody”, 2024.
- Vasiliy A. Laptev, Daria R. Feyzrakhmanova, “Application of Artificial Intelligence in Justice: Current Trends and Future Prospects”, Human-Centric Intelligent Systems, vol. 4, pp. 394–405, 2024.
- Sabine Wehnert, Shipra Dureja, Libin Kutty, Viju Sudhi, Ernesto William De Luca, “Applying BERT Embeddings to Predict Legal Textual Entailment”, The Review of Socionetwork Strategies, vol. 16, pp. 197–219, 2022.
- Benjamin M. Chen, Zhiyu Li, David Cai, Elliott Ash, “Detecting the influence of the Chinese guiding cases: a text reuse approach”, Artificial Intelligence and Law, vol. 32, pp. 463–486, 2024.

REFERENCES

- Masha Medvedeva, Martijn Wieling, Michel Vols, “Rethinking the field of automatic prediction of court decisions”, *Artificial Intelligence and Law*, vol. 30, pp. 45–65, 2022.
- Ha-Thanh Nguyen, Minh-Phuong Nguyen, Thi-Hai-Yen Vuong, Minh-Quan Bui, Minh-Chau Nguyen, Tran-Binh Dang, Vu Tran, Le-Minh Nguyen, Ken Satoh, “Transformer-Based Approaches for Legal Text Processing”, *The Review of Socionetwork Strategies*, vol. 16, pp. 221–243, 2022.
- Andrea Tagarelli, Andrea 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. 29, pp. 349–369, 2021.
- Masha Medvedeva, Michel Vols, Martijn Wieling, “Using machine learning to predict decisions of the European Court of Human Rights”, *Artificial Intelligence and Law*, vol. 27, pp. 341–367, 2019.