A Review on AI Applications in Modern Judicial Systems: Current Capabilities and Challenges

Lavina Rani[1], Amita Rani[2]

{24001901004lavina@dcrustm.org[1], amitamalik.cse@dcrustm.org[2] }

Computer Science & Engineering Department, DCRUST, Murthal[1,2]

Abstract. A substantial number of pending cases, such as 44 million unresolved cases in India only, are burdening court systems throughout the world, emphasizing the critical need for scalable technology support. Thanks to expansions in artificial intelligence (AI) and natural language processing (NLP) techniques, which have become a potent facilitator of legal automation. This paper reviews some important legal AI applications and analyzes current datasets, approaches, and research advancements across multiple jurisdictions. The paper studies mainly six legal domain tasks that have gained more researchers' interest in this area, and discusses issues and challenges that researchers face while researching in these specific legal areas, and provides a future perspective.

Keywords: Artificial Intelligence, legal judgment prediction, knowledge extraction, document summarization, legal case retrieval, legal question-answering.

1 Introduction

AI has shown incredible promise in legal work in recent years as AI approaches progress. A human legal expert must possess the necessary specialist knowledge of legal ideas in order to apply the law to a particular legal situation and render a judicial ruling [1]. The problem of "too many cases and too few judges" is still a major concern on a global scale. For example, just 56 family court judges in New York City oversee almost 200,000 cases annually, according to Laurie S. Kohn 2025 [2]. China's Supreme Court implemented docket diversion procedures to give priority to simpler matters in response to similar demands (Agency 2020). Over 44 million cases are still waiting for a decision in India, according to the National Judicial Data Grid (2023) [3]. At the current pace, Justice Katju (2019) estimated that it would take about 450 years to settle all of these matters [4,5,6]. This demonstrates the increasing significance of AI in handling routine, straightforward cases, freeing up judges' time for dealing with more complex legal issues.

According to preliminary estimates, computerized summarization can lessen the period needed for each case examination by as much as 60%, thus saving a substantial amount of time and money for legal firms and court institutions. These models help speed up case groundwork, mainly for attorneys overseeing huge libraries of case law, by minimizing the human labour required to extract important legal information [7]. By improving accuracy and lowering cognitive burden, AI technologies have the potential to support legal practitioners.

IICTDS 2025, December 19-20, Chandigarh, India
Copyright © 2026 EAI
DOI 10.4108/eai.19-12-2025.2366225

The contributions of this paper are:

- This paper highlights areas in the legal domain that are currently being explored more.
- The role of Artificial Intelligence in six broad legal areas such as Legal Judgements Prediction, Legal Knowledge Extraction, Legal Text/Document Summarisation, and Legal Question-Answering.
- The paper also gives a fine lookup of datasets that are publicly available in each particular area.
- Finally, the paper briefly mentions the areas that have little or have not been explored yet, and future directions in the legal field.

1.1 Inclusion-Exclusion Criteria:

This paper analyses only those papers that are relevant to AI in the legal area. Papers are collected using keyword-based search, keywords like "AI in Legal Domain", Legal Text Classification", "Legal AI", "Legal Prediction", "Legal Summarization", "Legal Document Translation", etc. Initially, 124 papers were downloaded, then 42 papers were selected based on their key area, innovative model, and features used in the paper. Most of the papers are mainly SCI/SCIE/ESCI/Scopus-indexed & are between the period of 2019 to 2025. To the best of our knowledge, we have considered all the possible scenarios where AI is being used in any form and based on that we have categorized them all in six broad areas namely: Legal knowledge extraction, legal summarization, legal information retrieval, legal prediction, legal classification and legal reasoning based. Upon analysing the downloaded papers, the authors observed the six broad categories as shown in Figure 1, which cover the maximum legal key research fields. For each category, 6-8 research papers are studied and concluded in literature analysis section 2.

The succeeding sections of the paper are described as follows: The goal and associated effort in a specific legal field are mentioned in Section 2, along with some model descriptions and obstacles. A fine examination of several datasets from diverse legal fields is provided in Section 3. The article is concluded in Section 4, which also offers an outlook on legal topics that artificial intelligence still has to address & its future scope.

2 Applications & Literature Analysis of AI in Legal Areas

There are numerous legal tasks and regions, but some legal domain tasks have garnered considerable interest through the use of Artificial Intelligence (ML/DL) Techniques. This section provides a review of mainly six legal domain tasks namely Legal Knowledge Extraction, Legal Document Summarization, Legal Information Retrieval, Legal Judgment Prediction, Legal Question-Answering, and Legal Document Classification, as shown in Figure 1. The goal is to identify and contextualize these tasks by highlighting new approaches and challenges in these legal fields. Additionally, Table 1 provide an overview of the pertinent literature by describing the model or approach used in that specific field of study.

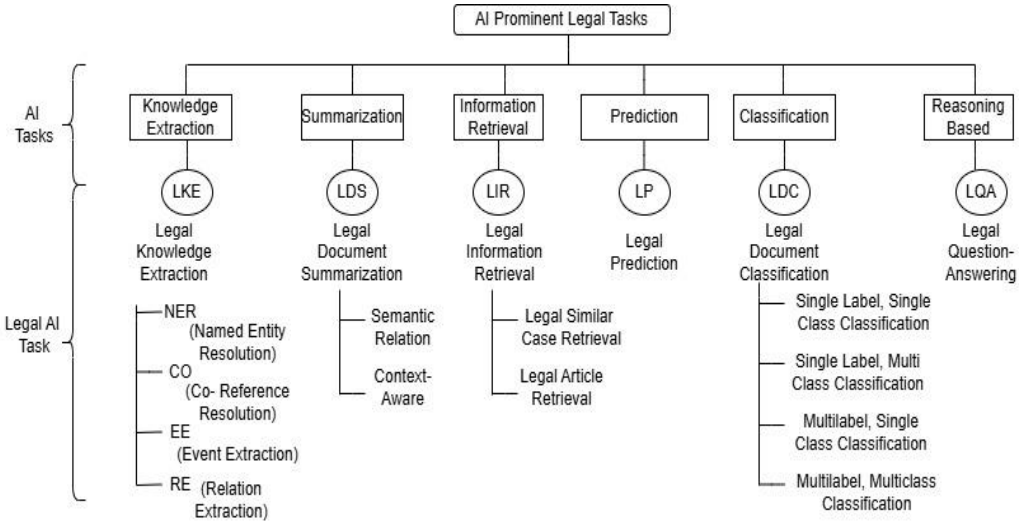


Fig. 1. AI-driven Legal Tasks

The articles & papers selected in the tabular representation (Table 1) have been chosen based on their applicability to the legal field, their recentness (to represent the advanced), and their usage of cutting-edge ideas or techniques. The literature analysis of these legal domain tasks is explored one by one as follows:

2.1 Legal Knowledge Extraction (LKE): Knowledge extraction from unstructured legal documents is the process of automatically identifying the legal entities along with their semantic relationships for various applications such as legal research, constructing knowledge graphs, document analysis, and case prediction. This extracted knowledge can then be used to enable advanced researches that facilitate reasoning and decision support in the legal domain. A variety of tasks are included in knowledge extraction from legal documents, such as identifying entities using Named Entity Recognition (NER), classifying relations among objects by means of Relation Extraction (RE), resolving expressions that refer to the similar object by means of Co-reference Resolution (CO), and classifying proceedings and their associated objects from unstructured legal texts using Event Extraction (EE). In order to recognize certain things referenced in text, such as people, locations, organizations, and roles (Judges, lawyers etc.), Named Entity Recognition (NER) is essential [8]. The information Extraction technique reduces repetitive tasks and facilitates the application of information in many contexts [9]. Due to the complication and specialization of legal language, an information extraction framework that can reliably perform a wide range of activities is required.

Information extraction using ontology is constructed by Denis et. Al. (2017) [10]. Legal Ontology specifies how legal entities are structured and related to each other. Using the inference mechanism, the proposed system incorporates linguistically informed rules into the legal events to efficiently perform the legal knowledge extraction task. Yabo et al. (2025) [11] present a particularly effective approach that involves leveraging pre-trained language models (PLMs) through fine-tuning techniques like LoRA, prefix tuning, and adapter-based methods—which have demonstrated superior performance and adaptability across heterogeneous textual domains. For the extraction of legal information and entities like lawyers, cases, Articles,

Sections and many more from Indonesian court decision documents, Eka Qadri et al. (2025) [12] establish an approach to train NLP models for this purpose.

2.2 Legal Document Summarization (LDS): The legal documents such as case descriptions, legal judgment documents, and others are generally written in complex legal language which are difficult to understand by general public and sometimes to legal professionals. It is challenging to ensure accuracy and consistency when summarizing these materials by hand since it receipts a portion of time and is prone to human error. Computerized summarization arrangements motivated by NLP and ML have become useful tools to expand the effectiveness and dependability of legal investigation and decision-making due to the ongoing growth of legal documents [13]. Legal papers are particularly difficult to summarize due to their intricate structure, complexity, and specialist terminology [14].

Sayyed Usman et al (2025) [4] present a Decision Intelligence framework inspired by Rhetorical Role Relatedness model in order to produce legal case document summaries after analysing them. In another research paper, in order to reduce the dimensionality of the TF-IDF embeddings, Alok Kumar et al (2025) [7] implement most popular dimensionality reduction approach, Principal Component Analysis (PCA), and Cosine Similarity Generation with TF-IDF and Name Entity Recognition method, which lays the groundwork for future advancements and in-depth exploration in the area of sentence importance prediction. In another research, to improve both accuracy and relevance, Vishnu Kesav et al. (2025) [8] fine-tuned two models—Google T5-small and Facebook BART—using rhetorical labels. These labels help the models process distinct legal components, such as facts, arguments, and conclusions. The team used LoRA adapters to ensure the fine-tuning was efficient and minimized computational costs, which also allows the models to differentiate between various legal roles for generating relevant output. A more advanced ensemble-based framework was developed by Maha Ahmed et al. (2025) [13] to enhance the summarization of Turkish legal documents. The framework successfully leverages both extractive accuracy and abstractive writing style by combining these transformer models namely GPT-3.5 Turbo, LED, BART-Large, and Long-T5, and employs three novel components (CASVM, WHSRF, and T5-Meta). The results are tested against 2,000 Turkish court decisions, showing the overall quality of the generated legal summaries.

2.3 Legal Article and Case Retrieval (LCR): LCR is the way of identifying similar cases based on a certain query and, in turn, saves a lot of time for Judges, lawyers, and legal professionals in searching for them. Deep Learning (DL) techniques & large language models (LLMs) are being used as a recent approach to handle the complexity and lengthy text of legal documents.

PromptCase, a legal case retrieval framework proposed by Yanran Tang et al (2023) [15], is basically an encoding scheme that uses prompt learning with LLMs and handles the challenges like legal context preservation & alignment of legal features. Nguyen Ba Thiem et al. (2024) [16] developed a document retrieval and QA system for Vietnam legal cases based on Retrieval Augmented Generation (RAG) and enhanced the information re-ranking & legal knowledge fusion methods. Chunyun Meng et al (2025) [14] proposed the HLAF-LCR (Hierarchical & Logic-Aware Framework for LCR) framework to case facts, reasoning among them, and ruling Structure for complex legal documents. Question-Context-Answer-based fine-tuning (QCATune) model proposed by Nhu Hai Phung et al. (2025) [17] efficiently finds the context-aware relationships between question, answer information, and contexts.

2.4 Legal Prediction (LP): LP has become a crucial area of the legal domain using AI that uses case data and legal precedents to predict decisions [19] of courts, companies, or organizations. Predictions can be of various kinds, such as Legal judgment prediction (LJP), bail prediction (Grant/Dismiss), insurance claim prediction (Accept/Reject), Legal precedent prediction, and Next Sequence Prediction. Based on factual descriptions, LJP automatically predicts judicial outcomes, such as relevant law articles [9], charges [20], and punishment terms [21]. It is frequently tackled as a challenge of text categorization [22]. Beginning with early mathematical models of legal decision-making (Kort 1963 [23]; Keown 1980 [24]), LJP research has developed along with the advancements in AI.

Souvik et al. (2022) [25] developed a methodology for predicting applicable law sections and subsections based on the facts provided and legal factors. In 2022, Lin Li et al. [20] introduced DCSCP (Double Layer Criminal System for Charge Prediction), a model for predicting criminal charges in Chinese law. It achieves highly detailed, multi-granularity inference by analysing the facts provided as input to the model. In recent research, Jiahui and Jianquan (2025) [21] leverage domain-specific legal keyword knowledge by embedding charge-related keywords and strengthening the semantic alignment among legal case facts and keywords, while Bin Wei et al. (2025) [26] propose a loss-minimization-based symbol–neural integration framework that embeds symbolic representations into neural network models.

2.5 Legal Question-Answering (LQA): In legal matters, if someone wants legal advice, they may consult any suitable and qualified law professional; otherwise, LQA systems are useful in the legal field that expedite legal research, enable comprehensive document analysis, and provide practitioners and the general public with conveniently available legal advice and information [27, 28, 29]. As expert assistants, Legal Question Answering (LQA) systems are capable of effectively identifying and summarizing the salient features of recently encountered legal situations. Previous studies on knowledge distillation have mostly focused on text categorization, generally ignoring generating tasks such as question answering.

Biralatei et al. (2019) [1] generate criminal law ontology from collections of legal official documents attached with certain legal rules using a partial-automated approach to leverage the logical thinking in LQA systems. For Legal Multiple-Choice Question Answering (LMCQA), Ying et al. (2021) [30] proposed an ensemble model that incorporates a Bilinear Attention mechanism with DeBERTa- a variant of the BERT transformer model. Their system can handle complicated problems and makes use of the information base for pre-training, collection, prediction, and response. This Model helps in the preparation of judicial exams. In a paper, Haiguang et al. (2022) [31] integrated the Bilinear Convolutional Neural Networks with Attention Networks to present a hybrid BCA model. This model is basically an enhancement of their team's previous model, developed in 2021, for the purpose of judicial exams using AI in Law. The BCA system first locally extracts features, then tries to extract the semantic relationship among questions with their suitable answer when there are multiple answers. The correct results are then refined using post-processing. Recently, Paolo Italiani et al (2025) [32], came up with a new approach, namely Ace-Attorney, which generates datasets suitable for LQA and supervised Small Language Models (SLM) utilizing Knowledge-Distillation Driven Large Language Model (LLM KD). Their method simulates a realistic retrieval-based environment, enabling the Selective Generative Paradigm to more effectively identify the correct documents for answering legal questions. These advancements collectively aim to support overburdened judicial systems and ensure timely access to justice.

2.6 Legal Document/Text Classification (LDC): Legal text categorization is more challenging than typical NLP jobs because of overlapping categories and sophisticated legal terminology, where little textual changes might affect the legal meaning. Automated text categorization is made more difficult by the particular vocabulary and context-dependent rules found in legal

Table 1. Related work in various legal areas

Ref	Focus	Model/ Method	Features	Technique/Technology Used	Findings	Limitations/Challenges
[10] LKE	Legal Domain in Brazil	Information extraction based on ontologies	Ontology-based Information Extraction; identify-formal charges, Acquittal, Conviction, Questioning	SWRL rules, OWL, SAURON system, ODomJurBR Ontology, POWLA data model, and ENI algorithm	Faster the IE process by running reasoner sentence by sentence , hence use less memory	Document size, computer memory to run IE Process
[12] LKE	Indonesian –(Theft cases)	Hierarchical Structure with BIO tagging scheme	Legal Entities Extraction, Named Entities Extraction; (First Instance, Appellate, Cassation)	NLP with Fleiss' Kappa computations, Label Studio annotation tools, and Python web scraping	Improve legal transparency, high annotation quality	No standard & detailed annotated corpora, Manual Annotation
[11] LKE	Chinese real industrial legal cases	AGUIE	Adaptive Gated Universal Information Extraction Model-Enhancement to Universal Information	Improvement to Gated Attention Unit (GAU)	Superior performance on legal texts	Poor performance on boundary identification, complex language structure

			Extraction (UIE) model			
[13] LDS	Indian Legal Document Summarization	Multi-Model Approach: Label Integration	Generates context-aware summaries for legal professionals; Rule-based approach; Prompt-Based Summarization.	Integration of Facebook BART, LoRA Adapters & Google T5-small with Rhetorical Labels, Prompting Strategies, and Adam Optimizer	Context Aware Summaries; improve efficiency & accuracy in decision-making & legal search;	Real-time legal Analysis & Summarization
[14] LDS	Turkish Legal Document Summarization consumer rights	Ensemble-Based Framework	Leverage both extractive precision and abstractive fluency.	Incorporates TurkishTokenizer, TurkishSentenceExtractor techniques, and a hybrid of four transformer models: LED, Long-T5, BART-Large, and GPT-3.5 Turbo.	Evaluates readability and sentence fluency; Effective for factual accuracy; balanced evaluation	Computationally complex; technical terminology of legal texts, phrase continuity not considered.
[4] LDS	Summarization of Indian Legal Documents: Criminal and Tax Law Decisions	Rhetorical Role Relatedness (RRR) Model	A Decision Intelligence framework in the legal domain.	TF-IDF, OpenNyAI NLP Pipeline module, BERTSUM RR module, BUILDNyAI model, LexRank summarization algorithm,	Identifies and summarizes five rhetorical roles in a legal brief document: issues, pertinent facts, argument summaries, court	Legal text's size, internal organization, temporal characteristics, specialist terminology, polysemy, and heterogeneity within legal texts

					decisions, and ratios.	
[7] LDS	Indian Legal Document Summarization	Multifaceted Approach	Integrating: Random Forest, logistic regression, gradient boosting& deep neural networks were implemented	PCA (Principal Components Analysis), TF-IDF with NER, and Cosine Similarity Generation	DNN model is more effective in identifying legally significant sentences than rest 4 models	Irregular Document Object Model Structure; SpaCy NER Inefficiency; Lack of Computational Power
[15] LCR	prompt based information retrieval legal case retrieval China & Canada	PromptCase+SAILER model	Extract legal facts & legal issues from original cases & encodes via proposed prompt-based encoding scheme using language models	BERT-based Encoder with prompts incorporation	Informative case representation Extract legal facts and legal issues	Input length limit, noisy & redundant information

Ref	Focus	Model/ Method	Features	Technique/Technology Used	Findings	Limitations/Challenges
[16] LCR	Vietnamese Legal Information Retrieval	LLM Re-ranker	Vietnamese-specific data processing, enhanced fusion	Ensemble Approach (BM25+Embedding), which includes keyword-based &	Enhanced fusion methods, active information re-ranking, reduce	Inefficient data processing, high retrieval

			methods, active information re-ranking for Vietnam QA System.	semantic-based search with Retrieval Augmented Generation (RAG)	the hallucination problem of large language models.	times, limit on input length
[17] LCR	Enhanced Legal Information Retrieval Vietnamese	QCATune (Question-Context-Answer-based fine-Tuning)	LCR (legal information retrieval) is improved by questions, context, and response linkages.	ChatGPT, WarmupLinear scheduler, AdamW optimizer, and LLM	High accuracy even with 50% data; Flexible to changes; adapts to low-resource languages & specialized domains	Scalability & Robustness Issues with limited data availability
[18] LCR	Multigranular Legal Information Fusion-Chinese	HLAF-LCR	The fact-reasoning-ruling structure of legal texts is modelled by the Hierarchical and Logic-Aware Framework for Legal Case Retrieval.	TF-IDF, BM25, and hierarchical dynamic masking technique	Fusion mechanism driven by logic and cross-hierarchical semantics	Lengthy text, high legal citations, deep domain-specific terminology, long-range logical dependencies
[20] LP	Prediction of Charges under Chinese Criminal Law	Double-layer Criminal System for Charge Prediction, or DCSCP	Conviction process simulation with	Word Mover's Distance (WMD), Python 3.6, Adam optimizer, first-order predicate	Macro-F1 higher than SVM and Qajudge & lower	Lack of labeled data, remote supervision, a greater propensity towards obtaining

			improved interpretation,	logic inference, and ELECTRA as a pre-training model	than other deep learning models	candidate charge while training
[25] LP	Section Prediction-Indian Income Tax Act of 1963	Multi-Label Classification-based methodology	ML and NLP approaches to forecast the relevant legal sections & subsections from the case description.	ML+NLP; ML Models (LR, DT, NB, CNN, SVM); Scikit-learn's Pipeline package (2 Transformers-TfIdf vectorizer, OneVsRest wrapper);	SVM with RBF (Radial Basis Function) shows better performance than other ML Models	Needs human expertise and efforts in analysing law articles
[36] LP	Judgment Prediction in 15 Different Indian Raw Criminal Cases	Chaotic Opposition Learning-based Pelican Optimization or COLPO Method	Bi-GRU and DeepMaxout NN are used with score-level fusion in a modified pelican optimization method. improved case facts' entropy and semantic similarity.	Bi-GRU and DeepMaxout Classifiers, Metaheuristic Optimization Techniques/Algorithm, and personalized ELMo word embedding	Achieves global optimization (enhances optimization algorithms), resolves the overfitting problem. COLPO algorithm	Problems include homonyms & context-dependent vocabulary; long legal text/documents, challenging case material extraction, lack of training data
[37] LP	Fault determination in traffic accidents for Korean legal system	Traffic Accident Criminal Sentencing Prediction Method based on large language models or TRACS-LLM	Predicts 3 things: 1. imprisonment Period; 2. Probation Possibility (grant or not); 3. Fine Amount	Combine the GQA model with ALBERT+Bi-LSTM;	Precise sentence prediction for traffic accident cases, with the lowest RMSE for predicting fines and imprisonment & maximum F1-score for	Sentiment analysis, Not Generalized for other language legal systems.

					predicting probation	
[21] LP	Prompt Learning & charge keyword Based judgement Criminal Law	A Prompt Learning approach using legal keywords integrated into the language model	Criminal Law of the Chinese People's Republic Court. Case Study on-Robbery and plunder (Confusing Charges)	Lawformer: to acquire language and contextual knowledge; TF-IDF: to extract multiple keywords; SVM: classifier; Prompt Templates	Better than baseline models with some interpretability feature.	Semantic correlation between legal keyword labels & legal fact descriptions. Input length limit, Results interpretability needs improvement
Ref	**Focus**	**Model/ Method**	**Features**	**Technique/Technology Used**	**Findings**	**Limitations/Challenges**
[26] LP	Prediction of Judgments in Chinese Private Lending Cases	Neuro-symbolic framework for LJP -A Hybrid Approach (Civil Cases)	Create a Chain-of-Thought (CoT) prompt and use LLMs (GPT-4) for Fact Elements. Removal from court cases	Deep neural networks (DNNs) with first-order logic (FOL) rules (for representation of legal knowledge)	Outperforms the MSJudge method in terms of predictive performance and judgment prediction interpretability.	Mislabelling, external calculation tools, Limited Legal Rules, implementation complexity, resource consumption in FOL rules.
[1] LQA	USA Bar Examination Questions	A System based on Legal NER and criminal law ontologies and rules (CLOR).	Systems of legal ontologies and rules that enable a legal question-answering	Stanford CoreNLP; Named Entity Recognition System (NERS); Ontology+Rule set (OWL axioms, SWRL rules)	Each question's generated ontologies depict legal ideas and the connections	Complex compound nouns, polysemy,legal NER, defeasible reasoning, scalability

			tool with a reasoning facility		between them in criminal law and process.	issues, and implicit information in legal text.
[31] LQA	National Judicial Examination of China	Bilinear Convolutional Neural Networks and Attention Networks, or BCA	Two modules: the Case-Analysis Module (CAM) clarifies the semantic differences between the query & its answers choices and the Knowledge-Driven Module (KDM) extracts local information.	BM+DAM+Random: (Bilinear Model+Data Augmentation Module); CNN, BERT, Prompt Templates; Softmax function; Data Augmentation Module (DAM)	The model's Recognition accuracy for the right response may be enhanced by using BM to analyse and forecast based on the question type and DAM.	Absence of outside expertise in the legal field, the model's integration of outside expertise, and the application of knowledge through multi-hop reasoning
[30] LQA	The Chinese Judicial Examination	DeBERTa-Att-LMCQA -	Determine intricate link between legal articles, queries, & response alternatives using DeBERTa & Attention Legal Multi-Choice Question Answering	LLM (BGE) for encoding; bilinear attention mechanism for feature extraction; softmax function for answer prediction; & cross-entropy loss.	Improved LQA performance and precise retrieval of legal articles	Requires domain-specific training & struggles with legal uncertainties & ambiguities

[32] LQA	Civil Rights litigation & Privacy Policies	The Ace-Attorney method	Automatically creating LQA-specific datasets & supervised SLM models via LLM Knowledge Distillation.	Python 3.8.10, PyTorch 1.12.1+cu113, HuggingFace Transformers, & SLM (Small Language Model)	Address scarcity of public LQA corpora; produce LQAspecific datasets & supervised SLMs through LLM KD.	Considered only a single dataset, lack of multilingual legal datasets
[34] LDC	Classification of Spanish court rulings/judgementwith several labels	A hybrid system that blends explanations with machine learning	Discuss the intricate multi-label multi-class classification of court rulings & the automatic explanation	ML+NLP+XAI XAI- Explainable Artificial Intelligence	Comprehensive entity detection method to extract pertinent judicial entities (jurisdiction, court, & case type) among others.	Manual Annotation by Lawyers
[36] LDC	Classification: Korean Sexual Offense Cases	On the task of classifying legal data, KLUE-BERT small LM outperforms GPT-3.5 & GPT-4.0.	The Crime Case Precedent Data is utilized for both training and evaluation in all categorization tests.	After testing a variety of small, medium, big, & very large language models, KLUE-BERT significantly outperformed them.	The most accurate models were KLUE-BERT & KPF-BERT.	An Imbalanced Dataset causes Bias, covering limited categories only, & contextual reasoning.
[33] LDC	Petition Classification of Legal Documents	2-layer Supervised learning Approach	Both document-level & word-level labels are applied to every document in the corpus.	Adam optimization technique, cross-entropy loss function, softmax activation function, & BiLSTM + SVM classifier	BiLSTM & BERT-CNN have high accuracy, BERT slightly higher	Annotation by Specialists

texts. The rapid organization, recovery, and examination of legal documents made conceivable through an effective categorization procedure facilitates legal research and decision-making. The core court in the state, the 15th Civil Court of the State of Alagoas, had a backlog of almost 140,000 lawsuits in June 2017 [33].

Francisco et al (2022) [34] proposed an ensembled model for processing and classifying Spanish legal documents using NLP and DL techniques, along with an explanation. They have collected legal judgments and annotated them with the help of E4Legal professional lawyers. Yuri D. R. et. Al (2025) [33] introduces a hybrid supervised learning architecture for petition classification. It involves a word-level deep neural network (e.g., BiLSTM or BERT) structured as an encoder–decoder Seq2Seq model, followed by a Support Vector Machine (SVM) to perform the final classification. Jeongmin et al (2025) [35] review the progression of various text/document classification techniques of legal domain, from traditional ML approaches to cutting-edge large language models (LLMs) and concluded that LDC, the performance of KLUE-BERT is better than generative models like GPT3.5 and GPT4.0.

Even if some may have built a strong model for any legal task but the overall performance of AI models depends upon the datasets to which they are trained. Next section discuss about the datasets and their features.

3 Datasets and their Characteristics

The fundamental building blocks of our artificial intelligence models are the datasets. A sizable corpus is required to train, test, and validate the machine learning and deep learning models. This corpus is referred to as a dataset when it is organized in a CSV or JSON file. Datasets that are publicly accessible across a range of legal areas are included in this section. Table 2 summarizes the structure and essential features of the datasets employed in different legal activities.

From Table 2, it has been noticed that most of the datasets available are Chinese; there are very less datasets in the Indian legal system. The biggest challenge in research in the legal area is the availability of a properly annotated legal dataset of a particular legal area. As we can see in table that most of the datasets that are publicly available are either of small in size or consist of data or files that are not properly annotated. It mandates the interference of legal professionals while collecting and preparing the dataset and to validate the outcome as produced by the model.

Table 2: Publicly available Datasets

Dataset [Ref]	Language	Area/Domain	Characteristics
Zalo2021 (Vietnamese NLP tasks) [17]	Vietnamese	Legal Text Retrieval (LTR) & Legal Question Answering (LQA)	Six JSON files with training, testing, and validation datasets are used for legal information or case retrieval.
Legal Case Retrieval dataset, or	Chinese	Criminal cases	There are roughly 43,000 candidate cases and 107 query cases,

LeCaRD [14, 15]			with 100 cases for each query.
COLIEE2020, COLIEE2021, COLIEE2023 [15, 17]	Canadian (English)	Statute law and case law	2020 saw 650 query cases (520 for training and 130 for testing), with 200 candidate instances for each question; 2023 saw cases from Canada's Federal Court, which uses a common law system.
CAIL2019-SCM (Similar Case Retrieval) [39]	Chinese	Chinese Case Law	8,964 triplets for case matching task
Case Law Collection [43]	American (English)	American judicial decision based	2,572 judgments over 12 queries
CAIL2018 (Chinese AI & Law challenge dataset) [20,21]	Chinese	Criminal cases	Three tasks: sentence prediction, legal provision suggestion, and crime prediction (LJP). 2.68 million criminal cases with 202 criminal charges and 183 criminal law articles.
CAIL 2019 [39]	Chinese	Criminal cases	3 tasks: reading comprehension, element recognition, and similar case matching
CAIL 2021 [39]	Chinese	Criminal cases	seven tasks: reading comprehension, case retrieval, judicial examination, judicial summary, argument comprehension, case tag prediction, and information extraction
CAIL 2022 [14]	Chinese	Criminal cases	8 tasks: judicial examination, event detection, document proofreading, case retrieval, legal public opinion summary, argument comprehension, information

					extraction, and interpretable case matching
CAIL 2023 [39]	Chinese	Criminal cases			7 tasks: judicial examination, conversational case retrieval, case retrieval, fact-finding, argument comprehension, information extraction, and judicial big data model
CAIL 2024 [39]	Chinese	Criminal cases			8 tasks: generating factual information in judgments, generating reasoning in judgments, identifying legal elements and points of contention, searching for and predicting the causes of second-instance reversals, generating legal consultation dialogues, conducting objective judicial exams, predicting multiple-crime judgments, and conducting subjective judicial exams.
CAIL 2025 [39]	Chinese	Criminal cases			6 Tasks: Objective questions for National Judicial Examination, Subjective questions for National Judicial Examination, agent-based courtroom simulation, agent-based mediation simulation, legal numerical computation sentencing factor identification & sentence prediction
JEC-QA [31, 40]	Chinese	National Judicial Examination of China			LMCQA dataset from the Chinese judicial examination. 21072 indefinite multiple-

			choice questions with four options for each question from the examinations of 2009 to 2017 & 30,371 practice questions from various websites
LexGLUE (Legal General Language Understanding Evaluation) [41]	English	Legal NLU tasks	Annotate multiple publicly available datasets & make them ready for various NLP tasks in various applications
AsyLex [42]	English	Refugee Law applications	59,112 documents on the assessment of refugee status in Canada between 1996 and 2022; NPL Tasks; Entity Extraction & Outcome Prediction tasks
GerLayQA [43]	German	civil-law system	21k laymen's legal questions & answers from lawyers
IDTheftCase	Indonesian	Judgment Corpus related to Theft cases	2687 records judicial rulings for knowledge extraction.
ACE04 (*Automatic Content Extraction 2004*) [26]	English, Arabic, Chinese	News, Speech	Entity Detection and Recognition (EDR), Entity Mention Detection (EMD), EDR Co-reference, Relation Detection and Recognition (RDR), Relation Mention Detection (RMD), and RDR given reference entities are the six domains for knowledge extraction. Training Data: Arabic (151,000 words), Chinese (307,000 characters, 154,000 words), and English (158,000 words)
ACE05 [11]	English, Arabic, Chinese	News, Speech	1,800 English, Arabic, and Chinese mixed-genre text files with annotations for

			entities, relationships, and events were used for knowledge extraction.
The 2003 Conference on Natural Language Learning, or CoNLL03 [11]	English, German	News	These datasets of judicial judgment papers are utilized for Relation Extraction (Knowledge Extraction) and Named Entity Recognition (NER) activities that are independent of language.
CivilSum [38]	Indian (English)	Summarization of Indian Court Decisions	23,350 legal case decisions taken from the Indian Supreme Court and other High Courts of India along with summaries written by legal professionals

4 Conclusion and Future Research Direction

In conclusion, this paper has provided an overview of the broad categories of legal systems on which AI is being utilized. Artificial Intelligence is impacting the legal field in terms of improved decision-making, reducing the workload of courts, lawyers, and legal professionals. The paper has covered six broad legal areas, such as legal knowledge extraction, summarization, Legal prediction, legal case retrieval, legal question-answering, and document classification. These areas have been selected based on studies that are being carried out. The paper has discussed thoroughly about models proposed, the features considered to achieve that legal task, the AI techniques used, and possible limitations of that approach or model. The paper has selected mostly those publications that have described new trends in machine learning & deep learning techniques such as neuro-symbolic models, domain-adaptive large and small language models (LLM/SLM), prompt-based techniques, ontology-based, and many more that continue to improve legal tasks to a great extent.

Despite significant advancements in AI techniques and technologies, there exist some potential/significant challenges in these legal areas such as:

(i) The main challenging task of knowledge extraction is to find the accurate semantic relation between entities. Advancements to the AI & NLP techniques may enhance the future of LKE by automatically linking complex legal documents, enabling predictive legal analytics, & intelligent legal research with the help of generative AI models.

(ii) Although AI models are performing well in producing summaries, even for legal case documents, the biggest challenge is now how to control generative AI models from hallucination. One solution is to use RAG (Recursive Augmented Generation) and legal

knowledge graphs, the problem is not resolved completely, which in turn opens the door in this direction.

(iii) The LCR task is especially difficult because of the complex nature of legal language, which demands the high level of understanding. Beyond the capabilities of conventional information retrieval methods, effective LCR depends on complex semantic comprehension, logical inference, and the capacity to match new cases with pertinent precedents.

(iv) Judgment prediction is a very challenging task in practical scenarios. The challenge is like multiple-defendants scenario, accurately predicting the term of penalty because it may be affected by various factors, like the previous record of the culprit, repeating the crime, it increases his penalty term; similarly, if the culprit surrenders, it decreases the term of penalty.

(v) The challenge is to build systems and datasets that apply to diverse legal jurisdictions & various languages beyond English to enhance LQA & improve the global scalability with a hallucination-free language model.

(vi) The complicated and unstructured nature of legal languages, linguistic variety among various jurisdictions, lack of availability of properly annotated legal datasets continue to pose significant challenges in the area of legal research.

Most of the models the theme based. That means they work only for particular case types. So there is need of models that covers all or most sections, cases and documents of the legal system. In future, research can be done in considering broader and realistic cases like multi-defendant cases, judge behaviour analysis, sentiment analysis of defendants' statements, context-aware contract analysis, etc. The goal of utilizing Artificial Intelligence in legal domain is not to replace judges or lawyers, but instead to provide AI facilities to legal professionals and to build a human-centred intelligent legal system and speeding up judicial decision-making process and advancing fair access to justice around the globe. In order to improve speed of judicial decision-making, accessibility, and consistency, the article highlights that AI seeks to supplement rather than replace legal personnel.

References

[1] Biralatei Fawei, Jeff Z. Pan, Martin Kollingbaum, Adam Z. and Wyner: A Semi-automated Ontology Construction for Legal Question Answering. New Generation Computing Vol. 37, pp. 453–478 (2019).

[2] Laurie S. Kohn: Justice Delayed By Design: The Harms of Our Protracted Divorce System. GWU Legal Studies Research Paper No 2025-32 70 Vill. L. Rev. 169 (2025).

[3] National judicial data grid. https://njdg.ecourts.gov.in.

[4] Sayyed Usman Ahmed, Nesar Ahmad: A Rhetorical Role Relatedness (RRR) framework for Legal Case Brief Generation. International Conference on Machine Learning and Data Engineering. Procedia Computer Science, Elsevier, Science Direct vol. 258, pp.3333–3344, (2025).

[5] Maansi Gupta, Nomesh B. Bolia: Factors affecting efficient discharge of judicial functions: Insights from Indian courts. Socio-Economic Planning Sciences Vol. 91, pp. 101755 (2024).

[6] Kshitiz Verma : Analysing HC-NJDG Data to Understand the Pendency in High Courts in India. arXiv:2307.10615 (2023).Top of Form

[7] Alok Kumar, Utsav Upadhyay, Gajanand Sharma, Varsha Arya, Wadee Alhalabi, Bassma Saleh Alsulami, Brij B. Gupta (2025): Integrating Extractive Techniques and Classification Methods for Legal Document Summarization. International Journal of Data Warehousing and Mining, Volume 21, Issue 1, (2025).

[8] Vishnu Kesav Omanakuttan, Rohith A, Uthara R, Geetha M: Indian legal text summarization using large language model. Sixth International Conference on Futuristic Trends in Networks and Computing Technologies (FTNCT06) with Organizing university venue: Graphic Era Hill University, Uttarakhand, India. Procedia Computer Science. Elsevier, Science Direct. vol. 259, pp. 1398–1406, (2025).
[9] Wang, L., Chen, S., Jiang, L., & Zhao, H.: Parameter-Efficient Fine-Tuning in Large Language Models: A Survey of Methodologies. Artificial Intelligence Review, 58, 227 (2025).
[10] Denis Andrei de Araujo, Sandro Jose´ Rigo, Jorge Luis Victo´ria Barbosa: Ontology-based information extraction for juridical events with case studies in Brazilian legal realm. Artif Intell Law 25:379–396 (2017) .
[11] Yabo Liu, Yatong Zhou, Kuo-Ping Lin: Adaptive gated universal information extraction for Chinese legal texts. Expert Systems with Applications. Volume 298, Part B, 1 March 2026, 129801 (2025).
[12] Eka Qadri Nuranti, Naili Suri Intizhami, Evi Yulianti, A. Muh. Iqbal Latief, Osama Iyad Al Ghozy, A systematically procedure to extracting legal entities from Indonesian judicial decisions, MethodsX, Vol. 15, 103610, ISSN 2215-0161, (2025).
[13] Maha Ahmed Abdullah Albayati, Kürşat Mustafa Karaoğlan, Oğuz Findik: Towards efficient multi-legal document summarization: An ensemble approach for Turkish law. Engineering Science and Technology, an International Journal. Elsevier, Science Direct. 70, 102138 (2025).
[14] Chunyun Meng, Cheng Tang, Yuki Todo, Weiping Ding: Multi-granular legal information fusion with adversarial compensation: A hierarchical and logic-aware framework for robust case retrieval. Knowledge-Based Systems, Elsevier, Vol. 325, 113964 (2025).
[15] Yanran Tang, Ruihong Qiu, and Xue Li: Prompt-based Effective Input Reformulation for Legal Case Retrieval. arXiv:2309.02962v1 [cs.IR] (2023).
[16] Nguyen Ba Thiem, Doan The Vinh, Pham Quang Tung, Tran Van Toan: Vietnamese Legal Information Retrieval in Question-Answering System. arXiv:2409.13699v1 [cs.IR] (2024).
[17] Nhu Hai Phung, Chi Thanh Nguyen, Minh-Tien Nguyen, Thu Ha Nguyen, Huu Loi Le, Truong-Phuc Nguyen (2025): A fine-tuning framework based on question, context, and answer relationships for enhancing legal information retrieval. Engineering Applications of Artificial Intelligence, Elsevier, Volume 159, Part B, 111570 (2025).
[18] Dozier C, Kondadadi R, Light M, Vachher A, Veeramachaneni S, Wudali R. Named entity recognition and resolution in legal text. In: Francesconi E, Montemagni S, Peters W, Tiscornia D, editors. Semantic processing of legal texts: where the language of law meets the law of language. Lecture notes in computer science, Springer.vol. 6036,pp. 27–43 (2010)
[19] TÜLAY TURAN AND ECİR UĞUR KÜÇÜKSİLLE: Summarization, Prediction, & Analysis of Turkish Constitutional Court Decisions With Explainable Artificial Intelligence & a Hybrid Natural Language Processing Method. IEEE Access Vol. 13, pp. 59766-59779 (2025).
[20] Lin Li, Lingyun Zhao, Peiran Nai, Xiaohui Tao: Charge prediction modeling with interpretation enhancement driven by double-layer criminal system. World Wide Web. Vol. 25, pp. 381–400 (2022).
[21] Jiahui Li, Jianquan Ouyang: A method of legal judgment prediction via prompt learning and charge keywords fusion. Artificial Intelligence and Law. Springer Nature B.V. (2025).
[22] Qihui Zhao: Legal judgment prediction via legal knowledge extraction and fusion, Journal of King Saud University Computer and Information Sciences. 37:31 (2025).
[23] Kort, F.: Simultaneous equations and Boolean algebra in the analysis of judicial decisions. Law and Contemporary Problems, Vol. 28, pp. 143–163 (1963).
[24] Keown, R.:Mathematical models for legal prediction. Computer Law Journal, Vol. 2, pp. 47–72 (1980).

[25] Souvik Sengupta, Vishwang Dave: Predicting applicable law sections from judicial case reports using legislative text analysis with machine learning. Journal of Computational Social Science 5:503–516 (2022) .
[26] Bin Wei, Yaoyao Yu, Leilei Gan, Fei Wu: An LLMs-based neuro-symbolic legal judgment prediction framework for civil cases. Artificial Intelligence and Law. Springer (2025).
[27] Zhong, Q., Fan, X., Luo, X., Toni, F.: An explainable multi-attribute decision model based on argumentation. Expert Systems with Applications vol. 117, pp. 42–61 (2019).
[28] Martinez-Gil J: A survey on legal question-answering systems. Computer Science Review 48, 100552 (2023).
[29] Martinez-Gil, Jorge: A survey on legal question answering systems. arXiv preprint arXiv:2110.07333 (2021).
[30] Ying Luo, Xudong Luo and Guibin Chen: DeBERTa-Att-LMCQA: A Hybrid Model of DeBERTa and Attention for Legal Multi-Choice Question Answering. Preprint submitted to Elsevier (2021).
[31] Haiguang Zhang, Tongyue Zhang, Faxin Cao, Zhizheng Wang, Yuanyu Zhang, Yuanyuan Sun, Mark Anthony Vicente: BCA: Bilinear Convolutional Neural Networks and Attention Networks for Legal Question Answering. Elsevier. AI Open 3, 172–181 (2022).
[32] Paolo Italiani, Gianluca Moro, Luca Ragazzi: Enhancing legal question answering with data generation and knowledge distillation from large language models. Artificial Intelligence and Law. Springer (2025).
[33] Yuri D. R. Costa, Hugo Oliveira, Valério Nogueira Jr., Lucas Massa, Xu Yang, Adriano Barbosa, Krerley Oliveira, Thales Vieira: Automating petition classification in Brazil's legal system:a two-step deep learning approach. Artificial Intelligence& Law 33:227–251 (2025).
[34] Francisco de Arriba-Pérez, Silvia García-Méndez, Francisco J. González-Castaño, Jaime González-González: Explainable machine learning multi-label classification of Spanish legal judgements. Journal of King Saud University – Computer and Information Sciences 34 (2022) 10180–10192 (2022).
[35] Jeongmin Lee: Legal text classification in Korean sexual offense cases: from traditional machine learning to large language models with XAI insights. Artificial Intelligence and Law. Springer (2025).
[36] G. Sukanya and J. Priyadarshini: Fine-Tuned Hybrid Deep Learning Model for Effective Judgment Prediction. Comput Model Engineering Sciences. 142(3), pp. 2923-2958 (2025).
[37] Hyunsik Min, Byeongjoon Noh: TRACS-LLM: LLM-based traffic accident criminal sentencing. Artificial Intelligence and Law. Springer (2025).
[38] Manuj Malik, Zheng Zhao, Marcio Fonseca, Shrisha Rao, and Shay B. Cohen.: CivilSum: A Dataset for Abstractive Summarization of Indian Court Decisions. In Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '24), Washington, DC, USA. ACM, New York, NY, USA, pp. 2241-2250 (2024).
[39] CAIL (Challenge of AI in Law) website. http://cail.cipsc.org.cn/.
[40] Zhong, H., Xiao, C., Tu, C., Zhang, T., Liu, Z., & Sun, M. : JEC-QA: A Legal-Domain Question Answering Dataset. Proceedings of the AAAI Conference on Artificial Intelligence, 34(05), 9701-9708 ((2020).
[41] Ilias Chalkidis, Abhik Jana, Dirk Hartung, Michael Bommarito, Ion Androutsopoulos, Daniel Martin Katz, Nikolaos Aletras: LexGLUE: A Benchmark Dataset for Legal Language Understanding in English. arXiv:2110.00976 (2021).
[42] Barale, Claire, Mark Klaisoongnoen, Pasquale Minervini, Michael Rovatsos, and Nehal Bhuta: Asylex: A dataset for legal language processing of refugee claims. In The 5th Natural Legal Language Processing Workshop 2023, pp. 244-257. Association for Computational Linguistics (ACL) (2023).

[43] Marius Büttner and Ivan Habernal. Answering legal questions from laymen in German civil law system. In Proceedings of the 18th Conference of the European Chapter of the ACL vol. 1, pp. 2015–2027 (2024).