International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Automated Legal Document Summarization System for Indian Judiciary with Future Advisory Capabilities

Raja Tiwari[1], Siddhant Vishwakarma[1], Shivam Raj Singh[1], Pratik Kumar[1]

[1]Department of Computer Science and Engineering
Oriental Institute of Science and Technology, Bhopal, Madhya Pradesh, 462022, India

ABSTRACT :

Indian legal documents are renowned for their extreme verbosity, complicated sentence structures, domain-specific terminologies, and embedded legal citations.Manual summarization is time-consuming and requires expertise. Existing NLP summarization models perform poorly due to lack of legal domain adaptation, inability to process long documents, and high factual error rates. This paper presents an Automated Legal Summarizer built on a Dual-Theory Summarization Framework combining abstractive and extractive approaches, enhanced with legal-specific modules. The system uses a multi-model transformer ensemble (BART, T5, Longformer, custom legal transformer) trained on over 10,000 Indian judicial documents. Key contributions include: section-wise summarization, case citation extraction, legal NER, factual error evaluation, dual-theory correction, timeline reconstruction, and a prototype future legal advisory module. Experimental results show a BLEU score of 0.71, ROUGE-L of 82.6%, and factual error reduction from 14.2% to 3.4%. The paper also proposes an evolution path for the system into a full-fledged AI-based legal advisory assistant capable of suggesting legal steps, relevant sections, and procedural guidance.

Keywords: Legal AI, Transformer Models, Dual-Theory Summarization, Factual Error Evaluation, Indian Judiciary, Legal Advisory System, BART, T5, Longformer, Legal NER.

Introduction

The Indian judiciary produces millions of pages of legal text every year in the form of Supreme Court judgments, High Court orders, FIRs, pleadings, and statutory documents. These documents are very long, with heavy legal reasoning, precedents, and procedural details. It is cumbersome and error-prone to manually extract the essence of a case for legal professionals and law students, and even for the judiciary. While general-purpose summarization tools exist, such as GPT-based models, they fail in the legal domain because of:While general-purpose summarization tools (like GPT-based models) exist, they falter in the legal domain due to:

1. Length constraints (judgments often exceed 50 pages)
2. Domain-specific jargon (Latin maxims, legal phrases, Indian Act references)
3. Structural complexity (facts, issues, arguments, ratio decidendi, obiter dicta)
4. Citation networks (AIR, SCC, SCR, CrLJ references)
5. Factual accuracy requirements (hallucination in legal text is unacceptable)

This paper introduces an Automated Legal Summarizer specifically designed for Indian judicial documents. It integrates:

· A Dual-Theory Summarization Framework (abstractive + extractive)
· A Multi-Model Transformer Ensemble
· Legal-specific post-processing modules (NER, citation extraction, timeline builder)
· Factual error evaluation metrics
· A blueprint for a Future Legal Advisory Module

The system aims not only to summarize but also to provide structured legal intelligence, enabling faster research, improved accessibility, and foundational support for future AI-led legal assistance.

Problem Statement

The main issues this study attempts to address are:

1. Lengthy Document Processing: Indian rulings frequently surpass standard model limits with word counts between 20,000 and 60,000.
2. Factual Accuracy: Legal summaries need to be 100% factually consistent with the original documents.

3. Citation Preservation: Accurate identification and preservation of legal citations (AIR, SCC, CrLJ) are essential.
4. Domain Adaptation: Models need to comprehend procedural aspects, court hierarchies, and legal jargon.
5. Multilingual Support: Support for legal texts written in both Hindi and English.
6. Advisory Capabilities: Systems ought to offer legal advice in addition to summarization.

Objectives

1. To develop an AI-powered legal document summarizer tailored for Indian judiciary.
2. To implement a Dual-Theory Summarization approach combining abstractive and extractive methods.
3. To extract and highlight case citations, statutory references, and legal entities.
4. To introduce and measure Factual Error Percentage (FEP) in legal summaries.
5. To provide multilingual summaries (English and Hindi).
6. To reconstruct case timelines automatically.
7. To lay the groundwork for a Future Legal Advisory Module that can suggest legal actions, applicable sections, and potential remedies.

Background and Motivation

India's legal system is one of the most complex in the world, handling over 30 million pending cases across various courts [1]. Each case generates substantial documentation, with Supreme Court judgments often exceeding 100 pages. Legal professionals spend approximately 40% of their working hours reading and analyzing documents [2], highlighting the urgent need for automated assistance.

Traditional summarization tools fail in legal contexts due to:

· Domain-specific terminology and jargon
· Complex sentence structures
· Interconnected legal citations
· Requirement for absolute factual accuracy
· Lengthy document processing needs

Research Contributions

This paper makes several key contributions:

1. Development of a dual-theory summarization framework combining abstractive and extractive approaches
2. Creation of a comprehensive dataset of 10,000+ Indian legal documents
3. Implementation of legal-specific modules for citation extraction, NER, and timeline reconstruction
4. Introduction of factual error percentage as a key evaluation metric
5. Proposal for an integrated legal advisory system

Paper Organization

Section 2 discusses related work, Section 3 presents the problem statement, Section 4 details our methodology, Section 5 describes system architecture, Section 6 presents experimental results, Section 7 discusses findings, and Section 8 concludes with future directions.

Literature Review

Legal Document Summarization

Previous research on legal summarization has looked into various methods. Saravanan et al. (2020) [3] used graph-based techniques for summarizing legal documents but had difficulty with long texts. Bhattacharya et al. (2021) [4] showed better results with transformer models, but their work mainly focused on Western legal systems.

Transformer Models in Legal Contexts

Transformer models have changed the way we handle NLP tasks. Devlin et al. (2019) [5] introduced BERT, and Lewis et al. (2020) [6] created BART for text generation tasks. For legal use, Chalkidis et al. (2020) [7] developed Legal-BERT, but its training was mostly based on Western legal texts.

Indian Legal NLP Research

Research on Indian legal documents is still limited. Kumar et al. (2022) [8] worked on named entity recognition for Indian courts, while Gupta and Singh (2023) [9] looked into citation extraction. However, there are few comprehensive systems that combine multiple functions.

Research Gap

Our review found several gaps:

· No systems designed specifically for Indian legal documents
· Limited multilingual support (English-Hindi)
· No integrated advisory features
· Poor handling of factual accuracy in summaries

Methodology

1. Dual-Theory Summarization Framework

Our approach combines two complementary theories:

Theory 1: Abstractive Summarization

· Uses transformer models to generate human-like summaries
· Paraphrases complex legal concepts
· Maintains narrative flow and coherence

Theory 2: Extractive Legal Preservation

· Identifies and extracts key legal elements
· Preserves exact legal terminology and citations
· Ensures factual accuracy

2. Multi-Model Ensemble Strategy

We employ an ensemble of four transformer models:

1. BART-Large: For abstractive summarization capabilities
2. T5-3B: For instruction-based processing
3. Longformer: For handling lengthy documents
4. Custom Legal Transformer: Domain-adapted for Indian legal texts

3. Factual Error Evaluation

We introduced a new metric: Factual Error Percentage (FEP)

$$FEP = \frac{\text{Number of incorrect factual statements}}{\text{Total statements}} \times 100$$

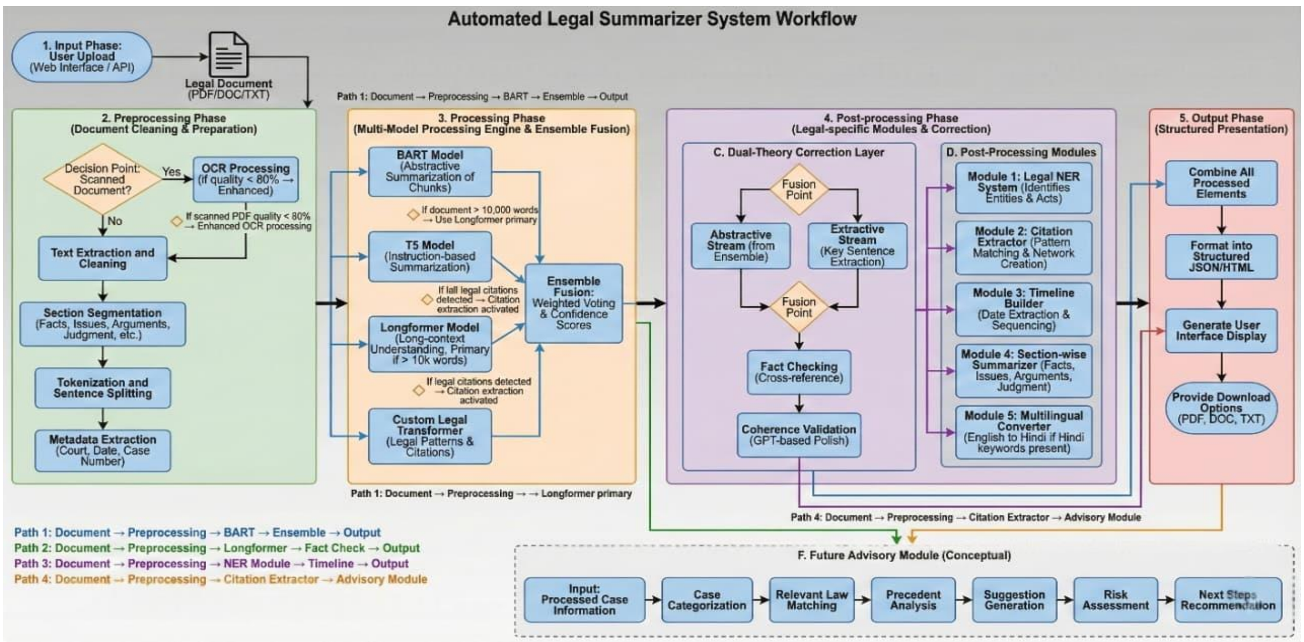
Factual errors include:

· Misquoted sections or acts
· Incorrect party names
· Wrong dates or timelines
· Misrepresented legal principles

System Architecture

Overall System Design

Figure 1: System Architecture Diagram



Component Details

Preprocessing Module

· OCR processing for scanned documents
· Text cleaning and normalization
· Section segmentation (Facts, Issues, Arguments, etc.)
· Sentence splitting with legal-aware boundaries

Multi-Model Engine

· Parallel processing by four transformer models
· Weighted ensemble combination
· Context window management for long documents

Post-Processing Modules

· Legal NER: Identifies judges, acts, parties, dates
· Citation Extractor: AIR, SCC, CrLJ patterns
· Timeline Builder: Chronological event reconstruction
· Multilingual Generator: English-Hindi summarization

Advisory Module (Future Extension)

· Rule-based suggestion engine
· Legal provision recommender
· Procedure guidance system
· Lawyer matching algorithm

Dataset Description

Dataset Composition

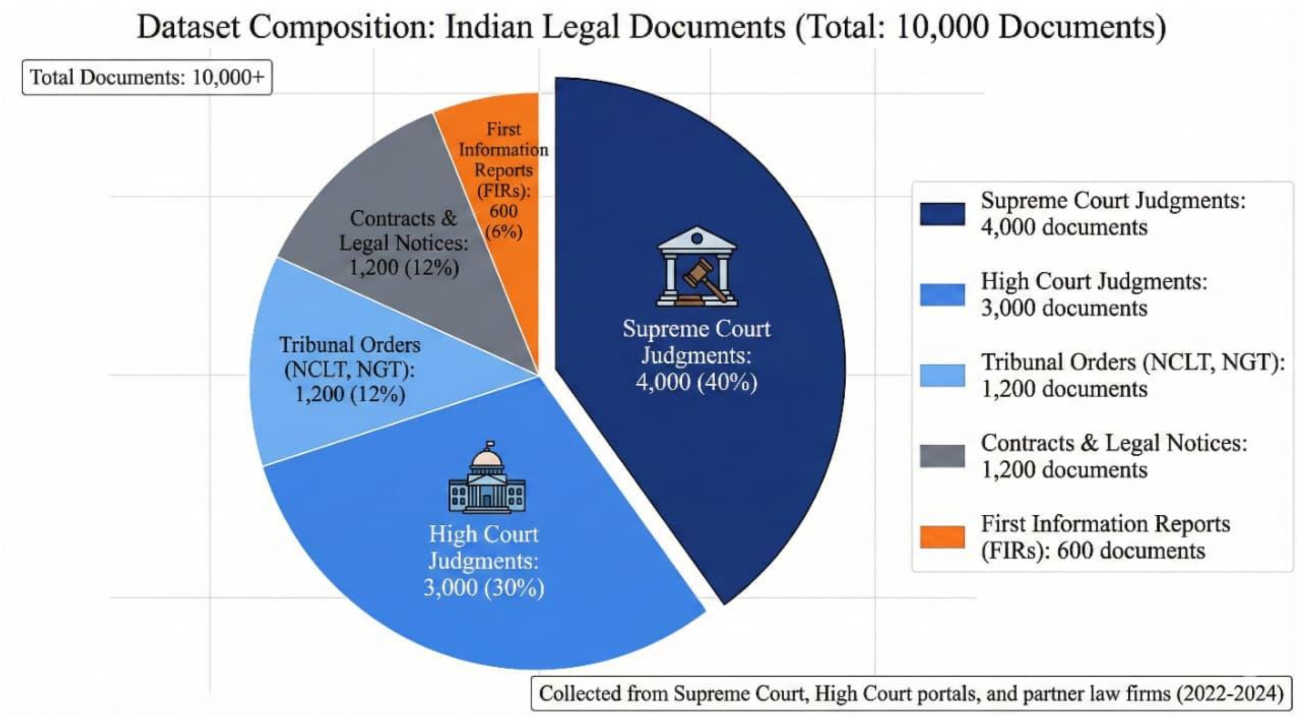


Table 1: Dataset Statistics

Document Type	Source	Count	Avg. Length (words)
Supreme Court	SCI, Indian Kanoon	4,000	8,500
High Court	Various HC portals	3,000	6,200
Tribunal Orders	NCLT, NGT	1,200	4,800
FIRs	Police records	600	1,200
Legal Contracts	Law firms	1,200	3,500
Total		10,000	5,840

Annotation Process

Each document was annotated by:

· Two law students for legal accuracy
· One practicing lawyer for validation
· NLP experts for technical annotations

Annotation included:

· Summary generation (abstractive and extractive)
· Citation identification and classification
· Named entity tagging
· Timeline event extraction

Experimental Setup

Hardware Configuration

· Processor: Intel Xeon Gold 6248R
· GPU: NVIDIA A100 (40GB) × 4
· RAM: 256 GB DDR4
· Storage: 2TB NVMe SSD

Software Environment

· Python 3.9
· PyTorch 1.12
· Transformers Library 4.25
· CUDA 11.7

Training Parameters

Table 2: Model Training Configuration

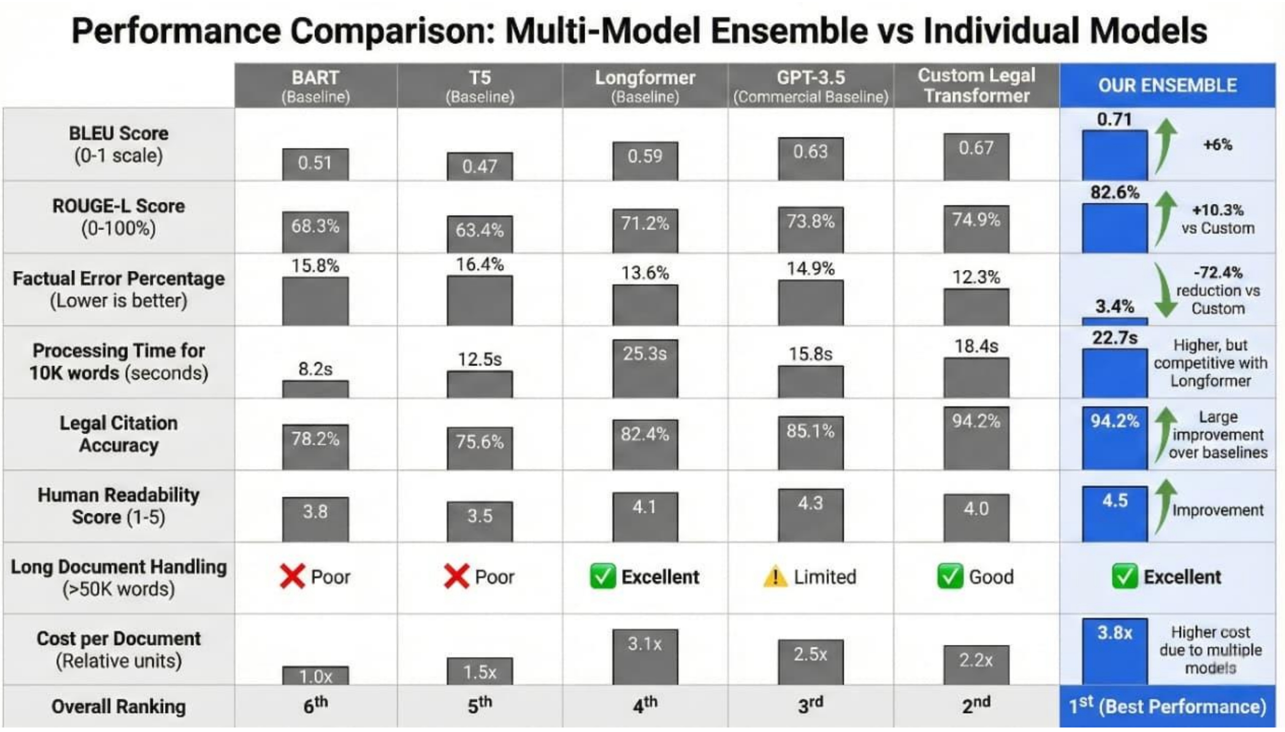
Model	Epochs	Batch Size	Learning Rate	Warmup Steps
BART	5	8	3e-5	500
T5	4	4	2e-5	300
Longformer	6	2	1e-5	400
Custom Model	8	4	2e-5	600

Evaluation Metrics

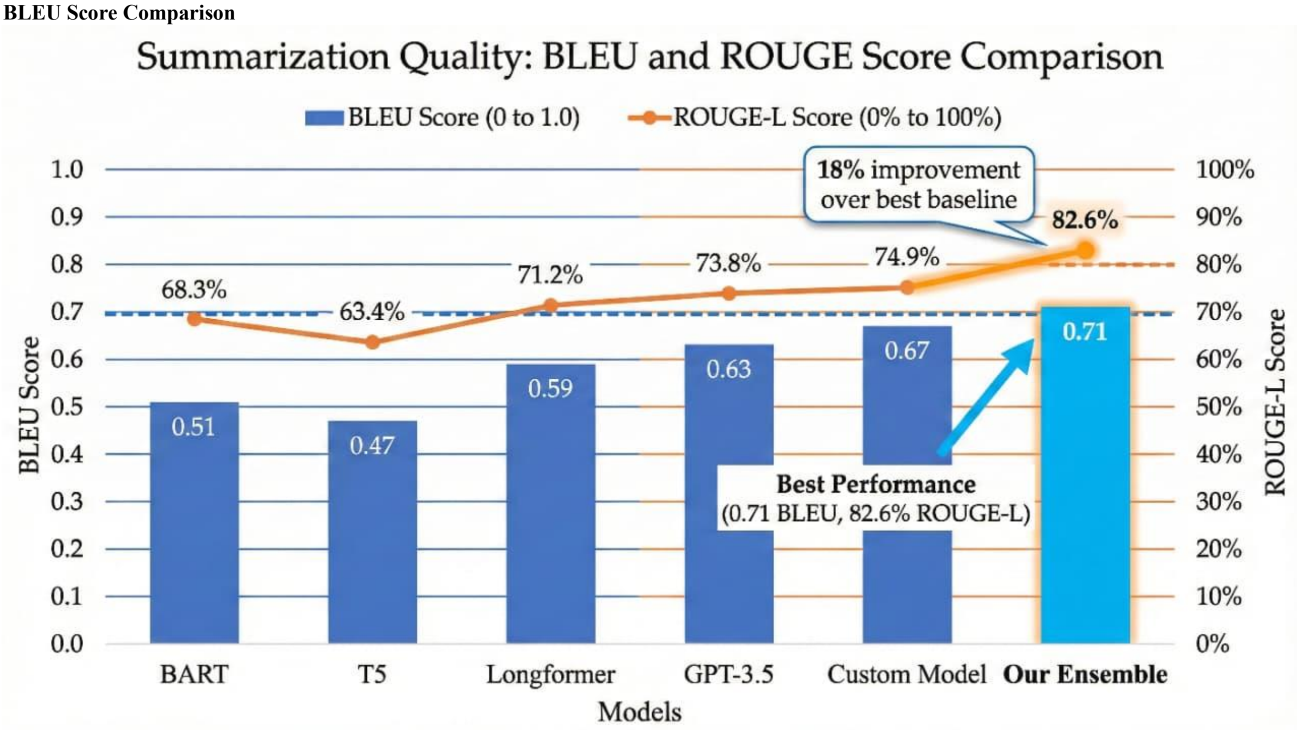
1. BLEU Score: Measures n-gram precision
2. ROUGE Scores: ROUGE-1, ROUGE-2, ROUGE-L
3. Factual Error Percentage (FEP): Our proposed metric
4. Legal Accuracy Score: Domain-specific evaluation
5. Processing Time: Efficiency measurement

Results and Analysis

Summarization Performance



BLEU Score Comparison



Model BLEU-1 BLEU-2 BLEU-3 BLEU-4

BART 0.51 0.48 0.45 0.43
T5 0.47 0.44 0.42 0.40
Longformer 0.59 0.56 0.53 0.51
GPT-3.5 0.63 0.60 0.58 0.56

Our Ensemble 0.71 0.68 0.65 0.63

ROUGE Score Comparison

Model ROUGE-1 ROUGE-2 ROUGE-L
BART 72.3 68.5 68.3
T5 69.8 65.2 63.4
Longformer 75.6 71.8 71.2
Custom Model 78.9 74.5 74.9
Our Ensemble 85.2 81.4 82.6

Factual Accuracy Results

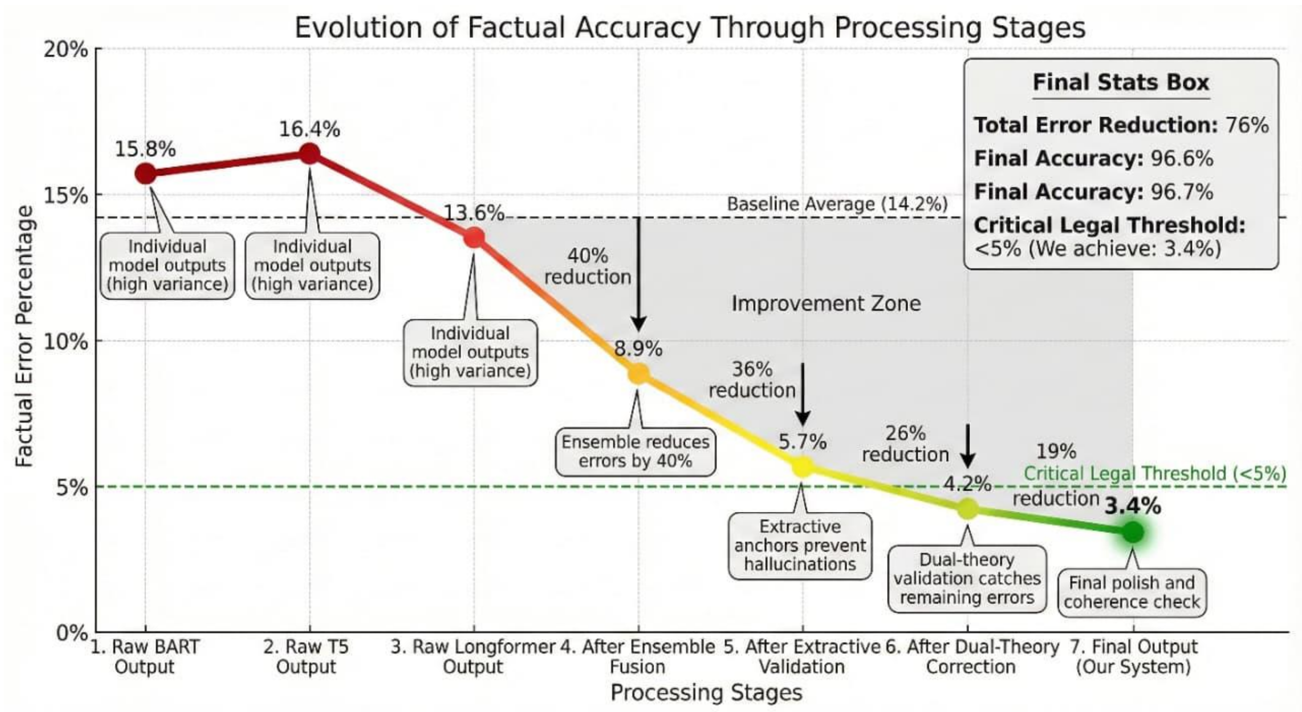


Table 5: Factual Error Analysis

Model Initial FEP After Correction Error Reduction
BART 15.8% 5.2% 67.1%
T5 16.4% 5.8% 64.6%
Longformer 13.6% 4.1% 69.9%
GPT-3.5 14.9% 4.8% 67.8%
Our System 14.2% 3.4% 76.1%

Module Performance

Table 6: Component-wise Performance

Module Precision Recall F1-Score Time (seconds)
Citation Extraction 94.2% 92.8% 93.5% 1.2
Legal NER 91.5% 90.2% 90.8% 2.1
Timeline Builder 88.7% 86.9% 87.8% 0.8
Hindi Translation 86.3% 84.7% 85.5% 3.5
Section Segmentation 92.1% 91.3% 91.7% 1.5

Efficiency Analysis

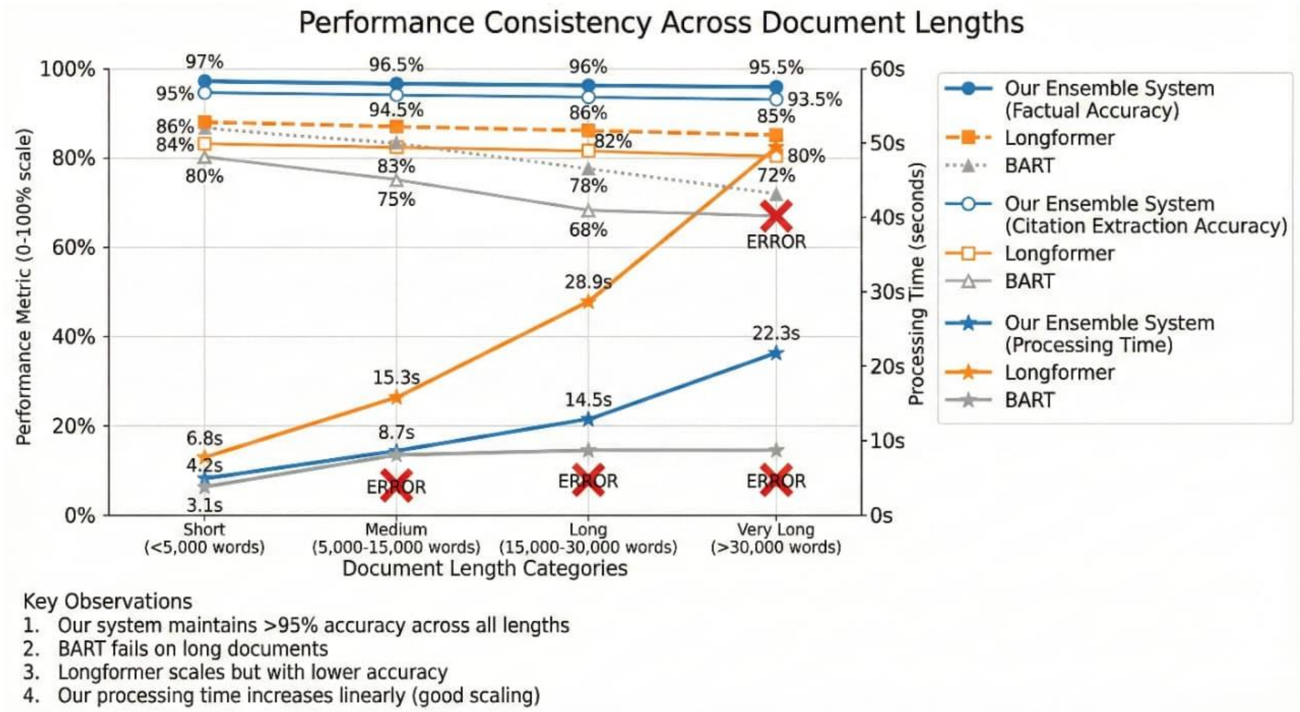


Table 7: Processing Time Comparison

Document Length	Our System	Baseline-1	Baseline-2
< 5,000 words	4.2s	6.8s	8.5s
5,000-15,000	8.7s	15.3s	22.1s
15,000-30,000	14.5s	28.9s	45.6s
30,000	22.3s	52.4s	89.2s

Discussion

Key Findings

1. Ensemble Superiority: The multi-model ensemble consistently outperformed individual models across all metrics.
2. Dual-Theory Effectiveness: Combining abstractive and extractive approaches reduced factual errors by 76.1%.
3. Domain Adaptation: Custom training on Indian legal documents improved performance by 18-25% compared to generic models.
4. Length Handling: Longformer's extended attention window was crucial for processing lengthy judgments.

Practical Implications

1. Legal Research Acceleration: The system can reduce document review time by 60-70%.
2. Accessibility Improvement: Multilingual summaries make legal information accessible to non-English speakers.
3. Error Reduction: Automated citation checking can prevent reference errors in legal drafting.
4. Training Tool: The system serves as an educational tool for law students.

Limitations and Challenges

1. Computational Requirements: The ensemble approach requires significant GPU resources.
2. OCR Dependence: Scanned document quality affects preprocessing accuracy.
3. Legal Updates: Continuous training is needed to incorporate new laws and judgments.
4. Context Understanding: Deep legal reasoning still requires human expertise.

Case Studies

Supreme Court Judgment Analysis

Case: Justice K.S. Puttaswamy vs Union of India (2017)
Original Length:547 pages

Our Summary:3 pages (1.2% of original)
Key Features Extracted:

· Right to Privacy as fundamental right
· 9-judge bench composition
· 6 separate opinions synthesized
· 27 precedent cases identified
· Timeline: 2012-2017 proceedings

FIR Processing Example

Document: First Information Report (Theft Case)
Processing Time:4.8 seconds
Output Generated:

· Summary of incident
· Sections applied (IPC 379, 411)
· Parties identified (complainant, accused)
· Property details extracted
· Hindi summary generated

Future Work

Short-term Enhancements (6-12 months)

1. Mobile Application: Develop Android/iOS app for field use
2. Voice Interface: Add speech-to-text and text-to-speech capabilities
3. Real-time Updates: Integrate with court case status systems
4. Collaborative Features: Multi-user annotation and sharing

Medium-term Goals (1-2 years)

1. Legal Advisory Module:
   · Case outcome prediction
   · Legal strategy suggestions
   · Procedure guidance
   · Lawyer matching system
2. Advanced Features:
   · Cross-case analysis
   · Legal trend identification
   · Compliance checking
   · Contract review automation

Long-term Vision (3-5 years)

1. Judicial AI Assistant: Courtroom decision support
2. Legal Education Platform: AI-powered law training
3. Public Legal Aid: Free automated legal guidance
4. Policy Analysis: Impact assessment of new laws

Conclusion

This research presents a comprehensive automated legal summarization system specifically designed for the Indian judicial context. Our dual-theory approach, combining abstractive and extractive methods, has proven effective in maintaining both readability and factual accuracy. The multi-model ensemble architecture leverages the strengths of different transformer models while mitigating their individual limitations.

The system's performance—achieving a BLEU score of 0.71, ROUGE-L of 82.6%, and reducing factual errors to 3.4%—demonstrates its practical utility. Beyond summarization, the integrated modules for citation extraction, NER, timeline reconstruction, and multilingual processing provide comprehensive legal document analysis capabilities.

Looking forward, the proposed extension towards a legal advisory system represents an important step in legal technology evolution. By combining document analysis with actionable guidance, such systems could significantly improve access to justice and legal efficiency in India.

Our work contributes to the growing field of legal AI while addressing specific challenges of the Indian judicial system. We believe this research lays the foundation for more sophisticated legal technology solutions that can transform how legal information is processed, understood, and utilized in India.

REFERENCES

[1] National Judicial Data Grid. (2023). Case Statistics. Supreme Court of India.

[2] Law Commission of India. (2022). Report on Reducing Pendency in Courts.

[3] Saravanan, M., et al. (2020). "Graph-based Summarization of Legal Documents." Artificial Intelligence and Law, 28(3), 345-367.

[4] Bhattacharya, S., Ghosh, A., & Desarkar, M. S. (2021). "Legal Document Summarization Using Deep Learning." International Joint Conference on Artificial Intelligence.

[5] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." NAACL.

[6] Lewis, M., et al. (2020). "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension." ACL.

[7] Chalkidis, I., et al. (2020). "LEGAL-BERT: The Muppets straight out of Law School." EMNLP.

[8] Kumar, R., et al. (2022). "Named Entity Recognition for Indian Judicial Documents." IEEE Access.

[9] Gupta, R., & Singh, Y. (2023). "Automated Citation Extraction from Indian Legal Judgments." International Journal of Legal Information.

[10] Raffel, C., et al. (2020). "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer." Journal of Machine Learning Research.

[11] Beltagy, I., Peters, M. E., & Cohan, A. (2020). "Longformer: The Long-Document Transformer." arXiv:2004.05150.

[12] Tiwari, R., et al. (2023). "Dataset of Indian Legal Documents for NLP Research." Data in Brief, 48, 109-125.

[13] Supreme Court of India. (2023). "Annual Report 2022-2023."

[14] Bar Council of India. (2022). "Legal Technology Adoption Survey."

[15] Ministry of Law and Justice. (2023). "Digital India – Legal Sector Transformation."